Occupancy Modeling in JAGS

Modeling the spatial distribution of a species[1]

A fundamental question in landscape ecology seeks to understand how landscape structure shapes variation in the abundance of species. We will use data from the Swiss Survey of Common Breeding Birds to model habitat occupancy by a common, resident bird in the Swiss Alps, the willow tit (*Parus montanus).* The data come from annual surveys of one km^2 quadrats distributed across Switzerland. Surveys are conducted during the breeding season on three separate days, but some quadrats have missing data so that the number of replicate observations is fewer than three. During each survey, an observer records every visual or acoustic detection of a breeding species and marks its location using a global positioning system or, in earlier years, a paper map. Because we are observing a resident species during the breeding season, we assume that the true state (occupied or unoccupied) does not change among sample dates. Occupancy data are summarized[2] in the file `Swiss BB data.csv`.

We want to understand the influence of forest cover and elevation on the distribution of the willow tit. We have data on the number of times a quadrat was searched (`number_visits`) and the number of times the species was detected (`number_detected`). We have data on forest canopy cover (% closure, column `forest`) as well as elevation in meters (column `elev`) for each quadrat.

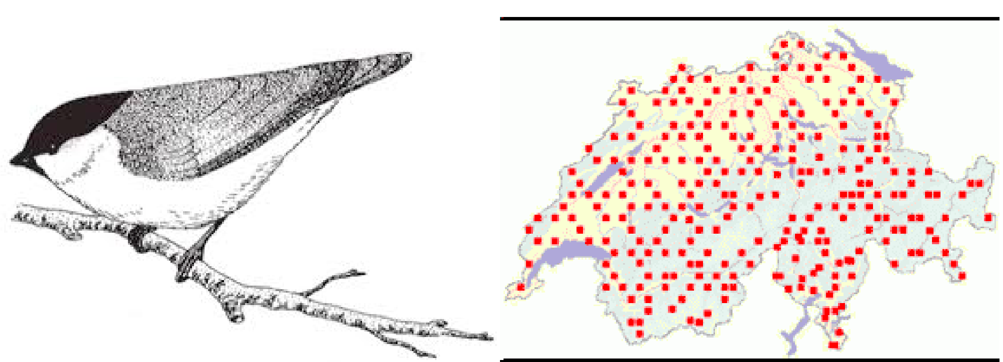

Figure 1: The willow tit (left) is one of 70 bird species that are surveyed annually for abundance in 267 1-km^2 sampling units distributed across Switzerland (right).

[1]Courtesy of Royle, J. A., and R. M. Dorazio. 2008. Hierarchical Modeling and Inference in Ecology: The Analysis of Data from Populations, Metapopulations, and Communities. Academic Press, London, UK.

[2]The original data(`wtmatrix.csv`) are also included in the zip file. There are a couple of cool `apply( )` tricks in the R code that were used to obtain the number of visits and the number of detections. It would be worth understanding these, but I didn't want you to spend an hour figuring out how to do this. Manipulating data is not the point of this lab.

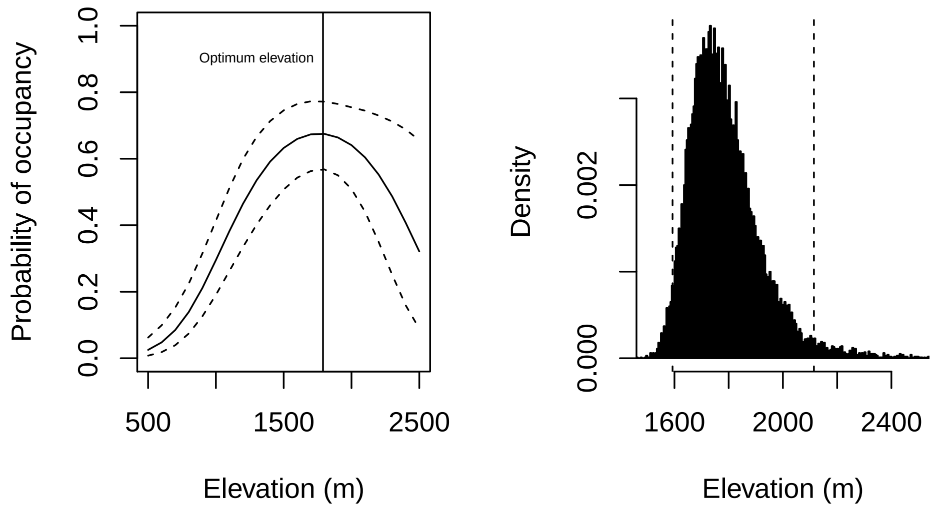

Figure 2: Probability of occupancy at the mean forest cover as a function of elevation (left panel) and the posterior density for the optimum elevation at the mean forest cover (right panel). Dashed give are .025 and .975 quantiles, also known as Bayesian credible intervals.

The challenge in these data is that failure to observe the bird can mean two different things: the bird was truly absent from the quadrat or the bird was present and unobserved. It follows that we must estimate the probability that the bird is present and the probability that we would detect the bird given that it is present.

Develop a model of the influence of forest cover and elevation on the distribution of willow tits. Your model should allow estimation of the optimum elevation of willow tit habitat at the mean forest cover. Diagram the network of knowns and unknowns. Write out a mathematical expression for the posterior and the joint distribution of the data and the parameters. Estimate the posterior distributions of the model parameters using JAGS. Check chains for convergence. Conduct and interpret posterior predictive checks using the mean and standard deviation as test statistics.

What can you conclude about the relative importance of elevation and forest cover in controlling the bird's distribution? Plot the probability of occupancy as function of elevation at the mean of forest cover. Estimate the posterior distribution of a derived parameter, the optimum elevation of the habitat for the bird at the mean forest cover. Your plots should resemble Figure 2.

You will need to standardize the covariates by subtracting the mean and dividing by the standard deviation for each observation in the elevation and forest cover data. A handy function for doing this is: `elev<-as.vector(scale(obs[,"elev"],center=TRUE))` where `obs[,"elev"]` is the unstandardized elevation data in meters. One more coding tip—you *must* give initial values of 1 to all unknown 0 or 1 states.