Outline of Project

Jack Thomas
Department of Mathematics and Statistics
University of Victoria
Victoria, BC, Canada

October 11, 2022

1 Introduction

One ecological problem we would like to solve is how to estimate the population size of wild animals that live in a measurable area. Ecological statistics is the sub-field of statistics that deals with such problems. The most common methods for monitoring species are the following:

1. Censuses in which it is possible to count all the individuals in the population in a single sample.
2. Presence-absence (or detect/non-detect) studies which can only say whether at least one individual was seen in a specific area in a single sample.
3. Mark-recapture-like studies in which animals can be identified, either from man-made tags or natural marks, so that they can be tracked over time.

A census would give the most information about a population if we were interested in how large the population was at a certain time. However, they are often impossible or prohibitively expensive to conduct on a population. Presence-absence studies are inexpensive and can generate large datasets but can only tell us whether at least one individual was in a site at a certain time. Mark-recapture studies provide a balance in information and resources required. An active field of research in ecological statistics concerns the methodology for combining information from multiple data sources to get better estimates for population parameters.

Blanc, Marboutin, Gatti, Zimmermann, and Gimenez [1] proposed a statistical model that combined mark-recapture and presence-absence data to estimate population sizes. The model attempted to estimate the population size of Eurasian lynx in the Jura Mountains, eastern France. Camera traps provided mark-recapture data on the lynx. Observers collected presence-absence data by an extensive sign survey for evidence of presence in the field [1]. There are concerns about Blanc et al. [1]. The method ignores locations of animals when working with spatial mark-recapture data. There was no assessment of the performance of the method either analytically or via a simulation study. The authors analyzed a dataset but provided no ground truth.

We believe that the method in [1] underestimates population sizes. In a comparison with other methods, the model estimated the population size of grizzly bears in a region of Alberta [2]. The approach of Blanc et al. [1] produced a point estimate of forty bears, despite other methods estimating four hundred bears or more. The main problem with the method is that it fails to account for animal movements when using spatial mark-recapture data. Our proposed work is to conduct a simulation study to explore the model and assess the performance and resultant population estimates. We hypothesize that the model will produce erroneous estimates in all but two extreme cases. First, when animals do not move between sites during the study, and second, when animals move throughout all the sites equally. This work is important because researchers may use this method and plan management strategies without realizing that the estimates are flawed. This could have grave consequences.

2 Literature Review

Summary, history, review of literature on the topic.

3 Model Specification

Explicitly outline model by [1].

4 Methods

Data was generated to feed into the supplementary code by [1].

4.1 Home-range centers

How the home-range centers were generated.

4.2 Capture-Recapture

How the capture-recapture data was generated.

4.3 Presence-Absence

How the presence-absence data was generated.

4.4 Simulation

Description of Canada HPC network, how parameters were specified and how the simulations were run. Link to supplementary code?

5 Results

Outline what the population estimates were. Graphs showing results.

6 Discussion

6.1 Case Study

Model ignores presence-absence data entirely. Model by [1] simply outputs the number of individuals captured. Credible proportion below 95% for all except high τ. Perfect correlation of the mean N to number of captured individuals n. Convergence of the mean N to the actual N actual with τ. Straight line graphs. Standard deviation bias decreased as τ increased. Model is invalid for all but one extreme case: when animals have a huge movement parameter and there is only one site (amount captured may as well be a census). When there is little movement, the model still outputs population estimate that matches the amount of individuals captured.

6.2 Analysis

Analysis of

$$\lambda = -\log(1 - \psi_i) \tag{1}$$

Higher abundance means higher occupancy, and vice versa. Data about occupancy pushes ψ_i down, data on abundance push λ up. Resulting estimates are based on the interplay between λ and ψ_i being on different scales and ignore occupancy data.

7 Future Work

Developing a model that combines presence-absence and mark-recapture data while taking into account animal movements.

References

[1] L. Blanc, E. Marboutin, S. Gatti, F. Zimmermann, and O. Gimenez. Improving abundance estimation by combining capture-recapture and occupancy data: Example with a large carnivore. *J. Appl. Ecol.*, 51(6):1733–1739, 2014. doi: 10.1111/1365-2664.12319.

[2] M. Jahid, H. N. Steeves, J. T. Fisher, S. J. Bonner, S. Muthukumarana, and L. L. Cowen. Shooting for abundance: Comparing integrated multi-sampling models for camera trap and hair trap data. *Environmetrics*, 2022. doi: 10.1002/env.2761.

A Data

Extra data showing that number of sampling occasions did not matter, other results.