Review of Mark loss can strongly bias demographic rates in multi-state models: a case study with simulated and empirical datasets

May 2, 2022

Dear authors, I enjoyed reading your paper very much. Your work is original in that no one has ever assessed the bias in transition parameters due to tag-loss in capture-recapture models. I have a few comments that I hope you'll find useful. Regards.

Main comments

- You have done an impressive work in gathering survival estimates and the magnitude of tag loss for hundreds of species. Congratulations on that. Tables S1 and S2 are awesome!
- I feel like lines 71-83 are too technical for an introduction. I would move this paragraph to Material and methods, and explain the problem with words here.
- Line 265: I find the use of ROPE disturbing because from the beginning of the paper we expect an assessment of bias. Why not simply consider posterior means/medians and compute bias in the frequentist way? I'm not asking to drop the ROPE metric, but to add a more traditional measure of bias (hopefully the simulations were saved and it will not take long).
- You're missing several important references. First, multistate models should not be called multistate CJS models. R. Cormack, G. Jolly and G. Seber developed models with a single alive state. Multistate models were developed by Neil Arnason and Carl Schwarz in a series of papers, and multistate models are sometimes referred to as the Arnason-Schwarz model. As of terminology, I guess multistate models or Arnason-Schwarz models do the job. The references that need to be cited are:
  - Arnason, A. N. (1972) Parameter estimates from mark-recapture experiments on two populations subject to migration and death, Researches on Population Ecology, 13, pp. 97-113.
  - Arnason, A. N. (1973) The estimation of population size, migration rates and survival in a stratified population, Researches on Population Ecology, 15, pp. 1-8.
  - Schwarz, C. J., Schweigert, J. F. & Arnason, A. N. (1993) Estimating migration rates using tag-recovery data, Biometrics, 49, pp. 177-193.

  Second, about tag-loss, there are two papers that need to be included in your paper:
  - Arnason, A. N., and K. H. Mills. 1987. Detection of handling mortality and its effects on Jolly–Seber estimates for mark-recapture experiments. Can. J. Fish. Aquat. Sci. 44: 64-73.
  - Juillet, C., Choquet, R., Gauthier, G. et al. A Capture–Recapture Model with Double-Marking, Live and Dead Encounters, and Heterogeneity of Reporting Due to Auxiliary Mark Loss. JABES 16, 88–104 (2011).

- Line 235: I've had a hard time identifying the model with tag-loss you're using. If I understand correctly, it is specified with Eqn 1. This equation should come with the relevant literature, unless you've developed this model yourself. I recommend having a specific section describing both models with and without tag loss, which are common to both the simulations and case study, so that the reader can easily goes back to it. In this new section, you should make it clear how the model with tag-loss you're using differs from existing models with double-marking. This new section could also include the literature review you do earlier in the paper, but if not, you should refer to it.

Minor comments

- Lines 31-33: I find the statement "little attention has been paid to the effect of common violations of the CMR model assumptions" a bit unfair. I understand that the abstract has to be punchy and attractive, but there are a lot of papers out there on the issues of heterogeneity, incorrect state assignment to individuals, and more broadly goodness-of-fit testing for multistate models. I would tone down this statement a little, and simply write that the tag-loss issue is tackled.
- I noticed that you use both multi-state and multistate, I would homogeneise throughout the paper.
- Line 99: Bugs is not a software program, you mean WinBUGS, JAGS, OpenBUGS, STAN, NIMBLE, etc. I would also keep MARK and E-SURGE for user-friendly software here. Please, add the relevant references too, folks have put a lot of efforts in developping these pieces of software, and their dedication has served the community well, the least we can do is to acknowledge their work.
- Sometimes you have initials for first names in the cited references. I guess we do not need them in the text, only in the list of references.
- Lines 104-108: You do a great job at reviewing the literature on tag loss in CJS models, and the applications of multistate models, but here we're missing what you expect to be new and/or different from the existing studies in terms of bias in demographic parameters. Actually, you have these predictions diluted in the paragraph lines 116-125, which is more about a specific case study. The general predictions should come with the simulation study.
- Why going for a Bayesian approach? You don't have random effects in the simulations, nor do you have prior information to incorporate. Also, bias is kind of a frequentist notion, it would have been more natural to go for maximum likelihood estimation. Plus, this would have allowed for more than 50 simulated datasets (for each scenario or combination of parameters), which I guess you had to keep under a limit due to the computational burden of MCMC methods. Estimation of bias requires more simulations in general. I'm not asking to go for a frequentist approach, just to provide the reader with some justifications.
- Line 132: What about the dead state? It is absorbing too ;-)
- Lines 125-126: Not sure there is a formal definition of bias in posterior distribution, isn't it?
- Line 193: What does 'To illustrate the simulations' mean? Some readers would be perfectly happy with simulations only. The case study brings less generality, because you don't know the truth (at least only 66% of it). I like the case study, it just needs to be better introduced: What does it add to the simulations?
- For the bat case study, you need to perform goodness-of-fit tests and cite the relevant literature. Bias in parameter estimates could be caused by transience or trap-dependence issues.
- Line 224: Bias on which parameter(s)?
- Lines 243-248: It is difficult to follow you blindly here without knowing for which parameters convergence was not reached. Actually, with only 50 simulations, we might expect some 'bias' in the results due only to lack of convergence, not tag loss. I encourage you to be more specific.

- Line 237: That's an old version of R ;-) I recommend updating JAGS, JagsUI and R.
- Line 274: Do we care about detection being biased?
- Lines 298-299: I am not sure I follow, it's one thing to quantify bias on survival, it is another to demonstrate bias on the relationship between a covariate and survival. You might have a bias in survival, but if this bias is the same for all individuals or time intervals, then I guess there is no problem to assess the effect of individual or temporal covariate. Am I missing something?
- Lines 311-328: The results are discussed with no reference whatsoever to the existing literature. What was found in other papers? Mobilizing the existing literature and comparing your results to previous findings will help you to emphasize what's new and original in your work.
- Lines 331-332: I've probably missed it, how can you be sure that you quantify bias in the case study with only two third of the individuals for which you know they've lost their tag?
- Line 345: Dispersal, not dispersion (I think).
- Lines 402-404: I applaud the authors for their efforts of making data and code available. I would deposit the code and data on GitHub/GitLab for versioning and also to make it easy to re-use your code (copying and pasting from a PDF to R can produce funky behaviors sometimes).
- Legend of Table 1: You write 'U=Univariate distribution)', you mean 'Uniform distribution' I guess. I also find it unclear the meaning of the minus something you have in the short-live species column. A Uniform(a,b) - 0.5 is Uniform(a-0.5,b-0.5), and Normal(mu,sigma)-0.3 is Normal(mu-0.3,sigma). Did you simulate directly from these 'translated' distributions, or did you simulate from Uniform(a,b) and Normal(mu,sigma) then substract something?
- Figure 1: Nice work!
- Figures 2 and 3: These figures cannot be read independently of the main text. I would remind the reader in the legend what the scenarios are, and what the numbers and letters on the axes refer to. Also I'd define ROPE.
- Line 266: I find 'for most parameters in most situations' to be a vague statement. I encourage you to provide some quantitative information so that the reader can make a judgment by herself/himself.
- Figure S2: I am not sure the relationship will remain significant once you account for uncertainty in both adult and juvenile survival probabilities. You're doing statistics on statistics here.