Stopover ecology and population dynamics of migratory shorebirds

by

Anna Maureen Tucker

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
May 5, 2019

Approved by

Conor P. McGowan, Chair, Associate Research Professor of Wildlife Sciences
Christopher A. Lepczyk, Professor of Wildlife Sciences
James B. Grand, Cooperative Research Unit Supervisor, U.S. Geological Survey
James E. Lyons, Research Ecologist, U.S. Geological Survey

Abstract

Migratory animals are declining worldwide, and for many species there are multiple threats to population stability from throughout the annual cycle. Animals performing long-distance migrations use stopover sites en route to replenish fat stores, and the congregation of individuals at migratory stopover sites affords an opportunity to track population health, particularly for species that are highly mobile and difficult to study at other times of year. However, analysis of these data presents several challenges due to the flow-through nature of stopover and uncertainties about how stopover site conditions influence both demographics and migratory behaviors. Here, I explore the use of quantitative methods and hierarchical modeling to lend clarity to our understanding of the role of stopover in the annual cycle of migratory birds, considering both an ecological perspective on the direct and indirect effects of stopover on demographics and migratory behavior, and a statistical perspective on advancing new and existing modeling frameworks for analyzing complex mark-recapture datasets. Using counts, trapping, and mark-resight monitoring data collected for three species of Arctic-breeding shorebirds in Delaware Bay, USA, I evaluate the ways that stopover site conditions and phenology of resource availability influence the timing and rate of mass gain during stopover, variation in site use among years, and population vital rates. Disentangling these responses is important for understanding how populations respond to changes in stopover conditions, and how to interpret changes in stopover passage population size. Understanding the relative role of stopover in the context of the annual cycle is important for informing conservation strategies that will be most effective at halting declines and promoting population stability.

Acknowledgments

I have had the immense privilege of the support of many mentors and friends, but only space to name a few of them here. First, I thank my advisor, Dr. Conor McGowan, for welcoming me into his lab when I showed up uninvited and told him I'd like to study shorebirds in Delaware Bay. Conor's professional and personal support has made this work possible. I'm also grateful to Dr. Geoff Hill for taking me in as a temporary foster student. I thank the many volunteers who collected the data used in these analyses and the many individuals that contribute to the ongoing monitoring of shorebirds via the Delaware Shorebird Project. In particular I thank Audrey DeRose-Wilson, Dr. Nigel Clark, and Jacquie Clark for welcoming me into the team, sharing their insights, and supporting my research endeavors.

I'm grateful to Dr. Lesley Bulluck for her continued mentorship and friendship, to Dr. Kim Derrickson for giving me my first field research experience, and to Ron Smith for planting the seed of curiosity to follow this path and for taking me along on my very first shorebird survey.

I am constantly lifted up by the love and unwavering support of my family. I am grateful to be sharing this journey with my husband, Ryan, who helps me be my best in life and science. Finally, I'm thankful for my dog, Cedar, for convincing me to get away from the computer every now and then to go for a walk in the woods.

Table of Contents

List of Tables

List of Figures

Chapter 1: General Introduction

Migratory animals are declining globally, and much research and conservation effort has been devoted to understanding the myriad threats faced by migratory species and how those threats interact throughout the annual cycle (Wilcove and Wikelski 2008, Flockhart et al. 2015, Rushing et al. 2017). The congregation of individuals during migratory stopover presents an opportunity to track population health (Warnock 2010), especially for species that are highly mobile and difficult to study at other points of the annual cycle, but there are several challenges inherent in the analysis of these data. Here, I explore the use of quantitative methods and hierarchical modeling to lend clarity to our understanding of the role of stopover in the annual cycle of migratory species, considering both an ecological perspective on the direct and indirect effects of stopover on demographics and migratory behavior, and a statistical perspective on advancing new and existing modeling frameworks for analyzing complex mark-recapture datasets.

The role of stopover in the annual cycle migratory species

Seasonal migrations are among the most studied phenomena in ecology (Milner-Gulland et al. 2011, Shaw 2016). A longstanding view has been that migration evolved as animals exploited new breeding areas at higher latitudes that allowed them to escape competition, avoid predation, and take advantage of resource pulses (Cox 1985, Alerstam et al. 2003, Zink and Gardner 2017), but a new hypothesis posits that migratory behaviors evolved so that animals could maintain breeding site fidelity and the many advantages it confers without having to

develop other adaptations for increasingly seasonal environments (Winger et al. 2018). In either case, this has led to migrations across extreme distances and the need for reliable stopover sites en route for individuals to stop and refuel before continuing their journeys (Alerstam 2011). Migratory species move great distances to maintain consistency in their environment and phenological synchrony and predictability is critical to the success of this strategy (Robinson et al. 2009, Miller-Rushing et al. 2010, Stutzman and Fontaine 2015).

Global change has been linked to population declines for many migratory species (Kirby et al. 2008, Pearce-Higgins et al. 2015). Declines may be attributable to conditions at breeding sites, nonbreeding sites, or during migration and the interactions between different points of the annual cycle (Morrison et al. 2016, Murray et al. 2017, Rushing et al. 2017). A full annual cycle perspective aids in identifying the relative influence of conditions experienced at different sites on changes in population growth rate and the ways in which those conditions interact or carry-over to influence success in subsequent periods (Hostetler et al. 2015, Marra et al. 2015). This leads to a fuller picture of the internal and external factors that govern population dynamics (Woodworth et al. 2017, Coates et al. 2018) and can also guide conservation effort toward sites where actions would have the greatest impact (Martin et al. 2007, Taylor and Stutchbury 2016).

Stopover and staging sites provide critical habitat for migratory species but are infrequently considered in full annual cycle studies, which tend to compare mainly breeding and stationary nonbreeding (wintering) sites (Hostetler et al. 2015). Conditions encountered during migration and stopover could have both direct effects on the probability of surviving the journey and indirect carry-over effects on success in subsequent seasons (Newton 2004, 2006; Norris and Marra 2007). Carry-over effects may be particularly important during pre-breeding migrations (Newton 2006). Some species rely on stored energetic reserves from previous time periods to

fuel reproduction, and these capital breeders may depend on resources acquired during migration (Jonsson 1997, Stephens et al. 2009). Resource availability during migration could also influence the timing of migration and arrival to the breeding site (Forchhammer et al. 2002, Both et al. 2005), and earlier breeding is generally associated with greater reproductive success (Kokko 1999). The relative importance of minimizing time or energy spent migrating are linked to strategies for maximizing reproductive output (Alerstam 2011).

That stopover sites provide important habitat for long distance migrants is widely accepted (Cimprich et al. 2006, Skagen 2006, Sawyer et al. 2009). A better understanding of the relative importance of stopover in the context of the annual cycle will aid in developing effective conservation strategies for declining species (Sheehy et al. 2011).

Practical considerations for analyzing monitoring data from migratory stopover

Migratory animals present challenges for monitoring because they are highly mobile and the connectivity among sites is not well established for many species (Webster et al. 2002, Dhanjal-adams et al. 2016). Remote tracking devices such as satellite tags or geolocators offer incredible insights to animal movements and site use but typically are limited in sample size due to cost and logistics of deployment and limited in temporal scope because of device lifespan or attachment method, although rapid technological advances are changing this (Bridge et al. 2011, Wilmers et al. 2015). Stable isotope analysis and genetic markers also offer ways to establish connectivity among breeding and wintering sites (Paxton et al. 2013, Rushing et al. 2014, Pekarsky et al. 2015), but on their own do not inform our understanding of demographic or behavioral responses to migratory conditions. One of the oldest methods of tracking individual animals over time remains one of the most useful—the attachment of individually-identifiable artificial tags (Silvy et al. 2012). These marks are typically inexpensive, easy to deploy, and last

a long time. When tags are field-readable, it becomes easy to harness citizen scientist or volunteer effort to collect re-encounters of marked individuals over large spatial and temporal scales (Gollan et al. 2012, Callaghan and Gawlik 2015). Such datasets may be susceptible to errors related to observer experience or skill, however, and ensuring data quality is important for all analyses (Tucker et al. 2019).

Monitoring populations during migratory stopover affords the benefit of having a large proportion of the population concentrated in a single place on a predictable annual schedule (Horton et al. 2019). This is a good opportunity to efficiently monitor species that are widely dispersed during other points of the year. With just a short period of monitoring effort each year, the resulting datasets can inform changes in migratory phenology (e.g., Salewski and Schaub 2007, Navedo et al. 2010, Kolzsch et al. 2016), population size (e.g., Masero et al. 2011, Lyons et al. 2015), demographics (e.g., Dinsmore and Collazo 2003, McGowan et al. 2011), or site use (e.g., Ydenberg et al. 2004, Woodworth et al. 2014).

The many challenges of analyzing stopover monitoring data can largely be attributed to the lack of population closure within the migratory period (Lyons et al. 2015, Drever and Hrachowitz 2017). In population ecology, closure refers to a state in which the true number of individuals present in the study area is fixed, and no changes occur due to births, deaths, immigration, or emigration (Williams et al. 2002). For many species we can reasonably assume that no births or deaths occur during the stopover period, which is typically very short relative to the average lifespan of the animal, but the continued arrival and departure of individuals from the site is an inherent characteristic of most stopovers. This flow-through nature of stopover sites precludes the use of many analysis methods that assume population closure. Additionally, use of the site by the larger population may vary among years, and relative use of different sites in the

flyway may shift directionally in response to conditions at one or more sites, e.g. habitat availability (Cohen et al. 2014, McCabe and Olsen 2015) or predation pressure (Ydenberg et al. 2004, Smith et al. 2012). Quantitative methods are needed to tease apart the drivers of apparent changes in counts conducted during stopover.

Hierarchical Bayesian modeling approaches offer flexibility in specifying model structures, borrowing information across years, and combining multiple sources of information in unified analyses that propagate estimation uncertainty (Kéry and Schaub 2012, Miller et al. 2019). This framework facilitates identification of ecological signals from datasets with extraneous noise such as observation error, irregular sampling, or missing data.

Shorebird stopover in Delaware Bay

The work presented in this dissertation focuses on shorebird stopover ecology in Delaware Bay, USA. Delaware Bay is located on the eastern coast of the United States between New Jersey and Delaware where the Delaware River outlets into the Atlantic Ocean. The bay shore is comprised of salt marshes, mudflats, and sandy beaches and supports the largest congregation of horseshoe crab (*Limulus polyphemus*) spawning activity in the Atlantic seaboard, likely due to relatively warmer waters and low wave activity on the bay shore beaches (Smith et al. 2009, Smith and Robinson 2015). Horseshoe crab spawning in this area occurs from April to July, with peak spawning activity typically in May and June (Smith et al. 2010). During this time, hundreds of millions of horseshoe crab eggs are deposited into small depressions on sandy beaches (Smith et al. 2002). Those eggs are a high-quality prey item for shorebirds because they are fat- and protein-rich and easily digestible (Tsipoura and Burger 1999, Gillings et al. 2007, Fraser et al. 2010). Due the large aggregation of this high-quality prey during the spring migration period, Delaware Bay is a globally important stopover site for migratory

shorebirds; hundreds of thousands of individuals of several species use this region as a stopover site in mid to late May (Clark et al. 1993).

Horseshoe crabs are commercially harvested, historically for use as fertilizer and currently for use as bait in other fisheries (Smith et al. 2016). They are also collected and subsequently released following blood collection, which has important uses in the biomedical industry (Smith et al. 2009). Increasing and unregulated horseshoe crab harvest in the late 1990s led to declines in spawning abundance in Delaware Bay, which has been implicated as a key contributor to concurrent population declines in shorebirds using this site, particularly the red knot (*Calidris canutus rufa*) (Baker et al. 2004, Niles et al. 2009). An adaptive management framework was developed in the late 2000s to determine harvest limits for horseshoe crabs, using linked population models for horseshoe crabs and red knot to predict the effects of management actions on the system (Smith et al. 2013, McGowan et al. 2015). Central to this framework are models that represent competing hypotheses about how horseshoe crab abundance impacts red knot demographics (McGowan et al. 2011b, a). Improving estimates of red knot demographic vital rates and refining our understanding of how the abundance and timing of horseshoe crab spawn impact red knots will be useful to the continued implementation of the adaptive management plan.

Among shorebirds that use Delaware Bay during stopover, red knot have experienced the most drastic population declines and are therefore thought to be especially reliant on horseshoe crab eggs (Baker et al. 2004, Niles et al. 2008). Their particular dependence on horseshoe crabs could be due to the extreme distance of their spring migration (some overwinter as far south as Tierra del Fuego, Argentina) or physiological constraints imposed by the shrinkage of digestive organs during this period to allow for larger flight muscles (van Gils et al. 2005a, b). In contrast,

ruddy turnstone (*Arenaria interpres*) and sanderling (*Calidris alba*) are similarly sized, have similar ecological needs, and overlap with red knot for much of their breeding and wintering ranges, but have not exhibited such a strong decline in numbers seen during spring stopover (Nettleship 2000, Baker et al. 2001, Macwhirter et al. 2002). Assessing the similarities and differences among these species in demographic and behavioral responses to stopover conditions may be instructive for understanding declines of Arctic-breeding shorebirds worldwide (Andres et al. 2012, Studds et al. 2017).

Here, I explore the use of stopover monitoring data for lending insights to the role of stopover in the context of the annual cycle, with a focus on migratory shorebirds in Delaware Bay. The next four chapters were prepared for submission to peer-reviewed journals and written in collaboration with co-authors; therefore, I use the third person plural throughout the remainder of this document. The second chapter was published in The Condor: Ornithological Applications in February 2019 and the third chapter is under review as written at Ecosphere as of March 2019. In Chapter 2, I investigate the potential for individual misidentification in resighting marked birds and explore the ways in which those errors could influence ecological inference about population dynamics. In Chapter 3, I examine interannual variation in arguably the most important part of migratory stopover: rapid mass gain due to deposition of fat. I explore the ways in which ecological mismatch with a key prey item influences mass gain dynamics, and how species traits are related to differences in their response to prey availability. In Chapter 4, I consider how the events and conditions encountered during stopover can influence both demographics and migratory behaviors. Disentangling these responses is important for understanding how populations respond to changes in stopover conditions, and how to interpret

changes in stopover passage population size. In Chapter 5, I take full annual cycle view of the role of stopover, and present an analysis to estimate both survival and recruitment of new breeding adults. This modeling framework facilitates addressing the question of whether stopover site conditions or breeding site conditions have a stronger effect on variation in population growth rate. Understanding the relative role of stopover is important for informing conservation strategies that will be most effective at halting declines and promoting population stability.

Literature Cited

Alerstam, T. (2011). Optimal bird migration revisited. Journal of Ornithology 152. doi: 10.1007/s10336-011-0694-1

Alerstam, T., A. Hedenstrom, and S. Akesson (2003). Long-distance migration: evolution and determinants. Oikos 103:247–260. doi: migration dispersion evolution adaptation strategie tactique contrainte comportement physiologie orientation

Andres, B. a., P. a. Smith, R. I. G. Morrison, C. L. Gratto-trevor, S. C. Brown, and C. a. Friis (2012). Population estimates of North American shorebirds, 2012. Wader Study Group Bulletin 119:178–194.

Baker, A., P. Gonzalez, R. I. G. Morrison, and B. A. Harrington (2001). Red Knot (Calidris canutus). In The Birds of North America. Cornell Lab of Ornithology, Ithaca. doi: 10.2173/bna.563

Baker, A. J., P. M. González, T. Piersma, L. J. Niles, I. D. L. S. do Nascimento, P. W. Atkinson, N. a Clark, C. D. T. Minton, M. K. Peck, and G. Aarts (2004). Rapid population decline in red knots: fitness consequences of decreased refuelling rates and late arrival in Delaware Bay. Proceedings. Biological sciences / The Royal Society 271:875–882. doi: 10.1098/rspb.2003.2663

Both, C., R. G. Bijlsma, and M. E. Visser (2005). Climatic effects on timing of spring migration and breeding in a long-distance migrant, the pied flycatcher Ficedula hypoleuca. Journal of Avian Biology 36:368–373. doi: 10.2307/3677791

Bridge, E. S., K. Thorup, M. S. Bowlin, P. B. Chilson, R. H. Diehl, R. W. Fléron, P. Hartl, R. Kays, J. F. Kelly, W. D. Robinson, and M. Wikelski (2011). Technology on the Move: Recent and Forthcoming Innovations for Tracking Migratory Birds. BioScience 61:689–698. doi: 10.1525/bio.2011.61.9.7

Callaghan, C. T., and D. E. Gawlik (2015). Efficacy of eBird data as an aid in conservation planning and monitoring. Journal of Field Ornithology 86:298–304. doi: 10.1111/jofo.12121

Cimprich, D., R. D. Sutter, D. N. Ewert, C. Duncan, M. Woodrey, B. Abel, D. W. Mehlman, and S. E. Mabey (2006). Conserving Stopover Sites for Forest-Dwelling Migratory Landbirds. The Auk 122:1281. doi: 10.1642/0004-8038(2005)122[1281:cssffm]2.0.co;2

Clark, E., L. J. Niles, and J. Burger (1993). Abundance and distribution of migrant shorebirds in Delaware Bay. The Condor 95:694–705.

Coates, P. S., B. G. Prochazka, M. A. Ricca, B. J. Halstead, M. L. Casazza, E. J. Blomberg, B. E. Brussee, L. Wiechman, J. Tebbenkamp, S. C. Gardner, and K. P. Reese (2018). The relative importance of intrinsic and extrinsic drivers to population growth vary among local populations of Greater Sage-Grouse: An integrated population modeling approach. The Auk 135:240–261. doi: 10.1642/AUK-17-137.1

Cohen, E. B., S. M. Pearson, and F. R. Moore (2014). Effects of landscape composition and configuration on migrating songbirds: Inference from an individual-based model. Ecological Applications 24:169–180. doi: 10.1890/12-1867.1

Cox, G. W. (1985). The Evolution of Avian Migration Systems Between Temperate and Tropical Regions of the New World. The American Naturalist 126:451–474.

Dhanjal-adams, K. L., M. Klaassen, S. Nicol, H. P. Possingham, and I. Chadès (2016). Setting conservation priorities for migratory networks under uncertainty. Conservation Biology:1–25. doi: 10.1111/cobi.12842.This

Dinsmore, S. J., and J. A. Collazo (2003). The influence of body condition on local apparent survival of spring migrant sanderlings in coastal North Carolina. Condor 105:465–473. doi: 10.1650/7226

Drever, M. C., and M. Hrachowitz (2017). Migration as flow: using hydrological concepts to estimate residence time of migrating birds from daily counts. Methods in Ecology and Evolution. doi: 10.1111/2041-210X.12727

Flockhart, D. T. T., J. B. Pichancourt, D. R. Norris, and T. G. Martin (2015). Unravelling the annual cycle in a migratory animal: Breeding-season habitat loss drives population declines of monarch butterflies. Journal of Animal Ecology 84:155–165. doi: 10.1111/1365-2656.12253

Forchhammer, M. C., E. Post, and N. C. Stenseth (2002). North Atlantic Oscillation timing of long- and short-distance migration. Journal of Animal Ecology 71:1002–1014. doi: 10.1046/j.1365-2656.2002.00664.x

Fraser, J. D., S. M. Karpanty, and J. B. Cohen (2010). Shorebirds Forage Disproportionately in Horseshoe Crab Nest Depressions. Waterbirds 33:96–100. doi: 10.1675/063.033.0111

Gillings, S., P. W. Atkinson, S. L. Bardsley, N. a. Clark, S. E. Love, R. a. Robinson, R. a. Stillman, and R. G. Weber (2007). Shorebird predation of horseshoe crab eggs in Delaware Bay: Species contrasts and availability constraints. Journal of Animal Ecology 76:503–514. doi: 10.1111/j.1365-2656.2007.01229.x

van Gils, J. A., S. R. De Rooij, J. Van Belle, J. Van Der Meer, A. Dekinga, T. Piersma, and R. Drent (2005a). Digestive bottleneck affects foraging decisions in red knots Calidris canutus. I. Prey choice. Journal of Animal Ecology 74:105–119. doi: 10.1111/j.1365-2656.2004.00903.x

van Gils, J. a, P. F. Battley, T. Piersma, and R. Drent (2005b). Reinterpretation of gizzard sizes of red knots world-wide emphasises overriding importance of prey quality at migratory stopover sites. Proceedings. Biological sciences / The Royal Society 272:2609–2618. doi: 10.1098/rspb.2005.3245

Gollan, J., L. L. De Bruyn, N. Reid, and L. Wilkie (2012). Can volunteers collect data that are comparable to professional scientists? a study of variables used in monitoring the outcomes of ecosystem rehabilitation. Environmental Management 50:969–978. doi: 10.1007/s00267-012-9924-4

Horton, K. G., B. M. Van Doren, F. A. La Sorte, E. B. Cohen, H. L. Clipp, J. J. Buler, D. Fink, J. F. Kelly, and A. Farnsworth (2019). Holding steady: Little change in intensity or timing of bird migration over the Gulf of Mexico. Global Change Biology 25:1106–1118. doi: 10.1111/gcb.14540

Hostetler, J. a., T. S. Sillett, and P. P. Marra (2015). Full-annual-cycle population models for migratory birds. The Auk 132:433–449. doi: 10.1642/AUK-14-211.1

Jonsson, K. I. (1997). Capital and Income Breeding as Alternative Tactics of Resource Use in Reproduction. Oikos 78:57–66.

Kéry, M., and M. Schaub (2012). Bayesian population analysis using WinBUGS: A hierarchical perspective. In. Academic Press, Waltham, Massachusetts.

Kirby, J. S., A. J. Stattersfield, S. H. M. Butchart, M. I. Evans, R. F. A. Grimmett, V. R. Jones, J. O'Sullivan, G. M. Tucker, and I. Newton (2008). Key conservation issues for migratory land- and waterbird species on the world's major flyways. Bird Conserv Int 18:S49--S73. doi: 10.1017/S0959270908000439

Kokko, H. (1999). Competition for early arrival birds in migratory birds. Journal of Animal Ecology 68:940–950. doi: 10.1046/j.1365-2656.1999.00343.x

Kolzsch, A., G. J. D. M. Muskens, H. Kruckenberg, P. Glazov, R. Weinzierl, B. A. Nolet, and M. Wikelski (2016). Towards a new understanding of migration timing: slower spring than autumn migration in geese reflects different decision rules for stopover use and departure. Oikos 125:1496–1507. doi: 10.1111/oik.03121

Lyons, J. E., W. L. Kendall, J. A. Royle, S. J. Converse, B. A. Andres, and J. B. Buchanan (2015). Population size and stopover duration estimation using mark-resight data and Bayesian analysis of a superpopulation model. Biometrics. doi: 10.1111/biom.12393

Macwhirter, R. B., P. A.-S. Jr., and D. E. Kroodsma (2002). Sanderling (Calidris alba). In The Birds of North America (A. F. Poole and F. B. Gill, Editors). 2nd edition. Cornell Lab of Ornithology, Ithaca, NY.

Marra, P. P., E. B. Cohen, S. R. Loss, J. E. Rutter, and C. M. Tonra (2015). A call for full annual cycle research in animal ecology. Biology letters 11:2015.0552. doi: 10.1098/rsbl.2015.0552

Martin, T., I. Chadès, P. Arcese, P. P. Marra, H. P. Possingham, and D. R. Norris (2007). Optimal conservation of migratory species. PLoS One 2:e751. doi: 10.1371/journal.pone.0000751

Masero, J. A., F. Santiago-Quesada, J. M. Sanchez-Guzman, A. Villegas, J. M. Abad-Gomez, R. J. Lopes, V. Encarnacao, C. Corbacho, and R. Morana (2011). Long lengths of stay, large numbers, and trends of the Black-tailed Godwit Limosa limosa in rice fields during spring migration. Bird Conservation … 21:12–24. doi: 10.1017/S0959270910000092

McCabe, J. D., and B. J. Olsen (2015). Landscape-scale habitat availability, and not local geography, predicts migratory landbird stopover across the Gulf of Maine. Journal of Avian Biology 46:395–405. doi: 10.1111/jav.00598

McGowan, C. P., J. E. Hines, J. D. Nichols, J. E. Lyons, D. R. Smith, K. S. Kalasz, L. J. Niles, A. D. Dey, N. a. Clark, P. W. Atkinson, C. D. T. Minton, and W. Kendall (2011a). Demographic consequences of migratory stopover: linking red knot survival to horseshoe crab spawning abundance. Ecosphere 2. doi: 10.1890/ES11-00106.1

McGowan, C. P., D. R. Smith, J. D. Nichols, J. E. Lyons, J. Sweka, K. Kalasz, L. J. Niles, R. Wong, J. Brust, M. Davis, and B. Spear (2015). Implementation of a framework for multi-species, multi-objective adaptive management in Delaware Bay. Biological Conservation 191:759–769. doi: 10.1016/j.biocon.2015.08.038

McGowan, C. P., D. R. Smith, J. A. Sweka, J. Martin, J. D. Nichols, R. Wong, J. E. Lyons, L. J. Niles, K. Kalasz, and J. Brust (2011b). Multispecies modeling for adaptive management of horseshoe crabs and red knots in the Delaware Bay. Natural Resource Modeling 24:117–156.

Miller-Rushing, A. J., T. T. Høye, D. W. Inouye, and E. Post (2010). The effects of phenological mismatches on demography. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences 365:3177–86. doi: 10.1098/rstb.2010.0148

Miller, D., K. Pacifici, J. S. Sanderlin, and B. Reich (2019). The recent past and promising future for data integration methods to estimate species' distributions. Methods in Ecology and Evolution 2019:22–37. doi: 10.1111/2041-210X.13110

Milner-Gulland, E. J., J. M. Fryxell, and A. R. E. Sinclair (Editors) (2011). Animal Migration: A synthesis. In. Oxford University Press, New York, New York, USA.

Morrison, C. A., R. A. Robinson, S. J. Butler, J. A. Clark, and J. A. Gill (2016). Demographic drivers of decline and recovery in an Afro-Palaearctic migratory bird population. Proceedings of the Royal Society B: Biological Sciences. doi: 10.1098/rspb.2016.1387

Murray, N. J., P. P. Marra, R. A. Fuller, R. S. Clemens, K. Dhanjal-Adams, K. B. Gosbell, C. J. Hassell, T. Iwamura, D. Melville, C. D. T. Minton, A. C. Riegen, et al. (2017). The large-scale drivers of population declines in a long-distance migratory shorebird. Ecography:867–876. doi: 10.1111/ecog.02957

Navedo, J. G., G. Orizaola, J. A. Masero, O. Overdijk, and J. M. Sanchez-Guzman (2010). Long-distance travellers stopover for longer: A case study with spoonbills staying in North Iberia. Journal of Ornithology 151:915–921. doi: 10.1007/s10336-010-0530-z

Nettleship, D. N. (2000). Ruddy Turnstone (Arenaria interpres). In The Birds of North America. Cornell Lab of Ornithology, Ithaca. doi: 10.2173/bna.537

Newton, I. (2004). Population limitation in migrants. Ibis 146:197–226. doi: 10.1111/j.1474-919X.2004.00293.x

Newton, I. (2006). Can conditions experienced during migration limit the population levels of birds? Journal of Ornithology 147:146–166. doi: 10.1007/s10336-006-0058-4

Niles, L. J., J. Bart, H. P. Sitters, A. D. Dey, K. E. Clark, P. W. Atkinson, A. J. Baker, K. a. Bennett, K. S. Kalasz, N. a. Clark, J. Clark, et al. (2009). Effects of Horseshoe Crab Harvest in Delaware Bay on Red Knots: Are Harvest Restrictions Working? BioScience 59:153–164. doi: 10.1525/bio.2009.59.2.8

Niles, L. J., H. P. Sitters, A. D. Dey, P. W. Atkinson, A. J. Baker, K. A. Bennett, R. Carmona, K. E. Clark, N. A. Clark, C. Espoz, P. M. Gonzalez, et al. (2008). Status of the red knot (Calidris canutus rufa) in the Western Hemisphere. Studies in Avian Biology.

Norris, D. R., and P. P. Marra (2007). Seasonal interactions, habitat quality, and population dynamics in migratory birds. The Condor 109:535–547.

Paxton, K. L., M. Yau, F. R. Moore, and D. E. Irwin (2013). Differential migratory timing of western populations of Wilson's Warbler (Cardellina pusilla) revealed by mitochondrial DNA and stable isotopes. The Auk 130:689–698. doi: 10.1525/auk.2013.13107

Pearce-Higgins, J. W., S. M. Eglington, B. Martay, and D. E. Chamberlain (2015). Drivers of climate change impacts on bird communities. Journal of Animal Ecology:n/a-n/a. doi: 10.1111/1365-2656.12364

Pekarsky, S., A. Angert, B. Haese, M. Werner, K. a. Hobson, and R. Nathan (2015). Enriching the isotopic toolbox for migratory connectivity analysis: a new approach for migratory species breeding in remote or unexplored areas. Diversity and Distributions:n/a-n/a. doi: 10.1111/ddi.12306

Robinson, R. A., H. Q. P. Crick, J. A. Learmonth, I. M. D. Maclean, C. D. Thomas, F. Bairlein, M. C. Forchhammer, C. M. Francis, J. A. Gill, B. J. Godley, J. Harwood, et al. (2009). Travelling through a warming world: Climate change and migratory species. Endangered Species Research 7:87–99. doi: 10.3354/esr00095

Rushing, C. S., J. A. Hostetler, T. S. Sillett, P. P. Marra, J. A. Rotenberg, and T. B. Ryder (2017). Spatial and temporal drivers of avian population dynamics across the annual cycle. Ecology 98:2837–2850. doi: 10.1002/ecy.1967

Rushing, C. S., T. B. Ryder, J. F. Saracco, and P. P. Marra (2014). Assessing migratory connectivity for a long-distance migratory bird using multiple intrinsic markers. Ecological Applications 24:445–456. doi: 10.1890/13-1091.1

Salewski, V., and M. Schaub (2007). Stopover duration of Palearctic passerine migrants in the western Sahara - Independent of fat stores? Ibis 149:223–236. doi: 10.1111/j.1474-919X.2006.00608.x

Sawyer, H., M. J. Kauffman, R. M. Nielson, and J. S. Horne (2009). Identifying and Prioritizing Ungulate Migration Routes for Landscape-Level Conservation. Ecological Applications 19:2016–2025. doi: 10.1890/08-2034.1

Shaw, A. K. (2016). Drivers of animal migration and implications in changing environments. Evolutionary Ecology. doi: 10.1007/s10682-016-9860-5

Sheehy, J., C. M. Taylor, and D. R. Norris (2011). The importance of stopover habitat for developing effective conservation strategies for migratory animals. Journal of Ornithology 152:161–168. doi: 10.1007/s10336-011-0682-5

Silvy, N. J., R. R. Lopez, and M. J. Peterson (2012). Techniques for marking wildlife. The Wildlife Techniques Manual Volume 1: Research:230–257.

Skagen, S. K. (2006). Migration Stopovers and the Conservation of Arctic-Breeding Calidridine Sandpipers. The Auk 123:313–322. doi: 10.1642/0004-8038(2006)123[313:MSATCO]2.0.CO;2

Smith, D. R., H. J. Brockmann, M. A. Beekey, T. L. King, M. J. Millard, and J. Zaldívar-Rae (2016). Conservation status of the American horseshoe crab, (Limulus polyphemus): a regional assessment. Reviews in Fish Biology and Fisheries. doi: 10.1007/s11160-016-9461-y

Smith, D. R., L. J. Brousseau, M. T. Mandt, and M. J. Millard (2010). Age and sex specific timing, frequency, and spatial distribution of horseshoe crab spawning in Delaware Bay: Insights from a large-scale radio telemetry array. Current Zoology 56:563–574.

Smith, D. R., C. P. McGowan, J. P. Daily, J. D. Nichols, J. a. Sweka, and J. E. Lyons (2013). Evaluating a multispecies adaptive management framework: must uncertainty impede effective decision-making? Journal of Applied Ecology 50:1431–1440. doi: 10.1111/1365-2664.12145

Smith, D. R., S. F. Michels, R. G. Weber, D. B. Carter, P. S. Pooler, R. E. Loveland, and M. L. Botton (2002). Horseshoe crab (Limulus polyphemus) reproductive activity on Delaware Bay beaches: Interactions with beach characteristics. Journal of Coastal Research 18:730–740.

Smith, D. R., M. J. Millard, and R. H. Carmichael (2009). Comparative Status and Assessment of Limulus polyphemus with Emphasis on the New England and Delaware Bay Populations. In Biology and Conservation of Horseshoe Crabs. pp. 361–386. doi: 10.1007/978-0-387-89959-6

Smith, D. R., and T. J. Robinson (2015). Horseshoe Crab Spawning Activity in Delaware Bay, USA, After Harvest Reduction: A Mixed-Model Analysis. Estuaries and Coasts 38:2345–2354. doi: 10.1007/s12237-015-9961-3

Smith, P. A., C. L. Gratto-Trevor, B. T. Collins, S. D. Fellows, R. B. Lanctot, J. Liebezeit, B. J. McCaffery, D. Tracy, J. Rausch, S. Kendall, S. Zack, and H. R. Gates (2012). Trends in Abundance of Semipalmated Sandpipers: Evidence from the Arctic. Waterbirds 35:106–119. doi: 10.1675/063.035.0111

Stephens, P. A., I. L. Boyd, J. M. McNamara, and A. I. Houston (2009). Capital breeding and income breeding: their meaning, measurement, and worth. Ecology 90:2057–2067.

Studds, C. E., B. E. Kendall, N. J. Murray, H. B. Wilson, D. I. Rogers, R. S. Clemens, K. Gosbell, C. J. Hassell, R. Jessop, D. S. Melville, D. A. Milton, et al. (2017). Rapid population decline in migratory shorebirds relying on Yellow Sea tidal mudflats as stopover sites. Nature Communications 8:1–7. doi: 10.1038/ncomms14895

Stutzman, R. J., and J. J. Fontaine (2015). Shorebird Migration in the Face of Climate Change. Phenological synchrony and bird migration: changing climate and seasonal resources in North America. Studies in Avian Biology.:145–159.

Taylor, C. M., and B. J. M. Stutchbury (2016). Effects of breeding versus winter habitat loss and fragmentation on the population dynamics of a migratory songbird. Ecological Applications 26:424–437. doi: 10.1890/14-1410

Tsipoura, N., and J. Burger (1999). Shorebird Diet During Spring Migration Stopover on Delaware Bay. The Condor 101:635–644.

Tucker, A. M., C. P. McGowan, R. A. Robinson, J. A. Clark, J. E. Lyons, A. Derose-wilson, R. Feu, G. E. Austin, P. W. Atkinson, and N. A. Clark (2019). Effects of individual misidentification on estimates of survival in long-term mark – resight studies. The Condor: Ornithological Applications 121:1–13. doi: 10.1093/condor/duy017

Warnock, N. (2010). Stopping vs. staging: The difference between a hop and a jump. Journal of Avian Biology 41:621–626. doi: 10.1111/j.1600-048X.2010.05155.x

Webster, M. S., P. P. Marra, S. M. Haig, S. Bensch, and R. T. Holmes (2002). Links between worlds: Unraveling migratory connectivity. Trends in Ecology and Evolution 17:76–83. doi: 10.1016/S0169-5347(01)02380-1

Wilcove, D. S., and M. Wikelski (2008). Going, going, gone: Is animal migration disappearing? PLoS Biology 6:1361–1364. doi: 10.1371/journal.pbio.0060188

Williams, B., J. Nichols, and M. Conroy (2002). Analysis and Management of Wildlife Populations. In. 1st edition. Academic Press, San Diego, California.

Wilmers, C. C., B. Nickel, C. M. Bryce, J. A. Smith, R. E. Wheat, and V. Yovovich (2015). The golden age of bio-logging how animal-borne sensors are advancing the frontiers of ecology. Ecology 96:1741–1753.

Winger, B. M., G. G. Auteri, T. M. Pegan, and B. C. Weeks (2018). A long winter for the Red Queen: rethinking the evolution of seasonal migration. Biological Reviews. doi: 10.1111/brv.12476

Woodworth, B. K., C. M. Francis, and P. D. Taylor (2014). Inland flights of young red-eyed vireos Vireo olivaceus in relation to survival and habitat in a coastal stopover landscape. Journal of Avian Biology 45:387–395. doi: 10.1111/jav.00276

Woodworth, B. K., N. T. Wheelwright, A. E. Newman, M. Schaub, and D. R. Norris (2017). Winter temperatures limit population growth rate of a migratory songbird. Nature Communications 8:1–9. doi: 10.1038/ncomms14812

Ydenberg, R. C., R. W. Butler, D. B. Lank, B. D. Smith, and J. Ireland (2004). Western sandpipers have altered migration tactics as peregrine falcon populations have recovered. Proceedings of the Royal Society B: Biological Sciences 271:1263–1269. doi: 10.1098/rspb.2004.2713

Zink, R. M., and A. S. Gardner (2017). Glaciation as a migratory switch. Science Advances:1–9. doi: 10.1126/sciadv.1603133

Chapter 2: Effects of individual misidentification on estimates of survival in long-term mark-resight studies

Abstract

All ecological measurements are subject to error; the effects of missed detection (false negatives) are well known, but the effects of mistaken detection (false positives) are less understood. Long term capture-recapture datasets provide valuable ecological insights and baselines for conservation and management, but where such studies rely on noninvasive re-encounters, such as field readable color bands, there is the potential to accumulate detection errors as the length of the study and number of tags deployed increases. We investigated the prevalence and effects of misreads in a 10-year dataset of Red Knots (Calidris canutus rufa) marked with field-readable leg flags in Delaware, USA. We quantified the effects of misreads on survival estimation via a simulation study and evaluated whether removal of individuals only reported once in a year (potential misreads) influenced survival estimation from both simulated datasets and our case study data. We found overall apparent error rates of 0.31% (minimum) to 6.6% (maximum). Observer-specific error rates and the variation among observers both decreased with the number of flags an observer recorded. Our simulation study showed that misreads lead to spurious negative trends in survival over time, particularly for long-term studies. Removing all records in which a flag was only recorded once in a sampling occasion reduced bias and eliminated spurious negative trends in survival but also reduced precision in survival estimates. Without data filtering, we found a slight decrease in Red Knot annual survival probability from 2008-2018 ($\beta = -0.043 \pm 0.03$), but removing all single-observation records

resulted in no apparent trend (β = -0.0074, ± 0.02). Spurious trends in demographic rates could influence inference about population trajectories and resultant conservation decision-making. Data filtering could eliminate errors, but researchers should carefully consider the trade-off between precision obtained by larger sample sizes and potential bias due to misreads in their data.

Introduction

Observation error is a well-studied problem in ecology. While the effect of missed detections (false negatives) is largely understood, the effects of incorrect detections (false positives) are more complex. False positives can occur in different ways depending on the data type and sampling methods, for example via species misidentification in occupancy surveys (Miller et al. 2011, Yu et al. 2014) or via genotyping error in noninvasive genetic sampling (Wright et al. 2009, Yoshizaki et al. 2009). The resulting effects on estimation vary among different types of errors. This issue is central to capture-recapture analyses, which rely on a core assumption that individuals are correctly identified when encountered and that there are no false positive detections (Lebreton et al. 1992, Williams et al. 2002). Many types of individual marks can be "encountered" without physical recapture, facilitating data collection on large spatial and temporal scales, including both artificial marks (e.g., dye marks, color bands, leg flags, patagial or ear tags (Silvy et al. 2012)) and natural marks (e.g., skin or pelage patterns, permanent scars (Beck et al. 2004) or genetic markers (Yoshizaki 2011)). However, with noninvasive encounters, also termed mark-resight studies, the risk of misidentifying individuals is greater than with physical recaptures because there is limited opportunity to confirm individual ID upon encounter.

Long-term datasets are valuable to ecology and conservation biology, and often rely on a large number of observers with varying experience and who spend varying lengths of time with the project (Bildstein 1998, Newman et al. 2003, Cohn 2008, Magurran et al. 2010, Conrad and Hilchey 2011, Tulloch et al. 2013). Both observer training and data quality control protocols are necessary to ensure accurate data collection. In some cases it may be relatively straightforward to remove impossible observations from the dataset before analysis; however, as both the study scale (spatial or temporal) and the number of deployed tags available for observation increases, initial filtering of the data becomes more difficult. Additionally, even if error rate is low and constant over time, the absolute number of errors may accumulate as the length of the study increases. These concerns apply particularly to data collected via citizen science, volunteer-based surveys, or by seasonal interns and field technicians, but even skilled professionals can make errors. Regardless of data collection protocol, as the number of observers and length of study increases, the potential for errors to appear in the data set also increases. Determining the observer-specific attributes that are associated with misreads can inform training protocols and aid in vetting data collected by observers with varying experience or training.

The potential for misidentifications to occur in mark-resight studies is understood by field researchers, and many marking protocols are designed to reduce the probability of misidentifications, such as excluding easily-confused letters from alphanumeric codes (Clark et al. 2005), or avoiding deployment of similar codes on individuals with nearby territories or home ranges. Several researchers have also estimated error rates for reading color band combinations and alphanumeric codes using both experimental methods and double-marking studies (Weiss et al. 1991, Burton 2000, Milligan et al. 2003, Lavers and Jones 2008, Mitchell and Trinder 2008, Roche et al. 2014), resulting inestimated error rates from 1-16%. These errors could bias

demographic estimates if not accounted for during either data processing or analysis (Schwarz and Stobo 1999, Bearhop et al. 2003, Morrison et al. 2011).

Some modeling approaches have been developed to explicitly account for individual misidentification during estimation, but most of these models were developed specifically for data collected via noninvasive genetic sampling or photographs (Link et al. 2010b, Morrison et al. 2011, Yoshizaki et al. 2011). In this context, misidentification results in the first detection of a new "ghost" individual that can only appear once in the data because it is not known to the researchers whether individuals are correctly identified upon first encounter (for example, the first photograph of a new individual or first sample of a unique genotype). However, with mark-resight studies it is almost certain that marks are perfectly identified upon first capture but that subsequent resightings may occur with error. Additionally, reports of non-existent marks are often easily identified and removed from the database, but false detection of existing marks may occur multiple times after the initial capture. We note, however, that in migratory stopover systems individuals marked elsewhere in the flyway are often encountered and it may not be known to all researchers which marks have and have not been deployed. When misidentification upon first encounter leads to ghosts, resulting survival estimates are negatively biased by as much as 25% (Morrison et al. 2011). If misidentification errors result in the false detection of a real individual that is no longer alive, however, annual survival estimates from earlier in the time series could be inflated, leading to an apparent negative trend in survival over time. The effects of these errors—incorrect observations of marks that are valid but may not be truly present because the individual has since died or permanently emigrated—has not been as thoroughly explored in the existing literature.

To address this type of false positive error, we evaluated the effect of misreads on analysis of long-term mark-resight datasets using a simulation study. Many monitoring programs that use field readable marks involve systematic searching for marked individuals throughout a defined sampling period, which often results in multiple observations of the same individuals within a sampling occasion. Multiple observations allow for more opportunities for misreads to occur and for the ability to confirm presence of individuals reported more than once. We propose a simple data filtering protocol that removes individuals recorded only once in a sampling period, and we evaluate the accuracy and precision of models that estimate survival using a both the raw and filtered dataset under varying levels of misread error.

We evaluated misread errors in the context of a long-term monitoring program for Red Knot (Calidris canutus rufa) during migratory stopover. The objectives of this study were first to estimate flag reading error rate in our dataset and determine whether observer experience was associated with misread errors. Second, we evaluated the effect of those misreads on estimation of apparent annual survival probability using a simulation study. Last, we evaluated the effect of our data filtering protocol, which removes potential misreads, on survival estimates by comparing estimates from both simulated data and Red Knot mark-resight data with and without data filtering.

Methods

Flag Deployment and Resighting

We estimated error rate in a mark-resight observations of Red Knot in Delaware, USA from 2008-2016. The Delaware Shorebird Project is a long-term volunteer-based research program designed to monitor the population status of migratory shorebirds that use Delaware Bay during spring stopover. Red Knot are long-distance migratory shorebirds that stop in

Delaware Bay *en route* to their Arctic breeding grounds each year (Baker et al. 2001). The mark recapture monitoring program in Delaware Bay has been an important component of numerous research and management studies since the 1990s (e.g., Atkinson et al. 2007, Gillings et al. 2009, McGowan et al. 2015).

Throughout the study, shorebirds were captured in mixed-species foraging flocks using cannon nets. At the time of capture, we collected biometric data for each individual and applied a USGS band along with an individually identifiable field-readable flag (Figure 1). Inscribed plastic leg flags were first deployed in Delaware Bay in 2004 and have been used thereafter. Flags used in our study were made of lime green Darvic PVC and were laser engraved with a unique three-character alphanumeric code, filled with black acrylic paint ('RAL 9005 Avkote KS Satin' from AVCO Ltd [www.avko.co.uk]) and varnished. Leg flag design and manufacturing are described in detail by Clark *et al.* (2005). With the intent of estimating future rates of misidentification in this system, before the 2008 field season 20% (280 individual flags) of the flags manufactured for that year were haphazardly selected and withheld from circulation.

During a three-week monitoring season in May each year (typically May 10 – June 1), trained observers visit beaches in Delaware to count the numbers of shorebirds and scan flocks with spotting scopes to record observations of Red Knot with leg flags. Most observers are volunteers who have widely varying backgrounds and experience with biological fieldwork in general and flag resighting specifically. Flag resighting occurs throughout the day in 30-minute time increments that are considered independent observation occasions, allowing for repeated detections of individuals throughout the day. After returning from the field, observers transcribe their resightings to a data sheet, which are then entered into a Microsoft Access database. After entry, resight records are printed from the database and compared to the transcribed data sheets

to ensure accurate data entry. These protocols minimize transcription errors from field data to digital data entry, but do not protect from transcription errors from field notebooks to data sheets, and do not eliminate observation errors that occur in the field.

Quantifying Error Rate

As a subset of alphanumeric codes were randomly withdrawn from use in 2008, for all analyses described below, we only used resighting data from 2009-2018 to estimate misread rates. We defined misidentification rate in two ways to establish a possible range of error rates. First, we determined the proportion of flag resightings from 2009-2018 that were either one of the withheld combinations or had not yet been deployed (i.e. known false detections). For this species and many other long-distance migrants, marks are deployed by researchers throughout the range and therefore the deployment date was not known for all flags resighted. Of the 8,135 unique flags resighted by observers on our project, 4,558 were deployed by our project and therefore had a known deployment date. The proportion of known false detections in the dataset served as our minimum error rate since we knew that reports of these flag codes were errors. To establish a maximum probable error rate we calculated the proportion of flags recorded by Delaware Shorebird Project volunteers that were only recorded once in a season (i.e. single-observation events). Because flags are frequently observed more than once in a season, but it is unlikely that the same incorrect flag is recorded more than once, these single-observation records were considered as possible misreads. This maximum probable error rate will likely be an overestimate since some individuals may have been transient during migration and only available to observe once in that year. However our focus in this study was annual survival probability and not within-year stopover dynamics. We calculated confirmed (impossible resightings) and

possible (single-observation) error rates at the population level as the proportion of all resightings in our resighting database and at the observer level as the proportion of each observer's total resightings.

Observer-specific Misread Rates

Identifying flag- and observer-specific attributes associated with flag misreads can inform data filtering and observer training to minimize errors, therefore, we were interested in observer-specific factors that might be associated with a greater proportion of potential misreads. We modeled observer-specific misread rates as an additive effect of a) number of years they had spent on the project and b) the total number of resightings they have contributed to the database across the whole study. Although those metrics are correlated within observers ($\rho = 0.63$), there is considerable variation as some volunteers participate for many years, but only for one or two days (many years, few resightings), while others might be new but join the project for the entire season (few years, many resightings). We used a beta-binomial generalized linear model to analyze apparent misread rates while accounting for overdispersion and unequal variance in our data. Overdispersion and unequal variance in the data are likely caused by observers coming to the project with widely varying previous background and training. Among-observer variance in misread rate decreased with observer experience on the project (Figure 2), so we modeled the variance in the probability of misreads (represented by the overdispersion parameter θ) as a function of the log-transformed total number of resightings. Beta-binomial models were specified in R and fit via maximum likelihood estimation using the bbmle package (R Core Team 2016, Bolker and R Development Core Team 2017).

We used the beta-binomial model to estimate the logit-linear relationship between observer-specific misread rate (*m*) and the logged total number of resightings each observer contributed to the project.

$$logit(m) = intercept + \beta * \log(total\ resightings)$$

Based on this relationship, we were interested in determining the number of resightings after which we would expect an observer's misread rate to be equal to the median rate among all observers, expecting observers with fewer observations than this threshold to have an above-average misread rate. We used median error rate instead of mean because distributions of observer-specific error rates were right-skewed, and therefore the median is a better metric of central tendency. We derived this metric as a data screening tool to define observer experience with the project. We calculated this threshold by setting *m* equal to the median error rate and rearranging the above equation to solve for the total number of resightings.

$$threshold\ resightings = e^{\left(\frac{logit(m) - intercept}{\beta}\right)}$$

Observers below this threshold were considered "inexperienced" and those above this threshold were considered "experienced" for the purposes of data filtering described below.

Effect of Flag Misreads on Survival Estimation

We used a simulation study to assess the effect of false detections arising from flag misreads on estimates of annual apparent survival probability. Capture histories were randomly generated to represent encounters of individual birds over a 5-year, 10-year, or 20-year study. We assumed all detections after first capture were field resightings and not physical captures, and that 250 individuals were newly marked each year. We generated capture histories with annual survival probability of ϕ = 0.8 and detection probability of p = 0.5, which are broadly

representative of our study system (McGowan et al. 2011, Méndez et al. 2018), by simulating a series of Bernoulli trials for survival and detection for each individual in each year following Kéry and Schaub (2012). To represent the Red Knot flag observation process in Delaware Bay, we allowed each individual to be detected multiple times in a given year. The number of times an individual was detected in a given year was randomly drawn from a Poisson distribution with a sampling intensity of $\lambda = 0.7$, which corresponds to a probability of being detected at least once in that year of 0.5 ($p = 1 - e^{-\lambda}$). Misidentification was a Binomial process with the number of trials equal to the number of detections of each individual in each year with probability of misidentification of 0, 0.005, 0.01, 0.05, or 0.1. Allowing for multiple detections in a year meant that misidentifications rarely resulted in non-detections of the real individual, which typically only occurred if an individual was only seen once and misidentified on that occasion.

There are two types of false positive errors that could result from flag misreads. The first type of false positive occurs when one flag is incorrectly recorded as another flag that exists in the database and the second occurs when a flag is incorrectly recorded as a flag that does not exist. For our simulations, we assumed that the second type of error would be scrubbed from the database before analysis and only considered the first type of error. We introduced errors in our simulated capture histories by randomly changing a subset of detections to non-detections and re-assigning them to individuals that had been marked prior to the year of reassignment but were not otherwise detected in that year. Therefore only real individuals had the opportunity to be reported as a result of a misread and additional fake individuals (i.e. marks that have not been deployed) cannot appear in the data. This approach results in false positive detections of both individuals that were alive but not detected and those that had died or emigrated from the study area. Errors were introduced at one of five error rates: 0, 0.005, 0.01, 0.05, or 0.1. These rates

were chosen to represent a range from low error rate (0.5%) to very high error rate (10%), and correspond with observed minimum and maximum apparent misread rates in our dataset. We simulated 1000 data sets under each of 15 scenarios (three possible study lengths, five possible error rates). We estimated apparent annual survival for each simulated data set using a Cormack-Jolly-Seber (CJS) model implemented in Program MARK using the RMark package for R (Laake et al. 2013, R Core Team 2016). For each dataset, we fit a model that estimated a linear trend in survival probability (ϕ) over time and a time-constant detection probability (p). We quantified the effects of misreads on resulting estimates by calculating the root mean squared error (RMSE) and relative bias (%) between model estimates and true data-generating values of ϕ and p, and by comparing estimated slope parameters of the trend in survival probability over time. RMSE was estimated as:

$$RMSE(\hat{\phi}) = \sqrt{\frac{\sum_{i=1}^{n}(\hat{\phi}_i - \phi)^2}{n-1}}$$

Where $\hat{\phi}_i$ is the survival estimate from a single replicate, ϕ is the true survival probability, and *n* is the number of replicates. Relative bias was estimated as:

$$Bias(\hat{\phi}) = \frac{\sum_{i=1}^{n}(\hat{\phi}_i - \phi)/\phi}{n}$$

Data Filtering to Minimize Errors

We evaluated the effect of four proposed data filtering methods on estimates of Red Knot apparent annual survival and estimates from simulated datasets. We estimated apparent annual survival for Red Knot first captured from 2008-2017 and resighted from 2009-2018 in Delaware, USA. All first encounters were physical captures during which the flag was deployed, and all

subsequent encounters were field resightings. We first filtered the data to remove all resightings of flags not deployed by researchers in Delaware, which removed all impossible records of withheld flags and those that were observed before their deployment date. Second, we removed both impossible flags and all records from inexperienced observers. We determined observers to be "inexperienced" based on the threshold method described above. Third, we removed both impossible flags and all observations of flags that were only observed once in a given year. Lastly, we removed all three types of known or potential errors.

We evaluated the fit of the Red Knot data to the fully time-dependent CJS model using the R2ucare package for R (Gimenez et al. 2018), which indicated potential transience and trap response. To account for these, we included two time-varying individual covariates. The first was a dummy variable that indicated whether each individual was detected in the previous year (0 = not seen, 1 = seen). This was included as a covariate on detection probability to account for a type of trap-response likely resulting from nonrandom temporary emigration (individuals skipping stopover in Delaware in some years). We accounted for transience by assigning each individual to an age class (first capture = 1, all subsequent captures = 2) and estimating the effect of this dummy age class on survival probability. This method of estimating survival after first capture separately from survival after subsequent encounters has been used to account for the presence of transients (individuals only present for one year) in capture-recapture models (Pradel et al. 1997).

We used RMark to fit a CJS model that estimated a linear trend and the effect of the dummy age class on annual apparent survival probability. Detection probability was modeled as a fixed effect of year and whether the individual was seen in the previous year. We compared the estimated slope of survival over time among the four data filtering methods to determine whether

they produced different ecological inferences. To provide some context for interpreting these results and more objectively evaluate the effects of removing observations from the data, we also fit the model to simulated datasets with and without data filtering. For each simulated dataset, we fit the CJS model both without data filtering and after removing all instances in which a flag was only detected once in a given year. We compared estimated trends in survival over time among scenarios and calculated RMSE and relative bias of estimates with and without removing single-observations.

Results

Quantifying Error Rate

Our flag resighting dataset contained 80,880 total recorded observations by 201 observers of 8,135 individual Red Knot from 2009-2018 (Table 1). The number of observers in a given year ranged from 36 to 53, with an average of 42 observers per year. The intensive survey effort results in many flags being observed more than once each year, with the average flag seen 5 times in total by 3 different observers in a year and at maximum seen 58 times by 23 different observers. There were 136 reports of withheld flags (one of the 280 flags removed from circulation in 2008) by 46 different observers and 116 reports of flags in a year before they were deployed by 37 different observers, giving a total of 252 impossible observations (0.31%). There were 5,374 observations that occurred only once in a given year in the database (6.6%). Therefore the range of potential misread errors in our resighting data was 0.31% at minimum and 6.6% at maximum.

Observer-specific Misread Rates

Distributions of observer-specific misread rates were all right-skewed, so we report among-observer medians and interquartile range (IQR). Withheld and not-yet-deployed flags were observed rarely. The median observer-specific rate of recording impossible flags was 0% (IQR: 0%, 0.24%), but the observer mean was 0.62%. The median observer-specific rate of single-observations was 7.4% of all observations (IQR: 4.2%, 12%).

The total number of resightings made by an observer was a significant predictor of both minimum and maximum possible misread rate (Figure 2). Both the probability of misreads and the among-observer variation in misread probability decreased as the number of total resightings logged increased ($\beta^{total}_{minimum}$ = -0.21 ± 0.08, $\beta^{total}_{maximum}$ = -0.19 ± 0.04), but the number of years an observer worked on the project was not a significant predictor of misread rates ($\beta^{years}_{minimum}$ = 0.039 ± 0.04, $\beta^{years}_{maximum}$ = 0.036 ± 0.02). Among-observer variation in rates of single-observation resightings decreased as the total number of resightings increased, as indicated by a positive relationship between the overdispersion parameter θ and the total number of resightings logged ($\alpha = 1.9 \pm 0.53$).

Using the predicted relationship between number of observations and apparent misread rate, we calculated the number of observations after which observers converged on the median error rate. Predicted maximum error rates for observers in our study converged on the median after 307 resightings. The median maximum error rate calculated only from experienced observers above this threshold was 6.4% (IQR: 4.3%, 7.6%) possible misreads per observation, while inexperienced observers below this threshold had a median single-observation rate of 8.8% (IQR: 3.1%, 15.3%).

Effect of Flag Misreads on Survival Estimation

Introducing random flag misreads into our simulated capture histories resulted in apparent negative trends in survival probabilities over time, particularly when the error rate was high (≥ 0.05, Figure 3). The effects of misreads were most pronounced in longer time series, with the RMSE of survival probability estimates for a 20-year study ranging from 0.007-0.066 and relative bias 0.5%-7.6%. RMSE and relative bias for all scenarios are listed in Appendix A Table 1. With low misread rates (0, 0.005, 0.01), precision of model estimates increased with study length, however for the higher error rates (0.05, 0.1), longer study lengths resulted in decreasing accuracy and precision of model estimates as errors had more chance to accumulate (Figure 4).

Data Filtering to Minimize Errors

We analyzed capture histories of 2,594 individual Red Knot from 2008-2018. Of the 25,226 total recorded resightings of these individuals, 2,850 were from inexperienced observers (defined for our study as observers with fewer than 300 total resightings) and 941 were single-observations. When only impossible resightings were removed, we found evidence of a negative trend in apparent annual survival probability over time (β = -0.043, CI: -0.094, 0.0081; Figure 5). Removing records from inexperienced observers had little effect on estimated trend (β = -0.034, CI: -0.084, 0.015), but removing single-observations resulted in no evidence of a trend in annual survival probability over the past 10 years (β = -0.0074, CI: -0.047, 0.032). Removing both inexperienced observers and single-observations had a similar effect as removing single-observations alone (β = -0.0059, CI: -0.046, 0.034). The differences in estimated survival probability with and without removal of single-observation records is most pronounced when comparing estimates for the first and last years included in this analysis. Without data filtering, estimated survival probabilities apparently declined from ϕ = 0.87 (95% CI: 0.84, 0.89) in 2008

to ϕ = 0.82 (95% CI: 0.77, 0.86) in 2017. After removing single-observation records, estimated survival probabilities were lower overall, with ϕ = 0.81 (95% CI: 0.78, 0.84) in 2008 and ϕ = 0.80 (95% CI: 0.76, 0.84) in 2017.

Removing single-observation records from the simulated datasets eliminated bias caused by misreads (Figure 4), but also decreased precision of estimates. When data filtering was applied to simulated data, no annual trend was detected in survival probabilities (Figure 3). For the 20-year study length, RMSE was reduced to 0.01 and 0.009 for error rates of 0.005 and 0.1, respectively, and relative bias ranged from -0.02% - 0.05%. However, for the 5-year study length, removing single-observations decreased the precision of survival estimates, with RMSE of ~0.09 for all error rate scenarios (Figure 4).

Discussion

We estimated the rate of individual misidentification in 10-year mark-resight dataset and found a minimum error rate of 0.31% and maximum of 6.6%. Our simulation study showed that introducing misreads into long-term mark-resight data results in spurious negative trends in annual survival probability when in reality survival is constant over time, but that those effects can be mitigated by removing all single-observation records from the data. The bias in survival estimates caused by misread errors increased with both the simulated error rate and the length of the study. Longer studies allow for more opportunities for misreads to occur, and individuals marked early in the study have a greater propensity to be falsely detected after they have died, inflating estimates of survival from previous years and leading to apparent negative trends.

Without data filtering, analysis of the Red Knot data indicated a slight decline in apparent annual survival from 2008-2018, but that trend was no longer detected when all single-observation records were removed. Our post-filtering estimates of Red Knot annual survival

probability agree with a recent meta-analysis of global shorebird survival rates that synthesized the existing literature and reported an average Red Knot annual survival probability of 0.801 ± 0.011 (Méndez et al. 2018). Along with many Arctic-breeding shorebirds, the *rufa* Red Knot is a species of conservation concern; accurate estimates of demographic rates are important both for understanding the ecological drivers of declines and for implementing effective conservation actions. Most mortality for long-distance migrants likely occurs during the migratory or nonbreeding period (Newton 2006). Survival estimation is a core component of full annual cycle population modeling (e.g., Flockhart et al. 2015, Rushing et al. 2017); positively biased survival estimates and spurious trends in those estimates could obscure our understanding of relative importance of each phase of the annual cycle for population stability.

Simulation Study Extensions

Our simulation study indicated that for misread error rates ≥ 0.05, survival estimates were positively biased for earlier years in the time series, resulting in apparent negative trends in survival probability over time. Other studies have also demonstrated the potential for errors in mark-recapture data to bias estimates of survival (Schwarz and Stobo 1999, Morrison et al. 2011), but here we also show that errors could result in spurious ecological inference about trends or, potentially, drivers of survival probability over time.

The inputs of our simulation could be tailored to the specifics of any study to empirically evaluate the effect of various levels of misidentification errors on inference about apparent annual survival. We encourage others to use a simulation approach to compare observed trends in survival estimates with those that would be expected given high error rates alone. Our simulation code is freely available online for this purpose. This exploratory analysis could

assuage concerns about suspected low misread rates, or alert researchers to potential bias in their estimates if error rates are presumed to be higher.

The way we simulated misreads was simplistic for the purposes of assessing general bias. In reality misreads are likely nonrandom and occur in several ways that can be difficult to explicitly model. The simplest type of misread occurs when a single character is misread and replaced by another, incorrect, character (e.g., AT2 recorded as 4T2, or *vice versa*). This type of misread is unlikely to occur randomly, as certain pairs or groups of characters are more likely to be mistaken for each other (e.g., A and 4, K and X). Another type of misread could occur where all individual characters are recorded correctly, but transcribed in the wrong order (e.g., AT2 recorded as A2T). This incorrect transcription could affect encounter histories in a more random way. The goal of our simulation work was not to capture and explore all the specific misread cases that apply to our system, but rather to generally explore the potential consequences of misread errors on estimation and ecological inference.

Minimizing Errors during Data Collection and Processing

The first, and most obvious, approach to reduce the effect of misreads is to take measures to reduce the probability of misreads occurring during data collection (Bearhop et al. 2003). Observer training and emphasis on data quality over quantity is key to ensuring robust mark-recapture data. For inscribed flags or bands, several papers have detailed the combinations of letters and numbers that are more likely to be misread, as well as the field conditions most likely to lead to errors (Burton 2000, Bearhop et al. 2003, Milligan et al. 2003, Clark et al. 2005, Mitchell and Trinder 2008). Here we also stress the importance of long-term participants instead of rapid turnover of new observers (whether they be volunteers or paid interns). In our study, the

number of resightings recorded in the field, but not necessarily the number of years with the project, was a good predictor of misread rates. Regardless of background or previous experience, all observers in our study converged on the median apparent misread rate after ~300 total observations. However, although estimated error rates from inexperienced observers were higher (8.8% compared to 6.4% from experienced observers), removing records from inexperienced observers had little to no effect on estimates of Red Knot survival. This may be a function of the larger variation in error rates among "inexperienced" observers, who have widely varying personal background and experience.

A second way to reduce the effect of misreads is through more extensive data filtering, as demonstrated in this paper. Analyzing observer-specific misread rate as a function of some metric for experience for any study (number of years, number of observations), as we did here, could help determine the point at which observers tend to the median, and therefore the appropriate study-specific threshold of excluding observations. We found that removing records from "inexperienced" observers had little effect on survival estimates, but estimating survival with and without those records could be a useful exercise in diagnosing whether observer experience is an important factor to consider.

Due to intensive resighting effort that occurs in Delaware during spring migration, we were able to identify instances when a flag was recorded only once in a year as possible misreads. Because the probability of the same false positive detection occurring more than once in the same year is low, most false detections will also be single observations. Removing those from the data will inevitably remove detections of real individuals, but will also remove all or nearly all false positives. Similar to the *ad hoc* method of dealing with transients first proposed by Pradel *et al* (1997), this is will reduce precision but ensure unbiased estimates of survival. We

do not propose removing all single-observation records as a rule of thumb, but as a potential data filtering option that could be considered if the occurrence of misreads is a concern and the primary goal of the analysis is to estimate annual survival probability. Researchers should carefully consider the trade-off between precision obtained by larger sample sizes and the number of potential errors due to misreads in their data.

Modeling Misidentification Errors

The Cormack-Jolly-Seber (CJS) model for estimating annual survival is the basis for a large suite of extensions to deal with more complex questions and data structures (see Williams et al. 2002). We investigated the effect of misreads on a simple application of the CJS model, and expect that biases identified here would carry through to more complex extensions. We only investigated the effect of misreads on estimation of apparent annual survival, but capture-recapture data are used to estimate a variety of other demographic parameters. Additionally, we simulated only one type of misread, where real marks are mistakenly observed, as opposed to the type of misread that results in the observation of a new, non-existent mark. In most mark-resight studies, the latter type is likely to be filtered from the data before analysis, and therefore is of less concern.

The type of misidentification that results in new, non-existent individuals is the type frequently encountered with noninvasive genetic sampling and photographic capture-recapture, and several model-based methods have been proposed to deal with these false positives in genetic and photographic mark-recapture. The $M_{t,\alpha}$ model, introduced by Lukacs and Burnham (2005) and extended by Link *et al* (2010) and Yoshizaki *et al* (2011), uses a latent multinomial to determine the probability that a given capture history is real. A key assumption of the $M_{t,\alpha}$

models is that false identifications always create new nonexistent individuals and that these ghost individuals can only be observed once. To deal with this problem of ghost capture histories, Morrison *et al.* (2011) described a method of "conditioning" natural mark capture-recapture data by removing all first detections of new individuals, which is conceptually similar to the *ad hoc* method to remove transients proposed by Pradel *et al.* (1997) and to our process of removing all single-observation events. Even though some real observations will be discarded in the process, it is likely that doing so will also eliminate all false detections, therefore reducing bias. For long-term mark-resight datasets, it is likely that both possible and impossible marks could be recorded incorrectly, and may be recorded more than once in the dataset. If filtering impossible marks from the data is not feasible, for example in a migratory flyway with multiple projects marking individuals, modeling the misread probability for these types of data becomes incredibly complex.

By modifying the likelihood of the existing CJS family of models, it could be possible to design an analysis framework that could directly estimate misread probability. Multievent models which directly account for uncertainty in state assignment (Pradel 2005, Choquet et al. 2009) could provide a basis for explicitly modeling for the probability of false positive detections. These complex models add additional parameters and associated variances, however, which could obfuscate ecological research and management support. Therefore addressing the problem first through data collection and screening protocols is preferable, and modeling false positives directly may only be necessary for very long studies with suspected high rates of false positive detections.

All human data recording is prone to error, and errors due to individual misidentifications can accumulate in mark-resight studies as the study length and number of deployed marks increases. Long-term mark-resight studies may be especially vulnerable to biased estimates of apparent survival as a result of these errors if misread rates are high. Observer training and retention protocols may help reduce error during data collection, but ecologists working with historical mark-resight datasets should also consider potential effects of misread errors when designing analyses. Exploratory data analyses and simulation studies such as those presented here can aid in estimating probable error rates in a dataset and evaluating potential bias with the planned modeling approach. If the potential for misreads is not considered in analysis long-term datasets collected by many observers, spurious trends caused by errors could influence ecological inference about population demographics and resulting conservation decision-making.

Literature Cited

Atkinson, P.W., A.J. Baker, K.A. Bennett, N.A. Clark, J.A. Clark, K.B. Cole, A. Dekinga, A. Dey, S. Gillings, P.M. González, K. Kalasz, C.D.T. Minton, J. Newton, L.J. Niles, T. Piersma, R.A. Robinson, and H.P. Sitters (2007). Rates of mass gain and energy deposition in red knot on their final spring staging site is both time- and condition-dependent. Journal of Applied Ecology 44: 885–895.

Baker, A.J., P.M. Gonzalez, R.I G. Morrison, and B.A. Harrington (2001). Red Knot (Calidris canutus). In The Birds of North America (A.F. Poole, Editor). Cornell Lab of Ornithology, Ithaca NY, USA. doi: 10.2173/bna.563

Bearhop, S., R.M. Ward, and P.R. Evans (2003). Long-term survival rates in colour-ringed shorebirds – practical considerations in the application of mark–recapture models. Bird Study 50: 271–279.

Beck, A., B. Ackerman, L. Barton, and C. Hartley (2004). Survival estimates for Florida manatees from the photo-identification of individuals. Marine Mammal Science 20: 438–463.

Bildstein, K.L. (1998). Long-term counts of migrating raptors: a role for volunteers in wildlife research. Journal of Wildlife Management 62: 435–445.

Bolker, B., and R Development Core Team (2017). bbmle: Tools for general maximum likelihood estimation. R package version 1.0.20.

Burton, N.H.K. (2000). Variation in sighting frequencies of colour-ringed Redshanks Tringa totanus according to ringing-scheme and ring colour. Wader Study Group Bulletin 91: 50–53.

Choquet, R., L. Rouan, and R. Pradel (2009). Program E-Surge: A software application for fitting multievent models. In Modeling Demographic Processes in Marked Populations (D.L. Thompson, E.G. Cooch, and M.J. Conroy, Editors) 845-865. doi: 10.1007/978-0-387-78151-8

Clark, N., S. Gillings, A.J. Baker, P.M. Gonzalez, and R. Porter (2005). The production and use of permanently inscribed leg flags for waders. Wader Study Group Bulletin 108: 38–41.

Cohn, J.P. (2008). Citizen Science: Can Volunteers Do Real Research? BioScience 58: 192–197.

Conrad, C.C., and K.G. Hilchey (2011). A review of citizen science and community-based environmental monitoring: Issues and opportunities. Environmental Monitoring and Assessment 176: 273–291.

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

Gillings, S., P.W. Atkinson, A.J. Baker, K.A. Bennett, N.A. Clark, K.B. Cole, P.M. González, K.S. Kalasz, C.D.T. Minton, L.J. Niles, R.C. Porter, I.D.L. Seranno, H.P. Sitters, and J.L.

Woods (2009). Staging Behavior in Red Knot (Calidris canutus) in Delaware Bay: implications for monitoring mass and population size. The Auk 126: 54–63.

Gimenez, O., J.D. Lebreton, R. Choquet, and R. Pradel (2018). R2ucare: An R package to perform goodness-of-fit tests for capture-recapture models. Methods in Ecology and Evolution 9: 1749-1754.

Laake, J.L., D.S. Johnson, and P.B. Conn (2013). marked: An R package for maximum-likelihood and MCMC analysis of capture-recapture data. Methods in Ecology and Evolution 4: 885–890.

Lavers, J.L., and I.L. Jones (2008). Assessing the type and frequency of band resighting errors for Razorbill Alca torda with implications for other wildlife studies. Marine Ornithology 36: 19–23.

Lebreton, J.D., K.P. Burnham, J. Clobert, and D.R. Anderson (1992). Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. Ecological Monographs 62: 67–118.

Link, W.A., J. Yoshizaki, L.L. Bailey, and K.H. Pollock (2010). Uncovering a latent multinomial: analysis of mark-recapture data with misidentification. Biometrics 66: 178–185.

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

Magurran, A.E., S.R. Baillie, S.T. Buckland, J.M. Dick, D.A. Elston, E.M. Scott, R.I. Smith, P.J. Somerfield, and A.D. Watt (2010). Long-term datasets in biodiversity research and monitoring: assessing change in ecological communities through time. Trends in Ecology & Evolution 25: 574–582.

McGowan, C.P., J.E. Hines, J.D. Nichols, J.E. Lyons, D.R. Smith, K.S. Kalasz, L.J. Niles, A.D. Dey, N.A. Clark, P.W. Atkinson, C.D.T. Minton, and W. Kendall (2011). Demographic consequences of migratory stopover: linking red knot survival to horseshoe crab spawning abundance. Ecosphere 2(6). doi: 10.1890/ES11-00106.1

McGowan, C.P., D.R. Smith, J.D. Nichols, J.E. Lyons, J. Sweka, K. Kalasz, L.J. Niles, R. Wong, J. Brust, M. Davis, and B. Spear (2015). Implementation of a framework for multi-species, multi-objective adaptive management in Delaware Bay. Biological Conservation 191: 759–769.

Méndez, V., J.A. Alves, J.A. Gill, and T.G. Gunnarsson (2018). Patterns and processes in shorebird survival rates: a global review. Ibis. doi: 10.1111/ibi.12586

Miller, D.A., J.D. Nichols, B.T. McClintock, E.H.C. Grant, L.L. Bailey, and L.A. Weir (2011). Improving occupancy estimation when two types of observational error occur: Non-detection and species misidentification. Ecology 92: 1422–1428.

Milligan, J.L., A.K. Davis, and S.M. Altizer (2003). Errors associated with using colored leg bands to identify wild birds. Journal of Field Ornithology 74: 111–118.

Mitchell, C., and M. Trinder (2008). On reading colour rings. Ringing and Migration 24: 11–14.

Morrison, T.A., J. Yoshizaki, J.D. Nichols, and D.T. Bolger (2011). Estimating survival in photographic capture-recapture studies: overcoming misidentification error. Methods in Ecology and Evolution 2: 454–463.

Newman, C., C.D. Buesching, and D.W. Macdonald (2003). Validating mammal monitoring methods and assessing the performance of volunteers in wildlife conservation—"Sed quis custodiet ipsos custodies?" Biological Conservation 113: 189–197.

Newton, I. (2006). Can conditions experienced during migration limit the population levels of birds? Journal of Ornithology 147: 146–166.

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

Pradel, R., J.E. Hines, J.D. Lebreton, and J.D. Nichols (1997). Capture-recapture survival models taking account of transients. Biometrics 53: 60–72.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Roche, E.A., C.M. Dovichin, and T.W. Arnold (2014). Field-readable alphanumeric flags are valuable markers for shorebirds: use of double-marking to identify cases of misidentification. Journal of Field Ornithology 85: 329–338.

Rushing, C.S., J.A. Hostetler, T.S. Sillett, P.P. Marra, J.A. Rotenberg, and T.B. Ryder (2017). Spatial and temporal drivers of avian population dynamics across the annual cycle. Ecology 98: 2837–2850.

Schwarz, C.J., and W.T. Stobo (1999). Estimation and effects of tag-misread rates in capture-recapture studies. Canadian Journal of Fisheries and Aquatic Sciences 56: 551–559.

Silvy, N.J., R.R. Lopez, and M.J. Peterson (2012). Techniques for marking wildlife. The Wildlife Techniques Manual Volume 1: Research (N.J. Silvy, Editor). 230–257.

Tulloch, A.I.T., H.P. Possingham, L.N. Joseph, J. Szabo, and T.G. Martin (2013). Realising the full potential of citizen science monitoring programs. Biological Conservation 165: 128–138.

Weiss, N.T., M.D. Samuel, D.H. Rusch, and F.D. Caswell (1991). Effects of resighting errors on capture-resight estimates for neck-banded Canada Geese. Journal of Field Ornithology 62: 464–473.

Williams, B., J. Nichols, and M. Conroy (2002). Analysis and Management of Wildlife Populations. 1st edition. Academic Press, San Diego, California.

Wright, J.A., R.J. Barker, M.R. Schofield, A.C. Frantz, A.E. Byrom, and D.M. Gleeson (2009). Incorporating genotype uncertainty into mark-recapture-type models for estimating abundance using DNA samples. Biometrics 65: 833–840.

Yoshizaki, J., C. Brownie, K.H. Pollock, and W.A. Link (2011). Modeling misidentification errors that result from use of genetic tags in capture-recapture studies. Environmental and Ecological Statistics 18: 27–55.

Yoshizaki, J., K.H. Pollock, C. Brownie, and R.A. Webster (2009). Modeling misidentification errors in capture-recapture studies using photographic identification of evolving marks. Ecology 90:3–9.

Yu, J., R.A. Hutchinson, and W.K. Wong (2014). A latent variable model for discovering bird species commonly misidentified by citizen scientists. AAAI Conference on Artificial Intelligence 500–507.

Tables

Table 2-1: Summary of red knot flag resightings 2009-2014.
The number of observers is the total number of unique observers that contributed resighting records in each year. Individual flags is the number of unique flags seen, total resightings is the total number of flag resightings in the database for that year, and mean resightings per flag is the total number resightings divided by the number of flagged individuals for each year. Confirmed misreads are records of withheld or not-yet-deployed flags. Possible misreads are flags that were only recorded a single time in that year.

Year	Number of observers	Individual flags	Total resightings	Average resightings per flag	Confirmed misreads	Possible misreads
2009	42	1917	10888	5.7	27 (0.25%)	581 (5.3%)
2010	41	1217	4144	3.4	13 (0.31%)	445 (10.7%)
2011	41	1897	8261	4.4	38 (0.46%)	634 (7.7%)
2012	39	1409	6781	4.8	26 (0.38%)	517 (7.6%)
2013	36	939	3879	4.1	19 (0.49%)	368 (9.5%)
2014	39	1336	6540	4.9	30 (0.46%)	422 (6.5%)
2015	42	2265	13755	6.1	49 (0.36%)	738 (5.4%)
2016	43	726	2346	3.2	6 (0.26%)	320 (13.6%)
2017	53	1831	11066	6.0	18 (0.16%)	622 (5.6%)
2018	46	2098	13220	6.3	26 (0.19%)	727 (5.5%)

Figures

Figure 2-1: Picture of a leg flag on a red knot.
Picture of a three-character lime green leg flag on a Red Knot in Delaware, U.S.A.
This flag was deployed in 2014 and the photograph was taken in May 2017. Photograph by Jean Hall.

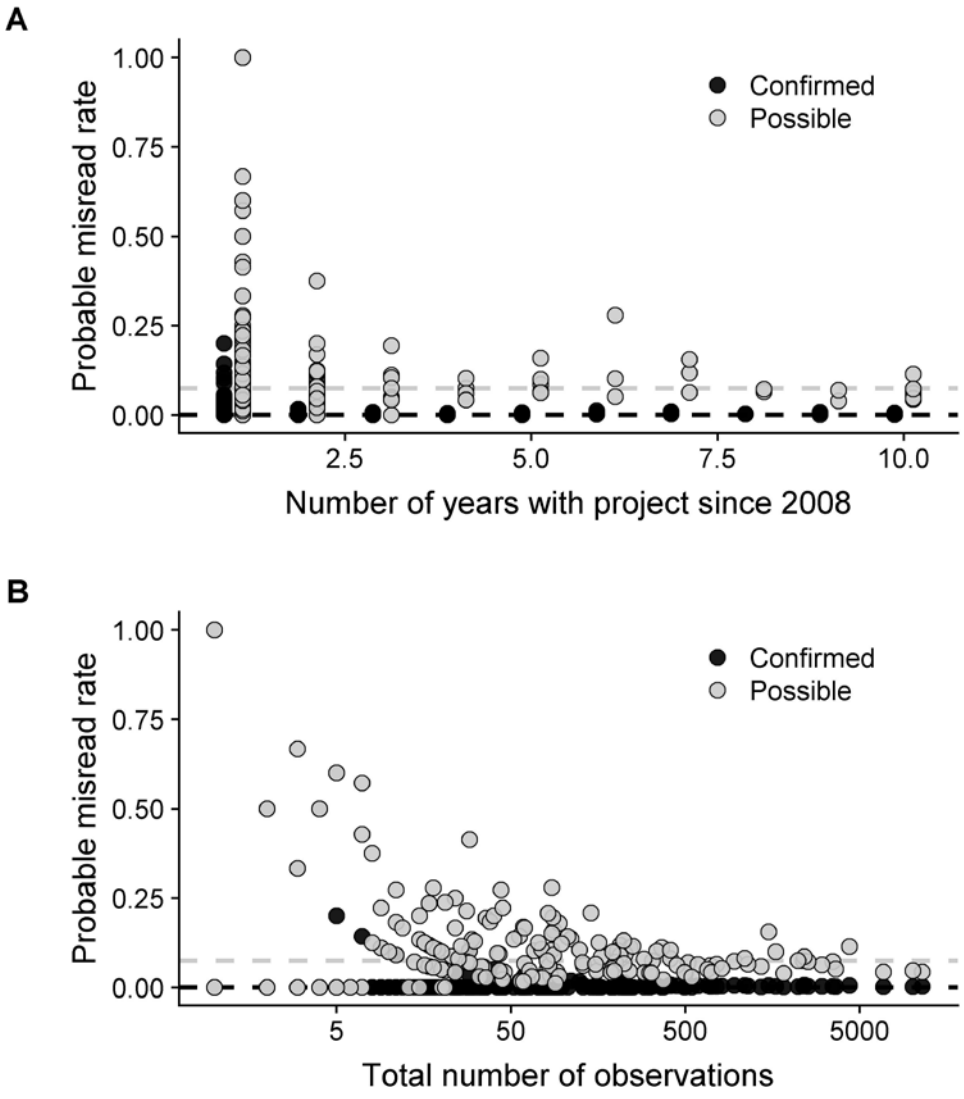

Figure 2-2: Observer-specific misread rates.
For each observer, the proportion of their total resightings that are confirmed (black) or possible (gray) misreads as a function of the number of years they have spent on the project (**A**) and the total number of resightings they have recorded (**B**). Total number of observations were modeled and are plotted on the log scale. Horizontal dashed lines indicate among-observer median.

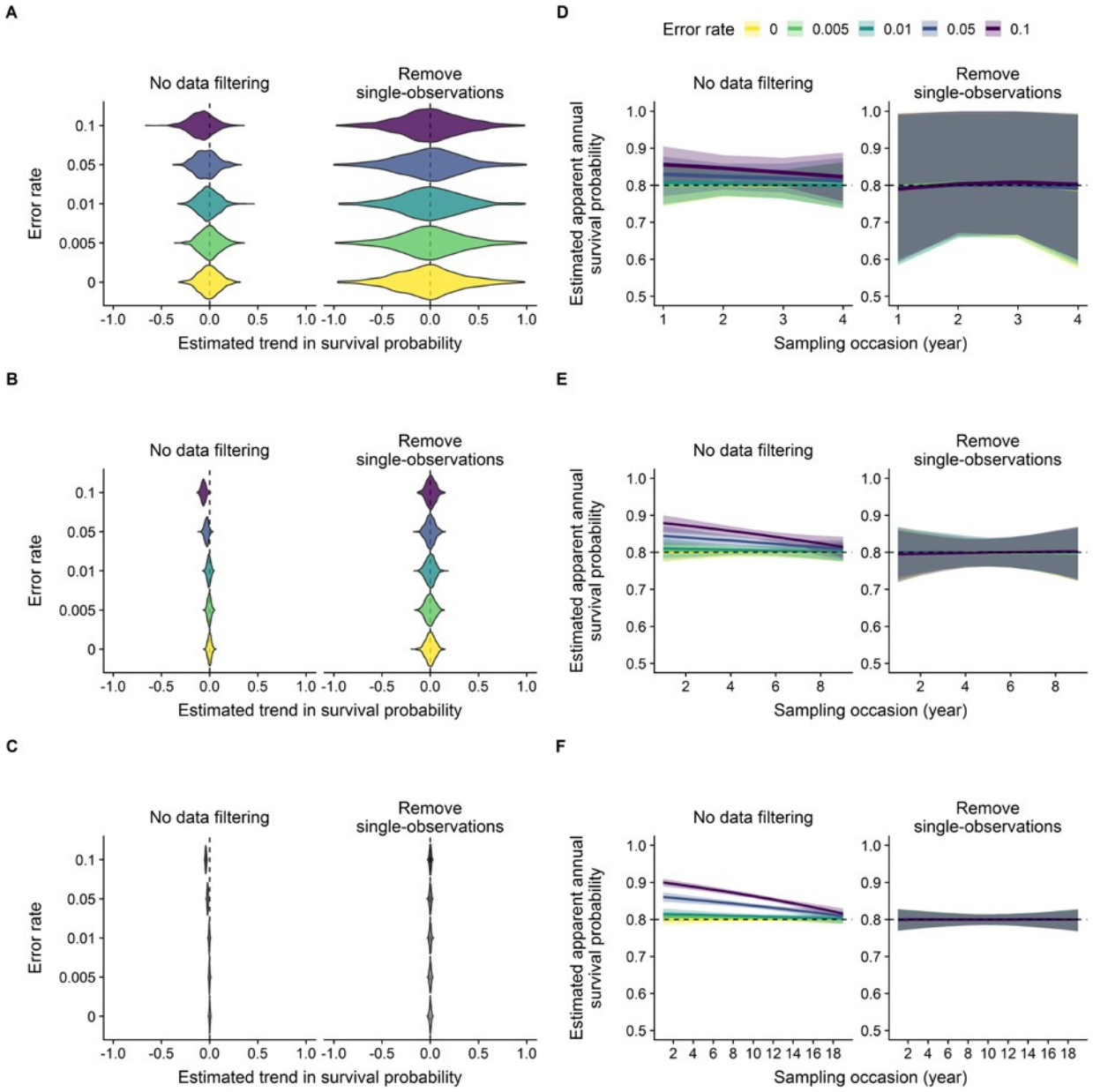

Figure 2-3: Results of simulation study.
Results are given for a 5-year study (**A, D**), 10-year study (**B, E**), and 20-year study (**C, F**). Misreads led to apparent negative trends in survival probability over time (**A-C**), particularly for long-term studies. Removing all single-observation records eliminated bias caused by misreads but also decreased precision of estimates, especially for short time series (**A, D**). Lines in **D**, **E**, and **F** represent median estimates from 1000 replications of each scenario and shaded regions represent 95% quantiles of estimates.

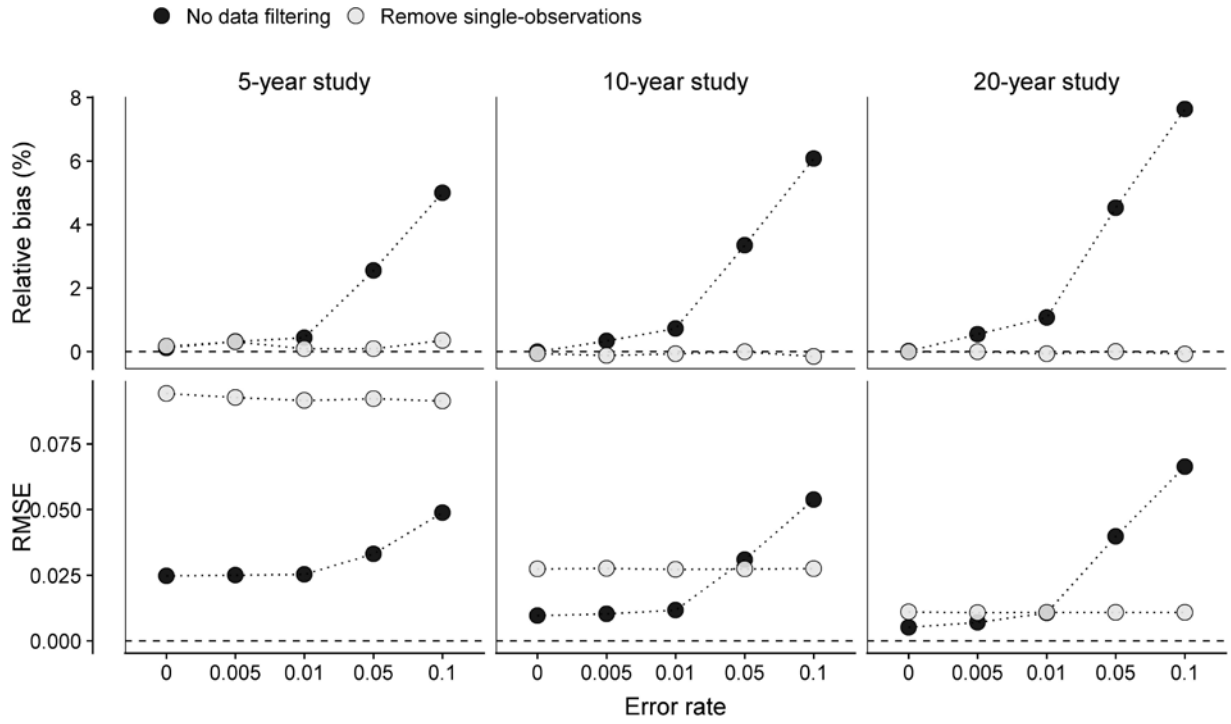

Figure 2-4: Effect of misreads on annual survival estimates
Root mean squared error (RMSE) and relative bias (%) of estimates under different study lengths and error rates. Without data filtering (black), bias induced by misread errors increased with study length and error rate. Filtering the data by removing all single observations (gray), effectively reduced bias but resulted in lower precision of estimates, particularly for shorter time series.

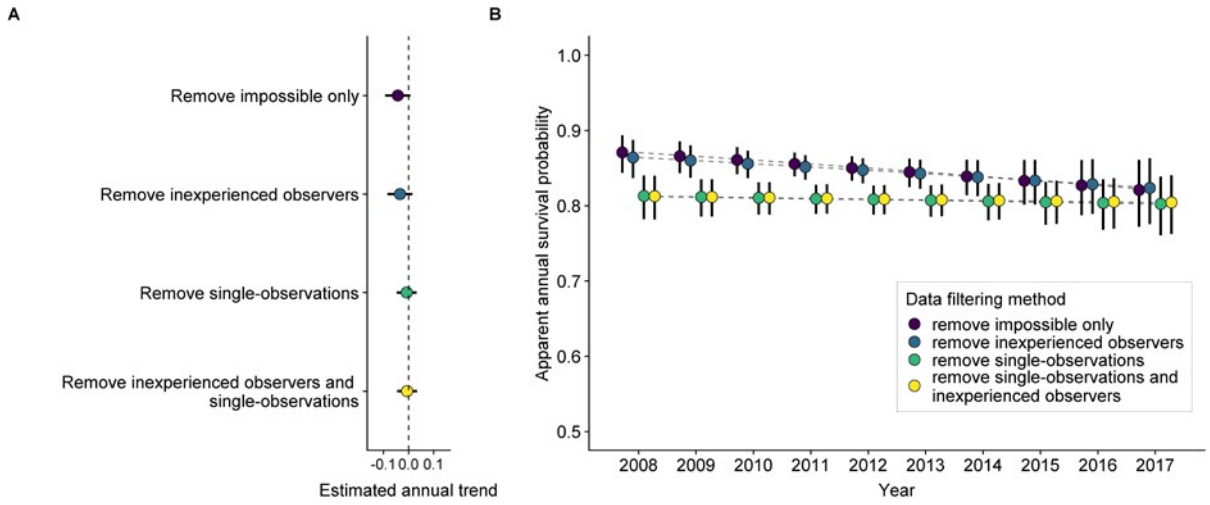

Figure 2-5: Effects of data filtering methods on estimates of red knots survival.
Effects of data filtering methods to reduce the effect of potential flag misread error on estimated trends (**A**) and annual apparent survival probability (**B**) of Red Knot
Removing only impossible observations or those from inexperienced observers led to an apparent negative trend in survival probability over the past 10 years, while removing single-observation records resulted in no trend in survival probability.

Supporting Information

Table S1: RMSE and bias of CJS model estimates under varying error rates and study lengths.

Study length (years)	Error rate	Survival probability		Detection probability	
		RMSE	Relative bias	RMSE	Relative bias
No data filtering					
5	0	0.025	0.113	0.018	0.231
5	0.005	0.025	0.320	0.019	-0.245
5	0.01	0.025	0.433	0.019	0.216
5	0.05	0.033	2.55	0.021	0.411
5	0.1	0.049	5.00	0.020	0.564
10	0	0.010	-0.006	0.008	0.004
10	0.005	0.010	0.334	0.009	-0.222
10	0.01	0.012	0.723	0.011	-0.950
10	0.05	0.031	3.344	0.022	-4.04
10	0.1	0.054	6.073	0.034	-6.56
20	0	0.005	0.008	0.005	-0.037
20	0.005	0.007	0.547	0.008	-1.40
20	0.01	0.011	1.070	0.016	-3.00
20	0.05	0.040	4.53	0.058	-11.5
20	0.1	0.066	7.63	0.092	-18.3
Remove single-observations					
5	0	0.094	0.164	0.425	-84.4
5	0.005	0.093	0.306	0.423	-84.1
5	0.01	0.092	0.086	0.424	-84.3
5	0.05	0.092	0.091	0.425	-84.5
5	0.1	0.091	0.352	0.424	-84.4
10	0	0.027	-0.062	0.425	-84.6
10	0.005	0.028	-0.126	0.425	-84.6
10	0.01	0.027	-0.067	0.426	-84.7
10	0.05	0.027	-0.004	0.426	-84.9
10	0.1	0.027	-0.157	0.425	-84.5
20	0	0.011	-0.006	0.425	-84.6
20	0.005	0.011	-0.006	0.425	-84.6
20	0.01	0.011	-0.070	0.425	-84.6
20	0.05	0.011	0.001	0.426	-84.7
20	0.1	0.011	-0.076	0.425	-84.6

Chapter 3: Foraging ecology mediates the effect of ecological mismatch on stopover mass gain dynamics

Abstract

Impacts of ecological mismatches should be most pronounced at points of the annual cycle when populations depend on a predictable and abundant, but aggregated, food resource. The degree to which species specialize on a key prey item should determine their sensitivity to mismatches. We evaluated the hypothesis that the effects of ecological mismatch on migratory stopover are mediated by a species' foraging ecology by comparing two similar long-distance migratory species that differ in their foraging strategies during stopover. We predicted that a specialist foraging strategy would make species more sensitive to effects of mismatch with a historically-abundant prey, while an active, generalist foraging strategy should help buffer against changing local conditions. We estimated arrival times, start of mass gain, and rate of mass gain during spring stopover in Delaware Bay, USA. At this site shorebirds feed on a temporally aggregated food resource (horseshoe crab *Limulus polyphemus* eggs), the timing of which is linked to water temperature; red knot (*Calidris canutus rufa*) specializes on these while the ruddy turnstone (*Arenaria interpres*) feeds more generally. We used a hierarchical nonlinear model to estimate the effect of mismatch between shorebird arrivals and timing of crab spawning on the timing and rate of mass gain over 22 years. In cooler years crabs spawned later, which was associated with later and faster mass gain for the knots. Turnstones exhibited less inter-annual variation in the timing and rate of mass gain than knots, and we found no relationship

between mass gain dynamics and the availability of horseshoe crab eggs for this generalist species. These results suggest that generalist foraging strategies may buffer migratory species against phenological mismatch. Less adaptable, specialist species are declining globally, and measures to mitigate the impact of future mismatch may be more feasible where species depend on spatially concentrated resources.

Introduction

Temporal shifts in phenology have the potential to cause mismatches between species needs and resource availability (Walther et al. 2002, Edwards et al. 2004, Petanidou et al. 2014). Migratory species depend on productive, seasonal habitats and may be susceptible to phenological mismatches if environmental changes at breeding, nonbreeding, or stopover sites occur asynchronously (Robinson et al. 2009, Thorup et al. 2017). The negative effects of these mismatches are most likely to be pronounced at points in the annual cycle when populations depend on predictable, aggregated food resources. Although theory predicts that mismatches should have negative population-level effects, empirical studies show mixed evidence for such effects (Stenseth and Mysterud 2002, Miller-Rushing et al. 2010). Understanding if and how species can adjust to shifting annual cues and resource availability is critical for predicting species' responses to climate change.

One mechanism by which animals may respond to climate change is through plasticity in the timing of annual events (e.g., migration, breeding). For example, many Neotropical migratory songbirds have advanced the timing of pre-breeding migration and hence breeding over the past few decades in response to warming spring temperatures (Winkler et al. 2002) and migratory ungulates and geese time northward movements to track the onset of spring across the landscape (Bischof et al. 2012, Si et al. 2015). However, other studies have shown that

individuals may not always be able to avoid mismatches, leading to insufficient resources during migration (Post et al. 2008) or the breeding season (Plard et al. 2014, Senner et al. 2016). Mismatches often occur because heterogeneous patterns of warming and cooling across the landscape lead to situations in which nonbreeding or migratory conditions are poor indicators of breeding site conditions (Robinson et al. 2009).

Ecological mismatches have been largely studied during the breeding season, but seasonal resource availability is critical during migration. Additionally, cues during the nonbreeding season influence the timing of migration and stopover, and if arrival to the breeding site is delayed individuals might miss the optimal resource window for reproduction (Kokko 1999). Pre-breeding migration typically occurs in a narrow time frame, and during stopover individuals have limited time to acquire fuel stores needed to continue migration (Nilsson et al. 2013). This narrow window must coincide with prey availability. Insufficient resources during migratory stopover could result in either delayed or underweight departure, both of which could have individual carry-over effects to the breeding season, with potential population-level demographic impacts (Newton 2006).

The degree to which mismatches negatively affect populations likely depends, in part, on foraging ecology and breadth of diet. A generalist foraging strategy makes a species less reliant on the availability of a single key prey item while temporal overlap with prey is critical for specialist foragers. Generalist species are better able to cope with novel or unpredictable environments (Mayr 1965, Sol et al. 2002), and therefore should be less sensitive to the availability of a specific prey item.

Shorebirds are time-constrained during spring migration and rely on plentiful food at stopover sites to quickly accumulate large fat stores, which are necessary to continue their

migration. Arctic-breeding shorebirds have some of the longest known migrations and have experienced widespread declines in the past several decades (Piersma et al. 2014, Pearce-Higgins et al. 2017, Wauchope et al. 2017). Degradation of stopover sites has been implicated as a key contributor to observed declines (Baker et al. 2004, Studds et al. 2017). These species rely on predictable aggregations of abundant prey at stopover sites *en route* to replenish fat stores needed to complete their migratory journey, therefore matching of arrival timing and food availability at those sites may be critical for population stability. In Delaware Bay, USA, a globally-important stopover site for several migratory shorebirds, the eggs of spawning horseshoe crabs (*Limulus polyphemus*) provide a critical, but temporally variable and ephemeral, food resource during spring migration (Tsipoura and Burger 1999). Declining horseshoe crab abundance due to past unregulated harvest has been suggested as the cause of observed declines in the numbers of birds seen in Delaware Bay during stopover (Baker et al. 2004), especially the red knot (*Calidris canutus rufa*), which was recently listed as Threatened under the United States Endangered Species Act (79 FR 73705 73748). As horseshoe crab and red knot populations recover with the aid of strict harvest regulations (McGowan et al 2015), the ability of birds to reach adequate departure mass may be important for population stability (Baker et al 2004). In addition to the overall abundance of eggs, timing of horseshoe crab spawn plays a key role in the availability of adequate food resources in Delaware Bay. Horseshoe crabs spawn once a year during the spring high tides in May and early June, coinciding with shorebird stopover in mid- to late-May (Niles et al. 2008). Timing of spawn is linked to lunar periods and water temperature, with the largest spawn typically on the full moon high tide when water temperature is ≥ 15°C (Smith et al. 2002, Smith and Michels 2006). Over the past 20 years, spring water temperatures in Delaware Bay have become slightly cooler, with an overall change in average May water temperature of -

0.01°C $year^{-1}$ from 1998-2018. Cooler temperatures have been anecdotally linked to observed delays in crab spawn and concerns about whether shorebirds are able to gain mass at the necessary rate.

Here, we evaluate the hypothesis that the effects of mismatch on stopover ecology differ by species' foraging strategy. We predicted that, when compared to a specialist, a generalist forager would show less among-year variation in the timing and rate of mass gain during stopover, and have no relationship between dynamics mass gain and the availability of a single prey item. To test our hypothesis we considered two long-distance migratory shorebirds, the red knot and the ruddy turnstone (*Arenaria interpres*). Both species breed in the Arctic with medium- to long- distance migrations and use Delaware Bay as a primary spring stopover. Red knots are specialist foragers that rely on horseshoe crab eggs during spring migration (Baker et al. 2001, Niles et al. 2008, Gillings et al. 2009), whereas ruddy turnstones are more active, generalist foragers that will dig aggressively for food, more readily consume other prey in addition to horseshoe crab eggs, and exploit rocky areas not typically used by red knot (Nettleship 2000). We predicted that red knot would show more variation in the timing and rate of mass gain than ruddy turnstone, and that mass gain dynamics for red knot, but not ruddy turnstone, would be dependent on the timing of horseshoe crab spawning. The results of this research will help advance our ability to assess and predict a population's vulnerability to climate change-induced mismatches in resource availability with annual cycle timing, especially for migratory species.

Methods

Study species

Northward migration for both red knot and ruddy turnstone begins in late April/early May, with most birds arriving to Delaware Bay in mid- to late-May and staying for 1-2 weeks, before continuing on to the central Canadian Arctic to begin breeding in early June (Smith et al. 2010). Evidence suggests that most Arctic-breeding shorebirds are income breeders and fuel stores acquired during migratory stopover are largely depleted during the final flight to the breeding grounds (Klaassen et al. 2001). The optimal temporal window for breeding in the Arctic is short and therefore punctual arrival is important (Meltofte et al. 2007).

Both red knot and ruddy turnstone primarily consume horseshoe crab eggs during stopover in Delaware Bay, but at other points of the annual cycle consume a variety of invertebrate prey, including crustaceans, mollusks, and insects (Nettleship 2000, Baker et al. 2001). Ruddy turnstone have a more diverse and opportunistic diet than red knot and will also scavenge carrion and human garbage (Gill 1986). When foraging for horseshoe crab eggs, red knot consume eggs on the sand surface or upper layer of sand (1-3 cm) while ruddy turnstone can actively dig for buried eggs increasing the effective resource availability (Nettleship 2000, Gillings et al. 2007).

Study site and field methods

Birds were captured in mixed foraging flocks using cannon nets during the spring stopover period between May 1 and June 5 in Delaware, USA from 1997-2018, weighed with a digital scale accurate to the nearest 0.1 g and marked with Incoloy USGS leg bands and—beginning in 2004—Darvic leg flags inscribed with unique field-readable alphanumeric codes. The number and timing of catches varied among years. From 2005-2018, beaches were also

surveyed to resight individually flagged birds. Trained observers visited study beaches in 3-day sampling periods to scan flocks for flagged individuals.

Timing of shorebird arrival to stopover site

We used mark-resight data from flagged birds to estimate shorebird arrival probabilities over the course of the stopover season for each year using Jolly-Seber models (Figure 1A). Details of model specification and goodness-of-fit tests for these are provided in the Supporting Information. We estimated sampling period-specific entry probabilities across the season for each species in each year and derived the cumulative probability of arriving during or before each sampling period. We assumed that flagged individuals were a representative sample of the stopover population and used this cumulative probability as an index of the proportion of the stopover population that had entered the region by each period. To determine the proportion of horseshoe crab spawn available to shorebirds, we defined a cutoff date for each shorebird species in each year as the period by which 95% of the population had arrived at the stopover site. To compare these dates to horseshoe crab spawning activity, and for clarity of presentation, we converted 3-day shorebird sampling periods to days since May 1 using the middle day of each sampling period. Resighting data were not available prior to 2005, so we used the median 95%-cutoff period for each species to assess horseshoe crab spawning availability for 1997-2004.

Availability of horseshoe crab eggs

Horseshoe crab spawning activity was surveyed across 13 beaches in Delaware each year from 2003-2017. Observers counted the number of females on the high tide line per one-meter quadrat during each lunar period's high tide (from 2 nights prior to the full or new moon to 2

nights post) totaling ~12 surveys per beach per year. Survey design and methodology are described in detail by Smith et al. (2002). The statewide average number of females m^{-2} is reported as an index of female spawning abundance (IFSA), which we use as a proxy for egg abundance (Zimmerman et al. 2016).

We used survey-specific IFSA from across the season to assess temporal patterns in spawning. For each year, we determined the proportion of cumulative total annual spawn that had occurred by each day. Horseshoe crab eggs accumulate on the beach and can remain available to birds on the mudflats and upper layer of sand, therefore we assumed that spawning activity that occurred before bird arrival would still be available to foraging birds. We used a loess smoothing function to estimate the predicted proportion of spawn that occurred by each date and defined relative horseshoe crab egg availability to shorebirds as the proportion of the total spawning activity that occurred by the 95%-arrival cutoff date for each species in each year.

Timing of horseshoe crab spawn is dependent on environmental cues, with water temperatures above 15°C thought necessary for substantial spawning activity to occur (Smith and Michels 2006). For each year, we determined the number of nights with average water temperatures above 15°C prior to the date by which 95% of birds had arrived, which we expected to explain some annual variation in horseshoe crab egg availability (Figure 1B). We obtained water temperature data collected by an offshore station in Lewes, DE (station ID 8557380) from the NOAA Center for Operational Oceanographic Products and Services. We used hourly water temperatures from May 1 – June 15 to estimate daily average water temperature for those dates from 1998-2018. Temperature data were unavailable from this station for 1997 or 2017. We estimated the relationship between the number of nights with water

temperature above 15°C with the relative amount of horseshoe crab spawn available to shorebirds using a linear model, with both variables scaled and centered for analysis.

Annual variation in stopover mass gain dynamics

We analyzed capture masses of 13,046 red knots and 12,476 ruddy turnstones from 239 catches (114 catches of red knots, 125 catches of ruddy turnstones) over 22 years in Delaware (Table S1). We only included catches with at least 25 individuals weighed. The number of individuals weighed on a single occasion ranged from 25-479 for red knot (median = 77) and from 25-584 for ruddy turnstone (median = 75). The number of catches in a year ranged from 1-9 for red knot and 1-10 for ruddy turnstone; we only estimated year-specific parameters for years with > 3 catches but all data were used to estimate global mean parameters.

Observed masses increased nonlinearly throughout the stopover period, so we modeled the mean mass of individuals of each species captured on a given day as a nonlinear effect of day of the stopover period (May 1 = day 1) using a four-parameter logistic curve:

$$mass_{day,t} = max_t + \frac{min_t - max_t}{1 + \left(\frac{day}{p_t}\right)^{r_t}} \quad (1)$$

where $mass_{day,t}$ is the mean mass of all birds captured on a given day in year *t*, max_t is the maximum average mass in year *t*, min_t is the minimum average mass in year *t*. The scaling parameters p_t and r_t dictate the shape of the curve; r_t is the maximum rate of mass gain in year *t* and p_t is the inflection point, *i.e.* the day on which the maximum rate of mass gain occurred. We fit the mass gain model separately for each species.

To assess among-year variation in mass gain dynamics, we used a hierarchical annual random effects model (Figure 1C) analyzed in a Bayesian framework to estimate overall means

and year-specific deviations for the maximum mass (max_t), minimum mass (min_t), maximum rate of mass gain (r_t), and inflection point (p_t). We estimated year-specific model parameters by defining a prior distribution for the global average value of each parameter and the variance around that average, which were used to define the normal distribution from which the year-specific values were drawn, for example:

$$\begin{aligned} Min &\sim Unif(80, 150) \\ \sigma_1 &\sim Unif(0, 50) \\ min_t &\sim Normal(Min, \sigma_1) \end{aligned} \quad (2)$$

where Min is the global average minimum mass, σ_{min} is the among-year standard deviation in minimum masses, and min_t is the estimated minimum mass in year *t*. All four mass gain model parameters were estimated using this general structure. We used weakly informative Uniform priors for the average minimum and maximum masses based on documented masses for each species (Nettleship 2000, Baker et al. 2001) to keep posterior sampling within reasonable values (minimum = 80 g, maximum = 300 g). Weakly informative Uniform priors were also used for the average rate and inflection point parameters (r_t: Unif(0, 20); p_t: Unif(0, 40).

We calculated the derivatives of each annual mass gain curve to estimate the mass gain start day, end day, and average rate of mass gain in that window for each species in each year (Figure 1C). We defined the start and end day of mass gain as the days with the maximum and minimum second derivative of the predicted curve, respectively (Figure S1). This method approximates the dates on which the population experienced the greatest changes in the rate of mass gain, which we interpret to correspond to the beginning and end of the period of peak refueling.

We assumed individual masses of birds weighed on a given day were normally distributed around daily predicted means with a catch-specific variance. Observed within-catch variance varied across the season and among years, due to continued influx of new (lighter) individuals as the season progressed, so we estimated this variance as a random effect of catch unconstrained by day or year. We assigned a Uniform (0, 50) prior distribution to the average and a Gamma (1, 0.01) prior distribution to the among-year variance. To assess our assumption that within-catch masses were normally distributed, we used posterior predictive checks to estimate the sum of squared residuals for the observed individual masses and a predicted ideal data set (Kéry 2010), which indicated that this assumption was reasonable with a Bayesian p-value of 0.50 for red knot and 0.49 for ruddy turnstone.

Red knot are slightly larger than ruddy turnstone, so to facilitate comparison of annual mass gain dynamics on a common scale we present the annual variation in model parameters using the coefficient of variation ($CV = {}^{\sigma}/_{mean}$) and report rates of mass gain scaled to average minimum mass. We evaluated our prediction that red knot (specialist) would exhibit greater among-year variation in mass gain dynamics than ruddy turnstone (generalist) by assessing the overlap of the posterior distributions for CV of the inflection point and rate parameters.

Effect of resource availability on mass gain dynamics

We specified covariates on growth model parameters to estimate the effect of the relative horseshoe crab egg availability on the timing and rate of mass gain by shorebirds. Because year-specific mass gain model parameters were assumed to be normally distributed around the global average, the effects of ecological covariates were estimated using linear models with normal errors and identity links. We simultaneously estimated the relationship between water

temperature and overall horseshoe crab abundance with the availability of horseshoe crab spawn and the effect of horseshoe crab egg availability on shorebird mass gain dynamics in an integrated analysis (Figure 1).

We predicted that both the timing and the rate of mass gain were associated with the relative amount of horseshoe crab eggs available during the stopover period for red knot (specialist), but not ruddy turnstone (generalist). We tested for an effect of both overall spawning abundance and the relative horseshoe crab egg availability on both the year-specific inflection point and maximum rate of mass gain. We also included an interaction term between abundance and availability of horseshoe crabs, which we interpret as the relative abundance of eggs available to the birds. We used indicator variable selection to simultaneously estimate parameter values and the probability that each covariate was an important predictor. For each covariate, we estimated a binary indictor variable γ. We defined a joint "slab-and-spike" prior on coefficient (β) estimates, such that the estimate would equal 0 if γ was equal to 0, and be drawn from an uninformative prior distribution otherwise.

$$\gamma_1 \sim Bernoulli(0.5)$$

$$\beta_1 \sim \begin{cases} \gamma_1 = 0 & Normal(0, 10^{-10}) \\ \gamma_1 = 1 & Normal(0, 10^{3}) \end{cases} \quad (3)$$

The proportion of MCMC iterations in which γ = 1 is interpreted as the inclusion probability for that covariate. We report covariate effects as the posterior means (β), the proportion of the posterior distribution with the same sign as the mean (*f*), and the inclusion probability (γ).

MCMC estimation

We fit all models using R 3.5.1, JAGS 4.3, and the jagsUI package (Plummer 2003, Kellner 2015, R Core Team 2016). We generated 3 MCMC chains of 100,000 iterations each

with 10,000 iterations in the adaptive phase and discarded the first 50,000 samples as burn-in values, which was sufficient to ensure convergence for all parameters ($\hat{R} < 1.1$). JAGS model code is provided in the Supporting Information.

Results

Timing of shorebird arrival to stopover site

We found consistent timing of arrival to the stopover site for both species (Figure 2). For both species, 95% arrival generally occurred during the seventh sampling period (May 26-28), sometimes occurring as early as the 6th sampling period (May 23-25) for red knot, but no later than the eighth sampling period for both species (May 29-31). See Supporting Information for details of all Jolly-Seber model results.

Availability of horseshoe crab eggs

The amount of horseshoe crab spawn available to the birds varied among years and was dependent on water temperature. The proportion of horseshoe crab spawn that occurred by peak shorebird arrival ranged from 0.23 (2003) to 0.74 (2004), with a median of 0.58 for both shorebird species, but there was not a detectable trend in degree of mismatch between horseshoe crab spawning and shorebird arrival over the past 20 years (slope = -0.05 (-0.06, 0.03)). Water temperature was a positive predictor of the proportion of spawn available to shorebirds (β = 0.64 (0.15, 1.16) in red knot model and β = 0.60 (0.12, 1.07) in ruddy turnstone model).

Annual variation in stopover mass gain dynamics

Red knots gained mass more quickly and over a shorter time period than ruddy turnstones (Table 1, Figure S2). Scaled to the average minimum mass, the maximum rate of mass gain for

red knot was equal to an increase of 10% of minimum average body mass per day, and that of ruddy turnstone was equal to an increase of 7.4% per day. Red knot mass gain also typically happened over fewer days than ruddy turnstone; the median number of days of the peak refueling period was 6.3 (2.5, 14.5) for red knot and 9.1 (6.5, 14.9) for ruddy turnstone (Table 1, Figure S2).

As predicted, red knots exhibited more among-year variation in year-specific model parameters than ruddy turnstones, measured as the among-year coefficient of variation (CV) of estimates (Table 1, Figure 3). This difference was most pronounced for the year-specific maximum rate; this parameter varied by 81% among years for red knot (95% CRI: 29%, 220%), but only 27% for ruddy turnstone (95% CRI: 0.15%, 45%); in only 3% of MCMC iterations was the maximum slope CV for knot lower than turnstone. Both species had similarly low among-year variation in the inflection point of the mass gain curve (red knot: CV = 0.13 (0.08, 0.20), ruddy turnstone: 0.12 (0.06, 0.25)).

The average rate of mass gain was positively associated with start day for both species, indicating faster average refueling rates in years when mass gain started later in the season (Figure 4). The start day of mass gain was also positively associated with the end day but was not associated with the predicted maximum average mass, indicating that the maximum average mass is not dependent on the day on which mass gain begins.

Effect of food availability on mass gain dynamics

In cooler years with later horseshoe crab spawning, red knots gained mass at a faster rate and the inflection point occurred later in the season (Table 2). The interaction term between statewide spawning abundance (IFSA) and overlap with shorebirds, i.e. the relative availability

of horseshoe crab eggs, was the strongest predictor of red knot maximum rate of mass gain, indicating that red knots gained weight at a slower rate in years with greater relative horseshoe crab egg availability. Importantly, this interaction was almost three times as likely to be included as a predictor of mass gain in knots (0.36) than turnstones (0.13).

Discussion

We evaluated the hypothesis that the effects of ecological mismatch on stopover mass gain dynamics are dependent on foraging strategy by comparing how yearly variation in the timing of food abundance affected the timing and rate of mass gain for a specialist and generalist species. We found support for our predictions that both the timing and rate of mass gain were linked to timing of food abundance for the specialist but not the generalist forager. Additionally, red knot (the specialist) exhibited more among-year variation in the rate of mass gain, suggesting more sensitivity to local conditions. Ruddy turnstone mass gain dynamics were more consistent across years and not associated with abundance or availability of horseshoe crab spawn, which aligns with our prediction that as generalist foragers they are less sensitive to availability of this food resource.

The hierarchical modeling approach used here allowed us to evaluate year-specific mass gain dynamics while borrowing information across years, which helped compensate for uneven sampling and sparse data in some years. Similarly, the integrated analysis of the effect of water temperature on the timing of horseshoe crab spawn allowed us to impute some missing data. However, this modeling approach propagates all uncertainty through each level of the modeling hierarchy and in years with sparse shorebird mass data this can result in substantial uncertainty in model parameter estimates. Wide credible intervals around these estimates may have contributed in part to low inclusion probabilities for the covariates associated with horseshoe crab abundance

and timing of spawn. Additionally, the horseshoe crab spawning survey used here measures spawning females, which we use as a proxy for egg abundance, rather than the abundance of horseshoe crab eggs directly throughout the season. However, we still found evidence of a relationship between mass gain dynamics—especially the maximum rate of mass gain—and the relative availability of horseshoe crab eggs for red knot and a much weaker relationship for ruddy turnstone. This indicates that these two species, similar in many other aspects of their ecology, are responding differently to changes in local conditions during migration. Ruddy turnstone seemed unaffected by annual variation in horseshoe crab egg availability, likely because their generalist foraging strategy allows them to exploit a variety of prey more easily. Red knot undergo extreme physiological changes during migration, including a reduction in size of internal organs to maximize rate of fat deposition and reduce other metabolic costs (van Gils et al. 2005). Red knot consume a variety of marine invertebrates and crustaceans during other points of the annual cycle (Baker et al 2001), but these physiological constraints during the migratory period may make them especially reliant on horseshoe crab eggs, an easily-digestible and fat-rich prey.

Our estimated minimum and maximum average mass correspond with the minimum and maximum published masses for these species (red knot: 125-205 g (Baker et al. 2001), ruddy turnstone: 84-190 g (Nettleship 2000)). Birds are arriving to this site close to their minimum possible mass and, after only a few weeks, departing close to their maximum possible mass. This underscores the importance of prey availability at this stopover site. Although this analysis indicated substantial uncertainty in maximum average mass due to sparse data towards the end of the season in most years, the overall average maximum mass for red knot predicted from this analysis (178 g) agrees with predictions of optimal red knot departure mass made by Baker et al.

(2004). That study, based on red knot physiology and optimal fuel load following Kvist et al. (2001), estimated that a departure mass of 180 g is necessary for sufficient fat stores to reach the Arctic breeding grounds in a single flight. This threshold mass of 180 g has been used in conservation management planning to assess flyway population health in Delaware Bay (e.g., McGowan et al. 2011), which this study reaffirms. We note, though, that this analysis focuses on the population while others have investigated departure weight on the individual level (Baker et al. 2004, McGowan et al. 2011), so a population average departure mass of 180 g will result in some individuals below that threshold, which could lead to negative individual carry-over effects. Additionally, for this study we mostly captured shorebirds in foraging flocks, which inherently biases our sampling to those birds still actively refueling and excludes those that have reached their departure masses and are preparing to continue migration. For these reasons we are cautious to interpret estimates of minimum or maximum average mass and focus instead on the dynamics of the timing and rate of mass gain, which should be unaffected by the above concerns.

In cooler years, when the relative availability of horseshoe crab eggs was below average, red knot had a later onset of mass gain and gained mass more rapidly. This indicates some plasticity in the rate of mass gain and suggests that birds may increase the rate of fat deposition in years with later food availability to avoid delayed departure from the stopover site. Optimal migration theory predicts strong selective pressure on optimizing fuel deposition rates during migratory stopover (Hedenström and Alerstam 1997), and evidence indicates that Arctic-breeding shorebirds operate primarily under time constraints during northward migration (Zhao et al. 2017). Our observation that red knots increased mass gain rates in mismatch years, as opposed to maintaining a lower rate of mass gain but delaying departure from the stopover site, aligns with such a time-constrained strategy. In addition to a narrow time window for breeding,

shorebirds are also constrained by a fixed single clutch size (usually 4 eggs) and have little flexibility in their breeding strategy to compensate for unfavorable conditions encountered during migration. The potential impacts of climate-induced changes in phenology for breeding behavior have been well-documented (e.g., Both et al. 2005; Winkler et al. 2002), but increased frequency of mismatches, and potential associated physiological costs, may be especially detrimental to species like shorebirds with limited capacity to compensate at other points of the annual cycle.

Refueling was slower in years with greater prey abundance, suggesting that a slower mass gain is preferable and there may be a physiological cost of rapid refueling. Other studies have shown that individual mass predicts timing of departure from the stopover site and possibly breeding success, with heavier birds choosing more advantageous tailwinds and having a greater probability of being observed during the following autumn (Duijns et al. 2017). We found no evidence for population-level delays or lack of mass gain in mismatch years, but carry-over effects from mismatches may potentially be acting on individuals to reduce survival probability or breeding success. Future work investigating individual-level responses to mismatch would help in assessing potential carry-over effects, but such studies are difficult because within-season individual recaptures are typically rare in stopover systems. Alternative methods of weighing birds without physical capture may be needed, such as that used by Delingat et al. (2006).

This work also provides a framework to evaluate population responses to changes in prey phenology at other stopover sites, such as inland wetlands, which may also be vulnerable to climatic change (e.g. Steen et al. 2018). Long-distance migrants rely on predictable resources *en route* and even when these linkages are simple and predictable, populations can be vulnerable to change (Runge et al. 2014). When availability of food resources is linked to environmental

conditions, changing phenology could result in mismatches between migrant stopover and resource pulses. In this system, cooler spring temperatures were associated with delayed prey availability, which affected mass gain dynamics for the specialist, but not the generalist species using the same site. Considering species traits such as foraging ecology is likely to be important in predicting population responses and understanding species persistence under global change (Foden et al. 2013, Pacifici et al. 2017). A diverse body of research has shown that species that specialize on habitat types, prey, or host plants are more vulnerable to changing environmental conditions, leading to disproportionate population declines (Clavel et al. 2011, Davey et al. 2012) while generalist species have greater adaptive capacity to expand into and cope with novel environments (Jonsen and Fahrig 1997, Stefanescu et al. 2011). Even with closely related species that share many ecological and life history traits, differences in degree of specialization may explain differences in distribution, population trends, and status. As the myriad effects of global climate change worsen, species traits will be important in predicting which species will persist and therefore the impacts of such changes on ecological communities.

Literature cited

Baker, A., P. Gonzalez, R. I. G. Morrison, and B. A. Harrington. 2001. Red Knot (Calidris canutus). Page The Birds of North America. Cornell Lab of Ornithology, Ithaca.

Baker, A. J., P. M. González, T. Piersma, L. J. Niles, I. D. L. S. do Nascimento, P. W. Atkinson, N. a Clark, C. D. T. Minton, M. K. Peck, and G. Aarts. 2004. Rapid population decline in red knots: fitness consequences of decreased refuelling rates and late arrival in Delaware Bay. Proceedings. Biological sciences / The Royal Society 271:875–882.

Bischof, R., L. E. Loe, E. L. Meisingset, B. Zimmermann, B. Van Moorter, and A. Mysterud. 2012. A Migratory Northern Ungulate in the Pursuit of Spring: Jumping or Surfing the Green Wave? The American Naturalist 180:407–424.

Clavel, J., R. Julliard, and V. Devictor. 2011. Worldwide decline of specialist species: Toward a global functional homogenization? Frontiers in Ecology and the Environment 9:222–228.

Davey, C. M., D. E. Chamberlain, S. E. Newson, D. G. Noble, and A. Johnston. 2012. Rise of the generalists: Evidence for climate driven homogenization in avian communities. Global Ecology and Biogeography 21:568–578.

Delingat, J., V. Dierschke, H. Schmaljohann, B. Mendel, and F. Bairlein. 2006. Daily stopovers as optimal migration strategy in a long-distance migrating passerine: the Northern Wheatear *Oenanthe oenanthe*. Ardea 94:593–605.

Duijns, S., L. J. Niles, A. Dey, Y. Aubry, C. Friis, S. Koch, A. M. Anderson, and P. A. Smith. 2017. Body condition explains migratory performance of a long-distance migrant. Proceedings of the Royal Society B.

Edwards, M., A. J. Richardson, and Martin Edwards & Anthony J. Richardson. 2004. Impact of climate change on marine pelagic phenology and trophic mismatch. Nature 430:881–884.

Foden, W. B., S. H. M. Butchart, S. N. Stuart, J. C. Vié, H. R. Akçakaya, A. Angulo, L. M. DeVantier, A. Gutsche, E. Turak, L. Cao, S. D. Donner, V. Katariya, R. Bernard, R. A. Holland, A. F. Hughes, S. E. O'Hanlon, S. T. Garnett, Ç. H. Şekercioğlu, and G. M. Mace. 2013. Identifying the World's Most Climate Change Vulnerable Species: A Systematic Trait-Based Assessment of all Birds, Amphibians and Corals. PLoS ONE 8.

Gill, R. E. 1986. What won't Turnstones eat? British Birds 79:401–402.

Gillings, S., P. W. Atkinson, A. J. Baker, K. a. Bennett, N. a. Clark, K. B. Cole, P. M. González, K. S. Kalasz, C. D. T. Minton, L. J. Niles, R. C. Porter, I. D. L. Serrano, H. P. Sitters, and J. L. Woods. 2009. Staging Behavior in Red Knot (Calidris canutus) in Delaware Bay: Implications for Monitoring Mass and Population Size. The Auk 126:54–63.

Gillings, S., P. W. Atkinson, S. L. Bardsley, N. a. Clark, S. E. Love, R. a. Robinson, R. a. Stillman, and R. G. Weber. 2007. Shorebird predation of horseshoe crab eggs in Delaware

Bay: Species contrasts and availability constraints. Journal of Animal Ecology 76:503–514.

van Gils, J. A., P. F. Battley, T. Piersma, and R. Drent. 2005. Reinterpretation of gizzard sizes of red knots world-wide emphasises overriding importance of prey quality at migratory stopover sites. Proceedings of the Royal Society B: Biological Sciences 272:2609–2618.

Hedenstrom, A., and T. Alerstam. 1997. Optimum Fuel Loads in Migratory Birds: Distinguishing Between Time and Energy Minimization. Journal of Theoretical Biology:227–234.

Jonsen, I. D., and L. Fahrig. 1997. Response of generalist and specialist insect herbivores to landscape spatial structure. Landscape Ecology 12:185–197.

Kellner, K. 2015. jagsUI: a wrapper around rjags to streamline JAGS analyses. Page R package version 1.1.

Kery, M. 2010. Introduction to WinBUGS for Ecologists. Academic Press.

Klaassen, M., A. Lindström A, H. Meltofte, and T. Piersma. 2001. Arctic waders are not capital breeders. Nature 413:794.

Kokko, H. 1999. Competition for early arrival birds in migratory birds. Journal of Animal Ecology 68:940–950.

Kvist, A., Lindström A., M. Green, T. Piersma, and G. H. Visser. 2001. Carrying large fuel loads during sustained bird flight is cheaper than expected. Nature 413:730–732.

Mayr, E. 1965. The nature of colonizations in birds. Page 47 The genetics of colonizing species. Academic Press, New York, New York, USA.

McGowan, C. P., J. E. Hines, J. D. Nichols, J. E. Lyons, D. R. Smith, K. S. Kalasz, L. J. Niles, A. D. Dey, N. a. Clark, P. W. Atkinson, C. D. T. Minton, and W. Kendall. 2011. Demographic consequences of migratory stopover: linking red knot survival to horseshoe crab spawning abundance. Ecosphere 2.

Meltofte, H., H. Meltofte, T. Piersma, T. Piersma, H. Boyd, H. Boyd, B. J. McCaffery, B. J. Mccaffery, I. Tulp, and I. Tulp. 2007. Effects of climate variation on the breeding ecology of Arctic shorebirds. Page Meddelelser Om Gronland Bioscience.

Miller-Rushing, A. J., T. T. Hoye, D. W. Inouye, and E. Post. 2010. The effects of phenological mismatches on demography. Philosophical Transactions of the Royal Society B: Biological Sciences 365:3177–3186.

Nettleship, D. N. 2000. Ruddy Turnstone (Arenaria interpres). Page The Birds of North America. Cornell Lab of Ornithology, Ithaca.

Newton, I. 2006. Can conditions experienced during migration limit the population levels of birds? Journal of Ornithology 147:146–166.

Niles, L. J., H. P. Sitters, D. Dey, Amanda, P. W. Atkinson, A. J. Baker, K. A. Bennett, R. Carmona, K. E. Clark, N. A. Clark, C. Espoz, P. M. Gonzalez, B. A. Harrington, D. E. Hernandez, K. S. Kalasz, R. G. Lathrop, N. Matus, Ricardo, C. D. T. Minton, R. I. G. Morrison, M. K. Peck, W. Pitts, R. A. Robinson, and I. L. Serrano. 2008. Status of the Red Knot (Calidris Canutus Rufa) in the Western Hemisphere (Calidris Canutus Rufa) in. Studies in Avian Biology 36:1–185.

Nilsson, C., R. H. G. Klaassen, and T. Alerstam. 2013. Differences in Speed and Duration of Bird Migration between Spring and Autumn. The American Naturalist 181:837–845.

Pacifici, M., P. Visconti, S. H. M. Butchart, J. E. M. Watson, F. M. Cassola, and C. Rondinini. 2017. Species' traits influenced their response to recent climate change. Nature Climate Change 7:205–208.

Pearce-Higgins, J. W., D. J. Brown, D. J. T. Douglas, J. A. Alves, M. Bellio, P. Bocher, G. M. Buchanan, R. P. Clay, J. Conklin, N. Crockford, P. Dann, J. Elts, C. Friis, R. A. Fuller, J. A. Gill, K. Gosbell, J. A. Johnson, R. Marquez-Ferrando, J. A. Masero, D. S. Melville, S. Millington, C. Minton, T. Mundkur, E. Nol, H. Pehlak, T. Piersma, F. Robin, D. I. Rogers, D. R. Ruthrauff, N. R. Senner, J. N. Shah, R. D. Sheldon, S. A. Soloviev, P. S. Tomkovich, and Y. I. Verkuil. 2017. A global threats overview for Numeniini populations: Synthesising expert knowledge for a group of declining migratory birds. Bird Conservation International 27:6–34.

Petanidou, T., A. S. Kallimanis, S. P. Sgardelis, A. D. Mazaris, J. D. Pantis, and N. M. Waser. 2014. Variable flowering phenology and pollinator use in a community suggest future phenological mismatch. Acta Oecologica 59:104–111.

Piersma, T., T. Lok, Y. Chen, C. J. Hassell, H.-Y. Yang, A. Boyle, M. Slaymaker, Y.-C. Chan, D. S. Melville, Z.-W. Zhang, and Z. Ma. 2014. Simultaneous declines in summer survival of three shorebird species signals a flyway at risk. Journal Of Applied Ecology:manuscript.

Plard, F., J. M. Gaillard, T. Coulson, A. J. M. Hewison, D. Delorme, C. Warnant, and C. Bonenfant. 2014. Mismatch Between Birth Date and Vegetation Phenology Slows the Demography of Roe Deer. PLoS Biology 12:1–8.

Plummer, M. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003):20–22.

Post, E., C. Pedersen, C. C. Wilmers, and M. C. Forchhammer. 2008. Warming, plant phenology and the spatial dimension of trophic mismatch for large herbivores. Proceedings of the Royal Society B: Biological Sciences 275:2005–2013.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Robinson, R. A., H. Q. P. Crick, J. A. Learmonth, I. M. D. Maclean, C. D. Thomas, F. Bairlein, M. C. Forchhammer, C. M. Francis, J. A. Gill, B. J. Godley, J. Harwood, G. C. Hays, B.

Huntley, A. M. Hutson, G. J. Pierce, M. M. Rehfisch, D. W. Sims, M. Bego??a Santos, T. H. Sparks, D. A. Stroud, and M. E. Visser. 2009. Travelling through a warming world: Climate change and migratory species. Endangered Species Research 7:87–99.

Runge, C. A., T. G. Martin, H. P. Possingham, S. G. Willis, and R. A. Fuller. 2014. Conserving mobile species. Frontiers in Ecology and the Environment 12:395–402.

Senner, N. R., M. Stage, and B. K. Sandercock. 2016. Ecological mismatches are moderated by local conditions for two populations of a long-distance migratory bird. Oikos 000.

Si, Y., Q. Xin, W. F. De Boer, P. Gong, R. C. Ydenberg, and H. H. T. Prins. 2015. Do Arctic breeding geese track or overtake a green wave during spring migration? Scientific Reports 5:1–6.

Smith, D. R., and S. F. Michels. 2006. Seeing the elephant: Importance of spatial and temporal coverage in a large-scale volunteer-based program to monitor horseshoe crabs. Fisheries 31:485–491.

Smith, D. R., P. S. Pooler, B. L. Swan, S. F. Michels, W. R. Hall, P. J. Himchak, and M. J. Millard. 2002. Spatial and temporal distribution of horseshoe crab (Limulus polyphemus) spawning in Delaware Bay: Implications for monitoring. Estuaries 25:115–125.

Smith, P. A., H. G. Gilchrist, M. R. Forbes, J. L. Martin, and K. Allard. 2010. Inter-annual variation in the breeding chronology of arctic shorebirds: Effects of weather, snow melt and predators. Journal of Avian Biology 41:292–304.

Sol, D., S. Timmermans, and L. Lefebvre. 2002. Behavioural flexibility and invasion success in birds. Animal Behaviour 63:495–502.

Steen, V., S. K. Skagen, and B. R. Noon. 2018. Preparing for an uncertain future: Migrating shorebird response to past climatic fluctuations in the Prairie Potholes: Migrating. Ecosphere 9.

Stefanescu, C., J. Carnicer, and J. Pen. 2011. Determinants of species richness in generalist and specialist Mediterranean butterflies : the negative synergistic forces of climate and habitat change. Ecography 34:353–363.

Stenseth, N. C., and A. Mysterud. 2002. Climate, changing phenology, and other life history traits: Nonlinearity and match-mismatch to the environment. Proceedings of the National Academy of Sciences 99:13379–13381.

Studds, C. E., B. E. Kendall, N. J. Murray, H. B. Wilson, D. I. Rogers, R. S. Clemens, K. Gosbell, C. J. Hassell, R. Jessop, D. S. Melville, D. A. Milton, C. D. T. Minton, H. P. Possingham, A. C. Riegen, P. Straw, E. J. Woehler, and R. A. Fuller. 2017. Rapid population decline in migratory shorebirds relying on Yellow Sea tidal mudflats as stopover sites. Nature Communications 8:1–7.

Thorup, K., A. P. Tøttrup, M. Willemoes, R. H. G. Klaassen, R. Strandberg, M. L. Vega, H. P.

Dasari, and M. B. Araújo. 2017. Resource tracking within and across continents in long-distance bird migrants. Science Advances 3.

Tsipoura, N., and J. Burger. 1999. Shorebird Diet During Spring Migration Stopover on Delaware Bay. The Condor 101:635–644.

Walther, G. R., E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. C. Beebee, J. M. Fromentin, O. Hoegh-Guldberg, and F. Bairlein. 2002. Ecological responses to recent climate change. Nature 416:389–395.

Wauchope, H. S., J. D. Shaw, Ø. Varpe, E. G. Lappo, D. Boertmann, R. B. Lanctot, and R. A. Fuller. 2017. Rapid climate-driven loss of breeding habitat for Arctic migratory birds. Global Change Biology 23:1085–1094.

Winkler, D. W., P. O. Dunn, and C. E. McCulloch. 2002. Predicting the effects of climate change on avian life-history traits. Proceedings of the National Academy of Sciences of the United States of America 99:13595–9.

Zhao, M., M. Christie, J. Coleman, C. Hassell, K. Gosbell, S. Lisovski, C. Minton, and M. Klaassen. 2017. Time versus energy minimization migration strategy varies with body size and season in long-distance migratory shorebirds. Movement Ecology 5:23.

Tables

Table 3-1: Mass gain model parameter estimates.
Model estimates for global average and among-year coefficient of variation (CV) in mass gain model parameters, reported as posterior mean and lower and upper bounds of the 95% credible interval (CRI).

Species	Model parameter	Among-year average			Among-year CV		
		Mean	LCRI	UCRI	Mean	LCRI	UCRI
Red knot	Minimum average mass (g)	116	109	123	0.097	0.049	0.161
	Maximum average mass (g)	179	170	189	0.068	0.026	0.115
	Inflection point (day)	22.1	20.8	23.6	0.127	0.078	0.195
	Maximum slope (g/day)	11.7	4.0	19.1	0.811	0.290	2.182
	Start day	18.7	12.7	20.9	--	--	--
	End day	24.4	22.6	26.9	--	--	--
	Average rate of mass gain	6.9	2.9	11.0	--	--	--
Ruddy turnstone	Minimum average mass (g)	98.9	96.3	101.6	0.032	0.008	0.064
	Maximum average mass (g)	160.7	153.8	168.7	0.034	0.002	0.080
	Inflection point (day)	24.0	23.0	25.1	0.115	0.055	0.251
	Maximum slope (g/day)	7.3	5.3	9.8	0.271	0.148	0.446
	Start day	18.7	16.6	20.3	--	--	--
	End day	27.3	25.9	28.9	--	--	--
	Average rate of mass gain	4.1	3.2	5.2	--	--	--

Table 3-2: Effects of horseshoe crab spawning on mass gain model parameters.
Parameter estimates are reported as the means and 95% credible intervals (LCRI, UCRI) of the posterior distributions. Inclusion probability (γ) is the proportion of MCMC iterations in which that covariate was included as a predictor. *f* is the proportion of the posterior distribution that is the same sign as the posterior mean.

Species	Model parameter	Covariate	Mean	LCRI	UCRI	γ	*f*
Red knot	Maximum slope	abundance	3.4	-9.6	15	0.21	0.72
	Maximum slope	timing	-4.1	-15	6.1	0.18	0.80
	Maximum slope	abundance x timing	-8.1	-22	3.6	0.36	0.92
	Inflection point	abundance	-0.35	-2.2	1.5	0.03	0.66
	Inflection point	timing	-1.4	-2.7	0.01	0.14	0.97
	Inflection point	abundance x timing	0.82	-1.2	2.9	0.04	0.80
Ruddy turnstone	Maximum slope	abundance	0.83	-1.7	3.9	0.13	0.75
	Maximum slope	timing	0.34	-1.9	2.9	0.09	0.63
	Maximum slope	abundance x timing	-0.69	-4.3	2.5	0.13	0.68
	Inflection point	abundance	-0.06	-1.3	1.2	0.06	0.55
	Inflection point	timing	-0.52	-1.7	0.63	0.08	0.83
	Inflection point	abundance x timing	0.38	-1.2	1.9	0.08	0.70

Figures

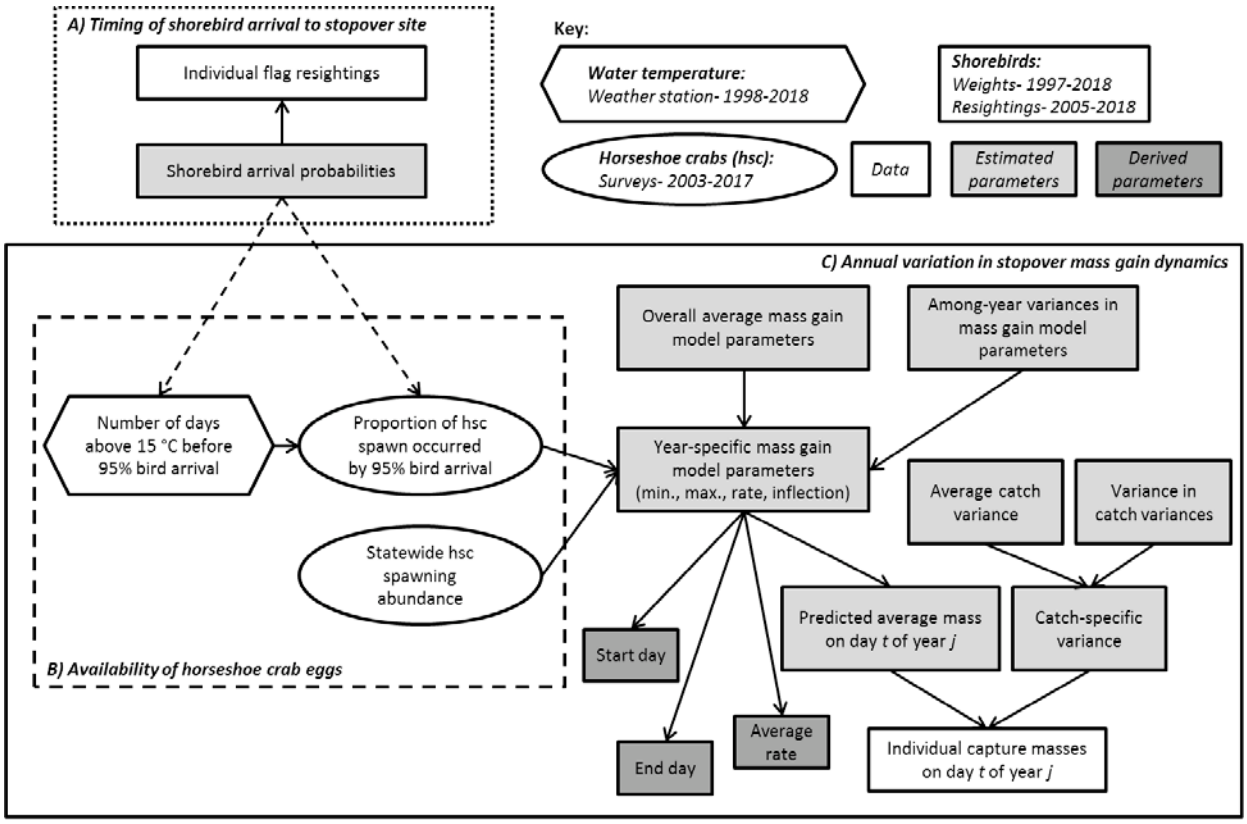

Figure 3-1: Hierarchical mass gain model structure.
This model was used to estimate the relationship between horseshoe crab egg availability and shorebird mass gain dynamics. Shorebird arrival probabilities were estimated using Jolly-Seber models (**A**). The relationship between water temperature and horseshoe crab spawn (**B**) and parameters governing shorebird mass gain dynamics (**C**) were estimated simultaneously in an integrated analysis.

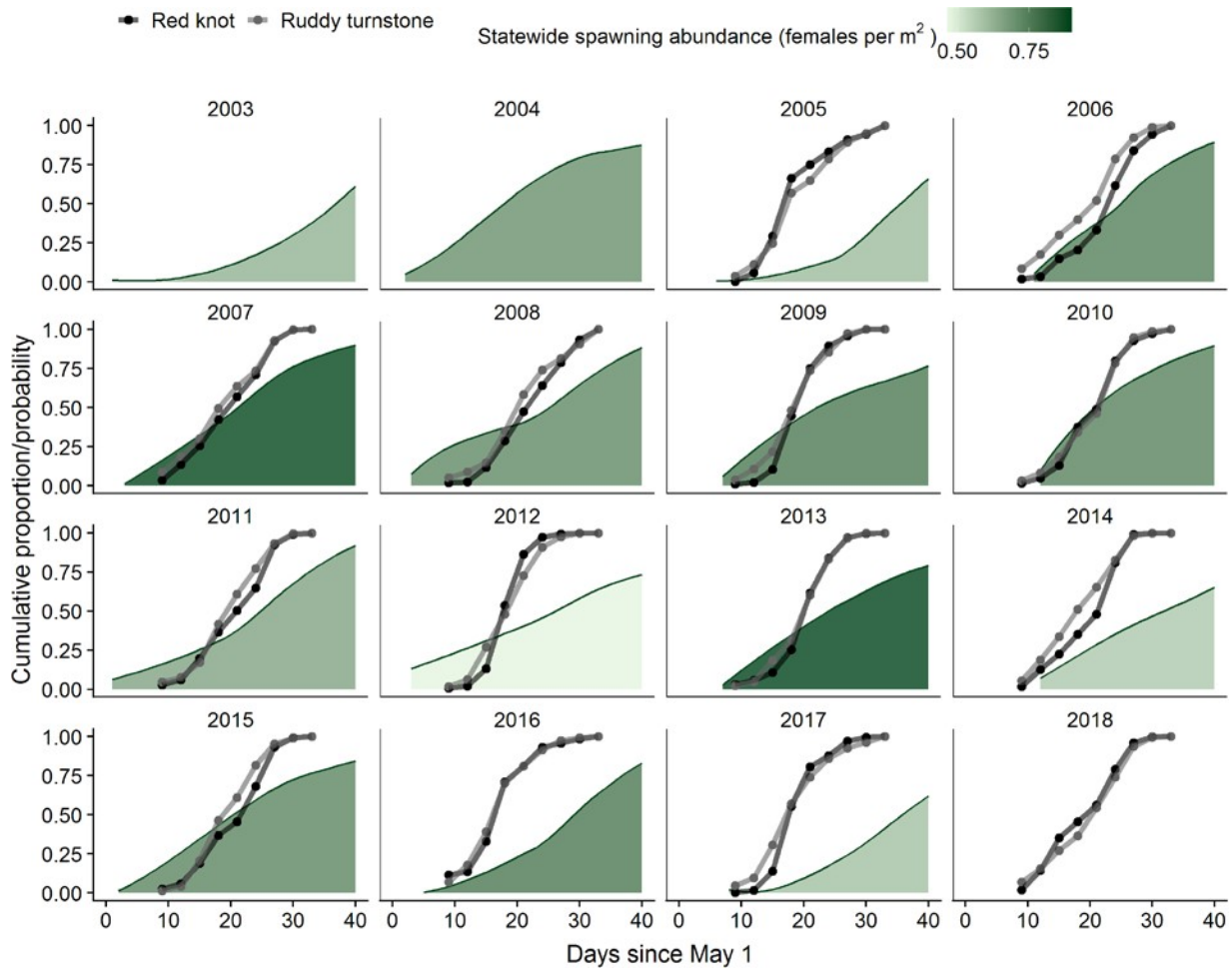

Figure 3-2; Timing of horseshoe crab spawn and shorebird arrivals.
Green lines show the cumulative proportion of horseshoe crab spawning activity that occurred by each day. The shade of green represents the total spawning abundance (IFSA) for that year. Black and gray lines indicate the estimated cumulative proportion of flagged birds arrived by each day (gray = ruddy turnstone, black = red knot).

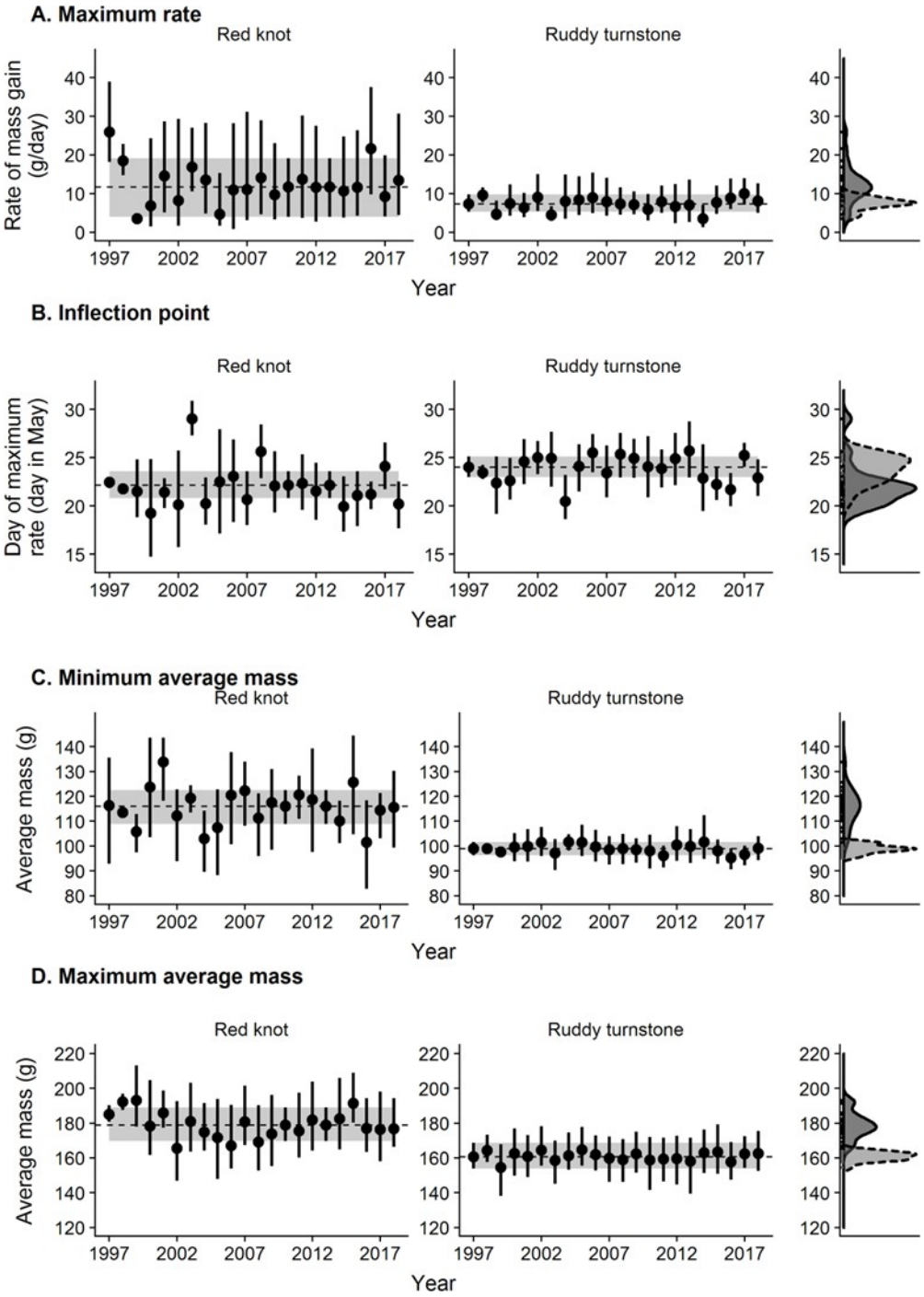

Figure 3-3: Annual variation in mass gain of red knot and ruddy turnstone.
Points and vertical lines show the posterior mean and 95% credible interval for each model parameter in each year. Dashed horizontal lines and shaded region represent the global average parameter value and its 95% credible interval. Kernel densities of estimated annual means are shown on the right (ruddy turnstone: dashed line, light gray, red knot: solid line, dark gray.

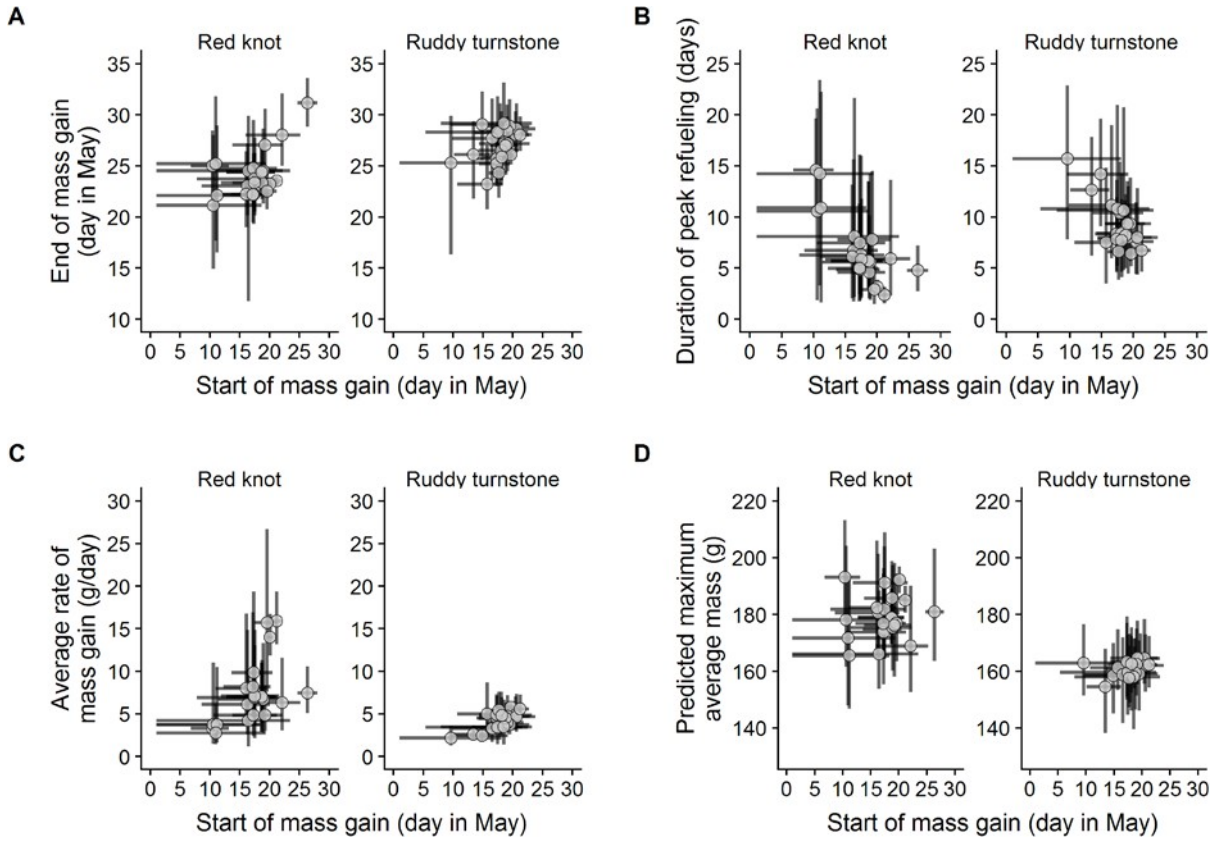

Figure 3-4: Association between timing and rate of mass gain.
Relationships between the day on which peak refueling began and the day on which it ended (**A**), the duration of the peak refueling period (**B**), the average rate of mass gain (**C**), and the predicted maximum average mass (**D**).

Chapter 4: Annual variation in use of a spring stopover site by three migratory shorebirds

Abstract

Understanding how resource availability during migration influences both annual demographic rates and migratory behaviors is critical for interpreting changes in passage population sizes at stopover sites and predicting population responses to future changes and conservation. We used an open robust design model to estimate within-year and between-year parameters related to survival and stopover site use, including temporary emigration and transience, for migratory birds during stopover and test for an effect of stopover site food availability on survival probability and the probability of returning in the next year. We analyzed mark-resight observations of red knot (*Calidris canutus rufa*), ruddy turnstone (*Arenaria interpres*), and sanderling (*Calidris alba*) using Delaware Bay during spring stopover from 2005-2018. For ruddy turnstone, the probability of returning to this site if present in the previous year was positively associated with the availability of a key prey item in the previous year, but annual survival probability was not associated with prey availability. For red knot, however, both apparent annual survival probability and the probability of returning to the site were associated with prey availability. Shorebird use of this site typically peaked during May 26-28, but the proportion of the population present during this period varied dramatically for red knot (range: 0.07-0.59) but less so for ruddy turnstone (0.24-0.43) and sanderling (0.18-0.54). This demonstrates that conditions at this site have species-specific effects on demographics and

migratory behavior, both of which should be considered in analyzing and interpreting monitoring data.

Introduction

Animals that perform long-distance seasonal migrations congregate at intermediate sites along the migratory route that provide annually predictable, abundant, and aggregated food resources (Newton 2006, Alerstam 2011). Such stopover sites are used by many migratory species including birds (Amano et al. 2010), butterflies (Gibbs et al. 2006), whales (Franklin et al. 2018), and bats (Mcguire et al. 2012). Many migratory routes consist of relatively connected habitats of varying quality (Sawyer et al. 2009, Albanese et al. 2012) however, some sites are especially important because they are the first available habitat after an ecological barrier (Moore et al. 1990, Buler and Moore 2011) or because they provide especially abundant and aggregated food resources (Clark et al. 1993). If conditions at key sites deteriorate, the ability of individuals to shift migratory schedules or use different stopover sites (given that other suitable sites exist) could be critical to population persistence (Peng et al. 2014, Piersma et al. 2014, Studds et al. 2017).

Migratory populations are declining globally (Wilcove and Wikelski 2008), and monitoring at stopover sites is important for understanding population responses to past events and predicting responses to future change or conservation efforts. Climate change and other human disturbances may alter both the overall abundance and phenology of prey availability along the migratory route (Thorup et al. 2017), and long-distance migrants may be particularly susceptible to negative effects of such changes if cues from distant nonbreeding sites become asynchronous with resource availability at stopover and breeding sites (Gwinner 1996, Robinson

et al. 2009, Both et al. 2010). Studies using remote tracking devices have shown variation in individual migratory behaviors for birds (Minton et al. 2011, Burger et al. 2012, Newstead et al. 2013, Verhoeven et al. 2019), suggesting some capacity for flexibility in migratory behaviors that could facilitate shifts in the proportion of the flyway population using a stopover site from year to year. The ability to exhibit migratory flexibility could be vital to species viability in response to changes in resource availability at stopover sites.

Shifts in migratory schedules or site use may be difficult to distinguish from changes in population size due to the difficulty of monitoring such wide-ranging species (Bart et al. 2007). For at least one species (semipalmated sandpiper *Calidris pusilla*) negative population trends measured during migratory stopover are contradicted by stable trends from the breeding grounds (Smith et al. 2012). In a given year, some proportion of the population may skip a site or pass through as transients instead of remaining to refuel (Schaub et al. 2004), and these processes of temporary emigration and stopover residency could vary due to conditions along the migratory route as well as prey availability and predation pressure at the stopover site (Ydenberg et al. 2004, Jonker et al. 2010). These factors may be important for understanding both changes in flyway population size and use of stopover sites in the flyway, as changing use of a stopover site could be misinterpreted as declining flyway population size (Bart et al. 2007).

The predictable annual aggregations of individuals at migratory stopover sites, sometimes called “staging areas,” make them useful for monitoring the population status of migratory species. Data collected during migratory stopover have contributed to decision-making in contexts including habitat protection (Cimprich et al. 2006, Xia et al. 2017), fisheries harvest limits (McGowan et al. 2015, Staton et al. 2017), and agriculture practices (Green et al. 2015, Lei

et al. 2018). Such decision processes depend on accurate assessments of both within-season and across-season population trends and site use.

The use of remote tracking devices such as light-level archival geolocators, satellite tags, GPS tags, and VHF transmitters has contributed greatly to our understanding of migratory movements and site use (Kranstauber et al. 2011, Bridge et al. 2013, Kays et al. 2015). However, the number of individuals tracked is often limited by cost of the devices and, for some, the ability to re-trap an individual to recover data stored on the tag (Bridge et al. 2013). Additionally, many existing tags have a limited lifespan due to battery capacity or attachment method. Although rapid advances in these technologies are reducing tag size and cost while increasing lifespan and data storage capacity (Bridge et al. 2011), use of cheaper methods to monitor individuals facilitates data collection for more individuals across longer timeframes.

Capture-recapture models provide a suite of tools for estimating demographic rates based on repeated encounters of individually-marked animals (Williams et al. 2002). More complex model extensions allow for the estimation of the probabilities of temporary emigration and transience, although the latter is often treated as a nuisance parameter (Kendall et al. 1997, Schaub et al. 2004). Several previous analyses of capture-recapture data collected during migratory stopover have focused either on estimating within-season dynamics for a single year (e.g., Schaub et al. 2004, Giavi et al. 2014, Lyons et al. 2015) or on estimating annual survival probabilities or return rates across several years (e.g., McGowan et al. 2011). Here, we present a modeling framework to simultaneously estimate both within-year and across-year dynamics of stopover site use (Kendall et al. 1997, Schwarz and Stobo 1997). We use this model to compare the responses of three long-distance migratory shorebirds to conditions in Delaware Bay, a spring stopover site that has received significant research and conservation attention in recent

years (e.g., Niles et al. 2009, McGowan et al. 2015), and we evaluate whether stopover conditions influenced survival probability or the probability of site use for three long-distance migrants. To understand the role of stopover sites in limiting population growth rate of migratory birds, it is important to consider both demographic and behavioral effects of stopover site conditions.

Methods

Study site and focal species

The Delaware Bay is a globally-important spring stopover site for several species of shorebirds due to the aggregation of spawning horseshoe crabs (*Limulus polyphemus*), the eggs of which provide a fat-rich and easily-digestible prey for shorebirds (Tsipoura and Burger 1999). An ongoing monitoring program was established in the late 1990s to track population status of migratory shorebirds and inform harvest regulations for horseshoe crabs (McGowan et al. 2015). We used monitoring data from three species during spring migration in Delaware Bay: red knot (*Calidris canutus rufa*), ruddy turnstone (*Arenaria interpres*), and sanderling (*Calidris alba*). All three of the focal species breed in the Arctic tundra and spend the nonbreeding season at sites along the coasts of North America, the Caribbean, and South America (Nettleship 2000, Baker et al. 2001, Macwhirter et al. 2002). Birds were captured during the spring stopover period between May 5 and June 5 in Delaware, USA from 2005-2018 and marked with Incoloy USGS leg bands and Darvic leg flags inscribed with unique field-readable alphanumeric codes (Clark et al. 2005). Beaches in Delaware Bay were surveyed by trained observers throughout the stopover season to re-sight individually flagged birds.

Open robust design model

Mark-recapture models use repeated encounters of individually-marked animals to estimate demographic parameters while accounting for imperfect detection (Williams et al. 2002). The robust design model (Pollock 1981) was developed to produced unbiased estimates of temporary emigration, which is often confounded with detection probability. The classic robust design model uses repeated sampling (secondary occasions) of marked individuals within a closed period (primary occasions) to estimate detection probability, allowing for the estimation of nonrandom temporary emigration from reencounters across primary periods. During migratory stopover, however, the assumption of closure is unreasonable, as individuals may arrive or depart at any point. The open robust design model allows for arrival and departure between secondary occasions within a year (Schwarz and Stobo 1997, Kendall and Bjorkland 2001). We developed an open robust design model for a migratory stopover system formulated using the m-array multinomial likelihood (Williams et al. 2002) and written in the BUGS model language. Use of the multinomial instead of an individual state-space formulation both increases computational efficiency and facilitates goodness-of-fit tests based on expected cell probabilities. The model is described below, and all parameters defined in Table 1.

Across years, we estimated apparent annual survival and the probability of temporary emigration while accounting for imperfect detection. We assumed that individuals only entered and departed from the stopover site once in a given year (primary periods), but that those entries or departures could occur during any secondary sampling period with in a year. For the analyses described in this paper we divided the stopover period into nine 3-day secondary sampling occasions following the shorebird monitoring data collection protocol, but the model structure can accommodate any number of primary or secondary periods. The likelihood for this model is

a product of three processes. Across k years, the likelihood L_1 describes the process of re-encountering individuals that were previously encountered ("released") in year *j* (R_j), and depends on the number of individuals last seen during year *j* and re-encountered in a given year *h* (m_{hj}) the apparent annual survival probability of year *j* (ϕ_j), the probability of being present at the stopover site if present in the previous year (γ_j^{II}), the probability of being present at the stopover site if not present in the previous year (γ_j^{OI}), and the probability of detecting an individual at least once during the year, if present (p_j^*). We note that other studies have used the notation γ_j'' and γ_j' where we use γ_j^{II} and γ_j^{OI}; in our notation the superscripts *I* and *O* indicate the state of being "in" or "out" of the stopover population, respectively.

$$L_1 = P\left(\left(\{m_{hj}\}\,|\{R_j\},\{\phi_j\},\{\gamma_j^{II}\},\{\gamma_j^{OI}\},\{p_j^*\}\right)\right)$$

Within years, the likelihood L_2 describes the process of newly encountering individuals in each year during one of l secondary occasions. In most Jolly-Seber models this likelihood describes new captures of unmarked individuals, but in our formulation this likelihood describes the first encounter of each marked bird in each year. The parameter ω_{jt} is the probability of entering the study area in year *j* in a secondary period before *t*, not being detected, and remaining in the study area until *t*.

$$\omega_{jt} = \begin{cases} t = 1 & \delta_{jt} \\ t > 1 & \omega_{jt-1}(1 - p_{jt-1})\psi_{jt-1}\tau_j + \delta_{jt} \end{cases}$$

The number of individuals newly-encountered in each secondary period (u_{jt}) depends on the probability of entering the study area during each secondary period (δ_{jt}), the period-specific detection probability (p_{jt}), the probability of entering in a previous period but not being detected ($_{\omega_{jt}}$), and the total number of individuals encountered in that year ($_{R_j}$).

$$L_2 = P\left(\left(\{u_{jt}\} \mid \{R_j\}, \{\delta_{jt}\}, \{\omega_{jt}\}, \{p_{jt}\}\right)\right)$$

The third likelihood describes the process of re-encountering individuals within a year. The number of individuals encountered in each secondary period *t* that have been previously encountered in secondary period *h* of year *j* (n_{jht}) depends on the year-specific stopover residency probability (τ_j), year- and period-specific stopover persistence probability (ψ_{jt}), year- and period-specific detection probability (p_{jt}), and the number of individuals encountered in the secondary sampling occasion *h* (r_{jh}).

$$L_3 = P\left(\left(\{n_{jht}\} \mid \{r_{jh}\}, \{\tau_j\}, \{\psi_{jt}\}, \{p_{jt}\}\right)\right)$$

The joint likelihood is a product of these three. All model code is available online at https://github.com/annamtucker/open-robust-design. We followed the construction of m-arrays as described by Kendall and Bjorkland (2001) (details in Supporting Information).

Within-year dynamics (secondary periods)

Transients are common at migratory stopover sites and failing to account for the presence of transients could negatively bias estimates of within-season stopover persistence (Schaub et al. 2004). We defined transients as individuals that were present in the study area for only a single secondary period and estimated an additional parameter, τ_j, the probability of remaining at the stopover site for at least two sampling periods in year *j*. To estimate the probability of transience we treated transients and residents as two age classes and created separate secondary period m-arrays for each group, one containing all first encounters and, if applicable, second encounters, and the second m-array containing all subsequent re-encounters within that year (Giavi et al. 2014). We estimated τ_j and ψ_{jt} directly and derived the stopover persistence probability for the

first m-array, $\psi_{jt}^{T} = \psi_{jt}\tau_j$. ψ_{jt}^{T} is therefore the probability of remaining at the stopover site after the first encounter (the product of persistence and residency), and ψ_{jt} is the probability of remaining at the stopover site after all subsequent encounters.

Arrival probabilities for each secondary occasion in each year were estimated following an uninformative Dirichlet prior distribution (Table 1). We estimated an overall average detection probability and allowed for random variation in each secondary period of each year. We estimated stopover persistence probability (ψ_{jt}) as a fixed effect of year and secondary period.

Between-year dynamics (primary periods)

We estimated apparent annual survival (ϕ_j), temporary emigration probabilities (γ_j^{II} and γ_j^{OI}), and stopover residency probability (τ_j) as a random effect of year to estimate temporal variability in realized rates and the effects of stopover site conditions. We included covariates related to prey availability and weather conditions encountered during stopover in the previous year. The relative availability of horseshoe crab eggs (a key shorebird prey during stopover at this site (Fraser et al. 2010)) depends on both the overall spawning abundance and the timing of horseshoe crab spawn, which is related to water temperature (Chapter 3). For each year, we estimated the horseshoe crab egg availability (hsc) as the product of the overall spawning abundance for that year (Zimmerman et al 2016) and the proportion of spawn that occurred by peak shorebird arrival. Previous analyses have shown that the dynamics of red knot mass gain during stopover is influenced by the availability of horseshoe crab eggs, with later and faster mass gain occurring in years when horseshoe crab spawning is delayed (Chapter 3). We were interested in determining whether there was physiological cost to this rapid refueling, and so also

included rate of mass gain as a covariate for red knot survival probability. We also estimated the effects of stochastic weather events on survival probability. During extreme weather events birds are unable to forage, horseshoe crab spawning is stalled by high wave activity, and egg deposits already present on beaches may be washed away (Nordstrom et al. 2006, Smith et al. 2011). We obtained storm data from the National Oceanic and Atmospheric Administration Severe Weather Inventory and Storm Events Database (https://www.ncdc.noaa.gov/stormevents/) and determined whether a severe event occurred in Delaware Bay during the stopover period (May 1-30) in each year. Severe weather events (as defined by NOAA) included coastal flood, hail, heavy rain, high surf, storm surge, or one of three high wind categories. We used a binary covariate for each year to indicate whether or not a severe weather event occurred during the stopover period. All covariates were scaled and centered for analysis.

To facilitate comparisons of environmental effects among species when including different covariates, we estimated covariate effects using standardized linear regression models (Mood 2010, Woodworth et al. 2017). For each draw from the posterior distribution, we standardized estimates of survival, site fidelity, and residency probabilities by subtracting the mean and dividing by the standard deviation and fit normal linear regression models with the standardized probabilities as the response variables. The standardized slope coefficients associated with each covariate are the effect of a 1 standard deviation change in the independent variable in units of standard deviation of the response variable. For each MCMC iteration we extracted the estimated slope coefficients and present these estimates as the posterior mean and the proportion of the posterior distribution that is the same sign as the mean (*f*). The models for each model parameter were:

$$\phi_j \sim hsc_j + storm_j \quad or \quad \phi_j \sim rate_j + storm_j$$

$$\gamma_j^{II} \sim hsc_j$$

$$\tau_j \sim hsc_j + water\ temp_j$$

Derived parameters

From the parameters described above we derived an estimate of the flyway population-wide probability of being present at the stopover site in each secondary occasion of each year. This parameter depends on the annual temporary emigration probabilities (γ_j^{II} and γ_j^{OI}), year-specific stopover residency probability (τ_j), year- and period-specific entry probabilities (δ_{jt}), and year- and period-specific stopover persistence probabilities (ψ_{jt}). The probability of being present in secondary period *t* given that an individual is available and a resident in year *j* is denoted a_{jt} and is a function of the probabilities of entry and persistence in that year.

$$a_{jt} = \begin{cases} t = 1 & \delta_{jt} \\ t > 1 & \delta_{jt} + a_{jt-1}\psi_{jt-1} \end{cases}$$

To account for transients that do not remain after the first occasion, the stopover residency probability is included for each period before the current period *t*, which can be calculated as:

$$z_{jt} = \tau_j(a_{jt} - \delta_{jt}) + \delta_{jt}$$

In each year, the probability of being available in year *j* is the sum of the probabilities of being present given that an individual was unavailable (γ_j^{OI}) or available (γ_j^{II}) in the previous year. The probability of being present in secondary occasion *t* of year *j* is therefore:

$$\pi_{jt} = \gamma_j^{II} z_{jt} + \gamma_j^{OI} z_{jt}$$

We can also estimate the probability of departure (d_{jt}) during each sampling occasion based on the probabilities of arrival and persistence:

$$d_{jt} = z_{jt}(1 - \psi_{jt})$$

Parameter identifiability and model assumptions

Some model parameters are not identifiable in the fully time-varying formulation of the open robust design model (Kendall and Bjorkland 2001). The entry probabilities for the first secondary occasions (δ_{j1} and δ_{j2}) are confounded with detection probability in the first occasion (p_{j1}) because new entries cannot be distinguished from newly-detected individuals, so we constrained $p_{j1} = p_{j2}$ to allow estimation of both entry probabilities. Similarly, persistence probability for the second to last occasion (ψ_{j1-1}) is confounded with entry and detection probabilities for the last occasion. Departures before the last period are not distinguishable from missed detections, and without information from subsequent periods the model cannot distinguish new entries from previous entries that were not detected. We constrained $p_{j1-1} = p_{j1}$ to allow estimation of entry and persistence for the second to last occasion.

There are also constraints on estimation of the temporary emigration parameters, γ_j^{II} and γ_j^{OI}. These parameters are not identifiable for the first year, because no animals were released before year 1, or the last year, because they are confounded with apparent survival probability, i.e., because there are no subsequent years of the study we cannot distinguish between deaths and temporary emigrants. Additionally, γ_2^{OI} is not identifiable because there are no marked animals present in year 2 that were not present in year 1. We constrained these parameters ($\gamma_1^{OI} = \gamma_2^{OI} = \mu^{\gamma^{OI}}$, $\gamma_k^{OI} = \gamma_{k-1}^{OI}$, $\gamma_1^{II} = \mu^{\gamma^{II}}$, and $\gamma_k^{II} = \gamma_{k-1}^{II}$,) and only present results from years with estimable parameters.

In this formulation of the open robust design model, we assume that within years all entries occur just prior to each sampling occasion and that all individuals that enter the site are available for detection. Alternatively, the model could be constrained to allow for uniform

entries, such that individuals could arrive and depart without being available for detection as described by Schwarz and Stobo (1997), but that formulation requires the assumption that persistence probabilities (Ψ_{jt}) are constant for all sampling occasions in year *j*. We also assume that any temporary emigration from the sampled area within years occurs randomly, and that there is no individual heterogeneity in estimated parameters. This model has the typical assumptions of mark recapture models that individuals are correctly identified, tags are not lost, and individuals are independent with respect to deaths, emigration, and detection. We additionally assume that no deaths occur during secondary sampling occasions and that survival probability is the same for all individuals encountered in a given year.

MCMC estimation and goodness of fit

For the primary period likelihood (L1) we pooled capture histories of birds first captured and marked in Delaware and all subsequent resightings; the first capture in these cases was physical capture and all subsequent encounters were resightings in the field. To estimate within-year processes (secondary period likelihoods L2 and L3), we only used resightings of individuals after the year of first capture, since birds are unobservable via resighting before first capture (before they are marked with a field-readable tag). Therefore, within-year dynamics were not estimable for the first year of monitoring (2005). Primary period m-arrays for each species are provided in the Supporting Information.

There is not a unified goodness-of-fit test for the open robust design model, so we assessed fit of the time-dependent Cormack-Jolly-Seber (CJS) model for the primary period and secondary period m-arrays separately using the R2ucare package for R (R Core Team 2016, Gimenez et al. 2018). Across primary periods (years), this procedure indicated evidence of both

transience and trap-dependence for red knot and ruddy turnstone, and trap-dependence alone for sanderling. The test components 2CT and 2CL are typically considered a test of trap-dependence, but could also result in a significant test statistic due to nonrandom temporary emigration (Choquet et al. 2009), in which the probability of being available for detection in period *t* depends on the availability of the individual in period *t-1*. For red knot and ruddy turnstone, we fit a version model that accounted for transients across years by splitting the primary period m-arrays as described above, however estimates of transience probability were extremely low (~0) and so we did not include this extra parameter in the models presented here.

We also assessed fit of the CJS model to the secondary period re-encounters within each year for each species, which indicated evidence of transience and trap-dependence within most years for red knot and ruddy turnstone but not for sanderling. Although re-encounters are field resightings and not physical captures, trap-dependence could arise in our data if individuals tend to remain at or return to feeding sites that receive high observer effort. In Delaware Bay, a small sheltered harbor has been created at the mouth of the Mispillion River, creating several sandy foraging beaches with low wave activity and high horseshoe crab spawning activity. This site consistently has the largest foraging flocks and receives the most observer effort, which could create a type of trap-dependence in observations. However, initial analyses indicated little difference in detection probabilities between birds that were or were not detected in the previous occasion, and so we did not model trap-dependence within years.

We assessed fit of the model as specified separately for each of the three likelihoods for each species using posterior predictive checks. For each MCMC iteration, we generated replicate m-arrays based on the parameter estimates and calculated the discrepancies between both the observed data and these replicated datasets and expected cell frequencies using the Freeman-

Tukey statistic (Conn et al. 2018). Graphical comparisons of observed and ideal data sets are provided in the Supporting Information.

The open robust design model was fit separately for each species using R 3.5.1 and Nimble 0.7.0 (R Core Team 2016, de Valpine et al. 2017). We generated 250,000 iterations of 3 chains, discarded the first 150,000 as burn-in values, and kept every 10th sample, for a total of 10,000 samples from the posterior distribution which was sufficient to ensure convergence for all parameters (all $\hat{R} < 1.1$ and confirmed with visual inspection of traceplots).

Results

Within-year dynamics

Estimates of within-year probabilities of arrival and persistence at the stopover site showed relatively consistent migration schedules for ruddy turnstone, but more interannual variation for red knot and sanderling (Figure 4, Figure S1). In most years there was a single sampling period with peak arrival probability for each species, but in some years red knot had multiple peaks indicating the delayed arrival of some birds, possibly those overwintering farther south (Atkinson et al. 2005). Stopover persistence probability was consistently high early in the season for all three species, and sharply declined in the last few sampling periods (typically after ~25 May) indicating departures to continue migration (Figure S1). The proportion of the population present typically peaked during the seventh sampling occasion, May 26-28, for all three species, but there was considerable variation among years (Figure 4). The estimated proportion present during the seventh sampling period varied most among years for red knot (range: 0.07 (2016) – 0.59 (2007), less so for sanderling (range: 0.18 (2013) – 0.54 (2015), and least for ruddy turnstone (range: 0.24 (2012) – 0.43 (2007))

Effects of stopover site conditions on survival and site use

We estimated high apparent annual survival probability for all three species (Table 1 Figure 2). Average apparent annual survival probability was 0.89 (0.84, 0.92) for red knot, 0.94 (0.89, 0.97) for ruddy turnstone, and 0.82 (0.72, 0.95) for sanderling across the 13 years of this study (Figure 2). Estimates of temporary emigration indicated nonrandom temporary emigration for all three species, with the probability of using the site if present last year (γ_j^{II}) consistently greater than the probability of using the site if not present last year (γ_j^{OI}), indicating high stopover site fidelity (Table 1, Figure 2). On average the probability of using the site in two consecutive years was ~0.75 for all three species. Residency probability (τ_j, the probability of remaining in the study area for at least two sampling periods) was generally high for most years for all three species with an average of ~0.65 and variance of ~0.005.

We found a demographic response to stopover site conditions for red knot and a behavioral response to stopover site conditions for ruddy turnstone. For ruddy turnstone, the probability of returning to this site if present in the previous year was positively associated with the availability of horseshoe crabs in the previous year (β = 0.27, *f* = 0.98, Table 3). We found no evidence of an association between horseshoe crab egg availability and ruddy turnstone annual survival probability. For red knot, however, apparent annual survival probability was positively associated with horseshoe crab egg availability (β = 0.37, *f* = 0.95) and negatively associated with the average rate of mass gain in the previous year (β = -0.33, *f* = 0.91). We also found some evidence that red knot apparent annual survival probability was lower in years following a severe weather event during stopover (β = -0.55, *f* = 0.87). We found no strong

associations between stopover site conditions and sanderling survival, temporary emigration, or residency (Table 3).

Discussion

Our application of the open robust design model to a migratory stopover system allows for integrated estimation of within-year dynamics of arrival and departure and between-year dynamics of transience, temporary emigration, and survival in a hierarchical Bayesian framework. We found differences between species in response to stopover site conditions; unfavorable conditions were associated with decreased survival for red knot but lower probability of site use for ruddy turnstone. Inter-annual variation in arrival, persistence, and transience led to variation in the proportion of the population present throughout the season, especially for red knot. This demonstrates that both the proportion of the flyway population using this stopover site and the proportion present during a given sampling period both vary among years, both of which should be considered in analyzing and interpreting monitoring data.

Migratory stopover sites are often treated as discrete islands of habitats, but in reality, there are usually relatively continuous stretches or connected networks of habitat that birds utilize along the migratory route, depending on local food and weather conditions in that year (Parent 1997, Gillespie and Fontaine 2017). This analysis shows that the average probability of returning to this stopover site in two consecutive years is ~75% for these species, indicating relatively high site fidelity but a ~25% probability of skipping this particular stopover site. Results from geolocator studies indicate that individual ruddy turnstone and red knot may skip Delaware Bay in some years (R. Porter pers. comm.), and that there is substantial intraspecific variation in migratory routes (Minton et al. 2013, Newstead et al. 2013). Additionally, we found

that for all species the probability of returning to the stopover site if absent in the previous year was low, which again indicates relatively fixed migratory behaviors for these species. If probability of returning to this site after a year of non-use is low, there may be the potential for shifting migratory behaviors away from of this site and towards use of other sites in the flyway. This underscores the importance of collaboration for flyway-wide monitoring and conservation action. In the Atlantic flyway, several shorebird monitoring programs collect similar counts and resighting data (Watts and Truitt 2015). Integrating the data from those studies in a single analysis will provide a fuller picture of population health for these species of high conservation concern.

This analysis provides, to our knowledge, the first estimates of annual survival probability for ruddy turnstone and sanderling using the Atlantic migratory flyway. A recent review paper gathered estimates of survival probability from studies of shorebirds worldwide (Méndez et al. 2018), and presented a meta-analytic average estimate of annual survival probability of 0.84 ± 0.037 for sanderling, 0.81 ± 0.020 for ruddy turnstone, and 0.80 ± 0.011 for red knot. Our estimate of average sanderling survival probability (0.82) agrees with the studies included in that review, but our estimates for both ruddy turnstone and red knot average (0.94 and 0.89, respectively), are higher. For ruddy turnstone, the only other available survival estimates were either from dead recoveries or return rates that do not account for imperfect detection, which likely underestimate the true survival probability. For red knot, our estimates agree with the two other available studies from this flyway, which both estimated an average annual survival probability close to 0.9 (McGowan et al. 2011, Schwarzer et al. 2012).

We estimated relatively high stopover site fidelity for all three species, which agree with previous studies on these species in other flyways (Metcalfe and Furness 1985, Harrington et al.

1988, Gudmundsson and Lindstrom 1992, Burton and Evans 1997, Buchanan et al. 2012). The probability of returning to the site was associated with food availability in the previous year, especially for ruddy turnstone, a relationship that Minias et al. (2010) also found for inland shorebird migrants in Europe. Stopover site fidelity is thought to be associated with migration distance (Alerstam 1979, Catry et al. 2004), and the prevailing thought is that the majority of *rufa* red knot use Delaware Bay during northward migration (Niles et al. 2008). However, our analysis indicates that less than 100% of the population returns to this site each year, presumably using other coastal sites along the Atlantic coast instead (Watts and Truitt 2015).

Previous work in this system showed that in cooler years when horseshoe crab spawning was delayed, red knot gained weight later in the season and at a faster rate (Chapter 3). Although other studies have found a positive association between rate of mass gain and survival probability for migratory shorebirds (Rakhimberdiev et al. 2018), we speculated that such extremely rapid weight gain (~13% of lean body mass per day) could have negative physiological effects, particularly since mass gain was slower in years of high resource availability. This analysis shows that an increased rate of mass gain was negatively associated with annual survival probability for red knot, which aligns with those previous predictions. Additionally, we found that a faster rate of mass gain was negatively associated with the probability of returning to the site, presumably because increased rates of mass gain occur in years of lower resource availability. This work provides further evidence that unfavorable stopover conditions can have negative demographic effects (Baker et al. 2004, McGowan et al. 2011, Piersma et al. 2014) and also highlights species differences in those effects. We found no evidence that annual survival of ruddy turnstone or sanderling were associated with horseshoe

crab egg availability, which may be attributable to their more diverse diet and foraging strategy during spring migration (Gill 1986, Castro and Suazo 2009).

Our analysis of migratory shorebirds during spring stopover showed several species-specific differences in responses to stopover site conditions. We found that ruddy turnstone were more likely to shift migratory behaviors away from use of this site when conditions were poor, but found little evidence that stopover site conditions influence survival probability. For red knot, we found that survival probability was associated with both horseshoe crab egg availability and severe weather events during stopover. These species differences align with previous analyses indicating that red knot are more sensitive to the availability of horseshoe crab eggs than ruddy turnstone, while the generalist foraging strategy used by ruddy turnstone makes them more adaptable to changing conditions (Chapter 3, Sol et al. 2002, Davey et al. 2012).

Red knot exhibited more inter-annual variation in timing of stopover than the other two species. It is unclear whether this variability is an artifact of sampling or indicative of true variation in migratory behaviors among years. For this study, we only sampled birds using the Delaware coast (southwest) of Delaware Bay, but shorebirds also forage and roost along the New Jersey coast (northeast). Anecdotally, red knot seem to exhibit high variability from year to year in their spatial distribution throughout the region, likely tied to the availability of their preferred prey, horseshoe crab eggs (Karpanty et al. 2006). It is possible that years with low proportions of the population present (e.g. 2016, Figure 2) indicate years in which spatial use of the region favored areas of the Bay that were not sampled by this study. Additionally, with this model we assume that all emigration from the study area within the season is permanent, i.e. departures to continue migration, and that any temporary emigration occurs randomly and is therefore captured by the detection process. If this assumption is violated, it could negatively bias our

estimates of persistence probability and, therefore, our estimates of the proportion of the population present (Kendall et al. 1997).

Understanding how stopover site conditions influence both demographic rates and migratory behaviors is critical for interpreting changes in passage population sizes at stopover sites and predicting population responses to future changes (Pomeroy et al. 2006, Smith et al. 2012, Piersma et al. 2014). In Delaware Bay, the adaptive resource management plan for horseshoe crab harvest depends on estimates of red knot passage population size (McGowan et al. 2015). More generally, predictive models for management decision-making with regards to stopover site for migratory species would greatly benefit if they accounted for inter-annual variation in site use as well as demographics. This work provides a framework to account for these flow-through dynamics while analyzing counts conducted during migration (Gillings et al. 2009, Dickie et al. 2014). This model could be applied to a variety of migratory species to improve population trend estimation and understand environmental effects on both demographics and migratory behaviors.

Literature Cited

Albanese, G., C. a. Davis, and B. W. Compton. 2012. Spatiotemporal scaling of North American continental interior wetlands: implications for shorebird conservation. Landscape Ecology 27:1465–1479.

Alerstam, T. 1979. Wind as Selective Agent in Bird Migration. Ornis Scandinavia 10:76–93.

Alerstam, T. 2011. Optimal bird migration revisited. Journal of Ornithology 152.

Amano, T., T. Székely, K. Koyama, H. Amano, and W. J. Sutherland. 2010. A framework for monitoring the status of populations: An example from wader populations in the East Asian-Australasian flyway. Biological Conservation 143:2238–2247.

Atkinson, P. W., A. J. Baker, R. M. Bevan, N. a. Clark, K. B. Cole, P. M. Gonzalez, J. Newton, L. J. Niles, and R. a. Robinson. 2005. Unravelling the migration and moult strategies of a long-distance migrant using stable isotopes: Red Knot Calidris canutus movements in the Americas. Ibis 147:738–749.

Baker, A., P. Gonzalez, R. I. G. Morrison, and B. A. Harrington. 2001. Red Knot (Calidris canutus). Page The Birds of North America. Cornell Lab of Ornithology, Ithaca.

Baker, A. J., P. M. González, T. Piersma, L. J. Niles, I. D. L. S. do Nascimento, P. W. Atkinson, N. a Clark, C. D. T. Minton, M. K. Peck, and G. Aarts. 2004. Rapid population decline in red knots: fitness consequences of decreased refuelling rates and late arrival in Delaware Bay. Proceedings. Biological sciences / The Royal Society 271:875–882.

Bart, J., S. Brown, B. Harrington, and R. I. Guy Morrison. 2007. Survey trends of North American shorebirds: Population declines or shifting distributions? Journal of Avian Biology 38:73–82.

Both, C., C. a M. Van Turnhout, R. G. Bijlsma, H. Siepel, A. J. Van Strien, and R. P. B. Foppen. 2010. Avian population consequences of climate change are most severe for long-distance migrants in seasonal habitats. Proceedings. Biological sciences / The Royal Society 277:1259–66.

Bridge, E. S., J. F. Kelly, A. Contina, R. M. Gabrielson, R. B. MacCurdy, and D. W. Winkler. 2013. Advances in tracking small migratory birds: A technical review of light-level geolocation. Journal of Field Ornithology 84:121–137.

Bridge, E. S., K. Thorup, M. S. Bowlin, P. B. Chilson, R. H. Diehl, R. W. Fléron, P. Hartl, R. Kays, J. F. Kelly, W. D. Robinson, and M. Wikelski. 2011. Technology on the Move: Recent and Forthcoming Innovations for Tracking Migratory Birds. BioScience 61:689–698.

Buchanan, J. B., J. E. Lyons, L. J. Salzer, R. Carmona, N. Arce, G. J. Wiles, K. Brady, G. E. Hayes, S. M. Desimone, G. Schirato, and W. Michaelis. 2012. Among-year site fidelity of

Red Knots during migration in Washington. Journal of Field Ornithology 83:282–289.

Buler, J. J., and F. R. Moore. 2011. Migrant-habitat relationships during stopover along an ecological barrier: Extrinsic constraints and conservation implications. Journal of Ornithology 152:101–112.

Burger, J., L. J. Niles, R. R. Porter, A. D. Dey, S. Koch, and C. Gordon. 2012. Migration and over-wintering of red knots (Calidris canutus rufa) along the Atlantic Coast of the United States. The Condor 114:302–313.

Burton, N. H. K., and P. R. Evans. 1997. Survival and winter site-fidelity of Turnstones Arenaria interpres and Purple Sandpipers Calidris maritima in northeast England. Bird Study 44:34–44.

Castro, M., and C. Suazo. 2009. Diet selection of sanderlings (Calidris alba) in Isla Guamblin national park in the Chilean fjords. Ornitología Neotropical 20:247–253.

Catry, P., V. Encarnac, A. Araujo, P. Fearon, A. Fearon, M. Armelin, and P. Delaloye. 2004. Are long-distance migrant passerines faithful to their stopover sites ? Journal of Avian Biology 34:170–181.

Cimprich, D., R. D. Sutter, D. N. Ewert, C. Duncan, M. Woodrey, B. Abel, D. W. Mehlman, and S. E. Mabey. 2006. Conserving Stopover Sites for Forest-Dwelling Migratory Landbirds. The Auk 122:1281.

Clark, E., L. J. Niles, and J. Burger. 1993. Abundance and distribution of migrant shorebirds in Delaware Bay. The Condor 95:694–705.

Clark, N., S. Gillings, A. Baker, P. Gonzalez, and R. Porter. 2005. The production and use of permanently inscribed leg flags for waders. Wader Study Group Bulletin 108:38–41.

Conn, P. B., D. S. Johnson, P. J. Williams, S. R. Melin, and M. B. Hooten. 2018. A guide to Bayesian model checking for ecologists. Ecological Monographs 0:1–17.

Davey, C. M., D. E. Chamberlain, S. E. Newson, D. G. Noble, and A. Johnston. 2012. Rise of the generalists: Evidence for climate driven homogenization in avian communities. Global Ecology and Biogeography 21:568–578.

Dickie, A. M., P. A. Smith, and H. G. Gilchrist. 2014. The Importance of Survey Timing on Shorebird Density Estimates at East Bay, Nunavut, Canada. Waterbirds 37:394–401.

Franklin, T., W. Franklin, L. Brooks, and P. Harrison. 2018. Site-specific female-biased sex ratio of humpback whales (Megaptera novaeangliae) during a stopover early in the southern migration. Canadian Journal of Zoology 96:533–544.

Fraser, J. D., S. M. Karpanty, and J. B. Cohen. 2010. Shorebirds Forage Disproportionately in Horseshoe Crab Nest Depressions. Waterbirds 33:96–100.

Giavi, S., M. Moretti, F. Bontadina, N. Zambelli, and M. Schaub. 2014. Seasonal survival probabilities suggest low migration mortality in migrating bats. PLoS ONE 9.

Gibbs, D., R. Walton, L. Brower, and A. K. Davis. 2006. Monarch Butterfly (Lepidoptera: Nymphalidae) Migration Monitoring at Chincoteague, Virginia and Cape May, New Jersey: A Comparison of Long-term Trends. Journal of the Kansas Entomological Society 79:156–164.

Gill, R. E. 1986. What won't Turnstones eat? British Birds 79:401–402.

Gillespie, C. R., and J. J. Fontaine. 2017. Shorebird stopover habitat decisions in a changing landscape. Journal of Wildlife Management 81:1051–1062.

Gillings, S., P. W. Atkinson, A. J. Baker, K. a. Bennett, N. a. Clark, K. B. Cole, P. M. González, K. S. Kalasz, C. D. T. Minton, L. J. Niles, R. C. Porter, I. D. L. Serrano, H. P. Sitters, and J. L. Woods. 2009. Staging Behavior in Red Knot (Calidris canutus) in Delaware Bay: Implications for Monitoring Mass and Population Size. The Auk 126:54–63.

Green, J. M. H., S. Sripanomyom, X. Giam, and D. S. Wilcove. 2015. The ecology and economics of shorebird conservation in a tropical human-modified landscape. Journal of Applied Ecology 52:1483–1491.

Gudmundsson, G. a., and a. Lindstrom. 1992. Spring migration of sanderlings Calidris alba through SW Iceland - Wherefrom and whereto. Ardea 80:315–326.

Gwinner, E. 1996. Circannual clocks in avian reproduction and migration. Ibis 138:47–63.

Harrington, B. A., J. M. Hagan, and L. E. Leddy. 1988. Site Fidelity and Survival Differences between Two Groups of New World Red Knots (Calidris canutus). The Auk 105:439–445.

Jonker, R. M., G. Eichhorn, F. van Langevelde, and S. Bauer. 2010. Predation danger can explain changes in timing of migration: The case of the Barnacle goose. PLoS ONE 5:4–11.

Karpanty, S. M., J. D. Fraser, J. Berkson, L. J. Niles, A. Dey, and E. P. Smith. 2006. Horseshoe Crab Eggs Determine Red Knot Distribution in Delaware Bay. Journal of Wildlife Management 70:1704–1710.

Kays, R., M. C. Crofoot, W. Jetz, and M. Wikelski. 2015. Terrestrial animal tracking as an eye on life and planet. Science 348:aaa2478.

Kendall, W. L., and R. Bjorkland. 2001. Using open robust design models to estimate temporary emigration from capture-recapture data. Biometrics 57:1113–1122.

Kendall, W. L., J. D. Nichols, and J. E. Hines. 1997. Estimating Temporary Emigration Using Capture-Recapture Data with Pollock' s Robust Design. Ecology 78:563–578.

Kranstauber, B., A. Cameron, R. Weinzerl, T. Fountain, S. Tilak, M. Wikelski, and R. Kays. 2011. The Movebank data model for animal tracking. Environmental Modelling and

Software 26:834–835.

Lei, W., J. A. Masero, T. Piersma, B. Zhu, H. Y. Yang, and Z. Zhang. 2018. Alternative habitat: The importance of the Nanpu Saltpans for migratory waterbirds in the Chinese Yellow Sea. Bird Conservation International 28:549–566.

Lyons, J. E., W. L. Kendall, J. A. Royle, S. J. Converse, B. A. Andres, and J. B. Buchanan. 2015. Population size and stopover duration estimation using mark-resight data and Bayesian analysis of a superpopulation model. Biometrics.

Macwhirter, R. B., P. A.-S. Jr., and D. E. Kroodsma. 2002. Sanderling (Calidris alba). Page *in* A. F. Poole and F. B. Gill, editors. The Birds of North America. Second edition. Cornell Lab of Ornithology, Ithaca, NY.

McGowan, C. P., J. E. Hines, J. D. Nichols, J. E. Lyons, D. R. Smith, K. S. Kalasz, L. J. Niles, A. D. Dey, N. a. Clark, P. W. Atkinson, C. D. T. Minton, and W. Kendall. 2011. Demographic consequences of migratory stopover: linking red knot survival to horseshoe crab spawning abundance. Ecosphere 2.

McGowan, C. P., D. R. Smith, J. D. Nichols, J. E. Lyons, J. Sweka, K. Kalasz, L. J. Niles, R. Wong, J. Brust, M. Davis, and B. Spear. 2015. Implementation of a framework for multi-species, multi-objective adaptive management in Delaware Bay. Biological Conservation 191:759–769.

Mcguire, L. P., C. G. Guglielmo, S. A. Mackenzie, and P. D. Taylor. 2012. Migratory stopover in the long-distance migrant silver-haired bat, Lasionycteris noctivagans. Journal of Animal Ecology 81:377–385.

Méndez, V., J. A. Alves, J. A. Gill, and T. G. Gunnarsson. 2018. Patterns and processes in shorebird survival rates: a global review. Ibis.

Metcalfe, N., and R. Furness. 1985. Survival, winter population stability and site fidelity in the turnstone Arenaria interpres. Bird Study 32:207–214.

Minias, P., K. Kaczmarek, R. Włodarczyk, T. Janiszewski, and R. Bargiel. 2010. Short communication Feeding conditions determine return rates to stopover sites of inland waders on autumn migration. Ibis 152:840–844.

Minton, C. D. T., K. Gosbell, P. Johns, M. Christie, M. Klaassen, C. Hassell, A. Boyle, R. Jessop, and J. W. Fox. 2011. Geolocator studies on Ruddy Turnstones Arenaria interpres and Greater Sandplovers Charadrius leschenaultii in the East Asian-Australasia Flyway reveal widely different migration strategies. Wader Study Group Bulletin 118:87–96.

Minton, C., K. Gosbell, P. Johns, M. Christie, M. Klaassen, C. Hassell, A. Boyle, R. Jessop, and J. Fox. 2013. New insights from geolocators deployed on waders in Australia. Wader Study Group Bulletin 120:37–46.

Mood, C. 2010. Logistic regression: Why we cannot do what We think we can do, and what we

can do about it. European Sociological Review 26:67–82.

Moore, F. R., P. Kerlinger, and T. R. Simons. 1990. Stopover on a Gulf Coast Barrier Island by Spring Trans-Gulf Migrants. Wilson Ornithological Society 102:487–500.

Nettleship, D. N. 2000. Ruddy Turnstone (Arenaria interpres). Page The Birds of North America. Cornell Lab of Ornithology, Ithaca.

Newstead, D. J., L. J. Niles, R. R. Porter, A. D. Dey, J. Burger, and O. N. Fitzsimmons. 2013. Geolocation reveals mid-continent migratory routes and Texas wintering areas of Red Knots *Calidris canutus rufa*. Wader Study Group Bulletin 120:53–59.

Newton, I. 2006. Can conditions experienced during migration limit the population levels of birds? Journal of Ornithology 147:146–166.

Niles, L. J., H. P. Sitters, D. Dey, Amanda, P. W. Atkinson, A. J. Baker, K. A. Bennett, R. Carmona, K. E. Clark, N. A. Clark, C. Espoz, P. M. Gonzalez, B. A. Harrington, D. E. Hernandez, K. S. Kalasz, R. G. Lathrop, N. Matus, Ricardo, C. D. T. Minton, R. I. G. Morrison, M. K. Peck, W. Pitts, R. A. Robinson, and I. L. Serrano. 2008. Status of the Red Knot (Calidris Canutus Rufa) in the Western Hemisphere (Calidris Canutus Rufa) in. Studies in Avian Biology 36:1–185.

Nordstrom, K. F., N. L. Jackson, D. R. Smith, and R. G. Weber. 2006. Transport of horseshoe crab eggs by waves and swash on an estuarine beach: Implications for foraging shorebirds. Estuarine, Coastal and Shelf Science 70:438–448.

Parent, A. H. F. and A. H. 1997. Effects of the Landscape on Shorebird Movements at Spring Migration Stopovers. The Condor 99:698–707.

Peng, H.-B., N. Hua, C.-Y. Choi, D. S. Melville, Y. Gao, Q. Zhou, Y. Chen, W. Xue, Q. Ma, W. Wu, C. Tang, and Z. Ma. 2014. Adjusting migration schedules at stopping sites: time strategy of a long-distance migratory shorebird during northward migration. Journal of Ornithology 156:191–199.

Piersma, T., T. Lok, Y. Chen, C. J. Hassell, H.-Y. Yang, A. Boyle, M. Slaymaker, Y.-C. Chan, D. S. Melville, Z.-W. Zhang, and Z. Ma. 2014. Simultaneous declines in summer survival of three shorebird species signals a flyway at risk. Journal Of Applied Ecology:manuscript.

Pollock, K. 1981. A capture-recapture sampling design robust to unequal catchability. The Journal of Wildlife Management 46:752–757.

Pomeroy, A. C., R. W. Butler, and R. C. Ydenberg. 2006. Experimental evidence that migrants adjust usage at a stopover site to trade off food and danger. Behavioral Ecology 17:1041–1045.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rakhimberdiev, E., S. Duijns, Karagicheva, Julia, C. J. Camphuysen, V. Castricum, A. Dekinga, R. Dekker, A. Gavrilov, J. ten Horn, J. Jukema, A. Saveliev, M. Soloviec, T. L. Tibbitts, J. A. van Gils, and T. Piersma. 2018. Fuelling conditions at staging sites can mitigate Arctic warming effects in a migratory bird. Nature Communications.

Robinson, R. A., H. Q. P. Crick, J. A. Learmonth, I. M. D. Maclean, C. D. Thomas, F. Bairlein, M. C. Forchhammer, C. M. Francis, J. A. Gill, B. J. Godley, J. Harwood, G. C. Hays, B. Huntley, A. M. Hutson, G. J. Pierce, M. M. Rehfisch, D. W. Sims, M. Bego??a Santos, T. H. Sparks, D. A. Stroud, and M. E. Visser. 2009. Travelling through a warming world: Climate change and migratory species. Endangered Species Research 7:87–99.

Sawyer, H., M. J. Kauffman, R. M. Nielson, and J. S. Horne. 2009. Identifying and Prioritizing Ungulate Migration Routes for Landscape-Level Conservation. Ecological Applications 19:2016–2025.

Schaub, M., F. Liechti, and L. Jenni. 2004. Departure of migrating European robins, Erithacus rubecula, from a stopover site in relation to wind and rain. Animal Behaviour 67:229–237.

Schwarz, C. J., and W. T. Stobo. 1997. Estimating Temporary Migration Using the Robust Design. Biometrics 53:178–194.

Schwarzer, A. C., J. a. Collazo, L. J. Niles, J. M. Brush, N. J. Douglass, and H. F. Percival. 2012. Annual survival of Red Knots (Calidris canutus rufa) wintering in Florida. The Auk 129:725–733.

Smith, D. R., N. L. Jackson, K. F. Nordstrom, and R. G. Weber. 2011. Beach characteristics mitigate effects of onshore wind on horseshoe crab spawning: Implications for matching with shorebird migration in Delaware Bay. Animal Conservation 14:575–584.

Smith, P. A., C. L. Gratto-Trevor, B. T. Collins, S. D. Fellows, R. B. Lanctot, J. Liebezeit, B. J. McCaffery, D. Tracy, J. Rausch, S. Kendall, S. Zack, and H. R. Gates. 2012. Trends in Abundance of Semipalmated Sandpipers: Evidence from the Arctic. Waterbirds 35:106–119.

Sol, D., S. Timmermans, and L. Lefebvre. 2002. Behavioural flexibility and invasion success in birds. Animal Behaviour 63:495–502.

Staton, B. A., M. J. Catalano, T. M. Farmer, A. Abebe, and F. S. Dobson. 2017. Development and evaluation of a migration timing forecast model for Kuskokwim River Chinook salmon. Fisheries Research 194:9–21.

Studds, C. E., B. E. Kendall, N. J. Murray, H. B. Wilson, D. I. Rogers, R. S. Clemens, K. Gosbell, C. J. Hassell, R. Jessop, D. S. Melville, D. A. Milton, C. D. T. Minton, H. P. Possingham, A. C. Riegen, P. Straw, E. J. Woehler, and R. A. Fuller. 2017. Rapid population decline in migratory shorebirds relying on Yellow Sea tidal mudflats as stopover sites. Nature Communications 8:1–7.

Thorup, K., A. P. Tøttrup, M. Willemoes, R. H. G. Klaassen, R. Strandberg, M. L. Vega, H. P.

Dasari, M. B. Araújo, M. Wikelski, and C. Rahbek. 2017. Resource tracking within and across continents in long-distance bird migrants. Science Advances 3:1–11.

Tsipoura, N., and J. Burger. 1999. Shorebird Diet During Spring Migration Stopover on Delaware Bay. The Condor 101:635–644.

de Valpine, P., D. Turek, C. J. Paciorek, C. Anderson-Bergman, D. T. Lang, and R. Bodik. 2017. Programming With Models: Writing Statistical Algorithms for General Model Structures With NIMBLE. Journal of Computational and Graphical Statistics 26:403–413.

Verhoeven, M. A., A. H. J. Loonstra, N. R. Senner, A. D. McBride, C. Both, and T. Piersma. 2019. Variation From an Unknown Source: Large Inter-individual Differences in Migrating Black-Tailed Godwits. Frontiers in Ecology and Evolution 7:1–9.

Watts, B. D., and B. R. Truitt. 2015. Spring migration of red knots along the Virginia barrier islands. Journal of Wildlife Management 79:288–295.

Wilcove, D. S., and M. Wikelski. 2008. Going, going, gone: Is animal migration disappearing? PLoS Biology 6:1361–1364.

Williams, B., J. Nichols, and M. Conroy. 2002. Analysis and Management of Wildlife Populations. First edition. Academic Press, San Diego, California.

Woodworth, B. K., N. T. Wheelwright, A. E. Newman, M. Schaub, and D. R. Norris. 2017. Winter temperatures limit population growth rate of a migratory songbird. Nature Communications 8:1–9.

Xia, S., X. Yu, S. Millington, Y. Liu, Y. Jia, L. Wang, X. Hou, and L. Jiang. 2017. Identifying priority sites and gaps for the conservation of migratory waterbirds in China's coastal wetlands. Biological Conservation 210:72–82.

Ydenberg, R. C., R. W. Butler, D. B. Lank, B. D. Smith, and J. Ireland. 2004. Western sandpipers have altered migration tactics as peregrine falcon populations have recovered. Proceedings of the Royal Society B: Biological Sciences 271:1263–1269.

Tables

Table 4-1: Average and annual variance estimates of survival, transience, and temporary emigration.
Estimates of the overall average and estimated temporal variance of γ^{II} (probability of being available if present in the previous year), γ^{OI} (probability of being available if not present in the previous year), ϕ (apparent annual survival probability), and τ (residency probability).

Species	Parameter	Average		Annual variance	
		Mean	95% CRI	Mean	95% CRI
Red knot	γ^{II}	0.76	0.7, 0.81	0.006	0, 0.023
	γ^{OI}	0.15	0.09, 0.26	0.016	0.002, 0.071
	ϕ	0.89	0.84, 0.92	0.003	0, 0.011
	τ	0.65	0.61, 0.7	0.004	0, 0.013
Ruddy turnstone	γ^{II}	0.76	0.72, 0.81	0.004	0.001, 0.013
	γ^{OI}	0.06	0.03, 0.11	0.004	0, 0.015
	ϕ	0.94	0.89, 0.97	0.001	0, 0.007
	τ	0.63	0.59, 0.67	0.005	0.001, 0.012
Sanderling	γ^{II}	0.74	0.64, 0.86	0.012	0, 0.058
	γ^{OI}	0.13	0.02, 0.46	0.178	0, 2.149
	ϕ	0.82	0.72, 0.95	0.002	0, 0.011
	τ	0.67	0.6, 0.75	0.006	0, 0.029

Table 4-2: Effects of stopover site conditions on shorebird demographics and behavior. Standardized effects of stopover site conditions on demographic rates and migratory behaviors. For each species, covariate effects are listed in order from strongest to weakest relationship. Parameter is the model parameter (response variable), covariate is the ecological predictor (rate of mass gain (red knot only), horseshoe crab egg availability, occurrence of severe weather, or water temperature), estimate is the posterior mean, LCRI and UCRI are the bounds of the 95% HPDI, and *f* is the proportion of the posterior distribution that is the same sign as the mean (*f* = 0.5 indicates no evidence of a relationship).

Species	Parameter	Covariate	Estimate	LCRI	UCRI	*f*
Red knot	ϕ	Hsc availability	0.37	-0.08	0.71	0.95
	ϕ	Rate of mass gain	-0.33	-0.74	0.12	0.91
	ϕ	Severe weather	-0.55	-1.49	0.44	0.87
	γ^{II}	Hsc availability	-0.10	-0.41	0.19	0.75
	τ	Hsc availability	0.05	-0.34	0.45	0.59
	τ	Water temperature	0.04	-0.4	0.46	0.58
Ruddy turnstone	γ^{II}	Hsc availability	0.27	0.02	0.48	0.98
	τ	Hsc availability	0.10	-0.13	0.33	0.81
	τ	Water temperature	-0.11	-0.36	0.16	0.81
	ϕ	Hsc availability	0.08	-0.46	0.57	0.63
	ϕ	Severe weather	0.15	-1.1	1.18	0.62
Sanderling	γ^{II}	Hsc availability	-0.16	-0.52	0.29	0.78
	τ	Hsc availability	0.14	-0.35	0.58	0.74
	ϕ	Hsc availability	-0.09	-0.64	0.5	0.63
	τ	Water temperature	0.03	-0.44	0.52	0.53
	ϕ	Severe weather	-0.04	-1.35	1.27	0.52

Figures

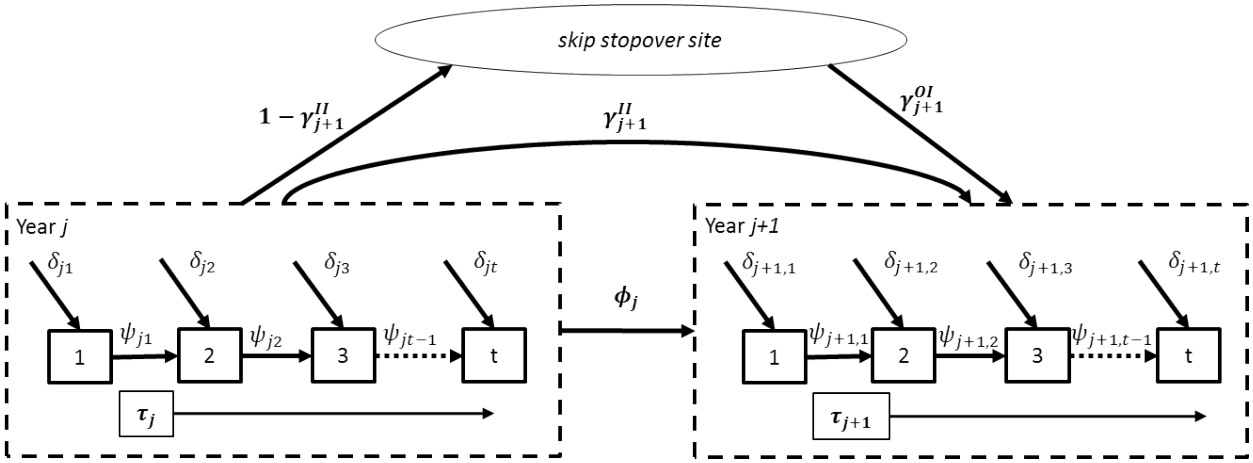

Figure 4-1. Conceptual diagram of the open robust design model.
Within a year (*j*, dashed boxes), individuals can arrive during any secondary sampling occasion *t* with probability δ_{jt}, remain as stopover residents with probability τ_j, and persist between sampling occasions with probability ψ_{jt}. Between years (primary periods), the probability of returning to the site is determined by the apparent annual survival probability ϕ_j and temporary emigration parameters γ_j^{II} and γ_j^{OI}, which describe the probability of returning to the site if present or not present in the previous year, respectively. The detection process is not depicted in this diagram, but individuals can be detected during any secondary sampling occasion with probability p_{jt} and are detected at least once in a year with probability p_j^*, which is derived based on estimates of p_{jt}, δ_{jt}, and ψ_{jt} (Table 1).

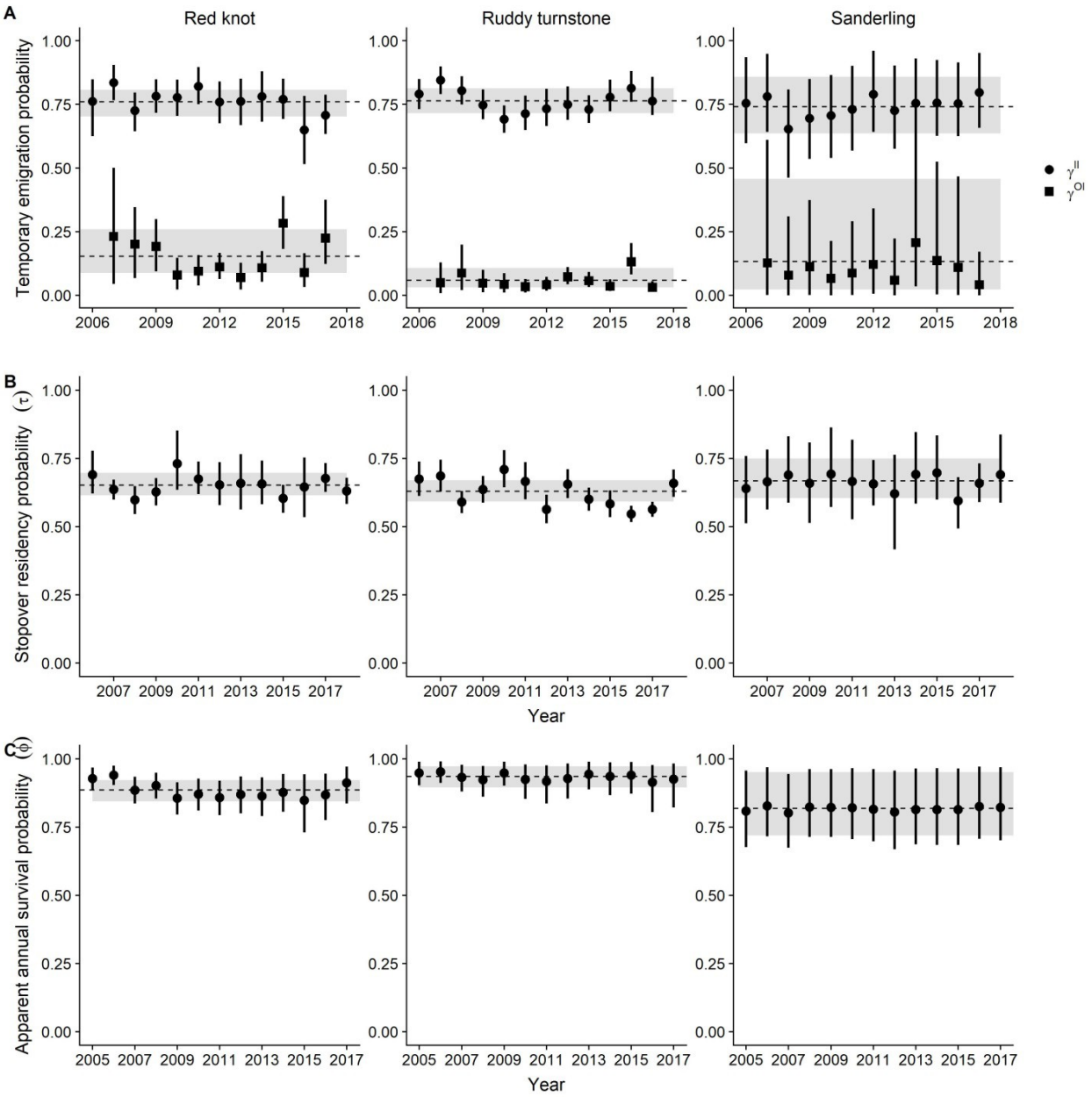

Figure 4-2. Annual use of the stopover site by three shorebird species.
For temporary emigration probabilities (A), circles represent the probability of using the site if present in the previous year (γ_j^{II}) and squares represent the probability of using the site if not present in the previous year (γ_j^{OI}). Stopover residency probability (B) is the probability of remaining at the stopover site for at least 2 sampling occasions (τ_j). Apparent annual survival probability (C) is the probability that an individual alive in year *j* survives to year *j+1*. Points and vertical lines indicate the posterior mean and 95% credible interval of each parameter for each year. Horizontal dashed lines and shaded gray region indicate the posterior mean and 95% credible interval of the overall average estimate.

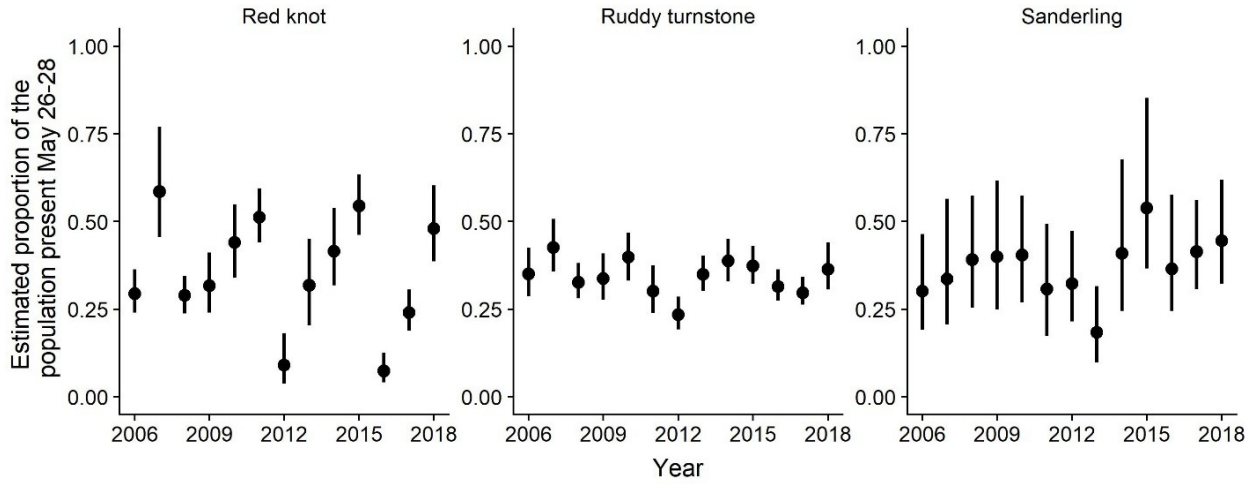

Figure 4-3. Annual variation in the probability of presence at the stopover site.
Estimated probability of being present in sampling period 7 (May 26-28) each year ($\pi_{j,7}$), which was on average the sampling period with the greatest probability of presence. The proportion of the stopover population present was calculated based on the arrival, persistence, and temporary emigration probabilities for each occasion of each year as described in text. Points represent posterior means and vertical lines 95% credible intervals.

Supporting Information

Estimates of arrival and departure

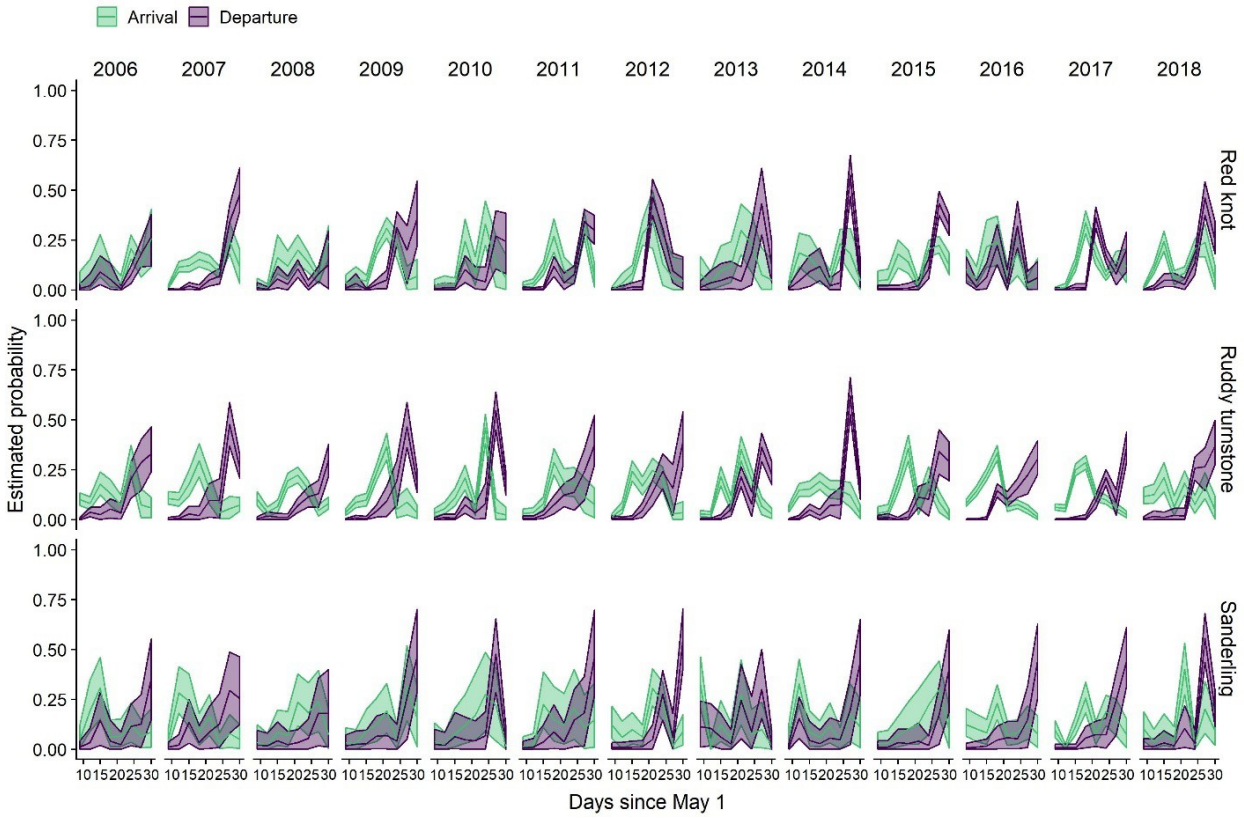

Figure S1. The estimated arrival and departure probabilities for each species in each year. Departure probability is derived for each occasion based on the previous arrival and persistence probabilities. Shaded regions indicate 95% credible intervals and lines indicate posterior means.

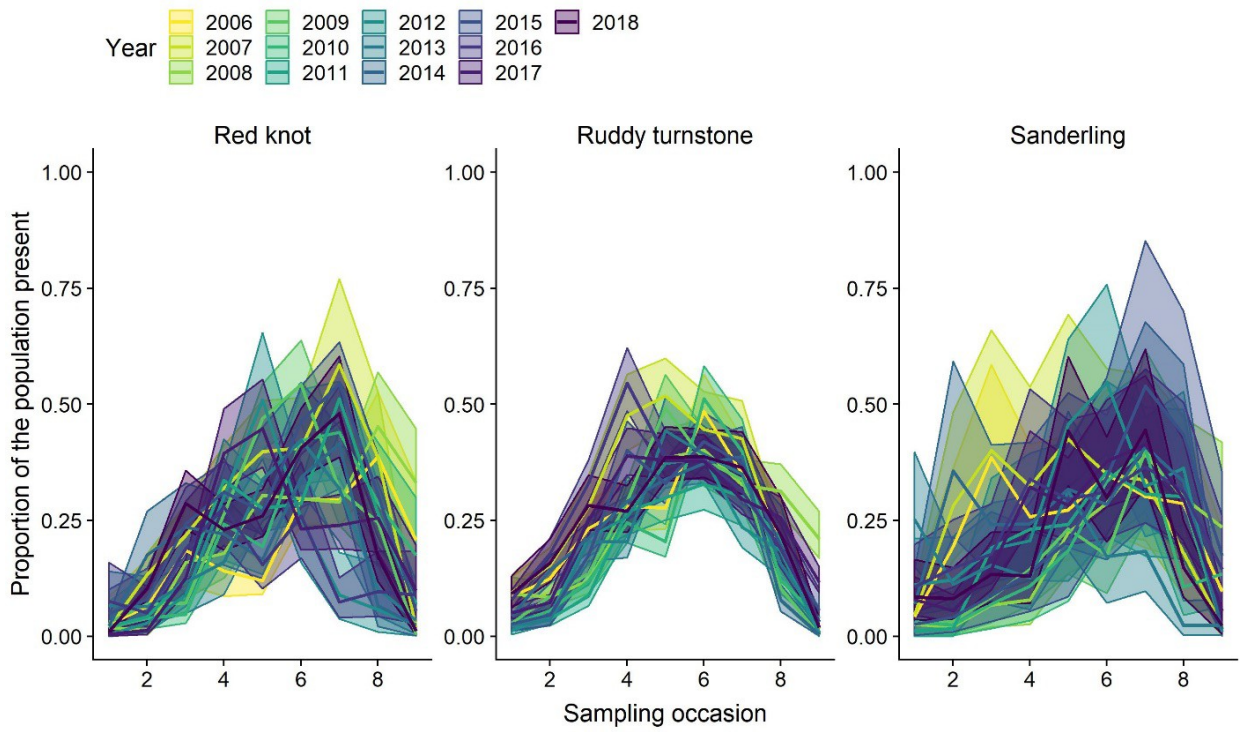

Figure S2. Estimated probability of being present in each sampling occasion of each year. Lines indicate posterior means and shaded regions show 95% credible intervals. The probability of presence (π_{jt}) is a derived parameter based on estimated probabilities of arrival, persistence, transience, and temporary emigration.

Table S1. Annual parameter estimates, reported as posterior means and 95% credible interval.

Species	Year	γ_j^{II}	γ_j^{OI}	ϕ_j	τ_j
Red knot	2005	--	--	0.93 (0.89, 0.97)	0.63 (0.56, 0.71)
	2006	0.76 (0.63, 0.85)	--	0.94 (0.90, 0.98)	0.69 (0.62, 0.78)
	2007	0.84 (0.77, 0.91)	0.23 (0.04, 0.50)	0.89 (0.84, 0.93)	0.64 (0.60, 0.67)
	2008	0.73 (0.64, 0.80)	0.20 (0.07, 0.35)	0.90 (0.85, 0.95)	0.60 (0.55, 0.65)
	2009	0.78 (0.72, 0.85)	0.19 (0.09, 0.30)	0.86 (0.80, 0.91)	0.63 (0.58, 0.68)
	2010	0.78 (0.70, 0.85)	0.08 (0.02, 0.15)	0.87 (0.81, 0.93)	0.73 (0.64, 0.85)
	2011	0.82 (0.75, 0.90)	0.09 (0.04, 0.16)	0.86 (0.79, 0.92)	0.67 (0.62, 0.74)
	2012	0.76 (0.68, 0.84)	0.11 (0.06, 0.17)	0.87 (0.80, 0.94)	0.65 (0.58, 0.74)
	2013	0.76 (0.67, 0.85)	0.07 (0.02, 0.13)	0.86 (0.79, 0.93)	0.66 (0.56, 0.77)
	2014	0.78 (0.68, 0.88)	0.11 (0.05, 0.17)	0.88 (0.81, 0.94)	0.66 (0.58, 0.74)
	2015	0.77 (0.69, 0.85)	0.28 (0.18, 0.39)	0.85 (0.73, 0.94)	0.60 (0.55, 0.65)
	2016	0.65 (0.52, 0.78)	0.09 (0.03, 0.17)	0.87 (0.78, 0.95)	0.64 (0.54, 0.75)
	2017	0.71 (0.63, 0.79)	0.22 (0.12, 0.37)	0.91 (0.84, 0.97)	0.68 (0.63, 0.73)
	2018	--	--	--	0.63 (0.58, 0.68)
Ruddy turnstone	2005	--	--	0.95 (0.90, 0.99)	0.67 (0.61, 0.74)
	2006	0.79 (0.73, 0.85)	--	0.95 (0.91, 0.99)	0.67 (0.61, 0.74)
	2007	0.84 (0.79, 0.90)	0.05 (0.01, 0.13)	0.93 (0.88, 0.98)	0.69 (0.63, 0.75)
	2008	0.80 (0.75, 0.86)	0.09 (0.02, 0.20)	0.92 (0.86, 0.97)	0.59 (0.55, 0.63)
	2009	0.75 (0.69, 0.81)	0.05 (0.01, 0.10)	0.95 (0.90, 0.99)	0.64 (0.59, 0.69)
	2010	0.69 (0.64, 0.75)	0.04 (0.01, 0.09)	0.93 (0.85, 0.98)	0.71 (0.64, 0.78)
	2011	0.71 (0.65, 0.78)	0.03 (0.01, 0.06)	0.92 (0.84, 0.98)	0.67 (0.60, 0.74)
	2012	0.73 (0.67, 0.81)	0.04 (0.02, 0.07)	0.93 (0.85, 0.98)	0.56 (0.51, 0.62)

	2013	0.75 (0.69, 0.82)	0.07 (0.04, 0.11)	0.94 (0.89, 0.99)	0.66 (0.60, 0.71)
	2014	0.73 (0.68, 0.79)	0.06 (0.03, 0.09)	0.94 (0.87, 0.99)	0.60 (0.56, 0.64)
	2015	0.78 (0.72, 0.85)	0.04 (0.02, 0.06)	0.94 (0.87, 0.99)	0.58 (0.53, 0.63)
	2016	0.81 (0.76, 0.88)	0.13 (0.08, 0.21)	0.91 (0.80, 0.98)	0.55 (0.52, 0.58)
	2017	0.76 (0.71, 0.86)	0.03 (0.02, 0.06)	0.93 (0.82, 0.98)	0.56 (0.54, 0.59)
	2018	--	--	--	0.66 (0.61, 0.71)
Sanderling	2005	--	--	0.81 (0.68, 0.96)	0.70 (0.57, 0.87)
	2006	0.75 (0.60, 0.94)	--	0.83 (0.72, 0.97)	0.64 (0.51, 0.76)
	2007	0.78 (0.64, 0.95)	0.13 (0, 0.61)	0.80 (0.67, 0.95)	0.66 (0.56, 0.78)
	2008	0.65 (0.46, 0.81)	0.08 (0, 0.31)	0.82 (0.71, 0.96)	0.69 (0.59, 0.83)
	2009	0.70 (0.54, 0.85)	0.11 (0, 0.37)	0.82 (0.71, 0.96)	0.66 (0.51, 0.81)
	2010	0.71 (0.54, 0.87)	0.07 (0, 0.21)	0.82 (0.71, 0.97)	0.69 (0.57, 0.86)
	2011	0.73 (0.57, 0.90)	0.09 (0, 0.29)	0.82 (0.70, 0.96)	0.66 (0.53, 0.82)
	2012	0.79 (0.64, 0.96)	0.12 (0.01, 0.34)	0.80 (0.67, 0.96)	0.66 (0.58, 0.74)
	2013	0.72 (0.58, 0.90)	0.06 (0, 0.22)	0.81 (0.69, 0.96)	0.62 (0.42, 0.76)
	2014	0.75 (0.61, 0.93)	0.21 (0.04, 0.67)	0.81 (0.68, 0.97)	0.69 (0.58, 0.85)
	2015	0.76 (0.63, 0.92)	0.14 (0, 0.52)	0.81 (0.68, 0.96)	0.70 (0.60, 0.83)
	2016	0.75 (0.62, 0.91)	0.11 (0, 0.47)	0.83 (0.71, 0.97)	0.59 (0.49, 0.68)
	2017	0.80 (0.66, 0.95)	0.04 (0, 0.17)	0.82 (0.70, 0.97)	0.66 (0.59, 0.73)
	2018	--	--	--	0.69 (0.59, 0.84)

Table S2. Primary period m-array for red knot

Release year (*j*)	Number released (	Number re-encountered in year *k* (m_{jk})													
		2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Never seen
2005	746	282	185	38	26	6	5	3	1	1	9	3	10	3	174
2006	848		524	80	33	12	7	6	3	2	4	1	4	6	166
2007	1210			488	183	29	30	18	5	7	22	2	8	9	409
2008	960				468	65	67	15	7	7	17	5	8	3	298
2009	978					333	171	33	12	19	27	2	9	4	368
2010	498						266	34	15	13	11	0	4	3	152
2011	846							319	73	49	51	8	11	12	323
2012	629								195	90	73	3	13	13	242
2013	358									146	64	7	10	6	125
2014	628										330	30	35	18	215
2015	934											211	182	71	470
2016	416												190	50	176
2017	1007													497	510

Table S3. Primary period m-array for ruddy turnstone

Release year (j)	Number released (R_j)	Number re-encountered in year k (m_{jk})													
		2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Never seen
2005	412	237	37	23	6	0	2	2	4	4	2	3	2	0	90
2006	549		349	57	14	4	3	2	3	2	0	5	4	1	105
2007	692			428	39	20	4	3	4	3	2	8	4	0	177
2008	995				512	72	30	13	12	8	5	18	8	0	317
2009	887					394	76	33	21	15	7	20	14	4	303
2010	732						341	71	30	13	6	11	3	0	257
2011	668							329	69	18	6	10	9	1	226
2012	661								351	58	18	18	7	0	209
2013	877									460	76	49	16	3	273
2014	829										446	112	20	3	248
2015	713											480	29	7	197
2016	1308												801	43	464
2017	1322													662	660

Table S4. Primary period m-array for sanderling

Release year (*j*)	Number released (	Number re-encountered in year *k* (m_{jk})													
		2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	Never seen
2005	50	20	4	1	1	0	0	0	0	0	0	0	0	0	24
2006	99	0	52	1	2	2	0	1	0	1	2	0	0	0	38
2007	108	0	0	31	10	2	1	1	0	2	1	1	1	0	58
2008	107	0	0	0	34	4	6	6	0	2	2	0	0	0	53
2009	95	0	0	0	0	29	6	7	1	2	1	1	0	0	48
2010	68	0	0	0	0	0	17	11	2	1	0	0	1	0	36
2011	129	0	0	0	0	0	0	64	3	3	1	0	0	0	58
2012	206	0	0	0	0	0	0	0	56	24	10	5	0	1	110
2013	82	0	0	0	0	0	0	0	0	31	9	3	0	0	39
2014	134	0	0	0	0	0	0	0	0	0	55	14	3	0	62
2015	136	0	0	0	0	0	0	0	0	0	0	58	12	2	64
2016	186	0	0	0	0	0	0	0	0	0	0	0	100	8	78
2017	313	0	0	0	0	0	0	0	0	0	0	0	0	129	184

Posterior predictive checks

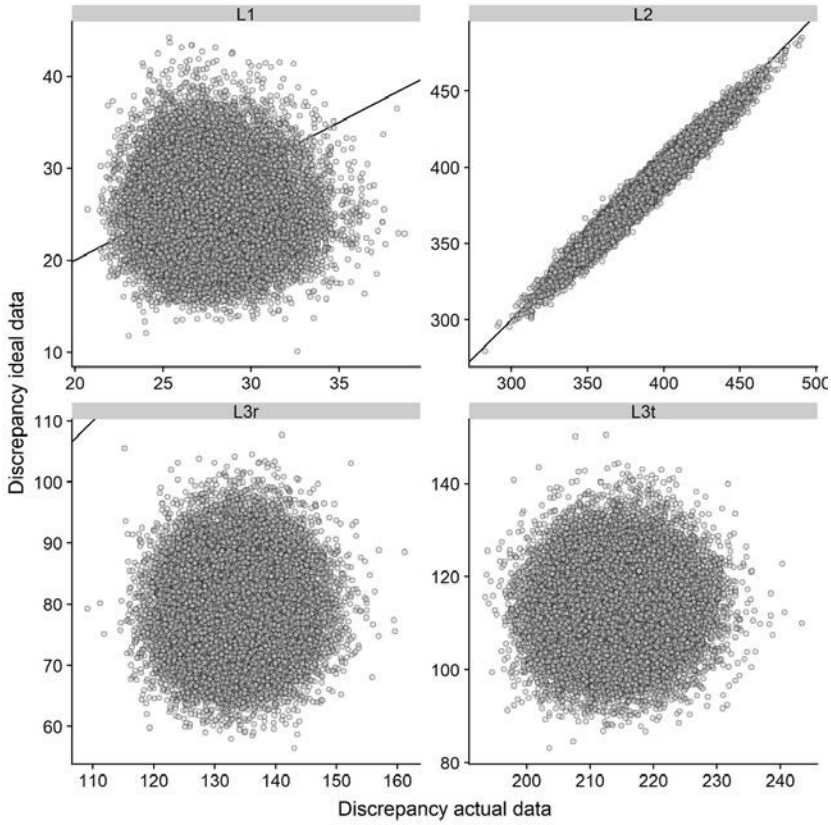

Figure S3. Posterior predictive checks for red knot using the Freeman-Tukey discrepancy statistic.

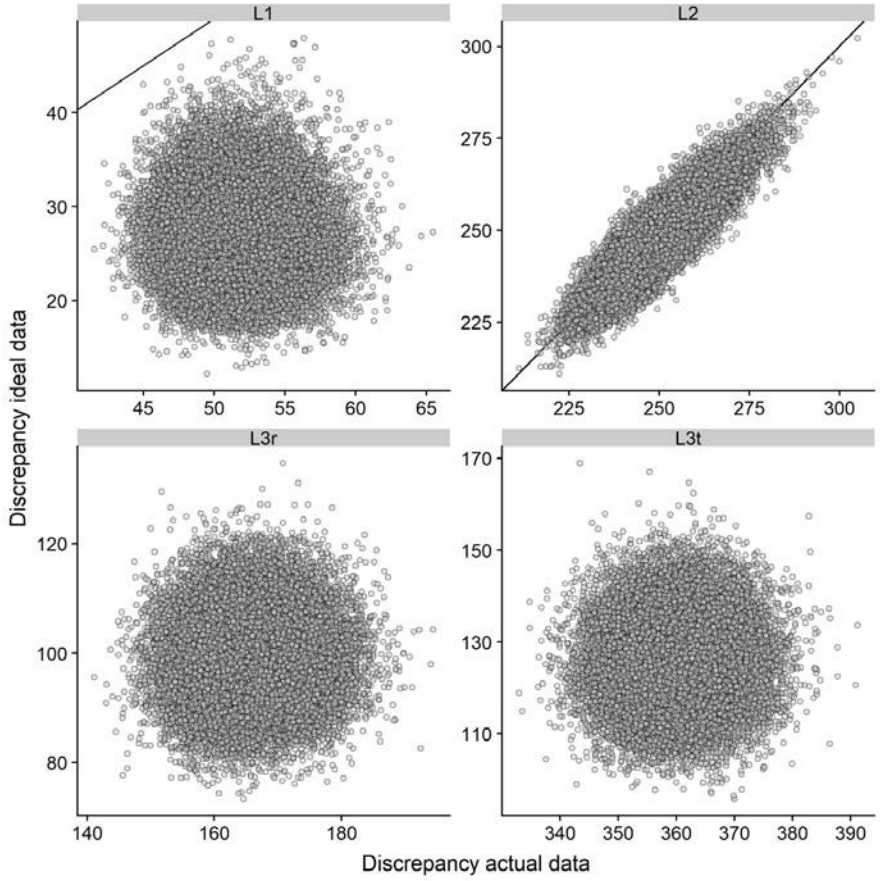

Figure S4. Posterior predictive checks for ruddy turnstone using the Freeman-Tukey discrepancy statistic.

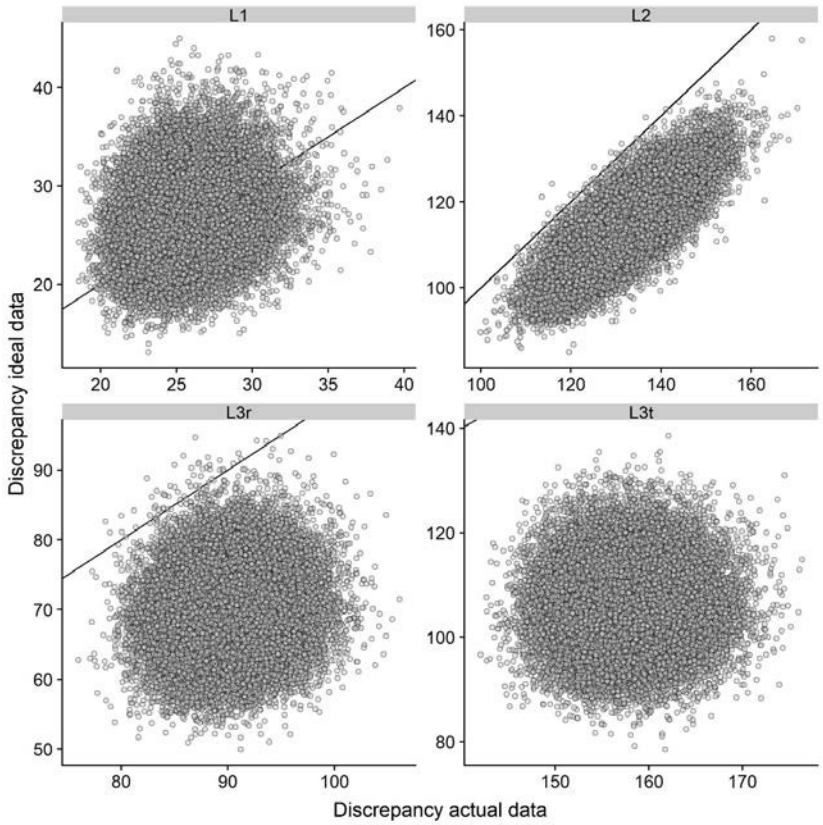

Figure S5. Posterior predictive checks for sanderling using the Freeman-Tukey discrepancy statistic.

Simulation study

We conducted a simulation study to confirm parameter identifiability and assess overall model performance. 100 datasets were stochastically simulated and fit to the open robust design model. We designed the simulation to be representative of our application to migratory shorebirds monitored during spring stopover. We simulated datasets for 10 years and 9 secondary occasions per year, with 500 new individuals marked per year. Values of the average survival probability, temporary emigration probabilities, transience probability, and detection probability were drawn from a Uniform distribution for each replicate to test model behavior across a range of possible parameter values (all data-generating values are given in Table 1). Year-specific probabilities on the logit scale were drawn from a normal distribution. For the simulation, we assumed that all new individuals marked in year *j* were first captured and marked prior to the first secondary occasion. For each simulation, we calculated the cell probabilities for the m-array of each of the likelihoods described above and used random draws from the multinomial distributions to generate pooled capture histories.

Simulated datasets were fit to the open robust design model using R 3.5.1 and Nimble 0.7.0 (R Core Team 2016, de Valpine et al. 2017). We generated three MCMC chains of 100,000 iterations each, discarded the first 50,000 as burn-in values, and kept every 5th sample, for a total of 10,000 samples from the posterior distribution. This was sufficient to ensure convergence of all parameters (all $\hat{R} < 1.1$ and confirmed with visual inspection of traceplots). To assess model performance, we calculated the mean relative bias and root mean square error of the posterior means of each parameter ($\hat{\theta}_j$ relative to the true data-generating values (θ_j) for *n* replicates.

$$bias = \frac{\sum_{j=1}^{n}(\hat{\theta}_j - \theta_j)/\theta_j}{n}$$

$$RMSE = \sqrt{\frac{\sum_{j=1}^{n}(\hat{\theta}_j - \theta_j)^2}{n}}$$

We also calculated the proportion of replicates in which the true value was within the 95% credible interval of the estimated posterior distribution. All code to simulate data and fit the model is available online at https://github.com/annamtucker/open-robust-design.

Table S5. Open robust design model parameters and parameterization. Weakly informative priors were provided for estimates of apparent annual survival probability (ϕ_j) and detection probability (p_{jt}) based on the literature and uninformative priors were provided for all other parameters.

Notation	Definition	Priors and parameterization
Estimated parameters		
ϕ_j	Probability of surviving from year *j* until year *j-1*, if alive in year *j*	$logit(\phi_j) = \mu^{\phi} + N(0,\sigma^{\phi})$ $ilogit(\mu^{\phi}) \sim Beta(5,2)$ $\sigma^{\phi} \sim Unif(0,10)$
γ_j^{II}	Probability of being present in the study area in year *j* if present in year *j-1*	$logit(\gamma_j^{II}) = \mu^{\gamma^{II}} + N(0,\sigma^{\gamma^{II}})$ $ilogit(\mu^{\gamma^{II}}) \sim Beta(3,3)$ $\sigma^{\gamma^{II}} \sim Unif(0,10)$
γ_j^{OI}	Probability of being present in the study area in year *j* if absent in year *j-1*	$logit(\gamma_j^{OI}) = \mu^{\gamma^{OI}} + N(0,\sigma^{\gamma^{OI}})$ $ilogit(\mu^{\gamma^{OI}}) \sim Beta(3,3)$ $\sigma^{\gamma^{OI}} \sim Unif(0,10)$
τ_j	Probability of remaining in the study area for more than one sampling occasion in year *j*	$logit(\tau_j) = \mu^{\tau} + N(0,\sigma^{\tau})$ $ilogit(\mu^{\tau}) \sim Beta(1,1)$ $\sigma^{\tau} \sim Unif(0,10)$
δ_{jt}	Probability of entering the study area after occasion *t-1* but just before occasion *t* in year *j*	$d_{j,t} \sim Gamma(1,1)$ $\delta_{j,t} = {d_{jt}}/{\sum_1^l d_{jt}}$
ψ_{jt}	Probability of remaining in the study area from occasion *t* until occasion *t-1* in year *j*, if present in occasion *t*	$logit(\psi_{jt}) = \mu_t^{\psi} + N(0,\sigma^{\psi})$ $ilogit(\mu_t^{\psi}) \sim Beta(1,1)$ $\sigma^{\psi} \sim Unif(0,10)$
p_{jt}	Probability of detection in occasion *t* of year *j*	$logit(p_{jt}) = \mu_j^{p} + N(0,\sigma^{p})$ $ilogit(\mu_j^{p}) \sim Beta(3,3)$ $\sigma^{p} \sim Unif(0,10)$
Derived parameters		
p_j^*	The probability of being detected at least once in primary period *j*, if present in the study area.	$p_j^* = \sum_{t=1}^{l} p_{jt}\,\omega_{jt}$
ω_{jt}	The probability of arriving before period *t* in year *j* and remaining in the study area undetected until period *t*	$\omega_{j1} = \delta_{j1}$ $\omega_{jt} = \omega_{jt-1}(1 - p_{jt-1})\psi_{jt-1} + \delta_{jt}$
a_{jt}	The probability of being present in the study area in period *t* of year *j*, if available in year *j*.	$a_{j1} = \delta_{j1}$ $a_{jt} = a_{jt-1}\psi_{jt-1} + \delta_{jt}$
π_{jt}	The proportion of the flyway population present in the study area in period *t* of year *j*.	$\pi_{jt} = \gamma_j^{II} a_{jt} + \gamma_j^{OI} a_{jt}$

Table S6. Data-generating input values, average bias and coverage of parameters in simulation study of 100 stochastically-generated datasets. For each replicate, an average parameter value was drawn from the Uniform distribution and year-specific values were drawn from a normal distribution on the logit scale. We assessed model performance by calculating the mean relative bias, root mean square error (RMSE), and proportion of the replicates in which the 95% credible interval contained the true value (posterior coverage).

Parameter	Simulated values	Average bias	RMSE	Posterior coverage
ϕ_j	Unif(0.6, 0.9)	0.01	0.06	0.95
γ_j^{II}	Unif(0.6, 0.9)	0.04	0.09	0.87
γ_j^{OI}	Unif(0.1, 0.5)	-0.01	0.12	0.95
τ_j	Unif(0.2, 0.5)	-0.02	0.06	0.94
δ_{jt}	Gamma(1,1)*	-0.01	0.13	0.93
ψ_{jt}	Unif(0.5, 0.95)	-0.02	0.31	0.63
p_{jt}	Unif(0.3, 0.7)	-0.02	0.15	0.75

*Secondary occasion arrival probabilities (δ_{jt}) were simulated and estimated for each year as:

$$d_{j,t} \sim Gamma(1,1), \quad \delta_{j,t} = {d_{jt}}/{\sum_1^I d_{jt}}$$

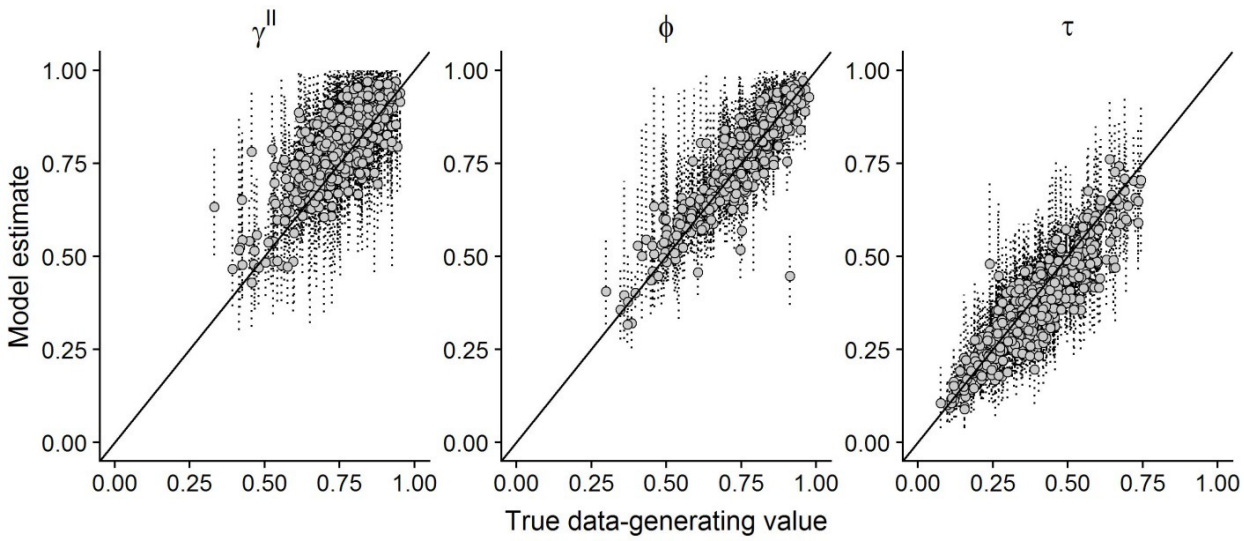

Figure S6. Model estimates and data-generating values for parameters describing among-year dynamics of survival (ϕ_j) and stopover site use (temporary emigration probabilities and stopover residency probability τ_j). The solid line indicates a perfect correlation between model estimates and true values. Gray points and vertical dashed lines show posterior means and 95% credible intervals from year-specific estimates from 100 replicated datasets.

Chapter 5: Estimating population growth rate and latent recruitment at a migratory stopover site using an integrated population model

Abstract

Consideration of the full annual cycle of migratory species can provide useful insight into ecological research and conservation efforts. Unfortunately, collecting data needed to statistically estimate demographic parameters is often logistically difficult. For example, for species that breed in remote areas monitoring is often conducted during migratory stopover or at nonbreeding sites and fecundity can be difficult to estimate directly. Integrated population models (IPMs) provide a method for inferring demographic rates of interest using data collected through existing monitoring efforts from other points of the annual cycle. Here we present an IPM that uses mark-resight and count data collected during migratory stopover to estimate population growth rate and recruitment for two species of Arctic-breeding shorebirds. This IPM includes an open robust design submodel to account for the flow-through dynamics of migratory stopover and adjust expected counts. We use this model to estimate adult survival probability, population growth rate, and recruitment rate using monitoring data collected during spring stopover in Delaware Bay, USA from 2005-2014. We found red knot and ruddy turnstone populations were most likely stable from 2005-2014; adult survival probabilities were consistently high (red knot: 0.90 (0.86, 0.93); ruddy turnstone: 0.93 (0.89, 0.96)) and average recruitment rate was 0.34 (0.01, 2.1) for red knot and 0.15 (0.001, 1.2) for ruddy turnstone. We found no strong association between vital rates and either Arctic conditions or stopover site conditions, but found that recruitment rate had the stronger effect on variation in population growth rate for both species.

Our approach demonstrates the utility of IPMs for understanding full annual cycle population dynamics, even when data are only available from one point of the annual cycle.

Introduction

Full annual cycle models help researchers identify sites or events during the annual cycle that have the strongest influence on variation in population growth rate (i.e., population dynamics)(Hostetler et al. 2015, Marra et al. 2015, Robinson et al. 2016). These models are useful for developing a better understanding of the intrinsic and extrinsic factors that govern population dynamics, particularly for migratory species. Identifying points that constrain population dynamics throughout the annual cycle is useful for developing effective conservation strategies and assessing the relative effects of habitat change or loss at nonbreeding and breeding sites (Flockhart et al. 2015, Robinson et al. 2016, Rushing et al. 2016, Taylor and Stutchbury 2016). Full annual cycle models can also help estimate carry-over effects between seasons, which may be important to long-term population stability (Norris et al. 2004, Norris and Marra 2007).

Integrated population models (IPMs) are a statistical method for combining multiple data sources for a fuller picture of population dynamics (Besbeas et al. 2002, Schaub and Abadi 2011). IPMs are lend themselves to full annual cycle applications because they are anchored on a description of the life cycle of the species, usually via a population matrix model (Caswell 2001). Integrating analyses of demographic data allows researchers to identify the demographic rates that have the strongest contribution to population growth rate, λ, (Robinson et al. 2014, Wilson et al. 2016, 2018) and the environmental covariates from throughout the annual cycle with the strongest indirect effects on λ (Rushing et al. 2017, Woodworth et al. 2017). Due to their

intrinsic consideration of the full life cycle of the species and flexibility at combining different sources of demographic information, IPMs are a good tool for supporting conservation decision-making and prioritization (Robinson et al. 2018, Zipkin and Saunders 2018), especially for species for which there are competing hypotheses about the most important drivers of observed declines (Morrison et al. 2016).

Arctic-breeding shorebirds (Charadriiformes: Charadrii) are declining globally and face many potential threats from throughout their complex annual cycle (Thomas et al. 2006, Andres et al. 2012, Hua et al. 2015). Population declines have been linked to both habitat loss and degradation at nonbreeding sites, particularly migratory stopover sites (Baker et al. 2004, Studds et al. 2017), and rapidly changing conditions on breeding grounds due to climate change (Wauchope et al. 2017). Some key stopover sites such as Delaware Bay, USA and the Yellow Sea in Asia are the focus of conservation management actions to ensure abundant food resources (McGowan et al. 2011b) and reduce habitat loss (Murray et al. 2014), respectively. A better understanding of the relative importance of stopover site and breeding site conditions will help predict the effectiveness of those conservation actions.

Conditions encountered at migratory stopover sites likely influence population dynamics via effects on adult survival, although empirical evidence is mixed. For example, variation in stopover site food resource availability has been linked to body mass gain during the stopover period and subsequent survival probability (Baker et al. 2004, Piersma et al. 2014, Duijns et al. 2017, Rakhimberdiev et al. 2018). However, McGowan et al. (2011a) found limited support for the hypothesis that departure mass at a stopover site strongly influences red knot *Calidris canutus rufa* survival and Senner et al. (2014) found that delays in migration of Hudsonian godwits *Limosa haemastica* had no negative effect on survival probability. Stopover site

conditions may also influence λ via carry-over effects on breeding success (Newton 2006). Although these species are mainly income breeders (Klaassen et al. 2001) that do not rely on stored energetic reserves for breeding, stopover site conditions could influence the timing of migration (Duijns et al. 2017, Rakhimberdiev et al. 2018) and therefore the timing of arrival to the breeding grounds. The breeding season in the Arctic is short and earlier breeding is associated with greater reproductive success (Weiser et al. 2018b).

The Arctic ecosystem and food web are experiencing drastic changes due to climate change (Legagneux et al. 2012), which may influence shorebird reproductive output due to changes in habitat availability (Wauchope et al. 2017, Flemming et al. 2019), timing of resource pulses (McKinnon et al. 2012), and predation pressure (Kubelka et al. 2018). Again, empirical evidence for an effect of breeding site conditions on reproductive output is mixed, with several studies finding neutral or positive effects of warmer summer conditions on hatching success and clutch size (Smith et al. 2010, Aharon-Rotman et al. 2015, Weiser et al. 2018a).However, others found evidence that warmer years led to asynchrony between nesting and food pulses, leading to decreased chick growth rates and body size (McKinnon et al. 2012) and negative effects on juvenile survival (van Gils et al. 2016). More compelling is the evidence that warmer summers are associated with increased predation pressure, possibly due to the disruption of the population cycles of lemmings, an alternative prey (Fraser et al. 2012, McKinnon et al. 2014, Aharon-Rotman et al. 2015, McGowan 2015) or increased predator density (Kubelka et al. 2018, Weiser et al. 2018a). Empirical estimates of fecundity are difficult to obtain for these species because they nest in remote area at very low densities and chicks are mobile within a day of hatching (Meltofte et al. 2008).

A unified analysis that jointly estimates adult survival probability and recruitment can identify the relative contribution of each vital rate to temporal variation in λ and estimate the indirect effects of stopover site conditions (e.g., food availability) and breeding site conditions (e.g., timing of snowmelt, predation pressure) on population growth rate. Here we present an integrated population model that uses mark-recapture data and counts conducted during northbound migratory stopover at Delaware Bay. This stopover site has been recognized as a globally important refueling site for migratory shorebirds. (Clark et al. 1993). Migratory stopover sites play a critical role in the annual cycle of long-distance migrants (Newton 2004, Alerstam 2011) and are the focus of many monitoring programs. Analysis of stopover population monitoring data can be complicated by the occurrence of transients (Salewski et al. 2007) and high turnover of individuals throughout the season (Lyons et al. 2015, Drever and Hrachowitz 2017). Additionally, the same proportion of the flyway population may not use a given site every year, which could bias population estimates that are made from counts at that site (Chapter 4, Bart et al. 2007, Smith et al. 2012).

We developed an IPM that accounts for interannual variation in site use using an open robust design mark-recapture model (Kendall and Bjorkland 2001, Kendall et al. 2019), which simultaneously estimates annual variation in survival and dynamics of site use (Chapter 4). We use a simulation study to assess the performance of our models and the consequences of ignoring flow-through dynamics when analyzing counts from a migratory stopover site. We apply this model to evaluate the key drivers of population dynamics for two species of migratory shorebirds. This modeling framework enables us to use existing long-term monitoring data from a single site to infer full annual cycle population dynamics for species of conservation concern.

Methods

Focal species, study site, and data collection

The Delaware Bay, located on the Atlantic coast of the United States between Delaware and New Jersey, is a globally-important spring stopover site for shorebirds in the Atlantic flyway, with hundreds of thousands of shorebirds stopping to refuel on beaches in this area during spring migration (Clark et al. 1993). Migrating birds rely on abundant resources present on the sandy beaches and mudflats throughout the Bay—and notably on the eggs of spawning horseshoe crabs *Limulus polyphemus* (Botton et al. 1994, Tsipoura and Burger 1999, Karpanty et al. 2006)—to replenish fat necessary to continue migration to their Arctic breeding grounds.

We analyzed population dynamics for red knot and ruddy turnstone (*Arenaria interpres*). Both of these species breed in the Arctic tundra and spend the nonbreeding season at sites along the Gulf Coast of North America, the Caribbean, and the Atlantic and Pacific coasts of South America and typically do not migrate north or breed until their second year (Nettleship 2000, Baker et al. 2001). These small- to medium-sized shorebirds feed in mixed flocks on sandy beaches and intertidal mudflats during the nonbreeding season, but ruddy turnstones also use rocky shores and manmade structures (e.g. breakwaters, jetties). Their diet during the nonbreeding season consists of a variety of invertebrates found on the sand surface or by probing the top layer of sediment (Gillings et al. 2007). Horseshoe crab eggs are a key fat-rich resource consumed by all three species during spring stopover in Delaware Bay (Botton et al. 1994, Gillings et al. 2007, Haramis et al. 2007).

Shorebirds were monitored during spring migration in Delaware Bay from 2005-2014. Monitoring effort consisted of capture and marking, resighting marked individuals, and aerial surveys to count the number of each species using the region each year. Birds were captured in mixed species foraging flocks primarily using cannon nets, and a subset of individuals were

marked with a field-readable Darvic leg flag inscribed with a unique field-readable alphanumeric code (Clark et al. 2005). Throughout the season in each year, observers visited survey beaches at least once within predefined 3-day sampling periods to scan flocks with spotting scopes and record observations of individual birds. We included 55,100 resightings of 4,932 marked red knot and 66,528 resightings of 5,421 marked ruddy turnstone in this analysis. We filtered the resighting data by removing all impossible records of flags that were either experimentally withheld or not yet deployed (Chapter 2).

Integrated population model

Matrix projection model

We used a two-stage matrix model in which the number of juveniles (superscript *J*) and adults (superscript *A*) alive in the spring of year *j* depends on the number of individuals of each age class in the previous year (N_j^J and N_j^A), the annual survival for each age class (ϕ_j^J and ϕ_j^A), and fecundity (f_j). We developed this model considering a spring stopover site (i.e., a pre-breeding census), and therefore the fecundity parameter is the product of the average number of offspring produced per breeding adult and first-winter survival of newly-hatched birds. This pre-breeding census allowed us to us a two-stage matrix despite delayed maturation and with a 2-year immature/juvenile stage.

$$\begin{bmatrix} N_{j+1}^J \\ N_{j+1}^A \end{bmatrix} = \begin{bmatrix} 0 & f_j \\ \phi_j^J & \phi_j^A \end{bmatrix} * \begin{bmatrix} N_j^J \\ N_j^A \end{bmatrix}$$

We considered a species in which individuals do not reproduce or undergo spring migration until their second spring, and therefore only animals in the adult age class are observable. We estimated a recruitment parameter ρ_j as the product of fecundity and juvenile survival:

$$\rho_j = f_j * \phi^J_{j+1}$$

For a system in which juveniles migrate in their first year and are distinguishable from adults, juvenile survival and fecundity could be estimated separately. This recruitment rate operates on a two-year lag; the expected number of new recruits in year *j* (i.e., individuals migrating north to breed for the first time, denoted with superscript *R*) is equal to the total number of adults in the population in year *j-2* ($N_j = N^A_j + N^R_j$, $N^R_j = N_{j-2}\rho_{j-2}$). The expected number of returning adults in year *j* is equal to the total number of adults in *j-1* multiplied by the annual survival probability ($N^A_j = N_{j-1}\phi_{j-1}$). To account for demographic stochasticity and other unmeasured variation in realized survival and recruitment, we assumed the true number of new recruits was drawn from a Poisson distribution around the expected value, and that the true number of surviving adults was drawn from a Binomial distribution with number of trials set as the number of individuals alive at year j-1 and the probability equal to the adult annual survival probability.

$$N^R_j \sim Poisson(N_{j-2} * \rho_{j-2})$$

$$N^A_j \sim Binomial(N_{j-1}, \phi_{j-1})$$

Open robust design model for mark-recapture data

During migratory stopover there is a high rate of turnover of individuals and we cannot assume that the entire passage population is present on any given day. Additionally, temporary emigration (skipping the stopover site) could make some individuals unavailable in some years. To account for these dynamics, we explicitly modeled temporary emigration and the entry and departure of individuals from the stopover site within a year using an open robust design model (ORD), described in detail in Chapter 4 (Figure 1). Under this model, primary sampling occasions (years, denoted *j*) are divided into secondary sampling occasions (denoted *t*). We

assume that individuals can arrive and depart only once within a year, but that those arrivals and departures can occur during any secondary sampling occasion. The ORD model consists of three likelihoods. The first (L1) describes the process of re-encountering individuals across years, which is conditioned on the number of individuals released in each year and depends on the annual probabilities of survival (ϕ_j), temporary emigration (γ_j^{II} and γ_j^{OI}), and detection (p_j^*). We estimate two temporary emigration parameters: γ_i^{II} is the probability of being present at the site in year *j* if present in year *j-1* and γ_j^{OI} is the probability of presence if absent in year *j-1*. If temporary emigration occurs randomly, then this parameter could be collapsed into a single parameter $\gamma_j = \gamma_j^{II} = \gamma_j^{OI}$. The second likelihood (L2) describes the process of encountering individuals for the first time in secondary period *t* of year *j*. It also conditions on the total number of individuals seen in each year and depends on the probabilities of arrival (δ_{jt}), persistence (ψ_{jt}), detection (p_{jt}), and transience (τ_j). We estimated the probability of remaining at the site for at least two sampling occasions in year *j* (τ_j) and derived transience probability as $1 - \tau_j$. The third likelihood (L3) describes the process of re-encountering individuals during secondary occasions of each year. It conditions on the total number of individuals encountered in occasion *t* of year *j* and depends on the year- and period-specific probabilities of persistence, transience, and detection.

For the primary period likelihood (L1) we pooled capture histories of birds first captured and marked in Delaware and all subsequent resightings; the first capture in these cases was physical capture and all subsequent encounters were resightings in the field. To estimate within-year processes (secondary period likelihoods L2 and L3), we only used resightings of individuals after the year of first capture, since birds are unobservable via resighting before first capture

(before they are marked with a field-readable tag). Therefore, within-year dynamics were not estimable for the first year of monitoring (2005).

State space model for counts

When counting large groups of individuals both undercounting and over-counting can occur, so we modeled the observed count (C_i) using a Normal distribution centered around the true number of birds present with an observation error σ_i^{obs}. The magnitude of counting error is often related to the number of animals being counted, and we assumed that the observation error was linearly related to the total number of birds available to be counted. We estimated a single coefficient of variation parameter (CV), and derived the observation error for each count. We also estimated the probability of being available for detection in secondary period *t* of year *j* as described in Chapter 4 and used this probability to estimate the number of birds available to be counted during each sampling occasion.

$$N_i^{avail} = N_{year(i)} * \pi_{year(i),occ(i)}$$

$$\sigma_i^{obs} = CV * N_i^{avail}$$

$$C_i \sim Normal(N_i^{avail}, \sigma_i^{obs})$$

Where π_{jt} is the probability of being present at the stopover site in period *t* of year *j*, based on estimates of arrival, persistence, transience, and temporary emigration from the open robust design sub-model (Chapter 4).

Each year aerial surveys are conducted in Delaware Bay to count the number of each species of shorebird present (Niles et al. 2008). Surveys are conducted by the New Jersey Endangered and Nongame Species Program and typically occur in late May to capture the peak of spring migration, but timing varies slightly among years due to weather conditions and other

logistical constraints. In 2009 the survey flight path was expanded to include some areas not previously sampled (Dey et al. 2011), so we estimated the counting error (CV) separately for surveys 2005-2008 and those 2009-2014. The number of surveys per season ranged from 1-5. For each count *i*, occurring in occasion *t* of year *j*, we estimated the number of birds available to be counted and the count was assumed to have normal error as described above. We assumed that only breeding adults were available for counting; though some juvenile knots are observed at this site in some years, they tend arrive late in the season and in very small numbers.

Simulation study

We used simulated data sets to verify that our model produced unbiased results and to determine the effects of failing to account for stopover flow-through dynamics in the capture-recapture submodel. We compared estimates from the model described above to those from a model that used a Cormack-Jolly-Seber (CJS) capture-recapture submodel, in which all observations of an individual in a given year are collapsed into one detection/non-detection event. The CJS model estimates apparent annual survival probability conditioned on the first capture of an individual while accounting for imperfect detection (Cormack 1964, Williams et al. 2002). The open robust design (ORD) formulation requires a high volume of resighting data throughout the stopover period that may not be available in some systems, while the CJS model requires less intensive resighting effort. We simulated 100 data sets under the ORD model for a 10-year study with 1000 new individuals marked per year using input parameter values listed in Table S1. Details of the data simulation and R code are available in the Supporting Information. We then estimated survival, recruitment, and population growth rate using each model (IPM-CJS and IPM-ORD). For each model we assessed estimation accuracy and precision by calculating

the average bias and root mean square error (RMSE) of estimates relative to the true data-generating value. We also calculated the posterior coverage, which is the proportion of replicates in which the 95% credible interval of the estimate contained the true data-generating value.

Effects of environmental variables on demographic rates

Stopover site conditions

Stopover sites could play a role in limiting adult survival of long-distance migratory species (Piersma et al. 2014) and it has also been proposed that carry-over effects from stopover could influence breeding success (Newton 2006). We included two covariates related to conditions encountered during stopover: relative horseshoe crab egg availability and the occurrence of severe storm events. The relative availability of horseshoe crab eggs (key shorebird prey during stopover at this site) depends on both the overall spawning abundance and the timing of horseshoe crab spawn, which is related to water temperature (Chapter 3). For each year, we estimated the horseshoe crab egg availability (hsc) as the product of the overall spawning abundance for that year (Zimmerman et al 2016) and the proportion of spawn that occurred by peak shorebird arrival. Even if horseshoe crabs are abundant and spawning is not delayed, severe weather events before or during stopover could result in delays in shorebird mass gain and departure. During extreme weather events birds are unable to forage, horseshoe crab spawning is stalled by high wave activity, and egg deposits already present on beaches may be washed away (Nordstrom et al. 2006, Smith et al. 2011). We obtained storm data from the National Oceanic and Atmospheric Administration Severe Weather Inventory and Storm Events Database (https://www.ncdc.noaa.gov/stormevents/) and determined whether a severe event occurred in Delaware Bay during the stopover period (May 1-30) in each year. Severe weather

events (as defined by NOAA) included coastal flood, hail, heavy rain, high surf, storm surge, or one of three high wind categories. We used a binary covariate for each year to indicate whether or not a severe weather event occurred during the stopover period.

Breeding site conditions

Environmental conditions in the Arctic, especially the timing of snowmelt, influence the timing of nesting and availability of food resources (Smith et al. 2010), and may also be associated with adult survival (McGowan et al. 2011a). We obtained Arctic weather data from four weather stations within the known breeding ranges of the three focal species from the Environment Canada Historical Weather Database (http://climate.weather.gc.ca/historical_data.html). Temperature and snow depth data was obtained from the Taloyoak, Cambridge Bay, Resolute, and Coral Harbour stations, which were chosen because they are within the focal species breeding ranges and had records for the length of our study. We extracted three metrics of breeding season weather conditions for each year: average temperature during June and July, snowmelt date, and snow depth on 10 June. We defined date of snow melt as the median first day with snow depth equal to 0 cm across all four weather stations each year. Snow depth was calculated as the average snow depth across all four stations on 10 June. Snow depth was assessed on 10 June to correspond with the time when most of the population is thought to have arrived at the breeding site and are preparing for nesting (Smith et al. 2010, McGowan et al. 2011a). Snowmelt date and mean temperature were positively correlated ($r = 0.70$) so we fit models separately with either snowmelt date or mean temperature. Snow depth on 10 June was not strongly correlated with either snowmelt date or mean temperature ($r = 0.10$ and $r = 0.31$, respectively).

We were also interested in ecological variables that could influence recruitment, specifically predation pressure. Lemmings are small arcivoline rodents that occur throughout the breeding range of these species and exhibit strong 3-4 yr population cycles (Fauteux et al. 2015). It has been proposed that lemming population cycles could influence shorebird reproductive output by saturating predators in years of high abundance, relieving predation pressure on shorebird nests and young (McKinnon et al. 2014). There is some evidence that these cycles have been disrupted in the past ~20 yr, which may be linked to concurrent shorebird population declines (Fraser et al. 2012, Nolet et al. 2013, Aharon-Rotman et al. 2015, McGowan 2015). We obtained estimates of summer densities of brown and collared lemmings brown (*Lemmus trimucronatus* and *Dicrostonyx groenlandicus*)from long-term study on within the breeding range of our focal species (Fauteux et al. 2015), and used this as a metric of lemming abundance in each year (Table S2). Densities from wet and mesic habitats were highly correlated ($r = 0.83$) and our focal shorebird species tend to nest in dryer habitats (Nettleship 2000, Macwhirter et al. 2002, Lathrop et al. 2018) so we used lemming densities from mesic habitat only.

Nonbreeding site conditions

Nonbreeding sites for these species include coastal areas as far north as the southern Atlantic and Gulf coasts of the United States and as far south as Tierra del Fuego in Argentina. Additionally, spatial and temporal use of these sites, and the relative use of different nonbreeding sites by migrants that stop in Delaware Bay is not well established (Nettleship 2000, Baker et al. 2001, Atkinson et al. 2006). The El Niño Southern Oscillation (ENSO) influences global weather patterns and coastal productivity; strong ENSO warm phases have been linked to decreased body mass and overwinter survival of shorebird and seabirds (Duffy 1993, Butler et al. 2001, O'Hara

et al. 2007). We used the Multivariate ENSO Index (MEI) from the National Oceanic and Atmospheric Administration's Earth System Research Laboratory (Wolter and Timlin 1993, 1998), available online at https://www.esrl.noaa.gov/psd/enso/mei/, to estimate the effect of ENSO events on apparent annual survival probability and recruitment. We assumed that if ENSO events had a strong effect on survival, such an effect would be detectable across all nonbreeding sites. We used the MEI historic ranks presented by NOAA to determine whether a moderate or stronger ENSO warm phase (ranks 56-69) occurred during each bimonthly period from 2004-2013. We assigned a dummy variable to each year that was equal to 1 if a strong El Niño event occurred in the months September-April (roughly corresponding with the nonbreeding season) following stopover and a 0 otherwise. For example, the year 2005 was considered an ENSO year because an ENSO event occurred between September 2005-April 2006.

Path analysis

We used path analysis to estimate the effects of environmental covariates on demographic rates and the direct and indirect effects of each vital rate and covariate on λ (Woodworth et al. 2017). This method involves fitting a series of linear models to the output of each MCMC iteration after scaling and centering all parameters by subtracting the mean and dividing by the standard deviation. For annual survival probability, we estimated the effect of horseshoe crab egg availability (hsc), storm events (storm), Arctic snow depth on 10 June (snow), and ENSO events (ENSO). For annual recruitment rate, we estimated the effect of horseshoe crab egg availability, ENSO events, Arctic snow depth, and either mean temperature (temp) or date of first snowmelt (snowmelt). Initial analyses indicated that snowmelt date had a

stronger relationship with recruitment than mean temperature, so the final models for each demographic rate were:

$$\phi_j \sim hsc_j + storm_j + snow_j + ENSO_j$$

$$\rho_j \sim hsc_j + snow_j + snowmelt_j + lemming_j + ENSO_{j+1}$$

The standardized regression coefficients from the above models represent the direct effect of each ecological covariate on each vital rate. Next, we estimated the direct effect of each vital rate on λ using a third linear regression model:

$$\lambda_j \sim \phi_j + \rho_{j-1}$$

The standardized regression coefficients from this model represent the direct effect of each vital rate on population growth rate. The indirect effect of covariate *X* on population growth rate via vital rate *Y* was then calculated as the product of the coefficient for the effect of *X* on *Y* and the coefficient for the effect of *Y* on population growth rate (Woodworth et al. 2017).

MCMC estimation and model fit

Integrated population models for each species were specified and fit using Nimble 0.7.0 and R 3.5.1 (R Core Team 2016, de Valpine et al. 2017). We sampled three MCMC chains of 100,000 iterations each, discarded 50,000 as burn-in values and kept every 5th sample. This resulted in a total of 10,000 draws from the posterior distribution, which was sufficient to ensure convergence of all parameters ($\hat{R}$ > 1.1 and confirmed with visual inspection of traceplots).

Unified goodness-of-fit tests have not been developed for IPMs, so we assessed model fit for each sub-model independently using posterior predictive checks (Kéry and Schaub 2012), which involves simulating expected data sets and comparing expected and observed data. The open robust design sub-model was assessed using the Freeman-Tukey diagnostic and the state-

space model was assessed using χ^2. Details of goodness-of-fit for the open robust design submodel are provided in Chapter 4. Graphical posterior predictive checks are provided in the Supporting Information.

Results

Simulation study and goodness of fit

The IPM-CJS model produced negatively biased estimates of recruitment rate (average bias = -0.21, Table 2). Use of the IPM-ORD greatly reduced bias (average bias = -0.06), although estimates of recruitment were still imprecise (RMSE = 0.52, Table 2). Posterior predictive checks indicated evidence of overdispersion for the ruddy turnstone mark-resight model, but adequate model fit for red knot and for the count models for both species (Supporting Information Figure S5).

Vital rate estimates

Our analysis indicates that both red knot and ruddy turnstone populations were most likely stable to increasing from 2005-2014, with an average population growth rate over this period 1.07 (0.94, 1.2) for red knot and 0.99 (0.93, 1.1) for ruddy turnstone. However, year-specific estimates of lambda were right-skewed for red knot, with most of the posterior mass < 1 (Supporting Information Figure S4). Adult apparent annual survival probabilities were consistently high for both species (Figure 2), with an average of 0.90 (0.86, 0.93) for red knot and 0.93 (0.89, 0.96) for ruddy turnstone. On average, recruitment rates were greater for red knot than ruddy turnstone (red knot: 0.34 (0.01, 2.1); ruddy turnstone: 0.15 (0.001, 1.2)).

Effects of environmental variables

We found no evidence of a significant effect of any of the environmental covariates included in this analysis on estimates of adult survival probability or recruitment (Figure 3). There was some evidence that severe weather events during stopover may be associated with survival and snowmelt date may be associated with recruitment (Figure 3), but these effects were not strong. Both adult survival probability and recruitment rate had a significantly positive association with population growth rate, but the magnitude of this relationship was greater for recruitment rate, indicating a stronger contribution to temporal variation in population growth rate (red knot: β^{ϕ} = 0.27 (0, 0.92), β^{ρ} = 0.89 (0.14, 1.2); ruddy turnstone: β^{ϕ} = 0.54 (0.07, 1.2), β^{ρ} = 0.76 (0.04, 1.3)).

Discussion

We developed an integrated population model to estimate latent recruitment and population growth rate from data collected at a single migratory stopover site. Our simulation study demonstrated that failing to account for within-year dynamics of arrivals and departures from the site underestimated recruitment, likely due to unmodeled variation in the proportion of the population available to be counted in a given year. Our open robust design submodel IPM returned more precise estimates of latent demographic rates than models without an open robust design. Application of our open robust design IPM to two migratory shorebirds showed that both species were most likely stable from 2005-2014 and that variation in recruitment rate had the stronger effect on variation in population growth rate. This extension presents another use of integrated population models to infer latent demographic rates, leading to a better understanding of population dynamics for species of conservation concern.

Previous analyses in this system indicated that food availability during stopover was positively associated with apparent annual survival probability for red knot, but not ruddy

turnstone (Chapter 4). We did not find significant evidence of that relationship in this analysis, possibly because of the additional uncertainty included in estimation of survival in the integrated model. Other researchers have also found that stopover sites are critical for population stability of long-distance migrants (Newton 2006, Rakhimberdiev et al. 2018), and in particular that the availability of horseshoe crab eggs is critical for red knot at this site (Baker et al. 2004, Niles et al. 2008, McGowan et al. 2011a). Ruddy turnstone are less reliant on horseshoe crab eggs than red knot because of their diverse diet and generalist foraging strategy (Chapter 3, Gill 1986). However, both species showed consistently high survival probabilities, and we found that variation in recruitment rate may be more influential for changes in population growth rate.

We found no evidence of an effect of Arctic conditions on recruitment or adult survival for either species. We note that our recruitment parameter combines information from two years: the fecundity from year *j* and juvenile survival from year *j+1*. Available data did not permit us to isolate these events, and by including covariates specific to breeding site conditions in year *j* we assumed that any potential relationships with Arctic conditions were driven by a direct effect on fecundity or carry-over effect on juvenile survival. This model structure may have prevented us from detecting effects of Arctic weather and snow conditions on annual reproduction. It is also possible that the spatial or temporal scale of the covariates we included to describe Arctic conditions did not adequately capture conditions experienced by these birds. Additionally, estimates of any latent demographic rate are typically imprecise (Kéry and Schaub 2012) and we may not have had enough precision in our estimates of recruitment to capture an underlying relationship with covariates.

Our analysis provides some of the first estimates of recruitment for shorebirds in this flyway. Many current estimates of fecundity for Arctic-breeding shorebirds come from ratios of

juvenile to adult birds observed at nonbreeding sites. Estimates of the number of juveniles per adult range from 0-0.4 (Boyd and Piersma 2001, Rogers et al. 2005, Meltofte et al. 2008) but unknown heterogeneity in spatial and temporal use of those sites and how populations might segregate by age make interpretation of those ratios difficult (Robinson et al. 2005). McGowan et al. (2011b) described a three-stage female-based matrix model to project red knot abundance in Delaware Bay under different horseshoe crab management actions and used published estimates from European populations to parameterize fecundity as 0.40 female young per female. With their estimated first-year and juvenile survival at 70% and 90% of adult survival, respectively, those estimates would result in a recruitment rate of ~0.2 female recruits per female which aligns our average estimates of recruitment. These parameters and this model play an integral role in the adaptive management for horseshoe crab harvest in Delaware Bay and our estimates from these data in this flyway could be important for that adaptive management modeling (McGowan et al. 2015). There is little to no information in the literature about recruitment rate of ruddy turnstone in this flyway, and this analysis indicates that although the current rate is low, this population is likely stable. For both species, however, we note that there is substantial uncertainty in our estimates of recruitment, with estimates near 0 within the 95% credible interval for most years.

The recruitment rate estimated here represents new adults joining the population, which we assume to be local recruits but could also represent immigrants. Our sampling population is defined by use of Delaware Bay during spring migration. For red knot, it is thought that nearly all of the Atlantic flyway population (*rufa* subspecies) uses this site (Niles et al. 2008). For ruddy turnstone, however, migratory connectivity is not as well-established. This species is common along the Atlantic coast during migration and the nonbreeding season (Nettleship 2000), and it is

plausible that our low estimates of recruitment could reflect declining use of this site. Indeed, the lack of strong relationships between Arctic conditions and estimated recruitment rate suggests that this parameter may influenced by either nonbreeding season effects on juvenile survival or usage of Delaware Bay by new adults.

Nonbreeding stationary sites may also play a role in limiting population dynamics through adult and juvenile survival. We could not determine the nonbreeding origin of individuals captured in our study, but did not find that ENSO events had a strong effect on either adult survival or recruitment rate. These species overwinter as far north as the Atlantic coast of North America and as far south as Tierra del Fuego and drawing generalizations about nonbreeding conditions is therefore difficult. Studies of ruddy turnstone in other flyways have found very high overwinter survival for adults (Metcalfe and Furness 1985, Burton and Evans 1997). Many shorebirds make local and regional movements during the nonbreeding season as food availability changes (Evans 1976, Dondua et al. 2016, Fonseca et al. 2017), a behavior that may allow them to cope with fluctuations in nonbreeding season conditions.

We formulated this model for a system in which individuals do not breed until their second year, and so juveniles are not observable during monitoring. If juveniles are observable and distinguishable from adults during monitoring or if demographic data is available (e.g. nest monitoring), it may be possible to separately estimate juvenile survival and fecundity. Studies of migratory songbirds have found that juvenile survival has a large contribution to population growth rate (Robinson et al. 2014, Rushing et al. 2017). Efforts to include other sources of data to allow for estimation of juvenile survival will be useful in creating a fuller picture of stage structure and stage-based variation in vital rates for these species.

Even when ecological mechanisms cannot be identified, information about which vital rate (e.g., adult survival, recruitment) has the stronger relationship with population growth rate can direct future research and conservation efforts. We found that recruitment rate had the stronger contribution to variation in population growth rate. For long-lived species adult survival rate is typically the most elastic demographic rate (Caswell 2001); however, survival probability was high (0.9) and fairly constant over time for both species and there may be limited opportunity to increase survival rates. If adult survival were to precipitously decline, as has been observed for similar species in other flyways (Piersma et al. 2014), we would expect even more drastic population declines. Long-term stopover monitoring programs such as this one are valuable for tracking population dynamics over time (Robinson et al. 2005), but future conservation efforts for these species may be more fruitful if directed toward refining estimates of annual fecundity and juvenile survival and understanding factors contributing to low realized recruitment.

One of the strengths of data integration models is the ability to combine information-rich data streams with relatively low-information data streams to obtain a fuller picture of ecological dynamics (Miller et al. 2019). The use of the open robust design submodel in this IPM allowed us to use additional information about within-season dynamics from the information-rich mark-resight dataset to adjust the expected count from aerial surveys. Failing to account for inter-annual variation in the flow of animals through the site led to negative bias in estimates of recruitment, but inclusion of information about the probability of presence during counts led to unbiased, if imprecise, estimates. This extension demonstrates a novel use of stopover site monitoring data in an IPM framework and could be applied to any system where animals are counted and marked during migration, including many taxa of birds (Buler and Moore 2011,

Crewe et al. 2016), migratory fish (Roni et al. 2012), and possibly some insects (Gibbs et al. 2006) and ungulates (Sawyer and Kauffman 2011). For species that are difficult to study throughout the annual cycle, estimates of many vital rates are lacking in the literature. This modeling framework demonstrates how integrated models can be used to understand full annual cycle population dynamics, even when data are only available from a single point of the annual cycle.

Literature cited

Aharon-Rotman, Y., M. Soloviev, C. Minton, P. Tomkovich, C. Hassell, and M. Klaassen. 2015. Loss of periodicity in breeding success of waders links to changes in lemming cycles in Arctic ecosystems. Oikos 124:861–870.

Alerstam, T. 2011. Optimal bird migration revisited. Journal of Ornithology 152.

Andres, B. a., P. a. Smith, R. I. G. Morrison, C. L. Gratto-trevor, S. C. Brown, and C. a. Friis. 2012. Population estimates of North American shorebirds, 2012. Wader Study Group Bulletin 119:178–194.

Atkinson, P. W., A. J. Baker, and K. A. Bennett. 2006. Using stable isotope ratios to unravel shorebird migration and population mixing: a case study with Red Knot Calidris canutus. Waterbirds around the world:960.

Baker, A., P. Gonzalez, R. I. G. Morrison, and B. A. Harrington. 2001. Red Knot (Calidris canutus). Page The Birds of North America. Cornell Lab of Ornithology, Ithaca.

Baker, A. J., P. M. González, T. Piersma, L. J. Niles, I. D. L. S. do Nascimento, P. W. Atkinson, N. a Clark, C. D. T. Minton, M. K. Peck, and G. Aarts. 2004. Rapid population decline in red knots: fitness consequences of decreased refuelling rates and late arrival in Delaware Bay. Proceedings. Biological sciences / The Royal Society 271:875–882.

Bart, J., S. Brown, B. Harrington, and R. I. Guy Morrison. 2007. Survey trends of North American shorebirds: Population declines or shifting distributions? Journal of Avian Biology 38:73–82.

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

Botton, M. L., R. E. Loveland, and T. R. Jacobsen. 1994. Site selection by migratory shorebirds in Delaware Bay, and its relationship to beach characteristics and abundance of horseshoe crab (Limulus polyphemus) eggs. The Auk 111:605–616.

Boyd, H., and T. Piersma. 2001. Changing balance between survival and recruitment explains population trends in Red Knots Calidris canutus islandica wintering in Britain, 1969-1995. Ardea 89:301–317.

Buler, J. J., and F. R. Moore. 2011. Migrant-habitat relationships during stopover along an ecological barrier: Extrinsic constraints and conservation implications. Journal of Ornithology 152:101–112.

Burton, N. H. K., and P. R. Evans. 1997. Survival and winter site-fidelity of Turnstones Arenaria interpres and Purple Sandpipers Calidris maritima in northeast England. Bird Study 44:34–44.

Butler, R. W., N. C. Davidson, and R. I. G. Morrison. 2001. Global-scale shorebird distribution in relation to productivity of near-shore ocean waters. The International Journal of Waterbird Biology 24:224–232.

Caswell, H. 2001. Matrix population models: construction, analysis, and interpretation. 2nd edition. Sinauer Associates, Inc., Sunderland, MA, USA.

Clark, E., L. J. Niles, and J. Burger. 1993. Abundance and distribution of migrant shorebirds in Delaware Bay. The Condor 95:694–705.

Clark, N., S. Gillings, A. Baker, P. Gonzalez, and R. Porter. 2005. The production and use of permanently inscribed leg flags for waders. Wader Study Group Bulletin 108:38–41.

Cormack, R. M. 1964. Estimates of survival from the sighting of marked animals. Biometrika 51:429–438.

Crewe, T. L., P. D. Taylor, and D. Lepage. 2016. Temporal aggregation of migration counts can improve accuracy and. Avian Ecology and Conservation 11.

Dey, A. D., L. J. Niles, H. P. Sitters, K. Kalasz, and R. I. G. Morrison. 2011. Update to the status of the Red Knot (Calidris canutus) in the Western Hemisphere, April 2011. Draft. Studies in Avian Biology 36.

Dondua, A., M. Robards, R. Porter, D. Solovyeva, and R. Bentzen. 2016. Large-scale movements of Dunlin breeding in Chukotka, Russia, during the non-breeding period. Wader Study 123.

Drever, M. C., and M. Hrachowitz. 2017. Migration as flow: using hydrological concepts to estimate residence time of migrating birds from daily counts. Methods in Ecology and Evolution.

Duffy, D. C. 1993. Stalking the Southern Oscillation: environmental uncertainty, climate change, and North Pacific seabirds. Pages 61–67 *in* D. S.-C. K. Vermeer, K.T. Briggs, K.H. Morgan, editor. The status, ecology, and conservation of marine birds of the North Pacific. Canadian Wildlife Service Special Publication, Ottawa.

Duijns, S., L. J. Niles, A. Dey, Y. Aubry, C. Friis, S. Koch, A. M. Anderson, and P. A. Smith. 2017. Body condition explains migratory performance of a long-distance migrant. Proceedings of the Royal Society B.

Evans, P. R. 1976. Energy Balance and Optimal Foraging Strategies in Shorebirds: Some Implications for Their Distributions and Movements in the Non-Breeding Season. Ardea 38–90:117–139.

Fauteux, D., G. Gauthier, and D. Berteaux. 2015. Seasonal demography of a cyclic lemming population in the Canadian Arctic. Journal of Animal Ecology 84:1412–1422.

Flemming, S. A., E. Nol, L. V Kennedy, P. A. Smith, and S. A. Flemming. 2019. Hyperabundant

herbivores limit habitat availability and influence nest site selection of Arctic- - breeding birds:1–12.

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

Fonseca, J., E. Basso, D. Serrano, and J. G. Navedo. 2017. Effects of tidal cycles on shorebird distribution and foraging behaviour in a coastal tropical wetland: Insights for carrying capacity assessment. Estuarine, Coastal and Shelf Science 198:279–287.

Fraser, J., S. Karpanty, J. Cohen, and B. Truitt. 2012. The Red Knot (Calidris canutus rufa) decline in the western hemisphere: is there a lemming connection? Canadian Journal of Zoology 16:13–16.

Gibbs, D., R. Walton, L. Brower, and A. K. Davis. 2006. Monarch Butterfly (Lepidoptera: Nymphalidae) Migration Monitoring at Chincoteague, Virginia and Cape May, New Jersey: A Comparison of Long-term Trends. Journal of the Kansas Entomological Society 79:156–164.

Gill, R. E. 1986. What won't Turnstones eat? British Birds 79:401–402.

Gillings, S., P. W. Atkinson, S. L. Bardsley, N. a. Clark, S. E. Love, R. a. Robinson, R. a. Stillman, and R. G. Weber. 2007. Shorebird predation of horseshoe crab eggs in Delaware Bay: Species contrasts and availability constraints. Journal of Animal Ecology 76:503–514.

van Gils, J. A., S. Lisovski, W. Meissner, T. Lok, A. Ożarowska, J. de Fouw, M. Y. Soloviev, T. Piersma, and M. Klaassen. 2016. Body shrinkage due to warming Arctic warming reduces red knot fitness in tropical wintering range. Science 352:819–821.

Haramis, G. M., W. a Link, P. C. Osenton, D. B. Carter, R. G. Weber, N. A. Clark, M. A. Teece, and D. S. Mizrahi. 2007. Stable isotope and pen feeding trial studies confirm the value of horseshoe crab Limulus polyphemus eggs to spring migrant shorebirds in Delaware Bay. Journal of Avian Biology 38:367–376.

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

Hua, N., K. Tan, Y. Chen, and Z. J. Ma. 2015. Key research issues concerning the conservation of migratory shorebirds in the Yellow Sea region. Bird Conservation International 25:38–52.

Kendall, W. L., and R. Bjorkland. 2001. Using open robust design models to estimate temporary emigration from capture-recapture data. Biometrics 57:1113–1122.

Kendall, W. L., S. Stapleton, G. C. White, J. I. Richardson, K. N. Pearson, and P. Mason. 2019. A multistate open robust design: population dynamics, reproductive effort, and phenology of sea turtles from tagging data. Ecological Monographs 89.

Kéry, M., and M. Schaub. 2012. Bayesian population analysis using WinBUGS: A hierarchical perspective. Academic Press, Waltham, Massachusetts.

Klaassen, M., A. Lindström A, H. Meltofte, and T. Piersma. 2001. Arctic waders are not capital breeders. Nature 413:794.

Kubelka, V., M. Salek, P. Tomkovich, Z. Vegvari, R. P. Freckleton, and T. Szekely. 2018. Global pattern of nest predation is disrupted by climate change in shorebirds. Science 683:680–683.

Lathrop, A. R. G., L. Niles, P. Smith, M. Peck, A. Dey, R. G. Lathrop, L. Niles, P. Smith, M. Peck, A. Dey, R. Sacatelli, and J. Bognar. 2018. Mapping and modeling the breeding habitat of the Western Atlantic Red Knot (Calidris canutus rufa) at local and regional scales. The Condor 120:650–665.

Legagneux, P., G. Gauthier, D. Bertaux, J. Bety, M.-C. Cadiuex, F. Bilodeau, E. Bolduc, L. McKinnon, A. Tarroux, J.-F. Therrien, L. Morissette, and C. J. Krebs. 2012. Disentangling trophic relationships in a High Arctic tundra ecosystem through food web modeling. Ecology 93:1707–1716.

Lyons, J. E., W. L. Kendall, J. A. Royle, S. J. Converse, B. A. Andres, and J. B. Buchanan. 2015. Population size and stopover duration estimation using mark-resight data and Bayesian analysis of a superpopulation model. Biometrics.

Macwhirter, R. B., P. A.-S. Jr., and D. E. Kroodsma. 2002. Sanderling (Calidris alba). Page *in* A. F. Poole and F. B. Gill, editors. The Birds of North America. Second edition. Cornell Lab of Ornithology, Ithaca, NY.

Marra, P. P., E. B. Cohen, S. R. Loss, J. E. Rutter, and C. M. Tonra. 2015. A call for full annual cycle research in animal ecology. Biology letters 11:2015.0552.

McGowan, C. P. 2015. Comparing models of Red Knot population dynamics. The Condor 117:494–502.

McGowan, C. P., J. E. Hines, J. D. Nichols, J. E. Lyons, D. R. Smith, K. S. Kalasz, L. J. Niles, A. D. Dey, N. a. Clark, P. W. Atkinson, C. D. T. Minton, and W. Kendall. 2011a. Demographic consequences of migratory stopover: linking red knot survival to horseshoe crab spawning abundance. Ecosphere 2.

McGowan, C. P., D. R. Smith, J. D. Nichols, J. E. Lyons, J. Sweka, K. Kalasz, L. J. Niles, R. Wong, J. Brust, M. Davis, and B. Spear. 2015. Implementation of a framework for multi-species, multi-objective adaptive management in Delaware Bay. Biological Conservation 191:759–769.

McGowan, C. P., D. R. Smith, J. A. Sweka, J. Martin, J. D. Nichols, R. Wong, J. E. Lyons, L. J. Niles, K. Kalasz, and J. Brust. 2011b. Multispecies modeling for adaptive management of horseshoe crabs and red knots in the Delaware Bay. Natural Resource Modeling 24:117–156.

McKinnon, L., D. Berteaux, and J. Bêty. 2014. Predator-mediated interactions between lemmings and shorebirds: A test of the alternative prey hypothesis. The Auk 131:619–628.

McKinnon, L., M. Picotin, E. Bolduc, C. Juillet, and J. Bêty. 2012. Timing of breeding, peak food availability, and effects of mismatch on chick growth in birds nesting in the High Arctic. Canadian Journal of Zoology 90:961–971.

Meltofte, H., T. T. Høye, and N. M. Schmidt. 2008. Effects of Food Availability, Snow and Predation on Breeding Performance of Waders at Zackenberg. Advances in Ecological Research 40:325–343.

Metcalfe, N., and R. Furness. 1985. Survival, winter population stability and site fidelity in the turnstone Arenaria interpres. Bird Study 32:207–214.

Miller, D., K. Pacifici, J. S. Sanderlin, and B. Reich. 2019. The recent past and promising future for data integration methods to estimate species' distributions. Methods in Ecology and Evolution 2019:22–37.

Morrison, C. A., R. A. Robinson, S. J. Butler, J. A. Clark, and J. A. Gill. 2016. Demographic drivers of decline and recovery in an Afro-Palaearctic migratory bird population. Proceedings of the Royal Society B: Biological Sciences.

Murray, N. J., R. S. Clemens, S. R. Phinn, H. P. Possingham, and R. A. Fuller. 2014. Tracking the rapid loss of tidal wetlands in the Yellow Sea. Frontiers in Ecology and the Environment 12:267–272.

Nettleship, D. N. 2000. Ruddy Turnstone (Arenaria interpres). Page The Birds of North America. Cornell Lab of Ornithology, Ithaca.

Newton, I. 2004. Population limitation in migrants. Ibis 146:197–226.

Newton, I. 2006. Can conditions experienced during migration limit the population levels of birds? Journal of Ornithology 147:146–166.

Niles, L. J., H. P. Sitters, D. Dey, Amanda, P. W. Atkinson, A. J. Baker, K. A. Bennett, R. Carmona, K. E. Clark, N. A. Clark, C. Espoz, P. M. Gonzalez, B. A. Harrington, D. E. Hernandez, K. S. Kalasz, R. G. Lathrop, N. Matus, Ricardo, C. D. T. Minton, R. I. G. Morrison, M. K. Peck, W. Pitts, R. A. Robinson, and I. L. Serrano. 2008. Status of the Red Knot (Calidris Canutus Rufa) in the Western Hemisphere (Calidris Canutus Rufa) in. Studies in Avian Biology 36:1–185.

Nolet, B. A., S. Bauer, N. Feige, Y. I. Kokorev, I. Y. Popov, and B. S. Ebbinge. 2013. Faltering lemming cycles reduce productivity and population size of a migratory Arctic goose species. Journal of Animal Ecology 82:804–813.

Nordstrom, K. F., N. L. Jackson, D. R. Smith, and R. G. Weber. 2006. Transport of horseshoe crab eggs by waves and swash on an estuarine beach: Implications for foraging shorebirds. Estuarine, Coastal and Shelf Science 70:438–448.

Norris, D. R., and P. P. Marra. 2007. Seasonal interactions, habitat quality, and population dynamics in migratory birds. The Condor 109:535–547.

Norris, D. R., P. P. Marra, T. K. Kyser, T. W. Sherry, and L. M. Ratcliffe. 2004. Tropical winter habitat limits reproductive success on the temperate breeding grounds in a migratory bird. Proceedings. Biological sciences / The Royal Society 271:59–64.

O'Hara, P. D., B. J. M. Haase, R. W. Elner, B. D. Smith, and J. K. Kenyon. 2007. Are population dynamics of shorebirds affected by El Nino/Southern Oscillation (ENSO) while on their non-breeding grounds in Ecuador? Estuarine, Coastal and Shelf Science 74:96–108.

Piersma, T., T. Lok, Y. Chen, C. J. Hassell, H.-Y. Yang, A. Boyle, M. Slaymaker, Y.-C. Chan, D. S. Melville, Z.-W. Zhang, and Z. Ma. 2014. Simultaneous declines in summer survival of three shorebird species signals a flyway at risk. Journal Of Applied Ecology:manuscript.

R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rakhimberdiev, E., S. Duijns, Karagicheva, Julia, C. J. Camphuysen, V. Castricum, A. Dekinga, R. Dekker, A. Gavrilov, J. ten Horn, J. Jukema, A. Saveliev, M. Soloviec, T. L. Tibbitts, J. A. van Gils, and T. Piersma. 2018. Fuelling conditions at staging sites can mitigate Arctic warming effects in a migratory bird. Nature Communications.

Robinson, O. J., C. P. McGowan, and P. K. Devers. 2016. Disentangling density-dependent dynamics using full annual cycle models and Bayesian model weight updating. Journal of Applied Ecology:670–678.

Robinson, O. J., V. Ruiz-Gutierrez, D. Fink, R. J. Meese, M. Holyoak, and E. G. Cooch. 2018. Using citizen science data in integrated population models to inform conservation. Biological Conservation 227:293464.

Robinson, R. A., C. A. Morrison, and S. R. Baillie. 2014. Integrating demographic data: Towards a framework for monitoring wildlife populations at large spatial scales. Methods in Ecology and Evolution 5:1361–1372.

Robinson, R. a, N. a Clark, R. Lanctot, S. Nebel, B. Harrington, J. a Clark, J. a Gill, H. Meltofte, D. I. Rogers, K. E. N. G. Rogers, B. J. Ens, C. M. Reynolds, and P. W. Long. 2005. Long term demographic monitoring of wader populations in non-breeding areas. Wader Study Group Bulletin 106:17–29.

Rogers, D. I., K. G. Rogers, and M. A. Barter. 2005. Measuring recruitment of shorebirds with telescopes: a pilot study of age proportions on Australian non-breeding grounds. Status and conservation of shorebirds in the East Asian-Australasian Flyway: Proceedings of the Australasian Shorebirds Conference:63–72.

Roni, P., T. Bennett, R. Holland, G. Pess, K. Hanson, R. Moses, M. McHenry, W. Ehinger, and J. Walter. 2012. Factors affecting migration timing, growth, and survival of juvenile coho salmon in two coastal Washington watersheds. Transactions of the American Fisheries

Society 141:890–906.

Rushing, C. S., J. A. Hostetler, T. S. Sillett, P. P. Marra, J. A. Rotenberg, and T. B. Ryder. 2017. Spatial and temporal drivers of avian population dynamics across the annual cycle. Ecology 98:2837–2850.

Rushing, C. S., T. B. Ryder, P. P. Marra, and C. S. Rushing. 2016. Quantifying drivers of population dynamics for a migratory bird throughout the annual cycle. Proceedings of the Royal Society B: Biological Sciences 283:20152846.

Salewski, V., M. Thoma, and M. Schaub. 2007. Stopover of migrating birds: Simultaneous analysis of different marking methods enhances the power of capture-recapture analyses. Journal of Ornithology 148:29–37.

Sawyer, H., and M. J. Kauffman. 2011. Stopover ecology of a migratory ungulate. Journal of Animal Ecology 80:1078–1087.

Schaub, M., and F. Abadi. 2011. Integrated population models: a novel analysis framework for deeper insights into population dynamics. Journal of Ornithology 152:1–11.

Senner, N. R., W. M. Hochachka, J. W. Fox, and V. Afanasyev. 2014. An exception to the rule: Carry-over effects do not accumulate in a long-distance migratory bird. PLoS ONE 9.

Smith, D. R., N. L. Jackson, K. F. Nordstrom, and R. G. Weber. 2011. Beach characteristics mitigate effects of onshore wind on horseshoe crab spawning: Implications for matching with shorebird migration in Delaware Bay. Animal Conservation 14:575–584.

Smith, P. A., H. G. Gilchrist, M. R. Forbes, J. L. Martin, and K. Allard. 2010. Inter-annual variation in the breeding chronology of arctic shorebirds: Effects of weather, snow melt and predators. Journal of Avian Biology 41:292–304.

Smith, P. A., C. L. Gratto-Trevor, B. T. Collins, S. D. Fellows, R. B. Lanctot, J. Liebezeit, B. J. McCaffery, D. Tracy, J. Rausch, S. Kendall, S. Zack, and H. R. Gates. 2012. Trends in Abundance of Semipalmated Sandpipers: Evidence from the Arctic. Waterbirds 35:106–119.

Studds, C. E., B. E. Kendall, N. J. Murray, H. B. Wilson, D. I. Rogers, R. S. Clemens, K. Gosbell, C. J. Hassell, R. Jessop, D. S. Melville, D. A. Milton, C. D. T. Minton, H. P. Possingham, A. C. Riegen, P. Straw, E. J. Woehler, and R. A. Fuller. 2017. Rapid population decline in migratory shorebirds relying on Yellow Sea tidal mudflats as stopover sites. Nature Communications 8:1–7.

Taylor, C. M., and B. J. M. Stutchbury. 2016. Effects of breeding versus winter habitat loss and fragmentation on the population dynamics of a migratory songbird. Ecological Applications 26:424–437.

Thomas, G. H., R. B. Lanctot, and T. Székely. 2006. Can intrinsic factors explain population declines in North American breeding shorebirds? A comparative analysis. Animal

Conservation 9:252–258.

de Valpine, P., D. Turek, C. J. Paciorek, C. Anderson-Bergman, D. T. Lang, and R. Bodik. 2017. Programming With Models: Writing Statistical Algorithms for General Model Structures With NIMBLE. Journal of Computational and Graphical Statistics 26:403–413.

Wauchope, H. S., J. D. Shaw, Ø. Varpe, E. G. Lappo, D. Boertmann, R. B. Lanctot, and R. A. Fuller. 2017. Rapid climate-driven loss of breeding habitat for Arctic migratory birds. Global Change Biology 23:1085–1094.

Weiser, E. L., S. C. Brown, R. B. Lanctot, H. R. Gates, K. F. Abraham, R. L. Bentzen, J. Bêty, M. L. Boldenow, R. W. Brook, T. F. Donnelly, W. B. English, S. A. Flemming, S. E. Franks, H. G. Gilchrist, M. A. Giroux, A. Johnson, S. Kendall, L. V. Kennedy, L. Koloski, E. Kwon, J. F. Lamarre, D. B. Lank, C. J. Latty, N. Lecomte, J. R. Liebezeit, L. Mckinnon, E. Nol, J. Perz, J. Rausch, M. Robards, S. T. Saalfeld, N. R. Senner, P. A. Smith, M. Soloviev, D. Solovyeva, D. H. Ward, P. F. Woodard, and B. K. Sandercock. 2018a. Effects of environmental conditions on reproductive effort and nest success of Arctic-breeding shorebirds. Ibis:608–623.

Weiser, E. L., S. C. Brown, R. B. Lanctot, H. R. Gates, K. F. Abraham, R. L. Bentzen, J. Bêty, M. L. Boldenow, R. W. Brook, T. F. Donnelly, W. B. English, S. A. Flemming, S. E. Franks, H. G. Gilchrist, M. Giroux, A. Johnson, L. V Kennedy, L. Koloski, E. Kwon, J. Lamarre, D. B. Lank, N. Lecomte, J. R. Liebezeit, L. Mckinnon, E. Nol, J. Perz, J. Rausch, M. Robards, S. T. Saalfeld, N. R. Senner, P. A. Smith, M. Soloviev, D. Solovyeva, D. H. Ward, P. F. Woodard, and B. K. Sandercock. 2018b. Life history tradeoffs revealed by seasonal declines in reproductive traits of Arctic-breeding shorebirds. Journal of Avian Biolgy:1–16.

Williams, B., J. Nichols, and M. Conroy. 2002. Analysis and Management of Wildlife Populations. First edition. Academic Press, San Diego, California.

Wilson, S., K. C. Gil-Weir, R. G. Clark, G. J. Robertson, and M. T. Bidwell. 2016. Integrated population modeling to assess demographic variation and contributions to population growth for endangered whooping cranes. Biological Conservation 197:1–7.

Wilson, S., J. F. Saracco, R. Krikun, D. T. T. Flockhart, C. M. Godwin, and K. R. Foster. 2018. Drivers of demographic decline across the annual cycle of a threatened migratory bird. Scientific Reports 8:7316.

Wolter, K., and M. S. Timlin. 1993. Monitoring ENSO in COADS with a seasonally adjusted principal component index.

Wolter, K., and M. S. Timlin. 1998. Measuring the strength of ENSO events - how does 1997/98 rank? Weather 53.

Woodworth, B. K., N. T. Wheelwright, A. E. Newman, M. Schaub, and D. R. Norris. 2017. Winter temperatures limit population growth rate of a migratory songbird. Nature Communications 8:1–9.

Zipkin, E. F., and S. P. Saunders. 2018. Synthesizing multiple data types for biological conservation using integrated population models. Biological Conservation 217:240–250.

Tables

Table 5-11. Integrated population model parameter notation and definitions.
We use *j* to denote years, which are the primary sampling occasions, and *t* to denote the secondary 3-day sampling periods within each year.

Notation	Definition
ϕ_j	Adult apparent annual survival probability – the probability that an adult bird that is alive in year *j* will survive until year *j+1*
ρ_j	Recruitment– the per capita number of new recruits entering the population in year *j+2*
τ_j	Transience probability—the probability that an individual will remain at the stopover site for more than one secondary sampling period in year *j*
γ_j^{II}	Temporary emigration probability—the probability that an individual that is alive and present in year *j-1* will stop in Delaware Bay in year *j* given that it has survived
γ_j^{OI}	Temporary emigration probability—the probability that an individual that is alive in year *j* and not present in year *j-1* will stop in Delaware Bay in year *j* given that it has survived
p_j^*	Primary period detection probability – the probability that an individual alive and present at the site in year *j* is detected at least once
$\psi_{j,t}$	Stopover persistence probability—the probability that an individual present in secondary sampling period *t* will remain in the study area until secondary period *t+1*
$\delta_{j,t}$	Entry probability—the probability that an individual that has not yet entered the stopover site will enter in time *t*
$p_{j,t}$	Secondary period detection probability—the probability that an individual present in period *t* of year *j* is detected
λ_j	Population growth rate – the proportional change in population size from year *j* to year *j+1*
π_{jt}	Proportion of the flyway population present in the study area in period *t* of year *j*

Table 5-2. Results of simulation study comparing IPMs.
We compared performance of an IPM using a Cormack-Jolly-Seber mark-recapture model (CJS) that ignores within-season dynamics with an IPM that uses an open robust design model (ORD) to estimate the proportion of the population available for counting. ϕ^* and ρ^* are the overall average estimates for survival and recruitment, and parameters with subscript *j* are year-specific estimates. Posterior coverage is the proportion of the MCMC estimates in which the 95% credible interval of the posterior distribution contained the true data-generating value.

Mark-recapture submodel	Parameter	Relative bias	RMSE	Posterior coverage
CJS	ϕ^*	-0.07	0.07	0.78
	ρ^*	-0.54	0.68	0.50
	ϕ_j	0.06	0.08	0.60
	ρ_j	-0.21	1.8	0.77
	λ_j	-0.05	1.1	0.80
ORD	ϕ^*	0.05	0.06	0.84
	ρ^*	0.05	0.07	1
	ϕ_j	0.05	0.09	0.78
	ρ_j	-0.06	0.52	0.94
	λ_j	0.01	0.45	0.93

Table 5-3. Estimates of average survival, recruitment, and population growth rate for two species of migratory shorebirds from 2005-2014.
Average survival probability and recruitment rate were estimated directly. The average and variance of population growth rate were calculated using the geometric mean and geometric standard deviation of year-specific estimates for each MCMC iteration.

Species	Parameter	Average		Annual variance	
		Mean	95% CRI	Mean	95% CRI
Red knot	ϕ	0.90	0.86, 0.93	0.003	0, 0.011
	ρ	0.34	0.01, 2.1		
	λ	1.07	0.94, 1.2	1.5	1.1, 2.5
Ruddy turnstone	ϕ	0.93	0.89, 0.96	0.001	0, 0.007
	ρ	0.15	0.001, 1.2		
	λ	0.99	0.93, 1.1	1.2	1.05, 0.7

Figures

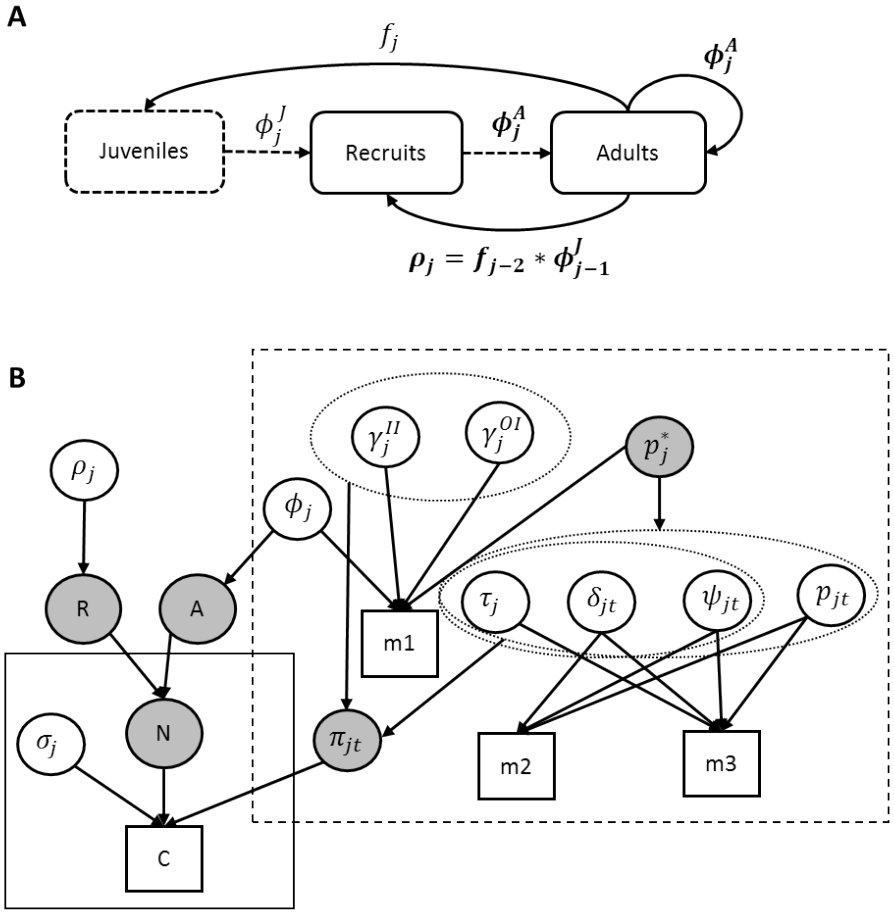

Figure 5-1. Conceptual diagram of the integrated population model.
This model is based on a two-stage life cycle with a pre-breeding census (**A**) in which individuals do not breed until their second year. We estimate a recruitment rate, ρ, which is a product of fecundity and juvenile survival and assume that new recruits have the same survival probability as returning adults. The integrated population model (**B**) uses an open robust design sub-model to estimate annual survival probability, ϕ, while accounting for temporary emigration and the proportion of the population available to be counted, π, based on estimated within-season dynamics of arrival and departure. In this diagram, white circles indicate estimated parameters, gray circles indicate derived parameters, and squares indicate data (C = counts, m1 = m-array for the primary periods, m2 = m-array for new encounters within secondary periods, and m3 = m-array for subsequent encounters within secondary periods). The open robust design sub-model components are in the dashed box and the state-space sub-model for counts is in the solid box. All parameters are defined in Table 1.

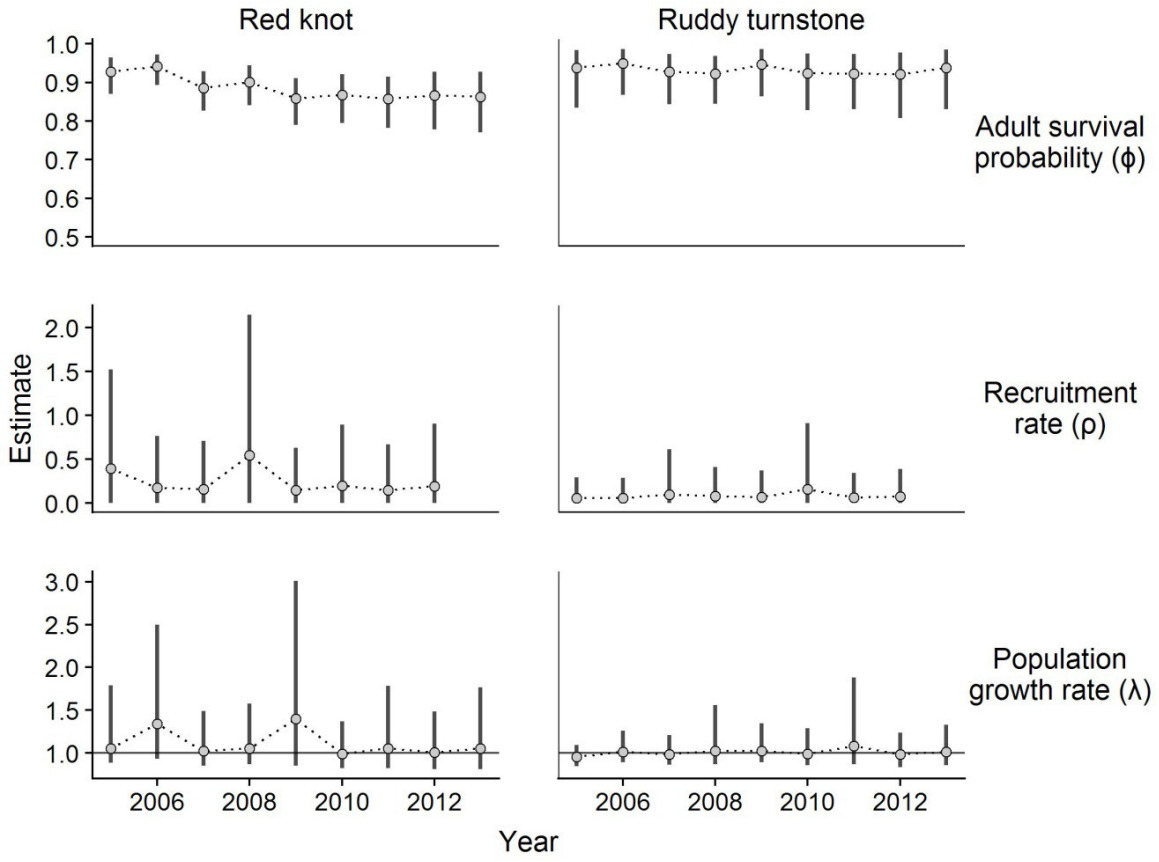

Figure 5-2. Estimates of survival, fecundity, and lambda over time for red knot and ruddy turnstone, 2005-2014.
Points represent posterior means and vertical lines represent 95% credible intervals. The horizontal shaded regions and dashed line represent the 95% credible intervals and posterior mean of the overall average estimate of each parameter.

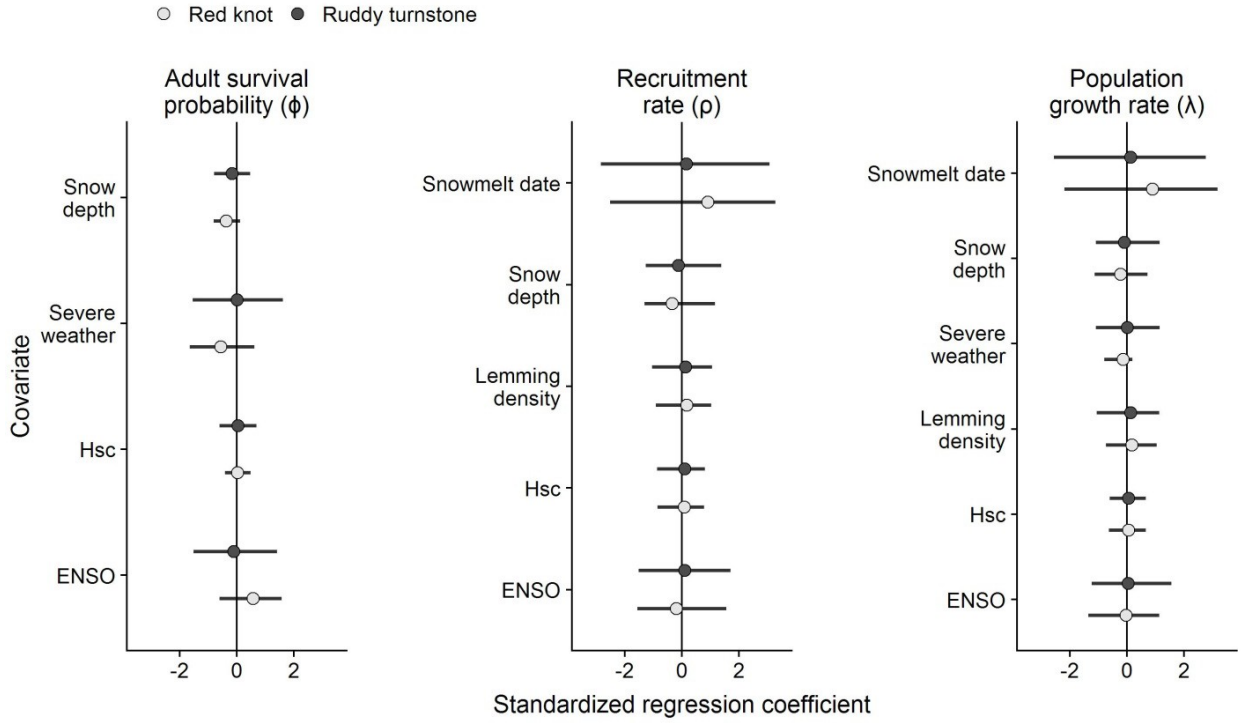

Figure 5-3. Direct and indirect effects of environmental variables on survival, recruitment, and population growth rate

Estimates are shown for red knot (light gray) and ruddy turnstone (dark gray). Points represent posterior means of the standardized regression coefficients and vertical lines represent 95% credible intervals. Direct and indirect effects were estimated using a series of standardized linear models, described in detail in text.

Supporting information

Simulation study

We simulated mark-resight and count data of a migratory population to compare performance of the integrated population models with either a Cormack-Jolly-Seber mark-recapture submodel or an open robust design submodel. We simulated 100 datasets for a 10-year study with 9 secondary sampling occasions and 1000 new individuals marked per year. For each replicate dataset, we drew true parameter values from the distributions described in Table S1. We first projected the population based on the input values of recruitment and adult survival probability. For each year, the number of new recruits (R) was drawn from a Poisson distribution and the number of surviving adults (S) was drawn from a binomial distribution.

$$R_j \sim Poisson(\rho_{j-2} * N_{j-2})$$

$$S_j \sim Binomial(N_{j-1}, \phi_{j-1})$$

$$N_j = R_j + S_j$$

We calculated the probability of presence in each occasion *t* of year *j* based on input values for entry ($\delta_{j,t}$), residency (τ_j), persistence ($\psi_{j,t}$), and temporary emigration (γ_j^{OI} and γ_j^{II}) probabilities. The probability of being present in secondary period *t* given that an individual is available and a resident in year *j* is denoted a_{jt} and is a function of the probabilities of entry and persistence in that year.

$$a_{jt} = \begin{cases} t = 1 & \delta_{jt} \\ t > 1 & \delta_{jt} + a_{jt-1}\psi_{jt-1} \end{cases}$$

To account for transients that do not remain after the first occasion, the stopover residency probability is included for each period before the current period *t*, which can be calculated as:

$$z_{jt} = \tau_j(a_{jt} - \delta_{jt}) + \delta_{jt}$$

In each year, the probability of being available in year *j* is the sum of the probabilities of being present given that an individual was unavailable (γ_j^{OI}) or available (γ_j^{II}) in the previous year. The probability of being present in secondary occasion *t* of year *j* is therefore:

$$\pi_{jt} = \gamma_j^{II} z_{jt} + \gamma_j^{OI} z_{jt}$$

We simulated a single count per year, and the counting occasion was randomly selected from all occasions. The number of individuals available to be counted was drawn from a binomial distribution and the counting error (σ_j^{obs}) was calculated based on the value of the counting coefficient of variation (CV) for that replicate.

$$N_j^{avail} \sim Binomial(N_j, \pi_{j,occ(j)})$$

$$\sigma_j^{obs} = CV * N_j^{avail}$$

$$Count_j \sim Normal(N_j^{avail}, \sigma_j^{obs})$$

Mark-recapture data was simulated by calculating the cell probabilities for the m-arrays as described by Kendall and Bjorkland (2001) and simulating draws from the multinomial distribution. For the primary period m-array, we first constructed three matrices to simplify the calculation of cell probabilities. The matrix G_j describes the probability of being either unavailable (row 1, column 1) or available but not detected (row 2, column 2) in year *j* given the state in the previous year.

$$G_j = \begin{bmatrix} 1-\gamma_j^{OI} & \gamma_j^{OI}(1-p_j^*) \\ 1-\gamma_j^{II} & \gamma_j^{II}(1-p_j^*) \end{bmatrix}$$

The matrix f_j describes the probability of being unavailable (column 1) or available (column 2) if available in the previous year.

$$f_j = \begin{bmatrix} 1-\gamma_j^{II} & \gamma_j^{II}(1-p_j^*) \end{bmatrix}$$

The matrix a_j describes the probability of being available if either unavailable (row 1) or available (row 2) in the previous year.

$$a_j = \begin{bmatrix} \gamma_j^{OI} \\ \gamma_j^{II} \end{bmatrix}$$

Use of these equations and matrix algebra shortens the equations needed to construct cell probabilities (Table S2). For the first year the number of individuals encountered during each occasion was drawn from a multinomial distribution with 1000 individuals released. For each subsequent year, the number released was equal to sum of the number encountered in that occasion plus 1000 new individuals.

Within years, the probability of encountering an individual for the first time in occasion *t* is calculated based on input parameter values. The probability of an individual entering either entering in occasion *t* or before period *t*, not being detected, and remaining in the study area is:

$$\omega_{jt} = \begin{cases} t = 1 & \delta_{jt} \\ t > 1 & \omega_{jt-1}(1 - p_{jt-1})\psi_{jt-1}\tau_j + \delta_{jt} \end{cases}$$

And the probability of first encountering an individual in occasion *t* is $\omega_{jt}p_{jt}$. The number of individuals first encountered each occasion was drawn from a multinomial distribution with a probability vector of $prL2_j = [\omega_{j1}p_{j1} \quad \cdots \quad \omega_{jl}p_{jl}]$ for *l* secondary sampling occasions. Subsequent reencounters within years were based on the probabilities of persistence and detection, where the probability of being next encountered in occasion *l* if last seen in occasion *t* equal to $prL3_{t,l} = \prod_t^l \psi_{jt} * \prod_t^l (1 - p_{jt}) * p_{jt}$. For each year, we iterated over occasions and drew the number of individuals next encountered in each subsequent sampling occasion from a multinomial distribution.

We fit two different formulations of the integrated population and were mainly interested in evaluating estimation of survival probability, recruitment, and population growth rate. The

two models differed in their mark-recapture submodel. For the first, we used a Cormack-Jolly-Seber submodel that only used encounters across years, i.e. only the primary period data, to estimate apparent annual survival probability. For this model we did estimated any within-year processes of arrival and departure, and therefore did not estimate the probability of presence during each secondary occasion. Additionally, this model does not account for nonrandom temporary emigration from the stopover site. In this model the counts for each year were estimated as:

$$\sigma_j^{obs} = CV * N_j$$

$$Count_j \sim Normal(N_j, \sigma_j^{obs})$$

The second model included the full open robust design model as described in text, which includes adjusted the expected count based on the estimated probability of presence during each occasion.

Models were specified and fit using Nimble and R. We generated 3 MCMC chains of 700,000 iterations each, discarded the first 200,000 as burnin and kept every 50th sample for a total of 10,000 samples from the posterior distribution. We calculated the relative bias, root mean square error, and posterior coverage of estimates relative to the true data-generating value.

Table S1. Simulation study input values

Parameter	Notation	Input
Initial population size	N_0	Unif(1000, 10000)
Adult annual survival probability	ϕ_j^A	Unif(0.6, 0.9)
Annual recruitment of new adults	ρ_j	Unif(0.5, 2)
Transience probability	τ_j	Unif(0.2, 0.5)
Temporary emigration probability – remaining	γ_j^{II}	Unif(0.6, 0.9)
Temporary emigration probability – returning	γ_j^{OI}	Unif(0.1, 0.5)
Stopover persistence probability	$\psi_{j,t}$	Unif(0.2, 0.9)
Arrival probability	$\delta_{j,t}$	Gamma(1,1)*
Detection probability	$p_{j,t}$	Unif(0.3, 0.7)
CV of counts	CV	0.1 or 0.5

*Secondary occasion arrival probabilities (δ_{jt}) were simulated for each year as:

$$d_{j,t} \sim Gamma(1,1),\ \ \delta_{j,t} = {d_{jt}}/{\sum_1^I d_{jt}}$$

Table S2. Multinomial cell probabilities for primary period m-array, using the matrices described above. Modified from Kendall and Bjorkland (2001).

Release year (j)	Number released	Number re-encountered in year k (m_{jk})			
		2	3	4	5
1	R_1	$\phi_1\gamma_2^{II}p_2^*$	$\phi_1 f_2\phi_2 a_3 p_3^*$	$\phi_1 f_2\phi_2 G_3\phi_3 a_4 p_4^*$	$\phi_1 f_2\phi_2 G_3\phi_3 G_4\phi_4 a_5 p_5^*$
2	R_2	-	$\phi_2\gamma_3^{II}p_3^*$	$\phi_2 f_3\phi_3 a_4 p_4^*$	$\phi_2 f_3\phi_3 G_4\phi_4 a_5 p_5^*$
3	R_3	-	-	$\phi_3\gamma_4^{II}p_4^*$	$\phi_3 f_4\phi_4 a_5 p_5^*$
4	R_4	-	-	-	$\phi_4\gamma_5^{II}p_5^*$

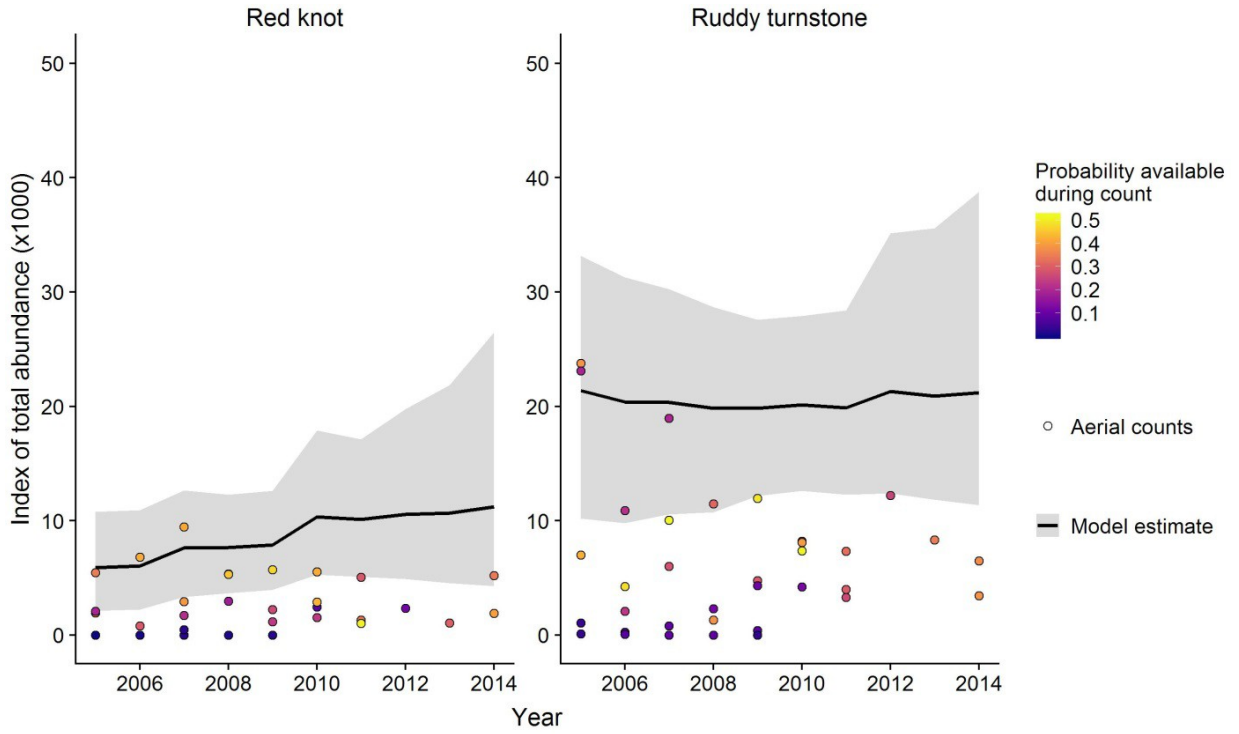

Figure S1. Aerial counts and estimated abundance indices for red knot and ruddy turnstone, 2005-2014. The solid line represents the model estimated overall abundance, which is an index of the total number of birds in the flyway population. The gray shaded region represents the 95% credible interval around this estimate. Points represent aerial survey counts conducted in each year, and the color of the points indicates the estimated probability of presence (i.e. estimated proportion of the population available) at the time of the count.

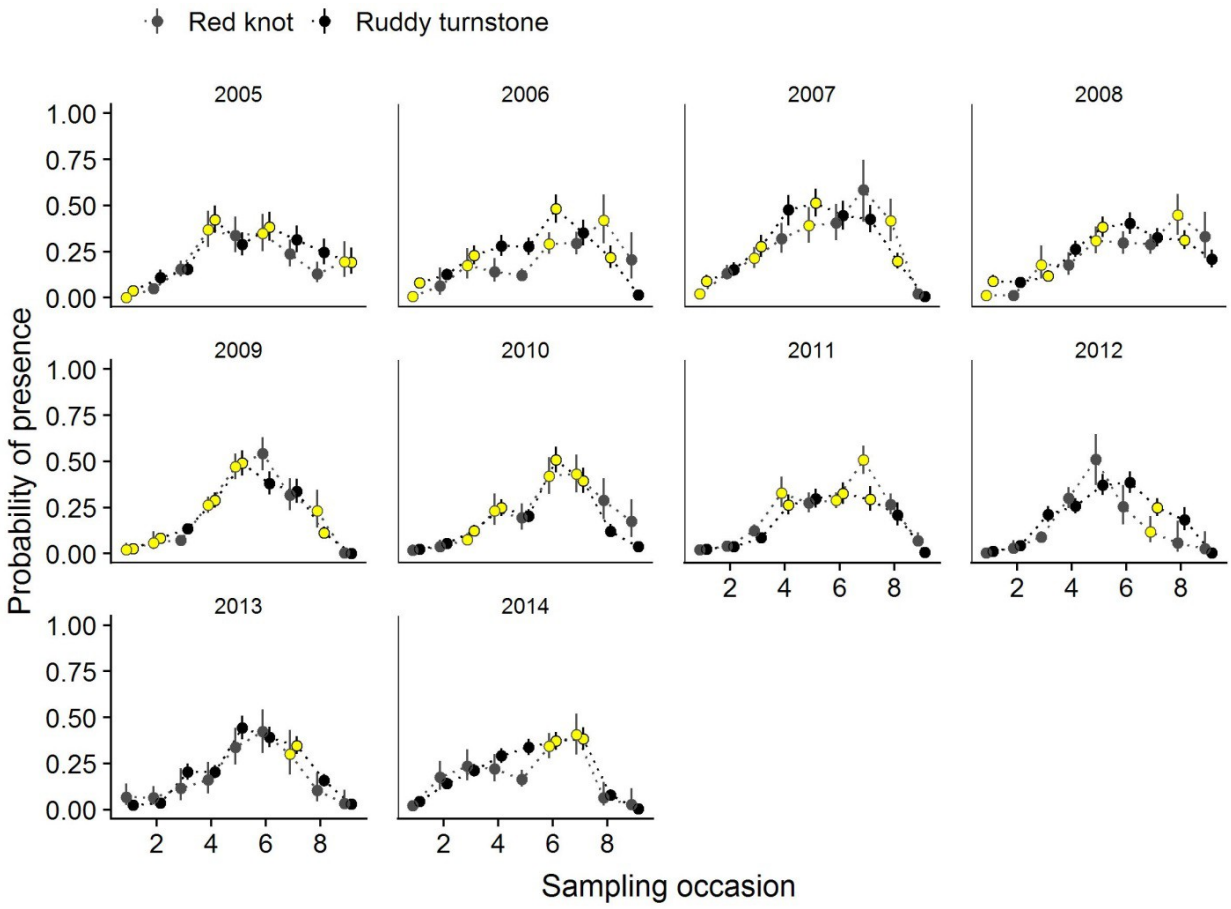

Figure S2. Estimated probability of presence during each sampling occasion of each year for red knot (gray) and ruddy turnstone (black). Points are highlighted in yellow for occasions during which an aerial count was conducted.

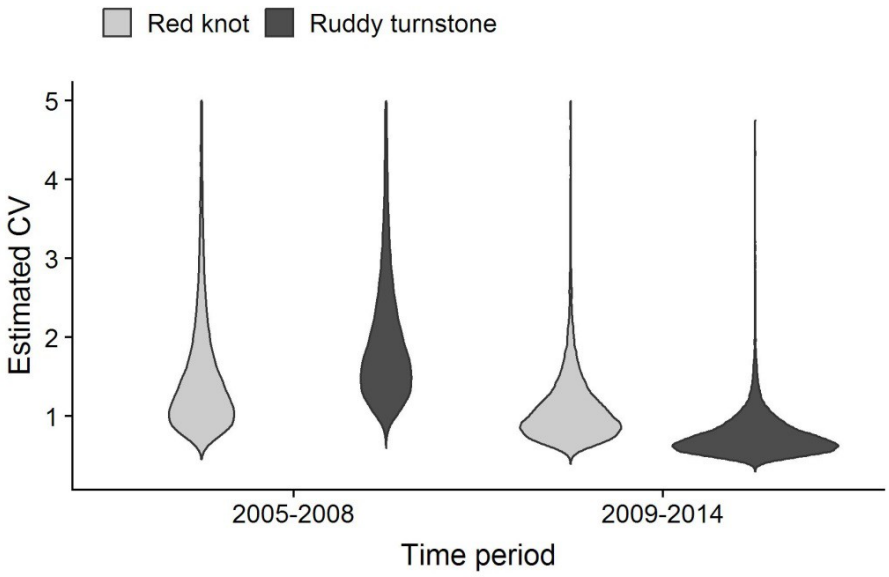

Figure S3. Estimated coefficient of variation (CV), used to estimate counting error in each year. The aerial survey flight path and main observer changed in 2009, so we estimated a separate CV for years 2005-2008 and 2009-2014. The CV was used to define the expected counting error, $\sigma^{obs} = CV * N^{avail}$

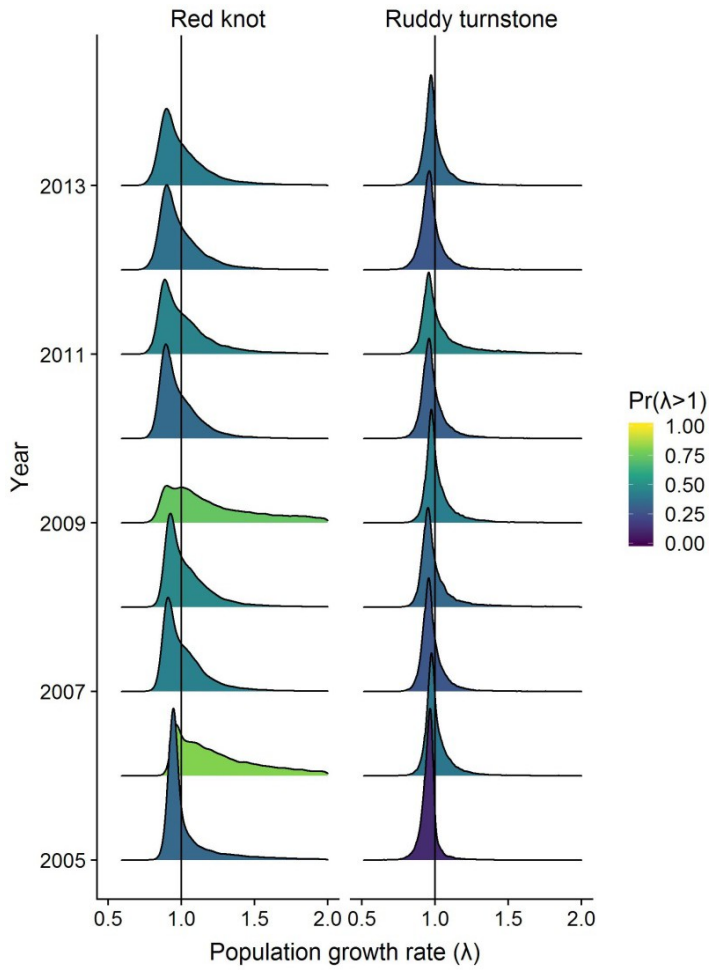

Figure S4. Estimated population growth rate for each year. The color of each plot indicates the probability that the population is increasing.

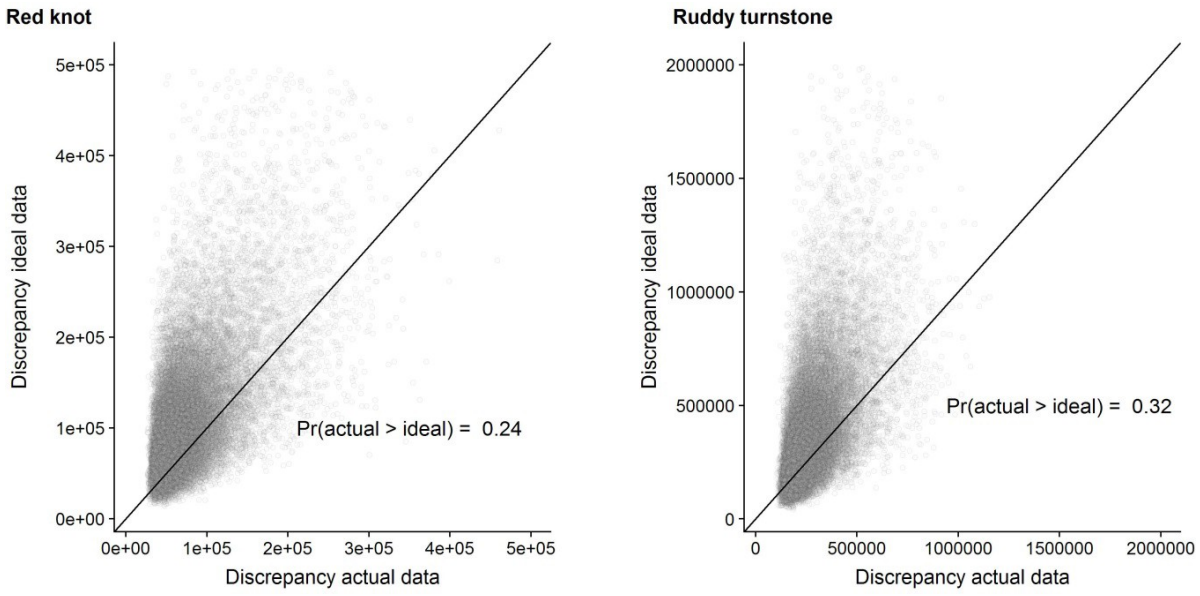

Figure S5. Discrepancy between expected values and observed data and a replicate ideal data set, used as a posterior predictive model check for state-space count model.

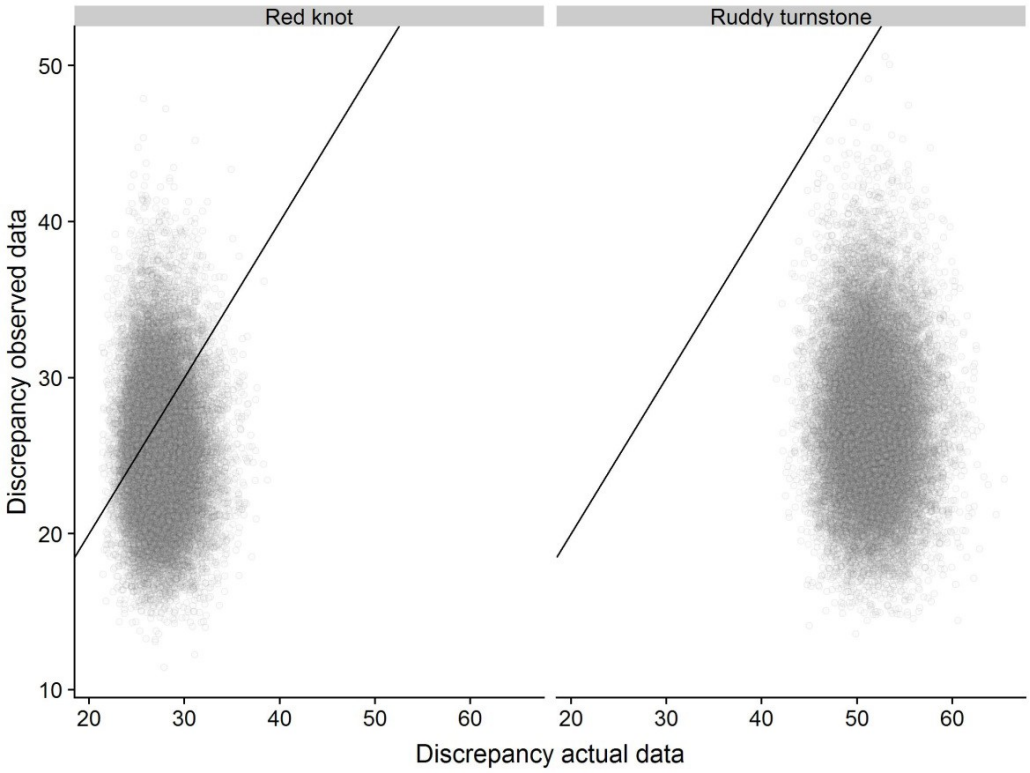

Figure S6. Discrepancy between expected values and observed data and a replicate ideal data set, used as a posterior predictive model check for the primary period likelihood of the open robust design submodel.