Compare predictions of subregion nest abundance between a few model types

2022-07-08

Here we fit 4 versions of linear models to data on total number of nests observed in each subregion and across the entire study area.

The first 3 model versions are of the form: response ~ year

\*the log model is a linear model fitted to the natural log-transformed number of nests

\*the lm model is a linear model fitted to the untransformed number of nests

\*the glmnb model is a negative binomial GLM fitted to the untransformed number of nests.

\*the glmnb.yr2 model is the same as glmnb, but has an extra term for the quadratic year effect.

For log, glmnb, and glmnb.yr2, the model predictions were back-transformed. Predictions from all model types are plotted against the raw data to evaluate which method provides the best estimates of nest abundance, paying particular attention to predictions for the beginning and end of the study period.

An advantage of the log model is that back transformed model coefficients can be interpreted as the annual proportional change in nest abundance, and valid CI can be calculated for this estimate. An advantage of both glmnb models is that the error distribution is best suited for count data, especially overdispersed count data.

In general the 4 methods yield similar results. However, for species X subregion combinations where the number of nests changes by >1 order of magnitude across the study period, the log and glmnb models tend to highly overestimate nests abundance at the higher end. Thus, the plain linear model seems to provide better estimates at higher abundances, but when observed abundances were near 0, the linear model often yields predictions <0 which are nonsensical for this application. The glmnb.yr2 model provides a much better fit to these species X subregion cases with large changes in abundance, at the relatively small cost of slightly more complicated interpretation of model coefficients.

A fifth option which may yield better end point estimates of nesting abundance (but will not provide an overall trend estimate with valid CI) would be a generalized additive model or other smoothing function. I have not yet formally evaluated this method.







