POPULATION MODELING TO INFORM MONITORING AND MANAGEMENT OF HERPETOFAUNA IN THE SOUTHEASTERN UNITED STATES

by

DANIELLE R. BRADKE

(Under the Direction of John C. Maerz)

ABSTRACT

Designing monitoring with adequate certainty to meet management objectives can be challenging for difficult-to-detect species. Many species of herpetofauna are particularly difficult to detect because of cryptic behavior and inaccessible habitat components. Despite challenges, monitoring is necessary to understand the status of populations and effectively apply and evaluate management. In this dissertation, I use population modeling to inform monitoring and management of diamond-backed terrapins (*Malaclemys terrapin*) and terrestrial salamanders (Genus *Plethodon*, Family Plethodontidae) in the southeastern United States. In chapter 2, I use population models and simulations to assess the ability of a common monitoring approach using seining capture-recapture methods to detect a change in diamond-backed terrapin survival, determine how detection of the change in survival varies under three thresholds of certainty that could be adopted by managers, and examine how the monitoring design can be improved to increase detection of the change in survival. In chapter 3, I investigate the capacity and practicality of using repeated counts of diamond-backed terrapins and N-mixture models to meet large-scale monitoring needs and inform management objectives in salt marsh environments. In chapter 4, I leverage an extensive five-year capture-recapture dataset of nesting diamond-backed terrapins collected as part of the Skidaway Audubon Diamondback Terrapin Rescue Project to obtain estimates of nesting frequency and values of clutch size for Georgia terrapins. Using this information, I update previous projections of population growth and persistence to help managers make more informed decisions about resource allocation. In chapter 5, I present a case study of terrestrial salamanders to demonstrate how using a hierarchical sampling and modeling approach can improve abundance estimates for species that spend substantial time unavailable for capture (e.g., below ground). Additionally, I demonstrate how integrated models can be used to leverage data from other studies to improve estimates of abundance and management effects. Results of this research can be used to improve monitoring and management for these species. This work can also be used as an example of how to address uncertainty in monitoring and how to design monitoring to meet management objectives in other challenging-to-detect species.

INDEX WORDS: abundance, capture-recapture, clutch size, count data, diamond-backed terrapin, Georgia, hierarchical model, integrated model, life-history, management, monitoring, N-mixture model, population viability analysis, reproductive frequency, robust-design, simulation, salamanders, survival, uncertainty

POPULATION MODELING TO INFORM MONITORING AND MANAGEMENT OF HERPETOFAUNA IN THE SOUTHEASTERN UNITED STATES

by

DANIELLE R. BRADKE

B.S., Grand Valley State University, 2013

M.S. Biology, Grand Valley State University, 2017

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2023

POPULATION MODELING TO INFORM MONITORING AND MANAGEMENT OF HERPETOFAUNA IN THE SOUTHEASTERN UNITED STATES

by

DANIELLE R. BRADKE

Major Professor:	John C. Maerz
Committee:	Clinton T. Moore
	Brian M. Shamblin
	Stacey L. Lance
	Brian A. Crawford

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
December 2023

DEDICATION

To Eric and Josie

ACKNOWLEDGEMENTS

I want to express my sincerest gratitude to my advisor, Dr. John Maerz, for his support throughout my PhD and for providing me with so many opportunities to pursue the research and career path that I am interested in. I also want to thank my committee members, Dr. Brian Crawford, Dr. Clint Moore, Dr. Brian Shamblin, and Dr. Stacey Lance, for their contributions to my dissertation and my development as a scientist. I especially appreciate their thoughtful comments and edits, assistance with the quantitative aspects of my project, and help collecting and processing genetic samples. I am also very grateful for the generous help provided by Heather Gaya with analyses and modeling emergencies, and Heather, Cyndi Carter, Angela Burrow, Anje Kidd-Weaver, and Angela Hsuing for being an amazingly supportive group of fellow researchers and friends that helped me stay focused and feel connected during my remote work. I am also thankful for current and past members of the Maerz lab who I have enjoyed getting to know and who have helped with data collection and other aspects of my research, including lab mates Dare Scott, Shelby Bauer, and Kirstine Grab who I am lucky to have as co-authors on chapters in this dissertation. I am thankful for all the help from Vanessa Terrell to keep me and my project running smoothly. I would also like to thank Katie Mascovich and Bo for letting me live with them while conducting fieldwork – this helped to make this project possible, and I am grateful to have also made a new friend. I feel very lucky to have had the opportunity to work with Mark Dodd and Ashley Raybould from the Georgia Department of Natural Resources. I appreciate their help with surveys, and I have learned a lot about terrapin conservation in Georgia from my discussions with Mark. I am also very grateful to the Georgia

Sea Turtle Center for help with collecting genetic samples and seining, especially Terry Norton, Michelle Kaylor, Jamie Gamby, and David Zailo. Thank you so much to the Skidaway Audubon Diamondback Terrapin Rescue Project for sharing their amazing dataset and hosting me for a visit to their project, and to Patricia Levasseur for helping me with questions about headcount surveys and for thoughtful discussions about terrapin monitoring.

I would like to acknowledge the support of funders including The Environmental Resources Network, the University of Georgia Graduate School, the Chelonian Research Foundation Justin Congdon and Nancy Dickson Research Fund, the Society for the Study of Amphibians and Reptiles, and the Diamondback Terrapin Working Group. Without this support, this research would not have been possible.

I am so incredibly grateful to my friends and family. Thank you to my mom and my sister for their encouragement during yet another degree (I swear this is the last one). I am so lucky to have their enthusiastic support in everything that I do. Thank you to Cali Wilson for being an incredible roommate and friend. I don’t know what I would do if I didn’t have you in my COVID bubble and as a friend throughout the past 4.5 years. Thank you to Josie for many long hours of company while I wrote my dissertation and for insisting on much needed outside breaks. And, finally, thank you to Eric Hileman for your unwavering support in me pursuing my goals and your belief in my abilities.

TABLE OF CONTENTS

APPENDICES

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

MONITORING POPULATIONS FOR EFFECTIVE MANAGEMENT

For ecology to aid wildlife conservation and management, the goal must be to reduce uncertainty about ecological processes and the effects of management on populations (Allen et al. 2011, Sells et al. 2018, Nichols 2019). Over the last several decades, major advances have been made in modeling studies of wildlife populations, improving the reliability of estimates and the power to test hypotheses (Schwarz and Seber 1999, Williams et al. 2002, Mazerolle et al. 2007, Iijima 2020), including the use of hierarchical sampling designs and analyses that account for imperfect detection by incorporating models of the observation process (Halstead et al. 2012, Kéry and Royle 2016, 2021, Kellner et al. 2023). These advances are key to effective wildlife monitoring, which is essential to make strong inferences and informed management decisions.

Monitoring and management go hand-in-hand. The goals of wildlife management often include growing and sustaining wildlife populations (or suppressing and eradicating populations for invasive species) while considering other social or economic interests. Monitoring is typically conducted to provide information to managers that they can use to achieve their goals. Effective wildlife monitoring requires thoughtful design linked to clear objectives (Block et al. 2001, Nichols and Williams 2006). Defining clear objectives gives managers the ability to address important gaps in the decision-making process (Nichols and Williams 2006). Common monitoring objectives include identifying if or when a management action should be taken, reducing uncertainty about a system to better understand how or why the system may or may not respond to management (as part of adaptive management), and evaluating whether management actions have been successful (Lyons et al. 2008). Each of these objectives is tied to specific management needs, fostering straightforward application of the knowledge obtained via monitoring.

Linking monitoring to stated objectives relies on selecting suitable indicators that can be measured or estimated over time. When monitoring is performed to evaluate management success, it is important to choose an indicator predicted to be sensitive to the specified management action and directly connected to the mechanism of change in the population (Gibbs et al. 1999, Lyons et al. 2008). For example, monitoring a population vital rate (e.g., survival, fecundity) can provide a more direct link to management effects than monitoring population size, which may lag the change in vital rate (Temple and Wiens 1989). Even when a strong indicator is selected for monitoring, however, the complexity of natural systems can make it difficult to obtain precise and accurate estimates and have confidence that a management action has been successful. Stochasticity and unmeasured environmental variables may create noise in the monitoring indicator, making management effects difficult to detect (Harwood and Stokes 2003, Kendall and Moore 2012). Imperfect observations of natural systems also contribute to the uncertainty inherent in monitoring (Harwood and Stokes 2003, Kendall and Moore 2012). Evaluating the power of a monitoring design to detect management effects is thus important to ascertain whether monitoring can meet its objectives—yet this is rarely done in practice (Block et al. 2001, Lindenmayer and Likens 2009).

The use of power analysis to evaluate study design is widely discussed in the scientific literature, despite it being uncommon practice in wildlife monitoring (Reynolds et al. 2011). Power analysis can be used to assess whether a given amount of sampling and the spatial and temporal distribution of sampling are sufficient to detect an effect of a specified size (Sewell et al. 2012). It can also be used to estimate the level of certainty likely to be obtained with a particular sampling design. While reducing uncertainty is often ideal, doing so typically involves collecting more or better data, which may be prohibitively costly or logistically infeasible. Thus, it can be necessary for managers to consider what level of uncertainty, or risk, they are willing to tolerate for decisions based on monitoring outcomes. In other words, managers may need to determine an acceptable tradeoff between making decisions in the face of uncertainty and delaying decisions until uncertainty is reduced.

Designing monitoring with adequate certainty is particularly challenging for rare and elusive species, especially when they occur in inaccessible environments. Data collection in these environments can be prohibitively expensive because of the need for innovative methods and technology (e.g., aerial surveys, cameras) or high labor costs. Thus, managers may find value in assessing the statistical power of a monitoring design to detect population changes, assessing their capacity to improve the power of the monitoring approach, and evaluating what level of uncertainty they are willing to tolerate in guiding decisions.

Scaling up the number of study sites is one means of increasing the power to detect changes in a population. Additionally, sampling multiple sites across a broader spatial extent may more accurately represent what is occurring at the larger landscape scale (i.e., over many populations) in response to management (Gurevitch et al. 2016). Otherwise, the application of monitoring results may reflect an understanding of one or a few uncharacteristic sites. However, spatially extensive wildlife monitoring is typically costly, time consuming, and logistically challenging (Jones 2011). This can be especially true for difficult to capture species in relatively inaccessible environments.

Using counts of unmarked animals, rather than more costly capture-recapture data, is a common means of scaling up the number of sites that can be monitored with available resources. Binomial N-mixture models (Royle 2004) in particular, have become a widespread tool to estimate abundance because they use repeated counts to correct for imperfect detection, similar to capture-recapture abundance models. Collecting count data is appealing because it may require less effort and training and be less time or resource intensive than methods requiring animals to be in-hand for individual identification. However, N-mixture models are less robust than capture-based models to assumption violations and, thus, are more prone to bias (Duarte et al. 2018, Knape et al. 2018, Link et al. 2018). For example, the assumptions of closure and constant detection are difficult to meet in many studies and violations of these assumptions can severely bias estimates.

Because of the bias documented in abundance estimates from N-mixture models, these models may not be suitable to estimate abundance in many scenarios. However, N-mixture models (and other methods that can account for variable detection: e.g., Poisson regression) may still provide a useful index of relative abundance (Barker et al. 2018). Thus, whether count data and N-mixture models are a functional solution for large-scale monitoring is likely context dependent, making it necessary to investigate under what, if any, circumstances these methods can be used to inform a specific monitoring objective for a given species.

POPULATION VIABILITY ANALYSES

In addition to estimating and monitoring survival and abundance parameters, monitoring life-history traits can be important for making effective management decisions. Life-history traits are directly linked to vital rates which drive population growth and persistence (Williams et al. 2002). For example, reproductive characteristics, including clutch size and reproductive frequency, determine recruitment rates of new individuals into a population (Cole 1954, Mills and Lindberg 2002). Population viability analysis (PVA) uses life-history traits and vital rates to predict population growth rates and assess the likelihood of population persistence over time (Beissinger and McCullough 2002). Projections from PVAs are often relied upon to allocate limited resources among competing management actions (Boyce 1992, Lindenmayer et al. 1993, Beissinger and Westphal 1998, Manlik et al. 2018). Useful projections of population growth and persistence depend on having accurate estimates of life history traits and vital rates. Inaccurate estimates of key life history traits and vital rates that result in overestimation of population viability can lead to inaction that results in adverse outcomes for priority species. Conversely, underestimation of population viability can result in the allocation of additional resources to unnecessary actions instead of allocating those resources toward other priorities.

Life history and vital rate parameters used in a PVA often come from multiple sources, including direct measurement and estimation of the focal population, transferred values from published studies of other populations or closely-related species, or expert opinion (Heppell et al. 2000, Etterson and Bennett 2006, McGowan et al. 2011). It is rare that PVAs can be parameterized entirely using data from the focal population due to the time and effort required to measure or estimate these parameters (Beissinger 2002, Etterson and Bennett 2006). Some parameters may be reasonably transferred from other populations or closely related species (Heppell 1998, Heppell et al. 2000, Schtickzelle et al. 2005); however, viability models may be sensitive to and biased by transferred parameters that vary temporally or spatially (Lindenmayer et al. 1993, Gurevitch et al. 2016). Therefore, it can be important to update initial viability models with improved estimates of parameters when available.

INTEGRATED MODELS

While there have been many advances in the estimate of abundance and vital rates that have led to improved monitoring and management of wildlife populations in recent history, one of the most recent and exciting advances is integrated modeling. Integrated models offer a framework for bringing together multiple data types to generate more credible demographic estimates and to allow for studying more populations over wider spatial scales and broader environmental gradients (Schaub et al. 2007, Zipkin and Saunders 2018, Zylstra et al. 2021). For example, data from less labor-intensive methods (e.g., counts) collected at many sites over a broad spatial scale can be integrated with data from more labor-intensive methods (e.g., capture-recapture) collected at a small number of sites to improve reliability of estimates and management applications. Integrating datasets can also produce the ability to disentangle and estimate parameters that would otherwise be confounded if relying on a single sampling methodology (Cole and McCrea 2016, Saunders et al. 2019).

In an integrated model, datasets may each have different observation processes and may inform unique variables or processes of interest, but the sub models associated with each dataset must share some parameters (e.g., availability, abundance, or survival; Schaub and Abadi 2011). These shared parameters link the datasets together, giving the statistical power of one dataset the ability to increase the statistical power of the other. For instance, a larger dataset that is collected over a broader spatial extent can help inform a smaller-scale dataset with more limited replication that is part of a different study (Schmidt et al. 2019, 2021).

Another use of integrated models is in improving or resurrecting old datasets collected under a poor sampling design. There are many existing datasets, some long term, that do not incorporate designs suited to analyses that explicitly model the observation process. This can be a problem because, if the important processes that affect observation of the study species are not adequately addressed, estimates of the processes of interest may be severely biased and inferences may be unreliable. Using integrated models, a dataset collected under an improved sampling design – either from a different study or from changing methods in an ongoing study – can be used to improve abundance estimates for a historical dataset with high bias and uncertainty.

STUDY SYSTEMS AND KNOWLEDGE GAPS

Diamond-backed terrapins (*Malaclemys terrapin*)

The diamond-backed terrapin (*Malaclemys terrapin*) is an excellent example of a difficult-to-monitor species. The estuarine habitat of terrapins is typically difficult to access, and daily and seasonal variation in terrapin habitat use presents additional challenges for detection. In much of their range, terrapin habitat consists of dendritic networks of tidal creeks within expansive salt marshes. On a daily basis, tide cycles affect whether terrapins are likely to be found in creeks or the high marsh, which is virtually inaccessible for monitoring. Throughout the active season, nesting migrations affect the location of mature females. Additionally, nesting activity can be difficult to monitor because of timing and locations of nesting events and long nesting seasons (extending nearly four months in the southern portion of the species' range), making it difficult to devote sufficient time and labor to monitor the entire season. Despite these challenges, monitoring is necessary to understand the status of populations and effectively apply and evaluate management.

Terrapin populations are threatened by multiple factors including vehicular road mortality during nesting, high nest depredation by subsidized predators, and coastal development (Isdell et al. 2015, Chambers and Maerz 2019, Maerz et al. 2019). Additionally, one of the greatest present-day threats to terrapin populations is commercial and recreational crab fisheries because of the tendency of terrapins to enter and drown in crab traps (Roosenburg 2004, Grosse et al. 2009, 2011, Isdell et al. 2015, Lovich et al. 2018, Chambers and Maerz 2019). As a result, they are listed as a high priority species or Species of Greatest Conservation Need in all states where they occur, along the Atlantic and Gulf coasts in the United States. Therefore, there is a need to be able to accurately and precisely estimate abundance, vital rates, and life-history traits to inform monitoring and management of this species.

In the state of Georgia, terrapin monitoring within their estuarine habitat has been mostly limited to annual capture-recapture surveys at 2-4 creeks and, thus far, has not been used to directly inform long-term, state-wide population management objectives of state managers (e.g., Bradke et al. in press., Crawford et al. 2018, GA DNR Coastal Resources Division unpublished data). Grosse et al. (2011) conducted the only intensive statewide study of terrapins to assess the effects of commercial crabbing pressure and road proximity on terrapin abundance. Capture-recapture within tidal creeks is time and labor intensive and, therefore, difficult to use over many sites and long periods of time. For example, sampling to estimate abundance in the 29 creeks studied by Grosse et al. (2011) took two years and cost ~ $250,000 to complete. Unmarked methods that do not require capture and individual identification have also been considered for monitoring terrapins in Georgia (e.g., Levasseur et al. 2019), but have not been tested in salt marsh tidal creek environments. Additionally, data and estimates on nesting frequency and clutch sizes are lacking in Georgia, where there are no published studies with this information, potentially impeding accurate population projections.

Terrestrial salamanders (Genus *Plethodon*, Family Plethodontidae)

Terrestrial salamanders (Genus *Plethodon*, Family Plethodontidae), like many amphibians and reptiles, are also challenging to monitor. Because they have permeable skin and a propensity for evaporative water loss, they are highly sensitive to their environment and spend substantial time below ground to mediate effects of stressful environmental conditions (Taub 1961, Feder 1983, Peterman and Semlitsch 2014, Muñoz et al. 2016). Consequently, observations can be low and are driven by factors affecting surface activity (i.e., availability), in addition to the ability of observers to detect individuals available at the surface (i.e., detection; Bailey et al. 2004*a*, *b*, Kéry and Schmidt 2008, O'Donnell and Semlitsch 2015).

Many past and ongoing studies of terrestrial salamanders (and other herpetofauna) do not explicitly address the processes affecting animal observations even though the use of hierarchical sampling designs are a compelling option (Halstead et al. 2012, Kéry and Royle 2016, 2021, Kellner et al. 2023). Methods and analyses that do not account for imperfect detection by incorporating models of the observation process, including animal availability, can introduce strong bias and uncertainty in estimates leading to poor inferences and ineffective conservation. However, questions and concerns exist regarding how to navigate the skillset needed to implement complex models and how to achieve comparability of past data to future data if sampling methods are changed during long-term studies.

RESEARCH OBJECTIVES

In chapter 2, I use population models and simulations to simultaneously evaluate multiple aspects of monitoring for diamond-backed terrapins in Georgia, USA. My objectives were to assess the ability of a common monitoring approach using seining capture-recapture methods to detect a change in diamond-backed terrapin survival, determine how detection of the change in survival varies under 3 thresholds of certainty that could be adopted by managers, and examine if and how the monitoring design can be improved to increase detection of the change in survival.

In chapter 3, my objectives were to investigate the capacity and practicality of using repeated counts of diamond-backed terrapins and N-mixture models to meet large-scale monitoring needs and inform management objectives in salt marsh environments. The usefulness to GA DNR of estimates from head count surveys depends on obtaining sufficient precision to place tidal creeks into reliable categories of abundance. Therefore, I evaluate the robustness of abundance estimates from head count surveys to understand how the timing of sampling within the active season, location of sampling along a given creek, and amount of survey effort (e.g., number of repeated surveys per visit) affect estimates.

In chapter 4, I leverage an extensive five-year capture-recapture dataset of nesting diamond-backed terrapins collected as part of the Skidaway Audubon Diamondback Terrapin Rescue Project to estimate nesting frequency and obtain clutch sizes for Georgia terrapins. The objectives of this study were to use this information to update previous projections of population growth and persistence to help managers make more informed decisions about resource allocation. Specifically, I aimed to determine whether and how much new nesting parameter estimates changed population projections under the same management scenarios considered in a previously published population viability analysis of diamond-back terrapins by Crawford et al. (2018).

In chapter 5, my objective was to present an encouraging perspective on the use of a hierarchical approach in designing or modifying sampling methods and analyzing current and historical datasets for amphibians and reptiles. First, I provide a brief overview of hierarchical sampling designs and models. Second, I describe how integrated models can be used to leverage hierarchical sampling designs to improve estimates from older sampling designs. Third, I present a case study where I attempt to estimate forest management effects on terrestrial salamander abundance using an older sampling design, a modified hierarchical design, and integrated abundance models that use older and new datasets. I also demonstrate how integrated models can be used to leverage data from other studies to further improve estimates of abundance and forest management effects. Finally, I offer advice on adapting sampling designs and suggest ways to overcome the understandable barriers to herpetologists who could benefit from using hierarchical and integrated approaches.

Collectively, my dissertation aims to use population modeling approaches to inform information gaps important for effective monitoring and management of herpetofauna in the southeastern United States. In addition to addressing needs for the focal species, this work can also be used as an example of how to address uncertainty in monitoring and how to design monitoring to meet management objectives in other challenging-to-detect species.

CHAPTER 2

EVALUATING UNCERTAINTY TO IMPROVE A COMMON MONITORING METHOD AND GUIDE MANAGEMENT DECISIONS FOR DIAMOND-BACKED TERRAPINS [1]

---

[1] Bradke, D.R., B.A Crawford, M. Kaylor, and J.C. Maerz. 2023. *Journal of Wildlife Management*.e22513. Reprinted here with permission of the publisher.

ABSTRACT

Designing monitoring with adequate certainty to evaluate management actions can be challenging, especially for elusive species in relatively inaccessible habitats. The diamond-backed terrapin (*Malaclemys terrapin*) is considered a high priority species for management in all states within its range. Among key threats to terrapin populations is bycatch mortality in crab pots, prompting states throughout the species' range to consider crabbing regulations and creating a need for a monitoring approach able to evaluate the success of any new crabbing regulations to benefit terrapins. Because terrapins occupy extensive and often difficult to access estuarine habitat, it is hard to collect sufficient data to estimate population responses accurately and precisely for management. To assist state managers with decisions regarding monitoring bycatch reduction regulations, we modeled 12 years of terrapin capture-recapture data. We used estimates from those models to simulate capture-recapture data pre- and post-implementation of bycatch reduction that resulted in a 0.20 absolute increase in mean apparent survival probability. Results indicated weak reliability of a monitoring approach using only manual seining of tidal creeks to detect a real management effect, with a positive management effect detected at the 95% certainty level only 34% of the time. When we considered 85% and 75% certainty thresholds, we detected a positive effect on survival among 61% and 75% of simulations, respectively. Low within-year recapture probability and the effect of tide amplitude on terrapin availability indicate there is low feasibility of improving monitoring precision at a single site, requiring monitoring of more sites to improve confidence in the detection of the management effect. The number of sites monitored depends on the acceptable level of certainty. We recommend that researchers and management entities that use seining assess the level of certainty they require to evaluate management actions and increase the number of sites sampled to meet that level of certainty.

Additionally, the use of multiple monitoring methods and integrated models should be explored to reduce uncertainty and to allow for easier monitoring of more populations over broader spatial scales.

INTRODUCTION

Effective wildlife monitoring requires thoughtful design linked to clear objectives (Block et al. 2001, Nichols and Williams 2006). Defining clear objectives gives managers the ability to address important gaps in the decision-making process (Nichols and Williams 2006). Common monitoring objectives include identifying if or when a management action should be taken, reducing uncertainty about a system to better understand how or why the system may or may not respond to management (as part of adaptive management), and evaluating whether management actions have been successful (Lyons et al. 2008). Each of these objectives is tied to specific management needs, fostering straightforward application of the knowledge obtained via monitoring.

Linking monitoring to stated objectives relies on selecting suitable indicators that can be measured or estimated over time. When monitoring is performed to evaluate management success, it is important to choose an indicator predicted to be sensitive to the specified management action and directly connected to the mechanism of change in the population (Gibbs et al. 1999, Lyons et al. 2008). For example, monitoring a population vital rate (e.g., survival, fecundity) can provide a more direct link to management effects than monitoring population size, which may lag the change in vital rate (Temple and Wiens 1989). Even when a strong indicator is selected for monitoring, however, the complexity of natural systems can make it difficult to obtain precise and accurate estimates and have confidence that a management action has been successful. Stochasticity and unmeasured environmental variables may create noise in the monitoring indicator, making management effects difficult to detect (Harwood and Stokes 2003, Kendall and Moore 2012). Imperfect observations of natural systems also contribute to the uncertainty inherent in monitoring (Harwood and Stokes 2003, Kendall and Moore 2012). Evaluating the power of a monitoring design to detect management effects is thus important to ascertain whether monitoring can meet its objectives—yet this is rarely done in practice (Block et al. 2001, Lindenmayer and Likens 2009).

The use of power analysis to evaluate study design is widely discussed in the scientific literature, despite it being uncommon practice in wildlife monitoring (Reynolds et al. 2011). Power analysis can be used to assess whether a given amount of sampling and the spatial and temporal distribution of sampling are sufficient to detect an effect of a specified size (Sewell et al. 2012). It can also be used to estimate the level of certainty likely to be obtained with a particular sampling design. While reducing uncertainty is often ideal, doing so typically involves collecting more or better data, which may be prohibitive. Thus, it can be necessary for managers to consider what level of uncertainty, or risk, they are willing to tolerate. In other words, managers may need to determine an acceptable tradeoff between making decisions in the face of uncertainty and delaying decisions until uncertainty is reduced.

Designing monitoring with adequate certainty is particularly challenging for rare and elusive species, especially when they occur in inaccessible environments. Data collection in these environments can be prohibitively expensive because of the need for innovative methods and technology (e.g., aerial surveys, cameras) or high labor costs. Thus, managers may find value in assessing the statistical power of a monitoring design to detect population changes, assessing their capacity to improve the power of the monitoring approach, and evaluating what level of uncertainty they are willing to tolerate.

The diamond-backed terrapin (*Malaclemys terrapin*) is an estuarine turtle distributed along the Atlantic and Gulf coasts in the United States. It is considered a high priority species or species of greatest conservation need in wildlife action plans in all states within its range, and it is listed as vulnerable on the International Union for Conservation of Nature Red List (Roosenburg et al. 2019). The estuarine habitat of terrapins makes monitoring particularly challenging. In much of their range, terrapin habitat consists of dendritic networks of tidal creeks within expansive salt marshes, which are difficult to access and sample. Additionally, terrapins spend time within creeks, where surveys are typically conducted and have reported low detection rates (Grosse et al. 2011, Crawford et al. 2018), and within the high marsh, where they are virtually undetectable (Harden et al. 2007). These low detection probabilities make it difficult to obtain precise or accurate estimates of population parameters.

In many localities, the greatest current threat to diamond-backed terrapin populations is bycatch mortality in commercial and recreational crab traps (Roosenburg 2004, Grosse et al. 2009, 2011, Isdell et al. 2015, Lovich et al. 2018, Chambers and Maerz 2019). Bycatch reduction devices (BRDs) attached to the entrances of crab traps can effectively exclude terrapins from traps (Chambers and Maerz 2019), prompting several states to require these devices and others to consider similar regulations. Population-level responses to these regulations have not been documented; however, an increase in terrapin survival and abundance in areas where BRDs are required is expected (Roosenburg 2004, Chambers and Maerz 2019). Managers in states where BRD regulations are being considered, such as Georgia and South Carolina, want to assess the effectiveness of BRDs. Seining, where nets are hand-pulled through tidal creeks, is a common monitoring approach used in these states to collect capture-recapture data on diamond-backed terrapins and detect effects of bycatch and vehicle strikes on terrapin survival or abundance

(Tucker et al. 2001, Cecala et al. 2009, King and Ludlam 2014, Witczak et al. 2014, Crawford et al. 2018). Studies using seining as the sole method have thus far relied on long-term data (>20 years; Dorcas et al. 2007) at few sites (1 to 4) or extensive data collection across many sites (>20 sites) over short periods (Grosse et al. 2011). It is unknown the extent to which seining methods alone have sufficient power to detect positive effects of management actions over temporal or spatial scales relevant to managers.

In this study, we demonstrate the value of using population models and simulations to simultaneously evaluate multiple aspects of monitoring for terrapins in Georgia, USA. Our objectives were to assess the ability of the current seining monitoring approach to detect a change in survival, determine how detection of the change in survival varies under 3 different thresholds of certainty that could be adopted by managers, and examine if and how the monitoring design can be improved to increase detection of the change in survival. We describe annual variation in diamond-backed terrapin survival probabilities, the effect of tidal amplitude on the probability that a terrapin is available for capture each year, and daily variation in and a behavioral effect on capture probabilities.

STUDY AREA

We collected data within 2 salt marsh tidal creeks adjacent to and on opposite sides of the Downing Musgrove Causeway, near Jekyll Island, in Glynn County, Georgia, during 2010 to 2022 (Figure 2.1). We sampled an approximate 400-m section of the creek to the northeast of the causeway and an approximate 800-m section of the creek to the southwest of the causeway, with the lengths of the sections sampled constrained by creek depth (i.e., our ability to pull the seines through the creek). Typical of coastal Georgia, these creeks were characterized by large fluctuations in tidal amplitude, with 2 approximately 2- to 3-m tide cycles/day. The creeks had high turbidity and were non-vegetated, with substrate consisting of unconsolidated muddy bottoms and occasional oyster beds. The marsh vegetation bordering the creeks was primarily smooth cordgrass (*Spartina alterniflora*). Dominant fauna observed in the high marsh included marsh periwinkles (*Littoraria irrorata*), fiddler crabs (*Uca* spp.), and wharf crabs (*Armases cinereum*), and dominant fauna observed within the tidal creeks included mud snails (*Ilyanassa obsoleta*), grass shrimp (*Palaemonetes* spp.), Atlantic blue crabs (*Callinectes sapidus*), Florida stone crabs (*Menippe mercenaria*), and a wide diversity of fish species, such as mullet (*Mugil* spp.), Atlantic menhaden (*Brevoortia tyrannus*), and Atlantic croakers (*Micropogonias undulatus*). Larger species such as bonnet head sharks (*Sphyrna tiburo*), spotted gar (*Lepisosteus oculatus*), and Atlantic stingrays (*Hypanus sabinus*) were also observed regularly. The creeks sampled were undeveloped; however, the causeway is the only road leading to Jekyll Island and, as such, is highly trafficked with many terrapin road mortalities reported each nesting season (Crawford et al. 2014*b*). The climate is subtropical, with mild winters and hot, humid summers. Thirty-year climate normal (i.e., 1991–2020) mean daily spring temperatures (Mar–May) averaged 20°C, mean daily summer temperatures (Jun–Aug) averaged 28°C, mean daily fall temperatures (Sep–Nov) averaged 22°C, mean daily winter temperatures (Dec–Feb) averaged 13°C, and annual precipitation averaged 115 cm (Palecki et al. 2021).

METHODS

Data collection

We conducted capture-recapture surveys annually from 2010 to 2022 (except for 2020) under a robust-design sampling scheme (Pollock 1982). Specifically, in 2010–2019 we surveyed each creek over 3 consecutive days and in 2021–2022 we surveyed each creek over 2 consecutive days, with survey dates occurring between 22 April and 3 May each year. We considered each year to be a primary period (*n* = 12) and each day sampled to be a secondary period (*n* = 34; Pollock 1982).

We captured diamond-backed terrapins with teams of approximately 8 people manually pulling 2 seines in tandem through the length of each creek at low tide (Figure 2.2; Grosse et al. 2011, Crawford et al. 2018, Roosenburg and Burke 2019). Each seine was a 10-m wide mesh net with bag at the center and had floats along the top, lead weights along the bottom, and a wood post on either side. Conducting these surveys at low tide maximized coverage of the creek by the seine to minimize the ability of terrapins to evade capture by swimming around the net or rapidly escaping into vegetation (Crawford et al. 2018). Conducting surveys at low tide also helped researchers pull the net forward with sufficient speed to keep captured terrapins from escaping. There can be substantial variation in tide amplitude among primary periods and, occasionally, secondary periods due to natural tidal cycles, wind, storm surges, and other factors. This variation in amplitude at low tide is thus likely important in explaining variation in detection of terrapins.

We used capture-recapture data from only males. Because of their smaller maximum size, male terrapins disproportionately enter and drown in crab pots; therefore, the effects of crabbing on survival are likely more apparent in males (Grosse et al. 2011). Additionally, males were more commonly captured during seining efforts, resulting in a larger sample size. We determined sex using body size, head size, and the cloacal position on the tail (Gibbons and Lovich 1990, Tucker et al. 1995, Underwood et al. 2013). We uniquely marked terrapins by filing notches or drilling holes into their marginal scutes (Cagle 1939).

Data analysis

We used Bayesian, robust-design Cormack-Jolly-Seber (CJS) models (Cormack 1964, Jolly 1965, Seber 1965) to estimate capture probability (p) for each secondary period (day), the probability that a terrapin was available for capture (γ) during each primary period (yr), and apparent survival probability (φ) between primary periods. We accounted for the 2-year interval between the 2019 and 2021 primary periods using a dummy covariate to fix p to 0 in 2020. We developed a candidate set of 8 models (Equations S1–S14, available in Appendix A) and used full-likelihood widely applicable information criterion (WAIC) to select the best model (Table A.1; Watanabe 2010, 2013). All models included a random year effect on apparent survival to account for anticipated annual variation in this parameter that we could not explain with available covariates. Models also included either a fixed effect of tidal amplitude or a random year effect on availability. For capture probability, models were either time-invariant, or included a random day effect, random year effect, or fixed tidal amplitude effect. Additionally, half of the models included a behavioral effect, where capture probability differed between the first capture and subsequent captures (recaptures) within a given year (i.e., a trap happy or trap shy effect). We used daily low tidal amplitude values for a tide effect on capture probability, and we used the mean amplitude of all low tides occurring during surveys for a given year for a tide effect on availability. We never included the tide effect on both capture probability and availability in the same model. We used uninformative prior distributions on all parameters (Table A.2).

We fit models using Markov chain Monte Carlo (MCMC) sampling in JAGS (Plummer 2003) called via the jagsUI package (Kellner 2021) in program R (4.2.1; R Core Team 2022). We used 3 MCMC chains with 120,000 iterations each and burn-in of 30,000. We assessed chain convergence visually using trace plots and with Gelman and Rubin (1992) diagnostic values (i.e., $\hat{R} < 1.1$). We tested the goodness of fit of the top-ranked model using the Freeman–Tukey statistic to calculate discrepancy values between observed versus expected recaptures and between recaptures predicted by the model versus expected recaptures (Freeman and Tukey 1950, Brooks et al. 2000, Rose et al. 2022). We compared discrepancy measures using a Bayesian *P*-value and visually using a scatter plot (Figure A.1).

Based on our knowledge of the system and on previously observed retrospective negative terminal bias in our dataset, φ for the last primary period appeared to be severely biased low. Retrospective negative terminal bias in φ is defined as negative bias in the final survival estimates of the time series identified by observing an increase in estimates when additional years of data are added to the series. To avoid propagating this bias forward into our simulations, we re-ran the model with the last primary period omitted from the random year effect on φ. Instead, a fixed effect on φ for the final primary period was included separately so that it did not affect the posterior distribution of the random year effect on φ (i.e., biasing mean φ low and the SD of φ high). This modification, which had negligible effects on all estimates other than the random effect, allowed us to propagate forward only the annual variation in φ from years unaffected by apparent terminal bias (Table A.2).

Simulations

We used simulation to generate datasets with a positive change in survival, so that we could assess the likelihood of detecting an increase in survival associated with management. Results of WAIC model selection clearly distinguished a single model from the others in the robust-design CJS model candidate set (Table A.1). Using this top-ranked model, we simulated 100 new capture-recapture datasets, assuming management (i.e., BRD regulations) resulted in a

0.20 absolute increase in mean apparent survival probability. To determine a plausible value for mean apparent survival post-implementation of crabbing regulations, we took the mean of the apparent survival estimates from 4 studies located at sites where crabbing is believed to be low or to not occur (Hart 1999, King and Ludlam 2014, Crawford et al. 2018; Table 2.1), for an average value of 0.77. Based on this estimate of survival without crabbing mortality, we used a value for the mean apparent survival pre-regulations of 0.57 to produce the 0.20 increase in mean survival. For each of the 100 simulated datasets, we simulated 15 years of data with 3 secondary periods per year, for 45 secondary periods. Additionally, we simulated that the regulations went into effect immediately after sampling on the eighth year, yielding 7 pre-regulation sampling intervals between primary occasions and 7 post-regulation sampling intervals.

For each of the 100 simulations, we drew a random MCMC iteration from the seining data model results and extracted the posterior parameter values (i.e., mean capture probability, behavioral effect, random day effect on capture probability; mean availability probability, tidal amplitude effect on availability probability; and standard deviation for the random year effect on apparent survival probability). We used these values, the mean pre- and post- regulation apparent survival values (0.57 and 0.77, respectively), and a starting population of 60 individuals. We replaced individuals in the population with an annual rate equal to 60 individuals × (1 – pre-regulation mean survival probability; i.e., we simulated recruitment of 26 new individuals per year). To simulate annual tidal amplitude covariate values, we took the mean low tide amplitude occurring during surveys for each of the 12 years seining was conducted and drew new values from a normal distribution with the same mean and standard deviation.

To evaluate if the simulated increases in survival would be detected using seining monitoring methods, we used the top-ranked robust-design CJS model fit to each of the 100 simulated datasets. This model assumed a study design with 8 years of pre-BRD monitoring followed by 7 years of post-BRD implementation monitoring, and therefore included a fixed effect of BRD, a binary covariate on apparent survival probability to estimate the change in mean survival associated with BRD regulations. For this binary covariate, the primary period sampling intervals 1–7 were coded as 0 and sampling intervals 8–14 were coded as 1. To increase processing speed relative to running the initial candidate set of models, we fit models to the simulated datasets using MCMC sampling via the nimble package (de Valpine et al. 2017, 2022) in program R (4.2.1; R Core Team 2022) with the same arguments as before (3 chains, 120,000 iterations, 30,000 burn-in).

We evaluated detectability of the survival increase under 3 different levels of certainty that could be adopted by managers for their monitoring program: 95%, 85%, and 75%. For the change in mean apparent survival to be considered accurate, the posterior distribution mean would be 0.20. The mean estimate, however, does not provide any information about how certain or confident managers could be in the change in survival probability. For this reason, we used the posterior distribution of the estimate to assess a manager's ability to detect the simulated change in survival at each level of certainty. For the increase in survival to be considered detected, the 25th, 15th, and 5th percentiles of the posterior distribution would be >0, corresponding to a 75%, 85%, or 95% probability that survival increased (Figure 2.3).

We also evaluated how the detection of the simulated management effect improved as we increased the number of sites monitored from 1 to 2, 4, 6, or 8 sites. For this analysis, we used the same 100 simulated datasets previously described, but here we combined them to investigate the effect of having 2, 4, 6, or 8 of these datasets to make inference, rather than only 1. Thus, for each number of sites (2, 4, 6, or 8), we randomly drew a corresponding number of simulated datasets. We combined the randomly drawn datasets into a single capture history and created a grouping variable to indicate the original simulated dataset (i.e., site). We completed this process, with replacement, 100 times to produce 100 combined datasets for each number of sites considered. For each of the newly combined datasets representing multiple sites, we again evaluated if the simulated increase in survival would be detected by fitting the top-ranked robust-design CJS model. We used the grouping variable to allow estimated parameters to vary between (or among) sites as a fixed effect, except for the change in mean apparent survival associated with management, which was estimated as a single parameter. We fit these models with the nimble package using 3 MCMC chains, 120,000 iterations, and 30,000 burn-in. For each number of sites considered, we calculated the percent of simulations where the effect of management was detected (evaluated using the robust-design CJS model) under each of the 3 levels of certainty. We then fit an exponential decay function to this percent of simulations with survival increase detected (y) for each certainty level (Equation 1; Pinheiro and Bates 2000, Pinheiro et al. 2022), where y is a function of starting value (y_0), asymptote (y_f; fixed to 100), rate constant (α), and the number of sites sampled (x):

$$y = y_f + (y_0 - y_f)e^{-\exp(\log \alpha)x} \tag{1}$$

Fitting the data to this function allowed us to estimate the percentage of the time that this management effect is expected to be detected when monitoring >8 sites. This analysis also allowed us to estimate the error around the percentage of the time that the management effect is expected to be detected as the number of sites sampled increases. For convenience and because the increased processing speed of nimble was not needed for this analysis, we fit this function using MCMC sampling in JAGS with 3 chains, 120,000 iterations, and 20,000 burn-in.

RESULTS

Seining data

From 2010–2022 we had 496 captures of 322 male terrapins. The top-ranked robust-design CJS model included a random year effect on apparent survival, a fixed effect of tide amplitude on availability, and random day and behavioral effects on capture probability (Table A.1; Figure 2.4). Overall, the probability of capture was relatively high ($\bar{x}$ = 0.64, 95% Bayesian credible interval [BCI] = 0.55–0.73), but if a terrapin was already captured once in a given year, its recapture probability was very low ($\bar{x}$ = 0.02, 95% BCI = 0.01–0.04; Figure 2.4C). Tidal amplitude had a negative effect on availability (β_{Tide} = −0.42, 95% BCI = −0.73 – −0.13), with a 0.20-m increase in tidal amplitude, relative to the mean amplitude during surveys, resulting in a 0.05 decrease in the probability that a terrapin was available for capture (Figure 2.4B). Posterior distribution means for annual apparent survival ranged from 0.57–0.92 (Figure 2.4A), with an overall mean annual survival probability across all years of 0.80 (95% BCI = 0.61–0.95; Table A.2). Trace plots and $\hat{R}$ values indicated successful chain convergence. A Bayesian P-value of 0.49 and the scatter plot of discrepancy values (Figure A.1) indicated adequate fit of the top-ranked model to the data.

Simulations

Posterior distribution means for the estimated change in survival associated with the BRD effect were positive in 90 out of 100 simulations. Thus 90% of the time a positive effect of management was detected if we relied on the mean estimate and ignored any uncertainty in this estimate. At the 95% certainty level, however, a positive effect of management on survival was detected in only 34 out of 100 simulations (Figure 2.5). Alternatively, when the 85% and 75% certainty thresholds were used, a survival increase was detected in 61 out of 100 and 75 out of

100 simulations, respectively (Figure 2.5). Therefore, 25% of the time the survival increase was not detected under any of the 3 certainty thresholds when only 1 site was monitored.

When we increased the number of simulated sites sampled, detection of increased survival improved across all uncertainty thresholds (Figure 2.6). As the number of sites increased from 1 site to 8 sites, mean detection of the management effect increased from 32% (95% BCI = 23–41%) to 95% (95% BCI = 92–98%) under 95% certainty, from 61% (95% BCI = 56–65%) to 98% (95% BCI = 97–99%) under 85% certainty, and from 74% (95% BCI = 65–82%) to 99% (95% BCI = 97–100%) under 75% certainty (Figure 2.6).

DISCUSSION

This study serves as a model for how to use existing data and simulation to assess the ability of a monitoring program to meet its objectives. In our case, we wanted to know whether the seining monitoring approach currently used for diamond-backed terrapins is sensitive enough to detect what we expect to be the likely effect of BRD regulations on adult male terrapin survival. Our analysis indicated that 66% of the time the current approach failed to detect simulated management success at a single site at a 95% level of certainty. Thus, we conclude that the current monitoring approach is not sensitive enough to evaluate management success at a single site if managers require a very high level of certainty.

While 95% confidence is a common criterion used to evaluate whether an effect is statistically significant, it may not always be practical to institute a monitoring program that can estimate management effects with this level of certainty. Many species of conservation concern are rare, data-limited, and difficult to sample, which tends to increase uncertainty when estimating population responses. Given that management decisions often need to be made regardless of monitoring precision, it may be beneficial to accept a lower threshold as the criterion for measuring when management is needed or if it is successful. Because the criterion adopted will likely affect the decision to continue or abandon the management approach, there is a tradeoff between the risk of falsely deeming a successful management action to be unsuccessful and vice versa (Shrader-Frechette and McCoy 1992, Di Stefano 2003). Incorrectly deeming a successful management action to be unsuccessful may lead to abandoning an effective management approach, and incorrectly deeming an unsuccessful management action successful may lead to wasting limited resources on continuing an approach that is not working. We considered 2 additional levels of certainty that managers may be willing to adopt: 75% and 85%. As expected, accepting higher uncertainty increased the likelihood of detecting simulated management effects. Even at the 75% certainty level, however, we failed to detect the simulated effect 25% of the time, indicating unreliability of the current monitoring approach to consistently meet its objective.

Although we were primarily interested in whether the precision of seining methods were sufficient to estimate a positive change in survival, the accuracy of our estimates may also be relevant to managers. Overall, posterior distribution mean estimates of the simulated change in survival were more often biased low than biased high (Figure 2.5); however, these mean estimates ranged from −0.17 to 0.47. Posterior distributions of the estimated change demonstrated that accuracy tended to increase with precision (Figure 2.5; Figure A.2). Specifically, estimates with wider credible intervals tended to underestimate the simulated increase in survival.

In addition to evaluating the statistical power of current monitoring, we sought to assess our capacity to improve monitoring precision by better understanding factors influencing the current system. We considered the potential to improve detection or availability by changing our monitoring approach (i.e., sampling during more optimal conditions) or design (i.e., adding additional sampling occasions). Improving this observation process could, in turn, increase the precision in estimates of the latent survival parameter that is the focus of monitoring (Williams et al. 2002). Our results revealed low feasibility of improving estimates when collecting data at a single site. The estimated mean daily capture probability was relatively high ($\bar{x}$ = 0.64, 95% BCI = 0.55–0.73), though within-year recapture probability was low ($\bar{x}$ = 0.02, 95% BCI = 0.01–0.04; Figure 2.4C). We can use the estimated daily capture probability to derive the probability of capturing an individual ≥1 time across all 3 secondary occasions within a year: $1 - (1 - 0.64)^3$ = 0.95, compared to the probability of capturing an individual at least once across 5 secondary occasions: $1 - (1 - 0.64)^5$ = 0.99. These results indicate that minimal gains in precision would be achieved by adding additional sampling days within a year. Additionally, model results suggest that the availability of terrapins for capture is largely influenced by tidal amplitude. Because we already conduct sampling during the lowest tide possible, this indicates that we have little capacity to improve availability. While the limited ability to improve capture probability and availability may be discouraging, identifying these limitations may help direct a more effective use of limited resources. For example, resources that may have been used to add labor-intensive survey days would be better allocated to adding monitoring sites or other priorities.

We assessed the effect of increasing the number of sites on our ability to detect a positive management effect (Figure 2.6). Detection of the survival increase surpassed 90% when 3 sites were sampled at the 75% certainty threshold, when 5 sites were sampled at the 85% threshold, or when 7 sites were sampled at the 95% threshold. Both bias and accuracy also improved as the number of sites sampled increased (Figure A.3). Thus, increasing the number of sites sampled may be an effective way to improve the current monitoring approach, with the recommended number of sites to sample dependent on the level of certainty required by managers.

In addition to improving precision of the estimated management effect, sampling multiple sites may more accurately represent what is occurring at the larger landscape scale (i.e., over many populations) in response to management (Gurevitch et al. 2016). While monitoring is often conducted on a small scale because of resource limitations, it can be dubious to base broad management decisions on results from a single site that might not be characteristic of the entire management area. Terrapin survival estimates, for instance, differ widely across sites regardless of crabbing pressure (and often have high uncertainty; Table 2.1).

Our conclusions regarding the capacity to improve monitoring at a single site rely on the candidate set of models considered. We chose our models based on our knowledge of the system and expectations for variation in the detection, availability, and survival probability parameters. Where we lacked explanatory variables (i.e., covariates) to directly explain variation in these parameters, we relied on random effects to allow parameters to vary as expected. It is possible that additional models not included in our candidate set could better explain variation in these parameters and identify other important variables that could be used to improve monitoring; however, we were unable to identify any additional data that we believed would be useful for explaining the parameters estimated here.

Our simulations assumed a starting population of 60 individuals and recruitment of 26 new individuals each year. This number of individuals directly influences the sample size (i.e., number of captures and recaptures) used to estimate the detection, availability, and survival probability parameters. In our existing dataset, which included 12 years of sampling, we had 496 captures of 322 unique male terrapins. Proportionally scaling up this number of individuals and captures to what could be expected with 15 years of sampling yields projected sample sizes of 620 captures and 403 individuals. These sample sizes fall within the range of values extracted from our 15-year simulated datasets, which ranged from 477–716 captures and 370–418 individuals across the 100 sites simulated. Therefore, we believe that these values chosen for our simulation were reasonable.

Our results also relied on the assumption that the effect of BRD regulations on terrapin survival is a 0.20 absolute increase in survival. We aimed to choose a plausible but large effect size to incorporate into our simulations. Had this effect been detected with high consistency, we had planned to conduct additional simulations with smaller effect sizes. Because the current sampling methods could not consistently detect such a large management effect, however, we deemed consideration of smaller effects to be unnecessary. While BRDs have demonstrably reduced the capture of terrapins when used (Chambers and Maerz 2019), there is a possibility that population-level survival increases could be minimal with management requiring BRDs, especially if compliance rates are low (Radzio et al. 2013) or if there are other factors that are reducing survival (Crawford et al. 2014*a*). Additionally, states may choose to require BRDs only in certain locations or at certain times (e.g., in New Jersey BRDs are required only in creeks <45.72 m wide; Chambers and Maerz 2019). It is also plausible that requiring BRDs could result in high survival increases in areas where crab traps currently catch a large proportion of the population (Roosenburg et al. 1997). Thus, the true effect size may be larger or smaller than what we simulated and could differ among sites.

In addition to increasing the number of sites monitored using capture-recapture methods, it may be useful to consider integrating additional monitoring methods to improve precision (Zipkin and Saunders 2018). Integrating multiple methodologies could also help resolve the issue of monitoring a small number of populations to address questions about broad-scale management, particularly if the added method is deployable on a large spatial scale. For instance, using boat trawls might improve capture probabilities for terrapins when tide amplitude is high, but this approach may be restricted to larger tidal creeks. Head counts of terrapins, which are less labor intensive than seining, can also provide information on abundance (Harden et al. 2007, Lovich et al. 2018, Levasseur et al. 2019), which could be used to assess population responses to management. While the head count method likely lacks sufficient sensitivity to detect effects of crabbing regulations on its own, combining >1 data source using an integrated model can improve accuracy and precision compared to models using a single data source (Zipkin and Saunders 2018). For terrapins, the application of an integrated modeling approach has already resulted in sufficient precision to detect the effectiveness of management directed at reducing road mortality of nesting adult females (Crawford et al. 2018). In this case, the same creek seining dataset used in the current study (through 2015) was integrated with a terrestrial capture-recapture dataset collected from adult females nesting along the Downing Musgrove Causeway. Because the focus of the current management requiring BRDs is expected to primarily affect male terrapins, we suggest that an approach integrating ≥2 types of data collected in aquatic habitat (e.g., repeated unmarked headcounts of terrapins [Levasseur et al. 2019] and capture-recapture data) should be explored.

MANAGEMENT IMPLICATIONS

Managers using the seining monitoring approach should consider reducing the level of certainty they use to evaluate management effectiveness, especially if they do not have additional resources to allocate to monitoring diamond-backed terrapins. Because we identified limited capacity to improve estimates via increased sampling effort at a single site or targeting better conditions, we recommend increasing the number of sites sampled if possible. Additionally, integrated models should be considered and evaluated as a means for bringing together multiple data sources to generate more credible estimates and to allow studying more populations over wider spatial scales and broader environmental gradients. Ultimately, modeling approaches used in this study are effective to transparently capture the level of certainty expected from different monitoring designs, but selecting a monitoring program will require managers to consider tradeoffs between the expected certainty of information, resource costs, and risk tolerance, among other factors.

TABLES

Table 2.1. Diamond-backed terrapin survival estimates reported in the literature. Threats listed include any anthropogenic threats reported by study authors that were likely influencing the study population. The estimates used for setting post-crabbing regulation mean survival probability for simulations of male terrapin capture-recapture datasets (this study) are indicated with an asterisk. These estimates were relevant for setting this simulation parameter because they include males and are from sites where crabbing is believed to be low or to not occur.

Location	Years	Sex	Survival estimate (95% CI)	Threats	Citation
Sandy Neck Beach and Great Marsh, Barnstable, MA	1980–1996	Both	0.83* (0.81–0.86)	None	Hart (1999)
Kiawah Island, SC:	1982–1998			High crabbing; boating (minor threat)	Tucker et al. (2001)
Oyster Creek		Males	0.79 (0.72–0.85)		
		Females	0.75 (0.66–0.82)		
Terrapin Creek		Males	0.83 (0.74–0.90)		
		Females	0.85 (0.70–0.93)		
Fiddler Creek		Males	0.82 (0.79–0.85)		
		Females	0.79 (0.75–0.83)		
Sandy Creek		Males	0.90 (0.55–0.99)		
		Females	0.97 (0.04–1.00)		
Narragansett Bay, RI	1990–2001	Females	Decreasing from 0.96 (0.93–0.98) to 0.94 (0.89–0.97)[a]	Subsidized predators	Mitro (2003)
Big Sable Creek complex, Everglades National Park, FL	2001–2003	Both	0.79* (0.60–0.91)[b]	None	Hart (2005)
North Inlet-Winyah Bay, SC	2006–2011	Both	0.78*[c]	Low crabbing	King and Ludlam (2014)
Kiawah Island, SC:	2003–2013	Both		High crabbing; habitat modification; road mortality; subsidized predators	Witczak et al. (2014)
Oyster Creek			0.82 (0.76–0.87)[d]		
Fiddler Creek			0.78 (0.72–0.83)[d]		
Sandy Creek			0.71 (0.66–0.76)[d]		
Stingray Slough			0.66 (0.50–0.79)[d]		
Terrapin Creek			0.61 (0.30–0.85)[d]		

Jekyll Island, GA	2010–2015			Road mortality; subsidized predators; habitat modification (i.e., roadside vegetation)	Crawford et al. (2018)
		Males	0.68* (SD=0.082)[e]		
		Females	0.72 (SD=0.075)[e]		
Jamaica Bay, NY:		Females	Negative trend:	Habitat modification; pollution; boating; poaching; crabbing; subsidized predators	Roe et al. (2022)
Ruler's Bar Hassock	2003–2018		0.92 (0.88–0.95) to 0.89 (0.86–0.92)		
John F. Kennedy Airport	2011–2018		0.92 (0.89–0.95)		

[a] Model-averaged estimates, which were primarily influenced by constant survival (top model) and linear trend (second-ranked model).

[b] Estimate from the top-ranked model. See Hart (2005) for estimates from the second- and third-ranked models, including sex-specific survival estimates from the third-ranked model.

[c] Mean model-averaged annual survival estimate. The top model included a time effect on survival and model-averaged estimates ranged from 0.69–1.00. See figure 3 in King and Ludlam (2014) for bar plot of year- and sex-specific model-averaged estimates with 95% CIs.

[d] Estimates from the top-ranked model. See Witczak et al. (2014) for sex-specific survival estimates (from fifth-ranked model) for comparison to sex-specific estimates reported in Tucker et al. (2001).

[e] Standard deviation reported in lieu of 95% confidence intervals. Female estimate represents a grand mean. See Crawford et al. (2018) for female survival estimates specific to their state (road crossing versus not crossing) and pre- versus post-management.

FIGURES

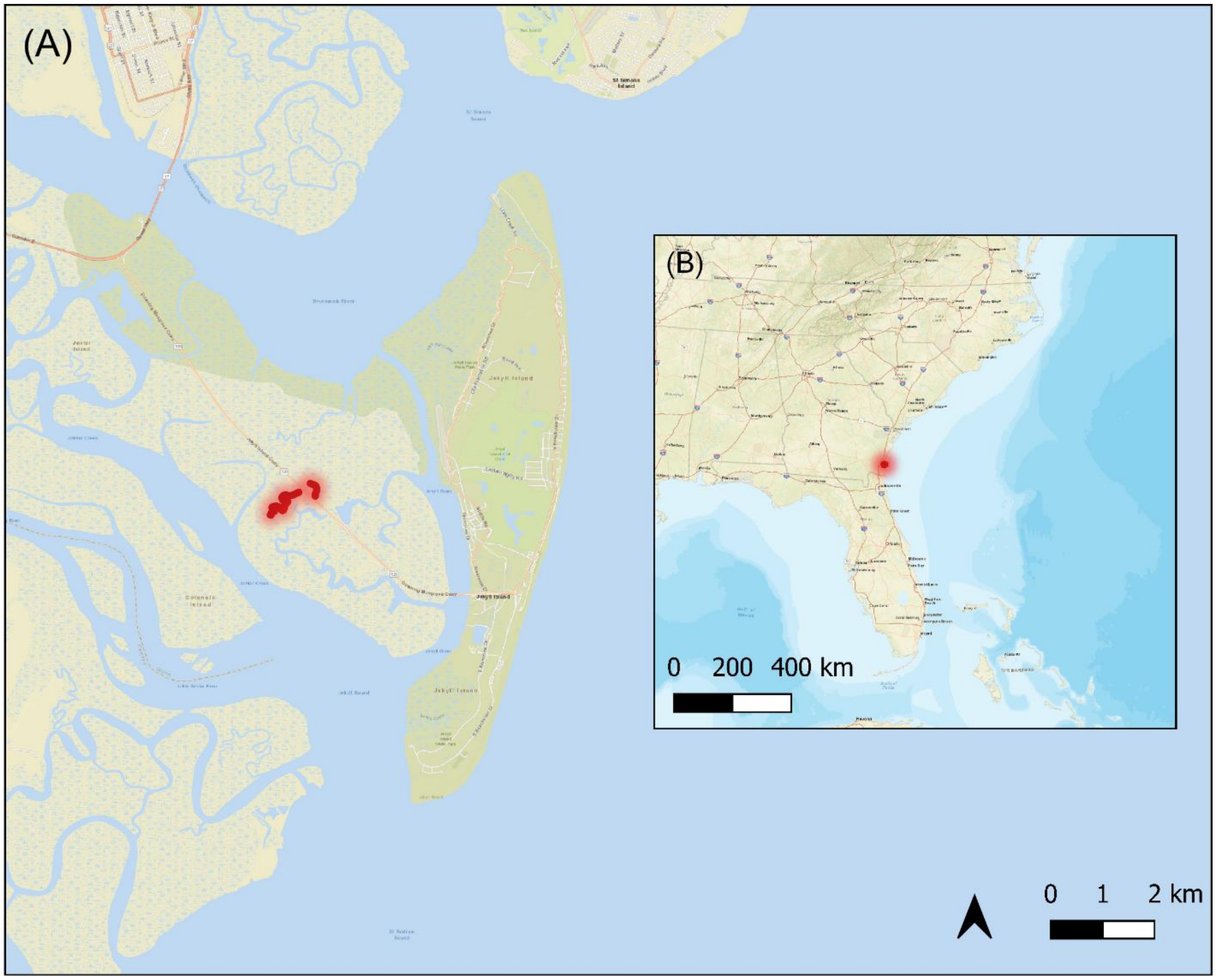


Figure 2.1. Location of 2 tidal creeks (red lines with glow effect) sampled for diamond-backed terrapins using capture-recapture methods annually from 2010 to 2022 (except for 2020). Creeks were adjacent to and on opposite sides of the Downing Musgrove Causeway leading to Jekyll Island in Glynn County, Georgia, USA. A) Overview of study site centered on Jekyll Island and surrounding salt marsh and B) inset displaying context of study location along the East Coast.

Figure 2.2. Seining method used for diamond-backed terrapin capture-recapture sampling conducted annually from 2010 to 2022 (except for 2020) within 2 salt marsh tidal creeks adjacent to the Downing Musgrove Causeway in Glynn County, Georgia, USA. Nets were staggered as needed to maximize coverage of the width of the creek.

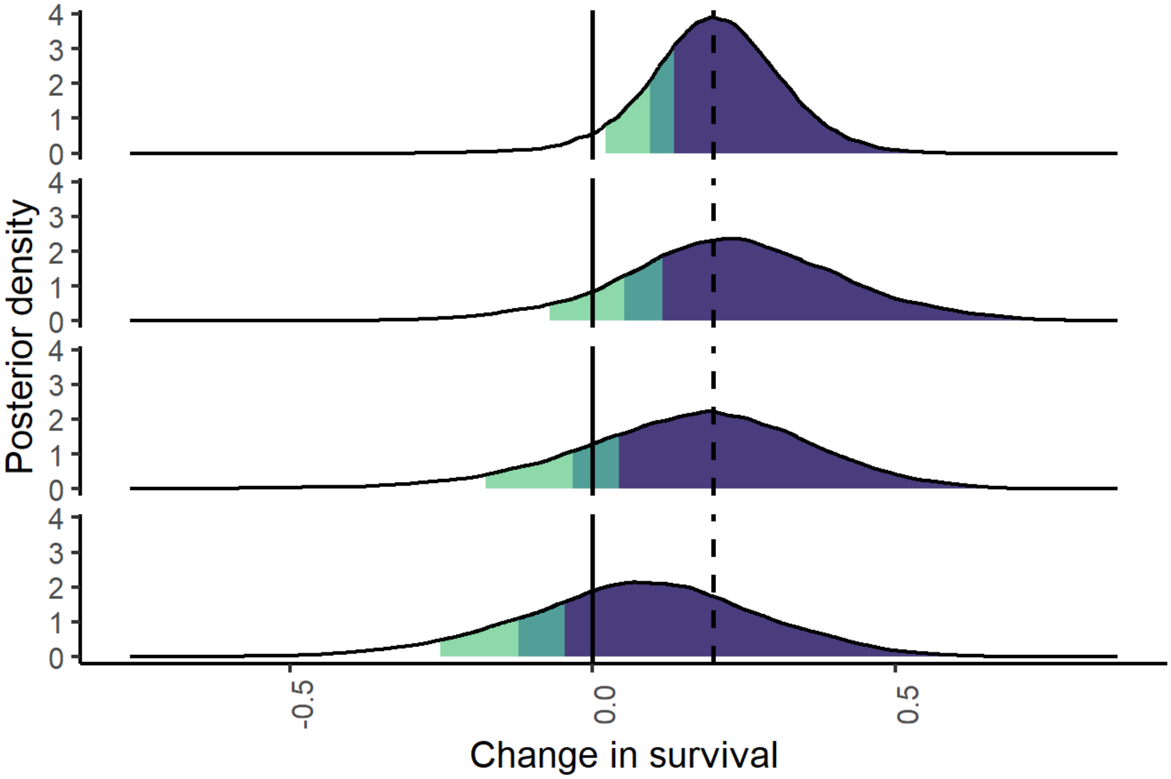


Figure 2.3. Examples of posterior distributions of the estimated change in mean apparent survival for 4 out of 100 simulated diamond-backed terrapin capture-recapture datasets, based on data collected in 2010–2022 near the Downing Musgrove Causeway in Georgia, USA. We estimated the change in survival using a robust-design Cormack-Jolly-Seber model. We selected simulations to exemplify 1 of 4 possible scenarios of detection or non-detection of the simulated management effect under the 3 levels of certainty considered. The shaded area (including all 3 colors) includes the upper 95% of the posterior distribution, the 2 darkest shades (dark green, purple) include the upper 85% of the posterior distribution, and the darkest shade (purple only) represents the upper 75% of the posterior distribution. The solid line is at no change (0), and dashed line is at the simulated change (0.20). Thus, from top to bottom, each plot displays the following probabilities that survival increased: >95%, >85%, >75%, and <75%.

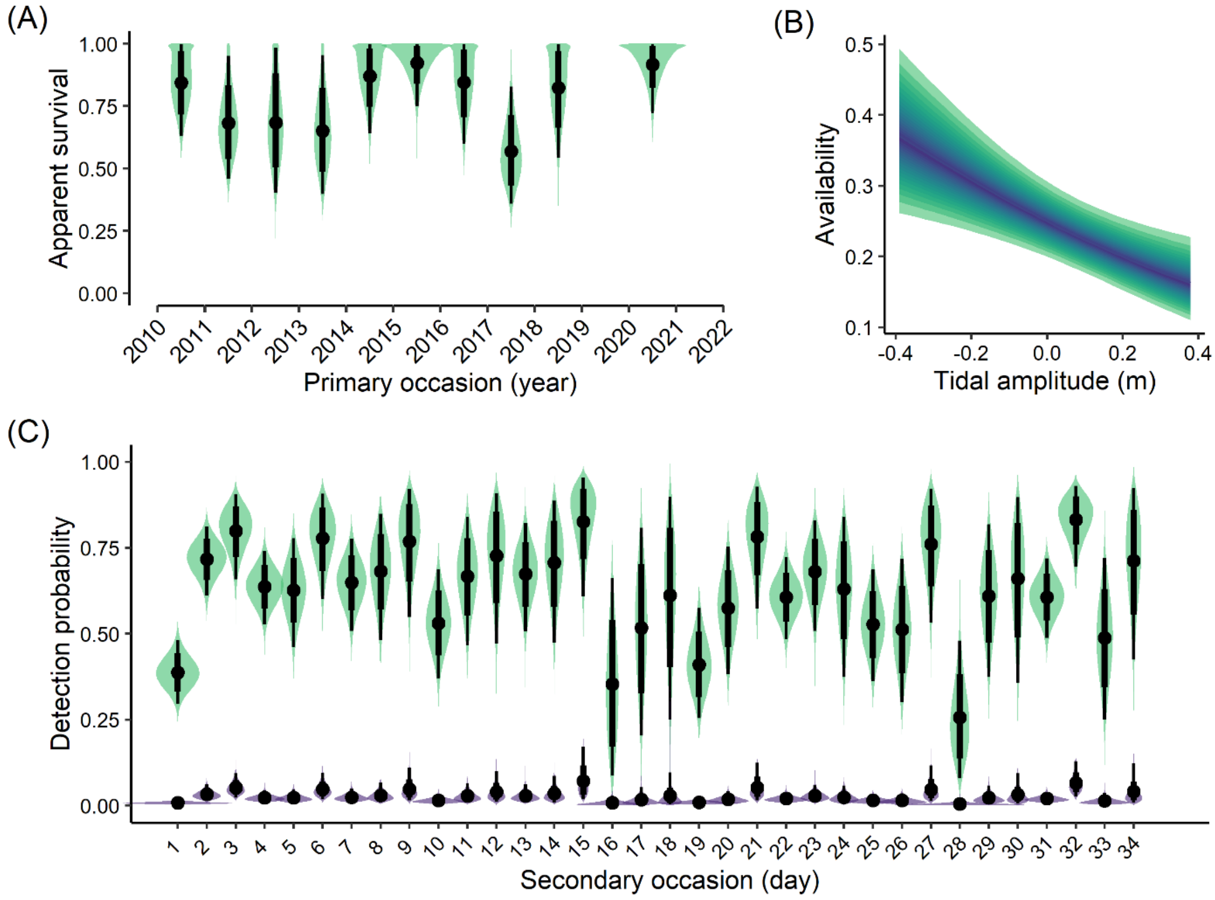


Figure 2.4. Estimates from the top-ranked robust-design Cormack-Jolly-Seber model fit to diamond-backed terrapin capture-recapture data collected from 2010 to 2022 (except for 2020) from tidal creeks in Glynn County, Georgia, USA. A) Apparent survival probabilities (φ) estimated for primary occasion intervals (between years). Violin plots (green) show full posterior distribution, with posterior mean (points), 75% Bayesian credible interval (thick uncertainty bars), and 95% Bayesian credible interval (thin uncertainty bars). B) Availability (γ) probability as a function of tidal amplitude, with posterior mean represented by dark purple line and Bayesian credible intervals increasing at 5% intervals represented by the multilayered ribbon surrounding the mean. C) Detection (i.e., capture and recapture) probabilities for each of the 34 secondary occasions (days) sampled. Upper estimates (green violins) represent capture probabilities and lower estimates (purple violins) represent within primary period recapture probabilities.

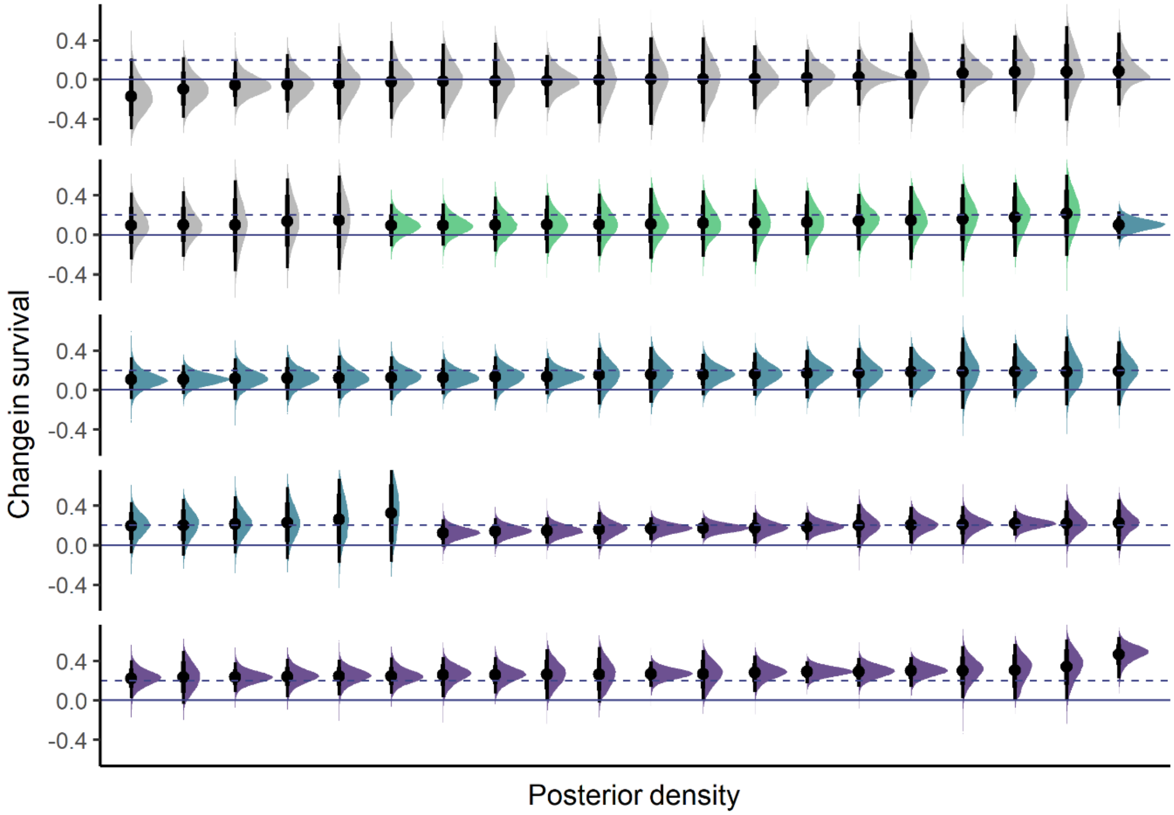


Figure 2.5. Posterior distributions of the estimated change in mean apparent survival for all 100 simulated diamond-backed terrapin capture-recapture datasets, based on data collected in 2010–2022 near the Downing Musgrove Causeway in Georgia, USA. We estimated the change in survival using a robust-design Cormack-Jolly-Seber model. These posterior distributions can be interpreted as results from 100 identical but independent simulated sites. Full posterior distributions are displayed, with posterior mean (points), 75% Bayesian credible interval (thick uncertainty bars), and 95% Bayesian credible interval (thin uncertainty bars). From top left to bottom right, results are sorted first by certainty threshold met (low to high) and then by posterior mean (low to high). We present posteriors of simulations where a positive management effect was not detected at any of the 3 certainty thresholds (grey), and where a positive effect was detected at only the lowest level of certainty (75%; green), the 2 lowest levels of certainty (75% and 85%; blue), and all 3 levels of certainty (75–95%; purple). The solid line is at no change (0), and the dashed line is at the simulated change (0.20).

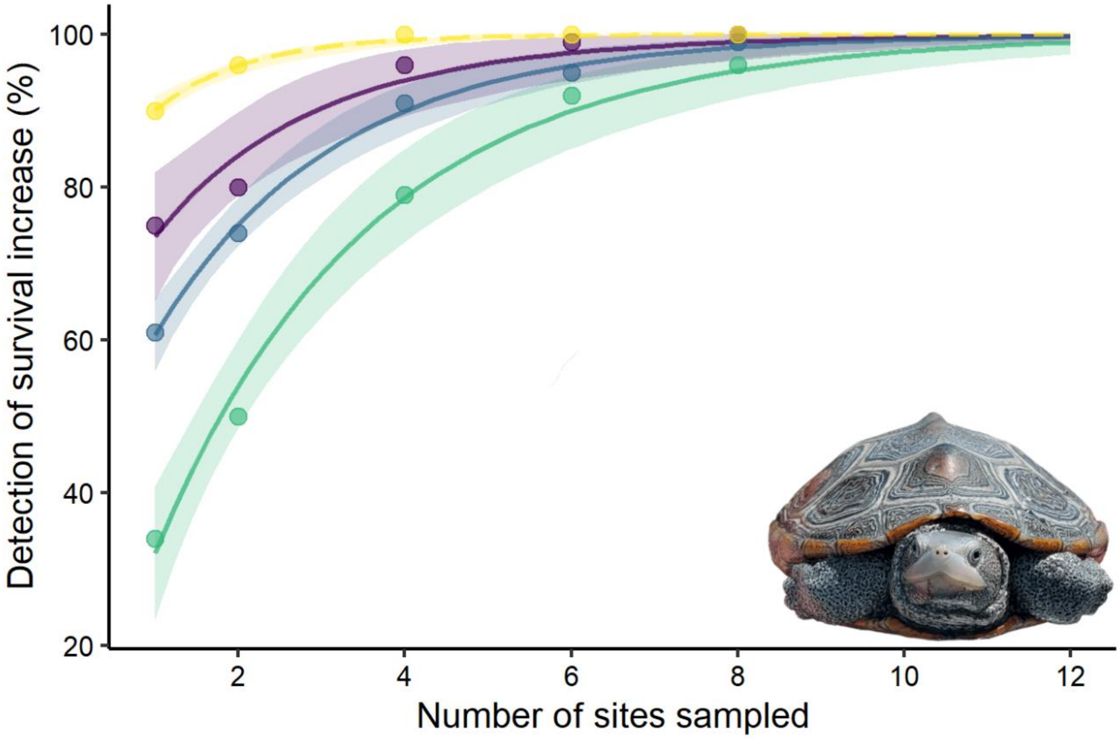


Figure 2.6. Percent of diamond-backed terrapin capture-recapture simulations with simulated management effect (i.e., survival increase) detected as a function of the number of sites (i.e., simulated capture-recapture datasets) sampled. Simulations were based on data collected in 2010–2022 near the Downing Musgrove Causeway in Georgia, USA. Points represent the percent detection at each number of sites considered (1, 2, 4, 6, or 8), lines represent means of the asymptotic regression functions fit to these points, and ribbons represent 95% Bayesian credible intervals. Colors differentiate the 3 certainty thresholds considered: 75% (purple), 85% (blue), 95% (green). Yellow (dashed line) represents the percent of simulations with posterior distribution mean for the estimated change in survival > 0.

CHAPTER 3

POTENTIAL AND LIMITATIONS OF UNMARKED POPULATION SURVEYS TO INFORM STATE-WIDE MANAGEMENT FOR DIAMOND-BACKED TERRAPINS [2]

---

[2] Bradke, D.R., M. Dodd, and J.C. Maerz. To be submitted to *Journal of Applied Ecology*.

ABSTRACT

Diamond-backed terrapins (*Malaclemys terrapin*) are challenging to monitor in their salt marsh habitat. However, large-scale monitoring is needed to understand the status of populations and effectively apply targeted management. The Georgia Department of Natural Resources (GA DNR) is interested in developing a long-term monitoring program that it can use for tracking relative levels of terrapin abundance among creeks throughout the state to trigger varying intensities of management actions. The objective of this study was to investigate the potential of repeated head count surveys and binomial N-mixture models to inform the state-wide monitoring and management objectives of the GA DNR. The unmarked repeated count method generated sufficient precision to categorize survey areas into three levels of relative abundance, while accounting for uncertainty in estimates. However, relative abundance estimates were sensitive to the timing of sampling within the active season and location of sampling along a given creek, indicating limitations of this method. Results also showed that as survey effort increased, the number of tidal creek areas that could be confidently assigned to higher abundance categories also increased. Therefore, if an objective is to minimize the number of sites requiring more intensive and potentially more restrictive management actions, a high level of survey effort is recommended. Because relative abundance estimates using head count methods varied with timing and location of sampling, survey methods would need to be standardized and consistent across years. If applied judiciously, a method using head count surveys and N-mixture models may be the most effective means of state or region-wide monitoring of terrapins across large numbers of sites. Additionally, if integrated with ongoing, long-term capture-recapture studies at a smaller number of sites, the capacity for head count surveys to inform terrapin management decisions across large areas and many sites could be enhanced.

INTRODUCTION

Spatially extensive wildlife monitoring is typically costly, time consuming, and logistically challenging (Jones 2011). This can be especially true for difficult to capture species in relatively inaccessible environments. Consequently, it is common for researchers to focus on a small number of sites to make inferences about what factors drive population changes, and for managers to make decisions or evaluate actions using data collected at only a few sites (Lindenmayer et al. 1993, Beissinger and Westphal 1998, Ellner et al. 2002, Gurevitch et al. 2016). However, drawing robust conclusions about population drivers or management effects of wide-ranging species necessitates monitoring many populations over a large spatial scale. Otherwise, the application of monitoring results may reflect an understanding of one or a few uncharacteristic sites.

Using counts of unmarked animals, rather than more costly capture-recapture data, is a common means of scaling up the number of sites that can be monitored with available resources. Binomial N-mixture models (Royle 2004) in particular, have become a widespread tool to estimate abundance because they use repeated counts to correct for imperfect detection, similar to capture-recapture abundance models. Collecting count data is appealing because it may require less effort, training, time, or resources than methods requiring animals to be in-hand for individual identification. However, N-mixture models are less robust than capture-based models to assumption violations and, thus, are more prone to bias (Duarte et al. 2018, Knape et al. 2018, Link et al. 2018). For example, the assumptions of closure and constant detection are difficult to meet in many studies and violations of these assumptions can severely bias estimates.

Because of the bias documented in abundance estimates from N-mixture models, these models may not be suitable to estimate abundance in many scenarios. However, N-mixture models (and other methods that can account for variable detection: e.g., Poisson regression) may still provide a useful index of relative abundance (Barker et al. 2018). Thus, whether count data and N-mixture models are a functional solution for large-scale monitoring is likely context dependent, making it necessary to investigate under what, if any, circumstances these methods can be used to inform a specific monitoring objective for a given species.

The diamond-backed terrapin (*Malaclemys terrapin*) is a unique estuarine habitat specialist that is challenging to monitor in its salt marsh habitat, which can be difficult to access and traverse. It is considered a high priority species or Species of Greatest Conservation Need (SGCN) in wildlife action plans throughout its range, including Georgia (Georgia Department of Natural Resources Wildlife Resources Division 2015), where available estuarine habitat consists of expansive tidal salt marshes. Terrapin populations are threatened by multiple factors including vehicular road mortality during nesting, high nest depredation by subsidized predators, and coastal development (Isdell et al. 2015, Chambers and Maerz 2019, Maerz et al. 2019). Additionally, one of the greatest present-day threats to terrapin populations is commercial and recreational crab fisheries because of the tendency of terrapins to enter and drown in crab traps (Roosenburg 2004, Grosse et al. 2009, 2011, Isdell et al. 2015, Lovich et al. 2018, Chambers and Maerz 2019).

The Georgia Department of Natural Resources (GA DNR) is interested in developing a long-term monitoring program that it can use to identify and track relative levels of terrapin abundance among creeks throughout the state and trigger varying intensities of management actions. Such management actions may range from recreational and commercial fisher awareness efforts to the requirement of bycatch reduction devices or restricting crabbing within specific areas of marshes. By implementing targeted, status-dependent management actions, the DNR aims to minimize the need for more time intensive, cost intensive, or restrictive actions while stabilizing and recovering terrapin populations. The persistence of low terrapin abundance in some areas might also prompt future actions such as efforts to reintroduce terrapins to areas of historic abundance. Using count data and N-mixture models to monitor terrapins in tidal creeks is unlikely to produce unbiased estimates of abundance because of expected assumption violations of both closure and constant detection. The closure assumption is difficult to meet because terrapins can easily move through creek survey areas even within short time periods. Additionally, individual heterogeneity in terrapin size and behavior (e.g. surfacing frequency and duration) will likely induce variable detection. However, if differences in detection probabilities among sites or surveys can be identified and accounted for using covariates in the N-mixture model and if assumption violations are similar in magnitude among sites and surveys, such a monitoring method may still be useful to the DNR as an index of relative abundance.

Following an intensive statewide two-year study among 29 tidal creeks to assess the effects of commercial crabbing pressure and road proximity on terrapin abundance (Grosse et al. 2011), terrapin monitoring methods in Georgia have consisted of annual capture-recapture surveys at 2-4 creeks (e.g., Bradke et al. in press., Crawford et al. 2018, GA DNR Coastal Resources Division unpublished data) and systematic sampling of nesting females along a single causeway to Jekyll Island (Crawford et al. 2014*b*, *a*, 2018). Capture-recapture within tidal creeks is time and labor intensive and, therefore, difficult to use over many sites and long periods of time. For example, sampling to estimate abundance in the 29 creeks studied by Grosse et al. (2011) took two years and cost ~ $250,000 to complete.

Unmarked methods that do not require capture and individual identification have also been considered for monitoring terrapins in Georgia. A study by Harden et al. (2009) showed a positive correlation between number of terrapins counted during a head count survey and the number subsequently captured when seining the creek, though there was a high amount of uncertainty in the relationship. Harden et al. (2009) recommended refinement of head count survey methods as a means to improve terrapin monitoring. For example, Harden et al. (2009) found that significantly more terrapins were observed during counts conducted at low tide. Lovich et al. (2018) subsequently used head counts at low tide and a single detection covariate, cloud cover, to show that terrapin abundance among tidal creeks within the Savannah Coastal Refuges Complex, Georgia, was negatively correlated with the number of commercial crab pots at the refuge unit scale. Though Harden et al. (2009) and Lovich et al. (2018) incorporated tide and cloud cover and Harden et al. (2009) included other detection related covariates, a limitation is that they did not incorporate repeated samples to address the issue of imperfect detection during counts. In both studies, the authors summed head counts from two passes within a creek. Recently, a study in Wellfleet Bay, Massachusetts demonstrated that repeated terrapin head counts within a relatively “closed” sample modeled under a binomial N-mixture model framework was potentially an improved method for monitoring terrapin abundance (Levasseur et al. 2019). The Wellfleet Bay study generated renewed interest in applying a similar methodology to meet terrapin monitoring objectives in Georgia; however, the use of repeated head counts has not been evaluated in the context of actual management objectives.

The objective of this study was to investigate the capacity and practicality of repeated terrapin counts to meet large-scale monitoring and future management objectives. The usefulness of estimates from head count surveys to GA DNR depends on obtaining sufficient precision to place tidal creeks into reliable categories of abundance. Though prior studies including the Wellfleet Bay study are encouraging, a key difference between Wellfleet Bay and the salt marshes of Georgia is that Wellfleet Bay is a relatively open bay with good, generally unobstructed observer visibility. In contrast, tidal creeks in Georgia are narrow and winding, creating more limited fields of view. Most of Georgia's tidal creeks require a boat to access, creating some logistical constraints and costs, and some of Georgia's tidal creeks are inaccessible by boat during low tide. The incorporation of land-based surveys of some tidal creeks (e.g., at bridge overpasses or boat launches and docks) could increase the capacity for monitoring a greater number of sites, but surveys from land may create additional limitations (e.g., site selection, field of view). Before the methodology is used, there is a need to evaluate the robustness of relative abundance estimates from head count surveys and to understand how the timing of sampling within the active season, location of sampling along a given creek, and amount of survey effort (e.g., number of repeated surveys per visit) affect estimates. Understanding how these factors affect which category of relative abundance a creek may be assigned to will inform optimal monitoring design and the ability of GA DNR to conduct informative monitoring given resource constraints and the agency's other management obligations.

METHODS

Data collection

We used a modified version of the protocol described by Levasseur et al. (2019). During 4 April – 13 July 2021 and 21 April – 13 July 2022 we conducted repeated counts of terrapins in tidal creeks at 58 total tidal creek sites spanning the entire Georgia coast, including multiple survey areas at a subset of tidal creek sites (119 total survey areas; Figure 3.1). Each site was surveyed 1 to 5 times per year (in at least one year) and included 1 to 6 survey areas, resulting in 326 surveys of 112 survey areas representing 56 sites in 2021 and 280 surveys of 83 survey areas representing 48 sites in 2022. Size of survey areas averaged 6,104 m^2 (range 739 – 21,552 m^2). To target times when terrapins would be present within creeks versus high marsh habitat, we conducted all surveys within approximately three hours of low tide when the water line was below the high marsh (delineated by the presence of smooth cordgrass; *Spartina alterniflora*; Harden et al. 2009). Like Levasseur et al. (2019), we conducted each survey from a fixed survey point. At each visit to a survey point, we defined our survey area – within a maximum distance of 100 m from the observer – using a range finder. We recorded the number of crab pots in and around the survey area. We recorded average wind speed (m/s; based on recording wind for approximately 1 minute) using a kestrel meter, salinity (ppt) using a traceable salinity pocket tester (Cole-Parmer model EW-19601-21), cloud cover (No clouds, partial cloud cover <50%, ~50% cloud cover, partial cloud cover >50%, overcast), and the perceived effect of waves and glare on the observer's ability to observe terrapins (no issue, minor issue, major issue). We performed 5-10 repeated scans of our defined survey area per observer, with one minute between each scan to allow for mixing by individuals (Levasseur et al. 2019), and we recorded the number of terrapins counted during each scan by each observer. Most surveys were conducted by a single observer, but a subset of surveys included two or three observers.

We used geographic information system (GIS) software and spatial layers to measure covariates that we hypothesized may explain terrapin abundance. These data included the size of each survey area (m^2) and the distance of survey area to nearest potential nesting habitat. To identify potential nesting habitat, we used a digital elevation model produced using lidar data for the National Oceanic and Atmospheric Administration (NOAA; National Oceanic and Atmospheric Administration (NOAA) Office for Coastal Management 2016). Based on overlaying the digital elevation model on top of satellite imagery of the Georgia coast we established that the minimum elevation where terrapins are known to nest was 1.5 m. Therefore, we deemed any land surrounding the marsh that was ≥ 1.5 m elevation to be potential nesting habitat and measured straight-line distance of the survey areas to this habitat. We also calculated the proportion of the shoreline that was armored within 100 m of survey areas. To do so, we considered the shoreline to be the perimeter of all land ≥ 1.5 m elevation and we determined the proportion of that land that was armored (i.e., classified as bulkhead, revetment, or offset) using shoreline armoring spatial data developed by the Skidaway Institute of Oceanography from The Georgia Wetlands Restoration Access Portal (Alexander 2015).

We determined if crabbing was known to occur within or near each survey area using observations of crab pots recorded during our data collection and using all additional records of crabbing within Georgia that we were able to obtain. These records included seining surveys (2007-2008; Grosse et al. 2011), drone surveys (2009; J. Maerz unpublished data), and crab pot surveys conducted by the Department of Natural Resources (2004-2009; GA DNR Coastal Resources Division, unpublished data). We combined all datasets to determine whether a site either had known crabbing activity or did not have known crabbing activity.

Model selection

Using a Bayesian framework, we fit a binomial N-mixture model (Royle 2004) to repeated count data to estimate detection probability (p) and abundance (N). We used a static model where we modeled each survey area-survey combination as a separate survey area. We modeled each count (y) at survey area (i) during scan (k) as a binomial outcome drawn from abundance (N_i) with detection probability ($p_{i,k}$) and we modeled N_i as a Poisson outcome with mean equal to expected abundance (λ_i).

$$N_i \sim Poisson(\lambda_i)$$

$$(y_{i,k}|N_i) \sim Binomial(N_i, p_{i,k})$$

Because our data likely violated N-mixture model assumptions (i.e., closure and constant detection) we considered N an estimate of relative abundance rather than absolute abundance.

To select variables important in explaining detection probability and relative abundance, we modeled count data from all surveys across both years and used indicator variable (IV) selection. We did not include a random effect of survey area on abundance to account for non-independence among multiple surveys at the same survey area, but we included a random effect of site to account for non-independence among all survey areas and surveys within the same creek. Survey area size was also included as a covariate to explain abundance. Additional abundance covariates that we included and assessed using IV selection were known crabbing activity, proportion of shoreline armoring, distance to potential nesting habitat, salinity, and day of year (including main effect and quadratic terms). We constrained λ_i using a log link and the above-described covariates and random effect. We also included four covariates to explain detection probability, which we assessed using IV selection. These covariates were average wind speed, cloud cover, perceived effect of waves, and perceived effect of glare. We constrained p using a logit link and these four covariates.

To implement IV selection, we used "slab and spike" priors (Hooten and Hobbs 2015, Lawson et al. 2022) and multiplied each beta coefficient associated with covariate x (β_x) by a binary IV (w_x). Each IV was drawn from a bernoulli distribution with probability $p.w_x$, which was given a uniform prior between 0 and 1. Beta coefficients were given normal priors with mean of 0 and variance equal to $(1 - w_x) \times (1 \div \text{Spike}) + w_x \times (1 \div \text{Slab})$, where Spike indicates the value used for spike precision (1,000) and Slab indicates the value used for slab precision (0.01 for abundance covariates or 0.37 for detection covariates). The proportion of Markov chain

Monte Carlo (MCMC) iterations where $w_x = 1$ indicated the amount of support for the covariate in the model. We derived model weights for all possible combinations of IV selection covariates (1,024 models), where weights were equal to the proportion of MCMC iterations in which the covariate combination appeared. Because we were not interested in all possible combinations of covariates, we re-distributed model weights to compare only models of interest (432 models). Models eliminated from model selection included all models with the quadratic term of day of year but not the associated main effect, because these models were not expected to be biologically meaningful. We also eliminated models that had both wind speed and wave effects or both cloud cover and glare effects because each of these pairs of variables were considered two different measurements of the same effect. We selected the top-ranked model (model with highest weight) for estimating relative abundance (described below).

We standardized covariates prior to analysis and fit the model in JAGS (Plummer 2003) with the jagsUI package (Kellner 2021) in program R (4.2.1; R Core Team 2022). We used vague prior distributions on all parameters not included in IV selection (i.e., mean detection probability, mean expected abundance, standard deviation of random site effect on abundance, beta coefficient for survey area size effect on abundance; Table B.1). We used 3 MCMC chains with 350,000 iterations each and burn-in of 50,000. We assessed chain convergence visually using trace plots and with Gelman and Rubin (1992) diagnostic values (i.e., $\hat{R} < 1.1$).

Estimation and categorization of relative abundance

To obtain estimates of relative abundance across all sites and surveys, we ran the top-ranked model identified by IV selection using vague prior distributions (Table B.1). We again used 3 MCMC chains with 350,000 iterations each and burn-in of 50,000 and assessed chain convergence visually using trace plots and using $\hat{R}$ values. We assessed goodness-of-fit by calculating Chi-square discrepancy statistics, computing a Bayesian *p*-value, and estimating c-hat (Kéry and Royle 2016).

To improve comparability and interpretability of relative abundance estimates, we scaled each estimate by 6,000 m^2, which was the approximate size of our average survey area. A creek that was approximately 30 m wide and surveyed 100 m in each direction lengthwise would be ~ 6,000 m^2 (Figure 3.2). We categorized estimates of relative abundance using three thresholds that could be used by the DNR to determine the intensity of management actions to implement within each creek. The lowest threshold was < 1 terrapin per 6,000 m^2, representing survey areas that were virtually without terrapins. The intermediate threshold was ≥ 1 but ≤ 10 terrapins per 6,000 m^2. The highest threshold was > 10 terrapins per 6,000 m^2, representing moderate to high relative abundance. To be placed in the intermediate or highest abundance categories, we required a 95% probability that relative abundance met or exceeded the respective thresholds. In other words, 95% of the posterior distribution had to be at or above the respective threshold.

In addition to the three categories described above, we also compared our estimates of relative abundance to estimates intended to represent a natural density of terrapins in creeks not under any crabbing pressure. To do this, we used estimates from Grosse et al. (2011), who estimated density of diamond-backed terrapins per 1 km of creek within 29 tidal creeks that experienced a range of crabbing pressure. To scale densities estimated by Grosse et al. (2011) to our estimates, we measured mean creek width of the Grosse et al. (2011) study sites using GIS software. We then multiplied the 1 km creek length used in the Grosse et al. (2011) study by the measured mean creek width (20 m), to infer an average study area of 20,000 m^2. Scaling the Grosse et al. (2011) mean density estimates (range 100 – 361 per 20,000 m^2 area of creek) to our mean study area yielded 30 – 108 individuals per 6,000 m^2. Thus, we quantified the number of surveys of our study areas with 95% of the posterior distribution ≥ 30 terrapins per 6,000 m^2.

Because 13 of our 58 sites overlapped with sites seined by Grosse et al. (2011), we also compared abundance estimates at these sites. Again, we re-scaled densities estimated by Grosse et al. (2011) to facilitate comparability between the studies. We approximated areas associated with 1 km length of creek for each of the 13 sites using the above mentioned GIS measurements of mean creek width (i.e., 1 km length x mean creek width) and re-scaled the Grosse et al. (2011) density estimates for each of these areas to 6,000 m^2.

Evaluating sensitivity of estimates to timing, location, and effort

To assess the sensitivity of relative abundance estimates to the timing of sampling within the active season we selected all survey areas visited ≥ 2 times within the same year and we quantified the percentage of time that each survey area was placed in each category of abundance. To evaluate whether relative abundance estimates varied based on location of sampling along the creek, we selected creeks (i.e., study sites) with more than one survey area and quantified the percentage of survey areas within a site surveyed on the same date that were placed in each category of abundance.

To assess the effect of survey effort on precision, we used a random subset of the full dataset, which included a single survey of each survey area each year. This sampling scheme is more representative of a plausible number of surveys for DNR biologists to conduct, given their time and labor constraints. Then we used either 10, 5, or 3 scans per survey to estimate relative abundance by fitting each of these three datasets to the top-ranked N-mixture model. We compared the proportion of sites placed into each relative abundance category across the three levels of survey effort (10, 5, or 3 scans). We used the same MCMC settings and model checking methods as described for the full-dataset analysis.

RESULTS

We conducted a total of 606 surveys across all survey areas, 71 of which had two or three observers. We detected terrapins at 235 out of 606 (39%) of surveys representing 38 out of 58 (66%) of our study sites. In total, we counted 2,202 terrapins (mean = 0.32 terrapins/scan; range = 0 – 17).

Model selection

The top-ranked model using IV selection included a random site effect and fixed effects of survey area size, distance to potential nesting habitat, salinity, and day of year (main effect and quadratic terms) to explain abundance and fixed effects of perceived effects of waves and glare to explain detection (Figure 3.3, Table B.2). Trace plots and $\hat{R}$ values for the IV selection model indicated successful convergence.

Estimation and categorization of relative abundance

Estimates had sufficient precision to categorize survey areas into relative abundance categories during each survey. Of all 606 surveys, 424 (70%) were in the < 1 terrapin per 6,000 m^2 category, 163 (27%) were in the 1 – 10 terrapins per 6,000 m^2 category, and 19 (3%) were in the > 10 terrapins per 6,000 m^2 category (Figure 3.4). Only one survey of one study area (surveyed twice in 2021 and once in 2022) had an estimated relative abundance ≥ 30 terrapins per 6,000 m^2. All relative abundance estimates within study sites overlapping Grosse et al. (2011) sites were lower in our study. Mean estimates for these sites within our study ranged from 0 – 39 individuals per 6,000 m^2 and re-scaled estimates from Grosse et al. (2011) ranged from 0 – 567 individuals per 6,000 m^2.

Variation of estimates with timing

Among survey areas visited ≥ 2 times within the same year, 66% of survey areas were consistently placed into the same category each time they were surveyed during the same year (58% < 1, 7% 1 – 10, and 1% > 10; Table 3.1). Of the remaining survey areas, 29% were placed into both the < 1 and the 1 – 10 categories and 5% were placed into both the 1 – 10 and > 10 categories (Table 3.1). Only one survey area was placed in all three categories across surveys within the same year (i.e., four surveys in 2021).

Variation of estimates with location

Among tidal creek sites with ≥ 2 survey areas visited on the same date, 72% had all survey areas placed within the same category (53% < 1, 19% 1 – 10, and 1% > 10; Table 3.2). Of the remaining study sites, 22% had one or more survey areas placed into each of the < 1 and the 1 – 10 categories, 3% had one or more survey areas placed into each of the 1 – 10 and > 10 categories, and 1% had one or more survey areas placed into each of the <1 and > 10 categories (Table 3.2). Four sites had survey areas that spanned all three categories across surveys on the same date, with one of these sites consisting of four survey areas and the other three sites consisting of five survey areas.

Trace plots and $\hat{R}$ values indicated successful chain convergence for the model used to evaluate sensitivity of estimates to timing and location (i.e., the top-ranked IV model fit to the full dataset). A Bayesian p-value of 0.14, c-hat of 1.04, and the scatter plot of discrepancy values (Figure B.1) indicated adequate fit of the model to the data.

Effect of survey effort on precision

As the number of scans of the survey area decreased, the ability to confidently place survey areas into the 1 – 10 and > 10 relative abundance categories also decreased (Table 3.3).

No survey areas shifted to a higher abundance category as less survey effort was used, but 14 survey areas shifted one category lower when scans decreased from 10 to five and 11 survey areas shifted one category lower when scans decreased from five to three (Figure 3.5; Table 3.3). Trace plots and $\hat{R}$ values indicated successful chain convergence for all three reduced dataset models. Bayesian p-values for the three models ranged from 0.10 to 0.22 and c-hat ranged from 1.09 to 1.14. Scatter plots of discrepancy values (Figure B.1) indicated adequate fit of all models to the data.

DISCUSSION

Results from this study demonstrate the potential of diamond-backed terrapin head count surveys conducted within tidal creeks to meet state-wide monitoring objectives. While this survey method demonstrated potential in Wellfleet Bay, Massachusetts (Levasseur et al. 2019), the tidal creek system of Georgia differs substantially from the bay environment. However, our results suggest that head count surveys may also have utility within salt marsh tidal creeks, which compose a large amount of the estuarine habitat used by terrapins (Roosenburg and Kennedy 2019). This unmarked repeated count method generated sufficient precision for us to categorize survey areas into three levels of relative abundance, while accounting for uncertainty in estimates, suggesting its potential to inform the implementation of management actions by the GA DNR.

The potential of this method is also bolstered by our identification of covariates that affect variation in detection, which suggests that using an N-mixture model may be an improvement over methods that do not account for imperfect and variable detection. Results of model selection indicated the importance of an observer’s perceived effects of waves and glare on detection probability, with waves producing a significant negative effect and glare producing a significant positive effect. The direction of the relationship between glare and detection was unexpected but reasonable because, while glare may be uncomfortable for the observer, the conditions created by glare can make terrapin silhouettes more visible against the water. Accounting for factors important in explaining detection probability can make this method a more reliable index of abundance than methods that do not account for such factors (Barker et al. 2018).

Although the precision of estimates and ability to categorize survey areas by level of abundance is promising, this method has limitations. Because survey areas lacked barriers to prevent migration and were small compared to the movement ability of terrapins (Harden et al. 2007, Lamont et al. 2021), study areas did not appear to be closed during surveys. Therefore, this survey method violates the assumption of closure implicit in binomial N-mixture models (Royle 2004). Data collected with closure violations result in estimates that are biased high but may be interpretable as the number of individuals associated with the survey area during the entire survey, rather than absolute abundance (Kery and Royle 2016). However, a second assumption of N-mixture models, difficult to meet in any wildlife monitoring study, was also likely violated; the “constant detection” assumption conditions that there is no unmodeled heterogeneity in detection probability (i.e., detection is constant or is fully explained by covariates; Martin et al. 2011, Kéry and Royle 2016, Barker et al. 2018, Duarte et al. 2018, Knape et al. 2018). Individual heterogeneity cannot be accounted for in N-mixture models because individuals are not uniquely identified, although heterogeneity among individuals is likely ubiquitous in wildlife monitoring (Kéry and Royle 2016, Veech et al. 2016). It is also likely that there is unaccounted for heterogeneity between replicate counts, surveys, and sites, despite our efforts to include the important, measurable factors that we expected to affect detection. Accordingly, we consider our estimates to represent relative abundance, acknowledging that they are likely biased. Notably, if violations of assumptions vary across space (e.g., different degrees of closure at different sites) estimates may not provide a valid index of abundance. While there is likely some variation in bias among sites, based on observations during surveys, we believe such variation is likely negligible.

Additional limitations of this unmarked survey method include that relative abundance estimates were sensitive to timing of sampling within the active season or location of sampling along a given creek. This is evidenced by inconsistencies identified in the assigned abundance category of a creek when sampling the same area on multiple dates and when sampling different areas along the same creek on the same date. The effect of day of year on mean expected abundance also corroborates that relative abundance is affected by timing. Thus, when and where sampling is conducted could affect management decisions.

We also assessed whether estimates were robust to amount of survey effort. The large range in tidal amplitude experienced in Georgia salt marshes during semidiurnal tide cycles creates a short (~ 6 hour) time window in which the water line is below the high marsh, limiting the number of sites that can be surveyed in a single day. Because sampling many sites is important to meet the monitoring objectives of the DNR, sampling efficiency is a priority. We found that as effort increased from 3 to 10 repeated counts of the survey area, the number of survey areas that were assigned to the higher abundance categories also increased. During our surveys, each scan of a survey area took an average of 69 seconds (range = 10 – 213 seconds/scan), with duration primarily affected by survey area size, but also impacted by other factors (e.g., waves). Thus, on average, with one-minute allotted between scans, a 3-scan survey takes 5.5 minutes, a 5-scan survey takes 9.8 minutes and a 10-scan survey takes 20.5 minutes.

This information can be useful in weighing the importance of conducting surveys over fewer days versus the importance of being able to confidently assign creeks to higher abundance categories, allowing less stringent actions at more sites.

Model selection indicated strong support for effects of salinity, distance to potential nesting habitat, and the day of year on relative abundance, but little support for the effects of known crabbing activity or proportion of shoreline armoring. The quadratic relationship observed between day of year and abundance is likely an artifact of nesting phenology and sampling design. Because of limited boat access, most surveys were conducted at sites that were accessible by land (n = 542 surveys by land vs 64 surveys by boat) – including bridges, fishing piers, public and private docks, uplands, and a few high marsh locations where the marsh was safely accessed by foot. Many land-accessible sites were on or adjacent to potential nesting habitat and often terrapins observed from these locations appeared to be females waiting for suitable conditions to nest (Feinberg 2004, Butler and Burke 2019). Conversely, most of the surveys conducted by boat were in more remote marsh locations (i.e., foraging habitat). Expected abundance peaked in late May, aligning with what we could expect for terrapin nesting activity in the region (Crawford et al. 2014*b*). Levasseur et al. (2019) similarly found that abundance varied seasonally in Wellfleet Bay, Massachusetts, consistent with phenological and spatiotemporal expectations for mating and nesting activity in the north (Levasseur et al. 2019, 2023). If future surveys are conducted by boat to target terrapins in their foraging habitat, as planned by GA DNR, we expect that the direction of the day-of-year versus abundance relationship will change, with abundance likely greatest early in the season (i.e., pre-nesting).

The positive relationship observed between expected abundance and distance to potential nesting habitat reflects the larger numbers of terrapins typically observed at sites accessed by boat (mean heads/scan = 0.61, range = 0-17) compared to sites accessed by land (mean heads/scan = 0.25, range = 0-9). The candidate model set we considered limited our ability to determine the biological mechanisms driving this relationship, but there are likely multiple contributing factors, including the nesting phenology of terrapins described above and the range of dates sampled, which included times outside of peak nesting season. Additionally, sites accessed by land included locations with higher human activity (e.g., boating, fishing) and locations near more development (e.g., creeks passing under or near roads) – factors that could affect habitat quality. Whether or not we observed crabbing activity in the sections of creek visible to us from our land-access survey points, these less-remote creeks also may be crabbed more heavily because of their convenient access. Future work could assess support for a model including an interaction term between distance to potential nesting habitat and day of year, if there are sufficient data to support the number of model parameters, to improve interpretation of the expected abundance and distance to potential nesting habitat relationship.

Counterintuitive in view of previous research (Harden and Southwood Williard 2019), we also observed that abundance increased with the average salinity of a site. While terrapins are adapted to brackish water environments, substantial energy is required for physiological osmoregulation and terrapins have been found to eat less and grow slower when in high salinity environments (Dunson 1985, Davenport and Ward 1993, Holliday et al. 2009, Southwood Williard et al. 2019, Ashley et al. 2021). Despite this, when other covariates were held at mean values, model predictions yielded an expected abundance of < 1 individual at salinities ≤ 30 ppt and > 1 individual at salinities ≥ 31 ppt. Additionally, we never observed terrapins at the five creeks with the lowest mean salinity records (range = 0.11 – 3.77 ppt) that otherwise appeared to be suitable habitat. Although the reason for this relationship is unknown, it may be important to control for the effect of salinity on abundance when making inferences about causal effects of suspected threats to terrapin populations.

Our inability to detect negative effects of crabbing or shoreline armoring on abundance differs from previous studies (Dorcas et al. 2007, Grosse et al. 2011, Isdell et al. 2015, Lovich et al. 2018). The measurement of crabbing activity used in our study was very coarse compared to these other studies. We could not account for amounts of crabbing effort (i.e., number of pots and duration of crabbing), locations of crab pots, and soak times which are factors that could influence the effect of crabbing on abundance. It is also possible that some creeks with no observed crabbing were being crabbed within areas not visible to us from our terrestrial survey points. Crabbers in Georgia are also known to move the location of their crab pots during the year, so it is possible that our surveys might not have coincided with crabbing at some sites. The lack of support for an effect of shoreline armoring may have been affected by the scale we considered (within 100 m of the survey area), which was smaller than the 1-km scale at which armoring was found to affect terrapins in Virginia (Isdell et al. 2015). It is also possible that our study sites did not have enough shoreline armoring to affect terrapin abundance. Among our study areas, the distribution of proportion of shoreline armoring at the 1-km spatial scale was highly skewed with 90% of values < 0.11 (range = 0 – 0.77), so we did not consider this scale in our analyses. Isdell et al. (2015) stratified sampling to survey a gradient of shoreline armoring and identified a threshold of 0.17 armored shore to be important at the 1-km scale. Therefore, our study areas appear to have generally less shoreline armoring than the study sites in Virginia.

Comparisons between this study and sites with naturally high terrapin densities (i.e., creeks without crabbing pressure) are difficult to make. No previous studies have used the same method to estimate relative abundance in the tidal creek environment; thus, comparisons should be interpreted cautiously. After re-scaling estimates from the largest spatial-scale capture-recapture study in Georgia to match our scale (individuals per 6,000 m^2), we found that only one of our study areas during a single survey exceeded the lower 95% confidence boundary for estimated abundance at sites with no crabbing pressure (Grosse et al 2011). Additionally, among 13 head count sites that overlapped with seining sites sampled by Grosse et al (2011), estimates using the head count method were consistently lower than estimates from seining. These studies were conducted 14 years apart and it is possible that abundance truly decreased within all 13 sites. However, N-mixture model estimates have been found to be biased low when detection is low, while capture-recapture model estimates were not (Couturier 2013). Because terrapins spend time below the surface of the water, may be hidden in vegetation within the high marsh, and can be otherwise difficult to observe, low detectability is a known issue. We also found that most of our surveys resulted in estimates in the < 1 individuals per 6,000 m^2 category. These low abundance estimates could be, in part, related to the sensitivity of estimates to location along the creek, and may not characterize abundance of the entire creek.

Because relative abundance estimates using head count methods are sensitive to timing and location of sampling, we recommend standardizing these variables across years. Although we sampled some creeks by boat, locations of sampling during our surveys were primarily dictated by accessibility by land so we could not control the locations of sampling along most creeks surveyed. It is not surprising that terrapin abundance varies along the length of a creek based on our previous observations during capture-recapture surveys, during which some sections of creek are observed to consistently have relatively more or less terrapins across sampling years. For sampling conducted by boat, locations could be selected to be more representative of terrapin abundance within the creek. While accessing creeks for head count surveys and other monitoring efforts, we observed that terrapins typically tended to be most abundant within the terminal ends of the creeks. Therefore, this location may best characterize the level of terrapin abundance within the creek for the monitoring objectives of GA DNR.

The limited amount of data collected by boat in this study restricted our ability to estimate optimal survey timing for monitoring terrapins in their foraging habitat. However, based on our knowledge of terrapin nesting phenology, conducting surveys within a short timeframe early in the season when females are most likely to be found within the medium and smaller tidal creeks potentially farther from nesting areas may be ideal. Future work to inform optimal monitoring should use a larger sample size of boat surveys to estimate optimal survey timing and assess the number of annual surveys and the number of sample locations along a creek needed to reduce the risk of mis-categorizing a site to an acceptable level.

Finally, integrating head count data with capture-recapture data should be explored as a means of improving estimates. Integrated models can improve accuracy and precision of estimates by jointly estimating parameters shared by more than one dataset (Zipkin and Saunders 2018, Frost et al. 2023). For example, in a study on western bluebirds (*Sialia mexicana*), Sanderlin et al. (2019) found that integrating presence-absence data collected at 149 sites with capture-recapture data collected at a subset of the same sites (n = 72) improved estimate precision, relative to modeling either of the datasets separately. A similar approach could be evaluated for diamond-backed terrapins in Georgia to estimate precision gains expected with integrating capture-recapture data collected at a subset of head count sites.

Conducting surveys by land accessible points may still be useful under some circumstances if there is an appropriate monitoring objective. This method is easy to learn and could harness volunteer effort to conduct many surveys over a short period, similar to programs implemented for other taxa. For birds, the Christmas Bird Count and Breeding Bird Survey have demonstrated how large amounts of volunteer-collected count data, while imperfect, can be used to monitor large-scale population trends and address other ecological questions (Bock and Root 1981, Hagan 1993, Sauer et al. 2003, Silvertown 2009). Volunteer-collected counts can also be combined with other data types via integrated models to improve reliability and inferential capacity (Hanks et al. 2011, Steger et al. 2017, Robinson et al. 2018, Sun et al. 2019).

Diamond-backed terrapins are challenging to monitor because of their difficult-to-access habitat and daily and seasonal variation in habitat use. On a daily basis, tide cycles affect whether terrapins are likely to be found in creeks or the high marsh, which is virtually inaccessible for monitoring. Throughout the active season, nesting migrations affect the location of mature females. Despite these challenges, monitoring is necessary to understand the status of populations and effectively apply targeted management. Capture-recapture studies of terrapins within their aquatic habitat are not feasible to conduct at broad spatial scales long-term. Therefore, if used judiciously, a method using head count surveys and N-mixture models may be the most effective means of monitoring state-wide.

This study can serve as a starting point for refining this method to meet the needs of the Georgia DNR. Because terrapin distributions within creeks tend to be clustered, we recommend surveying non-randomly within each creek to monitor areas that best characterize abundance. Conducting surveys by boat is likely necessary to target such locations. We also recommend conducting surveys over short and consistent time periods within the active season. Pre-nesting season is likely ideal, but this should be evaluated with more boat survey data. Other states with similar salt marsh tidal creek habitat may benefit from using this same sampling approach if objectives are the same. However, if habitat, sampling conditions, or monitoring objectives differ, we recommend re-evaluating the methodology under the new context.

TABLES

Table 3.1. Abundance categories (< 1, 1 – 10, and/or > 10) occupied by each survey area among all visits within the same year (for survey areas visited ≥ 2 times). Sample sizes (n) indicate the number survey areas with the respective number of visits. Data were collected 4 April – 13 July 2021 and 21 April – 13 July 2022 within 119 total survey areas spanning the Georgia coast.

			Relative abundance categories occupied					
	Visits	**n**	**< 1 only**	**1 – 10 only**	**> 10 only**	**< 1 & 1 – 10**	**1 – 10 & > 10**	**< 1, 1 – 10, & > 10**
2021	**All**	91	**55%**	**8%**	**1%**	**31%**	**4%**	**1%**
	2	22	36%	23%	5%	27%	9%	0%
	3	16	69%	0%	0%	31%	0%	0%
	4	52	58%	4%	0%	33%	4%	2%
	5	1	100%	0%	0%	0%	0%	0%
2022	**All**	73	**62%**	**7%**	**0%**	**26%**	**5%**	**0%**
	2	2	100%	0%	0%	0%	0%	0%
	3	44	89%	0%	0%	11%	0%	0%
	4	1	0%	0%	0%	100%	0%	0%
	5	26	15%	19%	0%	50%	15%	0%
Total both years		**164**	**58%**	**7%**	**1%**	**29%**	**5%**	**1%**

Table 3.2. Abundance categories (< 1, 1 – 10, and/or > 10) occupied by each tidal creek site across all survey areas within the same site sampled on the same date (for sites with ≥ 2 survey areas). Sample sizes (n) indicate the number sites with the respective number of survey areas. Data were collected 4 April – 13 July 2021 and 21 April – 13 July 2022 within 119 total survey areas spanning the Georgia coast.

		Relative abundance categories occupied						
Number of survey areas	**n**	**< 1 only**	**1 – 10 only**	**> 10 only**	**< 1 & 1 – 10**	**< 1 & >10**	**< 1 & 1 – 10**	**< 1, 1 – 10, & > 10**
2	125	57%	25%	1%	16%	0%	2%	NA
3	25	48%	0%	0%	44%	0%	8%	0%
4	16	44%	6%	0%	38%	6%	0%	6%
5	3	0%	0%	0%	0%	0%	0%	100%
6	2	0%	0%	0%	50%	0%	50%	0%
All	**171**	**53%**	**19%**	**1%**	**22%**	**1%**	**3%**	**2%**

Table 3.3. Percentage of survey areas assigned to each relative abundance category based on amount of sampling effort. Sample sizes (n) indicate the number survey areas sampled each year. Data were collected 4 April – 13 July 2021 and 21 April – 13 July 2022 within 119 total survey areas spanning the Georgia coast.

		Abundance category		
	Survey effort	**< 1**	**1 – 10**	**>10**
2021 (n = 112)				
	10 scans	74%	21%	4%
	5 scans	81%	16%	3%
	3 scans	84%	13%	3%
2022 (n = 83)				
	10 scans	69%	27%	5%
	5 scans	71%	27%	2%
	3 scans	78%	22%	0%
Total both years				
	10 scans	72%	24%	5%
	5 scans	77%	21%	3%
	3 scans	82%	17%	2%

FIGURES

Figure 3.1. Locations of diamond-backed terrapin count survey areas (orange points) conducted during 2021 – 2022 in Georgia, USA. Orange line delineates the Georgia state boundary.

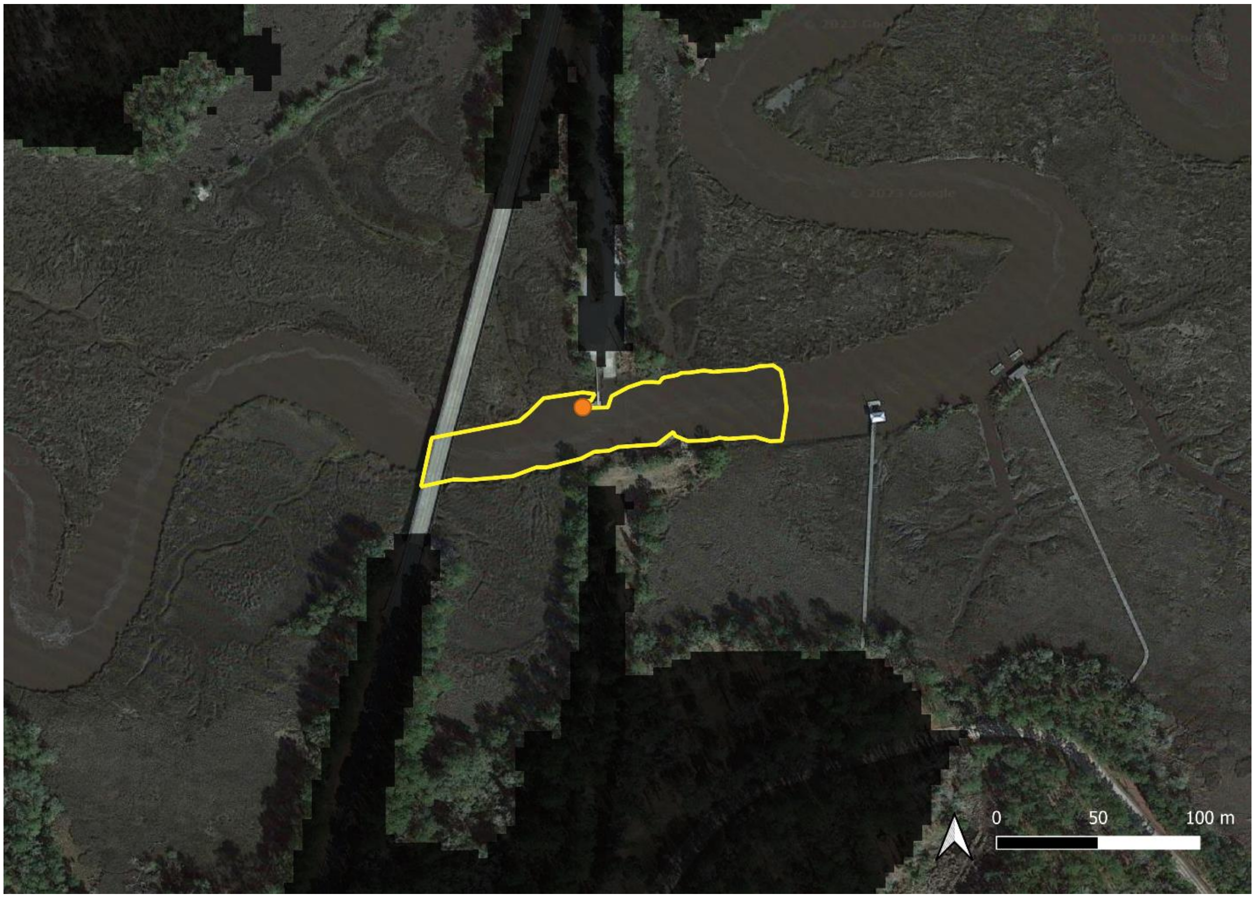


Figure 3.2. Survey area of 6,089 m^2 (i.e., approximately average sized survey area) located in Townsend, Georgia, USA. Yellow line delineates the survey area and orange point indicates location of observer during surveys. Dark shaded area indicates land ≥ 1.5 m elevation, which was considered to be potential nesting habitat.

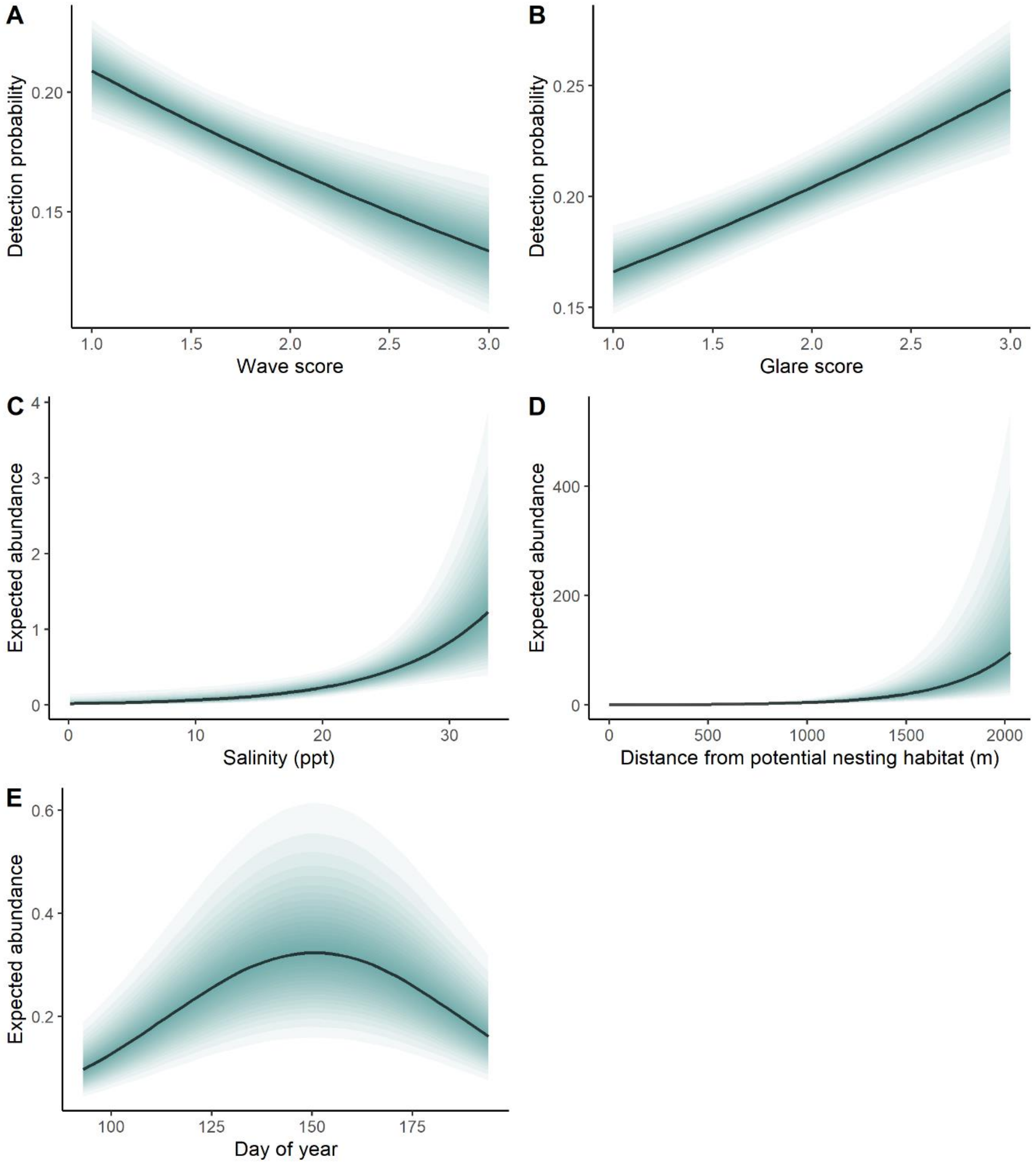


Figure 3.3. Estimated effects of binomial N-mixture model covariates selected using indicator variable selection. Effects were estimated using the top-ranked model, which included fixed effects of perceived effects of **(A)** waves and **(B)** glare (wave score and glare score, respectively) to explain detection probability (p) and a random site effect and fixed effects of survey area size, **(C)** salinity (ppt), **(D)** distance to potential nesting habitat (m), and **(E)** day of year (main effect and quadratic terms) to explain expected abundance (λ). Posterior medians are represented by dark green lines and Bayesian credible intervals increasing at 5% intervals are represented by the multilayered ribbon surrounding the median. Data modeled included 606 surveys of 119 survey areas spanning the Georgia coast during 2021 and 2022.

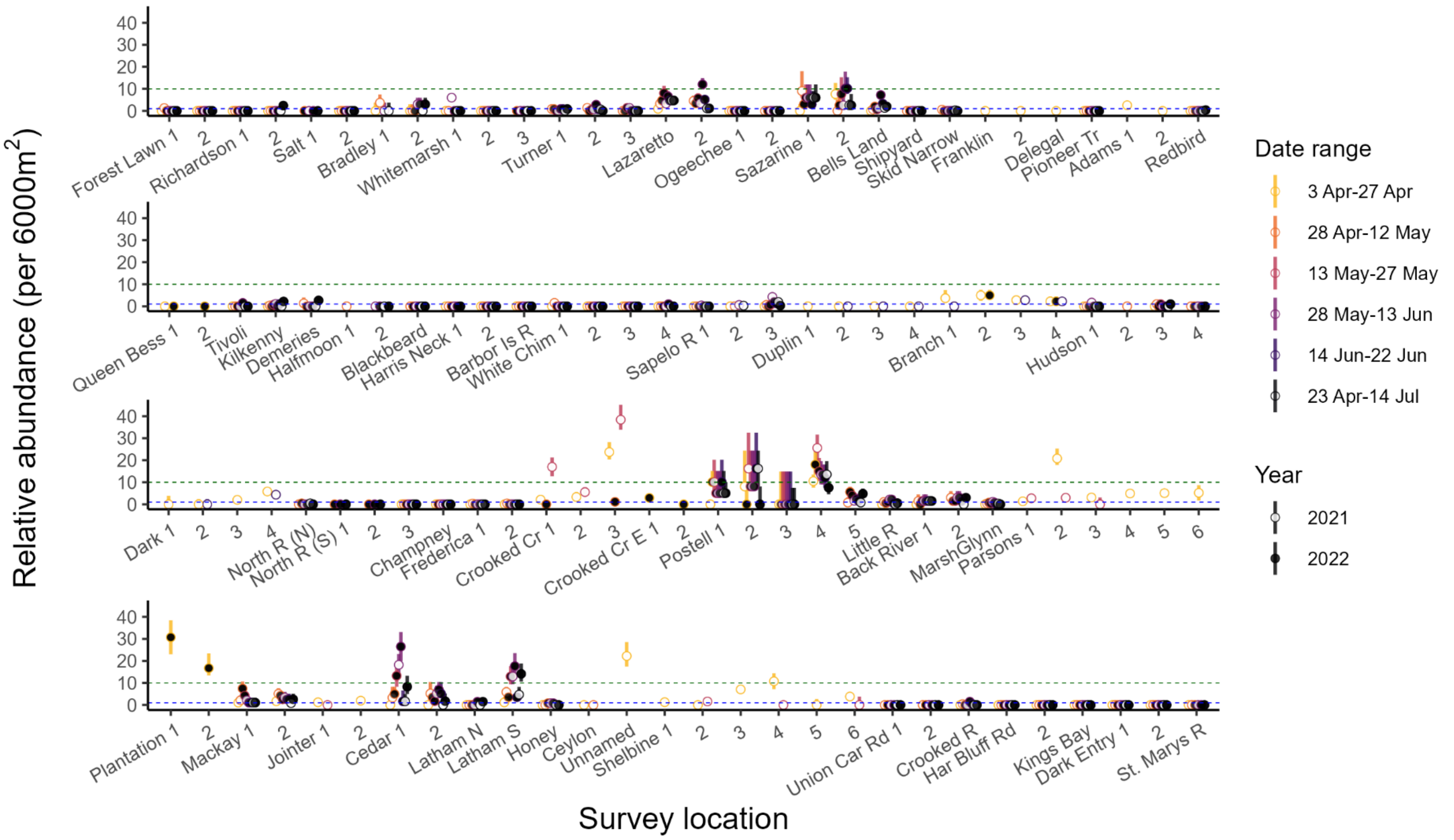


Figure 3.4. Diamond-backed terrapin relative abundance estimates for 606 surveys of 119 survey areas representing 58 tidal creek sites in Georgia, USA sampled during 4 April – 13 July 2021 and 21 April – 13 July 2022. Points represent median estimates and error bars represent 95% Bayesian credible intervals. Timing of surveys is identified by color and point fill. Each tidal creek site is labeled by name with numbers indicating multiple survey areas at the same site. Blue dashed line represents the relative abundance threshold of ≥ 1 terrapin per 6,000 m^2 and the green dashed line represents the relative abundance threshold of > 10 terrapins per 6,000 m^2.

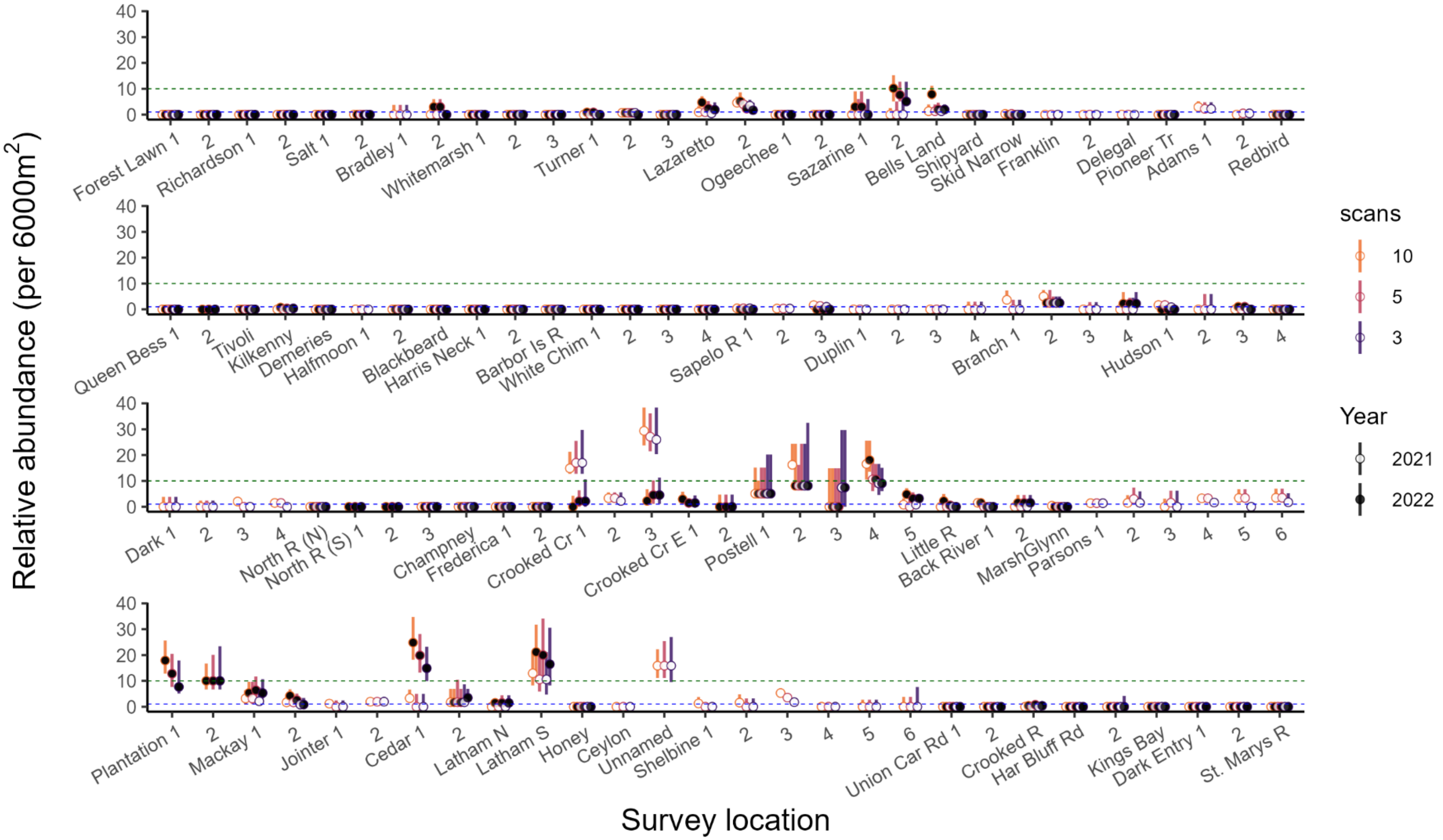


Figure 3.5. Diamond-backed terrapin relative abundance estimates modeled using three different levels of survey effort (10 scans, 5 scans, or 3 scans per survey) for 195 surveys of 119 survey areas within 58 tidal creek sites in Georgia, USA. The full dataset was randomly subsampled to include one survey per year per study area (n = 112 survey areas sampled in 2021 and n = 83 survey areas sampled in 2022). Points represent median estimates and error bars represent 95% Bayesian credible intervals. The number of scans is identified by color and survey year is identified by point fill. Each tidal creek site is labeled by name with numbers indicating multiple survey areas at the same site. Blue dashed line represents the relative abundance threshold of ≥ 1 terrapin per 6,000 m^2 and the green dashed line represents the relative abundance threshold of > 10 terrapins per 6,000 m^2.

CHAPTER 4

LEVERAGING COMMUNITY SCIENCE TO IMPROVE ESTIMATES OF NESTING FREQUENCY AND POPULATION PROJECTIONS FOR DIAMOND-BACKED TERRAPINS [3]

---

[3] Bradke, D.R., D.C. Scott, C.T. Moore, P. Miller, C. McInerney, J. Olsen, L. Isenhour, D. Cordo, P. Field, J. Howard, P. Howard, H. Lieberman, L. Munro, T. Ramsden, M. Schwickrath, and J.C. Maerz. To be submitted to *Chelonian Conservation and Biology*.

ABSTRACT

Informative projections from population viability analyses (PVAs) depend on having accurate estimates of life history traits and vital rates. However, because of data limitations, it is often necessary to incorporate surrogate values from studies and populations that may be uncharacteristic of the focal population or based on limited data or expert opinion. Because predictions made with PVAs are often used to make management decisions, it can be important to update initial viability models with more appropriate parameters when available. Diamond-backed terrapins (*Malaclemys terrapin*) have a large range, resulting in latitudinal variability in life history traits among populations. While reproduction is one of the most well-studied aspects of this species, data and estimates on nesting frequency and clutch sizes are lacking, especially in Georgia where there are no published studies. A previously published PVA for terrapins in Georgia used clutch size and nesting frequency values from studies in South Carolina and Maryland, which may be biased or underestimate fecundity of Georgia terrapins. In this study we leveraged an extensive five-year capture-recapture dataset of nesting terrapins collected as part of the Skidaway Audubon Diamondback Terrapin Rescue Project to obtain estimates of nesting frequency and values of clutch size for Georgia terrapins. We then updated the estimates from the previously published PVA of different potential management strategies. To estimate true nesting frequency, we developed a 3-level hierarchical model to account for imperfect detection. We estimated the mean annual number of nests per female to be 17% higher than previously assumed (mean = 2.33, 95% Bayesian credible interval = 2.17 – 2.70). Importantly, at least 34% of nesting females nested three or more times within a season with one individual confirmed to have nested six times in a single season. Such variation and high rates of individual nesting have not been reported for wild terrapin populations previously. Mean clutch size (8.55, range = 1 –

16) was 24% higher than assumed in the prior PVA. Both population growth and persistence projections increased under the updated population viability model, with population persistence increasing by 19% – 111% across nine potential management scenarios. These more optimistic results of current and potential management actions can aid decisions about resource allocation for terrapins or other conservation projects. Additionally, the Skidaway Audubon Diamondback Terrapin Rescue Project can serve as a model for integrating successful community-based conservation with rigorous research.

INTRODUCTION

Life-history traits are directly linked to vital rates which drive population growth and persistence (Williams et al. 2002). For example, reproductive characteristics, including clutch size and reproductive frequency, determine recruitment rates of new individuals into a population. (Cole 1954, Mills and Lindberg 2002). Population viability analysis (PVA) uses life-history traits and vital rates to predict population growth rates and assess the likelihood of population persistence over time (Beissinger and McCullough 2002). Projections from PVAs are often relied upon to allocate limited resources among competing management actions (Boyce 1992, Lindenmayer et al. 1993, Beissinger and Westphal 1998, Manlik et al. 2018). Useful projections of population growth and persistence depend on having accurate estimates of vital rates. Inaccurate estimates of key life history traits and vital rates that result in overestimation of population viability can lead to inaction that results in adverse outcomes for priority species. Conversely, underestimation of population viability can result in the allocation of additional resources to unnecessary actions instead of allocating those resources toward other priorities. Life history and vital rate parameters used in a PVA often come from multiple sources, including direct measurement and estimation of the focal population, transferred values from published studies of other populations or closely-related species, or expert opinion (Heppell et al. 2000, Etterson and Bennett 2006, McGowan et al. 2011). It is rare that PVAs can be parameterized entirely using data from the focal population due to the time and effort required to measure or estimate these parameters (Beissinger 2002, Etterson and Bennett 2006). Some parameters may be reasonably transferred from other populations or closely related species (Heppell 1998, Heppell et al. 2000, Schtickzelle et al. 2005); however, viability models may be sensitive to and biased by transferred parameters that vary temporally or spatially (Lindenmayer et al. 1993, Gurevitch et al. 2016). Therefore, it can be important to update initial viability models with improved estimates of parameters when available.

Community science or citizen science initiatives are one possible solution to overcoming logistical and economic limitations of data collection (Dickinson et al. 2010, Theobald et al. 2015). For example, citizen science was used to collect resight data on bearded vultures (*Gypaetus barbatus*) to estimate survival rates needed for population viability analysis of this imperiled species (Schaub et al. 2009). Likewise, long-term data collected by citizen scientists has been used to estimate viability of monarch butterflies (*Danaus plexippus plexippus*) via count-based PVA (Schultz et al 2017). Harnessing the capacity of community-based monitoring programs can generate data that are difficult for researchers and management organizations to achieve with limited resources.

Diamond-backed terrapins (*Malaclemys terrapin*) inhabit estuarine ecosystems along the Atlantic and Gulf coasts of the U.S., resulting in a broad latitudinal and longitudinal range and highly variable life history traits (Buhlmann et al. 2009, Lovich and Hart 2019, Lovich et al. 2019). They are listed as a high priority species or Species of Greatest Conservation Need in all states within their range. Threats to terrapin populations include bycatch in commercial and recreational crab pots, adult female mortality on roads while attempting to nest, high rates of nest depredation from mesopredators, shore armoring, and agricultural, residential, and commercial land use adjacent to salt marsh habitats (Isdell et al. 2015, Chambers and Maerz 2019, Maerz et al. 2019). Diamond-backed terrapins, like other Chelonians, exhibit characteristics of a 'slow life-history' with delayed sexual maturity and long lifespans. Accordingly, conservation focused on mitigating threats to adult female survival (e.g., road mortality) may be most effective for improving population viability (Mitro 2003, Crawford et al. 2018). However, because threats to nests (e.g., mesopredators) are extremely high in some cases (Roosenburg 1991, Butler et al. 2004) and because it may be easier to target nest survival with management (e.g., via predator removal or nest protection), conservation efforts that focus on this stage may also be able to increase the probability of population persistence (Feinberg and Burke 2003, Crawford et al. 2014*a*, Maerz et al. 2019).

Although diamond-backed terrapins exhibit nest site fidelity (Roosenburg 1991, Sheridan et al. 2010), nesting can be difficult to observe because of timing (e.g., nests laid outside of typical work hours) and duration of nesting events, which can be as brief as 15 minutes (Roosenburg 1994). Vegetation and nest location (e.g., private property) can also contribute to low individual detection rates of nesting females. Additionally, nesting seasons can extend nearly four months in the southern portion of the species' range (Butler and Burke 2019), making it difficult to devote sufficient time and labor to monitor the entire season. As a result, there are few published studies of terrapin nesting frequency, and those available do not account for imperfect detection of nesting females. Clutch size data are also lacking for many populations. Although relatively easier to obtain, collecting these data typically requires excavating nests, which may be viewed as invasive. Alternatively, x-ray can be used (e.g., Kern et al. 2016, Donini and Selman 2022), but requires access to machinery. Currently, there are no published studies that include nesting frequency and clutch size information for terrapins in Georgia, necessitating that these values are borrowed from populations elsewhere or assumed for PVAs.

Between 2009 – 2015, a study was conducted on and around the Jekyll Island Causeway in Georgia, to estimate population vital rates for diamond-backed terrapins, assess the efficacy of management conducted during the study, and use a PVA to project the efficacy of different future management scenarios that could be adopted (Crawford et al. 2018). Despite high effectiveness of flashing signage which alerted drivers to nesting females crossing the road and increased survival of crossing adult females from 23.5% to 53.2%, the PVA predicted a population growth rate (λ) < 1 under current management (i.e., flashing signage and a 22-m nest box intended to block females from crossing and provide a safe place to nest). More aggressive management scenarios to increase nest success and proportion of female hatchlings (i.e., adding additional nest boxes/fences, predator removal, and vegetation management) were also considered, but still yielded predictions of $\lambda < 1$. While this study used a rigorous approach to estimate many of the focal population survival parameters included in the PVA, nesting frequency (2 nests/year; Roosenburg 1991) and clutch size (6.9 eggs per clutch; Zimmerman 1992), were based on limited information from other studies and populations.

The Skidaway Audubon Diamondback Terrapin Rescue Project is a community science and conservation project established to protect nests laid by terrapins at the Terrapin Point Golf Course on Skidaway Island, Georgia. Because terrapins often nest in and around sand traps, nests were susceptible to destruction by golfers, maintenance machinery, and predators. To alleviate these threats, Diamondback Terrapin Rescue Project volunteers monitor the course and move nests to protected nest boxes where they are hatched. High nest-site fidelity of terrapins to the sand traps coupled with high volunteer survey effort appears to have led to higher-than-normal individual detection rates and observations of as many as 6 nests per year by a single female. These observations suggested that the nesting frequency assumed in diamond-backed terrapin PVAs, including Crawford et al. (2018), may be low and result in estimates of population growth and persistence that are biased low. Thus, the objectives of this study were to estimate nesting frequency for terrapins on Skidaway Island, GA using this unique, high-effort, high-detection nesting dataset and assess how updated nesting frequency and clutch size parameters change predictions of diamond-backed terrapin population viability. Specifically, we wanted to know whether and how much these new parameter estimates changed population growth rates and persistence probabilities projected under the management scenarios considered by Crawford et al. (2018).

METHODS

Skidaway Audubon Diamondback Terrapin Rescue Project data collection

We collected nesting data on mature female diamond-backed terrapins at the Terrapin Point Golf Course on Skidaway Island, Georgia during April through August of 2018 – 2022. Trained volunteers encountered terrapins by patrolling five golf holes twice a day around the diurnal high tides – when most nesting activity occurs (Crawford et al. 2014*b*). Signage and information in golf carts also prompted phone calls to report terrapin encounters by golfers, residents, or maintenance staff. Volunteers watched each terrapin until the terrapin completed nesting, recorded the date, time, and location, and excavated and moved the eggs to a nest box for incubation. Excavating nests allowed volunteers to determine the number of eggs laid. Volunteers took photos of the carapace and plastron of each terrapin observed, which were used to individually identify terrapins and produce individual nesting histories. In 2018, this identification method was successfully piloted at one golf hole, and from 2019 – 2022 it was adopted across the entire survey area. We included 2018 data only in calculating clutch sizes but not for our model of individual nesting frequency.

Occasionally volunteers encountered adult female terrapins on land but did not observe them nesting (263 of 1,574 adult female observations between 2019 – 2022). When this occurred, volunteers still recorded the date and time of the observation and photographed the female. It is likely that these female terrapins nested either before or after they were encountered. Therefore, we assumed that any female encountered on land nested that day, unless the same female was encountered again within a week and no clutch was detected during the first encounter. If we encountered the same female a second time within a week and we did not observe her laying a nest during the first encounter, we assumed that the first encounter was a "false crawl" or abandoned nesting attempt (Szerlag-Egger and McRobert 2007). If we encountered the same female a second time within a week and we observed her laying a nest during the first encounter and not the second encounter (this only occurred two times), we treated the second observation as a false crawl. Any false crawl was not included in our analysis of nesting frequency. In only five instances did we observe a female laying eggs twice within one week. In four of these instances the female nested twice within two days and in one instance the female nested twice within three days. In these five cases that did not appear to be true inter-nesting intervals, but rather two, partial clutches, we treated each as a single clutch and used the first date of the two observations as the nesting date.

Nesting Frequency Analysis

While naïve nesting frequencies at Terrapin Point Golf Course suggest relatively high individual encounter rates of nesting terrapins, we know that detection is still imperfect at this site. Some nests are detected during additional patrols each morning and evening, which are conducted specifically to locate nests laid outside of the high tide patrol window. These clutches are recorded and moved to nest boxes in the same manor described above, but they cannot be associated with the female that laid them. An unknown number of additional nests likely go undetected, especially clutches that are laid in areas where females or signs of nesting are more difficult to observe (e.g., adjacent private yards or other vegetated areas).

To estimate true nesting frequency, we developed a three-level hierarchical model under a Bayesian framework. This model accounts for imperfect detection of nesting events and reflects our understanding of the nesting process in this species. The three levels include two state models specifying an individual-specific nesting process and an observation model specifying the detection process. The first state model assumes that adult female terrapins leave the pool of nesting individuals (or the nesting population) and cease nesting activities at variable dates during a given nesting season. This individual variation in nesting season length likely occurs as females exhaust energy stores at different times throughout the summer. The second state model assumes individual variation in nesting season start date. Additionally, it assumes a cyclical nesting probability with variable average inter-nesting intervals (i.e., days between nesting events) among individuals. We fit this 3-level hierarchical model to individual daily capture histories. These capture histories were formatted as binary data (i.e., detected “1” vs undetected “0”) – identical to those used in traditional capture-recapture models – with each individual-year combination forming a separate record in the capture history.

The detection process, which indicates if individual (i) was observed on day (j), was specified as a Bernoulli outcome with probability equal to the product of capture probability (p_i) and current state ($z_{i,j}$) as nesting (1) or not nesting (0; Eq. 1). We constrained p_i so that $p_i = \overline{p}$

(i.e., we assumed all individuals had the same capture probability). The nesting process, which indicates if individual (i) nested on day (j), was specified as a Bernoulli outcome with probability equal to the product of nesting probability ($\gamma_{i,j}$) and current state ($s_{i,j}$) as remaining in (1) or not remaining in (0) the nesting population (Eq. 2). This state of remaining in or not remaining in the nesting population was also specified as a Bernoulli outcome with a probability equal to the product of the individual's probability of remaining in the nesting population during the previous sampling interval ($\varphi_{i,j-1}$) and the individual's state at the previous sampling occasion ($s_{i,j-1}$; Eq. 3). We constrained $\varphi_{i,j}$ so that $\varphi_{i,j} = \overline{\varphi}$ (i.e., we assumed a constant probability of remaining in the nesting population). This final level of the model is analogous to a Cormack-Jolly-Seber (CJS) model and therefore conditioned on first capture. To allow z states (as nesting or not nesting) to be estimated prior to first capture, we supplied the model with known s states of 1 prior to first capture, thereby assuming all individuals are part of the nesting population at the start of the capture history.

$$y_{i,j} \sim Bernoulli(\mathrm{p}_i \times z_{i,j}) \quad \text{(Eq. 1)}$$

$$z_{i,j} \sim Bernoulli(\gamma_{i,j} \times s_{i,j}) \quad \text{(Eq. 2)}$$

$$s_{i,j} \sim Bernoulli(\varphi_{i,j-1} \times s_{i,j-1}) \quad \text{(Eq. 3)}$$

We constrained nesting probability ($\gamma_{i,j}$) using a logit-sine function (Eq. 4). This function was based on a sinusoidal regression where individual and time-specific nesting probability is a function of a constant (C), amplitude (α), angular frequency (ω_i), time (T_j), and phase shift (Δ_i).

$$logit(\gamma_{i,j}) = C + \alpha \sin(\omega_i T_j + \Delta_i) \quad \text{(Eq. 4)}$$

We derived the angular frequency and phase shift parameters from individual nesting season start days (f_i) and individual average inter-nesting intervals (i_i; Eq. 5, 6).

$$\omega_i = \frac{2\pi}{i_i} \quad \text{(Eq. 5)}$$

$$\Delta_i = \frac{\pi}{2} - \left(\frac{f_i}{i_i}\right) \times 2\pi \qquad \text{(Eq. 6)}$$

Nesting season start days (f_i) and average inter-nesting intervals were specified as Poisson outcomes with means equal to a population mean start day and population mean inter-nesting interval (I), respectively (Eq. 7, 8).

$$f_i \sim Poisson(F) \qquad \text{(Eq. 7)}$$

$$i_i \sim Poisson(I) \qquad \text{(Eq. 8)}$$

We fixed nesting probability to 0 for all days prior to the individual's estimated start day (i.e., $j < f_i$). Additionally, because the shortest observed inter-nesting interval (i.e., period between two clutches laid by the same individual, excluding the five instances of clutches laid ≤ 3 days apart) was nine days, we fixed nesting probability to 0 if an individual's nesting state was equal to 1 any time during the previous eight days (i.e., $\sum_{j=1}^{8} \mathrm{z}_{i,j} > 0$). To derive estimates of total annual nests for each individual, we summed the z states for each row in the capture history (i.e., each individual-year combination). We report all estimates from the nesting frequency model based on the median values of posterior distributions.

We fit the model using Markov chain Monte Carlo (MCMC) sampling using the nimble package (de Valpine et al. 2017, 2022) in program R (4.2.1; R Core Team 2022). We used 3 MCMC chains with 500,000 iterations each and burn-in of 100,000 iterations. We assessed chain convergence visually using trace plots and with Gelman and Rubin (1992) diagnostic values (i.e., $\hat{R} < 1.1$). We evaluated goodness-of-fit using a posterior predictive check, in which we compared data predicted under the model to the capture-history dataset. Specifically, we compared the distribution of mean predicted detections of terrapins with the mean actual detections in the capture history (i.e., the count of all 1s divided by the count of all 1s and 0s in the capture history).

Model validation

We validated that all parameters were estimable in our model using simulation. To do this, we simulated capture-history data using our model parameterized with values that produced distributions of nesting events and inter-nesting intervals like our actual data. We then fit our model to the simulated data and evaluated whether posterior distributions (95% BCIs) of the parameters included the values used to simulate the data.

Population viability analysis

Using our estimates of nesting frequency and data on clutch sizes collected for the Skidaway Island Terrapin Point Golf Course population, we updated the previously published population viability analysis for the population of diamond-backed terrapins near the Jekyll Island Causeway in Georgia, USA (Crawford et al. 2018). Except for nesting frequency and clutch size, we used the same PVA model as Crawford et al. (2018) and examined the same nine scenarios, which included no management, status-quo management, and different combinations of three additional management actions. The three management actions were as follows: roadside barriers (i.e., nest boxes/fences) to prevent nesting females from crossing the road, predator removal to improve nest survival; and roadside vegetation management to improve nest survival and create favorable conditions for development of female hatchlings.

The PVA was a stage-based model that included six classes (hatchlings, juveniles, and adults for each sex) and incorporated variable states within the adult female class to allow population dynamics to differ across the nine management scenarios. Adult females were grouped into three possible states that affected their survival: crossing the road to nest, not crossing the road to nest, and not nesting on the road (i.e., not in the study area). Females in either of the two states on the road (crossing, not crossing) were further divided into three different habitats that occurred along the road and had different rates of nest survival.

Adult survival probabilities, entry probabilities, site fidelities, and road crossing probability used in the PVA were all estimated by Crawford et al. (2018) using an integrated model. Using the same methods as Crawford et al. (2018), we incorporated estimates of adult survival – including mean survival probabilities, standard deviations, road crossing effects, and effects of flashing signage to alert drivers (in applicable scenarios) – in the PVA by randomly drawing samples from the posterior distributions of these parameters for each model iteration. We incorporated the estimated probabilities of females nesting along the road and crossing the road in a similar manner as adult survival. Other parameters were included as constant values (see methods and Table 1 in Crawford et al. 2018 for additional details and a full list of PVA parameters, respectively).

To incorporate the Skidaway Island estimates of nesting frequency in the PVA, we randomly selected an MCMC iteration from our nesting frequency analysis for each iteration in the PVA. From these MCMC iterations, we drew estimates of individual total annual nests, with a sample size equal to the number of nesting females in the study area (specific to the management scenario, iteration, and year). These randomly drawn individual nests were summed to obtain the total number of nests laid along the road for the year. We also used the clutch size data from Skidaway Island in the PVA by sampling from all observed clutch sizes.

We ran each management scenario for 50 years and with 1,000 iterations and estimated persistence probability and mean population growth rate using the same methods as Crawford et al. (2018). This included implementing a quasi-extinction threshold of 50 individuals to determine persistence and calculating the mean population growth rate using the annual growth rates during the last 30 years. We also ran the Crawford et al. (2018) PVA using the same data and parameters originally used and compared estimates of population growth and persistence to estimates from our updated version.

RESULTS

From 2019 – 2022, the Skidaway Audubon Diamondback Terrapin Rescue Project had 1,574 captures of 293 adult females among the five golf holes patrolled at Terrapin Point Golf Course. After accounting for false nesting events, we had 1,528 nesting events that we could associate with the 293 individuals. The earliest date a terrapin or nest was encountered was 25 April (2020) and the latest date was 17 August (2020 and 2021). Observed annual nesting season lengths (i.e., from first observation date to last observation date) were 109 days in 2019, 114 days in 2020, 111 days in 2021, and 107 days in 2022. The mean annual number of nests observed per female was 2.15 (range = 1 – 6), with ≥ 3 annual nests per female in 244 of 711 total annual nesting histories (34.3%; Figure 4.1). The mean interval between observed nesting events was 23 days and the mode was 17 days (Range = 9 – 67 days; Figure 4.2). Volunteers encountered 345 additional nests within the survey area that could not be associated with the female that laid them. From 2018 – 2022, the mean number of eggs per clutch among all nests observed was 8.55 (range = 1 – 16; Figure 4.2).

Nesting Frequency Analysis

We estimated the mean annual number of nests per female to be 2.33 (95% Bayesian credible interval [BCI] = 2.17 – 2.70), with a range of 1 to 6 nests (Figure 4.1) indicating that accounting for detection probability increased the estimate of nest production by 8% over the naïve estimate. We estimated the mean interval between nesting events to be 19 days (95% BCI = 17.99 – 19.41), compared to the naïve estimate of 23 days. Capture probability was 0.91 (95%

BCI = 0.79 – 0.99) and the daily probability of remaining in the nesting population was 0.996 (0.995 – 0.997). Nesting probability peaked approximately 14 days after the earliest observed nest and followed a declining cyclical pattern (Figure 4.3). Our goodness-of-fit evaluation indicated that the model tended to overestimate the number of captures compared to our actual data (i.e., mean predicted detections = 0.020 [95% BCI = 0.019 – 0.022] and mean actual detections = 0.018). This indicates that either estimated detection probability or the estimated nesting events may be slightly overestimated.

Model validation

Based on simulation, all parameters were estimable in our model. Posterior distributions (95% BCIs) of all parameters estimated included the values used to simulate the data.

Population viability analysis

Under the baseline scenario, which did not include any management for diamond-backed terrapins, we predicted that the population would decrease over the next 50 years and estimated a 14% probability that the population would persist over the same period (Figure 4.4). Scenarios with management actions were all more optimistic, with mean projected population growth rates > 0.9 and upper 95% Bayesian credible intervals > 1.0 (Figure 4.4). Persistence probability estimates for these eight scenarios with management ranged from 0.74 to 0.91. Population growth rates and persistence probabilities followed similar patterns across management scenarios to those projected by Crawford et al. (2018), but both metrics were consistently higher under our updated PVA, with increases in persistence probabilities greater than increases in population growth. Across the eight scenarios with management actions, the original PVA predicted population growth rates ranging from 0.90 – 0.95, while the updated PVA predicted growth rates ranging from 0.92 – 0.98. For these same eight scenarios, the original PVA yielded persistence probabilities ranging from 0.54 – 0.76 , while the updated PVA yielded persistence probabilities ranging from 0.74 – 0.91 (Figure 4.4). Additionally, under the no management scenario, the growth rate increased from 0.81 to 0.84 and persistence probability increased from 0.07 to 0.14 (Figure 4.4).

DISCUSSION

Achieving high detection of nesting turtles is challenging, making it difficult to accurately account for reproductive output in population viability analyses. This study demonstrates the capacity for a community-based conservation project to contribute new knowledge about important life history traits in a Species of Greatest Conservation Need. We reported higher nesting frequencies than previously described for diamond-backed terrapins anywhere in their range. Estimated annual number of nests per female were higher than the grand mean of 1.5 clutches per year (population means ranging from 1 – 2 clutches per year) across 11 populations range-wide reviewed by Lovich et al. (2019). We estimated that 34.3% (95% BCI = 34.3% – 90.7%) of terrapins nested ≥ 3 times per year and that 14.6 % (95% BCI = 14.6% – 51.1%) nested ≥ 4 times per year, with one individual confirmed to have nested six times in a single year. The highest nesting frequency previously reported in wild diamond-backed terrapins was three clutches per year (Roosenburg 1991, Roosenburg and Dunham 1997); however, among captive terrapins reared at an experimental farming facility in North Carolina, female terrapins were reported to lay as many as five clutches per year (Hildebrand and Hatsel 1926).

The Diamondback Terrapin Rescue Project also obtained a large sample of clutch sizes (n = 1,934) during their efforts to protect nests, which requires that eggs be excavated and moved to nest boxes. We observed larger clutch sizes (mean = 8.55, range = 1 – 16) than previously reported in southern populations. Among 14 datasets from wild terrapin populations within the southern states of South Carolina, Louisiana, Florida, and Texas, clutch sizes ranged from 4 to 7.2 eggs per clutch (Lovich et al. 2019, Donini and Selman 2022).

The other main objective of this study was to improve projections of population growth and persistence for a managed population of diamond-backed terrapins near Jekyll Island, Georgia. The Terrapin Point Golf Course on Skidaway Island, also in Georgia, is closer to the Crawford et al. (2018) PVA focal population than the studies that were previously referenced to obtain nesting frequency (Maryland; Roosenburg 1991) and clutch size (South Carolina; Zimmerman 1992). Additionally, clutch size data used in this study are based off a much larger sample size (n = 1,934 nests) compared to Zimmerman (1992; n = 25 nests) and the individual nesting frequencies used represent estimates of true nesting frequencies versus the naïve frequency reported by Roosenburg (1991). Both population growth and persistence estimates increased under the updated model, with changes in projected persistence being most notable. For scenarios with management actions, most persistence projections under the updated model were > 80%, whereas all projections were previously < 80% (Figure 4.4).

This new information can help managers make decisions about resource allocation. For example, support for devoting resources to continuing current management or instituting more intensive management actions may increase because of increased confidence that management will improve population viability. Conversely, the increase in projected population viability under the status quo management scenario, may give managers enough confidence in their current strategy to refrain from allocating more resources to more intensive management of this population. Instead they may choose to focus their additional resources on other terrapin populations or other conservation problems. Managers will likely need to incorporate their budget and risk tolerance with updated estimates of population growth and persistence to choose among management scenarios.

Our model of nesting frequency is a unique approach to estimate the number of true nesting events among individuals in a population that could be applied to other turtle or wildlife populations. The sinusoidal regression component is based on our observation that individual terrapins appear to have relatively regular cycles of nesting events during the nesting season. For example, we often observed the same individual nesting multiple times with approximately the same inter-nesting interval. Longer inter-nesting intervals were sometimes noted to be approximately twice as long as the regular, shorter interval length observed for that individual. Therefore we suspected that some of these longer intervals included undetected nesting events at their midpoint. We also built into our model a process to account for some terrapins exhausting resources and leaving the nesting population earlier than others.

While using nesting events and clutch size from the Skidaway Island population as surrogate estimates for the Jekyll Island population PVA is likely an improvement, there are factors that could cause differences in these reproductive characteristics between these two populations. Because the Skidaway Island population does not have high road mortality of nesting females like the Jekyll Island population (Crawford et al. 2014*b*, *a*), higher survival may lead to more larger females in the population. Larger females are known to be more fecund than smaller females (Montevecchi and Burger 1975, Seigel 1980, Roosenburg and Dunham 1997, Lovich et al. 2019, Donini and Selman 2022); thus, clutch sizes and number of clutches per female may be lower in the Jekyll Island population. However, these differences are small in comparison to differences between populations at the northern versus southern portions of the range. The relationships between nesting frequency versus plastron length and between clutch size versus plastron length in the Skidaway Island population were highly variable and plastron length only explained a small amount of the variation in each of these measures of fecundity (See Figure C.1 in Appendix C vs. Figure 6.2 in Lovich et al. 2019). Additionally, there is a greater body size difference with a larger change in latitude and females of the same size produce larger clutches in the north than in the south (i.e., latitude better explains variation in clutch size than body size in this species; Lovich et al. 2019).

Community-based and citizen science initiatives are a valuable way to collect large amounts of data for monitoring and addressing other ecological questions (Silvertown 2009, Sullivan et al. 2009, McKinley et al. 2017). Moreover, non-scientist involvement in research projects can improve public interest and support for conservation (Silvertown 2009, Shirk et al. 2012, McKinley et al. 2017). In addition to data collection and nest protection activities, the Skidaway Audubon Diamondback Terrapin Rescue Project also engages members of the broader public in learning about diamond-backed terrapin conservation during multiple events each year where terrapin hatchlings from the protected nests are released into the marsh. Others conducting diamond-backed terrapin research have also demonstrated effective use of public education and outreach as an element of their studies (Heinrich et al. 2019). For example, conducting terrapin head-starting in K-12 classrooms in Maryland is used to simultaneously teach children about science and the environment, provide data to terrapin research, and cultivate public interest in conserving the species (Heinrich et al. 2019). Projects such as these can serve as models for how to successfully integrate research and conservation.

FIGURES

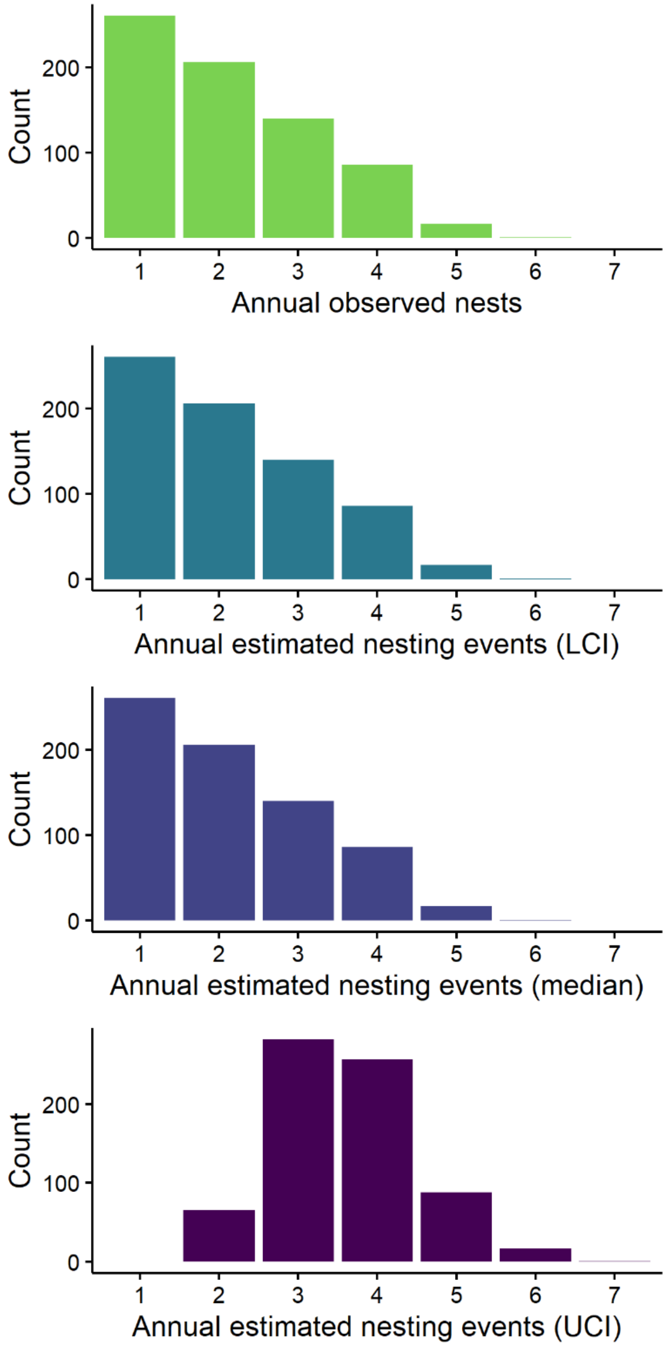


Figure 4.1. Distribution of number annual of nesting events for diamond-backed terrapins at Terrapin Point Golf Course on Skidaway Island, Georgia from 2019-2022. Top plot (green) shows the observed nesting events, and lower three plots show estimated true nesting events, including lower 95% Bayesian credible intervals (LCI; blue), median estimates (light purple), and upper 95% Bayesian credible intervals (UCI; dark purple).

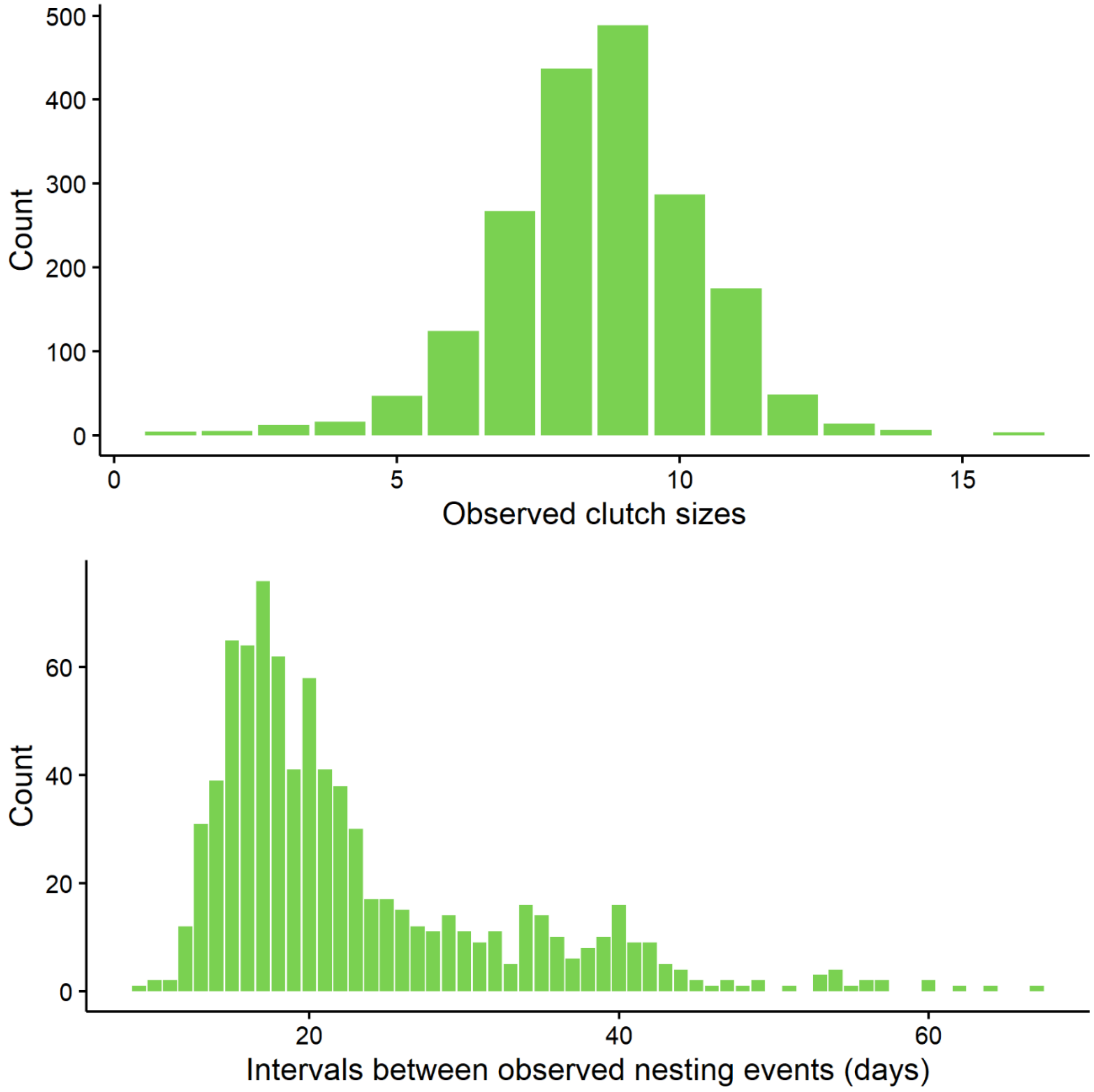


Figure 4.2. Distribution of clutch sizes (top) and intervals (number of days) between observed nesting events (bottom) for diamond-backed terrapins at Terrapin Point Golf Course on Skidaway Island, Georgia. Clutch size data collected from 2018-2022 and nesting interval data collected from 2019-2022.

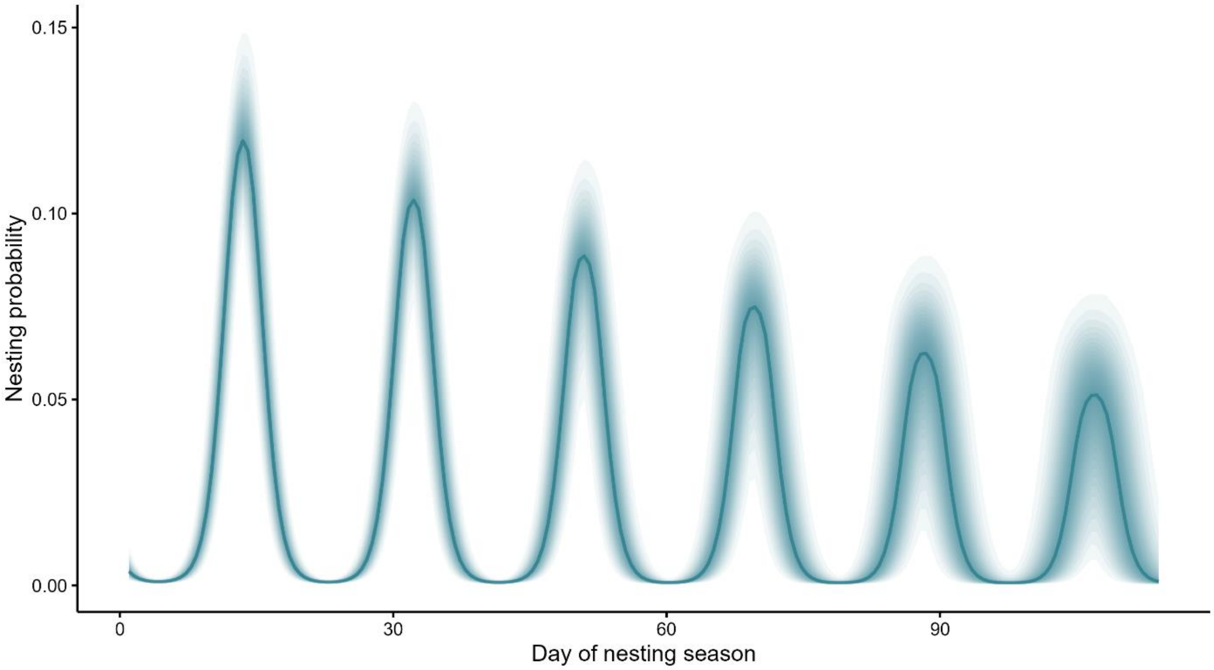


Figure 4.3. Average estimated nesting probability as a function of day of the nesting season. Mean individual nesting probability ($\gamma_{\bar{i}}$) on day (j) is a logit-sine function of a constant (C), amplitude (α), mean individual angular frequency ($\omega_{\bar{i}}$), time (T_j), and mean individual phase shift ($\Delta_{\bar{i}}$). This probability was multiplied by the mean individual daily probability of remaining in the nesting population ($\varphi_{\bar{i}}$) exponentiated to the number of days.

$logit(\gamma_{\bar{i},j}) = C + \alpha \sin(\omega_{\bar{i}} \omega T_j + \Delta_{\bar{i}}) \times \varphi_{\bar{i}}{}^{j}$

The blue line represents the posterior median and the multilayered ribbon surrounding the mean displays Bayesian credible intervals increasing at 5% intervals. Estimates are based on obervations of nesting diamond-backed terrapins at Terrapin Point Golf Course on Skidaway Island, Georgia from 2019-2022.

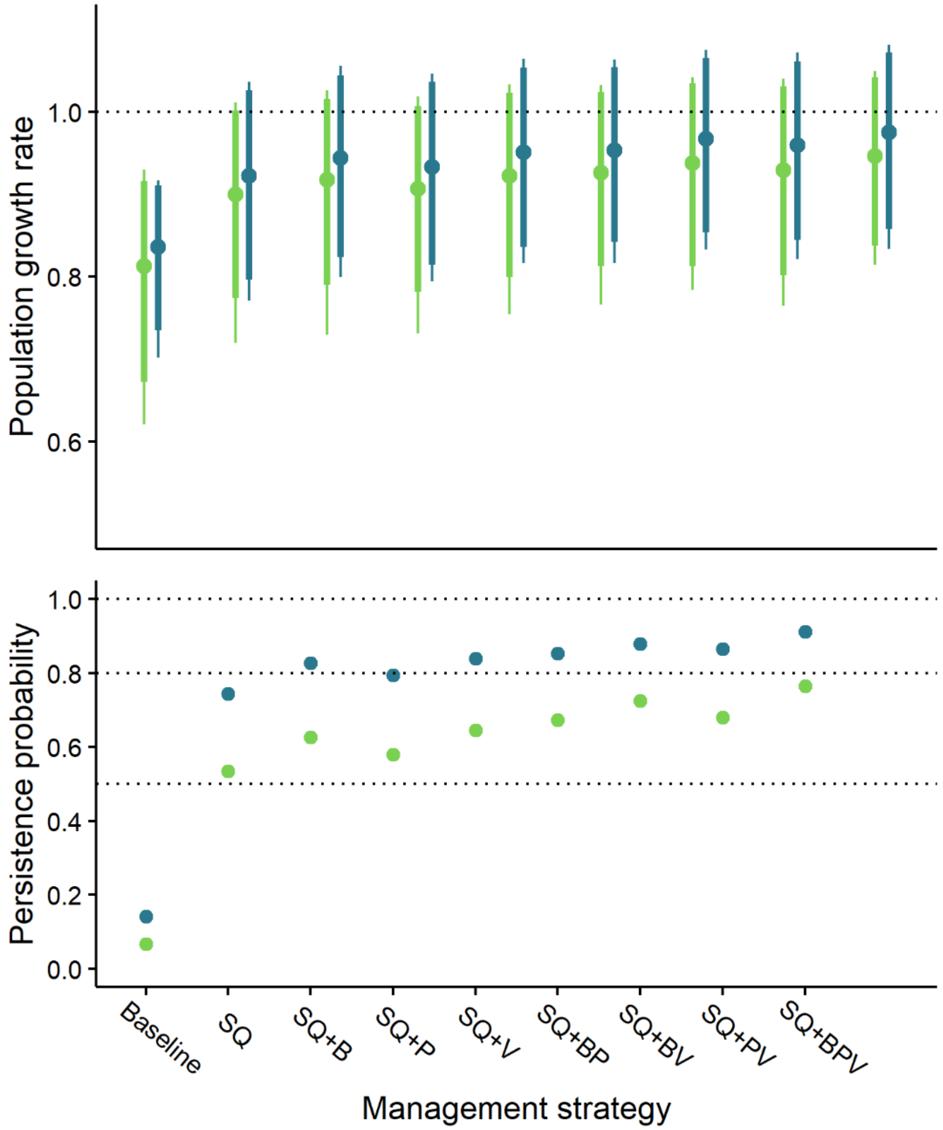


Figure 4.4. PVA metrics for a diamond-backed terrapin population on Jekyll Island, GA under baseline (no management), status quo (SQ), and seven additional management scenarios. The additional management scenarios include different combinations of 3 actions: barriers to prevent road crossing (B), predator management (P), and vegetation management (V). Projected population growth rates (top plot) and population persistence probabilities (bottom plot) are displayed for a previously published PVA (Crawford et al. 2018; green) and an updated version of the same PVA (blue), using estimates of nesting frequency and data on clutch sizes from Terrapin Point Golf Course on Skidaway Island, Georgia from 2019-2022. Thick error bars and thin error bars represent 90% and 95% Bayesian credible intervals and points represent mean estimates.

CHAPTER 5

AN ENCOURAGING PERSPECTIVE ON USING A HIERARCHICAL APPROACH TO IMPROVE STUDIES OF HERPETOFAUNA [4]

---

[4] Bradke, D.R., S.M. Bauer, K.M. Grab, and J.C. Maerz. To be submitted to *Herpetologica*.

ABSTRACT

Hierarchical sampling designs and modeling approaches have effectively raised the bar for studies of wildlife populations. Previous approaches that do not explicitly address the processes affecting animal observations are less defensible than they once were and can lead to poor inferences and ineffective conservation. Many wildlife ecologists – particularly those who study amphibians or reptiles – have been slow or resistant to adapt their approaches because of their concerns about the skillset needed to implement complex models or the comparability of past data to future data. We empathize with these concerns and present an encouraging perspective on using hierarchical designs and integrated models to improve studies of herpetofauna. To demonstrate the value of adopting a hierarchical sampling design, we provide a brief overview of hierarchical designs and models and describe how integrated models can be used to improve estimates from older sampling designs. Then we present a case study where we use these approaches to estimate forest management effects on terrestrial salamander abundance. We show that our original sampling methods, which did not effectively address the processes involved in observing these animals, resulted in low-biased abundance estimates and a reduced ability to detect effects of management. By updating our sampling methodology to accommodate a more realistic hierarchical model, we improved estimates of abundance and detected a clear, negative effect of forest management on salamander abundance with high confidence. We also show that a historical dataset can be integrated with data collected under an unrelated study to improve abundance estimates and the ability to make inference in the original study. We conclude by offering advice on adapting sampling designs and suggesting ways to overcome barriers to using hierarchical and integrated approaches. We recommend collaborating with other researchers with population modeling and software coding skills as a means to move forward with new or existing

studies. Understanding and moving toward hierarchical study designs and analyses will improve research to the benefit of conserving amphibians and reptiles.

INTRODUCTION

For ecology to aid wildlife conservation and management, the goal must be to reduce uncertainty about ecological processes and the effects of management (Allen et al. 2011, Sells et al. 2018, Nichols 2019). Over the last several decades, major advances have been made in the analysis of abundance and survival data, improving the reliability of estimates and the power to test hypotheses (Schwarz and Seber 1999, Williams et al. 2002, Mazerolle et al. 2007, Iijima 2020). Key among these advances is the use of hierarchical sampling designs and analyses that account for imperfect detection by incorporating models of the observation process, including animal availability (Halstead et al. 2012, Kéry and Royle 2016, 2021, Kellner et al. 2023). These advances have raised the bar on what should be acceptable when estimating how animal abundances or vital rates vary in response to environmental gradients or management actions. Once common approaches are now less defensible, and the biases associated with those approaches can lead to poor inferences and ineffective conservation (Howard and Maerz 2021).

Many wildlife ecologists – particularly those who study amphibians or reptiles – have been slow or resistant to adapt their approaches. We empathize with that struggle. The learning curve to implement hierarchical models, particularly within a Bayesian framework, can be daunting to many who could benefit from their use. Additionally, those conducting long-term studies believe that changing their sampling methodology will compromise comparability of past data to future data (Henry et al. 2008, Banks-Leite et al. 2014, Schmidt et al. 2021). Some of these concerns are misplaced and, as we argue in this paper, all can be overcome without learning complex modeling or losing hard earned data. In fact, the approaches we advocate in this paper could improve ongoing studies or resurrect old datasets. Understanding and moving toward hierarchical study designs and analyses will improve research to the benefit of conserving amphibians and reptiles.

In this paper we aim to present an empathetic and encouraging perspective on the use of a hierarchical approach in designing or modifying sampling methods and analyzing current and historical datasets for amphibians and reptiles. First, we provide a brief overview of hierarchical sampling designs and models. Second, we describe how integrated models can be used to leverage hierarchical sampling designs to improve estimates from data collected under older sampling designs. Third, we present a case study where we estimate forest management effects on terrestrial salamander abundance using an older sampling design, a modified hierarchical design, and integrated abundance models that use older and new datasets. We also demonstrate how integrated models can be used to leverage data from other studies to improve estimates of abundance and reduce uncertainty of forest management effects. Finally, we offer advice on adapting sampling designs and suggest ways to overcome the understandable barriers to herpetologists who could benefit from using hierarchical and integrated approaches.

SECTION 1 – OVERVIEW

Hierarchical models

We are confident that every ecologist thinks hierarchically when designing their field studies, whether they recognize it explicitly or incorporate it into their study design. For example, assume you are interested in measuring the effect of some factor on the abundance of a species of interest. Anyone who has worked with amphibians or reptiles is aware that several factors besides the abundance of the animals will affect the number of animals that you will observe. Amphibians and reptiles are often hidden in inaccessible habitats where you cannot sample them, and their availability for capture will depend on factors that strongly affect their activity. This is why nearly all ecologists measure weather and other variables when they sample. Even when an amphibian or reptile is available for capture, other factors such as visual obstructions, the animal's body size, or observer experience might affect whether the animal is detected. Researchers may try to "control" for some of these effects by incorporating them as covariates in their analyses or standardizing sampling conditions, or they may convince themselves they are unimportant and ignore them. However, explicitly addressing the observation process in sampling design and using hierarchical models is a better, more transparent approach.

Models are representations of hypotheses. Hypotheses lead to predictions that can be observed, so ideally a model incorporates all processes that affect observations, not just the processes of interest. Hierarchical models do precisely this. Hierarchical models are composed of multiple levels [sub models] that separately model the observed data and the unobservable variables (e.g., abundance) or processes of interest (e.g., changes in abundance over time or response to some environmental change; Kéry and Royle 2016). The observation process itself may also have multiple sub models, depending on the ecology of the species and how an area is sampled. For example, there can be separate models to explain the availability of animals for capture (e.g., states of hibernation) and the ability of researchers to detect available animals.

To use a hierarchical model, field sampling methods need to be carefully designed so that observation processes can – ideally – be separated from each other and from the unobservable variables or processes of interest. In a simple world when all individuals are available for capture, a simple two-level model and design that uses repeated counts collected over a period when abundance is not expected to change could estimate observer detection of an individual and true abundance (e.g., a binomial N-mixture model, hereafter N-mixture model; Royle 2004). However, for most herpetofauna, field sampling methods need to be designed to disentangle detection from availability to generate unbiased estimates of abundance and factors that are hypothesized to affect abundance. There are multiple sampling designs that can be used to separate availability from detection. One of the most common is the robust-design approach (Pollock 1982, MacKenzie et al. 2003). In a robust design, there are multiple primary survey events (i.e., primary occasions) among which availability may change. Each primary occasion is comprised of multiple subsamples (i.e., secondary occasions) collected over a brief period, across which availability is not expected to change. For herpetofauna, the latter assumption of availability can be unrealistic. Numbers of individuals available for capture can change quickly and substantially from one day to the next, throughout a day, or with rapid shifts in weather. An alternative approach to using a robust design is to adopt a sampling method that can be used to estimate detection within a single survey (no secondary periods) and requires individuals only be captured or counted once during that sampling event (i.e., removal sampling, distance sampling, or double-observer sampling; Table 5.1), and then to conduct this type of sampling on at least two occasions, between which availability may change (Chandler et al. 2011). In fact, collecting data under a range of conditions while measuring relevant covariates (e.g., weather) can lead to better models of availability and subsequent estimates of abundance. Data collected using these sampling designs can be analyzed using a three-level hierarchical model with sub models that separately estimate detection, availability, and abundance (Chandler et al. 2011, Kéry and Royle 2021).

Integrated models

There are many existing datasets, some long term, that did not incorporate sampling designs suited to hierarchical model analysis. This understandably raises questions about whether older datasets can be modeled to reduce bias and uncertainty and whether investigators should modify sampling designs for ongoing studies. The answer to both questions is yes. Integrated models combine multiple datasets collected using different sampling methodologies in a single model (Schaub and Abadi 2011, Zipkin and Saunders 2018, Hostetter et al. 2019). There are multiple potential benefits to an integrated modeling approach, including the ability to disentangle and estimate variables that would otherwise be confounded (Cole and McCrea 2016, Saunders et al. 2019) and the potential to increase accuracy and precision of estimates (Schaub et al. 2007, Zipkin and Saunders 2018). The integrated datasets may each have different observation processes and may inform unique variables or processes of interest, but the sub models associated with each dataset must share some parameters (e.g., availability, abundance, or survival; Figure 5.1; Schaub and Abadi 2011). These shared parameters link the datasets together, giving the statistical power of one dataset the ability to increase the statistical power of the other. For instance, a larger dataset that is collected over a broader spatial extent can help inform a smaller-scale dataset with more limited replication that is part of a different study (Schmidt et al. 2019, 2021). This means that a dataset collected under a hierarchical design – either from a different study or from changing methods in an ongoing study – can be used to improve abundance estimates for a historical dataset that did not use a sampling methodology capable of separating availability and detection.

SECTION 2 - CASE STUDY USING TERRESTRIAL SALAMANDERS

Case study methods

From 2018 – 2023, we conducted a study (Project 1) in the Coweeta Basin of the southern Appalachian Mountains, North Carolina, USA to estimate terrestrial salamander (Genus *Plethodon*, Family Plethodontidae) abundance and determine if abundance was affected by a set of forest management actions, which included vegetation removal and prescribed fire, started between the 2018 and 2019 active seasons. From 2018 – 2020, we collected count data using a robust-design sampling approach. Under this approach, we conducted sampling for three consecutive nights three times per year (May, July, and late September/early October). This study included 18, 25-m^2 study plots, with 9 plots distributed across a similar elevation (~950 m) in one watershed that underwent forest management and 9 plots distributed in an adjacent reference watershed with similar elevation, aspect, slope, climate, and initial tree composition (Figure 5.2). On each sample night, we conducted time- and area-constrained visual area searches of each plot. A single observer searched the entire plot at least four times and for a minimum of 15 minutes. When observers were finding salamanders, the time to complete a survey was often longer than 15 minutes, and when observers were not finding salamanders, observers usually completed more than four full searches of the plot within 15 minutes. The observer did not turn any cover objects or rake through leaf litter. Every salamander observed was hand captured and placed in an individually labeled Ziploc bag. When a salamander was observed in a burrow, the observer collected as much information on the animal as could be initially observed and then attempted to lure the salamander from the burrow by wiggling a small twig under low light in front of the salamander. All salamanders that were captured were measured (snout-vent length [nearest mm], total length [nearest mm], and wet mass [nearest 0.01

g]). We determined sex by the presence of secondary sexual characteristics (enlarged nasal cirri, presence of a mental gland, or swollen vent for males and the presence of developing eggs for females). After all salamanders were measured, they were released into the plot. In the field or using published weather products, we measured variables that we hypothesized would affect detection (ground visibility), availability (temperature, vapor pressure deficit, and precipitation), and abundance (10-year climate average active season vapor pressure deficit) during surveys or using geographic information system (GIS) software and spatial layers (see detailed methods in Appendix D).

Our intent was for each night (secondary period) of the 3-consecutive-night primary period to be sufficiently close together in time that we could assume constant availability and abundance within the primary period; however, weather conditions often varied enough from night to night and were so influential in salamander activity that availability was highly variable among secondary nights. Consequently, our dataset did not have the capacity to disentangle availability and detection as intended. Instead, this dataset was ultimately more appropriately modeled using an N-mixture model with detection and availability combined into a single parameter, which as we will demonstrate, was not ideal. This dataset also had additional limitations that could affect our power to detect management effects, which included having little pre-management data (1 year) and a limited number of replicates (i.e., 9 unmanaged and 9 managed study plots).

In 2021, we decided to modify our field sampling approach for the Coweeta study so that we could separate detection from availability. We used this approach from 2021 through 2023. Additionally, in 2021 and 2022 we conducted a second, unrelated project (Project 2) on terrestrial salamanders at 24 sites across the Southern Appalachian Mountains in north Georgia, western North Carolina, and eastern Tennessee. In this second project, there were 15, 25-m^2 plots at each study site for a total of 360 study plots (Figure 5.2).

To meet the objectives of both projects, we needed to develop a sampling approach that could address the flaws in our original Coweeta forest management study design and logistically accommodate a more spatially extensive and intensively replicated project. To achieve this, we modified the previous N-mixture sampling to accommodate a three-level hierarchical removal model (Chandler et al. 2011, Kery and Royle 2020). Under this approach, a single observer still conducted time- and area-constrained visual area searches of each plot. The searches were divided into four consecutive sessions of searching the entire plot at least once for a minimum of 3.75 minutes each time, which we defined as a “pass”. Observers completed four passes, which totaled to a minimum of four complete searches for at least 15 minutes, which was the same total effort as our prior N-mixture design. However, under the new methods, we recorded the number of salamanders captured during each pass. We measured salamanders and the variables expected to affect detection, availability, and abundance the same as described above. This new sampling methodology was better suited to disentangle availability and detection because availability could be assumed to be relatively constant within a 15-minute sample and detection could be estimated during this period of constant availability using a removal modeling framework.

For the Coweeta forest management study, we continued sampling the same plots, but conducted sampling on only one night, three times per year (May, July, and October) during 2021 – 2022 and one night each month April – October in 2023. For Project 2, we sampled each site twice, approximately four weeks apart and usually under different weather conditions, between late May and late July of the same year.

Although the second project was conducted to meet a different objective than the first project, both projects shared key elements. In both projects, we needed to account for effects of weather on availability and vegetation obstruction on observer detection. These shared parameters allowed us to link the two datasets together in an integrated model and evaluate the ability of the second project to improve abundance estimates for the Coweeta forest management study. Additionally, the large number of plots and samples in the second project (675 samples of 360 plots) produced a wide array of weather conditions across which to better evaluate effects of weather on availability. Finally, the sites in the second study were intentionally stratified across a range of elevations and hydroclimate, which we hypothesized would be an important determinant of abundance (Figure 5.2; Hocking et al. 2021). Therefore, we could use the larger dataset for the second project to inform our estimates of abundance in the Coweeta forest management project where we lacked sufficient spatial variation between watersheds to otherwise include such effects.

To compare abundance estimates and the power to detect management effects under original and updated field sampling methods and under an integrated modeling approach, we first evaluated the Coweeta data. Specifically, we estimated abundance parameters using (1) only the first 3 years of data for the Coweeta forest management study using the N-mixture model, (2) only the last 3 years of data using the hierarchical removal model, and (3) both Coweeta datasets combined via integration of the two models (Figure 5.1A). Next, we integrated the original N-mixture data and model from the Coweeta forest management study with the second project hierarchical removal data and model (Figure 5.1B). This demonstrates the potential for using two discrete studies with different methods to improve estimates in the original forest management study. Finally, we integrated all three datasets and models to estimate abundance and forest management effects in the Coweeta study. Details of all models are included in Appendix D.

Case study results

Annual estimates of abundance for the first 3 years of the Coweeta forest management study using the N-mixture model were 36%-80% lower than estimates for the same period when the original N-mixture dataset was integrated with data collected using the new hierarchical removal sampling design (either additional years of data from the same sites or data collected for the second project; Figure 5.3). The precision of abundance was higher for the N-mixture analysis, but this was because the estimates were so severely biased low. Even when using coefficient of variation (CV) to assess precision, most abundance estimates had greater dispersion in each of the integrated models compared to the N-mixture model. In other words, the precision in the N-mixture model analysis was also biased and unreliable.

Using the N-mixture model and first 3 years of data only, mean capture probability (detection and availability combined) was 0.13 (95% BCI = 0.08 – 0.18). Like estimates of abundance, this estimate appeared to be biased – in this case the estimate was biased high compared to models that separately estimated detection and availability. When the original N-mixture dataset was integrated with data collected using the new hierarchical removal sampling design, mean detection probability estimates ranged from 0.43 to 0.49 and mean availability estimates ranged from 0.08 to 0.16.

The three-level hierarchical removal design and all integrated model approaches produced more negative estimates of forest management on abundance with greater precision compared to the N-mixture analysis (Figure 5.4). Using only the N-mixture dataset and analysis on data from years 1-3, we found little support for any effect of forest management on salamanders (Figure 5.4). The mean estimate of the effect was negative but negligible and the posterior distribution of estimates broadly overlapped zero, with only an estimated 60% chance of a negative effect of forest management. In contrast, the three-level hierarchical design using a similar amount of data from years 4-6 estimated a larger negative effect and more precise estimate of forest management on terrestrial salamander abundance and a 98% chance that forest management had a negative effect. Integrating the original N-mixture data from years 1-3 with the modified hierarchical design and analysis from years 4-6 also increased the estimated negative effect of forest management and substantially improved the precision of the estimate (Figure 5.4). Integrating the Coweeta N-mixture dataset from years 1-3 with the data and model from the separate project also increased the estimated effect of forest management and improved the precision, though not as much as other models. Integrating all three datasets had the greatest effect on reducing uncertainty about the effect of forest management. In other words, the precision in estimating forest management effects with a limited number of plots and relatively short study duration was improved substantially by both modifying methods later in the study and integrating data from other studies.

Case study discussion

This case study demonstrates the value of adapting sampling methods to facilitate adopting a hierarchical modeling approach for estimating abundance. Our results suggest that abundance estimates using our original sampling methods, which could not separate detection from availability, were biased low. Even though we included variables in our model expected to influence both detection and availability, our abundance estimates likely did not reflect both above ground and below ground individuals. Instead, they likely represented only the surface-active proportion of the population or some proportion in between surface-active individuals and all individuals. This biased metric and underestimation of terrestrial salamander abundance could have implications in representing their importance in ecosystem processes. Additionally, this metric was not useful for evaluating forest management effects because our original sampling methods were unable to confidently detect an effect of management on abundance and had low precision.

The updated methodology and three-level hierarchical removal modeling approach substantially improved our estimates of abundance and our ability to estimate a management effect. This approach was able to detect a negative effect of management with high certainty. Integrating data collected under updated methods with the original data also substantially improved our ability to detect a management effect, compared to using the original data on its own. Additionally, integrating the extensive, broadscale southern Appalachian Mountain dataset with the smaller Coweeta dataset improved precision of abundance estimates relative to integrating the two Coweeta datasets. By integrating the broadscale dataset, we were also able to account for the effect of climate (i.e., 10-year average active season vapor pressure deficit) on abundance, which was not possible with only two sites (i.e., watersheds) included in the Coweeta forest management study. Because we know climate affects abundance, being able to control for this variable gives us additional confidence in the validity of the management effect on abundance.

We recognize that the change in sampling methodology is confounded with the duration of forest management activities. This raises the question of whether our ability to detect an effect of management under the new methodology is a result of an increasing effect of management with time versus increased power of the new methodology. However, results from the integrated model using the first three years of Project 1 data combined with the data from Project 2 show an

85% probability of a negative effect of management on expected abundance, suggesting that the negative effect of management estimated with the removal sampling approach is primarily a consequence of improvements to the sampling methodology.

We acknowledge that alternative approaches exist for estimating abundance from data collected under a methodology incapable of separating observation process components. One of these approaches is to use informative prior distributions under a Bayesian modeling framework (Dixon et al. 2005, Hooten et al. 2007). Using this approach, the model can be given prior knowledge about detection and/or availability that is used in producing estimates (Wade 2000). However, this approach assumes that independent information about detection and/or availability, either empirical or from expert judgement, is available to use for this purpose. Additionally, it can be difficult to construct informative prior distributions even if information is available, and some people may be critical of the use of informative priors to influence estimates (Gelman et al. 1995). Using hierarchical and integrated models allows the data to be used directly to extricate these parameters. Another approach to addressing count data that were not collected using a hierarchical design is to use post-hoc methods to adjust, or ‘correct’, each count for imperfect detection (Schmidt et al. 2019). This method may be better than using unadjusted count data to represent an estimate of abundance but requires additional, potentially costly, data to estimate detection rates. Incorporating additional data with an integrated model is likely to be more efficient and result in better estimates than post-hoc adjustment methods.

Other data types could also be integrated with historical data to facilitate the separate estimation of detection and availability. Telemetry data has potential to inform availability (including pit-tag telemetry for animals that cannot be fit with external transmitters; Connette and Semlitsch 2012, O’Donnell et al. 2016, Carter 2023). Additionally capture-recapture data can be a powerful way to disentangle the components of the observation process (Bailey et al. 2004*c*, *b*, Muñoz et al. 2016). The trade-off with collecting these data is that they are typically more time intensive than unmarked methods, placing greater restrictions on the number of sites that can be monitored or used to test hypotheses. However, integrating telemetry or mark-recapture at a smaller number of sites with unmarked methods conducted across more sites may be an effective way to balance the larger sample size achievable with unmarked methods and the precision and accuracy associated with methods that require individual identification.

An additional consideration in updating sampling methods and integrating the resulting datasets is to ensure that the shared parameters used to link the datasets are equivalent (Frost et al. 2023). For example, detection probabilities from one sampling method may not directly correspond to detection probabilities from another sampling method. In our case, we were able to share detection probabilities across our sub models by linking detection probability associated with all four removal passes (four, 3.75 person-minute visual area searches) with the detection probability from our original field methodology (one 15 person-minute visual area search). In other cases, the amount of effort or type of survey may result in detection probabilities that are not equivalent. Fortunately, in that case, availability may still be a useful link between sub models. For instance, if our removal sampling dataset were integrated with data from other terrestrial salamander projects in the southern Appalachian Mountains collected with different protocols (e.g., transect surveys), the probability that individuals are surface active could still be interpreted the same way across datasets, even if detection probabilities could not. Finally, while the focus of our case study was on estimating abundance, implementing a hierarchical model-based sampling approach can be useful to account for detection and availability in estimating other variables of interest, such as occupancy (DiRenzo et al. 2022) or survival (Howard and Maerz 2021).

SECTION 3 - OVERCOMING BARRIERS TO USING HIERARCHICAL MODELS

Designing field sampling

An essential step to adopting a hierarchical modeling approach is to critically design field sampling such that the observation process(es) can be separated from the unobservable state(s) or process(es) of interest. For terrestrial salamanders and many other herpetofauna this includes determining what, if any, sampling structure can be practically implemented that will disentangle availability from detection. Robust-design methods or repeated surveys that use one of the other model-based methodologies described in Table 5.1 are all potential options for estimating abundance, but not all will be practical for many species of herpetofauna. The removal sampling design in our case study worked well for plethodontids in the southern Appalachian Mountains. However, such an approach was not practical for a concurrent project on stream salamanders, where habitat disturbance from turning cover objects within streams precludes making repeated samples without altering availability or abundance. In addition to determining a hierarchical model-based sampling methodology, other design factors may need to be considered to develop useful models of detection and availability. These include spatial and temporal variables important in explaining each observation component (Table 5.2). One possible way to deal with such variables is to standardize sampling to minimize variability and, ideally, maximize detection and availability; however, such control of field conditions (e.g., weather) is usually impractical if not impossible. An alternative approach is to intentionally sample across a gradient of each important variable to estimate their effects on availability or detection and control for them in estimating abundance.

Identifying datasets to use in an integrated model

As we demonstrated in our case study, designing or redesigning methods for new or ongoing studies and using integrated models is advantageous over continuing to use designs that will result in biased estimates. But what can be done for historical datasets? To address limitations of historical datasets, it may be possible to identify other datasets to integrate with the historical data. Other researchers and practitioners may have data from overlapping study sites or within a comparable region that could address the limitations of the historical data via an integrated model. This collaborative approach is a potentially cost-efficient way to increase the statistical power of data that previously lacked the precision or accuracy needed to meet original objectives. Likewise, we urge those with datasets collected under hierarchical designs to make them publicly available so that others may use them for this purpose.

Collaboration

Finally, we recognize that specialized knowledge is needed to analyze data using hierarchical and integrated models. Despite a growing number of software packages and tutorials offering guidance, there is still a considerable learning curve that is the major barrier to adopting these approaches. Most individuals and entities conducting monitoring do not have the skillsets necessary to perform such analyses or the time to efficiently learn, code, and execute the models. However, population modeling and software coding skills are becoming a common focus in graduate-level training and, thus, the number of researchers with these skills is growing rapidly. Rather than learn these skills, we recommend collaborating with others. We believe that such collaborations are a better option than continuing to use methods that are difficult to defend and more likely to lead to erroneous conclusions and poor management. Ideally, such collaboration could begin with the design (or re-design) of sampling methods, but this should not be a requirement for revisiting old datasets. Ultimately, it should be our shared goal to provide the best information with the least bias and lowest uncertainty to have a more rigorous understanding of reptile and amphibian ecology and to do better in guiding their conservation.

Table 5.1. Potential methods for estimating detection for use in hierarchical models. Repeated surveys using these methods can be used to separately estimate both availability and detection. Some of the listed requirements may be able to be relaxed in some cases but may require additional data (see listed citations). Additionally, all methods can be modified (e.g., 3 or more observers) or used in combinations (e.g., distance sampling + removal sampling) to strengthen inference.

Method	**Requirements**	**Citations**
Double observer sampling		
Independent	Two observers	
	Observers do not influence one another's detection of animals	
	Ability to identify which observer(s) observed which individuals - (i.e., which animals were observed by observer 1 only, which were observed by observer 2 only, which were observed by both observers)	
	Closed population (i.e., animals are available for detection by both observers)	Williams et al. (2002), Moore et al. (2004)
Dependent	Two observers	
	Observer 2 does not influence detection of observer 1, but notes any animals missed by observer 1	
	Ability to identify which observer(s) observed which individuals - (i.e., which animals were observed by observer 1 and confirmed by observer 2, and which were observed only by observer 2)	
	Closed population (i.e., animals are available for detection by both observers)	Cook and Jacobson (1979), Nichols et al. (2000)
Distance sampling	Conducted using point or line transects	
	Detection on the transect is perfect and detection decreases with distance from the transect	
	Neither the placement of the transect nor the presence of the observer affects animal locations (animals are distributed independently and recorded at their original location)	
	Perpendicular distance from the transect to each animal can be measured	Buckland et al. (2015)

Removal sampling	Multiple passes of the survey area (≥ 3 unless detection is very high) with individuals detected each pass not available for detection (or ignored if detected) on subsequent passes	
	Equal sampling effort during each pass is typically assumed	
	The number of individuals captured should tend to decrease across passes	
	Closed population (i.e., availability does not change between passes)	White et al. (1982)

Table 5.2. Potential design factors for designing field sampling to account for detection and availability. Important factors to a particular study may be accounted for by standardizing across surveys or by designing the study to sample over the range of variability of the factor, such that its affects can be estimated and accounted for.

Observation process	Potential design factor
Detection	Observer skill level Survey gear (e.g., net mesh size) Capture technique (e.g., traps, visual survey) Survey effort Individual heterogeneity (e.g., body size, color morph) Habitat complexity (e.g., ground cover, number of cover objects)
Availability	Temporal variation/phenology Topography Aspect Elevation Weather (rainfall, temperature, wind, humidity, vapor pressure deficit) Disturbance Survey area size Body size Sex (e.g., females guarding nests for a part of the year) Ground cover

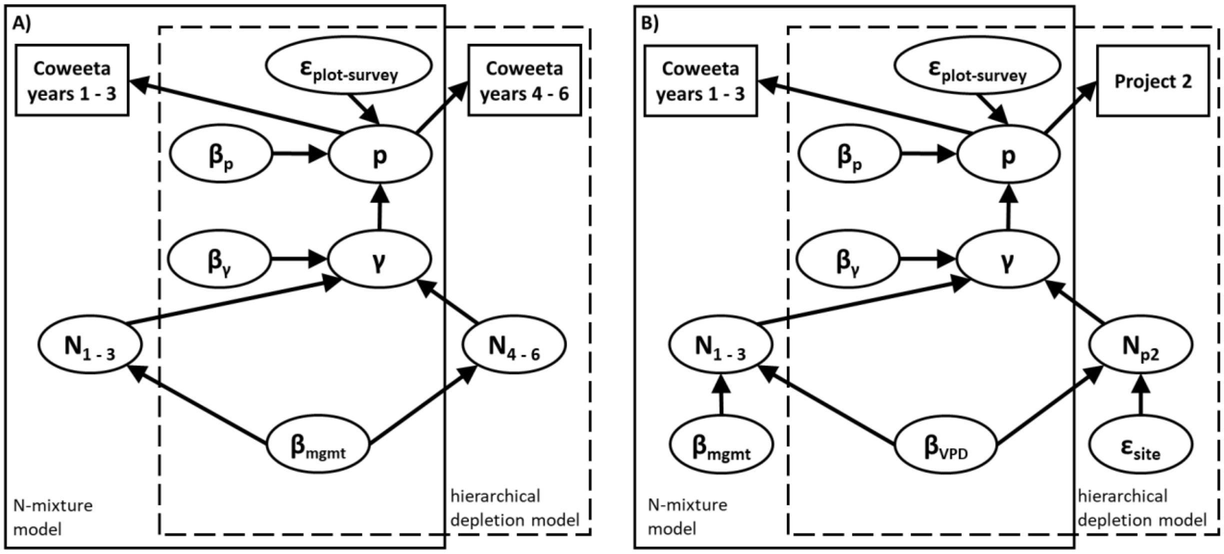
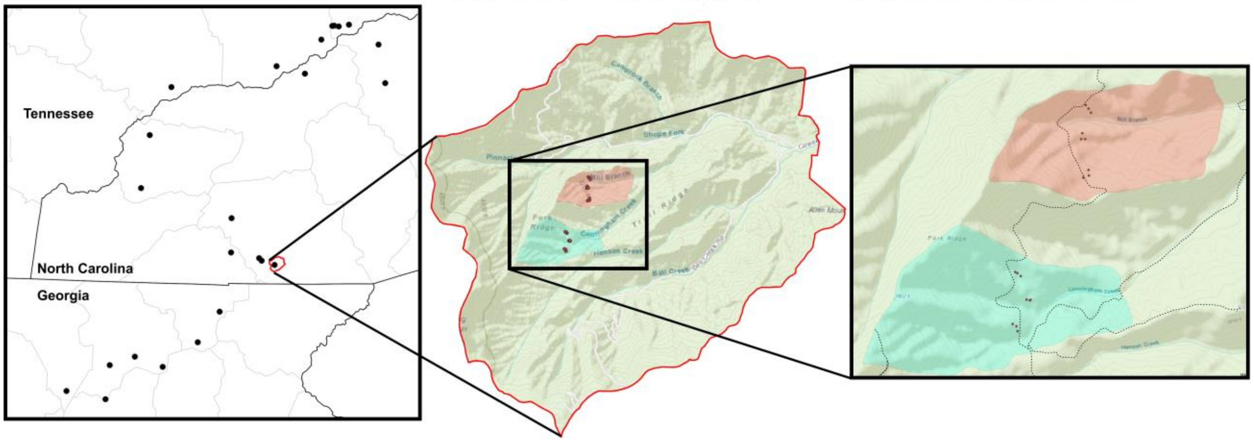


Figure 5.1. Directed acyclic graphs of two integrated models developed for the case study of terrestrial salamander abundance. (A) The integrated model of the first 3 years of data (N-mixture model) and the last 3 years of data (hierarchical removal model) for the Coweeta forest management study. (B) The integrated model of the first 3 years of data (N-mixture model) for the Coweeta forest management study and the broadscale southern Appalachian Mountain study data (hierarchical depletion model). A third model integrating all three datasets (not shown) was also developed. Ovals indicate estimated parameters and rectangles indicate data supplied to the model. Solid arrows represent dependencies between parameters. Notations: *Coweeta years 1 - 3* robust design model data; *Coweeta years 4 – 6* and *Project 2* hierarchical depletion model data; p detection probability; β_p detection covariate (ground visibility); $\varepsilon_{plot\text{-}survey}$ random effect of plot-survey on detection; γ availability; β_γ availability covariates (weather variables); N abundance; β_{mgmt} management covariate on abundance; β_{VPD} climate covariate on abundance; ε_{site} random site effect on abundance.

a) Regional distribution of Project 1 and 2 study sites b) Coweeta Basin, USFS Coweeta Hydrologic Lab c) Project 1 paired watersheds with study plots



c) Mean daily vapor pessure deficit, 2011-2020 d) Mean daily precipitation, 2011-2020 e) Mean annual precipitation, 1991-2020

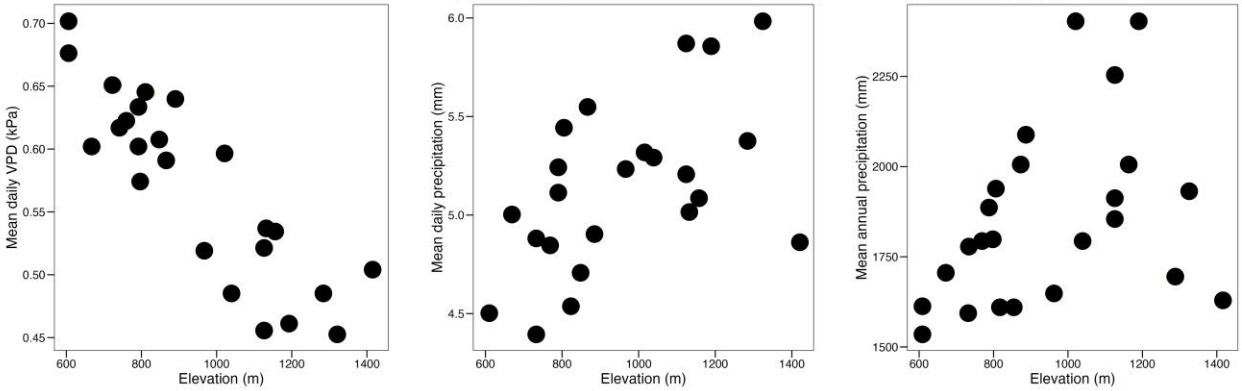


Figure 5.2. Maps showing (a) the locations of study sites for Project 1 and Project 2 with the Coweeta Basin, which was the location for Project 1 and one study site for Project 2; (b) the Coweeta basin within the U.S.F.S. Coweeta Hydrologic Laboratory highlighting the two study watersheds of Project 1 (the red basin is the watershed that was treated and the blue basin is the reference watershed); and (c) close-up of the distribution of study plots within the two Project 1 study watersheds. Figures (d, e, and f) show the distribution of study sites in Project 2 relative elevation and measures of hydroclimate including (d) 1-km grid Daymet estimates of mean daily vapor pressure deficit between April 1 and October 31 from 2011-2020 (Thornton et al. 2022; https://doi.org/10.3334/ORNLDAAC/2131), (e) 1-km grid Daymet estimates of mean daily vapor precipitation between April 1 and October 31 from 2011-2020 (Thornton et al. 2022; https://doi.org/10.3334/ORNLDAAC/2131), and (f) 800-m grid PRISM estimates mean annual precipitation from 1991-2020 (PRISM Climate Group, Oregon State University; https://prism.oregonstate.edu/normals).

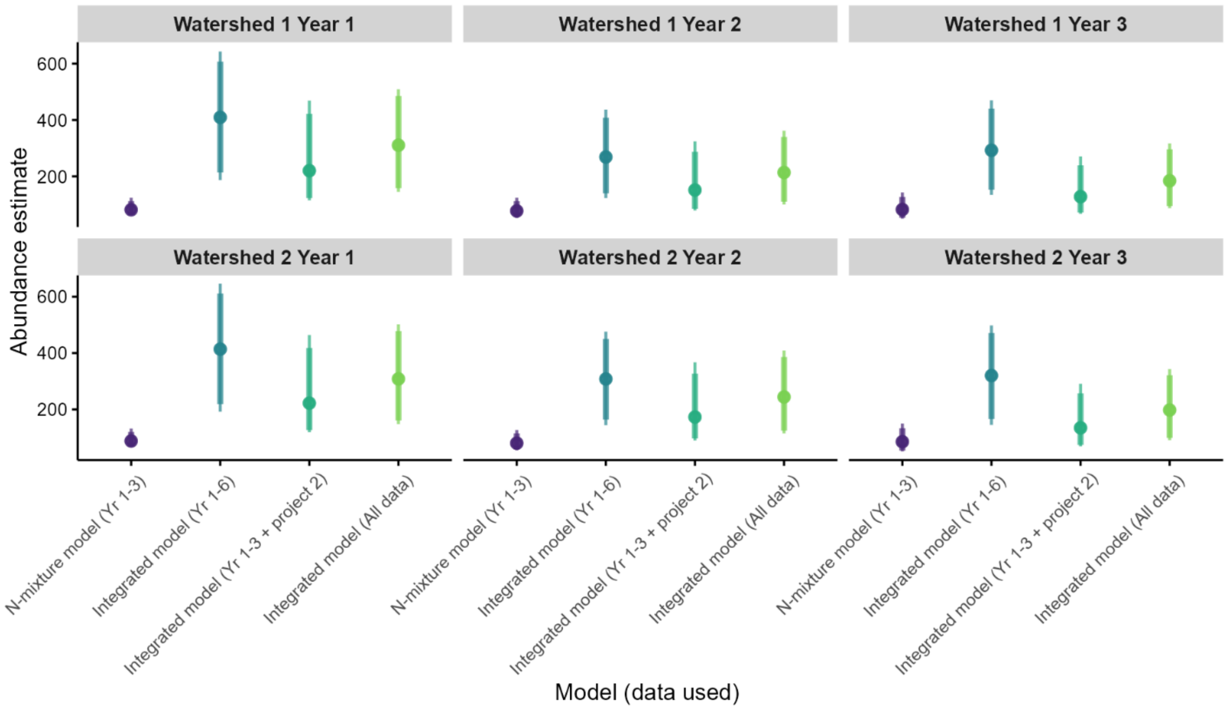


Figure 5.3. Annual estimates of terrestrial salamander abundance in each of the two study sites (i.e., watersheds) for the first 3 years of the Coweeta forest management study. Estimates were generated using four different combinations of datasets and modeling approaches. The first model (years 1-3 of the Coweeta data) included data collected under the original study design, in which availability and detection components were confounded. The second model (years 1-6 of the Coweeta data) integrated data collected under the original study design and data collected under an updated removal sampling approach. The third model (years 1-3 of the Coweeta data and data from a second project) integrated data collected under the original study design with a broadscale southern Appalachian Mountain dataset collected under the updated removal sampling approach. The fourth model (all data) integrated original and updated study design data from Coweeta with updated study design data from the broadscale project. Thick error bars and thin error bars represent 90% and 95% Bayesian credible intervals and points represent mean estimates.

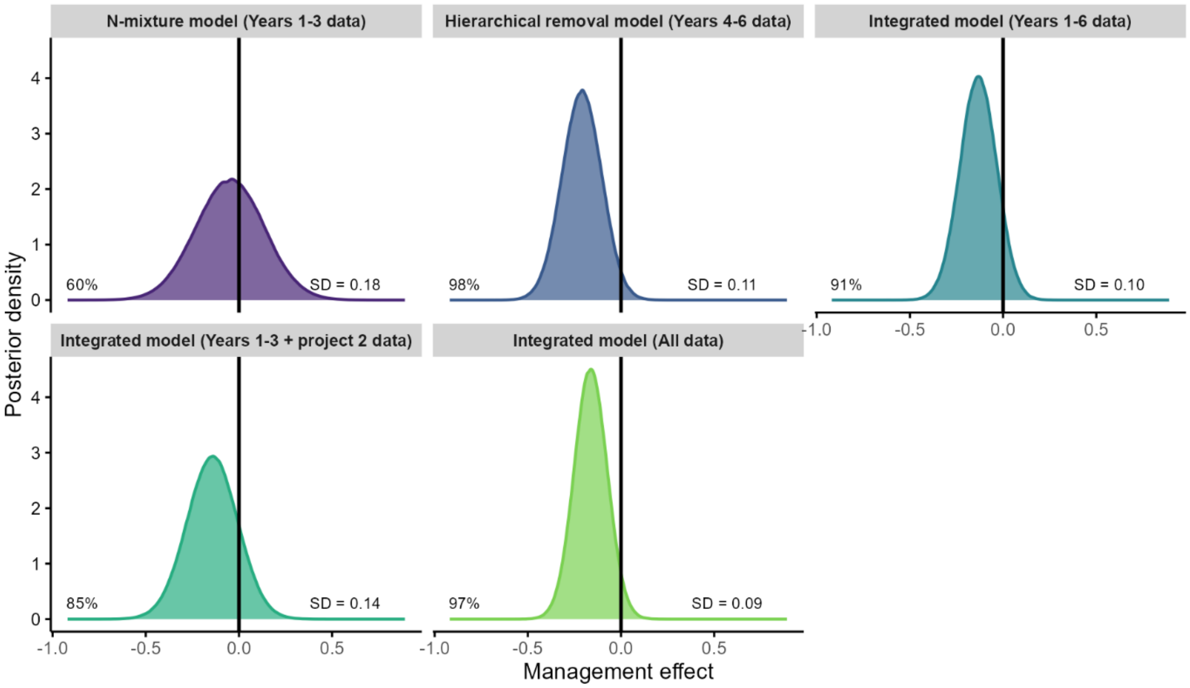


Figure 5.4. Posterior distributions (i.e., estimates) of the effect of forest management on expected abundance of terrestrial salamanders in the Coweeta Basin, North Carolina. Estimates were generated using five different combinations of datasets and modeling approaches. The first model (years 1-3 of the Coweeta data) included data collected under the original study design, in which availability and detection components were confounded. The second model (years 4-6 of the Coweeta data) used data collected under an updated removal sampling approach, in which availability and detection components were both estimable. The third model (years 1-6 of the Coweeta data) integrated data collected under the original study design and data collected under the updated removal sampling approach. The fourth model (years 1-3 of the Coweeta data and data from a second project) integrated data collected under the original study design with a broadscale southern Appalachian Mountain dataset collected under the updated removal sampling approach. The fifth model (all data) integrated original and updated study design data from Coweeta with updated study design data from the broadscale project. Black vertical lines are at 0 (i.e., no management effect), percentages indicate respective probabilities that there was a negative effect of management on expected abundance. Standard deviation (SD) of the estimates is also indicated to provide a comparison of precision across models.

CHAPTER 6

CONCLUSIONS

The goal of this dissertation was to use population modeling approaches to fill information gaps important for effective monitoring and management of herpetofauna in the southeastern United States. The results of this work can, and are, being used to design future monitoring of diamond-backed terrapins. In my final chapter, I broadened the focus to present an approach to improve monitoring and management of herpetofauna more generally. Collectively, this work demonstrates methods to address uncertainty in monitoring and how to design monitoring to meet management objectives in challenging-to-detect species.

In chapter 2, I showed how simulation can be combined with existing data to assess the ability of a monitoring program to meet its objectives. I argue that managers should consider reducing the level of certainty they use to evaluate management effectiveness when resources are limited and species are difficult to monitor. In the case of monitoring diamond-backed terrapins using seining capture-recapture methods, requiring 95% confidence to assess effectiveness of crabbing regulations on terrapin survival is likely impractical, especially when monitoring only one or a few sites. If managers use this level of certainty to inform future management decisions, they face a relatively high risk of deeming a successful management action to be unsuccessful, which may lead to abandoning an effective management approach. However, if managers have additional resources to devote to monitoring, scaling up the number of sites (rather than increasing effort at one site) is likely to be the most effective means of improving accuracy and precision using seining capture-recapture monitoring. Additionally, sampling more sites may more accurately represent what is occurring at the larger landscape scale (i.e., over many populations) in response to management. Ultimately, the modeling approaches used in this chapter are effective to transparently capture the level of certainty expected from different monitoring designs, but selecting a monitoring program will require managers to consider tradeoffs between the expected certainty of information, resource costs, and risk tolerance, among other factors.

In chapter 3, I found that diamond-backed terrapin head count surveys conducted within tidal creeks have the potential to meet state-wide monitoring objectives in Georgia, but there are important limitations of this method to consider. I propose that because this method generated sufficient precision for us to categorize survey areas into three levels of relative abundance, while accounting for uncertainty in estimates, it may be informative to the implementation of management actions by the Georgia DNR. However, the relative abundance estimates generated using this method were sensitive to timing of sampling within the active season and location of sampling along a given creek; therefore, I recommend standardizing these factors across years based on selecting timing and sampling locations most likely to be representative of terrapin abundance within these creeks. If managers are aiming to minimize the number of sites requiring more intensive and potentially more restrictive management actions, a high level of survey effort (i.e., 10 repeated scans during each survey) is also recommended. Considering the challenging nature of monitoring this species and the need to understand the status of populations to apply targeted management, I conclude that, if used judiciously, a method using head count surveys and N-mixture models may be the most effective means of monitoring state-wide. I do stress, however, that because binomial N-mixture models make assumptions that are difficult to meet (or even approximate), the usefulness of this method is context dependent and should be reevaluated if considered for different objectives. Results and recommendations from this chapter are currently being used by the Georgia DNR to design state-wide monitoring.

In chapter 4, I (and the Skidaway Audubon Diamondback Terrapin Rescue Project) demonstrated the capacity for a community-based conservation project to contribute new knowledge about important life history traits in a Species of Greatest Conservation Need. We recorded higher nesting frequencies than previously described for diamond-backed terrapins range-wide and observed larger clutch sizes than previously reported in southern populations. Additionally, I developed a novel model of nesting frequency to remove detection bias in the estimation of the number of nesting events among individuals in a population, which could be applied to other turtle or wildlife populations. This higher fecundity improved projections of population growth and persistence for a managed population of diamond-backed terrapins near Jekyll Island, Georgia, with increases in projected persistence being most notable. I propose that managers use this information to help them decide whether to continue devoting resources to current management practices as well as whether additional resources would be best spent on instituting more intensive management actions for the current focal population versus spending any additional resources on other populations or conservation problems. Additionally, the Skidaway Audubon Diamondback Terrapin Rescue Project can serve as a model for integrating successful community-based conservation with rigorous research.

In chapter 5, I proposed and demonstrated methods for herpetologists to consider for designing or re-designing sampling under a hierarchical approach to improve estimates and inference in species that spend substantial time unavailable for capture. Results of the case study on terrestrial salamanders used in this chapter confirmed the value of such an approach in substantiating previously undetectable effects of management on populations. Additionally, I demonstrated how an integrated model can be used to leverage a dataset collected for a different study to improve abundance estimates and the ability to make inference from a limited historical dataset. I emphasize that field sampling methods need to be carefully designed so that observation processes can – ideally – be separated from each other and the unobservable variable(s) or process(es) of interest. To resurrect historical datasets that were not collected in this manner, researchers can try to identify other datasets to integrate with the historical data. Recognizing that specialized knowledge is needed to analyze data using hierarchical and integrated models, I believe that collaborating with others who have these skills is key for expanding the use of these approaches across studies of amphibians or reptiles, which will ultimately benefit the conservation of these species.

REFLECTIONS AND FUTURE DIRECTIONS

The role of uncertainty in monitoring and management is a central aspect of my dissertation. Accounting for and communicating uncertainty is essential for transparency and honesty in monitoring. Reducing uncertainty about ecological processes and the effects of management is ideal and leads to stronger inferences and more informed management decisions. However, sometimes resources may not be available to address uncertainty in this way. Regardless, management decisions must be made and continuing status quo management (which may mean doing nothing at all) is still a decision. Therefore, another consideration should be the level of uncertainty allowable by managers in their decision-making process. Requiring 95% certainty is common but is arbitrary and may not be practical.

Key among the knowledge I have gathered from this work is the importance of assessing the usefulness of a monitoring methodology in the context of its objectives and the resource restrictions of managers. While capture-recapture is often considered a better monitoring approach than collecting unmarked data, it may not be feasible to implement capture-recapture methods on their own at the scale necessary to be useful to managers. Thus, focusing on such a methodology as the only means to inform management may be short-sighted. Despite my own initial skepticism of using count data and N-mixture models to monitor terrapins, it seems that this methodology may be important to improve the conservation of terrapin populations state-wide.

While integrated models are only used in my final chapter, the potential of integrated models to improve monitoring is a theme that occurs throughout my dissertation. These models could and should be explored to address problems that remain for monitoring diamond-backed terrapins. In particular, integrating head count data with capture-recapture data should be assessed for its effectiveness in improving survival and abundance estimates. As demonstrated in my final chapter, integrated models can improve accuracy and precision of estimates, resulting in better detection of management effects. In addition to producing a strong monitoring indicator for detecting management effects, improving vital rate estimates using integrated models could lead to more reliable population projections in population viability analyses, which I showed the importance of in chapter four.

Ideally, future work to inform optimal monitoring of diamond-backed terrapins should also re-evaluate the head count method after data are collected under the conditions recommended in chapter three so that this approach can be further refined. Specifically, the number of sites needed to retain adequate precision when using this approach should be evaluated. This could then be compared to the numbers of both head count sites and capture-recapture sites that would need to be monitored under an integrated approach. Obtaining better data on crabbing intensity and practices within the monitored creeks would also be useful for making inference about any effects of crabbing on current abundance and population changes over time.

Future research could also investigate ways to best leverage volunteer effort in land-based head counts of diamond-backed terrapins. Chapter four of this dissertation demonstrated the great potential of volunteer-based data collection to improve terrapin monitoring. While boat-based count surveys are more likely to inform the objectives of the Georgia DNR, land-based surveys would be relatively easy to implement at a large-scale with the help of volunteers interested in terrapin conservation. The limitations of using land accessible survey points I identified in chapter four (e.g., nesting phenology, disturbance-level) could be a good starting point for determining how to best use land-based survey data.

Finally, improving the ability of researchers to share and identify datasets that could be integrated with deficient historical datasets may facilitate advances in amphibian and reptile conservation. One way to do this could be the creation of a public repository to share hierarchical-sampling-methodology-based and other datasets that could inform other projects. If past datasets can be resurrected without requiring additional field sampling, this could have large potential for making fast gains in understanding the ecology of many species and populations of conservation need.

REFERENCES

Alexander, C. 2015. Armored shore. Skidaway Institute of Oceanography, University of Georgia. <https://geospatial.gatech.edu/G-WRAP/>.

Allen, C. R., J. J. Fontaine, K. L. Pope, and A. S. Garmestani. 2011. Adaptive management for a turbulent future. Journal of Environmental Management 92:1339–1345.

Ashley, E. A., A. K. Davis, V. K. Terrell, C. Lake, C. Carden, L. Head, R. Choe, and J. C. Maerz. 2021. Effects of Salinity on Hatchling Diamond-Backed Terrapin (*Malaclemys terrapin*) Growth, Behavior, and Stress Physiology. Herpetologica 77:45–55.

Bailey, L. L., T. R. Simons, and K. H. Pollock. 2004*a*. Comparing Population Size Estimators for Plethodontid Salamanders. Journal of Herpetology 38:370–380.

Bailey, L. L., T. R. Simons, and K. H. Pollock. 2004*b*. Estimating detection probability parameters for plethodon salamanders using the robust capture-recapture design. Journal of Wildlife Management 68:1–13.

Bailey, L. L., T. R. Simons, and K. H. Pollock. 2004*c*. Spatial and temporal variation in detection probability of plethodon salamanders using the robust capture-recapture design. Journal of Wildlife Management 68:14–24.

Banks-Leite, C., R. Pardini, D. Boscolo, C. Righetto Cassano, T. Püttker, C. Santos Barros, and J. Barlow. 2014. Assessing the utility of statistical adjustments for imperfect detection in tropical conservation science. Journal of Applied Ecology 51:849–859.

Barker, R. J., M. R. Schofield, W. A. Link, and J. R. Sauer. 2018. On the reliability of N-mixture models for count data. Biometrics 74:369–377.

Beissinger, S. R., and D. R. McCullough. 2002. Population Viability Analysis. University of Chicago Press, Chicago.

Beissinger, S. R., and M. I. Westphal. 1998. On the Use of Demographic Models of Population Viability in Endangered Species Management. The Journal of Wildlife Management 62:821–841.

Beissinger, Steven R. 2002. PVA: past, present, future. Pages 5–17 *in* S.R. Beissinger and D. R. McCullough, editors. Population Viability Analysis. University of Chicago Press, Chicago.

Block, W. M., A. B. Franklin, J. P. Ward, J. L. Ganey, and G. C. White. 2001. Design and implementation of monitoring studies to evaluate the success of ecological restoration on wildlife. Restoration Ecology 9:293–303.

Bock, C. E., and T. L. Root. 1981. The Christmas bird count and avian ecology. Studies in Avian Biology 17–23.

Boyce, M. S. 1992. Population Viability Analysis. Source: Annual Review of Ecology and Systematics 23:481–506.

Bradke, D. R., B. A. Crawford, M. Kaylor, and J. C. Maerz. in press. Evaluating uncertainty to improve a common monitoring method and guide management decisions for diamond-backed terrapins. Journal of Wildlife Management.

Brooks, S. P., E. A. Catchpole, and B. J. T. Morgan. 2000. Bayesian animal survival estimation. Statistical Science 15:357–376.

Buhlmann, K. A., T. S. B. Akre, J. B. Iverson, D. Karapatakis, R. A. Mittermeier, A. Georges, A. G. J. Rhodin, P. P. Van Dijk, and J. W. Gibbons. 2009. A Global Analysis of Tortoise and Freshwater Turtle Distributions with Identification of Priority Conservation Areas. Chelonian Conservation and Biology 8:116–149.

Butler, J. A., C. Broadhurst, M. Green, and Z. Mullin. 2004. Nesting, Nest Predation and Hatchling Emergence of the Carolina Diamondback Terrapin, *Malaclemys terrapin centrata*, in Northeastern Florida. The American Midland Naturalist 152:145–155.

Butler, J. A., and R. Burke. 2019. Reproductive Behavior and Ecology. Pages 81–91 *in* W. M. Roosenburg and V. S. Kennedy, editors. Ecology and Conservation of the Diamond-backed Terrapin. Johns Hopkins University Press, Baltimore, MD.

Cagle, F. R. 1939. A system of marking turtles for future identification. Copeia 1939:170–173.

Carter, C. M. 2023. Impact of introgression on the distribution of traits across an Appalachian salamander hybrid zone. University of Gerogia.

Cecala, K. K., J. W. Gibbons, and M. E. Dorcas. 2009. Ecological effects of major injuries in diamondback terrapins: implications for conservation and management. Aquatic Conservation: Marine and Freshwater Ecosystems 19:421–427.

Chambers, R. M., and J. C. Maerz. 2019. Bycatch in blue crab fisheries. Pages 231–244 *in* W. M. Roosenburg and V. S. Kennedy, editors. Ecology and Conservation of the Diamond-backed Terrapin. Johns Hopkins University Press, Baltimore, MD.

Chandler, R. B., J. A. Royle, and D. I. King. 2011. Inference about density and temporary emigration in unmarked populations. Ecology 92:1429–1435.

Cole, D. J., and R. S. McCrea. 2016. Parameter redundancy in discrete state-space and integrated models. Biometrical Journal 58:1071–1090.

Cole, L. C. 1954. The Population Consequences of Life History Phenomena. The Quarterly Review of Biology 29:103–137.

Connette, G. M., and R. D. Semlitsch. 2012. Successful use of a passive integrated transponder (PIT) system for below-ground detection of plethodontid salamanders. Wildlife Researchv

39:1–6.

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

Couturier, T. 2013. Estimating Abundance and Population Trends When Detection Is Low and Highly Variable: A Comparison of Three Methods for the Hermann’s Tortoise. Population Ecology 77:454–462.

Crawford, B. A., J. C. Maerz, N. P. Nibbelink, K. A. Buhlmann, and T. M. Norton. 2014*a*. Estimating the consequences of multiple threats and management strategies for semi-aquatic turtles. Journal of Applied Ecology 51:359–366.

Crawford, B. A., J. C. Maerz, N. P. Nibbelink, K. A. Buhlmann, T. M. Norton, and S. E. Albeke. 2014*b*. Hot spots and hot moments of diamondback terrapin road-crossing activity. Journal of Applied Ecology 51:367–375.

Crawford, B. A., C. T. Moore, T. M. Norton, and J. C. Maerz. 2018. Integrated analysis for population estimation, management impact evaluation, and decision-making for a declining species. Biological Conservation 222:33–43.

Davenport, J., and J. F. Ward. 1993. The effects of salinity and temperature on appetite in the diamondback terrapin *Malaclemys terrapin* (Latreille). Herpetological Journal 3:95–98.

Dickinson, J. L., B. Zuckerberg, and D. N. Bonter. 2010. Citizen Science as an Ecological Research Tool: Challenges and Benefits. Annual Review of Ecology, Evolution, and Systematics 41:149–172.

DiRenzo, G. V., D. A. W. Miller, and E. H. C. Grant. 2022. Ignoring species availability biases occupancy estimates in single-scale occupancy models. Methods in Ecology and Evolution 13:1790–1804.

Dixon, P. M., A. M. Ellison, and N. J. Gotelli. 2005. Improving the precision of estimates of the frequency of rare events. Ecology 86:1114–1123.

Donini, J., and W. Selman. 2022. Clutch Size, Clutch Frequency, and Egg Characteristics of Diamond-Backed Terrapins (*Malaclemys terrapin*) in Southwestern Louisiana. Southeastern Naturalist 21:235–245.

Dorcas, M. E., J. D. Willson, and J. W. Gibbons. 2007. Crab trapping causes population decline and demographic changes in diamondback terrapins over two decades. Biological Conservation 137:334–340.

Duarte, A., M. J. Adams, and J. T. Peterson. 2018. Fitting N-mixture models to count data with unmodeled heterogeneity: Bias, diagnostics, and alternative approaches. Ecological Modelling 374:51–59.

Dunson, W. A. 1985. Effect of Water Salinity and Food Salt Content on Growth and Sodium Efflux of Hatchling Diamondback Terrapins (*Malaclemys*). Physiological Zoology 58:736–747.

Ellner, S. P., J. Fieberg, D. Ludwig, and C. Wilcox. 2002. Precision of Population Viability Analysis. Biology 16:258–261.

Etterson, M. A., and R. S. Bennett. 2006. On the Use of Published Demographic Data for Population-Level Risk Assessment in Birds. Human and Ecological Risk Assessment: An International Journal 12:1074–1093.

Feder, M. E. 1983. Integrating the Ecology and Physiology of Plethodontid Salamanders. Herpetologica 39:291–310.

Feinberg, J. A. 2004. Nest predation and ecology of terrapins, *Malaclemys terrapin terrapin*, at the Jamaica Bay wildlife refuge. Pages 5–12 *in* C. W. Swarth, W. M. Roosenburg, and E.

Kiviat, editors. Conservation and ecology of turtles of the mid-Atlantic region: a symposium. Bibliomania!, Salt Lake City, Utah.

Feinberg, J. A., and R. L. Burke. 2003. Nesting Ecology and Predation of Diamondback Terrapins, *Malaclemys terrapin*, at Gateway National Recreation Area, New York. Journal of Herpetology 37:517–526.

Freeman, M. F., and J. W. Tukey. 1950. Transformations related to the angular and the square root. Source: The Annals of Mathematical Statistics 21:607–611.

Frost, F., R. McCrea, R. King, O. Gimenez, E. Zipkin, R. McCrea RMcCrea, and lancasteracuk Ruth King. 2023. Integrated Population Models: Achieving Their Potential. Journal of Statistical Theory and Practice 17:6.

Gelman, A., and D. B. Rubin. 1992. Inference from iterative simulation using multiple sequences. Statistical Science 7:457–472.

Georgia Department of Natural Resources Wildlife Resources Division. 2015. Georgia State Wildlife Action Plan. Social Circle, GA.

Gibbons, J. W., and J. E. Lovich. 1990. Sexual Dimorphism in Turtles with Emphasis on the Slider Turtle (*Trachemys scripta*). Herpetological Monographs 69 (4): 397–404.

Gibbs, J. P., H. L. Snell, and C. E. Causton. 1999. Effective monitoring for adaptive wildlife management: lessons from the Galapagos Islands. The Journal of Wildlife Management 63:1055–1065.

Grosse, A. M., J. Van Daniel Dijk, K. L. Holcomb, and J. C. Maerz. 2009. Diamondback terrapin mortality in crab pots in a Georgia tidal marsh. Chelonian Conservation and Biology 8:98–100.

Grosse, A. M., J. C. Maerz, J. Hepinstall-Cymerman, and M. E. Dorcas. 2011. Effects of roads and crabbing pressures on diamondback terrapin populations in coastal Georgia. The Journal of Wildlife Management 75:762–770.

Gurevitch, J., G. A. Fox, N. L. Fowler, and C. H. Graham. 2016. Landscape demography: population change and its drivers across spatial scales. Quarterly Review of Biology 91:459–485.

Hagan, J. M. 1993. Decline of the Rufous-Sided Towhee in the Eastern United States. The Auk 110:863–874.

Halstead, B. J., G. D. Wylie, P. S. Coates, P. Valcarcel, and M. L. Casazza. 2012. 'Exciting statistics': the rapid development and promising future of hierarchical models for population ecology. Animal Conservation 15:133–135.

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

Harden, L. A., N. A. Diluzio, J. W. Gibbons, and M. E. Dorcas. 2007. Spatial and thermal ecology of diamondback terrapins (*Malaclemys terrapin*) in a South Carolina salt marsh. Journal of the North Carolina Academy of Science 123:154–162.

Harden, L. A., S. E. Pittman, J. W. Gibbons, and M. E. Dorcas. 2009. Development of a rapid-assessment technique for diamondback terrapin (*Malaclemys terrapin*) populations using head-count surveys. Applied Herpetology 6:237–245.

Harden, L. A., and A. Southwood Williard. 2019. Osmoregulation. Pages 111–125 *in* W. M. Roosenburg and V. Kennedy, editors. Ecology and Conservation of the Diamond-backed Terrapin. Johns Hopkins University Press, Baltimore, MD.

Hart, K. M. 1999. Declines in diamondbacks: terrapin population modeling and implications for management. Duke University, Durham, North Carolina.

Hart, K. M. 2005. Population biology of diamondback terrapins (*Malaclemys terrapin*): defining and reducing threats across their range. Duke University, Durham, North Carolina.

Harwood, J., and K. Stokes. 2003. Coping with uncertainty in ecological advice: lessons from fisheries. Trends in Ecology & Evolution 18:617–622.

Heinrich, G. L., T. J. Walsh, and W. Williams. 2019. Conservation through environmental education. Pages 245–252 *in* W. M. Roosenburg and V. S. Kennedy, editors. Ecology and Conservation of the Diamond-backed Terrapin. Johns Hopkins University Press, Baltimore, MD.

Henry, P. Y., S. Lengyel, P. Nowicki, R. Julliard, J. Clobert, T. Čelik, B. Gruber, D. S. Schmeller, V. Babij, and K. Henle. 2008. Integrating ongoing biodiversity monitoring: Potential benefits and methods. Biodiversity and Conservation 17:3357–3382.

Heppell, S. S. 1998. Application of Life-History Theory and Population Model Analysis to Turtle Conservation. Copeia 367–375.

Heppell, S. S., H. Caswell, L. B. Crowder, and L. B. Crowder. 2000. Life Histories and Elasticity Patterns: Perturbation Analysis for Species with Minimal Demographic Data. Ecology 81:654–665.

Hildebrand, S. F., and C. Hatsel. 1926. Diamond-back terrapin culture at Beaufort, N.C. Economic Circular No. 60.

Hocking, D. J., J. A. Crawford, W. E. Peterman, and J. R. Milanovich. 2021. Abundance of montane salamanders over an elevational gradient. Ecology and Evolution 11:1378–1391.

Holliday, D. K., A. A. Elskus, and W. M. Roosenburg. 2009. Impacts of multiple stressors on growth and metabolic rate of *Malaclemys terrapin*. Environmental Toxicology and

Chemistry 28:338–345.

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

Hooten, M. B., C. K. Wikle, R. M. Dorazio, and J. A. Royle. 2007. Hierarchical Spatiotemporal Matrix Models for Characterizing Invasions. Biometrics 63:558–567.

Hostetter, N. J., B. Gardner, T. S. Sillett, K. H. Pollock, and T. R. Simons. 2019. An integrated model decomposing the components of detection probability and abundance in unmarked populations. Ecosphere 10.

Howard, J. S., and J. C. Maerz. 2021. Review and Synthesis of Estimated Vital Rates for Terrestrial Salamanders in the Family Plethodontidae. Ichthyology & Herpetology 109:929–939.

Iijima, H. 2020. A Review of Wildlife Abundance Estimation Models: Comparison of Models for Correct Application. Mammal Study 45:177–188.

Isdell, R. E., R. M. Chambers, D. M. Bilkovic, and M. Leu. 2015. Effects of terrestrial–aquatic connectivity on an estuarine turtle. Diversity and Distributions 21:643–653.

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

Jones, J. P. G. 2011. Monitoring species abundance and distribution at the landscape scale. Journal of Applied Ecology 48:9–13.

Kellner, K. 2021. jagsUI: A Wrapper Around “rjags” to Streamline “JAGS” Analyses. <https://cran.r-project.org/package=jagsUI>.

Kellner, K. F., A. D. Smith, J. A. Royle, M. Kéry, J. L. Belant, and R. B. Chandler. 2023. The unmarked R package: Twelve years of advances in occurrence and abundance modelling in ecology. Methods in Ecology and Evolution 14:1408–1415.

Kendall, W. L., and C. T. Moore. 2012. Maximizing the utility of monitoring to the adaptive management of natural resources. Pages 74–98 *in* R. A. Gitzen, J. J. Millspaugh, A. B. Cooper, and D. S. Licht, editors. Design and analysis of long-term ecological monitoring studies. Cambridge University Press.

Kéry, M., and J. A. Royle. 2016. Applied hierarchical modeling in ecology: analysis of distribution, abundance and species richness in R and BUGS: Volume1: Prelude and Static Models. Academic Press.

Kéry, M., and J. A. Royle. 2021. Kéry, Marc, and J. Andrew Royle. Applied hierarchical modeling in ecology: Analysis of distribution, abundance and species richness in R and BUGS: Volume 2: Dynamic and advanced models. Academic Press.

Kéry, M., and B. R. Schmidt. 2008. Imperfect detection and its consequences for monitoring for conservation. Community Ecology 9:207–216.

King, P., and J. P. Ludlam. 2014. Status of diamondback terrapins (*Malaclemys terrapin*) in north inlet-Winyah Bay, South Carolina. Chelonian Conservation and Biology 13:119–124.

Knape, J., D. Arlt, F. Barraquand, Å. Berg, M. Chevalier, T. Pärt, A. Ruete, and M. Żmihorski. 2018. Sensitivity of binomial N-mixture models to overdispersion: The importance of assessing model fit. Methods in Ecology and Evolution 9:2102–2114.

Lamont, M. M., D. Johnson, and D. J. Catizone. 2021. Home ranges and Movements of Two Diamondback Terrapins (*Malaclemys terrapin macrospilota*) in Northwest Florida. Estuaries and Coasts 44:1484–1491.

Lawson, A. J., P. G. R. Jodice, T. R. Rainwater, K. D. Dunham, M. Hart, J. W. Butfiloski, P. M. Wilkinson, K. W. McFadden, and C. T. Moore. 2022. Hidden in plain sight: Integrated population models to resolve partially observable latent population structure. Ecosphere 13:e4321.

Levasseur, P., R. Prescott, M. Faherty, and C. Sutherland. 2023. Factors affecting spatiotemporal patterns of nest site selection and abundance in diamondback terrapins. Ecology and Evolution 13:e9866.

Levasseur, P., S. Sterrett, and C. Sutherland. 2019. Visual head counts: a promising method for efficient monitoring of diamondback terrapins. Diversity 11:101.

Lindenmayer, D. B., R. C. Lacy, and V. C. Thomas. 1993. Population Viability Analysis as a Tool in Wildlife Conservation Policy: With Reference to Australia. Environmental Management 17:745–758.

Lindenmayer, D. B., and G. E. Likens. 2009. Adaptive monitoring: a new paradigm for long-term research and monitoring. Trends in Ecology & Evolution 24:482–486.

Link, W. A., M. R. Schofield, R. J. Barker, and J. R. Sauer. 2018. On the robustness of N-mixture models. Ecology 99:1547–1551.

Lovich, J. E., J. W. Gibbons, and K. M. Greene. 2019. Life history with emphasis on geographic variation. Pages 63–80 *in* W. M. Roosenburg and V. S. Kennedy, editors. Ecology and conservation of the diamond-backed terrapin. Johns Hopkins University Press, Baltimore, MD.

Lovich, J. E., and K. M. Hart. 2019. Taxonomy: a history of controversy and uncertainty. Pages 37–50 *in* W. M. Roosenburg and V. S. Kennedy, editors. Ecology and Conservation of the Diamond-backed Terrapin. Johns Hopkins University Press, Baltimore, MD.

Lovich, J. E., M. Thomas, K. Ironside, C. Yackulic, and S. R. Puffer. 2018. Spatial distribution of estuarine diamond-backed terrapins (*Malaclemys terrapin*) and risk analysis from commercial blue crab (*Callinectes sapidus*) trapping at the Savannah Coastal Refuges Complex, USA. Ocean & Coastal Management 157:160–167.

Lyons, J. E., M. C. Runge, H. P. Laskowski, and W. L. Kendall. 2008. Monitoring in the context of structured decision-making and adaptive management. The Journal of Wildlife Management 72:1683–1692.

MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. Ecology 84:2200–2207.

Maerz, J. C., R. A. Seigel, and B. A. Crawford. 2019. Conservation in Terrestrial Habitats: Mitigating Habitat Loss, Road Mortality, and Subsidized Predators. Pages 201–220 *in* W. M. Roosenburg and V. S. Kennedy, editors. Ecology and Conservation of the Diamond-backed Terrapin. Johns Hopkins University Press, Baltimore, MD.

Manlik, O., R. C. Lacy, and W. B. Sherwin. 2018. Applicability and limitations of sensitivity analyses for wildlife management. Journal of Applied Ecology 55:1430–1440.

Martin, J., J. A. Royle, D. I. Mackenzie, H. H. Edwards, M. Kéry, and B. Gardner. 2011. Accounting for non-independent detection when estimating abundance of organisms with a Bayesian approach. Methods in Ecology and Evolution 2:595–601.

Mazerolle, M. J., L. L. Bailey, W. L. Kendall, J. A. Royle, S. J. Converse, and J. D. Nichols. 2007. Making Great Leaps Forward: Accounting for Detectability in Herpetological Field Studies. Source: Journal of Herpetology 41:672–689.

McGowan, C. P., M. R. Ryan, M. C. Runge, J. J. Millspaugh, and J. Fitts Cochrane. 2011. The role of demographic compensation theory in incidental take assessments for endangered species. Biological Conservation 144:730–737.

McKinley, D. C., A. J. Miller-Rushing, H. L. Ballard, R. Bonney, H. Brown, S. C. Cook-Patton, D. M. Evans, R. A. French, J. K. Parrish, T. B. Phillips, S. F. Ryan, L. A. Shanley, J. L. Shirk, K. F. Stepenuck, J. F. Weltzin, A. Wiggins, O. D. Boyle, R. D. Briggs, S. F. Chapin, D. A. Hewitt, P. W. Preuss, and M. A. Soukup. 2017. Citizen science can improve conservation science, natural resource management, and environmental protection. Biological Conservation 208:15–28.

Mills, L. S., and M. S. Lindberg. 2002. Sensitivity analysis to evaluate the consequences of conservation actions. Pages 338–366 *in* S. R. Beissinger and D. R. McCullough, editors. Population Viability Analysis. University of Chicago Press, Chicago.

Mitro, M. G. 2003. Demography and viability analyses of a diamondback terrapin population. Canadian Journal of Zoology 81:716–726.

Montevecchi, W. A., and J. Burger. 1975. Aspects of the Reproductive Biology of the Northern Diamondback Terrapin *Malaclemys terrapin terrapin*. Source: The American Midland Naturalist 94:166–178.

Muñoz, D. J., K. Miller Hesed, E. H. Campbell Grant, and D. A. W. Miller. 2016. Evaluating within-population variability in behavior and demography for the adaptive potential of a dispersal-limited species to climate change. Ecology and Evolution 6:8740–8755.

Nichols, J. D. 2019. Confronting uncertainty: Contributions of the wildlife profession to the broader scientific community. Journal of Wildlife Management 83:519–533.

Nichols, J. D., and B. K. Williams. 2006. Monitoring for conservation. Trends in Ecology and Evolution 21:668–673.

NOAA Coastal Services Center Coastal Inundation Digital Elevation Model. 2012. NOAA Office for Coastal Management Sea Level Rise Data: Current Mean Higher High Water

Inundation Extent. <http://www.csc.noaa.gov/digitalcoast/tools/slrviewe>. Accessed 22 Dec 2022.

O’Donnell, K. M., and R. D. Semlitsch. 2015. Advancing Terrestrial Salamander Population Ecology: The Central Role of Imperfect Detection. Source: Journal of Herpetology 49:533–540.

O’Donnell, K. M., F. R. Thompson, and R. D. Semlitsch. 2016. Prescribed fire alters surface activity and movement behavior of a terrestrial salamander. Journal of Zoology 298:303–309.

Peterman, W. E., and R. D. Semlitsch. 2014. Spatial variation in water loss predicts terrestrial salamander distribution and population dynamics. Oecologia 176:357–369.

Pinheiro, J., D. Bates, and R Core Team. 2022. _nlme: Linear and Nonlinear Mixed Effects Models_. R package version 3.1-157. <https://cran.r-project.org/package=nlme>.

Pinheiro, J. C., and D. M. Bates. 2000. Mixed-effects models in S and S-PLUS. Springer, New York.

Plummer, M. 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. Pages 1–10 *in*. Proceedings of the 3rd International Conference on Distributed Statistical Computing (DSC 2003). Vienna, Austria.

Pollock, K. H. 1982. A capture-recapture design robust to unequal probability of capture. The Journal of Wildlife Management 46:752–757.

R Core Team. 2022. R: A Language and Environment for Statistical Computing. Vienna, Austria. <https://www.r-project.org/>.

Radzio, T. A., J. A. Smolinsk, and W. M. Roosenburg. 2013. Low use of required terrapin bycatch reduction devices in a recreational crab pot fishery. Herpetological Conservation and Biology 8:222–227.

Reynolds, J. H., W. L. Thompson, and B. Russell. 2011. Planning for success: identifying effective and efficient survey designs for monitoring. Biological Conservation 144:1278–1284.

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 227:361–368.

Roe, M. T., J. D. Anadón, and R. L. Burke. 2022. Survival rates of two nesting aggregations of diamondback terrapins at neighboring sites in New York City. The Journal of Wildlife Management 86:e22257.

Roosenburg, W. M. 1991. The diamondback terrapin: population dynamics, habitat requirements, and opportunities for conservation. Pages 227–234 *in*. New Perspectives in the Chesapeake System: A Research and Management Partnership. Proceedings of a Conference. CRC Publ. No. 137.

Roosenburg, W. M. 1994. Nesting habitat requirements of the diamondback terrapin: a geographic comparison. Wetlands Journal 6:9–12.

Roosenburg, W. M. 2004. The impact of crab pot fisheries on terrapin (*Malaclemys terrapin*) populations: where are we and where do we need to go. Pages 23–30 *in*. Conservation and ecology of turtles of the mid-Atlantic region: a symposium. Salt Lake City, UT.

Roosenburg, W. M., P. J. Baker, R. Burke, M. E. Dorcas, and R. C. Wood. 2019. *Malaclemys terrapin*. The IUCN Red List of Threatened Species e.T12695A507698.

Roosenburg, W. M., and R. L. Burke. 2019. Capture, measurement, and field techniques. Pages 7–25 *in* Willem M. Roosenburg and V. S. Kennedy, editors. Ecology and conservation of the diamond-backed terrapin. Johns Hopkins University Press, Baltimore, MD.

Roosenburg, W. M., W. Cresko, M. Modesitte, and M. B. Robbins. 1997. Diamondback terrapin (*Malaclemys terrapin*) mortality in crab pots. Conservation Biology 11:1166–1172.

Roosenburg, W. M., and A. E. Dunham. 1997. Allocation of reproductive output: Egg- and clutch-size variation in the diamondback terrapin. Copeia 1997:290–297.

Roosenburg, W. M., and V. S. Kennedy, editors. 2019. Ecology and Conservation of the Diamond-backed Terrapin. Johns Hopkins University Press, Baltimore, MD.

Rose, J. P., R. Kim, E. J. Schoenig, P. C. Lien, and B. J. Halstead. 2022. Integrating growth and survival models for flexible estimation of size-dependent survival in a cryptic, endangered snake. Ecology and Evolution 12:e8799.

Royle, J. A. 2004. N-Mixture Models for Estimating Population Size from Spatially Replicated Counts. Biometrics 60:108–115.

Sanderlin, J. S., W. M. Block, B. E. Strohmeyer, V. A. Saab, and J. L. Ganey. 2019. Precision gain versus effort with joint models using detection/non-detection and banding data. Ecology and Evolution 9:804–817.

Sauer, J. R., J. E. Fallon, and R. Johnson. 2003. Use of North American Breeding Bird Survey Data to Estimate Population Change for Bird Conservation Regions. Source: The Journal of Wildlife Management 67:372–389.

Saunders, S. P., M. T. Farr, A. D. Wright, C. A. Bahlai, J. W. Ribeiro, S. Rossman, A. L. Sussman, T. W. Arnold, and E. F. Zipkin. 2019. Disentangling data discrepancies with integrated population models. Ecology 100:e02714.

Schaub, M., and F. Abadi. 2011. Integrated population models: A novel analysis framework for deeper insights into population dynamics. Journal of Ornithology 152:S227–S237.

Schaub, M., O. Gimenez, A. Sierro, and E. Arlettaz. 2007. Use of Integrated Modeling to Enhance Estimates of Population Dynamics Obtained from Limited Data. Conservation Biology 21:945–955.

Schaub, M., R. Zink, H. Beissmann, F. Sarrazin, and R. Arlettaz. 2009. When to end releases in reintroduction programmes: demographic rates and population viability analysis of bearded vultures in the Alps. Journal of Applied Ecology 46:92–100.

Schmidt, J. H., J. H. Reynolds, K. L. Rattenbury, L. M. Phillips, K. S. White, D. Schertz, J. M. Morton, and H. S. Kim. 2019. Integrating distance sampling with minimum counts to improve monitoring. The Journal of Wildlife Management 83:1454–1465.

Schmidt, J. H., J. H. Reynolds, K. S. White, D. T. Schertz, J. M. Morton, and H. S. Kim. 2021. Integrating Distance Sampling and Minimum Count Data: A Reply to Becker and Herreman. The Journal of Wildlife Management 85:411–417.

Schtickzelle, Nicolas, Michiel F WallisDeVries, M. Baguette Schtickzelle, N Schtickzelle, M. Baguette, M F WallisDeVries, and D. Vlinderstichting. 2005. Using surrogate data in population viability analysis: the case of the critically endangered cranberry fritillary butterfly. Oikos 109:89–100.

Schultz, C. B., L. M. Brown, E. Pelton, and E. E. Crone. 2017. Citizen science monitoring demonstrates dramatic declines of monarch butterflies in western North America. Biological Conservation 214:343–346.

Schwarz, C. J., and G. A. F. Seber. 1999. Estimating Animal Abundance: Review III. Statistical Science 14:427–456.

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

Seigel, R. A. 1980. Nesting Habits of Diamondback Terrapins (*Malaclemys terrapin*) on the

Atlantic Coast of Florida. Source: Transactions of the Kansas Academy of Science 83:239–246.

Sells, S. N., S. B. Bassing, K. J. Barker, S. C. Forshee, A. C. Keever, J. W. Goerz, and M. S. Mitchell. 2018. Increased scientific rigor will improve reliability of research and effectiveness of management. The Journal of Wildlife Management 82:485–494.

Sewell, D., G. Guillera-Arroita, R. A. Griffiths, and T. J. C. Beebee. 2012. When is a species declining? Optimizing survey effort to detect population changes in reptiles. PLOS ONE 7:e43387.

Sheridan, C. M., J. R. Spotila, W. F. Bien, and H. W. Avery. 2010. Sex-biased dispersal and natal philopatry in the diamondback terrapin, *Malaclemys terrapin*. Molecular Ecology 19:5497–5510.

Shirk, J. L., H. L. Ballard, C. C. Wilderman, T. Phillips, A. Wiggins, R. Jordan, E. Mccallie, M. Minarchek, B. V Lewenstein, M. E. Krasny, and R. Bonney. 2012. Public Participation in Scientific Research: a Framework for Deliberate Design. Ecology and Society 17:29.

Shrader-Frechette, K. S., and E. D. McCoy. 1992. Statistics, costs and rationality in ecological inference. Trends in Ecology & Evolution 7:96–99.

Silvertown, J. 2009. A new dawn for citizen science. Trends in Ecology & Evolution 24:467–471.

Southwood Williard, A., L. A. Harden, T. T. Jones, and S. R. Midway. 2019. Effects of temperature and salinity on body fluid dynamics and metabolism in the estuarine diamondback terrapin (*Malaclemys terrapin*). Journal of Experimental Biology. 222: jeb202390.

Di Stefano, J. 2003. How much power is enough? Against the development of an arbitrary convention for statistical power calculations. Functional Ecology 17:707–709.

Steger, C., B. Butt, and M. B. Hooten. 2017. Safari Science: assessing the reliability of citizen science data for wildlife surveys. Journal of Applied Ecology 54:2053–2062.

Sullivan, B. L., C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling. 2009. eBird: A citizen-based bird observation network in the biological sciences. Biological Conservation 142:2282–2292.

Sun, C. C., J. A. Royle, and A. K. Fuller. 2019. Incorporating citizen science data in spatially explicit integrated population models. Ecology 100.

Szerlag-Egger, S., and S. P. McRobert. 2007. Northern diamondback terrapin occurrence, movement, and nesting activity along a salt marsh access road. Chelonian Conservation and Biology 6:295–301.

Taub, F. B. 1961. The Distribution of the Red-Backed Salamander, Plethodon C. Cinereus, within the Soil. Ecology 42:681–698.

Temple, S. A., and J. A. Wiens. 1989. Bird populations and environmental changes: can birds be bio-indicators? American Birds 43:260–270.

Theobald, E. J., A. K. Ettinger, H. K. Burgess, L. B. DeBey, N. R. Schmidt, H. E. Froehlich, C. Wagner, J. HilleRisLambers, J. Tewksbury, M. A. Harsch, and J. K. Parrish. 2015. Global change and local solutions: Tapping the unrealized potential of citizen science for biodiversity research. Biological Conservation 181:236–244.

Thornton, M. M., R. Shrestha, Y. Wei, P. E. Thornton, S-C. Kao, and B. E. Wilson. 2022. Daymet: monthly climate summaries on a 1-km grid for North America, Version 4 R1. ORNL DAAC, Oak Ridge, Tennessee, USA. https://doi.org/10.3334/ORNLDAAC/2131

Tucker, A. D., N. N. Fitzsimmons, and J. W. Gibbons. 1995. Resource partitioning by the estuarine turtle *Malaclemys terrapin*: Trophic, spatial, and temporal foraging constraints. 51:167–181.

Tucker, A. D., J. W. Gibbons, and J. L. Greene. 2001. Estimates of adult survival and migration for diamondback terrapins: conservation insight from local extirpation within a metapopulation. Canadian Journal of Zoology 79:2199–2209.

Underwood, E. B., S. Bowers, J. C. Guzy, J. E. Lovich, C. A. Taylor, J. Whitfield Gibbons, and M. E. Dorcas. 2013. Sexual Dimorphism and Feeding Ecology of Diamond-backed Terrapins (*Malaclemys terrapin*). Herpetologica 69:397–404.

de Valpine, P., C. Paciorek, D. Turek, N. Michaud, C. Anderson-Bergman, F. Obermeyer, C. Wehrhahn Cortes, A. Rodrìguez, D. Temple Lang, and S. Paganin. 2022. NIMBLE: MCMC, particle filtering, and programmable hierarchical modeling.

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.

Veech, J. A., J. R. Ott, and J. R. Troy. 2016. Intrinsic heterogeneity in detection probability and its effect on N-mixture models. Methods in Ecology and Evolution 7:1019–1028.

Wade, P. R. 2000. Bayesian Methods in Conservation Biology. Conservation Biology 14:1308–1316.

Watanabe, S. 2010. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. Journal of Machine Learning Research 11:3571–3594.

Watanabe, S. 2013. A widely applicable Bayesian information criterion. Journal of Machine Learning Research 14:867–897.

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

Witczak, L. R., J. C. Guzy, S. J. Price, J. W. Gibbons, and M. E. Dorcas. 2014. Temporal and spatial variation in survivorship of diamondback terrapins (*Malaclemys terrapin*). Chelonian Conservation and Biology 13:146–151.

Zimmerman, T. D. 1992. Latitudinal reproductive variation of the salt marsh turtle, the diamondback terrapin (*Malaclemys terrapin*). University of Charleston.

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

Zylstra, E. R., L. Ries, N. Neupane, S. P. Saunders, M. I. Ramírez, E. Rendón-Salinas, K. S. Oberhauser, M. T. Farr, and E. F. Zipkin. 2021. Changes in climate drive recent monarch butterfly dynamics. Nature Ecology & Evolution 2021 5:10 5:1441–1452.

APPENDIX A

SUPPLEMENTAL INFORMATION FOR CHAPTER TWO

Model selection

We used a hierarchical robust-design Cormack-Jolly-Seber (CJS) model to estimate capture, availability, and apparent survival probabilities. We specified the observation process, which indicates if individual (i) was observed in secondary period (j), as a Bernoulli outcome with probability equal to the product of its capture probability ($p_{i,j}$) and its current state ($s_{i,j}$) as available (1) or unavailable (0; Eq. 1), with availability regulated by a Bernoulli outcome with a probability equal to the product of availability probability ($\gamma_{i,t}$) and its current state ($z_{i,t}$) as alive (1) or dead (0; Eq. 2). We specified the state process as a Bernoulli outcome with a probability equal to the product of the individual's apparent survival probability during the previous sampling interval ($\varphi_{i,t-1}$) and its state at the previous sampling occasion ($z_{i,t-1}$; Eq. 3).

$$y_{i,j} \sim Bernoulli(\mathrm{p}_{i,j} \times s_{i,j}) \tag{1}$$

$$s_{i,t} \sim Bernoulli(\gamma_{i,t} \times z_{i,t}) \tag{2}$$

$$z_{i,t} \sim Bernoulli(\varphi_{i,t-1} \times z_{i,t-1}) \tag{3}$$

For each of our eight candidate models, we modeled apparent survival probability as a logit-linear function of year (ε_t; Eq.4), with year modeled as a random effect.

$$logit(\varphi_{i,t}) = \mu + \varepsilon_t \tag{4}$$

We modeled availability probability as a logit-linear function of year (ε_t; Eq. 5) , with year modeled as a random effect, or of mean tide amplitude of the low tides occurring during sampling for a given year ($Tide_t$; Eq. 6).

$$logit(\gamma_{i,t}) = \mu + \varepsilon_t \tag{5}$$

$$logit(\gamma_{i,t}) = \mu + \beta_{tide} \times Tide_t \tag{6}$$

We modeled capture probability as either time-invariant (Eq. 7) or as a logit-linear function of year (ε_t; Eq. 8), day (ε_j; Eq. 9), or daily tide amplitude ($Tide_j$; Eq. 10), with year and day treated as random effects. We also combined each of the previous four models additively a with behavioral (recapture) effect by using a binary covariate (c) that indicated if the individual was previously captured (1) or not (0) in the current primary period (Eq. 11–14).

$$logit(p_{i,j}) = \mu \tag{7}$$

$$logit(p_{i,j}) = \mu + \varepsilon_t \tag{8}$$

$$logit(p_{i,j}) = \mu + \varepsilon_j \tag{9}$$

$$logit(p_{i,j}) = \mu + \beta_{tide} \times Tide_j \tag{10}$$

$$logit(p_{i,j}) = \mu + \beta_c \times c_{i,j} \tag{11}$$

$$logit(p_{i,j}) = \mu + \varepsilon_t + \beta_c \times c_{i,j} \tag{12}$$

$$logit(p_{i,j}) = \mu + \varepsilon_j + \beta_c \times c_{i,j} \tag{13}$$

$$logit(p_{i,j}) = \mu + \beta_{tide} \times Tide_j + \beta_c \times c_{i,j} \tag{14}$$

In summary, our model set included all possible combinations of models for availability and detection probability, excluding those models containing a tide effect in both parameters, yielding 8 models.

Table A.1. Model selection for candidate robust-design Cormack-Jolly-Seber models used to estimate diamond-backed terrapin capture probability (p), availability (γ), and apparent survival (φ). WAIC is the mean WAIC, WAIC SE is the standard error of WAIC, and ΔWAIC is the difference between the mean WAIC of each model and the WAIC of the top-ranked model. Our model set included all possible combinations of models for availability and detection probability, excluding those models containing a tide effect in both parameters, yielding 8 models. Analysis was conducted on capture-recapture data collected during 2010 – 2022 (except for 2020) from two salt marsh tidal creeks adjacent to the Downing Musgrove Causeway in Glynn County, Georgia, USA.

Model description	**WAIC**	**WAIC SE**	**ΔWAIC**
Φ (year) γ(tide) p(day + c)	3736.37	162.67	0.00
Φ (year) γ(tide) p(c)	3925.76	165.83	189.39
Φ (year) γ(year) p(tide + c)	3980.93	166.70	244.56
Φ (year) γ(tide) p(year + c)	4019.41	166.42	283.04
Φ (year) γ(tide) p(.)	5266.02	178.69	1529.64
Φ (year) γ(tide) p(day)	5271.98	176.91	1535.61
Φ (year) γ(year) p(tide)	5380.31	176.99	1643.94
Φ (year) γ(tide) p(year)	5398.23	179.49	1661.86

Table A.2. Parameters, priors, and posterior estimates (posterior mean, standard deviation, upper and lower 95% Bayesian credible intervals, and effective sample size (ESS) for each parameter), from (A) top-ranked robust-design Cormack-Jolly-Seber model and from (B) modified top model (last primary period omitted from the random year effect on φ) used for simulations. U denotes a uniform prior distribution and N denotes a normal prior distribution. Analysis was conducted on capture-recapture data collected during 2010 – 2022 (except for 2020) from two salt marsh tidal creeks adjacent to the Downing Musgrove Causeway in Glynn County, Georgia, USA.

A)

Parameter	Description	Prior	Posterior				
			Mean	SD	Lower	Upper	ESS
p_{mean}	Mean daily capture probability	U(0,1)	0.64	0.04	0.55	0.73	13,972
p_{SD}	Standard deviation of daily variation in p	U(0,5)	0.84	0.20	0.50	1.28	33,690
β_c	Effect of previous capture in the same primary period on p	N(0,0.37)	-4.40	0.32	-5.06	-3.79	147,798
γ_{mean}	Mean annual apparent availability probability	U(0,1)	0.25	0.03	0.20	0.30	1,133
β_{Tide}	Effect of tide amplitude on γ	N(0,0.37)	-0.42	0.15	-0.73	-0.13	3,491
Φ_{mean}	Mean annual apparent survival probability	U(0,1)	0.80	0.09	0.61	0.95	1,929
Φ_{SD}	Standard deviation of annual variation in Φ	U(0,5)	1.80	0.85	0.60	4.01	10,778

B)

Parameter	Description	Prior	Posterior				
			Mean	SD	Lower	Upper	ESS
p_{mean}	Mean daily capture probability	U(0,1)	0.64	0.04	0.55	0.73	12,719
p_{SD}	Standard deviation of daily variation in p	U(0,5)	0.84	0.20	0.51	1.28	12,425
β_c	Effect of previous capture in the same primary period on p	N(0,0.37)	-4.40	0.32	-5.06	-3.80	21,732

γ_{mean}	Mean annual apparent availability probability	U(0,1)	0.25	0.03	0.20	0.31	8,792
β_{Tide}	Effect of tide amplitude on γ	N(0,0.37)	-0.44	0.15	-0.74	-0.15	177,762
Φ_{mean}	Mean annual apparent survival probability	U(0,1)	0.79	0.08	0.60	0.92	8,405
Φ_{SD}	Standard deviation of annual variation in Φ (Excluding final year)	U(0,5)	1.37	0.81	0.29	3.56	15,625
β_{2022}	Fixed effect to allow final year (2022) to differ from Φ_{mean}	N(0,0.37)	-2.20	0.79	-3.71	-0.58	12,330

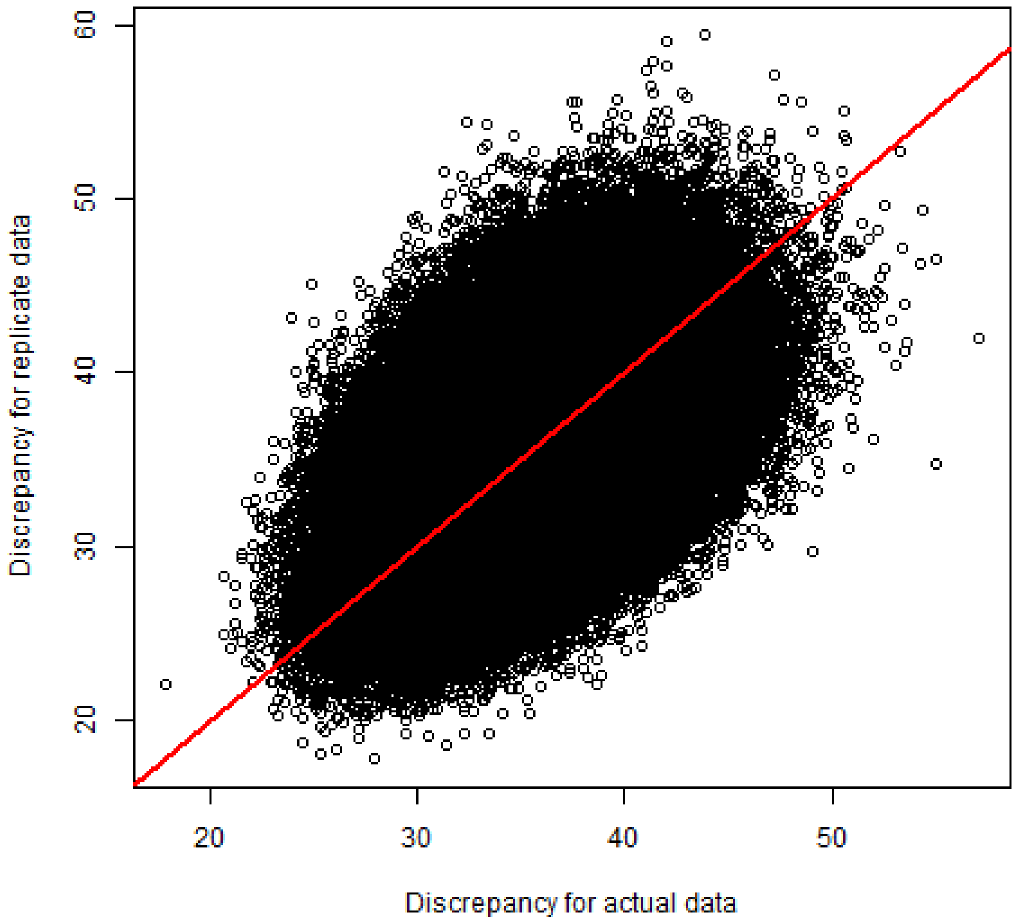


Figure A.1. Goodness-of-fit test for the top-ranked model. Discrepancy values between observed vs expected recaptures and between recaptures predicted by the model vs expected recaptures were calculated using the Freeman–Tukey statistic (Freeman and Tukey 1950).

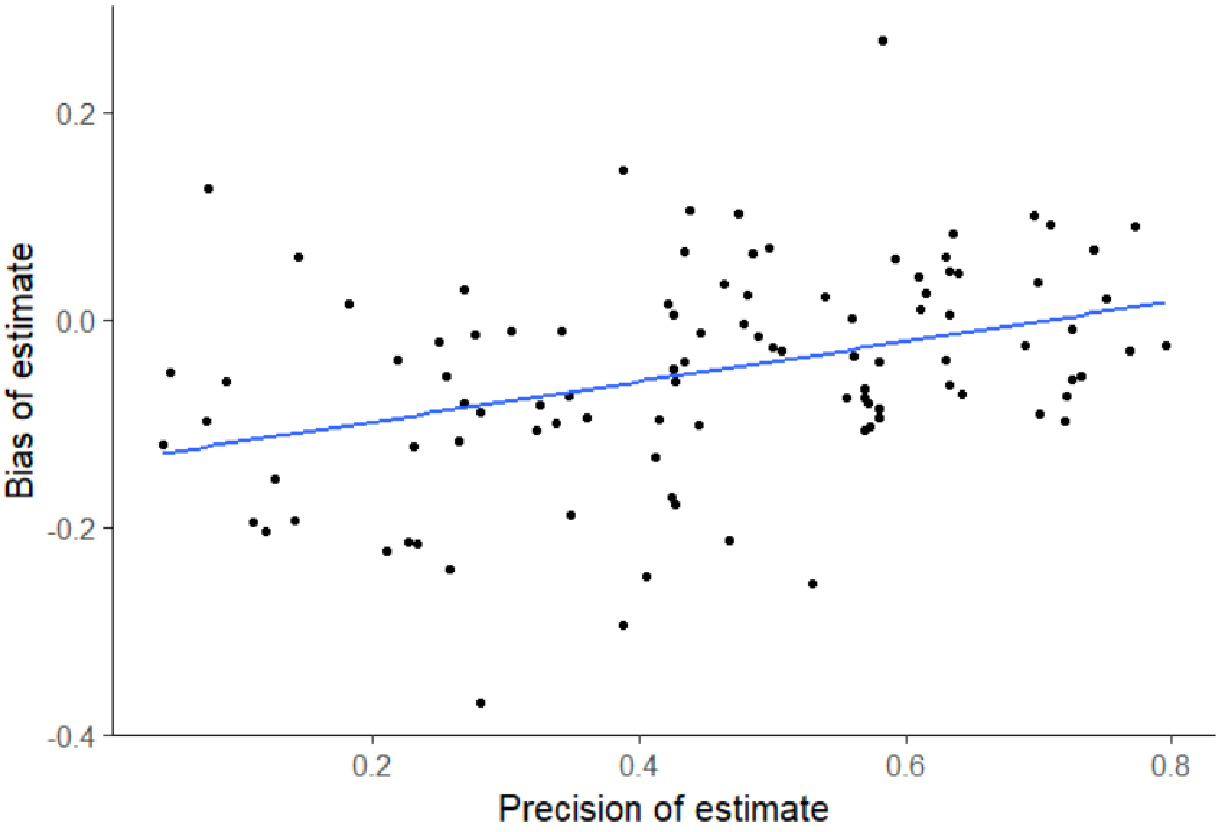


Figure A.2. Relationship between bias and precision of estimated change in mean apparent survival for 100 simulations. Precision is represented by 1 minus the range of the 95% Bayesian credible interval (BCI; 1 – (2.5th–97.5th percentile of the posterior distribution)). Bias was calculated by subtracting the posterior distribution mean from the 0.20 change in survival simulated. A linear trend line (blue) was fit to the data.

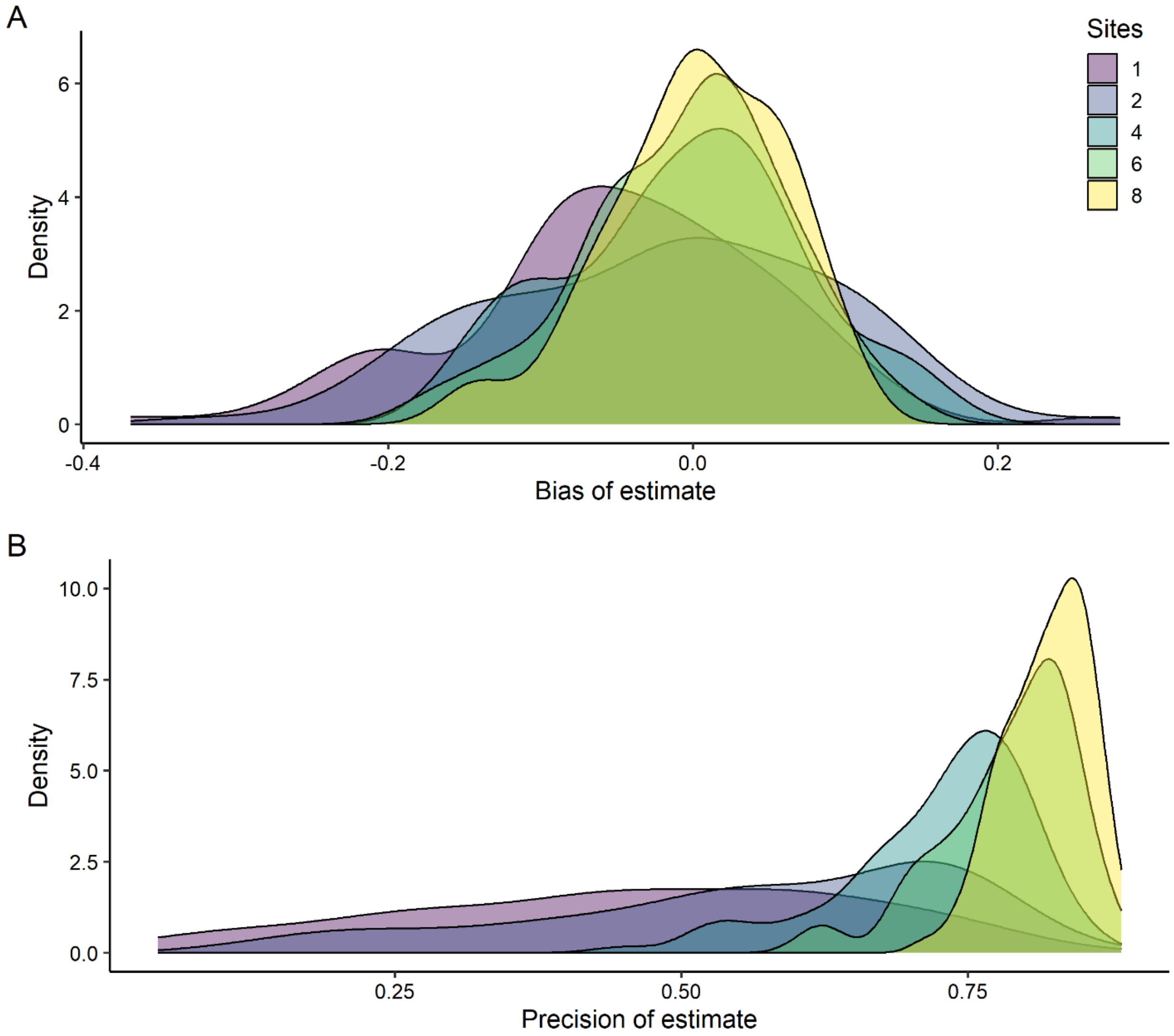


Figure A.3. Density plots of A) bias and B) precision of estimated change in mean apparent survival for 100 simulations based on sampling 1, 2, 4, 6, or 8 sites. Precision is represented by 1 minus the range of the 95% Bayesian credible interval (BCI; 1 – (2.5th–97.5th percentile of the posterior distribution)). Bias was calculated by subtracting the posterior distribution mean from the 0.20 change in survival simulated.

APPENDIX B

SUPPLEMENTAL INFORMATION FOR CHAPTER THREE

Table B.1. Parameters, priors, and posterior estimates (posterior mean, standard deviation, upper and lower 95% Bayesian credible intervals, and effective sample size (ESS) for each parameter) from the top-ranked binomial N-mixture model. **(A)** Estimates using the full dataset (606 surveys of 119 survey areas spanning the Georgia coast during 2021 and 2022) and estimates using a subset of the full dataset (randomly selected single survey to each survey area each year) with three different amounts of survey effort: **(B)** 10 repeated counts per survey, **(C)** 5 repeated counts per survey, or **(D)** 3 repeated counts per survey. U denotes a uniform prior distribution and N denotes a normal prior distribution.

A) All data

			Posterior				
Parameter	**Description**	**Prior**	**Mean**	**SD**	**Lower BCI**	**Upper BCI**	**ESS**
p_{mean}	Mean detection probability	U(0,1)	0.19	0.01	0.17	0.20	391989
β_{wave}	Coefficient for perceived effect of waves on *p*	N(0,0.37)	-0.17	0.05	-0.26	-0.08	4372
β_{glare}	Coefficient for perceived effect of glare on *p*	N(0,0.37)	0.16	0.04	0.09	0.23	11363
λ_{mean}	Mean expected abundance	U(0,100)	0.34	0.12	0.16	0.62	18473
λ_{SD}	Standard deviation random site effect on abundance	U(0,5)	2.19	0.31	1.67	2.88	11014
β_{area}	Coefficient for survey area size effect on λ	N(0,0.01)	1.17	0.15	0.88	1.47	25485
$\beta_{salinity}$	Coefficient for salinity effect on λ	N(0,0.01)	1.02	0.37	0.33	1.78	2101
β_{nest}	Coefficient for distance from potential nesting habitat effect on λ	N(0,0.01)	0.95	0.14	0.68	1.23	50387
β_{DOY}	Coefficient for day of year main effect on λ	N(0,0.01)	0.07	0.05	-0.03	0.16	74428
β_{DOY^2}	Coefficient for day of year quadratic term effect on λ	N(0,0.01)	-0.33	0.05	-0.43	-0.23	20147

B) 1 survey and 10 scans

			Posterior				
Parameter	**Description**	**Prior**	**Mean**	**SD**	**Lower BCI**	**Upper BCI**	**ESS**
p_{mean}	Mean detection probability	U(0,1)	0.21	0.02	0.17	0.25	15915
β_{wave}	Coefficient for perceived effect of waves on p	N(0,0.37)	-0.31	0.12	-0.55	-0.09	2543
β_{glare}	Coefficient for perceived effect of glare on p	N(0,0.37)	0.25	0.08	0.09	0.41	8870
λ_{mean}	Mean expected abundance	U(0,100)	0.26	0.12	0.08	0.56	2140
λ_{SD}	Standard deviation random site effect on abundance	U(0,5)	2.49	0.44	1.76	3.51	1256
β_{area}	Coefficient for survey area size effect on λ	N(0,0.01)	1.05	0.25	0.58	1.56	2592
$\beta_{salinity}$	Coefficient for salinity effect on λ	N(0,0.01)	1.03	0.46	0.20	2.01	5336
β_{nest}	Coefficient for distance from potential nesting habitat effect on λ	N(0,0.01)	1.52	0.34	0.88	2.21	5300
β_{DOY}	Coefficient for day of year main effect on λ	N(0,0.01)	0.25	0.11	0.03	0.47	16077
β_{DOY^2}	Coefficient for day of year quadratic term effect on λ	N(0,0.01)	-0.28	0.12	-0.51	-0.05	8873

C) 1 survey and 5 scans

			Posterior				
Parameter	**Description**	**Prior**	**Mean**	**SD**	**Lower BCI**	**Upper BCI**	**ESS**
p_{mean}	Mean detection probability	U(0,1)	0.26	0.03	0.19	0.32	544
β_{wave}	Coefficient for perceived effect of waves on p	N(0,0.37)	-0.48	0.17	-0.81	-0.16	1877
β_{glare}	Coefficient for perceived effect of glare on p	N(0,0.37)	0.27	0.10	0.07	0.47	1112
λ_{mean}	Mean expected abundance	U(0,100)	0.28	0.13	0.09	0.59	6630
λ_{SD}	Standard deviation random site effect on abundance	U(0,5)	2.22	0.41	1.55	3.14	32012
β_{area}	Coefficient for survey area size effect on λ	N(0,0.01)	0.70	0.26	0.20	1.23	6258
$\beta_{salinity}$	Coefficient for salinity effect on λ	N(0,0.01)	0.92	0.42	0.14	1.82	13276
β_{nest}	Coefficient for distance from potential nesting habitat effect on λ	N(0,0.01)	1.19	0.33	0.57	1.87	27899
β_{DOY}	Coefficient for day of year main effect on λ	N(0,0.01)	0.30	0.14	0.02	0.58	10667
β_{DOY^2}	Coefficient for day of year quadratic term effect on λ	N(0,0.01)	-0.30	0.14	-0.59	-0.02	39748

D) 1 survey and 3 scans

			Posterior				
Parameter	**Description**	**Prior**	**Mean**	**SD**	**Lower BCI**	**Upper BCI**	**ESS**
p_{mean}	Mean detection probability	U(0,1)	0.32	0.06	0.21	0.43	4910
β_{wave}	Coefficient for perceived effect of waves on *p*	N(0,0.37)	-0.28	0.23	-0.74	0.17	9898
β_{glare}	Coefficient for perceived effect of glare on *p*	N(0,0.37)	0.30	0.14	0.02	0.58	45837
λ_{mean}	Mean expected abundance	U(0,100)	0.22	0.12	0.06	0.51	7231
λ_{SD}	Standard deviation random site effect on abundance	U(0,5)	2.45	0.48	1.66	3.54	4753
β_{area}	Coefficient for survey area size effect on λ	N(0,0.01)	0.68	0.30	0.11	1.28	118839
$\beta_{salinity}$	Coefficient for salinity effect on λ	N(0,0.01)	0.90	0.47	0.04	1.90	38931
β_{nest}	Coefficient for distance from potential nesting habitat effect on λ	N(0,0.01)	1.27	0.38	0.55	2.06	157016
β_{DOY}	Coefficient for day of year main effect on λ	N(0,0.01)	0.35	0.17	0.02	0.68	146811
$\beta_{DOY^\wedge 2}$	Coefficient for day of year quadratic term effect on λ	N(0,0.01)	-0.43	0.18	-0.79	-0.09	22030

Table B.2. Results of indicator variable (IV) model selection for candidate binomial N-mixture models of diamond-backed terrapin mean expected relative abundance (λ) and capture probability (p). A random site effect (site) and fixed effect of survey area size (area) were included in all models of λ. Additional fixed effects of known crabbing activity (crab), proportion shoreline armoring (armor), distance to potential nesting habitat (nesting), salinity, and day of year (DOY; main effect and quadratic term) were considered with IV selection to explain abundance. Fixed effects of average wind speed (wind), cloud cover, perceived effect of waves (waves), and perceived effect of glare (glare) were considered with IV selection to explain detection. Model weights were calculated as the proportion of MCMC iterations in which each covariate combination appeared and re-distributed to compare only models of interest (i.e., eliminating models with quadratic term but no associated main effect and all models with two different measurements of the same effect: both wind speed and wave effects or both cloud cover and glare effects), yielding 432 models. Only models with weight ≥ 0.01 (after rounding) are shown. Data were collected 4 April – 13 July 2021 and 21 April – 13 July 2022 within 119 total survey areas spanning the Georgia coast.

Model description	**Model weight**
λ(site + area + nesting + salinity + DOY + DOY^2) p(waves + glare)	0.45
λ(site + area + nesting + DOY + DOY^2) p(waves + glare)	0.18
λ(site + area + nesting + salinity + DOY + DOY^2) p(wind + glare)	0.09
λ(site + area + nesting + salinity + crab + DOY + DOY2) p(waves + glare)	0.09
λ(site + area + nesting + crab + DOY + DOY2) p(waves + glare)	0.06
λ(site + area + nesting + DOY + DOY^2) p(wind + glare)	0.05
λ(site + area + nesting + salinity + DOY + DOY^2) p(glare)	0.03
λ(site + area + nesting + salinity + crab + DOY + DOY2) p(wind + glare)	0.02
λ(site + area + nesting + salinity + armor + DOY + DOY2) p(waves + glare)	0.01
λ(site + area + nesting + DOY + DOY^2) p(glare)	0.01
λ(site + area + nesting + crab + DOY + DOY2) p(wind + glare)	0.01

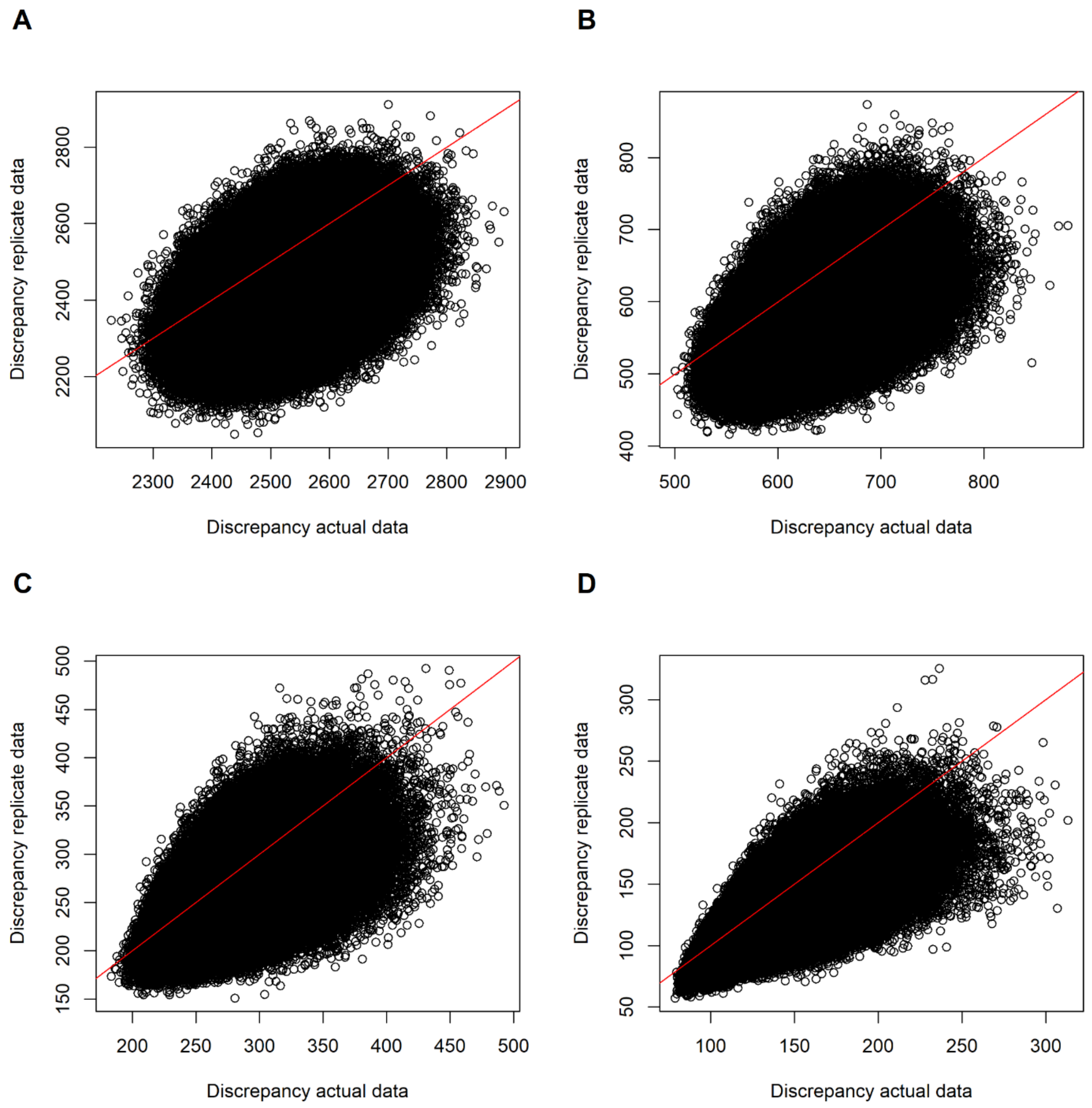


Figure B.1. Posterior predictive check for goodness-of-fit using Chi-square discrepancy for the top-ranked binomial N-mixture model **(A)** using the full dataset (606 surveys of 119 survey areas spanning the Georgia coast during 2021 and 2022) and using a subset of the full dataset (randomly selected single survey to each survey area each year) with three different amounts of survey effort: **(B)** Ten repeated counts per survey, **(C)** Five repeated counts per survey, or **(D)** Three repeated counts per survey.

APPENDIX C

SUPPLEMENTAL INFORMATION FOR CHAPTER FOUR

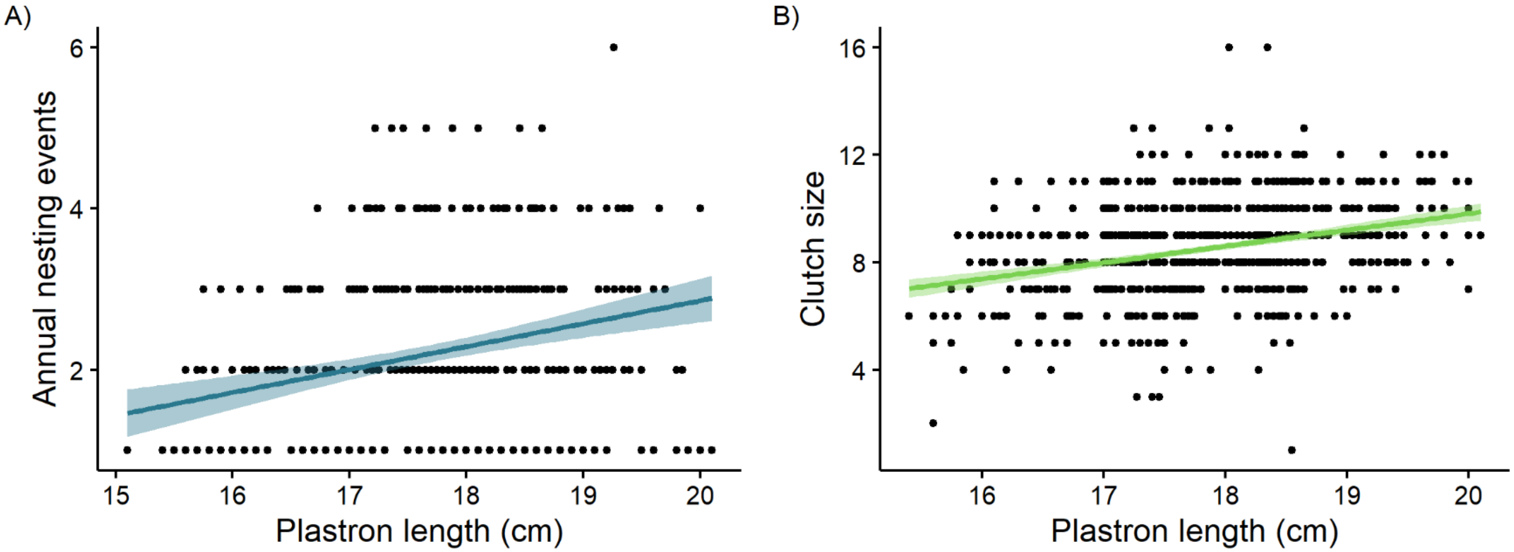


Figure C.1. Linear models of A) annual nesting events vs. plastron length and B) clutch size vs. plastron length for diamond-backed terrapins at the Terrapin Point Golf Course on Skidaway Island, Georgia. Solid lines indicate estimates and shaded areas represent 95% confidence intervals. The fitted regression model for annual nesting events was y = -2.84 + 0.28x ($F_{1,401}$ = 27.62, $p < 0.001$) and for clutch size was y = -2.32 + 0.61x ($F_{1,756}$ = 83.67, $p < 0.001$). Plastron length accounted for 6.21% of variation in annual nesting events and 9.85% of variation in clutch size.

APPENDIX D

SUPPLEMENTAL INFORMATION FOR CHAPTER FIVE

Additional field methods

We measured weather variables that we expected to affect availability (i.e., temperature and relative humidity) using a kestrel meter. We used our measurements of temperature and relative humidity to calculate vapor pressure deficit. Additionally, we estimated ground visibility on a six-category scale (100%, 70-75%, 50-75%, 25-50%, 10-25%, or 0% obstructed by vegetation) to use as a variable to explain detection. We also used geographic information system (GIS) software and spatial layers to measure additional variables that we expected to have important effects on availability (precipitation the day of sampling and within 5-days prior to sampling) and abundance (10-year active season climate Vapor Pressure Deficit). We accessed daily rainfall estimates from NOAA and calculated 1-day and 5-day rainfall totals leading up to the night of each sample. To obtain vapor pressure deficit climate data, we used Daymet 2011-2020 data (Thornton et al. 2022; https://doi.org/10.3334/ORNLDAAC/2131) to generate estimates of mean temperature and relative humidity during April – October. Using the Daymet mean daily temperature and relative humidity, we calculated mean daily vapor pressure deficit for each site.

Extended Methods - Analyses

For all models we did the following:

We standardized all covariates (except for one binary covariate indicating whether or not management was applied) prior to analysis. We fit the model using Markov chain Monte Carlo (MCMC) sampling in JAGS (Plummer 2003) called via the jagsUI package (Kellner 2021) in program R (4.2.1; R Core Team 2022). We used three MCMC chains with 500,000 iterations each and burn-in of 100,000 and ran additional iterations if necessary. We assessed chain convergence visually using trace plots and with Gelman and Rubin (1992) diagnostic values (i.e., R-hat < 1.1).

Coweeta: years 1 – 3

We fit a binomial N-mixture model (Royle 2004) using a Bayesian framework to estimate capture probability (p) and expected abundance (λ). Here, capture probability was the product of availability and detection because we were unable to separately account for availability with this dataset. We modeled abundance (N) at site (i) during year (t) as a Poisson outcome with mean equal to expected abundance ($\lambda_{i,t}$) and we modeled each count (y) at site (i) during survey (j) (nested within year t) as binomial outcome drawn from ($N_{i,\,t}$) with detection probability ($p_{i,j}$) as follows:

$$N_{i,t} \sim Poisson(\lambda_{i,t})$$

$$y_{i,j(t)} \sim Binomial(N_{i,t}, p_{i,j})$$

We constrained expected abundance to be a log-linear function of management, a binary covariate indicating if the site was managed (1) or unmanaged (0). To allow abundance to vary annually, we also modeled a year-specific intercept (μ_t). Thus, our model of expected abundance was:

$$log(\,\lambda_{i,t}) = \mu_t + \beta_{Management} \times Management_{i,t}$$

We constrained capture probability to be a logit-linear function of factors expected to affect detection if available for capture – ground visibility (vis) – and factors expected to affect availability – air temperature (temp), vapor pressure deficit (vpd), precipitation during one day prior to sampling (rain1), and precipitation during the 5 days prior to sampling (rain5).

Additionally, because preliminary analysis indicated a lack of model fit, we included a random effect of plot-survey ($\varepsilon_{i,j}$) to account for random variation in detection across sites and visits. Thus, our model of detection probability was:

$$logit(p_{i,j}) = \mu + \beta_{vis} \times vis_{i,j} + \beta_{temp} \times temp_{i,j} + \beta_{vpd} \times vpd_{i,j} + \beta_{rain1} \times rain1_{i,j} + \beta_{rain5} \times rain5_{i,j} + \varepsilon_{i,j}$$

Coweeta: years 4 – 6

We fit a three-level hierarchical removal model (Chandler et al. 2011, Kery and Royle 2020) using a Bayesian framework to estimate capture probability (*p*), availability (*γ*), and expected abundance (*λ*). We modeled "super population" abundance (*M*; including both above and below ground individuals) at site (*i*) during year (*t*) as a Poisson outcome with mean equal to expected abundance ($\lambda_{i,t}$). We modeled the available (i.e., above ground) population (*N*) at site (*i*) during survey (*j*) (nested within yead *t*) as a binomial outcome drawn from ($M_{i,t}$) with probability ($\gamma_{i,j}$). We modeled the detection process as multinomial, where counts were drawn from ($N_{i,j(t)}$) with multinomial probabilities ($\boldsymbol{\pi}_{i,j}$) that were a function of capture probability ($p_{i,j}$) and number of removal passes.

$$M_{i,t} \sim Poisson(\lambda_{i,t})$$

$$N_{i,j(t)} \sim Binomial(M_{i,t}, \gamma_{i,j})$$

$$y_{i,j(t)} \sim Multinomial(N_{i,j(t)}, \boldsymbol{\pi}_{i,j})$$

$$\boldsymbol{\pi}_{i,j} = \{p_{i,j}, (1 - p_{i,j})p_{i,j}, (1 - p_{i,j})^2 p_{i,j}, (1 - p_{i,j})^3 p_{i,j}, (1 - p_{i,j})^4\}$$

Identical to the model of years 1 – 3, we constrained expected abundance to be a log-linear function of management, a binary covariate indicating if the site was managed (1) or unmanaged (0). To allow abundance to vary annually, we also modeled a year-specific intercept (μ_t). Thus, our model of expected abundance was:

$$log(\lambda_{i,t}) = \mu_t + \beta_{Management} \times Management_{i,t}$$

We constrained availability to be a logit-linear function of air temperature (temp), vapor pressure deficit (vpd), precipitation during one day prior to sampling (rain1), and precipitation during the 5 days prior to sampling (rain5):

$$logit(\gamma_{i,j}) = \mu + \beta_{temp} \times temp_{i,j} + \beta_{vpd} \times vpd_{i,j} + \beta_{rain1} \times rain1_{i,j} + \beta_{rain5} \times rain5_{i,j}$$

Finally, we constrained capture probability to be a logit-linear function of ground visibility (vis):

$$logit(p_{i,j}) = \mu + \beta_{vis} \times vis_{i,j}$$

Broadscale southern Appalachian Mountain data

We fit the same three-level hierarchical removal model as we did to Coweeta years 4 – 6 data, with a few differences. Because each site was sampled 1 – 2 times in the same year, modeled a single super population size per site (M_i) and we did not include a year-specific intercept in our model of expected abundance (λ_i). However, because this study included many sites over a broad spatial extent, we included a random site effect on abundance, which we specified as a site-specific intercept, to account for inherent differences among sites not otherwise captured by plot-level covariates. Additionally, the greater number of sites within this dataset allowed us to include an effect of 10-year mean active season vapor pressure deficit, to explain differences in abundance among plots. Finally, we also included a random effect of plot-survey ($\varepsilon_{i,j}$) to account for random variation in detection across sites and visits because preliminary analysis indicated a lack of model fit (the model of Coweeta years 4 – 6 data appeared to adequately fit the data without this additional random effect).

Integrated model: ***Coweeta years 1 – 3 and years 4 – 6***

We integrated the two above-described Coweeta sub models by linking multiple parameters. Both models shared a common intercept, effect of visibility, and random effect of plot-survey on detection probability. Additionally, both models shared a common intercept and temperature, precipitation, and vapor pressure deficit effects on availability. The effect of management on expected abundance was also shared between the two sub models.

Integrated model: ***Coweeta years 1 – 3 and Broadscale southern Appalachian Mountain data***

We integrated the sub model for Coweeta years 1 – 3 and the sub model for the Broadscale southern Appalachian Mountain dataset by linking multiple parameters, similar to the integrated model for Coweeta years 1 – 3 and years 4 – 6. As before, both models shared a common intercept, effect of visibility, and random effect of plot-survey on detection probability. Additionally, both models shared a common intercept and temperature, precipitation, and vapor pressure deficit effects on availability. However, unlike the Coweeta years 1 – 3 and years 4 – 6 integrated model, the effect of management on expected abundance was only included in the Coweeta sub model. Additionally, both models shared an effect of the 10-year mean active season vapor pressure deficit on expected abundance.

Integrated model: ***Coweeta years 1 – 3, Coweeta years 4 – 6, and Broadscale southern Appalachian Mountain data***

Finally, we integrated all three datasets and sub models. All models shared a common intercept, effect of visibility, and random effect of plot-survey on detection probability. Additionally, all models shared a common intercept and temperature, precipitation, and vapor pressure deficit effects on availability. All models also shared an effect of the 10-year mean active season vapor pressure deficit on expected abundance. Lastly, Coweeta years 1 – 3 and years 4 – 6 sub models shared an effect of management on expected abundance.