**Lab Assignment 4 – Optimizing Scale in Wildlife-Habitat Analyses**

**WFSC 570 Wildlife Habitat Analysis**

**Due October 15, 2024**

For this assignment, you will analyze some presence-absence data by creating a multi-covariate pseudo-optimized multi-scale wildlife-habitat model. You will receive landscape data measured at multiple spatial scales, extract some additional landscape data across the same range of spatial scales, use the pseudo-optimization procedure to identify the scale of effect for each covariate, and then fit a final model where each covariate is represented at its scale of effect.

These data are inspired by the Madrean alligator lizard (*Elgaria kingii*). These really cool lizards are found throughout southeast and central Arizona and are associated with a variety of habitats from grasslands up into montane conifer forest. These data represent presence-absence surveys across 171 sites and we will assume that the hypothetical observers were able to perfectly detect alligator lizards at sites where they were present (but we know that this is an unreasonable assumption in reality, particularly for Madrean alligator lizards!).

A lizard on the ground

Description automatically generated

*Madrean alligator lizard from Mt. Lemmon*

Your data for this lab include a single CSV file (*ELKI\_data\_scale\_lab.csv*) with various columns:

* *ELKI*
  + A binary covariate denoting the presence or absence of Madrean alligator lizards at each site (again, we are assuming perfect detection)
* *x* and *y*
  + The UTM coordinates for each site in NAD83 Zone 12 N (EPSG code = 26912)
* *For\_100* through *For\_1200*
  + Proportion of circular buffers (100 – 1200 m radii) that is forest land cover from NLCD classes 41, 42, 43 (<https://www.mrlc.gov/data/legends/national-land-cover-database-class-legend-and-description>)
* *Grass\_100* through *Grass\_1200*
  + Proportion of circular buffers (100 – 1200 m radii) that is grassland/herbaceous land cover (NLCD class 71)
* *Water\_100* through Water*\_1200*
  + Proportion of circular buffers (100 – 1200 m radii) that is open water land cover (NLCD class 11), woody wetland land cover (NLCD class 90), or emergent herbaceous wetland land cover (NLCD class 95)
* *PREC\_100* through PREC*\_1200*
  + Mean precipitation from PRISM climate data across circular buffers (100 – 1200 m radii)

You also have a .tif file (*Final\_DEM\_for\_scale\_lab.tif*) which is the digital elevation model (DEM) for the study area. This file is also in NAD83 Zone 12 N (the same coordinate reference system as the UTM coordinates for each site).

**Complete the following questions/tasks:**

**Task I: Measure the elevation at each site and calculate the mean elevation around each site using circular buffers with radii ranging from 100-1200 m**

To complete this task, first read your CSV file (*ELKI\_data\_scale\_lab.csv*) into R to create a data frame containing all your data (both your response variable [*ELKI*] and your covariate data). Next, read in the DEM .tif file using the *rast*() function in terrra. You will need to use the *extract*() function within terrra to extract the elevation from each site and add this to your data frame containing the rest of your covariate data.

Next, you will need to measure the elevation surrounding each site at multiple spatial scales. Use the same range of spatial scales as you have for the other landscape covariates.

buffer\_sizes <- seq(100,1200,by = 100)

Use the approach we used in lab where you buffer each site by a circle of the appropriate radius (i.e., scale) and then use the *extract*() function to calculate the mean elevation within each buffer. You will first need to create a SpatVector point object representing your sites in order to use terra’s *buffer*() function to create your buffers. Here is an example from lab of how we used this same procedure. You will need to modify this code with the correct object names but this should give you a good template to work from.

for(i in 1:length(buffer\_sizes)){

buff\_i <- buffer\_sizes[i]

cat("Starting buffer size =",buff\_i,"\n")

site\_buffers\_i <- buffer(site\_data\_SV,

width = buff\_i)

elev\_buffer\_i <- extract(DEM,

site\_buffers\_i,

fun = "mean")

HYAR\_data$TMP <- elev\_buffer\_i$DEM

colnames(HYAR\_data)[which(colnames(HYAR\_data)=="TMP")] <- paste0("Elev\_",buff\_i)

}

**Task II: Calculate topographic position index (TPI) at each of your spatial scales**

Remember that TPI is the difference between the elevation at a site and the average elevation surrounding that site. You have just calculated the average elevation around each site at a range of scales (i.e., 100-1200 m buffers). To complete this task, you will simply need to difference the elevation at each site from the average elevation around that site. Add these new TPI columns to your data frame containing the rest of your data.

**Task III: Determine the scale of effect of each landscape feature**

Use the pseudo-optimization procedure from lab to identify the scale of effect of each covariate (forest, grassland, water, precipitation, elevation, and TPI). You can modify this code from lab:

Elev\_scales <- data.frame(Covariate = "Elev",

Scale = seq(100,1200,by = 100),

AIC = NA)

for(i in 1:nrow(Elev\_scales)){

cov\_i <- paste0(Elev\_scales$Covariate[i],"\_",Elev\_scales$Scale[i])

data\_i <- HYAR\_data\_final[,c("HYAR",cov\_i)]

model\_i <- glm(data\_i[,1] ~ data\_i[,2], data = data\_i, family = binomial)

Elev\_scales$AIC[i] <- AIC(model\_i)

}

Elev\_scales

plot(AIC ~ Scale, Elev\_scales,

type = "b",

xlab = "Scale (buffer radius in meters)",

main = "Scale of effect for Elevation")

**Answer the following questions:**

1. **What is the scale of effect for each covariate?**
2. **What is the most common scale of effect across these covariates?**
3. **Look at the binomial GLM for each covariate at its scale of effect. Does this covariate have a significant (p < 0.05) effect on the presence of Madrean alligator lizards at this scale?**

**Task IV: Fit a multi-covariate pseudo-optimized multi-scale model**

Once you have identified the scale of effect for each covariate, take those covariates at their scales of effect and combine them into a multi-covariate pseudo-optimized multi-scale model. Your model should include all six covariates (forest, grassland, water, precipitation, elevation, and TPI).

**Answer the following questions:**

1. **What covariates show significant (p < 0.05) relationships with Madrean alligator lizard presence in this multi-scale model?**
2. **Which covariates have the strongest effect(s) on Madrean alligator lizard presence?**
3. **Did the significance (or non-significance) of any covariates change when you went from the single-covariate model to this multi-covariate model?**

**Bonus Question (5 points):**

Lets say that you decided to not consider the role of spatial scale when investigating the relationships between Madrean alligator lizard presence and these landscape features. You assume that because Madrean alligator lizards are small that they only respond to landscape features immediately around the locality at which you find them. You therefore only measured each covariate using 100-m radius buffers and fit a multi-covariate single-scale model. Fit this single-scale model. Does it appear to be a better performing model than your multi-scale model? Do any of your inferences change when you use this single-scale model?

Your assignment is to work in pairs, complete the following tasks and answer the following questions as a pair, and submit a single written report for each pair describing how you completed the tasks and your answers to the questions. Reports should be written using complete sentences and paragraph structure. Also include your R script (either as a separate file and copied-and-pasted into the end of your report).

**This assignment will be due on D2L by the beginning of lab (2:00) on Tuesday October 8, 2024.**