Review of "Estimating abundance from counts in large data sets of irregularly-spaced plots using spatial basis functions"

Overview

This manuscript presents a model-based methodology for estimating population sizes from an irregularly-sampled set of plots. The authors address the issue of how to introduce a finite population correction in a model-based context and deal with overdispersion. The work is well-motivated by a real-world application in counting seals in Alaska.

I found this to be a well-written, clearly-motivated, practical, careful piece of work. I have a few substantive questions that the authors should be able to readily address.

Major comments

1. I'm unclear on the possibility of a selection bias here. The authors say that the camera is turned off where seals are necessarily absent, but I think they then go on to estimate abundance in those locations from the model rather than treating them as known zeros. It seems to me there is a selection bias possibility that the model may estimate non-zero abundance based on smoothing from nearby areas that have seals even though the counts are known to be zero (perhaps just very close to zero?) in those areas where the camera was turned off. Provided there are data near to the "camera-off" areas that have low abundance, any bias is probably limited, because low abundance will be inferred. But if there were an area of high abundance next to an area where the camera was off, there would be an overestimate for the unsampled area from the smoothing. It would be helpful if the authors can comment on this issue. Perhaps I'm misunderstanding something?

2. The motivating goals listed in Section 1.3 and the interest in the actual number of seals rather than the mean intensity, as well as the interest in frequentist performance, are all nicely laid out, and I thank the authors for their clear presentation of the methodological motivation.

3. The authors use maximum likelihood for estimating basis function coefficients. In many contexts this would be done with some penalization (penalized splines, Bayesian estimation etc.) and doing unpenalized maximization carries a risk of having very high uncertainty. In this context, the authors use a limited number of knots and their estimation does account for uncertainty, so perhaps this is not a major concern, but a sentence or two recognizing this as a potential issue would be warranted. I guess the main concern comes as the number of coefficients gets large relative to the number of observations, in which case asymptotic justification for their variance estimator on page 10 breaks down.

4. On page 10, line 23, the authors decompose the prediction variance into two terms, which I interpret as (1) the variability of the true number of seals around the mean (which is of course based on the intensity function) and (2) the uncertainty from estimating the basis coefficients and therefore in estimating the intensity function. On page 11, line 20, the authors describe the two terms as (1) the variance of predicting the intensity surface given

the regression parameters, and (2) the variance in estimating the regression parameters. Is my interpretation of the first term off-base? I don't understand what the authors mean by the variance of predicting the intensity surface. If the regression model is correct (which I think is an assumption that is implicitly being made here), then the intensity surface is just a deterministic function of the coefficients and there is no variability when conditioning on the regression parameters. The variability comes because of the realization of the actual counts given the intensity. On a related note, in the application, I think it would be very valuable to show the variance decomposition for the actual prediction variance. My intuition is that the second term may often be more important than the first, following my point #3 above.

5. Page 12, top: Why is weighted least squares rather than ordinary least squares the right thing to do here? I worry about robustness to outliers even in unweighted form. Is there a mathematical justification for the weighting - the authors merely state that they want to weight values with large expectation more. More generally, I would have thought that estimation for overdispersion in Poisson models is well-developed in the literature, yet the authors treat this as an open area for investigation. Can the authors confirm that this is still an open question and give the reader a few citations to point to the current understanding of this problem? Is the linear regression estimator completely new or have others proposed this? Given the simplicity of the linear regression estimator, if this has not been proposed before, is there a reason for that?

Minor comments

1. Abstract, line 29: "unsampled area" => "unsampled areas"

2. Page 1, line 50: I didn't follow why "extensions to count data have been difficult". There is lots of work on spatial GLMs where the likelihood is Poisson. Is there some difficulty in going from maps developed based on count data to abundance estimates? Are the issues involved basically those given in Section 1.3 of the manuscript?

3. Page 3, line 24: "days" => "day's"

4. Page 8, line 50: It might be helpful to the reader to explicitly call this a block-wise coordinate descent algorithm.

5. Page 9, line 7: As far as I can tell, R's optim() does not provide for constraints when using Nelder-Mead, only when using BFGS (see the help information on the 'lower' and 'upper' arguments). Can the authors clarify how they imposed the constraints? Furthermore, I believe the constraints in optim() are constants and wouldn't allow for a constraint such as $\rho_C > \rho_F$.

6. Page 9, line 51: "Reimann" => "Riemann"

7. Page 11, line 30: "Overdisperion" => "Overdispersion" !!!

8. Page 18, line 46: A side note that I believe similar convergence issues happen with spatial models with Tweedie (continuous observations with zero inflation) likelihoods as well, though I don't have a good citation offhand.

9. Fig. 4: It would be helpful to use different line types to distinguish the two solid lines in panel A.

10. In Fig. 2, dark is high abundance, while in Figs. 3 and 7, dark is low. It would be good to have the color ramp be consistent in direction. I'd also suggest color as a way of allowing the reader to more easily see the variations than a greyscale, particularly in Figs. 2 and 7, but I'll leave this comment as being one of personal preference.