SDM Homework 3

Michael W Belitz

3/18/2020

1) Download the point records. We'd Suggest doing that on GBIF

I downloaded point records of *Chaetodipus californicus* using GBIF's api. Code shown below.

```r
mouse_occ <- occ_search(scientificName = "Chaetodipus californicus",
          fields = c("gbifID", "basisOfRecord", "catalogNumber", "collectionCode",
                     "collectionID", "recordedBy", "continent", "countryCode",
                     "stateProvince", "county", "locality", "year", "month", "day",
                     "decimalLatitude", "decimalLongitude", "coordinateUncertaintyInMeters",
                     "establishmentMeans", "scientificName", "class", "order",
                     "family", "genus", "specificEpithet", "infraspecificEpithet",
                     "taxonomicStatus"), limit = 500000)

mouse_occ_df <- mouse_occ$data
```

2) Clean the records using tools in R. Your approach can be your own but show your code.

Below is the code I used to clean the data. I removed points without latitude and longitude data, and I also removed data with large coordinate uncertainty, given the small home range of this species. I then removed all points that fell 1.5 latitudinal degrees outside of the range map provided by the IUCN. The large buffer (~160 km) around the range map will include for species that are found in areas outside the expert range map, while not including related species that are misidentified outside of the range.

```r
# remove occurrence points with NA values
mouse_occ_df_clean <- mouse_occ_df %>%
  filter(!is.na(decimalLatitude)) %>%
  filter(!is.na(decimalLongitude))

# remove occurrence points with large coordinate uncertainty (5 km)
mouse_occ_df_clean2 <- mouse_occ_df_clean %>%
  filter(coordinateUncertaintyInMeters < 5000)

# read in range map
rm <- shapefile("redlist_species_data_a4f9b948-1974-4534-bb9f-936c1b3cdede/data_0.shp")

# make everything a simple feature
rm_sf <- st_as_sf(rm)

# Visualize where the points fall on a map
na <- rnaturalearth::ne_countries(continent = "North America", returnclass = "sf")

ggplot() +
```
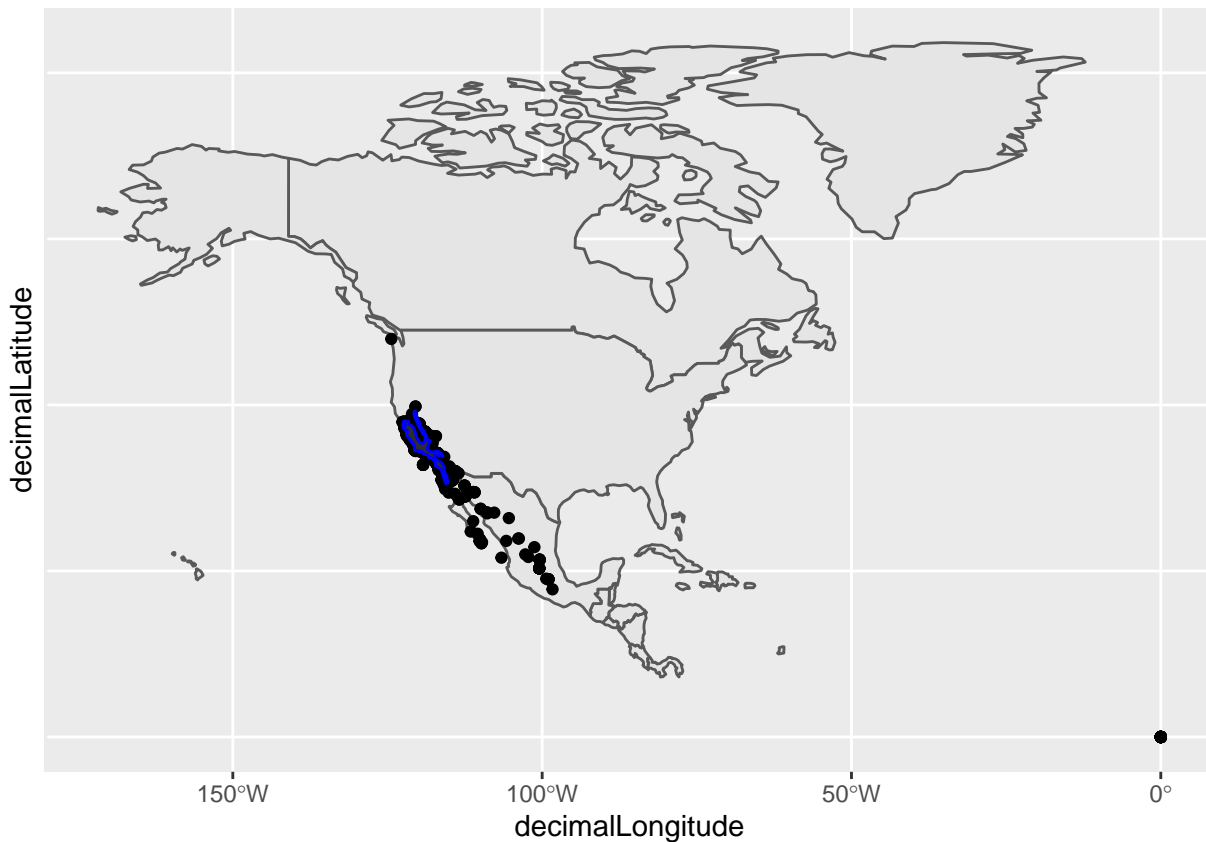
```
geom_sf(na, mapping = aes()) +
geom_point(mouse_occ_df_clean, mapping = aes(x = decimalLongitude, y = decimalLatitude)) +
geom_sf(rm_sf, mapping = aes(), color = "blue", alpha = 0.2)
```



```
# I'm going to clip occurrence points outside of 1.5 latitudinal degrees of the range map
rm_sf_buffer <- st_buffer(rm_sf, dist = 1.5)
```

```
## dist is assumed to be in decimal degrees (arc_degrees).
```

```
occ_sf <- st_as_sf(x = mouse_occ_df_clean, coords = c("decimalLongitude", "decimalLatitude"))
st_crs(occ_sf) <- "+proj=longlat +datum=WGS84 +no_defs"
```
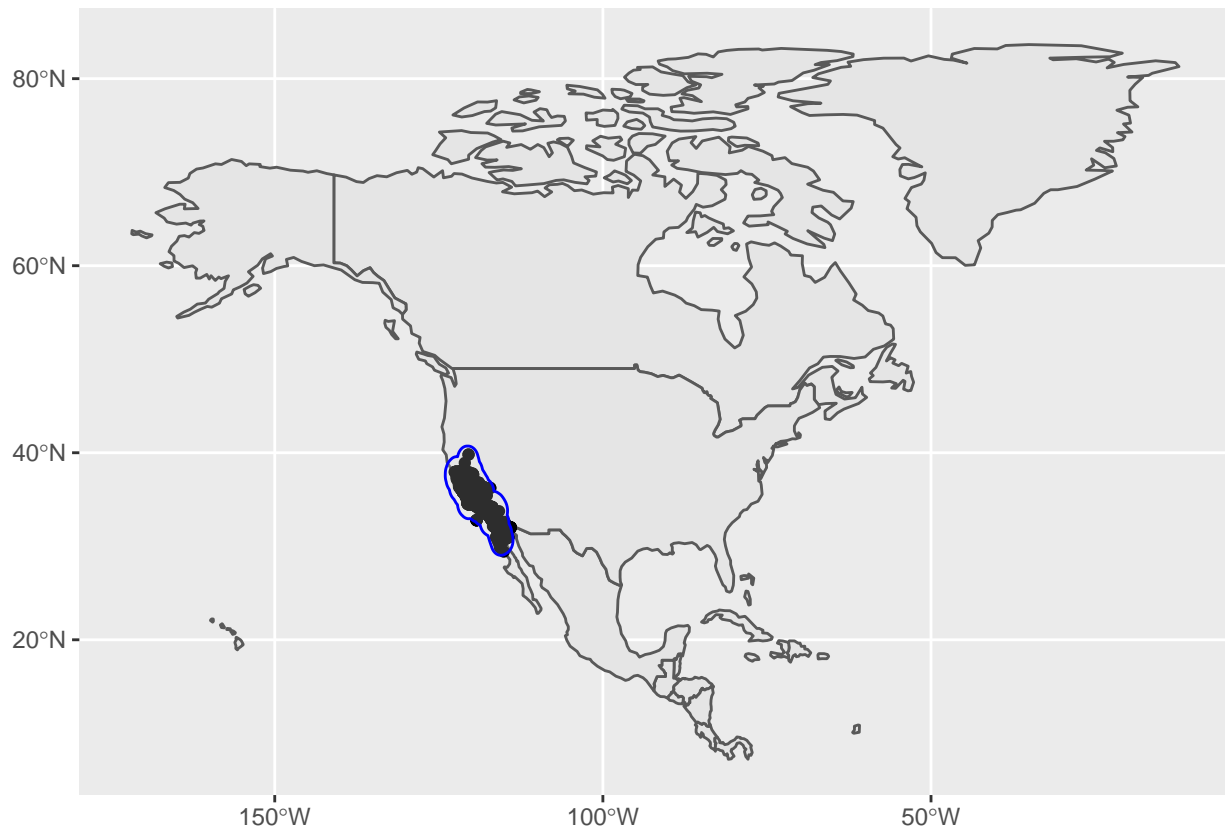
```
ggplot() +
  geom_sf(na, mapping = aes()) +
  geom_sf(occ_sf, mapping = aes()) +
  geom_sf(rm_sf_buffer, mapping = aes(), color = "blue", alpha = 0.2)
```
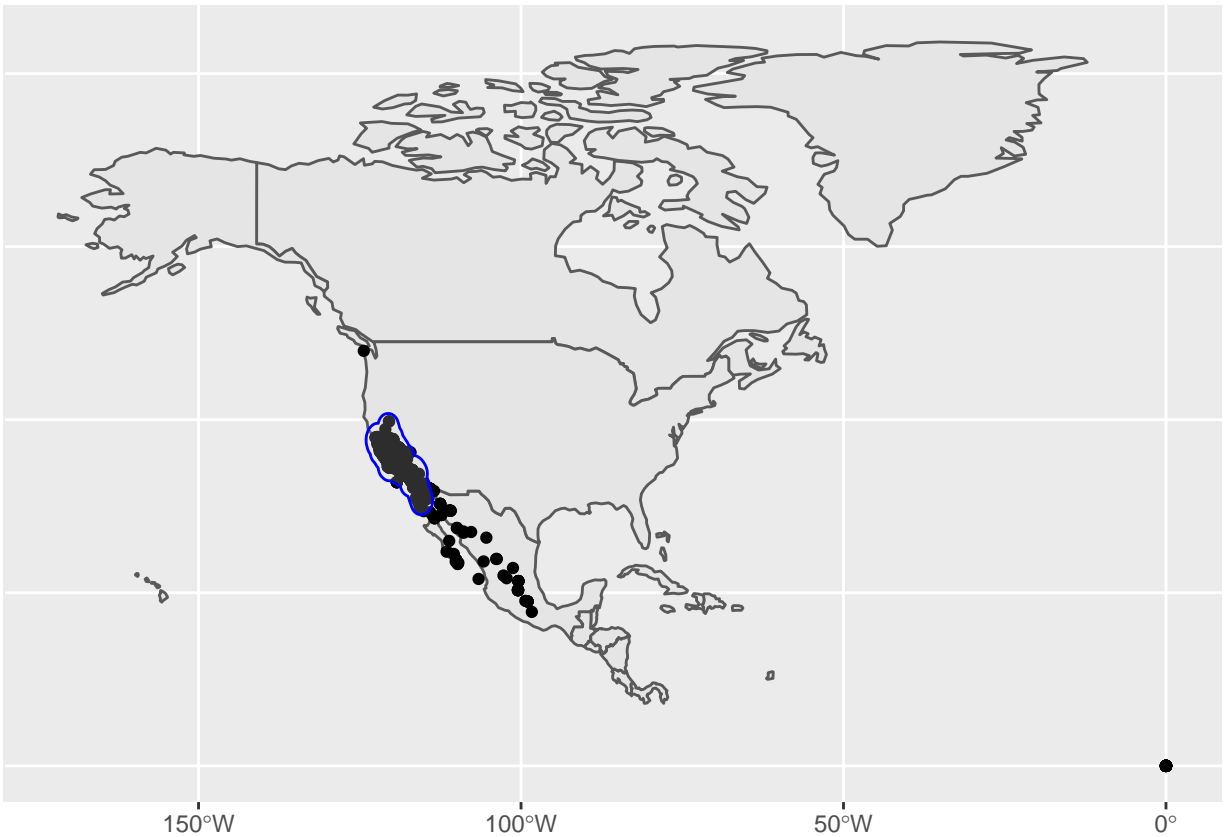
```r
# keep points contained by range map buffer
occ_df_intersect <- st_intersection(occ_sf, rm_sf_buffer)
```

```
## although coordinates are longitude/latitude, st_intersection assumes that they are planar
```

```r
ggplot() +
  geom_sf(na, mapping = aes()) +
  geom_sf(occ_df_intersect, mapping = aes()) +
  geom_sf(rm_sf_buffer, mapping = aes(), color = "blue", alpha = 0.2)
```

```
# remove points in the ocean
occ_df_intersect <- st_intersection(occ_df_intersect, na)
```

although coordinates are longitude/latitude, st_intersection assumes that they are planar

```
ggplot() +
  geom_sf(na, mapping = aes()) +
  geom_sf(occ_df_intersect, mapping = aes()) +
  geom_sf(rm_sf_buffer, mapping = aes(), color = "blue", alpha = 0.2)
```
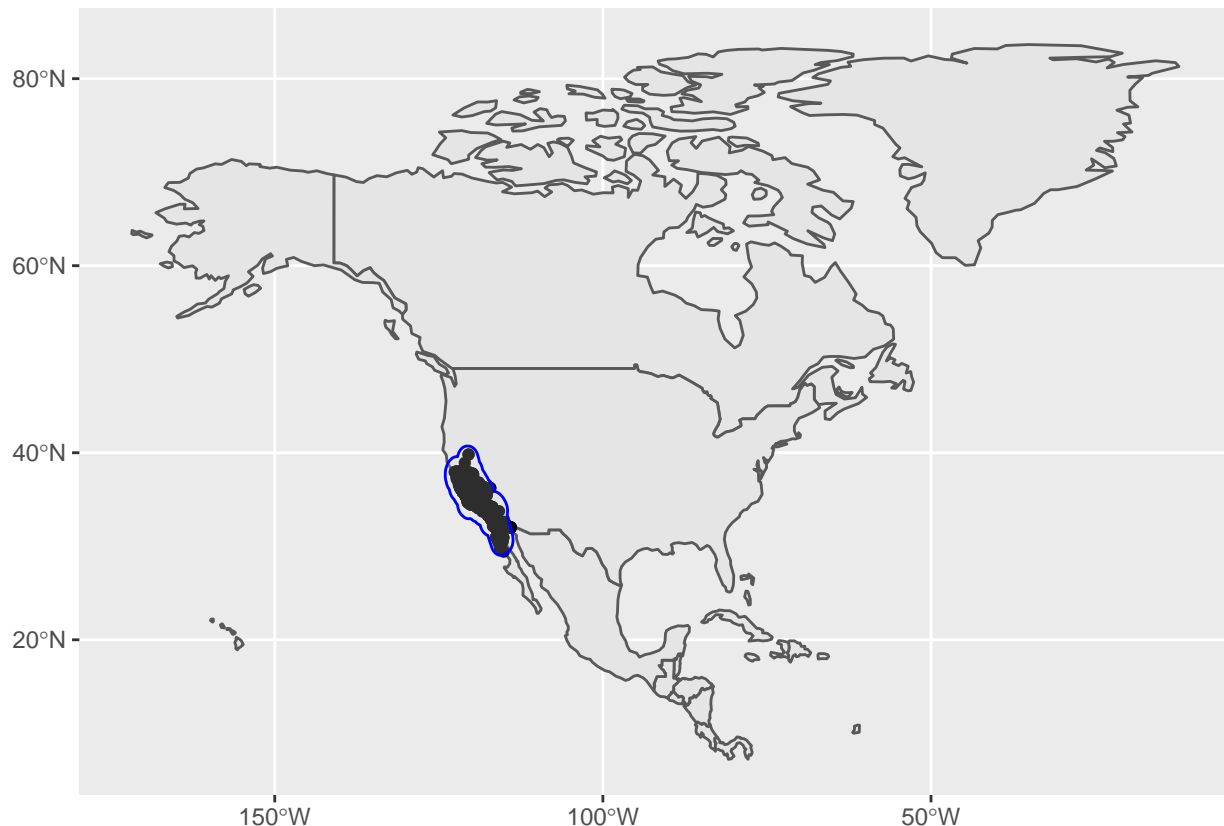
3) Decide on the accessible area of the species, using whatever criteria you think make sense based on the species biology. Write your explanation for accessible area.

Due to the small home range size of up to 56.74 ha and maximum distance moved by an individual moved during a 24-h period of 850-m suggested by trapping studies (Spiegel et al., 1995), I expect *C. californicus* to have limited mobility and therefore have restricted accessible area. I include some area still accessible to the pocket mouse, without including too big of an area outside of the know occurrences of the species. I decided on the accessible area by creating a convex hull around the occurrence points and adding a buffer of approximately 10-km to the convex hull. This buffer was then cropped to include exclude the ocean areas. This relatively small area still includes lots of climate heterogeneity as it includes different ecoregions and elevational ranges.

Reference: Spiegel, L., M. Small, J. Tom, T. Dao. 1995. An occurrance of long range movement in the California pocketmouse (*Chaetodipus californicus*). The Southwestern Naturalist, 40: 225-226.

See code below for how I decided on accessible area

```
# first I'm going to get Lat and Long coordinates out of the sf geometry
occ_df_intersect <- occ_df_intersect %>%
  mutate(decimalLongitude = st_coordinates(occ_df_intersect )[,1]) %>%
  mutate(decimalLatitude = st_coordinates(occ_df_intersect )[,2])

# create convex hull
occ_convex_hull <- chull(occ_df_intersect$decimalLongitude, occ_df_intersect$decimalLatitude)
```

```
## generate the end points of polygon.
CoordsPoly = occ_df_intersect[c(occ_convex_hull, occ_convex_hull[1]), ] %>%
  as.data.frame() %>%
  dplyr::select(decimalLongitude, decimalLatitude)# closed polygon

## convert this polygon coordinate matix to SpatialPolygon, so that buffering cound be done.
SpPoly = SpatialPolygons(list(Polygons(list(Polygon(CoordsPoly)), ID=1)))

## let's add a buffer to the convex hull. Because the mouse appears to have limited
## dispersal ability, I'm going to keep the accessible area relatively small ~10 km
Buff_SpPoly = buffer(SpPoly, width = 0.1, dissolve = F)

## See whether the convex hull is generated correctly or not
## Convert back to sf
Buff_SpPoly_sf <- st_as_sf(Buff_SpPoly)
st_crs(Buff_SpPoly_sf) <- "+proj=longlat +datum=WGS84 +no_defs"

ggplot() +
  geom_sf(na, mapping = aes()) +
  geom_sf(occ_df_intersect, mapping = aes()) +
  geom_sf(Buff_SpPoly_sf, mapping = aes(), color = "blue", alpha = 0.2) +
  coord_sf(xlim = c(-125, -110), ylim = c(29, 40))
```
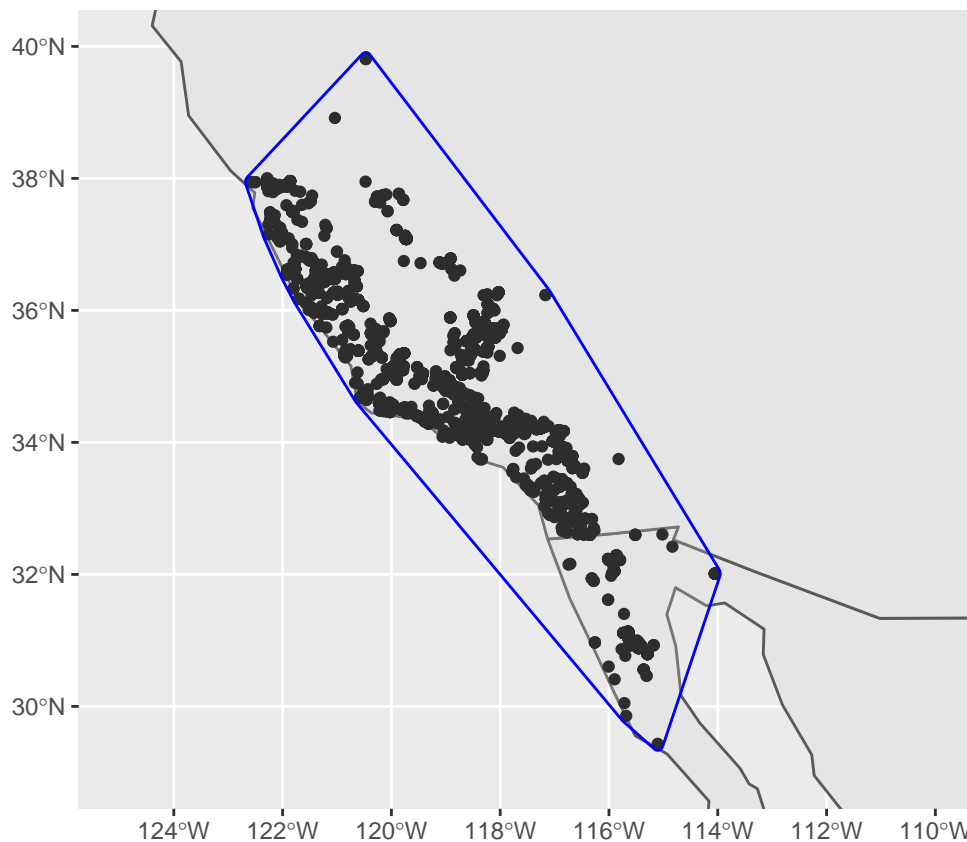


```
# crop buffer to terrestrial land
buff_crop <- st_intersection(x = Buff_SpPoly_sf, y = na)
```

```
## although coordinates are longitude/latitude, st_intersection assumes that they are planar

## Warning: attribute variables are assumed to be spatially constant throughout all
## geometries
```

```
ggplot() +
  geom_sf(na, mapping = aes()) +
  geom_sf(occ_df_intersect, mapping = aes()) +
  geom_sf(buff_crop, mapping = aes(), color = "blue", alpha = 0.2) +
  coord_sf(xlim = c(-125, -110), ylim = c(29, 40))
```
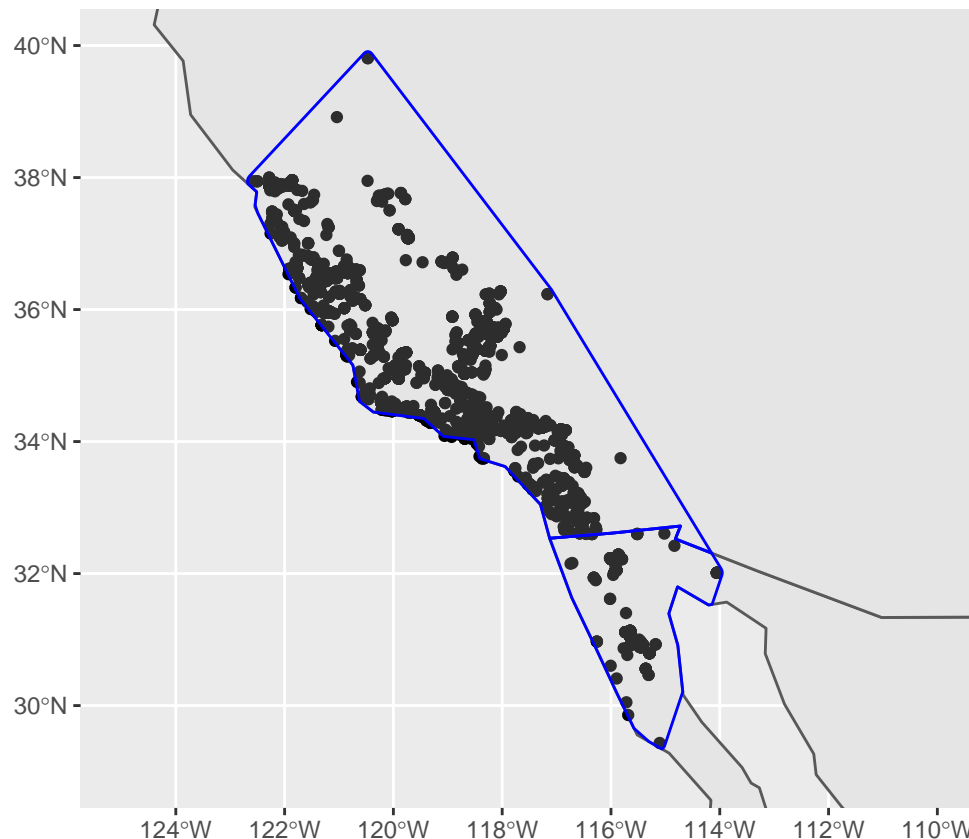

4) Download relevant layers for your model - We'd suggest a bioclimatic model for this case so likely worldclim layers

I downloaded all 19 bioclimatic layers from worldclim.

5) Explain your choice for layers in your model.

I plan to use Bioclimatic variables that are relevant to *Chaetodipus californicus. C. californicus* seems to be sensitive to both elevation and to be strongly associated with chapparral. Chaparral is found in regions with hot, dry summers, and mild, wet winters. I want the bioclim layers in my model to reflect this plant community. Additionally, I may be including layers that as of now are correlated, but I would remove correlated variables using model selection techniques before running my final ENM.

I intend to include the followign 8 variables:

1 & 2) BIO1 and BIO12 - Annual mean temperature and annual precipitation layers

These variables may broadly explain the niche of *C. californicus.*

3 & 4) BIO4 and BIO 15 - Temperature Seasonality & Precipitation Seasonality layers

These variables may be useful in describing the variance in temperature and precipitation that are representitive of the seasonality characteristic of chaparral ecosystems.

5 & 6) BIO5 and BIO6 - Max Temperature of Warmest Month and Min Temperature of Coldest month

C. californicus can initiate torpor if necessary, but avoids high elevation (>2400m), so Min temperature of Coldest month could help determine the species niche, while max temperature of warmest months could help delineate true dessert areas outside of the chaparral ecosystem.

7 & 8) BIO16 and BIO17 Precipitation of wettest quarter and precipitation of driest quarter

Using quarterly precipitation values would get at the seasonality we would expect with wet winters and dry summers. Non-chapparal ecosystems could experience different precipitation patterns than chapparal ecosystems.

6) Clip your layers to the accessible area you choose in #3.

Clip all 8 layers described above and found in the Layers directory.

```r
# read in list of files
filelist <- list.files("Layers/", full.names = TRUE)

for(i in 1:length(filelist)){
  # read in raster
  r <- raster(filelist[i])

  # change buffer from sf object to spatial
  buff_crop_sp <- as_Spatial(buff_crop)

  #crop file to extent of the buffer
  r_crop <- crop(r, buff_crop_sp)

  # mask the cropped file
  r_mask <- mask(r_crop, mask = buff_crop_sp)

  # save files
  r_names <- names(r)
  FileName = paste("Clipped_Layers", "/", r_names, ".tiff", sep = "")

  writeRaster(r_mask, FileName, "GTiff", overwrite = TRUE)

  #plot and print to check loop
  plot(r_mask)
  print(i)
}
```
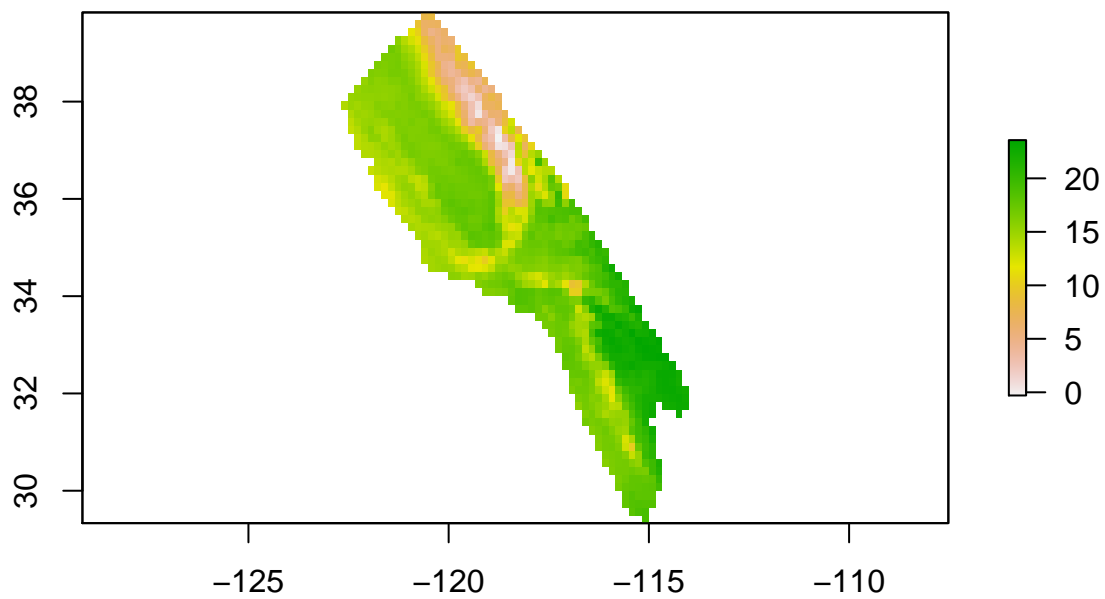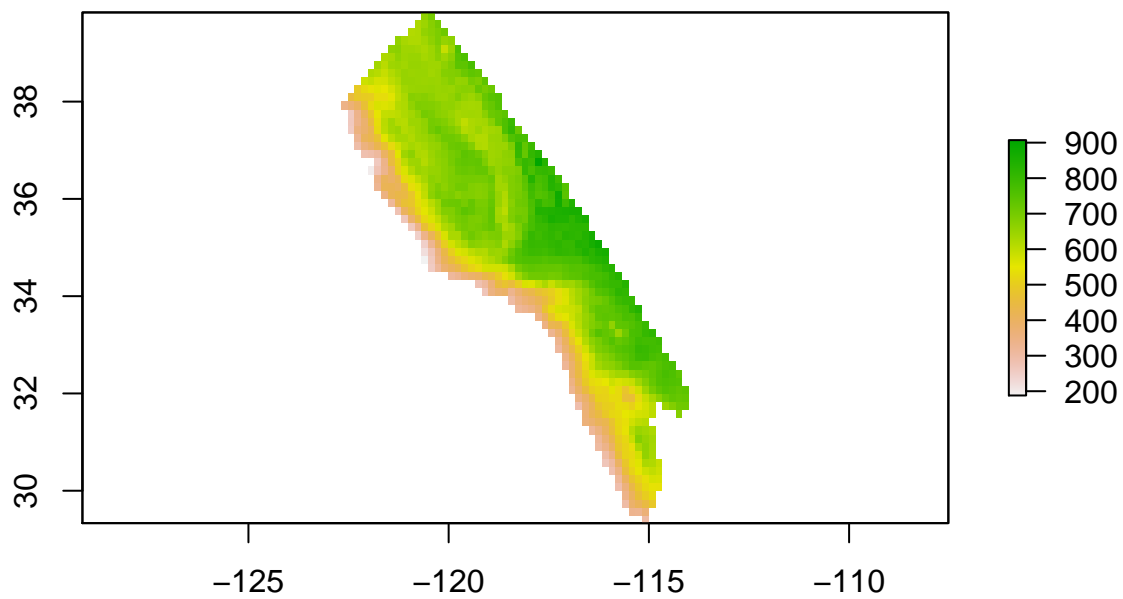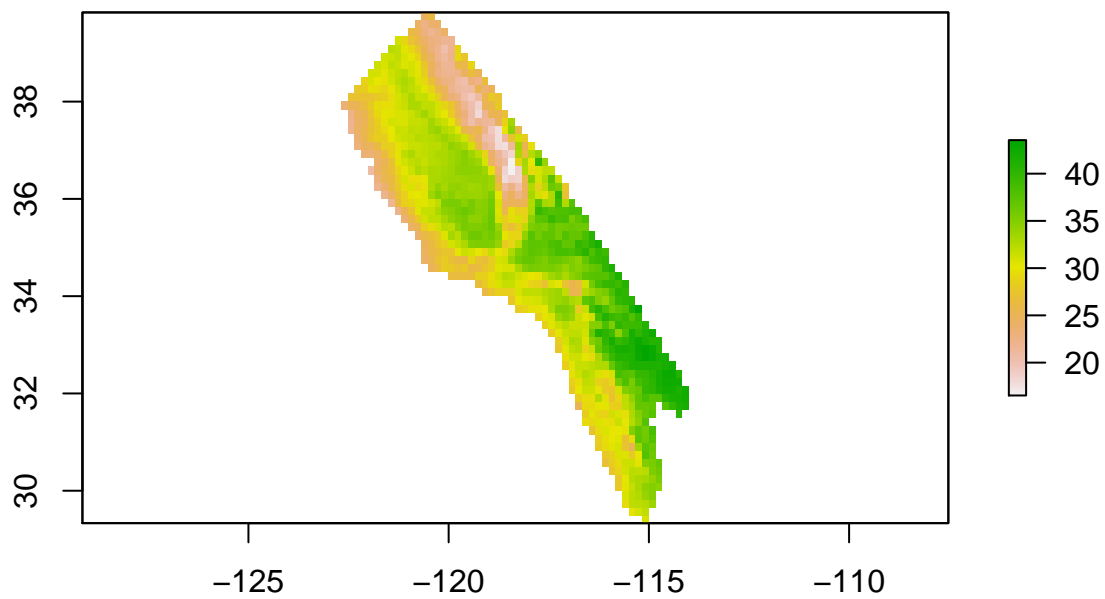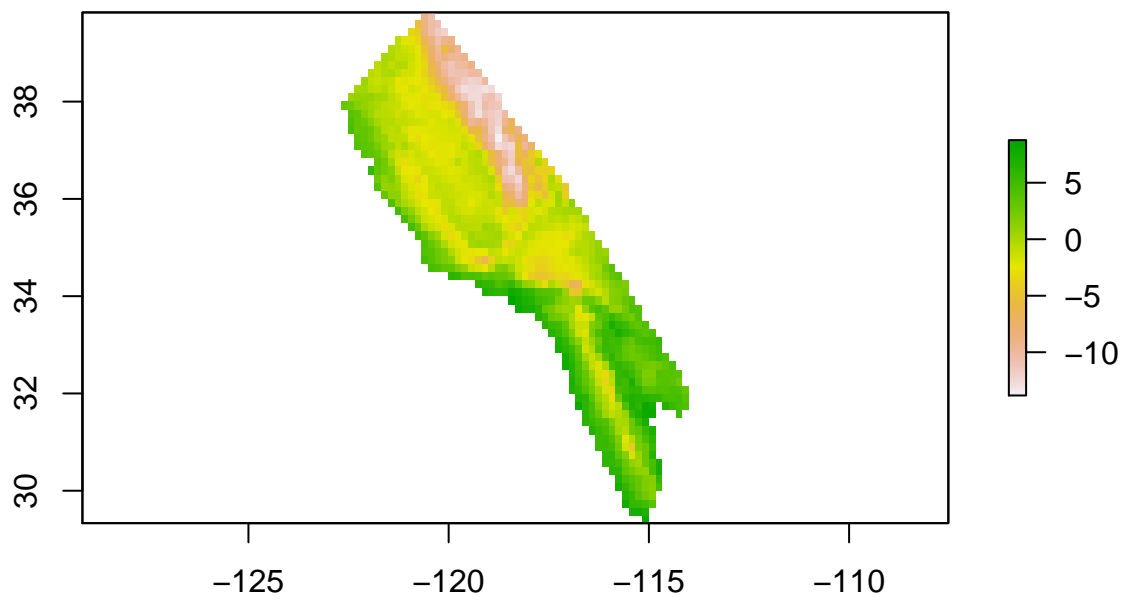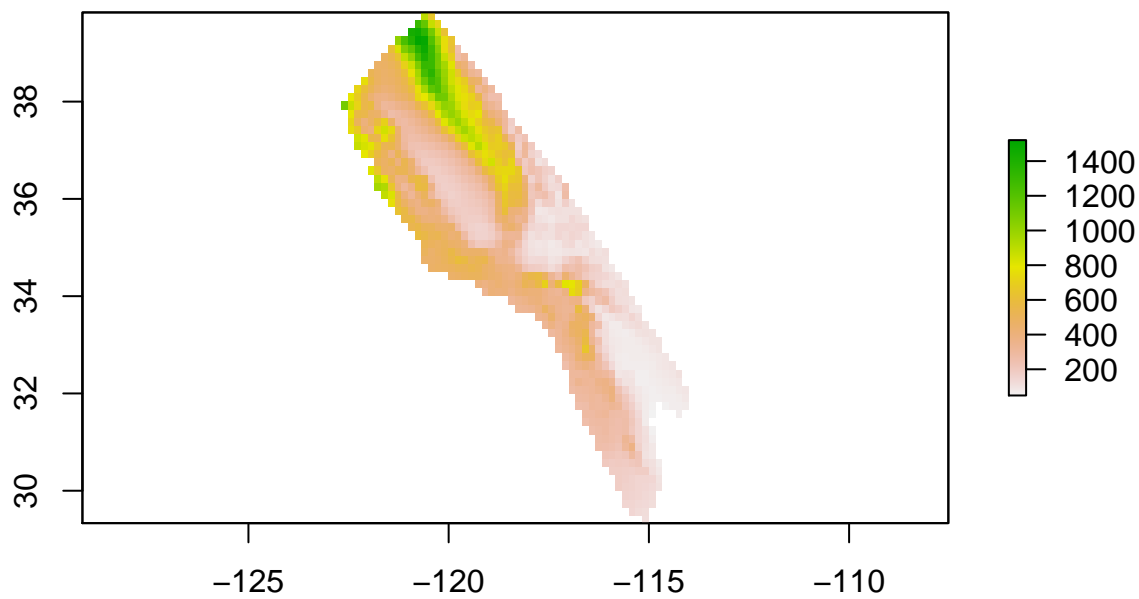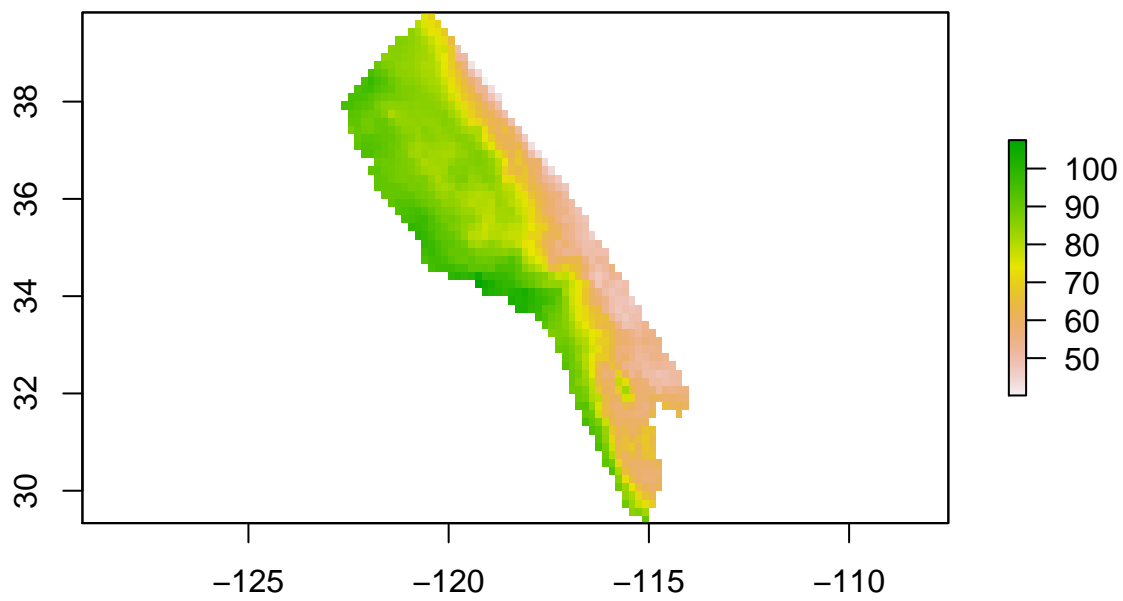
```
## [1] 1
```

[1] 2

```
## [1] 3
```

```
## [1] 4
```

```
## [1] 5
```

```
## [1] 6
```

```
## [1] 7
```

```
## [1] 8
```