

# Modeling dynamics of the Nēnē population in Hawai‘i Volcanoes National Park

N. Thompson Hobbs<sup>1</sup>, Kathleen Misajon<sup>2</sup>, and Ryan Monello<sup>2</sup>

December 15, 2018

<sup>1</sup>Conservation Science Partners Inc., Fort Collins, CO and Natural Resource Ecology Laboratory,  
Colorado State University, Fort Collins, CO

<sup>2</sup>National Park Service, Hawaii Volcanoes National Park, Hawaii National Park, HI

# Contents

<b>1</b>	<b>Overview</b>	<b>3</b>
<b>2</b>	<b>Data</b>	<b>3</b>
2.1	Capture histories . . . . .	3
2.2	Covariate data . . . . .	4
2.2.1	Age classes . . . . .	4
2.2.2	Age and sex . . . . .	4
2.2.3	Locations . . . . .	5
2.2.4	Total counts . . . . .	5
<b>3</b>	<b>Modeling framework</b>	<b>5</b>
3.1	Jolly-Seber state-space model . . . . .	5
3.1.1	Formulation . . . . .	5
3.1.2	Data augmentation . . . . .	7
3.1.3	Defining sampling occasions . . . . .	9
3.1.4	Including individual variation in model parameters . . . . .	9
3.2	Integrated population model . . . . .	14
3.2.1	Overview . . . . .	14
3.2.2	Population model for count and age-classification data . . . . .	16
3.2.3	Integration of Jolly-Seber model and population model . . . . .	18
3.3	Model implementation . . . . .	18
<b>4</b>	<b>Unresolved issues</b>	<b>18</b>

## 1 Overview

This document outlines procedures for modeling dynamics of the population of Nēnē in Hawai‘i Volcanoes National Park. The goal of the modeling is to use National Park Service mark-resight, survey, mortality, and spatial data to estimate the Nēnē population size and age and sex specific vital rates over time. Estimates will include uncertainty about the ecological processes governing population growth, uncertainty arising from variation in vital rates among individuals, as well as uncertainty in the data used to fit models. Models will be used to forecast the future trajectory of the population and to evaluate alternatives for population monitoring and management. Inference will be made supporting Nēnē conservation in Hawai‘i Volcanoes National Park and other important conservation areas on Hawai‘i island.

The purpose of this document is to provide a basis for discussion among project participants, a discussion that will be used to guide future model development, implementation, and inference. The document provides information on data formats required for analysis and technical details on the initial modeling framework.

## 2 Data

### 2.1 Capture histories

The data include capture histories of marked individuals visually observed over time. We will initially structure models using a yearly interval between observation times, but multiple observation times per year will be considered, as discussed below (subsection 3.1.3). We assume that an observation of each individual at each capture occasion can be described by a four element vector  $\mathbf{y}_{i,t}$  containing three zeros and a single one. The position of the one indicates the type of observation, that is  $\mathbf{y}_{i,t} = (1, 0, 0, 0)'$  indicates individual  $i$  was observed alive at time  $t$ ;  $\mathbf{y}_{i,t} = (0, 1, 0, 0)'$  indicates it was observed dead,  $\mathbf{y}_{i,t} = (0, 0, 1, 0)'$  indicates that an animal previously marked was not observed at time  $t$ , and  $\mathbf{y}_{i,t} = (0, 0, 0, 1)'$  indicates the animal was not in the study at time  $t$ . A full history of captures for individual  $i$  consists of a matrix  $(\mathbf{Y}_i)$  with four rows and  $t = 1, \dots, T$  columns (Figure 1 where  $T$  is the total number of years in the study).

Table 1: Example format of capture history data for a single individual. The individual was marked in year two, not observed in year three, observed alive in years four and five, and not observed thereafter.

Individual $i$	Year, $t$									
Observation	1	2	3	4	5	.	.	.	.	$T$
Alive	0	1	0	1	1	0	0	0	0	0
Dead	0	0	0	0	0	0	0	0	0	0
Unobserved	0	0	1	0	0	0	0	0	0	0
Out of study	1	0	0	0	0	0	0	0	0	0

The leftmost one in row one of the matrix occurs during the year the individual was first marked and the leftmost ones in row four indicate “left censoring” of the data. A capture history for the entire population consists of an  $4 \times T \times M$  array  $\mathbf{Y}$  where  $M$  is the total number of animal marked during the  $T$  years of the study.

## 2.2 Covariate data

### 2.2.1 Age classes

Covariates will include an  $M \times T$  matrix  $\mathbf{S}$  of stages  $s_{i,t} = 1$  if individual  $i$  was less than one year when observed,  $s_{i,t} = 2$  if the individuals was greater than or equal to one year old. We will assume that there is no error in classification of individuals into these categories. Missing data resulting from failure to classify an individual known to be present will be entered as NA.

### 2.2.2 Age and sex

We will use an  $M \times 4$  data frame  $\mathbf{A}$  containing information on age and sex. The element  $a_{i,1}$  contains the individual’s known age at the time of last capture and contains NA if the observed individuals’s age was not known at its last capture. Known ages will presumably come from individuals that are young of the year when they were marked. Elements  $a_{i,2}$  will contain the total number of occasions between the first and last capture of the individual. Note that this will equal the age at last capture for birds that were marked during their first year of life and will be less than the bird’s age for those of unknown age at last capture. There should be no missing data in this column. Elements  $a_{i,3}$  will contain the individual’s sex if known and NA otherwise. Elements  $a_{i,4}$  will contain the location where the individual was first marked. It is important that these locations differentiate between sites in the park and outside the park to support inference on emigration.

### 2.2.3 Locations

Locations will be recorded using codes for sites where surveys were conducted. These codes will populate the  $M \times T$  matrix  $\mathbf{L}$ . The site code will be entered in  $l_{it}$  when an individual is observed at that site and will be NA otherwise. A separate vector  $\mathbf{xy}$  will give the spatial coordinates (preferably UTM's) of the centroid of each survey site.

### 2.2.4 Total counts

The total number of individuals in each age class will be recorded site, month, year, and observer (Table 2).

Table 2: Example structure of record for count data.

year	month	site	observer	young of year	adults
2018	6	lagoon	Monello	67	204

## 3 Modeling framework

The modeling framework will include two components, a Jolly-Seber state-space model (Royle and Dorazio, 2008; Royle, 2009a,b; Kéry and Schaub, 2012), and a discrete time, state-space population model. The two components will be combined in an integrated population model (Schaub et al., 2007; McCrea et al., 2010; Schaub and Abadi, 2011). We will give an overview of the initial approach to modeling here, occasionally identifying directions for work that will occur later in the project.

### 3.1 Jolly-Seber state-space model

#### 3.1.1 Formulation

The classic Jolly-Seber model supports inference on survival, additions, and total population size from capture history data<sup>1</sup>. The model works this way. Recall that the total number of marked individuals =  $M$  and define  $N$  as the total population size of animals that were ever alive during the  $T$  years of observations<sup>2</sup>. Presume for the moment that capture histories include *all* of the individuals in the population, which is to say that  $M = N$ . We will relax this heroic assumptions subsequently.

---

<sup>1</sup>Recruitment is traditionally defined as the number of young of the number of individual s born into the population that survive to their first birthday. Recruitment defined this way requires the assumption that the population is closed to immigration. In open populations, recruitment includes surviving young of the year as well as immigrants. We will use the term “additions” to include recruits and immigrants.

<sup>2</sup> $N$  is often referred to as the super-population.

We notate the true state of the individual at time  $t$  using a three element vector  $\mathbf{z}_{i,t} = (1, 0, 0)'$  if it is alive at time  $t$ ;  $\mathbf{z}_{i,t} = (0, 1, 0)'$  if it is dead at time  $t$ , and  $\mathbf{z}_{i,t} = (0, 0, 1)'$  if it is not in the study. Transitions among states are governed by the matrix  $\mathbf{S}$ :

$$\mathbf{S} = \begin{pmatrix} \phi & 0 & \gamma_t \\ 1 - \phi & 1 & 0 \\ 0 & 0 & 1 - \gamma_t \end{pmatrix}, \quad (1)$$

where  $\phi$  is the probability that an individual that is alive at time  $t - 1$  survives to time  $t$  and  $\gamma_t$  is the probability that an individual enters the study at time  $t$ . The probabilities of state change for individual  $i$  at time  $t$  are given by the product  $\mathbf{z}_{i,t} = \mathbf{S}\mathbf{z}_{i,t-1}$ . So, for example an animal that is alive at time  $t - 1$  will be alive at time  $t$  with probability  $\phi$  and dead with probability  $1 - \phi$ <sup>3</sup>.

We cannot perfectly observe the true state of the individual because of ambiguity about the ones in row three in the capture history, which could arise because the animal is dead and unobserved or alive and unobserved. This requires a way to probabilistically relate the true state to what we are able to observe. We do so as follows. We notate an attempted observation<sup>4</sup> of animal  $i$  at time  $t$  using the vector  $\mathbf{y}_{i,t}$  as described above (subsection 1). We define an observation matrix as

$$\mathbf{D} = \begin{pmatrix} p & 0 & 0 \\ 0 & q & 0 \\ 1 - p & 1 - q & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (2)$$

where  $p$  is the probability of observing animal  $i$  at time  $t$  conditional on it being alive and  $q$  is the probability of observing animal  $i$  at time  $t$  conditional on it being dead.

We are now equipped to model the population using

$$\mathbf{y}_{i,t} \mid \mathbf{z}_{i,t}, \mathbf{D} \sim \text{multinomial}(1, \mathbf{D}\mathbf{z}_{i,t}) \quad (3)$$

$$\mathbf{z}_{i,t} \mid \mathbf{z}_{i,t-1}, \mathbf{S} \sim \text{multinomial}(1, \mathbf{S}\mathbf{z}_{i,t-1}) \quad (4)$$

$$\mathbf{z}_{i,1} \mid \gamma_1 \sim \text{multinomial}(1, (\gamma_1, 0, 1 - \gamma_1)') \quad (5)$$

---

<sup>3</sup>It is perhaps reassuring that the probability that an animal that is dead remains dead is 1.

<sup>4</sup>Notation follows the definitions established the data section 2, above.

<sup>5</sup>. Initial conditions are specified as  $\mathbf{z}_{i,1} \mid \gamma_1 \sim \text{multinomial}(1, (\gamma_1, 0, 1 - \gamma_1)')$ . The population size at time  $t$  is simply  $n_t = \sum_{i=1}^M z_{1,i,t}$ . It might be helpful to understand this model to realize that a random draw from multinomial distribution with a sample size of one, like those above, returns a vector of length equal to the probability vector (e.g.  $\mathbf{Dz}_{it}, \mathbf{Sz}_{i,t-1}$ ) containing zeros and a single one. The position of the one in the vector depends probabilistically on  $\mathbf{Dz}_{it}$ , and  $\mathbf{Sz}_{i,t-1}$ .

The Bayesian posterior and joint distribution for this model are:

$$[\phi, \gamma, p, q, \mathbf{Z} \mid \mathbf{Y}] \propto \prod_{i=1}^M \prod_{t=2}^T [y_{i,t} \mid \mathbf{Dz}_{i,t}] [\mathbf{z}_{i,t} \mid \mathbf{Sz}_{i,t-1}] [\mathbf{z}_{i,1} \mid \gamma_1] \quad (6)$$

$$\times [\phi] [\gamma_t] [p] [q]. \quad (7)$$

The model accounts our failure to know the true state of unobserved animals that were in the study and were not observed on any subsequent occasion.

### 3.1.2 Data augmentation

Of course, we do not know the total population size  $N$  because in reality  $M < N$ . How can we find the population size during each year  $n_t$  and the super-population  $N$ ? We offer some intuition for how we obtain this estimate here. For statistical details see Royle et al. (2007) and Royle (2009a). Imagine for simplicity that we sample and mark a population and observe it on two subsequent occasions and that we do not observe any dead animals. There are four possible capture histories for each individual. The matrix

$$\mathbf{Y}_i = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad (8)$$

---

<sup>5</sup>At the risk of getting ahead of ourselves, all of the parameters in this model can be modeled as functions of covariates, which we will describe in sub-section 3.1.4. This will require a  $t$  subscript on the matrices  $\mathbf{D}$  and / or  $\mathbf{S}$ .

gives the case when the individual is marked on occasion one and observed on occasions two and three. A marked individual observed only on occasion three is given by

$$\mathbf{Y}_i = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (9)$$

and a marked individual observed only on occasion two is

$$\mathbf{Y}_i = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}. \quad (10)$$

The case of an individual who is *never* observed is

$$\mathbf{Y}_i = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{pmatrix}. \quad (11)$$

It is important to understand that this “out of study” capture history (equation 11) represents *an individual that is a member of the population but that was not marked*. It is easy to see that estimating the total population requires us to find the number of these capture histories of unmarked individuals. The sum of the “out of study” capture histories and  $M$  is the true population size.

The Bayesian approach to this problem is to *augment* the observed capture histories with a large number  $n_z$  of all “out of study” capture histories such that the data array now has dimensions  $4 \times T \times (M + n_z)$ . We can specify a model capable of estimating  $N$  as

$$\begin{aligned} [\phi, \gamma, p, q, \mathbf{Z} \mid \mathbf{Y}] &\propto \prod_{i=1}^{M+n_z} \prod_{t=2}^T \overbrace{[\mathbf{y}_{i,t} \mid \mathbf{D}\mathbf{z}_{i,t}]}^{\text{observation}} \overbrace{[\mathbf{z}_{i,t} \mid \mathbf{S}\mathbf{z}_{i,t-1}]}^{\text{survival and additions}} \overbrace{[\mathbf{z}_{i,1} \mid \gamma_1]}^{\text{initial conditions}} \\ &\times [\phi] [\gamma_t] [p] [q]. \end{aligned} \quad (12)$$



A sensible person would be entitled to view this as voodoo. However, the key to understanding how data augmentation works is to realize that the  $M$  capture histories contain *all* of the information that is required to estimate  $\phi$ ,  $\gamma_t$ ,  $p$ , and  $q$  and hence,  $\mathbf{z}_{i,t}$ . Once we know these parameters we can approximate the population size as  $n_t = \sum_{i=1}^N z_{1,i,t}$  and  $N = \sum_{t=1}^T n_t$ . It may help to see that  $\gamma_t$  will get smaller as the somewhat arbitrary quantity  $n_z$  gets larger and to understand the term labeled addition in equation 12. This term requires that an individual is “available” to be added to the population if and only if its true state is “out of study.” An “available” individual is added to the population with probability  $\gamma_t$ .

Note that the equivariance property of Markov chain Monte Carlo means that quantities that are functions of random variables become random variables with their own marginal posterior distributions (Hobbs and Hooten, 2015). It follows that we will obtain distributions of  $n_t$  and  $N$  reflecting all uncertainty in the data, parameters, and process.

### 3.1.3 Defining sampling occasions

It is necessary to align events in the model of unobserved survival and additions with the timing of observations (equation 12). Observations of the Nēnē population in Hawai‘i Volcanoes National Park cluster during two time intervals, the breeding season (approximately December - February ????) and the subsequent summer (June and July ????). We will initially model the population using the summer observation period as a single, annual sampling occasion because most birds are congregated in areas searched during that period. Next, we will include both observation periods in the data used to fit models, scaling survival to the time between periods (e.g.,  $\phi^{\frac{1}{2}}$  = the probability of survival for six months) and estimating separate detection probabilities for each period. It will probably be important to model individual variation in survival and detection probability among individuals at each sampling period. For example, the location where an animal was first marked will likely influence detection probability during the breeding season because some areas where animals were marked are too remote to sample routinely. We discuss approaches for representing individual variation in the subsequent sections.

### 3.1.4 Including individual variation in model parameters

We now expand the model to exploit other information about the population and the conditions where it lives. The model above assumes survival detection probabilities are constant across indi-

viduals and time. We relax that assumption by including data on age and sex and well as other covariates to predict variation among individuals. We will emphasize understanding variation in survival, however, the same approaches can be applied to probabilities of detection. how survival probability varies among individuals, locations, and years.

The simplest way to model individual heterogeneity is to know it exists without trying to explain it with covariates. For example, we relax the assumption we made above that survival probability is constant among individuals by allowing each individual to have its own survival probability drawn from a distribution of survival probabilities,

$$\phi_i \sim \text{beta}(\alpha, \beta) \quad (13)$$

$$\alpha \sim \text{gamma}(.001, .001) \quad (14)$$

$$\beta \sim \text{gamma}(.001, .001). \quad (15)$$

We make inference on the population average survival by approximating the marginal posterior distribution of the derived quantity  $\mu_\phi = \frac{\alpha}{\alpha + \beta}$ . We could model detection probabilities in the same fashion. The posterior and joint distributions including individual heterogeneity in survival is

$$[\phi, \alpha, \beta, \gamma, p, q, \mathbf{Z} \mid \mathbf{Y}] \propto \prod_{i=1}^{M+n_z} \prod_{t=2}^T [\mathbf{y}_{i,t} \mid \mathbf{D}\mathbf{z}_{i,t}] [\mathbf{z}_{i,t} \mid \mathbf{S}\mathbf{z}_{i,t-1}] [\mathbf{z}_{i,1} \mid \gamma_1] \quad (16)$$

$$\times [\phi_i \mid \alpha, \beta] [\gamma_t] [p] [q] [\alpha] [\beta]. \quad (17)$$

**Age class model** It is also possible to represent individual variation using information on differences among individuals. For example, we assume that the age class of all marked individuals is known at each capture occasion. The age class of individual  $i$  at time  $t$   $s_{i,t}$  equals one if the bird was born that year and two if it has survived beyond its first year. We will call a member of age class one a juvenile and age class two an adult.

It is plausible that probability of survival of juveniles is lower than the survival of adults, that the variance in juvenile survival is greater than the variance among adults, and that adult and juvenile survival are correlated, at least weakly. We require a model with the flexibility to represent these characteristics. Define  $\phi_1$  as the survival probability of juveniles and  $\phi_2$  as the survival probability

of adults. We can model their joint distribution over time as

$$\text{logit}(\boldsymbol{\phi}_t) \sim \text{multivariate normal}(\text{logit}(\boldsymbol{\mu}_\phi), \Sigma) \quad (18)$$

$$\Sigma = \begin{pmatrix} \sigma_{\phi_1}^2 & \sigma_{\phi_1} \sigma_{\phi_2} \rho \\ \sigma_{\phi_1} \sigma_{\phi_2} \rho & \sigma_{\phi_2}^2 \end{pmatrix} \quad (19)$$

where  $\boldsymbol{\mu}_\phi$  is a two element vector containing the mean juvenile and adult survival probability and  $\Sigma$  is a variance covariance matrix composed of the standard deviations (on the logit scale) of juvenile ( $\sigma_{\phi_1}$ ) and adult ( $\sigma_{\phi_2}$ ) survival and  $\rho$  is the correlation between adult and juvenile survival.

The posterior and factored joint distribution becomes

$$[\boldsymbol{\gamma}, \boldsymbol{\phi}, p, q, \mathbf{Z}, \boldsymbol{\mu}_\phi \rho, \boldsymbol{\sigma}_\phi \mid \mathbf{Y}, \mathbf{X}] \propto \prod_{i=1}^{M+n_z} \prod_{t=2}^T [\mathbf{y}_{i,t} \mid \mathbf{D} \mathbf{z}_{i,t}] [\mathbf{z}_{i,t} \mid \mathbf{S} \mathbf{z}_{i,t-1}] [\mathbf{z}_{i,1} \mid \boldsymbol{\gamma}_1] \quad (20)$$

$$\times [\text{logit}(\boldsymbol{\phi}_t) \mid \text{logit}(\boldsymbol{\mu}_\phi), \Sigma] \quad (21)$$

$$\times [\boldsymbol{\gamma}] [p] [q] [\rho] [\boldsymbol{\mu}_\phi] [\boldsymbol{\sigma}_\phi]. \quad (22)$$

Note that the subscript  $s_{i,t}$  in equation 20 is the age class of individual  $i$  at time  $t$ . The value of  $s_{i,t}$  determines which element of the vector  $\boldsymbol{\phi}_t$  is used as the survival probability for individual  $i$  at time  $t$ .

We assume that the age class (juvenile or adult) of all observed individuals is known. However, age classes of the augmented records are unknown. Missing age class data are *imputed* (Gelman and Hill, 2009, Chapter 25). Data imputation treats missing observations as random variables drawn from distributions that are informed by the data we have in hand. For example, to assign age classes to augmented records, we will use

$$\psi_1 \sim \text{uniform}(0, 1) \quad (23)$$

$$\psi_2 = 1 - \psi_1 \quad (24)$$

$$s_{i,t} \sim \text{categorical}(\boldsymbol{\psi}). \quad (25)$$

The individual's age class will be incremented once at time  $t + 1$  if it was initially assigned  $s_{i,t} = 1$

at time  $t$ . A more detailed example of data imputation is provided in when we discuss age as a covariate, below.

This formulation (equation 20) makes the initial, simplifying assumption that each age class has a single survival probability that varies among years. This assumption can be relaxed by modeling heterogeneity in survival as a beta distributed random variable with mean  $\phi_{s_{i,t}}$ . Alternatively, this variation can be modeled using covariates, as described next.

Additional individual and time specific covariates (e.g., sex, location of first capture, population size at time  $t-1$ , and weather variables believed to influence survival  $\mathbf{x}_{it}$ ) could be added to model variation among individuals and years by defining

$$g(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{x}_{i,t}, s_{i,t}) = \text{inverse logit}(\alpha_{s_{i,t}} + \mathbf{x}_{it}'\boldsymbol{\beta}) \quad (26)$$

$$\text{logit}(\phi_{i,t}) \sim \text{normal}(\text{logit}(g(\boldsymbol{\alpha}, \boldsymbol{\beta}, \mathbf{x}_{i,t}, s_{i,t})), \sigma_\phi^2) \quad (27)$$

$$\boldsymbol{\alpha} \sim \text{multivariate normal}(\text{logit}(\boldsymbol{\mu}_\alpha), \boldsymbol{\Sigma}) \quad (28)$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_{\alpha_1}^2 & \sigma_{\alpha_1}\sigma_{\alpha_2}\rho \\ \sigma_{\alpha_1}\sigma_{\alpha_2}\rho & \sigma_{\alpha_2}^2 \end{pmatrix} \quad (29)$$

where  $\boldsymbol{\alpha}$  is a two element vector giving the mean survival of each age class when the standardized covariates are zero,  $\boldsymbol{\mu}_\alpha$  is a two element vector of the means of the distributions of  $\boldsymbol{\alpha}$  and  $\boldsymbol{\Sigma}$  is the variance covariance matrix for  $\boldsymbol{\alpha}$ . The vector of covariates need not be the same for the two age classes.

It is possible to model variation in the detection probability in a similar fashion to the model of survival (e.g. Royle, 2009a).

**Age specific model** Until now we have modeled the population with two age classes. This neglects information that we have on the specific ages of some individuals, those of known age at time of first capture. We might be able to improve our inference by specifically modeling how survival changes as birds age.

Individuals first marked as adults will have unobserved ages. What do we know about their age? We know that the individual is *at least* as old as the number of sampling occasions between the first time it was captured and the last time it was captured, which we will notate<sup>6</sup> as  $a_i^{\text{n.cap}}$ .

---

<sup>6</sup>We use superscripts for clarity here, i.e.,  $a_i^{\text{n.cap}}$  and  $a_i^{\text{last}}$ . However note that  $a_i^{\text{last}}$  corresponds to  $a_{i,3}$  in the data

We also know the ages of many individuals on their last capture, which we will notate as  $a_i^{\text{last}}$ . By definition, the bird's age on its *first* capture is  $a_i^{\text{initial}} = a_i^{\text{last}} - a_i^{\text{n.cap}}$ , which of course, will be missing for individuals for which age of last capture is missing. For individuals with non-missing data  $a_i^{\text{initial}} = 0 + 1$  because birds of known age at last capture were young of the year when they were first marked<sup>7</sup>. How can we approximate  $a_i^{\text{last}}$  when observations of age at last capture are missing?

The Bayesian approach treats all unobserved quantities including missing data in exactly the same way: We learn about the quantities that are unobserved by treating them as random variables drawn from a distribution and we use the observations that we have to inform us about the parameters of that distribution, including uncertainty. Thus, we can model the ages of individuals at first capture in the presence of missing data using

$$a_i^{\text{last}} \sim \text{Poisson}(\eta)T(1, ) \quad (30)$$

$$a_i^{\text{initial}} = \max(a_i^{\text{n.cap}}, a_i^{\text{last}} - a_i^{\text{n.cap}}). \quad (31)$$

Equation 30 specifies that age at last capture is a Poisson distributed random variable left-truncated at age one with mean  $\eta$ . The posterior distribution of  $\eta$  is informed by the non-missing data for age at last capture. The distribution is truncated because we presume that individuals of age one at last capture have known ages that should be excluded<sup>8</sup>. When data are missing, i.e.,  $a_i^{\text{last}} = \text{NA}$  then the value for  $a_i^{\text{last}}$  is drawn from the truncated Poisson distribution with mean  $\eta$ . We know that the individual's age at last capture cannot be less than  $a_i^{\text{n.cap}}$ , so we compute  $a_i^{\text{initial}}$  as  $a_i^{\text{n.cap}}$  or  $a_i^{\text{last}} - a_i^{\text{n.cap}}$  depending on which is greater.

Age at time of capture  $a_{i,t}$  is required to model the effect of age on survival. We obtain  $a_{i,1}$  using

$$z_{i,1} \sim \text{Bernoulli}(\gamma_1) \quad (32)$$

$$a_{i,1} = a_i^{\text{init}} z_{i,1} \quad (33)$$

---

and  $a_i^{\text{n.cap}}$  corresponds to  $a_{i,2}$ . See subsection 2.2.2.

<sup>7</sup>We will define the age of birds during their first year as 1 rather than 0 to permit zeros to be used for “ages” of recruits that have not been added to the population.

<sup>8</sup>This is a simple example of modeling data that are missing not at random.

then we specify  $a_{i,t}$  for  $t = 2, \dots, T$  using

$$a_{i,t} = a_i^{\text{init}}(1 - z_{i,t-1}) + a_{i,t-1} + z_{i,t-1} \quad (34)$$

where  $a_{i,t}$  is set at zero when  $z_{i,t} = 0$ . The value of age  $a_{i,t}$  when the state indicator switches from  $z_{i,t} = 1$  to  $z_{i,t} = 0$  is the lifespan of the bird.

This approach allows inference based on observed ages when they are present and modeled ages when data on ages are missing. Filling in the missing observations of age from the allows us to examine the effects of age and sex on survival and include these effects in the overall Jolly-Seber model (equation 20). There are three promising approaches. We could include age as a covariate in an inverse logit model of survival. This approach has some limitations because it forces survival to change monotonically with increasing age. A better model might be a Hill function that allows a relatively constant survival until a threshold age is exceeded. Thereafter survival begins to decline asymptotically to zero. A third, proven alternative is to represent effects of age survival using a proportional hazard model. All three approaches will be explored.

## 3.2 Integrated population model

### 3.2.1 Overview

It is possible to combine the Jolly-Seber model with difference equations describing growth of Nënë population in an integrated population model (Abadi et al., 2010; Schaub et al., 2007; Schaub and Abadi, 2011). The idea of integrated population models is simple. We have a deterministic model of population dynamics

$$N_t = g(N_{t-1}, \phi, f, x_{t-1}), \quad (35)$$

where  $N_t$  is the true, unobserved abundance of the population at time  $t$ , and  $g()$  is function that computes population size at  $t$  based the population size at  $t - 1$ , survival  $\phi$ , recruitment  $f$ , and ancillary data  $x_{t-1}$  (e.g, harvest, emigration, weather) explaining variation in rate of population growth. All of the quantities in this basic model can be scalars or vectors. We develop a hierarchical model to capture our uncertainty about the process of population change and uncertainty in the

data:

$$\begin{aligned}
y_t &\sim y_t \mid N_t, \theta, \sigma_{\text{observation}}^2 \text{ Observation model} \\
N_t &\sim N_t \mid g(N_{t-1}, \phi, f, x_{t-1}), \sigma_{\text{process}}^2 \text{ Process model}
\end{aligned} \tag{36}$$

$$[f][\phi][\theta][\sigma_{\text{process}}^2][\sigma_{\text{process}}^2] \text{ Parameter model} \tag{37}$$

where  $y_t$  is an observation of population size,  $\theta$  are parameters in a model relating the observations of population size to the unobserved, true size;  $\sigma_{\text{observation}}^2$  is the variance in the fit of data model, and  $\sigma_{\text{process}}^2$  is the variance in the model of population processes. An integrated population model informs parameters in equation 36 using demographic data and models, like the mark-resight dataset and Jolly-Seber model described above (subsection 3.1).

Integrated population models are valuable for four reasons. We can exploit data that do not enter the Jolly-Seber model, for example ratios of juveniles to adults, total counts, and data on immigrants or harvest. In so doing, we can use all the data in hand to inform parameters and latent states, thereby improving our inferences (Schaub et al., 2007; Schaub and Abadi, 2011). Properly constructed, these models support true forecasts of the state of the population that can be used to evaluate the ability of management actions to meet goals for the future population. We can use integrated population models to specifically represent the effects of emigration and immigration as well as planned additions to the population. Mathematical analysis of the model can reveal synthetic quantities of interest, for example long term population growth rate and age structure.

### 3.2.2 Population model for count and age-classification data

We will use a two stage, discrete time population model to portray dynamics of juveniles and adults in the Nënë population. We will start with the basic model

$$\text{logit}(\phi_t) \sim \text{multivariate normal}(\text{logit}(\mu_\phi), \Sigma) \quad (38)$$

$$\mathbf{n}_{0,t} \sim \text{Poisson}(f_t \phi_{2,t-1}^{\frac{1}{2}} n_{2,t-1}) \quad (39)$$

$$\mathbf{n}_{1,t} \sim \text{binomial}(n_{1,t-1}, \phi_{1,t-1}) \quad (40)$$

$$\mathbf{n}_{2,t} \sim \text{binomial}(n_{2,t-1}, \phi_{2,t-1}) \quad (41)$$

$$n_{1,t} = \mathbf{n}_{0,t} + \frac{m_{1,t-1}}{p_{t-1}} \quad (42)$$

$$n_{2,t} = \mathbf{n}_{2,t} + \mathbf{n}_{1,t} + \frac{m_{2,t-1}}{p_{t-1}} \quad (43)$$

$$f_t \sim \text{Poisson}(\mu_f)$$

$$v_t \sim \text{binomial}\left(\sum_{i=1}^2 n_{i,t}, p_t\right) \quad (44)$$

$$u_t \sim \text{binomial}\left(c_t, \frac{n_{1,t}}{\sum_{i=1}^2 n_{i,t}}\right) \quad (45)$$

with symbols defined in Table 3. This formulation assumes that eggs are laid approximately six months after census, which occurs in early summer. The term for recruitment  $f_t \phi_{2,t-1}^{\frac{1}{2}}$  includes adult survival because females must survive six months from census to egg laying. We choose discrete distributions for the latent states  $\mathbf{n}$  to allow for demographic stochasticity in population growth. Binomial and Poisson distributions for the data and process models represent a starting point for model fitting that may have overly restrictive variances. The binomial will be mixed with the beta distribution and / or the Poisson mixed with the gamma distribution if model checking reveals additional variance terms are needed.



Table 3: Definitions of parameters, latent states, and data used in Nënë population model (equations 38 to 45).

Parameter	Definition
$f_t$	Number of offspring produced per adult in population that survive to census at time $t$ . We assume that model census occurs approximately six months after eggs are laid.
$\mu_f$	Mean number of offspring produced per adult.
$\phi_t$	Two element vector of mean survival probabilities of age classes $\mu_{\phi_1}$ = probability that a juvenile born at time $t$ survives to its first birthday at time $t + 1$ , $\mu_{\phi_2}$ = probability that an adult survives from time $t$ to time $t + 1$ .
$p_t$	Probability of detection of an individual conditional on it being present in the population at time $t$
$\mathbf{n}_{0,t}$	Number of new juveniles at time $t$ .
$\mathbf{n}_{1,t}$	Number of juveniles at time $t-1$ surviving to adults at time $t$
$\mathbf{n}_{2,t}$	Number of adults at time $t-1$ surviving to adults at time $t$
$\mu_\phi$	Two element vector of mean survival probability for juveniles and adults.
$\Sigma$	Variance - covariance matrix for annual survival probability
Latent state	
$\mathbf{n}_t$	Two element vector of the true number of individuals in two age classes at time $t$ , juveniles ( $n_{1,t}$ ), adults ( $n_{2,t}$ ) computed as a derived quantity from data on immigration and the vector $\mathbf{n}_t$ .
Data	
$v_t$	Total count of birds at time $t$
$u_t$	Number of juveniles classified at time $t$
$c_t$	Number of individuals classified into age classes at time $t$
$\mathbf{m}_t$	Two element vector of the observed number of immigrants at time $t$ containing number of juveniles ( $m_{1,t}$ ) and adults ( $m_{2,t}$ ).

The posterior and joint distribution are

$$n_{1,t} = \mathbf{n}_{0,t} + m_{1,t-1} \quad (46)$$

$$n_{2,t} = \mathbf{n}_{1,t} + \mathbf{n}_{2,t} + m_{2,t-1} \quad (47)$$

$$[\mathbf{f}, \Phi, \mathbf{n}, \mu_\phi, \Sigma, \mu_f \mid \mathbf{v}, \mathbf{u}] \propto n_{2,t} \prod_{t=1}^T \left[ v_t \mid \sum_{i=1}^2 n_{i,t}, p_t \right] \left[ u_t \mid c_t, \frac{n_{1,t}}{\sum_{i=1}^2 n_{i,t}} \right] \quad (48)$$

$$\begin{aligned} & \times \prod_{t=2}^T [\mathbf{n}_{0,t} \mid f_t \phi_{2,t} n_{2,t}] [f_t \mid \mu_f] \\ & \times [\mathbf{n}_{1,t} \mid n_{1,t-1}, \phi_{1,t-1}] [\mathbf{n}_{2,t} \mid n_{2,t-1}, \phi_{2,t-1}] \end{aligned} \quad (49)$$

$$\times [\phi_t \mid \mu_\phi, \Sigma] \quad (50)$$

$$\times [\mathbf{n}_1] [\mu_\phi] [\Sigma] [\mu_f].$$

### 3.2.3 Integration of Jolly-Seber model and population model

We will initially fit the two models separately using the marginal posterior distributions of survival and detection parameters obtained from the Jolly-Seber models as prior distributions for those parameters in the population model. We will then fit the two models simultaneously using a joint likelihood specified as the product of probabilities of capture history, count, and classification data conditional on model parameters and latent states (Figure 1). This formulation requires the assumption that data are conditionally independent. Simulation studies have revealed that inference is not sensitive to this assumption (Ahrestani et al., 2017).

### 3.3 Model implementation

Implementing these models will be achieved step-wise, beginning with models with minimal detail and complexity and gradually working toward the full models described here.

Marginal posterior distributions of all unobserved quantities will be approximated using the Markov chain Monte Carol algorithm implemented in R (R Core Team, 2017) and JAGS (Plummer, 2003, 2012). Convergence will be verified by visual inspection of trace plots and by the diagnostics of Brooks and Gelman (1988) and Heidelberger and Welch (1983). Model fit will be checked using posterior predictive checks (Hobbs and Hooten, 2015). Model selection and model weighting will be accomplished using indicator variable selection and posterior predictive loss (Hooten and Hobbs, 2015). The ability of alternative management actions to meet goals for the future specified by decision makers and stakeholders will be evaluated using procedures described in Raiho et al. (2015); Hobbs et al. (2015) and Ketz et al. (2016).

## 4 Unresolved issues

- **Timing of births relative to census:** Discrete time population model require specifying the time when model census occurs, which conventionally coincides with the timing of counts and observations of capture histories. The definition of model structure depends in fundamental way on the timing of census relative to the time when births and emigration occurs (Noon and Sauer, 1992). The relative timing of these events needs to be discussed. Hobbs' experience says that this is the place where *many* population models go wrong.
- **Number of age classes:** The stage-based Jolly-Seber model and the population model were

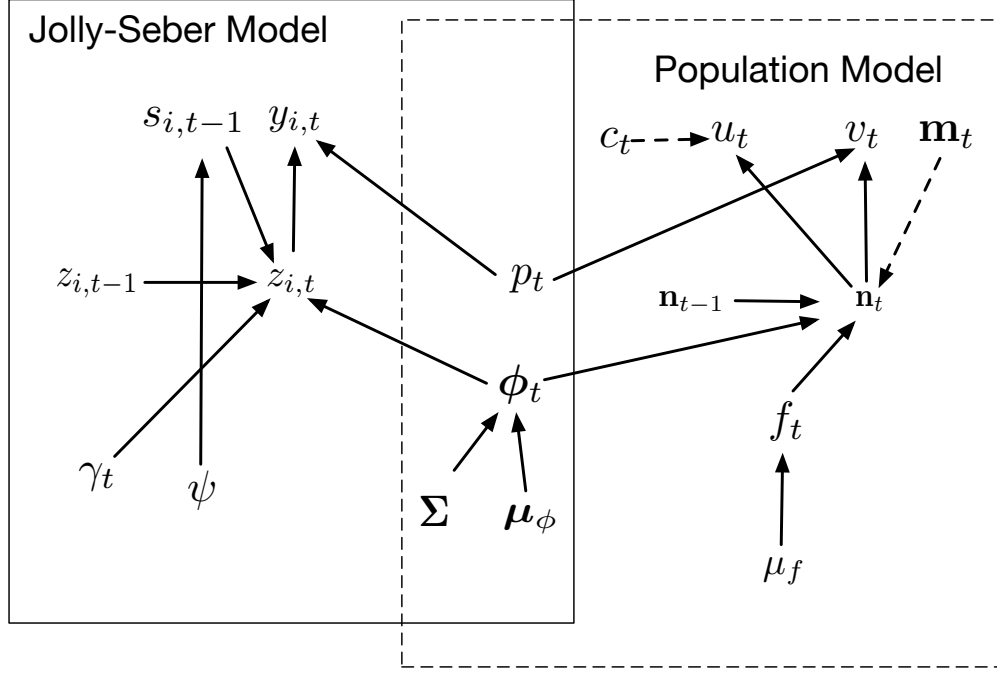


Figure 1: Directed acyclic graph of integrated population model based on the joint likelihoods for the Jolly-Seber model (equations 26 - 28) and the population model (equations -26 - 28 ). Data, parameters, and latent states for the Jolly-Seber model are shown in the solid-line box; for the matrix population model, in the dashed-line box. Data are found in the top row of the figure. Latent states occupy the middle ground and parameters occupy the lower portion of the figure (with the exception of  $p_t$ ). Solid lines represent stochastic relationships and dashed lines represent deterministic relationships. The subscript  $t$  reads “at time  $t$ ”. Definitions of symbols in the Jolly-Seber model are as follows:  $y_{i,t}$  is the observed state of individual  $i$ ,  $z_{i,t}$  is true state of individual  $i$ ;  $\gamma_t$  probability of addition of new individual;  $\phi_{i,t}$  is the probability of survival of individual  $i$ ;  $s_{i,t}$  is the observed age class of individual  $i$ ,  $\psi$  is the probability that individual  $i$  is a juvenile;  $\phi_t$  is a vector of juvenile and adult survival probabilities;  $\mu_\phi$  is the mean survival probability;  $\Sigma$  is a variance-covariance matrix for juvenile and adult survival probability, and  $p_t$  is the probability of detection. Definitions of symbols in the population model are as follows:  $c_t$  is the number of individuals classified;  $u_t$  is the number of classified individuals that are juveniles;  $v_t$  is the total count of population;  $\mathbf{m}_t$  is a two element vector of observed number of adult and juvenile immigrants added to the population at time  $t$ ;  $\mathbf{n}_t$  is a two element vector of number of the true, unobserved number of juveniles and adults;  $f_t$  number of offspring produced per adult;  $\mu_f$  is mean of number of offspring produced per adult;  $\mu_\phi$  is a two element vector of survival probabilities for each age class;  $\Sigma$  is a variance-covariance matrix for juvenile and adult survival probability; and  $p_t$  is the probability of detection.

formulated assuming two age classes could be reliably distinguished in the field: juveniles and adults. This will need to be adjusted if yearlings are identifiable or if sexes can be visually discriminated. We need to discuss if it is reasonable to model recruitment as a function of the number of adults in the population, which is necessary with a two age class model.

- **Spatial structure in the population.** Is inference on the population as a whole adequate, ignoring spatial structure, or is inference desired on spatially discrete subpopulations?
- **Missing marks.** The Jolly-Seber model requires the assumption that no individual loses its mark and that marks are not overlooked on re-sighting. We will need to model the way that marks are lost or missed if these assumptions are not plausible.
- **Data formatting.** Formatting the data for analysis will be a reasonably challenging task. Hobbs will evaluate this challenge. It may be that he can accomplish the task without unduly compressing the time he can commit to analysis.
- **Additional covariates.** There are probably weather covariates that should be included in the survival model. Covariates influencing detection probability should be discussed.
- **Likelihood for mark recapture data.** Computational challenges for achieving convergence of the integrated population model may necessitate reformulating the mark-recapture component of model as Cormack-Jolly-Seber model with a multinomial likelihood (Kéry and Schaub, 2012).

## Literature Cited

- Abadi, F., O. Gimenez, R. Arlettaz, and M. Schaub, 2010. An assessment of integrated population models: bias, accuracy, and violation of the assumption of independence. *Ecology* **91**:7–14.
- Ahrestani, F. S., J. F. Saracco, J. R. Sauer, K. L. Pardieck, and J. A. Royle, 2017. An integrated population model for bird monitoring in North America. *Ecological Applications* **27**:916–924.
- Brooks, S. P. and A. Gelman, 1988. Alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics* **7**:434–455.

- Gelman, A. and J. Hill, 2009. Data analysis using regression and multilevel / hierarchical modeling. Cambridge University Press, Cambridge, UK.
- Heidelberger, P. and P. Welch, 1983. Simulation run length control in the presence of an initial transient. *Operations Research* **31**:1109–1044.
- Hobbs, N. T., C. Geremia, J. Treanor, R. Wallen, P. J. White, M. B. Hooten, and J. C. Rhyen, 2015. State-space modeling to support management of brucellosis in the Yellowstone bison population. *Ecological Monographs* **85**:2–28.
- Hobbs, N. T. and M. B. Hooten, 2015. Bayesian models: A statistical primer for ecologists. Princeton University Press, Princeton New Jersey, USA.
- Hooten, M. B. and N. T. Hobbs, 2015. A guide to Bayesian model selection for ecologists. *Ecological Monographs* **85**:3–28.
- Kéry, M. and M. Schaub, 2012. Bayesian population analysis using WinBUGS: A hierarchical perspective. Academic Press, Waltham, MA, USA.
- Ketz, A. C., T. L. Johnson, R. J. Monello, and N. T. Hobbs, 2016. Informing management with monitoring data: The value of Bayesian forecasting. *Ecosphere* **7**:e01587–n/a.
- McCrea, R. S., B. J. T. Morgan, O. Gimenez, P. Besbeas, J. D. Lebreton, and T. Bregnballe, 2010. Multi-site integrated population modelling. *Journal of Agricultural Biological and Environmental Statistics* **15**:539–561.
- Noon, B. R. and J. R. Sauer, 1992. Population models for passerine birds: Structure, parameterization, and analysis, pages 441–464. Elsevier, London, UK.
- Plummer, M., 2003. JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. *DSC Working Papers* <http://www.ci.tuwien.ac.at/Conferences/DSC-2003/>, *Proceedings of the 3rd International Workshop on Distributed Statistical Computing, March 20-22, 2003, Technische Universität Wien, Vienna, Austria* .
- Plummer, M., 2012. rjags: Bayesian graphical models using MCMC. R package version 3-9.

- R Core Team, 2017. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- Raiho, A., M. B. Hooten, S. Bates, and N. T. Hobbs, 2015. Forecasting the effects of fertility control on overabundant ungulates. *PLOS ONE* **10**:e0143122. doi:10.1371/journal.pone.0143122.
- Royle, J. A., 2009a. Analysis of capture-recapture models with individual covariates using data augmentation. *Biometrics* **65**:267–274.
- Royle, J. A., 2009b. Analysis of capture-recapture models with individual covariates using data augmentation. *Biometrics* **65**:267–274.
- Royle, J. A. and R. M. Dorazio, 2008. Hierarchical modeling and inference in ecology: the analysis of data from populations, metapopulations, and communities. Academic Press, London, UK.
- Royle, J. A., R. M. Dorazio, and W. A. Link, 2007. Analysis of multinomial models with unknown index using data augmentation. *Journal of Computational and Graphical Statistics* **16**:67–85.
- Schaub, M. and F. Abadi, 2011. Integrated population models: A novel analysis framework for deeper insights into population dynamics. *Journal of Ornithology* **152**:227–237.
- Schaub, M., O. Gimenez, A. Sierro, and R. Arlettaz, 2007. Use of integrated modeling to enhance estimates of population dynamics obtained from limited data. *Conservation Biology* **21**:945–955.