MAR 580: Advanced Population Modeling

Fall 2023

Homework Assignment 2

State-space modeling

Due: 10/26/23, 13:00 pm (Eastern Time)

Please provide Gavin with a brief report containing your solutions to the tasks below, and also include your R scripts (or markdown), .cpp file, and any additional files needed to run your assignment. Remember that in addition to performing the technical analyses, communicating those analyses through interpretation and discussion is an important component.

Review of state-space modeling

1. Use the LLM of your choice (e.g. Chat-GPT4) to produce a brief (1-2 paragraph) overview of state-space models, their estimation and application to population dynamics model, and the advantages and disadvantages of using state-space models for these applications. Based on your reading and knowledge from the course material, provide a critical review of the LLM-generated response, including correcting any necessary statements, and validate any literature that the model provided.

In your report, include what model you used to generate the response, the prompt you provided, as well as the generated answer and your review.

Stage-structured state-space model, Steller sea lions in Southeast Alaska

The BRS (births, recruitment, survival) model in the Newman et al. (2014) is similar to our previously encountered stage-structured model for Steller sea lions, keeping track of two state variables (pups and nonpups) over time.

$$\begin{bmatrix} N_{pups,t+1} \\ N_{np,t+1} \end{bmatrix} = \begin{bmatrix} 0 & f \\ \phi_p & \phi_{np} \end{bmatrix} \begin{bmatrix} N_{pups,t} \\ N_{np,t} \end{bmatrix}$$

where: `f` is the 'fecundity' relating the number of non-pups to the number of pups in the following year,
ϕ_p is pup survival rate,
ϕ_{np} is the non-pup survival rate.

As in the book, each annual update of this process model could be considered as the outcome of a series of bernoulli trials for the 3 population processes, which equate to the outcome of 3 binomial random variables (the number of births, the number of pups surviving, and the number of nonpups surviving).

Alternatively, we could consider annual process errors on the outcome of the state vector, where:

$$\begin{bmatrix} \ln(N_{pups,t+1}) \\ \ln(N_{np,t+1}) \end{bmatrix} = \ln(\begin{bmatrix} 0 & f \\ \phi_p & \phi_{np} \end{bmatrix} \begin{bmatrix} N_{pups,t} \\ N_{np,t} \end{bmatrix}) + \begin{bmatrix} \psi_{pups,t} \\ \psi_{np,t} \end{bmatrix}$$

with $\psi_{X,t} \sim N(0, \tau^2)$

Therefore modeling the annual process errors as iid normals.

We have count data between 1961-2019 from the Forrester Island rookery complex in Southeast Alaska, for both pups and non-pups. Assume that the pup counts are unbiased estimates of the number of pups, but there is an unknown fraction of non-pups that are not hauled out at a given time (assume this fraction is constant over time).
The observation model for this is then:

$$\ln(Y_{pups,t}) = \ln(N_{pups,t}) + \epsilon_{p,t}$$
$$\ln(Y_{np,t}) = \ln(q * N_{np,t}) + \epsilon_{np,t}$$

with $\epsilon_{p,t} \sim N(0, \sigma_p^2)$ and $\epsilon_{np,t} \sim N(0, \sigma_{np}^2)$

Additionally, we have estimates from mark recapture experiments for pup survival (mean = 0.6, sd = 0.05), and the fraction of nonpups hauled out (mean = 0.3, sd =0.1).

2. Inspect the available data for pups and non-pups. Provide a list of 3-5 summaries of the data, including any conclusions about how the properties of the data might facilitate (or not) application of the state-space model described above, including estimation of the magnitude of process and observation error.
3. Develop a TMB model that can fit the state-space model described above to the available Forrester Island sea lion data.
4. Run the model in 2., estimate parameters and time series of pup and non-pup abundance. Diagnose, evaluate, and plot the results, and summarize your findings in terms of the population dynamics and any implications arising from estimation.
5. Conduct a stochastic projection of the population through to 2030. Summarize the population status.
6. Using the same .cpp program, fit (a) an observation error only model (no process error), and (b) a process error only model. Summarize the results and comment on the implications for the estimates of population parameters, the population dynamics, and the prediction of population size in 2030.
7. **BONUS** Explore the behavior of the model when some parameters are fixed at various initial values rather than estimated, and comment on the estimation performance of models that assume separate or the same values for one or both of the osbervation and process error variance terms.
8. **BONUS BONUS** The above models all consider the annual process errors as indpenedent for pups and non-pups. Explore using your model results whether this assumption makes sense, and perhaps fit a model that models covariance in these annual terms, and comment on implications of any results.