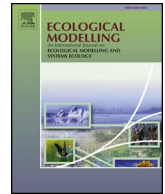




ELSEVIER

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Fitting N-mixture models to count data with unmodeled heterogeneity: Bias, diagnostics, and alternative approaches

Adam Duarte^{a,*}, Michael J. Adams^b, James T. Peterson^c

^a Oregon Cooperative Fish and Wildlife Research Unit, Department of Fisheries and Wildlife, Oregon State University, 104 Nash Hall, Corvallis, OR, United States of America

^b U.S. Geological Survey, Forest and Rangeland Ecosystem Science Center, Corvallis, OR, United States of America

^c U.S. Geological Survey, Oregon Cooperative Fish and Wildlife Research Unit, Department of Fisheries and Wildlife, Oregon State University, Corvallis, OR, United States of America

ARTICLE INFO

Keywords:

Abundance
Count data
Detection
Goodness of fit
Monitoring
N-mixture model

ABSTRACT

Monitoring animal populations is central to wildlife and fisheries management, and the use of N-mixture models toward these efforts has markedly increased in recent years. Nevertheless, relatively little work has evaluated estimator performance when basic assumptions are violated. Moreover, diagnostics to identify when bias in parameter estimates from N-mixture models is likely is largely unexplored. We simulated count data sets using 837 combinations of detection probability, number of sample units, number of survey occasions, and type and extent of heterogeneity in abundance or detectability. We fit Poisson N-mixture models to these data, quantified the bias associated with each combination, and evaluated if the parametric bootstrap goodness-of-fit (GOF) test can be used to indicate bias in parameter estimates. We also explored if assumption violations can be diagnosed prior to fitting N-mixture models. In doing so, we propose a new model diagnostic, which we term the quasi-coefficient of variation (QCV). N-mixture models performed well when assumptions were met and detection probabilities were moderate (i.e., ≥ 0.3), and the performance of the estimator improved with increasing survey occasions and sample units. However, the magnitude of bias in estimated mean abundance with even slight amounts of unmodeled heterogeneity was substantial. The parametric bootstrap GOF test did not perform well as a diagnostic for bias in parameter estimates when detectability and sample sizes were low. The results indicate the QCV is useful to diagnose potential bias and that potential bias associated with unidirectional trends in abundance or detectability can be diagnosed using Poisson regression. This study represents the most thorough assessment to date of assumption violations and diagnostics when fitting N-mixture models using the most commonly implemented error distribution. Unbiased estimates of population state variables are needed to properly inform management decision making. Therefore, we also discuss alternative approaches to yield unbiased estimates of population state variables using similar data types, and we stress that there is no substitute for an effective sample design that is grounded upon well-defined management objectives.

1. Introduction

Abundance is a population state variable of considerable interest to wildlife and fisheries managers, as it is frequently used to inform state-dependent decision making (Williams et al., 2002). Accordingly, natural resource monitoring programs often concentrate on tracking the distribution and abundance of populations within and among landscapes. Yet a census is rarely, if ever, executed afield because at least a portion of a population inevitably remain undetected, particularly when monitoring rare or elusive species (Thompson, 2004). Hence, monitoring data need to be corrected for incomplete detection to properly inform management decision making.

N-mixture models represent a class of hierarchical models that are increasingly used to correct monitoring count data by explicitly modeling population-level processes along with the underlying detection process (Royle, 2004a). This approach models the number of individuals counted (C) at sample unit i during survey occasion t as binomial random variables, $C_{it} \sim \text{Binomial}(N_i, p)$, where N_i is abundance at sample unit i and p is detection probability (i.e., the probability of detecting an individual in the population given it is available for detection). By treating abundance as an independent random variable generated from a statistical distribution (i.e., Poisson, negative binomial, zero-inflated Poisson, etc.), the model is able to estimate detection probability. Perhaps the primary appeal of this approach over

* Corresponding author.

E-mail address: adam.duarte@oregonstate.edu (A. Duarte).

<https://doi.org/10.1016/j.ecolmodel.2018.02.007>

Received 30 October 2017; Received in revised form 4 January 2018; Accepted 13 February 2018
0304-3800/ © 2018 Elsevier B.V. All rights reserved.

alternative abundance estimation techniques, such as capture-recapture models, is that N-mixture models only require relatively inexpensive, convenient to collect, and less invasive spatially and temporally replicated count data of unmarked individuals. Also, explanatory variables that may influence abundance and detection probability can be included using generalized linear models (GLMs). Extensions to the original parameterization continue to be developed, increasing the overall flexibility and usefulness of this approach (reviewed in Dénes et al., 2015).

Like all estimators, N-mixture models require a set of assumptions that should be considered prior to fitting the model. Specifically, these models assume population closure (i.e., no immigration, emigration, births, or deaths) across replicate survey occasions within a primary occasion (i.e., within a sampling season), individuals are not double counted during a single survey occasion, individuals are detected independently from each other, individual heterogeneity in detection probability is negligible, and the data match the distributional assumptions of the model. New developments have enabled managers to relax and at times even overcome some of these basic assumptions. Still, it is evident that N-mixture models yield imprecise estimates with small sample sizes and unstable estimates when detection probabilities are low (Royle, 2004a; McIntyre et al., 2012; Couturier et al., 2013; Yamaura, 2013; Veech et al., 2016; Barker et al., 2017) or when heterogeneity in abundance or detection probability is not modeled (Kéry and Royle, 2016; Veech et al., 2016 – but see Kéry, 2017).

Diagnostics to identify when parameter estimates from N-mixture models are likely to be biased are largely unexplored. In practice, potential sample size issues can be avoided by monitoring a greater number of sample units or increasing the number of primary occasions (e.g., Yamaura et al., 2011). Furthermore, sample size requirements can be evaluated via data simulation. Nonetheless, diagnosing potential problems associated with low detectability and unmodeled heterogeneity in model parameters is much more difficult. A common recommendation is to avoid fitting N-mixture models when detectability is too low or when there is evidence unmodeled heterogeneity is too high. While this recommendation is statistically valid, it is of little real value to managers because detection probability is often unknown *a priori*. That is, managers typically only have the count data in hand, and variation in those data is complicated by heterogeneity in both abundance and detection probability. Parametric bootstrapping procedures can be implemented as a goodness-of-fit (GOF) test when fitting models using a maximum likelihood framework, presumably allowing managers to evaluate how well the model fits the data. Yet, parametric bootstrapping procedures can suffer from substantial sample error, and GOF tests for hierarchical models, in general, are considered underdeveloped and in need of further evaluation (Kéry and Royle, 2016). Nevertheless, although not the original intention and never formally evaluated, the parametric bootstrap GOF test for N-mixture models is often portrayed as assessing the validity in the parameter estimates derived from the models.

Given count data of unmarked individuals are increasingly used to inform management decision making, it is imperative that potential biases and diagnostics are thoroughly evaluated. Here, we begin by briefly describing types of heterogeneity that may arise in count data using common sample designs. Using these scenarios as a foundation, we simulate replicated count data across a gradient of detection probability to assess the bias in estimated abundances at various types and extents of unmodeled heterogeneity in the count data. We also evaluate the performance of the parametric bootstrap GOF test when used to diagnose bias in parameter estimates. Furthermore, we explore the extent to which assumption violations can be diagnosed prior to fitting N-mixture models. In doing so, we propose a new diagnostic to identify bias in parameter estimates, which we term the quasi-coefficient of variation (hereafter, QCV). Finally, we discuss alternative procedures to estimate population state variables that can be implemented using similar monitoring data when assumption violations are suspected but

cannot be resolved.

2. Methods

2.1. General description of the scenarios

Heterogeneity in count data can arise from process variation (i.e., true variation in abundance) and variation in sample error (i.e., variation in the detectability of individuals). Here, we focus on five scenarios that represent common sample designs used to monitor animal populations. Scenario 1 represents the ideal sample design, where all statistical assumptions are met. This scenario serves as a baseline and allows for the examination of bias associated with variable detectability. Scenario 2 represents a sample design in which random heterogeneity in abundance is present. This represents a case where individuals enter (i.e., are recruited and immigrate) and leave (i.e., die and emigrate) the sample units across survey occasions within a primary occasion. Scenario 3 also represents a case where there is heterogeneity in abundance, but the heterogeneity is unidirectional (i.e., decreasing), which can occur if survey protocols are invasive (i.e., trap or net captures, flipping cover objects [rocks, logs or cover boards], electrofishing, etc.) and individuals leave the sample units or perish after each survey occasion within a primary occasion. Although we focus on a decrease in abundance here, the same patterns in bias can be expected if abundances increase across survey occasions. Scenario 4 represents a sample design in which random unmodeled heterogeneity in detection probability is present. Like scenario 2, this can be associated with an assortment of factors, but our rationale was to capture a case where explanatory variables that may influence detection probability were omitted, such as habitat, behavior, etc. Last, scenario 5 represents a sample design where detection probability decreases across survey occasions within a primary occasion. This pattern could be associated with sample designs that cause individuals to become more vigilant and, by extension, less detectable after each survey occasion, changing environmental conditions that influence detectability (i.e., rain obscuring tracks or feces during sign surveys), animal activity patterns that vary in synchrony with the mating season, etc. Again, although we focus on a downward trend the same patterns can be expected if detectability increases across survey occasions.

2.2. Simulating population and monitoring data

We simulated population and monitoring count data using the scenarios outlined above as a foundation. For each scenario, data were simulated across a range in detection probability (ranging from 0.1 to 0.9 by 0.2), number of sample units (25, 50, or 100), number of survey occasions (3, 6, or 12), and type and extent of heterogeneity in true abundance or detection probability (see Sections 2.2.1–2.2.5 for specific details). This resulted in the evaluation of 837 combinations, which were each simulated for 1000 iterations within program R (R Core Team, 2016). The simulation code is provided in Appendix A.

2.2.1. Scenario 1

For scenario 1, true abundance for each sample unit was generated assuming a Poisson process, where mean true abundance among sample units (λ) was randomized by sampling from a uniform distribution ranging from 1 to 50. Observation error was then simulated to generate count data using a binomial process. We used this same approach when simulating data for scenarios 2–5, except we induced heterogeneity in abundance or detection probability.

2.2.2. Scenario 2

For scenario 2, we induced random heterogeneity in true abundance. Specifically, we treated the initial true abundance for each sample unit as the mean true abundance, and generated a true abundance for each survey occasion by sampling from a gamma distribution

Download English Version:

<https://daneshyari.com/en/article/8846078>

Download Persian Version:

<https://daneshyari.com/article/8846078>

[Daneshyari.com](https://daneshyari.com)