Homework 1 – Foraging heights of Hawaiian birds

   

In this assignment we will analyze data from a study by Rankin et al. (2018). The authors investigated how invasive rats (*Rattus rattus*) affect the ecology of Hawaiian birds (shown above: ʻApapane, Hawaiʻi ʻAmakihi, and ʻIʻiwi, which are native, and the Japanese White-eye, which is introduced). The authors were specifically focused on whether competition for food (fruits, seeds, arthopods, etc.) would cause forest birds to shift their foraging to higher locations in the canopy when rats are present. The rationale is that the rats overlap in diet with the birds, and the rats can consume arboreal prey (they can climb trees), but the rats cannot access the whole canopy (rats did not climb higher than 12 m in the study). To test the role of the rats the authors performed experiments in kipuka (forest fragments) on Hawai'i island, where they experimentally removed rats from some of kipuka, while leaving other kipuka as unmanipulated controls. The authors also hypothesized that fragment size would alter the effect of rats: kipuka vary greatly in area, and smaller kipuka tend to have shorter trees and lower arthropod biomass.

The columns of the data file are: SPECIES (bird species: APAP, HAAM, IIWI, and JAWE are abbreviations of the four species depicted above), Kipuka (which kipuka the observation was taken in), Area_ha (area of the kipuka in hectares), Rat_Removal (experimental treatment), DATE (date of observation), OBSERVER (which observer recorded the observation), Year, tot.arth.bm (total arthopod biomass, measured via sticky traps at different heights in the canopy), dietary.grouping (whether the bird species is primarily a frugivore, insectivore, or nectarivore), and foraging.ht.m (foraging height of the observed bird, in meters). Each row of the dataset is one observation of a foraging bird.

(1) Start with some data exploration. How many observations are there of each bird species in each experimental treatment? Make a boxplot or violinplot to show the distribution of foraging heights of each bird species in each experimental treatment. What have you learned so far?

(2) Now make a plot showing the mean foraging height of each bird species in each treatment, and include error bars displaying +/- one standard error of the mean. What is the meaning of the standard error of the mean? How does this plot differ from the plot in #1?

(3) Fit a linear model that tests whether the bird species differ in foraging height, whether the rat removal treatment affects foraging height, and whether the effect of the rat removal differs between the bird species. Make plots to assess the

distribution of the residuals. Report F-tests for the terms of the model. Create an effects plot displaying the fitted effects. How do you interpret the results? What are the magnitudes of the effects?

(4) So far we have examined the effect of the rat removal, but we have not considered the role of kipuka size, or the role of food availability (which is influenced in part by kipuka size). Make a new model that adds kipuka area and arthopod biomass as predictors, and include interactions that test whether the treatment effect depends on these additional predictors.

Kipuka area and arthopod biomass are both continuous predictors – before you add them to the model, make some scatterplots to assess whether these predictors should be transformed when including them in the model. When assessing predictors for transformation, it doesn't matter if the predictors are normally distributed (this only matters for the response variable) – what matters is whether the predictors are very skewed, such that a few outlying points will have a large influence on a fitted regression line. For skewed predictors, a log or square root transformation will generally help.

Report F-tests and effects plots. How do the results of this model differ from the model in #3? How do you interpret the results at this stage?

(5) One thing we have not accounted for in our models is the fact that the study design involves repeated measurements of the same kipuka (and potentially the same birds) over time. We will learn how to properly account for repeated measures, by adding a random effect for kipuka, later in the course. However, there is a different additional predictor that may be helpful in testing the treatment effect. Add DATE as a factor to the linear model from #4. How does the inclusion of the date of sampling alter the model results? Why do you think that is?