Distance sampling analysis of parrot road-survey data

Francisco V. Dénes – 20/03/2021

Analysis description:

Distance sampling (DS) models were fit for each combination of country, ecoregion and species (henceforth case). Maximum detection distance was fixed at 500 m for all species. While this value may not be optimal for some species and/or habitat types, it encompasses most of the detections (only 58 detections occur farther than 500 m for the entire dataset). More importantly, having a single maximum distance allows for comparisons among cases.

Because the number of individuals in a group can influence detection, for each case we evaluated potential correlation between group size and detection distance, using the Spearman correlation test.

We binned distance data for each case to facilitate fitting of detection functions, using three sets of breaks, 25 m, 50m and 100 m. For each binning setup, we fitted DS models with a half-normal key function. We compared models with no adjustment terms, and with cosine, Hermite polynomial and simple polynomial adjustments, up to order 5. For models where group size was correlated with detection distances, we also fit a DS model with a half-normal key function and a group size covariate.

AIC selection was used to compare models within a distance break set, but it cannot be used to compare models fit to data with different binning setups. Thus, we performed chi-square goodness-of-fit tests to compare the best models from each binning setup to identify the best overall model for each case.

Abundance (N) was calculated by dividing the counts by the estimated detection probability. Density was calculated by dividing N by the length (in km) surveyed for each case.

Results of the best model for each case are shown in the “output\_table.csv” file. For each case, the table includes the following information:

* *Country, ecoregion and species*
* *Km*: total km surveyed
* *Count*: sum of all individuals recorded
* *Encounters\_(<=500m)*: number of independent encounters within the 500 m maximum distance
* *Group\_sizeXdistance\_cor\_test\_p*: p-value of the Spearman correlation test for group size and detection distance
* *DS\_cutpoints:* detection distance binning setup of the best model
* *Detection\_probability*: estimated detection probability within 500 m
* *Detection\_function*: the key function of the best model
* *GoF\_DS\_Xsq*: p-value of the chi-square test of the best model. Values under 0.05 indicate model has a poor fit (in a few cases even the best model had poor fit).
* *N\_(count/p)*: estimated abundance (count divided by detection probability) within 500 m maximum distance (or 1km-wide strip centered on the road).
* *D\_(N/Km)*: estimated density (N divided by Km). This is density / km, but since the maximum detection distance was 500 m (and thus the width surveyed was 1 km), it is also the density / km2.
* *Group\_size\_beta*: DS model coefficient for group size, when the best model included the covariate.

For each case, I also created figures with the histogram of detection distances, a group size X detection distance bi-plot, and the DS model overlaid with the histogram of detection distances using the binning setup of the model (for the supplement).

Analysis was done in R using the “Distance” package:

R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Miller DL, Rexstad E, Thomas L, Marshall L, Laake JL (2019). “Distance Sampling in R.” Journal of Statistical Software, 89(1), 1-28. doi: 10.18637/jss.v089.i01 (URL: https://doi.org/10.18637/jss.v089.i01).