

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE GEOGRAFÍA E HISTORIA



**MÁSTER EN TECNOLOGÍAS DE LA INFORMACIÓN
GEOGRÁFICA**

**TRABAJO FIN DE MÁSTER
CURSO 2015-2016**

**DISTRIBUCIÓN ACTUAL Y POTENCIAL DE
ESPECIES LEÑOSAS ALIMENTICIAS EN
ÁFRICA PARA LA FORMULACIÓN DE
ESTRATEGIAS DE CONSERVACIÓN**

José Manuel Macías Barredo

Septiembre 2016

Tutor/a: María Dolores Agúndez Leal

Co- tutor/a: María Pilar García Rodríguez

Resumen: Los modelos de nicho ecológico tienen un amplio rango de aplicaciones biológicas y constituyen una de las herramientas más útiles para respaldar la toma de decisiones de conservación de especies. En este estudio se ha generado un modelo de distribución para la actualidad y otro para evaluar el impacto del cambio climático de cuatro especies vegetales elegidas por su capacidad de proveer alimento incluso en regiones con factores climáticos extremos; *Adansonia digitata*, *Balanites aegyptiaca*, *Boscia senegalensis* y *Ziziphus mauritina*. Para construir estos modelos se necesitan dos fuentes de información; ocurrencia de las especies y variables climáticas. Estas variables deben de carecer de correlación y colinearidad y poseer la mayor devianza explicada posible. La modelación del hábitat en la actualidad de las especies se generó con Maxent, un software basado en la aproximación de la máxima entropía que toma como *inputs* las presencias de las plantas, las variables climáticas elegidas de cada especie y el número de puntos de control adecuado a la extensión del área de estudio. La evaluación de los resultados demuestra la eficacia del proceso, el estadístico AUC no bajo de 0,926, lo que indica una buena discriminación entre presencia/ausencia y la ganancia de entrenamiento regularizada mínima fue de 1,358, cuatro veces más que en los puntos de ausencia, demostrando una aproximación correcta al clúster de presencias.

Palabras clave: Modelado de distribución de especies, África, Maxent, Cambio Climático.

Abstract: Ecological niche models have a wide range of biological applications and they are useful tools to support conservation decision-making. In this study was created a current distribution model and a climate change assessment model of four vegetal species valued for their ability to provide food even in extreme environment conditions; *Adansonia digitata*, *Balanites aegyptiaca*, *Boscia senegalensis* y *Ziziphus mauritiana*. In order to build these models, two resources are needed; presence records and climatic variables. These variables must lack of correlation and colinearity and have the greatest deviance as possible. The species current habitat model was generated with Maxent, a software based on the maximum-entropy approach which takes as inputs plants occurrence, bias free variables and the suitable number of background points for the study area extent. The results assessment prove the model efficiency, AUC minimum value was 0.926 which indicates a great classification between presence/absence and the smallest regularized training gain was 1.358, four times bigger than absence's, showing a suitable approach to presence cluster.

Keywords: Species Distribution Modelling, Africa, Maxent, Climate Change.

1. INTRODUCCIÓN	4
2. MATERIALES	5
2.1. Área de estudio	5
2.2. Fuentes de datos de presencia y ráster climáticos	6
3. METODOLOGÍA	9
3.1. Depurado de datos de presencia y ráster climáticos	10
3.2. Modelación y análisis de distribución de especies	13
<u>3.2.1. Preparación de datos de entrada</u>	13
<u>3.2.2. Creación del modelo de nicho ecológico</u>	14
<u>3.2.3. Impacto del cambio climático en la distribución de especies</u>	18
<u>3.2.4. Ejecución del modelo en Maxent.</u>	22
<u>3.2.5. Representación gráfica de las modelaciones</u>	24
4. RESULTADOS	26
4.1. Modelación de la distribución de las especies	27
4.2. Modelación del impacto del cambio climático en la distribución de especies	31
5. CONCLUSIONES	36
6. BIBLIOGRAFÍA	37
ANEXO 1. PROCESO ESTADÍSTICO DE CALIBRACION DEL MODELO	43

1. INTRODUCCIÓN.

Definir áreas prioritarias es el objetivo de la conservación de la biodiversidad. Los modelos de nicho ecológico pueden mejorar la toma de decisiones en cuanto la gestión de la conservación de especies (Gomes y Brito, 2015).

Adansonia digitata, *Balanites aegyptiaca*, *Boscia senegalensis* y *Ziziphus mauritiana* son especies vegetales con ocurrencia en el continente africano, su selección para la creación de los modelos se determina por ser una fuente de alimento estable y segura, ya que pueden desarrollarse en ambientes diversos (Assogbadjo y Loo, 2011) (http://www.worldagroforestry.org/treedb/AFTPDFS/Balanites_aegyptiaca.PDF, 2016) (Khalafalla et al., 2011) (Kalinganire y Koné, 2011).

A. digitata es una especie con presencia natural en regiones semiáridas de África, incluyendo la mayoría de los países del sur del Sahara exceptuando Liberia, Uganda, Djibuti, Burundi y la República Centrafricana. El límite norte de su rango son las regiones semidesérticas de matorrales, suele crecer en tierras bajas, bosques maduros y en la sabana. Es importante para humanos y animales en las áreas secas de África, porque proporciona comida y medicina, siendo sus hojas, fruta y semillas, una alta fuente de proteínas, minerales y vitaminas. Soporta temperaturas de hasta 42°C pero está restringida en zonas de heladas (Assogbadjo y Loo, 2011).

La distribución ecológica de *B. aegyptiaca* es amplia, sin embargo su máximo desarrollo se da en terrenos de limo y arcilla de llanuras aluviales y en ecosistemas de sabana y matorral tropical. La pulpa de sus frutos es comestible y se consume tanto seca como fresca, el ganado también se alimenta de sus hojas y brotes. La decocción de sus raíces se utiliza para el tratamiento de la malaria y dolores estomacales (http://www.worldagroforestry.org/treedb/AFTPDFS/Balanites_aegyptiaca.PDF, 2016).

B. senegalensis es un matorral perenne nativo del Sahel y la sabana Sahariana, presente en Mauritania, Senegal, Mali, Níger, Nigeria, Camerún, Sudán, Etiopía, Somalia y Kenia. La especie está adaptada a estas regiones de sequía, tolera temperaturas de hasta 45°C y sobrevive con 100 mm de lluvia anuales. Su importancia radica en el mantenimiento de la economía agrícola rural y es un gran revulsivo contra las hambrunas, ya que sus semillas se consumen en forma de harina y sus hojas se utilizan para endulzar platos. También es conocido su uso medicinal contra enfermedades como cólicos y eczemas y como agente clarificador de aguas turbias (Khalafalla et al., 2011).

Otra especie que contribuye a la seguridad alimenticia es *Ziziphus mauritiana*, se encuentra ampliamente distribuida por los subtrópicos y trópicos africanos, en regiones de áridas a semiáridas de todos los países del Sahel oeste y en Sudán, Kenia, Mozambique y Angola. Es una especie mayoritariamente cultivada por sus características nutritivas. Su principal producto es la pulpa del fruto, que se consume fresca, seca o en zumo y sus hojas, semillas y corteza tienen utilidad medicinal (Kalinganire y Koné, 2011).

El propósito de este trabajo es crear un modelo de nicho ecológico que permita determinar la distribución presente y futura de las especies a partir de una base de datos puntuales georeferenciados repartidos por el territorio continental africano, teniendo en cuenta las condiciones climáticas actuales y su proyección de futuro.

2. MATERIALES.

2.1. Área de estudio.

El área de actuación de este trabajo viene determinada por la selección de las especies, en función de la capacidad de proveer alimento para las comunidades nativas de los hábitats en los que están presentes algún ejemplar objeto de este estudio. Aunque también pueden llegar a desarrollarse en otros lugares del mundo el hábitat autóctono de la selección de especies se encuentra en África, por ese motivo el estudio se centra en el conjunto del territorio africano (Figura 1). La escala elegida es a nivel continental, para conocer la potencialidad de desarrollo en nuevas regiones, libres hasta el momento de distribución alguna. Se ha buscado información puntual de las especies con localización en el continente y climática con el fin de crear modelos que cubran la totalidad de la superficie de estudio.



Figura 1. El territorio africano como área de estudio del proyecto

2.2. Fuentes de datos de presencia y ráster climáticos.

En la fase inicial del proyecto se requiere recopilar la mayor cantidad de información que ayude a establecer un modelo de predicción para la biodiversidad africana. Para ello se necesita conseguir información de la presencia de la especie y condiciones climáticas actuales y futuras en el área de estudio.

Si bien es cierto, las bases de datos de biodiversidad libres no son muy abundantes, en este caso se optó por el banco de datos de GBIF (Global Biodiversity Information Facility). Esta organización nació en 1999 con el ánimo de promover la información sobre biodiversidad de libre acceso y permite conocer cualquier tipo de vida en la Tierra animando y ayudando a las instituciones públicas a hacer públicos los datos, según unos estándares comunes. Las bases de datos de presencias comprenden datos de todo el mundo, expresados en coordenadas geográficas y el sistema de referencia de coordenadas (SRC) mundial WGS84. (Figura 2). La especie *Balanites Aegyptiaca (L.) Delile* se completó con la muestra utilizada en Weber et al., 2009, que cumple con los requerimientos en cuanto al área de actuación de este estudio (Figura 1).

Las condiciones climáticas presentes y futuras necesarias para el trabajo se descargaron de WorldClim.org; una base de datos de capas climáticas para uso cartográfico y creación de modelos espaciales de todo el mundo (Tabla 1). Los datos son a escala mundial vienen dados en coordenadas geográficas en el sistema de referencia de coordenadas (SRC) WGS84 para 18 variables bioclimáticas expresando la temperatura en °C*10 y en milímetros la precipitación (Tabla 2). Además se utilizó la serie de datos más actualizada (versión 1.4) y con la mayor precisión posible (30 segundos geográficos, ~1 km).

El formato de descarga para la actualidad es .BIL, un formato genérico compatible con la mayoría del software de sistemas de información geográfica. Sin embargo para el futuro los datos vienen dados en GeoTIFF. Ambos formatos se incluyen dentro del grupo de datos ráster, que consisten en una matriz de celdas o píxeles organizados en filas y columnas donde cada celda contiene un valor que representa información (<http://support.esri.com/other-resources/gis-dictionary>, 2016).

Los datos de clima futuro se generan con Modelos de Clima Globales (GCMs) y cada uno de ellos ofrece resultados diferentes en función de los escenarios de emisión de gases de efecto invernadero que asuman. Para permitir agilizar los procesos y lidiar con las limitaciones de memoria, las capas se dividen en un número limitado de unidades espaciales (píxeles), cometiendo errores del orden de 2 a 3 grados geográficos. Para solucionarlo un procedimiento válido es comparar las relaciones de los datos climáticos observados a grandes escalas (e.g. presión atmosférica a 1000 metros) y los datos a escala local (e.g. precipitación), esta relación se aplica al nuevo GCM teniendo en cuenta un mayor peso de las variables a gran escala y que las relaciones se mantienen en el futuro.

Los datos de futuro disponibles en WorldClim se produjeron interpolando cambios discretos entre la actualidad y los valores predichos a una resolución de 1 km, asumiendo que los cambios del clima tienen una correlación espacial muy alta. Estos cambios luego son aplicados a los datos del periodo actual para conseguir Modelos de Clima Globales (GCMs).

La base cartográfica de los límites políticos del territorio africano es de tipo vectorial. Los datos vectoriales se basan en un sistema de coordenadas para representar las características geográficas como puntos, líneas y polígonos. Cada punto está representado por un único par de coordenadas mientras que líneas y polígonos se representan como listas de vértices de inicio y fin, los vértices en los que se unen varias líneas o arcos se denominan nodos (support.esri.com/other-resources/gis-dictionary, 2016). En este caso la cartografía vectorial empleada es de tipo polígono, en formato shapefile (.shp), representada con el SRC WGS 84 y coordenadas geográficas y descargada de ArcGIS.com.

<i>Adansonia Digitata L.</i>	 GBIF	http://doi.org/10.15468/dl.kqrtwa
<i>Balanites Aegyptiaca (L.) Delile</i>	 GBIF Weber et al., 2009	http://doi.org/10.15468/dl.w4vitd
<i>Boscia Senegalensis (Pers.) Lam.</i>	 GBIF	http://doi.org/10.15468/dl.pkvpfp
<i>Ziziphus Mauritiana Lam.</i>	 GBIF	http://doi.org/10.15468/dl.mg0bkf
Condiciones climáticas actuales (~1960-1990)	 WorldClim	http://www.worldclim.org/current
Condiciones climáticas futuras (media para 2041- 2060)	 WorldClim	http://www.worldclim.org/cmip5_30s

Tabla1. Fuentes de descarga de datos.

VARIABLE	LITERAL
BIO1	Temperatura Media Anual
BIO2	Rango Medio Diurno (Media mensual (max temp - min temp))
BIO3	Isotermalidad (BIO2/BIO7) (* 100)
BIO4	Estacionalidad de la Temperatura (desviación estándar*100)
BIO5	Temperatura Máxima del Mes más Cálido
BIO6	Temperatura Mínima del Mes más Frío
BIO7	Rango de Temperatura Anual (BIO5-BIO6)
BIO8	Temperatura Media del Cuarto más Húmedo
BIO9	Temperatura Media del Cuarto más Seco
BIO10	Temperatura Media del Cuarto más Cálido
BIO11	Temperatura Media del Cuarto más Frío
BIO12	Precipitación Anual
BIO13	Precipitación del Mes más Húmedo
BIO14	Precipitación del Mes más Seco
BIO15	Estacionalidad de la Precipitación (Coeficiente de Variación)
BIO16	Precipitación del Cuarto más Húmedo
BIO17	Precipitación del Cuarto más Seco
BIO18	Precipitación del Cuarto más Cálido
BIO19	Precipitación del Cuarto más Frío

Tabla 2. Variables Bioclimáticas.

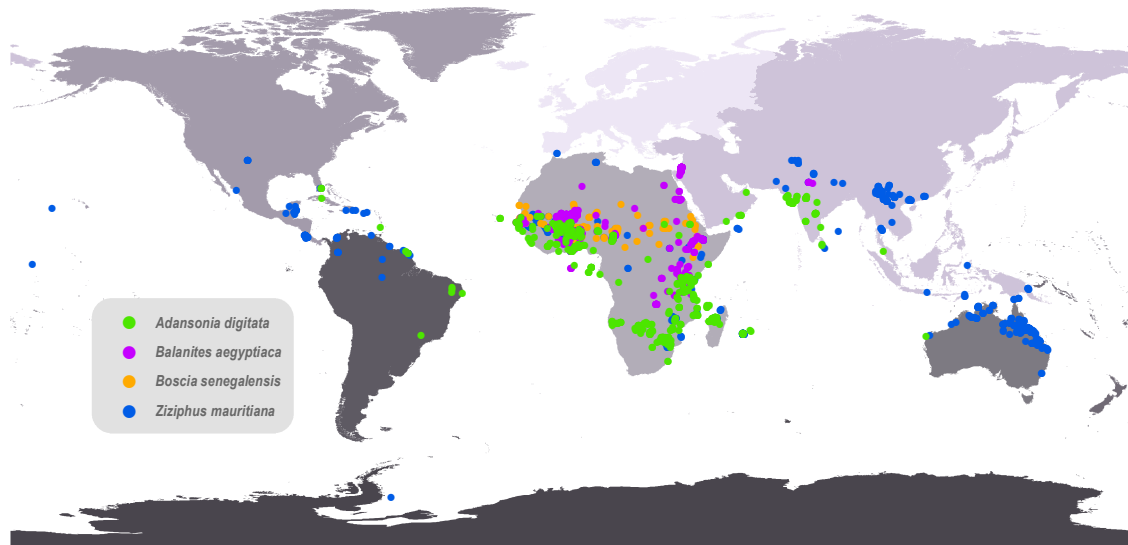


Figura 2. Datos originales de distribución de especies.

3. METODOLOGÍA.

3.1. Depurado de datos de presencia y ráster climáticos

Una vez que ya están descargados todos los datos necesarios para el estudio se procede a examinarlos con el fin de tratarlos de la manera más eficiente posible. La representación geográfica de todas las presencias de las especies objeto de este estudio demuestra que muchas de las coordenadas son erróneas porque aparecen en zonas aisladas de los océanos Atlántico, Índico y los mares Mediterráneo y Rojo, entonces, es necesario eliminar esta información que añade ruido al estudio y sesga su precisión.

El software de sistemas de información geográfica ArcGIS ofrece la posibilidad de iterar operaciones en distintas capas, con el fin de automatizar procesos y su correspondiente ahorro de tiempo, la herramienta Model Builder es un lenguaje de programación visual para crear flujos de geoprocésamiento representados como un diagrama que une secuencias de procesos. Otra herramienta de ArcGIS que comprende el proceso de depuración de presencias es *Recorte*, que permite eliminar entidades que no se encuentran dentro de los límites de la capa de recorte y mantener la parte interior.

Cada una de las especies constituye una capa vectorial de tipo punto y el área de estudio otra capa vectorial, pero, de tipo polígono, usando Model Builder se recorta cada presencia con el límite del continente africano, obteniendo de este modo, registros de especies en zonas de tierra en África (Figura 3).

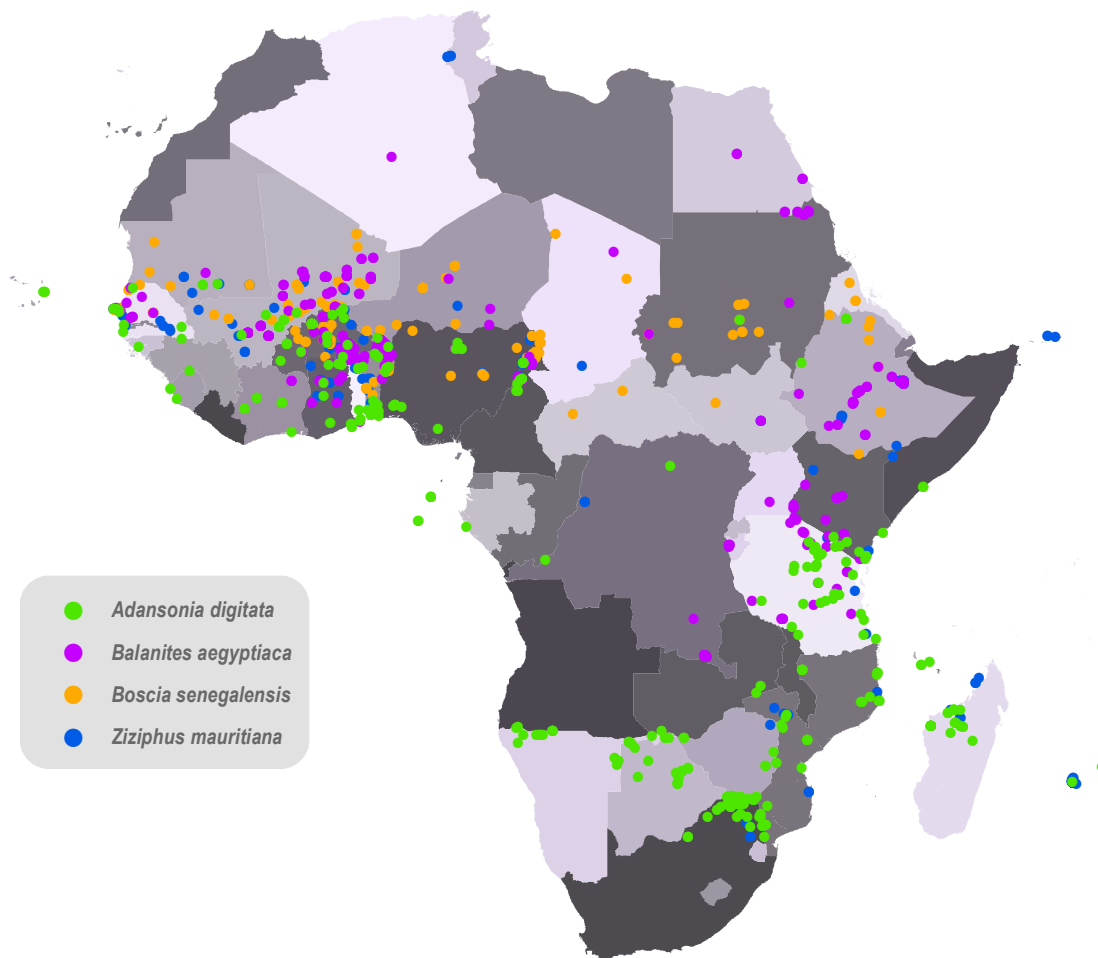


Figura 3. Distribución espacial de las especies susceptibles de estudio.

Las bases de datos de cada una de las especies están compuestas por 224 campos, de los cuales muchos de ellos se encuentran vacíos, otros sólo contienen información para algunos registros y también los hay que contienen datos para toda la serie.

Después de comprender la información que proporciona la base de datos se debe realizar un filtro para depurar los datos que resulten poco convenientes para el trabajo. En primer lugar se deben eliminar los campos vacíos, una tarea tediosa, teniendo en cuenta la gran cantidad que suponen (126 campos en el caso de *Adansonia digitata L.*) en cada una de las cuatro bases de datos de las especies.

Para automatizar el proceso se ha encontrado una solución rápida, sencilla y eficiente, recurriendo al lenguaje de programación Python, del que utilizamos su funcionalidad a través de una librería de análisis de datos llamada Pandas, de acceso libre, que complementa a Python en el área de análisis de datos. La manera más simple de instalar esta librería es con un administrador de paquetes, también de acceso libre llamada Anaconda, que se usa para ingeniería de datos.

Con ayuda de código de libre acceso, se da solución a nuestro problema con un código simple mostrado en la Figura 4, en este ejemplo, lee la base de datos de presencias de la especie *Adansonia digitata*, elimina las columnas que cumplan que todos sus registros contengan “NA”, es decir, que estén vacíos y crea un nuevo archivo .csv en el que todas las columnas tienen como mínimo un registro con datos.

```
from pandas.io.parsers import read_csv
data = read_csv('ADANSONIA_DIGITATA.csv')
filtered_data = data.dropna(axis='columns', how='all')
filtered_data.to_csv('ADANSONIA_DIGITATA_D.csv')
```

Figura 4. Código para eliminar columnas vacías

El proceso se repite con los datos del resto de especies y los resultados varían en cada una de ellas, así, después de este primer depurado, los datos contienen 98, 79, 79 y 111 columnas y 332, 485, 165 y 192 registros para *Adansonia digitata*, *Balanites aegyptiaca*, *Boscia senegalensis* y *Ziziphus mauritiana* respectivamente.

Ahora que ya se han filtrado las presencias de interés y eliminado la información sin utilidad para éstas, se consultó a varios expertos en agricultura para el desarrollo en África con el fin de conocer los campos de las bases de datos más adecuados para el trabajo, desechando aquellos que los expertos veían conveniente por su falta de aplicación con el objetivo de este trabajo (Tabla 3).

Una cantidad de puntos de presencia es crucial para obtener unos resultados de modelación convincentes, pero sobre todo influye en gran medida que los puntos de presencia estén distribuidos por todo el territorio en cuestión. Teniendo en cuenta las condiciones climáticas del continente africano se considera que la cantidad y dispersión de las presencias es la adecuada, ya que se obtienen mayores cantidades de registros en el centro y el este, coincidiendo con zonas de gran idoneidad de las especies de este trabajo.

IDENTIFICADOR	LITERAL
gbifID	Identificador único dado por GBIF
class	Clase taxonómica
countryCOD	Código del país de pertenencia
county	Región de pertenencia
decimalLat	Distancia angular entre el objetivo y el paralelo del Ecuador
decimalLon	Distancia angular entre el objetivo y el meridiano de Greenwich
eventDate	Fecha de registro

IDENTIFICADOR	LITERAL
eventRemar	Observaciones del registro
family	Familia botánica
genus	Género taxonómico
kingdom	Reino taxonómico
locality	Municipio
phylum	División taxonómica
scientific	Nombre científico
specificEP	Epíteto específico
year	Año de registro

Tabla 3. Relación de campos elegidos para las bases de datos.

Los datos ráster que contienen información ambiental deben abarcar la misma área de actuación, tener el mismo número de píxeles y del mismo tamaño, pero, los datos en crudo de los que se dispone tienen una escala mundial por lo que es necesario adecuarlos al área de interés del estudio. Se recurre de nuevo a la funcionalidad de ArcGIS, Model Builder para, como en el caso anterior, crear un proceso de iteración de la herramienta *Recorte*, utilizando como información de entrada los ráster originales y como área de interés la extensión de la cartografía base vectorial del territorio africano (Figura 5), suponiendo un ahorro de tiempo considerable en comparación con la forma manual de la operación.

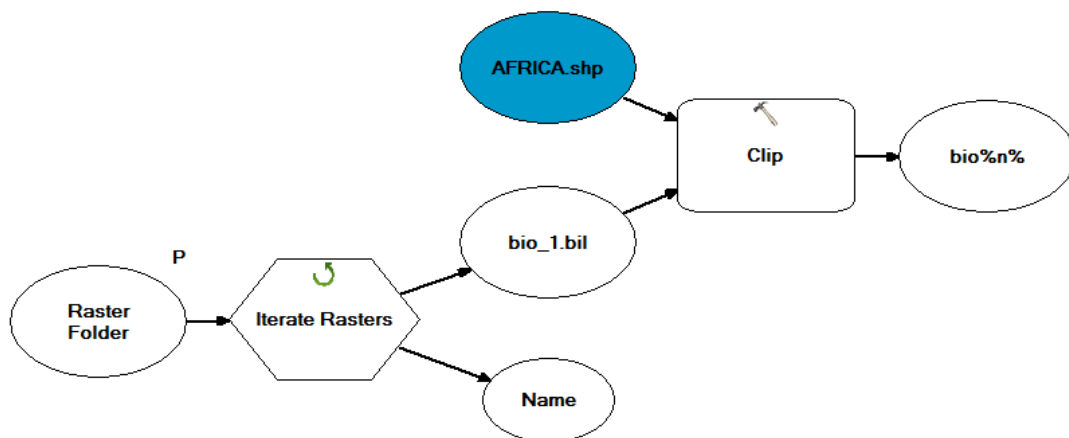


Figura 5. Esquema Model Builder iterador para el *Recorte* de ráster

3.2. Modelación y análisis de distribución de especies.

Los datos de presencia disponibles no cubren todo el rango de distribución natural de las especies por lo que es necesario un programa de modelación que permita aproximar el rango total de distribución para identificar las áreas probables de presencia de las especies.

Maxent es un software de modelación de nicho ecológico que identifica sitios con ambientes similares a aquellos donde ya se ha encontrado la presencia de una especie, como áreas de posible incidencia (Scheldeman y van Zonneveld, 2010).

Los datos requeridos para llevar a cabo el modelo son los valores de nicho de cada especie y ráster climáticos, que cubran la totalidad de la superficie potencial de presencia, luego para cada celda se calculan las similitudes de los valores ambientales en un píxel específico, el modelo calcula la idoneidad de presencia en cada celda del ráster.

La entropía es una medida de posibilidades en la selección de un evento, más entropía equivale a más posibilidades, más dispersas en el espacio y relacionadas con una distribución uniforme. Cuando se usa Maxent, un espacio “x” representa un conjunto de celdas con valores ambientales discreto que cubren el área de estudio, (x_1, x_2, \dots, x_m) . La función de distribución de la idoneidad en el área de estudio “P(x)” se implementa con un eficiente algoritmo (Phillips et al., 2004) basado en la aproximación de máxima entropía, idéntico a la distribución de Gibbs:

$$P(x) = \exp (c1*f1(x)+ c2*f2(x)+ \dots) / Z \quad \text{Ec.1}$$

Donde “P(x)” es la función de idoneidad, “c1, c2...” son constantes, “f1, f2,…” son funciones para cada variable ambiental y “Z” es una constante que asegura que la idoneidad llegue como máximo hasta 1 en todo el área de estudio (Moreno et al., 2011).

3.2.1 Preparación de datos de entrada.

Antes de ejecutar Maxent se deben preparar los datos de entrada para que los resultados sean aceptables; el .csv de presencias debe contener tres campos fijos que se corresponden con la especie, longitud y latitud, en este orden estrictamente, aunque se permiten columnas adicionales, que serán ignoradas por el programa. Se recomienda que los datos ráster de entrada estén en formato .ascii, para que el proceso de modelación de datos se ejecute sin interrupciones (Scheldeman y van Zonneveld, 2010), sin embargo, los datos originales vienen dados en formatos .BIL y GeoTIFF para las condiciones actuales y futuras respectivamente, es por tanto necesario usar de nuevo Model Builder para convertir al formato adecuado utilizando la herramienta *Raster a ASCII* (Figura 6).

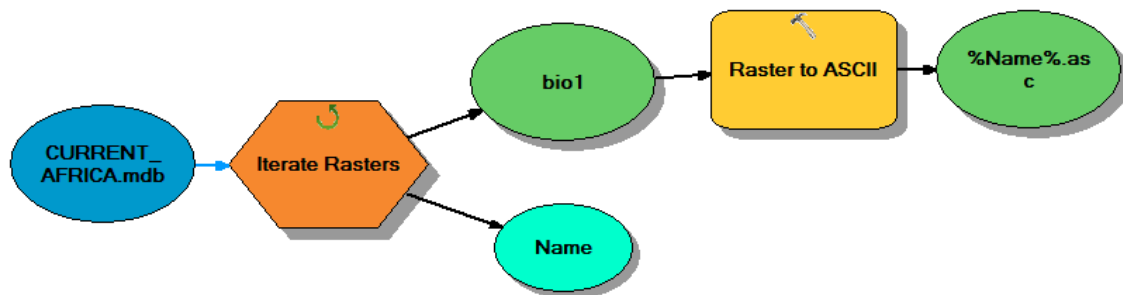


Figura 6. Esquema Model Builder iterador para la conversión de ráster a .asc

3.2.2. Creación del modelo de nicho ecológico.

De las 19 variables bioclimáticas de las que se disponen, no todas ellas son adecuadas para generar un modelo óptimo, pudiendo algunas resultar de más utilidad para una especie que para otras. Deben pasar unos filtros estadísticos que ofrezcan confianza en el uso de los resultados.

El procedimiento estadístico más común para medir la fortaleza de las relaciones entre dos variables continuas es la correlación de Pearson, la muestra aleatoria comprende pares de puntos (y_{i1}, y_{i2}) , entonces se dice que existe correlación cuando modificando los valores de una de las variables, la segunda también cambia con la misma magnitud que la primera (Quinn y Keough, 2002) como demuestra la ecuación 2.

$$r_{Y_1Y_2} = \frac{\sum_{i=1}^n [(y_{i1} - \bar{y}_{i1})(y_{i2} - \bar{y}_{i2})]}{\sqrt{\sum_{i=1}^n (y_{i1} - \bar{y}_{i1})^2 \sum_{i=1}^n (y_{i2} - \bar{y}_{i2})^2}} \quad \text{Ec. 2}$$

Donde y_{i1} , y_{i2} son los valores de dos variables, \bar{y}_{i1} y \bar{y}_{i2} son las medias de la muestra para cada una de las variables en cuestión.

El método de los pares de puntos no ponderados usando medias aritméticas (UPGMA) es generalmente el más recomendado como estrategia para la creación del *clustering* acumulado (Quinn y Keough, 2002), se opta por ejecutar el algoritmo *hclust()* en R, ya que cumple con las condiciones descritas (cran.r-project.org, 2016).

Si las variables objeto del estudio están correlacionadas, se dice que los datos están afectados por colinearidad, estadístico que aplicado a estudios de ecología mide la capacidad de que dos variables o más sean capaces de explicar una tercera o, lo que es lo mismo, que cumplen con un modelo de regresión lineal múltiple. La falta de colinearidad es muy difícil de obtener en datos biológicos y produce un efecto perjudicial para el modelo; pequeños cambios en los datos, como añadir o quitar una de las variables pueden cambiar la distribución considerablemente. Esto significa que no se podría confiar en las estimaciones del modelo resultante (Quinn y Keough, 2002).

Una buena forma de evaluar la colinearidad es con el factor de la varianza de inflación (VIF), que se define como el inverso de la diferencia de 1 y la correlación, valores por encima de 10 sugieren una fuerte colinearidad (Quinn y Keough, 2002) (Ec.3).

$$VIF = \frac{1}{1 - r^2} \quad \text{Ec. 3}$$

Otro elemento que ayuda a refinar el modelo es la devianza, un estadístico logarítmico que indica el valor de explicación de cada variable objeto del estudio en el modelo resultado, por lo que ayuda a una mejor elección de las variables una vez conocida la correlación entre ellas. La máxima verosimilitud es un método estadístico que consiste en estimar parámetros que maximicen la probabilidad de que ocurra la muestra tomada para el estudio, de este modo se define la devianza como la diferencia logarítmica del modelo menos ajustado y el más ajustado (Quinn y Keough, 2002) (Ec. 4).

$$\text{Devianza} = -2 (\log\text{-verosimilitud reducida}) - (\log\text{-verosimilitud completa}) \quad \text{Ec. 4}$$

Con la inestimable ayuda de una experta en la materia se utilizó un código escrito en el lenguaje de programación R que permitió conocer los estadísticos antes expuestos y necesarios para calibrar el modelo con el que se quiere conocer las distribuciones actual (~1960-1990) y futura (media para 2041-2060) de las cuatro especies vegetales que se estudian.

El análisis de correlación se ejecutó con una muestra del 20 % de los píxeles del área de actuación no nulos. Para crear esta muestra se utilizó la herramienta de ArcGIS, *Crear puntos aleatorios*, estableciendo un número de puntos determinado, en este caso fueron 8.864.949, en los que posteriormente se extrajeron los valores de las 19 variables climáticas a examinar con un sencillo código en R creado a partir de códigos de acceso libre (<https://cran.r-project.org>, 2016).(Figura 8).

La devianza explicada toma como *input* una tabla con las presencias recogidas para cada especie y 5000 ausencias tomadas aleatoriamente del área de estudio en las que se extrajeron los valores de las 19 variables climáticas con el código de R propuesto anteriormente (Tabla 4).

Una vez calculadas la correlación y la devianza de las variables climáticas, se escogen las variables no correlacionadas y las que mejor expliquen la distribución de cada especie para la calibración del modelo.

El resultado del cálculo de la correlación se ofrece en forma de dendrograma, un gráfico cuyo método se basa en una agrupación jerárquica acumulativa, que calcula las diferencias de una matriz de pares de objetos, agrupando aquellos con mayor similitud (Quinn y Keough, 2002) (Figura 7).

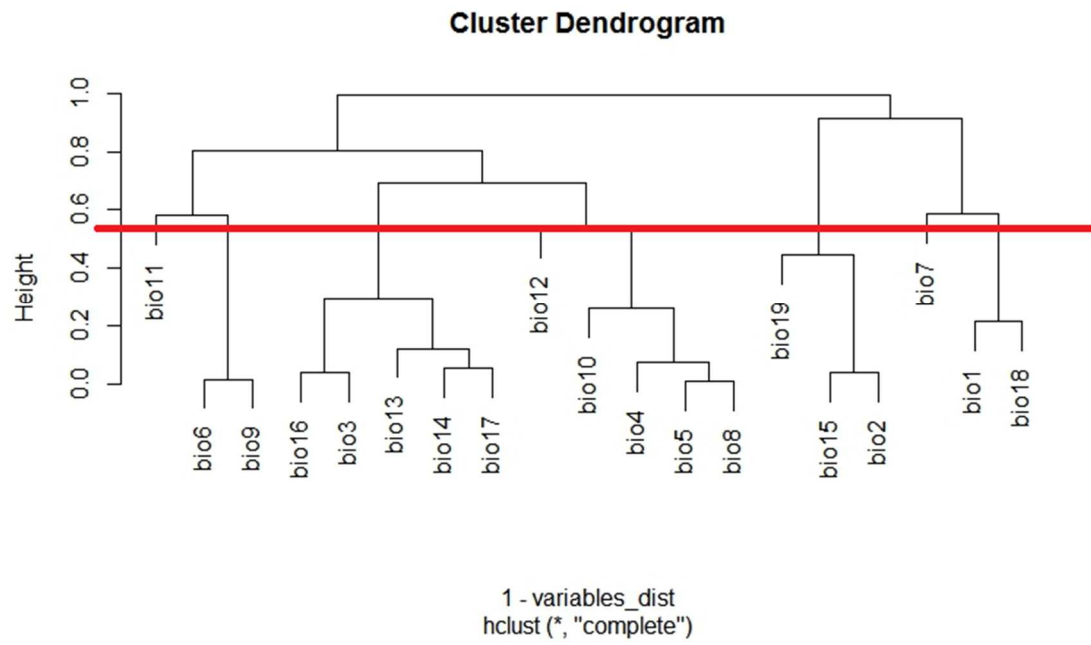


Figura 7. Agrupación de variables en función de su correlación

```
#####EXTRACCIÓN DE DATOS AMBIENTALES#####
#Lectura de un shapefile con puntos aleatorios
muestra <- readOGR(dsn = "C:/Users/acer/Documents/TFM/CORRELACION/RANDOM", layer = "random")
#Lista de capas ambientales
lista_variables <- list.files(path="C:/Users/acer/Documents/TFM/CORRELACION/CURRENT_ASCII", pattern='*.asc', full.names=TRUE)
#Combinación de marcos de datos entre el shapefile y los archivos raster
variables <- stack(lista_variables)
#Extracción de los valores ambientales en los puntos
valores <- extract(variables, muestra)
#Exportación de resultados
write.table(valores, "C:/Users/acer/Documents/TFM/CORRELACION/valoresthree.txt", sep=" ")
```

Figura 8. Extracción de los valores climáticos en los puntos aleatorios

obs	bio1	bio2	bio3	bio4	bio5	bio6	bio7	bio8	bio9	bio10	bio11	bio12	bio13	bio14	bio15	bio16	bio17	bio18	bio19
0	247	269	228	521	164	4	107	280	26	60	26	163	76	1529	369	155	214	255	228
0	290	343	222	105	39	0	136	84	2	61	4	144	50	4760	423	140	283	336	261
0	277	309	240	208	51	3	91	124	11	78	11	123	61	2728	373	174	199	303	240
0	212	307	119	104	13	1	41	35	9	9	27	129	37	7346	396	55	341	218	307
0	222	232	206	1354	239	1	73	569	7	569	7	110	72	1071	286	134	152	232	206
0	275	313	237	658	212	0	128	494	0	134	0	141	54	2872	393	136	257	272	237
0	215	290	125	0	0	0	0	0	0	0	0	170	49	6512	378	35	343	140	140
0	267	329	182	13	6	0	88	11	0	11	0	156	48	5897	413	91	322	329	202
0	285	320	259	695	205	0	122	499	0	67	1	135	58	2449	398	167	231	271	260
0	287	336	220	98	41	0	141	84	0	38	0	156	52	4578	423	127	296	321	236

Tabla 4. Input de la especie *Adansonia Digitata* para el cálculo de la devianza

El análisis muestra nueve grupos de variables, algunos de ellos con una única variable, de aquellos con dos o más se elegirán las más propicias según su nivel de explicación del modelo por especies (devianza), los valores más altos de este estadístico indican mayor grado de explicación, en la Tabla 5 se resaltan los valores adecuados de las variables de cada agrupación por especie.

Estas variables también tienen que ser contrastadas bajo el punto de vista de la colinearidad, aquellas con un valor mayor de 10 deben ser desestimadas. El proceso consiste en ejecutar el código R del VIF con las variables libres de correlación y con mayor devianza, eliminando las que no ofrezcan un resultado válido, el proceso se repite hasta conseguir que todos los valores sean los óptimos. En la primera ejecución la única especie que cumplía con los requerimientos de la colinearidad fue *Ziziphus mauritiana*, en las demás especies se eliminan las variables con mayor valor y menos relación con la supervivencia de la especie, de este modo para *Adansonia digitata* se elimina la variable 15 (Estacionalidad de la precipitación), porque puede soportar ambientes muy secos, para *Balanites aegyptiaca* se elimina la 3 (Isotermalidad), porque tiene flexibilidad para adaptarse a condiciones cambiantes y la misma se elimina para *Boscia senegalensis* porque las altas temperaturas no son impedimento para su desarrollo. Después de la segunda ejecución todas las variables carecían de colinearidad. (Tabla 6)

El modelado de generalización lineal (GLM) de regresión logística utilizado en este trabajo necesita datos de presencia, localizaciones en las que se ha contrastado que existe la especie y datos de ausencia, puntos en los que se ha verificado que no existe la especie. Sólo se disponen de datos de presencia por lo que se deben determinar puntos de control que den consistencia al estudio (Hijmans et al., 2016).

Los puntos de control, background points o pseudo-ausencias no intentan predecir la distribución en puntos de ausencia, sino que pretenden caracterizar el área de estudio, pese a que los datos de presencia establecen las condiciones bajo las que las especies son más idóneas (Phillips et al., 2005). Para este caso particular se eligió una muestra de 100.000 puntos de control, dispersos de forma aleatoria por toda el área de estudio. Cuantos más background points se dispongan, mejores resultados ofrecerá el modelo. Se comprobó que utilizando más de 100.000 no hay variaciones en la calidad y además el modelado comienza a ralentizarse.

Una vez calibrado el modelo, se introducen los datos elegidos en la interfaz del programa Maxent, este proceso tiene dos fases; 1) en la primera se calculan las distribuciones de las especies para las condiciones actuales, tomando como input las presencias, las variables climáticas adecuadas y el número de puntos de control, 2) posteriormente se hace la modelación del cambio climático en la distribución de las especies utilizando como datos de entrada todos los anteriores y las mismas variables climáticas pero para las situaciones de futuro.

3.2.3. Impacto del cambio climático en la distribución de especies.

El cambio climático es cada vez más evidente y consecuentemente las áreas geográficas que corresponden a biomas, ecosistemas y nichos ecológicos de las especies están cambiando, lo que afectará a la distribución natural de muchas especies (Scheldeman y van Zonneveld, 2011), amenazando su viabilidad debido a la reducción de áreas propicias para su desarrollo.

DESCRIPCION	VARIABLE	CLUSTER	ADANSONIA	BALANITES	BOSCIA	ZIZIPHUS
Temperatura Media Anual	bio1	1	0.01855234	0.12750271	0.14856574	0.03453192
Precipitación del Cuarto más Cálido	bio18	1	0.11900341	0.08470078	0.13078851	0.07291335
Estacionalidad de la Temperatura (desviación estándar*100)	bio4	2	0.09900641	0.18281417	0.10600711	0.10056167
Temperatura Máxima del Mes más Cálido	bio5	2	0.08446434	0.13481093	0.10936973	0.09023719
Temperatura Media del Cuarto más Húmedo	bio8	2	0.08240277	0.13218896	0.10797571	0.09009865
Temperatura Media del Cuarto más Cálido	bio10	2	0.10573717	0.06229327	0.05779898	0.04013556
Temperatura Media del Cuarto más Frío	bio11	3	0.00389664	0.00297538	0.00262179	0.03928497
Precipitación Anual	bio12	4	0.04706192	0.0916777	0.03844458	0.09561581
Isotermalidad (BIO2/BIO7) (* 100)	bio3	5	0.04981792	0.18175256	0.14774238	0.08503608
Precipitación del Mes más Húmedo	bio13	5	0.14253502	0.11887	0.12872462	0.04761151
Precipitación del Mes más Seco	bio14	5	0.1473397	0.16093776	0.14709507	0.10806483
Precipitación del Cuarto más Húmedo	bio16	5	0.04601082	0.13890144	0.13510864	0.12366101
Precipitación del Cuarto más Seco	bio17	5	0.09084215	0.15954401	0.12747213	0.09295935
Rango Medio Diurno (Media mensual (max temp - min temp))	bio2	6	0.05125007	0.02313771	0.07167778	0.01217931
Estacionalidad de la Precipitación (Coeficiente de Variación)	bio15	6	0.05648068	0.0273749	0.07564216	0.006203
Precipitación del Cuarto más Frío	bio19	7	0.0229682	0.05042404	0.08366	0.02897451
Temperatura mínima del Mes más Frío	bio6	8	0.01139022	0.01286664	0.03124748	0.03681061
Temperatura Media del Cuarto más Seco	bio9	8	0.01457727	0.01516514	0.03817421	0.03162332
Rango de Temperatura Anual (BIO5-BIO6)	bio7	9	0.04117682	0.09046483	0.18112704	0.00135369

Tabla 5. Elección de las variables más representativas para cada especie

PRIMERA EJECUCIÓN

bio7 1.732094	bio9 2.131438	bio10 2.588443	bio11 1.50971	bio12 3.033398	bio14 5.155827	bio15 10.814492	bio18 4.933199	bio19 4.10798
bio1 51.640568	bio3 34.301031	bio4 4.340999	bio7 1.648966	bio9 2.078596	bio11 1.854257	bio12 3.892013	bio15 22.758926	bio19 2.631815
bio1 53.184674	bio3 35.625978	bio5 2.999617	bio7 1.704517	bio9 1.821299	bio11 1.776263	bio12 3.895082	bio15 22.951898	bio19 2.635729
bio2 5.478549	bio4 3.853388	bio6 1.750756	bio7 1.656911	bio11 1.820785	bio12 2.650377	bio16 6.448318	bio18 5.999149	bio19 4.08834

SEGUNDA EJECUCIÓN

bio7 1.719716	bio9 2.104063	bio10 2.49596	bio11 1.477209	bio12 1.9567	bio14 2.972884		bio18 1.697296	bio19 1.405714
bio1 3.305167		bio4 3.689207	bio7 1.648559	bio9 2.008788	bio11 1.824219	bio12 3.207994	bio15 7.91909	bio19 2.514728
bio1 3.249115		bio5 2.456301	bio7 1.688656	bio9 1.812156	bio11 1.746596	bio12 3.211233	bio15 7.590151	bio19 2.503764
bio2 5.478549	bio4 3.853388	bio6 1.750756	bio7 1.656911	bio11 1.820785	bio12 2.650377	bio16 6.448318	bio18 5.999149	bio19 4.08834

Adansonia digitata
Balanites aegyptiaca
Boscia senegalensis
Ziziphus mauritiana

Tabla 6. Descarte de variables para eliminar la colinearidad. Ejecución del VIF.

Muchas especies viven en regiones protegidas, de este hecho depende su supervivencia, sin embargo puede que en el futuro su hábitat más propicio no lo esté y terminen desapareciendo (Velásquez-Tibatá et al., 2012).

El Grupo Intergubernamental de Expertos sobre el Cambio Climático (IPCC), se creó en 1988 con la finalidad de proporcionar evaluaciones integrales del estado de los conocimientos científicos, técnicos y socioeconómicos sobre el cambio climático, sus causas, posibles repercusiones y estrategias de respuesta (<http://www.ipcc.ch/>). La generación de distribuciones de las especies vegetales en condiciones futuras se hizo con modelos climáticos globales (GCMs) procedentes del quinto informe de evaluación de este organismo.

Los escenarios socioeconómicos y de emisiones se utilizan en la investigación del clima para ofrecen predicciones realistas de cómo evolucionarán en el futuro las formas de crear energía, los usos del suelo, la tecnología, el nivel socioeconómico y la cantidad de gases de efecto invernadero presentes en la atmósfera. Al conjunto de escenarios que contienen las predicciones de emisiones, su concentración y de usos del suelo se les llaman Trayectorias de Concentración Representativa (Representative Concentration Pathways, RCPs), éstas deben basarse en escenarios previamente publicados, desarrollados independientemente y ser representativos del total de las publicaciones al respecto en términos de emisiones y sus concentraciones (van Vuuren et al. 2011). Existen cuatro RCPs definidas por su potencia radioactiva total, una medida de las emisiones antrópicas de gases de efecto invernadero de todas las fuentes expresada en vatios por metro cuadrado hasta el año 2100 (Tabla 7).

RCPs	Descriptivo de las condiciones bajo el escenario
RCP 8.5	Escenario más pesimista, en el futuro no incluye cambios de leyes para reducir las emisiones.
RCP 6	Fuerte dependencia de los combustibles fósiles y emisiones de metano estables.
RCP 4.5	Reducción de emisiones ambiciosa, programas de reforestación y reducción de cultivos.
RCP 2.6	Escenario más optimista, producción de bio-energía y reducción de un 40% de las emisiones de metano.

Tabla 7. Trayectorias de Concentración Representativa (sei-international.org, 2016)

Los Modelos Sistémicos Terrestres (Earth System Models, ESMs) son modelos climáticos globales con la capacidad de representar procesos biogeoquímicos que interactúan con el clima y lo alteran incluyendo las emisiones antrópicas de gases de efecto invernadero (Flato, 2011).

Tras consultar al experto Javier Pórtoles de la Fundación para la Investigación del Clima, se eligieron los escenarios RCP 8.5, que supone un aumento de la temperatura global de más de 4°C hasta 2100 a y el RCP 4.5 que predice un aumento de 2°C (Chaalalia et al., 2016) (Gomes y Brito, 2015) (Chou et al. 2014). En cuanto a los modelos climáticos globales (GCMs) para conocer las distribuciones de las especies hasta el año 2050, se recomienda analizar cinco modelos, ya que resulta un número adecuado para contrastar los resultados y

ofrecer datos del impacto del cambio climático objetivos, todos ellos están dentro del grupo de los ESMs, valorados por la adecuación de sus modelos a las últimas condiciones terrestres.

El primero de los modelos elegidos fue el HadGEM2-ES, usado por una gran cantidad de instituciones en todo el mundo para la predicción el tiempo y para la investigación sobre el clima. Está compuesto por dos GCMs uno atmosférico y otro para los océanos y dos componentes sistémicos terrestres, uno del ciclo del carbono y otro de la composición química de la troposfera (<https://verc.enes.org/models/earthsystem-models/metoffice-hadley-centre/hadgem2-es>) (Araya et al., 2015) (Karunaratne et al., 2014) (Gomes y Brito, 2015) (McSweeney y Jones, 2016) (Déqué, et al., 2016).

El IPSL-CM5A-LR_IP, desarrollado por el instituto Pierre Simon Laplace, incluye 5 componentes que representan el sistema climático de la Tierra y su ciclo del carbono; atmósfera, NEMO (océano, biogeoquímica del océano y hielo de los mares), superficies y vegetación continental, químicos presentes en la atmósfera y un modelo de catástrofes (<https://verc.enes.org/models/earthsystem-models/ipsl/ipslesm>, 2016) (Araya et al., 2015) (Gomes y Brito, 2015) (McSweeney y Jones, 2016)

MIROC-ESM es un Modelo Sistémico Terrestre basado en el Modelo Climático Global MIROC (Model for Interdisciplinary Research on Climate), está compuesto por un modelo atmosférico de circulación general, un componente de aerosoles, otro de un GCM de los océanos y hielo de los mares, un modelo de la superficie terrestre y componentes para la químicos de la atmósfera, detritos de plancton y ecosistemas de los océanos (Watanabe et al., 2011) (Araya et al., 2015) (Gomes y Brito, 2015)

La contribución alemana al Quinto Informe de Evaluación del Grupo Intergubernamental de Expertos sobre el cambio climático (IPCC) es el modelo MPI-ESM-LR que acopla componentes de la atmósfera, el océano y la superficie terrestre a través del intercambio de energía, potencia, agua e importantes gases como el dióxido de carbono (<http://www.mpimet.mpg.de/en/science/models/mipi-esm.html>, 2016) (Déqué, et al., 2016) (Araya et al., 2015)

El último modelo utilizado fue el NorESM1-M, desarrollado en Oslo, su desarrollo empezó bajo el ahora finalizado proyecto NorClim y continuó como parte del proyecto EarththClim participado por diez instituciones de prestigio. Está formado por cuatro componentes; el atmosférico, el de la superficie terrestre, el del hielo de los océanos y el componente oceánico (verc.enes.org) (Bentsen et al., 2012) (Araya et al., 2015) (Gomes y Brito, 2015) (McSweeney y Jones, 2016) (Déqué, et al., 2016).

3.2.4. Ejecución del modelo en Maxent.

Conocidos todos los datos de entrada de Maxent sólo queda ejecutarlo para obtener las distribuciones de las especies y el impacto del cambio climático sobre ellas. En la figura 9 se muestra el aspecto de la interfaz y los parámetros de ejecución del proceso; 1) cuadro de diálogo de la muestra de presencias de la especie, 2) elección de las 19 variables climáticas sin correlación, sin colinearidad y máxima devianza explicada, 3) elección del formato de datos logístico, ya que mejora la calibración del modelo, de este modo, los valores de salida más diferentes se corresponden mejor con grandes diferencias de idoneidad (Phillips y Dudík, 2008), 4) directorios de salida y proyecciones del cambio climático, 5) selección del número de *background points* 6) ejecución del modelo.

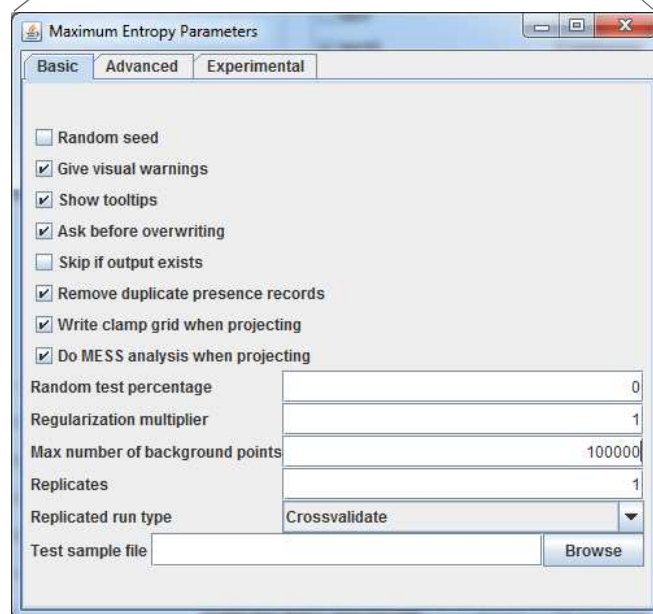
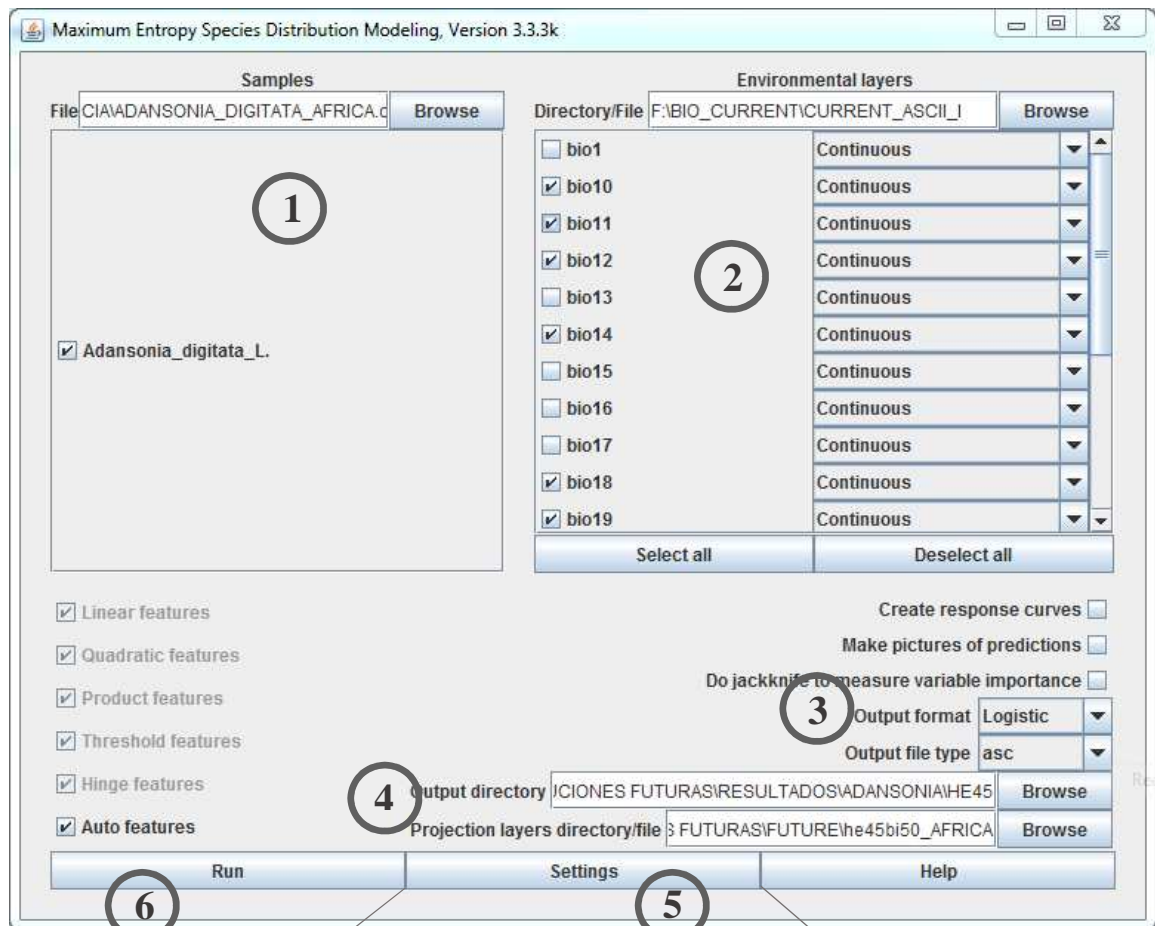


Figura 9. Proceso de ejecución en Maxent

El resultado principal de la ejecución de Maxent es un ráster que contiene valores de idoneidad de la especie en cada uno de los píxeles que lo forman. Estos datos necesitan ser representados gráficamente en un sistema de información geográfica, de forma que se

elimine la visibilidad de los píxeles con una idoneidad inferior a un umbral definido por Maxent y permanezcan los que están por encima de éste. La elección del umbral es objeto de discusión ya que hay disparidad de opiniones entre cuál sería el mejor valor de corte, para determinarlo, este estudio recopiló referencias de otros autores que plantean problemáticas parecidas, en cuanto a la metodología se refiere, a las que aquí se exponen (Hoveka et al., 2015) (De Cauwer, et al., 2014) (Khanum et al., 2013) (Cameron et al., 2008) (Escalante et al., 2013) (Redon y Luque, 2010).

El valor del percentil diez de las presencias de entrenamiento (*10 Percentil Training Presence*) significa que el 10% de los registros de cada especie con los menores valores en el modelo predicho estarán en regiones de ausencia del modelo con umbral y las regiones de presencia abarcarán el otro 90% de los registros, que tienen los valores más altos del modelo (Cameron et al., 2008). Se optó por este umbral porque no se disponían datos de ausencia (Escalante et al., 2013) y porque ofrece un resultado ecológicamente significativo en comparación con otros valores de umbral mucho más restrictivos (Redon y Luque, 2010).

3.2.5. Representación gráfica de las modelaciones.

Las capas de distribuciones resultado de Maxent se procesan en el software de Sistemas de Información Geográfica ArcGIS para su representación gráfica. Todos los conjuntos de datos geográficos tienen un sistema de coordenadas que se usa a través de ArcGIS para mostrar, medir y transformar los datos. Los ráster resultantes de Maxent tienen un sistema de referencia de coordenadas desconocido por lo que es necesario definirlo para poder aplicar los umbrales de presencia/ausencia y cambiar sus valores decimales para hacerlos más legibles. Con la herramienta *Define Projection* se consigue adecuar los datos a un marco geográfico. Con *Ráster Calculator* se utiliza *Set Null* para establecer nulos los valores por debajo del umbral acordado y posteriormente multiplicarlos por cien y redondearlos, iterando el proceso para todos los ráster susceptibles de esta operación (Figura 10).

Además se clasifican sus valores de cada capa en tres clases siguiendo el método de cortes naturales de Jenks, que busca la clasificación de los datos en función de los puntos de valle del histograma (<http://support.esri.com/other-resources/gis-dictionary>, 2016)

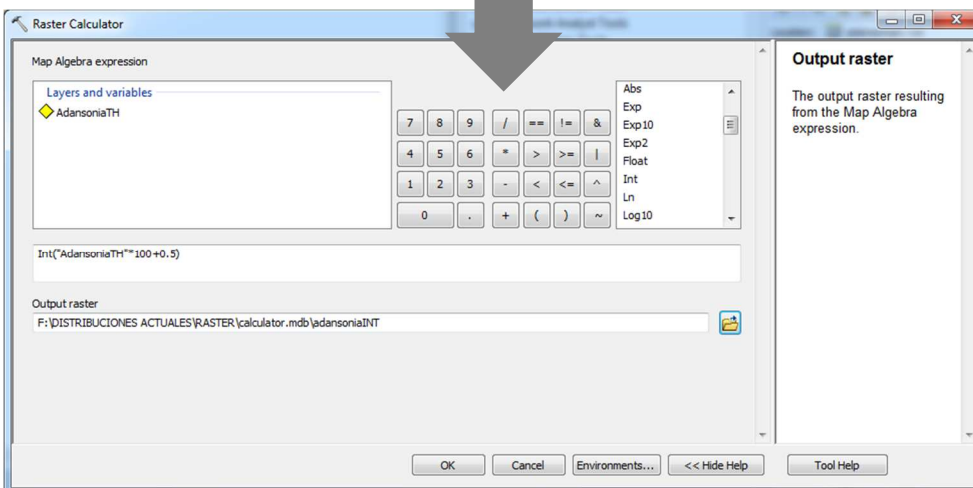
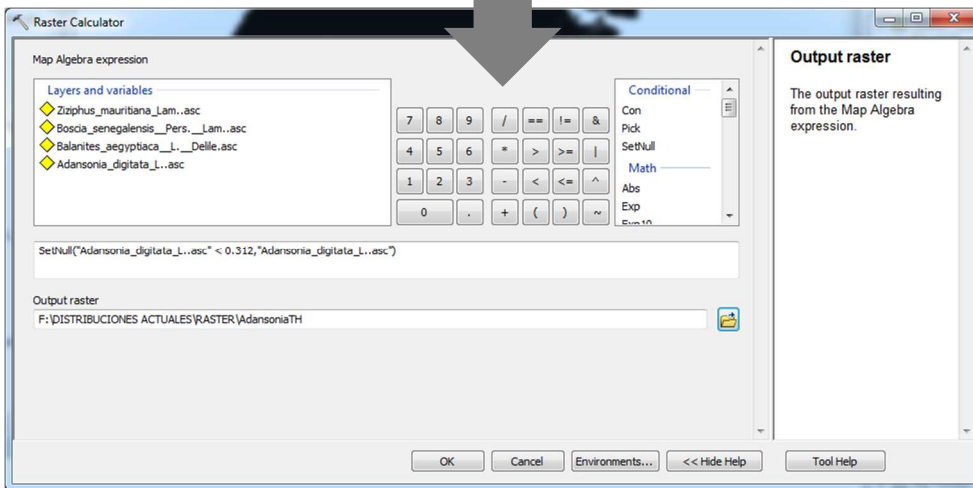
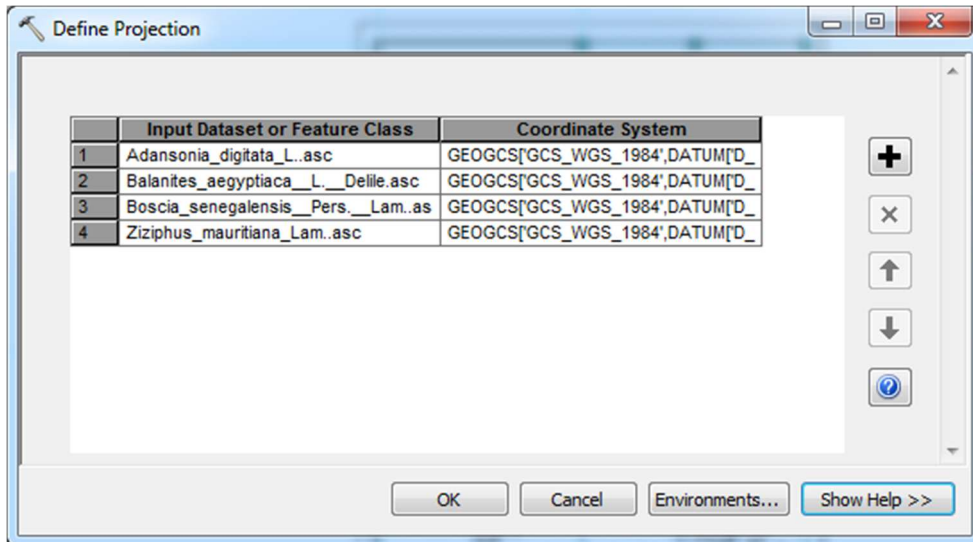


Figura 10. Flujo de representación gráfica de ráster

4. RESULTADOS.

Con el fin de comprobar la objetividad de los resultados se realiza un proceso de evaluación del modelo con una serie de estadísticos que el propio Maxent ofrece. El modelado se considera un problema de clasificación de instancias, tanto positivas como negativas, un clasificador asigna un valor real a cada instancia, a la cual se le aplica un umbral para predecir la pertenencia a una clase {Y, N}. La *sensibilidad* de un clasificador para un umbral determinado es la parte de todas las instancias positivas clasificadas como Y, mientras que la *especificidad* es la parte de todas las instancias negativas que se han clasificado N. La *sensibilidad* es también conocida como el índice verdadero positivo y representa la ausencia de error de omisión, *1-especificidad* es el índice falso positivo y representa el error de comisión. (Phillips et al., 2005)

La curva de la característica operativa del receptor (ROC) se obtiene representando *1-especificidad* en el eje *x* y la *sensibilidad* en el eje de las *y*, para todos los posibles umbrales. Las ausencias son la fuente de las instancias negativas, con las que se mide la *especificidad*, no es posible construir una curva ROC cuando sólo se disponen de presencias. Este caso es el que se nos presenta en este estudio, para solucionarlo se utilizó un método de clasificación diferente, en vez de presencias/ ausencias, se tomaron presencias/ pseudo-ausencias.

En una predicción continua como la que se define en este modelo, la curva ROC contiene un punto para cada instancia, el área bajo la curva (AUC) se determina uniendo estos puntos con líneas rectas. El AUC se interpreta como la probabilidad de que las instancias positivas y negativas se hayan clasificado de forma correcta y es un buen indicador del grado de bondad del modelo (Phillips et al., 2005). En muestras de datos no sesgadas un valor de AUC alto indica que los lugares con mayor idoneidad predicha tienden a ser áreas de presencia conocida y las localizaciones con predicciones del modelo bajas tienden a ser áreas donde no se conoce la presencia de las especies (Hijmans et al., 2016).

Una medida adicional de la conducta del modelo es la ganancia de entrenamiento regularizada (*Gain*) que describe la adecuación de los datos de presencia de la distribución comparados con una distribución uniforme (Gormley et al., 2011). La ganancia se define como la probabilidad logarítmica media e indica la proximidad del modelo alrededor de la muestra de presencias, si la ganancia es igual a 2, significa que la probabilidad media de las muestras de presencia es $\exp(2)$, aproximadamente 7.4 veces mayor que un píxel aleatorio de pseudo-ausencia (Phillips, 2006).

4.1. Modelación de la distribución de las especies.

Todas las especies tienen un valor de AUC cercano a 1, lo que permite deducir que el modelo ha conseguido agrupar de manera eficiente las presencias y pseudo-ausencias de cada especie, en especial de *Boscia senegalensis*, que posee el valor más alto de entre las cuatro seguida por *Ziziphus mauritiana*, *Balanites aegyptiaca* y *Adansonia digitata* respectivamente, en cuanto a la ganancia de entrenamiento regularizada el valor más alto lo obtuvo de nuevo *B. senegalensis* con un valor de 1.603, cinco veces mayor que un punto de *background*. En general el estadístico explicativo ganancia no bajó de 1.358, lo que da idea de lo óptimo que resulta el modelo prediciendo las distribuciones (Figura 11).

Cabe destacar que hay variables climáticas que influyen más que otras en la predicción de las distribuciones, por ejemplo para *A. digitata* la precipitación del mes más seco y la temperatura media del cuarto más cálido fueron las variables que más contribuyeron (74%) a la explicación de la distribución, algo entendible si se tiene en cuenta que es una especie que puede soportar ambientes muy secos y restringida en zonas de heladas (Assogbadjo y Loo, 2011).

La temperatura media anual y la estacionalidad de la temperatura (desviación estándar*100) fueron las variables que más peso tuvieron en la creación de la distribución de *B. aegyptiaca* con un 71.4 % del total, coincidiendo con las características de pervivencia de la planta; puede crecer en condiciones diversas pero sobre todo produce frutos incluso en épocas de sequía. (http://www.worldagroforestry.org/treedb/AFTPDFS/Balanites_aegyptiaca.PDF, 2016).

B. senegalensis constituye un recurso muy valorado contra la hambruna debido a su habilidad de crecer en desiertos con unas condiciones extremadamente secas y cálidas (Khalafalla et al., 2011). Esto explica que las variables que más representen esta especie sean la máxima temperatura del mes más cálido y rango anual de temperatura con un 59.2%, un porcentaje menor que las dos primeras variables con más contribución en las anteriores especies. La temperatura media anual gana protagonismo en este modelado y representa un 20.5% del total.

En cambio *Z. mauritiana* la precipitación del cuarto más húmedo, la estacionalidad de la temperatura (desviación estándar*100) y la precipitación anual explican el 75.9% de su distribución, ya que se desarrolla en zonas semiáridas con periodos de lluvias aislados.

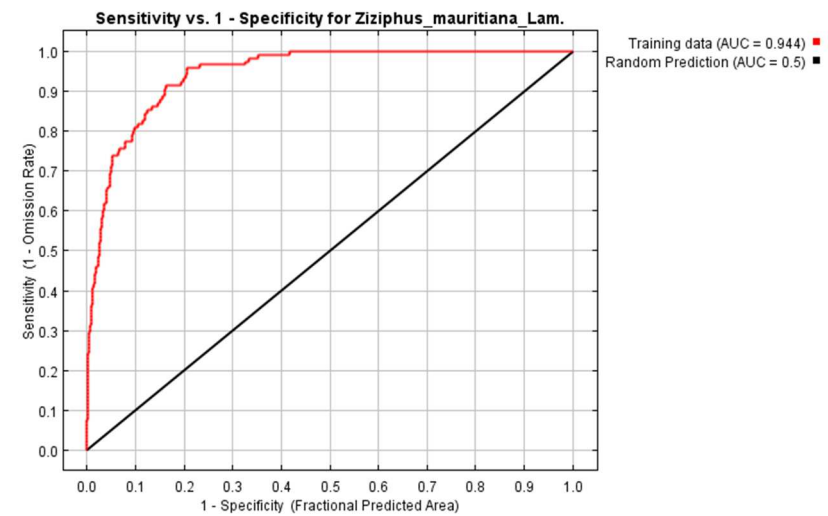
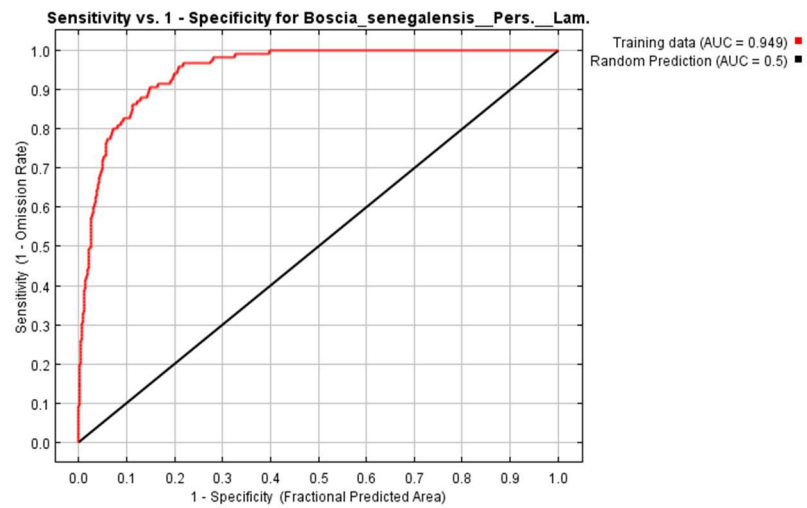
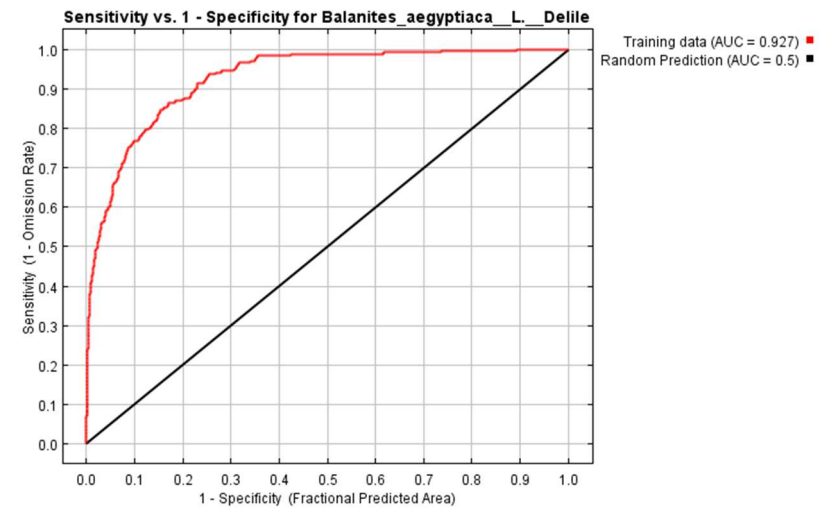
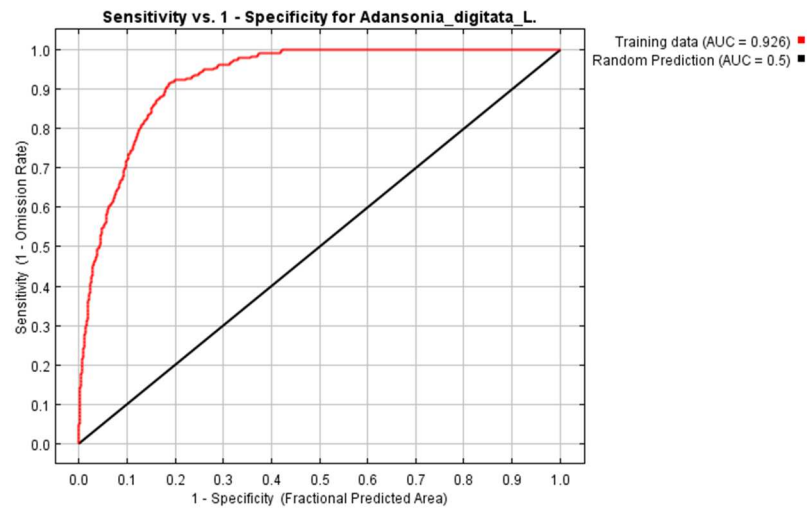


Figura 11. Curvas ROC de la modelación de las especies

La figura 12 muestra que todas las especies tienen en común el centro continental propicio para su desarrollo, países como Senegal, Gambia, Guinea-Bissau, Costa de Marfil, Burkina Faso, Mali, Ghana, Togo, Benín, Nigeria, Níger y Chad aglutinan la mayor confluencia entre especies, coincidiendo con zonas de pastizal tropical, sabana, zonas con matorral y zonas semidesérticas.

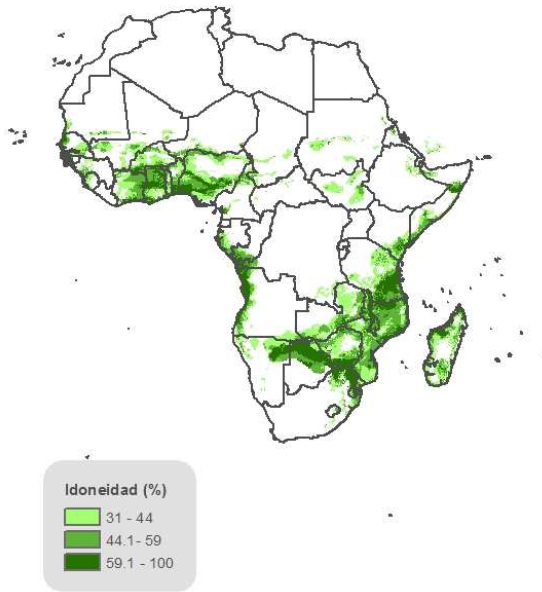
En el análisis especie por especie a nivel del conjunto del continente africano, las distribuciones varían, *A. digitata* crecería en la zona oeste del continente, Gabón, Congo, Angola, sureste, Botsuana, Zimbabue, Mozambique, Malawi, Tanzania y la isla de Madagascar, en ecosistemas de bosques tropicales secos y pastizales tropicales.

B. aegyptiaca comparte distribución con el *A. digitata* en todas sus macro zonas excepto la isla de Madagascar, pero además se desarrolla en la parte este/ sur-este, Sudán, Sudán del Sur, Etiopía, Somalia, Kenia, Uganda, Ruanda Burundi, Malawi, Mozambique, Botsuana y Sudáfrica, aunque con una probabilidad de idoneidad más baja que en la zona oeste-central.

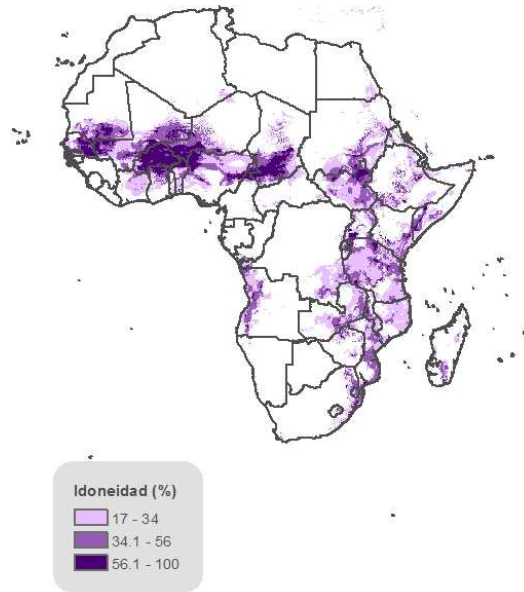
B. senegalensis es un caso particular, las zonas más idóneas de su distribución probable se dan en el centro continental, en regiones desérticas o semidesérticas y de pastizal tropical, el resto se dispersa por el este en zonas aisladas de Mozambique, Kenia, Somalia y Etiopía.

Después de *A. digitata* y *B. aegyptiaca*, *Z. mauritina* es la especie con una distribución más dispersa, se encuentra en el centro del continente, oeste, este e isla de Madagascar al igual que *A. digitata* exceptuando Botsuana, Sudáfrica, Mozambique y Malawi, sus ecosistemas de desarrollo son el bosque tropical estacionalmente seco, regiones semidesérticas, pastizales tropicales y bosque tropical.

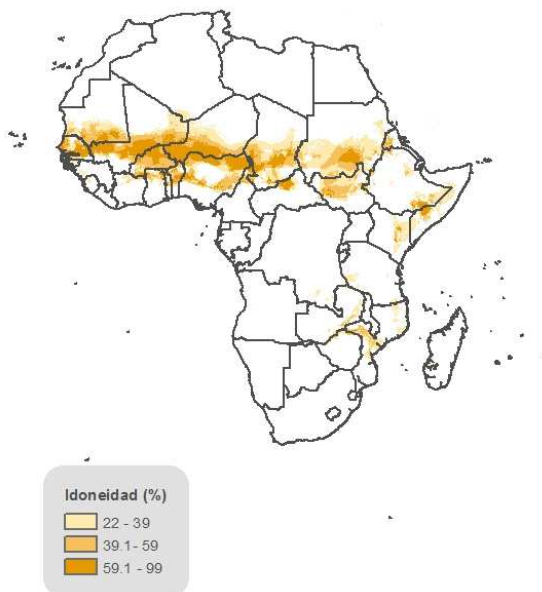
Adansonia digitata



Balanites aegyptiaca



Boscia senegalensis



Ziziphus mauritiana

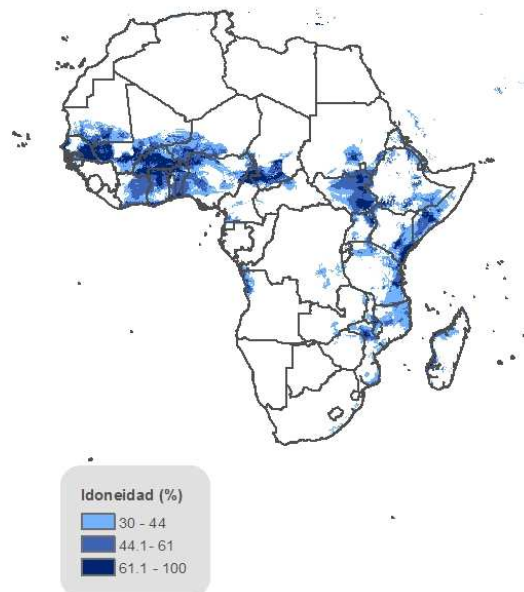


Figura 12. Distribuciones de las cuatro especies

4.2. Modelación del impacto del cambio climático en la distribución de especies.

El modelo creado sirve para proyectar predicciones de las distribuciones con las condiciones climáticas esperadas del futuro, en este estudio hasta 2050, proporcionando información valiosa sobre el impacto del cambio climático en la configuración de los nuevos hábitats de las especies. La calibración del modelo de nicho ecológico realizada para conseguir la modelización de las especies en el espacio temporal actual sirve para las proyecciones de cambio climático por lo que los valores de los estadísticos evaluadores del modelo, AUC (Figura 10), ganancia de entrenamiento regularizada (Gain) y contribución de las variables son los mismos.

Las distribuciones resultado de la proyección futura dan resultados diversos en función de la especie objeto de análisis. *A. digitata* se predice en zonas del desierto del Sahara y la península arábiga, hecho contradictorio con las estimaciones de aumento de temperaturas, aunque, si bien es cierto, se prevé que la precipitación aumente en zonas de la región (www.ipcc.ch/, 2016). En una posición opuesta se encuentra la predicción de *B. aegyptiaca*, que estima la persistencia de la especie en todo el continente menos en la parte noreste. *B. senegalensis* en cambio tendría una mayor adaptabilidad en regiones no desérticas aunque evita los bosques tropicales del centro del continente. El norte de África, Sahel y sur continental son regiones restringidas para *Z. mauritiana*, evitando los pastizales que las abundan, en cambio el resto del continente africano sería un óptimo hábitat para la especie (Figuras 13, 14, 15 y 16).

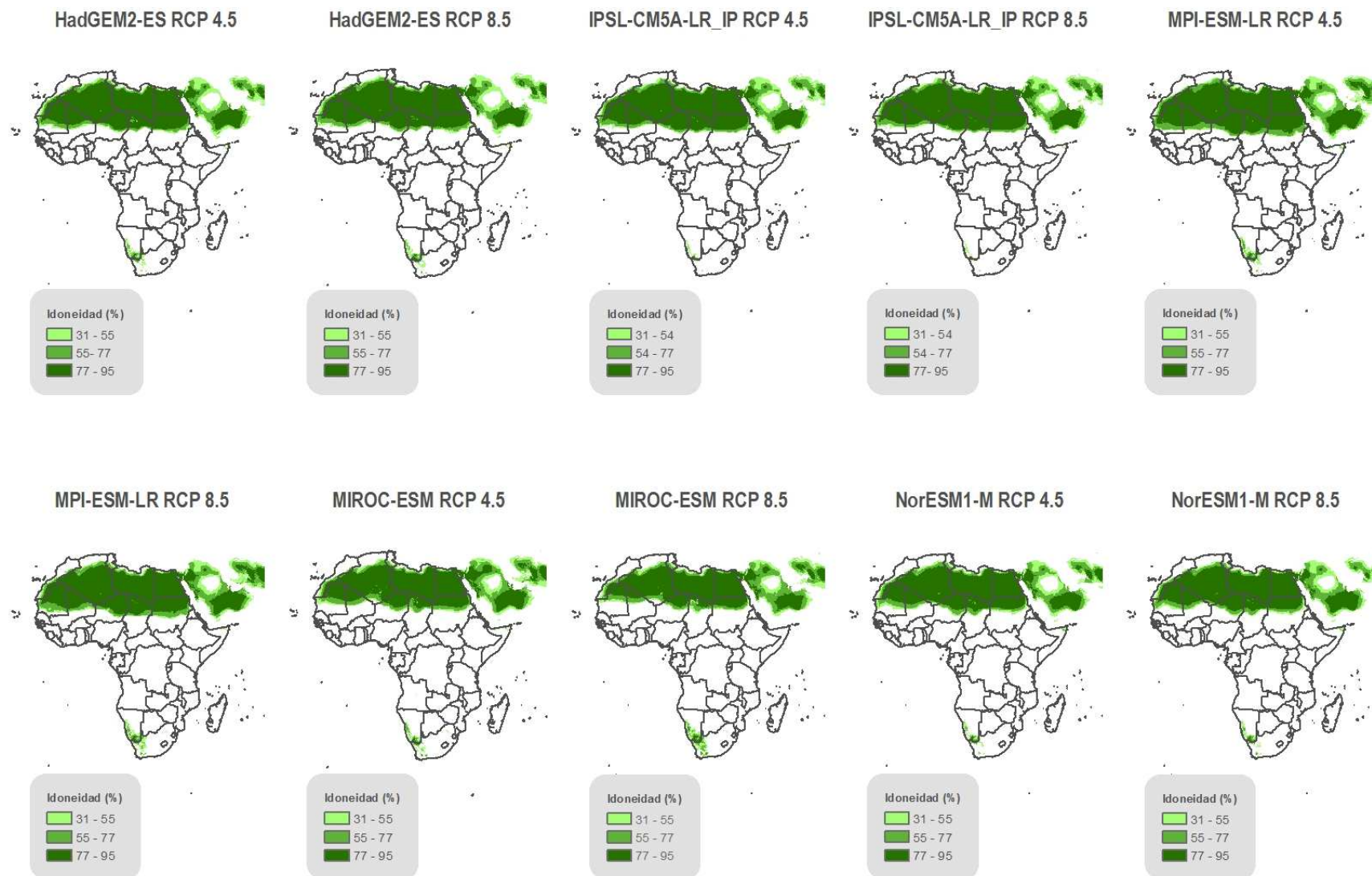


Figura 13. Evaluación del impacto del cambio climático para *Adansonia digitata* con cinco GCMs.

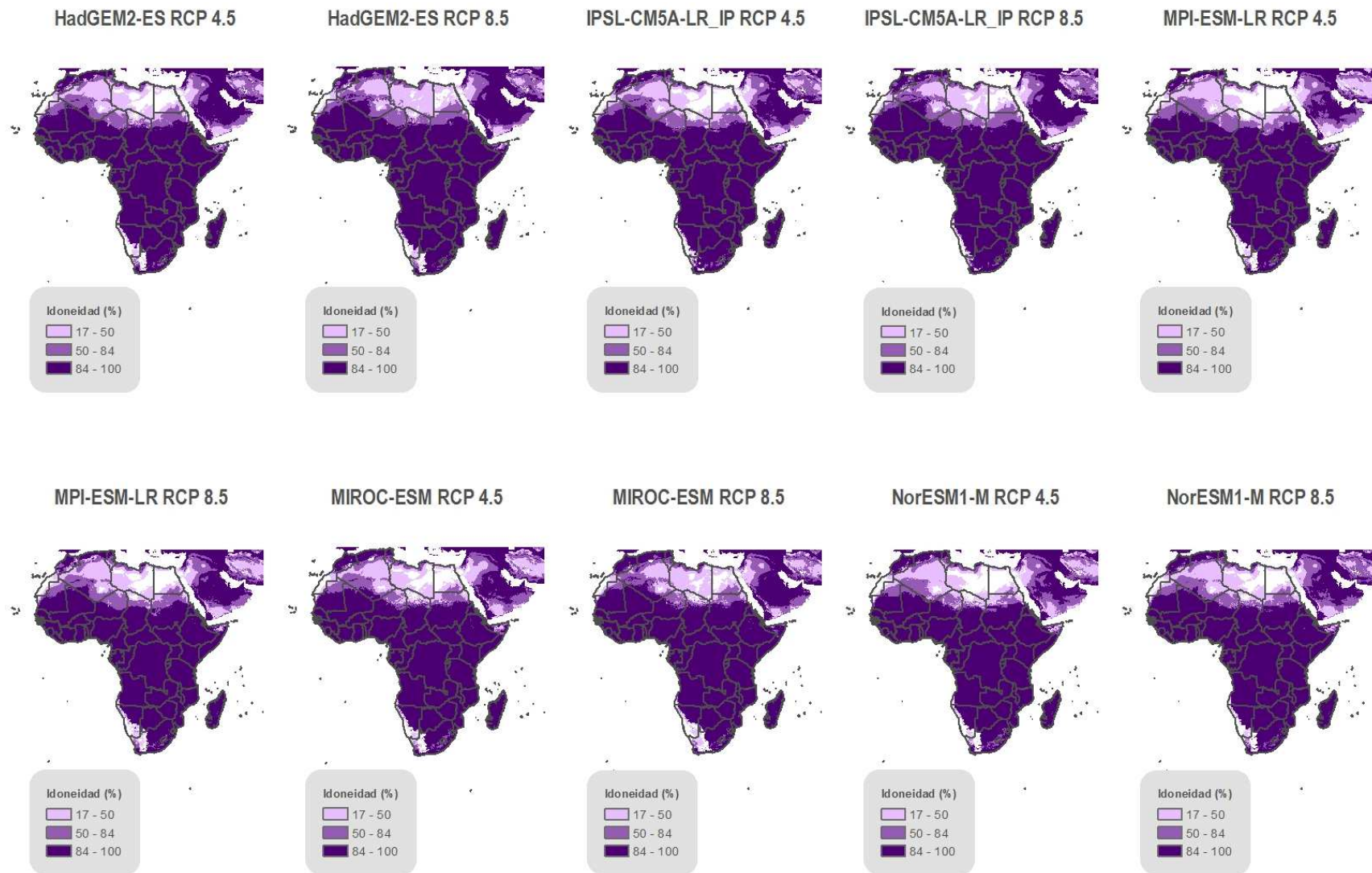


Figura 14. Evaluación del impacto del cambio climático para *Balanites aegyptiaca* con cinco GCMs.

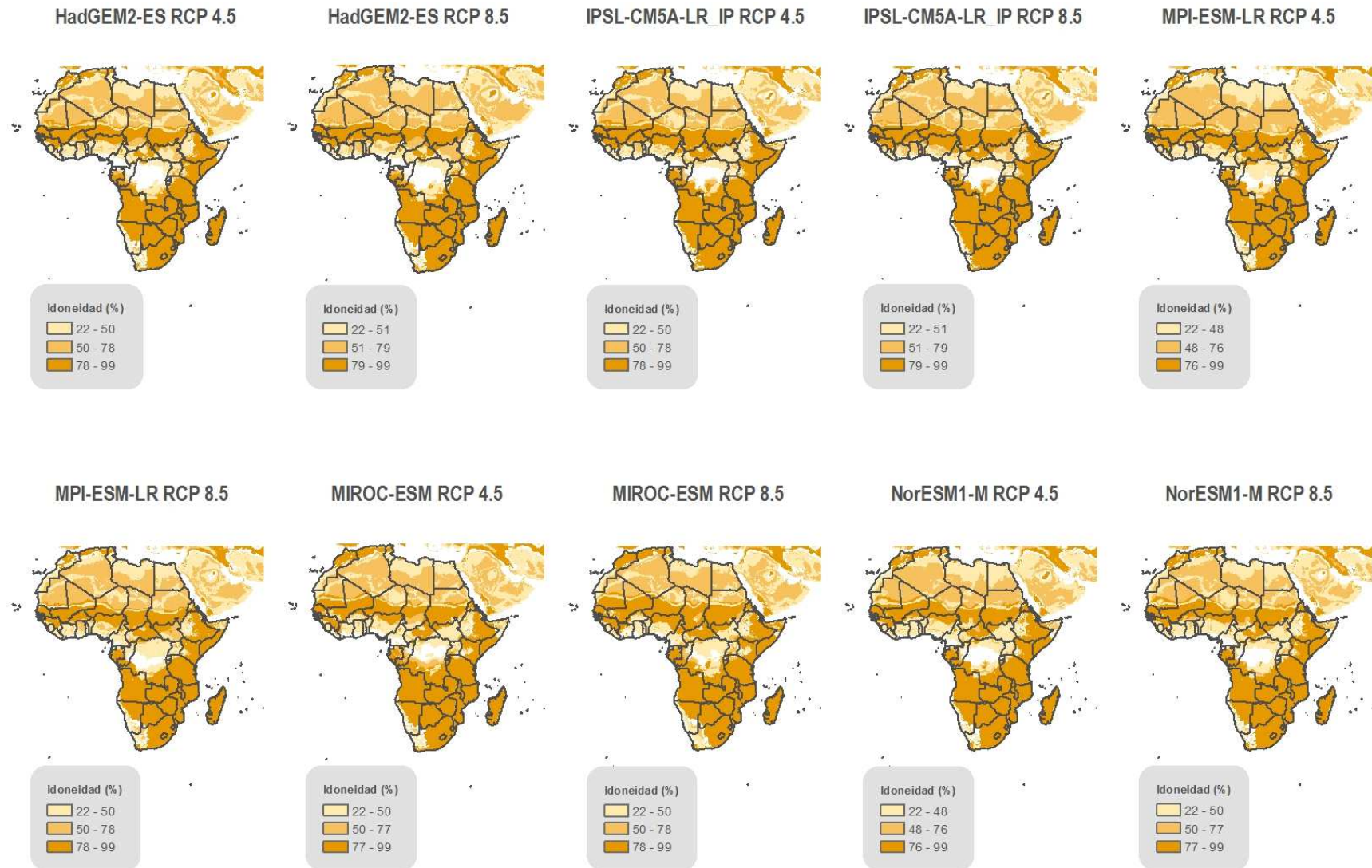


Figura 15. Evaluación del impacto del cambio climático para *Boscia senegalensis* con cinco GCMs.

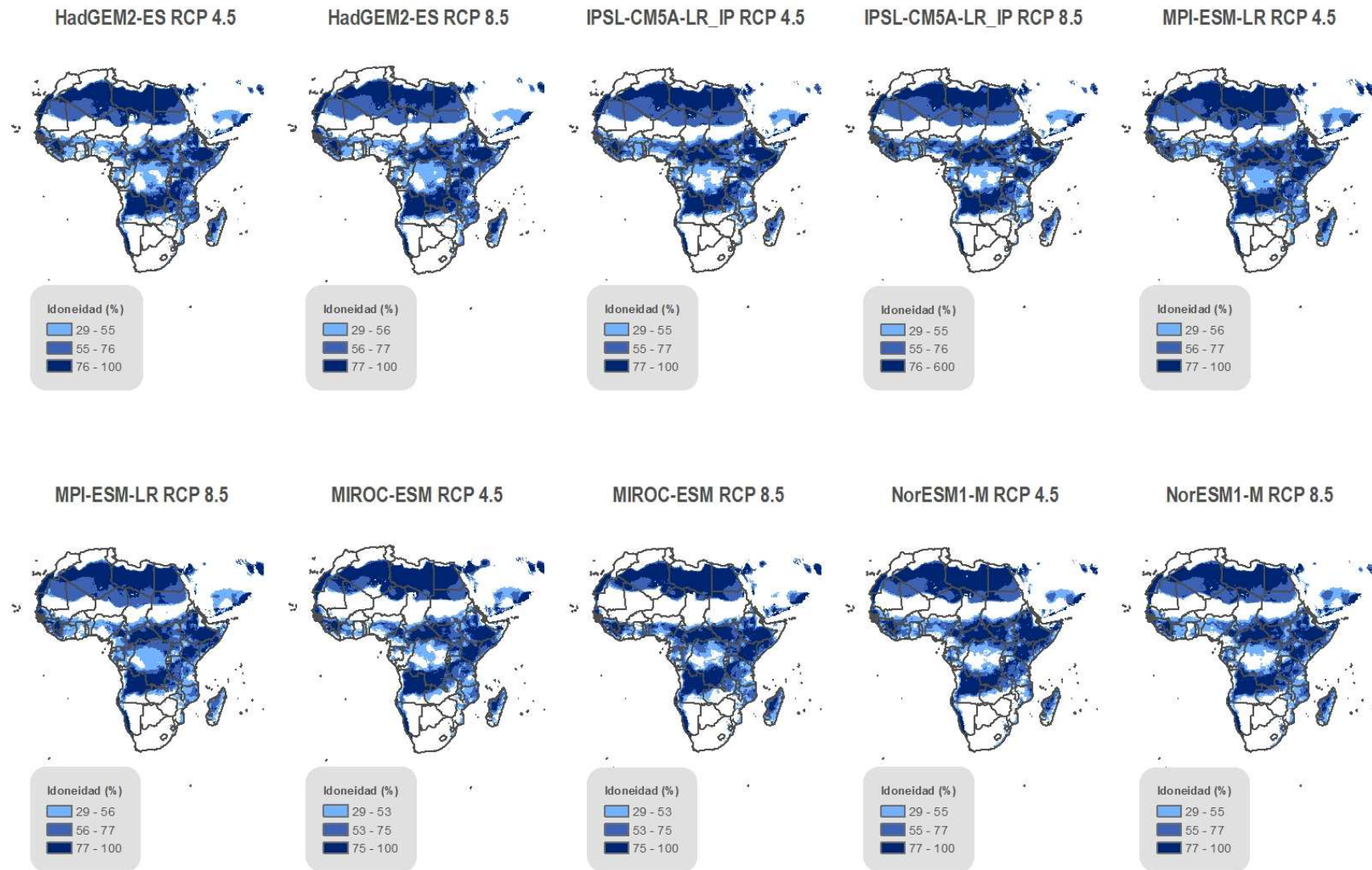


Figura 16. Evaluación del impacto del cambio climático para *Ziziphus mauritiana* con cinco GCMs.

5. CONCLUSIONES.

Este documento propone la evaluación de las distribuciones de especies en la actualidad y el impacto del cambio climático en ellas, con vistas a determinar las estrategias de conservación más adecuadas. El modelo generado se ha calibrado con las presencias registradas en la base de datos de GBIF y el conocimiento de expertos. Se demostró un desempeño óptimo del mismo a la hora de determinar las áreas más idóneas en el periodo actual, sin embargo las predicciones del impacto establecen unas distribuciones con alta incertidumbre.

El cambio climático tiene consecuencias en la presencia de la biodiversidad, en este estudio se revela un gran impacto, con el aumento en todos los casos de estudio de la superficie idónea para el desarrollo de las especies incluso en zonas en las que en la actualidad sus condiciones extremas no lo permiten. Según el modelo, la supervivencia de todas las especies estaría asegurada hasta el año 2050 en el continente africano, pero el interés de esta evaluación radica en conocer regiones concretas de presencia que ayuden al desarrollo de las comunidades.

Se subraya el potencial de esta estrategia para evaluar la seguridad alimentaria global y regional, las restricciones climáticas y el cambio de hábitat de especies en el contexto del cambio climático (Ramírez-Villegas et al., 2013).

Con motivo de difundir los resultados obtenidos, se creó una aplicación web que permite visualizar la idoneidad de cada especie en lugares concretos, comparar las distribuciones entre especies y conocer las características más reseñables de cada una de ellas. (josemaciasbarredo.neocities.org).

6. BIBLIOGRAFÍA.

ASSOGBADJO A.E.; LOO J. *Adansonia digitata*, African baobab. Conservation and Sustainable Use of Genetic Resources of Priority Food Tree Species in sub-Saharan Africa. 2011 Bioversity International (Rome, Italy).

ARAYA A.; HOOGENBOOM G.; LUEDELING E.; HADGU KIROS M.; KISEKKA ISAYA; MARTORANO LUCIETA G. Assessment of maize growth and yield using crop models under present and future climate in southwestern Ethiopia. 2015. *Agricultural and Forest Meteorology*. Volumes 214–215, Pages 252–265, doi: 10.1016/j.agrformet.2015.08.259

BENTSEN M. ; BETHKE I. ; DEBERNARD J. B. ; IVERSEN T. ; KIRKEVÅG A. ; SELAND Ø.; DRANGE H. ; ROELANDT C. ; SEIERSTAD I. A.; HOOSE C. ; KRISTJÁNSSON J. E. The Norwegian Earth System Model, NorESM1-M – Part 1: Description and basic evaluation of the physical climate. *Geosci. Model Dev.*, 6, 687-720, doi: 10.5194/gmd-6-687-2013

CAMERON A. (a); PHILLIPS S. (b); GOLDEN C. (a); HIJMANS R. J. (c); LOUIS E. (d); WRIGHT P. C. (e); KREMEN C. (a). Identifying corridors for Lemur dispersal, for two future climate change scenarios. 2015. (a) Department of Environmental Sciences, Policy and Management, 137 Mulford Hall, University of California, Berkeley, CA 94720-3114, USA.(b) AT&T Labs-Research, 180 Park Avenue, Florham Park, NJ 07932, USA. (c) Department of Environmental Science and Policy, University of California, Davis Agriculture, Biogeography, Ecological modeling, Health, Spatial data science. (d) Center for Conservation and Research, Henry Doorly Zoo, Omaha, NE, 68107, USA. (e) Department of Anthropology, State University of New York, Stony Brook, NY, 11794, USA.

CHAALALIA AURÉLIE; BEAUGRAND GRÉGORY; RAYBAUD VIRGINIE; LASSALLE GÉRALDINE; SAINT-BÉATA BLANCHE; LE LOC'H FRANÇOIS; BOPP LAURENT; TECCHIO SAMUELE; SAFI GEORGES; CHIFFLET MARINA; LOBRY JÉRÉMY; NIQUIL NATHALIE. From species distributions to ecosystem structure and function: a methodological perspective. 2016. *Ecological Modelling*. Volume 334, Pages 78–90, doi: 10.1016/j.ecolmodel.2016.04.022

CHOU SIN CHAN; LYRA ANDRÉ; MOURÃO CAROLINE; DEREZYNSKI CLAUDINE; PILOTTO ISABEL; GOMES JORGE; BUSTAMANTE JOSIANE; TAVARES PRISCILA; SILVA ADAN; RODRIGUES DANIELA; CAMPOS DIEGO; CHAGAS DIEGO; SUEIRO GUSTAVO; SIQUEIRA GRACIELLE; MARENGO JOSÉ. Assessment of Climate Change over South America under RCP 4.5 and 8.5 Downscaling Scenarios. 2014. *American Journal of Climate Change*, 3, 512-527, doi: 10.4236/ajcc.2014.35043.

DE CAUWER VERA; BART MUYS; RASMUS REVERMANN; ANTONIO TRABUCCO. Potential, realised, future distribution and environmental suitability for *Pterocarpus angolensis* DC in southern Africa. 2014. *Forest Ecology and Management*. Volume 315, 1 March 2014, Pages 211–226, doi: 10.1016/j.foreco.2013.12.032.

ESCALANTE TANIA; RODRÍGUEZ-TAPIA GERARDO; LINAJE MIGUEL; ILLOLDI-RANGEL PATRICIA; GONZÁLEZ-LÓPEZ RAFAEL. Identification of areas of endemism from species distribution models: threshold selection and Nearctic mammals. 2013. *TIP*. Volume 16, Issue 1, 2013, Pages 5–17, doi: 10.1016/S1405-888X(13)72073-4.

FLATO, GREGORY M. Earth system models: an overview *WIREs Clim Change* 2011, 2:783–800. doi: 10.1002/wcc.148.

GOMES VALE CANDIDA; BRITO JOSÉ CARLOS. Desert-adapted species are vulnerable to climate change: Insights from the warmest region on Earth. 2015. *Global Ecology and Conservation*. Volume 4, Pages 369–379, doi: 10.1016/j.gecco.2015.07.012.

GORMLEY ANDREW M.; FORSYTH DAVID M.; GRIFFIOEN PETER; LINDEMAN MICHAEL; RAMSEY DAVID S.L.; SCROGGIE MICHAEL P. ; WOODFORD LUKE. Using presence-only and presence–absence data to estimate the current and potential distributions of established invasive species. 2010. *Journal of Applied Ecology*, doi: 10.1111/j.1365-2664.2010.01911.x

HIJMANS ROBERT J.; ELITH JANE Species distribution modeling with R. 2016. (a) Department of Environmental Science and Policy, University of California, Davis Agriculture, Biogeography, Ecological modeling, Health, Spatial data science. (b) School of BioSciences, Building 122. The University of Melbourne, Victoria 3010 Australia.

HOVEKA L.N.; BEZENG B.S.; YESSOUFOU K.; BOATWRIGHT J.S.; VAN DER BANK M. Effects of climate change on the future distributions of the top five freshwater invasive plants in South Africa. 2015. *South African Journal of Botany*. Volume 102, January 2016, Pages 33–38, doi:10.1016/j.sajb.2015.07.017

KALINGANIRE A.; KONÉ B. *Ziziphus mauritiana*, ber. Conservation and Sustainable Use of Genetic Resources of Priority Food Tree Species in sub-Saharan Africa. 2011 *Bioversity International* (Rome, Italy).

KHALAFALLA, M. M.; DAFFALLA, H. M.; ABDELLATEF, E.; AGABNA, E.; EL-SHEMY, H. A. Establishment of an in vitro micropropagation protocol for *Boscia senegalensis* (Pers.) Lam. Ex Poir. 2011. *Journal of Zhejiang University Science Biomedicine & Biotechnology* 12 (4): 303–312, doi:10.1631/jzus.B1000205. PMC 3072594. PMID 21462387.

KARUNARATNE A.S.; WALKER S.; AZAM-ALI S.N. Assessing the productivity and resource-use efficiency of underutilised crops: Towards an integrative system. 2014. *Agricultural Water Management*. Volume 147, Pages 129–134, doi: 10.1016/j.agwat.2014.08.002

KHANUM RIZWANA; MUMTAZ A.S.; KUMAR SUNIL. Predicting impacts of climate change on medicinal asclepiads of Pakistan using Maxent modeling. 2013. *Acta Oecologica*. Volume 49, May 2013, Pages 23–3, doi: 10.1016/j.actao.2013.02.007

MCSWEENEY CAROL F.; JONES RICHARD G. How representative is the spread of climate projections from the 5 CMIP5 GCMs used in ISI-MIP? 2016. *Climate Services*. Volume 1, Pages 24–29, doi: 10.1016/j.cliser.2016.02.001

MICHEL DÉQUÉ; SANDRO CALMANTI; OLE BØSSING CHRISTENSEN; ALESSANDRO DELL AQUILA; CATHRINE FOX MAULE; ANDREAS HAENSLER; GRIGORY NIKULIN; CLAAS TEICHMANN. A multi-model climate response over tropical Africa at +2 °C. 2016. *Climate Services*, doi: 10.1016/j.cliser.2016.06.002

MORENO ROBERTO; ZAMORA A RICARDO; MOLINA JUAN RAMÓN; VASQUEZ ANGÉLICA; HERRERA MIGUEL ÁNGEL. Predictive modeling of microhabitats for endemic birds in South Chilean temperate forests using Maximum entropy (Maxent). 2011. *Ecological Informatics*. Volume 6, Issue 6, Pages 364–370, doi: 10.1016/j.ecoinf.2011.07.003

PHILLIPS STEVEN J.; ANDERSON ROBERT P.; SCHAPIRE ROBERT E. Maximum entropy modeling of species geographic distributions. 2005. *Ecological Modelling*. Volume 190, Issues 3–4, 25, Pages 231–259, doi: 10.1016/j.ecolmodel.2005.03.026

PHILLIPS STEVEN J.; DUDÍK MIROSLAV. Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. 2008. *Ecography*, doi: 10.1111/j.0906-7590.2008.5203.x

PHILLIPS STEVEN J.; DUDÍK MIROSLAV; SCHAPIRE ROBERT E. A Maximum Entropy Approach to Species Distribution Modeling. 2004. *Proceedings of the Twenty-First International Conference on Machine Learning*, pages 655-662.

QUINN GERRY P.; KEOUGH MICHAEL J. *Experimental Design and Data Analysis for Biologists*. 2002. ISBN-13 978-0-511-07812-5. Cambridge University Press. The Edinburgh Building, Cambridge, United Kingdom.

RAMIREZ-VILLEGAS JULIAN; JARVIS ANDY; LÄDERACHD PETER. Empirical approaches for assessing impacts of climate change on agriculture: The EcoCrop model and a case study with grain sorghum. 2013. *Agricultural and Forest Meteorology* 170 67–78.

REDON M.; LUQUE S. Presence-only modelling for indicator species distribution: biodiversity monitoring in the French Alps. 6th Spatial Analysis and Geomatics international conference (SAGEO 2010). Nov 2010. Université de Toulouse, Toulouse, France. 1, p. 42-p. 55. <hal-00558859>

SCHELDEMAN, XAVIER; VAN ZONNEVELD MAARTEN. 2011. *Manual de Capacitación en Análisis Espacial de Diversidad y Distribución de Plantas*. Bioversity International, Roma, Italia. 186 pp. ISBN 978-92-9043-908-0

VAN VUUREN DETLEF P.; EDMONDS JAE; MIKIKO KAINUMA; RIAHI KEYWAN; THOMSON ALLISON; HIBBARD KATHY; HURTT GEORGE C.; KRAM TOM; KREY VOLKER; LAMARQUE JEAN-FRANCOIS; MASUI TOSHIHIKO; MEINSHAUSEN MALTE; NAKICENOVIC NEBOJSA; SMITH STEVEN J.; ROSE STEVEN K. The representative concentration pathways: an overview. 2011. *Climatic Change*, 109: 5, doi: 10.1007/s10584-011-0148-z

VELÁSQUEZ-TIBATÁ JORGE; SALAMAN PAUL; GRAHAM CATHERINE H. Effects of climate change on species distribution, community structure, and conservation of birds in protected areas in Colombia. 2012. *Reg Environ Change*, doi: 10.1007/s10113-012-0329-y

WATANABE S.; HAJIMA T.; SUDO K.; NAGASHIMA T.; TAKEMURA T.; OKAJIMA H.; NOZAWA T.; KAWASE H.; ABE M.; YOKOHATA T.; ISE T.; SATO H.; KATO E.; TAKATA K.; EMORI S.; KAWAMIYA M. MIROC-ESM 2010: model description and basic results of CMIP5-20c3m experiments. 2011. *Japan. Geosci. Model Dev.*, 4, 845-872, doi: 10.5194/gmd-4-845-2011

WEBER JOHN C.; SOTELO CARMEN. Correlations and clines in tree growth and wood density of *Balanites aegyptiaca* (L.) Delile provenances in Niger. 2009. *New Forests* 39:39–49, doi: 10.1007/s11056-009-9153-8

Balanites aegyptiaca Report [en línea]:

< http://www.worldagroforestry.org/treedb/AFTPDFS/Balanites_aegyptiaca.PDF>

[Consulta: agosto de 2016].

Código Python para la eliminación de columnas vacías [en línea]:

< <https://gist.github.com/aculich/fb2769414850d20911eb>>

[Consulta: abril de 2016].

Global Climate Data [en línea]:

< <http://www.worldclim.org/>>

[Consulta: mayo de 2016].

Grupo Intergubernamental de Expertos sobre el Cambio Climático (IPCC) [en línea]:

< <http://www.ipcc.ch/>>

[Consulta: agosto de 2016].

IPSL Earth System Model for the 5th IPCC report [en línea]:

< <https://verc.enes.org/models/earthsystem-models/ipsl/ipslesm>>

[Consulta: agosto de 2016].

Límites de los países del mundo [en línea]:

< <https://www.arcgis.com/home/item.html?id=3864c63872d84aec91933618e3815dd2>>

[Consulta: agosto de 2016].

Manuales y descarga del software estadístico R [en línea]:

< <https://cran.r-project.org>>

[Consulta: agosto de 2016].

Met Office Hadley Centre ESM [en línea]:

<<https://verc.enes.org/models/earthsystem-models/metoffice-hadley-centre/hadgem2-es>>

[Consulta: agosto de 2016].

New Earth system model of Max Planck Institute for Meteorology [en línea]:

< <http://www.mpimet.mpg.de/en/science/models/mpi-esm.html>>

[Consulta: agosto de 2016].

Norwegian Earth System Model [en línea]:

< <https://verc.enes.org/models/earthsystem-models/ncc/noresm>>

[Consulta: agosto de 2016].

Software de Sistemas de Información Geográfica. Definiciones [en línea]:

< <http://support.esri.com/other-resources/gis-dictionary>>

[Consulta: mayo de 2016].

Software de Sistemas de Información Geográfica. Descarga de datos [en línea]:

< <https://www.arcgis.com/home/index.html/>>

[Consulta: mayo de 2016].

Stockholm Environment Institute [en línea]:

< <https://www.sei-international.org/>>

[Consulta: agosto de 2016].

AGRADECIMIENTOS

El autor agradece a su familia todo el apoyo recibido durante la realización de este proyecto, el valor de la perseverancia inculcado ha sido clave en la consecución de este documento.

Quisiera agradecer al Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria (INIA) por ofrecerme la oportunidad de participar en un proyecto tan apasionante como éste y en especial a María Jesús Serra, María Dolores Agúndez Leal y Ricardo Alía por sus buenos consejos y alentadoras palabras.

Una mención especial a Javier Pórtoles de la Fundación para la Investigación del Clima, por sus comentarios constructivos y su predisposición a compartir sus conocimientos.

Este documento ha sido realizado en colaboración con la Agencia Española de Cooperación Internacional para el Desarrollo (AECID), en el marco de la Acción de Cooperación para el desarrollo 2014/ACDE/003520 “Acceso, conservación, gestión y valorización de especies leñosas alimenticias en Níger”. Su contenido es responsabilidad exclusiva de INIA y no refleja necesariamente la opinión de la AECID.

ANEXO

```
#####ANÁLISIS DE CORRELACIÓN y COLINEARIDAD#####
```

```
library(HH)
```

```
# Análisis de correlación
```

```
variables_tabla<-read.csv("C:/Users/acer/Documents/TFM/ESTADISTICOS/RESULTADOS_CORRELACION/aleatorio.csv")
```

```
variables_correlacion<-cor(variables_tabla) # cálculo de las correlaciones entre variables
```

```
variables_dist<-abs(as.dist(variables_correlacion)) #converting negative values into their absolute values and calculating the distance among variables
```

```
variables_cluster<-hclust(1-variables_dist) # cuanto más pequeña sea la distancia mayor será la correlación , por lo que se usa el 1-distancia para construir el dendrograma
```

```
plot(variables_cluster, horiz=TRUE)
```

```
cutree(variables_cluster,h=0.25) #la correlación más alta admitida es 0.75 (1-0.75)=0.25
```

```
#Análisis de colinearidad
```

```
variables_tabla<-na.omit(tabla_completa) #omitimos las celdas vacías de la tabla
```

```
variables_tabla<-variables_tabla[sample(nrow(variables_tabla),4431157),]#se trabaja con la mitad de la muestra
```

```
write.csv(variables_tabla, "C:/Users/acer/Documents/TFM/ESTADISTICOS/RESULTADOS_CORRELACION/muestra.csv")
```

```
variables_tabla<-muestra
```

```
selected_variables<-
```

```
data.frame(bio11=variables_tabla$bio11,bio9=variables_tabla$bio9,bio15=variables_tabla$bio15,bio12=variables_tabla$bio12,bio5=variables_tabla$bio5,bio19=variables_tabla$bio19,bio7=variables_tabla$bio7,bio1=variables_tabla$bio1) #Ejemplo para Boscia Senegalensis
```

```
# VIF must be below 10
```

```
vif(selected_variables)
```

```
#####ANÁLISIS DE DEVIANZA#####
```

```
##ANÁLISIS DE DEVIANZA POR CADA VARIABLE DE FORMA INDIVIDUAL (combinado para cada grupo y cada variable)##
```

```
#Lectura de datos
```

```

complete_table<-("C:/Users/acer/Documents/TFM/ESTADISTICOS/RESULTADOS_DEVIANZA/boscia.csv")

#Definición de la función que se usará
vartest0.glm <- function(spp,clmobj)
{
  ql.matrix<-matrix(ncol=1,nrow=dim(clmobj)[2])
  for (i in 1:(dim(clmobj)[2]))
  {
    tmp <- glm(spp ~ clmobj[,i] + I((clmobj[,i])^2), na.action=na.omit, family=binomial)
    ql.matrix[(i),1] <- (1-(tmp$deviance/tmp$null.deviance))
  }
  return(ql.matrix)
}

#EJECUCIÓN DEL ANÁLISIS
test<-vartest0.glm(ziziphus$obs,boscia[, -1])
colnames(test)<-"species"
test

```

*Script facilitado por María Jesús Serra

Estadísticos empleados en la fase de creación del modelo de nicho ecológico.

