

## Informe final\* del Proyecto JE015

### Implementación de herramientas computacionales para la predicción espacial de la abundancia de las especies\*

<b>Responsable:</b>	Dr. Enrique Martínez Meyer
<b>Institución:</b>	Universidad Nacional Autónoma de México Instituto de Biología
<b>Correo electrónico:</b>	emm@ibiologia.unam.mx
<b>Fecha de inicio:</b>	15 de mayo de 2012
<b>Fecha de término:</b>	13 de octubre de 2016
<b>Principales resultados:</b>	Informe final, Cartografía, Base de datos
<b>Forma de citar** el informe final y otros resultados:</b>	Martínez Meyer E., Rojas Soto O., Yáñez Arenas C., Landgrave Ramírez R. y J. J. Giddings Soto. 2016. Implementación de herramientas computacionales para la predicción espacial de la abundancia de las especies. Universidad Nacional Autónoma de México. Instituto de Biología. Informe final SNIB-CONABIO Proyecto No. JE015. Ciudad de México.

#### Resumen:

La abundancia de las especies dentro de sus áreas de distribución presentan patrones geográficos consistentes en los diferentes grupos biológicos. En general, la mayor parte de los sitios de presencia de las especies tienen tamaños poblacionales relativamente pequeños, mientras que existen muy pocos sitios en donde los tamaños poblacionales son mayores en órdenes de magnitud. Aunque desde hace décadas se ha buscado la explicación a estos patrones en las características de los rangos de distribución y de los paisajes, investigaciones recientes desarrolladas por este grupo de trabajo han encontrado que estos patrones se encuentran mejor explicados por las características del nicho ecológico de las especies en el espacio ambiental. Específicamente, descubrimos que las abundancias de las poblaciones se relacionan inversamente con la distancia que guardan con el centroide del nicho ecológico de la especie, en donde teóricamente existen las condiciones óptimas para el desarrollo de las poblaciones. Los modelos de nicho ecológico en su estado actual tienen muy poca o nula capacidad de capturar esta relación y, por lo tanto, los mapas resultantes, incluso de algoritmos que producen modelos probabilísticos, difícilmente se pueden interpretar como mapas que reflejen los patrones de abundancia de las especies. En este proyecto proponemos desarrollar cuatro herramientas computacionales que permitan estimar el centroide del nicho de las especies y a partir del cálculo de las distancias de los puntos de registro, modelar los patrones de abundancia relativa o absoluta de las especies: un programa independiente, una extensión para ArcView, otra para ArcGis y un módulo del programa R. Adicionalmente, proponemos un muestreo sistemático en campo que permita calibrar y validar la modelación de la abundancia de tres especies de cactáceas. Consideramos que esta aportación ofrece un paso adelante en la modelación de la distribución geográfica de las especies.

Centroide del nicho, abundancia, distribución geográfica, centro-abundante, Baja California.

- 
- \* El presente documento no necesariamente contiene los principales resultados del proyecto correspondiente o la descripción de los mismos. Los proyectos apoyados por la CONABIO así como información adicional sobre ellos, pueden consultarse en [www.conabio.gob.mx](http://www.conabio.gob.mx)
  - \*\* El usuario tiene la obligación, de conformidad con el artículo 57 de la LFDA, de citar a los autores de obras individuales, así como a los compiladores. De manera que deberán citarse todos los responsables de los proyectos, que proveyeron datos, así como a la CONABIO como depositaria, compiladora y proveedora de la información. En su caso, el usuario deberá obtener del proveedor la información complementaria sobre la autoría específica de los datos.

**Proyecto JE015: Implementación de herramientas  
computacionales para la predicción espacial de la abundancia de  
las especies**

**INFORME FINAL**

Participantes

*Equipo científico*

Responsable: Dr. Enrique Martínez Meyer – Instituto de Biología, UNAM  
(emm@ib.unam.mx)

Dr. Octavio Rojas Soto – Instituto de Ecología, A.C.

Dr. Carlos Yañez Arenas – Instituto de Ecología, A.C.

*Programadores*

Fís. Rosario Landgrave Ramírez - Instituto de Ecología, A.C.

Lic. John J. Giddings Soto – Programador Independiente

Mayo 2016

## ÍNDICE GENERAL

RESUMEN.....	1
INTRODUCCIÓN.....	1
DESARROLLO DEL PROYECTO.....	5
1. Herramientas para la modelación de abundancia.....	5
a. Preparación de los datos de entrada.....	6
b. Modelación.....	7
c. Generación de resultados.....	7
d. Validación estadística.....	8
2. Trabajo de campo.....	9
3. Modelación de distribución de la abundancia de plantas columnares.....	17
a. Preparación de los insumos.....	17
b. Cálculo de la distancia al centroide.....	21
c. Estimación de la abundancia.....	23
d. Validación estadística.....	27
EL FUTURO.....	30
LITERATURA CITADA.....	30
ANEXOS.....	33
Anexo 1. Instructivo para el uso del programa <i>Nichosfera</i> para Java v2.1.....	34
Anexo 2. Instructivo para el uso de la extensión <i>Nichosfera</i> para ArcView 3.x..	38
Anexo 3. Instructivo para el uso de la extensión <i>Nichosfera</i> para ArcGis 10.2 o superior.....	42
Anexo 4. Instructivo para el uso de <i>Nichosfera</i> para R.....	45
Anexo 5. Regresión y mapeo de incertidumbre por medio de un proceso de <i>bootstrap</i> de la predicción de abundancias por el método de Distancia al Centroides (DC) para R.....	48

## RESUMEN

En este proyecto desarrollamos la primera versión del sistema de modelación geográfica de la distancia al centroide del nicho ecológico para modelar la abundancia de las especies, que denominamos *Nichosfera*. Este enfoque tiene sus bases conceptuales en la teoría del nicho ecológico multidimensional, que propone que la abundancia de una especie está determinada por la estructura interna de su nicho (*sensu* Hutchinson). El método que utiliza *Nichosfera* para el desarrollo de los modelos de abundancia es el de la distancia al centroide del nicho (DCN), que fue conceptualizado y desarrollado por este grupo de trabajo.

Implementamos *Nichosfera* en cuatro herramientas computacionales: un programa en Java, un código de R, una extensión para ArcView 3.x y una extensión para ArcGis 10.2 o superior. Como parte de este proyecto, también realizamos un muestreo a lo largo de toda la Península de Baja California para generar bases de datos de los registros de la presencia, ausencia y abundancia de tres especies de plantas columnares endémicas o cuasi-endémicas a los desiertos de esta región, que fueron: el Cirio (*Fouquieria columnaris*), el Cardón (*Pachycereus pringlei*) y la Pitaya agria (*Stenocereus gummosus*). Finalmente, usando los datos obtenidos en el campo y las herramientas desarrolladas aquí, presentamos un ejercicio de modelación para estas tres especies.

## INTRODUCCIÓN

Una regla universal en la biogeografía es que la distribución de la abundancia de los individuos dentro de sus áreas de presencia no es homogénea (Brown 1984). De hecho, se han identificado patrones geográficos consistentes en los diferentes grupos biológicos, desde microorganismos hasta plantas y vertebrados, por ejemplo: en la mayor parte de los sitios de presencia de cualquier especie los tamaños poblacionales son relativamente pequeños, mientras que existen relativamente pocos sitios en donde la cantidad de individuos es órdenes de magnitud mayor. También se ha observado la existencia de un efecto de autocorrelación espacial en los tamaños poblacionales; es decir, las poblaciones cercanas entre sí tienden a mantener tamaños poblacionales más parecidos que con poblaciones más distantes (Brown 1984). Asimismo, algunas especies mantienen poblaciones más numerosas hacia el centro geográfico de la distribución y menores hacia

las periferias; a este último se le ha denominado la hipótesis centro-abundante y en biogeografía es casi un dogma (Sagarin & Gaines 2002).

Desde hace décadas se ha buscado la explicación a estos patrones, principalmente abordándola desde el punto de vista de las características geográficas de las distribuciones de las especies. En particular, la hipótesis centro-abundante ha sido cuestionada y en un meta-análisis los autores concluyeron que la evidencia empírica indica que sólo en el 39% de los estudios revisados se encuentra un patrón-centro abundante (Sagarin & Gaines 2002). Además de la geografía, también se ha sugerido que las características ecológicas de los sitios, en combinación con la historia natural y procesos evolutivos de las especies, explican estos patrones (Brown 1995, Sagarin *et al.* 2006). Brown (1995) propuso que los tamaños poblacionales a nivel local reflejan qué tanto las condiciones de los sitios satisfacen los requerimientos ecológicos de la especie; es decir, la magnitud con la que las condiciones locales forman parte del nicho ecológico de la especie. Esta idea implícitamente sugiere que el nicho ecológico multidimensional de las especies (*sensu* Hutchinson 1957) no es un hipervolumen uniforme, sino que tiene una estructura interna en donde hay condiciones mejores en algunas regiones de ese híperespacio que en otras.

Maguire Jr. (1973) fue el primero en proponer que el hipervolumen del nicho ecológico tiene una estructura interna determinada por la adecuación de las especies a las condiciones del ambiente. Él planteó que existen condiciones óptimas, subóptimas y marginales dentro del hipervolumen que afectan la adecuación y, por ende, los tamaños poblacionales. Esta estructura interna del nicho está determinada por la posición relativa de las condiciones del ambiente para la especie, y sugirió que el óptimo se encontraría hacia el centroide del nicho cuando la respuesta de la especie a las variables del ambiente fuera con tendencia a una función normal (Fig. 1).

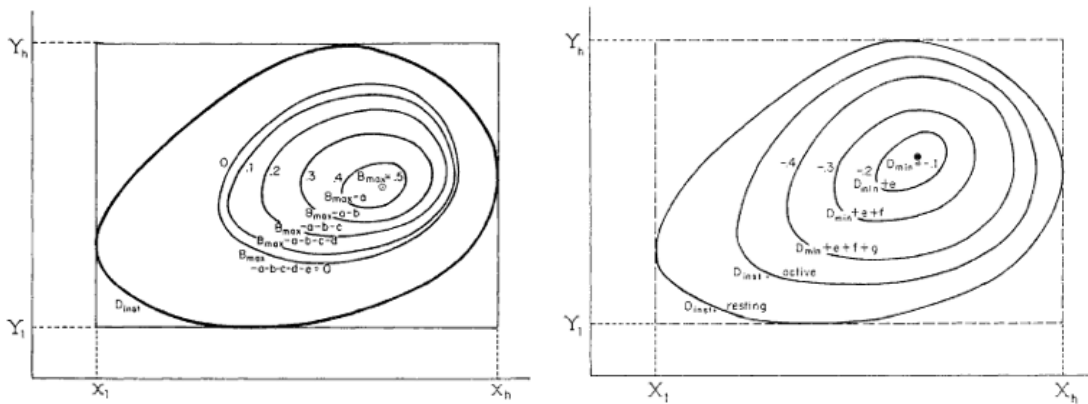


Figura 1. Estructura interna del nicho ecológico de una especie hipotética en donde la tasa de natalidad (izquierda) es máxima hacia el centroide y disminuye gradualmente conforme se aleja de éste. La tasa de mortalidad (derecha) presenta un patrón opuesto, lo que provoca que la densidad poblacional máxima se encuentre en las inmediaciones del centroide del nicho ecológico. Los ejes  $X$  y  $Y$  representan gradientes de dos variables fundamentales cualquiera del nicho ecológico que afectan la adecuación de la especie hipotética. Tomado de Maguire Jr. (1973).

Investigaciones recientes desarrolladas por este grupo de trabajo han encontrado que la distribución geográfica de la abundancia de las especies es mejor explicada por las características del nicho ecológico de las especies en el espacio ambiental que por las características del paisaje (Martínez-Meyer *et al.* 2013). En esos estudios comprobamos empíricamente que las abundancias de las poblaciones se relacionan inversamente con la distancia que guardan con el centroide del nicho ecológico multidimensional de la especie (Yáñez-Arenas *et al.* 2012, Martínez-Meyer *et al.* 2013), como lo sugirió Maguire Jr. (1973).

Por otro lado, un campo de rápido desarrollo en la biogeografía ecológica es el modelado de nichos ecológicos y distribuciones de especies (Peterson *et al.* 2011), en donde datos georeferenciados de presencia de una especie y un conjunto de capas ambientales en formato de sistemas de información geográfica (SIG) son analizados mediante métodos numéricos que buscan relaciones no aleatorias entre la presencia de la especie y las condiciones ecológicas de una región, reconstruyendo así el nicho ecológico de la especie y proyectándolo de nuevo en el espacio geográfico para producir un mapa que

refleja la distribución de esas condiciones propicias para el desarrollo de poblaciones de la especie; lo que se conoce como su distribución potencial (Peterson *et al.* 2011).

Recientemente, se ha explorado la capacidad de dichos modelos para predecir los patrones geográficos de abundancia (VanDerWal *et al.* 2009, Tôrres *et al.* 2012). Los resultados indican que los modelos de nicho ecológico en su estado actual tienen cierta (VanDerWal *et al.* 2009, Oliver *et al.* 2012, Tôrres *et al.* 2012) o nula (Jiménez-Valverde *et al.* 2009) capacidad de capturar la relación que existe entre las condiciones ecológicas y la adecuación de una especie y, por lo tanto, no es conveniente interpretar los mapas resultantes –incluso de algoritmos que producen modelos probabilísticos– como información que refleje los patrones de abundancia de las especies. Además, independientemente de la capacidad intrínseca que tengan los diversos métodos de modelación de nichos ecológicos para aproximar los patrones de abundancias, las pruebas han sido exclusivamente empíricas, sin ningún apoyo teórico detrás.

Por lo anterior, una aportación significativa sería, por un lado, un mejor entendimiento de los condicionantes de la abundancia de las especies y, por otro, en el campo de la modelación biogeográfica, la posibilidad de reconstruir la distribución geográfica de la abundancia de las especies. Los estudios preliminares que hemos realizado (Yáñez-Arenas *et al.* 2012, Escalante & Martínez-Meyer 2013, Martínez-Meyer *et al.* 2013) demuestran que existen las bases conceptuales y herramientas tecnológicas para lograrlo. Es por ello que en este proyecto desarrollamos una serie de programas para modelar la abundancia de las especies a lo largo de su distribución geográfica con base en la distancia al centroide del nicho, como una implementación de la teoría del nicho ecológico. Asimismo, reportamos los resultados del muestreo en campo para obtener datos sobre la distribución y abundancia de tres especies de plantas columnares endémicas o cuasi-endémicas a la Península de Baja California, las cactáceas *Pachycereus pringlei* (Cardón) y *Stenocereus gummosus* (Pitaya agria) y la fouquieriacea *Fouquieria columnaris* (Cirio).

Sin duda, éste es un campo vasto y fértil para desarrollar investigación y generar desarrollos tecnológicos en el campo de la biogeografía y la informática de la biodiversidad. Esta contribución es apenas un inicio que pretende estimular la curiosidad de los interesados y facilitar los procedimientos básicos de análisis para fomentar la

investigación. Las herramientas desarrolladas y presentadas aquí no las consideramos productos finales, pues en cada una de las etapas de la implementación de la modelación de la abundancia de las especies (i.e., preparación de datos de entrada, modelación, generación de resultados y validación) sabemos que es necesario profundizar en nuestra investigación y mejorar nuestras técnicas de manipulación y análisis de datos. La abundancia de las especies es un fenómeno complejo y multifactorial, en donde el nicho ecológico contribuye a una parte de la explicación. De esta forma, se exhorta a los interesados en este tema que utilicen profusamente estas herramientas para identificar sus fortalezas y debilidades, realicen críticas constructivas y sugerencias que permitan mejorarlas. De igual forma, es importante tomar los resultados de los análisis con toda la mesura y objetividad necesarios, con ojos críticos y siempre conscientes de sus alcances para evitar el abuso en el que frecuentemente se incurre con herramientas de uso sencillo.

## **DESARROLLO DEL PROYECTO**

El proyecto dio inicio en abril de 2012 y tiene tres componentes: (1) el desarrollo de herramientas computacionales que implementan el método de la distancia al centroide del nicho (DCN) para modelar los patrones geográficos de abundancias, (2) el trabajo de campo para obtener datos de distribución y abundancia de tres especies vegetales (Cardón, Pitaya agria y Cirio) en la Península de Baja California, y (3) el desarrollo de modelos de distribución de la abundancia de las tres especies de las plantas para las que se obtuvieron datos de campo.

### ***1. Herramientas para la modelación de abundancia***

El método de la distancia al centroide del nicho (DCN) para modelar los patrones geográficos de la abundancia es un desarrollo teórico-práctico creado por nuestro grupo de trabajo (Martínez-Meyer *et al.* 2013), con base en la teoría del nicho ecológico (Hutchinson 1957, Maguire Jr. 1973) y que ha sido implementado para algunas especies con datos de campo (Yáñez-Arenas *et al.* 2012, Ureña-Aranda *et al.* 2015) o extraídos de la literatura o bases de datos (Escalante & Martínez-Meyer 2013, Martínez-Meyer *et al.* 2013).



El procedimiento involucra cuatro etapas: (a) preparación de los datos de entrada, (b) modelación, (c) generación de resultados y (d) validación (Fig. 2). A continuación se describe cada una de las etapas.

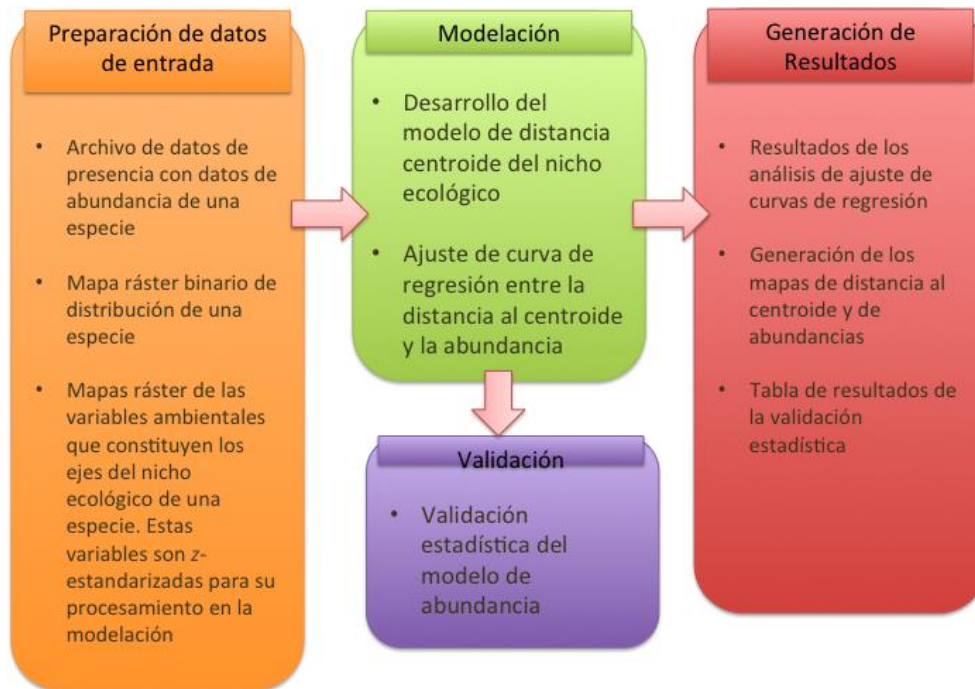


Figura 2. Procedimiento para la modelación de abundancias bajo el método de la distancia al centroide del nicho ecológico (Martínez-Meyer *et al.* 2013).

*a) Preparación de los datos de entrada.*

Los datos que sirven de insumo para la modelación de los nichos ecológicos son: (1) un archivo (en formato .csv) que contenga los datos de presencia y abundancia de una especie, con los siguientes campos: especie, longitud, latitud y abundancia; (2) un mapa ráster con la distribución potencial binaria (presencia = 1, ausencia = 0) de la especie. Este mapa debió haber sido generado previamente con algún método convencional de modelación de nichos (*e.g.*, GLM, GAM, GARP, MaxEnt, Bioclim, etc.) y procesado para convertirlo en binario; y (3) un conjunto de capas ráster que conforman las variables del nicho de la especie. Se sugiere que sean las mismas que se usaron para generar el mapa de distribución potencial.

*b) Modelación.*

La etapa de la modelación implica dos procesos: (1) la estimación de la distancia euclidiana multidimensional al centroide del nicho ecológico y (2) la identificación de la curva que mejor ajuste a la relación de la distancia al centroide y la abundancia. Para ello, el mapa de la distribución binaria de la especie funciona como máscara, ya que se extraen los pixeles de presencia (*i.e.*, los de valor 1) y los valores de cada variable son  $z$ -estandarizados (*i.e.*, media = 0, d.e. = 1); de esta forma el centroide multidimensional es por definición 0. Posteriormente, se calcula la distancia euclidiana multidimensional de cada uno de los pixeles de presencia al centroide con la fórmula:

$$Dc = \sqrt{\sum_{j=1}^n (\bar{\mu}_j - a_{ij})^2}$$

En donde  $Dc$  es la distancia al centroide,  $\bar{\mu}_j$  es la media de la variable  $j$ , y  $a_{ij}$  es el valor del pixel  $i$  de la variable  $j$ .

Una vez obtenida la distancia al centroide de cada uno de los pixeles predichos como presente en el mapa de distribución potencial, se asigna el valor de distancia al centroide a cada uno de los registros de presencia del archivo de puntos que tiene asociado un dato de abundancia registrado en campo. De esta manera, para cada registro de las especies se tiene un valor de abundancia y un valor de distancia al centroide.

Esta información se analiza por medio de una regresión bajo diferentes funciones de ajuste, para identificar la que maximice la relación entre las dos variables. Es decir, se prueban diferentes curvas para identificar aquella que explique mejor la relación por medio del coeficiente de determinación ( $R^2$ ). Finalmente, usando la ecuación de la regresión de mejor ajuste se estima la abundancia en cada uno de los pixeles del mapa resultando en un mapa de abundancia modelada para la especie.

*c) Generación de resultados.*

Existen diferentes resultados generados de estos análisis. En primer lugar se producen dos mapas, uno corresponde a la distancia al centroide del nicho de la especie y

otro a las abundancias estimadas. También se presentan los coeficientes de determinación para las regresiones probadas y los resultados de la validación estadística.

*d) Validación estadística.*

Para evaluar la capacidad predictiva de las estimaciones de abundancia se implementó un proceso automatizado de *bootstrap* que consiste en generar un número determinado de simulaciones de modelos de regresión entre la abundancia y la distancia al centroide, pero usando particiones aleatorias de los datos de abundancia. En cada iteración se utiliza un porcentaje de los datos para generar predicciones de la abundancia y el resto se usa para validar esas predicciones. A partir de las predicciones de todos los modelos de regresión se calcula la varianza, la desviación estándar y los intervalos de confianza. De esta forma se obtiene para cada pixel un valor de incertidumbre en las predicciones.

Existen técnicas de regresión que incorporan la estructura geográfica de los datos para considerar el efecto que puede tener la autocorrelación espacial en la regresión, como el estadístico *I* de Moran, pero su implementación es para regresiones lineales (Rangel et al. 2010). En el caso de las especies analizadas en este estudio ninguna mostró una relación lineal entre la distancia al centroide y la abundancia, por lo que no fue posible integrar este análisis en las regresiones.

En este proyecto desarrollamos cuatro herramientas para la modelación de la abundancia con el fin de que tengan una amplia difusión y sean exploradas y utilizadas por usuarios con diferentes perfiles y habilidades técnicas: un programa independiente (“stand-alone”) desarrollado en lenguaje Java, un grupo de herramientas implementadas en el programa estadístico R, una extensión para el sistema de información geográfica ArcView 3.x programada en lenguaje Avenue y otra para ArcGis 10.2 (o superior) programada en lenguaje Phyton. De forma genérica hemos denominado a estas herramientas como *Nichosfera*, que es un término acuñado por el M. en C. Alberto González Gallina, del INECOL, durante un curso de modelado de nichos que dictamos el Dr. Octavio Rojas y el Dr. Enrique Martínez Meyer, y que se resume en una palabra el concepto de nicho multidimensional de Hutchinson (1957). Estas herramientas tienen como finalidad hacer un

análisis de la estructura interna de este hipervolumen abstracto que aquí denominamos *Nichosfera*.

Aunque todas estas herramientas siguen el protocolo descrito anteriormente, cada una hace ciertas funciones. El programa en Java calcula la distancia al centroide a partir del mapa binario de la especie y las variables ambientales, calcula regresiones entre la distancia al centroide y la abundancia en los registros de la especie bajo las tres funciones que con base en nuestra experiencia son las más frecuentes: lineal, logarítmica y exponencial. Genera un mapa de distancia al centroide y reporta los resultados de las regresiones.

El conjunto de scripts en R también estima las distancias al centroide y genera el mapa de distancias al centroide; además, está implementada la técnica de *bootstrap* para la validación y estimación de los límites de confianza del modelo de abundancias.

El más simple de todos es la extensión para ArcView 3.x, ya que sólo calcula las distancias al centroide y genera un mapa de ello. Finalmente, la extensión para ArcGis 10.2 (o superior) calcula las distancias al centroide y hace estimaciones de abundancia bajo un modelo lineal. Cada una de las herramientas tiene un instructivo para su uso incluidos en la sección de Anexos.

Es importante resaltar que ninguna de las herramientas genera un mapa de la abundancia modelada debido a que las regresiones que calcula son sólo un conjunto limitado dentro de un espectro más amplio. Consideramos que automatizar la estimación de la abundancia puede resultar contraproducente porque los usuarios pueden fácilmente incurrir en el error de quedarse con las curvas que ajustan las herramientas. Sin embargo, lo recomendable es que los usuarios realicen una exploración más amplia de la relación entre la distancia al centroide y la abundancia, incluso con métodos no lineales, por ejemplo Modelos Aditivos Generalizados (GAM, por sus siglas en inglés); sin embargo, esto no lo hemos explorado aún. En su estado actual, las herramientas desarrolladas en este proyecto permiten automatizar la parte más laboriosa del proceso de la modelación.

## **2. Trabajo de campo**

Nuestro trabajo de campo para la obtención de datos de las tres especies columnares se restringió a la región continental de la Península de Baja California; es decir, no visitamos las islas del Golfo de California ni las pequeñas zonas de lado continental en las

que el Cardón y la Pitaya agria mantienen poblaciones. El trabajo de campo consistió en recorrer la Península de Baja California en toda su extensión, visitando sitios para verificar la presencia y cuantificar la abundancia de las tres especies de plantas columnares (Cirio, *Fouquieria columnaris*; Pitaya agria, *Stenocereus gummosus*; y Cardón, *Pachycereus pringlei*).

El diseño para la selección de sitios de muestreo de las tres especies se hizo con base en una regionalización de la Península de Baja California, generada mediante la clasificación de 19 parámetros bioclimáticos (realizado generosamente por el Dr. Oswaldo Téllez) en 20 categorías o regiones climáticas. Posteriormente, se generaron 400 puntos de forma aleatoria estratificada a la proporción de cada una de estas regiones (Fig. 3). De esos puntos se seleccionaron todos aquellos que se encontraran dentro de una distancia de 2 km de un camino transitable en vehículo (carreteras pavimentadas y terracerías), resultando así en 149 sitios que sirvieron de guía para el muestreo (Fig. 4).

Es importante hacer notar que en el diseño del muestreo estratificado que usamos se incluyeron sitios en todos los dominios climáticos, sin embargo la proporción de puntos no fue exacta por limitaciones de acceso a algunos sitios (por ejemplo, por la ausencia de caminos transitables). Los datos entonces presentan un nivel de sesgo ambiental que creemos que es relativamente bajo (al menos al nivel de los 20 dominios climáticos), ya que los modelos de distribución potencial tuvieron alto desempeño y significancia en las pruebas de validación (ver la sección 3. Modelación de distribución de la abundancia de plantas columnares). Sin embargo, reconocemos que existe un sesgo geográfico debido a que los muestreos fueron cercanos a las carreteras. Esta proximidad a las carreteras puede tener efectos sobre la abundancia de las especies, más que sobre su presencia (que es lo que se usa para los modelos de distribución potencial). A pesar de que en el trabajo de campo hicimos los conteos en sitios con aparente bajo nivel de perturbación, es imposible saber en una visita de unos minutos si la población de cada sitio estaba impactada en su demografía.

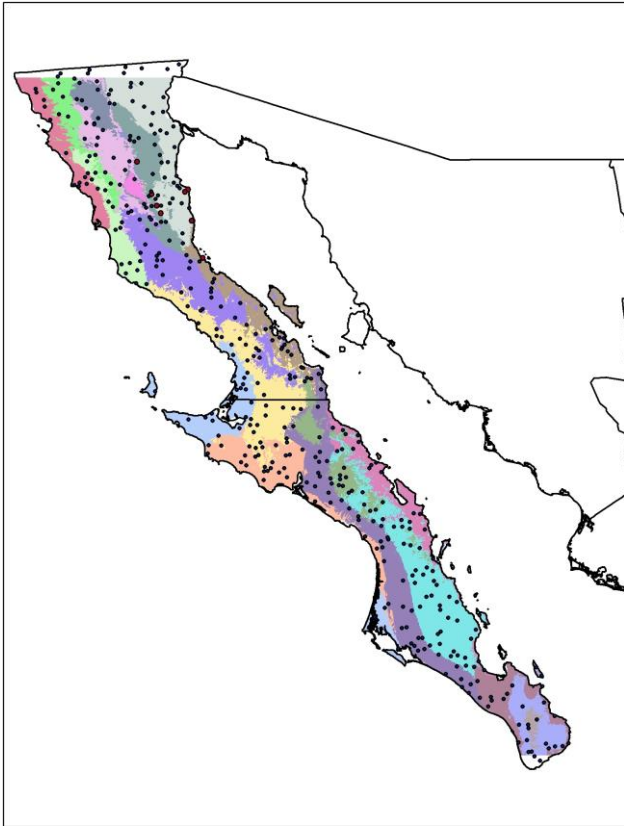
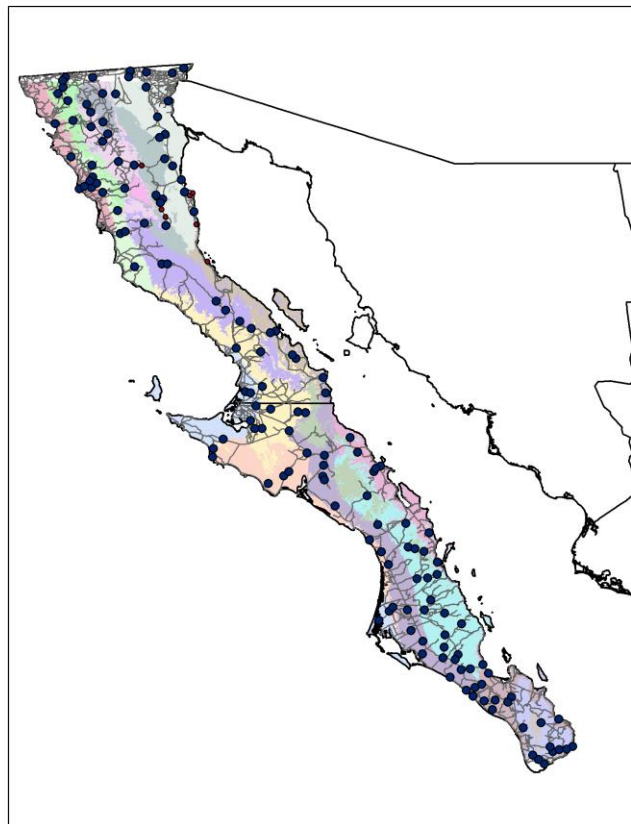


Figura 3. Mapa de la Península de Baja California en donde se muestran las 20 regiones climáticas (polígonos de colores) y 400 puntos aleatorios estratificados.

Figura 4. Mapa de la Península de Baja California en donde se muestran los 149 sitios cercanos a las vías de comunicación que fueron seleccionados para el muestreo.



Realizamos tres salidas al campo: del 18 al 25 de noviembre de 2012, del 4 al 12 de abril y del 29 de agosto al 17 de septiembre de 2013 en las que visitamos un total de 194 localidades. De éstas, en 139 registramos al menos un individuo de alguna de las tres especies y en 55 sitios no registramos a ninguna de las especies. Es importante resaltar que los registros de ausencia son muy importantes, pues son sitios que se visitaron y en donde se comprobó que no se encuentra la especie en cuestión. Los registros de ausencia son poco comunes en las bases de datos y sabemos que pueden ser muy útiles para la modelación del nicho ecológico de las especies (Peterson *et al.* 2011).

En los sitios en donde sí registramos la presencia de alguna de las especies, implementamos un diseño de muestreo sugerido por el Dr. Ricardo Rodríguez Estrella, del Centro de Investigaciones Biológicas del Noroeste (CIBNOR), que consistió en trazar dos cuadrantes de 50 m de largo por 5 m de ancho dispuestos en forma de “L”, cubriendo una superficie total de 500 m<sup>2</sup>. Dentro de cada uno de estos cuadrantes contamos a todos los individuos de las tres especies (Fig. 5).

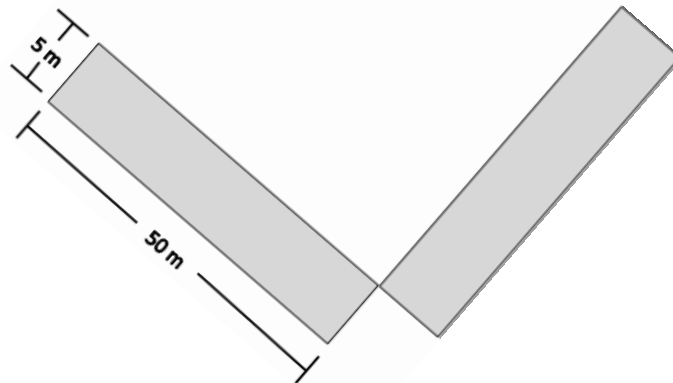


Figura 5. Diseño de muestreo para el conteo de organismos de las tres especies de plantas columnares.

Una vez en los sitios con presencia de las especies, observamos que había individuos de distintos tamaños, por lo que establecimos conteos diferenciados por clase de tamaño: (1) Plántulas, que para Cirio y Cardón definimos como individuos menores a 1 m de longitud y para la Pitaya agria establecimos que además de ser menores a 1 m no tuvieran ramificaciones. (2) Jóvenes, que para Cirio y Cardón fueron individuos entre 1 y 3 m de longitud, y en el caso del Cardón que no tuvieran ramificaciones; para la Pitaya agria fueron individuos entre 1 y 2 m y con un máximo de tres ramificaciones. (3) Adultos, que para las tres especies fueron individuos mayores a tres metros y ramificados para el caso del Cardón y la Pitaya agria. Los resultados de los conteos por clase de tamaño se presentan en el Cuadro 1. Las tres especies presentan características morfológicas muy particulares que las hacen fácilmente distinguibles de otras cactáceas o fouquieriaceas aún en sus etapas de plántulas, por lo que consideramos que el error de omisión en los conteos de cada grupo de edad es muy bajo.

Cuadro 1. Número de individuos de Cirio, Cardón y Pitaya agria, y número de sitios en donde se registró la presencia de cada una de las especies en los muestreos en la Península de Baja California.

	<b>Plántulas</b>	<b>Jóvenes</b>	<b>Adultos</b>	<b>Todos</b>	<b># Sitios de presencia</b>
<b>Cirio</b>	19	55	91	165	27
<b>Cardón</b>	145	492	615	1252	103
<b>Pitaya</b>	291	261	531	1083	121

Los patrones de distribución de presencia y abundancia de las tres especies fueron específicos. El Cirio es la especie con la distribución más restringida, acotada a la región central de la Península de Baja California, mientras que tanto la Pitaya agria y el Cardón tienen una distribución amplia, solo ausentes en la región norte de la Península (Figs. 6-8).

La abundancia de las tres especies fue heterogénea, encontrándose algunos sitios de muy alta concentración de individuos. Para el Cirio, la zona de mayor abundancia fue la región central del Valle de los Cirios (Fig. 6).



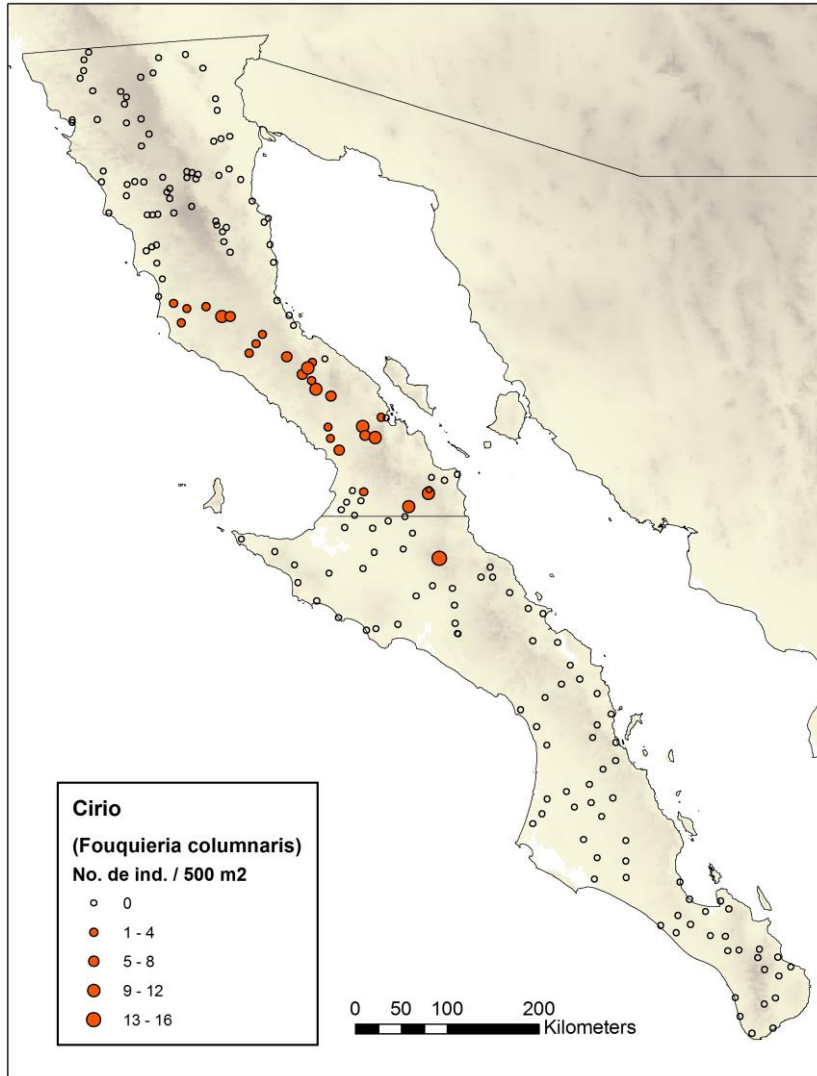
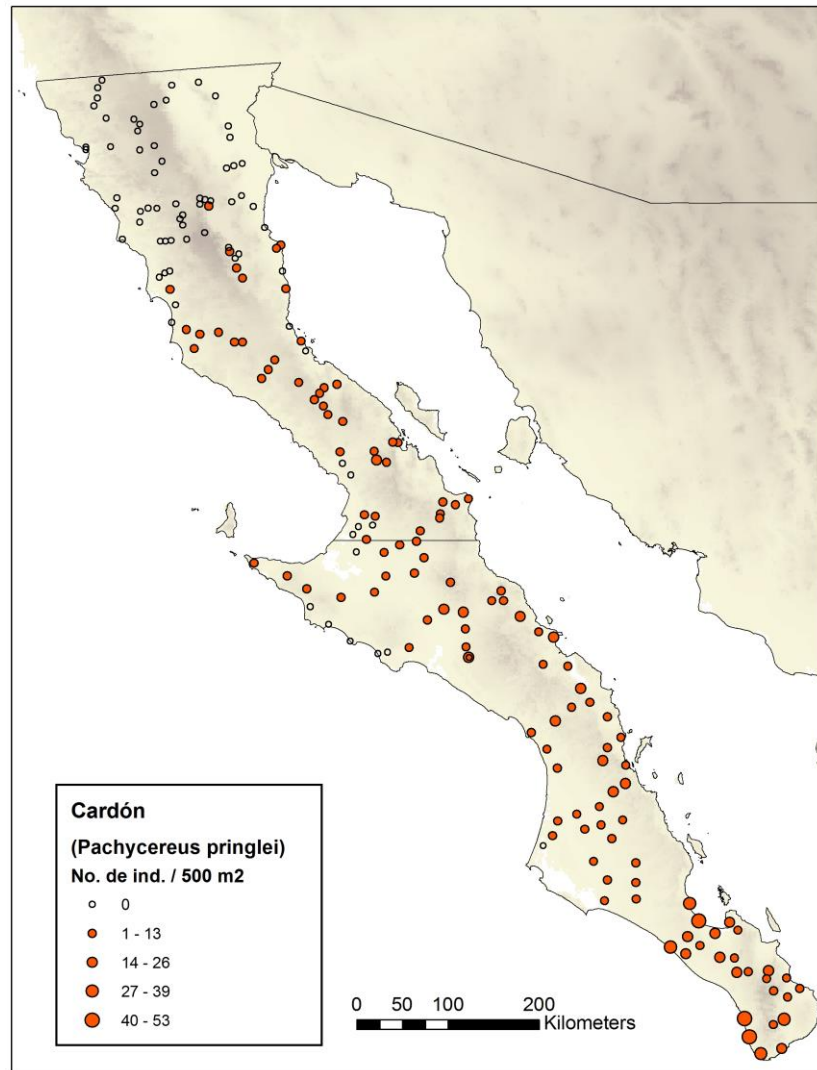


Figura 6.  
Distribución y  
abundancia del Cirio  
(*Fouquieria  
columnaris*) en la  
Península de Baja  
California.

En el Cardón observamos algunos aspectos muy interesantes. Aunque las zonas con muy altas densidades de individuos se encontraron en la región sur de la Península, muy cerca de La Paz y cerca de Todos Santos, BCS (Fig. 7), los ejemplares más altos y aparentemente vigorosos se encontraron al norte, en la región del Valle de los Cirios e inmediaciones del Desierto del Vizcaíno.

Figura 7. Distribución y abundancia del Cardón (*Pachycereus pringlei*) en la Península de Baja California.



Finalmente, la Pitaya agria fue la especie que presentó mayor dificultad para el registro de su abundancia debido a su forma de crecimiento. Esta especie presenta un crecimiento rastrero en sitios costeros con vientos fuertes o muy arenosos; es decir, las ramas crecen y se extienden sobre la superficie del suelo, mientras que en sitios menos agrestes las ramas crecen verticalmente. Igualmente, en algunos sitios crece en grupos formando cúmulos de organismos, dificultando mucho el conteo individual. Es por ello, que consideramos que no todos los conteos de esta especie son del todo confiables, así que

advertimos que deben ser usados con cautela. Dicho esto, los sitios de mayor abundancia de Pitaya agria se observaron en Baja California Sur, a la altura de La Paz, pero del lado de la costa del Pacífico (Fig. 8).

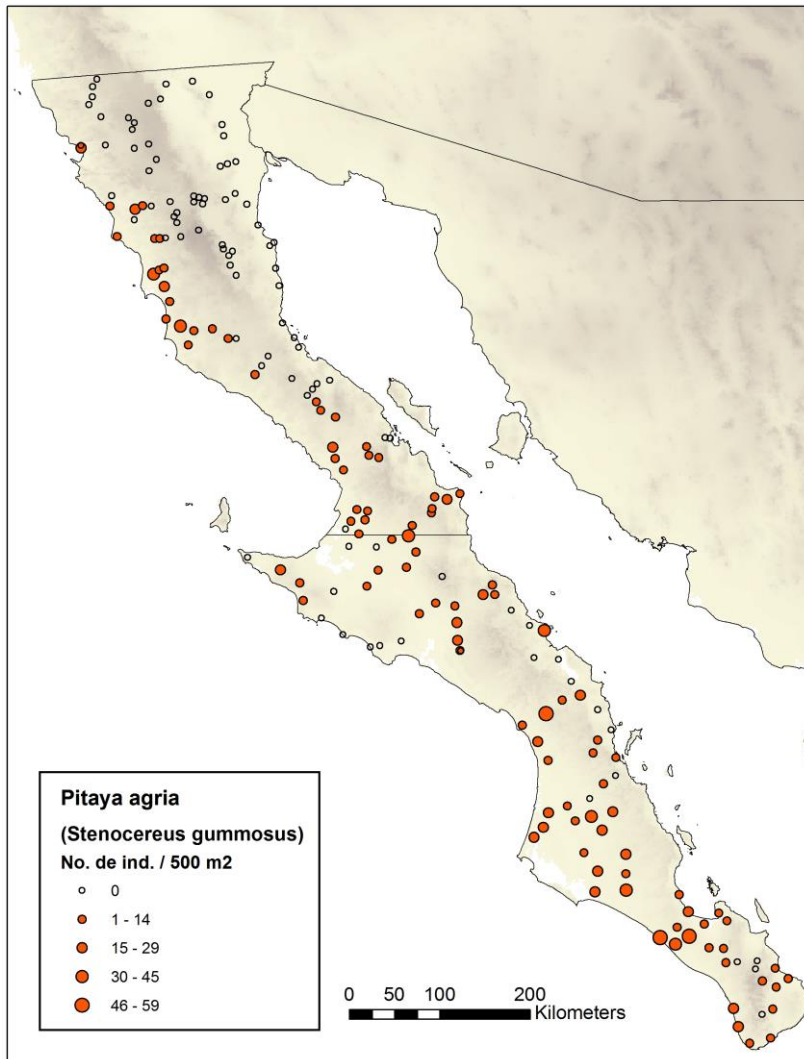


Figura 8. Distribución y abundancia de la Pitaya agria (*Stenocereus gummosus*) en la Península de Baja California.

### ***3. Modelación de distribución de la abundancia de plantas columnares***

La última sección de este informe corresponde un ejercicio de modelación de la distribución geográfica de la abundancia de las tres especies de plantas columnares, Cirio, Cardón y Pitaya agria, utilizando las herramientas desarrolladas durante este proyecto.

El proceso de modelación consistió de las siguientes etapas: (a) preparación de los insumos, (b) cálculo de las distancias al centroide, (c) estimación de la abundancia y (d) validación estadística del modelo y estimación de la incertidumbre. A continuación se describe cada una de ellas.

#### *a) Preparación de los insumos.*

Los insumos para cada una de las especies corresponden a: (i) base de datos de registros de presencia y abundancia, (ii) mapa de distribución binaria, y (iii) capas ambientales que definen las variables del nicho.

(i) Las bases de datos de las especies fueron analizadas para ver el comportamiento de los datos de abundancia. Observamos que los datos de las plántulas y los jóvenes mostraban patrones más erráticos pues en varios sitios de presencia de la especie no se encontraban estas clases de edad, y en otros sitios los números eran muy altos comparados con los adultos. En cambio, los adultos mostraron un comportamiento más estable. Asimismo, considerando la historia de vida de las especies en el contexto de la teoría del nicho, pensamos que la abundancia de adultos es un mejor reflejo de las condiciones ambientales para la especie que la abundancia total, pues aunque las tres especies pueden producir gran cantidad de vástagos, como son especies muy longevas la condición de llegar a adulto reproductivo es un indicativo de que las condiciones del ambiente les son favorables. Por lo anterior, los resultados que se presentan son producto de los análisis y la modelación para adultos únicamente. De esta forma se prepararon los archivos en formato .csv para cada una de las especies, con cuatro campos: especie, longitud, latitud y abundancia (conteo de todos los adultos en 500 m<sup>2</sup>). Cabe mencionar que los sitios muestreados estuvieron separados por al menos 5 km y la resolución espacial de los análisis fue de 1 km; por lo tanto cada punto de muestreo corresponde solamente a un pixel, independientemente del número de individuos registrados en cada uno de ellos.

(ii) Los mapas de distribución binaria fueron generados modelando la distribución de cada una de las especies bajo dos métodos, GARP (Stockwell y Noble 1999) en su versión Desktop 1.1.6, y MaxEnt v3.3.3k (Phillips *et al.* 2006). Para ello, se usaron los mismos datos de campo (todos los registros, no solo los adultos), pero sin el dato de abundancia. Las capas ambientales que usamos para generar los modelos fueron de dos tipos, climáticas y topográficas. Las capas climáticas provinieron de Cuervo-Robayo *et al.* (2013), que son superficies climáticas desarrolladas para México bajo el método de interpolación “thin-plate spline”, que es el mismo implementado para WorldClim (Hijmans *et al.* 2005), la base de datos climática más popular en el campo de la modelación de nichos. La base de datos que usamos también consta de las mismas 19 variables bioclimáticas que WorldClim, pero la cobertura temporal es mayor (de 1910 a 2009) y la densidad de estaciones climáticas para México también. Las variables topográficas usadas fueron elevación, pendiente e índice topográfico, obtenidas de la base de datos Hydro1k (USGS 2001). Todas las variables ambientales fueron re-escaladas a 0.01 grados (aprox. 1 km<sup>2</sup>) bajo el método de interpolación del vecino más cercano porque la resolución nativa de las capas climáticas y topográficas no era la misma (0.00833 y 0.01, respectivamente) y los modelos de nicho exigen una misma resolución y extenso de las capas ambientales. El área de análisis (M) fue la Península de Baja California para el Cirio y se incluyó una porción de Sonora para el caso del Cardón y la Pitaya agria, ya que ambas especies tienen poblaciones en la región costera continental.

Para los modelos GARP, los puntos de presencia se dividieron en una proporción de 75:25 para calibración y validación, respectivamente. Se generaron 100 modelos y se escogieron los 10 mejores usando como criterio los 20 modelos con menor omisión (la opción “soft” en el programa) y de éstos, los 10 más cercanos a la mediana del tamaño del área predicha (eje de la comisión), siguiendo las sugerencias de Anderson *et al.* (2003). Los 10 mejores modelos fueron posteriormente sumados en un sistema de información geográfica y se obtuvo un mapa de consenso. Posteriormente, el mapa de consenso fue convertido a binario utilizando como valor de corte aquél a partir del cual se predecían el 95% de los puntos de presencia, tanto de entrenamiento como de validación.

En MaxEnt se usó la misma división de los puntos de presencia, es decir, el mismo 75% de los puntos que se usaron para la calibración de los modelos GARP se usaron para la

calibración de los modelos MaxEnt y el 25% restante para la validación. En MaxEnt se dejaron los parámetros por defecto para generar el modelo de nicho y el formato de la salida fue en escala logística. El mapa de salida fue convertido a binario en un SIG siguiendo el mismo procedimiento que con GARP.

Los modelos resultantes del procesamiento con GARP y MaxEnt fueron evaluados usando ese 25% restante de los puntos se dejaron fuera de la calibración mediante pruebas de ROC-parcial (Peterson *et al.* 2008). Para ello usamos la herramienta desarrollada por Barve (2008), generando un *bootstrap* de 1000 iteraciones, con una partición de esos datos de validación al 50% y un error de omisión aceptado de 1%. Los resultados de las pruebas de validación indicaron que tanto los modelos de GARP y de MaxEnt para las tres especies fueron estadísticamente significativos ( $P < 0.001$ ; Cuadro 2).

Cuadro 2. Resultados de las pruebas de ROC-parcial para los modelos de nicho generados con GARP y Maxent para Cirio, Cardón y Pitaya.

	<b>GARP</b>		<b>MaxEnt</b>	
	Promedio de AUC Ratio	<i>P</i> -valor	Promedio de AUC Ratio	<i>P</i> -valor
<b>Cirio</b>	1.94	< 0.001	1.93	< 0.001
<b>Cardón</b>	1.56	< 0.001	1.70	< 0.001
<b>Pitaya agria</b>	1.69	< 0.001	1.74	< 0.001

Posteriormente, los mapas binarios generados con GARP y MaxEnt para cada especie fueron sobrepuestos entre sí en un SIG para identificar las zonas de coincidencia entre los dos métodos de modelación. El área de sobreposición que contuviera el 95% de los registros de las especies fue seleccionada y conformó el mapa binario final para la modelación de la distancia al centroide (Figs. 9-11).

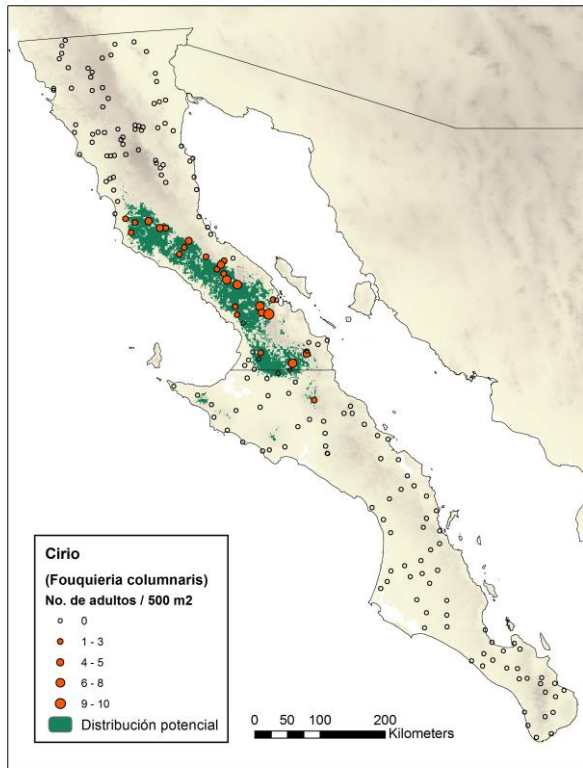
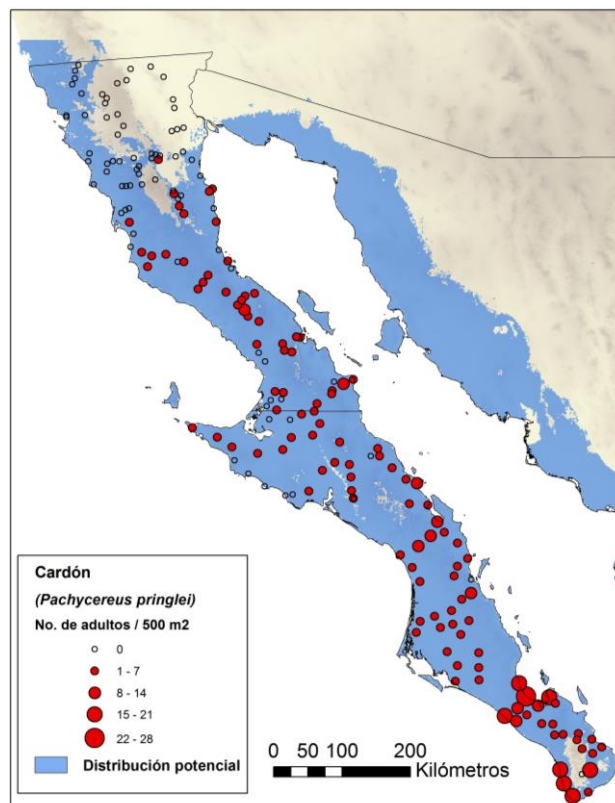


Figura 9. Mapa binario representando la distribución potencial del Cirio (*Fouquieria columnaris*) en la península de Baja California.

Figura 10. Mapa binario representando la distribución potencial del Cardón (*Pachycereus pringlei*).



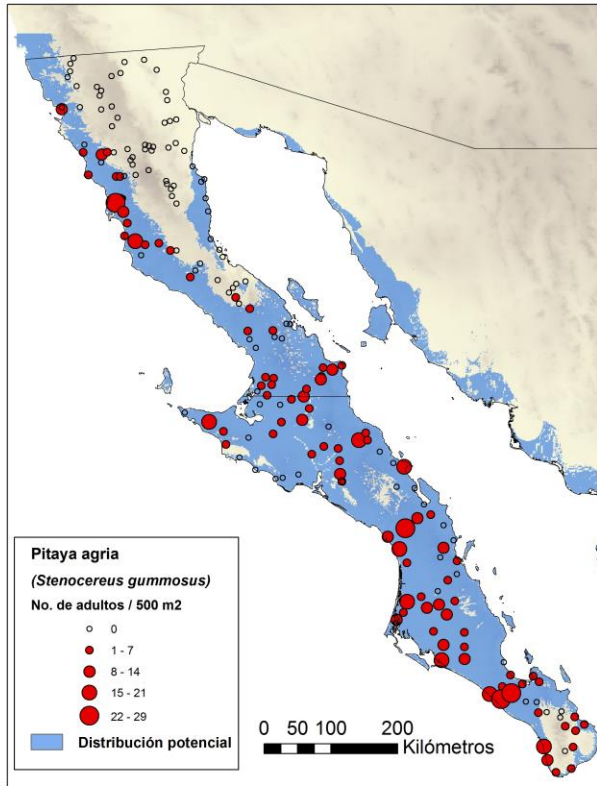


Figura 11. Mapa binario representando la distribución potencial de la Pitaya agria (*Stenocereus gummosus*).

(iii) Finalmente, las capas ambientales que conformaron los ejes del nicho ecológico para calcular las distancias al centroide fueron las mismas 19 variables ambientales (Cuervo-Robayo *et al.* 2013) y las tres topográficas (USGS 2001) usadas para construir los modelos de nicho.

*b) Cálculo de la distancia al centroide.*

Se realizó el procedimiento para el cálculo de las distancias al centroide en cada una de las especies, ya automatizado en los programas que acompañan a este informe. Este procedimiento consiste en: (1) usar el mapa binario de distribución potencial para definir el área de análisis (píxeles de presencia), (2) realizar una *z*-estandarización de las variables en esos píxeles, (3) calcular las distancias euclidianas al centroide, y (4) proyectar esos valores en el espacio geográfico para producir un mapa para cada especie (Figs. 12-14).

La utilidad del mapa de distancias al centroide radica en que se puede usar como un índice de abundancia relativa cuando no se cuenta con buena información de abundancia a través de la distribución de la especie, pero sólo en los casos en los que exista una relación (inversa) entre la abundancia y la distancia al centroide (Yañez-Arenas *et al.* 2012).



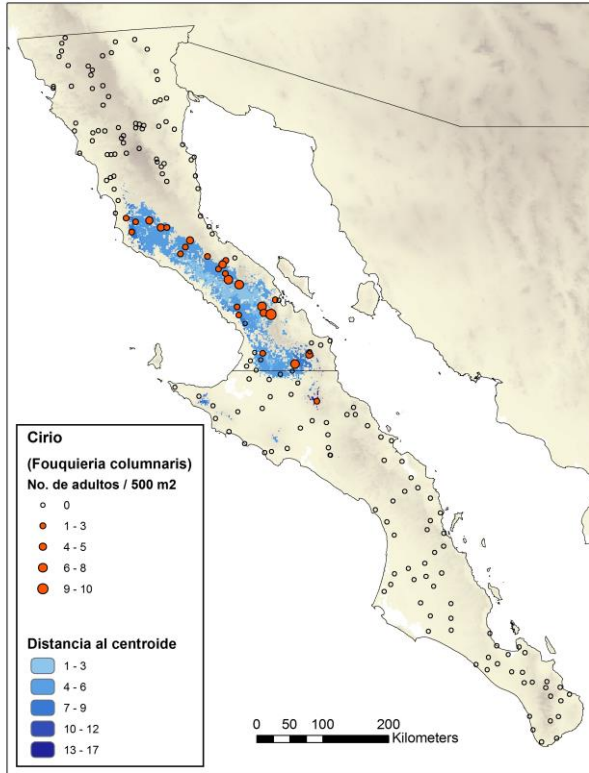
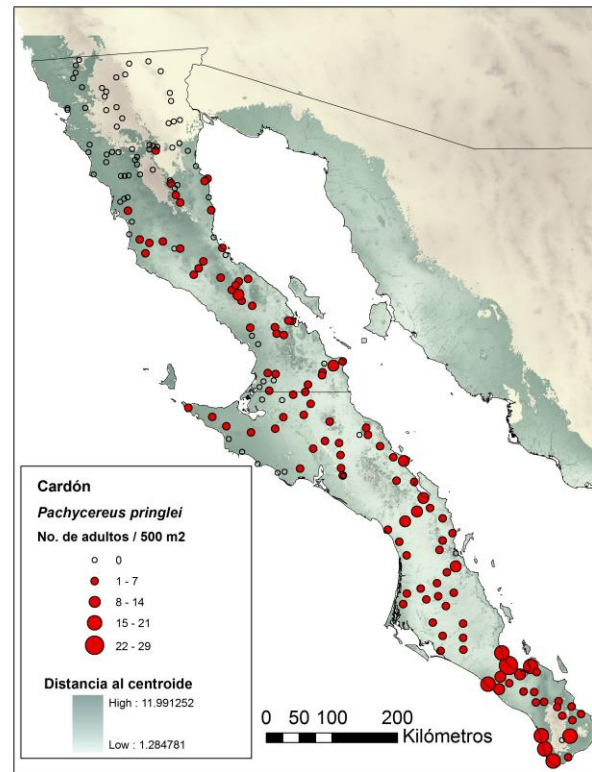


Figura 12. Distancias al centroide del nicho ecológico del Cirio (*Fouquieria columnaris*).

Figura 13. Distancias al centroide del nicho ecológico del Cardón (*Pachycereus pringlei*).



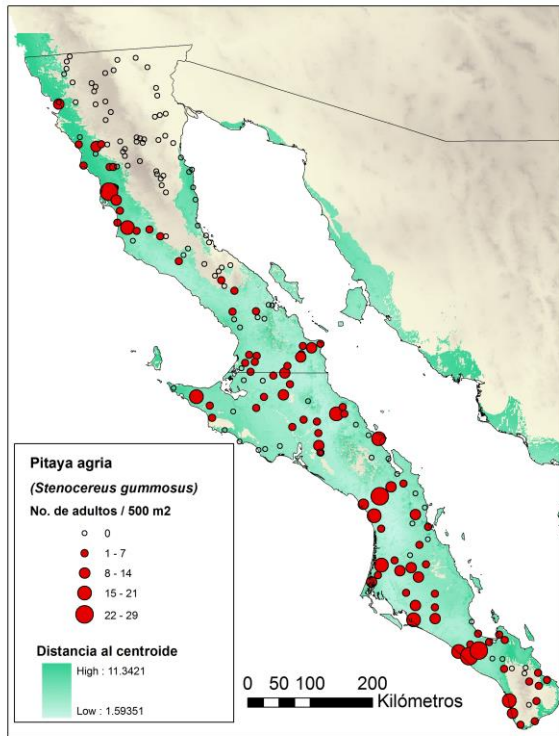


Figura 14. Distancias al centroide del nicho ecológico de la Pitaya agria (*Stenocereus gummosus*).

c) *Estimación de la abundancia.*

El paso final fue la estimación de los patrones espaciales de la abundancia de los individuos adultos de las especies a partir de la relación de ésta con la distancia al centroide del nicho. Para ello, se realizó una regresión lineal, una logarítmica y una exponencial entre estas dos variables para cada una de las especies para identificar la que mejor ajustara (maximizando la  $R^2$  y la significancia), como ha sido observado en otras especies (Martínez-Meyer *et al.* 2013). De todas las combinaciones entre especie y método de regresión sólo se encontró una relación significativa, que fue el caso del Cirio con una regresión logística (Cuadro 3). Debido a esto, la información se analizó con mayor profundidad en el programa estadístico SPSS v1.9 (IBM 2010) para determinar si con otras funciones se encontraba un mejor ajuste.

Cuadro 3. Resultados (coeficiente de determinación  $R^2$  y valor de significancia  $P$ ) de las regresiones entre la abundancia de adultos (ind. / 500 m<sup>2</sup>) y la distancia al centroide para el Cirio (*Fouquieria columnaris*), el Cardón (*Pachycereus pringlei*) y la Pitaya agria (*Stenocereus gummosus*).

	Lineal		Logarítmica		Exponencial	
	R <sup>2</sup>	P-valor	R <sup>2</sup>	P-valor	R <sup>2</sup>	P-valor
<b>Cirio</b>	0.097	0.113	<b>0.192</b>	<b>0.022</b>	0.014	0.746
<b>Cardón</b>	0.017	0.169	0.014	0.216	0.002	0.609
<b>Pitaya</b>	0.001	0.797	0.002	0.718	0.006	0.497

Las funciones adicionales aplicadas (además de la lineal, logarítmica y exponencial) fueron: inversa, cuadrática, cúbica, potencia, compuesta, de crecimiento, S y logística. Los resultados de esos análisis mostraron que sólo para el Cirio se encontró un mejor ajuste bajo una regresión cuadrática ( $R^2 = 0.286$ ,  $P = 0.017$ ; Fig. 15).

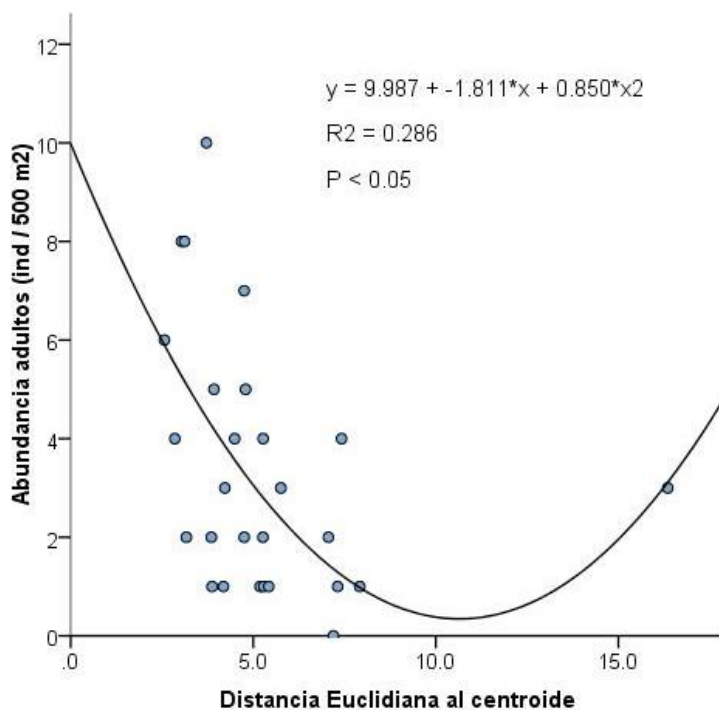
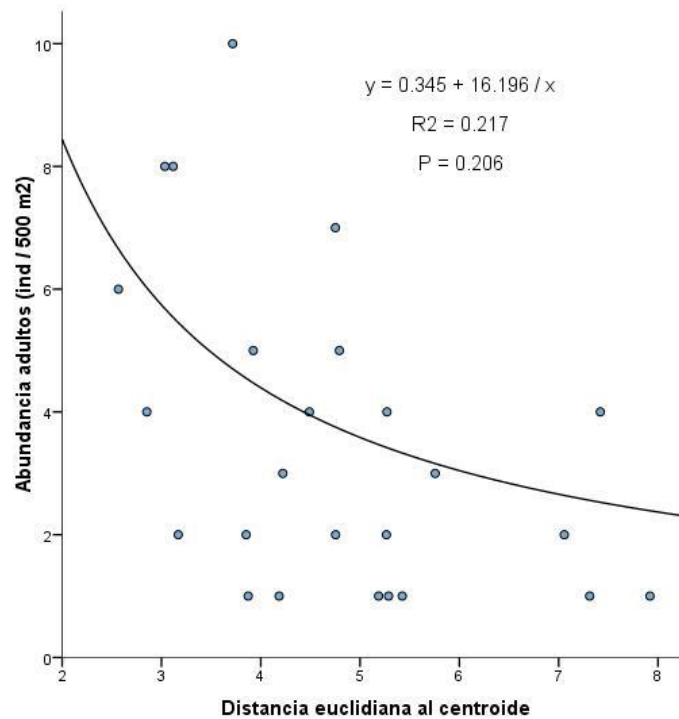


Figura 15. Relación cuadrática entre la distancia euclidiana al centroide del nicho y la abundancia de individuos adultos del Cirio (*Fouquieria columnaris*).

Es interesante observar que hay un dato que aunque válido en términos de su abundancia y distancia al centroide, es extremo y hace que el mejor ajuste sea cuadrático. Al quitar ese dato en la regresión, la relación cambia a inversa, el ajuste baja (de  $R^2 = 0.286$  a  $R^2 = 0.216$ ) y la relación no es estadísticamente significativa ( $P = 0.206$ ) (Fig. 16).

Figura 16. Relación inversa entre la distancia euclidiana al centroide del nicho y la abundancia de individuos adultos del Cirio (*Fouquieria columnaris*).



En los casos del Cardón y la Pitaya agria no hubo ninguna curva que mostrara un ajuste significativo ( $P < 0.05$ ). Para el Cardón y la Pitaya agria, el ajuste más alto fue bajo una regresión cúbica, pero no fue una relación fuerte ni significativa ( $R^2 = 0.045$ ,  $P = 0.145$ ; Fig. 16 y  $R^2 = 0.03$ ,  $P = 0.495$ , respectivamente; Figs. 17 y 18).

Figura 17. Relación cúbica entre la distancia euclidiana al centroide del nicho y la abundancia de individuos adultos del Cardón (*Pachycereus gramei*).

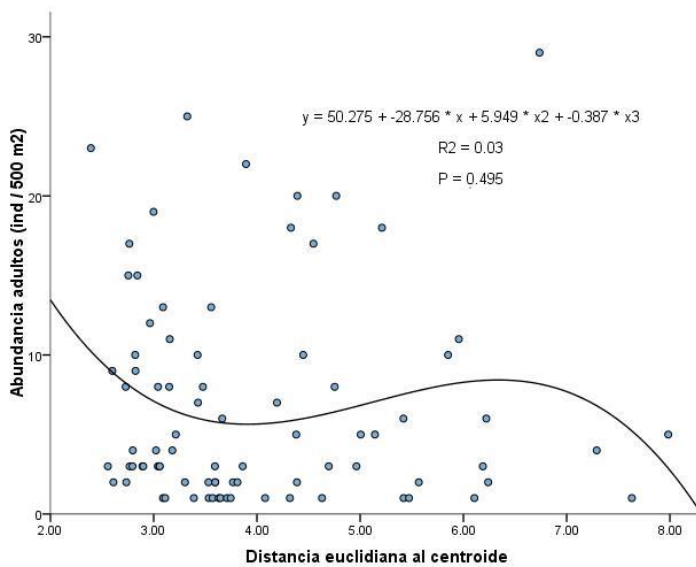
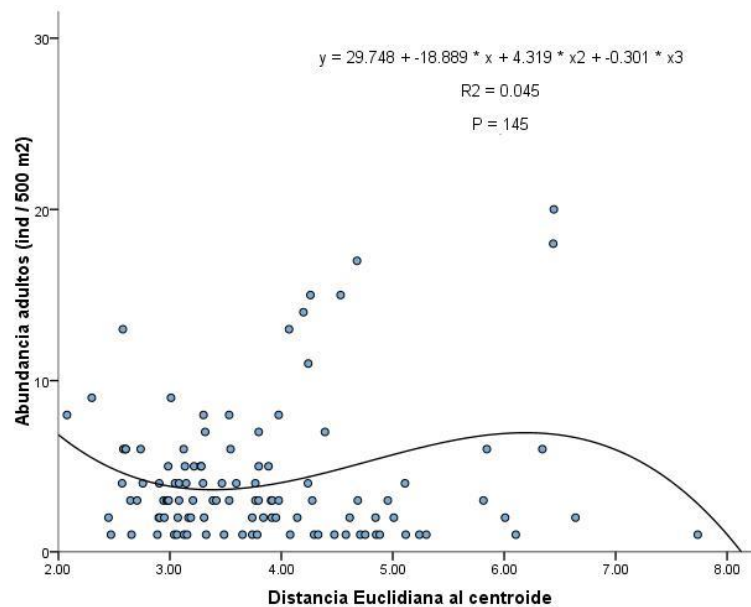


Figura 18. Relación cuadrática entre la distancia euclidiana al centroide del nicho y la abundancia de individuos adultos de la Pitaya agria (*Stenocereus gummosus*).

Estos resultados indican que para el Cirio las variables ambientales usadas para modelar la distancia al centroide del nicho ecológico explican cerca del 30% de la variabilidad de los datos de abundancia bajo el modelo cuadrático. Sin embargo, para el Cardón y la Pitaya agria, el nicho ambiental de estas especies, definido por las variables utilizadas, no explican la variación observada de la abundancia. Por lo anterior, solo fue posible generar el mapa de abundancias para el Cirio (Fig. 19).

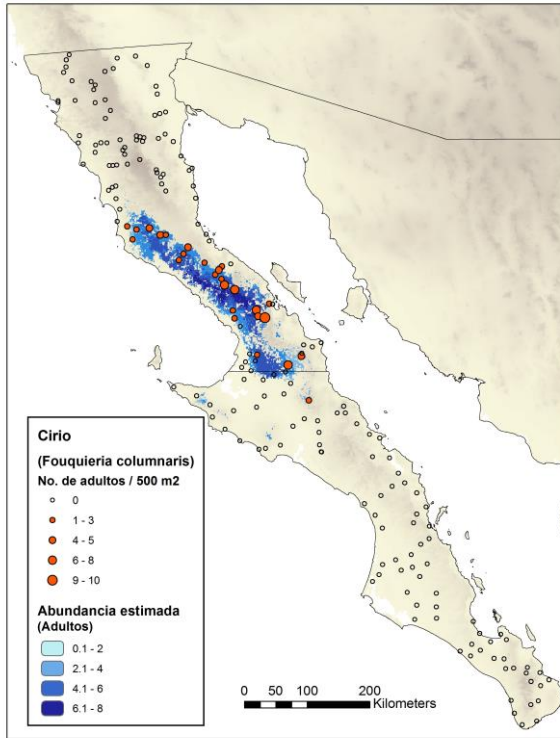


Figura 19. Mapa de distribución de la abundancia estimada del Cirio (*Fouquieria columnaris*).

d) *Validación estadística.*

La capacidad predictiva del modelo de abundancia del Cirio se evaluó mediante un proceso de *bootstrap* generando 100 regresiones entre la distancia al centroide y la abundancia, con particiones aleatorias de los datos de abundancia en una proporción de 70% de los puntos para generar predicciones y el 30% restante para validarlas. Los resultados de este proceso mostraron una  $R^2$  promedio = 0.188, con un rango que va de 0.031-0.513 y una desviación estándar = 0.129 (Fig. 20). A partir de las predicciones de todos los modelos de regresión se calculó y mapeó la incertidumbre en las predicciones, expresada como los intervalos de confianza al 95% para cada uno de los píxeles (Fig. 21).

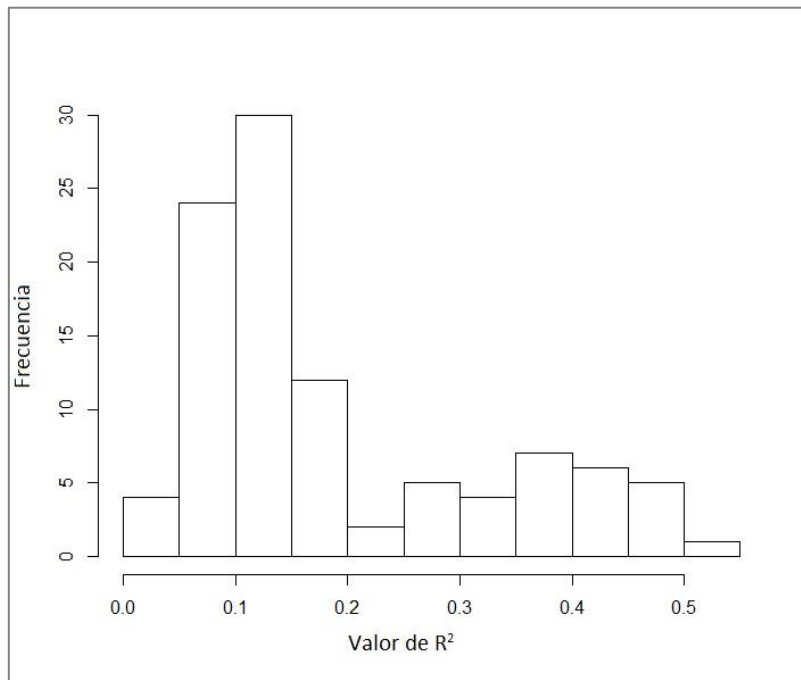
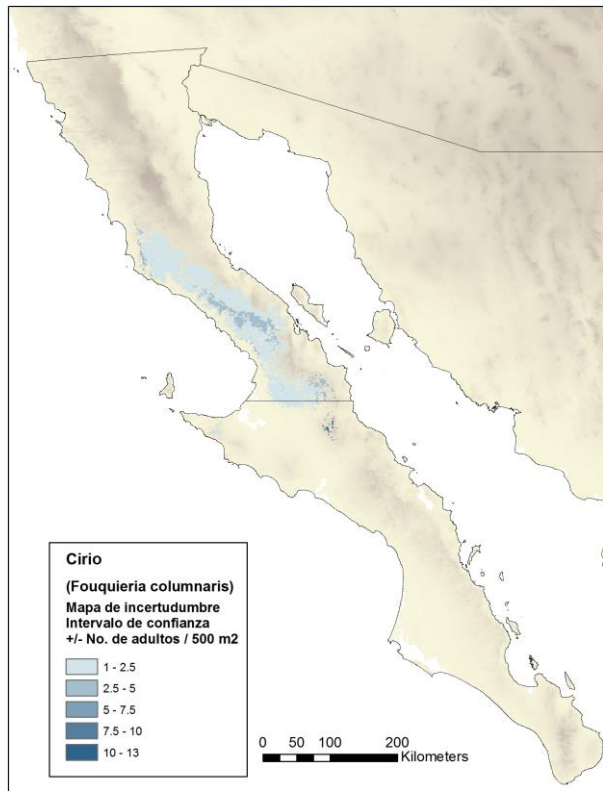


Figura 20. Frecuencia en la variación de los valores del coeficiente de determinación ( $R^2$ ) de 100 regresiones de la distancia al centroide y la abundancia del Cirio (*Fouquieria columnaris*) obtenidas mediante un proceso de *bootstrap* para la validación del modelo de abundancia.

Figura 21. Mapa de incertidumbre (intervalos de confianza al 95%: ind / 500 m<sup>2</sup>) de la abundancia estimada del Cirio (*Fouquieria columnaris*). Los valores más bajos corresponden a menor incertidumbre.



Los resultados obtenidos son sorprendidos pero interesantes. Es notable que a pesar de que los modelos de nicho generados fueron altamente significativos ( $P < 0.001$ ) y además altamente precisos, pues los errores de omisión de los datos de validación fueron bajos y no hubo una sobreestimación notoria del área predicha (Figs. 10-11) (Peterson *et al.* 2011), las variables que explican la distribución no explican los patrones de abundancia.

Encontramos tres posibles explicaciones a esta observación: (1) Que hay una o más variables del nicho que no estamos considerando en el análisis y que son las más influyentes en determinar la abundancia, por ejemplo, las concentraciones de nutrientes u otras características físicas y químicas en el suelo. Para probar esto sería interesante incorporar algunas variables edáficas para investigar su influencia; el problema es que la información edáfica a escalas amplias es muy poco precisa. Se ha mencionado que a escalas locales las variables que más influyen en la distribución y abundancia de las especies son del tipo Eltoniano (*sensu* Soberón 2007), es decir, aquellas que están ligadas dinámicamente con los procesos demográficos de las especies, como son las de recursos (e.g., nutrientes) o las interacciones bióticas (e.g., depredadores, dispersores de semillas, etc.). Este tipo de variables no fueron incorporadas en los análisis porque no existen datos para crearlas en el formato necesario para la modelación. (2) Otra posible explicación a esta falta de relación del nicho ambiental con la abundancia es que la Península de Baja California impone una restricción espacial a la distribución de las especies, funcionando más como una isla. Para el caso del Cardón y de la Pitaya agria, su límite de distribución sureño termina con la Península, literalmente, se les acaba la tierra. De hecho, en el sur de la Península se encontraron abundancias altas para las dos especies. Esto puede provocar que el centroide del nicho existente (*sensu* Peterson *et al.* 2011) no corresponda con el óptimo de estas dos especies. (3) Finalmente, existe la posibilidad de que las abundancias registradas se vean afectadas por algún tipo de perturbación que no detectamos en el campo. Como se mencionó anteriormente, los conteos de individuos se hicieron en sitios a una distancia no mayor de 2 km de las carreteras y terracerías que recorrimos; es posible entonces que para algunos sitios las abundancias observadas estén afectadas por las actividades humanas, a pesar de que procuramos evitar los sitios aparentemente impactados. Para no incurrir en este posible problema sería necesario hacer los muestreos en sitios con el menor impacto posible, lo que implicaría alejarse más de los sitios bajo



influencia humana (i.e., carreteras y poblados), con el respectivo aumento en el esfuerzo y costo en tiempo y recursos. Es necesario hacer un conjunto de análisis diferentes para determinar si alguna de estas tres posibilidades ofrece una explicación a estos resultados.

Los resultados presentados en este proyecto son una oportunidad importante para explorar los factores que determinan los patrones de abundancia de las especies estudiadas. En esta primera exploración analítica observamos que las variables macroclimáticas son las más importantes para determinar la distribución de las tres especies, pero no su abundancia. Sin duda, éste es un campo fértil para la investigación biogeográfica y esperamos que los datos que aquí se presentan sirvan para motivar la investigación en esta área del conocimiento.

## **EL FUTURO**

Como se mencionó anteriormente, los productos generados en este proyecto son apenas el principio de un desarrollo científico y bioinformático en el área de la biogeografía, particularmente en el campo de la modelación de nichos ecológicos (Peterson *et al.* 2011). La modelación de la abundancia es un campo que está generando gran interés en la comunidad científica (Jiménez-Valverde *et al.* 2009, VanDerWal *et al.* 2009, Jiménez-Valverde 2011, Kulhanek *et al.* 2011, Oliver *et al.* 2012; Tôrres *et al.* 2012), por lo que en la medida que avancemos en el entendimiento de los factores que determinan los patrones espaciales de la abundancia, podremos generar herramientas cada vez más precisas y, por ende, más útiles. Estas son las aspiraciones de este grupo de trabajo.

## **LITERATURA CITADA**

- Anderson, R.P., D. Lew & A.T. Peterson. 2003. Evaluating predictive models of species' distributions: Criteria for selecting optimal models. *Ecological Modelling* 162: 211-232.
- Barve, N. 2008. Tool for Partial-ROC. Biodiversity Institute. Lawrence, KS. ver 1.0.
- Brown, J.H. 1984. On the relationship between abundance and distribution of species. *The American Naturalist* 124: 255-79.
- Brown, J.H. 1995. *Macroecology*. University of Chicago Press, Chicago.
- Cuervo-Robayo, A. P., O. Téllez-Valdés, M.A. Gómez-Albores, C.S. Venegas-Barrera, J. Manjarrez, & E. Martínez-Meyer. 2013. An update of high-resolution monthly climate surfaces for Mexico. *International Journal of Climatology*. doi: 10.1002/joc.3848.

- Escalante, T. & E. Martínez-Meyer. 2013. Ecological niche modeling and Wildlife Management Units (UMA's): An application to deer in Campeche, Mexico. *Tropical and Subtropical Agroecosystems* 6: 183-191.
- Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones & A. Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965-1978.
- Hutchinson, G.E. 1957. Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology* 22: 415-427.
- IBM. 2010. IBM SPSS Statistics Version 1.9. New York, USA.
- Jiménez-Valverde, A. 2011. Relationship between local population density and environmental suitability estimated from occurrence data. *Frontiers of Biogeography* 3: 59-61.
- Jiménez-Valverde, A., F. Diniz, E.B. de Azevedo & P.A.V. Borges. 2009. Species distribution models do not account for abundance: the case of arthropods on Terceira Island. *Annales Zoologici Fennici* 46: 451-464.
- Kulhanek, S.A., B. Leung & A. Ricciardi. 2011. Using ecological niche models to predict the abundance and impact of invasive species: application to the common carp. *Ecological Applications* 21: 203–213.
- Maguire Jr., B. 1973. Niche response structure and the analytical potentials of its relationship to the habitat. *American Naturalist* 107: 213-246.
- Martínez-Meyer, E., D.F. Díaz-Porras, A.T. Peterson & C. Yáñez-Arenas. 2013. Ecological niche structure and rangewide abundance patterns of species. *Biology Letters* 9: 20120637.
- Oliver, T.H., S. Gillings M. Girardello, G. Rapacciuolo, T.M. Brereton, G.M. Siriwardena, D.B. Roy, R. Pywell & R.J. Fuller. 2012. Population density but not stability can be predicted from species distribution models. *Journal of Applied Ecology* 49: 581-590.
- Peterson, A.T., M. Papeş, & J. Soberón. 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modelling. *Ecological Modelling* 213: 63-72.
- Peterson, A.T., J. Soberón, R.G. Pearson, R.P. Anderson, E. Martínez Meyer, M. Nakamura & M.B. Araújo. 2011. *Ecological Niches and Geographic Distributions*. Monographs in Population Biology. Princeton University Press, Princeton, N.J. 314 p.
- Phillips S.J., R.P. Anderson & R.E. Schapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecological Modelling* 190: 231-259.
- Rangel, T.F.L.V.B, J.A.F. Diniz-Filho, & L.M. Bini. 2010. SAM: a comprehensive application for Spatial Analysis in Macroecology. *Ecography* 33: 46-50.
- Sagarin, R.D. & S.D. Gaines. 2002. The 'abundant centre' distribution: to what extent is it a biogeographical rule? *Ecology Letters* 5: 137–147.
- Sagarin, R.D., S.D. Gaines & B. Gaylord. 2006. Moving beyond assumptions to understand abundance distributions across the ranges of species. *Trends in Ecology & Evolution* 21: 524-530.
- Soberón, J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. *Ecology Letters* 10: 1115–1123.

- Stockwell, D.R.B. & D.P. Peters. 1999. The GARP modelling system: Problems and solutions to automated spatial prediction. *International Journal of Geographic Information Systems* 13: 143-158.
- Tôrres, N.M., P. De M. Júnior, T. Santos, L. Silveira, A.T. de Almeida-Jácomo & J. A.F. Diniz-Filho. 2012. Can species distribution modelling provide estimates of population densities? A case study with jaguars in the Neotropics. *Diversity & Distributions* 18: 615-627.
- Ureña-Aranda C.A., O. Rojas-Soto, E. Martínez-Meyer, C. Yáñez-Arenas, R. Landgrave Ramírez & A. Espinosa de los Monteros. 2015. Using Range-wide abundance modeling to identify key conservation areas for the micro-endemic Bolson Tortoise (*Gopherus flavomarginatus*). *PlosOne* 10(6): e0131452.
- VanDerWal, J., Shoo, L.P., Johnson, C.N. & Williams, S.E. 2009. Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. *The American Naturalist* 174: 282-291.
- Yáñez-Arenas, C, E. Martínez-Meyer, S. Mandujano & O. Rojas-Soto. 2012. Modelling geographic patterns of population density of the white-tailed deer in central Mexico by implementing ecological niche theory. *Oikos* 121: 2081-2089.
- USGS. 2001. HYDRO1k elevation derivative database, U.S. geological survey. Available at: <http://edcdaac.usgs.gov/gtopo30/hydro/>.

## ANEXOS

1. Instructivo para el uso del programa *Nichosfera* para Java.
2. Instructivo para el uso de la extensión *Nichosfera* para ArcView 3.x.
3. Instructivo para el uso de la extensión *Nichosfera* para ArcGis 10.2 o superior.
4. Instructivo para el uso de *Nichosfera* para R.
5. Procedimiento de *bootstrap* para la validación de los modelos de abundancia para R.

# Anexo 1. Instructivo para el uso del programa *Nichosfera* para Java v2.1

Desarrollado por John Giddings

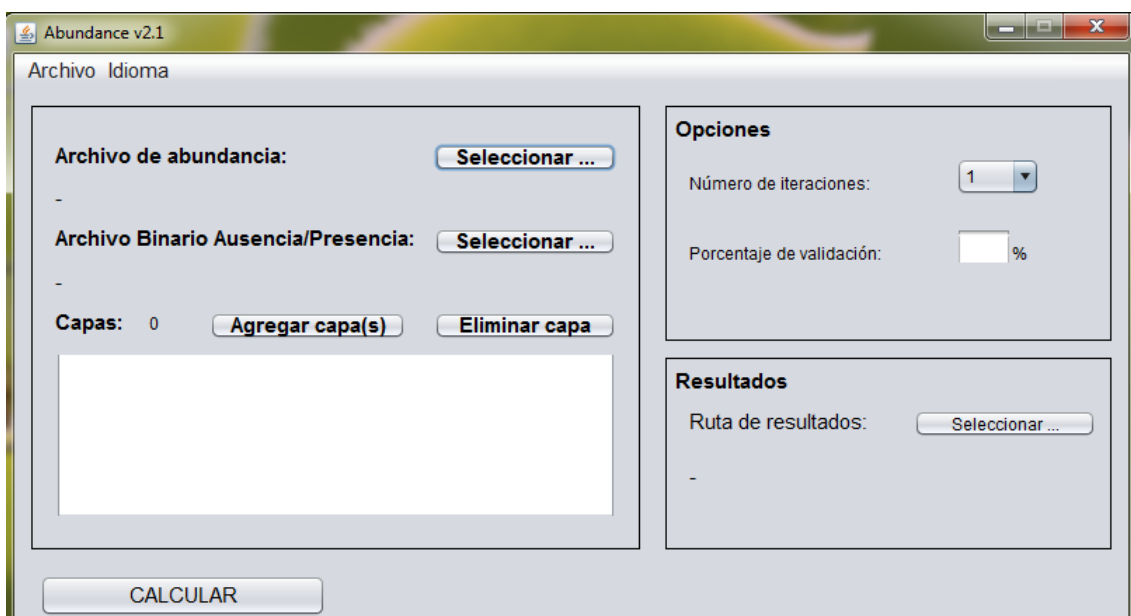
(john.giddings@gmail.com)

Requerimientos del sistema:

- Sistema Operativo Windows, Mac, Linux
- Tener instalado Java 1.6 o superior

En este documento se describen los pasos a seguir para usar el programa *Nichosfera* para Java v2.1 para el cálculo de las distancias al centroide del nicho de cualquier especie. Para una descripción detallada del método y sus bases teóricas consultar Martínez-Meyer *et al.* (2013) y Yañez-Arenas *et al.* (2012).

Para hacer funcionar el programa en Windows, hacer doble clic en el nombre archivo *Nichosfera.exe*. Para Mac y Linux, abrir el archivo por línea de comandos desde la Terminal. Si el archivo no abre, verificar que esté instalado Java 1.6 o superior.



Una vez abierto, los datos de entrada son los siguientes:

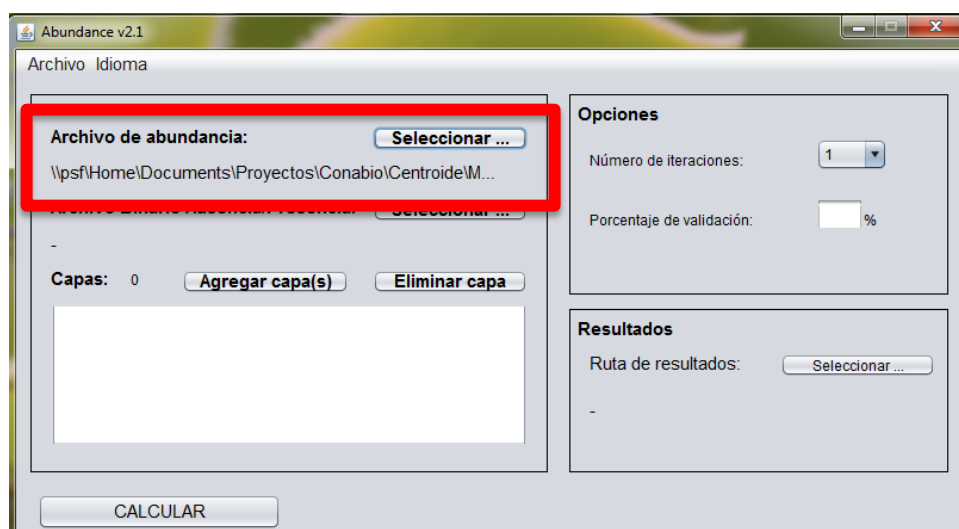
1. Archivo de registros de presencia y abundancia de una especie. Este archivo debe de estar en formato ascii separado por comas (.csv) y debe contener cuatro columnas de la siguiente forma:

Especie,Longitud,Latitud,Abundancia

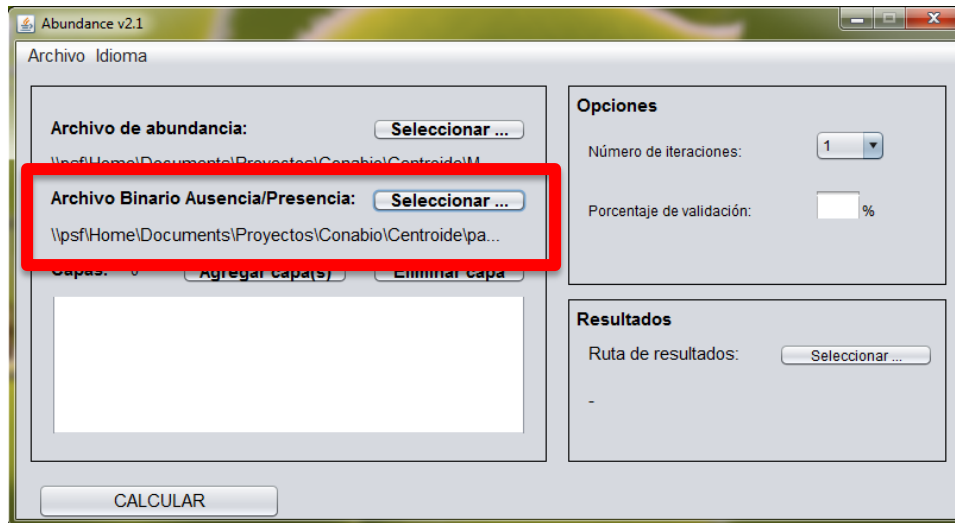
Cirio,-115.64811,30.08662,11

Cirio,-114.2977,29.32943,3

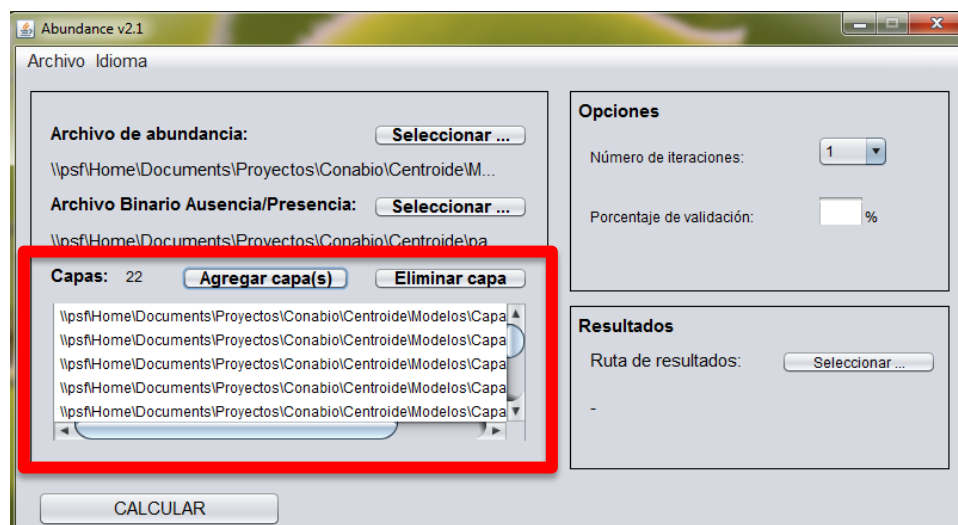
....



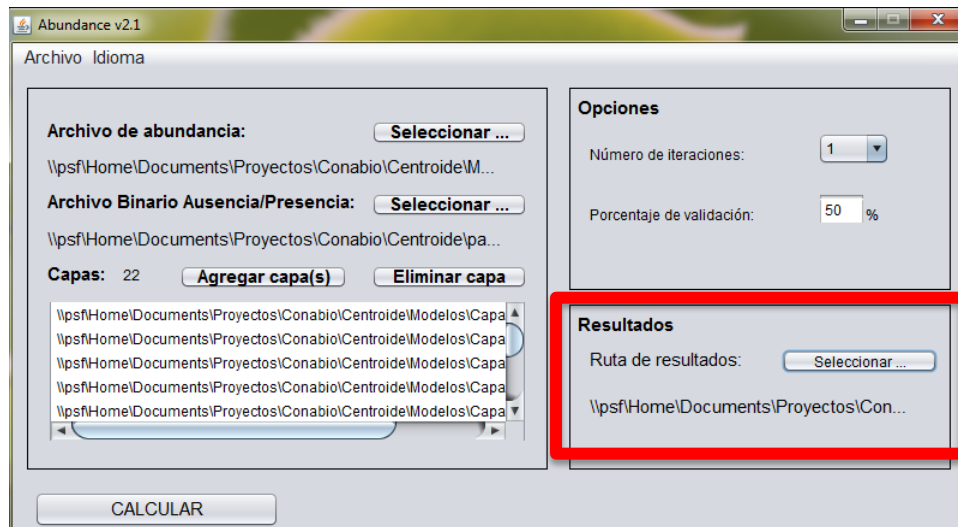
2. Mapa ráster de distribución potencial de la especie (1 = presencia, 0 = ausencia) en formato ascii-grid (.asc). Este mapa debió haber sido generado previamente con algún método convencional de modelación de nichos (e.g., GLM, GAM, GARP, MaxEnt, Bioclim, etc.) y procesado para convertirlo en binario.



3. Capas ambientales que conforman los ejes del nicho ecológico de la especie en formato ascii-grid (.asc). Generalmente, estas variables son las mismas que se usaron para generar el modelo de distribución potencial.



4. Especificar el directorio en donde se guardarán los resultados.



5. Una vez ingresados los datos, dar clic en “Calcular” para producir el modelo.

Cuando ha concluido el proceso, en la parte inferior derecha se encuentra el resultado de las regresiones lineal, logarítmica y exponencial. El directorio de salida contiene dos archivos, uno que es el mapa de las distancias al centroide (*nombre.asc*) y el archivo de los registros de presencia/abundancias (*nombre.csv*) con una columna adicional que contiene las abundancias estimadas.

## REFERENCIAS

- Martínez-Meyer, E., D.F. Díaz-Porras, A.T. Peterson & C. Yáñez-Arenas. 2013. Ecological niche structure and rangewide abundance patterns of species. *Biology Letters* 9: 20120637.
- Yáñez-Arenas, C, E. Martínez-Meyer, S. Mandujano & O. Rojas-Soto. 2012. Modelling geographic patterns of population density of the white-tailed deer in central Mexico by implementing ecological niche theory. *Oikos* 121: 2081-2089.

*Citar como: Giddings, J., R. Landgrave, C. Yáñez-Arenas, O. Rojas-Soto & E. Martínez-Meyer. 2014. Nichosfera para Java v2.1. Distrito Federal, México.*



## **Anexo 2. Instructivo para el uso del programa *Nichosfera* para ArcView**

### **3.x**

Desarrollado por Rosario Landgrave

(rosario.landgrave@inecol.mx)

Requerimientos del sistema:

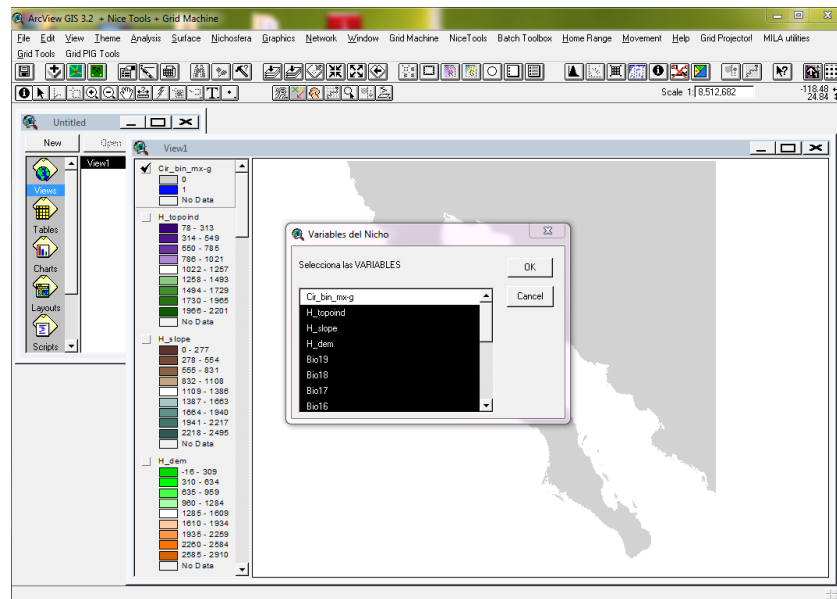
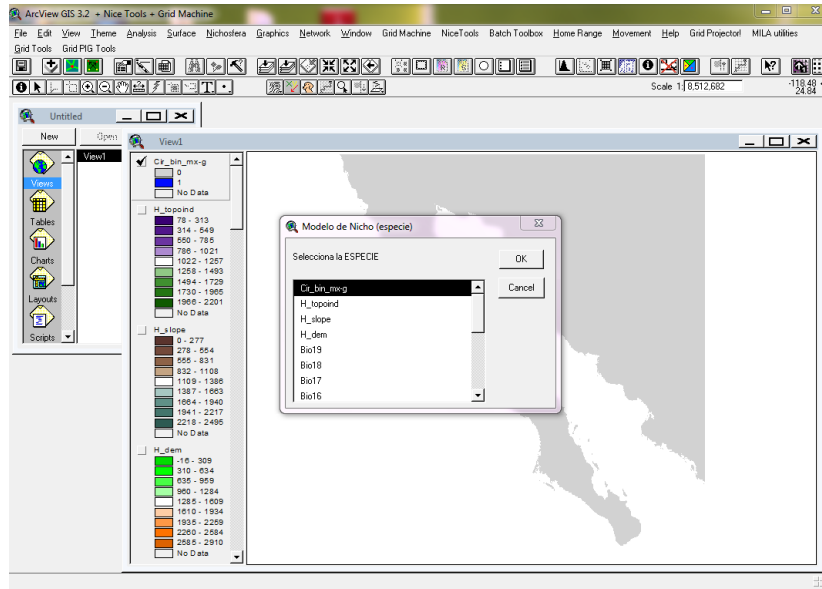
- Sistema Operativo Windows
- Tener instalado ArcView 3.x con la extensión Spatial Analyst

En este documento se describen los pasos a seguir para usar el programa *Nichosfera* para ArcView 3.x para el cálculo de las distancias al centroide del nicho de cualquier especie. Para una descripción detallada del método y sus bases teóricas consultar Martínez-Meyer *et al.* (2013) y Yañez-Arenas *et al.* (2012).

1. Guardar el archivo *Nichosfera.avx* en /ESRI/AV\_GIS30/ARCVIEW/EXT32/
2. Abrir ArcView y activar la extensión *Nichosfera*. Una vez activada, aparecerá en el menú la opción *Nichosfera*. Nota: Es necesario que la extensión Spatial Analyst esté instalada y activada.
3. Los datos que sirven de insumo son: (1) un mapa raster binario (1 = presencia, 0 = ausencia) en formato ArcInfo grid de la distribución potencial de la especie. Este mapa debió haber sido generado previamente con algún método convencional de modelación de nichos (e.g., GLM, GAM, GARP, MaxEnt, Bioclim, etc.) y procesado para convertirlo en binario. (2) Capas ambientales que conforman los ejes del nicho ecológico de la especie en formato ArcInfo grid. Generalmente, estas variables son las mismas que se usaron para generar el modelo de distribución potencial. Los dos conjuntos de datos tienen que ser cargados en la vista.

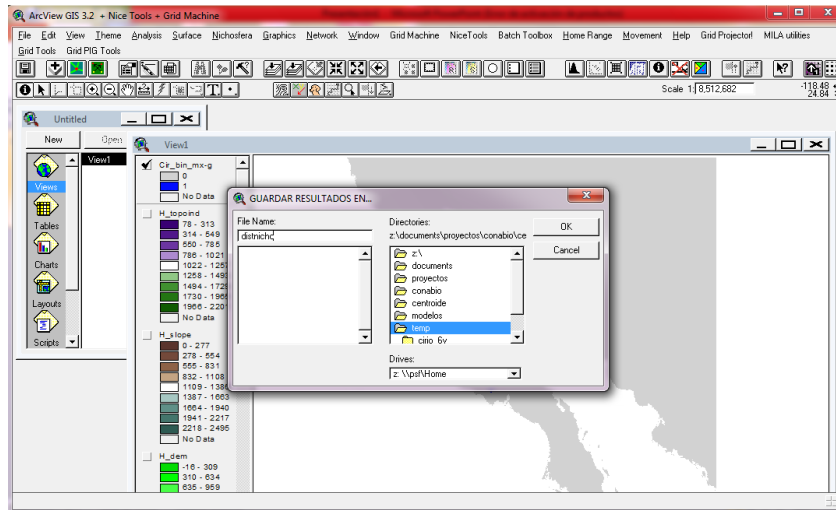
4. Ya con los datos cargados en la vista, en la barra de menú seleccionar Nichosfera-Distancia al centroide del nicho para que abra la utilería.

5. La primera ventana solicita que se seleccione el mapa binario de la especie.

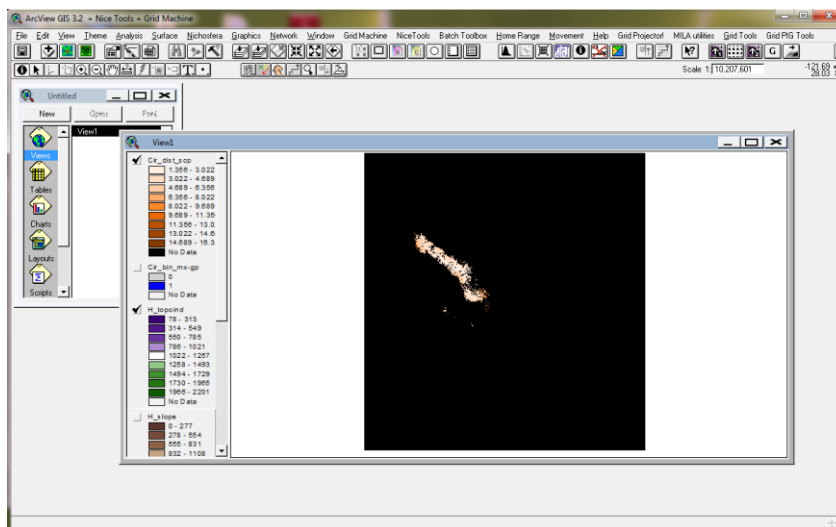


6. Una vez seleccionado, la siguiente ventana solicita seleccionar las variables del nicho.

7. La última ventana solicita el nombre de archivo de salida. El usuario puede seleccionar el directorio en donde guardarlo.



8. Finalmente, el mapa generado con las distancias al centroide puede ser cargado a la vista.



## REFERENCIAS

Martínez-Meyer, E., D.F. Díaz-Porras, A.T. Peterson & C. Yáñez-Arenas. 2013. Ecological niche structure and rangewide abundance patterns of species. *Biology Letters* 9: 20120637.

Yáñez-Arenas, C, E. Martínez-Meyer, S. Mandujano & O. Rojas-Soto. 2012. Modelling geographic patterns of population density of the white-tailed deer in central Mexico by implementing ecological niche theory. *Oikos* 121: 2081-2089.

*Citar como: R. Landgrave, J. Giddings, C. Yáñez-Arenas, O. Rojas-Soto & E. Martínez-Meyer. 2014. Nichosfera para ArcView 3.x. Xalapa, México.*

## Anexo 3. Instructivo para el uso del programa *Nichosfera* para ArcGis

### 10.2 o superior

Desarrollado por Rosario Landgrave

(rosario.landgrave@inecol.mx)

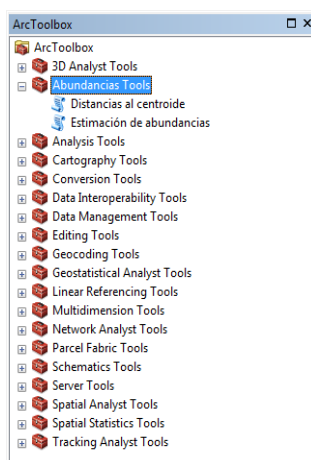
Requerimientos del sistema:

- Sistema Operativo Windows
- Tener instalado ArcGis 10.2 o superior con la extensión Spatial Analyst

En este documento se describen los pasos a seguir para usar el programa *Nichosfera* para ArcGis 10.2 o superior para el cálculo de las distancias al centroide del nicho de cualquier especie. Para una descripción detallada del método y sus bases teóricas consultar Martínez-Meyer *et al.* (2013) y Yañez-Arenas *et al.* (2012).

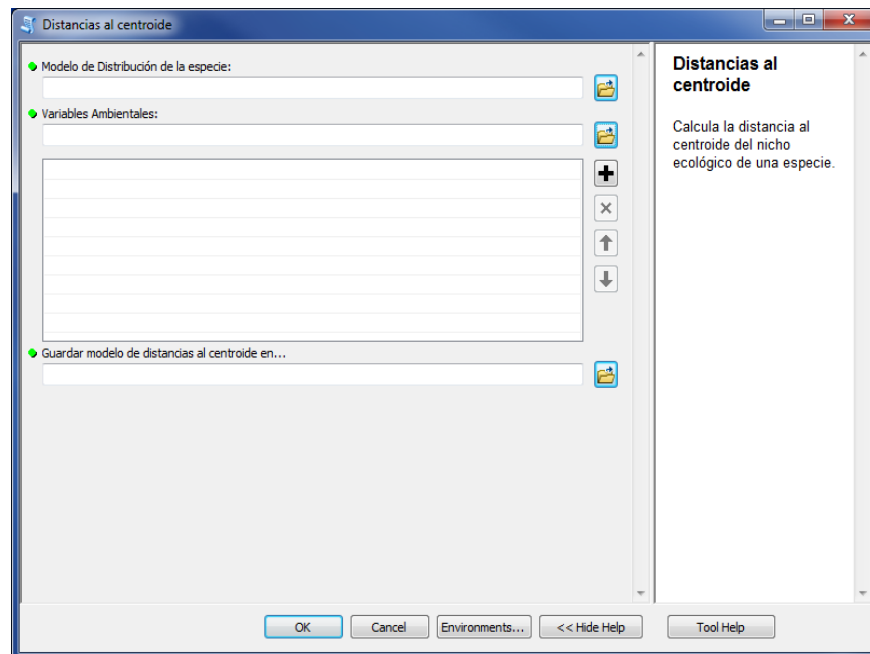
La extensión *Nichosfera.tbx* contiene dos scripts: (1) “distancias al centroide”, que sirve únicamente para calcular y generar un mapa de las distancias al centroide y (2) “estimación de abundancias” que calcula las distancias al centroide y además estima las abundancias (linealmente).

Para instalar la herramienta se abre ArcMap y se abre la ventana de “Toolbox”. Dar clic al botón derecho y aparecerá un recuadro en donde se debe indicar la ruta donde se encuentra el script, posteriormente aceptar y estará listo para usarse.



El script “Distancias al centroide” requiere como insumos: (1) un mapa ráster binario (1 = presencia, 0 = ausencia) de la distribución potencial de la especie. Este mapa debió haber sido generado previamente con algún método convencional de modelación de nichos (e.g., GLM, GAM, GARP, MaxEnt, Bioclim, etc.) y procesado para convertirlo en binario. (2) Las capas ambientales que conforman los ejes del nicho ecológico de la especie. Generalmente, estas variables son las mismas que se usaron para generar el modelo de distribución potencial. Este script funciona para las versiones de ArcGis 10.2 y superior.

Al ejecutar el script se despliega una ventana de diálogo, en la que el usuario introduce los parámetros necesarios:

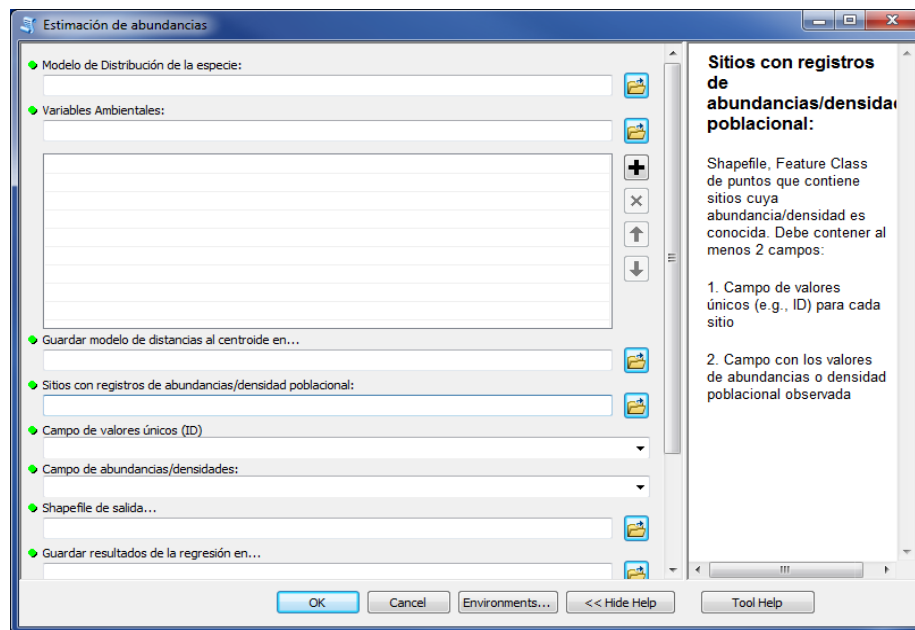


Las ventanas del lado derecho contienen la descripción de cada uno de los parámetros que requiere el script.

El script “Estimación de abundancias” requiere como insumos los dos conjuntos de datos que el script de “Distancias al centroide”, es decir, el mapa ráster binario de la distribución potencial de la especie y las capas ambientales que conforman los ejes del nicho ecológico de la especie. Además, un archivo en formato shapefile o feature class con

los registros de presencia y abundancia de la especie. Este script funciona sólo para la versión ArcGIS 10.2 o superior debido a que uno de los productos de salida es un archivo PDF. Al ejecutar el script se despliega una ventana de diálogo en la que el usuario introduce los parámetros necesarios:

Al igual que en el script anterior, las ventanas del lado derecho contienen la descripción de cada uno de los parámetros que requiere el script.



## REFERENCIAS

Martínez-Meyer, E., D.F. Díaz-Porras, A.T. Peterson & C. Yáñez-Arenas. 2013. Ecological niche structure and rangewide abundance patterns of species. *Biology Letters* 9: 20120637.

Yáñez-Arenas, C, E. Martínez-Meyer, S. Mandujano & O. Rojas-Soto. 2012. Modelling geographic patterns of population density of the white-tailed deer in central Mexico by implementing ecological niche theory. *Oikos* 121: 2081-2089.

*Citar como: R. Landgrave, J. Giddings, C. Yáñez-Arenas, O. Rojas-Soto & E. Martínez-Meyer. 2014. Nichosfera para ArcGIS 10.x. Xalapa, México.*

## Anexo 4. Instructivo para el uso de *Nichosfera* para R

Desarrollado por Enrique Martínez Meyer

(emm@ib.unam.mx)

Requerimientos del sistema:

- Sistema Operativo Windows, Mac, Linux
- Tener instalado R 3.0 o superior

En el presente documento se describen los comandos de un *script* en el programa estadístico R desarrollado para el cálculo de la distancia al centroide del nicho de cualquier especie. Consultar Martínez-Meyer *et al.* (2013) y Yañez-Arenas *et al.* (2012) para obtener mayores detalles conceptuales y teóricos sobre éste método.

Los insumos para la función son dos: (1) capas ambientales en formato ascii ráster que resuman la variables ambientales (*e.g.*, las superficies climáticas del proyecto WorldClim; Hijmans *et al.* 2005) que forman parte del nicho de la especie de interés. (2) Modelo binario (0 – ausencia, 1 – presencia) que represente la distribución de la especie en formato ascii ráster. Este debe haber sido generado previamente con algún algoritmo de modelado de nicho (*e.g.*, GLM, GAM, GARP, MaxEnt, Bioclim, etc.).

En R es necesario tener instalados los paquetes “raster” y “rgdal” para que funcione:

```
install.packages (“raster”)
```

```
install.packages (“rgdal”)
```

Si ya se tienen instalados, sólo es necesario activar las librerías:

```
library(raster)
```

```
library(rgdal)
```

Una vez instaladas las librerías se ejecuta la función de la siguiente manera:



1.- Copiar y pegar los siguientes comandos en la consola de R.

```
Nichosfera<-function(nombre)  
  
  {  
  
    r_binario = choose.files(caption="Select ASCII file: ")  
      rbin <- raster(r_binario)  
        plot(rbin)  
      capas = choose.files(caption="Select ASCII file: ")  
        stk = stack(capas)  
      msh = mask(stk, rbin, maskvalue = 0)  
        media = cellStats(msh,mean)  
          num = msh - media  
            desv = cellStats(msh,sd)  
              pred_est = num/desv  
                r_dc = sqrt(sum(pred_est^2))  
                  setwd(choose.dir())  
                    writeRaster(r_dc, nombre, 'ascii')  
                      }  
  
    # Nichosfera("especie")
```

2. En la última línea **#Nichosfera("especie")** borrar **#** y cambiar “especie” por el nombre que se quiera asignar al modelo de distancias al centroide. Por ejemplo, si se están calculando las distancias al centroide del nicho del jaguar: **Nichosfera("jaguar\_dc")**. Seleccionar *enter* o ejecutar la última línea.

3. La función automáticamente abrirá una primera ventana para buscar el modelo binario de distribución de la especie en formato ascii. Se busca el modelo, se selecciona y al abrirlo se genera un gráfico del modelo en una ventana dentro de R.

4. Aparece una segunda ventana para buscar las capas ambientales que conforman las dimensiones del nicho de la especie (también en formato ascii). Se pueden seleccionar todas las coberturas al mismo tiempo y se presiona el botón *abrir*.
5. Finalmente se genera un gráfico del modelo de distancias al centroide del nicho y aparece una tercera ventana en la cual se debe buscar una ruta y seleccionar una carpeta donde guardar el resultado. Una vez seleccionada la carpeta, se acepta. Se guarda el modelo en esta carpeta en formato ascii y en la consola R se presentan algunos metadatos del modelo.

## REFERENCIAS

- Hijmans, R.J., S.E. Cameron, J.L. Parra, P.G. Jones & A. Jarvis, A. 2005. Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* 25: 1965-1978.
- Martínez-Meyer, E., D.F. Díaz-Porras, A.T. Peterson & C. Yáñez-Arenas. 2013. Ecological niche structure and rangewide abundance patterns of species. *Biology Letters* 9: 20120637.
- Yáñez-Arenas, C, E. Martínez-Meyer, S. Mandujano & O. Rojas-Soto. 2012. Modelling geographic patterns of population density of the white-tailed deer in central Mexico by implementing ecological niche theory. *Oikos* 121: 2081-2089.

*Citar como: Yáñez-Arenas, C. J. Giddings, R. Landgrave, O. Rojas-Soto & E. Martínez-Meyer. 2014. Nichosfera para R. Xalapa, México.*

## **Anexo 5. Regresión y mapeo de incertidumbre por medio de un proceso de *Bootstrap* de la predicción de abundancias por el método de Distancia al Centroide (DC) para R.**

Desarrollado por Carlos A. Yañez Arenas  
(lichoso@gmail.com)

En el presente documento se describen los comandos de un *script* en el programa estadístico R desarrollado para evaluar la capacidad predictiva de los modelos de abundancia por medio del método de distancia al centroide del nicho ecológico (DC). Consultar Martínez-Meyer *et al.* (2013) y Yañez-Arenas *et al.* (2012) para obtener mayores detalles conceptuales sobre el método.

El *script* funciona con una base de datos en formato ".csv" la cuál debe ser extraída a partir de la tabla de atributos de un ráster con las distancias al centroide del nicho (DC) que tenga como extensión geográfica la distribución potencial de una especie. La tabla de atributos de este ráster por lo general está compuesta por cuatro columnas que son: 1) *value* = un valor de identificación (p. ej. 1, 2, 3 ...), 2) *LONG* = la coordenada longitud, 3) *LAT* = la coordenada latitud y 4) *DC* = la distancia al centroide del nicho calculada. Las filas son cada pixel del ráster y el número dependerá de la extensión de la distribución de la especie y la resolución del ráster. A la tabla de atributos de este ráster de DC se le debe agregar una quinta columna que tenga valores de abundancia/densidad poblacional de la especie en cuestión. Estos valores de preferencia deben obtenerse en campo para que tengan el mismo sesgo de muestreo. También, de preferencia deben ser tomados aleatoriamente y tratando de abarcar la mayor cantidad de combinaciones ambientales posibles en las que la especie se encuentra, es decir se debe tener una buena caracterización del nicho. Los datos de abundancia/densidad deberán estar geo-referenciados para poder ubicarlos en sus respectivas filas de la tabla del ráster. Al resto de las filas (que equivalen a cada pixel del ráster) se les deberá asignar el valor de NA (sin datos).

La base de datos en formato ".csv" descrita en el párrafo anterior es la que se necesita para correr el *script* de R. A continuación se describirán los comandos de éste *script*:

1.- Se lee el archivo, es decir la base de datos mencionada previamente en formato ".csv". Se crea un objeto que en este ejemplo se llama "datos" y que contiene la información de la base.

```
datos <- read.csv("/ruta/base.csv", header = TRUE, sep = ",")
```

Se adjunta y se le indica que enliste el nombre de cada columna (variables) contenida en el objeto creado (éstas deberán ser las mismas que constituyen la base original).

```
attach(datos)
```

```
names(datos)
```

2.- Se genera un objeto donde se seleccionarán únicamente los renglones que poseen datos de abundancia/densidad poblacional. Por ejemplo, supóngase se tienen 50 localidades con una estimación de abundancia o densidad poblacional. En el siguiente comando se indica que se seleccionen estos 50 datos y las columnas 4 y 5 (DC y Densidad). Se adjunta ésta sub-muestra y se revisa que sean las columnas indicadas.

```
datos_O <- datos[1:50, 4:5]
```

```
attach(datos_O)
```

```
names(datos_O)
```

3.- Se crean dos objetos que serán matrices y contendrán toda la información de la tabla de atributos original (el número de renglones o filas será el total de píxeles en el ráster de distancias al centroide).

```
SXX <- numeric(dim(datos)[1])
```

```
SX <- numeric(dim(datos)[1])
```

4.- Se desarrolla un proceso automatizado de remuestreos que se repetirá 100 veces (el usuario decide cuantas iteraciones) y que consiste en lo siguiente: a) del conjunto de observaciones (en este caso son 50) con registro de abundancia/densidad poblacional se selecciona aleatoriamente en cada repetición una sub-muestra de 70% datos (en este caso 35, pero el usuario puede decidir la proporción de datos de calibración y validación y si quiere que el remuestreo sea con o sin remplazo), b) de esta sub-muestra se crean dos objetos "Y" y "X", el primero contiene los valores de abundancia/densidad mientras que el segundo contiene las distancias al centroide del nicho, c) se aplica un modelo de regresión lineal y se guarda en un objeto que en el presente *script* es nombrado "M", d) se generan las predicciones para el modelo de regresión y se guardan en las matrices (SXX y SX) creadas en el punto número 3.

```

for(i in 1:1000){
MUESTRA <- sample(1:50, size = 35, replace = FALSE)
Y <- datos_O[MUESTRA,2]
X <- datos_O[MUESTRA, 1]
M <- lm(Y ~ X)
PRED <- predict(M, data.frame(X = datos$DC))
SXX <- SXX + PRED^2
SX <- SX + PRED}

```

5.- A partir de las predicciones de todos los modelos de regresión se calcula la varianza, la desviación estándar, la media, el coeficiente de variación y los intervalos de confianza.

```

VAR <- (SXX - SX^2/i) / (i - 1)
SD <- sqrt(VAR)
MEDIA <- SX/i
CV <- SD/MEDIA
IC95 <- SD*qt(0.975, 8)

```

6.- Finalmente, se escribe toda la información generada en una tabla (reemplazar “ruta/nombre de archivo” por la ruta en donde está la carpeta de destino y el nombre del archivo que se desee) que contendrá el mismo número de renglones que la tabla original de atributos del ráster de distancias, pero que además tendrá calculado para cada pixel los parámetros descritos en el punto anterior.

**write.table(data.frame(datos, MEDIA, VAR, SD, IC95), "ruta/nombre de  
archivo.csv", sep = ",", row.names = FALSE)**

## REFERENCIAS

Martínez-Meyer, E., Díaz-Porras, D. F., Peterson, A. T. & Yañez-Arenas, C. (2013). Ecological niche structure and rangewide abundance patterns of species. *BiolLett* 9.

Yañez-Arenas, C., Martínez-Meyer, E., Mandujano, S. & Rojas-Soto, O. (2012). Modelling geographic patterns of population density of the white-tailed deer in central Mexico by implementing ecological niche theory. *Oikos* 121, 2081–2089.

*Citar como: Yañez-Arenas, C. J. Giddings, R. Landgrave, O. Rojas-Soto & E. Martínez-Meyer. 2014. Regresión y Mapeo de Incertidumbre en la Predicción de Abundancias por el Método de Distancia al Centroides (DC) para R. Xalapa, México.*