Using geodiversity to improve SDMs for data poor species

Beth Gerstner

1/2/2023

Project: Using geodiversity to improve SDMs for data poor scientific\_name

Description: This code searches for candidate scientific\_name to use by running spThin over an entire suite of scientific\_name from BioModelos, saving scientific\_name datasets as seperate files (full scientific\_name occurrence datasets and spatially thinned datasets). Looking for both data rich and data poor primate scientific\_name. There has to be enough occurrence records to be considered rich after using spThin to remove sampling bias from the datasets.

Authors: Beth E. Gerstner

Date: 12/29/22

spThin test

Read in full occurrence dataset for Bassaricyon neblina. THis species is not part of BioModelos and instead is from the literature.

occ\_full <- read.csv("/mnt/ufs18/home-048/gerstn11/geodiversity/bassaricyon\_neblina\_occs\_clean.csv")  
occ\_full %>% count(scientific\_name)

## scientific\_name n  
## 1 Bassaricyon neblina 33

Set working directory for thinned records and run thinning function.Takes the master scientific\_name list for endemic scientific\_name, splits them into seperate scientific\_name files and thins each scientific\_name dataset. Saves as it’s own file.

setwd("/mnt/ufs18/rs-008/plz-lab/DATA/neotropical\_frugivores/andes\_geodiv/occurrence\_records/")  
  
for(i in unique(occ\_full$scientific\_name)){  
 scientific\_name1=occ\_full[occ\_full$scientific\_name==i,]  
 f\_name <- str\_replace(i, " ", "\_")  
 write.csv(scientific\_name1, paste(f\_name,".csv",sep=""))  
 thinned\_dataset\_full <-thin( loc.data = scientific\_name1,   
 lat.col = "latitude", long.col = "longitude",   
 spec.col = "scientific\_name",   
 thin.par = 10, reps = 100,   
 locs.thinned.list.return = TRUE,   
 write.files = TRUE,   
 max.files = 5,   
 out.dir = paste(f\_name,"\_thinned\_full/", sep=""), out.base = paste(f\_name,"\_thinned", sep=""),   
 write.log.file = TRUE,  
 log.file = paste(f\_name,"\_thinned\_full\_log\_file.txt", sep=""))   
}

Perform same task as above, but run for all species.

occ\_full <- read.csv("/mnt/ufs18/home-048/gerstn11/geodiversity/records\_primatespp\_depured\_private\_rm\_2022.csv")  
occ\_full %>% count(species)

## species n  
## 1 14  
## 2 Alouatta palliata 164  
## 3 Alouatta seniculus 1362  
## 4 Aotus brumbacki 56  
## 5 Aotus griseimembra 233  
## 6 Aotus jorgehernandezi 8  
## 7 Aotus lemurinus 263  
## 8 Aotus nancymaae 157  
## 9 Aotus vociferans 51  
## 10 Aotus zonalis 59  
## 11 Ateles belzebuth 224  
## 12 Ateles fusciceps 84  
## 13 Ateles geoffroyi 85  
## 14 Ateles hybridus 240  
## 15 Cacajao melanocephalus 58  
## 16 Cacajao ouakary 18  
## 17 Callimico goeldii 18  
## 18 Cebuella pygmaea 138  
## 19 Cebus albifrons 490  
## 20 Cebus apella 308  
## 21 Cebus capucinus 943  
## 22 Cebus malitiosus 8  
## 23 Cebus versicolor 40  
## 24 Cheracebus lucifer 382  
## 25 Cheracebus lugens 65  
## 26 Cheracebus medemi 22  
## 27 Lagothrix lagothricha 428  
## 28 Lagothrix lagotricha 120  
## 29 Leontocebus fuscus 91  
## 30 Leontocebus nigricollis 495  
## 31 Pithecia hirsuta 404  
## 32 Pithecia milleri 52  
## 33 Plectorucebus caquetensis 46  
## 34 Plecturocebus caquetensis 100  
## 35 Plecturocebus discolor 27  
## 36 Plecturocebus ornatus 224  
## 37 Saguinus geoffroyi 310  
## 38 Saguinus inustus 34  
## 39 Saguinus leucopus 802  
## 40 Saguinus oedipus 474  
## 41 Saimiri cassiquiarensis 428  
## 42 Saimiri sciureus 86  
## 43 Sapajus apella 633

Takes the master scientific\_name list for endemic scientific\_name, splits them into seperate scientific\_name files and thins each scientific\_name dataset. Saves as it’s own file.

for(i in unique(occ\_full$species)){  
 species1=occ\_full[occ\_full$species==i,]  
 f\_name <- str\_replace(i, " ", "\_")  
 write.csv(species1, paste(f\_name,".csv",sep=""))  
 thinned\_dataset\_full <-thin( loc.data = species1,   
 lat.col = "latitude", long.col = "longitude",   
 spec.col = "scientific\_name",   
 thin.par = 10, reps = 100,   
 locs.thinned.list.return = TRUE,   
 write.files = TRUE,   
 max.files = 5,   
 out.dir = paste(f\_name,"\_thinned\_full/", sep=""), out.base = paste(f\_name,"\_thinned", sep=""),   
 write.log.file = TRUE,  
 log.file = paste(f\_name,"\_thinned\_full\_log\_file.txt", sep=""))   
}

```