

eBird trends and models

Laura Graham

26 January 2016

Basic idea

Using Bayesian methods we can correct for biases in citizen science data; extract trends (either long-term or seasonal) in species' occupancies; create deterministic models which test hypotheses around what factors affect species' occupancy; and use these deterministic models as priors on hybrid species' distribution models.

Data

For now, eBird data. Will need to think about (a) which species to focus on (I am thinking North American resident hummingbird species) and (b) what the biases are in general for citizen science [e.g. @Boakes2010; @Bird2013; @Isaac2014a], for the eBird data [e.g. @Sullivan2009], and for the species under study.

Dynamic occupancy models

Dynamic occupancy models [@MacKenzie2003; @Kery2010] model species occupancies by modelling the observation process, the initial occupancy and the extinction and colonisation probabilities. These are modelled using repeat observations in many sites within a closed sampling period (I think in our case this will be by month perhaps?). This has the benefit that covariates on the observation process can correct for imperfect detection due to a range of factors. Using these models, we can create corrected time series of species occupancy in the area under study.

Species' trend analysis

The first steps should be to investigate using time series analysis methods whether there is a significant trend in species occupancy through time. Understanding whether a species is declining, increasing or stable in occupancy can help to inform national biodiversity indicators. We can analyse these trends based on species functional traits and phylogenetic history. This will involve creating time series for a large number of species. We could approach this in two ways - categorical response variable (declining, increasing, stable), and/or continuous response variable (slope of trend).

Next, we can analyse whether there are also seasonal trends in species occupancy. These can help to understand species' relationships with environment and look at synchrony (? I'm less sure about this).

Testing ecological hypotheses

We can use the same dynamic occupancy models to use corrected time series to test ecological hypotheses [along similar lines to @Shoemaker2015a]. This can be done by including covariates on the estimations of initial occupancy, extinction and colonisation. At the moment I am mostly thinking about testing hypotheses of what the drivers of change are (e.g. climate/weather, habitat/land-use change, resource availability etc.). Because the covariates are separate on extinction and colonisation it means that it's possible to tease apart the drivers on each of these.

Creating hybrid SDMs

Deterministic models of species–environment relationships can be incorporated into Bayesian SDMs as priors [Golding2016a]. For example if we can have a good model of a species’ relationship with temperature, this relationship can be used as a prior on the mean function and we can then use this in conjunction with other key variables in a correlative model. This will require using an independent dataset.

References