

This is an impressive data set, rigorous and well-explained analysis, and well-written paper. Almost every concern or question for further consideration that I jotted down while reading early sections of the paper was ultimately addressed in more detail in the Discussion. I have only a few big picture issues for you to consider, if they seem relevant.

I presume many of the volunteers might be hunters who enjoy the opportunity to work their hunting dogs during the pre-season, but perhaps also identify good places to go hunting after the hunting season opens? If this is the case, and depending on whether hunting is additive or compensatory, and depending on whether the decision to go hunting is dependent on how many broods you observe during the survey, there might be some potential unintended feedback in this monitoring system. In Michigan, most young American Woodcock are banded with the help of hunters who appreciate the opportunity to train their dogs during the non-hunting season, but if these same hunters go back and hunt areas that had lots of broods, the resulting harvest and survival rates might be biased. Not sure if this is a possibility in Norway's ptarmigan surveys, but if so, perhaps you should include some Discussion of possible impacts (does the embargo you mention on line 187 related to this, and does it last until the hunting season is over?).

To me, this still qualifies as an IPM if you don't include the wee little bit of telemetry data, as your stage-structured surveys alone are sufficient to estimate fecundity, survival, and abundance. However, if the telemetry data are age-structured, or even if they are adults only, it would be possible to have age-structured survival estimates in your model (e.g. equation on line 257 could be  $S_j * D_j + S_a * D_a$ ), although I suspect HY vs AHY survival would be only partially identifiable. In North America, meta-analyses of radiomarked waterfowl indicate that they survive less well than birds receiving only legbands, so it might be worth using the telemetry data more as a prior, with understanding that it provides a lower bound on true survival of unmarked ptarmigan.

Line 288: Is it feasible to list the key feature(s) of this custom half-normal distribution?

For Figure 5b, I wondered if this was indicative of a cost of reproduction, or of correlated sampling errors in your model (i.e. if you miss high on Recruitment, you will underestimate Survival, and vice versa). You hint that this might be the case later in the discussion when describing the traceplots, but a more formal way to address this might be by looking at correlations between S and R across the sims.list within a single model run. If there is an inherent sampling correlation within sites, it seems like that might also show up as a process correlation among sites. Ultimately, it might require simulations with 0 process correlation between R and S to address this, and I'm NOT recommending you do that for this manuscript, but it may be worth exploring this a bit more and adding additional caveats if warranted.

One of the benefits of a CAR approach to spatial variation might be the ability to predict ptarmigan populations and population trajectories in currently unmonitored areas, especially if you included appropriate landscape covariates. Not suggesting you do this, but you might add it to the already very thoughtful "suggestions for future work".

Some minor wording issues:

Line 19: "structured"

Line 72: “last decade” or “last few decades”

Line 78: omit “the” or add program or scheme after Swiss Biodiversity Monitoring

Line 84: “allows for greatly reduced costs and extended”, or insert subject (conservationists, analysts) between allows and to.

Line 113: change final “as” to “an”

Line 138-140: As a population, willow ptarmigan are plural, but as a species it is singular. So change remain to remains, and later “ptarmigan is considered a sentinel species that is sensitive...”

Line 148: “cycles”

Line 222: “switches”

Line 223: “from” instead of for

Fig. 1 legend: Not clear that “area” refers to the 41 reporting districts unless you refer to the text – recommend adding to the figure legend too, e.g. by adding “(41 reporting districts)” to the end of the second sentence.

Line 284: change on to one

Line 349: omit “the” in front of northern.

Line 374: When you first provide estimates of the random effects, recommend you remind readers of the scale they are measured on, e.g. SD = 0.169 on the logit scale (log scale for Recruitment).

Line 551: “in areas”

Line 566: “are managed”

Line 625: “observer”

Best,

Todd W. Arnold