# MSM4PCOD Task 3C

Eiren Jacobson

11 January 2023

## Summary

In this project, we are interested in detecting trends in metapopulations of marine mammals where disturbance and sampling may not be uniformly distributed. How should we collect data to maximize the ability to detect declines? We propose to investigate this question using a simulation tool. Briefly, the simulation includes population simulation via an individual-based population model, data generation via simulated line-transect, capture-recapture, and passive acoustic surveys, and trend analysis via an integrated population model. As a case study, we simulate a hypothetical metapopulation loosely based on Cuvier's beaked whales (Ziphius cavirostris) in Southern California, with the eventual goal of building a framework that could be adapted for any species/region/scenario.

**NOTE:** All code is available at https://github.com/eirenjacobson/MSM4PCoD.git

## Task 3C.1: Population Simulation

The population simulation takes a simple two-box structure, meant to emulate a metapopulation with two regions each containing a subpopulation. Each region has its own carrying capacity and internal dynamics, but animals may move between regions (explained below). The populations are assumed to have Pella-Tomlinson type density-dependent fecundity (Brandon et al. 2007). For simplicity, we assume that the two regions are the same size (A = 2,500 km2) and start with the same carrying capacity (K = 100).

The script initPop.R initializes a population, given parameters S0 (calf survival), S1 (juvenile survival), S2 (subadult and adult survival), AFR (age at first reproduction), AJU (age at which become juvenile subject to S1), ASA (age at which become subadult subject to S2), fmax (maximum fecundity), K (initial carrying capacity), and z (degree of compensation for the Pella-Tomlinson response). For the case study, these values are set as detailed in Table 1.

| Parameter | Value |
|---|---:|
| S0 | 0.80 |
| S1 | 0.85 |
| S2 | 0.95 |
| AFR | 10.00 |
| AJU | 3.00 |
| ASA | 5.00 |
| AMAX | 50.00 |
| fmax | 0.20 |
| K | 100.00 |

The script projPop.R takes the output of initPop.R and projects the populations forward in time, according to the above parameters plus the number of years to project (here nyears is set to 100). Carrying capacities are set for each year and region, so that disturbance can be simulated in one or both regions. Fecundity and

survival are simulated stochastically. Individuals animals are tracked over time (this is necessary for the mark-recapture simulation, explained below).

Density-dependent movement occurs. If the density of animals in one region is higher than the density of animals in another region, some animals will move to the region with lower density. The simulation assumes that the animals do not correctly perceive the habitat quality; i.e., the density each year Nt/K is evaluated based on K in year 1, if K changes, it is assumed that the animals are not aware of this. However, population dynamics (esp. changes in fecundity) do respond to the new K. This is called an ecological trap. For the purposes of the case study, this design is meant to emulate the hypothesized scenario in Southern California, where disturbance due to Naval training and testing is thought to overlap with high-quality habitat.

Only juvenile animals are allowed to move between regions. The simulation includes a connectivity parameter c that controls how much redistribution of animals occurs. The parameter takes values between 0 and 1. When c = 1, animals redistribute so that the densities in the two regions are equal. When $0 < c < 1$, "surplus" animals in one region will move to the other region with probability c. The script redistributePop.R evaluates the ratio of N/K in each of the two regions and moves individuals (tracked using unique IDs) between regions.

For the case study, we simulate two scenarios: one where K is constant across regions and time and others where K decreases in one region but not the other (e.g., by 50%; Fig. 1).
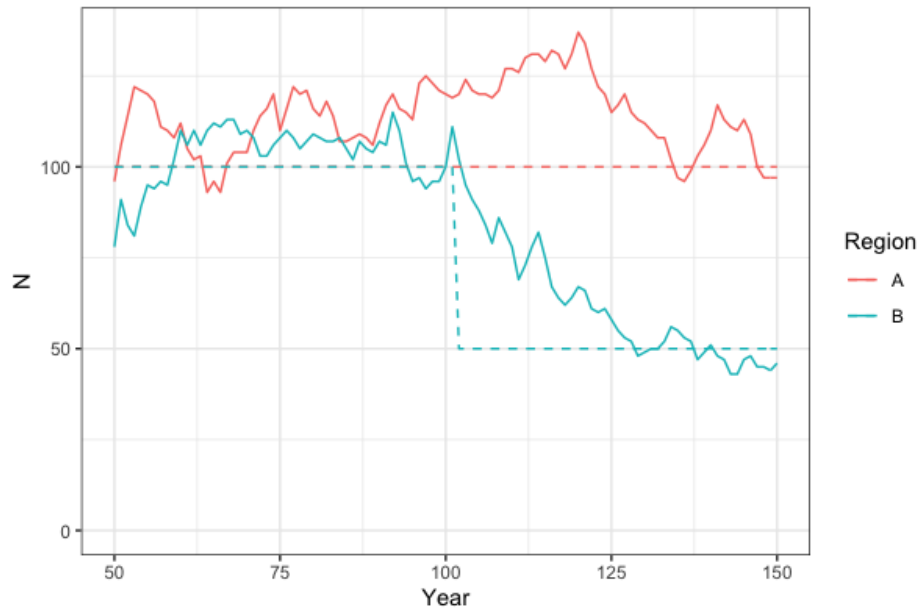


Figure 1: Simulated metapopulation showing number of animals (vertical axis) in each of two subpopulations (red indicates region A, blue indicates region B) over a 100 year period (horizontal axis). The carrying capacity in each region is shown by a dashed line.

## Task 3C.2: Survey Simulation

We simulate three different types of data collection and analysis: photographic capture-recapture surveys for estimates of adult survival, visual line-transect surveys for estimates of abundance, and fixed passive acoustic monitoring for estimates of abundance.

The script simCapRecap.R simulates photographic capture-recapture surveys. The surveys are simulated to occur only in the disturbed region (e.g., region B in Fig. 1). Parameters of the simulation include survey years and effort (which affects capture probability).

TODO: need to get relationship of time spent on the water to capture probability, either from Curtis paper or the last chapter of Dave Moretti's thesis.

The script simLineTransect.R simulates line-transect surveys with properties similar to line-transect surveys carried out by the SWFSC (Moore and Barlow 2013). Parameters of the simulation include survey years and effort (total length of line transects). It is assumed that the detection function is estimated from data pooled across multiple surveys (not just from the simulated data).

TODO: currently simLineTransect includes only binomial variance. Need to fix this to include overdispersion according to Distance book. Should also switch from using length of transect to using CV directly, then explain methods for conversion.

TODO: Add acoustic survey simulation (need CV estimates from Megan) add to above figure

## Task 3C.3: Integrated Population Model

We constructed an integrated population model to jointly analyze multiple simulated datasets. The integrated population model is implemented in Nimble and fit with MCMC. The file nimbleIPM.R contains the model code, while the file runIndSim.R contains code to generate data (as described above) and fit the Nimble model.
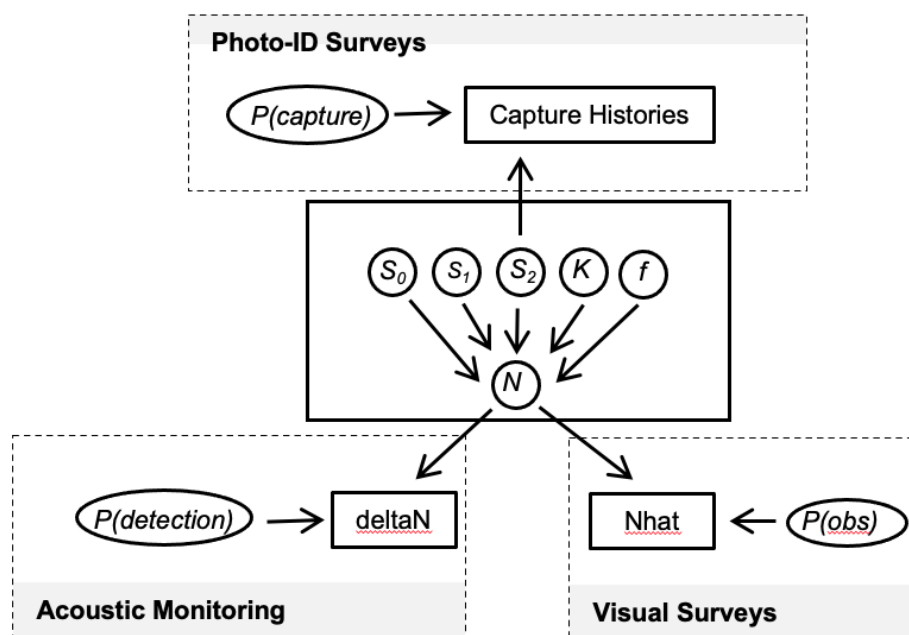


Figure 2: Directed acyclic graph of model structure. Parameters are shared between the population process model (central box) and each of the observational submodels (peripheral boxes).

TODO: re-make model DAG in R

TODO: second version of the model where metapopulation structure is known

TODO: In Nimble, can you compile a model before you show it any data? This would make the simulation process faster because you wouldn't need to recompile the model

TODO: quantification of power, histogram/density plot of estimated trends would be more informative (where power would just be the proportion of those trends that are on the right side of zero)

TODO: Look at new IPM book, look into more efficient JS

**Future work and/or for discussion**

- Expand simulation to include group size simulation/estimation. Bernoulli for detecting groups, then likelihood for group size that is poisson or negative binomial.

- Could include any number of survey modalities, each with a frequency/inter-survey-interval and CV associated with resulting estimates of abundance, and where it happens.

- Could use open population capture recapture to estimate N and survival.

- When both capture-recapture and fractional information (e.g., calf or juvenile ratio) are used, data will not be independent – think about whether/how to handle this in the likelihood.

- More complex simulation could introduce environmental variation, fluctuation in carrying capacity from year to year. Squid fishery or CalCOFI data on variability/fluctuations as indicator of environmental variability? E.g., percentage change in carrying capacity from year to year?

- Currently only individuals are detected. Could add group size simulation. Bernoulli for detecting groups, then likelihood for group size that is poisson or neg bin.

- TODO: re-read Charlotte's paper and see how they did inference on local v regional?