N-mixture models with auxiliary populations and for large population abundances

by

Matthew R. P. Parker
B.Sc., University of Victoria, 2016

A Thesis Submitted in Partial Fulfilment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Mathematics and Statistics

N-mixture models with auxiliary populations and for large population abundances

by

Matthew R. P. Parker
B.Sc., University of Victoria, 2016

Supervisory Committee

Dr. L. E. Cowen, Supervisor
(Department of Mathematics and Statistics)

Dr. J. Zhou, Departmental Member
(Department of Mathematics and Statistics)

Supervisory Committee

Dr. L. E. Cowen, Supervisor
(Department of Mathematics and Statistics)

Dr. J. Zhou, Departmental Member
(Department of Mathematics and Statistics)

ABSTRACT

The key results of this thesis are (1) an extension of N-mixture models to incorporate the additional layer of obfuscation brought by observing counts from a related auxiliary population (rather than the target population), (2) an extension of N-mixture models to allow for grouped counts, the purpose being two-fold: to extend the applicability of N-mixtures to larger population sizes, and to allow for the use of coarse counts in fitting N-mixture models, (3) a new R package allowing the easy application of the new N-mixture models, (4) a new R package allowing for optimization of multi-parameter functions using arbitrary precision arithmetic, which was a necessary tool for optimization of the likelihood in large population abundance N-mixture models, as well as (5) simulation studies validating the new grouped count models and comparing them to the classic N-mixtures models.

Table of Contents

List of Tables

List of Figures

Acknowledgements

I would like to thank:

My supervisor, Dr. Laura Cowen, for her tireless support, and going above and beyond all expectations as a mentor and supervisor, and for her encouragement throughout my studies.

Dr. Julie Zhou, for her advice, and for all of her help navigating my way from an undergraduate student having taken just a single statistics course, through to the end of my MSc degree.

Dr. Rob James, for his friendship and mentorship in equal measure.

My family, friends, and June Sun, for their understanding and support, despite my frequent absences and preoccupations.

Dr. Belaid Moa, for his frequent assistance with Compute Canada resources throughout my work on this thesis.

Compute Canada, for providing computing resources without which this work would have been impossible.

and all of the VADA program faculty, for the career training, the personal engagement, and for affording me the opportunity to focus on my research by funding me with a substantial scholarship.

Chapter 1

Introduction

There are many reasons to study populations as they evolve through time. Ecologists may be interested in changing population abundances, for example to detect population decline or colony collapse for species of special concern. Disease analysts require methods for estimating total numbers of infected individuals from observed counts in open population situations. Government and health care policy makers require current, best possible estimates of human populations needing special care and services, such as populations of homeless individuals, or the population of persons suffering from undiagnosed depression. The populations of interest to researchers is thus diverse, and methods applicable to one field of study are not always applicable to others.

Any methods proposed to deal with such a diverse set of populations must necessarily be sufficiently flexible in construction. Parametric formulations allow for such flexibility both through distributional choices, and through parameter tuning. The N-mixture model, introduced by Royle (2004), is a candidate method satisfying the requirement of flexibility. However, the N-mixture model as first given, is only applicable to closed populations (those for which individuals may neither enter nor leave the population during the period of study). The limitation to closed populations was lifted by Dail and Madsen (2011), when they extended the framework to allow for population dynamics such as births, deaths, immigration, and emigration.

In the discrete count time series literature, such as in Fernández-Fontelo et al. (2016), models similar to open population N-mixtures have been developed. These models are called Integer Autoregressive order 1, or INAR(1). The mathematical framework is identical to that of N-mixtures, where the number of sites in the model is taken to be one. In this way, the INAR(1) models can be seen as a subset of N-

mixture models. Because of this relation, advances in INAR(1) research can often be extended to further research surrounding N-mixtures. Therefore, although INAR(1) models are indeed a candidate method for studying diverse populations, we will in this work consider N-mixture models only.

The form of data expected for N-mixture models is a set of discrete counts of observed individuals, one count per observation site ($i = 1, 2, \cdots, U$) per sampling occasion ($t = 1, 2, \cdots, M$), so that there are UM total data points (Royle, 2004). The observation sites are considered to be independent, and the counts are assumed to be accurate (in the sense that there is no double-counting of individuals, and no false-counts). However, the observations are not a census, as under counting is assumed to take place. This under counting is caused by imperfect detection, and often modelled with binomial thinning (for example, Belant et al. (2016), DiRenzo et al. (2019)). Under binomial thinning, the observed counts n_{it} are related to the true population N_{it} by the binomial thinning operator: $n_{it} = p \circ N_{it}$ (Fernández-Fontelo et al., 2016). In this way the observed count n_{it} is a binomial random variable with probability of success p (called the probability of detection), and with maximum value N_{it}. Population dynamics can be accommodated in many ways, and a common formulation is the one given by Dail and Madsen (2011). In this formulation, apparent survival from sampling occasion $t-1$ to t is also considered to be a result of binomial thinning, with survival probability ω: $S_{it} = \omega \circ N_{it-1}$. New population members, or apparent recruitments, are often modelled as a Poisson random variable, with recruitment parameter γ: $G_{it} = \text{Poisson}(\gamma)$. Then the true population size can be written $N_{it} = S_{it} + G_{it}$. The final parameter in the Dail and Madsen (2011) model is the initial population parameter λ, which represents the mean abundance at first sampling occasion. Usually the associated distribution for this parameter is also Poisson, so that $N_{i1} = \text{Poisson}(\lambda)$.

Chapter 2 deals with a moderately sized population of Ancient Murrelets, for which the size of the population causes large computation times for N-mixture model fitting. The classical N-mixture models are extended to incorporate use of auxiliary populations in estimation of the target population. This allows population abundance and population dynamics estimation for difficult to count populations, for which a known link exists to the auxiliary population. The computational complexities encountered in Chapter 2 motivated development of the new grouped count models which we will introduce in Chapter 3, allowing the same framework to be applied to much larger populations.

Chapter 2 is based on the manuscript titled "Estimating population abundance using counts from an auxiliary population and N-mixture models", by Matthew R. P. Parker, Vivian Pattison, and Laura L.E. Cowen, which has been submitted for publication.

Chapter 3 introduces grouped count N-mixture models, an extension of N-mixture models which use coarse counts data rather than exact counts data. This can be useful for applying N-mixtures to aggregate counts such as public health data, where data suppression can prevent the reporting of exact counts (Matthews et al., 2016). Another useful application of the new models is to artificially coarsen count data by counting groups instead of individuals. We will show in Chapter 5 that this artificial coarsening improves computation times over the original N-mixtures models.

Chapter 4 contains a discussion of arbitrary precision optimization, the relevance to grouped count maximum likelihood methods, an R package developed to aid in multi-parameter function optimization with arbitrary precision, and a second R package for the application of grouped count N-mixture models.

Chapter 5 provides a simulation study validating the grouped count models and comparing their performance to that of the classical N-mixtures models.

Chapter 2

Estimating population abundance using counts from an auxiliary population and N-mixture models

2.1 Introduction

Statistical methods of population abundance estimation play a critical role in species conservation efforts. Knowledge of current population levels, as well as both current and past population trends can be used to inform important policy and decision making processes. Accurate estimates of wildlife population abundances and trends are often essential for understanding ecosystems and for managing wildlife. Greater precision of these estimates increases confidence in study results, and is of particular importance when repeated studies are infeasible.

Developing less labour-intensive methods of producing accurate population estimates is an important goal of population biologists. This goal is being realized as computationally intensive methods become available, made feasible by the advancement of computing technologies. Evaluating the applicability and accuracy of these methods is vital in assessing their strengths and their shortcomings. This is especially true when these methods are applied to new situations and populations.

New statistical techniques should aim to improve on established methods in both accuracy and precision of estimates, while remaining cognisant of the importance of minimizing the cost of data collection. Improving the precision of estimates often involves modelling with covariates to reduce unexplained variability, or estimating

the probability of detection. N-mixture modelling methods make use of both of these potential improvements.

N-mixture models were originally developed to produce population estimates using only replicated count data on closed-populations where there are no births/deaths or immigration/emmigration (Royle, 2004). Count data suitable for N-mixture modelling are often inexpensive to collect compared with full population censuses or tagging studies. N-mixture models are being widely used for estimating population sizes and trends for many different groups of animals, such as birds (Lyons et al., 2012), reptiles (Ward et al., 2017), and large mammals (Belant et al., 2016). For N-mixture models, counts are replicated at a number of sites U and number of sampling occasions M, giving a total of UM observations (Royle, 2004). Sites are assumed to be independent, and spatially distinct. One of the key features of N-mixture models is their flexibility in model form. They accommodate a variety of detection characteristics, and allow for parameter dependency on covariates. As discussed by Dail and Madsen (2011), who extend these models to open-populations, population dynamics can be modelled in many ways, since different distributional choices for apparent recruitment and apparent survival are viable. These models are computationally very expensive, as they treat population abundance as a nuisance variable, requiring summations over very large ranges (to simulate summation to infinity). Due to these large summations, computation time grows rapidly with population abundance. Alternatively, N-mixtures can be framed as hidden Markov models (Cowen et al., 2017), however the computation issues remain. Barker et al. (2018) raised concerns regarding identifiability of abundance (N) and probability of detection (p) in closed-population N-mixture models. We address these concerns in the discussion.

Obtaining data to make population estimates can be labour-intensive and expensive when they involve extensive mark-recapture studies, or they can be inaccurate if the species involved is hard to count. For example, many seabird species are burrow-nesting, and nocturnal while on land, making individuals challenging to count (Major and Chubaty, 2012). The ability to link an auxiliary population which is more easily counted with one which would otherwise be difficult to count improves this situation.

We develop a new model to estimate population size using data from an auxiliary population under the framework of hidden Markov models. We apply this model to a case study of a colonial nesting seabird (Ancient Murrelets) where counts of the chick population serves as an auxiliary population when estimating the population of adult breeding pairs.

2.2 Materials and Methods

2.2.1 N-mixture Formulation

Data are in the same form as for N-mixture models (replicated counts over sites and sampling occasions) except that the counts are of the auxiliary population rather than the population of interest. We focus on the open-population models only, as the closure assumption is not valid for our population of study. The observed counts are assumed to be the result of a binomial thinning on the current population (Fernández-Fontelo et al., 2016) with detection probability p_{it} for site i and time t. Other thinning operators can be considered, however we focus on the binomial thinning operator as we expect detection to be a simple binomial process. In open-population N-mixture models, population dynamics are accounted for by considering the population as two distinct groups. The population abundance at site i and time t is given by $N_{it} = S_{it} + G_{it}$ (Nichols et al., 2000), where S_{it} is the number of individuals at site i who have survived from time $t-1$ to t, and G_{it} is the number of individuals recruited for site i at time t. We model S_{it} with apparent survival probability ω_{it} as Binomial(N_{it-1}, ω_{it}), and G_{it} as Poisson(γ_{it}) where γ_{it} is the recruitment rate into site i at time t. Initial abundance for each site i is modelled as a Poisson random variable with parameter λ_i, representing the initial mean abundance. Population abundance N_{it} is assumed to have the Markov property, so that $P[N_{it} = k|N_{i1}, N_{i2}, ..., N_{it-1}] = P[N_{it} = k|N_{it-1}]$. Each of the distributions mentioned are a model choice, and other choices do exist. For example a population that shows over-dispersion of counts may be better modelled with a negative binomial distribution rather than a Poisson distribution, however, there has been discussion in the literature around problematic behaviour of the negative binomial (see for example Kéry 2017). The model parameters (λ_i, ω_{it}, γ_{it}, and p_{it}) can be given additional covariate structure using appropriate link functions (such as the logit link for probability parameters: $logit(p) = \beta_0 + \sum_j \beta_j x_j$, where x_j are the covariates) (Dail and Madsen, 2011). A simple model of this form would have each parameter constant over time and across sites ($\lambda_i = \lambda$, $\omega_{it} = \omega$, $\gamma_{it} = \gamma$, and $p_{it} = p$).

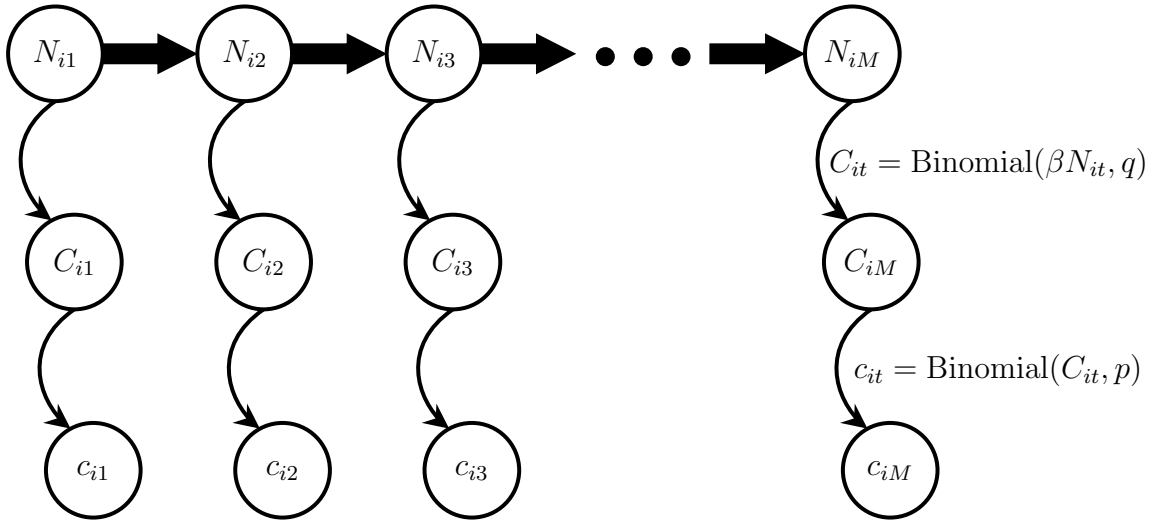
A link needs to be established between the auxiliary and target population. In the case of the Ancient Murrelets, the link between the chicks and the adult breeding pairs is the clutch size: $\beta = 2$. Consider a population N_{it} of breeding pairs. Each breeding pair from N_{it} will have a potential clutch size of two; however not all of those

eggs will be laid, cared for, and survive to become observable chicks. If we let q be the probability for each potential chick to survive and become an observable chick, and if we make the assumption that each chick's survival is independent (as suggested by Gaston, 1990), then the auxiliary chick population can be formulated as $C_{it} = \sum_{j=1}^{\beta N_{it}} Bernoulli(q)$. This is equivalent to a binomial thinning, $C_{it} = q \circ (\beta N_{it})$. The observed counts are, by the N-mixture formulation, also due to a binomial thinning. If we let c_{it} be the observed auxiliary counts, and p the probability of detection, then $c_{it} = p \circ C_{it} = p \circ q \circ (\beta N_{it}) = pq \circ (\beta N_{it})$. Thus the probability of detection p and auxiliary survival probability q (related to productivity) are confounded. By using this auxiliary population we lose the ability to obtain estimates of probability of detection, and instead our estimates will be of the inseparable product $\pi = pq$. We will refer to this π as the auxiliary probability of detection, not to be confused with the probability of detection p.

The entire process can be viewed as a layered hidden Markov model (Oliver et al., 2004) with two hidden layers. Under this view, each site produces an independent unobserved Markov chain of length equal to the number of sampling occasions, which taken together form the first layer. Transitions occur through the open N-mixture population dynamics for the adult breeding pairs, $N_{it} = S_{it} + G_{it}$. The first layer has symbols equal to the unobserved target population abundance (N_{it}). The second layer is formed by the link from target population to unobserved auxiliary population (C_{it}) as symbols. In our study this link is a binomial thinning with thinning probability q. The third and final layer of the hidden Markov chain is formed by detection thinning on the unobserved auxiliary population. This produces observed auxiliary counts (c_{it}) as symbols. The layered hidden Markov process is illustrated for a single site in Figure 2.1. Since sites are independent, the process is identical for each site.

We are interested in estimating the population of breeding pairs N_{it}, from the observed auxiliary counts c_{it}. It is important to note that for the proposed model, the population parameters being estimated are for the adult breeding pair population, and not for the auxiliary population (as the auxiliary population represents surrogate counts for the breeding pairs, and not a population in the sense of open-population N-mixture models). The detection thinning process $\pi \circ \beta N_{it}$ is still binomial in nature, and directly analogous to the detection thinning process of the usual N-mixture models. This allows us to use the standard N-mixture model fitting software `unmarked` (Fiske and Chandler, 2011) for our model fitting, being careful to correctly identify π from the fitted model parameters. When using the auxiliary population to estimate

Figure 2.1: Layered hidden Markov chain for adult breeding population N_{it} with clutch size β. Illustrates the link to the auxiliary chick population C_{it}, and to the observed chick counts c_{it} for a single site i. Transitions between N_{it-1} and N_{it} are through the open N-mixture population dynamics for the adult breeding pairs, $N_{it} = S_{it} + G_{it}$.



breeding pairs, we will be over-estimating by a factor of β, and so we need to correct our estimates by dividing them by β. We performed a simulation study to illustrate the viability of making population abundance estimates using auxiliary populations, showing that the estimated populations closely match the actual populations (see Appendix A, Web Appendix A and Appendix A, Web Figure 1).

2.2.2 Ancient Murrelet Case Study

We apply our open-population N-mixture model to a colony of the Ancient Murrelet *Synthliboramphus antiquus*, which is a colonial burrow-nesting seabird that nests in forests along shorelines of the North Pacific Ocean. They produce highly precocial chicks that, unlike almost all other seabird species, leave the nest at several days old and run to the ocean (Gaston, 1990). Since 1990, the Laskeek Bay Conservation Society (LBCS) has been collecting long-term count data of the Ancient Murrelet chicks at several locations on East Limestone Island (Gaston, 2013).

The Canadian Wildlife Service (CWS) is responsible for seabird monitoring in Canada, and conducts colony surveys to produce abundance estimates at most colonies. The CWS mapped the East Limestone Island colony boundary by running transects through the colony, and by exploration between transects (Rodway et al., 1988). The mean burrow occupancy rate was found using quadrats along those transects. Total breeding pairs for the colony were estimated from the mean burrow occupancy rate multiplied by the total colony area (Rodway et al., 1988). Since the CWS survey occurs approximately every 10 years, only two abundance estimates were made in the period from 1990 to 2006, one for 1995, and the other for 2006 (Lemon, 2007).

Clutch size is the number of eggs laid by a breeding pair. Gaston (1990) found that for the Ancient Murrelet, clutch size is almost always two. Rarely a clutch size of one was found, and this was thought to be due to the second egg being stashed in another nest, or being laid outside of the burrow. Clutch sizes of 3-4 were also observed in small quantities, however, these were likely due to multiple breeding pairs leaving their eggs in the same burrow, as evidenced by shell colour and markings.

In this study we estimated the Ancient Murrelet total breeding population for East Limestone Island, and compared our results to those of the CWS surveys. Estimation was done in two stages. First we used N-mixture models to estimate the annual abundances of an auxiliary population (the number of chicks in the LBCS sampling sites). From those estimates we then used clutch size and area expansion to estimate

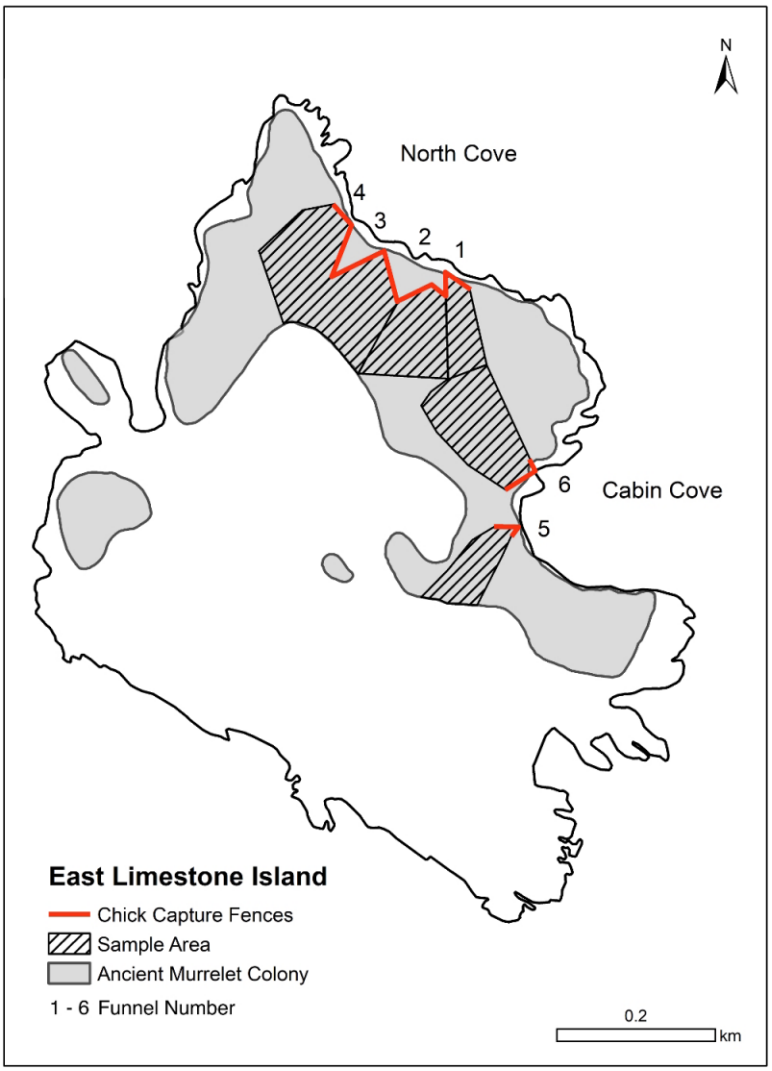
the colony's total population of breeding pairs. Our method provides the opportunity to interpolate the population abundance between CWS survey years. The LBCS data are formed from effectively sampling a large area of the colony, and allow for variability in counts to be accurately assessed both between sites, and over sampling occasions. Therefore our estimates are likely to be more representative of the colony as a whole, and of its trend over time, than the CWS transect based estimates.

To collect the annual counts of Ancient Murrelet chicks, plastic fences were set up at specific observation sites within the colony (see Figure 2.2). The fences intercept the chicks as they run downhill, and funnel them to collection points where they are counted as they leave the colony (Gaston, 1990; Gaston and Smith, 2001). Each counting site is referred to as a funnel. Two areas of the colony were sampled, North Cove and Cabin Cove, containing 4 and 2 funnels respectively. The counts were conducted nightly between 22:00 and 02:30, beginning on the 7th of May each year, and ending after 2 consecutive nights with zero captured chicks (ranging from early to late June). For our study we used data collected from 1990 to 2006, because during this period the number and location of trapping funnels were kept constant (Gaston and Descamps, 2011). The chick count data for this period are available in Appendix A, Web Table 1.

The sampled area is not consistent between funnels; to account for this, funnel fence coordinates were plotted on a map of the island, polygons for each site were delineated, and the proportion of the colony sampled by each trapping funnel was estimated using `ArcGIS` (Esri, 2014). When the chicks leave their burrows, they orient themselves using slope and the sound and light from the ocean (Gaston et al., 1988). Therefore, in order to draw the boundaries for the sampling locations, chicks were assumed to always travel downhill and towards the ocean. The funnel fences were used as the shoreward limits of the sample polygons, while topographies (such as ridge lines) were used to define the sides of the sample polygons. The uphill limit was defined by the inland colony boundary estimated by CWS (Rodway et al., 1988). The horizontal area of each polygon was calculated and then corrected using the mean slope of each polygon in `ArcGIS`. The calculated funnel areas were summed to give area estimates for North Cove (A_N) and Cabin Cove (A_C). Areas sampled (A_N and A_C) and total colony area ($A_{T,t}$) were used to extrapolate estimates to total colony abundance. Total colony area ($A_{T,t}$) was calculated by the previous colony surveys conducted by CWS in 1995 and 2006 (Lemon, 2007).

Open-population N-mixture models were fit to the Ancient Murrelet chick count

Figure 2.2: Map of East Limestone Island Ancient Murrelet colony. Capture funnels are labelled by number; 1-4 are in the area called North Cove, 5-6 in the area called Cabin Cove. The estimated area sampled by each funnel is a shaded polygon. This figure appears in color in the electronic version of this article.



data using the R package `unmarked` (Fiske and Chandler, 2011), which produces maximum likelihood parameter estimates. We allowed the model parameters to have various covariate structures. Initial abundance (λ) and auxiliary probability of detection (π) were allowed to vary by geographic location (North Cove and Cabin Cove, see Figure 2.2). This allowed the model to account for differing population densities at the two coves. We considered models in which the breeding pair survival probability and recruitment rate were each allowed to vary by location and time. Recruitment rate is taken to be independent of number of chicks in previous years. This choice is made for two reasons, the first is that new adults can join the colony from neighbouring colonies, and the second is that chicks do not immediately enter the breeding population the following year (and may choose to leave the colony before maturing). We chose to only fit models in which the auxiliary probability of detection π was not time dependent, as the method of count collection did not change over the period of study. Although there is possible overlap in funnel collection areas, we make the assumption that sites are independent. Chicks from a burrow near a funnel border may one year go down funnel A, and next year go down funnel B, however the possible overlap in funnel collection areas is small, therefore the deviation from independence should have little effect.

Bayesian information criterion (Wit et al., 2012), BIC, was used to rank the models, and this ranking was then used to select the best model. This approach showed little support for models allowing cove dependence of population dynamics γ or ω. Akaike information criterion, (Akaike, 1973), AIC, was also considered, and gave very similar rankings to BIC. In calculation of BIC, sample size was conservatively taken to be the number of sites ($U = 6$) rather than the product of number of sites with number of time replicates ($UM = 102$). The conservative sample size was chosen due to large expected overlap in the identity of breeding pairs from year to year, so that time replicates should not be considered independent (this would not be the case if we were studying, for example, a one-way population flow such as a migration).

The fitted models resulted in estimates of the auxiliary population abundance for each site and each year, $\hat{C}_{it}$. We note that the package `unmarked` does not provide SE estimates for $\hat{C}_{it}$. The 95% confidence intervals were calculated using *confint()* in the package `unmarked`, and are empirical Bayes estimates using the empirical probability distribution calculated by *ranef()* (also in the package `unmarked`). The auxiliary population estimates ($\hat{C}_{it}$) and their confidence intervals were summed over geographic location (i.e. $\hat{C}_{cove,t} = \sum_{i \in cove} \hat{C}_{it}$), producing one estimate and one confidence in-

terval for each time and each cove ($\hat{C}_{N,t}$ for North Cove, and $\hat{C}_{C,t}$ for Cabin Cove). These were then divided by the clutch size $\beta = 2$, resulting in the estimated number of breeding pairs within the sample areas ($\hat{N}_{cove,t}$).

The estimated numbers of breeding pairs $\hat{N}_{T,t}$ in the whole colony were calculated using an area expansion of the estimated number of breeding pairs in the two sampled locations (North Cove and Cabin Cove). It was necessary to assume that the sampled area abundances ($\hat{N}_{N,t}$ and $\hat{N}_{C,t}$) were proportionally representative of the total colony abundances. This is similar to the assumption made by the CWS for their 2006 survey (Lemon, 2007), in which mean burrow density across transects and mean occupancy rate for the colony were used to estimate the population of the entire colony. Our method allows for different population densities for North and Cabin Cove, which should produce more accurate area expansion estimates of total colony abundance than using a single colony mean. The total annual colony abundance estimates, $\hat{N}_{T,t}$, were calculated as in Equation 2.1. A_N and A_C are respectively the North Cove and Cabin Cove sampling areas. $A_{T,t}$ is the total colony area for year t. See Appendix A, Web Appendix B for details of the short derivation for Equation 2.1.

$$\hat{N}_{T,t} = \frac{A_{T,t}(\hat{C}_{N,t} + \hat{C}_{C,t})}{(A_N + A_C)\beta} \tag{2.1}$$

The `unmarked` package relies on the approximate Hessian matrix computed by the optimization algorithm `optim` in order to determine standard errors for parameter estimates. However, the approximate Hessian matrix method has several shortcomings. The method fails to provide estimates in cases where the estimated Hessian is not invertible, or when parameter estimates are near the boundary of the parameter space, and the method provides worse estimates for flatter likelihood functions. For this reason, we developed a parametric bootstrap estimator for standard errors for the parameter estimates. See Appendix A, Web Appendix C for the parametric bootstrap algorithm.

2.2.3 Case Study Results

We fit several open-population models, and we refer to the various models using notation similar to Lebreton et al. (1992). The notation involves a list of parameters with subscripts representing time dependence t, cove dependence c, and no subscript referring to a constant parameter. For example, model $\{\lambda_c, \gamma_t, \omega_t, p_c\}$ refers to the

Table 2.1: Model parameters, log-likelihood (ℓ), number of parameters (K), Akaike information criterion (AIC), ΔAIC, Bayesian information criterion (BIC), and ΔBIC. Parameter covariates are indicated by subscript (t for time dependence, c for cove dependence).

Model Parameters	K	$\boldsymbol{\ell}$	**AIC**	Δ_{AIC}	**BIC**	Δ_{BIC}
λ_c, γ_t, ω_t, π_c	36	-482.44	1036.9	0	1029.4	0
λ_c, γ_c, ω_t, π_c	22	-501.92	1047.8	10.9	1043.3	13.9
λ_c, γ, ω_t, π	20	-509.45	1058.9	22.0	1054.7	25.3
λ_c, γ, ω_t, π_c	21	-509.43	1060.9	24.0	1056.5	27.1
λ_c, γ, $\omega_{c,t}$, π	21	-524.19	1090.4	53.5	1086.0	56.6
λ_c, γ, $\omega_{c,t}$, π_c	22	-528.62	1101.2	64.3	1096.7	67.3
λ_c, $\gamma_{c,t}$, ω, π_c	22	-540.54	1125.1	88.2	1120.5	91.1
λ_c, $\gamma_{c,t}$, ω_c, π_c	23	-539.75	1125.5	88.6	1120.7	91.3

model where λ is cove dependent, γ is time dependent, ω is time dependent, and p is cove dependent. The final parameter estimates of λ, ω, γ, and p are used to derive estimates of the population abundance for each site and time. This can be done through an iterative process using the parameter estimates, or through empirical Bayes estimation (Royle, 2004; Dail and Madsen, 2011).

Of the fitted models, the best performing model as selected by BIC had the initial abundance parameter λ and auxiliary detection probability π both cove dependent (Table 2.1), and apparent recruitment parameter γ and apparent survival probability ω were both time dependent. Parameter estimates for the model $\{\lambda_c, \gamma_t, \omega_t, \pi_c\}$ are given in Table 2.2, along with bootstrap standard error estimates.

The estimated number of chicks in the observation area as estimated by the $\{\lambda_c, \gamma_t, \omega_t, \pi_c\}$ model are plotted per year in Figure 2.3. Shown in the same figure are the total estimated breeding pairs for the entire colony, 901 for 1995 and 761 for 2006. These were calculated using Equation 2.1, and the area data contained in Appendix A, Web Table 3. The error bars for our estimates are 95% posterior density confidence bands, and do not account for error in the estimates of total colony and site areas, similar to the estimates of the CWS which also did not include those sources of error (Lemon, 2007). The estimates are compared with the CWS survey estimates for those years, which are shown in Figure 2.3 with one standard deviation error bars. The 95% confidence band for our 1995 estimate falls below the CWS estimate, whereas the 95% confidence band for our 2006 estimate lies above the CWS estimate.

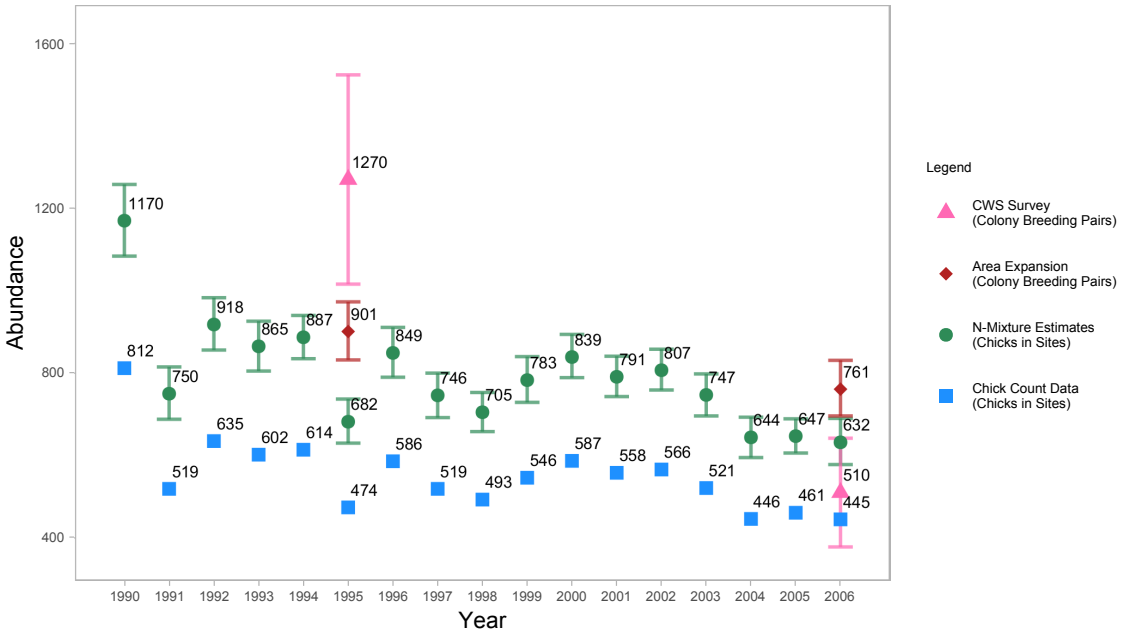
A parametric bootstrap of 1000 samples was used to produce standard error es-

Table 2.2: Parameter estimates for model $\{\lambda_c, \gamma_t, \omega_t, \pi_c\}$. Standard error estimates (SE) were obtained using 1000 samples of a parametric bootstrap.

Parameter	**Estimate**	**SE**
$\lambda_{NorthCove}$	159.49	6.37
$\lambda_{CabinCove}$	262.17	22.19
ω_{1990}	0.64	0.02
ω_{1991}	1.00	0.00
ω_{1992}	0.88	0.07
ω_{1993}	1.00	0.00
ω_{1994}	0.77	0.02
ω_{1995}	1.00	0.00
ω_{1996}	0.88	0.01
ω_{1997}	0.94	0.01
ω_{1998}	0.95	0.07
ω_{1999}	1.00	0.00
ω_{2000}	0.94	0.01
ω_{2001}	0.97	0.03
ω_{2002}	0.90	0.04
ω_{2003}	0.87	0.02
ω_{2004}	1.00	0.00
ω_{2005}	0.76	0.18

Parameter	**Estimate**	**SE**
$\pi_{NorthCove}$	1.00	0.00
$\pi_{CabinCove}$	0.63	0.06
γ_{1990}	0.00	0.00
γ_{1991}	27.11	2.49
γ_{1992}	9.97	8.97
γ_{1993}	3.67	0.94
γ_{1994}	0.00	0.00
γ_{1995}	27.12	2.46
γ_{1996}	0.00	0.00
γ_{1997}	0.00	0.00
γ_{1998}	18.17	7.13
γ_{1999}	9.97	1.55
γ_{2000}	0.00	0.00
γ_{2001}	6.69	3.48
γ_{2002}	3.67	5.33
γ_{2003}	0.00	0.00
γ_{2004}	0.00	0.00
γ_{2005}	22.20	18.12

Figure 2.3: Annual population abundance estimates for model $\{\lambda_c, \gamma_t, \omega_t, \pi_c\}$ for Ancient Murrelet chicks and breeding pairs on East Limestone Island 1990-2006. For comparison, the chick count data is shown as well as the Canadian Wildlife Service (CWS) survey estimates (Lemon, 2007). Error bars for CWS Survey show standard errors, while all other error bars show 95% confidence bands. This figure appears in color in the electronic version of this article.



timates for the population parameters. The simulated true population abundance distributions were plotted against the estimated population abundance distributions (see Appendix A, Web Figure 2). The simulation shows that the abundance estimates are stable, and that that the estimated populations closely match the actual populations (for populations generated with the same parameters as those estimated for model $\{\lambda_c, \gamma_t, \omega_t, \pi_c\}$).

2.3 Discussion

Due to the relative ease and affordability of collecting count data, we studied a novel application of N-mixture modelling that could be employed to other similar populations for which a link to an auxiliary population exists. The choice of model distributions, as well as parameter covariates, allow an enormous variety of populations to be studied. We show how one population (Ancient Murrelet chicks) can be used to obtain estimates for a separate but related population (Ancient Murrelet breeding pairs) for which it is more difficult to obtain the count data. Other populations may have their own unique links to auxiliary populations, and similar methods can be employed to study those. One possible example is the breeding population of grey seals in Ireland (Ó Cadhla et al., 2013). The adult grey seals are difficult to count, whereas the pups remain near haul-out locations while young, making them much easier to count from aerial photos of the haul-out sites. In this case the auxiliary population would be the pups, where the known link is that there is at most one pup per female seal per breeding season.

The term "population" can refer to any number of geographic scales. Our methods are applicable to the study of populations globally, regionally, or locally, dependent on the sampling methodology. The population of interest in this study is specifically the population of breeding pairs of Ancient Murrelets local to the East Limestone Island colony. To extend this methodology to a larger population, such as breeding pairs in all colonies of Haida Gwaii, we would lose accuracy and precision as all of our sites are representative of only the East Limestone Island colony. We could combat this loss of fidelity by collecting chick counts for sites located in the other colonies as well. This would have the added benefit of allowing for comparison of population trends between colonies, and of reducing variability of the colony estimates where population dynamics are similar between colonies.

We found a decrease in the number of breeding pairs for the Ancient Murrelet

colony on East Limestone Island, from the year 1995 to 2006, as shown in Figure 2.3. Our estimates agree with those of the CWS colony surveys in that they both show a downward trend in population. However, our estimates show a significantly less extreme rate of population decline than that reported by the CWS for the period 1995 to 2006. An advantage of our model is that it can be applied to estimating total breeding pairs between CWS survey years. In this way the N-mixture models could be used to supplement the survey estimates, providing updated estimates more often than the CWS survey. Comparing the N-mixture estimates with the survey estimates provides a method of checking model adequacy, and of double-checking the survey results.

The CWS breeding pairs estimate for 1995 was 1270 (SE=254), which is larger than the N-mixtures based estimate of 901 (95% CI=832, 973). In 2006 the CWS estimate was 510 (SE=132). Considering that this is a colony wide estimate, the 2006 CWS estimate seems low when compared with the observed number of chicks from the sampling sites alone (445). The N-mixtures estimate for 2006 was 761 (95% CI=696, 831), substantially larger than the CWS survey would indicate, and better reflects the observed trend in chick counts. This discrepancy could indicate that the colony had shrunk in area; however, the colony area required to produce this result is $A_{T,t} = 84080m^2$. This is close to two thirds of the reported colony area for 2006. Since the CWS survey includes colony area estimates, such a large decrease in colony size is not likely to have gone unnoticed, making this explanation implausible.

Barker et al. (2018) raised concerns about identifiability of abundance in N-mixture models. It is important to note that the Barker paper dealt specifically with the closed-population formulation of N-mixture models, whereas we deal with the open-population formulation. Barker et al. (2018) showed that if the probability of detection p is allowed to vary with time, then the models become over-specified. We do not consider in this paper any models with time varying p, and so this is not a concern for us. Barker et al. (2018) make note that data generated from an N-mixture process can be indistinguishable from data generated from a model where N is not identifiable (or not even a parameter of the model). Barker et al. (2018) considered the limiting case of a Binomial(N, p) random variable, which converges to Poisson(Np) when $N \to \infty$, $p \to 0$ and Np is kept constant. While this situation does produce unidentifiable N and p, this is not a reasonable situation in cases where probability of detection is not expected to be very small. In our case we expect our auxiliary probability of detection π to be very high, since the trapping funnels force

most of the individuals towards the point of counting. Further, Barker et al. (2018) indicate that the issues are most problematic when data quality is low, or when the count data is sparse. In this study we have high confidence in the data quality, as the methods of data collection leave little room for error, preclude double-counting, and are consistent through time. We are also not dealing with sparse data, as the counts per site (funnel) are always large ($n_{it} \geq 20$). Thus, while the issues brought forward by Barker et al. (2018) are valid, they do not likely apply to this investigation.

Due to boundary estimates for some parameters, several of the fitted models had singular estimated Hessian matrices, resulting in parameter standard errors that could not be estimated using the methods employed by `unmarked`. Parameter standard errors were thus calculated using a parametric bootstrap (Efron and Tibshirani, 1993), with code available from Parker et al. (2018).

N-mixture methods provide a powerful tool for simultaneously estimating abundance and probability of detection. Open-population formulations allow estimation of abundance trends over time, including estimates of apparent survival and apparent recruitment. Novel use of auxiliary populations (such as the population of newly hatched chicks, with unobserved adults) allow N-mixture models to be applied to populations for which it is otherwise difficult to obtain accurate counts, at the cost of losing identifiability of probability of detection. Comparing the results of N-mixture models to currently established methods (such as the CWS survey methods) is a vital component of validating new modelling techniques, and uncovering their strengths and potential weaknesses.

Supplementary Materials

Web Appendices, Tables, and Figures referenced in Sections 2.2.1, 2.2.2, and 2.2.3 are available in Appendix A.

Chapter 3

Extending N-mixture models for large populations using grouped count distributions

3.1 Introduction

N-mixture models suffer from computational complexity, which prevents researchers from applying these models to larger populations, and complicates model fitting with many covariates for population dynamics parameters (which increases computing time for the optimization of the likelihood).

The primary factor influencing computation times are the population sizes. N-mixture models make use of i infinite sums computed to remove the unknown population sizes N_{it} from the likelihood. In practice a sufficiently large upper bound K on the summations is chosen such that the change in likelihood upon increasing K is negligible (Royle, 2004). The value of K required to satisfy this condition is proportional to the actual population sizes, and so larger population sizes require larger values of K. The computation time for maximizing the likelihood increases with increasing K, due both to K being the upper bound on summations (for both closed and open population likelihood functions), and to the transition probability matrices in the model likelihood being of dimension $K + 1 \times K + 1$ (for the open populations likelihood). Strictly speaking, it is the choice of K which dictates the computation times, and not the population abundance; however, since larger population sizes require larger choices of K, we will refer to the computation time as being dependent

on population size.

The summations in the likelihood function can be replaced by the product over time of discrete state matrices and transition probability matrices by casting the problem as a hidden Markov model (see for example Cowen et al. (2017)). While this can remove some computational burden, this does not solve the computation time issue, as both the discrete state matrices and the transition probability matrices grow in dimension as K increases.

Thus, N-mixture models can be easily fit to populations with small true abundances N_{it} (for example: each $N_{it} < 100$), but can be impractical to fit for large true abundances. Of course we do not know the true population size beforehand, and instead have the observed counts n_{it}. These counts being small does not guarantee that N_{it} are also small, since $N_{it} = \frac{n_{it}}{p_{it}}$, and p_{it} could be very small.

We define grouped counts as the counts data resulting from counting the number of observed groups rather than the number of observed individuals (akin to counting by fives or counting by tens). Grouped counts data has been discussed extensively in the literature, often in terms of grouping continuous random variables into discrete bins, such as in Grundy (1952). Other authors, for example Dempster et al. (1977), discuss grouped counts as resulting from missing data, where the exact counts are unknown. Heitjan and Rubin (1991) discuss the broad subject of coarse counts. They define coarse counts as count data arising from a situation where data are neither entirely present nor entirely missing. Grouped counts are a special case of coarse counts data where the coarsening is non-stochastic. In our situation, we are dealing with discrete distributions such as the binomial distribution, and either non-missing data which we would like to transform into grouped counts, or collected grouped counts data which we would like to treat appropriately in the model likelihood.

We will use the grouped counts framework of Heitjan and Rubin (1991), summarized here. Suppose a random variable of interest X is known to have distribution $f(X;\theta)$, depending on some set of parameters θ. We would like to know the likelihood of observing a grouped count $y = Y(X)$. The grouped variable Y is degenerate in terms of X, since multiple values of x will map to a single value of y. The likelihood can be calculated as in Equation 3.1, where $\mathbb{I}_{Y(x)=y}$ is the indicator function equal to 1 when $Y(x) = y$, and 0 otherwise. Importantly, the parameters θ are recoverable from the likelihood via likelihood maximization, making parameter estimation and inference possible.

$$\mathscr{L}(\theta; y) = \int_x \mathbb{I}_{Y(x)=y} f(x; \theta) dx \tag{3.1}$$

The main idea behind our proposed grouped count model is to change the unit of measurement from 1 count = 1 individual to 1 count = 1 group containing r individuals. In changing the units of measurement, we reduce the population we are estimating from N_{it} individuals to N_{it}/r groups of r individuals. One difficulty arises in reducing the N_{it}/r values to discrete values. There are many ways of doing this, such as using rounding, ceiling, or floor functions. Call the method of reduction $R(x; r)$, so that the counts x are reduced by a factor r, and then converted to discrete values. In this way, $R(X; r)$ is analogous to the grouped variable Y of Heitjan and Rubin (1991). The key difference is that our grouping may occur post data collection, so that in actuality we may be aware of the true counts, and rather choose to use a reduction to groups for computational benefit. Of course we can also use these grouped count models on data which was truly collected as course counts, where we are not aware of the exact counts. Throughout this paper, we will consider $R(x; r)$ to use the rounding function, although similar derivations will work for other functions.

3.2 Grouped count distributions (rBinom, rPois)

Converting from full counts n_{it} to grouped counts $n_{R,it} = R(n_{it}; r)$ changes the probability distributions involved in the likelihood function for the N-mixture models. Therefore it is not sufficient to simply reduce the counts with $R(x; r)$, and use those in maximizing the original likelihood. We will need to find the new likelihood corresponding to the reduced counts, and this requires defining two new probability distribution functions (as per Equation 3.1), one corresponding to the Binomial distribution, and the other to the Poisson distribution. We note that alternate formulations of the likelihood involving other parameter distributional choices are reasonable, and similar grouped count distributions could be easily derived for those as well.

3.2.1 Distribution: rBinom

Let $X \sim \text{Binom}(N, p)$, with $N > r$, and let $Y = R(X; r)$. We want to find the distribution of Y. Essentially we are binning r values of X and assigning them to a single value of Y. From Equation 3.1, we have $P[Y = y] = \sum_x \mathbb{I}_{Y(x)=y} \text{Binom}(x; N, p)$. This can easily be computed as the difference between two binomial cumulative dis-

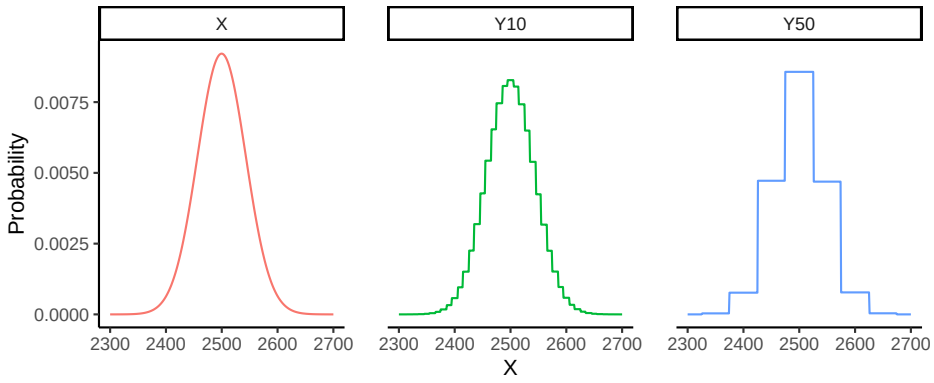


Figure 3.1: Probability distribution functions for:
$X \sim \text{Binom}(N = 10000, p = 0.25)$, $Y_{10} = R(X; r = 10)$, $Y_{50} = R(X; r = 50)$. Y_{10} and Y_{50} are examples of rBinom distributed random variables.

tribution functions, with the evaluation points dependent on choice of r. We will call the distribution of Y a reduced binomial distribution, $Y \sim \text{rBinom}(N, p, r)$. See Figure 3.1 for a comparison of the probability distribution functions of $X \sim \text{Binom}(N = 10000, p = 0.25)$, $Y_{10} \sim \text{rBinom}(N = 10000, p = 0.25, r = 10)$, and $Y_{50} \sim \text{rBinom}(N = 10000, p = 0.25, r = 50)$. Notice that as r increases, the distributions become coarser (in the sense that there are fewer available states which Y_r may obtain). Table 3.1 shows an example of the probability distribution of rBinom, and its calculation. As can be seen in the Table 3.1 example, not every group necessarily has the same size. For example, $y = 0$ is mapped to by $y = R(x; 2)$ for exactly the value $x = 0$, while $y = 1$ is mapped to for both $x = 1$ and $x = 2$.

3.2.2 Distribution: rPois

Similar to the definition of rBinom, rPois can be found using Equation 3.1. Letting $X \sim \text{Pois}(\lambda)$ be our random variable of interest, and $Y = R(X; r)$ be the reduction function applied to X, we have $P[Y = y] = \sum_x \mathbb{I}_{Y(x)=y} \text{Pois}(x; \lambda)$. We will call the distribution of Y a reduced Poisson distribution, $Y \sim \text{rPois}(\lambda, r)$. See Figure 3.2 for a comparison of the probability distribution functions of $X \sim \text{Pois}(\lambda = 2000)$, $Y_{10} \sim \text{rPois}(\lambda = 2000, r = 10)$, and $Y_{50} \sim \text{rPois}(\lambda = 2000, r = 50)$.

Table 3.1: Comparison of probability distribution of $X \sim \text{Binom}(N = 19, p = 0.25)$ with $Y_1 = R(X; r = 2) \sim \text{rBinom}(N = 19, p = 0.25, r = 2)$, and with $Y_2 = R(X; r = 3) \sim \text{rBinom}(N = 19, p = 0.25, r = 3)$. For $X \sim \text{Binom}(N, p)$, and $Y = R(X; r)$, rBinom is $\text{rBinom}(y; N, p, r) = \sum_x \mathbb{I}_{R(x;r)=y} \text{Binom}(x; N, p)$.

X	0	1	2	3	4	5	6	7	8	9
$P[X = x]$	0.004	0.027	0.080	0.152	0.202	0.202	0.157	0.097	0.049	0.020
$P[Y_1 = y]$	0.004		0.107		0.354		0.359		0.146	
$Y_1 = R(X, r = 2)$	0		1		2		3		4	

X	10	11	12	13	14	15	16	17	18	19
$P[X = x]$	0.007	0.002	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
$P[Y_1 = y]$	0.027		0.002		<0.001		<0.001		<0.001	<0.001
$Y_1 = R(X, r = 2)$	5		6		7		8		9	10

X	0	1	2	3	4	5	6	7	8
$P[X = x]$	0.004	0.027	0.080	0.152	0.202	0.202	0.157	0.097	0.049
$P[Y_2 = y]$	0.031			0.434			0.457		
$Y_2 = R(X, r = 3)$	0			1			2		

X	9	10	11	12	13	14	15	16	17	18	19
$P[X = x]$	0.020	0.007	0.002	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
$P[Y_2 = y]$	0.076			0.002			<0.001			<0.001	
$Y_2 = R(X, r = 3)$	3			4			5			6	

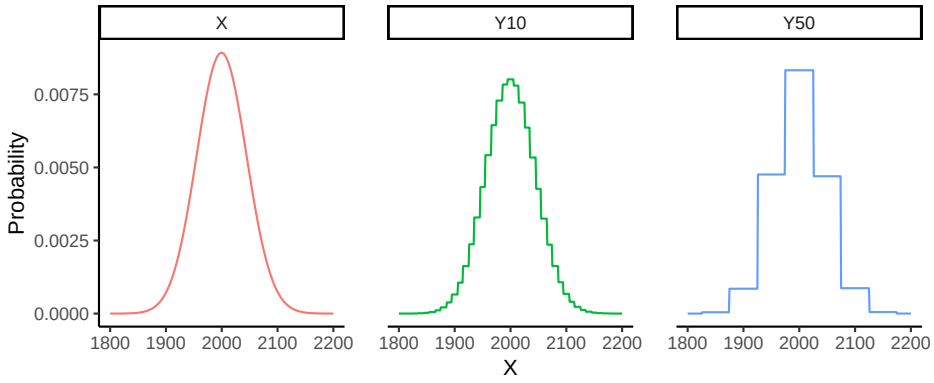


Figure 3.2: Probability distribution functions for: $X \sim \text{Pois}(\lambda = 2000)$, $Y_{10} = R(X; r = 10)$, $Y_{50} = R(X; r = 50)$. Y_{10} and Y_{50} are examples of rPois distributed random variables.

3.3 Grouped Count Likelihood Function

The likelihood functions from Royle (2004) and from Dail and Madsen (2011) can be readily updated to incorporate grouped counts. To do so, the Poisson and binomial distributions must be swapped out for their reduced count analogues. As well, the bounds of summation need to be changed to account for the grouped counts. The closed population model of Royle (2004) is shown in Equation 3.2, and the grouped count analogue is shown in Equation 3.3. Likewise, the open population model of Dail and Madsen (2011) is shown in Equation 3.4, and the grouped count analogue is shown in Equation 3.5.

In the situation where site populations are unbalanced, for example if a single site has a significantly larger population than other sites, or if some of the counts data has been suppressed, there is a simple adjustment which can be made to the likelihood function. By replacing r by r_i in the likelihood we can adjust the size of the counted groups to be site dependent. This allows, for example, sites with small observed abundances to be counted exactly ($r_i = 1$) while sites with large observed abundances can be counted using groups ($r_i > 1$). In a similar way, spikes in observed counts (or suppressed counts) over time could be dealt with by replacing r with r_{it}, allowing the group size to be both site and time dependent.

3.3.1 Closed Population Models (Royle 2004)

Definitions	
U	Number of sites
M	Number of sampling occasions
N_i	Total (observed plus unobserved) population at site i
n_{it}	Observed population at site i and sampling occasion t
p	Probability of detection
λ	Initial site abundance parameter
$N_{R,i}$	Reduced count (or grouped count) of N_i
r	Reduction factor (group size variable)
$R(x;r)$	Reduction function

Original Full Counts Model Likelihood

$$\mathscr{L}^{\text{closed}} = \prod_{i=1}^{U} \left\{ \sum_{N_i=max_t n_{it}}^{\infty} \left(\prod_{t=1}^{M} \text{Binom}(n_{it}; N_i, p) \right) \text{Pois}(N_i; \lambda) \right\} \tag{3.2}$$

Grouped Counts Model Likelihood

$$\mathscr{L}_R^{\text{closed}} = \prod_{i=1}^{U} \left\{ \sum_{N_{R,i}=max_t R(n_{it};r)}^{\infty} \left(\prod_{t=1}^{M} \text{rBinom}(n_{it}; rN_{R,i}, p, r) \right) \text{rPois}(N_{R,i}; \lambda, r) \right\} \tag{3.3}$$

3.3.2 Open Population Models (Dail and Madsen 2011)

Definitions	
U	Number of sites
M	Number of sampling occasions
N_{it}	Total (observed plus unobserved) population at site i, time t
n_{it}	Observed population at site i and sampling occasion t
p	Probability of detection
γ	Apparent recruitment parameter
ω	Apparent survival parameter
λ	Initial site abundance parameter
$N_{R,it}$	Reduced count (or grouped count) of N_{it}
r	Reduction factor (group size variable)
$R(x;r)$	Reduction function

Original Full Counts Model Likelihood

$$\mathscr{L}^{\text{open}} = \prod_{i=1}^{U} \left[\sum_{N_{i1}=n_{i1}}^{\infty} \cdots \sum_{N_{iM}=n_{iM}}^{\infty} \left\{ \left(\prod_{t=1}^{M} \text{Binom}(n_{it}; N_{it}, p) \right) \text{Pois}(N_{i1}; \lambda) \prod_{t=2}^{M} P_{N_{it-1},N_{it}} \right\} \right] \tag{3.4}$$

$$P_{a,b} = \sum_{c=0}^{min\{a,b\}} \text{Binom}(c; a, \omega) \text{Pois}(b-c; \gamma)$$

Grouped Counts Model Likelihood

$$\mathscr{L}_R^{\text{open}} = \prod_{i=1}^{U} \left[\sum_{N_{R,i1}=R(n_{i1};r)}^{\infty} \cdots \sum_{N_{R,iM}=R(n_{iM};r)}^{\infty} \left\{ \left(\prod_{t=1}^{M} \text{rBinom}(n_{it}; rN_{R,it}, p, r) \right) \text{rPois}(N_{R,i1}; \lambda, r) \prod_{t=2}^{M} P_{N_{R,it-1},N_{R,it}} \right\} \right] \tag{3.5}$$

$$P_{a,b} = \sum_{c=0}^{min\{a,b\}} \text{rBinom}(c; a, \omega, r) \text{rPois}(b-c; \gamma, r)$$

3.4 Selection of K using K_R from grouped count models

Consider that K is the upper bound of the summations in $\mathscr{L}^{\text{closed}}$ (or $\mathscr{L}^{\text{open}}$), a full counts likelihood. Let K_R be an upper bound on the summations in $\mathscr{L}_R^{\text{closed}}$ (or $\mathscr{L}_R^{\text{open}}$), a grouped counts likelihood, for which the model parameters have converged. We can estimate the optimal choice for K using the known value of K_R. Because we have access to the full counts data, we can calculate the maximum loss of precision due to grouping using Equation 3.6. A first estimate for K can then be calculated from Equation 3.7.

$$L_r = max_{i,t}|(rR(n_{it}; r) - n_{it})| \tag{3.6}$$

$$\hat{K} = rK_R + L_r \tag{3.7}$$

Thus, the procedure for finding a suitable value of K for fitting the full counts model is:

1. fit grouped count models for increasing values of K_R
2. choose a value of K_R for which the parameter estimates have stabilized
3. calculate L_r from Equation 3.6, and $\hat{K}$ from Equation 3.7
4. fit successive full counts models, starting with $K = \hat{K}$, increasing K at each iteration, until the model parameters have converged.

Since computation time increases with K, the grouped count estimates will compute significantly faster than their full count model counterparts, so that finding the initial value $K = \hat{K}$ will be more efficient when using the reduced count models. Note that the amount to increase K by in each iteration of step 4 is a choice to be made by the researcher. Smaller increments can mean more models require fitting before parameter convergence, but can result in a smaller final value of K. We recommend increasing K by r at each iteration as a reasonable trade-off.

To illustrate this process, we consider a population with a closed population N-mixtures structure. We will determine an appropriate value for K using the reduced counts K_R method. The population and observed counts were generated with $\lambda =$

Table 3.2: Example simulated population/observation set for $\lambda = 300$ initial individuals, $p = 0.35$ probability of detection, $U = 5$ sites, $M = 10$ sampling occasions, and $r = 10$ reduction factor.

True Population:

$U \backslash M$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	328	328	328	328	328	328	328	328	328	328
2	310	310	310	310	310	310	310	310	310	310
3	268	268	268	268	268	268	268	268	268	268
4	304	304	304	304	304	304	304	304	304	304
5	292	292	292	292	292	292	292	292	292	292

Observed Counts:

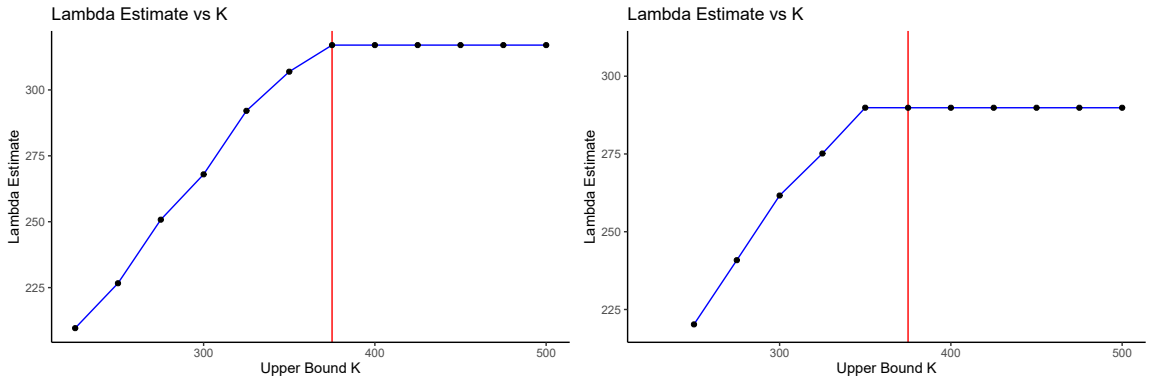
$U \backslash M$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	114	117	122	131	112	100	114	138	113	114
2	116	96	105	101	109	107	101	111	102	98
3	99	90	92	93	103	82	100	78	100	81
4	92	130	112	110	114	114	113	99	109	108
5	99	102	110	105	106	103	101	104	92	115

Coarse/Grouped Counts ($r = 10$):

$U \backslash M$	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	11	12	12	13	11	10	11	14	11	11
2	12	10	11	10	11	11	10	11	10	10
3	10	9	9	9	10	8	10	8	10	8
4	9	13	11	11	11	11	11	10	11	11
5	10	10	11	11	11	10	10	10	9	12

300, $p = 0.35$ for 5 sites and 10 sampling occasions. The population and observed counts are shown in Table 3.2, along with the grouped counts when the group size is $r = 10$.

Using a reduction factor $r = 10$, the model was fit for increasing K, and the resulting $\hat{\lambda}$ values were plotted against K. The plot is shown in Figure 3.3a. The value of K_R after which the estimates stabilized is $K_R = 37$. We calculated $L_r = 5$ from Equation 3.6, and we calculated $\hat{K} = 375$ from Equation 3.7. We then fit successive full counts models, beginning with $\hat{K} = 375$. We note that the full counts model parameters converged for $K = 375$. For the purpose of comparison with Figure 3.3a, Figure 3.3b shows $\hat{\lambda}$ against K for each value of K_R considered (rather than just the two successive values needed to show convergence).



(a) Grouped count N-Mixture model with reduction factor $r = 10$, $K_R = R(K; r)$.

(b) Full count N-Mixture model (reduction factor $r = 1$).

Figure 3.3: Plots of $\hat{\lambda}$ versus K. For this data, $\lambda = 300$, $p = 0.35$, and there were 5 sites, and 10 sampling occasions. Vertical red lines indicate $\hat{K} = 375$

3.5 Discussion

Grouped count models have several distinct uses. Grouping can be used to decrease computation times for exploratory analysis, such as for choosing an appropriate upper bound K for later fitting the full counts model. In the case where full counts models are computationally intractable, grouped counts can also be used to reduce the computational complexity to levels appropriate to the hardware and time limitations of the study. Thus, when full counts N-mixture models are not computationally feasible, grouped count models can still give population parameter estimates and population abundance estimates. As well, grouped count models can be applied directly to grouped count data.

Grouped count data can occur in publicly available public health data where exact counts would be considered a data privacy risk (see Matthews et al. (2016) for a discussion on count suppression). For example, if aggregated health data results in a table cell containing exactly one record, then it becomes feasible to link this data with other sources to identify the individual represented in the aggregate data, thus presenting a serious risk of data privacy breach. As discussed in Matthews et al. (2016), common solutions to this problem include suppressing small counts (so that only the counts above a certain threshold are reported), or rounding of all counts to the chosen group size. Both of these solutions provide coarse count data to which grouped count N-mixture models can be easily applied.

In the case of suppressing only small counts, observations with small counts could be treated with an appropriate group size (with r chosen so that all suppressed counts would map to a grouped count $R(n_{it}; r)$ of 0). All other observations could then be treated with $r = 1$, since exact counts are available. In the case for which all counts have been rounded according to the chosen group size, the data is already coarsened, and can be immediately used in a grouped count N-mixtures model (with r set to be the known group size).

Grouped count models do however have a large disadvantage. The precision of model estimates decreases as r increases, so that there is a trade off between model precision and computation time. The reduction parameter r should be chosen as small as possible in order to avoid loss of information in the likelihood. Choosing $r = 1$ is equivalent to fitting a full counts model, rather than a grouped counts model. In contrast, the feasibility of the computation of the likelihood is directly related to the size of r. If r is too small, then a large population can cause an intractable likelihood. In

general, the smaller p_{it} is, the worse the performance improvements of using grouped count models will be (since the required upper bound K may remain large). Despite this, reducing the counts via grouping will always improve computation time, and will be especially useful for detection probabilities not close to zero.

It is important to note that the parameter values found for the grouped counts model are not necessarily the same as those that will be found for the full counts model (as there is loss of precision from the grouping). See Chapter 5 for how bias can be introduced to the estimates by choosing r too large for the observed counts. As well, when using the selection of K from K_R method proposed in Section 3.4, the value $\hat{K}$ will not necessarily be equal to the minimum value of K needed for parameter convergence in the full counts models. Thus, care should be taken to always fit the full counts model for at least two successive values of K, so as to verify convergence of the parameter estimates.

It is well known that as N grows large, the limiting distributions of both binomial and Poisson are Gaussian (see for example Bain and Engelhardt (1992), chapter 7.3 'The Central Limit Theorem', pp. 236–238). For this reason, it may be beneficial to use continuous models in the large population regime. Future work should compare the discrete distribution grouped count models approach with potential continuous population model approaches.

Chapter 4

Arbitrary precision arithmetic, multi-parameter optimization, and R packages

4.1 Introduction

Computational precision is generally limited by the hardware on which software is being run, in particular the physical size of computer registers determines the number of bits available to store floating point numbers. Multiple registers can be used in conjunction to increase the number of bits available for numerical accuracy (such is the case for double precision floats, which make use of two registers to double the available bits). A 32-bit computer system groups memory into chunks of 32 bits, allowing 32 bits to be used in defining a single precision float (and 64 bits for a double precision float) (Bryant and O'Hallaron, 2011). However, there is a trade-off between computational precision and computational speed (Bailey and Borwein, 2013). In general, increasing the precision of computations comes at the cost of increased computing time. For a detailed discussion of floating point arithmetic and computer precision, see Goldberg (1991).

When calculating likelihood functions for N-mixture models, computer precision becomes a necessary consideration when the population size becomes large. The point at which the population is too large for the current level of precision is case specific, as it is a function of population size, number of sampling occasions and sampling sites, and on the particular values of the observed counts. The root of the precision problem

stems from the use of probability distribution functions, which by their nature produce smaller probabilities as the number of possible states increases. For example, as $N \to \infty$, Binomial$(x; N, p)$ shrinks to zero for all values of x. This problem is exacerbated in the N-mixture likelihood, as the product over sampling occasions can cause the product of very small probabilities, which result in even smaller probabilities. These small probabilities may then experience integer underflow.

Integer underflow occurs when a sufficiently small magnitude number is used in a computer calculation, such that the small number must be treated as zero due to the limits of precision in the calculation (Coonen, 1981). When calculating the N-mixture likelihood for large population sizes, this results in a likelihood function producing zero uniformly for all possible parameter inputs. When this occurs, the likelihood function appears flat, and so optimization algorithms cannot succeed in finding solutions which maximize the likelihood.

4.2 Arbitrary Precision Arithmetic

Arbitrary precision arithmetic (APA) is a method of overcoming the hardware limitation of register size on arithmetic precision in computer computations. There are many APA implementations available, such as MPFR (Fousse et al., 2007) and ARPREC (Bailey et al., 2002). APA uses software to create virtual registers of arbitrary size. These virtual registers allow for an incredible increase in precision by grouping many physical registers together to achieve the desired level of precision. The limit on precision in APA is thus moved from physical register size to available physical memory.

In general, APA can be hundreds or thousands of times slower than arithmetic on double precision floats (Bailey and Borwein, 2013). There are several factors influencing the increase in computation times associated with APA. Increasing the number of bits used in a calculation obviously increases the number of bits which need to be checked and updated with every calculation. As well, APA calculations take longer than they would without APA, due to the computational overhead of using software to manage the virtual registers.

An implementation of APA available as a package in the programming language R is `Rmpfr` (Maechler and Heiberger, 2020), and we use this package as a base from which to implement the algorithms necessary for maximizing multi-parameter functions.

4.3 Arbitrary Precision Optimization

Optimization of functions with arbitrary precision parameters is no different in mathematical theory than optimizing a function without arbitrary precision parameters. Thus, the same algorithms which have been developed for functional optimization can be used to optimize arbitrary precision functions. However, the implementation of these algorithms is practically different, in that APA must be used.

The R package `Rmpfr` (Maechler and Heiberger, 2020) contains a function called `optimizeR()`, which is an implementation of an APA optimization algorithm for optimizing functions of a single parameter (functions of the form $f(x)$, with x a real valued scalar). Since the N-mixtures likelihood is a multi-parameter function (of the form $f(\underline{x})$, with $\underline{x}$ a real valued vector), multi-parameter optimization algorithms are necessary. To our knowledge, there are no multi-parameter optimization algorithm implementations currently available in R for arbitrary precision.

We provide a new R package, `optimizeAPA` (Parker, 2020a), which implements the Davidon-Fletcher-Powell (DFP) algorithm (Davidon, 1959). The algorithm is implemented both as a non-APA version through the function `optim_DFP_NAPA()`, and as an APA version through the function `optim_DFP_APA()`. The key innovation of this new R package is the combination of the DFP algorithm with APA, which provides a unique tool for research involving the optimization of multi-parameter functions where very high precision is required.

4.4 Grouped Count N Mixture Model Fitting

Several points addressed in this manuscript necessitate the development of a new N-mixtures R package. Grouped count N-mixture models make use of different probability distributions than full count N-mixture models, and so possess new likelihood functions which need to be implemented. Large population sizes cause integer underflow in the N-mixture likelihood functions (whether grouped counts or full), requiring the use of arbitrary precision arithmetic to compute. Large population sizes also cause substantial increases in computation times for N-mixture model fitting, and this is heavily exacerbated by the necessity of using APA to avoid the integer underflow issues. To help alleviate the computation times involved, both grouped count methods and parallel computing methods can be employed.

N-mixture models explicitly assume that each site is statistically independent of

every other site. This leads to the separability of the likelihood into one independently calculable function per site. Parallel computing is then especially useful for speeding up the process of optimizing N-mixture models, since site calculations can be split easily across processors. While the gains from using parallel processing can be high, the reduction in computing time is heavily hardware and case dependent (and can indeed be slower for small population sizes and small numbers of sampling sites).

To address the needs for a new N-mixtures R package, we provide the package `redNMix` (Parker, 2020b). The package `redNMix` is based on the algorithms of the R package `unmarked` (Fiske and Chandler, 2011), which provides functions for fitting both closed and open population N-mixture models. `redNMix` achieves each of the points mentioned which require a new model fitting package. Grouped count (and full count) N-mixture models can be fit either with or without arbitrary precision arithmetic, and on compatible machines, parallel processing can be used to increase computation speeds. The `redNMix` package enables researchers to make easy use of the new grouped count N-mixture models, and to apply both the grouped and the full counts N-mixture models to the study of large populations, which would otherwise be inaccessible due to integer underflow problems.

Chapter 5

Simulation study for grouped count N-mixture models

5.1 Introduction

The inception of N-mixture models in Royle (2004) came with a simulation study investigating the properties of the new models. The original simulation study showed the estimated abundances to be slightly biased, with the bias lessening for larger p. The mean estimate was smaller than the true abundance for considered cases, however the median and mode were both very close to the true abundances.

We performed a set of simulations, to illustrate the properties of the grouped count models in comparison with the properties of the original full count models. For the small abundance regime, arbitrary precision arithmetic was unnecessary in optimizing the likelihood, which greatly reduced the computational burden. A similarly large set of simulations for very large abundance would not be feasible to conduct due to the combined increases in computational demands of large population likelihood calculations and of using arbitrary precision arithmetic.

5.2 Simulation Study

Here we wish to compare and contrast the original N-mixture models with the new grouped count N-mixture models. The original simulation study by Royle (2004) was illustrated in terms of the distribution of the error (estimated total abundance $\hat{N} = \sum_{i=1}^{U} \hat{N}_i$ minus the true total abundance N). We used relative error for our

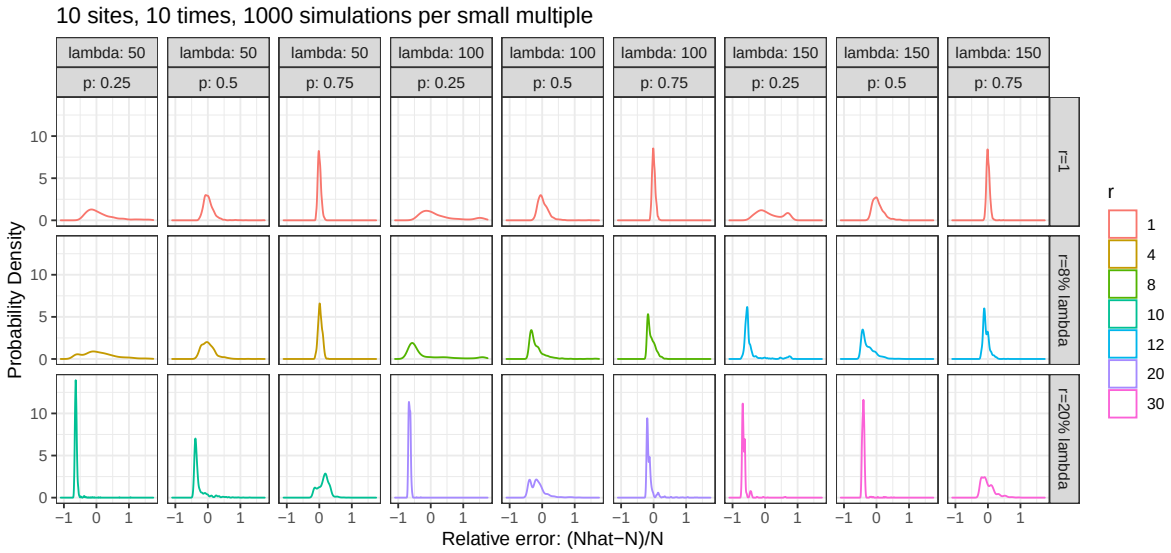


Figure 5.1: Results of fitting grouped count N-mixtures models to randomly generated population/observation pairs. Each small multiple is a distribution plot of the relative error $\frac{\hat{N}-N}{N}$ for 1000 fitted models.

set of simulations, this allowed the results to be comparable across different population parameters. We chose our parameters from $U = (5, 10, 15)$ sampling sites, $M = (5, 10, 15)$ sampling occasions, and $\lambda = (50, 100, 150)$ mean site abundance parameter. For probability of detection, we considered $p = (0.25, 0.5, 0.75)$. This choice of p was in contrast to the choice made by Royle (2004), who consider only $p = (0.25, 0.5)$. We made this choice because many population abundance applications (such as estimating undiagnosed depression, or Ancient Murrelet chick counts as described in Chapter 2) can assume relatively high rates of detection. The reduction parameter was chosen proportional to λ, with $r = (1, 0.08\lambda, 0.2\lambda)$.

Figure 5.1 illustrates the simulation results via the density distributions of the relative error $\frac{\hat{N}-N}{N}$ for 10 sites, and 10 sampling occasions. Similar figures are available in Appendix B for each of the nine combinations of U and M. Each small multiple shown in the figures accounts for 1000 simulated population/observation pairs, for a total of 243,000 simulations. The columns contain different parameter combinations for p and λ, while the rows contain different values for the reduction r.

The median relative error, median$(\frac{\hat{N}-N}{N})$, is shown in Figure 5.2, illustrating the effect on error of p, λ, and r. For $r = 1$, the only significant error occurs for $p = 0.25$. Only p and r appear to have large impact on the median error, with decreasing p and increasing r both causing increases in error. The full data table summarizing the

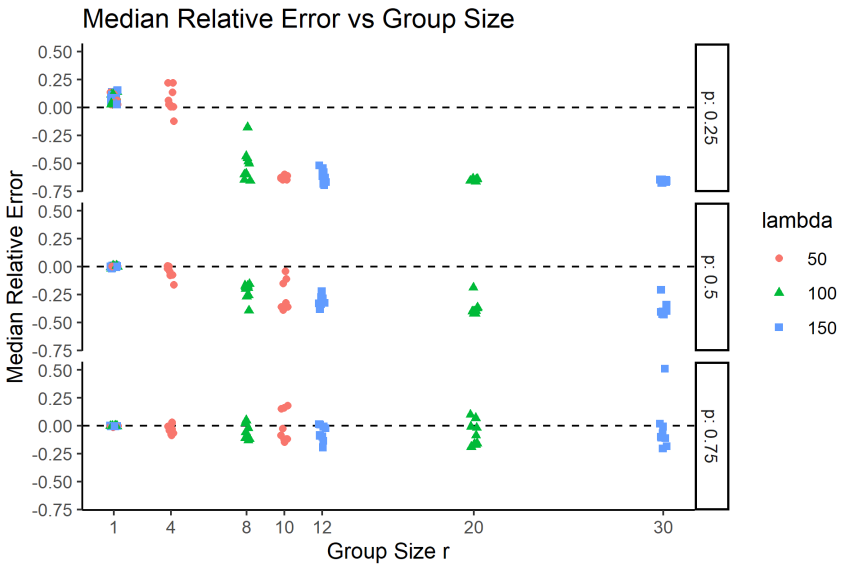


Figure 5.2: Median relative error across 243,000 simulated models.

simulated sampling distributions is available in Appendix B Table B.1.

5.3 Analysis of Computation Time

It is of particular interest to determine the effects of using grouped count models on computation time. To illustrate the potential gains over the original N-mixtures models, the computation times for the simulation study were investigated. There is high variability in computing time, however the general trend is very clear. Increasing the reduction parameter r drastically reduces the computing time required, as well as reducing the variability in computing time (Figure 5.3). It is important to note that since there are only three design points for each parameter, we can only investigate linear trends in this analysis. To investigate the higher order effects of each parameter on computation time would require more design points. This would be time consuming; however, it would also be desirable to increase accuracy and precision for future applications, and so should be considered for future work.

The 243,000 fitted models of the simulation study were split randomly into a training data set of 121,500 models and a testing data set of 121,500 models. A simple multiple linear regression model using ordinary least squares was fit to the training data, and was used to determine the effect size of each parameter on computation time. The testing data was used in predicting computation times. The linear regres-

Table 5.1: Covariate effect sizes, standard errors, and p-values for the linear regression model with log(computation time) as response variable.

Covariate	**Coefficient**	**Standard Error**	**p-value**
log(Sites)	0.848	0.002	< 0.001
log(Times)	0.585	0.002	< 0.001
$\hat{p}$	-0.836	0.005	< 0.001
$\log(r)$	-1.470	0.002	< 0.001

sion model is given in Equation 5.1, with $\varepsilon \sim N(0, \sigma^2)$. Note that λ is not present in the computation time model. This is because K was kept constant at 300 for the simulation studies. Future simulations should investigate the impact of allowing K to be directly proportional to λ.

$$\log(\text{computation time}) = \beta_0 \log(\text{sites}) + \beta_1 \log(\text{times}) + \beta_2 \log(r) + \beta_3 \hat{p} + \varepsilon \quad (5.1)$$

We use a log transform on computation time as computation time is strictly positive, and the range of a simple regression model is the whole real line. This in turn necessitated the same log transform of r, number of sites, and number of sampling occasions, as each was expected to be nearly linearly predictive of computation time.

Increasing number of sites and number of sampling occasions were both found to be positively correlated with computing time, while increasing probability of detection and increasing group size were found to be negatively correlated with computing time (Table 5.1). The variability of computation time explained by this model is 96.1% (adjusted R-squared of 0.961). To assess normality of the residuals, both a histogram (Figure 5.4a) and a qqplot (Figure 5.4b) were investigated. The residuals do not appear to follow a normal distribution, as there is some evidence of non-normality in the peak and skewness of the residual histogram, and the qqplot shows a sharp deviation in the lowest quantiles. However, the histogram indicates that this is not a large departure from the normality assumption required of ordinary least squares regression, and so is not a cause for concern regarding the validity of the model.

The computation time model can be useful in deciding the feasibility of running a particular grouped or full count model. For a given choice of r, a best guess estimate of p can be used to calculate the expected computing time for the model. To illustrate the feasibility of making these computation time estimates, we used the testing data

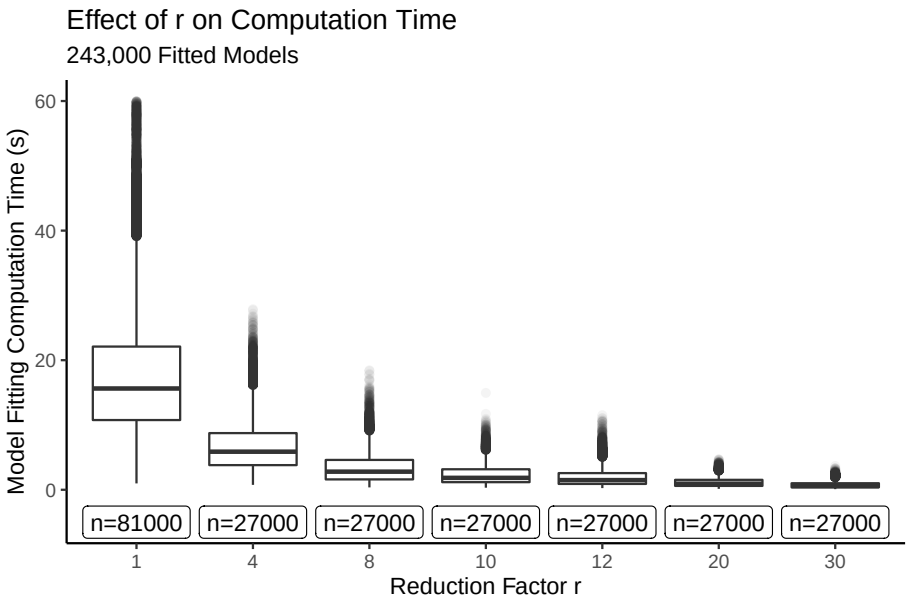
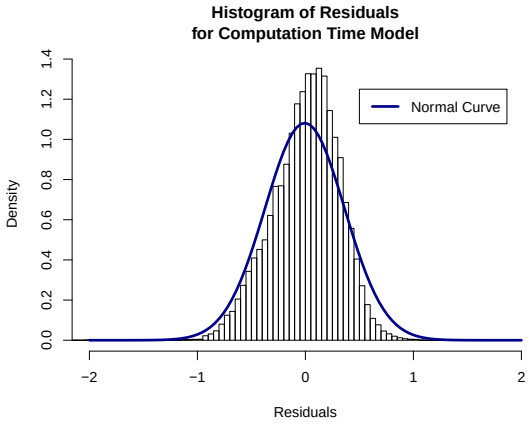
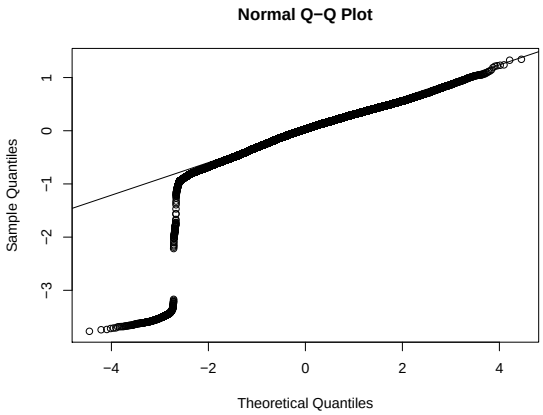


Figure 5.3: Computing time in seconds is plotted against increasing reduction parameter r. A total of 243,000 models are shown in these boxplots, with the sample size n shown beneath each boxplot.



(a) Histogram of residuals. Blue line indicates a normal distribution curve.



(b) Normal quantile plot.

Figure 5.4: Residual analysis for the computation time regression model

set to compute predicted and observed computation times (Figure 5.5).

5.4 Discussion

We performed a large simulation study with 243,000 simulations. Figure 5.1 illustrates the simulation results for 10 sites, and 10 sampling occasions. Similar figures are available in Appendix B for each of the nine combinations of U and M, with all parameter combinations. We note that $r = 1$ in the simulation study corresponds to the original full counts models, and so we can compare the full counts models with the grouped counts models directly from the simulation results.

Comparing the results shown in the Appendix B Figures 1-9, we see that increasing the number of sites and the number of sampling occasions both have the general effect of reducing variability in the estimates. Ideally the bias and variance of the estimates would tend to zero as the number of sampling sites and sampling occasions become large, which would show N-mixture models to be asymptotically consistent. This has not been shown in the N-mixture literature.

The interplay between parameter values is complex; however, some general conclusions can be drawn. Referring to Figure 5.2, it is clear that the relative error of $\hat{N}$ is heavily affected by the true probability of detection, with larger p reducing the bias substantially. The choice of r also impacts error, but in a more subtle way. If r is chosen to be too large in comparison with the observed counts n_{it}, then the re-

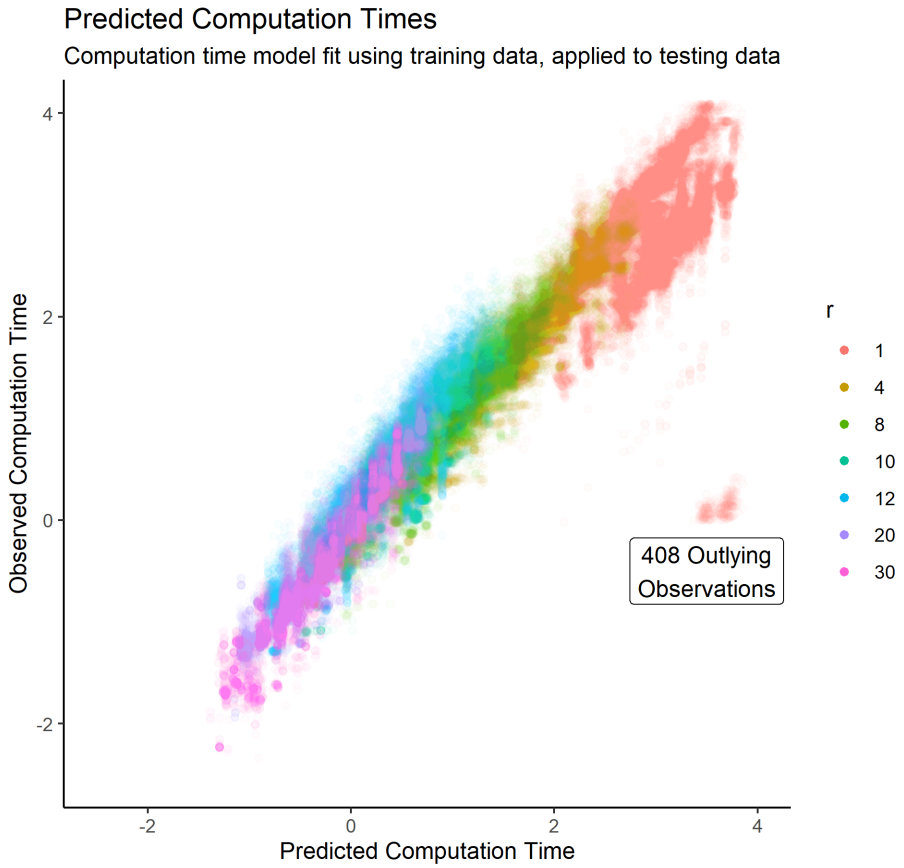


Figure 5.5: Observed computing time in seconds is plotted against predicted computing time. A total of 121,500 models are plotted here.

duction parameter introduces heavy bias. Thus choosing r should be done based on the observed counts, rather than on the assumed population size. This is important because it limits the effectiveness of the grouped count models when the probability of detection is small. The smaller the probability of detection, the smaller the observed counts, and so the larger the loss of information through grouping. When the probability of detection is $p = 0.75$, then little to no bias is introduced to the estimates for most considered values of r (with some small bias present for $r = 30$). In contrast we see that for $p = 0.25$ the relative error becomes large for $r > 4$, and the population abundance is consistently underestimated. When $p = 0.5$, the relative error is smaller than for $p = 0.25$, but still substantial for $r > 4$.

We were interested in the effect on computation times of using grouped count models for varying group sizes r. We fit a computation time model using 121,500 simulations. The computation time model, as shown in Equation 5.1 and Table 5.1, was used to compute predicted computation times for the testing data, which were then compared to the observed computation times (Figure 5.5). The predicted computing times closely match the observed computing times, with the exception of a small island of observations where the observed computing time is much smaller than the predicted (approximately 0.3% of the testing models lay in this island, labelled in Figure 5.5 as the "408 Outlying Observations"). It is unusual that this group contains only models for which $r = 1$, and that the group is well separated from other observations. Asides from $r = 1$, no other patterns were apparent among those 408 testing models belonging to the outlying group.

There are several important limitations to using the computation time model. The computing times will certainly be dependent on the speed of the computing system being used, and so the results may need to be calibrated to the specific systems employed by other researchers. All of our models were fit using Compute Canada resources, in particular the Cedar computing cluster. As well, a large factor in computing time is the choice of upper bound K. For these models, K was kept fixed at 300, which would not be a reasonable choice in the case of much larger populations. This limits the applicability of the computation time model to those where the true site abundance is less than 300. The computation time predictions were made using prior knowledge of $\hat{p}$, so that the accuracy of the estimates would be negatively impacted by a lack of prior knowledge. This should not dissuade the use of the computation time model, as in general a reasonable hypothesis should be formed as to the probability of detection for the study at hand. For example, in the Ancient Murrelet

study from Chapter 2, it was known that the probability of detection would be very high due to the sampling methodology. In a situation where no prior knowledge exists of the probability of detection, it would be reasonable to calculate a range of computing times, say for $p = 0.25, 0.5$, and 0.75. A worst case scenario can be calculated for the smallest reasonable expected value of p.

We conclude by summarizing the comparison between the full counts models ($r = 1$) and the grouped count models ($r > 1$). As seen in Figure 5.5, the computation time is much improved by use of grouped count models, so that even a small increase in r provides large decreases in computing time. The grouped counts models perform very well when r is chosen appropriately, with little to no bias in the estimates over the full counts models. However, when r is chosen too large, the grouped count models perform significantly worse in terms of bias (Figure 5.2). This is not surprising, as when r is large with respect to the observed counts, most of the information in the observed data is lost due to grouping. The effect of choosing r too large is exacerbated by small p, since smaller p causes smaller observed counts. We confirm the results of Royle (2004), that the full counts models perform worse for small p (as seen in Figure 5.2, $r = 1$), and so it may be prudent for future N-mixtures applications to focus more heavily on the moderate to large detection probability regime.

Chapter 6

Conclusions

N-mixture models provide a flexible parametric framework from which many different populations can be investigated. In this work we have extended N-mixture models in three key ways. In Chapter 2, we extended N-mixture models to allow the use of an auxiliary population as count data for estimating the population abundance of a separate, related population. This work has many potential applications in ecology. For example, many species have difficult to count adults and easy to count offspring, making population abundance estimates easier through the use of an auxiliary population. In Chapter 3, we extended N-mixture models to incorporate coarse counts. This has two immediate applications, the first is to artificially coarsen the count data to improve computation speeds, and the second is to allow the use of counts data which are collected as coarse counts in fitting N-mixture models. In either case, the switch to grouped counts changes the probability distributions involved, and precludes use of the original likelihood functions in model fitting. Thus, new grouped count models are introduced. In Chapter 4, we discuss extending N-mixture models to be applicable to very large populations, and the necessity of using arbitrary precision arithmetic, and arbitrary precision optimization for such applications.

The applications involving large populations are plentiful. An example in ecological research requiring large population considerations is the study of colonies of Ancient Murrelet seabirds larger than the one situated on East Limestone Island. These types of ecological studies enable population monitoring for difficult to census animal populations, and such monitoring provides invaluable evidence for policy making and conservation efforts. Many applications also exist in the field of health sciences and disease analytics, where human populations can be very large. Two such potential applications include estimating the size of the homeless population in a city

or geopolitical region, and estimating the number of undiagnosed individuals suffering from depression in a province or territory. Another potential application would be in estimating the degree of under-reporting in the spread of infectious diseases, such as the recent COVID-19 outbreak (Velavan and Meyer, 2020). These health science applications can help inform health care policy and decision making, and ultimately improve our ability to identify unmet service demands and to provide care or outreach for those in need.

In Chapter 4, we introduced two new R packages. First we introduced `optimizeAPA` (Parker, 2020a), for multi-parameter function arbitrary precision optimization. The package provides a necessary tool for researchers in all fields of science needing arbitrary precision optimization for multi-parameter functions. In particular, we applied `optimizeAPA` in the second R package, `redNMix` (Parker, 2020b). `redNMix` is a new package for model fitting making use of the new grouped count models. We also note that the auxiliary population models can be fit using `redNMix`, taking care to recognize that the probability of detection is confounded for auxiliary populations. This allows the auxiliary count models to incorporate the grouped count methods, either to use coarse counts for the auxiliary population, or to increase computation speeds for large observed counts.

In Chapter 5, we ran an extensive set of simulations, both to verify the validity of the new grouped count models, and to compare them against the original full count models ($r = 1$). We found that increasing the group size r greatly reduces the computing time needed to fit the N-mixture models. The simulations also showed that the choice of r is extremely important. If r is chosen to be too large in comparison with the observed counts, then it will induce heavy bias to the estimates. However if r is chosen moderately, then it has no great impact on bias. The bias induced by using grouped counts is most pronounced for small probability of detection, and decreases rapidly for increasing p. Both the full count models and the grouped count models perform worse when p is small, and so extra care should be taken if using N-mixture models for small detection probabilities.

Future work on grouped counts models should focus heavily on the large population regime, where arbitrary precision arithmetic is necessary in computing the model likelihoods. Extensive simulation studies for such large populations would prove to be infeasible even with the use of grouped counts, and so in the absence of increased computing power, smaller investigations may necessarily suffice.

We have successfully extended N-mixture models, incorporating both auxiliary

populations and grouped counts. These extensions increase the applicability of N-mixture models to many more populations of study. Using auxiliary count models we can now use separate populations with parametric relations to the population of interest to make abundance estimates. Using grouped count models we can fit N-mixture models to coarsely collected data, or we can artificially coarsen our data to improve computing times for large population studies. Further, we can combine the two extensions easily to make use of both auxiliary populations and grouped counts.

Appendix A

Supplementary Materials for Chapter 2

A.1 Web Appendix A

A.1.1 Auxiliary Population Viability (Model Definition)

The purpose of this simulation is to verify the validity of the model

$$\begin{aligned} N_{i1} &\sim Poisson(\lambda) \\ N_{it} &= S_{it} + G_{it} \\ S_{it} &\sim Binomial(N_{it}, \omega) \\ G_{it} &\sim Poisson(\gamma) \\ C_{it} &= \sum_{k=1}^{\beta N_{it}} Bernoulli(q) \\ c_{it} &= Binomial(C_{it}, p) \end{aligned}$$

as an N-mixture model with auxiliary probability of detection $\pi = pq$. Here $\beta = 2$ is the clutch size, N_{it} is the number of breeding pairs per site i and sampling occasion t, C_{it} are the auxiliary population of chicks, and c_{it} are the observed chick counts. p is the probability of detection for a chick, and q is the probability of a potential chick surviving to be observable. S_{it} and G_{it} describe the population dynamics of the adult breeding pair population in terms of survival probability ω and recruitment parameter γ. These can be equivalently thought of as the survival and recruitment of individual breeding females (rather than of breeding pairs).

A.1.2 Auxiliary Population Viability (Algorithm)

First we chose a suitable set of population parameters, chosen to be commensurate with the population of study. We set the number of sites $U = 6$, number of sampling occasions $M = 17$, initial mean abundance $\lambda = 150$, mean recruitment $\gamma = 10$, survival probability $\omega = 0.85$, and probability of detection $p = 0.8$. We also chose to use a productivity of $\alpha = 1.5$, as studies have shown that Ancient Murrelet productivity is around 1.5 (Gaston 1990, Gaston 2007). It is important to note that since q and p are confounded, this choice of α should not impact the simulation results.

1. (a) Generate breeding pairs from Poisson(λ), for all sites i and $t = 1$
   (b) Generate breeding pairs for $t > 1$ using Binomial(ω)+Poisson(γ)
2. Use breeding pairs to generate auxiliary population $C_{it} = \sum_{k=1}^{N_{it}}(Bin(\beta, q = \alpha/\beta))$. In this way, each adult breeding pair produces 0, 1, or 2 chicks, with mean chicks per pair of α.
3. Generate observed chick counts from the potential chicks using $c_{it} = Bin(C_{it}, p)$.
4. Fit open population N-mixture model to the observed chick counts
5. Switch from auxiliary estimates to breeding pair estimates by correcting by the factor β.
6. Repeat steps (1) to (5) 1000 times.

See Web Figure 1 for results.

A.2 Web Appendix B

Area and Productivity Expansion:

$\hat{C}_{N,t}$	estimated chick abundance for North Cove at time t
$\hat{C}_{C,t}$	estimated chick abundance for Cabin Cove at time t
$\hat{N}_{T,t}$	Total estimated breeding pair abundance at time t
A_N	North Cove sampling area
A_C	Cabin Cove sampling area
$A_{T,t}$	Total colony area at time t
β	Clutch size

The Cabin Cove and North Cove estimates were converted to densities using their respective sampling areas ($\hat{C}_{N,t}/A_N$ and $\hat{C}_{C,t}/A_C$), and their relative area proportions were calculated ($A_N/(A_N + A_C)$ and $A_C/(A_N + A_C)$). The total annual colony abundance estimates, $\hat{N}_{T,t}$, were then calculated as in Web Equation A.1.

$$\begin{aligned}\hat{N}_{T,t} &= \frac{A_{T,t}}{A_N}\frac{\hat{C}_{N,t}}{\beta}\frac{A_N}{A_N + A_C} + \frac{A_{T,t}}{A_C}\frac{\hat{C}_{C,t}}{\beta}\frac{A_C}{A_N + A_C} \\ &= \frac{A_{T,t}}{\beta}\frac{\hat{C}_{N,t}}{A_N + A_C} + \frac{A_{T,t}}{\beta}\frac{\hat{C}_{C,t}}{A_N + A_C} \\ &= \frac{\hat{C}_{N,t} + \hat{C}_{C,t}}{\beta}\frac{A_{T,t}}{A_N + A_C}\end{aligned} \tag{A.1}$$

A.3 Web Appendix C

Parametric Bootstrap Algorithm:

1. Generate a bootstrap sample using the estimated population parameters of model $\{\lambda_c, \gamma_t, \omega_t, \pi_c\}$
   (a) Generate an initial sample abundance at sampling occasion $t = 1$ for each site, drawn from a Poisson distribution with mean λ_i
   (b) Generate sample abundances for sampling occasions $t > 1$ using binomial survival probability ω_{it} and Poisson recruitment with parameter γ_{it}
2. Generate a bootstrap observation using binomial thinning (auxiliary probability of detection π_{it}) on the bootstrap sample

3. Fit an open population N-mixture model to the bootstrap observation to obtain the bootstrap parameter estimates

4. Repeat above steps (1) to (3) 1000 times

5. Calculate the standard deviation of the 1000 sets of bootstrap parameter estimates to obtain standard error estimates for each of the parameters

See Web Figure 2 for results.

A.4 Web Table 1

Table A.1: Ancient Murrelet chick counts for 1990 to 2006. These data were collected by the Laskeek Bay Conservation Society for the North Cove and Cabin Cove areas of East Limestone Island (Rock and Pattison, 2006).

Cove	**Funnel**	**1990**	**1991**	**1992**	**1993**	**1994**	**1995**	**1996**	**1997**	**1998**
North Cove	1	32	22	40	45	51	36	44	27	35
	2	134	61	93	84	92	65	105	86	69
	3	173	114	150	107	121	97	129	114	102
	4	142	95	115	116	107	93	121	114	92
Cabin Cove	5	148	93	90	103	94	58	65	65	53
	6	183	134	147	147	149	125	122	113	142
Total		812	519	635	602	614	474	586	519	493

Cove	**Funnel**	**1999**	**2000**	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**
North Cove	1	20	43	37	45	47	38	38	54
	2	84	87	68	77	76	67	58	59
	3	140	140	147	145	134	103	112	94
	4	123	123	115	116	95	104	100	90
Cabin Cove	5	63	67	63	59	52	37	43	41
	6	117	127	128	124	117	97	110	107
Total		546	587	558	566	521	446	461	445

A.5 Web Table 2

Table A.2: Total colony breeding pair estimates based on a clutch size of two and an N-mixture model applied to the auxiliary population of chicks for the years 1995 and 2006.

Year	**Abundance Estimate**	**95% CI**
1995	901	(832,973)
2006	761	(696,831)

A.6 Web Table 3

Table A.3: Total colony areas estimated by the Canadian Wildlife Service for 1995 and 2006 (survey methods available in Rodway et al., 1988). North Cove and Cabin Cove areas were estimated from polygons on a topographic map in the software `ArcGIS`.

Source	**Location**	**Year**	**Area (m^2)**
CWS	Total Colony ($A_{T,t}$)	1995	137600.0
		2006	125500.0
`ArcGIS` Estimates	North Cove (A_N)	1995/2006	34189.9
	Cabin Cove (A_C)	1995/2006	17906.9

A.7 Web Figure 1

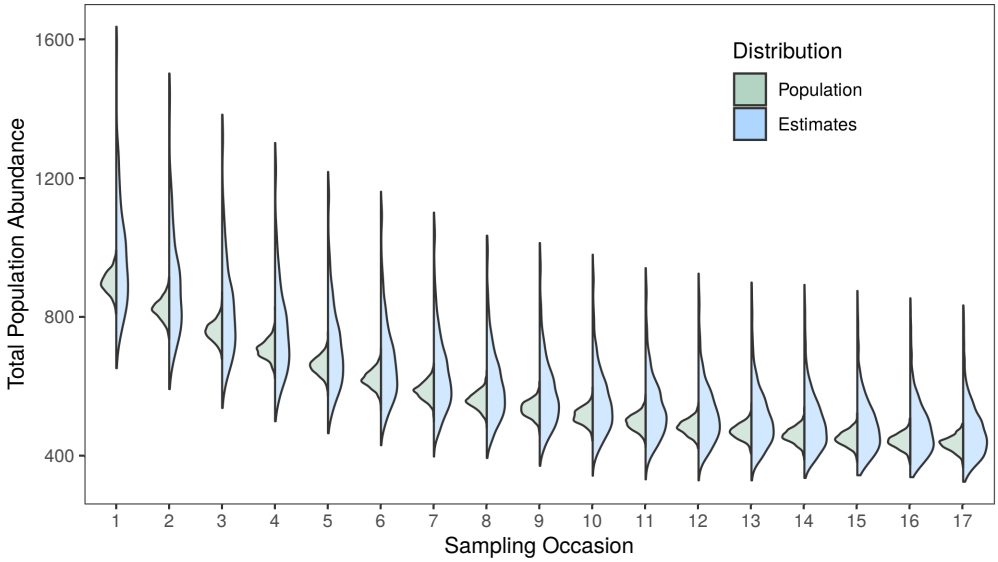


Figure A.1: Simulation study verifying validity of using an auxiliary chick population to estimate an adult population using their clutch size of $\beta = 2$ as a population link. Left hand distributions show the actual breeding pair population size distributions, right hand distributions show the distributions of the β-corrected N-mixture estimates of breeding pairs.

A.8 Web Figure 2

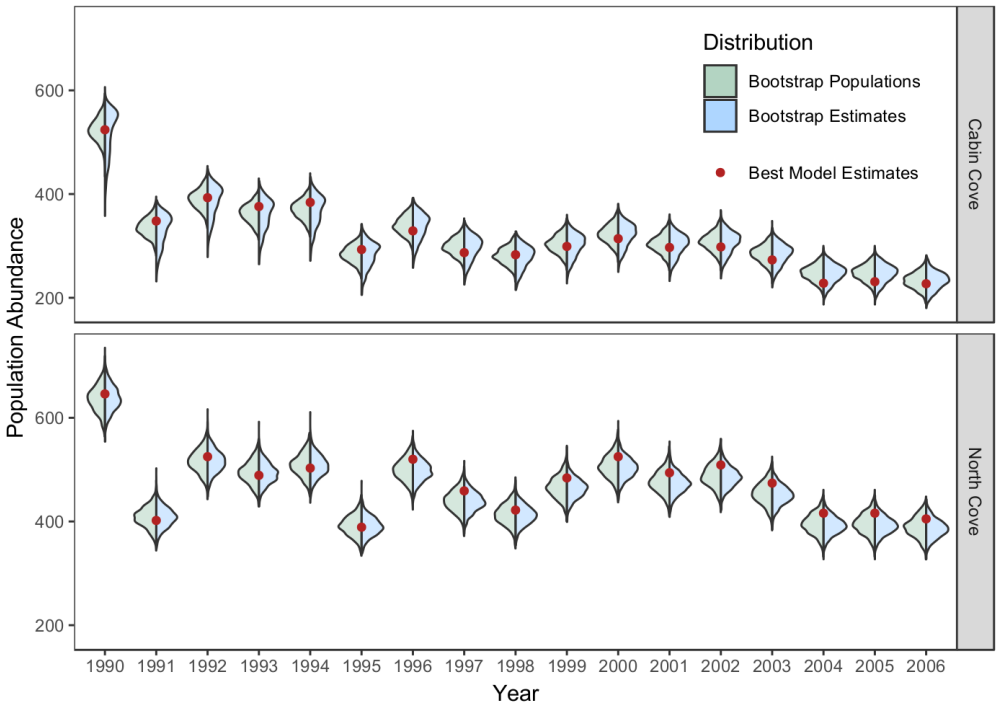


Figure A.2: Bootstrap population distributions by year and separated by Cabin Cove (top) and North Cove (bottom). Left-hand distributions are the actual bootstrap population size (C_{ct}, where c denotes cove and t denotes sampling occasion) generated from the parameter estimates of model $\{\lambda_c, \gamma_t, \omega_t, \pi_c\}$. Right-hand distributions are the bootstrap population size estimates ($\hat{C}_{ct}$) estimated from the thinned bootstrap populations (c_{ct}). Dots show the Ancient Murrelet chick population size estimates given by the model $\{\lambda_c, \gamma_t, \omega_t, \pi_c\}$.

Appendix B

Supplementary Materials for Chapter 5

B.1 Simulation Study Figures and Tables

B.1.1 Figure 1

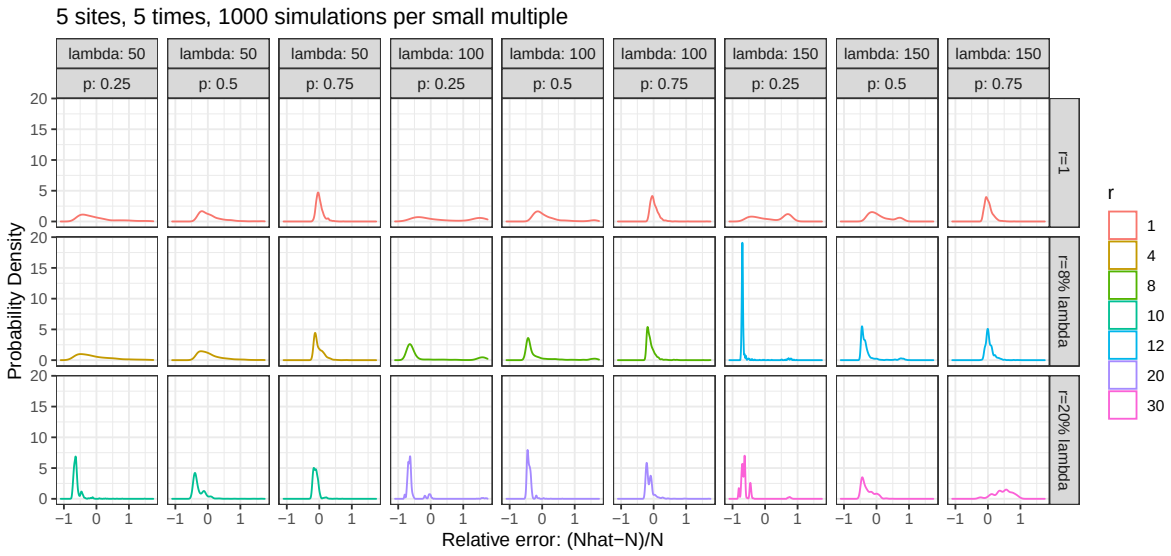


Figure B.1: Results of fitting grouped count N-mixtures models to randomly generated population/observation pairs. Each small multiple is a distribution plot of the relative error $\frac{\hat{N}-N}{N}$. Each small multiple represents 1000 simulations. For this set of simulations, $U = 5$, $M = 5$.

B.1.2 Figure 2

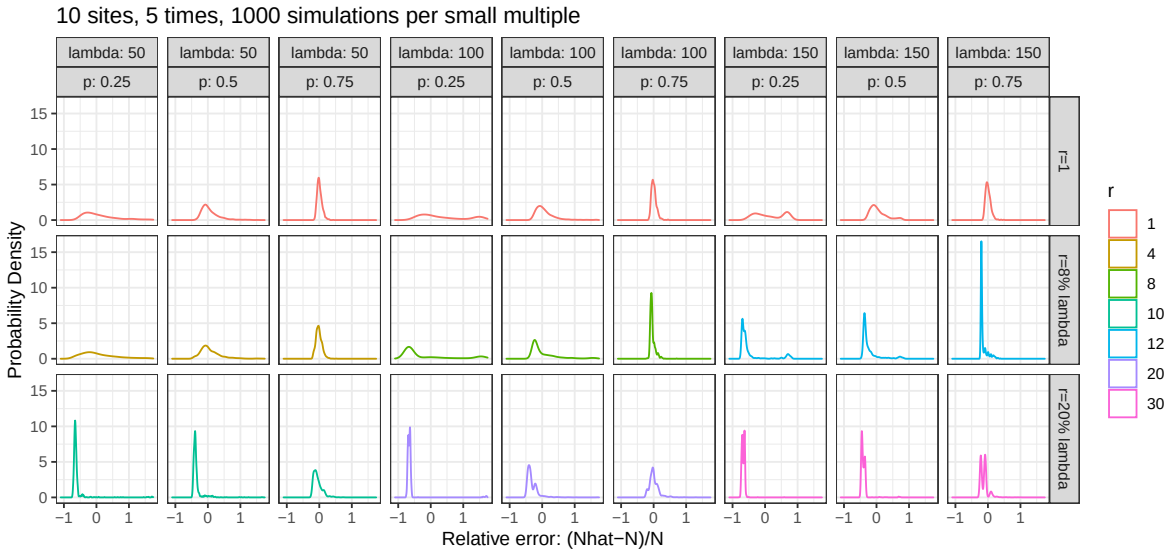


Figure B.2: Results of fitting grouped count N-mixtures models to randomly generated population/observation pairs. Each small multiple is a distribution plot of the relative error $\frac{\hat{N}-N}{N}$. Each small multiple represents 1000 simulations. For this set of simulations, $U = 10$, $M = 5$.

B.1.3 Figure 3

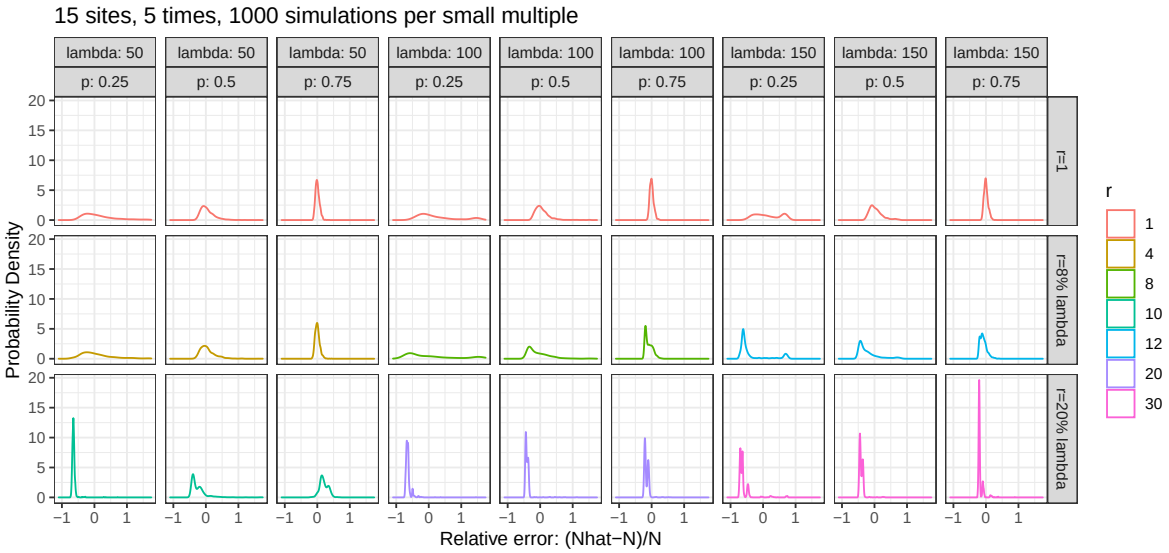


Figure B.3: Results of fitting grouped count N-mixtures models to randomly generated population/observation pairs. Each small multiple is a distribution plot of the relative error $\frac{\hat{N}-N}{N}$. Each small multiple represents 1000 simulations. For this set of simulations, $U = 15$, $M = 5$.

B.1.4 Figure 4

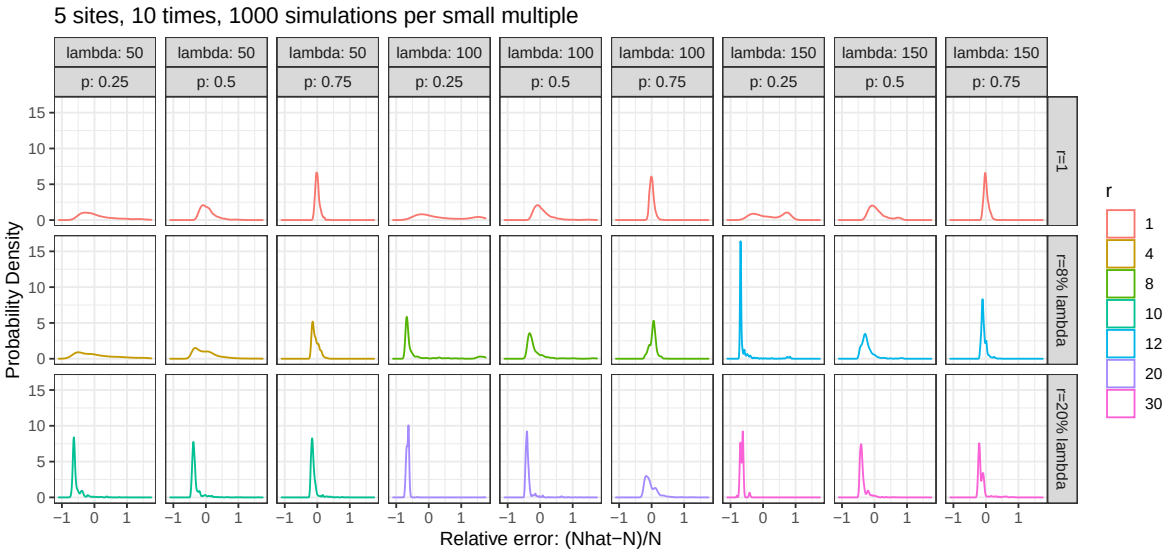


Figure B.4: Results of fitting grouped count N-mixtures models to randomly generated population/observation pairs. Each small multiple is a distribution plot of the relative error $\frac{\hat{N}-N}{N}$. Each small multiple represents 1000 simulations. For this set of simulations, $U = 5$, $M = 10$.

B.1.5 Figure 5

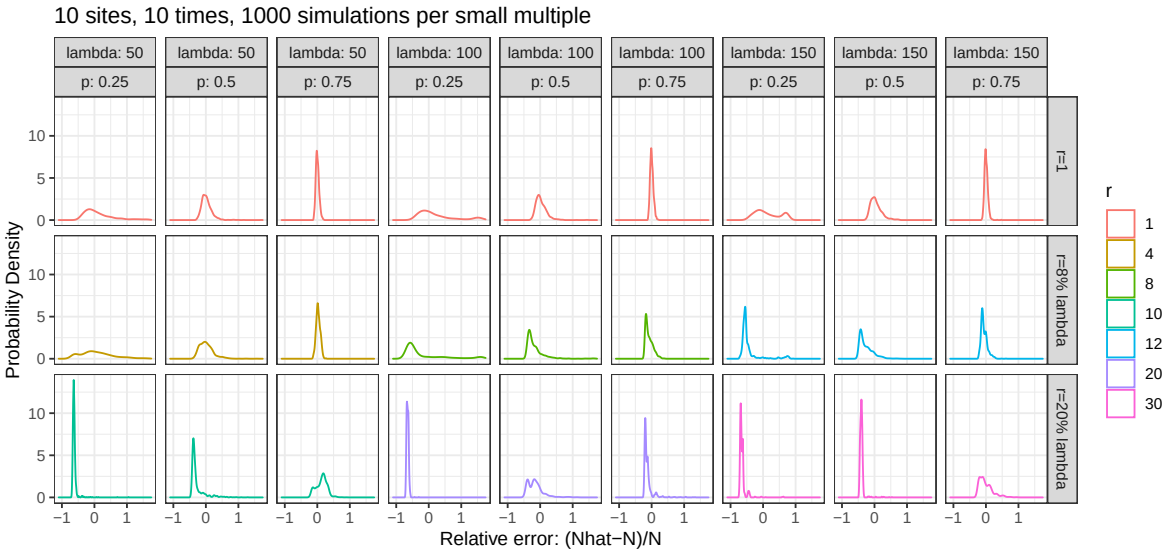


Figure B.5: Results of fitting grouped count N-mixtures models to randomly generated population/observation pairs. Each small multiple is a distribution plot of the relative error $\frac{\hat{N}-N}{N}$. Each small multiple represents 1000 simulations. For this set of simulations, $U = 10$, $M = 10$.

B.1.6 Figure 6

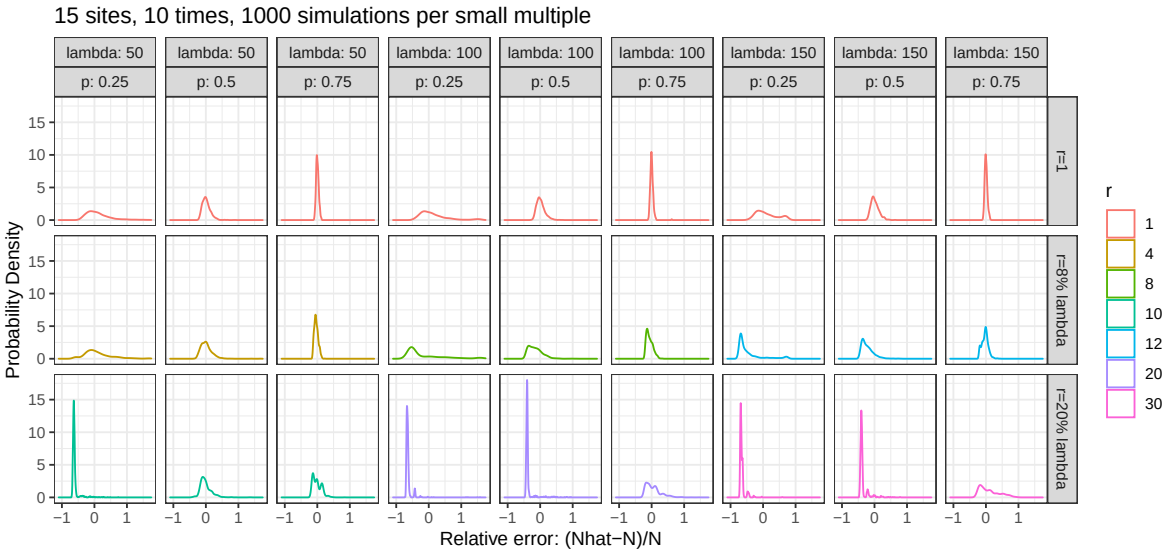


Figure B.6: Results of fitting grouped count N-mixtures models to randomly generated population/observation pairs. Each small multiple is a distribution plot of the relative error $\frac{\hat{N}-N}{N}$. Each small multiple represents 1000 simulations. For this set of simulations, $U = 15$, $M = 10$.

B.1.7 Figure 7

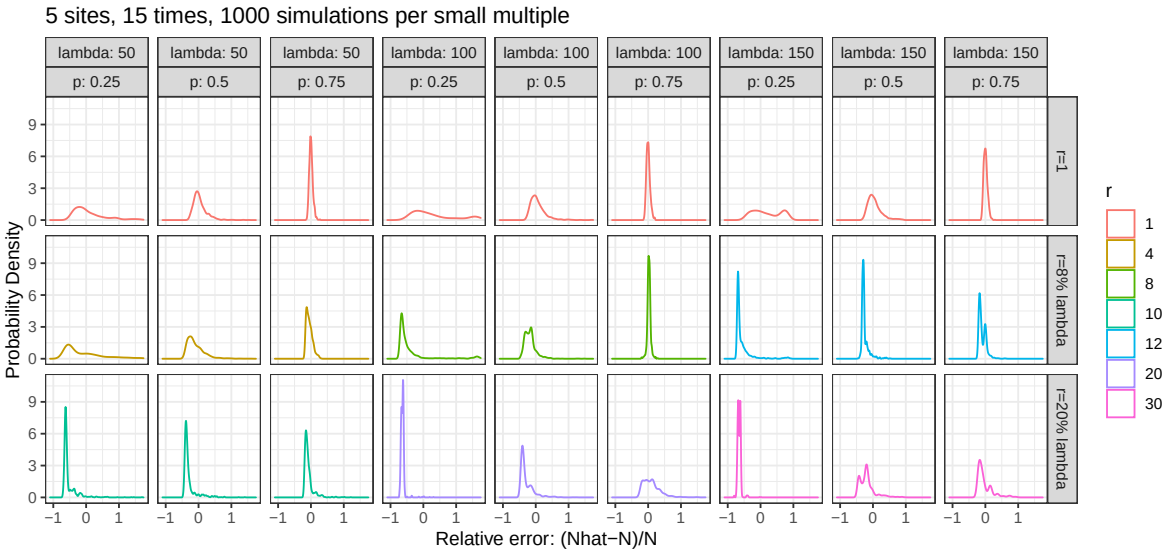


Figure B.7: Results of fitting grouped count N-mixtures models to randomly generated population/observation pairs. Each small multiple is a distribution plot of the relative error $\frac{\hat{N}-N}{N}$. Each small multiple represents 1000 simulations. For this set of simulations, $U = 5$, $M = 15$.

B.1.8 Figure 8

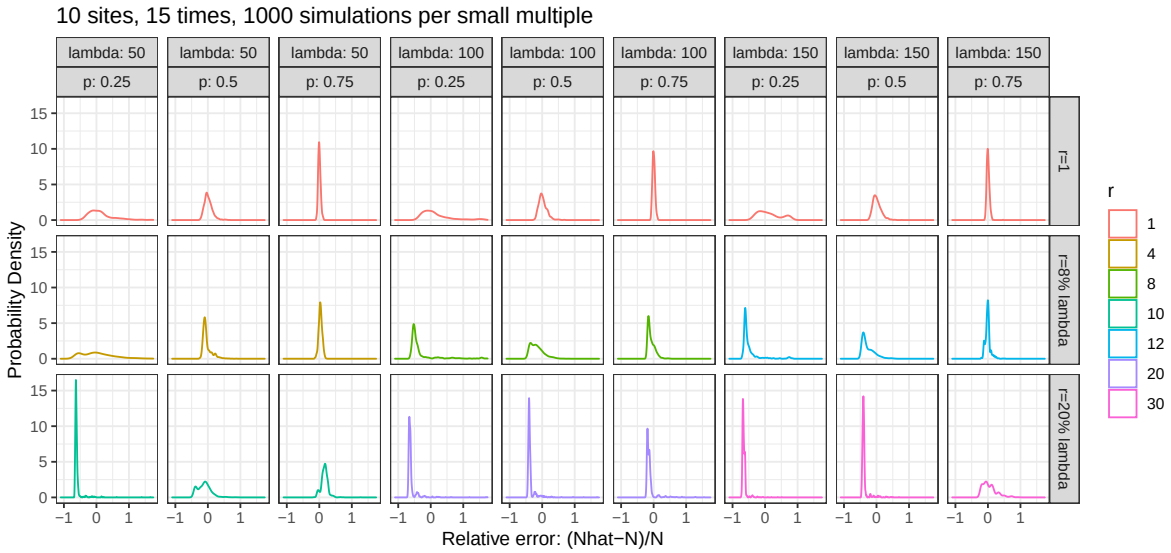


Figure B.8: Results of fitting grouped count N-mixtures models to randomly generated population/observation pairs. Each small multiple is a distribution plot of the relative error $\frac{\hat{N}-N}{N}$. Each small multiple represents 1000 simulations. For this set of simulations, $U = 10$, $M = 15$.

B.1.9 Figure 9

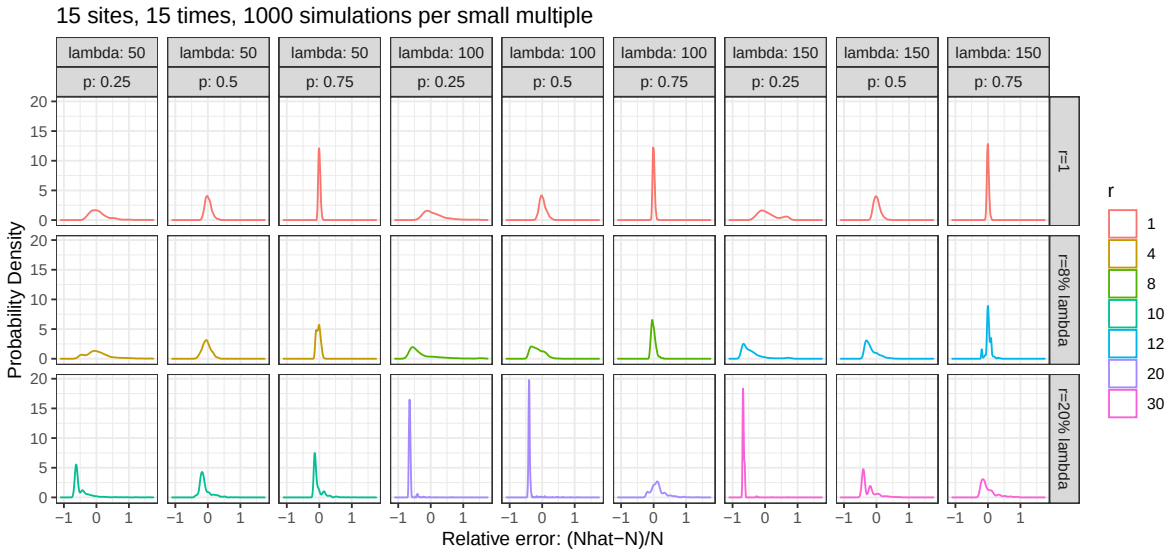


Figure B.9: Results of fitting grouped count N-mixtures models to randomly generated population/observation pairs. Each small multiple is a distribution plot of the relative error $\frac{\hat{N}-N}{N}$. Each small multiple represents 1000 simulations. For this set of simulations, $U = 15$, $M = 15$.

B.1.10 Table 1

Table B.1: Summary of sampling distribution and relative error for grouped count N-mixture models simulation study. 243,000 models are summarized here. U is the number of sampling sites, M is the number of sampling occasions, λ is the site abundance parameter, p is the probability of detection, r is the group size, $\bar{N}$ is the expected abundance per simulation, calculated as $U \times \lambda$, q1 is the 1st quartile, Mean is the mean of $\hat{N}$, Median is the median of $\hat{N}$, q3 is the third quartile, and MRE is the median relative error which is calculated as median($\frac{\hat{N}-N}{N}$).

U	M	λ	p	r	$\bar{N}$	q1	Mean	Median	q3	MRE
5	5	50	0.25	1	250	166.77	548.06	286.78	1200.70	0.1452
5	5	50	0.25	4	250	161.80	550.02	306.21	1217.13	0.2201
5	5	50	0.25	10	250	80.02	164.78	90.41	106.26	−0.6346
5	5	50	0.50	1	250	204.09	327.12	252.21	327.42	0.0042
5	5	50	0.50	4	250	197.00	338.41	249.75	339.61	−0.0078
5	5	50	0.50	10	250	146.87	203.04	160.23	210.76	−0.3534
5	5	50	0.75	1	250	230.20	252.37	245.94	269.96	−0.0132

U	M	λ	p	r	$\bar{N}$	q1	Mean	Median	q3	MRE
5	5	50	0.75	4	250	214.48	245.88	232.18	267.27	−0.0766
5	5	50	0.75	10	250	207.52	224.16	218.69	237.04	−0.1185
5	5	100	0.25	1	500	323.20	719.55	572.53	1244.21	0.1400
5	5	100	0.25	8	500	165.40	428.80	200.67	481.03	−0.5934
5	5	100	0.25	20	500	154.61	229.37	181.96	187.45	−0.6451
5	5	100	0.50	1	500	414.84	572.73	490.24	635.01	−0.0162
5	5	100	0.50	8	500	277.62	464.27	309.63	465.69	−0.3933
5	5	100	0.50	20	500	274.97	298.54	296.09	316.58	−0.4194
5	5	100	0.75	1	500	462.09	508.99	493.70	542.80	−0.0111
5	5	100	0.75	8	500	412.37	456.65	434.93	478.75	−0.1298
5	5	100	0.75	20	500	396.46	437.26	418.39	457.02	−0.1623
5	5	150	0.25	1	750	497.55	869.84	846.73	1279.46	0.1359
5	5	150	0.25	12	750	218.60	279.29	225.26	237.49	−0.6948
5	5	150	0.25	30	750	230.56	302.97	274.92	283.32	−0.6440
5	5	150	0.50	1	750	629.36	832.61	755.31	1007.93	0.0013
5	5	150	0.50	12	750	417.59	539.34	456.88	526.62	−0.3824
5	5	150	0.50	30	750	413.76	539.85	482.47	630.06	−0.3429
5	5	150	0.75	1	750	692.38	761.66	744.34	812.69	−0.0057
5	5	150	0.75	12	750	724.41	783.43	765.56	823.27	0.0132
5	5	150	0.75	30	750	982.46	1096.59	1131.66	1252.97	0.5114
5	10	50	0.25	1	250	191.70	484.46	286.82	635.22	0.1368
5	10	50	0.25	4	250	163.30	496.86	307.80	726.27	0.2223
5	10	50	0.25	10	250	87.79	138.70	96.43	117.65	−0.6169
5	10	50	0.50	1	250	221.88	280.25	254.73	296.76	0.0185
5	10	50	0.50	4	250	174.70	267.51	233.27	299.87	−0.0739
5	10	50	0.50	10	250	147.49	178.47	157.32	177.23	−0.3658
5	10	50	0.75	1	250	235.64	251.85	249.24	266.35	−0.0052
5	10	50	0.75	4	250	210.38	235.23	229.59	254.64	−0.0845
5	10	50	0.75	10	250	207.22	221.33	217.30	227.34	−0.1444
5	10	100	0.25	1	500	385.80	701.50	558.36	1077.93	0.1162
5	10	100	0.25	8	500	157.20	341.58	174.51	270.53	−0.6469
5	10	100	0.25	20	500	157.78	176.35	177.85	190.60	−0.6443
5	10	100	0.50	1	500	437.31	544.02	497.89	587.01	−0.0065

U	M	λ	p	r	$\bar{N}$	q1	Mean	Median	q3	MRE
5	10	100	0.50	8	500	332.39	424.21	368.81	438.50	−0.2696
5	10	100	0.50	20	500	277.27	322.96	296.09	318.99	−0.4010
5	10	100	0.75	1	500	474.34	504.74	501.29	530.12	0.0015
5	10	100	0.75	8	500	474.06	519.03	536.15	557.65	0.0493
5	10	100	0.75	20	500	418.10	487.24	457.04	537.64	−0.0855
5	10	150	0.25	1	750	563.96	892.61	856.97	1286.97	0.1541
5	10	150	0.25	12	750	225.00	312.54	236.88	273.24	−0.6840
5	10	150	0.25	30	750	233.67	262.71	267.21	286.00	−0.6495
5	10	150	0.50	1	750	664.77	810.18	760.78	900.06	0.0099
5	10	150	0.50	12	750	489.39	579.74	550.01	621.88	−0.2735
5	10	150	0.50	30	750	411.86	485.71	445.39	483.55	−0.3970
5	10	150	0.75	1	750	705.71	747.99	740.00	783.26	−0.0066
5	10	150	0.75	12	750	657.41	697.99	681.43	717.54	−0.0900
5	10	150	0.75	30	750	595.94	654.95	630.85	686.15	−0.1834
5	15	50	0.25	1	250	190.33	399.02	258.31	432.60	0.0269
5	15	50	0.25	4	250	117.95	370.07	218.33	435.62	−0.1237
5	15	50	0.25	10	250	94.97	127.70	96.43	117.65	−0.6110
5	15	50	0.50	1	250	223.60	261.32	249.13	284.32	−0.0027
5	15	50	0.50	4	250	178.38	232.96	209.33	257.00	−0.1634
5	15	50	0.50	10	250	147.49	179.47	157.62	177.25	−0.3596
5	15	50	0.75	1	250	236.78	250.79	249.63	263.12	−0.0010
5	15	50	0.75	4	250	213.91	237.61	233.89	257.45	−0.0662
5	15	50	0.75	10	250	207.25	231.17	217.38	247.08	−0.1199
5	15	100	0.25	1	500	399.62	667.52	552.78	889.58	0.0976
5	15	100	0.25	8	500	165.46	309.89	205.28	286.10	−0.5988
5	15	100	0.25	20	500	171.57	183.09	183.26	195.33	−0.6354
5	15	100	0.50	1	500	448.96	530.10	504.73	581.05	0.0095
5	15	100	0.50	8	500	356.84	415.98	407.31	451.08	−0.1909
5	15	100	0.50	20	500	296.09	360.33	318.98	399.49	−0.3733
5	15	100	0.75	1	500	477.19	501.75	499.14	522.80	−0.0036
5	15	100	0.75	8	500	501.25	512.78	516.55	525.77	0.0239
5	15	100	0.75	20	500	452.35	553.65	537.63	617.50	0.0691
5	15	150	0.25	1	750	616.28	890.02	834.86	1239.97	0.1139

U	M	λ	p	r	$\bar{N}$	q1	Mean	Median	q3	MRE
5	15	150	0.25	12	750	236.87	331.13	249.30	345.12	−0.6681
5	15	150	0.25	30	750	234.24	262.04	257.54	285.96	−0.6582
5	15	150	0.50	1	750	677.89	787.72	758.85	860.83	0.0045
5	15	150	0.50	12	750	513.41	563.20	537.36	573.42	−0.2865
5	15	150	0.50	30	750	483.55	609.86	595.93	659.10	−0.2083
5	15	150	0.75	1	750	713.78	749.65	747.08	780.67	0.0014
5	15	150	0.75	12	750	609.40	676.88	633.62	741.31	−0.1344
5	15	150	0.75	30	750	626.53	729.09	659.45	808.69	−0.1116
10	5	50	0.25	1	500	379.98	891.28	557.91	1075.39	0.1200
10	5	50	0.25	4	500	375.79	923.87	563.61	1215.75	0.1375
10	5	50	0.25	10	500	166.48	316.39	178.23	194.08	−0.6468
10	5	50	0.50	1	500	437.19	558.27	500.95	593.57	−0.0039
10	5	50	0.50	4	500	437.03	588.91	500.30	620.01	0.0021
10	5	50	0.50	10	500	289.65	333.12	305.17	325.01	−0.3895
10	5	50	0.75	1	500	473.11	504.78	498.83	529.93	−0.0036
10	5	50	0.75	4	500	455.57	493.73	486.99	526.68	−0.0207
10	5	50	0.75	10	500	425.44	473.31	455.18	500.86	−0.0861
10	5	100	0.25	1	1000	779.78	1372.52	1115.53	2044.31	0.1202
10	5	100	0.25	8	1000	307.32	848.76	347.17	1182.51	−0.6552
10	5	100	0.25	20	1000	309.91	375.01	339.65	363.92	−0.6630
10	5	100	0.50	1	1000	875.33	1080.58	997.01	1187.83	−0.0012
10	5	100	0.50	8	1000	758.17	1042.92	840.42	1168.40	−0.1567
10	5	100	0.50	20	1000	574.68	692.16	631.01	762.94	−0.3658
10	5	100	0.75	1	1000	950.53	1009.44	997.57	1055.75	−0.0039
10	5	100	0.75	8	1000	912.37	963.84	942.78	982.13	−0.0585
10	5	100	0.75	20	1000	913.78	1011.69	995.29	1084.16	−0.0163
10	5	150	0.25	1	1500	1094.24	1710.54	1628.88	2479.25	0.0853
10	5	150	0.25	12	1500	457.94	802.27	555.43	687.55	−0.6288
10	5	150	0.25	30	1500	444.72	496.15	486.16	551.03	−0.6775
10	5	150	0.50	1	1500	1302.89	1573.34	1479.05	1756.95	−0.0183
10	5	150	0.50	12	1500	936.45	1164.76	989.53	1199.77	−0.3428
10	5	150	0.50	30	1500	820.79	897.08	862.22	949.67	−0.4255
10	5	150	0.75	1	1500	1417.11	1509.46	1488.08	1583.30	−0.0063

U	M	λ	p	r	$\bar{N}$	q1	Mean	Median	q3	MRE
10	5	150	0.75	12	1500	1183.03	1279.16	1215.27	1282.94	−0.1943
10	5	150	0.75	30	1500	1192.04	1321.53	1335.13	1390.39	−0.1066
10	10	50	0.25	1	500	405.92	682.33	516.78	719.75	0.0368
10	10	50	0.25	4	500	381.64	668.43	525.64	743.45	0.0653
10	10	50	0.25	10	500	174.01	192.93	183.97	192.95	−0.6333
10	10	50	0.50	1	500	455.76	513.58	500.38	550.91	0.0000
10	10	50	0.50	4	500	428.46	508.77	494.42	564.49	−0.0148
10	10	50	0.50	10	500	304.22	367.08	324.43	370.92	−0.3580
10	10	50	0.75	1	500	477.51	501.03	499.43	521.06	−0.0012
10	10	50	0.75	4	500	487.52	513.66	514.52	538.96	0.0228
10	10	50	0.75	10	500	494.84	567.59	587.19	628.87	0.1545
10	10	100	0.25	1	1000	814.59	1226.37	1039.05	1430.22	0.0400
10	10	100	0.25	8	1000	404.39	836.87	504.38	1031.96	−0.5014
10	10	100	0.25	20	1000	327.77	349.92	350.91	372.51	−0.6522
10	10	100	0.50	1	1000	913.88	1029.28	1000.88	1114.13	−0.0061
10	10	100	0.50	8	1000	659.91	833.48	737.93	903.36	−0.2566
10	10	100	0.50	20	1000	638.00	846.81	814.94	941.31	−0.1873
10	10	100	0.75	1	1000	965.20	1003.75	999.18	1037.37	−0.0013
10	10	100	0.75	8	1000	827.05	925.11	892.25	996.92	−0.1100
10	10	100	0.75	20	1000	795.24	879.37	836.60	894.70	−0.1713
10	10	150	0.25	1	1500	1243.81	1700.49	1561.53	2187.80	0.0429
10	10	150	0.25	12	1500	619.39	847.39	689.22	810.42	−0.5427
10	10	150	0.25	30	1500	464.98	538.11	503.79	556.42	−0.6683
10	10	150	0.50	1	1500	1380.98	1569.88	1525.05	1703.15	0.0113
10	10	150	0.50	12	1500	858.66	1111.56	1005.73	1263.54	−0.3253
10	10	150	0.50	30	1500	850.30	900.19	890.79	932.46	−0.4057
10	10	150	0.75	1	1500	1443.31	1501.36	1494.22	1552.64	−0.0016
10	10	150	0.75	12	1500	1303.17	1414.39	1374.70	1489.84	−0.0816
10	10	150	0.75	30	1500	1234.00	1495.79	1399.71	1672.15	−0.0658
10	15	50	0.25	1	500	434.11	627.91	532.12	667.52	0.0648
10	15	50	0.25	4	500	336.39	602.92	501.55	702.40	0.0077
10	15	50	0.25	10	500	182.54	200.81	191.48	194.76	−0.6245
10	15	50	0.50	1	500	461.16	507.54	500.21	542.74	−0.0038

U	M	λ	p	r	$\bar{N}$	q1	Mean	Median	q3	MRE
10	15	50	0.50	4	500	440.81	485.40	459.00	504.05	−0.0774
10	15	50	0.50	10	500	365.19	445.06	445.89	505.00	−0.1096
10	15	50	0.75	1	500	483.00	502.36	501.00	520.80	−0.0016
10	15	50	0.75	4	500	498.30	517.07	517.15	537.72	0.0337
10	15	50	0.75	10	500	546.04	579.66	578.44	605.94	0.1632
10	15	100	0.25	1	1000	848.63	1156.90	1024.43	1299.55	0.0277
10	15	100	0.25	8	1000	467.21	694.22	518.43	624.55	−0.4836
10	15	100	0.25	20	1000	337.59	403.48	365.60	383.85	−0.6412
10	15	100	0.50	1	1000	930.77	1020.88	998.58	1098.37	−0.0005
10	15	100	0.50	8	1000	666.76	834.95	800.23	948.68	−0.1987
10	15	100	0.50	20	1000	592.19	643.65	613.70	637.98	−0.3957
10	15	100	0.75	1	1000	969.20	1003.19	999.63	1035.04	0.0009
10	15	100	0.75	8	1000	835.14	922.64	891.00	986.96	−0.1166
10	15	100	0.75	20	1000	815.16	880.73	836.26	879.93	−0.1568
10	15	150	0.25	1	1500	1248.74	1677.25	1576.70	2010.83	0.0427
10	15	150	0.25	12	1500	570.84	770.55	619.20	770.89	−0.5853
10	15	150	0.25	30	1500	469.88	529.82	506.55	552.22	−0.6671
10	15	150	0.50	1	1500	1380.24	1523.81	1495.01	1629.81	−0.0049
10	15	150	0.50	12	1500	882.69	1111.11	991.53	1272.32	−0.3362
10	15	150	0.50	30	1500	851.93	933.91	890.88	933.54	−0.4020
10	15	150	0.75	1	1500	1454.02	1505.04	1498.31	1552.77	−0.0011
10	15	150	0.75	12	1500	1410.72	1480.34	1471.37	1531.37	−0.0081
10	15	150	0.75	30	1500	1318.67	1569.83	1493.96	1730.52	−0.0088
15	5	50	0.25	1	750	594.34	1122.23	810.12	1241.94	0.0774
15	5	50	0.25	4	750	555.95	1100.69	762.61	1185.69	0.0267
15	5	50	0.25	10	750	248.45	272.20	264.23	279.87	−0.6473
15	5	50	0.50	1	750	672.12	788.31	751.47	868.04	0.0032
15	5	50	0.50	4	750	663.86	795.16	752.39	870.54	0.0077
15	5	50	0.50	10	750	451.44	580.93	511.23	631.53	−0.3222
15	5	50	0.75	1	750	715.37	754.21	750.26	786.81	−0.0030
15	5	50	0.75	4	750	707.03	750.94	745.26	788.78	−0.0045
15	5	50	0.75	10	750	838.67	899.76	888.04	967.73	0.1810
15	5	100	0.25	1	1500	1192.07	1866.50	1549.66	2268.70	0.0328

U	M	λ	p	r	$\bar{N}$	q1	Mean	Median	q3	MRE
15	5	100	0.25	8	1500	619.51	1618.41	1232.97	2355.34	−0.1777
15	5	100	0.25	20	1500	477.41	542.58	521.45	549.15	−0.6562
15	5	100	0.50	1	1500	1365.52	1591.12	1510.73	1728.39	0.0091
15	5	100	0.50	8	1500	988.37	1416.38	1243.54	1664.37	−0.1699
15	5	100	0.50	20	1500	832.43	907.88	873.99	933.93	−0.4195
15	5	100	0.75	1	1500	1432.38	1502.56	1492.06	1559.50	−0.0038
15	5	100	0.75	8	1500	1218.21	1373.69	1339.35	1497.85	−0.1098
15	5	100	0.75	20	1500	1176.44	1264.26	1230.77	1339.14	−0.1892
15	5	150	0.25	1	2250	1835.75	2578.19	2434.75	3587.13	0.0870
15	5	150	0.25	12	2250	851.83	1448.39	939.58	1334.03	−0.5826
15	5	150	0.25	30	2250	681.86	928.55	805.20	837.11	−0.6470
15	5	150	0.50	1	2250	2006.71	2343.04	2253.58	2551.44	0.0000
15	5	150	0.50	12	2250	1245.68	1769.94	1501.65	1993.08	−0.3297
15	5	150	0.50	30	2250	1231.83	1336.00	1282.19	1412.16	−0.4278
15	5	150	0.75	1	2250	2149.98	2253.01	2236.69	2337.39	−0.0046
15	5	150	0.75	12	2250	1892.40	2073.85	2034.09	2207.03	−0.0963
15	5	150	0.75	30	2250	1753.42	1845.17	1792.16	1833.63	−0.2042
15	10	50	0.25	1	750	640.63	897.35	778.83	980.93	0.0430
15	10	50	0.25	4	750	618.42	917.11	768.40	1006.75	0.0269
15	10	50	0.25	10	750	264.11	298.98	278.93	292.18	−0.6306
15	10	50	0.50	1	750	694.65	763.66	745.55	817.63	−0.0018
15	10	50	0.50	4	750	665.56	760.06	739.71	825.02	−0.0087
15	10	50	0.50	10	750	662.96	750.75	720.80	804.73	−0.0396
15	10	50	0.75	1	750	724.04	751.67	748.69	778.80	−0.0011
15	10	50	0.75	4	750	688.42	726.22	717.47	757.59	−0.0413
15	10	50	0.75	10	750	662.83	745.59	732.14	812.94	−0.0237
15	10	100	0.25	1	1500	1262.45	1715.48	1560.21	1944.74	0.0359
15	10	100	0.25	8	1500	683.70	1312.48	823.59	1710.33	−0.4518
15	10	100	0.25	20	1500	487.93	554.56	516.90	549.07	−0.6570
15	10	100	0.50	1	1500	1386.97	1526.55	1496.96	1626.52	−0.0033
15	10	100	0.50	8	1500	1016.64	1276.48	1221.27	1461.50	−0.1830
15	10	100	0.50	20	1500	867.73	946.84	891.12	932.34	−0.4015
15	10	100	0.75	1	1500	1451.08	1505.70	1498.90	1550.14	−0.0022

U	M	λ	p	r	$\bar{N}$	q1	Mean	Median	q3	MRE
15	10	100	0.75	8	1500	1287.77	1408.95	1383.41	1513.79	−0.0815
15	10	100	0.75	20	1500	1296.99	1569.68	1483.04	1713.60	−0.0078
15	10	150	0.25	1	2250	1904.10	2482.79	2307.25	2961.83	0.0269
15	10	150	0.25	12	2250	699.34	1251.63	869.46	1344.97	−0.6140
15	10	150	0.25	30	2250	694.87	786.23	737.23	809.16	−0.6749
15	10	150	0.50	1	2250	2073.92	2276.96	2228.79	2443.96	−0.0061
15	10	150	0.50	12	2250	1438.52	1758.21	1655.75	1962.78	−0.2606
15	10	150	0.50	30	2250	1298.78	1451.68	1344.04	1417.51	−0.3985
15	10	150	0.75	1	2250	2176.30	2247.63	2241.90	2311.47	−0.0020
15	10	150	0.75	12	2250	2016.08	2181.06	2195.82	2323.26	−0.0215
15	10	150	0.75	30	2250	1860.99	2483.96	2313.32	2961.09	0.0171
15	15	50	0.25	1	750	660.07	851.83	765.80	918.24	0.0257
15	15	50	0.25	4	750	610.07	845.63	754.04	941.03	0.0055
15	15	50	0.25	10	750	278.93	412.76	298.96	448.34	−0.5992
15	15	50	0.50	1	750	705.85	761.37	751.73	806.84	0.0018
15	15	50	0.50	4	750	652.11	723.69	717.30	784.72	−0.0454
15	15	50	0.50	10	750	597.66	680.34	641.57	719.08	−0.1496
15	15	50	0.75	1	750	725.56	750.80	750.26	774.32	−0.0006
15	15	50	0.75	4	750	688.52	729.07	727.78	768.71	−0.0270
15	15	50	0.75	10	750	642.02	715.56	671.85	732.67	−0.1155
15	15	100	0.25	1	1500	1294.48	1687.17	1535.71	1919.78	0.0255
15	15	100	0.25	8	1500	650.69	1142.91	846.33	1368.25	−0.4356
15	15	100	0.25	20	1500	504.23	548.35	533.34	559.57	−0.6472
15	15	100	0.50	1	1500	1420.26	1532.10	1507.69	1630.77	0.0049
15	15	100	0.50	8	1500	1041.35	1283.25	1245.12	1489.49	−0.1721
15	15	100	0.50	20	1500	885.71	927.51	893.69	934.08	−0.4003
15	15	100	0.75	1	1500	1461.52	1503.19	1499.87	1541.37	−0.0012
15	15	100	0.75	8	1500	1406.92	1490.00	1473.56	1560.68	−0.0200
15	15	100	0.75	20	1500	1491.50	1696.84	1650.21	1796.73	0.0980
15	15	150	0.25	1	2250	1950.84	2460.88	2284.30	2846.84	0.0254
15	15	150	0.25	12	2250	736.07	1296.74	1087.57	1514.24	−0.5183
15	15	150	0.25	30	2250	715.14	767.13	745.30	795.59	−0.6696
15	15	150	0.50	1	2250	2115.78	2283.57	2255.74	2420.62	0.0018

U	M	λ	p	r	$\bar{N}$	q1	Mean	Median	q3	MRE
15	15	150	0.50	12	2250	1552.15	1878.80	1747.22	2099.34	−0.2230
15	15	150	0.50	30	2250	1331.36	1718.43	1415.09	1892.62	−0.3787
15	15	150	0.75	1	2250	2192.05	2250.27	2244.34	2306.20	−0.0009
15	15	150	0.75	12	2250	2201.47	2272.86	2273.07	2367.85	0.0100
15	15	150	0.75	30	2250	1858.85	2250.16	2031.75	2476.04	−0.0982

Bibliography

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In *Proceeding of the Second International Symposium on Information Theory*, pages 267–281, Budapest: Akadémiai Kiadó.

Bailey, D. and Borwein, J. (2013). High-precision arithmetic: Progress and challenges. *http://www.davidhbailey.com/dhbpapers/hp-arith.pdf*.

Bailey, D., Hida, Y., Li, S., and Thompson, B. (2002). ARPREC: An Arbitrary Precision Computation Package.

Bain, L. J. and Engelhardt, M. (1992). *Introduction to Probability and Mathematical Statistics*. Brooks/Cole, Cengage Learning, 2nd edition.

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

Belant, J. L., Bled, F., Wilton, C. M., Fyumagwa, R., Mwampeta, S. B., and Beyer, D. E. (2016). Estimating lion abundance using N-mixture models for social species. *Scientific Reports* **6,** 35920.

Bryant, R. E. and O'Hallaron, D. R. (2011). *Computer Systems: A Programmer's Perspective*. Prentice Hall, Boston, 2nd edition.

Coonen, J. (1981). Underflow and the denormalized numbers. *Computer* **14,** 75–87.

Cowen, L. E., Besbeas, P., Morgan, B. J. T., and Schwarz, C. J. (2017). Hidden Markov models for extended batch data. *Biometrics* **73,** 1321–1331.

Dail, D. and Madsen, L. (2011). Models for estimating abundance from repeated counts of an open metapopulation. *Biometrics* **67,** 577–587.

Davidon, W. (1959). Variable metric method for minimization. Technical Report ANL-5990, 4252678.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society. Series B (Methodological)* **39,** 1–38.

DiRenzo, G. V., Che-Castaldo, C., Saunders, S. P., Grant, E. H. C., and Zipkin, E. F. (2019). Disease-structured N-mixture models: A practical guide to model disease dynamics using count data. *Ecology and Evolution* **9,** 899–909.

Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap.* Chapman and Hall/CRC, Boca Raton.

Esri (2014). ArcGIS ArcMap Release 10. Redlands, CA: Environmental Systems Research Institute. *https://www.esri.com/en-us/home*.

Fernández-Fontelo, A., Cabaña, A., Puig, P., and Moriña, D. (2016). Under-reported data analysis with INAR-hidden Markov chains. *Statistics in Medicine* **35,** 4875–4890.

Fiske, I. and Chandler, R. (2011). unmarked: An R package for fitting hierarchical models of wildlife occurrence and abundance. *Journal of Statistical Software* **43,** 1–23.

Fousse, L., Hanrot, G., Lefèvre, V., Pélissier, P., and Zimmermann, P. (2007). MPFR: A multiple-precision binary floating-point library with correct rounding. *ACM Transactions on Mathematical Software* **33,** 13–es.

Gaston, A. J. (1990). Population parameters of the Ancient Murrelet. *The Condor* **92,** 998–1011.

Gaston, A. J. (2007). Changes in numbers and breeding biology of Ancient Murrelets at East Limestone Island, 1990-2006. In Gaston, A. J., editor, *Laskeek Bay Research 15*, pages 39–54. Laskeek Bay Conservation Society, Queen Charlotte City, B.C.

Gaston, A. J. and Descamps, S. (2011). Population change in a marine bird colony is driven by changes in recruitment. *Avian Conservation and Ecology* **6,** 5.

Gaston, A. J., Jones, I. L., Noble, D. G., and Smith, S. A. (1988). Orientation of Ancient Murrelet, *Synthliboramphus antiquus*, chicks during their passage from the burrow to the sea. *Animal Behaviour* **36,** 300 – 303.

Gaston, A. J. and Smith, J. L. (2001). Changes in oceanographic conditions off northern British Columbia (1983-1999) and the reproduction of a marine bird, the Ancient Murrelet (*Synthliboramphus antiquus*). *Canadian Journal of Zoology* **79,** 1735–1742.

Gaston, A. J. (ed). (2013). Laskeek Bay Research 17. In *Laskeek Bay Conservation Society Scientific Report*. Laskeek Bay Conservation Society, Queen Charlotte City, B.C.

Goldberg, D. (1991). What every computer scientist should know about floating-point arithmetic. *ACM Computing Surveys (CSUR)* **23,** 5–48.

Grundy, P. M. (1952). The fitting of grouped truncated and grouped censored normal distributions. *Biometrika* **Vol. 39,** pp. 252–259.

Heitjan, D. F. and Rubin, D. B. (1991). Ignorability and coarse data. *The Annals of Statistics* **19,** 2244–2253.

Kéry, M. (2017). Identifiability in N-mixture models: A large-scale screening test with bird data. *Ecology* **99,** 281–288.

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

Lemon, M. J. F. (2007). East Limestone Island Ancient Murrelet Colony Survey, June 2006. In Gaston, A. J., editor, *Laskeek Bay Research 15*, pages 67–86. Laskeek Bay Conservation Society, Queen Charlotte City, B.C.

Lyons, J. E., Royle, J. A., Thomas, S. M., Elliott-Smith, E., Evenson, J. R., Kelly, E. G., Milner, R. L., Nysewander, D. R., and Andres, B. A. (2012). Large-scale monitoring of shorebird populations using count data and N-mixture models: Black Oystercatcher (*Haematopus bachmani*) surveys by land and sea. *The Auk* **129,** 645–652.

Maechler, M. and Heiberger (2020). Rmpfr: R MPFR - Multiple Precision Floating-Point Reliable. *http://rmpfr.r-forge.r-project.org/*.

Major, H. and Chubaty, A. (2012). Estimating colony and breeding population size for nocturnal burrow-nesting seabirds. *Marine Ecology Progress Series* **454,** 83–90.

Matthews, G. J., Harel, O., and Aseltine, R. H. (2016). Privacy protection and aggregate health data: A review of tabular cell suppression methods (not) employed in public health data systems. *Health Services and Outcomes Research Methodology* **16,** 258–270.

Nichols, J. D., Hines, J. E., Lebreton, J. D., and Pradel, R. (2000). Estimation of contributions to population growth: A reverse-time capture-recapture approach. *Ecology* **81,** 3362–3376.

Ó Cadhla, O., Keena, T., Strong, D., Duck, C., and Hiby, L. (2013). Monitoring of the breeding population of grey seals in Ireland, 2009 – 2012. In *Irish Wildlife Manuals No. 74*. National Parks and Wildlife Service, Department of the Arts, Heritage and the Gaeltacht, Dublin, Ireland.

Oliver, N., Garg, A., and Horvitz, E. (2004). Layered representations for learning and inferring office activity from multiple sensory channels. *Computer Vision and Image Understanding* **96,** 163–180.

Parker, M. (2020a). optimizeAPA: An R package for optimizing arbitrary precision multi parameter functions. *https://github.com/mrparker909/optimizeAPA*.

Parker, M. (2020b). redNMix: An R package for implementing both grouped count N-mixture models, and large population N-mixture models. *https://github.com/mrparker909/redNMix*.

Parker, M., Pattison, V., and Cowen, L. (2018). Mrparker909/ANMU-Code: ANMU-Code Release v1.0. Zenodo, doi: 10.5281/zenodo.1472416.

Rock, J. and Pattison, J. (2006). East Limestone Island field station: report on the 2006 field season. Queen Charlotte, B.C. *www.laskeekbay.org/field-season-summaries/*.

Rodway, M. S., Lemon, M. J. F., and Kaiser, G. W. (1988). British Columbia seabird colony inventory: Report #1 - East Coast Moresby Island. Pacific and Yukon Region, B.C.

Royle, J. A. (2004). N-mixture models for estimating population size from spatially replicated counts. *Biometrics* **60,** 108–115.

Velavan, T. P. and Meyer, C. G. (2020). The COVID-19 epidemic. *Tropical Medicine & International Health.* doi:10.1111/tmi.13383.

Ward, R. J., Griffiths, R. A., Wilkinson, J. W., and Cornish, N. (2017). Optimising monitoring efforts for secretive snakes: a comparison of occupancy and N-mixture models for assessment of population status. *Scientific Reports* **7,** 18074.

Wit, E., Heuvel, E., and Romeijn, J. (2012). 'All models are wrong...': an introduction to model uncertainty. *Statistica Neerlandica* **66,** 217–236.