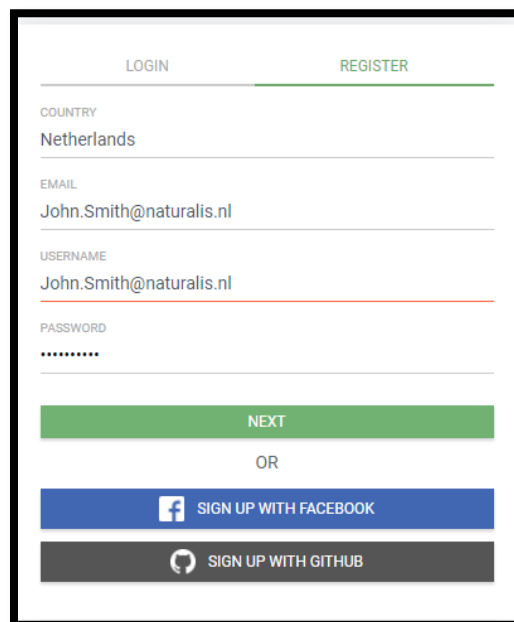


Exercise: Model your chosen species' habitat suitability under present and future climate conditions

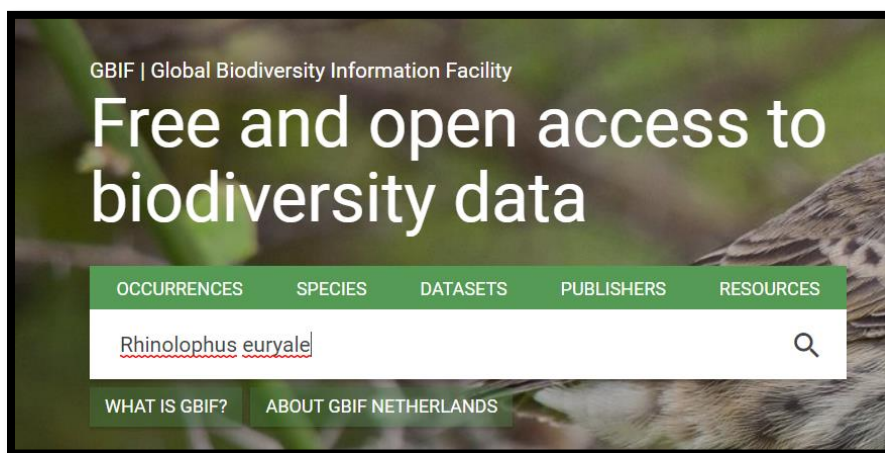
Step 1: Choose species and download occurrence data

1. Choose your species and go to the GBIF portal gbif.org.
2. To download data you have to create an account → Create an account → login as user.



The screenshot shows the GBIF registration form. At the top, there are two tabs: 'LOGIN' and 'REGISTER', with 'REGISTER' being the active tab. Below the tabs, there are four input fields: 'COUNTRY' with the value 'Netherlands', 'EMAIL' with the value 'John.Smith@naturalis.nl', 'USERNAME' with the value 'John.Smith@naturalis.nl', and 'PASSWORD' with a masked value '.....'. Below these fields is a green 'NEXT' button. Underneath the 'NEXT' button is the text 'OR'. Below 'OR' are two buttons: a blue button with the Facebook logo and the text 'SIGN UP WITH FACEBOOK', and a dark grey button with the GitHub logo and the text 'SIGN UP WITH GITHUB'.

3. Search for species of interest



4. In the example you see the data for *Rhinolophus euryale* (**mediterranean horseshoe bat**)



5. Click on occurrences to be taken to database showing all available records
6. Click Under the filters on the left hand side you can select → 'Basis of record' → 'Specimen' (if you want to download specimens only)

7. You can also filter out only results with spatial coordinates

The screenshot shows the GBIF Occurrences interface. On the left, the 'Filters' panel is visible, with the 'Location' section circled in orange. It contains three radio buttons: 'No preference', 'Including coordinates' (which is selected), and 'Without coordinates'. Below this, there is a checkbox for 'Include records where coordinates are flagged as suspicious'. At the bottom of the filters, there is a map and a 'Range' section with 'South' and 'North' input fields and a slider.

The main panel on the right displays a table of occurrences. The table has three columns: 'Scientific Name', 'Country', and 'Coordinates'. The data is as follows:

Scientific Name	Country	Coordinates
Rhinolophus euryale Blasius, 1853	Spain	39.2N, 0.7W
Rhinolophus euryale Blasius, 1853	Spain	39.3N, 0.8W
Rhinolophus euryale Blasius, 1853	Spain	38.9N, 0.3W
Rhinolophus euryale Blasius, 1853	Spain	40.0N, 0.0E
Rhinolophus euryale Blasius, 1853	Spain	39.1N, 1.2W
Rhinolophus euryale Blasius, 1853	Spain	38.9N, 0.3W
Rhinolophus euryale Blasius, 1853	Spain	39.9N, 0.4W
Rhinolophus euryale Blasius, 1853	Spain	39.3N, 0.6W
Rhinolophus euryale Blasius, 1853	Spain	38.7N, 0.5W
Rhinolophus euryale Blasius, 1853	Spain	39.3N, 0.8W
Rhinolophus euryale Blasius, 1853	Spain	39.1N, 0.6W
Rhinolophus euryale Blasius, 1853	Italy	41.9N, 13.2E
Rhinolophus euryale Blasius, 1853	Spain	39.2N, 0.7W
Rhinolophus euryale Blasius, 1853	Spain	38.9N, 0.3W
Rhinolophus euryale Blasius, 1853	Spain	39.9N, 0.4W
Rhinolophus euryale Blasius, 1853	Spain	39.7N, 0.5W
Rhinolophus euryale Blasius, 1853	Spain	40.2N, 0.2W
Rhinolophus euryale Blasius, 1853	Spain	38.7N, 0.5W
Rhinolophus euryale Blasius, 1853	Spain	38.8N, 0.1E
Rhinolophus euryale Blasius, 1853	Spain	39.1N, 0.6W

8. Download the occurrences as a csv file.
9. Once the download is complete, you will have to unzip the file.
10. Next, we will open the csv file in excel for some minor formatting.
 - a. The csv files can be a bit tricky with excel.
 - b. Open Blank Workbook.

- c. Go to **DATA** tab.
 - d. Click button **From Text** in the *General External Data* section.
 - e. Select your CSV file.
 - f. Follow the *Text Import Wizard*. (in step 2, select the delimiter of your text)
11. We must swap the Longitude and Latitude columns, as MAXENT requires the data format specifically ordered to species name, longitude, latitude.
 12. Save the file as a .csv, and it will be ready for use in MAXENT.

	A	B	C	D	E
1	species	longitude	latitude		
2	Rhinolophus euryale	13.815125	45.617693		
3	Rhinolophus euryale	16.917221	48.133888		
4	Rhinolophus euryale	9.298056	42.665001		
5	Rhinolophus euryale	15.45	41.783333		
6	Rhinolophus euryale	15.45	41.783333		
7	Rhinolophus euryale	15.45	41.783333		
8	Rhinolophus euryale	15.45	41.783333		
9	Rhinolophus euryale	15.45	41.783333		
10	Rhinolophus euryale	20.75	48.516666		
11	Rhinolophus euryale	20.75	48.516666		
12	Rhinolophus euryale	-0.866667	37.616669		
13	Rhinolophus euryale	-1.05	38.200001		
14	Rhinolophus euryale	18.51	47.706		
15	Rhinolophus euryale	2.20559	42.85152		
16	Rhinolophus euryale	2.20559	42.85152		
17	Rhinolophus euryale	2.6732	46.6346		
18	Rhinolophus euryale	2.20559	42.85152		
19	Rhinolophus euryale	-0.49727	43.06749		
20	Rhinolophus euryale	5.77462	43.35856		
21	Rhinolophus euryale	2.20559	42.85152		
22	Rhinolophus euryale	2.20559	42.85152		
23	Rhinolophus euryale	2.20559	42.85152		
24	Rhinolophus euryale	3.4345	43.75316		
25	Rhinolophus euryale	2.6732	46.6346		

13. Plot the downloaded records on a map – See course manual day 2, ARCGIS.

Step 2: Download Climate Variables

WorldClim - Global Climate Data

Free climate data for ecological modeling and GIS

Contact

WorldClim

WorldClim is a set of global climate layers (gridded climate data) with a spatial resolution of about 1 km². These data can be used for mapping and spatial modeling.

The new **Version 2.0** is now available (current climate only --- more coming soon)

The old version is **Version 1.4**.
For this version you can get data for past, current and future climates.

[Read more](#)

1. Go to worldclim.org
2. Select old version
3. Download bioclimatic variables (generic grid format) at 5 minute resolution

WorldClim 1.4: Current conditions (~1960-1990)

If you need the highest resolution (**30 arc-seconds (~1 km)**) then you can **download by tile**. See the [Methods](#) page for more info on how these data were generated, and [this page](#) for info on details about the data (such as units).

Generic grid format

variable	10 minutes	5 minutes	2.5 minutes	30 seconds
minimum temperature (°C * 10)	tmin 10m	tmin 5m	tmin 2.5m	tmin 30s
maximum temperature (°C * 10)	tmax 10m	tmax 5m	tmax 2.5m	tmax 30s
average temperature (°C * 10)	tavg 10m	tavg 5m	tavg 2.5m	tavg 30s
precipitation (mm)	prec 10m	prec 5m	prec 2.5m	prec 30s
bioclimatic variables	bio 10m	bio 5m	bio 2.5m	bio1-9, 10-19

4. Unzip bioclimes into their own folder. E.g. “C:/SDM/Bioclimes”
5. Download future bioclimes (bi) at 5 minutes resolution.
 - a. Choose between **2050** or **2070**
 - b. Choose representative concentration pathways (**RCPs**)
 - c. Choose **climate model**

2050

GCM	code	rcp26	rcp45	rcp60	rcp85
ACCESS1-0 (#)	AC		tn, tx, pr, bi		tn, tx, pr, bi
BCC-CSM1-1	BC	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi
CCSM4	CC	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi
CESM1-CAM5-1-FV2	CE		tn, tx, pr, bi		
CNRM-CM5 (#)	CN	tn, tx, pr, bi	tn, tx, pr, bi		tn, tx, pr, bi
GFDL-CM3	GF	tn, tx, pr, bi	tn, tx, pr, bi		tn, tx, pr, bi
GFDL-ESM2G	GD	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi	
GISS-E2-R	GS	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi
HadGEM2-AO	HD	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi
HadGEM2-CC	HG		tn, tx, pr, bi		tn, tx, pr, bi
HadGEM2-ES	HE	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi
INMCM4	IN		tn, tx, pr, bi		tn, tx, pr, bi
IPSL-CM5A-LR	IP	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi
MIROC-ESM-CHEM (#)	MI	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi
MIROC-ESM (#)	MR	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi
MIROC5 (#)	MC	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi
MPI-ESM-LR	MP	tn, tx, pr, bi	tn, tx, pr, bi		tn, tx, pr, bi
MRI-CGCM3	MG	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi
NorESM1-M	NO	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi	tn, tx, pr, bi

E.g. 2050

RCP 4.5

HadGEM2-AO

6. Unzip bioclimes into their own folder. E.g. “C:/SDM/Bioclimes_Future”

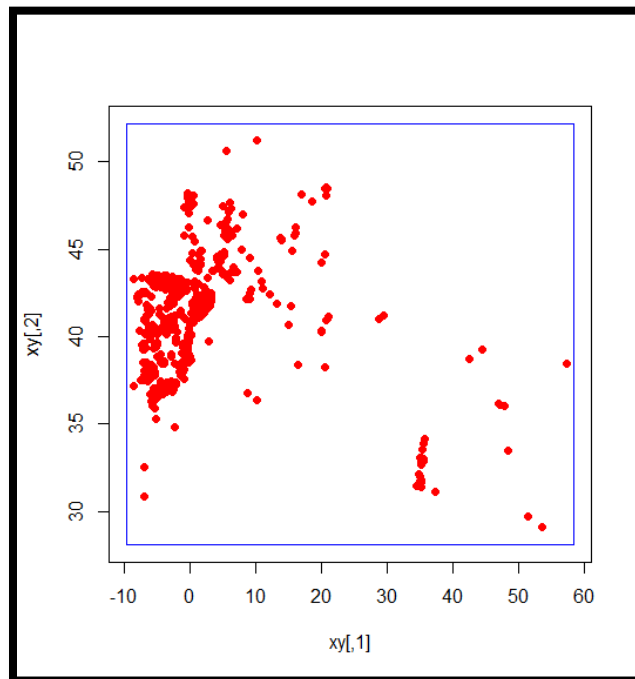
Step 3: Variable Selection and Clipping

1. In working directory create folder 'MAXENT' and inside create 'Species' and 'Results' folders
2. Copy csv of species occurrence to Species folder
3. Open rstudio and load "RScript_SDM_Workshop_VariableSelection_Clipping.R"
4. Load packages and functions
5. Load species occurrence file

```
> sp <- read.csv("G:/SDM_Course/MAXENT/Species/Rhinolophus_euryale.csv",header=T)
#load csv of occurrence
> head(sp) #check table looks correct
  species longitude latitude
1 Rhinolophus euryale 13.815125 45.61769
2 Rhinolophus euryale 16.917221 48.13389
3 Rhinolophus euryale 9.298056 42.66500
4 Rhinolophus euryale 15.450000 41.78333
5 Rhinolophus euryale 15.450000 41.78333
6 Rhinolophus euryale 15.450000 41.78333
> sp_shp <- sp #rename table
> coordinates(sp_shp) <- ~longitude+latitude #convert table to points shapefile
> proj4string(sp_shp) <- CRS("+proj=longlat +ellps=WGS84 +datum=WGS84 +no_defs")
#define projection WGS1984
```

1. Create bounding box around points

```
> bbox <- extent(sp_shp) #create bounding box of points
> bbox <- bbox+2 #increase border by 2 degrees so we do not truncate data
> plot(bbox, col='blue') #plot bounding box
> plot(sp_shp,add=T,pch=19,col='red') #add points to make sure they are enclosed
```



2. Load bioclim rasters and put them into raster stacks

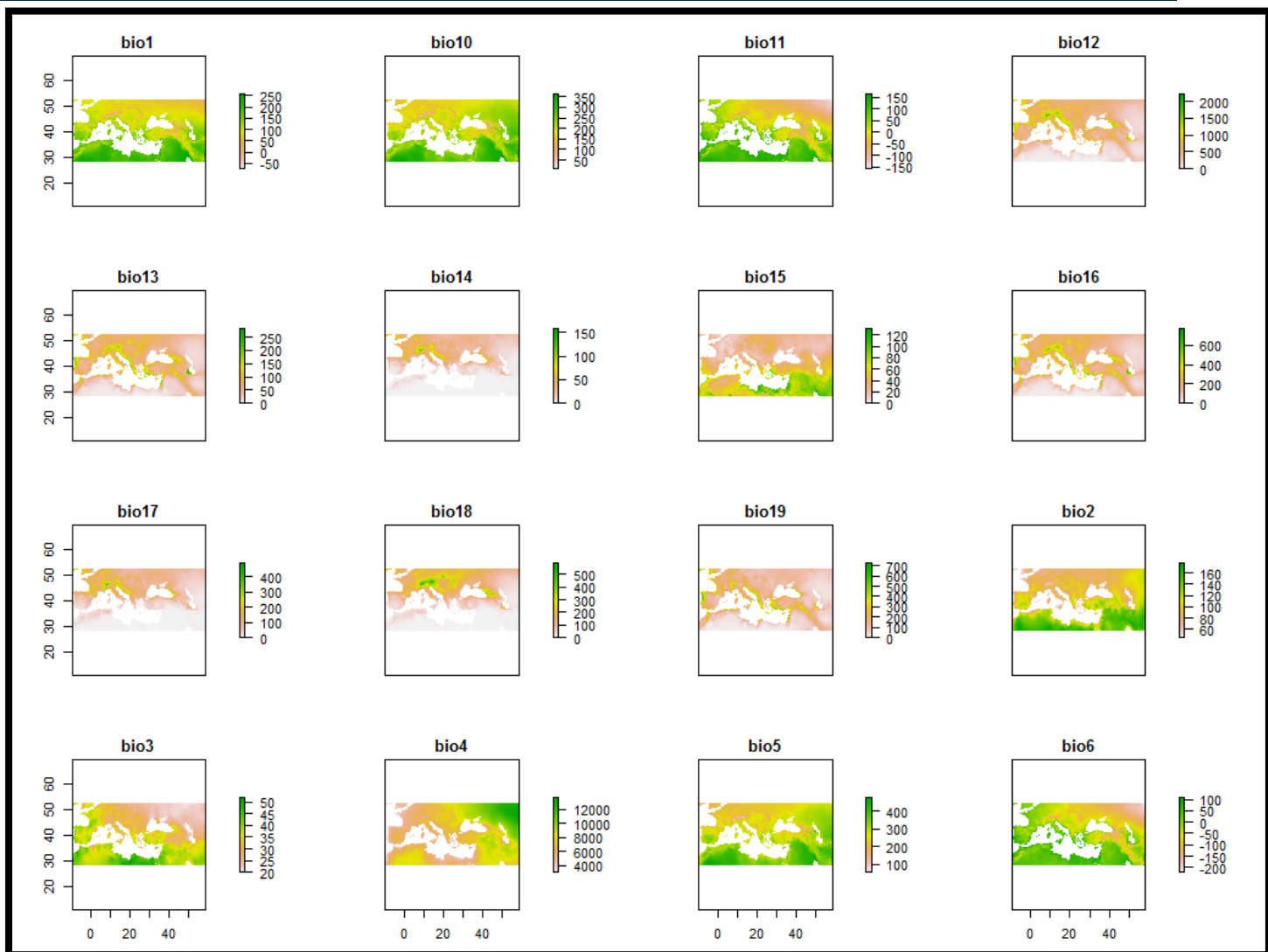
Folder from Step 2.4

Folder from Step 2.6

```
> bioclims <- stack(list.files("G:/SDM_Course/Bioclims",  
+                             pattern="*.bil$", full.names=TRUE,  
+                             ignore.case=TRUE))  
#create a stack of present bioclims  
> bioclims_fut <- stack(list.files("G:/SDM_Course/Bioclims_Future",  
+                                 pattern="*.tif$", full.names=TRUE,  
+                                 ignore.case=TRUE))  
#create a stack of future bioclims
```

3. Clip present bioclim rasters by bounding box and create new raster stack

```
> bioclims_bbox <- stack() #create empty raster stack  
> for(i in 1:19){  
+   ras <- crop(bioclims[[i]],bbox) #clip rasters by bbox  
+   bioclims_bbox <- stack(bioclims_bbox,ras) #stack rasters together  
+ } #loop through all 19 bioclims and clip each one and add to stack  
> plot(bioclims_bbox) #check extent
```



4. Variable selection with spearman's correlations

```
> bioclims_corr <- as.data.frame(bioclims_bbox, xy=T)
#convert raster stack to a table
> bioclims_corr <- na.omit(bioclims_corr[3:21])
#remove NA values (ocean) and remove coordiantes

> sc <- flattenSquareMatrix(cor.prob(bioclims_corr)) #Table summary correlations
> sc1 <- sc[with(sc, order(cor)), ]
> sc2 <- sc1[sc1$cor>=0.7|sc1$cor<=-0.7,] #High correlations
>
> sc2[with(sc2, order(i)), ] #list correlated variables
  i      j      cor p
37 bio1 bio18 -0.8096250 0
11 bio1 bio14 -0.7706305 0
29 bio1 bio17 -0.7500979 0
67 bio1 bio3  0.7257459 0
92 bio1 bio5  0.8179779 0
1  bio1 bio10 0.8602442 0
106 bio1 bio6 -0.8766742 0
154 bio1 bio9  0.9025477 0
2  bio1 bio11 0.9361634 0
38 bio10 bio18 -0.8761039 0
12 bio10 bio14 -0.8622974 0
30 bio10 bio17 -0.8583215 0
5  bio10 bio12 -0.8127567 0
```

E.g. Choose between 'Annual mean temp' or 'Min temp of coldest month'

1. Select the variables which are most representative of important climatic conditions for your species – use your intuition.

Temperature		Moisture	
Bio1	Annual Mean Temperature	Bio12	Annual Precipitation
Bio2	Mean Diurnal Range (Mean of monthly (max temp - min temp))	Bio13	Precipitation of Wettest Month
Bio3	Isothermality (BIO2/BIO7) (* 100)	Bio14	Precipitation of Driest Month
Bio4	Temperature Seasonality (standard deviation *100)	Bio15	Precipitation Seasonality (Coefficient of Variation)
Bio5	Max Temperature of Warmest Month	Bio16	Precipitation of Wettest Quarter
Bio6	Min Temperature of Coldest Month	Bio17	Precipitation of Driest Quarter
Bio7	Temperature Annual Range (BIO5-BIO6)	Bio18	Precipitation of Warmest Quarter
Bio8	Mean Temperature of Wettest Quarter	Bio19	Precipitation of Coldest Quarter
Bio9	Mean Temperature of Driest Quarter		
Bio10	Mean Temperature of Warmest Quarter		
Bio11	Mean Temperature of Coldest Quarter		

- Remove the correlated variables unwanted in the SDM and repeat code from step 3.9. Continue to remove variables until there are no correlations greater than 0.7. Approx. 5-12 variables.

```
> bioclims_corr2 <- subset(bioclims_corr, select=-c(bio18,bio19.....etc.....))
#subset out unwanted bioclims
> sc <- flattenSquareMatrix(cor.prob(bioclims_corr2)) #Table summary correlations
> sc1 <- sc[with(sc, order(cor)), ]
> sc2 <- sc1[sc1$cor>=0.7|sc1$cor<=-0.7,] #High Correlations
> sc2[with(sc2, order(i)), ] #list correlated variables
[1] i j cor p
<0 rows> (or 0-length row.names)
```

Continue until output is empty
(i.e. no correlations over 0.7)

Put names of unwanted
bioclims here

- Subset raster stacks so they only include your selected variables

```
> list.names <- c("bio1","bio10","bio11","bio12","bio13","bio14",
+               "bio15","bio16","bio17","bio18","bio19","bio2",
+               "bio3","bio4","bio5","bio6","bio7","bio8","bio9")
#list of names in same order as rasterstacks
> select.names <- names(bioclims_corr2) #list of selected rasters from step 3.11
> numbers <- match(select.names,list.names)
#numbers corresponding to chosen variables
> pres_bioclims <- bioclims[[numbers]] #select only uncorrelated variables
> fut_bioclims <- bioclims_fut[[numbers]] #select only uncorrelated variables
> clip_bioclims <- bioclims_bbox[[numbers]] #select only uncorrelated variables
```

Raster stacks of present global, future global and
present clipped for selected-uncorrelated bioclims

- Create new folders in MAXENT folder

```
> dir.create("G:/SDM_Course/MAXENT/ClimatePresent") #create folder
> dir.create("G:/SDM_Course/MAXENT/ClimateFuture") #create folder
> dir.create("G:/SDM_Course/MAXENT/ClimateTraining") #create folder
```

- Save selected bioclims in '.asc' format (required by MAXENT)

```
> for(i in 1:nlayers(pres_bioclims)){
+   ras_sc=scale(clip_bioclims[[1]],center=TRUE, scale=TRUE) #scale variables
+   writeRaster(ras_sc,
+   file=paste("G:/SDM_Course/MAXENT/ClimatePresent/",select.names[i],".asc",sep=""))
+ } #save selected present variables in MAXENT folder for use in SDM

> for(i in 1:nlayers(fut_bioclims)){
+   ras_sc=scale(fut_bioclims[[1]],center=TRUE, scale=TRUE) #scale variables
+   writeRaster(ras_sc,
+   file=paste("G:/SDM_Course/MAXENT/ClimateFuture/",select.names[i],".asc",sep=""))
+ } #save selected future variables in MAXENT folder for use in SDM

> for(i in 1:nlayers(clip_bioclims)){
+   ras_sc=scale(clip_bioclims[[1]],center=TRUE, scale=TRUE) #scale variables
+   writeRaster(ras_sc,
+   file=paste("G:/SDM_Course/MAXENT/ClimateTraining/",select.names[i],".asc",sep=""))
+ } #save selected clipped present variables in MAXENT folder for use in SDM
```

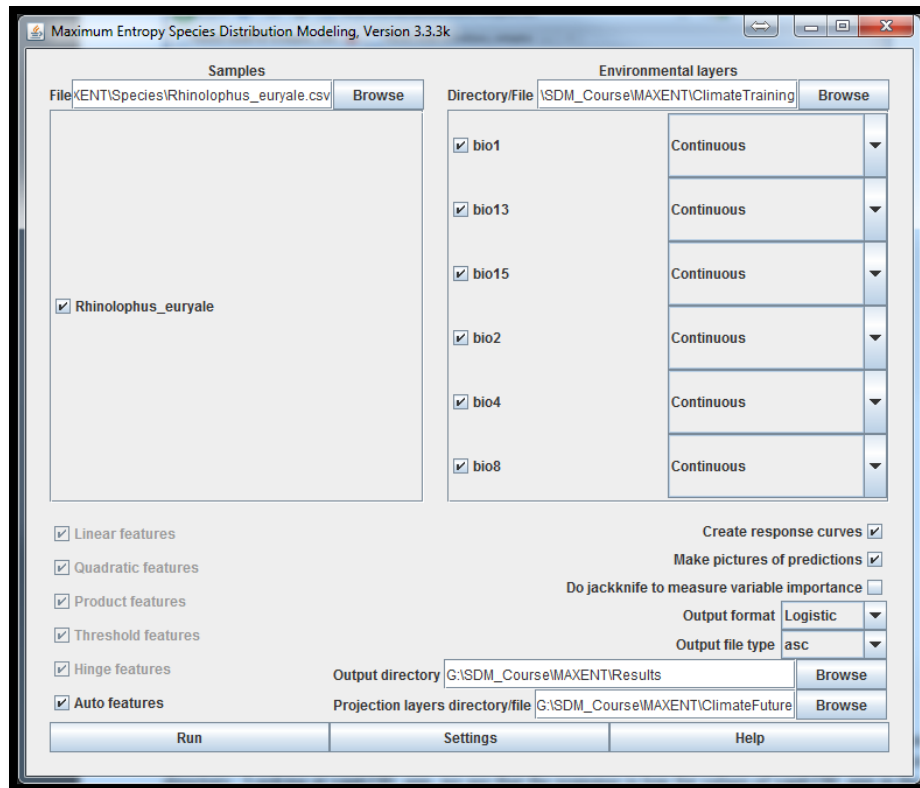
Step 4: Species Distribution Models (SDMs)

Species Distribution Models (SDMs), also known as Ecological Niche Models (ENMs), predict the presence and absence of species by interpolating identified relationships between collection data, stored in Natural History Museums and Herbaria, and environmental data. In this section we identify the relationships between species presence records and environmental data with a distribution modelling application, and subsequently interpolate the relationships to the research area of interest; here Borneo. We start with downloading the MaxEnt application from <http://www.cs.princeton.edu/~schapire/maxent/>. You can also download the MaxEnt tutorial from this website. Besides MaxEnt there are many other algorithms and applications such as GARP (<http://nhm.ku.edu/desktopgarp/>), BioMapper (<http://www2.unil.ch/biomapper/>), and Generalized Dissimilarity Modelling (<http://www.biomaps.net.au/gdm/>), amongst others! Most modelling algorithms are also available in R (R Development Core Team 2014). MaxEnt is a Java application, so you need Java to be installed on your computer (<http://www.java.com>). You open MaxEnt by clicking the Maxent.bat file.

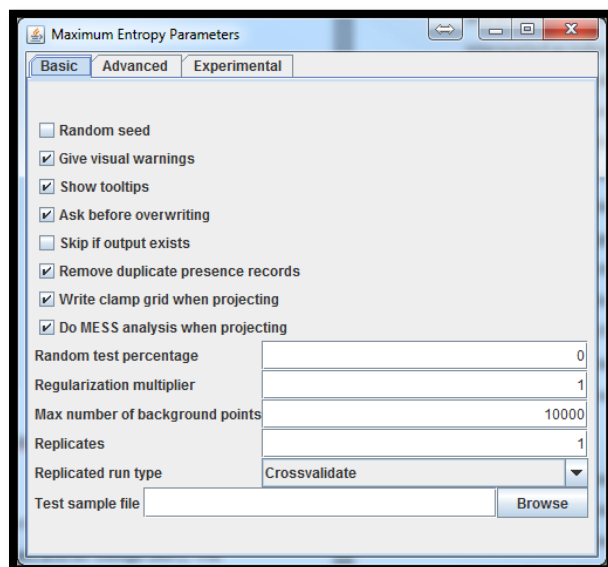
Maxent uses the maximum entropy algorithm which is defined as follows:

MaxEnt, or the maximum entropy method for species' distribution modelling, estimates the most uniform distribution ("maximum entropy") across the study area, given the constraint that the expected value of each environmental predictor variable under this estimated distribution matches its empirical average (average values for the set of species' presence records) (Phillips et al. 2006).

1. We start with downloading the MaxEnt application from <http://www.cs.princeton.edu/~schapire/maxent/>
2. Open MaxEnt software
3. Use the Browse buttons to select the 'Samples File', the 'Environmental layers Directory File', and the 'Output directory'
 - a. Make sure that in the projections layers you include both 'ClimatePresent' and 'ClimateFuture'. We want to see the projected distribution at the global scale for both the present and the future. Separate the two directories with a comma, e.g.
'G:\SDM_Course\MAXENT\ClimatePresent,G:\SDM_Course\MAXENT\ClimateFuture'
 - b. Make sure create response curves is selected



4. Click the button 'Settings' and check the option 'Remove duplicate presence records', set the 'Random test percentage' to zero, and set the 'Max number of background points' to 10,000. Make sure there are more background points than presence records.



5. Click the button 'Run' and create a Maxent model for your species.

Validating SDMs

The most widely applied method to validate SDMs is the Area Under the Curve (AUC) of the Receiver Operator Curve (ROC) (Fielding and Bell 1997, McPherson et al. 2004, Raes and ter Steege 2007). The advantage of the AUC value over other measures of model accuracy (i.e. Cohen's kappa, sensitivity, specificity) is that it is a) threshold independent, and b) prevalence insensitive.

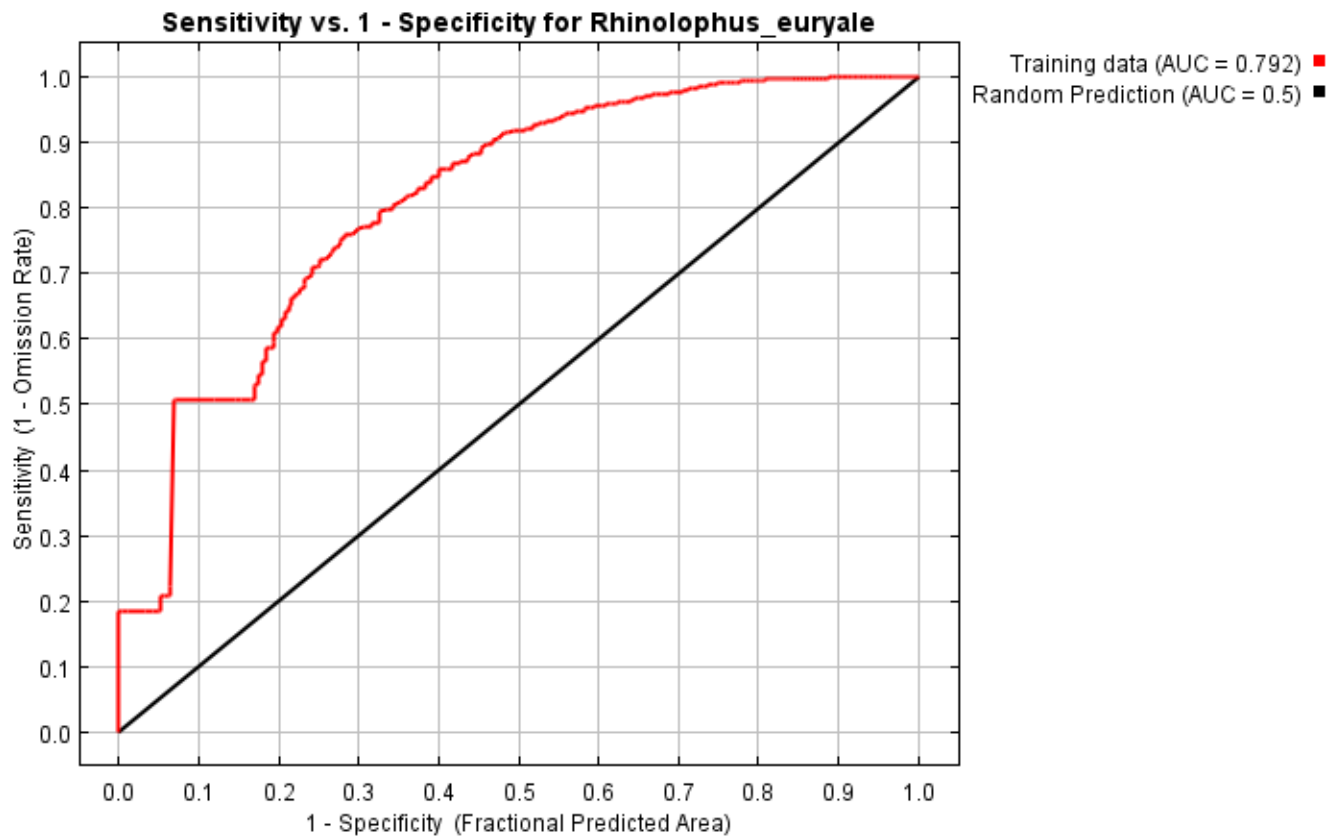
Setting a threshold means that continuous MaxEnt values, running from 0-1, do not have to be converted to discrete presence/absence values. There are several techniques to set thresholds (Liu et al. 2005), but this is not required for the AUC value. Prevalence is the proportion of the data representing species' presence, or presences / (presences + absences). The fact that the AUC value is relatively insensitive to prevalence is of special relevance because when absences are lacking, which is often the case, they are replaced by pseudo-absences, or background points. A sufficiently large sample of pseudo-absences is needed to provide a reasonable representation of the environmental variation exhibited by the geographical area of interest, typically 1,000-10,000 points. These large numbers of pseudo-absences automatically result in low prevalence values. The number of records by which a species is represented in herbaria and natural history museums range from 1 to 150-200 records. Even when a species is represented by 200 unique presence-only records and 1000 pseudo-absences are used, prevalence is only 16.7% (200/1200).

AUC values range from 0 to 1, with a value of 0.5 indicating model accuracy not better than random, and a value of 1.0 indicating perfect model fit (Fielding and Bell 1997). An AUC value can be interpreted as indicating the probability that, when a presence site (site where a species is recorded as present) and an absence site (site where a species is recorded as absent) are drawn at random from the population, the presence site has a higher predicted value than the absence site (Phillips et al. 2006). SDMs with an AUC value of 0.7 are considered to be reliable, values over 0.8 as good.

A major drawback of using pseudo-absences instead of true absences, however, is that the maximum achievable AUC value indicating perfect model fit, is no longer 1, but $1-a/2$ (where a is the fraction of the geographical area of interest covered by a species' true distribution, which typically is not known). Nevertheless, random prediction still corresponds to an AUC value of 0.5. Therefore, standard thresholds of AUC values indicating SDM accuracy (e.g. the threshold of $AUC > 0.7$ that is often used), do NOT apply (Raes and ter Steege 2007). Therefore be very cautious when SDMs based on presence-only data are validated with AUC values. This problem can be solved by testing against a null-model. This procedure is described in detail in Raes and ter Steege 2007, but this goes beyond this practical.

6. Open the your_species.html – file in the folder 'MAXENT/Results'.

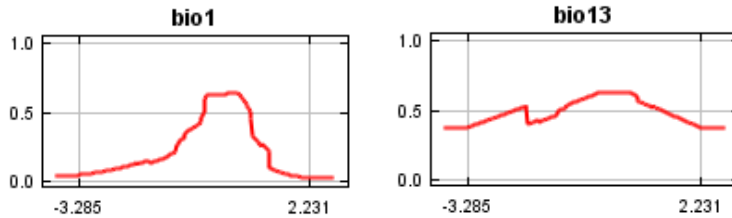
7. Check the AUC value



8. Go to the threshold table in the your_species.html file. This table contains a list of thresholds. Among the most widely used ones are: '10 percentile training presence', 'Equal training sensitivity and specificity', and 'Maximum training sensitivity plus specificity'.

Cumulative threshold	Logistic threshold	Description	Fractional predicted area	Training omission rate
1.000	0.065	Fixed cumulative value 1	0.789	0.007
5.000	0.163	Fixed cumulative value 5	0.595	0.045
10.000	0.262	Fixed cumulative value 10	0.483	0.086
0.182	0.020	Minimum training presence	0.906	0.000
11.015	0.272	10 percentile training presence	0.468	0.099
29.042	0.434	Equal training sensitivity and specificity	0.266	0.267
27.072	0.387	Maximum training sensitivity plus specificity	0.282	0.242
1.618	0.096	Balance training omission, predicted area and threshold value	0.748	0.010
6.158	0.173	Equate entropy of thresholded and original distributions	0.571	0.055

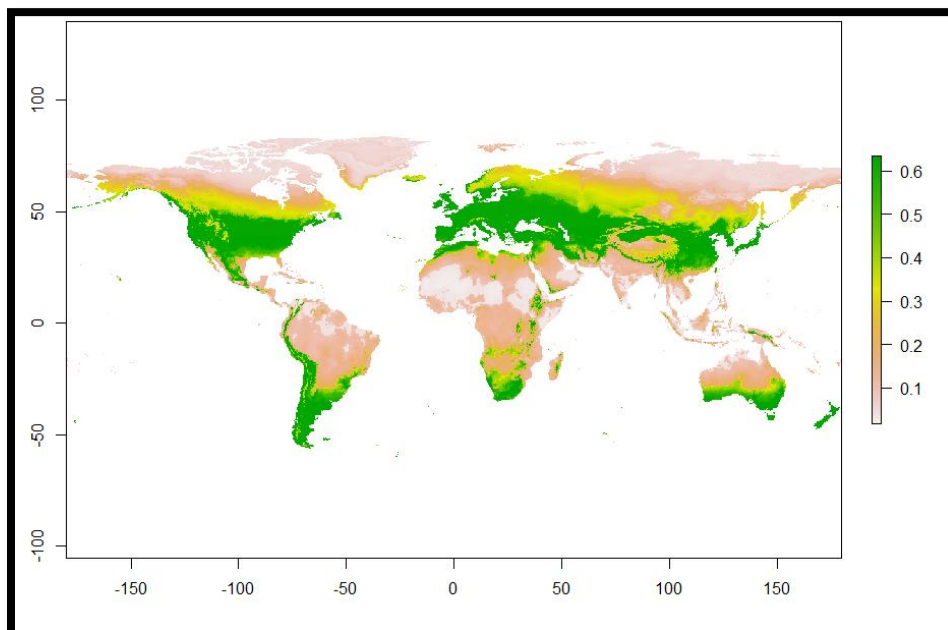
9. Check the contribution of the different variables in the model and the response curves. The variables are scaled so do not have a clear ecological meaning, and you can only see the general broad patterns, we should return the variables to their original values but this is not necessary for this workshop.



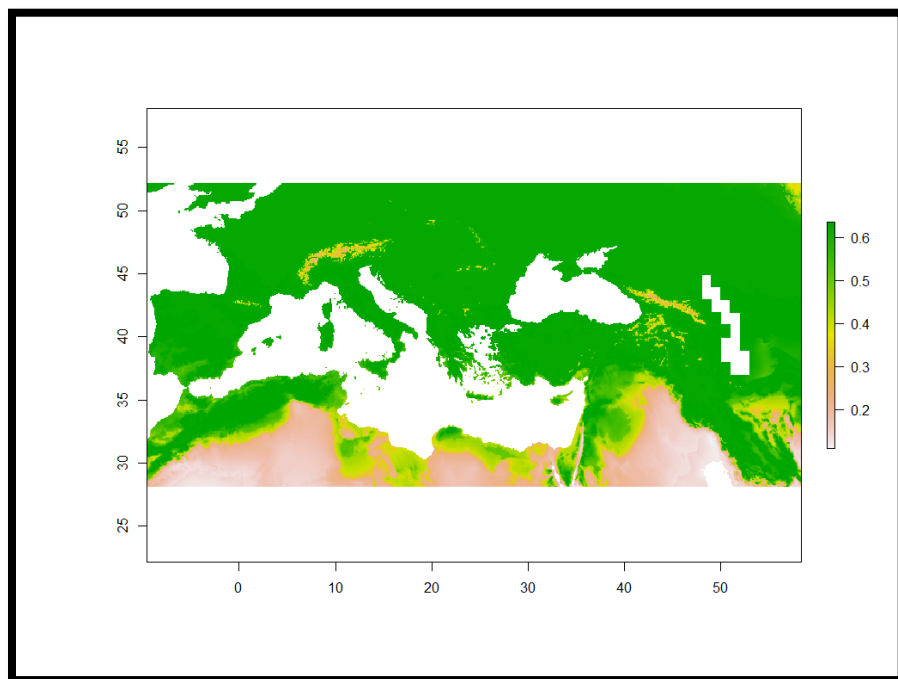
Variable	Percent contribution	Permutation importance
bio1	88.7	94.6
bio8	6.1	0
bio2	4.3	0
bio13	0.9	4.2
bio15	0	0.9
bio4	0	0.4

10. Load, plot and clip projection rasters for present.

```
> pres<-raster("G:/.../MAXENT/Results/Rhinolophus_euryale_ClimatePresent.asc")  
# load present global distribution  
> plot(pres) #visualize
```



```
> pres_clip <- crop(pres,bbox) #clip to training area
> plot(pres_clip) #visualize
```



We can see from the global figure and the clip that when only looking at climate variables there is much greater suitable habitat when compared to the actual occurrence of the species. Why do you think this is the case? What other factors have we not modelled that mean the distribution of the species is more restricted?

11. Load, plot and clip projection rasters for future.

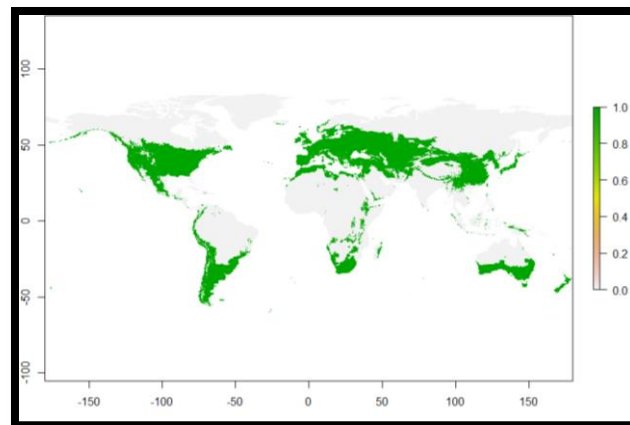
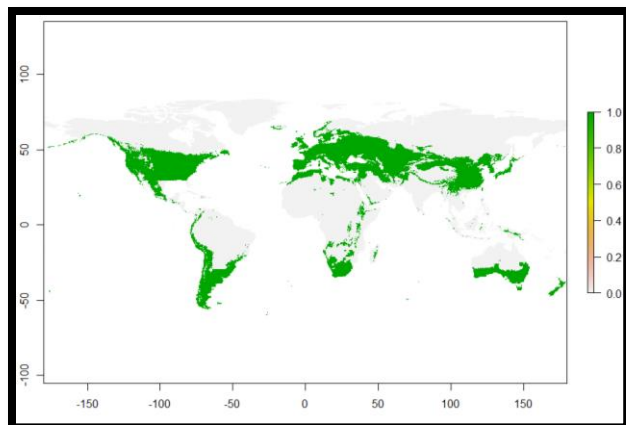
```
> fut <- raster("G:/.../MAXENT/Results/Rhinolophus_euryale_ClimateFuture.asc")
# load future global distribution
> plot(fut) #visualize (not shown)
> fut_clip <- crop(fut,bbox) #clip to training area
> plot(fut_clip) #visualize (not shown)
```

12. Convert habitat suitability maps to thresholded binary maps (0,1) using the threshold 'Maximum training sensitivity plus specificity'. Step 4.8.

```
> th <- 0.387 #define threshold
> m <- c(0, th,0, th, 1, 1)
#matrix everything before 'th' as 0 and everything after 'th' as 1
> bin_mat <- matrix(m, ncol=3, byrow=TRUE) #convert to correct matrix
> pres_bin <- reclassify(pres, bin_mat) #reclassify by matrix present
> fut_bin <- reclassify(fut, bin_mat) #reclassify by matrix future
```



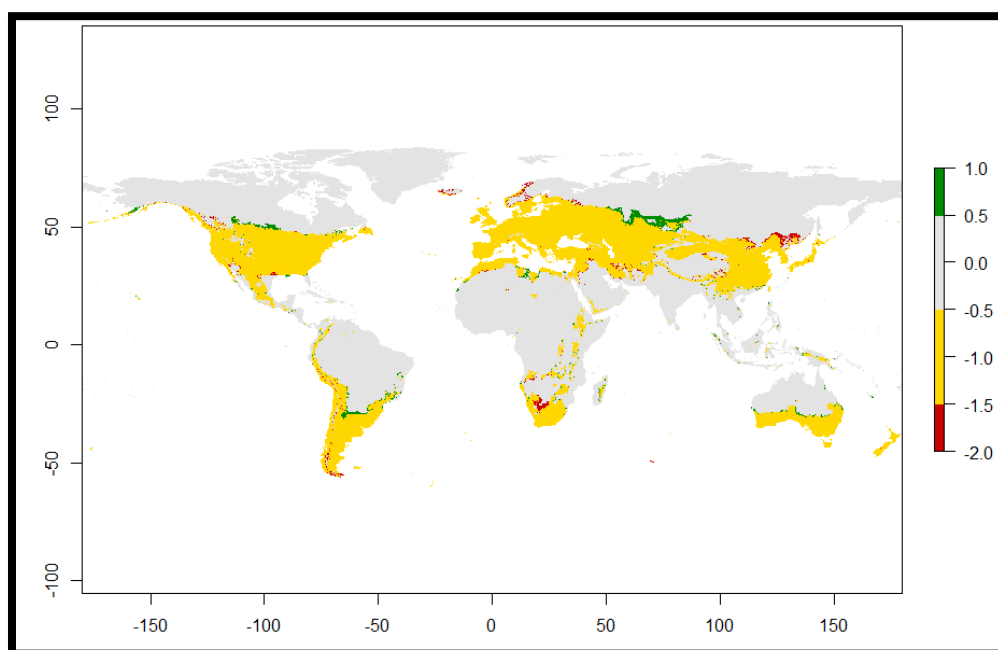
```
> plot(pres_bin)
> plot(fut_bin)
```



Very little difference between present and future distribution – suggests that under scenario RCP 4.5 for 2050 not a strong effect of climate change – the distribution of *Rhinolophus euryale* is more affected by other factors.

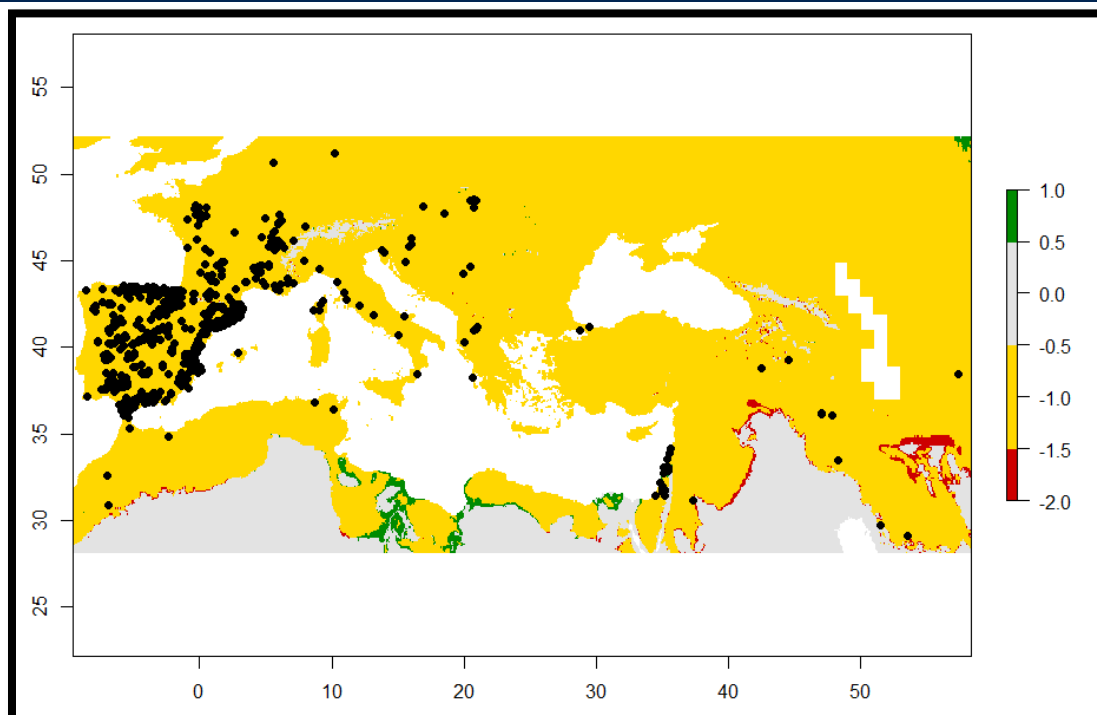
- Calculate differences between present and future in habitat suitability at global scale. Grey is never suitable, yellow is remains suitable, red is lost, and green is gained.

```
> range_change<-BIOMOD_RangeSize(CurrentPred=pres_bin,
+                               FutureProj=fut_bin,SpChange.Save=NULL)
#calculate range change between two maps
> col.lst <- c("red3", "gold", "grey89", "green4") #define plot colours
> plot(range_change$Diff.By.Pixel,col=col.lst) #plot range change
```



14. Calculate differences between present and future in habitat suitability at occurrence scale. Grey is never suitable, yellow is remains suitable, red is lost, and green is gained.

```
> range_change_clip<-BIOMOD_RangeSize(CurrentPred=pres_bin_clip,
+                                     FutureProj=fut_bin_clip,SpChange.Save=NULL)
#calculate range change between two maps
> col.lst <- c("red3", "gold", "grey89", "green4") #define plot colours
> plot(range_change_clip$Diff.By.Pixel,col=col.lst) #plot range change
> plot(sp_shp,add=T, col="black", pch=19) #overlay points
```



As we expected the climatic envelope of the mediterranean horseshoe bat is much wider than the actual occurrence and therefore climate change does not appear to be the greatest threat. There are other factors as to why this species is threatened and on the redlist (see below).

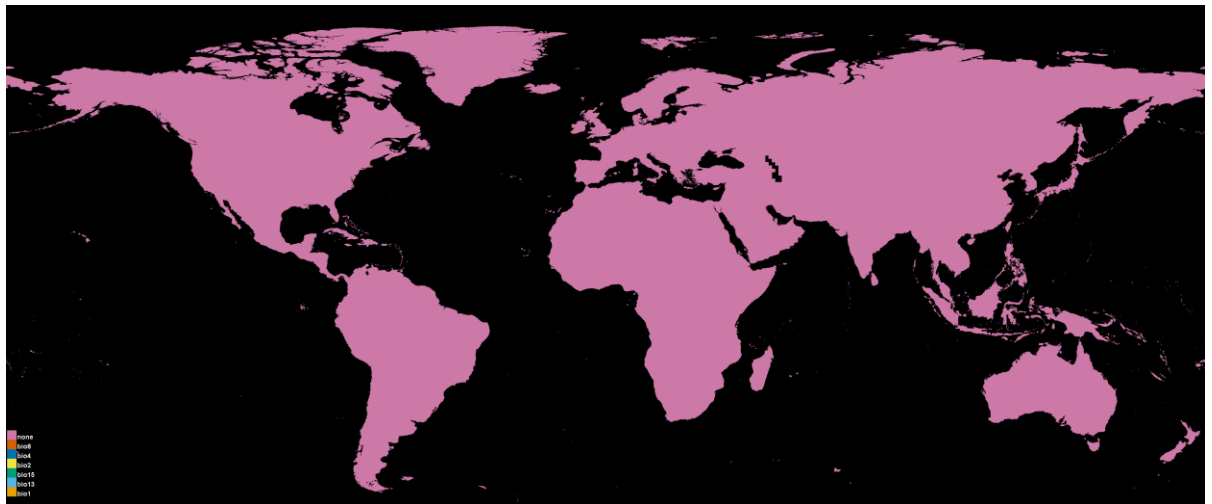
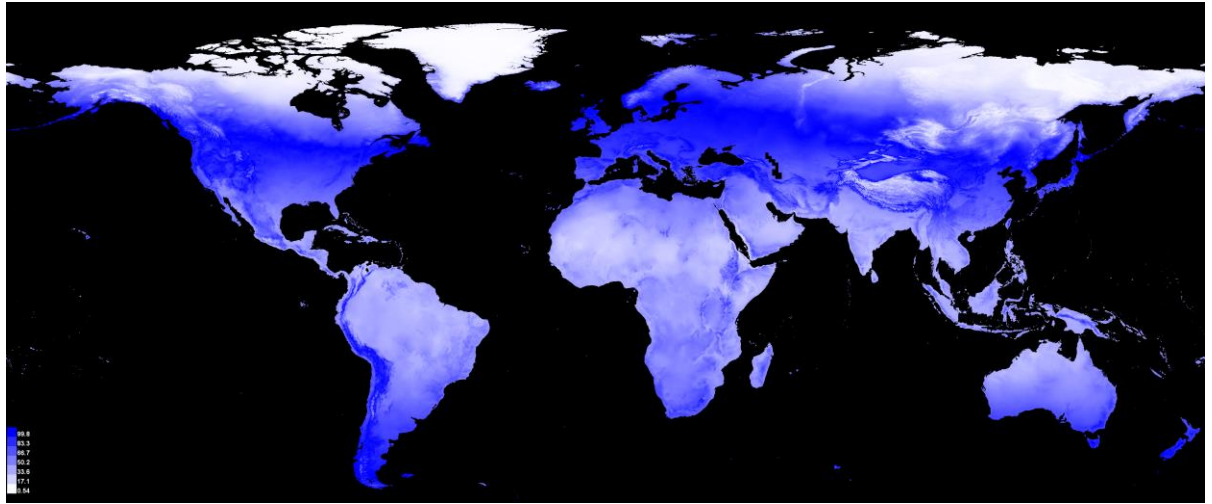
Threats [top]

Major Threat(s):

Threats include loss of foraging habitat, and disturbance and loss of underground habitats. On a landscape scale, fragmentation and loss of linear elements such as hedgerows and riparian vegetation is a problem because such elements are used as landscape references for commuting. Foraging habitats are lost due to intensive agriculture, urbanization and large infrastructures. The species' strong dependence upon caves for roosting makes it particularly sensitive to cave disturbance, such as that from caving or tourism. Tourist disturbance of caves affects the species in a number of range states. The use of organochlorine pesticides is believed to have contributed to the earlier dramatic decline of the species in France (Brosset *et al.* 1988). In North Africa, threats include habitat loss due to agriculture (livestock) and human disturbance.

15. Are you predicting into unknown conditions?

- a. The following two pictures compare the environmental similarity of variables in ClimateFuture to the environmental data used for training the model. In the first picture (MESS), areas in red have one or more environmental variables outside the range present in the training data, so predictions in those areas should be treated with strong caution. The second picture (MoD) shows the most dissimilar variable, i.e., the one that is furthest outside its training range. For details, see Elith et al., *Methods in Ecology and Evolution*, 2010)



Step 4: Report

- A 2-3 page report on a selected species
- Introduction – about the species and its present distribution (from GBIF data)
- Methodology
 - Which settings on MAXENT?
 - Variable selection
 - Which variables and why?
- Model Output
 - Present + Future Distribution Map
 - Model performance – AUC (Good model or not?)
 - Variable Importance Table (What is driving the distribution patterns?)
- Response to future scenario
 - Future Distribution Change Map
- Biological interpretation
 - How is the distribution expected to change?
 - What could this mean for the biology of the species?
 - Is your model useful? Be critical!
 - What are the limitations?
 - What information is missing?