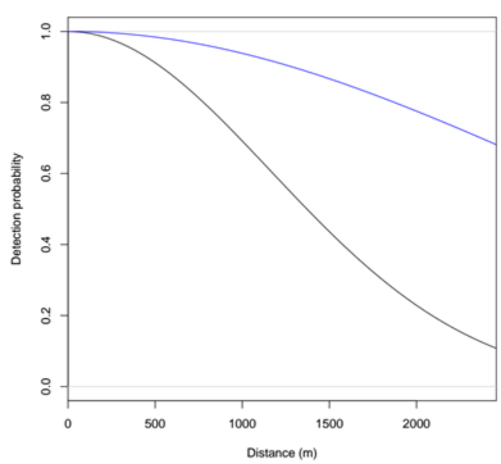
Spatial Capture-Recapture (SCR) Models Fitted to Sonobuoy Data for NMFS

Ben Stevenson

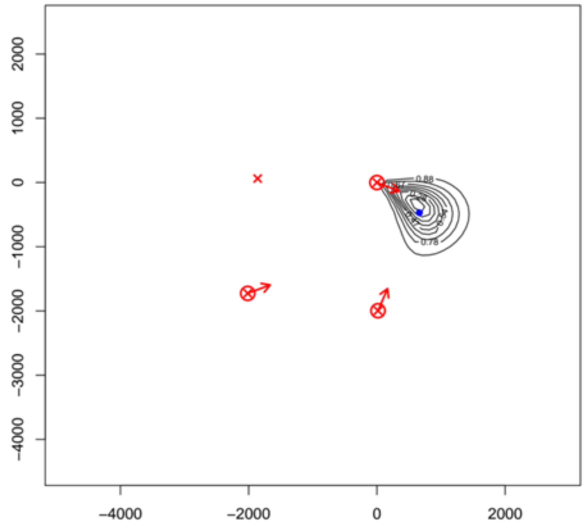
April 2017

Different cue ("Species") were analysed separately as I assumed some would be more easily detected than others; by default SCR estimates the same detection function for all cues, which is not appropriate if some are different to others. There were fairly distinct differences between cue types. The ones ending in "1" (dn1, tone1, upX1, etc) appeared to be harder to detect than the others. For example, see the figure **est-detfn.pdf**. The black line is the estimated detection function for calls upY1, the blue line is for upY5.



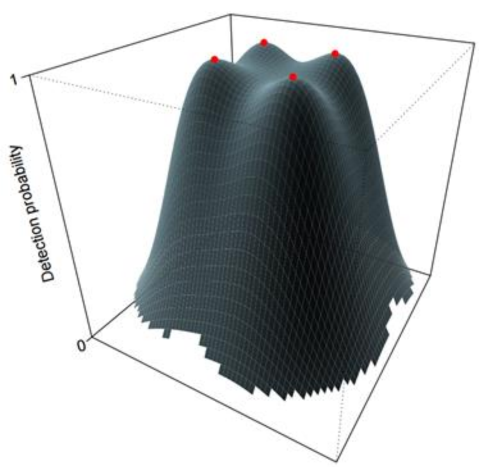
est.detfn.pdf

Unlike distance sampling, spatial capture-recapture does not assume that the locations of cue sources are known. We don't need to accurately estimate the distance from each buoy to each detection. Instead, for each cue, SCR estimates how likely it was for the cue to have originated in each point in space. It does this based on which detectors did (and didn't) detect it, and potentially additional bonus information. In our case the bonus information is the estimated bearings to the cue sources. For example, see figure **est-loc.pdf** for a plot showing this likelihood across space for one particular cue's source location. Red crosses are the buoy locations. A circled red cross means that the buoy in question detected this particular cue. The arrows are the estimated bearings. Essentially, the contours are (sort of) an estimated probability density function (PDF) for where this particular cue source was located. SCR accounts for the uncertainty of where the cue source location is using this PDF, which is why we don't need to know it precisely. The blue dot is the true location of the boat when it emitted the cue (we wouldn't normally know this when we fit an SCR model).



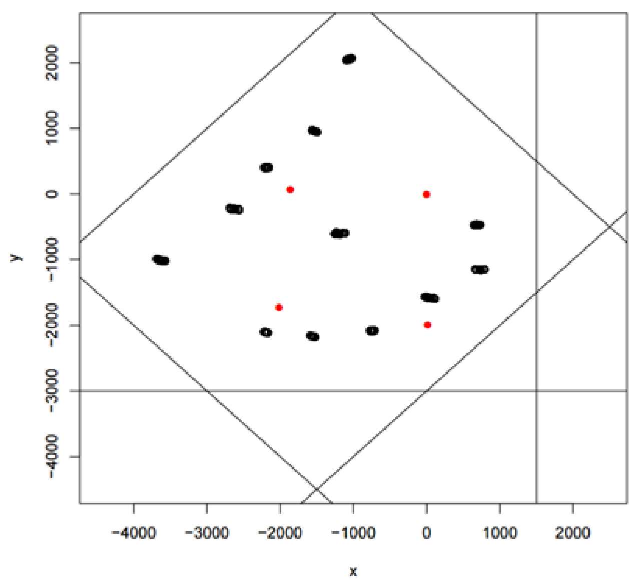
est.loc.pdf

Also see figure **est-detsurf.pdf** for a plot of detectability across the survey area. That is, for each point in space, it shows the probability that at least one buoy would detect a call that was emitted there. This one was generated from a model fitted to upY1 calls. The red dots are the buoy locations. There are lumps in this surface at each buoy location, because if an upY1 call is emitted right below a buoy then it is almost certainly detected, but if one is emitted halfway between two the probability is slightly lower. For the more easily detected calls, this surface is just a flat plane at a probability of 1 (i.e., it estimates that any call within the survey area is detected with certainty).



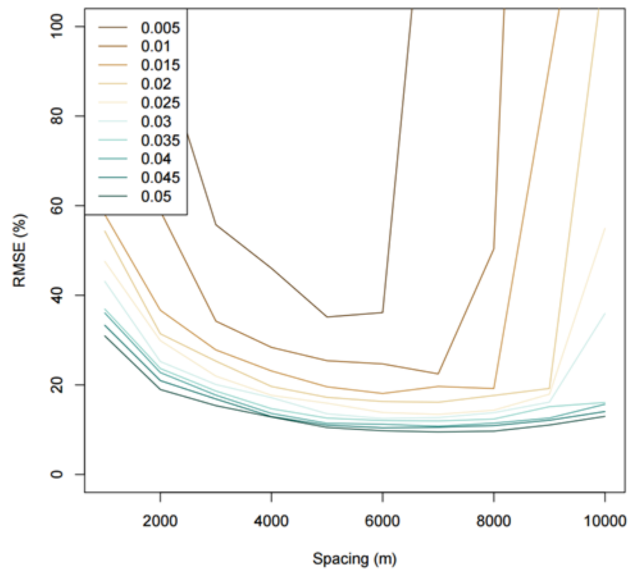
est.detsurf.pdf

The main goal of fitting an SCR model is often to estimate animal or cue density (animals or cues emitted per hectare). In our case, it's a little hard to define exactly what the true cue density is. There were about 30 of each cue emitted -- but to get density we need to know the area in which the cue locations are uniformly scattered. I arbitrarily drew a hexagon-shaped "mask" area around the boat locations (see **mask-area.pdf**) so that it looks like the cue locations are uniformly scattered within it. With upY1 calls, for example, this resulted in an abundance estimate of 22 cues, which isn't a million miles away from the true value of 29 cues. This final step is pretty arbitrary because the cue density estimate it is heavily dependent on the mask area. In practice, for real data of cues detected from actual whales, we don't have to worry about this -- whale cue locations are naturally scattered all over the place near and far, and are not selected by the surveyors driving around in a boat (who don't have the time to scatter their cues near and far all over the place).



mask-area.pdf

The results so far only consider two-buoy simulations. The results were better than I expected. Plot **twobuoy-sim.pdf** shows spacing between the microphones vs root mean square error (RMSE) for various call densities. RMSE is a measure of estimator performance (the smaller the better). For ecology, anything less than 20% is usually considered pretty good. I simulated data using a halfnormal detection function with sigma set at 3000 m, as this was fairly consistent with the blue whale detection functions from your report.



Twobuoy-sim.pdf

Estimation is better at higher densities, which is unsurprising -- larger sample sizes give more precise estimates. Estimation seems to be best at a spacing of around 6000 m, as long as the underlying call density is sufficiently high. Previous studies have shown that if you are using a grid of detectors for an SCR survey, spacing them 2*sigma apart tends to be a good idea, so we're pretty consistent with that. It wasn't obvious to me that we'd see this agreement with only two detectors (these studies used large grids of 50--100 detectors).

The meaning of the density values in the legend are a little difficult to explain; suffice to say that a density of 0.05 (the largest) is roughly equivalent to detecting 500 individual calls across the two microphones. So, for example, to get a RMSE around 20% or less at a spacing of 6000 m, you need to be analysing at least around 150 detected calls.

So, for the question "can we get away with only using two buoys", my answer is "yes, probably", but there are a few points:

1) This all assumes that the estimated bearings are half decent. A lot of the spatial information about where cues originate comes from these. Usually in SCR we rely solely on which detectors detect each call. Basically, I think we can get away with only two detectors in this case because, in comparison to standard SCR, we are trading the spatial information we'd otherwise have from more detectors for spatial information provided by the bearings.
2) I'd still feel more comfortable using four buoys, mostly to improve model robustness and model selection. I have a feeling that the choice of detection function may affect the resulting call density estimates somewhat, and with only two detectors I think you may have reduced ability to adequately decide which to use.