PROBLEM STATEMENT

For bird point count data with:

- **distance bins 0-25m., 25-50m, 50-100m, 100m+;**
- **time-to-first-detection rounded to the nearest minute, up to 10min;**
- **unrecorded mixed habitats in distance bins of many/most individual points;**
- **repeat surveys in a subset of years**

can either meaningful estimates of annual density (and long-term trend), or at least some form of densities and trend correcting for imperfect detection, be generated, and how can I test the validity and uncertainty about such estimates?

IF I can't analyze the data with a distance sampling and/or time-to-removal approach, can I get meaningful estimates with N-mixture models using the subset of repeat-survey years to estimate detection probability?

NETWORK'S DESIRED END-POINT is annual density estimates and long-term trend for "common" species, analyzed separately for each park-species.

SAMPLING

- **In each park, ~ 20-30 survey plots are somewhat randomly located**—no habitat stratification (Fig. 1) *[Relevant for distance sampling questions]*
- **Point count surveys are 10 minutes. Observer records the first detection of each bird (removal sampling). Observer records the species and a distance bin (no information on sex or auditory vs. visual) and minute of detection (e.g., within 1st min., in 1-2min., in 2-3min., etc.). Distance bins are 0-25m., 25-50m, 50-100m, 100+ (unlimited distance).** *[Relevant for distance sampling questions]*
- **Point count surveys since 2010.** From 2010-2017 they mostly had a rotating panel with half the plots surveyed once each year. In 2018 they surveyed all the plots once. **Starting in 2019 they surveyed every plot twice each year** (second round of surveys within 1 month of first). *[Relevant for N-mixture questions]*
- **A few covariates are recorded for each point count—date, observer name, start time, wind (collapsed to 2 categories), noise (dropping this b/c a lot of missing data).** I have coarse information about **habitat (forested, shrub, open, or mixed) and understory density (LiDAR) within 50m, 100m, and 150m of point center**, but it's unclear what metrics calculated from the LiDAR would be meaningful for detection or occupancy.

Below are my main concerns/questions about analyzing the data with the different approaches. I realize there are also options to combine the approaches and that by combining approaches, some assumptions can be relaxed.

DISTANCE SAMPLING QUESTIONS

- **Because the point count locations (may) have different habitat types and/or roads running through them, the data may often violate the assumption that animals are distributed independent of distance-from-observer.** How do I determine if/when this is a problem? (Since the detection function is fit to data summed across plots in a park, perhaps it's likely any bin-habitat biases at individual points may be cancelled out so it's okay as long as we don't see a big bin-habitat bias overall in the park?)
- **For some park-species, the detection density (number of detections/area) in the 0-25m bin is lower than in the 25-50m bin** (Fig. 2), possibly b/c the birds are avoiding researchers**.** When is this "enough" of an issue that we may need to combine the first two distance bins to meet assumption of monotonic decline in detection density? If two combine the first two bins we only end up with two known-distance bins (0-50m, 50-100m), and we can't fit a distance function to just two known-distance bins.
- **Can we use the 100m+ distance bin in analyses for estimating density, or do we need to drop it?** (Fig. 3) For most park-species, 20-30% of detections occur in this distance bin.
- **Detection functions for auditory detections are going to be different than those for visual detections, but the data don't distinguish the mode of detection.** If the network says that most detections were auditory, would we just mention that as a caveat in the analysis, but still try to analyze the data with this approach?

- **With max 3 distance bins there are not enough degrees of freedom for tests of the fit of the distance functions.** When I googled what people have done in this situation, **I often see that they just eyeball it to see if the distance function is a good enough fit… is that an acceptable alternative?** NOTE: When a hazard rate function has (apparently) fit the data better than a half-normal distance function, sometimes the estimated detection probability is tiny (e.g. 2%) and I don't trust it even though it may "look" like a good distance function fit.
- **It seems that a common practice with distance sampling is to estimate detection probability for each sampling event and then use that as an offset in GLMM. How do you incorporate the uncertainty in the estimated detection probability,** to ensure it is reflected in the 95%CI for the trend estimate? Do you bootstrap from the distribution of the detection probability each year, or does this not matter?

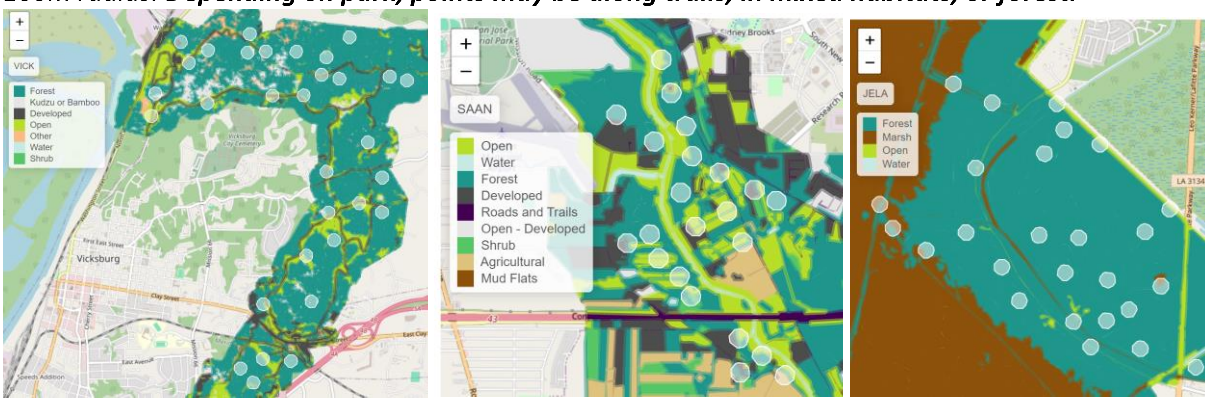
TIME-TO-REMOVAL QUESTIONS

- **Time-to-removal patterns for these data typically look like the examples in Fig. 4. They don't fit well with the assumption of constant call rate**, which should lead to an exponential decline in first-detections over time. Is there a way to still use these data in this type of analysis?

N-MIXTURE MODEL QUESTIONS FROM ELLEN & JP

- **Since we have repeat data for only 2019-present, it seems that the detection model is fit just to this subset of data years and then applied to all years, is that right? And if that is the case, would one possible option be to have the network do repeat surveys every few years so the detection model is fit to a subset of data and the network doesn't have to do repeat surveys every year?**
- **Kery et al. suggested that even when N-mixture models have unreliable estimates of abundance, they are still superior to GLMMs** b/c N-mixture allows covariates separately for detection and abundance. They said that when N-mixture model assumptions are not met (e.g., b/c we have unmodeled heterogeneity in detection), the abundance estimates may be very wrong but trend estimates should be fairly robust so we can just think of it as relative abundance instead of absolute abundance. Would you agree? (I don't)
- **Simulation studies suggest that the GOF tests for N-mixture models often perform poorly.** I've seen a few suggested alternatives but nothing that really stands out as an agreed "best practice" for N-mixture GOF tests. **What have you used?**
- When I plot my N-mixture estimates of annual abundance (y-axis is estimated abundance, x-axis is year), they seem to either follow the pattern of my GLMM results for relative abundance or to have weird occasional wild jumps. But **I feel like a lot of the "textbook examples" of how great N-mixture models are show N-mixture estimates that exhibit much less year-to-year variability compared to fitting data to GLMMs. Is the latter what we should "typically expect"** when we compare results from fitting N-mixture models to results from fitting the same data to GLMMs? (I'm asking this b/c I'm looking for a data-driven way to "figure out" if model assumptions were violated or not.)
- **Using pcount() to get estimated density—Can you actually use the point count circle area as the denominator here?** We are really counting birds in the "superpopulation" of the site, which includes birds whose home ranges just partially overlap the point count circle, right? So the survey area really is not the correct area to divide by?
- **Our network does not record if a detection was auditory/visual or male/female—is this okay for N-mixture models?** The models assume individuals have similar detection probability, right?
- **When we use estimated detection probability to get a "corrected" abundance estimate (w/95%CI) for each year, then put the annual abundance estimates in a GLMM to estimate trend, what is the best way to account for the uncertainty in the annual estimates?** I assume we need to bootstrap--or does this not matter?
- When we are doing estimates for individual species, we have some points that are good habitat, that may or may not have birds, and some points that are completely unsuitable and would never have any birds of that species (process vs. structural 0's). **By including all the "bad" points we will have artificially reduced variance (they are always 0) and the potential for badly mis-estimating the effects of various covariates. How can we spot/correct this in a repeatable and credible manner?**

Fig. 1 – *Example (partial) maps of parks showing point count locations. White circles are point count locations with 100m radius.* ***Depending on park, points may be along trails, in mixed habitats, or forest.***



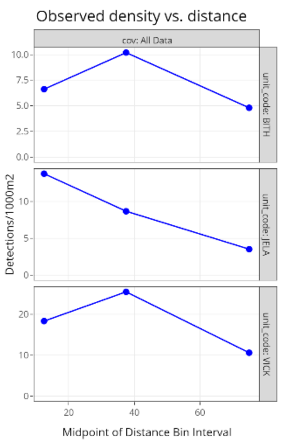


Fig. 2 – *At BITH and VICK, red-bellied woodpecker* ***has a lower detection density in the 0-25m than in the 25-50m bin, violating the assumption of monotonic decline in detection probability with increasing distance.*** *Each figure is a park. The three blue points from left-to-right are detection density in distance bins 0-25m, 25-50m, 50-100m.* ***If we combine the first two distance bins we are left with only two known distance bins—how can we analyze such data with a distance sampling approach?***

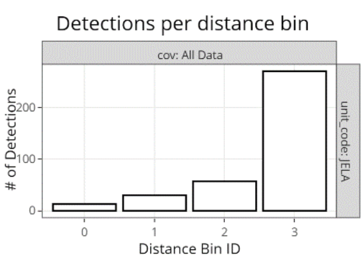


Fig. 3 – *American crow at JELA--# of detections per distance bin (x-axis). Distance Bin 0 is 0-25m from point center, Bin 1 is 25-50m, Bin 2 is 50-100m, Bin 3 is 100m+.* ***Out of the total 370 AMCR detections at JELA, 270 detections are in Bin 3 (100m+). Dropping that 100m+ bin from analyses would drop a lot of the data, but since we don't have a bin cutoff distance I don't think we can use it for estimating AMCR DENSITY.*** *Should these still be analyzed with distance sampling? (Also occurs w/other species like black-crested titmouse)*

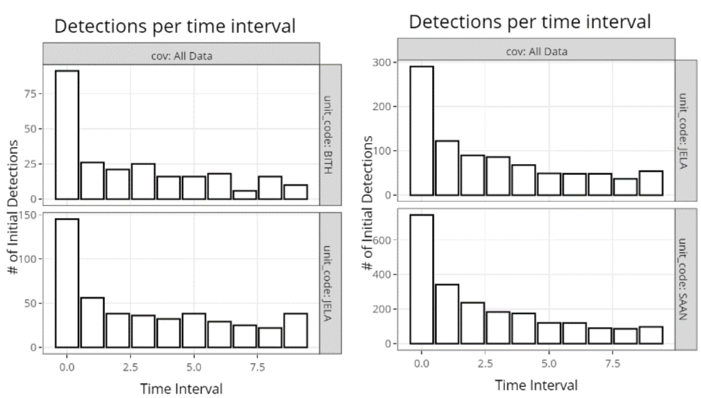


Fig. 4—*Time-to-removal patterns often don't show the exponential decline expected under the assumption of constant call rate. Even if we have enough data to fit a 2-call-rate model, the # of first detections don't taper enough from minutes 2-10 to fit removal functions well. Examples: Left-side plots are for tufted titmouse, right-side are for northern cardinal. X-axis is number of first detections, y-axis is # of minutes since start of point count. Each bar represents a one-minute interval.*