Comments on "Extending density surface models to include multiple and double-observer survey data" by D. Miller and coauthors

May 7, 2021

1 Basic reporting

Overall the writing was of high quality (see Specific Comments for a few grammar errors and typos I caught), the introduction and literature cited were well referenced, and the figures were great. I was able to access the data and code via David Miller's github site. I suggest this material be archived to a publicly available repository after manuscript acceptance.

2 Validity of findings

I found that the methods, data, and results to be reliable.

3 Experimental design

The statistical methods developed are sound and the text identifies how the research fills into existing knowledge gaps. Code and data provide a useful template from which readers can replicate the author's results and exemplify how they can apply these methods to their own research.

4 General Comments

Miller and colleagues describe a framework for modeling animal count data from multiple types of surveys, including strip transects, point transects, and various distance sampling setups. Specifically, they describe how a common spatial model can be assumed for animal density, and how this can be integrated with different types of observation

processes. The framework uses a two-stage framework, whereby inference is first made about detection probability, and those results are incorporated as offsets into spatial models. Importantly, uncertainty about detection probability is propagated into abundance estimates via a variance propagation method. They show how this framework can be used by applying it to fin whale and fulmar data sets.

I thought the manuscript was well written and represented competent research. In addition, the software extensions to the `dsm` R package should allow other researchers to apply these types of analyses to their own data sets. I view the latter as the manuscript's primary strength. I only have a few very minor comments, mostly on grammar, below. However, the authors might want to reflect on whether the fulmar surfaces should be added together instead of averaged.

5 Specific comments

1. Line 55, for a very recent example of combining various detection methods via integrated likelihood, see Conn et al. 2021 (https://doi.org/10.1371/journal.pone.0251130). Just FYI, I'm not suggesting the authors need to cite this as it uses very similar ideas to the Sigourney paper.

2. Lines 58-62. Various grammar fixes needed here

3. Line 81. Is is worth saying that for line transects it is usually perpendicular distance from the transect line rather than distance when the sampler detects the animal?

4. Lines 150-151. This looks like a subsection heading with a formatting issue?

5. Lines 159-160. Worth talking about this in discussion... e.g. future extensions to software that can partially accommodate detection heterogeneity via point independence

6. Lines 202-206. It would be slightly clearer if things were given subscripts here, maybe with "For each $b \in \{1, 2, \cdots, B\}$:" (not a big deal though)

7. Lines 216-217. The bird example is an interesting one. In this case, wouldn't you add the two (flying and non-flying) rather than take a mean?

8. Lines 237-246. This seems to be the first time availability is mentioned in this paper. Is there capacity in `dsm` to put in availability and it's associated variance? If so, is it a new or existing feature? It seems like it might be a useful topic to discuss earlier in the paper (e.g. combining platforms with and without availability issues)

9. Line 272. Were there "unknown species" recorded?

10. Lines 274-275. Grammar

11. Line 277. A bit of a fourth-decimal-place question, but were x and y lat-long or in projected space?

12. Line 286. Wow! I guess this gives the GAMs flexibility to capture extremely fine scale patterns if they exist, but it's quite a bit larger than default `mgcv` values. Just curious if this is something the authors often do in DSMs?

13. Discussion. Currently there's a paragraph about the fin whale example here but nothing about the fulmar example. Should lessons from the fulmar example be covered here to balance things out?