

Methods in Ecology and

A flexible Bayesian approach for estimating survival probabilities from age-at-harvest data

Journal:	<i>Methods in Ecology and Evolution</i>
Manuscript ID	MEE-22-08-656
Wiley - Manuscript Type:	Research Article
Date Submitted by the Author:	21-Aug-2022
Complete List of Authors:	Skelly, Brett; West Virginia Division of Natural Resources; West Virginia University, Clipp, Hannah; West Virginia University, Division of Forestry and Natural Resources Landry, Stephanie; West Virginia University, Division of Forestry and Natural Resources, Wildlife and Fisheries Resources Program; Utah State University Rogers, Rich; West Virginia Division of Natural Resources Phelps, Quinton; Missouri State University Anderson, James; Clemson University, James C. Kennedy Waterfowl and Wetlands Conservation Center Rota, Christopher; West Virginia University, Division of Forestry and Natural Resources, Wildlife and Fisheries Resources Program
Keywords:	age-at-harvest, age-specific survival, age-structured populations, Bayesian analysis, fisheries populations, sampling scheme, survival, wildlife populations
Abstract:	<p>1. Understanding survival rates is critical for the sustainable harvest of wildlife and fisheries populations. In particular, age- and stage class-specific survival rates, often derived from harvest data, are needed to inform a suite of population models used to estimate abundance and track population trends over time. However, current established techniques for estimating survival rates using age-at-harvest methods require restrictive assumptions about model parameters or incorporate potentially unknown parameters that must nevertheless be specified as fixed and known quantities within the model.</p> <p>2. Using a Bayesian approach, we developed a flexible age-at-harvest model that incorporates either age- or stage-specific structured populations, while accounting for uncertainty in age structure, population growth rates, and capture probability. In our model, survival rates can vary by age or stage class, as well as by environmental covariates, and both population growth rates and capture</p>

	<p>probability can be specified as fixed and known, if auxiliary data are available, or these parameters can be specified as informative priors, allowing for the incorporation of expert opinion.</p> <p>We evaluated our model with simulations and empirical data from harvested bobcats (*Lynx rufus*) and American paddlefish (*Polyodon spathula*).</p> <p>3. Simulations of age-at-harvest data produced unbiased estimates of survival rate when population growth rates and capture probabilities were centered on the data generating parameter, and even when capture probabilities were not centered on the data generating parameter.</p> <p>Using empirical harvest data, our proposed age-at-harvest model produced congruent species survival rates to those reported in the literature within the same geographic regions.</p> <p>4. Age- and stage class-specific survival estimates are necessary for certain widely used population models, but collecting these estimates in traditional ways, such as with known-fate or capture-mark recapture methods, tends to be cost prohibitive.</p> <p>In addition, current age-at-harvest methods have restrictive assumptions or require auxiliary data to obtain realistic estimates of age or stage class survival rates.</p> <p>Thus, the model outlined here will be transformative for practitioners who collect age-at-harvest data and need to estimate age- or stage-specific survival rates, while realistically accounting for uncertainty in model parameters that cannot be achieved using established techniques.</p>

1 RH: *Bayesian survival model for age-at-harvest data*

2 **A flexible Bayesian approach for estimating survival probabilities**
3 **from age-at-harvest data**

4 Brett P. Skelly^{1,2*}, Hannah L. Clipp¹, Stephanie M. Landry^{1,3}, Rich Rogers⁴, Quinton Phelps⁵, James T.
5 Anderson⁶, and Christopher T. Rota¹

6 ¹*Division of Forestry and Natural Resources, West Virginia University, Morgantown, WV, USA;* ²*West*
7 *Virginia Division of Natural Resources, Elkins, WV, USA;* ³*Department of Wildland Resources, Utah State*
8 *University, Logan, UT, USA;* ⁴*West Virginia Division of Natural Resources, Romney, WV, USA;* ⁵*Department*
9 *of Biology, Missouri State University, Springfield, MO, USA;* ⁶*James C. Kennedy Waterfowl and Wetlands*
10 *Conservation Center, Belle W. Baruch Institute of Coastal Ecology and Forest Science, Georgetown, SC, USA*

11 *Correspondence author: Brett P. Skelly; P.O. Box 67, Elkins, WV, USA; brett.p.skelly@wv.gov

Abstract

1. Understanding survival probabilities is critical for the sustainable harvest of wildlife and fisheries populations. In particular, age- and stage class-specific survival probabilities, often derived from harvest data, are needed to inform a suite of population models used to estimate abundance and track population trends over time. However, current established techniques for estimating survival probabilities using age-at-harvest methods require restrictive assumptions about model parameters or incorporate potentially unknown parameters that must nevertheless be specified as fixed and known quantities within the model.
2. Using a Bayesian approach, we developed a flexible age-at-harvest model that incorporates either age- or stage-structured populations, while accounting for uncertainty in age structure, population growth rates, and relative capture probability. In our model, survival probabilities can vary by age or stage class, as well as by environmental covariates, and both population growth rates and capture probability can be specified as fixed and known, if auxiliary data are available, or these parameters can be specified as informative priors, allowing for the incorporation of expert opinion. We evaluated our model with simulations and empirical data from harvested bobcats (*Lynx rufus*) and American paddlefish (*Polyodon spathula*).
3. Simulations of age-at-harvest data produced unbiased estimates of survival probability when population growth rates and capture probabilities were centered on the data-generating parameter. Even when capture probabilities were not centered on the data-generating parameter and the simulated age distribution was different from the stable age distribution, we still obtained unbiased estimates of survival probability, although departure from the stable age distribution increased uncertainty in those estimates. In our simulations, there was consistently bias in estimates of survival probability when the prior distribution for population growth rate was not centered on the data-generating value. When incorporating empirical harvest data, our proposed age-at-harvest model produced congruent species survival probabilities to those reported in the literature within the same geographic regions.
4. Age- and stage class-specific survival estimates are necessary for certain widely-used population models, but collecting these estimates in traditional ways, such as with known-fate or capture-mark recapture methods, tends to be cost prohibitive. In addition, current age-at-harvest methods have restrictive assumptions or require auxiliary data to obtain realistic estimates of age or stage class survival probabilities. Thus, the model outlined here will be transformative for practitioners who collect age-at-harvest data and need to estimate age- or stage-specific survival probabilities, while realistically

accounting for uncertainty in model parameters that cannot be achieved using established techniques.

Keywords

age-at-harvest, age-specific survival, age-structured populations, Bayesian analysis, fisheries populations, sampling scheme, survival, wildlife populations

Introduction

Estimating species-specific survival is an important aspect of wildlife and fisheries management and conservation, with diverse applications (e.g. Skalski et al., 1998; Buckland et al., 2004; Rota et al., 2014). Estimates of survival probability can be used to better understand animal populations and to aid in conservation planning. For example, estimating survival probabilities is a critical component of integrated population monitoring and is necessary to perform a population viability analysis, a tool frequently used to aid in conservation decision-making for wildlife populations (Beissinger & McCullough, 2002). By incorporating survival estimates into population models (e.g. Crouse et al., 1987; Lehman et al., 2018), we can identify the relative significance of different ages or stages on overall population dynamics, identify the greatest threats to an at-risk population, and prioritize science-based conservation actions. Estimates of survival probability are also instrumental in integrated population models, where numerous independent data sources are used to jointly estimate demographic parameters that could not be estimated from any single dataset (Schaub & Kéry, 2021).

Given the importance of survival probability in wildlife and fisheries sciences, there is plentiful literature describing statistical methods for estimating this demographic parameter. Perhaps the simplest approaches are constructing life tables (Caughley, 1977; Roseberry & Woolf, 1991) and calculating return rates. More advanced mark-recapture methods are based on live encounters (e.g. Cormack-Jolly-Seber models), dead recoveries (e.g. band recovery models), or both (e.g. joint capture-recapture models; Barker, 1997; Williams et al., 2002). Known-fate models, wherein an individual is marked with a tracking device and regularly re-located to determine its status (i.e. alive or dead), permit estimates of survival probability with classical methods such as Kaplan-Meier, proportional hazard, or parametric survival models (Efron, 1988).

Although mark-recapture and known-fate methods for estimating survival probability are well-established, they usually require costly study design, in terms of funding, time, and effort (e.g. the expenses associated with capturing, marking, and monitoring individuals over long periods of time). An alternative approach to estimating survival probability is to collect age-at-harvest data (frequently called catch-at-age data in fisheries

science). For example, many wildlife management agencies in the United States record the ages of harvested white-tailed deer (*Odocoileus virginianus*) at hunter check stations, while many fisheries managers collect large samples of fish, which they sort by age. Given the high cost of mark-recapture or known-fate survival studies, data from annual harvest efforts and routine sampling often represent a substantial proportion of data for a focal species. Furthermore, even for well-studied species, age-at-harvest data are usually collected much more routinely than formally designed survival studies, and can be used to regularly update knowledge regarding current demographic trends. Thus, age-at-harvest data can potentially inform managers about species of interest by allowing them to obtain close to real-time estimates of survival and mortality rates in the absence of more expensive sources of data (e.g. mark-recapture studies).

As with mark-recapture and known-fate methods, there is a wide scope of literature describing methods for estimating survival probability from age-at-harvest data. Foundational approaches to estimating survival probability from age-at-harvest data were described by Chapman and Robson (1960). Additionally, regressing log count in each age class against age is a frequently used method to estimate survival probability. These methods assume constant survival probability among ages, no trends in recruitment, and equal detectability among ages (Smith et al., 2012). However, assuming equal detectability among ages often requires discarding data when age classes are under-represented (Allen & Hightower, 2010; Smith et al., 2012). Many advancements address these assumptions to various extents. For example, Thorson and Prager (2011) relax the assumption of constant mortality and detectability; Schnute and Haigh (2007) relax the assumptions of constant detectability and allow recruitment anomalies; and Udevitz & Ballachey (1998) and Udevitz & Gogan (2012) allow variation in survival probability among ages and population growth rates among years. While these methods represent important advancements, they tend to relax only one or two assumptions at a time (e.g. Udevitz & Ballachey (1998) do not allow unequal capture probability among ages) and / or require additional parameters to be included as fixed and known (e.g. Thorson & Prager (2011) require assumptions of parameters of a Lorenzen mortality-at-age model). These methods may therefore still require restrictive assumptions, and may not fully account for uncertainty in parameters that are assumed fixed and known. Alternatively, many approaches incorporate auxiliary data when estimating survival probability from age-at-harvest data (Broms et al., 2010; Conn et al., 2008; Gove et al., 2002). Such approaches no longer require many of the restrictive assumptions of basic age-at-harvest models, and uncertainty in auxiliary data can be appropriately propagated through construction of joint likelihoods that simultaneously account for multiple demographic processes. However, appropriate data are likely to be available for only the most well-studied species, and integrated population models often require similar effort as mark-recapture or known-fate methods for estimating survival probability.

In this paper, we aim to take a middle approach between age-at-harvest models with restrictive assumptions

and costly integrated population models. We seek to develop a method for estimating survival probability from age-at-harvest data when auxiliary data are not available. However, we also aim to relax the restrictive assumptions required to do so and to simultaneously allow a full accounting of uncertainty in all model parameters. In our model, we account for the effects of both population growth rate and unequal capture probability among age or stage classes, while accommodating either age- or stage-structured populations. Specifically, by adopting a Bayesian approach, we allow flexible definitions of the likelihood and are able to account for uncertainty in age structure, population growth rate, and capture probability by assigning prior distributions to those parameters. We further allow survival probability to vary by age or stage class and as a function of environmental variables. Thus, this model allows for substantial flexibility, such that users can choose to adhere to standard assumptions by fixing those parameters in the model or choose to relax assumptions and assign prior distributions for parameters for which they have data. In this paper, we describe applicable sampling schemes, present our statistical model, evaluate theoretical properties with a simulation study, and apply our proposed approach to two real-world case studies by estimating survival probabilities from bobcat (*Lynx rufus*) and American paddlefish (*Polyodon spathula*) harvest data.

Methods

Sampling Scheme

Our method requires counts of individuals segregated by age or stage class. We envision such data may arise from numerous sampling schemes. For example, biologists at hunter check stations for harvested deer may record the number of fawns (0.5 years old), yearlings (1.5 years old), and adults (2.5+ years old). If there is interest in allowing survival probability as a function of covariates, count data may be segregated among groups. For example, biologists could segregate counts of harvested deer by sex and / or geographic location.

While auxiliary information on age structure, recruitment, and / or population growth rates may help improve precision and accuracy of survival estimates, such data are not required. Depending on the amount of information, these parameters can be assumed fixed and known, allowed to vary within bounds assigned by the researcher, or assigned informative prior distributions.

Statistical Model

The basic structure from our model is derived from Udevitz & Ballachey (1998), though we modify nearly every aspect of their model to relax assumptions and allow for greater generality. We assume the number of

individuals within each age or stage class, denoted \mathbf{y}_i , is a multinomial random variable:

$$\mathbf{y}_i \sim \text{Multinomial}(\boldsymbol{\theta}_i, H_i),$$

where H_i is the total number of individuals harvested within group i (which may represent sex, time period, geographical region, etc.) and $\boldsymbol{\theta}_i$ is the probability a harvested individual within group i belongs to each age or stage class. We further model $\boldsymbol{\theta}_i$ as a function of population growth rate, age- or stage-specific survival probabilities, capture probabilities, and age or stage distribution, which we describe in detail below.

We can derive an expression for $\boldsymbol{\theta}_i$ by considering the age or stage structure of a population and the relative probability of harvesting and reporting an individual within each age or stage class. Allow \mathbf{A}_i to represent any age- or stage-structured (i.e., Leslie or Lefkovitch) matrix projection model with S age or stage classes (Caswell, 2001), and allow $s \in 1, 2, \dots, S$ to index each discrete age or stage class. Within projection matrix \mathbf{A}_i , parameters p_{is} describe the probability an individual within stage class s survives 1 time step and are the principle estimand of our model. Note that unlike Udevitz & Ballachey (1998), our formulation does not require all individuals in age or stage class S to die. The number of animals in any age or stage class $s > 1$ at time t , denoted n_{ist} , can be calculated as:

$$n_{ist} = N_{i,t-1} \mathbf{A}_{i[s,]} \mathbf{C}_i,$$

where $N_{i,t-1}$ is the total abundance at time $t - 1$, $\mathbf{A}_{i[s,]}$ represents row s of projection matrix \mathbf{A}_i , and \mathbf{C}_i is a conformable vector describing the age or stage distribution of individuals within group i , such that $\sum \mathbf{C}_i = 1$. We define the finite rate of increase as $\lambda_{it} = \frac{N_{it}}{N_{i,t-1}}$ (Caswell, 2001). Without information on age- or stage-specific reproductive rates, we can obtain the number of animals in age or stage class 1 at time t as:

$$\begin{aligned} n_{i1t} &= N_{it} - \sum_{s=2}^S N_{i,t-1} \mathbf{A}_{i[s,]} \mathbf{C}_i \\ &= N_{i,t-1} \left(\lambda_{it} - \sum_{s=2}^S \mathbf{A}_{i[s,]} \mathbf{C}_i \right). \end{aligned}$$

Alternatively, if there is information on reproductive rates, we can obtain the number of animals in age or stage class 1 at time t as:

$$n_{i1t} = N_{i,t-1} \mathbf{A}_{i[1,]} \mathbf{C}_i,$$

where $\mathbf{A}_{i[1,]}$ is calculated as a function of survival and reproduction, depending on whether \mathbf{A}_i is a pre- or post-birth pulse projection matrix (Mills, 2007).

Differential susceptibility to harvest may cause the observed number of animals in each stage class to differ from age or stage distribution \mathbf{C}_i . Therefore, the expected number of animals harvested in age or stage class s at time t is the product of age- or stage-specific relative capture probability α_{is} and age- or stage-specific abundance:

$$E(y_{ist}) = \alpha_{is}n_{ist},$$

and the probability a harvested individual belongs to age or stage class s at time t is

$$\theta_{ist} = \frac{E(y_{ist})}{\sum_{s=1}^S E(y_{ist})}. \quad (1)$$

Note that calculation of θ_i does not depend on knowledge of $N_{i,t-1}$, since this term cancels from the numerator and denominator in equation 1; α_{is} also cancels if it is assumed constant across age or stage classes. Additionally, when α_{is} varies among age or stage classes, only estimates of *relative* capture probability are necessary (i.e. we do not need to estimate *absolute* capture probability).

From this expression for θ_i , there is considerable flexibility in specifying model components. If there is sufficient information, \mathbf{C}_i , λ_{it} , and α_{is} can all be specified as fixed and known. Alternatively, prior distributions can be induced on these variables, with uncertainty quantified via the variance of the prior distribution.

Additional flexibility arises when defining the age or stage distribution \mathbf{C}_i , which may be assumed to follow the stable age or stage distribution if there is no other information available. Caswell (2001) provides convenient expressions for calculating stable age distributions (Caswell, 2001, p. 87) and stable stage distributions (Caswell, 2001, p. 186) as a function of age- or stage-specific survival probabilities and λ_t . These expressions assume an individual can remain in each age or stage class s for a fixed maximum duration d_s . Assuming an animal remains in age or stage class s for up to d_s time steps, and assuming an individual survives an individual time step with probability p_{is} , the probability an individual survives and remains within stage class s is calculated as:

$$z_{is} = \frac{1 - p_{is}^{d_s-1}}{1 - p_{is}^{d_s}} p_{is},$$

The probability an individual graduates to the next age or stage class is calculated as:

$$g_{is} = \frac{p_{is}^{d_s}(1 - p_{is})}{1 - p_{is}^{d_s}}.$$

Note that if $d_s = 1$, the probability that an individual remains in stage class s is 0 and the probability an

individual graduates to the next age or stage class is p_{is} . Using these definitions for z_{is} and g_{is} , we can obtain the stable stage distribution. First, we calculate a vector \mathbf{w}_i that is proportional to the stable age or stage distribution. We fix the first element of \mathbf{w}_i at 1:

$$w_{i1} = 1.$$

Then, for all $s > 1$:

$$w_{is} = \frac{g_{i,s-1}}{\lambda_{it} - z_{is}} w_{i,s-1}.$$

Finally, the stable age or stage distribution is obtained by scaling \mathbf{w}_i :

$$\mathbf{C}_i = \frac{\mathbf{w}_i}{\sum_{s=1}^S w_{is}},$$

where S is the total number of age or stage classes. Note that the expression for calculating the stable age or stage distribution above assumes specific dwell times within age or stage classes. However, there is substantial flexibility in specifying transition probabilities. For instance, dwell times in size-structured models could be a function of survival and growth rates, or dwell times could be assumed to follow a specific distribution.

There is also considerable flexibility in specifying age- and stage-specific survival probabilities. Survival probability can be assumed constant or to vary by age or stage class. Additionally, age- or stage-specific survival probabilities can be modeled as a linear function of covariates associated with group i (e.g. sex, time period, geographical region). A list and depiction of the model components, including inputs, parameters, and derived parameters, can be found in Table 1 and Figure 1.

Finally, if overdispersion is present, the probability of capturing an individual within each age or stage class can be modeled as a Dirichlet random variable:

$$\boldsymbol{\theta}_i \sim \text{Dirichlet}(\boldsymbol{\alpha}_i \circ \mathbf{n}_{it}), \quad (2)$$

where $\boldsymbol{\alpha}_i \circ \mathbf{n}_{it}$ represents element-wise multiplication of the vectors of age- or stage-specific capture probability and (relative) abundance. Note that this multinomial-Dirichlet model is the multivariate equivalent of a beta-binomial model.

Simulation study

We used a simulation study to evaluate theoretical properties of our age-at-harvest model. We were interested in (1) evaluating bias and precision in estimates of survival probability over realistic sample sizes when

prior knowledge of relative capture probabilities and population growth rates were unbiased but included uncertainty; (2) evaluating how estimates of survival probability varied when prior knowledge of relative capture probabilities and population growth rates were increasingly biased; and (3) evaluating how estimates of survival probability varied when the stage distribution departed from the stable stage distribution.

To do this, we simulated a realization of age-at-harvest data with five stage classes. We assumed the probability of surviving to the next time step was constant ($p = 0.8$) across all stages. At the end of each time step, individuals graduated to the next stage class, except for individuals in the oldest stage class, which could survive and remain within the oldest stage class. We assumed no knowledge of reproduction, but instead assumed a constant population growth rate of $\lambda = 1.01$. We assumed individuals in the youngest stage class were half as likely to be captured as individuals in older stage classes. We therefore fixed $\alpha_1 = 0.5$, and fixed $\alpha_s = 1$ for all remaining stage classes > 1 .

We simulated data where stage classes were distributed according to the stable stage distribution, and where the stage distribution departed from the stable stage distribution. For both scenarios, we first calculated the stable stage distribution \mathbf{C} as described above. For simulation scenarios assuming departure from the stable stage distribution, we further simulated a realized stage distribution \mathbf{C}_r as a Dirichlet random variable:

$$\mathbf{C}_r \sim \text{Dirichlet}(\mathbf{C} \times 100)$$

For each of the age distribution scenarios, we evaluated five different sample size scenarios of $n = 50$, $n = 100$, $n = 250$, $n = 500$, and $n = 1,000$. The smaller sample sizes may reflect species with lower harvest rates (e.g. bobcat harvest records), while the larger sample size scenarios may reflect abundant, well-sampled species (e.g. white-tailed deer reported to hunter check stations). We simulated age-at-harvest data by assuming counts of individuals by stage class were a multinomial random variable:

$$y_{sim} \sim \text{Multinomial}(\boldsymbol{\theta}_{sim}, n)$$

where $\boldsymbol{\theta}_{sim}$ is calculated as described above using the population structure and parameter values described in this section.

For all simulated scenarios, we quantified uncertainty in population growth rates and relative capture probability by inducing both a uniform and a normal prior distribution for λ and α_1 . To evaluate the influence of biased prior estimates of relative capture probability and population growth rates on estimates of survival probability, we varied the bounds of the prior distribution around λ and α_1 . We evaluated three different levels of prior bias for each parameter: (1) prior distributions centered on the truth (i.e. unbiased

prior information); (2) prior distributions not centered on the truth but within bounds of the uniform prior distribution, or within 95% of the probability density of the normal prior distribution; and (3) outside the bounds of the uniform prior distribution or outside 95% of the probability density of the normal prior distribution. This led us to assume the prior distributions listed in Table 2. Note that the standard deviation of the normal prior distributions listed in Table 2 is such that 95% of the probability density falls within the limits of the lower and upper bounds of the uniform distribution for each particular scenario. Note also that values of $\alpha > 1$ correspond to *greater* capture probability of stage class 1 relative to other stage classes.

We assumed a uniform(0,1) prior distribution for survival probability over all 2 age distributions \times 5 sample sizes \times 2 prior distributions \times 3 capture probabilities \times 3 population growth rates = 180 different simulation scenarios. We simulated 1,000 datasets for each scenario, recording the posterior mean survival probability from each scenario. We fit models in JAGS (Plummer, 2003) using the jagsUI package (Kellner, 2019) in program R (R Core Team, 2020). We used the autojags function within jagsUI to ensure all models achieved convergence (assuming convergence is achieved when $\hat{R} < 1.1$).

Bobcat survival study

To demonstrate the utility of our method, we used our age-at-harvest model to estimate bobcat survival probability in West Virginia, USA. To obtain updated estimates of bobcat vital rates, the West Virginia Division of Natural Resources collected carcasses from legally harvested bobcats during the 2014 / 2015 and 2015 / 2016 trapping years (Landry, 2017). Bobcats were sexed, and teeth were pulled from carcasses and aged by counting rings of cementum annuli (Crowe, 1972). Our data therefore consisted of counts of bobcat individuals by age, sex, and year collected (Table 3).

We assumed a six stage-class model (Table 3). We assumed individuals remain within the first five stage classes for 1 year ($d_1, d_2, \dots, d_5 = 1$), while individuals in the oldest stage class (aged 5+) can survive and remain within that stage class. We assumed survival probability varied by stage class, but fixed survival probability in stage class 5 (age 4 – 5) equal to stage class 6 to ensure survival probability was identifiable (Conn et al., 2008). We additionally allowed survival probability to vary among sex and year. When survival probability varied by year, we also allowed λ to vary by year. To find a parsimonious model of bobcat survival probability, we evaluated eight models with various combinations of sex, year, and stage-specific survival probability (Table 4). We assumed logistic(0, 1) prior distributions for all survival coefficients.

It is common for trappers to release bobcat kittens, and surveys indicate that $\sim 63\%$ of kittens that are trapped are released alive (Landry, 2017). We therefore assume a normal($\mu = 0.37, \sigma = 0.026$) prior distribution for relative capture probability of kittens aged 0 – 1 year, which puts 95% of the prior probability

density for capture probability between 0.32 and 0.42.

Population growth rates of bobcats within West Virginia are unknown. Based on trends in reported harvests between 2005 and 2016 (Landry, 2017), we assumed a normal($\mu = 1.01$, $\sigma = 0.026$) prior distribution on λ , which assumes a slightly growing population on average but puts 95% of prior density between 0.96 (a declining population) and 1.06 (an increasing population).

Finally, we assumed a stable stage distribution. Because kittens are likely to be released alive, empirical stage structure data are likely to be biased. Given the lack of other information regarding bobcat stage structure, a stable stage distribution is the best available approximation. The adequacy of this assumption would likely be apparent following model checking.

We evaluated model fit using posterior predictive checks (Kéry, 2010). We calculated the G-statistic (defined as $G = 2 \sum_i O_i \log(O_i/E_i)$, with O_i representing observed counts and E_i representing expected counts; Udevitz & Ballachey (1998)) from observed counts (G_{obs}) and from simulated counts (G_{sim}), and calculated a Bayesian p -value as $Pr(G_{sim} \geq G_{obs})$. We assume adequate fit if $(0.05 < p < 0.95)$ (Gelman et al., 2014).

We fit models in JAGS (Plummer, 2003) using the jagsUI interface (Kellner, 2019) to program R (R Core Team, 2020). We ran models for 11,000 iterations, discarding the first 1,000 iterations as burn-in and keeping every iteration thereafter. This was enough to achieve adequate convergence ($\hat{R} < 1.1$ for all parameters Gelman et al. (2014)).

Paddlefish survival study

We also used our age-at-harvest model to estimate American paddlefish survival probability in the Mississippi (rkm 587 – 322, rkm 313 – 0, rkm 1535 – 1353), Ohio (rkm 1160 – 1579), and Missouri (rkm 233 – 0) rivers, USA. Paddlefish were collected from these rivers using hobbled (6.1-m height, 60.1-m length, 127 mm bar mesh) floating monofilament gill nets (Scholten & Bettoli, 2007).

The left dentary of each paddlefish was removed, cleaned, and air dried. Dentaries were sectioned (~0.635 mm thick) using a low-speed water saw. Multiple thin sections were obtained from each dentary to ensure reliable age estimates (Scarnecchia et al., 1997). Dentary sections were aged under a dissecting scope (10 – 40x) using transmitted light. All sections were independently aged by two observers. If observers disagreed on age, the structure was viewed in concert to resolve discrepancies. In cases where a consensus could not be achieved, the sample was removed from analysis. In total, <3% of all sampled paddlefish were discarded due to a lack of agreement on age. Our data therefore consisted of counts of paddlefish individuals by age and river reach (Table 5).

We assumed a five stage-class model (Table 5). We assumed survival probability varied by stage class, but fixed survival probability in stage class 5 (age 18+) equal to stage class 4 (age 17) to ensure survival probability was identifiable (Conn et al., 2008). We additionally allowed survival probability and λ to vary by river reach. To find a parsimonious model of paddlefish survival probability, we evaluated four models with various combinations of river reach and stage-specific survival probability (Table 6). We assumed logistic(0, 1) prior distributions for all survival coefficients.

Sampling paddlefish across all life stages (i.e. stage classes) is imperative for unbiased demographic representation (Ricker, 1975; Scholten & Bettoli, 2007). However, capturing young paddlefish has proven difficult since the onset of paddlefish population evaluations (Allen, 1911; Phelps et al., 2009; Purkett, 1961; Stockard, 1907) because the gear used to sample paddlefish have a low probability of capturing individuals in the youngest stage class. We therefore assume a normal($\mu = 0.05$, $\sigma = 0.026$) prior distribution for relative capture probability of paddlefish in the youngest stage class. To ensure positive relative capture probability, we truncated the prior distribution below at 0.

Population growth rates of paddlefish within the Mississippi River are unknown. Given the lack of information on population growth rates, we assume a normal($\mu = 1.0$, $\sigma = 0.026$) prior distribution on λ , which assumes a stable population on average but puts 95% of prior density between 0.95 (a declining population) and 1.05 (an increasing population).

Finally, we assumed a stable stage distribution. Because paddlefish within the youngest stage class are poorly sampled, empirical stage structure data are likely to be biased. Given the lack of other information regarding paddlefish stage structure, a stable stage distribution is the best available approximation.

We evaluated model fit with posterior predictive checks (Kéry, 2010) using identical techniques as described in the bobcat survival analysis. Preliminary analyses indicated poor fit when assuming counts by stage class were a multinomial random variable. We therefore assumed the multinomial-Dirichlet model from equation 2.

We fit models in JAGS (Plummer, 2003) using the jagsUI interface (Kellner, 2019) to program R (R Core Team, 2020). We ran models for 11,000 iterations, discarding the first 1,000 iterations as burn-in and keeping every iteration thereafter. This was enough to achieve adequate convergence ($\hat{R} < 1.1$ for all parameters Gelman et al. (2014)).

Results

Simulation study

Based on our simulations, estimates of survival probability were unbiased when the prior distributions for λ and α_1 were centered on their data-generating values (Figure 2), and there appeared to be no advantage to assuming either a normal or uniform prior distribution for these parameters. In addition, there was substantial gain in precision when moving from 50 to 1,000 samples. The model appeared to be more sensitive to errors in estimates of population growth rate compared to relative capture probability. When the prior distribution for λ was centered on its data-generating parameter, the model could produce unbiased estimates of survival probability even with biased prior distributions for α_1 . Provided that 95% of the prior density for α_1 contained its data-generating value (i.e. the ‘uncentered’ scenario), estimates of survival probability were largely unbiased. In the uncentered α_1 scenario, sample size and form of the prior distribution (i.e. normal vs. uniform) did not have a strong influence on accuracy of estimates of survival probability. When 95% of the prior density for α_1 did not contain its data-generating value (i.e. the ‘outside’ scenario), unbiased estimates of survival probability only occurred at the larger sample sizes and when assuming a normal prior distribution for both α_1 and λ . In contrast, estimates of survival probability were always biased if prior distributions of λ_t were not centered on its data-generating value. The greater the prior bias in estimates of λ , the greater the bias in resulting estimates of survival probability. These same patterns held regardless of whether the stage distribution was distributed as the stable stage distribution. The primary effect of departure from the stable stage distribution was to introduce additional uncertainty into estimates of survival probability (Figure 2).

Bobcat survival

We fit our age-at-harvest model with data obtained from 524 harvested bobcats (Table 3). The top model demonstrated a strong effect of year on bobcat survival probability (Table 4). Although the 2nd ranked model included an effect of sex and year, 95% credible intervals of the slope coefficient for sex overlapped 0. Since only the top models exhibited reasonable fit, and the effect of sex was not strongly different from 0, we draw inference exclusively from the top model, which included an effect of year only. Mean survival probability over all years was 0.62 (95% CI = 0.58 – 0.65). Survival probability was greater (mean regression coefficient = 0.46, 95% CI = 0.17 – 0.76) in the 2014 / 2015 harvest season (0.67; 95% CI = 0.62 – 0.72) relative to the 2015 / 2016 harvest season (0.56; 95% CI = 0.51 – 0.62).

Paddlefish survival

We fit our age-at-harvest model with data obtained from 424 American paddlefish (Table 5). All models exhibited reasonable fit, and all models were within 1 Δ DIC unit of each other (Table 6). The top model included an effect of both age and river reach, and 95% credible intervals for the age 17+ slope coefficient in this model did not overlap 0. While the direction and magnitude of estimated slope coefficients were consistent across the remaining models, all 95% credible intervals of slope coefficients in the remaining models overlapped 0. Therefore, while acknowledging substantial model selection uncertainty, we report results only from the top model and pertaining to age.

We found 0.93 (95% CI = 0.8 – 0.99) annual survival probability of 0 – 8 year old fish; 0.96 (95% CI = 0.83 – 1) annual survival probability of 9 – 12 year old fish; and 0.96 (95% CI = 0.8 – 1) annual survival probability of 13–16 year old fish. Slope coefficients for the 9 – 12 and 13 – 16 year stage classes included credible intervals that overlapped 0, suggesting no strong difference in survival probability among the the youngest stage classes. However, we found 0.6 (95% CI = 0.19 – 0.91) annual survival probability of 17+ year old fish, and 95% credible intervals of the slope coefficient did not overlap 0, suggesting lower survival probability among the oldest stage classes of fish.

Discussion

We demonstrate the utility of a novel age-at-harvest model that can estimate survival probability without many of the restrictive assumptions of established approaches. To our knowledge, this is the first such model that can simultaneously account for variation in relative capture probability, population growth or reproductive rates, age or stage structure, and survival probability among age / stage groups or as a function of covariates, all while fully accounting for uncertainty in each of these processes. Our simulations do demonstrate that estimates of survival probability will be biased if poor estimates of population growth or reproductive rates are supplied. Therefore, while our proposed model may not require ancillary data, we caution that poor estimates of population growth rates obtained from the literature or expert opinion may yield inaccurate estimates of survival probability.

We expect that this model will be most useful when estimating survival probabilities for species that are primarily monitored through harvest records or for which counts of individuals by age can be obtained. Given the bias that is likely to occur if estimates of population growth rates or reproduction are poor, we do not advocate this model as a replacement for more rigorous mark-recapture or known-fate survival studies. Instead, we advocate use of this model as a method for deriving estimates of survival probability when

estimates from other rigorously designed studies are not available. In such circumstances, we envision that many of the required prior distributions can be provided as expert opinion. The Bayesian paradigm then allows for a full accounting of uncertainty when estimating survival probability from such prior information. While this may lead to imprecise estimates of survival probability, such an approach is probably a more accurate reflection of uncertainty relative to established age-at-harvest models that relax one assumption at a time, or assume auxiliary parameters are fixed and known.

Alternatively, our age-at-harvest likelihood could be substituted into integrated population models (Schaub & Kéry, 2021) for species with auxiliary data. For example, Gove et al. (2002) describe an integrated population model that uses an age-at-harvest likelihood in addition to likelihoods derived for auxiliary data. In such situations, estimates of population growth rates, reproduction, relative capture probability, etc. could be supplied as auxiliary data, potentially improving accuracy and precision of survival probability *and* other model parameters.

Our simulations indicate that accurate estimates of population growth rates are critical for unbiased estimates of survival probability from age-at-harvest data. We note that if good estimates of population growth rates are unavailable, estimates of reproductive rates could instead be substituted into our model. We expect accurate estimates of survival probability will still depend on accurate estimates of reproduction, but such estimates may be more widely available than population growth rates for some species. For example, estimates of reproductive rates could be derived from the primary literature and used to construct prior distributions for this parameter. Our simulations further demonstrate that the proposed model does not add additional bias to the survival estimate with moderate departures from the stable stage distribution (Figure 2). We expect that our model will produce accurate estimate of survival even when the population is not at a stable age or stage distribution. However, violating this model assumption does result in increased uncertainty around the survival estimate.

Importantly, our estimates of bobcat survival probability are congruent with other estimates reported in the literature. For example, a recent known-fate analysis of bobcats within West Virginia suggested annual survival probabilities of approximately 0.74 between 2018 and 2020 (Edwards et al., 2021). Other regional studies suggest annual survival probabilities between 0.61 and 0.84 in West Virginia (Fox, 1980) and Massachusetts (Fuller et al., 1995), respectively. This suggests that despite uncertain prior knowledge on population growth rates, relative capture probability of juvenile bobcats, and assumption of stable stage distributions, our model appears to produce reasonable estimates of bobcat survival probability.

Our estimate of American paddlefish annual survival probability was also similar to other published accounts. For example, Glassic et al. (2020) report annual survival probability of 0.92 and 0.82 for females and males, respectively, in the Missouri River. Similarly, Boone & Timmons (1995) and Donabauer et al.

(2009) both report annual survival probability of 0.91 in the absence of commercial or recreational harvest. We note that output from our American paddlefish model may not be much different than what could be obtained using established methods. However, it relaxes assumptions of established methods, incorporates uncertainty into model components that allow greater complexity, allows inference on other factors that may have influenced survival probability (even if we found no difference in survival probability along different stretches of the Mississippi River), and probably represents a more honest accounting of uncertainty relative to established methods.

Here, we have presented a Bayesian model for estimating survival probability from age-at-harvest data. We demonstrated that it can produce reasonable estimates of survival probability, while relaxing many of the assumptions of more established techniques. Our approach is highly flexible, both in terms of model components (e.g. specification of age or stage structure, reproduction, or population growth) and likelihood structure (e.g. straightforward implementation of multinomial-Dirichlet model for overdispersed data). We expect our proposed model will be useful to a wide array of practitioners seeking to understand demographic parameters of wildlife and fisheries populations from age-at-harvest data.

Acknowledgments

This paper is dedicated to the memory of Rich Rogers. We wish to recognize his more than 37 years of service as a Wildlife Biologist for the West Virginia Division of Natural Resources. He made significant impacts throughout his career, contributing greatly to wildlife conservation and management, as well as serving on graduate student committees and furbearer technical committees. Rich was a true outdoorsman, scientist, mentor, and friend to all who knew him, and his research and knowledge will continue to help guide wildlife management. This paper was a product of an independent study class, WMAN 695, supervised by C.T.R. We thank West Virginia University (WVU) and the WVU Division of Forestry and Natural Resources for access to resources while writing this paper. H.L.C. was supported by the National Science Foundation Graduate Research Fellowship under Grant No. DGE-1102689. C.T.R. was supported in part by the USDA National Institute of Food and Agriculture, McIntire Stennis project number WVA00818. We thank the West Virginia Division of Natural Resources for funding the bobcat project to obtain data used in this paper (PR grant R-48). We also thank the Missouri Department of Conservation and Kentucky Department of Fish and Wildlife Resources for providing paddlefish age information.

Authors' Contributions

C.T.R conceived the ideas and developed methodology; S.L., R.R., Q.P., and J.T.A. contributed data; B.P.S., H.L.C, and C.T.R. performed simulations and statistical analyses; B.P.S., H.L.C., and C.T.R. wrote and revised the manuscript. All authors contributed to drafts and gave final approval for publication.

Data Availability Statement

All data and R code used in simulations and the two real-world examples are available from the Zenodo data repository at <https://doi.org/10.5281/zenodo.7011470>.

441 **Tables**

Table 1: List of key components of our proposed model to estimate survival probability from age-at-harvest data, including parameters, states, definitions, inputs, and considerations.

Parameters	State	Definitions	Inputs	Considerations
Age-at-harvest data (y)	Fixed and known	Counts of harvested individuals segregated by age or stage class	Count data collected by biologists	Option to further segregate count data into groups representing sex, time period, geographical location, etc. if survival probability is thought to be a function of those covariates
Population growth rate (λ)	Can be assumed fixed and known, or assigned informative prior distributions	Rate of change in total population over a defined time period	Calculations of change in abundance or index data during a defined time period	Sources for abundance or index data and their associated biases
Relative capture probability (α)	Can be assumed fixed and known, or assigned informative prior distributions	Probability of individuals in a particular age or stage class being harvested given the opportunity, relative to individuals in other age or stage classes	Data on harvest selectivity by hunters or anglers	Sources for harvest selectivity based on equipment used, regulations, and hunter or angler preferences; often cases where selectivity favors older age classes or is biased against younger age classes
Age/Stage distribution (C)	Can be assumed fixed and known, or assigned informative prior distributions	Expected number of individuals in each age or stage class	Estimated number of individuals in each age or stage class	Biological consideration for pre- or post-birth pulse data collection and associated changes in age or stage class abundance
Age/Stage class duration (d)	Fixed and known	Maximum duration for which an individual can remain in each age or stage class	Data determined by biologists	Biological or management considerations for grouping ages into stage classes

Table 2: Prior distributions for relative capture probability (α_1) of stage class 1 and population growth rate (λ) used in simulation scenarios. Centered, Uncentered, and Outside refer to the simulation scenario corresponding to prior distributions being centered on the data-generating value of the parameter; prior distributions not centered on data-generating values of the parameter, but the data-generating value falls within 95% of prior probability density; and the data-generating value of the parameter falling outside of 95% of prior probability density.

Parameter	Simulation scenario		
	Centered	Uncentered	Outside
α_1	Uniform(0.25, 0.75)	Uniform(0.45, 0.95)	Uniform(0.65, 1.15)
	Normal($\mu=0.50$, $\sigma=0.13$)	Normal($\mu=0.70$, $\sigma=0.13$)	Normal($\mu=0.90$, $\sigma=0.13$)
λ	Uniform(0.960, 1.060)	Uniform(0.985, 1.085)	Uniform(1.020, 1.120)
	Normal($\mu=1.010$, $\sigma=0.026$)	Normal($\mu=1.036$, $\sigma=0.026$)	Normal($\mu=1.070$, $\sigma=0.026$)

Table 3: Counts by age, sex, and year of bobcat (*Lynx rufus*) carcasses obtained by trappers during 2014 – 2016 in West Virginia, USA.

Sex	Year	Age 0-1	Age 1-2	Age 2-3	Age 3-4	Age 4-5	Age 5+
Male	2014/15	32	27	28	24	21	20
Female	2014/15	28	30	36	17	14	23
Male	2015/16	24	42	22	11	5	8
Female	2015/16	23	45	20	8	5	11

Table 4: Model selection and Bayesian p-values for bobcat (*Lynx rufus*) survival models in West Virginia, USA, 2014 – 2016.

Model	Bayesian p-value	DIC
Year	0.08	120.46
Sex + Year	0.07	122.33
Age + Year	0.07	124.68
Age + Sex + Year	0.06	126.36
Constant	0.00	133.81
Sex	0.00	135.79
Age	0.00	136.10
Age + Sex	0.00	137.99

Table 5: Counts by age and river reach of American paddlefish (*Polyodon spathula*) obtained within the Mississippi River, USA.

River	Age 0-8	Age 9-12	Age 13-16	Age 17	Age 18+
Lower	2	67	75	6	5
Middle	30	65	55	2	1
Upper	31	41	20	12	12

Table 6: Model selection and Bayesian p-values for American paddlefish (*Polyodon spathula*) survival models in the Mississippi River, USA.

Model	Bayesian p-value	DIC
Age + River	0.69	72.07
River	0.68	72.24
Constant	0.70	72.26
Age	0.72	72.29

442 **Figures**

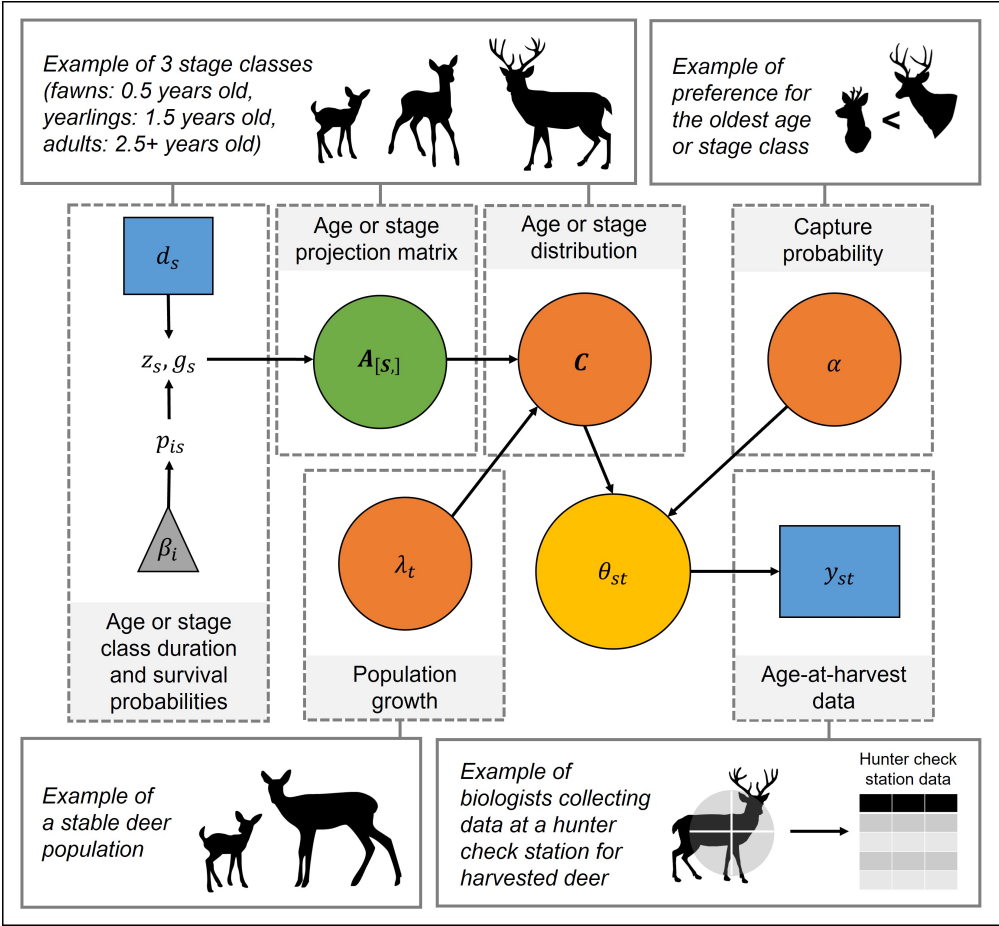


Figure 1: Conceptual diagram of the components of our proposed model to estimate survival probability from age-at-harvest data. The parameters of the model are color-coded such that: blue boxes represent available data that has been collected or determined (d_s = fixed maximum duration for which an individual can remain in each age or stage class s ; y_{st} = the number of harvested individuals within each age or stage class s during time t); the gray triangle represents covariate data (β_i = predictor variable influencing age- or stage-specific survival probabilities); the green circle indicates a parameter derived from data (A_i = an age- or stage-structured matrix projection model with S age or stage classes); orange circles represent parameters that can be treated as either fixed and known from available data or estimated using a prior distribution (λ_t = population growth rate at time t ; C = the age or stage distribution ; α = relative capture probability); and the yellow circle is a derived parameter and the process of interest (θ_{st} is the probability a harvested individual belongs to age or stage class s during time t). Boxes corresponding with certain parameters give corresponding real-world examples for white-tailed deer, a popular game species for which age-at-harvest data are often collected at hunter check stations.

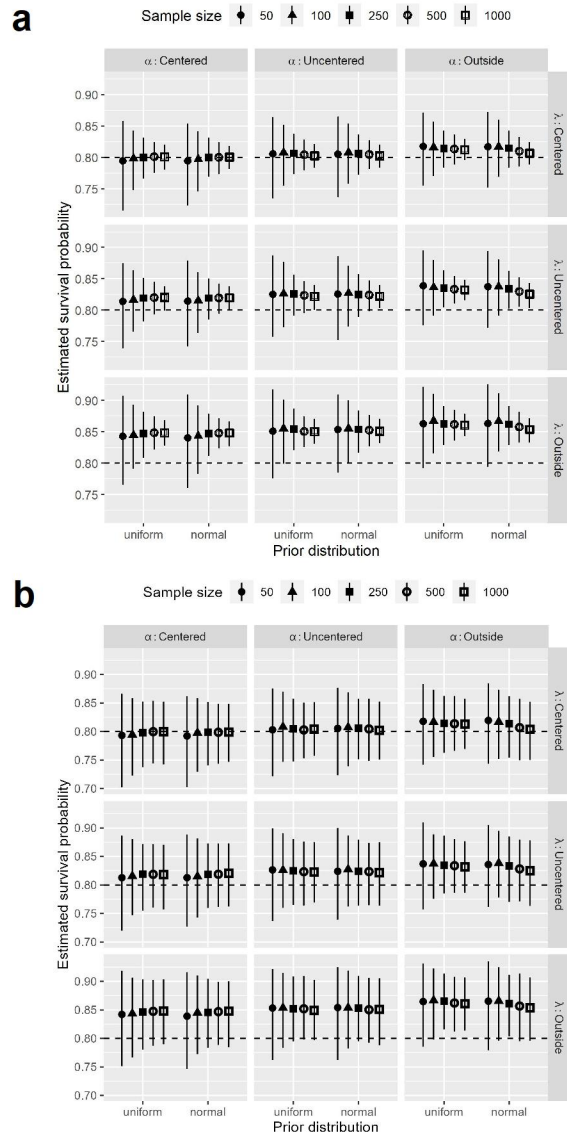


Figure 2: Mean estimates of survival probability (with 95% CI) obtained from simulated age-at-harvest data. Panel (a) represents simulations that follow a stable stage distribution and panel (b) represents simulations where there is a departure from the stable stage distribution. Each scenario had 1,000 replicates from which the mean estimates and 95% CI are determined from the posterior means. Centered, Uncentered, and Outside refer to whether the prior distribution for the associated parameter is centered on its data-generating value; the data-generating value of the associated parameter is within 95% of prior density; or the data-generating value of the associated parameter is outside 95% of prior density, respectively. The horizontal dashed line is the data-generating value for survival probability. Data-generating values for α_1 and λ are 0.5 and 1.01, respectively. Prior distributions for each scenario are listed in Table 2.

References

Allen, M. S., & Hightower, J. E. (2010). Fish population dynamics: Mortality, growth, and recruitment. In W. Hubert & M. Quist (Eds.), *Inland fisheries management in north america* (pp. 43–77). American Fisheries Society.

Allen, W. F. (1911). Notes on the breeding season and young of *Polyodon spathula*. *Journal of the Washington Academy of Sciences*, 1(10), 280–282.

Barker, R. J. (1997). Joint modeling of live-recapture, tag-resight, and tag-recovery data. *Biometrics*, 53(2), 666–677.

Beissinger, S. R., & McCullough, D. R. (2002). *Population Viability Analysis*. University of Chicago Press.

Boone, E. A., & Timmons, T. J. (1995). Density and natural mortality of paddlefish, *Polyodon spathula*, in an unfished Cumberland River subimpoundment, South Cross Creek Reservoir, Tennessee. *Journal of Freshwater Ecology*, 10(4), 421–431.

Broms, K., Skalski, J. R., Millspaugh, J. J., Hagen, C. A., & Schulz, J. H. (2010). Using statistical population reconstruction to estimate demographic trends in small game populations. *The Journal of Wildlife Management*, 74(2), 310–317.

Buckland, S. T., Newman, K. B., Thomas, L., & Koesters, N. B. (2004). State-space models for the dynamics of wild animal populations. *Ecological Modelling*, 171(1-2), 157–175. <https://doi.org/10.1016/j.ecolmodel.2003.08.002>

Caswell, H. (2001). *Matrix Population Models* (2nd ed.). Sinauer Associates, Inc.

Caughley, G. (1977). *Analysis of Vertebrate Populations*. John Wiley & Sons.

Chapman, D., & Robson, D. (1960). The analysis of a catch curve. *Biometrics*, 16(3), 354–368.

Conn, P. B., Diefenbach, D. R., Laake, J. L., Ternent, M. A., & White, G. C. (2008). Bayesian analysis of wildlife age-at-harvest data. *Biometrics*, 64(4), 1170–1177.

Crouse, D. T., Crowder, L. B., & Caswell, H. (1987). A stage-based population model for loggerhead sea turtles and implications for conservation. *Ecology*, 68(5), 1412–1423.

Crowe, D. (1972). The presence of annuli in bobcat tooth cementum layers. *The Journal of Wildlife*

Management, 36, 1330–1332.

Donabauer, S. B., Stoeckel, J. N., & Quinn, J. W. (2009). Exploitation, Survival, Reproduction, and Habitat Use of Gravid Female Paddlefish in Ozark Lake, Arkansas River, Arkansas [Proceedings Paper]. In G. Paukert, CP and Scholten (Ed.), *Paddlefish management, propagation, and conservation in the 21st century: Building from 20 years of research and management* (Vol. 66, p. 123+). American Fisheries Society.

Edwards, J., Rota, C. T., & Belcher, K. (2021). *Final report: Bobcat home range, habitat selection and survival in eastern West Virginia*. West Virginia University.

Efron, B. (1988). Logistic regression, survival analysis, and the kaplan–meier curve. *Journal of the American Statistical Association*, 83(402), 414–425.

Fox, L. (1980). *The age structure, reproduction, winter food habits, distribution and harvest of West Virginia bobcats during 1977 – 78* [MS Thesis]. State University of New York College of Environmental Science; Forestry. Syracuse, NY, USA.

Fuller, T. K., Berendzen, S. L., Decker, T. A., & Cardoza, J. E. (1995). Survival and cause-specific mortality rates of adult bobcats (*Lynx rufus*). *The American Midland Naturalist*, 134(2), 404–408.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis, Third edition*. CRC Press.

Glassic, H. C., Guy, C. S., Rotella, J. J., Nagel, C. J., Schmetterling, D. A., & Dalbey, S. R. (2020). Fort Peck paddlefish population survival and abundance in the Missouri River. *Journal of Applied Ichthyology*, 36(5), 559–567.

Gove, N. E., Skalski, J. R., Zager, P., & Townsend, R. L. (2002). Statistical models for population reconstruction using age-at-harvest data. *The Journal of Wildlife Management*, 66(2), 310–320.

Kellner, K. (2019). *jagsUI: A Wrapper Around 'rjags' to Streamline 'JAGS' Analyses*. R package version 1.5.1. <https://cran.r-project.org/package=jagsUI>

Kéry, M. (2010). *Introduction to WinBUGS for ecologists*. Elsevier.

Landry, S. M. (2017). *Bobcat Population Ecology in West Virginia* [MS Thesis]. West Virginia University, Morgantown, WV, USA.

Lehman, C. P., Rota, C. T., Raithel, J. D., & Millsbaugh, J. J. (2018). Pumas affect elk dynamics in absence

- of other large carnivores. *Journal of Wildlife Management*, 82(2), 344–353.
- Mills, L. S. (2007). *Conservation of Wildlife Populations: Demography, Genetics, and Management*. Blackwell Publishing.
- Phelps, Q. E., Tripp, S. J., Gravey, J. E., Herzog, D. P., Ostendorf, D. E., Ridings, J. W., Crites, J. W., & Hrabik, R. A. (2009). Ecology and Habitat Use of Age-0 Paddlefish in the Unimpounded Middle Mississippi River. In C. Paukert & G. Scholten (Eds.), *Paddlefish management, propagation, and conservation in the 21st century building from 20 years of research and management* (pp. 423–440). American Fisheries Society.
- Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *Proceedings of the 3rd International Conference on Distributed Statistical Computing*, 1–10. <http://www.ci.tuwien.ac.at/Conferences/DSC-2003/>
- Purkett, C. A. J. (1961). Reproduction and early development of the Paddlefish. *Transactions of the American Fisheries Society*, 90(2), 125–129.
- R Core Team. (2020). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. <https://www.r-project.org/>
- Ricker, W. (1975). *Computation and Interpretation of Biological Statistics of Fish Populations*.
- Roseberry, J. L., & Woolf, A. (1991). A comparative evaluation of techniques for analyzing white-tailed deer harvest data. *Wildlife Monographs*, 117, 3–59.
- Rota, C. T., Millsbaugh, J. J., Rumble, M. A., Lehman, C. P., & Kesler, D. C. (2014). The role of wildfire, prescribed fire, and mountain pine beetle infestations on the population dynamics of black-backed woodpeckers in the Black Hills, South Dakota. *PLoS ONE*, 9(4), e94700.
- Scarnecchia, D., Ryckman, L., & Lee, J. (1997). Capturing and tagging wild age-0 and age-1 paddlefish in a Great Plains reservoir. *North American Journal of Fisheries Management*, 17, 800–802.
- Schaub, M., & Kéry, M. (2021). *Integrated Population Models 1st Edition; Theory and Ecological Applications with R and JAGS*. Elsevier.
- Schnute, J. T., & Haigh, R. (2007). Compositional analysis of catch curve data, with an application to *Sebastes maliger*. *ICES Journal of Marine Science*, 64(2), 218–233.
- Scholten, G. D., & Bettoli, P. W. (2007). Lack of size selectivity for paddlefish captured in hobbled gillnets.

Fisheries Research, 83(2-3), 355–359. <https://doi.org/10.1016/j.fishres.2006.11.005>

Skalski, J. R., Smith, S. G., Iwamoto, R. N., Williams, J. G., & Hoffmann, A. (1998). Use of passive integrated transponder tags to estimate survival of migrant juvenile salmonids in the Snake and Columbia rivers. *Canadian Journal of Fisheries and Aquatic Sciences*, 55(6), 1484–1493.

Smith, M. W., Then, A. Y., Wor, C., Ralph, G., Pollock, K. H., & Hoenig, J. M. (2012). Recommendations for catch-curve analysis. *North American Journal of Fisheries Management*, 32(5), 956–967.

Stockard, C. R. (1907). Observations on the Natural History of *Polyodon spathula*. *The American Naturalist*, 41(492), 753–766.

Thorson, J. T., & Prager, M. H. (2011). Better catch curves: Incorporating age-specific natural mortality and logistic selectivity. *Transactions of the American Fisheries Society*, 140(2), 356–366.

Udevitz, M. S., & Ballachey, B. E. (1998). Estimating survival rates with age-structure data. *The Journal of Wildlife Management*, 62(2), 779–792.

Udevitz, M. S., & Gogan, P. J. P. (2012). Estimating survival rates with time series of standing age-structure data. *Ecology*, 93(4), 726–732.

Williams, B. K., Nichols, J. D., & Conroy, M. J. (2002). *Analysis and Management of Wildlife Populations*. Academic Press.

Reviewer 1

In this paper, Skelly and coauthors propose using a multinomial model for analyzing age-at-harvest data from fish and wildlife populations to estimate survival. Conditioning on the total number of harvested animals, multinomial cell probabilities are constructed in terms of demographic processes (e.g., survival, finite rate of pop growth λ) as well as some nuisance parameters such as differential sampling probabilities among age/stage classes. Their approach is most closely related to the approach of Udevitz and Ballachey [1998], but uses a Bayesian approach incorporating prior distributions on unknown parameters - an approach that in some senses generalizes previous work in the sense that one can get estimates with fewer assumptions and data types, but also possibly obscures which types of data are needed to enable reliable estimation. They illustrate their modeling approach with a small simulation study and by analyzing bobcat and paddlefish data.

I was a bit conflicted about this paper. On the one hand, age-at-harvest data are often available from fish and wildlife agencies and it's advantageous to try to use it to the best of our ability even if few other data are available to put them in context. In that sense, putting prior distributions on things and turning the crank to get usable estimates (best available science) makes sense to do, and the approach the authors outline is certainly a reasonable way to do it. On the other hand, their model makes a large number of assumptions, and I'm not the sure that the manuscript in its present form is self-aware as it should be. For instance, their simulation study doesn't address many common assumption violations that could potentially bias estimates, such as different types of selectivity functions, violations of stable stage structure induced by time-varying survival and fecundity, etc. So it's hard to recommend this approach generally, I would think, or to give biologists enough indications of when it might fail for their specific problem. This kind of comes to a head in the paddlefish example. Looking at those data, I would say that many of their assumptions are likely to be violated, and thus would have trouble recommending the authors' approach on those data. For instance, in catch curve analysis, one often eliminates data prior to the age at which maximum catches occur in order to eliminate the issue of increasing selectivity at younger ages (this wasn't the approach used here, and I suspect that that there is a large selectivity issue that isn't being taken care of in this analysis). Further, the large fall off in harvested paddlefish at older ages suggest that survival may really take a hit when maximum selectivity is reached - but in the model survival is age-independent.

I'm not really sure what to suggest at this point. Perhaps conducting a more thorough simulation study and removing the paddlefish example would be a reasonable option. Whatever approach is taken, the authors should try to do a better job of highlighting the limitations (i.e., not just the strengths) of their proposed approach.

We thank Reviewer 1 for their thoughtful and thorough review. We have incorporated many of their suggestions into our revised manuscript and consequently believe we have a much stronger paper. In particular, we conducted additional simulations to evaluate violations of stable stage structure, and we added additional structure to the paddlefish model that was able to detect a change in survival probability at the older stage classes.

We also respectfully point out that we had incorporated some of their suggestions in the previous draft. For example, we explicitly simulate selectivity by allowing the youngest stage class to have a lower capture probability relative to older stage classes, and we include such selectivity in both the bobcat and

paddlefish models. We also attempted to give a clear-eyed evaluation of our model – strengths and weaknesses both. For example, we clearly indicated that poor estimates of growth rates lead to biased estimates of survival probability. We attempt to clarify many of these statements throughout our revised manuscript.

Line 29, I'm interested in what the conditions for which p not centered on true value arrive at unbiased estimates of survival

We indicate that prior distributions for population growth rate that are not centered on the data-generating value lead to biased estimates of survival probability.

Line 44 and elsewhere. No comma needed preceding ampersand in citations

Citations fixed throughout the revised manuscript.

*Line 128 - since the multinomial index is conditional on animals being harvested, aren't we talking about the "relative" probability that an individual is in a given age or stage class *given* that it is harvested & reported? The way the explanation is currently written it comes off as if it's the probability that an individual in the population is "captured" which isn't really the case here.*

We agree and instead state that "theta is the probability a harvested individual ... belongs to each age or state class".

Line 132 - speaking of reporting, it might be good to define the probability of harvest as consisting of the probability of harvesting and reporting an animal. Mandatory check stations are one thing, but some U.S. state agencies rely on some level of goodwill and the hunting public voluntarily stopping, which seems more likely for hunters wanting to get an age on an older age class deer for instance (or not display to the world the tiny deer they just shot).

We revise the manuscript as suggested.

Line 139 - So λ is assumed to be time invariant?

Depending on the data available, λ need not be time invariant. We include a time index on λ to indicate that it can indeed vary with time. Moreover, we allow λ to vary with time when estimating bobcat survival (see also our response below).

Lines 158-163 -The material starting here could be clearer. For instance in the "if an individual can remain in stage class s for duration d_s " - this could be stated more exactly. From the sentence, it's not clear whether d_s is the expected time in a stage class or a limit - e.g., an animal can stay in s for at most duration d_s . But this formulation for z_s and p_s isn't looking very intuitive based on either of these... In a size structured model, for instance, transition probabilities from one stage to the next depend on both survival and growth, so in this context what is p_s ? There are also some issues of whether transition processes are memory independent (i.e., exponential, Markovian), or whether dwell times should have some type of distribution (e.g., normal, semi-Markov).

These are all excellent points. We attempt to clarify as suggested by the reviewer. For example, we state that d_s is the maximum time an individual can remain in a stage class, and we make it clear that the expressions provided by Caswell are for very specific transition processes between stage classes. To try to make the expression more intuitive, we give an example for a specific value of d_s . We also indicate that many other types of transitions probabilities are possible and borrow suggestions made by the reviewer.

I think for models like this it makes sense to ask the question: “Where is this model the most likely to go wrong?” and then target simulations to cover some of those cases. I’m not confident that the scenarios considered in Table 1 cover those bases very well. In particular, I think one of the most likely assumption violations is to specify a λ prior centered on 1.0, when in fact a population is decreasing or increasing. Another scenario would be annual changes in λ induced by variation in recruitment and first year survival (this messes up the stable age structure and constant λ assumptions. Yet another, particularly germane for fish populations but also sometimes wildlife, is an increasing (e.g. Beverton-Holt) or dome-shaped (e.g. Ricker) selectivity function. And one more - unequal survival among stage classes...or, at least for the first age class.

We agree that simulations should ask “where is the model most likely to go wrong?”. We attempted to do this by asking how prior distributions for α (relative capture probability) and λ (population growth rate) that were very different from data-generating parameters influenced estimates of survival probability. We believe we demonstrate that the model does a poor job of estimating survival probability when poor prior estimates of population growth rates are supplied. We also believe we demonstrate that the model is relatively robust to poorly specified priors on relative capture probability (what we believe the reviewer means by selectivity), as long as the prior for population growth rate is accurate.

We agree that it is important to evaluate how the model performs when assumptions of stable age distribution are not met. We therefore included another set of simulations where we evaluate bias in estimates of survival probability when age distribution is different from the stable age distribution. We demonstrate that this adds additional uncertainty to estimates of survival probability, but no additional bias.

Line 228. I’m not sure what the authors mean by logistic(0,1) priors

A logistic distribution is similar to a normal distribution, but with fatter tails. A nice property of this distribution is that if $y \sim \text{logistic}(0, 1)$, then $\text{logit}^{-1}(y) \sim \text{uniform}(0, 1)$, which makes it useful for specifying prior distributions on the log odds scale for probability parameters.

Line 229. Change double negative?

Revised as suggested.

Lines 229-232. Is there any reason to suspect that susceptibility to traps changes with age? (e.g., animals becoming more wary?)

This is not unreasonable to suggest, but we have no data regarding relative capture probability of older-aged individuals. Given the lack of data on relative capture probability of older individuals, we continue to assume equal capture probability of individuals >1 year old. However, we are willing to explore models with lower relative capture probability of older individuals if the AE or reviewers strongly feel that we should.

Line 242. Maybe don’t italicize ‘log’

Revised as suggested.

Lines 259-260. The problem with dropping the records is that this will tend to occur for older animals... could the authors give the reader a sense of the percent or records that were dropped in this manner?

This is a good point. We added a statement in the revised manuscript to indicate that less than 3% of the sampled paddlefish were discarded due to a lack of agreement on age.

Line 314. On a philosophical level, does it make sense to fit models with constant lambda and varying survival?

Good point! We allowed population growth rate of our bobcat model to vary by year. This did very little to influence estimated survival probability.

Table 3. Blank line in table

Blank line removed from the table.

Table 4. There certainly looks to be increasing selectivity here through the 13-17 age group. Amazing how quickly number drop off in the Age 18+ group, too!! This seems like a difficult data set to model without making changes to underlying model structure.

In the lower reach of the Mississippi, captures did indeed peak at the 13–17 age group. However, captures peaked at the 9–12 age class in the middle and upper reaches, so we're not sure one can say there is increasing selectivity through the 13–17 age group. We also agree there is a big drop-off in counts in the 18+ age group. This suggestion led us re-evaluate our assumption of equal survival probability between age 13–17 and age 18+. We added an additional stage group composed only of age 17 individuals and assumed equal survival probability between age 17 individuals and age 18+ individuals. This indeed improved our ability to capture the change in survival for the older age classes, and we now found lower survival probability of age 17 + individuals relative to other stage classes.

Lines 319-323. I would be extremely hesitant to recommend this model / estimate for management advice. Looking at Table 4, it really looks like there are some other things going on that aren't modeled (e.g., selectivity). Mortality then looks like it must really increase in the Age 13-17 level to explain the huge drop off in the Age 18+ numbers. So, we potentially have a situation with age-specific p , and age specific a , and probably some river-specific effects, which are probably too many effects to deal with age-at-harvest data alone. I suspect that the GOF test initially picks up on this lack-of-fit, and that trying to counter it by throwing in overdispersion dilutes the ability of the model to select more complicated models. Ultimately, this dataset appears to violate too many model assumptions for it to be a good illustration of the author's proposed model.

We hope our modifications above are enough to make the reviewer more comfortable with this model. Making changes based on their suggestions, we were able to estimate age-specific survival probability. Additionally, our original model assumed substantially lower relative capture probability for the youngest stage class.

Please note that we do not disagree with any specific points, but we do respectfully disagree with the conclusion. Yes, this is a noisy dataset and the overdispersed model may have diluted the ability to pick up on any river-specific effects. Nonetheless, the model was still able to recover a very reasonable estimate of survival probability, including a drop-off at the oldest stage class. Such noisy data are very commonly collected by natural resource agencies and often represent the best data available for many species. We do not believe current approaches for estimating survival from age-at-harvest data adequately or explicitly account for the myriad assumptions required to analyze such data, which led us to develop this approach. Our inability to find a river-specific effect and the relatively wide credible intervals in survival probability (despite the reasonable central tendency) represent honest uncertainty given the noisy dataset.

Line 352-353. “We expect . . . ”. I’m not sure this is true - the abundance-based age-at-harvest models don’t require constant lambda, and it’s not an issue with catch curves either. I’d suggest deleting or being more specific.

We delete as suggested.

Lines 351-357. It would be interesting to have priors on survival, λ , and fecundity at-age, because one is a function of the other two! Another philosophical quandry...

This is an interesting point. Indeed, if any two of these variables are fixed and known, the third can be solved for. However, we respectfully fail to see the philosophical quandary. We see no reason why prior information can’t be brought to bear on all 3 variables. For example, although reproduction is implied in our model, information on reproductive rates could in principle be included in the model, potentially improving estimates of survival probability.

Lines 358-368. I’m not entirely sure what point/points are being made here. Don’t provide prior distributions for relative capture probability when you have good information on current age/stage structure because relative capture probabilities can then be estimated from the model?

Given this reviewer’s confusion and the results from our new simulation that accurate estimates are survival probability are robust to modest departures from the stable stage distribution, we delete this paragraph.

Reviewer 2

The manuscript is well-written, though there were a number of redundancies that could be eliminated, especially in the introduction. My comments are all editorial. I do suggest you look at the works of Buckland et al. (2004) on their state-space modelling approach. There seems to be an issue with the Bayesian p-values for the bobcat model. Not sure what is going on there as no code was supplied. For the record, Caughley (1977) demonstrated that one either needs to know the rate of increase when analyzing age-structure data or assume a stable age distribution...so, not much changes here. Perhaps demonstrating an integration of your model with ancillary data to compensate for either of these assumptions would be a better hook.

We thank Reviewer 2 for their positive assessment and helpful suggestions. We made the following checks / edits in response to your comments.

- We incorporate Buckland et al. (2004) into our paper as suggested.
- Thank you for the cautionary note regarding Bayesian p-values. See our response to your comment below.
- Our data and R code are now public and accessible from Zenodo at the following link: <https://doi.org/10.5281/zenodo.7011470> (also provided in our Data Availability Statement).

- L 45 Technically, not true. Abundance estimation is separate and apart from survival estimation.

Good point. We modified the sentence to specify that “estimates of survival probability can be used to better understand animal populations...” rather than animal population size and trends.

- L 56 suggest “parameter” so as not to confuse “probability” and “rate”.

Revised as suggested.

- L 65 Usually “catch-at-age” data.

Revised as suggested.

- L 136 This is essentially a state-space formulation using the Leslie or Lefkovitch matrix model.

Correct. We state this in the revised manuscript.

Table 3. These are quite low for bpv's, and suggest underdispersion, which in my experience. Rarely occurs in harvest data. I would check these models.

Thank you for this suggestion. We carefully checked our code and found no error. Following Gelman et al. (2013, *Bayesian Data Analysis* 3), we calculated the Bayesian p-value as:

$$p_B = \Pr(T(y^{rep}, \theta) \geq T(y, \theta) | y)$$

In words, this is the probability the test statistic obtained from replicated data (y^{rep}) is greater than the test statistic obtained from observed data (y). Therefore, relatively small p -values indicate the test statistic obtained from replicated data tends to be smaller than the test statistic obtained from observed data. Since our test statistic gets larger the farther y departs from its expected value, we believe relatively low p -values indicate over-dispersion, since replicated data are closer to their expected values (on average) than observed data. This interpretation puts our Bayesian p -values more in line with the reviewer's expectations.

Reviewer 3

General comments:

1. The GitHub was not accessible at the link provided.

Our data and R code are now public and accessible from Zenodo at the following link:
<https://doi.org/10.5281/zenodo.7011470> (also provided in our Data Availability Statement).

2. The authors promote this approach as being useful to practitioners due to the need to estimate survival for management purposes from limited data and resources. I, too, see this need. However, the paper could be improved to make this work more accessible to practitioners. A figure that depicts the approach to show this method in a more graphical would really help people understand this. Along with that, a table that explains the variables, definitions, possible inputs, and considerations would be an improvement.

We added a color-coded conceptual figure (Figure 1) to depict our model approach in a more graphical way. We also created a new Table 1 with a list of model parameters, definitions, possible inputs, and considerations.

3. Along the lines of making this useful for practitioners, the jump from 50 to 1000 samples is huge. 1000 samples is already a pretty unreasonable sample size in many cases (e.g., tooth analysis is

expensive for managers and it takes a lot of time to process that many teeth). 10,000 samples is beyond useful for consideration. Better example sample sizes for agencies are in the Table 2 and 4 in the manuscript showing the actual data for bobcat and paddlefish. I recommend using something like 50, 100, 250, 500, 1000. It would be more useful to see between 50 and 1000 where the credible intervals really start to shrink.

We modified our sample size scenarios in the simulation study as suggested.

Specific comments:

Line 62: Please include parametric survival models here as an option as well. “KM, proportional hazards, or parametric survival models”

We added parametric survival models to the sentence.

Line 64: “costly study design”? rephrase or explain.

We added clarification to “costly study design”, specifying that we mean “in terms of funding, time, and effort (e.g. the expenses associated with capturing, marking, and monitoring individuals over long periods of time)”.

Line 210: not sure where this math is coming from, please label the numbers to be more meaningful.

We added labels to provide context to the numbers as follows: “We assumed a uniform(0,1) prior distribution for survival probability over all 2 age distributions \times 5 sample sizes \times 2 prior distributions \times 3 capture probabilities \times 3 population growth rates = 180 different simulation scenarios.”

Line 228, 266: binomial instead of logistic

Respectfully, logistic is the appropriate distribution. See also our response to reviewer 1 above.

Lines 229-232: what about selectivity in other age classes for bobcats in particular? What effect would this have on model performance. See Allen ML, Roberts NM, Van Deelen TR. 2018 Hunter selection for larger and older male bobcats affects annual harvest demography. R. Soc. open sci. 5: 180668. <http://dx.doi.org/10.1098/rsos.180668>

For our data, we had evidence that there was selectivity bias against bobcat kittens (Landry 2017), and there was no evidence for selectivity bias towards the oldest age class. In Allen et al. (2018), the selection of older, male bobcats appeared to be primarily driven by hunters compared to trappers, and our study used bobcats harvested by trappers only. However, if managers wanted to vary selectivity for other age classes of bobcats, our model allows for that, and our simulations study shows that survival probability can still be estimated accurately even if capture probability was inaccurate so long as the population growth was accurate.

Interestingly, this is the opposite suggestion provided by reviewer 1. Given the lack of data and no clear agreement on how selectivity should operate on older bobcats, we assume equal selectivity for bobcats >1 year old. However, we will evaluate the impacts of differential selectivity on bobcats of the AE requires this for publication.

Line 273, 283: delete “Note”

We deleted both of those instances of “Note”.

Reviewer 4

The manuscript presents a novel way to analyses age-at-harvest data for the estimation of survival probabilities. The new model extends the toolset of possibilities to analyses such data and adds flexibility. A simulation study confirms that the model works when assumptions are met and the model is applied in two case studies. The manuscript is generally well written and the topic is of relevance, because the age-at-harvest data are relatively frequently collected, but rarely used to estimate survival. By providing more flexible models population ecologists may make more often use of these data.

One main finding of the manuscript is that the model is sensitive to the assumption about the population growth rate. This sensitivity is consistent with already existing models, and was the reason why integrated approaches have been developed. Hence, although the proposed model offers increased flexibility, it still 'suffers' from the same basic problem. Hence, stressing that the novel model does not require auxiliary data may be a bit exaggerated.

We thank reviewer 4 for their positive assessment of our paper. We agree with the reviewer's final point, and edit the first paragraph of the discussion to indicate prior knowledge or expert opinion regarding population growth rate must be accurate in order to obtain unbiased estimates of survival probability.

Otherwise I have two main and a couple of minor points.

1. I had to read the model several times to really understand it. First I thought that the model is circular, as values for the Leslie matrix (A) have to be given, and at the end the parameters of that matrix are estimated. Now I realize that although A is part of the multinomial, the parameters are not fixed, but estimated. Hence, there is no circularity. Still, the presentation of the model should improve, it must become more clear at once, what is fixed and what is estimated. It may help to state explicitly that the parameters of A will be estimated. It would also be helpful to provide commented code of the model.

We thank the reviewer for this suggestion. We explicitly state at the beginning of the model description that the parameters of A are to be estimated, which we hope improves clarity. We also provide a link to our Zenodo repository with our annotated code. Finally, following reviewer 3's suggestion, we include a table of model parameters (Table 1) and a figure (Figure 1) depicting our model, which we hope further improves clarity.

Then, I was not sure about the calculation of the stable stage distribution (L 160 ff). Survival probabilities are needed for its calculation and it is unclear to me, whether they are fixed, or also estimated as part of the model. Again, some clarification is necessary.

As above, we attempt to clarify by explicitly stating that survival probability is estimated. Furthermore, by adding the figure depicting our model and providing code, we expect the readers will be able to see the stable stage distribution is a function of survival probability, which is a parameter to be estimated.

2. The proposed method uses the stable age/stage distribution (C in the model). However, populations are usually not in a stable age/stage distribution. I was wondering, how sensitive the proposed model is with respect to the violation of this assumption.

We conducted another set of simulations and found that departure from the stable stage distribution added additional uncertainty to model parameters, but did not introduce additional bias.

Minor points:

- 'survival rates' and 'survival probabilities' are both used in the manuscript. I suggest to be consistent and to use 'probabilities' only.

Revised as suggested.

- L 102, 'unequal capture probability'. In which sense / with respect to what?

Unequal capture probabilities are in respect to other age/stage classes. We have revised this in our manuscript to be more clear.

- Formulas in general: sometime you explicitly used indices (e.g. t for year), sometimes not. I think the formulas do become more clear, if the indices are consistently and always used. For example in L 126-128 H is a scalar, but to fit with the formulas that follow, it should have a year index.

This is a good suggestion, and consistent with comments from reviewer 1. We add indices throughout and generalize beyond just indexing by time.

- L 128: 'theta is the probability of capturing an individual...' Why 'capturing'? The individuals are harvested (or found dead), but not captured.

Following similar suggestions by reviewer 1, we clarify that this is the probability a harvested individual belongs in each of the observed stage classes.

- L 138: 'age or stage distribution'. Is this the stable age/stage distribution, or is it the age/stage distribution at a specific point of time (year)? Please be more specific.

Following the indexing suggestion above, we provide more details about the stage distribution.

- L 139: index t is lacking on λ

We added an index t to λ .

- L 146: here again you use the term 'capture probability', which I find confusing.

We more clearly define the parameter θ without the term capture probability. This allows us to clarify that here we are referring to relative capture probabilities within each stage class.

- L 169: you state that survival can vary by age or stage class. But what about temporal variation? This should also be possible and indeed in the bobcat example, year is a variable.

Given the reviewer's helpful suggestion of adding group indices, we are able to clarify that group identity can be defined by time or other variables.

- Bobcat study: it is written that 'we assumed that individuals remain within the first five stage classes for 1 year, while individuals in the oldest age can survive and remain within that stage class'. Hence this fixes d_s (L 159), but please be more specific about the values chosen for d_s .

We further clarify d_s as suggested.

- Results of bobcat study: the best model had temporally variable survival probabilities, hence I suggest to report these estimates in addition.

We added the estimates of the individual annual survival probabilities in addition to the mean survival probability across both years.

- Paddlefish study, L 279: which values were used for the stable stage distribution?

We're a little confused by this comment. Stable stage distribution was estimated as a function of survival and population growth rates. Perhaps the additional clarity we add above regarding model structure will clarify this question for the reviewer.

- figure 1: the results of the simulation study (1000 replicates) are presented here, and it is written that mean estimates and 95% CI are shown. Are these the means across the 1000 posterior means? Please clarify in the caption.

Yes, these are the means across the 1,000 posterior means. We have clarified this in the figure caption as suggested.