

# Collecting inputs for the biogeography model

Patrick McKenzie

June 6, 2017

This document walks through the process of collecting the two types of inputs for the biogeography model: a single environmental matrix, and a “current distribution” matrix for each species.

Two key functions used are `sdm_function()` and `conv_sdm_rasters_threshold()`. Both of these are included in the “R” folder. The functions use the MaxEnt species distribution modeling program to produce a distribution map for each species.

## Choosing the species and the geographic area

For maxent species distribution modeling, we chose 14 species of *Pedicularis*, a large genus of flowering plants concentrated in mid-southern China, from across its phylogeny. The geographic region from which data was collected was bounded by longitudes of 80 and 112 degrees and latitudes of 18 and 40 degrees.

The following code defines the species we’re using and creates directories for maxent files.

```
species <- c("salviiflora","muscoides","trichoglossa","ingens","axillaris","armata","rhinanthoides","re")
for (i in species) {
  dir.create(paste0("example_species_dist/species_dist_maps/",i))
}
```

Then we run MaxEnt for each species using `sdm_function()` to produce the distribution maps.

```
dir <- paste0("example_species_dist/species_dist_maps/",species)
speciesnames <- paste0("Pedicularis ",species)
for (i in 1:length(species)) {
  sdm_function(
    directory.path = dir[i],
    coords = c(80,112,18,40),
    species.name = speciesnames[i],
    climate.map.path = "example_species_dist/climate_map/wc2-5"
  )
}
```

## Creating the environmental matrix

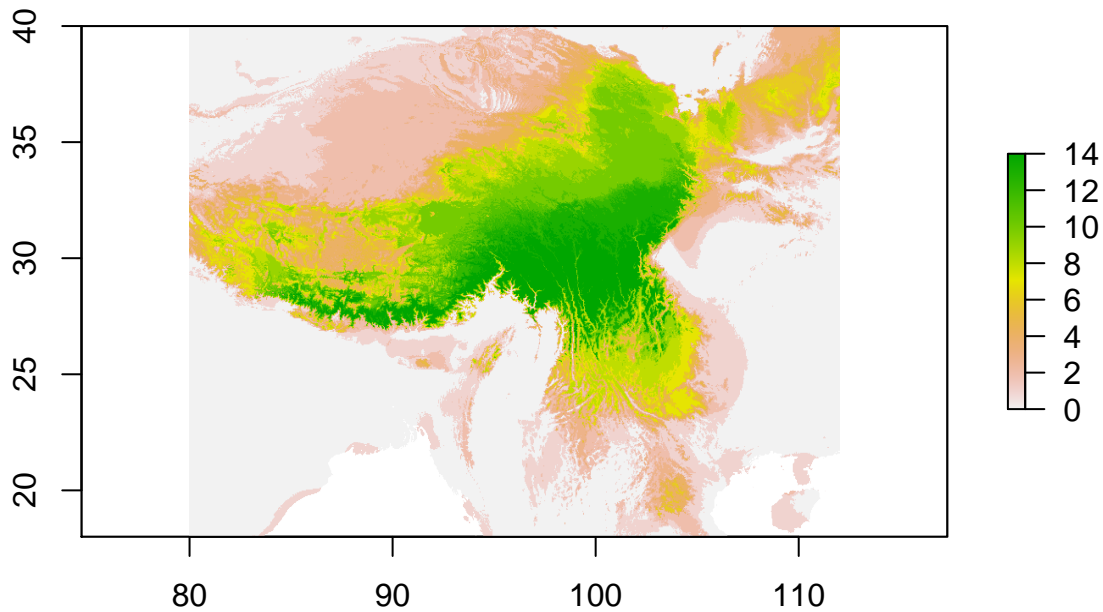
To make a single environmental matrix for the phylogeography method, we read in each species’ maxent output raster individually and apply `conv_sdm_rasters_threshold()`, which uses a threshold value to assign either 1 or 0 to each cell in the maxent output rasters. We then sum all raster layers and use `as.matrix()` to convert the summed distribution rasters to a matrix. This was then saved as a .csv file.

```
rasterfilenames <- paste0(gsub(" ","_",speciesnames),".asc")
specieslists <- list()
for (i in 1:length(species)) {
  raster_path <- paste0("example_species_dist/species_dist_maps/",species[i],"/outputs/",rasterfilenames[i])
  results_file <- paste0("example_species_dist/species_dist_maps/",species[i],"/outputs/maxentResults.csv")
  specieslists[[i]] <- conv_sdm_rasters_threshold(raster_path = raster_path,
    results_file = results_file,
    "balanced")
}
```

```

}
environmental_matrix <- sum(stack(specieslists))
write.csv(as.matrix(environmental_matrix),"example_species_dist/phyloproject_inputs/environmental_matrix.csv")

```



## Current species ranges

We want to save matrices for the individual species' ranges as inputs for the biogeography model. Again, we use `conv_sdm_rasters_threshold()`. However, we now use the most conservative threshold, "10pct," and we write each one as a .csv file.

```

rasterfilenames <- paste0(gsub(" ","_",speciesnames),".asc")
for (i in 1:length(species)) {
  raster_path <- paste0("example_species_dist/species_dist_maps/",species[i],"/outputs/",rasterfilenames[i])
  results_file <- paste0("example_species_dist/species_dist_maps/",species[i],"/outputs/maxentResults.csv")
  specieslayer <- conv_sdm_rasters_threshold(raster_path = raster_path,
                                           results_file = results_file,
                                           "10pct")
  write.csv(as.matrix(specieslayer),paste0("example_species_dist/phyloproject_inputs/",species[i],".csv"))
}

```

Now we can see the current ranges for each species.

