Homework 5 – Binomial GLMs

Let's consider some data on the prevalence of avian pox among birds on Hawai'i island. In the attached paper the researchers investigated the prevalence of the disease at twelve sites in windward rainforests, over a range of altitudes, in four bird species – three endemic ('Apapane, Hawai'i 'Amakihi, 'I'iwi) and one introduced (Japanese White-eye). Individual birds were assayed for active pox infection (lesions), as well as chronic malaria infection (birds that survive malaria become chronically infected). 'I'iwi is highly susceptible to malaria, while 'Apapane and Hawai'i 'Amakihi are moderately susceptible, and the White-eye is resistant.

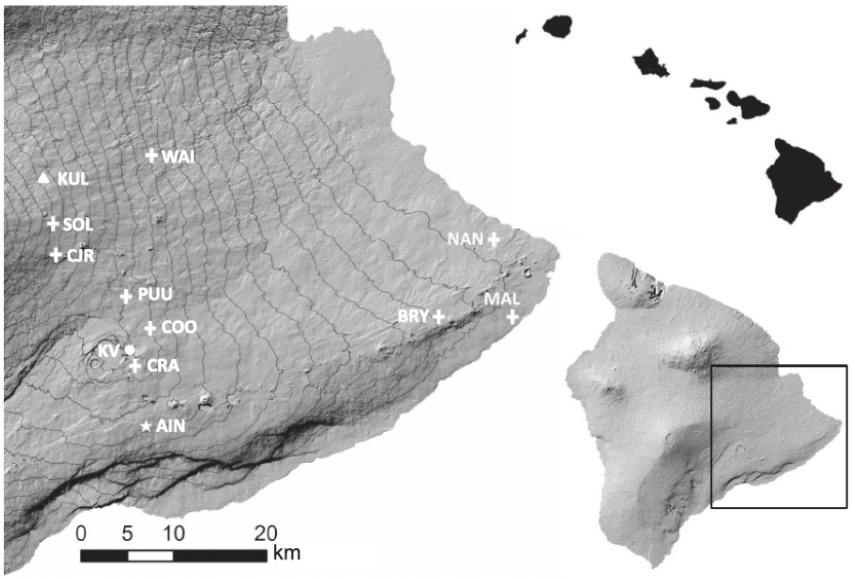

FIG. 1. Birds were captured from 12 sites that were part of four separate studies: the Biocomplexity (crosses), Kilauea Volcano (KV circle), Ainahou (AIN star), and Kulani (KUL triangle) study sites were located on the eastern slope of Mauna Loa and Kilauea Volcanoes on the Island of Hawai'i. Biocomplexity study sites are Bryson's (BRY), Malama Ki (MAL), and Nanawale (NAN) at low elevation, Crater (CRA), Cooper (COO), Pu'u (PUU), and Waiakea (WAI) at mid elevation, and C. J. Ralph (CJR) and Solomon (SOL) at high elevation. See Methods/*Study species and area* and Appendix S1: Table S1 for additional information on study sites.

1. The twelve sites have been coded as low/medium/high elevation. First, summarize the number of observations of each bird species at each elevation. This will be important context for what we can and cannot ask about avian pox patterns. Also, what is the prevalence of chronic malaria in the bird species at the difference elevations? Note that some birds do not have a known malaria status (code = 2). Only use birds with code = 0 or 1 to quantify the prevalence of malaria.

2. Create a model where the presence/absence of pox can differ between species and between elevations. Allow the effect of elevation to differ between species. Make an effects plot of the fitted model and perform likelihood ratio tests on the predictors. Provide a verbal explanation of what the model tells us.

Look at summary() of the model, and note that there is an NA for one of the coefficients. Why do you think this is?

3. Create a subset of the data that only includes birds with known malaria status (0 or 1). Also, you'll want to exclude 'I'iwi from this model (look at your answer to #1 to see why). Create a model where pox prevalence varies by species, elevation, and malaria status, and where the effect of elevation and the effect of malaria status can differ between species. Make effects plots and perform likelihood ratio tests on the predictors. Provide your interpretation of what the results mean.

Disclaimer: the authors of the study used a generalized linear mixed model, which included a random effect for Site. This is indeed a better way to analyze these data, because it accounts for the fact that there may be large differences between sites in the same elevation group, a form of non-independence in the data. We'll learn more about this later in the course.