



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
FACULTAD DE CIENCIAS
ECOLOGÍA

Sobre los mecanismos que determinan la
distribución geográfica de una especie:
idoneidad del hábitat, dispersión y barreras bióticas

TESIS

QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS

PRESENTA:

LUIS ALFREDO OSORIO OLVERA

TUTOR:

DR. MANUEL JESÚS FALCONI MAGAÑA
FACULTAD DE CIENCIAS

COMITÉ TUTOR:

DR. JORGE SOBERÓN MAINERO
FACULTAD DE CIENCIAS; BIODIVERSITY INSTITUTE, KU

DR. JUAN SERVANDO NÚÑEZ FARFÁN
INSTITUTO DE ECOLOGÍA

MÉXICO, CD. MX. MAYO 2018

CIUDAD UNIVERSITARIA, CD. MX., MAYO 2018



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
POSGRADO EN CIENCIAS BIOLÓGICAS
FACULTAD DE CIENCIAS
ECOLOGÍA

Sobre los mecanismos que determinan la
distribución geográfica de una especie:
idoneidad del hábitat, dispersión y barreras bióticas

TESIS

QUE PARA OPTAR POR EL GRADO DE:
DOCTOR EN CIENCIAS

PRESENTA:

LUIS ALFREDO OSORIO OLVERA

TUTOR:

DR. MANUEL JESÚS FALCONI MAGAÑA
FACULTAD DE CIENCIAS

COMITÉ TUTOR:

DR. JORGE SOBERÓN MAINERO
FACULTAD DE CIENCIAS; BIODIVERSITY INSTITUTE, KU

DR. JUAN SERVANDO NÚÑEZ FARFÁN
INSTITUTO DE ECOLOGÍA

MÉXICO, CD. MX. MAYO 2018

POSGRADO EN CIENCIAS BIOLÓGICAS
FACULTAD DE CIENCIAS
DIVISIÓN ACADÉMICA DE INVESTIGACIÓN Y POSGRADO

OFICIO FCIE/DEP/436/2018

ASUNTO: Oficio de Jurado

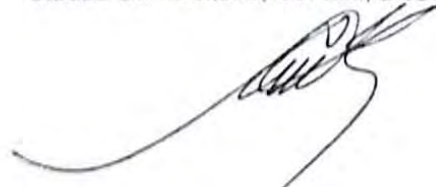
Lic. Ivonne Ramírez Wence
Directora General de Administración Escolar, UNAM
P r e s e n t e

Me permito informar a usted que en la reunión ordinaria del Comité Académico del Posgrado en Ciencias Biológicas, celebrada el día **26 de febrero de 2018**, se aprobó el siguiente jurado para el examen de grado de **DOCTOR EN CIENCIAS** del (la) alumno (a) **OSORIO OLVERA LUIS ALFREDO** con número de cuenta **511010644** con la tesis titulada: "**Sobre los mecanismos que determinan la distribución geográfica de una especie: idoneidad del hábitat, dispersión y barreras bióticas**", realizada bajo la dirección del (la) **DR. MANUEL JESÚS FALCONI MAGAÑA**:

Presidente:	DR. LUIS ANTONIO SÁNCHEZ GONZÁLEZ
Vocal:	DR. PABLO PADILLA LONGORIA
Secretario:	DR. JUAN SERVANDO NÚÑEZ FARFÁN
Suplente:	DR. CARLOS ALBERTO YÁÑEZ ARENAS
Suplente:	DR. MIGUEL ÁNGEL MARTÍNEZ MORALES

Sin otro particular, me es grato enviarle un cordial saludo.

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPIRITU"
Ciudad Universitaria, Cd. Mx., a 03 de mayo de 2018



DR. ADOLFO GERARDO NAVARRO SIGÜENZA
COORDINADOR DEL PROGRAMA



AGNS/VMVA/ASR/lpp

AGRADECIMIENTOS

- Al Consejo Nacional de Ciencia y Tecnología por la beca otorgada para llevar a cabo mis estudios de Doctorado en Ciencias Biológicas (beca 245648).
- Al Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) por el apoyo económico otorgado (proyectos IN112715 e IN116018).
- Al Posgrado en Ciencias Biológicas y a la Facultad de Ciencias de la UNAM por coadyuvar en mi formación y crecimiento académico.
- A mis tutores, el Dr. Manuel Jesús Falconi Magaña y al Dr. Jorge Soberón Mainero (mis padres académicos y amigos), por su apoyo invaluable a lo largo de todos estos años de mi formación académica, cada uno de sus comentarios y consejos fueron realmente constructivos para mi desarrollo profesional y personal. Quiero externarles que, sin su infinito amor a la investigación y a la enseñanza de la ciencia, esta tesis no se hubiera logrado, sus ideas están embebidas en cada capítulo de este manuscrito.
- Al Dr. Juan Núñez Farfán por cada una de sus observaciones y comentarios, sin duda esta tesis no hubiera sido lo mismo sin ellos.
- Al Dr. A. Townsend Peterson por su guía académica (sin ser tutor oficial) y la invitación que me realizó para participar en su increíble grupo de investigación.
- A cada uno de los miembros del comité revisor de esta tesis: Dres. Luis Antonio Sánchez González, Pablo Padilla Longoria, Miguel Ángel Morales Martínez y Carlos Yáñez Arenas. Les agradezco su tiempo, cooperación y cada una de sus atinadas observaciones para esta tesis.
- Al Dr. Enrique Martínez Meyer por sus valiosas ideas sobre los nichos y la abundancia y también por ser un gran amigo.
- A mis colegas de la Escuela Gualdalupana de Nichos: Jorge Soberón, Town Peterson, Enrique Martínez Meyer, Andrés Lira Noriega, Octavio Rojas, Fabricio Villalobos, Carlos Yáñez Arenas.
- A mis amigos del Departamento de Matemáticas de la Facultad de Ciencias: Álvaro, Cristóbal, Luis Miguel, Peregrino, Gama, MaFer, Nashiely y Rosbelfa por los buenos momentos compartidos.
- A mis amigos del Instituto de Biología: Ángela "Angelita" Nava, Delia "Bolu" Basanta, Claudia "Yaya" Moreno Arzate, Anny Meneses, Edith "Piliberta" Calixto, Julián Velasco, Abi Haro, Jorge García, Carlos "Flacarlos" Muñoz, Carlos "Chicken" Luna, Gloria Ponce y Ángela Cuervo, por los divertidos momentos compartidos y su apoyo académico.
- A mis amigos Javier, Sac, Augusto, Lenin y Adrián por brindarme su amistad durante todo este tiempo.
- A mis amigos y compañeros de CONABIO: Mariana Munguía, Oliver López, Julián Equihua, Fernando Prado, Tere, Wolke Tobon y Óscar Godínez.
- A mis primos: Yolis, Sandy, Nicté y Lita por hacer más llevadero este trayecto.
- A mis tíos: Beto, Pedro y Lulú por su apoyo moral.

AGRADECIMIENTOS A TÍTULO PERSONAL

- Agradezco a mis padres, Alfredo y Chelín, por su apoyo incondicional a lo largo de mi proyecto de vida, por ser un ejemplo de trabajo, honradez y amor.
- A Lauri (mi hermana) por ser una gran amiga quien siempre ha visto por mí, te quiero mucho.
- A mis sobrinitos, Julián y Sebastián, por ser la tierna e inocente alegría de mi núcleo familiar.
- A Francisco por su amistad y por cuidar de mis amados sobrinos.
- Finalmente, quiero agradecer a Rusby por su amor y cariño; tu compañía es lo más agradable y dulce que tengo, te adoro con todo mi corazón y pensamiento.

Índice general

§0.1. Resumen general	4
§0.2. Abstract	6
§0.3. Introducción general	7
1. Los nichos ecológicos y las áreas de distribución	11
§1.1. Resumen	11
§1.2. Introducción	11
§1.3. Artículo: Diferencias conceptuales entre modelación de nichos y áreas de distribución.	12
2. El proceso de dispersión: de los organismos a las distribuciones geográficas	18
§2.1. Introducción	18
§2.2. Causas y consecuencias de la dispersión (un contexto ecológico)	19
§2.3. La dispersión en los diferentes grupos de organismos	20
§2.4. La dispersión en el contexto biogeográfico y filogenético:	23
§2.4.1. Los procesos, factores ecológicos en la distribución de las especies	24
§2.4.2. Los factores y procesos históricos en la distribución de las especies	27
§2.4.3. Sobre el papel de la ecología, la dispersión y la vicarianza en los patrones filogenéticos	27
§2.4.4. Reflexiones para la modelación de estos fenómenos	29
§2.5. Conclusiones	31
§2.6. Artículo: Sobre la relación entre idoneidad del hábitat y la abundancia poblacional bajo diferentes escenarios de dispersión.	32
3. Abundancia y posición en el espacio de nicho: procesos y medidas	42
§3.1. Introducción	42
§3.2. Modelo matemático	44
§3.2.1. Modelando la idoneidad del hábitat	45
§3.2.2. Kernel de dispersión	47
§3.3. Método	47
§3.3.1. Método para determinar los efectos de la métrica en la relación y del tiempo que llevan las poblaciones establecidas en las localidades	47
§3.3.2. Método para estudiar el efecto de la estructura espacial de las idoneidades en la relación	48
§3.4. Resultados	48
§3.4.1. Efecto de la métrica y del tiempo de inicio de la dispersión.	48
§3.4.2. Efecto de la estructura espacial de las idoneidades	49

§3.4.3. Barrera biológica a la dispersión	51
§3.5. Discusión	52
§3.5.1. Métrica de distancia	53
§3.5.1.1. Relación entre la salida cruda (raw output) de Maxent y la distancia al centroide del nicho	54
§3.5.2. Estructura de las idoneidades en el espacio geográfico.	54
§3.5.3. Barrera biológica a la dispersión (efecto Allee)	54
§3.6. Artículo: A Theoretical Explanation for the Relationship Between Population Abun- dance and Niche Structure, and When to Expect Exceptions.	55
4. Herramientas informáticas para la modelación de la distribución de las especies	94
§4.1. Introducción	94
§4.1.1. Manual de NicheToolBox (en inglés).	97
§4.1.2. Artículo: Resumen del seminario de NicheToolbox.	145
§4.2. Paquete de R para modelar la distribución usando modelos basados en proceso.	159
5. Discusión general y conclusiones generales	162
6. Apéndices	181
§6.1. Apéndice 1	
Artículo en colaboración publicado: Predictible invasion dynamics in North American populations of the Euroasian collared dove <i>Streptopelia decaocto</i>	181
§6.2. Artículo en colaboración publicado: Complex groundwater flow systems as traveling agent models	191
§6.3. Artículo en colaboración (enviado): Assessing sustainability in North America's ecosys- tems using criticality and information theory	206
§6.4. Nota de respuesta al artículo de Dallas et al. (2017) en bioRxiv.	230

§0.1. Resumen general

Esta tesis tiene como interés principal estudiar el efecto del movimiento y las barreras bióticas a la dispersión en la distribución y supervivencia de las especies con el fin de avanzar en el entendimiento del proceso de colonización de una especie, bajo condiciones de idoneidad diferenciada del hábitat. Para ello, se utilizan modelos matemáticos que describen la abundancia de una especie en un espacio geográfico, en términos de la idoneidad del hábitat, su capacidad de movimiento y las barreras biológicas a este. Se considera un espacio geográfico puntual en el que las condiciones ambientales en cada punto o parche se reflejan en la tasa de reproducción de las especies.

La estructura general de los modelos que se estudian es la siguiente:

$$\begin{aligned} \frac{dN_{i,j}}{dt} &= R_{i,j}(\bar{N}_j, s_{i,j}) + \sum_{u \neq j}^l \Phi_{i,u} N_{i,u} + h_{i,j}(N_{i,j}, z^j) \\ \frac{dz^j}{dt} &= T_1(N_{1,j}, N_{2,j}, \dots, N_{n,j}, z^j) - dz^j + T_2(\bar{N}_1, \bar{N}_2, \dots, \bar{N}_l, z^j) \end{aligned} \quad (1)$$

donde $i = 1, \dots, n$ representa el estadio i de la especie de interés N y $j = 1, \dots, l$ es el parche j . La razón de crecimiento del estadio i de la especie N en el parche j se denota con $R_{i,j}$; esta depende de su densidad $N_{i,j}$, de la idoneidad $S_{i,j}$ para el estadio i en el parche j y de las densidades de las otras clases de edad en el parche j ; \bar{N}_j es el vector de las densidades de cada clase de edad en el parche j ; $\sum_{u \neq j}^l \Phi_{i,u} N_{i,u}$ es el flujo neto de inmigrantes al parche j y $h_{i,j}$ modela el efecto sobre el estadio i de la interacción con la especie z^j en el parche j .

Las formas específicas de las funciones $h_{i,j}$, $\Phi_{i,j}$, $R_{i,j}$, $S_{i,j}$ dependen de los mecanismos de interacción y dispersión. Las simulaciones numéricas y los métodos estadísticos son la base del análisis de este sistema.

En términos generales, la tesis está estructurada de la siguiente manera:

1. Capítulo 1: se da una breve introducción a los conceptos de nicho ecológico y de área de distribución. Cabe mencionar que esto dio lugar a la publicación de una nota científica en la Revista Mexicana de Biodiversidad.
2. Capítulo 2: Se da una revisión del procesos de dispersión y su importancia en una escala biogeográfica. Asimismo, se estudia la relación entre abundancia e idoneidad del hábitat bajo diferentes escenarios de dispersión, sobre esto se publicó un artículo de investigación en la Revista Mexicana de Biodiversidad.
3. Capítulo 3: se analiza la correlación entre abundancia y posición en el espacio de nicho y de manera específica se estudia el efecto de los estados transitorios, la estructura espacial de las idoneidades y el efecto Allee. Por último se compara el desempeño de las diferentes métricas para medir la correlación (Mahalanobis, Euclidena y las salidas del algoritmo de modelación de nichos MaxEnt). De este capítulo se deriva un artículo de investigación el cual fue enviado y está en proceso de arbitraje en la revista American Naturalist. El artículo tiene como apéndice el análisis geométrico del punto de equilibrio del modelo (3.1).
4. Capítulo 4: Se presentan algunas de las herramientas informáticas que desarrollé para analizar nichos y distribuciones geográficas. Específicamente, se presenta el manual del paquete de R «nichotoolbox» (escrito en inglés) y un tutorial del paquete «ntb», el cual es un software para resolver el modelo basado en proceso (ec. 3.1) que se introdujo en el capítulo 3.

§0.2. Abstract

This dissertation focuses on the effect of migration -including biotic barriers against movement- on the species distribution and its survival in order to provide insights on the colonization process of a species under differentiated habitat suitability. To do this, mathematical models that take into account the habitat suitability of the species and its movility are applied to analyze species abundance in a geographical space. A discretized geographic space is considered where the environmental conditions of each patch are reflected on the reproduction rate of the species. The general structure of the models that are studied is the following:

$$\begin{aligned} \frac{dN_{i,j}}{dt} &= R_{i,j}(\bar{N}_j, s_{i,j}) + \sum_{u \neq j}^l \Phi_{i,u} N_{i,u} + h_{i,j}(N_{i,j}, z^j) \\ \frac{dz^j}{dt} &= T_1(N_{1,j}, N_{2,j}, \dots, N_{n,j}, z^j) - dz^j + T_2(\bar{N}_1, \bar{N}_2, \dots, \bar{N}_l, z^j) \end{aligned} \quad (2)$$

where $i = 1, \dots, n$ denotes the stage of the species of interest and $j = 1, \dots, l$ is the j th patch. The growth rate of the stage i in the patch j is denoted by $R_{i,j}$ and it depends on its density $N_{i,j}$, the suitability $S_{i,j}$ and the densities of each stage in the patch j ; each entry of the vector \bar{N}_j is density of the stage i in the patch j ; $\sum_{u \neq j}^l \Phi_{i,u} N_{i,u}$ the net flow of immigrants into the patch j and $h_{i,j}$ models the effect of the interaction with the species z^j on the stage i , in the patch j .

The specific forms of functions $h_{i,j}$, $\Phi_{i,j}$, $R_{i,j}$, $S_{i,j}$ depends on the mechanisms of interaction and dispersion. Numerical simulations and statistical methods are the tools for analyzing this system. In general terms, this dissertation is structured as follows:

1. Chapter 1: A brief introduction to the concepts of ecological niche and distribution areas. This chapter derived in scientific note published in the Revista Mexicana de Biodiversidad.
2. Chapter 2: A review of the dispersion processes and their importance on a biogeographic scale. The relationship between abundance and habitat suitability under different dispersion scenarios is studied; this study was published as a research article in the Revista Mexicana de Biodiversidad.
3. Chapter 3: In this chapter the correlation between abundance and position in the niche space is analyzed and specifically: 1) the effect of the transitory states; 2) the spatial structure of the suitabilities; 4) the Allee effect. Finally, the performance of the different metrics to measure the relationship is compared (Mahalanobis distance, Euclidean distance and the outputs of the MaxEnt niche modeling algorithm). It is worth to note that from this chapter derives a research article which was submitted to the peer-reviewed scientific journal American Naturalist. In the appendix of this paper, the geometric analysis of the equilibrium point of the model is made (ec. 3.1).
4. Chapter 4: The computer tools that I have developed to analyze species niches and geographical distributions are described; the manual of R package 'nichetoolbox' (English version) and the tutorial of the «ntb» package (which is a software to solve the process based model of the chapter 3, ec. 3.1) are presented.

§0.3. Introducción general

La distribución de las especies está determinada por una serie de factores ecológicos, históricos y geográficos que la convierten en un proceso muy complejo [?, 3, 51]. Las propiedades poblacionales como la reproducción, mortalidad y dispersión determinan la dinámica metapoblacional de una especie, y por ende los sitios donde ésta será capaz de invadir. Desde el trabajo clásico de Grinnell (1917) [62] se ha considerado que las preferencias ambientales de una especie, llamadas su “nicho”, determinan a través de los parámetros demográficos de cada población, las posibilidades de que una especie colonice exitosamente una localidad. Factores ecológicos como las interacciones con otras especies también determinan la supervivencia de la especie colonizadora. Estos factores se relacionan dinámicamente con los ambientales y operan a diferentes escalas espaciales y temporales [152]. Bajo este marco teórico, el proceso de dispersión permitirá la llegada de la especie colonizadora a los sitios accesibles [152]; los nichos ecológicos junto con las interacciones con las especies locales, delimitarán aquellos sitios en los que la especie podrá mantener poblaciones aun sin que haya entrada de inmigrantes. Las ideas anteriores han sido abordadas por varios autores bajo diversas consideraciones [120, 70, 71, 131, 124, 145, 153, 115, 109]. Por ejemplo, Panetta y colaboradores (1991) [120] estiman el nicho ecológico para determinar los sitios potenciales de invasión de una herbácea, enfocándose solamente en el conjunto de condiciones ambientales propicias para el desarrollo de la especie invasora. Por otra parte Holt y Gaines (1992) [70] hacen un análisis matemático de las condiciones en las que una especie podrá adaptarse a ambientes fuera de su nicho ecológico, haciendo notar que la dispersión tiene un papel determinante en la evolución del nicho. Holt et al. (1997) [71] hacen un análisis de la relación entre la amplitud de la distribución geográfica y la abundancia de una especie bajo el supuesto de que el proceso de colonización ha concluido y por lo tanto la especie no presenta dispersión en su dinámica poblacional. Pulliam (2000) [131] analizando la relación entre idoneidades diferenciales del hábitat y la distribución de una especie que presenta dispersión y competencia, mostró que es posible observar poblaciones en sitios que presentan condiciones fuera del nicho a las cuales denominó poblaciones sumidero (sink populations). Schurr et al. (2007)[145] estudian la relación entre área de distribución y abundancia en el éxito de colonización de un grupo de plantas leñosas. Soberón (2010) [153] siguiendo las ideas de Vandermeer (1973) [172] esquematiza en un modelo matemático la forma en que la tasa de crecimiento poblacional per cápita en un parche j está determinada por el nicho, las interacciones con las otras especies presentes en el parche y el proceso de dispersión (ecuación 3).

$$\frac{1}{N_{i,j}} \frac{dN_{i,j}}{dt} = r_{i,j}(\vec{e}_j) \pm \varphi_{i,j}(N_j; \vec{R}_{i,j}; \vec{e}_j)e + \psi_{i,j}(\vec{N}_i; \mathbf{T}). \quad (3)$$

donde $N_{i,j}$ es la densidad poblacional de la especie i en la celda j ; $r_{i,j}$ es la tasa intrínseca de crecimiento; \vec{e}_j vector de variables ambientales (variables que definen el nicho fundamental); la función $\varphi_{i,j}(N_j; \vec{R}_{i,j}; \vec{e}_j)$ es un término de densidad-dependencia que expresa los efectos de las interacciones con otras especies y el uso del recurso; finalmente el término $\psi_{i,j}(\vec{N}_i; \mathbf{T})$ expresa los movimientos de los individuos de la especie N_i de otras celdas hacia j de acuerdo a una matriz de transición \mathbf{T} . Nótese que el efecto de las interacciones $\varphi_{i,j}(N_j; \vec{R}_{i,j}; \vec{e}_j)$ puede ser negativo o positivo dependiendo de que si las interacciones que se establecen son negativas (por ejemplo: depredación y competencia) o positivas (mutualismos).

Una forma de representar las ideas expuestas hasta el momento es por medio del llamado diagrama BAM, denominado así por sus siglas en inglés (Biotic, Abiotic and Movements; figura 1) [152]. Bajo este enfoque, es interesante estudiar cómo es que se relaciona el proceso de dispersión, con los

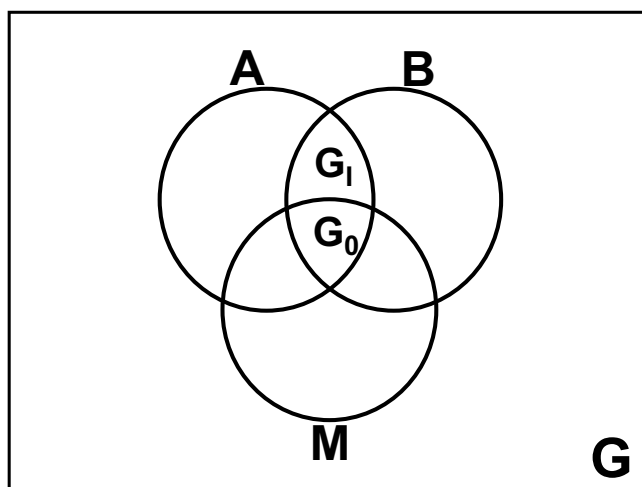


Figura 1: Diagrama BAM. La distribución geográfica de una especie depende principalmente de 3 factores: abióticos (A), bióticos (B) y los movimientos de la especie (M). G es el espacio geográfico; G_0 representa los sitios ocupados por la especie; G_1 es el área invadible y son los sitios donde la especie podría colonizar exitosamente, sin embargo no ha podido llegar.

factores bióticos y abióticos para determinar si una especie podrá llegar a un sitio determinado y colonizarlo exitosamente. Por ejemplo, habrá lugares donde la dispersión no sea una limitante, por lo que todos los sitios son accesibles, siendo A y B los factores que determinarán la colonización. Por otro lado, habrá sitios con condiciones de A o/y B adecuadas para el desarrollo de la especie, sin embargo, las limitaciones en la dispersión no permiten la llegada de ésta a esos sitios favorables [152, 8, 144]. Ejemplos de lo anterior, son la falta de dispersores de semillas o barreras geográficas que impiden el paso a aquellos sitios favorables (figura 2).

En esta tesis se estudian las relaciones dinámicas entre B, A y M mediante modelos matemáticos que consideran explícitamente el efecto de las condiciones ambientales (A) en la tasa intrínseca de crecimiento; el efecto de la dispersión (M) y de las restricciones bióticas (B) a esta en la distribución geográfica de las especies. Para ello, en el Capítulo 1 se da una breve introducción a los conceptos generales de «nicho ecológico» y de «áreas de distribución». En el capítulo 2 se muestra una revisión del proceso de «dispersión» y sus consecuencias a nivel biogeográfico, así como se muestra un artículo donde se estudia el efecto que tiene la dispersión en la relación entre la posición en el espacio de nicho y la abundancia poblacional de una especie que se dispersa en el espacio geográfico; esto se aborda utilizando un modelo a tiempo discreto con estructura de edades que integra al proceso de dispersión y el efecto de la posición en el espacio de nicho en la matriz de proyección demográfica. En el capítulo 2 se explora: 1) el efecto de la métrica en las correlaciones ente abundancia y posición respecto al centroide del nicho: específicamente se comparan a la distancia Euclidiana con la de Mahalanobis; 2) el efecto del tiempo que tiene establecida una población en una localidad dada, es decir se estudia la relación entre abundancia y posición en el espacio de nicho para diferentes tiempos de inicio del proceso de dispersión; 3) el efecto de una barrera biológica a la dispersión, para ello se introduce una limitante al crecimiento poblacional donde se supone que para que una población crezca en un parche dado, es necesario superar un umbral de abundancia poblacional

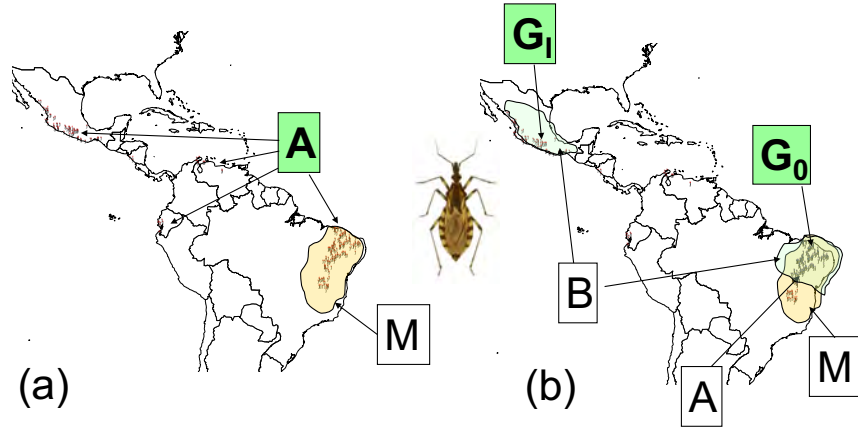


Figura 2: Área de distribución potencial y distribución geográfica de *Triatoma brasiliensis*. (a) A la derecha se muestra el área de distribución potencial de la especie (sitios señalados por las flechas); a pesar de haber sitios con condiciones favorables para el crecimiento de la especie en México, Ecuador y Venezuela, no todos son accesibles debido a una limitación en su dispersión. (b) Nótese que los factores bióticos también limitan la distribución (figura prestada por J. Soberón).

mínima N_0 por debajo del cual no hay crecimiento (efecto Allee). Para ello se utiliza un modelo a tiempo continuo con crecimiento denso-dependiente (logístico) [82].

En los capítulos 2 y 3 de la tesis se utilizan especies virtuales para analizar la relación entre abundancia y posición en el espacio de nicho. A continuación se da una breve descripción general de como se modela el nicho fundamental de las especies virtuales utilizadas en estos capítulos.

Se asumió que el nicho fundamental de la especie j se mide en términos de la idoneidad del hábitat S^j y que a su vez éstos dependen de un vector de las condiciones ambientales E de la siguiente forma:

1. Sea E un vector de n variables ambientales tales que:

$$E = \{(e_1, e_2, \dots, e_n) \mid e_1^{min} \leq e_1 \leq e_1^{max}; e_2^{min} \leq e_2 \leq e_2^{max}; \dots; e_n^{min} \leq e_n \leq e_n^{max}\} \quad (4)$$

2. La función de idoneidad del hábitat S^j para la especie j con $j = 1, \dots, m$, es una función del vector de variables ambientales y está definida en el intervalo cerrado $[0, 1]$:

$$S^j : E \rightarrow [0, 1] \quad (5)$$

3. Finalmente se asume que la tasa intrínseca de crecimiento r^j de la especie j es una función que depende de S^j :

$$[0, 1] \xrightarrow{r^j} \mathbb{R}^+ \quad (6)$$

En el capítulo 4 se da un panorama general de las herramientas informáticas que desarrollé con ayuda de algunos colaboradores durante el doctorado. La primera herramienta es una aplicación web llamada «NicheToolBox» en la cual se pueden realizar varios de los análisis que se utilizan frecuentemente para modelar el nicho y las distribuciones geográficas de las especies, como por ejemplo:

1. Obtención de datos georeferenciados de presencias de las especies. Estos datos son obtenidos de GBIF mediante su API (Application Programming Interface).
2. Limpieza de los datos de presencia duplicados.
3. Visualización del nicho n-dimensional utilizando como variables las bios de WorldClim [64].
4. Análisis de Cluster de los datos ambientales extraídos de los puntos de presencia (se usa el algoritmo de k-means).
5. Análisis de correlación bivariada de Pearson
 - a) Se puede descargar la gráfica de correlaciones de todas las bios
 - b) Tabla de correlaciones. En esta se muestran los valores numéricos de sus correlaciones
 - c) Un algoritmo que ayuda a seleccionar las variables que representan mejor los datos de acuerdo a un umbral de correlación seleccionado por el usuario (de las correlaciones más altas se eligen solo una las variables que están más correlacionadas).
6. Análisis de Componentes Principales
 - a) Se muestra el scree plot de los componentes. Este gráfico representa el porcentaje de varianza explicada por cada componente principal.
 - b) El biplot. Se puede graficar cualquier par de componentes principales.
 - c) El triplot
 - d) Los resultados generales de la transformación de los datos en componentes principales.
7. Algunas de las métricas de evaluación de los modelos de nicho
 - a) Contiene una opción para hacer ROC Parcial [125].
 - b) Métricas de evaluación de los modelos usando la matriz de confusión.
8. Workflow de todo lo que ha realizado el usuario dentro de la aplicación. Este workflow se puede descargar en los formatos: pdf, html y word (actualmente en periodo de mantenimiento).

La aplicación se encuentra alojada en uno de los servidores de CONABIO, la url del sitio es:

<http://shiny.conabio.gob.mx:3838/nichetoolb2/>

La otra herramienta informática es un paquete de R llamado «ntb» para realizar modelos de distribución geográfica utilizando modelos basados en proceso. En general las funciones que están programadas en el paquete se derivan de los análisis realizados en el capítulo 3 de esta tesis. En el apéndice 1 de la tesis se muestra un artículo de colaboración donde se analiza la relación entre centralidad en el espacio de nicho y dinámica de invasión de la paloma *Streptopelia decaocto*.

Capítulo 1

Los nichos ecológicos y las áreas de distribución

§1.1. Resumen

Los términos de modelación de nichos ecológicos y el de modelación de áreas de distribución están relacionados de manera estrecha; no se puede modelar a un nicho ecológico sin tener datos geográficos, tampoco se puede entender el área de distribución de una especie, sin considerar como influyen las condiciones ambientales de los sitios donde está presente la especie. Pese a que ambos términos están definidos en espacios con características diferentes, de manera frecuente son tratados como sinónimos; lo anterior resta claridad conceptual en su estudio. La dualidad de Hutchinson es una hipótesis sobre la correspondencia (relación) entre el espacio geográfico y el espacio de nicho, que puede ser utilizada para entender las diferencias entre ambos tipos de modelación. En el presente capítulo se muestran algunas definiciones de nicho ecológico y como éste en conjunto con otras variables como las barreras geográficas a la dispersión determinan las áreas de distribución. Así mismo se muestra una nota científica que fue escrita para aclarar las diferencias conceptuales entre la modelación de nichos y áreas de distribución.

§1.2. Introducción

El concepto de nicho ecológico, probablemente es uno de los más controversiales en la ecología debido a que se le asocia con más de una definición [10, 126]; en términos generales el nicho ha sido estudiado y definido bajo dos grandes enfoques [128]: a) el primero tiene que ver con el efecto de las condiciones ambientales en la distribución de las especies y b) el segundo está centrado en estudiar el papel que desempeñan las especies en el ambiente que las rodea. El primer enfoque está asociado a la visión de Grinnell [62] quien definió al «*nicho*» de una especie como el conjunto de requerimientos ambientales y de hábitat que delimitan su área de distribución. Por otra parte, el segundo está relacionado a la definición de Elton [40] quien conjeturó que el «*nicho*» es papel funcional de una especie en la comunidad. Si bien ambas definiciones parecieran tener poca relación, estas son complementarias; Hutchinson [76] utilizó ambos enfoques y definió al *nicho* como un espacio abstracto y multidimensional donde tanto las variables bióticas como abióticas constituyen los ejes

del espacio de nicho; siguiendo este orden de ideas, también propuso que existen subconjuntos en este espacio (Caja §1.2.1) y que existe una correspondencia 1 a 1 entre cada punto en el espacio de nicho y el espacio geográfico. Esta relación de correspondencia entre el espacio de nicho \mathbf{E} y el espacio geográfico \mathbf{G} es conocida como la «Dualidad de Hutchinson» [30].

Caja §1.2.1: Conceptos de nicho *sensu* Hutchinson

- **Nicho fundamental:** Es el conjunto de condiciones ambientales donde la tasa intrínseca de crecimiento es positiva.
- **Nicho realizado:** Son las condiciones abióticas donde la especie puede sobrevivir limitadas por la las interacciones bióticas y el área accesible.

Es importante notar que las variables bióticas y abióticas operan al mismo tiempo, solo que en escalas temporales y espaciales diferentes. Esto ha sido conceptualizado por autores como Soberón [150] quien propone que las variables Grinnellianas son aquellas variables ambientales que esencialmente no son interactivas (por ejemplo: temperatura, precipitación y topografía) y que son relevantes para comprender las propiedades ecológicas y de distribución de las especies a una escala geográfica gruesa. Desde otro ángulo, las variables bióticas o Eltonianas operan en escalas locales y son el resultado de la dinámica de las interacciones de tipo consumidor-recurso [181, 150]. En conjunto con otras variables como las capacidades de dispersión y los eventos estocásticos ambientales y geológicos, determinan el área de distribución de una especie. Las ideas mencionadas anteriormente, están representadas de manera esquemática en el diagrama BAM [149], denominado así por sus siglas en ingles (Biotic, Abiotic and Movements).

Con lo abordado hasta el momento, es preciso remarcar que el área de distribución de una especie, entonces es una mezcla de los tres factores del BAM, constreñida y moldeada por aquellas combinaciones de B, A y M donde la razón de crecimiento de cada población de la especie es positiva.

§1.3. Artículo: Diferencias conceptuales entre modelación de nichos y áreas de distribución.



Nota de opinión

Diferencias conceptuales entre modelación de nichos y modelación de áreas de distribución

*Conceptual differences between ecological niche modeling and species distribution modeling*Jorge Soberón^{a,*}, Luis Osorio-Olvera^b y Townsend Peterson^a^a Departamento de Ecología y Biología Evolutiva e Instituto de Biodiversidad, Universidad de Kansas, 1345 Jayhawk Blvd., Lawrence, KS 66045, Estados Unidos^b Facultad de Ciencias, Universidad Nacional Autónoma de México, Circuito Exterior s/n, Cd. Universitaria, 04510 Ciudad de México, México

Introducción

La disciplina, que puede considerarse parte de la macroecología o la biogeografía, de modelar las áreas de distribución de las especies utilizando los así llamados «modelos de nicho», ha crecido explosivamente en los últimos 20 años (Guisan et al., 2013). Una característica curiosa de esta nueva disciplina, nueva relativamente, ya que hay antecedentes desde la década de 1970 (Soto, Giddings y Gómez, 1996), es el predominio de resultados prácticos sobre la claridad conceptual (McInerny y Etienne, 2012). La abundancia de datos de presencias (Gaiji et al., 2013), la disponibilidad de datos de coberturas climáticas mundiales (e.g., Hijmans, Cameron, Parra, Jones y Jarvis, 2005) o nacionales (e.g., Cuervo-Robayo et al., 2014), y la variedad de programas de cómputo que correlacionan presencias con coberturas ambientales para estimar mapas (Franklin, 2009) han permitido una explosión en el número de aplicaciones, algunas de las cuales pueden adolecer de problemas conceptuales.

Un ejemplo de lo anterior es la extendida costumbre de considerar «modelación de nichos (ENM)» y «modelación de áreas de distribución (SDM)» como sinónimos (Elith y Graham, 2009; Elith y Leathwick, 2009). Estos 2 términos no son equivalentes y percibirlos como tales lleva a confusión y problemas de interpretación de resultados. La aplicación de métodos correlativos a puntos de presencias de una especie para identificar variables ambientales asociadas permite sin duda crear mapas que se pueden interpretar de diversas maneras, incluyendo distribuciones, potenciales o actuales, de una especie. Pero independientemente del aspecto estrictamente semántico, existe una clara diferencia entre modelar nichos o modelar distribuciones (Peterson y

Soberón, 2012; Warren, 2012), y el propósito de esta opinión es discutir y ejemplificar esta diferencia. Insistimos en que el punto no es meramente semántico, sino conceptual y está en la raíz de las disciplinas de ENM y SDM.

La «dualidad de Hutchinson»

Una forma clara y gráfica de discutir la diferencia entre ENM y SDM es utilizar la llamada «dualidad de Hutchinson» (Colwell y Rangel, 2009), que es una hipótesis propuesta por Hutchinson (1957) sobre la correspondencia entre el espacio geográfico y el espacio de nicho, abstracto y multidimensional. En la figura 1 se ilustra esta dualidad.

Esta correspondencia se hace evidente y se puede operacionalizar (Hijmans, Phillips, Leathwick, Elith y Hijmans, 2015) cuando se impone una retícula de resolución dada, con una proyección específica, sobre la región de interés y se procede a obtener los valores de las variables usadas para el espacio de nicho mediante coberturas ambientales en formato «raster» (Hijmans et al., 2015; Peterson et al., 2011). Si bien a cada punto de la retícula geográfica (que llamaremos G) corresponde un único punto del espacio ambiental (que llamaremos E), lo contrario no es verdad necesariamente porque puede haber 2 regiones del planeta con valores de las variables ambientales iguales o similares. Esta posibilidad en la práctica desaparece cuando se usan suficientes variables, con la precisión adecuada (Aiello-Lammens, Boria, Radosavljevic, Vilela y Anderson, 2015). En este caso se establece una relación 1:1 entre los puntos del espacio G y los del espacio E. La modelación correlativa de nichos permite entonces crear un objeto en E que es posteriormente «proyectado» a G, como se describe abajo. Es importante reconocer, no obstante, que estos espacios tienen topologías muy diferentes, que puntos muy lejanos en el espacio geográfico pueden ser muy similares en el espacio ambiental y que puntos muy

* Autor para correspondencia.

Correo electrónico: jsoberon@ku.edu (J. Soberón).

La revisión por pares es responsabilidad de la Universidad Nacional Autónoma de México.

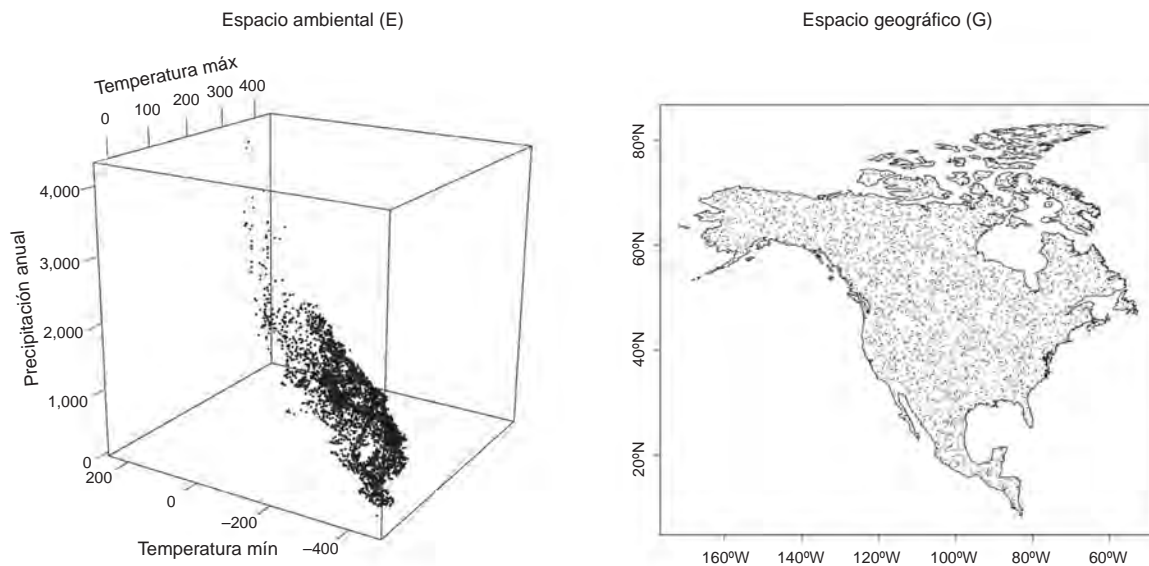


Figura 1. La dualidad de Hutchinson en 3 variables climáticas. Sobre el espacio geográfico se colocaron 4,000 puntos con coordenadas al azar (mapa, denotado por G) y se extrajeron los valores de 3 variables bioclimáticas correspondientes a cada punto (dispersograma en 3 dimensiones, denotado por E). En este caso existe una correspondencia 1 a 1 entre puntos de E y G.

cercanos en el espacio geográfico pueden ser muy diferentes en el ambiental.

La primera y obvia razón, entonces, por la cual SDM y ENM son diferentes, es que las áreas de distribución son subconjuntos del espacio G, mientras que los nichos son objetos en E. A despecho de que exista una relación 1:1 entre G y E, se trata de objetos esencialmente diferentes, que «habitan» espacios muy distintos, con propiedades topológicas diferentes. Sin embargo, existe una razón aún más fundamental para establecer la diferencia entre ENM y SDM.

El diagrama BAM

Entre los factores que determinan causalmente el área de distribución de una especie, se encuentran sus tolerancias ambientales (su nicho fundamental), la presencia o ausencia de otras especies (polinizadoras, dispersoras, competidoras, depredadoras, etc.) y las posibilidades de dispersión en períodos de tiempo y desde áreas originales relevantes (Cain, 1944; Gaston, 2003; Good, 1931; Udvardy, 1969). Dicho en otras palabras, un área de distribución, en general, depende de muchos factores, más que el simple nicho fundamental de la especie. Una manera muy simplificada de expresar lo anterior es el llamado «diagrama BAM» (Soberón y Peterson, 2005), ilustrado en la figura 2.

El diagrama BAM muestra que una especie tiene altas probabilidades de ser observada en aquellas regiones del planeta donde: *i*) existan las condiciones abióticas necesarias, simbolizadas por «A»; *ii*) existan las condiciones bióticas favorables («B»), en otras palabras, el medio biótico de la especie (McGill, Enquist, Weiher y Westoby, 2006) y *iii*) hayan estado accesibles a la dispersión de la especie desde la región de su origen histórico, o bien desde alguna región definida en un momento temporal relevante («M»), esto es, la región del planeta accesible a la dispersión (Svenning y Skov, 2004), lo cual es un asunto

esencialmente histórico (Hortal, Lobo y Jiménez-Valverde, 2012). Por ejemplo, qué partes de Norteamérica han estado accesibles desde el último máximo glacial, a partir del Eje Neovolcánico. La zona denotada por G_O , o área ocupada es la intersección de favorabilidad biótica y abiótica, y de accesibilidad histórica. La especie puede ser observada/recolectada en esa región.

Por definición, las variables ambientales en G_O constituyen el nicho realizado (N_R). La zona G_I es una región con condiciones favorables pero que se ha mantenido fuera de las capacidades de dispersión de la especie. La región A es favorable a la especie si se hace abstracción de las interacciones y por lo tanto su

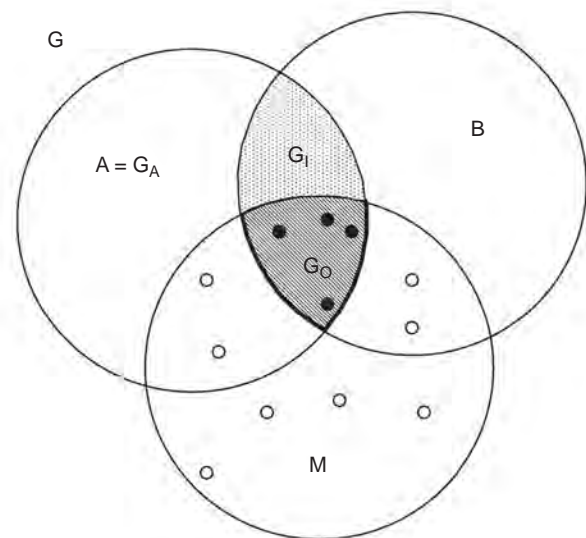


Figura 2. El diagrama BAM es una representación idealizada del espacio geográfico G, donde se ilustran diferentes regiones. Los puntos rellenos representan observaciones de poblaciones «fuente», denotadas por G_+ y que por ser observaciones dependen de la calidad del muestreo. Los círculos abiertos son poblaciones «sumidero». Cualquier algoritmo correlativo trabaja sobre estos puntos, abiertos y llenos.

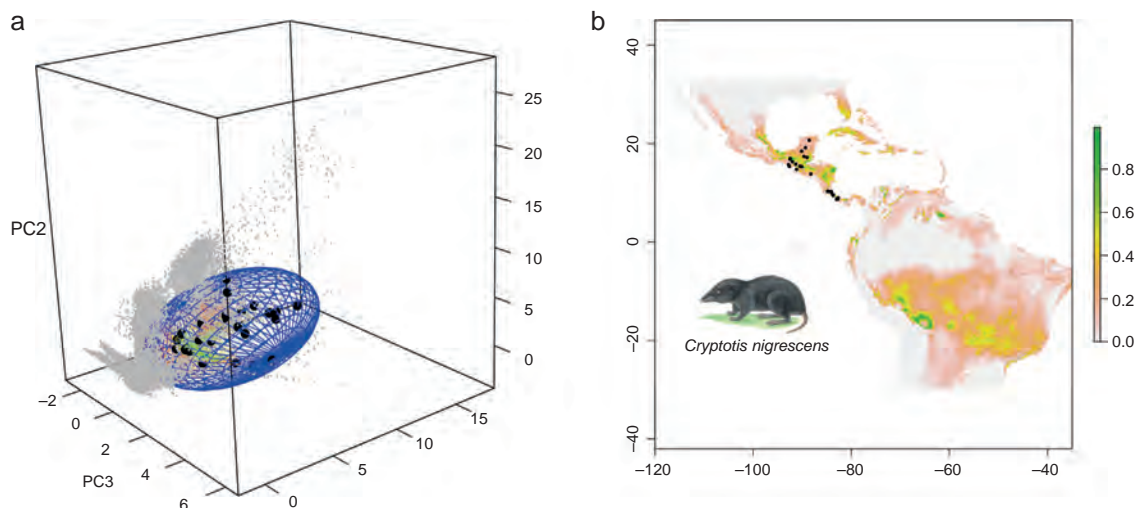


Figura 3. Modelo de nicho (elipsoide, fig. 3a) del complejo *C. nigrescens* y su proyección en la geografía, o modelo de distribución (mapa, fig. 3b). Los puntos grises corresponden a los sitios con ambientes fuera del nicho de la especie. Los círculos en negro, tanto en el espacio de nicho como en el mapa, son las observaciones procedentes de GBIF. Los puntos en el espacio de nicho son más idóneos mientras más cerca estén del centroide de la elipse. En la versión a color de la figura, un color verde indica cercanía al centroide (alta idoneidad ambiental) y un color rojo, lejanía del centroide (baja idoneidad ambiental).

medio ambiente corresponde a la idea de «nicho fundamental» de Hutchinson (1957), aunque más específicamente, a la parte del nicho fundamental que existe en una región dada, en un período dado (Jackson y Overpeck, 2000) y que simbolizamos con N^*_F , o el nicho existente. Las relaciones entre nichos y áreas se describen en la tabla 1 y se detallan en Peterson et al. (2011).

Resulta entonces que cuando se utilizan algoritmos correlativos para modelar distribuciones de especies, el algoritmo combina las presencias y las capas ambientales para generar un objeto (tabla, fórmula, conjunto) en el espacio E, que es posteriormente proyectado a G como consecuencia de la dualidad de Hutchinson (Guisan y Zimmermann, 2000; Peterson et al., 2011). Sin embargo, en los modelos correlativos el investigador muy raramente incluye explícitamente las capas B y M; algunas excepciones son Araújo y Luoto (2007) y Heikkinen, Luoto, Virkkala, Pearson y Korber (2007). Así, en una modelación correlativa, ni las interacciones ni los factores históricos y de dispersión son tomados en cuenta explícita y directamente (Barve et al., 2011). En este caso se modelan los ambientes A, y la proyección a G es una distribución potencial (Peterson, Papeş y Soberón, 2015). Vemos entonces que SDM se refiere a modelar objetos en G. Estos pueden ser considerados áreas potenciales (A o G_I), si se ignoran los efectos de B y M, o áreas ocupadas (G_O), si se cuenta con información sobre B y M. A su vez, ENM se refiere a modelar condiciones ambientales, que se aproximan más o menos a N^*_F o a N_R dependiendo del algoritmo, de la

configuración del BAM y de la calidad de las observaciones (Qiao, Soberón y Peterson, 2015; Saupe et al., 2012). Obtener aproximaciones al N_F es otro problema y no lo discutiremos en esta breve nota.

Para ilustrar lo anterior, usamos como ejemplo un complejo de 3 especies de musaraña, *Cryptotis nigrescens sensu lato* (el complejo incluye *C. mayensis*, *C. merriami* y *C. nigrescens*; fig. 3), con 31 observaciones obtenidas del Global Biodiversity Information Facility. Usando un algoritmo de elipsoides de volumen mínimo (Osorio-Olvera, Barve, Barve y Soberón, 2016) estimamos un ENM para *C. nigrescens*. Este ENM está en el

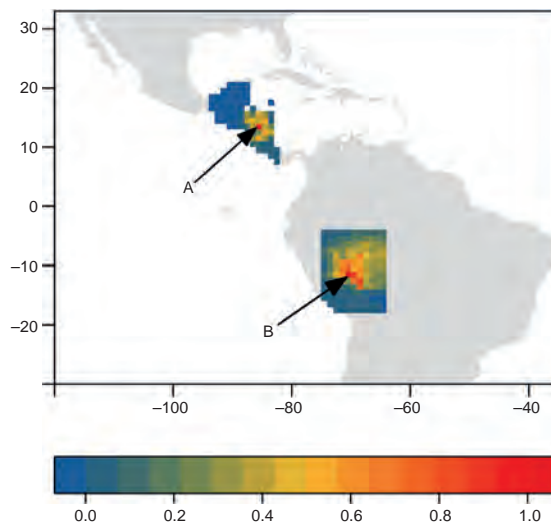


Figura 4. Dos predicciones del área de distribución a partir de 2 puntos de inicio diferentes. A, es un inicio cercano al área ocupada de la especie (puntos negros en la figura 3b); B, es un inicio muy lejano al área original, simulando la introducción de la especie a Brasil. La escala de colores representa la densidad relativa (respecto al máximo, en rojo) de individuos. En la versión en blanco y negro, las densidades máximas se encuentran hacia el centro de las 2 regiones del mapa.

Tabla 1
Relaciones entre nichos y áreas.

Espacio G	Espacio E
A (área potencial)	N^*_F (nicho existente)
G_O (área ocupada)	N_R (nicho realizado)
G_I (área invadible)	Nicho realizado acotado por movimientos
Observaciones	Modelo de nicho. Probablemente intermedio entre N^*_F y N_R

espacio E, representado por el elipsoide azul de la figura 3a. La proyección del modelo en el espacio geográfico (fig. 3b) muestra 2 áreas favorables a la especie, una en Centroamérica, cercana a los puntos de observaciones, y otra en Sudamérica, donde la especie no existe. Las hipótesis sobre M se pueden explorar usando modelos basados en procesos (Barve et al., 2011; Osorio-Olvera, Falconi y Soberón, 2016; Schurr et al., 2012). Una hipótesis que restringiera la M a la región donde el complejo de *C. nigrescens* ha existido históricamente produciría un SDM en Centroamérica. Otra hipótesis basada en una hipotética invasión de la especie a Brasil crearía un SDM completamente distinto, en el occidente de Brasil (fig. 4).

Conclusión

En el trabajo científico es esencial utilizar rigurosamente la terminología. Considerar que los modelos de nicho (ENM) y los modelos de áreas (SDM) son lo mismo, o que la terminología es «neutral» (Elith y Leathwick, 2009), conduce únicamente a dificultar la claridad conceptual y la comunicación entre científicos. Es obvio que existe una íntima e interesante relación entre nichos y áreas de distribución, que es lo que permite modelar las segundas mediante los primeros. Sin embargo, no se gana nada considerando los 2 procesos de modelación como sinónimos. Sugerimos que se evite hacerlo y se usen los términos SDM y ENM según resulten apropiados al objeto de cada estudio.

Agradecimientos

La inspiración para realizar este trabajo se debe en buena medida a nuestros colegas y estudiantes de los múltiples cursos sobre modelación de nicho en los que hemos participado; en particular, pero sin ser exhaustivos, agradecemos a Enrique Martínez-Meyer, Andrés Lira-Noriega, Octavio Rojas, Fabricio Villalobos, Pilar Rodríguez, Carlos Yáñez-Arenas y Manuel Falconi. Agradecemos el apoyo moral de Blitzi Soberón y el apoyo económico parcial de PAPIIT-UNAM IN112715 (2015) y del Posgrado en Ciencias Biológicas de la UNAM para LOO y el de National Science Foundation 1458640 para JS.

Referencias

- Aiello-Lammens, M. E., Boria, R. A., Radosavljevic, A., Vilela, B. y Anderson, R. P. (2015). *spThin*: an R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, 38, 541–545.
- Araújo, M. B. y Luoto, M. (2007). The importance of biotic interactions for modelling species distributions under climate change. *Global Ecology and Biogeography*, 16, 743–753.
- Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., et al. (2011). The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, 222, 1810–1819.
- Cain, S. A. (1944). *Foundations of plant geography*. New York: Harper & Brothers.
- Colwell, R. K. y Rangel, T. F. (2009). Hutchinson's duality: the once and future niche. *Proceedings of the National Academy of Sciences USA*, 106, 19644–19650.
- Cuervo-Robayo, A. P., Téllez-Valdés, O., Gómez-Albores, M. A., Venegas-Barrera, C. S., Manjarrez, J. y Martínez-Meyer, E. (2014). An update of high-resolution monthly climate surfaces for Mexico. *International Journal of Climatology*, 34, 2427–2437.
- Elith, J. y Graham, C. (2009). Do they? How do they? Why do they differ? On finding reasons for differing performances of species distribution models. *Ecography*, 32, 66–77.
- Elith, J. y Leathwick, J. (2009). Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology, Evolution and Systematics*, 40, 677–697.
- Franklin, J. (2009). *Mapping species distributions: spatial inference and prediction*. Cambridge: Cambridge University Press.
- Gajji, S., Chavan, V., Ariño, A. H., Otegui, J., Hobern, D., Sood, R., et al. (2013). Content assessment of the primary biodiversity data published through GBIF network: status, challenges and potentials. *Biodiversity Informatics*, 8, 94–172.
- Gaston, K. (2003). *The structure and dynamics of geographic ranges*. Oxford: Oxford University Press.
- Good, R. D. (1931). A theory of plant geography. *The New Phytologist*, 30, 149–171.
- Guisan, A., Tingley, R., Baumgartner, J. B., Naujokaitis-Lewis, I., Sutcliffe, P. R., Tulloch, A. I. T., et al. (2013). Predicting species distributions for conservation decisions. *Ecology Letters*, 16, 1424–1435.
- Guisan, A. y Zimmermann, N. (2000). Predictive habitat distribution models in ecology. *Ecological Modelling*, 135, 147–186.
- Heikkinen, R. K., Luoto, M., Virkkala, R., Pearson, R. G. y Korber, J. H. (2007). Biotic interactions improve prediction of boreal bird distributions at macro-scales. *Global Ecology and Biogeography*, 16, 754–763.
- Hijmans, R. J., Cameron, S., Parra, J., Jones, P. G. y Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, 25, 1965–1978.
- Hijmans, R. J., Phillips, S., Leathwick, J., Elith, J. y Hijmans, M. R. J. (2015). Package 'dismo'. *Circles*, 9, 1.
- Hortal, J., Lobo, J. M. y Jiménez-Valverde, A. (2012). Basic questions in biogeography and the (lack of) simplicity of species distributions: putting species distribution models in the right place. *Natureza & Conservacao*, 10, 108–118.
- Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22, 415–427.
- Jackson, S. T. y Overpeck, J. T. (2000). Responses of plant populations and communities to environmental changes of the late Quaternary. *Paleobiology*, 26(Supl.), 194–220.
- McGill, B. J., Enquist, B. J., Weiher, E. y Westoby, M. (2006). Rebuilding community ecology from functional traits. *Trends in Ecology and Evolution*, 21, 179–185.
- McInerney, G. J. y Etienne, R. S. (2012). Ditch the niche—is the niche a useful concept in ecology or species distribution modelling? *Journal of Biogeography*, 39, 2096–2102.
- Osorio-Olvera, L., Barve, V., Barve, N. y Soberón, J. (2016). Nichetoolbox: from getting biodiversity data to evaluating species distribution models in a friendly GUI environment, R package version 0.2.0.0. *GitHub*.
- Osorio-Olvera, L., Falconi, M. y Soberón, J. (2016). Sobre la relación entre idoneidad del hábitat y la abundancia poblacional bajo diferentes escenarios de dispersión. *Revista Mexicana de Biodiversidad*, 87, 1080–1088.
- Peterson, A. T., Papeş, M. y Soberón, J. (2015). Mechanistic and correlative models of ecological niches. *European Journal of Ecology*, 1, 28–38.
- Peterson, A. T. y Soberón, J. (2012). Species distribution modeling and ecological niche modeling: getting the concepts right. *Natureza & Conservacao*, 10, 1–6.
- Peterson, A. T., Soberón, J., Pearson, R. G., Anderson, R., Martínez-Meyer, E., Nakamura, M., et al. (2011). *Ecological niches and geographic distributions: monographs in population biology*. Princeton: Princeton University Press.
- Qiao, H., Soberón, J. y Peterson, A. T. (2015). No silver bullets in correlative ecological niche modelling: insights from testing among many potential algorithms for niche estimation. *Methods in Ecology and Evolution*, 6, 1126–1136.
- Saupe, E. E., Barve, V., Myers, C. E., Soberón, J., Barve, N., Hensz, C. M., et al. (2012). Variation in niche and distribution model performance: the need for

- a priori assessment of key causal factors. *Ecological Modelling*, 237-238, 11–22.
- Schurr, F. M., Pagel, J., Cabral, J. S., Groeneveld, J., Bykova, O., O'Hara, R. B., et al. (2012). How to understand species' niches and range dynamics: a demographic research agenda for biogeography. *Journal of Biogeography*, 39, 2146–2162.
- Soberón, J. y Peterson, A. T. (2005). Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodiversity Informatics*, 2, 1–10.
- Soto, M., Giddings, L. y Gómez, M. (1996). Algunos usos de bioclimas: un sistema especializado de información geográfica. *Boletín de Investigaciones Geográficas Núm Especial*, 4, 63–83.
- Svenning, J. C. y Skov, F. (2004). Limited filling of the potential range in European tree species. *Ecology Letters*, 7, 565–573.
- Udvardy, M. (1969). *Dynamic zoogeography, with special reference to land animals*. New York: van Nostrand Reinhold.
- Warren, D. L. (2012). In defense of 'niche modeling'. *Trends in Ecology & Evolution*, 27, 497–500.

Capítulo 2

El proceso de dispersión: de los organismos a las distribuciones geográficas

Resumen

El proceso de dispersión es uno de los mecanismos más importantes por medio del cual los organismos interactúan con su ambiente biótico y abiótico. Por medio de este proceso los organismos consiguen alimento, pareja y logran escapar de sus depredadores. Más aún, la dispersión junto con los procesos biogeográficos como la vicarianza han sido señalados como los mecanismos principales que dan origen a los patrones de distribución de las especies; la forma en que se relacionen la dispersión y la vicarianza determinará los niveles de flujo genético entre poblaciones y el grado en que estas compartirán trayectorias evolutivas o divergirán gracias a fuerza evolutivas como la deriva genética y la selección natural. En este documento doy una breve revisión de ¿Por qué ocurre la dispersión (en un contexto ecológico)? ¿Cuáles son sus consecuencias? ¿Cómo es que esta se da en los diferentes grupos de organismos en los que se ha estudiado? Finalmente, abordo el papel de la dispersión, la vicarianza y geodispersión en el contexto de la distribución de las especies, además de que doy algunas reflexiones con las que se busca entender como se generan patrones filogenéticos a partir de las relaciones entre el ambiente, las interacciones bióticas y los procesos de aislamiento genético.

§2.1. Introducción

En términos generales, la *dispersión* es el movimiento de los organismos hacia sitios que se encuentran fuera de su lugar de nacimiento [168, 79]. Aunque la definición anterior opera a nivel individuo, este proceso usualmente se estudia a nivel poblacional y se puede dividir en dos formas: emigración e inmigración; a) la *emigración*, ocurre cuando los individuos de una población se mueven hacia otros sitios fuera de los límites territoriales de ésta; b) la *inmigración*, es la entrada de individuos provenientes de una región externa a los límites territoriales la población.

Debido a las particularidades en que se presenta el proceso de dispersión, se han propuesto diversos conceptos relacionados a este (Caja §2.1.1), sin embargo, en todos ellos, la dispersión puede verse como un proceso demográfico con la capacidad de cambiar la estructura, distribución y abundancia de una población. Como consecuencia de lo anterior, la dispersión es uno de los mecanismos principales capaces de provocar cambio en la estructura genética de las poblaciones (*i.e.* cambio en la frecuencias de los alelos [89, 16]), lo cual a escalas de tiempo evolutivo puede dar origen a nuevas especies. Asimismo, las limitaciones en este proceso (*i.e.* barreras bióticas a la dispersión como lo es el efecto Allee) junto con la vicarianza, determinan los rangos geográficos y los patrones de distribución de las especies [152, 171, 151, 153]. Siguiendo este orden de ideas, no extraña el por qué el proceso de dispersión sea señalado como uno de los más importantes a nivel ecológico, evolutivo y biogeográfico [60, 79, 121, 105].

En este documento doy una breve revisión de ¿Por qué se da la dispersión (en un contexto ecológico)? ¿Cuáles son sus consecuencias? ¿Cómo es que esta se da en los diferentes grupos de organismos en los que se ha estudiado? Finalmente, abordo el papel de la dispersión, la vicarianza y geodispersión en el contexto de la distribución de las especies, además de que se dan algunas reflexiones con las que se busca entender como se generan patrones filogenéticos a partir de las relaciones entre el ambiente, las interacciones y los procesos de aislamiento genético.

Caja §2.1.1: Conceptos relacionadas a dispersión

- **Dispersión efectiva:** ocurre cuando los inmigrantes contribuyen a la reproducción local.
- **Dispersión natal:** es el movimiento de los individuos hacia otra localidad fuera del lugar de nacimiento y su subsecuente reproducción en ese sitio.
- **Dispersión reproductiva:** es el movimiento de un área a otra después de la primera temporada reproductiva.
- **Vagilidad:** Es la capacidad intrínseca que tienen los taxones de dispersarse [85].

§2.2. Causas y consecuencias de la dispersión (un contexto ecológico)

El movimiento de los organismos es uno de los mecanismos más frecuentes por el cual estos interactúan con el ambiente biótico y abiótico [81, 160]. Quizá esta sea la razón por la que se presenta en todos los organismos desde las bacterias hasta los animales, incluyendo aquellos que tienen hábitos sésiles [123]. Para los organismos que no son sésiles, el movimiento puede determinar un gran número de eventos trascendentes en su supervivencia, como la probabilidad de encontrar depredadores, competidores, pareja y recursos. De esta manera, el proceso de dispersión se tiene que ver como un evento multicausal que no depende solamente de las condiciones externas [26, 103], como aquellas relacionadas con las características ambientales (nicho climático, estructura del paisaje) sino que también depende de las características intrínsecas de los individuos, es decir, de como éstos perciben su ambiente [160].

Respecto a la percepción del ambiente se pueden encontrar diferencias sutiles a diferentes escalas, por ejemplo, para ciertos grupos de mamíferos como los ungulados, se ha documentado que la

percepción del riesgo de depredación cambia en función del tamaño de su manada; en grupos numerosos los individuos se sienten más seguros y por lo tanto, tienden a moverse menos dentro de los sitios donde obtienen su alimento [49, 69, 47, 63]. Por otro lado, la percepción sobre la cantidad de recursos también se modifica como función de la densidad poblacional, ya que al haber más individuos se incrementa el efecto de denso-dependencia [49, 69]; si ésta es demasiado fuerte, provocará que algunos individuos migren hacia otros sitios en búsqueda de alimento y/o pareja: por ejemplo, Paradis *et al.* [121] encontraron correlaciones negativas entre la migración y la abundancia poblacional de algunas aves británicas; de manera similar Cittadino y colaboradores [26] hallaron tasas de migración mayores para *Akodon azarae* en los ciclos poblacionales donde se presentaban abundancias altas.

Una ventaja de la dispersión, es que generalmente reduce la probabilidad de interactuar con los parientes, anulando así el riesgo de endogamia y competencia con la progenie «*kin selection*» [32]. Por otro lado, es pertinente señalar que también hay costos relacionados con esta; los individuos que se dispersan quedan expuestos a nuevos depredadores y son más susceptibles a experimentar alguna descompensación fisiológica asociada al gasto energético del desplazamiento, además de que muchos de los territorios explorados probablemente tengan características climáticas ajenas al nicho ecológico de los migrantes [56]. Algunos autores han mencionado que existe un *trade-off* entre los factores que empujan a un individuo a desplazarse (*i.e.* la competencia) y los beneficios que esto conlleva [186]. Cote y Clobert [32] comparando la cantidad de migrantes de *Lacerta vivipara* como función de la competencia entre individuos emparentados «*kin competition*» y no emparentados, mostraron que el número de migrantes es mayor cuando hay competencia entre parientes. Siguiendo estas líneas de investigación algunos autores han sugerido hipótesis que mencionan a la competencia entre parientes como uno de los mecanismos de evolución de la dispersión (Caja §2.3.1)[5]. Una posible razón por la que los individuos emparentados sean los que migran más, es que estos al ser más parecidos entre sí, tienen requerimientos fisiológicos y de alimento más similares, lo que genera una mayor competencia entre ellos y su subsecuente exclusión competitiva [43, 53].

Con todo lo anterior, se espera que los individuos respondan de manera variada a condiciones ambientales similares, y como consecuencia de ello, éstos presenten patrones de movimiento diferentes tanto a nivel individual como a nivel de especie [143, 160].

§2.3. La dispersión en los diferentes grupos de organismos

Si bien la decisión de dispersarse de un individuo depende de los costos y beneficios que obtiene, hay factores como la edad y el sexo que son determinantes en esta decisión. En cuanto al sexo, se ha encontrado que los diversos grupos de organismos presentan una asimetría en la migración, por ejemplo: en las aves se ha observado de manera frecuente que la dispersión se presenta más en las hembras que en los machos [60, 79, 27]. En los mamíferos monogámicos no hay un sesgo sexual [156]; sin embargo, en los poligámicos, la dispersión es más frecuente en los machos, tal es el caso de algunos felinos como *Puma concolor* [159], *Panthera tigris* [59] y *Panthera leo* [37] (*ver* Tabla 2.1). Greenwood [61, 28] en una revisión de trabajos sobre dispersión en los mamíferos y aves, encontró que de un total de 65 especies de mamíferos: 45 especies (correspondientes a 23 familias) presentan dispersión sesgada hacia los machos, 5 especies (4 familias) a las hembras y 15 (7 familias) no hay sesgo. Por otro lado de un total de 30 especies de aves: 3 (1 familia) presentan sesgo hacia los machos, en 21 hacia las hembras (11 familias) y en 6 (5 familias) no hay sesgo. Estos patrones se han asociado a la conducta; por ejemplo: los machos de los mamíferos poligámicos suelen ser más agresivos y territoriales que las hembras, por lo que obligan a otros machos a migrar más lejos en

Cuadro 2.1: Algunas distancias de dispersión por unidad de tiempo en diversas especies

Especie	Distancia	Sexo	Referencia
<i>Puma Concolor</i>	41.6-187.6 km / año	M	[159]
	4.42-73.5 km / año	H	
<i>Panthera tigris</i>	165 km / año	-	[97, 59]
	104-132 km / año	H	
<i>Panthera leo</i>	117-343 km / año	M	[37]
	50-128 km / año	H	
<i>Anolis sagrei</i>	3.18 m ± 0.37 / 4 horas	M	[18].
	1.25 m ± 0.40 / 4 horas	H	
<i>Plethodon cinereus</i>	4.05 ± 1.19 m/día	M	[92]
	1.74 ± 0.28 m/día	H	
<i>Salvelinus fontinalis</i>	195.4 ± 36.5 m/ 7 días	M	[75]
	81.5 ± 23.4 m/ 7 días	H	
<i>Spodoptera exempta</i>	Cientos de km/noche	-	[52]
<i>Thaumetopoea pityocampa</i>	300 m - pocos km / día	H	[84]
<i>Schistocerca gregaria</i>	5-200 km / día ^o	-	[33]
<i>Centaurea corymbosa</i>	0.2-2 m en su ciclo de vida	-	[29]

búsqueda de territorio y hembras (ver «hipótesis de la subordinación social» [25], Caja §2.3.1). Por otro lado, en aves es frecuente que las hembras se muevan distancias grandes en búsqueda de la pareja que les ofrezca el mejor sitio de anidación [34].

En el caso de los reptiles, anfibios y peces no se ha determinado un sesgo en la dispersión por el sexo; sin embargo en algunas especies se ha reportado que los machos se dispersan distancias mayores que las hembras, por ejemplo: en la especie *Crocodylus johnstoni*, los machos viajan distancias 3 veces mayores que las hembras; asimismo, los machos de la especie *Anolis sagrei* viajan con mayor frecuencia y distancias mayores [18]. Este patrón se repite en la salamandra *Plethodon cinereus* donde el macho viaja distancias más lejanas que las hembras [92]. En el salmónido *Salvelinus fontinalis*, el macho viaja 2.5 veces más lejos que las hembras [75]

En los insectos como la mariposa blanca *Pieris rapae crucivora*, la dispersión ocurre solo en las hembras después del apareamiento [68]. Las hembras de *Cotesia glomerata* abandonan su sitio de nacimiento cuando las oportunidades de apareamiento conllevan a la endogamia [139]. En un estudio realizado con marcadores mitocondriales, Kerdellué y colaboradores [84] determinaron que los machos de *Thaumetopoea pityocampa* viajan más que las hembras.

Otro factor que influye en la intensidad y la distancia máxima de dispersión es el tamaño corporal; en un metanálisis donde se examinaron un total de 795 especies (320 dispersores pasivos y 475 activos) incluyendo bacterias, hongos, plantas y animales, Jenkins et al. [78] determinaron que existe una relación positiva entre el tamaño corporal y la distancia que viajan los organismos cuyos movimientos son activos. De manera opuesta, los dispersores pasivos presentaron correlaciones negativas entre el tamaño de su cuerpo y la distancia que viajan, lo anterior se puede deber a que estos últimos son transportados por agentes externos como el viento o algunos vectores. Cabe señalar que dentro de las especies migratorias esta relación entre tamaño corporal y distancia no se preserva.

Algunos ecólogos distinguen dos tipos de la dispersión; la *dispersión natal* y la *dispersión re-*

productiva (Caja §2.1.1). En los peces, aves y mamíferos estos mecanismos se suelen presentar en diferentes etapas del ciclo de vida; la primera ocurre en las etapas tempranas y la segunda sucede cuando se alcanza la madurez sexual. En la mayoría de las especies la distancia recorrida a causa de la dispersión natal es mayor que la reproductiva (los juveniles se dispersan más lejos que los adultos) [61].

En diferentes taxa es común encontrar algunas adaptaciones especializadas para las dispersión, tal es el caso de los insectos de los órdenes Orthoptera y Hemiptera, quienes presentan adaptaciones morfológicas respecto al tamaño de sus alas; por ejemplo, es común que los migrantes tengan alas más largas y fuertes (este dimorfismo es evidente en los áfidos) [123]. Organismos sésiles como las plantas tienen mecanismos especializados para la dispersión a largas distancias (LDD) y a cortas distancias (SDD)[143]. Así, los géneros *Leontodon*, *Heterotheca*, *Senecio* de la familia Asteracea producen dos tipos de semillas: una especializada para LDD (la cual generalmente es una semilla pequeña y con un papus bien desarrollado) y otra mejor adaptada para la SDD (sin papus) [123, 23]. Las semillas transportadas por vectores (generalmente viajan distancias largas) presentan algunas adaptaciones extras como cubiertas seminales gruesas que les permiten sobrevivir al tracto digestivo de sus dispersores.

Caja §2.3.1: Hipótesis de dispersión

Se presentan algunas de las hipótesis que se han propuesto para explicar el proceso de dispersión, estas son de tres tipos: genético, ecológica y etológica.

- **Hipótesis de la depreciación de la descendencia:** Sugiere que la dispersión ocurre como mecanismo que evita la competencia entre parientes (*kin competition*) [87, 137].
- **Hipótesis de la adecuación de los residentes** (etológica): En esta los residentes (padres) maximizan su adecuación promoviendo que los descendientes migren (la dispersión está controlada por los padres) [57, 2].
- **Hipótesis de la adecuación de los migrantes** (genética): La emigración evolucionó a partir de una ganancia de adecuación a través de los migrantes (la dispersión está controlada por los genes de los migrantes) [2, 57].
- **Hipótesis de la cohesión social** (etológica): Los individuos que logran socializar son los que migran menos [12, 9].
- **Hipótesis del polimorfismo genético-etológico:** La agregación es el estímulo que impulsa a los individuos a dispersarse. La presencia de diferentes genotipos en la población, está determinada por los distintos niveles de agresividad de los individuos [24, 86].
- **Hipótesis de subordinación social** (etológica): Hay un relación causal entre el incremento de la agresión por parte de los individuos dominantes y la dispersión de los subordinados [25, 178].
- **Hipótesis de la dispersión presaturada-saturada** (ecológica): Sugiere que la dispersión (migración) ocurre en dos circunstancias: a) justo antes de que la población alcance su capacidad de carga (presaturation dispersal); b) cuando se ha superado la capacidad de carga (saturation dispersal) [91].
- **Hipótesis de la dispersión adaptativa y no adaptativa** (evolutiva): La dispersión es: a) adaptativa cuando no depende de la densidad y los individuos se dispersan se establecen y reproducen exitosamente y b) no adaptativa, cuando los organismos se dispersan forzamente (es análoga a la dispersión saturada) [155, 13].

§2.4. La dispersión en el contexto biogeográfico y filogenético:

Hasta el momento se han revisado algunas de las causas y consecuencias de la dispersión en un contexto ecológico, sin embargo, este proceso va más allá. La dispersión junto con los procesos biogeográficos como la vicarianza han sido señalados como los mecanismos principales que dan origen a los patrones de distribución de las especies. Más aún, la forma en que se relacionen estos dos procesos determinará si una población se aislará genéticamente o no. Lo anterior es muy importante debido a que los niveles de flujo genético entre poblaciones influyen en el grado en que estas compartirán trayectorias evolutivas o divergirán gracias a fuerza evolutivas como la deriva genética

y la selección natural [183, 148]. A continuación se darán algunos argumentos de como es que estos procesos interactúan para generar los patrones de distribución de las especies y los patrones filogenéticos. Para ello, primero se hará una distinción del papel de los procesos ecológicos y los procesos históricos en la distribución de las especies; posteriormente se darán elementos que nos ayuden a entender como estos pueden ser integrados para comprender los patrones filogenéticos.

§2.4.1. Los procesos, factores ecológicos en la distribución de las especies

Dentro de los factores ecológicos que determinan la distribución de las especies encontramos a: 1) a las interacciones bióticas, 2) el nicho ecológico y 3) la dispersión [11]. La forma en que se relacionan los procesos bióticos como la dinámica poblacional, la competencia intra-específica y las interacciones con otras especies (positivas, negativas o neutras, Figura 2.1) con el nicho ecológico, determina gran parte de las presiones de selección a las que una especie estará sometida y por tanto, influyen en la dirección e intensidad de la dispersión a escalas ecológicas de tiempo (*ver* [19]).

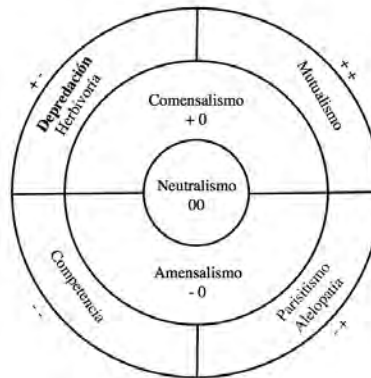


Figura 2.1: Modelo coordinado de los tipos interacciones entre dos especies. Se pueden distinguir tres grandes grupos: las positivas (+ +), las neutras (00) y negativas (– – ó + –). Reproducido de Odum & Barret [112].

En términos generales, el nicho ecológico (*sensu* Grinnell [62]) delimita los rangos de las variables ambientales en los cuales la especie puede sobrevivir. En ausencia de otros factores como las interacciones bióticas negativas o eventos geológicos catastróficos (*i.e.* temblores, erupciones), las poblaciones que se desarrollen en sitios con características ambientales inherentes a su nicho ecológico podrán sobrevivir en ellos, siempre y cuando puedan obtener recursos [17]. Queda claro que a pesar de que puede haber muchos sitios de alta calidad o con características ambientales adecuadas, no todos éstos pueden ser invadidos, para ello (en un contexto de factores ecológicos) se deben de cumplir por lo menos 3 condiciones:

1. Que los sitios sean accesibles para la especie, es decir, que los individuos puedan llegar por dispersión a los sitios favorables [8].
2. La cantidad de migrantes debe ser suficientemente grande para que se pueda dar la reproducción y/o se supere el umbral mínimo poblacional (efecto Allee *ver* [82]) que hace que la población sea viable. Fenómenos como la estocasticidad demográfica y ambiental pueden llevar fácilmente a la extinción a las poblaciones pequeñas (*ver* [98, 88, 106]).
3. Es indispensable que la tasa de mortalidad asociada a las interacciones bióticas negativas (Figura 2.1), sea menor que la tasa intrínseca de crecimiento (*ver* [113]).

Una forma de representar las ideas expuestas hasta el momento es por medio del llamado diagrama **BAM**, denominado así por sus siglas en inglés (Biotic, Abiotic and Movements; figura 2.2) [152]. Cada círculo en el diagrama **BAM** está directamente relacionado con uno de los factores mencionados. La letra **B** simboliza a las interacciones bióticas (intra o inter específicas); la **A** al nicho ecológico de la especie (*sensu* Grinnell [62]); y la **M** a los sitios accesibles para la especie; dicho de otra manera, los lugares donde la especie es capaz de llegar por medio de la dispersión. Diferentes configuraciones en el diagrama **BAM** develan la importancia de cada uno de los procesos ecológicos mencionados en la distribución de una especie (Figura 2.2). Así, habrá especies en las que los factores bióticos sean los que delimiten su distribución, algunas en las que las barreras al movimiento lo hagan y otras en las que los 3 círculos sean determinantes.

Existen diversos ejemplos en la naturaleza en los que una determinada configuración del **BAM** delimita la presencia o ausencia de una especie en una localidad dada (Figura 2.2). Para el caso donde dominan las interacciones bióticas y la dispersión (M), ha sido documentado cómo la depredación puede ser responsable de la extinción local de poblaciones y/o especies, tal es el caso de la introducción de la serpiente *Boiga irregularis* en la isla Guam, la cual produjo la extinción de por lo menos de 7 especies de aves nativas de la isla [138]; las limitaciones en la vagilidad (dispersión, Figura 2.2b) de estas aves quizá sean la razón por la que no pudieron escapar a islas cercanas con características climáticas similares como las islas Marianas del Norte (a ~ 210 km de distancia).

Un caso donde la competencia y la depredación son determinantes en la distribución de las especies y la **M** no representa una limitante, queda ejemplificado en los experimentos que realizó Paine [118, 119, 116, 117] en los que la extracción selectiva de un depredador (*Pisaster ochraceous*) provocaba el rápido crecimiento de su presa principal (*Mytilus californianus*) y la extinción local de varias especies por exclusión competitiva. Al comienzo del experimento de Paine, habían 15 especies diferentes de anémonas, esponjas, nudibranquios y macroalgas, que competían por un espacio en las rocas intermareales, sin embargo, después algunos años de la remoción de *Pisaster* solo encontraron 8 especies sobrevivientes. Con este experimento queda en evidencia que las relaciones resultantes de una configuración particular del **BAM** son dinámicas y pueden cambiar con el tiempo [152], promoviendo o disminuyendo la diversidad de especies presentes en una localidad; antes de la remoción de *P. ochraceous*, el factor biótico no representaba una limitante en la distribución de las especies del gremio de *Mytilus californianus*, sin embargo después de ello, la **B** se convirtió en la limitante principal que impedía la colonización de las rocas intermareales (Figura 2.2c); la presencia del depredador permitía la coexistencia de todas las especies y disminuía la restricción biótica de las especies excluidas competitivamente (Figura 2.2d).

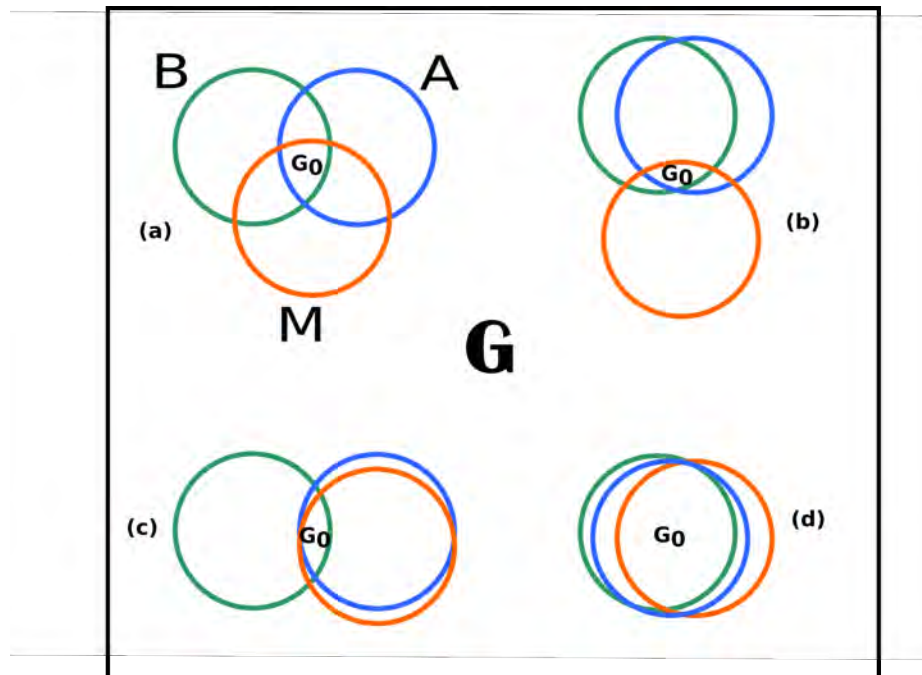


Figura 2.2: Diagrama BAM y algunas de sus configuraciones. (a) La representación clásica del BAM donde **B** simboliza a los factores bióticos (en verde); la **A** a los abióticos (en azul y es el nicho); y **M** son los sitios o áreas accesibles para la especie (en naranja). **G** es el espacio geográfico; **G₀** representa los sitios ocupados por la especie (ocurre en $B \cap A \cap M$). (b) Una configuración donde la **M** determina la distribución (el caso de las aves endémicas de Guam). (c) En esta configuración la **B** es el factor determinante de la distribución (el ejemplo de Paine después de la remoción de *P. ochraceous*). (d) Hay condiciones favorables tanto de **B**, **A** y **M** (el caso de las especies intermareales antes de la remoción de *P. ochraceous*).

§2.4.2. Los factores y procesos históricos en la distribución de las especies

Hasta el momento se ha revisado el papel de los factores ecológicos en la distribución de las especies, sin embargo, hay otro tipo de factores que determinan sus rangos geográficos; estos son los factores históricos. En estos se estudia cómo se han dado los eventos de dispersión a través de la historia y se resalta la importancia de los procesos geológicos y los cambios climáticos, en la partición de una población en subconjuntos de poblaciones aisladas (no hay flujo genético entre ellas). La idea anterior es conocida como «*vicarianza*» y los eventos de este tipo que conllevan a la partición de las poblaciones son llamados «*eventos vicariantes*» [73]. Un caso particular de evento vicariante que ha llamado mucho la atención de diversos biogeógrafos, es la aparición de los Andes, lo cual se cree ha sido el promotor principal de la diversificación de distintos grupos de plantas y animales en esa región; los gradientes de altitud derivados de la aparición de las cadenas montañosas dieron como resultado una gran variedad de condiciones climáticas que facilitaron la especiación en diferentes grupos taxonómicos [74, 4, 177]. En contra parte, los eventos geológicos también pueden promover la dispersión, tal es el caso de la geodispersión que es un tipo de dispersión facilitada por la desaparición de una barrera geográfica [142]. Lo anterior puede tener como consecuencia, la expansión del área de distribución de varios linajes y generalmente ocurre de manera masiva (incluso biotas completas) [180]. Ejemplos de eventos de geodispersión son la aparición del Istmo de Panamá (hace 3-4 millones de años) y la colisión de Asia e India. El surgimiento del Istmo de Panamá tuvo como consecuencia el intercambio masivo de mamíferos entre América del Norte y América del sur [175]. Análogamente, durante la colisión entre India y Asia, India tomó un trozo de la Península Arábiga y transportó consigo parte de su flora y fauna a Asia [31].

§2.4.3. Sobre el papel de la ecología, la dispersión y la vicarianza en los patrones filogenéticos

En un contexto evolutivo, la dispersión y la vicarianza pueden ser precursores de la especiación [104] y originar así patrones filogenéticos; toda vez que una población se dispersa (ya sea por un proceso geológico o ecológico) y queda aislada por algún evento vicariante, el flujo de genes que mantenía la estructura genética de la población original cesa; ahora las poblaciones separadas por la barrera geográfica (también conocidas como poblaciones alopátricas), estarán expuestas a presiones de selección inherentes a los sitios de aislamiento. La selección natural producto de las condiciones ambientales y/o ecológicas diferenciadas (ver Sección §2.4.1) puede causar la acumulación de cambios genéticos adaptativos entre poblaciones alopátricas y de este modo, la evolución. La especiación que ocurre bajo estas condiciones es conocida como especiación alopátrica.

Un ejemplo donde la especiación alopátrica fue el mecanismo de diversificación, es el de las aves de la familia Vangidae en Madagascar (Figura 2.3), donde se mostró que las tasas de especiación de esta familia fueron más altas que las de sus parientes de los continentes Africano y Asiático. Las fuentes nuevas de alimento y las nuevas presiones de selección en Madagascar, provocaron la aparición de nuevas especies de vándidos cuyas modos de forrajeo y morfología del pico difieren bastante de sus parientes continentales. Este patrón de radiación endémica y disparidad ecomorfológica es similar al de otros grupos de aves y animales endémicos de la isla ([135], ver [136, 174]). Reddy y colaboradores [135] atribuyen las altas tasas de diversificación en Madagascar a la falta de competidores (en este caso la **B** no representó una restricción a la colonización, ver Figura 2.2b) y al efecto fundador. La especiación que originó a los vándidos de Madagascar se considera un fenómeno de radiación adaptativa por la rápida especiación y adaptación a las nuevas condiciones ambientales (nichos ecológicos) [54]. A diferencia de los que ocurrió en Madagascar, se

han reportado tasas bajas de diversificación para estos grupos en las regiones continentales de Asia y África donde la competencia con otros linajes es fuerte (la **B** es la limitante, Figura 2.2c).

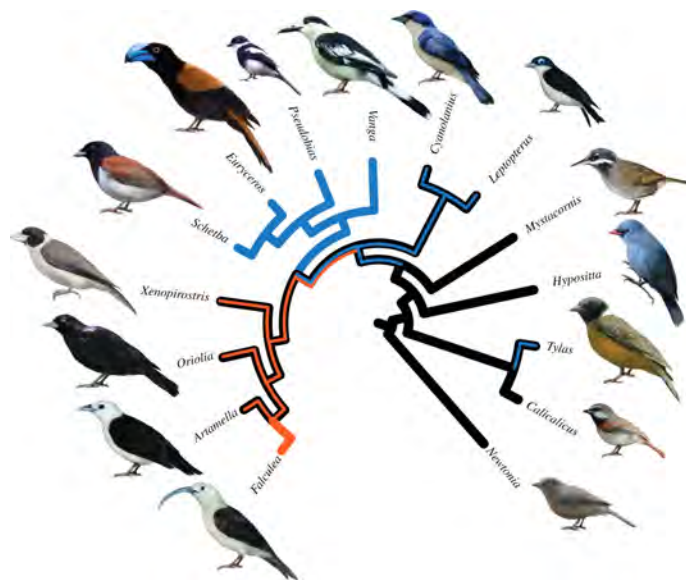


Figura 2.3: Especies de la familia Vangidae presentes en Madagascar. Se puede observar la gran variedad de formas del pico entre los integrantes de esta familia. Tomado de Reddy *et al.* [135].

La radiación adaptativa de las lagartijas neotrópicas del género *Anolis* es un ejemplo de como la relación dinámica de los fenómenos de dispersión (oceánica) y adaptación a nuevas condiciones ecológicas, pueden generar patrones filogenéticos a grandes escalas; la gran diversidad de especies que presenta este género en las islas del Caribe (~ 150 especies) se ha correlacionado con los eventos tempranos de colonización y con algunas adaptaciones que le confirieron la capacidad de explorar nuevos nichos ecológicos. Una de estas adaptaciones fue la aparición de cojinetes dactilares; algunos estudios han encontrado que lagartos con un cojinete más grande tienen mayor capacidad de adherencia en superficies lisas, lo cual les confiere un aumento en la capacidad de movimiento (incrementa la **M**) [77]. Lo anterior está relacionado directamente con los hábitos arborícolas de las especies de *Anolis* y se considera que el significado funcional y ecológico de este atributo fenotípico, es una innovación clave «*key innovation*» en la diversificación de las especies que conforman este género [90, 94]; un dato interesante es que esta adaptación morfológica también produjo cierto grado de evolución convergente entre las especies de las Antillas Mayores, de forma que es posible observar adaptaciones similares en cuanto a la ecología, etología y vagilidad de estas especies de *Anolis* [93, 95].

Otro caso donde se muestra la relación dinámica entre los fenómenos de vicarianza, dispersión y radiación endémica es el de la Mariposa del género *Calisto* en el Caribe (constituido por ~ 44 especies: 11 en Cuba, 1 en Puerto Rico, 1 en Anegada, 1 en Jamaica, 2 en las Bahamas y 28 especies en la Española). En un estudio realizado por Matos y colaboradores [102] se encontró que a pesar de que el último ancestro del género se originó en la parte continental de Sudamérica (Gaarlandia) en los inicios de Oligoceno (31 ± 5 millones de años), la mayor tasa de diversificación ocurrió hace 14 millones de años en el Caribe donde los arreglos geológicos favorecieron la aparición de nichos

ecológicos vacíos. Los eventos vicariantes en Cuba y la Española explican gran parte la distribución geográfica de este género. Por otro lado, la dispersión ecológica a largas distancias (LDD) es el proceso por el cual se dio la colonización de esta especie en Jamaica y las Bahamas.

§2.4.4. Reflexiones para la modelación de estos fenómenos

Una pregunta de interés general (planteada en el examen de candidatura) está relacionada con la posibilidad de integrar el fenómeno de vicarianza en los modelos de distribución que he presentado en la tesis. Considero que es posible incluirlos de manera indirecta por medio de los Kernels de dispersión; de modo esquemático, en los modelos propuestos en la tesis, el espacio geográfico se asume discreto conformado por una red finita de parches o hábitats completamente conectados (red completamente conexa) a través de un proceso de dispersión y con una dinámica propia en cada nodo, determinada por el nicho ecológico (*sensu* Grinnell) y procesos poblacionales. Un fenómeno vicariante se refleja en el modelo por una desconexión entre algunos nodos, lo cual puede producir incluso, el aislamiento de algunos de ellos. Basta cambiar el Kernel de dispersión o la dinámica local en los parches para modelar un fenómeno de ese tipo. A continuación planteo algunas ideas de lo que se debería considerar:

1. El primer paso sería tratar de entender la dinámica poblacional y la estructura espacial de las abundancias poblacionales como función del nicho ecológico y la capacidad de dispersión de la especie cuando no hay barreras a la dispersión, dicho de otro modo, estudiar diferentes Kernels de dispersión cuando no hay barreras que eviten el flujo de migrantes.
2. Posteriormente se debería explorar como cambia la distribución geográfica de las abundancias ahora en función de los cambios en la conectividad de la red; estos cambios en la estructura de la red representarían indirectamente el efecto de la vicarianza, ya que inhiben el flujo de genes entre poblaciones.
3. Una cuestión derivada del punto anterior, sería ver cuales son las estructuras topológicas de la red capaces de generar un cierto patrón de distribución. De este modo se podrían hacer conjeturas sobre los sitios clave donde un evento vicariante producirá de manera efectiva un aislamiento total de una o varias poblaciones (no todos los eventos vicariantes producen aislamiento); a manera de ilustración de las ideas anteriores considere los siguientes esquemas:
 - a) Una red de parches donde todo es accesible y por lo tanto, una población que comienza a dispersarse es capaz de invadir todos los nodos o parches (Figura 2.4a).
 - b) Casos donde hay una barrera a la dispersión:
 - 1) En el primero la barrera ocurre en el parche que conecta a la población inicial con un nodo que es el punto de acceso a los demás parches; en este es claro que no habrá invasión a los demás sitios; la población quedará aislada (Figura 2.4b).
 - 2) La barrera ocurre entre dos nodos que no son vitales para la conectividad de los parches, de tal forma que indirectamente hay intercambio de migrantes entre todos los parches de la red (Figura 2.4c).

De forma esquemática también se puede plantear el efecto de las interacciones con otras especies en los modelos de distribución; de hecho Soberón [153] muestra el siguiente esquema de modelo:

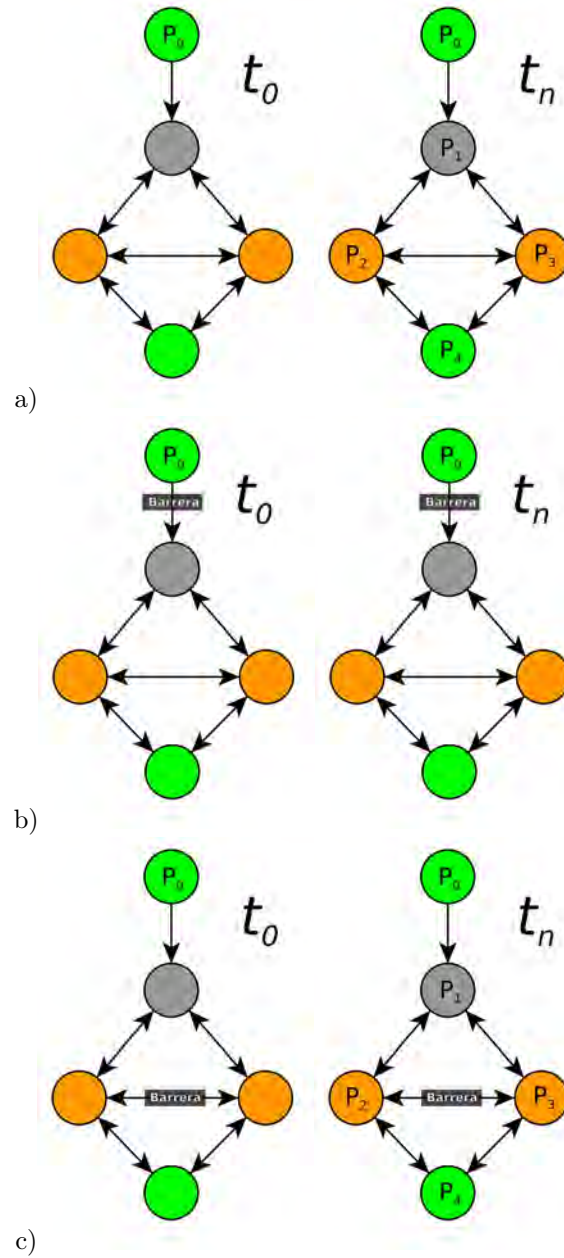


Figura 2.4: Algunas barreras a la dispersión y sus posibles patrones de distribución después de un tiempo t_n . a) No hay barreras a la dispersión; después de un tiempo t_n , la población invade todos los parches. b) Una barrera impide la conexión entre el parche de la población inicial y el nodo que permite la conexión con los demás (no hay un incremento en el área de distribución de la especie). c) El flujo entre dos nodos se detiene debido a la presencia de una barrera, sin embargo, gracias al flujo indirecto de migrantes entre los demás parches, se genera el mismo patrón que cuando no hay barreras (caso a).

$$\frac{1}{N_{i,j}} \frac{dN_{i,j}}{dt} = r_{i,j}(\vec{e}_j) \pm \varphi_{i,j}(N_j; \vec{R}_{i,j}; \vec{e}_j)e + \psi_{i,j}(\vec{N}_i; \mathbf{T}). \quad (2.1)$$

donde $N_{i,j}$ es la densidad poblacional de la especie i en el nodo j ; $r_{i,j}$ es la tasa intrínseca de crecimiento; \vec{e}_j vector de variables ambientales (variables que definen el nicho *sensu* Grinnell); la función $\varphi_{i,j}(N_j; \vec{R}_{i,j}; \vec{e}_j)$ es un término de denso-dependencia que expresa los efectos de las interacciones con otras especies y el uso del recurso; finalmente el término $\psi_{i,j}(\vec{N}_i; \mathbf{T})$ expresa los movimientos de los individuos de las especie N_i de otros nodos hacia j de acuerdo a una matriz de dispersión \mathbf{T} . Nótese que $\varphi_{i,j}(N_j; \vec{R}_{i,j}; \vec{e}_j)$, puede estar sumando o restando dependiendo de si las interacciones que se establecen son negativas o positivas (Figura 2.1). En este modelo una especie i , que se desarrolle en un nodo j con condiciones ambientales (el nicho *sensu* Grinnell) no favorables y cuya tasa de mortalidad asociada a las interacciones bióticas sea mayor que la suma de migrantes y el reclutamiento, se extinguirá.

Un modelo de evolución, además de tener en cuenta los factores mencionados anteriormente, tendría que considerar el impacto de alguna innovación o cambio en alguna característica fenotípica en la distribución o rango de una especie (*i.e.* los cojinetes de *Anolis*). Con fines ilustrativos, consideremos hipotéticamente una población en la que algunos individuos poseen una variante que facilita la dispersión; esta variante por supuesto tiene un gasto energético (*trade-off*) que provoca una disminución en la efectividad al realizar cierta actividad Q y por lo tanto, tiene una disminución de la adaptación (fitness). Para analizar esto debemos separar el tamaño de la población en dos clases: la que posee la innovación y los que no la poseen e introducir un parámetro e que indique el grado en que se posee esta innovación; en el modelo se deberá agregar una función $b(e)$ que decrece con e e indica el efecto negativo en la actividad Q . En términos más precisos, e es la estrategia y se buscaría determinar la estrategia evolutivamente estable (ESS) como función de las interacciones y el nicho. En este modelo dependiendo de como se den la relaciones $\mathbf{B} \cap \mathbf{A} \cap \mathbf{M}$ dentro de la estructura de los parches, se fijará o no esta innovación. Supóngase que esta innovación le confiere habilidad para escapar de los depredadores; si en el parche hay una gran presión de selección asociada a la depredación (el factor \mathbf{B}) es muy probable que la estrategia evolutivamente estable sea la de la innovación. Por otro lado, si tal presión de selección no es tan importante, los individuos que tenga el atributo fenotípico posiblemente dejen menos descendencia, y por lo tanto esta innovación vaya desaparecido por deriva genética y selección natural.

§2.5. Conclusiones

Con todo lo anterior, es evidente que los procesos ecológicos junto con los climáticos y geológicos pueden determinar las distribuciones geográficas, los patrones de evolución y extinción. En cuanto a la ecología hemos visto como las relaciones bióticas pueden llevar a la extinción local (el caso de *Boiga irregularis*) o promover la diversidad de especies (*i.e.* *Pisaster*), también vimos que la competencia intraespecífica entre parientes conlleva a la dispersión (el caso de las hembras de *Cotesia glomerata*) la cual puede tener como consecuencia la colonización de nuevos nichos ecológicos. En cuanto a procesos históricos se puede notar que los eventos de dispersión (*i.e.* dispersión oceánica y geodispersión) son capaces de producir migraciones (a veces masivas) de especies a nuevos territorios (esto produce un incremento en el área de distribución de las especies); la vicarianza es un evento aislador que puede promover fenómenos de evolución alopátrica (*i.e.* las mariposas del género *Calisto*). Por último, aunque no es parte de los objetivos inmediatos de este texto, no

quiero dejar pasar la oportunidad de mencionar el impacto que los eventos de cambio climático pueden tener en la distribución de las especies, ya que se ha visto que cambios en la temperatura han causado extinciones masivas; un ejemplo de esto ocurrió durante la glaciación en el continente Gondwana (hace aproximadamente 440-550 millones de años), la cual provocó la segunda extinción de especies marinas más grande en la historia de la tierra y donde una tercera parte de las familias de braquiopodos se extinguió [14].

§2.6. Artículo: Sobre la relación entre idoneidad del hábitat y la abundancia poblacional bajo diferentes escenarios de dispersión.



Ecología

Sobre la relación entre idoneidad del hábitat y la abundancia poblacional bajo diferentes escenarios de dispersión

On the relationship between habitat suitability and population abundance under different dispersal scenarios

Luis Alfredo Osorio-Olvera^{a,*}, Manuel Falconi^b y Jorge Soberón^c

^a Facultad de Ciencias, Universidad Nacional Autónoma de México, Circuito exterior s/n, Cd. Universitaria, 04510, Ciudad de México, México

^b Departamento de Matemáticas, Facultad de Ciencias, Universidad Nacional Autónoma de México, Circuito exterior s/n, Cd. Universitaria, 04510, Ciudad de México, México

^c Biodiversity Institute, University of Kansas, Dyche Hall, 1345 Jayhawk, Boulevard, Lawrence, KS, 66045, Estados Unidos

Recibido el 26 de febrero de 2015; aceptado el 7 de abril de 2016

Disponible en Internet el 17 de agosto de 2016

Resumen

Trabajos recientes basados en modelación del nicho para estimar la distribución de las especies han aportado evidencias de que la abundancia de una especie decrece conforme aumenta la distancia al centroide del nicho climático de la especie. Esta relación empírica se ha determinado utilizando modelos correlativos que relacionan datos de presencia/ausencia de una especie con factores abióticos, bajo el supuesto de que las especies se encuentran en equilibrio con su ambiente, pero sin tomar en cuenta explícitamente procesos poblacionales. La falta de un mecanismo que relacione explícitamente procesos poblacionales con el nicho de la especie dificulta la interpretación de los resultados y la obtención de conclusiones generales. En este trabajo se explora la relación entre la idoneidad del hábitat y la abundancia bajo diferentes escenarios de dispersión, utilizando una especie virtual. Se encontró que, en un escenario sin dispersión, hay una clara correlación entre la abundancia y la idoneidad del hábitat, pero en aquellos en los que se incluye el proceso de dispersión, a medida que se incrementa la proporción de individuos que se dispersan, la correlación entre abundancia e idoneidad del hábitat disminuye. Estos resultados apoyan la idea de que existe una relación entre idoneidad del hábitat y las abundancias, sin embargo, también muestran que esta se encuentra modulada por la dispersión.

Derechos Reservados © 2016 Universidad Nacional Autónoma de México, Instituto de Biología. Este es un artículo de acceso abierto distribuido bajo los términos de la Licencia Creative Commons CC BY-NC-ND 4.0.

Palabras clave: Modelos de distribución de especies; Modelos híbridos; Núcleo de dispersión; Especie virtual

Abstract

Recent studies based on niche modeling to estimate the distribution of species have provided evidence that the abundance of a species decreases as the distance to the centroid of the climatic niche of the species increases. This empirical relationship is determined using correlative models that relate the presence/absence of a species with abiotic factors under the assumption that species are in equilibrium with its environment, but without taking into account population processes explicitly. The lack of a mechanism that explicitly relates population processes to the niche hinders to interpret results and obtain general conclusions. In this paper the relationship between habitat suitability and abundance under different dispersal scenarios is explored using a virtual species. A clear correlation between abundance and habitat suitability in scenarios without dispersal was found, on the other hand, this correlation decreased as the proportion of individuals who are dispersed increases. These results support the idea that there is a relationship between habitat suitability and abundances, however, they also show that this relationship is modulated by dispersal.

All Rights Reserved © 2016 Universidad Nacional Autónoma de México, Instituto de Biología. This is an open access item distributed under the Creative Commons CC License BY-NC-ND 4.0.

Keywords: Species distribution models; Hybrid models; Dispersal Kernel; Virtual species

* Autor para correspondencia.

Correo electrónico: luismurao@comunidad.unam.mx (L.A. Osorio-Olvera).

La revisión por pares es responsabilidad de la Universidad Nacional Autónoma de México.

Introducción

Entender los patrones de distribución de las especies es uno de los tópicos más importantes de la ecología (Sagarin, Gaines y Gaylord, 2006; VanDerWal, Shoo, Johnson y Williams, 2009). Los patrones de distribución están determinados por factores ecológicos, evolutivos y geográficos, que hacen que el estudio de la distribución de las especies sea un problema complejo (Cain, 1994; Gaston, 2003; Udvary, 1969). Desde el trabajo clásico de Grinnell (1917), se ha considerado que las preferencias ambientales de una especie, denominadas «nicho ecológico», determinan, a través de los parámetros demográficos de cada población, las posibilidades de que una especie colonice exitosamente una localidad. En conjunto con el nicho, parámetros demográficos como la reproducción, la mortalidad y la dispersión, determinan la dinámica metapoblacional y por ende la abundancia y la distribución a diferentes escalas espaciales (Guisan y Thuiller, 2005; Lira-Noriega, Soberón y Milller, 2013; Marion et al., 2012; Schurr et al., 2007; Thuiller et al., 2008).

La forma en la que se relaciona el nicho ecológico con la abundancia es un tema que ha llamado la atención (Siqueira, Bini, Cianciaruso, Oliveira y Trivinho-Strixino, 2009; VanDerWal et al., 2009). En términos generales, se espera que la reproducción y la sobrevivencia sean altas en localidades ubicadas al interior del nicho de una especie y por ende, se alcancen abundancias poblacionales relativamente altas, y lo contrario en localidades periféricas (de Moraes y Viveiros, 2012; Maguire, 1973; Martínez-Meyer, Díaz-Porras, Townsend y Yáñez-Arenas, 2013; Siqueira et al., 2009; VanDerWal et al., 2009). Esta idea fue sugerida por Hutchinson (1957), quien expresa la posibilidad de que exista una región «óptima» dentro del nicho, es decir, con condiciones óptimas para la sobrevivencia de la especie, y regiones con condiciones menos favorables en la periferia. Maguire (1973) es el primero en proponer explícitamente la idea de que diferentes regiones del espacio de nicho corresponden a diferentes valores de la tasa intrínseca de crecimiento y de la abundancia poblacional. Una idea derivada de lo anterior es la hipótesis del «centroide del nicho», la cual plantea que las mayores abundancias se alcanzan en el centro del nicho y conforme una población esté más alejada de este, tenderá a presentar menores abundancias (Martínez-Meyer et al., 2013).

Se ha empezado a acumular evidencia empírica de que la posición de una población en el espacio de nicho determina su abundancia (de Moraes y Viveiros, 2012; Escalante y Martínez-Meyer, 2013; Martínez-Meyer et al., 2013; Siqueira et al., 2009; Torres et al., 2012; VanDerWal, 2009; Yáñez-Arenas, Martínez-Meyer, Mandujano y Rojas-Soto, 2012; Yáñez-Arenas, Guevara, Martínez-Meyer, Mandujano y Lobo, 2014). En estos trabajos se utilizan algoritmos de estimación de nichos para obtener una aproximación al nicho fundamental (medido como «idoneidad de hábitat»), y luego se establece una correlación entre la idoneidad de hábitat con datos de abundancia tomados en campo o de la literatura. Otra forma en la que se ha analizado la relación entre el nicho y la abundancia es identificando el centroide del nicho, calculado a partir de los valores promedio de distintas variables climáticas, y estableciendo una correlación entre la idoneidad de hábitat y la distancia a ese

centroide. En ninguno de estos dos casos se incluyen explícitamente procesos poblacionales como natalidad, mortalidad o migración, etc. Por tal motivo, si bien estos procedimientos pueden revelar la existencia de una relación entre abundancia y posición en el espacio de nicho, los resultados son correlaciones que no permiten profundizar en el entendimiento de los mecanismos que determinan esta relación (Dormann, 2009; Pagel y Schurr, 2012).

Se han propuesto diversos métodos para integrar aspectos demográficos en el análisis de los patrones de distribución y abundancia. Los modelos espacio-temporales, por ejemplo, combinan medidas del nicho con parámetros demográficos como la dispersión y el crecimiento poblacional. Se les denomina modelos acoplados o híbridos y pueden dar estimaciones de abundancia en un espacio geográficamente explícito (Catterall, Cook, Marion, Butler y Hulme, 2012; Nenzén, Swab, Keith y Araújo, 2012; Pagel y Schurr, 2012). Estos modelos se han utilizado frecuentemente para estimar la distribución de las especies en escenarios de cambio climático (i.e., Franklin, 2010), y para estudiar la relación entre demografía y la dinámica del área de distribución de algunas especies invasoras (Franklin, 2010; Gallien, Münkemüller, Albert, Boulangeat y Thuiller, 2010; Jongejans et al., 2008; Marion et al., 2012; Pagel y Schurr, 2012; Thuiller et al., 2008).

Trabajos como el de Jongejans et al. (2008) utilizan un modelo híbrido que incluye mecanismos poblacionales y dispersión para estimar la dispersión de una planta, pero sin referencia alguna al concepto de nicho. A su vez, Kearny y Porter (2009) explican cómo los modelos mecanísticos basados en procesos biofísicos ecofisiológicos, se pueden utilizar para estimar la distribución de una especie, pero no muestran cómo relacionarlos con los procesos demográficos y de dispersión. Buckley et al. (2010) proponen un modelo híbrido en el cual la tasa de crecimiento es función del ambiente, sin incluir al proceso de dispersión. La relación entre abundancia y nicho se menciona en un par de líneas, pero no se analiza, y finalmente, Pagel y Schurr (2012) utilizan un modelo híbrido sin estructura de edades para analizar cómo la tasa de crecimiento, la capacidad de carga y la densidad poblacional se relacionan con la probabilidad de presencia de una especie.

El objetivo de este trabajo es analizar la relación entre la abundancia y la posición en el espacio de nicho utilizando un modelo híbrido, que incorpora la estructura de edades de la población y que integra algunos parámetros poblacionales básicos, así como el proceso de dispersión. Con esta aproximación novedosa se pretende contribuir a la teoría que estudia la relación entre abundancia poblacional y la posición en el espacio del nicho, de tal forma que permita profundizar en el entendimiento de los mecanismos que determinan esta relación.

Materiales y métodos

Para estudiar la relación entre la idoneidad del hábitat y las abundancias poblacionales bajo diferentes escenarios de dispersión se utilizó una especie virtual. Se asumió que el nicho fundamental de la especie se mide en términos de la idoneidad del hábitat, la cual está definida por una función $f(x, y)$ donde

x y y son las variables ambientales que definen las condiciones donde la especie puede sobrevivir. Se considera que la tasa óptima de reproducción de la especie se obtiene en un cierto punto (\bar{x}, \bar{y}) del nicho fundamental de modo tal que $f(x, y)$ es menor que $f(\bar{x}, \bar{y})$ para todo $(x, y) \neq (\bar{x}, \bar{y})$.

En este trabajo se utilizó como función $f(x, y)$ a la función de densidad de probabilidad normal bivariada (ecuación 1) (Broennimann et al., 2012; Saupé et al., 2012), en donde la tasa óptima de reproducción $f(\bar{x}, \bar{y})$ se alcanza en las medias de las variables independientes x y y . La función de distribución multinormal bivariada es la siguiente:

$$f(x, y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp \left[-\frac{1}{2(1-\rho^2)} \left(\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} - \frac{2\rho(x-\mu_x)(y-\mu_y)}{\sigma_x\sigma_y} \right) \right], \quad (1)$$

donde ρ es la correlación entre la variable x y la variable y ; μ_x y μ_y son las medias de cada una de las variables independientes; σ_x y σ_y son las varianzas de ambas. Este método es equivalente a utilizar un método envolvente tipo distancia de Mahalanobis (Farber y Kadmon, 2003).

La función de la ecuación 1 es una función de densidad de probabilidad y por lo tanto esta se reescaló para que sus valores se encuentren en el intervalo cerrado [0,1]. La función de reescalamiento está dada por la expresión:

$$S(x, y) = \frac{f(x, y)}{\max [f(x, y) | x, y \in \mathbb{R}]}. \quad (2)$$

Dicha función de reescalamiento (ecuación 2) representa la denominada idoneidad del hábitat (Nenzén et al., 2012; Helaouet, Beaugrand y Edwards, 2013), en la que el valor de $S(x, y) = 1$ ocurre en los niveles de las variables ambientales donde se alcanza la tasa máxima de crecimiento poblacional.

Modelado de la dinámica espacial de crecimiento de la especie

Para analizar la relación entre abundancia poblacional y posición en el espacio de nicho bajo diferentes escenarios de dispersión, se utilizó un modelo híbrido de dinámica discreta de la forma dada por:

$$N_{t+1}(lon, lat) = S(x(lon, lat), y(lon, lat)) M N_t(lon, lat) + \Delta(lon, lat), \quad (3)$$

donde N_{t+1} es el vector que contiene las abundancias de la población en cada uno de los estadios en las coordenadas (lon, lat) al tiempo $t+1$. La matriz $M = (a_{i,j})$ es la matriz de proyección poblacional de tipo Lefkovich (Caswell, 2000) y contiene: 1), las proporciones que pasan de un estadio a otro (elementos a_{ij} con $i \neq j$); 2), la proporción de individuos que se mantiene en el mismo estadio (elementos de la diagonal principal) y 3), las fertilidades de los estadios (primer renglón de la matriz). Cada elemento a_{ij} es una variable aleatoria con una

Tabla 1
Matriz de proyección poblacional de *Astragalus scaphoides* (NR=no reproductiva).

	Latente	Pequeña (NR)	Grande (NR)	Reproductiva
Latente	0.50	0.03	0.04	5.00
Pequeña (NR)	0.40	0.45	0.04	0.00
Grande (NR)	0.00	0.24	0.37	0.16
Reproductiva	0.10	0.06	0.52	0.80

distribución normal $N(\mu_{i,j}, \sigma)$. El vector Δ representa el flujo neto de inmigrantes de cada estadio al parche con coordenadas (lon, lat) . Con $\Delta_{k,r} N_{lon,lat}^i$ se denota a los inmigrantes del estadio i al parche (lon, lat) provenientes del parche (k, r) ; es decir, $\Delta_{k,r} N_{lon,lat}^i = \Delta_{k,r}^{lon,lat} N_{k,r}^i - \Delta_{lon,lat}^{k,r} N_{lon,lat}^i$, donde $\Delta_{\gamma}^v N_{\gamma}^i$ es el número de individuos de la clase i que migran del parche γ al parche v ; por tanto la inmigración neta Δ está dada por:

$$\Delta(lon, lat) = \left(\sum_{k,r} \Delta_{k,r} N_{lon,lat}^1, \sum_{k,r} \Delta_{k,r} N_{lon,lat}^2, \dots, \sum_{k,r} \Delta_{k,r} N_{lon,lat}^n \right). \quad (4)$$

$S(x, y)$ es una matriz diagonal de $n \times n$ dimensiones con las idoneidades de cada estadio (ecuación 5). Las variables x, y dependen de la posición (lon, lat) , por lo tanto,

$$S(x(lon, lat), y(lon, lat)) = \begin{pmatrix} S_{lon,lat}^1 & \cdot & \cdot & \cdot & 0 \\ 0 & S_{lon,lat}^2 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & S_{lon,lat}^n \end{pmatrix}. \quad (5)$$

De la estructura del modelo (ecuación 3) se sigue que hay un valor umbral mínimo de idoneidad S_0 por debajo del cual la supervivencia de la especie es nula; es decir, en ausencia de dispersión, si $S^i(x, y) \leq S_0$ para todo $i = 1, \dots, n$ entonces la población en (x, y) tiende a extinguirse debido a que su tasa intrínseca de crecimiento λ será menor a 1. El valor S_0 se calcula como el inverso de la tasa máxima intrínseca de crecimiento

$$S_0 = \frac{1}{\lambda_{max}}, \quad (6)$$

Nótese que el valor máximo de la tasa intrínseca de crecimiento λ_{max} se alcanza en el valor de idoneidad $S(\bar{x}, \bar{y}) = 1$.

Parametrización del modelo híbrido

La matriz M que se utilizará en las simulaciones es la correspondiente a la especie *Astragalus scaphoides* (tabla 1) tomada de la base de datos demográficos COMPADRE (COMPADRE Plant Matrix Database, 2015). Los estadios son latente, pequeña, grande y reproductiva. Para la dispersión se considerarán dos

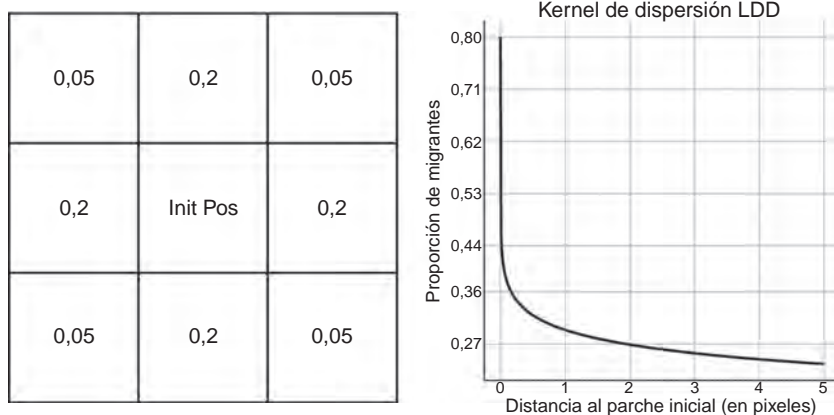


Figura 1. Kernels de dispersión. Imagen izquierda, dispersión a las celdas contiguas (SDD). Imagen derecha, dispersión a celdas distantes (LDD) como función de la distancia d al parche inicial S .

kernels de dispersión que indican el flujo neto de inmigrantes (esto corresponde al término Δ de las ecuaciones 3 y 4). El primer kernel se refiere a una dispersión a celdas cercanas (kernel SDD) y el segundo a la dispersión a celdas lejanas (kernel LDD). El porcentaje de la población de un parche que se dispersa a los parches contiguos SDD, lo denotamos con α y la que se dispersa a los más lejanos LDD, lo denotamos con β . Se considera que la dispersión ocurre solo en el estadio de semilla. El kernel SDD se da entre las 8 celdas contiguas a cada parche, con una proporción de 0.2α colonizadores hacia cada una de las 4 celdas laterales y una proporción de 0.05α hacia cada celda diagonal (fig. 1a). La dispersión a celdas no contiguas está determinada por un kernel (LDD) de dispersión exponencial del tipo siguiente:

$$P(d) = a \exp \left[-d \frac{c}{b} \right] \quad \text{con } a, b, c \in [0, 1], x \leq D_{\max}, \quad (7)$$

donde D_{\max} es la distancia máxima a la que puede viajar un individuo y $P(d)$ es la fracción de migrantes que viajan a una distancia d . La proporción máxima de la población que se desplaza está determinada por a . La razón entre c y b define que tan grande es la proporción de individuos que se desplazan a una distancia d . Los valores de los parámetros elegidos fueron $a=0.8$, $b=0.8$, $c=0.1$ y $D=2$. La distribución del kernel dados estos parámetros se muestra en la figura 1.

Para definir el nicho fundamental de la especie virtual, se eligieron como variables ambientales la temperatura media anual y la precipitación en el mes más húmedo, por ser las más importantes para el desarrollo y supervivencia de varias especies (Bartlein, Prentice y Webb III, 1986; Woodward, 1987). Para analizar el proceso de dispersión se eligió como «espacio geográfico» una zona al noroeste de México. En este espacio la especie virtual inicia el proceso de dispersión, y se definió tomando como base el nicho de la especie, es decir los sitios del espacio geográfico donde la especie puede subsistir. Los valores de temperatura media anual (bio1) y de precipitación en el mes más húmedo (bio13) fueron obtenidos de la página de Worldclim (Hijmans, Cameron, Parra, Jones y Jarvis, 2005) a una resolución de 30 segundos de arco. Aquí suponemos que la tasa óptima de reproducción de la especie virtual se alcanza en las medias (μ_x , μ_y) de bio1 y bio13 en el espacio geográfico, los

cuales fueron 24.2°C y 260 mm (este es el centroide del nicho). La correlación de las variables nichos fue de $\rho = -0.77$ con una desviación de $\sigma_T = 3.89$ para la bio1 y $\sigma_P = 9.85$ para bio13. El valor de la tasa máxima intrínseca de crecimiento (λ_{\max}) de la matriz \mathbf{M} en condiciones de idoneidad máxima $S(\mu_x, \mu_y) = 1$ es de $\lambda_{\max} = 1.6$.

Utilizando la función rasterToPoints del paquete «raster» de R (Hijmans, 2015), se extrajeron los valores de bio1 y bio13 para el espacio geográfico en cuestión, y se aplicaron las ecuaciones (1) y (2) para obtener la idoneidad $S(x, y)$ en cada una de las celdas (fig. 2). Se eligieron aleatoriamente 10 celdas cuya idoneidad del hábitat fuera mayor al valor de S_0 (en este caso $S_0 = 0.625$). En estas celdas se colocaron las poblaciones iniciales de la especie virtual y constituyeron el área inicial de distribución (puntos negros en el espacio geográfico, fig. 3).

Los escenarios

Se puso a prueba la hipótesis del centroide del nicho (la mayor abundancia se encuentran más cercanos al centroide del nicho)

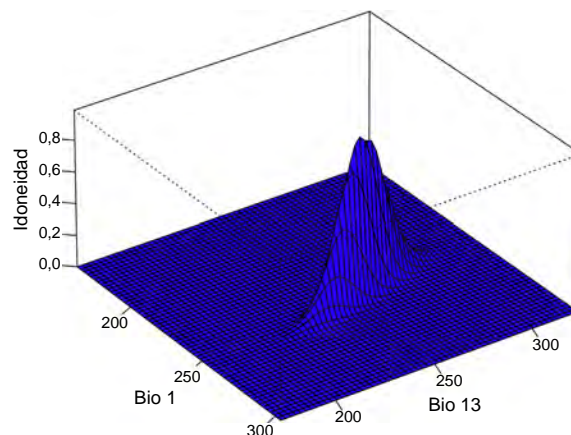


Figura 2. Nicho fundamental de la especie virtual. La función de distribución $f(x, y)$ que se utilizó para definir al nicho fundamental de la especie virtual, fue la función multinormal bivariada; se eligió como variable x a la temperatura media anual (bio1) y como variable y a la precipitación en el mes más húmedo (bio13).

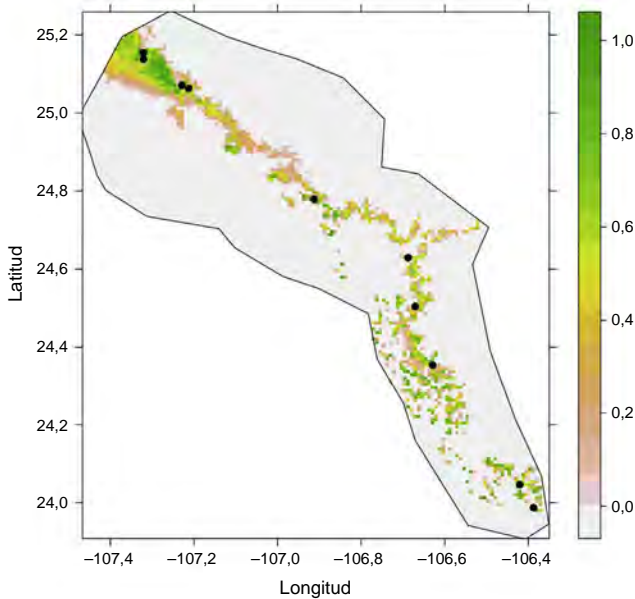


Figura 3. Mapa de idoneidad del hábitat. Los puntos negros muestran los sitios donde se colocaron a las 10 poblaciones iniciales de la especie virtual.

bajo distintos escenarios de dispersión: *a*) sin dispersión de la especie, i.e. $\alpha = 0 = \beta$; *b*) con dispersión; se simularon un total de 36 escenarios de dispersión, los cuales son resultado de combinar diferentes proporciones de α y β . Los valores posibles de α y β son [0.01, 0.04, 0.06, 0.8, 0.11, 0.14]. Para realizar las simulaciones se utilizó el paquete *demoniche* de R (Nenzén et al., 2012).

Se realizaron regresiones lineales entre el índice de idoneidad y la abundancia poblacional, usando los coeficientes de determinación. La comparación de las r^2 permitió comparar los resultados entre los diferentes escenarios de dispersión.

Resultados

Modelo sin dispersión

Las mayores abundancias se encontraron en los valores de idoneidad de hábitat S cercanos al óptimo de la especie (i.e. centroide del nicho). La correlación r fue positiva y la r^2 fue de 0.95 (fig. 4). La variación no explicada se debe a la estocasticidad introducida en los parámetros demográficos del modelo.

Modelo con dispersión

En general, los valores que tomó la abundancia fueron más altos conforme la proporción de individuos con dispersión a celdas cercanas SDD fue mayor. Por ejemplo, el número máximo de individuos que alcanzó un parche en el escenario LDD = 0.01 SDD = 0.01 fue de 3,500 individuos, mientras que en el escenario de mayor dispersión (LDD = 0.01 SDD = 0.08) fue de 120,000 individuos (fig. 5).

La relación idoneidad del hábitat S y abundancia fue positiva para todos los escenarios. Sin embargo, se observó que conforme incrementó la proporción de α individuos que viajaban a distancias cortas (SDD), los coeficientes de determinación r^2 resultaron considerablemente más bajos (fig. 6). En estos escenarios, conforme incrementó la dispersión, la relación entre abundancia e idoneidad se hace más débil, por lo tanto fue posible encontrar abundancias altas en sitios con idoneidad baja (fig. 6).

Las relaciones que se encontraron entre la idoneidad del hábitat y la abundancia poblacional para el escenario donde la proporción de individuos susceptibles a dispersarse a las largas distancias es de $0.01 \leq \beta = 0.06$, con proporciones de $0.01 \leq \alpha \leq 0.14$ individuos susceptibles dispersarse a cortas distancias, muestran una correlación r positiva en todos los casos,

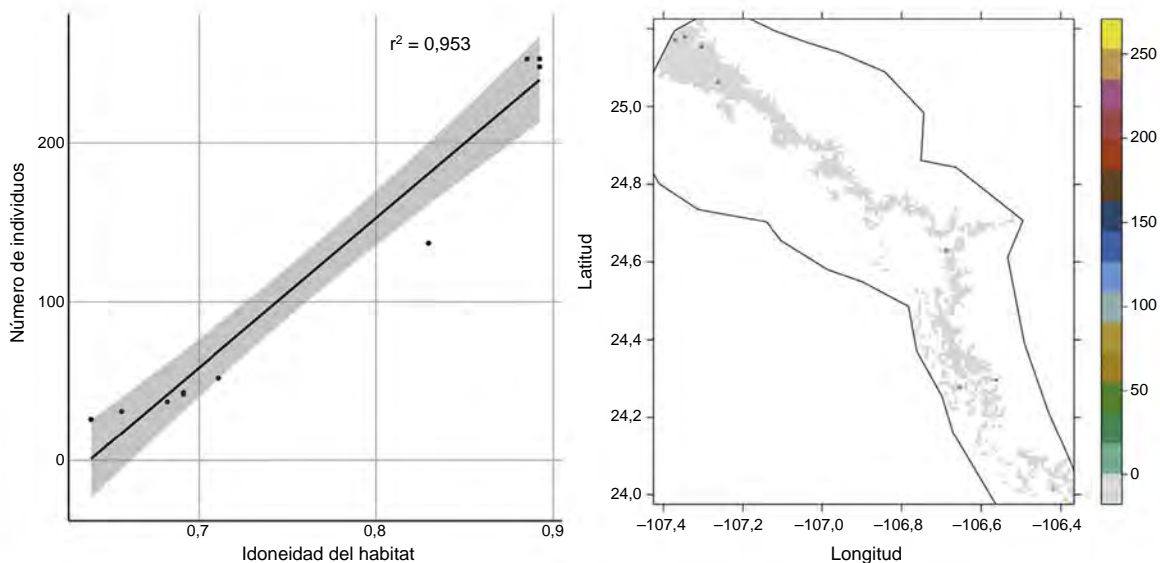


Figura 4. Relación entre idoneidad del hábitat y la abundancia para el escenario donde no hay dispersión. Imagen izquierda, la correlación entre estas variables fue de $r^2 = 0.95$, la banda alrededor de la línea de tendencia representa el intervalo de confianza de la estimación a un nivel de confianza del 95%. Imagen derecha, proyección de las abundancias en el espacio geográfico.

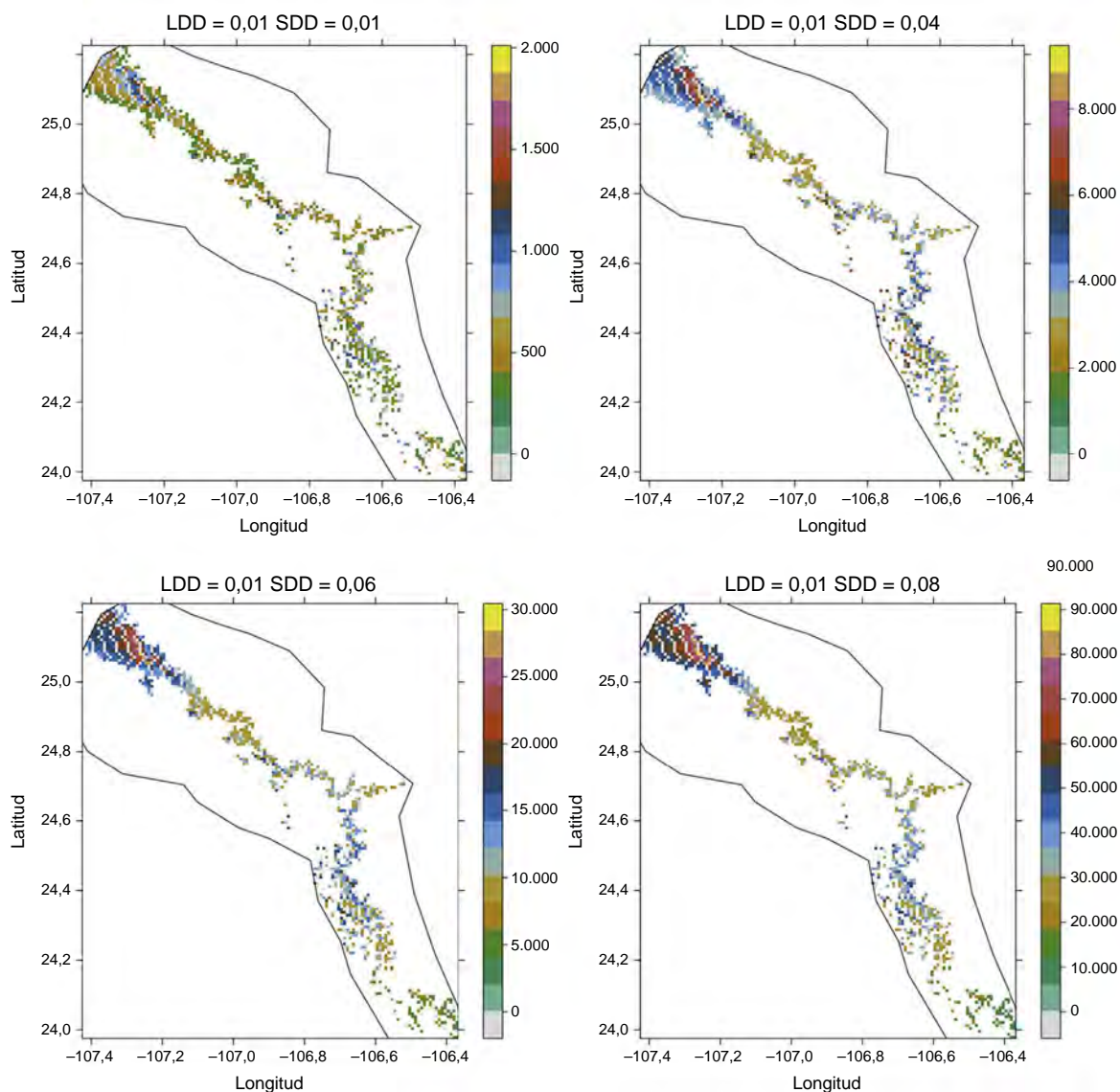


Figura 5. Proyección al espacio geográfico de las abundancias poblacionales para diferentes escenarios de dispersión a cortas distancias SDD ($0.01 \leq \alpha \leq 0.08$) y valores de dispersión a largas distancias LDD fijos ($\beta = 0.01$).

sin embargo también muestran que tanto r como r^2 disminuyen al incrementar las proporciones α y β (fig. 6).

En términos generales las r^2 obtenidas en las simulaciones para todos los escenarios de dispersión disminuyeron conforme se incrementaron las proporciones de individuos con dispersión a celdas cercanas SDD (α) y de individuos dispersión a celdas no contiguas LDD (β) (fig. 7).

Discusión

En este trabajo analizamos la relación entre la abundancia de una especie virtual y su posición en el espacio de nicho, utilizando un modelo «híbrido» con estructura de edades que integra procesos demográficos y de dispersión. A pesar de que en la literatura hay trabajos en los que se utilizan modelos híbridos y mecanicistas para estudiar la distribución y abundancia de las

especies, la diferencia principal radica en que el nuestro integra diferentes escenarios de dispersión.

El principal resultado que encontramos es que existe una correlación positiva entre abundancia poblacional y la idoneidad de hábitat, es decir, se confirma la hipótesis del centroide del nicho. Otro resultado fundamental es que esta correlación está regulada por la relación que guardan los parámetros demográficos respecto a los de la dispersión.

Los resultados obtenidos muestran que la intensidad de la correlación positiva entre abundancia e idoneidad del hábitat depende de la capacidad de dispersión de la especie en cuestión: la correlación entre abundancia e idoneidad del hábitat es casi perfecta en un escenario de no dispersión; sin embargo, cuando se incluye el factor dispersión, la correlación declina abruptamente pudiéndose encontrar poblaciones con abundancias mayores en sitios donde la calidad del hábitat es baja. Este resultado concuerda con lo señalado por Pulliam (2000), quien

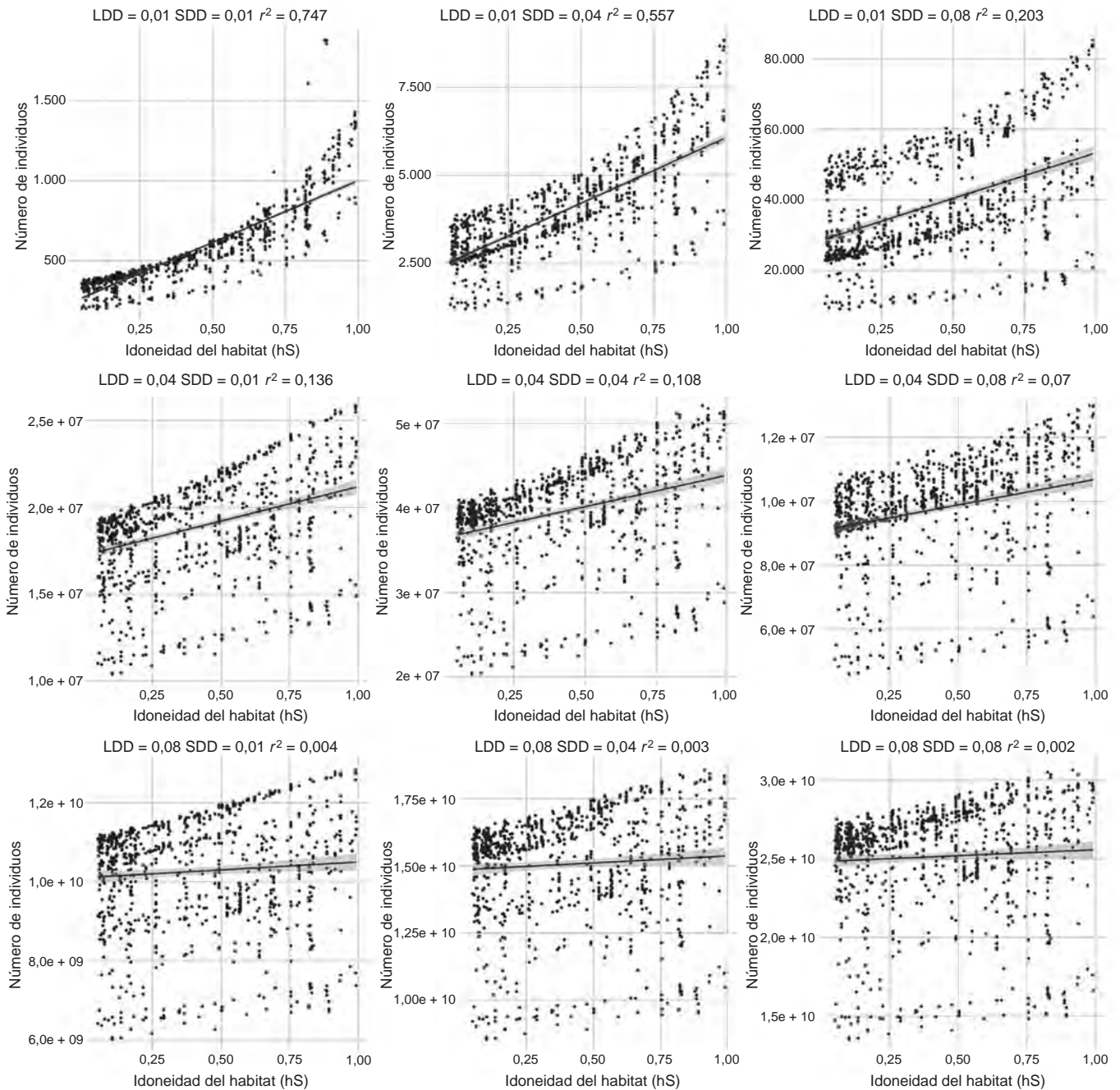


Figura 6. Correlaciones entre idoneidad del hábitat y las abundancias poblacionales para diferentes escenarios de dispersión a cortas distancias SDD y a largas distancias LDD. La banda alrededor de las líneas de tendencia representa el intervalo de confianza de la estimación a un nivel de confianza del 95%.

afirmó que las especies frecuentemente se encuentran en sitios donde su tasa de crecimiento es negativa (Pulliam, 2000; Schurr et al., 2012). A este fenómeno se le ha denominado dinámica fuente-sumidero (Pagel y Schurr, 2012; Pulliam, 1988, 2000; Schurr et al., 2012) y depende en gran medida de si las respuestas densodependientes afectan la dinámica de dispersión de la especie.

Nuestros resultados refuerzan las predicciones teóricas informales derivadas de la teoría del nicho de Hutchinson (Maguire, 1973; Schurr et al., 2012) y da elementos teóricos a los resultados empíricos como los de VanDerWal et al. (2009), Escalante y Martínez-Meyer (2013) y Martínez-Meyer et al. (2013), ya que

dan una explicación plausible en términos poblacionales a las correlaciones que se obtienen entre abundancia y distancia en el espacio de nicho. Además, sugieren algunas hipótesis acerca de las condiciones en las que los procesos demográficos y de dispersión pueden hacer que la correlación entre abundancia y distancia al centroide del nicho sea baja, por ejemplo, se espera que las especies cuyas tasas de migración son altas, muestren una correlación baja entre su abundancia y su nicho.

Asimismo, nuestro trabajo complementa estudios como el de Holt, Lawton, Gaston y Blanchette (1997) en el que se analiza la relación entre la abundancia y el rango de distribución de la especie. A diferencia del enfoque de Holt et al., que se basa en la

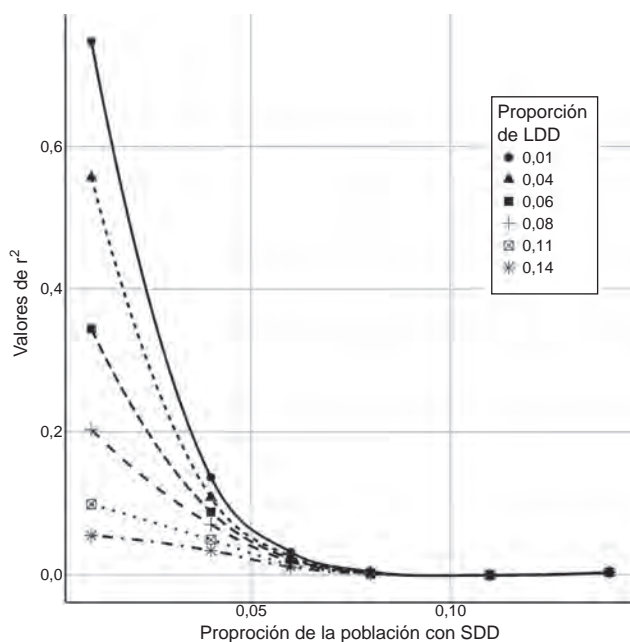


Figura 7. Valores de r^2 para todos los escenarios de dispersión a grandes y a cortas distancias.

idea de que el proceso de dispersión de la especie ha concluido y no consideran el efecto de la migración-inmigración sobre las abundancias locales, nosotros mostramos cómo el proceso de dispersión y el nicho pueden determinar la abundancia de las poblaciones en el rango geográfico.

Finalmente, si bien el enfoque de los modelos híbridos es prometedor, existe una limitante práctica para incluir el término de dispersión de estos modelos, ya que la estimación de los kernels de dispersión es difícil, los datos son escasos y kernels con colas largas son matemáticamente difíciles de manejar (Barve et al., 2011). A pesar de que en este trabajo se enfatizó el papel de la dispersión en la relación entre abundancia y posición en el espacio de nicho, otra limitante importante para la comprensión de esta relación son las interacciones, de manera que la ecuación de abundancia de una especie en un espacio geográficamente explícito debe ser una función de su nicho fundamental, su capacidad de dispersión y las interacciones que tenga con las demás especies (Soberón, 2010). El trabajo que presentamos aquí pretende ser una base teórica sobre la cual avanzar incluyendo este otro tipo de factores, lo cual permitirá entender mejor la compleja red de factores que determinan la geometría y la estructura interna del área de distribución de una especie.

Agradecimientos

Se agradece al Posgrado en Ciencias Biológicas por la formación académica de Luis Osorio. Asimismo se agradece al PAPIIT-UNAM IN112715 (2015) por el apoyo parcial brindado para la realización de esta investigación. Jorge Soberón agradece al proyecto NSF 1208472. Agradecemos a Andrés Lira-Noriega y a los revisores anónimos cuyos comentarios nos ayudaron a mejorar la presentación del trabajo.

Referencias

- Bartlein, P. J., Prentice, I. C. y Webb, T., III. (1986). Climatic response surfaces from pollen data for some eastern North American Taxa. *Journal of Biogeography*, 13, 35–57.
- Barve, N., Barve, V., Jiménez-Valverde, A., Lira-Noriega, A., Maher, S. P., Peterson, A. T., et al. (2011). The crucial role of the accessible area in ecological niche modeling and species' distribution modeling. *Ecological Modelling*, 222, 1810–1819.
- Broennimann, O., Fitzpatrick, M. C., Pearman, P. B., Petitpierre, B., Pellissier, L., Yoccoz, N. G., et al. (2012). Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, 21, 481–497.
- Buckley, L., Urban, M., Angilletta, M., Crozier, L., Rissler, L. y Sears, M. (2010). Can mechanism inform species' distribution models? *Ecology Letters*, 13, 1041–1054.
- Cain, S. A. (1944). *Foundations of plant geography*. New York: Harper & Brothers.
- Caswell, H. (2000). *Matrix population models: construction, analysis and interpretation*. Sunderland, Massachusetts: Sinauer Associates.
- Catterall, S., Cook, A. R., Marion, G., Butler, A. y Hulme, P. E. (2012). Accounting for uncertainty in colonization times: a novel approach to modelling the spatio-temporal dynamics of alien invasions using distribution data. *Ecography*, 35, 901–911.
- COMPADRE Plant Matrix Database (2015). Max Planck Institute for demographic research (Germany). (consultado 19 Sep 2015) Disponible en: www.compadre-db.org.
- De Moraes, M. y Viveiros, C. (2012). Does environmental suitability explain the relative abundance of the tailed tailless bat, *Anoura caudifer*? *Natureza a Conservacao*, 10, 221–227.
- Dormann, C. F. (2009). Response to comment on methods to account for spatial autocorrelation in the analysis of species' distributional data: a review. *Ecography*, 32, 379–381.
- Escalante, T. y Martínez-Meyer, E. (2013). Ecological niche modeling and wildlife management units (UMAS): an application to deer in campeche, México. *Tropical and Subtropical Agroecosystems*, 16, 183–191.
- Farber, O. y Kadmon, R. (2003). Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecological Modeling*, 160, 115–130.
- Franklin, J. (2010). Moving beyond static species' distribution models in support of conservation biogeography. *Diversity and Distributions*, 16, 321–330.
- Gallien, L., Münkemüller, T., Albert, C. H., Boulangeat, I. y Thuiller, W. (2010). Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions*, 16, 331–342.
- Gaston, K. J. (2003). *The structure and dynamics of geographic ranges*. Oxford series in ecology and evolution. New York: Oxford University Press.
- Grinnell, J. (1917). The niche-relationships of the California thrasher. *The Auk*, 34, 427–433.
- Guisan, A. y Thuiller, W. (2005). Predicting species' distribution: offering more than simple habitat models. *Ecology Letters*, 8, 993–1009.
- Helaouet, P., Beaugrand, G. y Edwards, M. (2013). Understanding long-term changes in species' abundance using a niche-based approach. *PloS one*, 8, e79186.
- Hijmans, R. J. (2015). Raster: geographic data analysis and modeling. R package version 2.4-18. (consultado 4 Nov 2015). Disponible en: <http://CRAN.R-project.org/package=raster>.
- Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G. y Jarvis, A. (2005). The WorldClim interpolated global terrestrial climate surfaces. (consultado el 01 Oct 2014). Disponible en: <http://www.worldclim.org/>.
- Holt, R. D., Lawton, J. H., Gaston, K. J. y Blanchette, C. A. (1997). On the relationship between range size and local abundance: back to basics. *Oikos*, 78, 183–190.
- Hutchinson, G. E. (1957). Concluding remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22, 415–427.
- Jongejans, E., Shea, K., Skarpaas, O., Kelly, D., Sheppard, A. W. y Woodburn, T. L. (2008). Dispersal and demography contributions to population spread of *Carduus nutans* in its native and invaded ranges. *Journal of Ecology*, 96, 687–697.

- Kearny, M. y Porter, W. (2009). Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, *12*, 334–350.
- Lira-Noriega, A., Soberón, J. y Miller, C. (2013). Process-based and correlative modeling of desert mistletoe distribution: a multiscale approach. *Ecosphere*, *4*, 1–23.
- Maguire, B., Jr. (1973). Niche response structure and the analytical potentials of its relationship to the habitat. *The American Naturalist*, *107*, 213–246.
- Marion, G., McInerny, G. J., Pagel, J., Catterall, S., Cook, A. R., Hartig, F., et al. (2012). Parameter and uncertainty estimation for process-oriented population and distribution models: data, statistics and the niche. *Journal of Biogeography*, *39*, 2225–2239.
- Martínez-Meyer, E., Díaz-Porras, D., Townsend, P. y Yáñez-Arenas, C. (2013). Ecological niche structure and range wide abundance patterns of species. *Biology Letters*, *9*, 1–5.
- Nenzén, H. K., Swab, R. M., Keith, D. A. y Araújo, M. B. (2012). Demoni-che - an R-package for simulating spatially-explicit population dynamics. *Ecography*, *35*, 577–580.
- Pagel, J. y Schurr, F. M. (2012). Forecasting species' ranges by statistical estimation of ecological niches and spatial population dynamics. *Global Ecology and Biogeography*, *21*, 293–304.
- Pulliam, H. R. (1988). Sources, sinks, and population regulation. *The American Naturalist*, *132*, 652–661.
- Pulliam, H. R. (2000). On the relationship between niche and distribution. *Ecology Letters*, *3*, 349–361.
- Sagarin, D. R., Gaines, S. y Gaylord, B. (2006). Moving beyond assumptions to understand abundance distributions across the ranges of species. *Trends in Ecology and Evolution*, *21*(9), 524–530.
- Saupe, E. E., Barve, V., Myers, C. E., Soberón, J., Barve, N., Hensz, C. M., et al. (2012). Variation in niche and distribution model performance: the need for a priori assessment of key causal factors. *Ecological Modelling*, *237*–238, 11–22.
- Schurr, F. M., Midgley, G. F., Rebelo, A. G., Reeves, G., Poschod, P. y Higgins, S. I. (2007). Colonization and persistence ability explain the extent to which plant species fill their potential range. *Global Ecology and Biogeography*, *16*, 449–459.
- Schurr, F. M., Pagel, J., Cabral, J. S., Groeneveld, J., Bykova, O., O'Hara, R. B., et al. (2012). How to understand species' niches and range dynamics: a demographic research agenda for biogeography. *Journal of Biogeography*, *39*, 2146–2162.
- Siqueira, T., Bini, L., Cianciaruso, M., Oliveira, F. y Trivinho-Strixino, S. (2009). The role of niche measures in explaining the abundance-distribution relationship in tropical lotic chironomids. *Hydrobiologia*, 1–10.
- Soberón, J. (2010). Niche and area of distribution modeling: a population ecology perspective. *Ecography*, *33*, 159–167.
- Thuiller, W., Albert, C., Araujo, M. B., Berry, P. M., Cabeza, M., Guisan, A., et al. (2008). Predicting global change impacts on plant species' distributions: future challenges. *Perspectives in Plant Ecology Evolution and Systematics*, *9*, 137–152.
- Torres, M. N., Júnior, P. D. M., Santos, T., Silveira, L., Jácomo-de Almeida, A. T. y Diniz-Filho, A. F. (2012). Can species' distribution modelling provide estimates of population densities? A case study with jaguars in the Neotropics. *Diversity and Distributions*, *18*, 615–627.
- Udvardy, M. D. F. (1969). *Dynamic zoogeography: with special reference to land animals*. New York: Van Nostrand Reinhold Comp.
- VanDerWal, J., Shoo, L., Johnson, C. y Williams, S. (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.
- Woodward, F. I. (1987). *Climate and plant distribution*. Cambridge: Cambridge University Press.
- Yáñez-Arenas, C., Guevara, R., Martínez-Meyer, E., Mandujano, S. y Lobo, J. M. (2014). Predicting species' abundances from occurrence data: effects of sample size and bias. *Ecological Modelling*, *294*, 36–41.
- Yáñez-Arenas, C., Martínez-Meyer, E., Mandujano, S. y 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.

Capítulo 3

Abundancia y posición en el espacio de nicho: procesos y medidas

Resumen

En este capítulo se consideran algunos casos de interés (no estudiados en el capítulo anterior) sobre la relación entre abundancia y posición en el espacio de nicho; de manera particular se estudia: 1) el efecto de la métrica en las correlaciones entre abundancia y posición respecto al centroide del nicho; específicamente se comparan a la distancia Euclidiana con la de Mahalanobis; 2) el efecto del tiempo que tiene establecida una población en una localidad dada; es decir se estudia la relación entre abundancia y posición en el espacio de nicho para diferentes tiempos de inicio del proceso de dispersión; 3) el efecto de una barrera a la dispersión, para ello se introduce una limitante al crecimiento poblacional donde se supone que para que una población crezca en un parche dado, es necesario superar un umbral de abundancia poblacional mínima N_0 por debajo del cual no hay crecimiento (efecto Allee). A diferencia del capítulo anterior donde se utilizó un modelo a tiempo discreto que no considera la denso-dependencia, en este capítulo se estudia la relación con un modelo a tiempo continuo con crecimiento denso-dependiente (logístico) [82].

Los resultados muestran que la métrica de distancia al centroide tiene un efecto importante en la correlaciones; en todos los escenarios estudiados se encontró una correlación positiva entre abundancia y posición en espacio de nicho, sin embargo, se determinó que la distancia de Mahalanobis tiene una mejor correlación que la Euclidiana. Asimismo en todos los escenarios se encontró que la correlación entre abundancia y posición en el espacio de nicho crece conforme las poblaciones alcanzan su estado estable; esto ocurre cuando las poblaciones en los parches se hacen más viejas. Finalmente, se determinó que el efecto Allee disminuye el área de distribución de la especie; de manera análoga se observó que las correlaciones entre abundancia y posición en el espacio de nicho disminuyen en función de la intensidad de dicho efecto.

§3.1. Introducción

Los patrones de distribución de las abundancias poblacionales de las especies en el espacio geográfico muestran una gran variación en todos los sitios donde estas se distribuyen [36, 101, 169].

Para explicarlos se han propuesto diversas hipótesis, por ejemplo, la hipótesis «centro abundante» la cual postula que las mayores abundancias se encuentran en el centro del área de distribución [17, 134]. Esta hipótesis ha sido criticada tanto desde el punto de vista empírico como teórico [71] debido a que pocos son los trabajos en los que se ha podido documentar esta relación *ver* [140, 141, 167]. Una propuesta surgida de la teoría del nicho de Hutchinson [76], es usar el índice de idoneidad del hábitat de los modelos de distribución de especies (MDE) para predecir la abundancia. Bajo este marco teórico, en la literatura hay varios trabajos donde se han encontrado correlaciones positivas entre la abundancia e idoneidad, sin embargo, la mayoría de ellas son bajas [170, 147, 165, 36, 184]; por ejemplo, Pearce y Ferrer [122] encontraron correlaciones significativas positivas pero bajas ($0.07 \leq r^2 \leq 0.5$) entre la abundancia y la idoneidad para 12 de las 44 especies de vertebrados y plantas que modelaron (el resto de las correlaciones no fueron significativas) utilizando regresiones logísticas. Análogamente Van der Wal *et al.* [170] observaron correlaciones positivas también bajas ($0.017 \leq r^2 \leq 0.486$) entre la abundancia y la salida logística de MaxEnt en 84 % de las 69 especies de vertebrados de bosque húmedo que analizaron. Asimismo Torres y colaboradores [165] reportaron correlaciones bajas para las abundancias del jaguar. Cabe señalar que son pocos los trabajos en donde se han reportado correlaciones relativamente altas, uno de ellos es el de Weber y Viveiros [36] quienes encontraron un coeficiente de determinación de $r^2 = 0.64$ para el murciélago *Anoura caudifer*.

Otra alternativa es correlacionar a la abundancia poblacional con una medida de posición respecto al centroide del nicho. Esta idea se deriva de los planteamientos propuestos por Maguire [96] y Brown [17] donde suponen que las condiciones óptimas para el desarrollo de la especie ocurren en las localidades que se encuentran más cercanas al centroide del nicho de la especie; como consecuencia de lo anterior, es muy probable que la tasa máxima de crecimiento ocurra en el centroide del nicho y por tanto, las poblaciones que se encuentren en este, presentarán las mayores abundancias.

Para cuantificar la distancia de un punto al centroide del nicho, se han utilizado frecuentemente a la distancia Euclidiana [101] y la distancia de Mahalanobis [44, 169]. Por ejemplo, Martínez-Meyer *et al.* en 2013 [101] utilizan distancias euclidianas como medida de distancia al centroide de *Toxostoma redivivum* obteniendo valores bajos de r^2 . Por otra parte, Yañez-Arenas *et al.* [185] utilizaron distancias de Mahalanobis y encontraron correlaciones muy altas ($r^2 = 0.902$) para *Odocoileus virginianus*; de forma similar Ureña *et al.* [169] reportan correlaciones de $r^2 > 0.6$ para la tortuga mexicana.

A pesar de que se ha encontrado que existe una correlación positiva entre la abundancia y la posición en el espacio de nicho para varios grupos de organismos, los valores de las correlaciones obtenidas en los trabajos citados varían ampliamente. Esto se puede deber, entre otros factores, a que estas métricas son aproximaciones estáticas a las abundancias que no capturan la dinámica espacio-temporal de las poblaciones [162, 146, 185, 169]; así a pesar de que en el la geografía hayan lugares con valores de idoneidad altos o bien muy cercanos al centroide de la especie, la accesibilidad a ellos puede estar limitada por barreras geográficas que impiden el acceso a dichas zonas [8] y/o por limitantes biológicas como las interacciones con otras especies [182]. Soberón y Peterson [152] ilustran de manera heurística estas ideas en el diagrama BAM donde sugieren que el área de distribución de una especie está determinada por 3 factores que interactúan a diferentes escalas: a) los factores abióticos que definen al nicho de una especie; b) los bióticos los cuales pueden estar determinados por las interacciones con otras especies o por la biología de la misma; y c) la capacidad de movimiento. Adicionalmente un factor que no ha sido bien explorado, pero que puede afectar las correlaciones entre abundancia poblacional y posición en el espacio de nicho tiene que ver con los factores históricos de las poblaciones, como su edad o dicho de otra manera, el tiempo que

tiene establecida una población en una localidad dada (*ver* [83]); siguiendo esta idea es plausible conjeturar que haya poblaciones cercanas al centroide de la especie con densidades más bajas que otras en una posición más alejada debido a que las primeras recién fundaron aquellos sitios cercanos a su centroide y están en proceso de crecimiento.

Con base en los planteamientos anteriores, en este capítulo se estudia la relación entre abundancia y posición en el espacio de nicho de una especie virtual que presenta dispersión en el espacio geográfico, para ello se considera:

1. el efecto de la métrica en la relación; específicamente se compara a la distancia Euclidiana con la de Mahalanobis como medidas de la distancia al centroide de la especie.
2. el tiempo que llevan las poblaciones establecidas en las localidades; se analizan las correlaciones entre abundancia y distancia al centroide para diferentes tiempos de inicio de la dispersión.
3. el efecto de la estructura espacial de las idoneidades en relación a los sitios donde comienza el proceso de dispersión, con los siguientes escenarios:
 - a) todos los sitios son accesibles para la especie.
 - b) una barrera biológica a la dispersión; para que una población se pueda establecer, es necesario superar un umbral mínimo de densidad (efecto Allee).

Una forma de estudiar los problemas anteriores es utilizar un modelo basado en procesos de una población estructurada espacialmente y cuya estructura de nicho está definida de modo tal que el resultado esperado es una relación directa entre la posición de nicho y la abundancia. Siguiendo las ideas de Vandermeer [172], Keitt et al. [83] y otros [153] quienes proponen la estructura general de las ecuaciones que se usan en este trabajo (un modelo logístico de crecimiento continuo en una retícula espacial discreta conectada por la dispersión) se analizan los tres casos mencionados. A pesar de que es un modelo simple en el sentido de que es una sola especie con ecología determinista y no hay evolución, el modelo es muy general y robusto; también, es complicado tanto numéricamente como analíticamente debido a que es un sistema de muchas ecuaciones diferenciales acopladas.

§3.2. Modelo matemático

En esta sección se describe el modelo matemático que se utilizó modelar la dinámica de crecimiento espacio-temporal de la especie virtual en la que se estudia el efecto de los 3 casos de interés en la relación entre posición en el espacio de nicho y la abundancia. También se explica como se define el nicho fundamental de la especie virtual y el Kernel de dispersión.

Se asume una región dividida por una retícula regular con $i = 1, 2, 3, \dots, n$ parches. Cada parche está caracterizado por variables ambientales que tienen un efecto directo en la tasa de crecimiento de las poblaciones en los parches i . Las poblaciones en los parches i crecen de manera sigmoidea y su tamaño depende de: una tasa intrínseca de crecimiento que a su vez es función de las condiciones ambientales en el parche i ; un factor de denso-dependencia constante, el cual mide la intensidad de la competencia intra-específica; y el flujo neto de migrantes hacia él. Se incorpora un umbral M_i para considerar el efecto de una barrera biológica a la dispersión (*ver* por ejemplo, [83, 173, 164]). El modelo es el siguiente:

$$\dot{x}_i = (r_i x_i - a x_i^2) \left(x_i - \frac{M_i}{r_i} \right) + \sum_h \delta_{hi} x_h - \sum_h \delta'_{ih} x_i, \quad (3.1)$$

donde \dot{x}_i es la razón de cambio poblacional del parche i en el instante t . La tasa intrínseca de crecimiento en el parche i se denota con r_i ; se considera que esta tasa depende de la tasa máxima de crecimiento de la especie r_{max} y de una función S_i que indica que tan idóneo es el ambiente en el parche i , esto es $r_i = r_{max} S_i$. El factor de denso-dependencia es a . δ_{hi} es la tasa de inmigración de los individuos provenientes de los parches h hacia la celda i . δ'_{ih} es la tasa de migrantes hacia otros parches. Nótese que si $M_i = 0$ el modelo tiene la forma:

$$\dot{x}_i = (r_i x_i - a x_i^2) + \sum_h \delta_{hi} x_h - \sum_h \delta'_{ih} x_i, \quad (3.2)$$

Asumiendo que no hay dispersión y que no hay efecto Allee ($\delta_{ih} = 0$, $\delta'_{hi} = 0$ y $M_i = 0$ para todos los parches), la ecuación 3.1 es el modelo logístico

$$\dot{x}_i = r_i x_i - a x_i^2, \quad (3.3)$$

el cual tiene dos estados de equilibrio [71] dados por

$$x_i^* = 0 \quad \text{y} \quad x_i^* = \frac{r_i}{a} \quad (3.4)$$

En general, el cero es inestable y el equilibrio en $\frac{r_i}{a}$ es estable. Por otra parte cuando las tasas de migración son diferentes de cero, no es posible encontrar las soluciones del sistema 3.1 por lo que éste se resuelve numéricamente. La forma en como se considera a la tasa de crecimiento (una función del nicho) muestra que en el espacio ambiental las poblaciones que se encuentren más cercanas al centroide de la especie tendrán tasas mayores de crecimiento que las que se encuentran en la periferia. Más adelante se discute que factores pueden afectar esta relación perfecta entre distancia al centroide del nicho y la abundancia.

§3.2.1. Modelando la idoneidad del hábitat

Para modelar el efecto de las características ambientales en la tasa de crecimiento se utilizó una especie virtual. Se asumió que el nicho fundamental de la especie se mide en términos de la idoneidad del hábitat, la cual está definida por una función $f(x_1, x_2, x_3)$ donde x_1 , x_2 y x_3 son las variables ambientales que definen las condiciones donde la especie puede sobrevivir. Se considera que la tasa óptima de reproducción de la especie se obtiene en un cierto punto $(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ del nicho fundamental de modo tal que $f(x_1, x_2, x_3)$ es menor que $f(\mu_1, \mu_2, \mu_3)$ para todo $(x_1, x_2, x_3) \neq (\mu_1, \mu_2, \mu_3)$. En este trabajo se utilizó como función $f(x_1, x_2, x_3)$ a la función de densidad de probabilidad normal multivariada (ecuación 3.5) [15, 144], en donde la tasa óptima de reproducción $f(\bar{x}_1, \bar{x}_2, \bar{x}_3)$ se alcanza en las medias de las variables independientes x_1 , x_2 y x_3 .

La función de distribución multinormal es la siguiente:

$$f(x_1, x_2, x_3) = \frac{1}{(2\pi)^3 |\Sigma|} \exp\left(-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})\right) \quad (3.5)$$

donde Σ es la matriz de covarianzas de las variables x_1 , x_2 y x_3 ; \mathbf{x} es el vector de variables ambientales y $\boldsymbol{\mu}$ es un vector que contiene las medias de cada una de las variables ambientales. El término $(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})$ es el cuadrado de la distancia de Mahalanobis [44].

La función de la ecuación 3.5 es una función de densidad de probabilidad y por lo tanto esta se reescaló para que sus valores se encuentren en el intervalo cerrado $[0, 1]$. La función reescalada aplicada a las condiciones ambientales en el parche i representa la denominada idoneidad del hábitat S_i y está dada por la expresión:

$$S_i(x_1^i, x_2^i, x_3^i) = \frac{f(x_1^i, x_2^i, x_3^i)}{\max [f(x_1^i, x_2^i, x_3^i) \mid x_1^i, x_2^i, x_3^i \in \mathbb{R}]} \quad (3.6)$$

En este trabajo se utilizó como variables de nicho (ambientales) a la temperatura máxima del mes más calido (bio5), la temperatura mínima del mes más frío (bio6) y precipitación anual (bio12). Los valores de estas variables para el espacio geográfico fueron los de WorldClim [64] a una resolución de 5 minutos de arco. Las medias que se eligieron para cada una de las variables fueron: $\mu_1 = 28.91$; $\mu_2 = 9.97$; y $\mu_3 = 1935.10$. La matriz de covarianzas es

$$\Sigma = \begin{pmatrix} 569.25 & 665.93 & 8537.93 \\ 665.93 & 833.70 & 11767.89 \\ 8537.93 & 11767.89 & 222306.05 \end{pmatrix} \quad (3.7)$$

El nicho fundamental de la especie virtual se muestra en la figura 3.1a. Al aplicar la función de idoneidad (ecuación 3.6) a los parches i (180 en total) del espacio geográfico que se eligió se obtiene el mapa de idoneidad del hábitat (figura 3.1b).

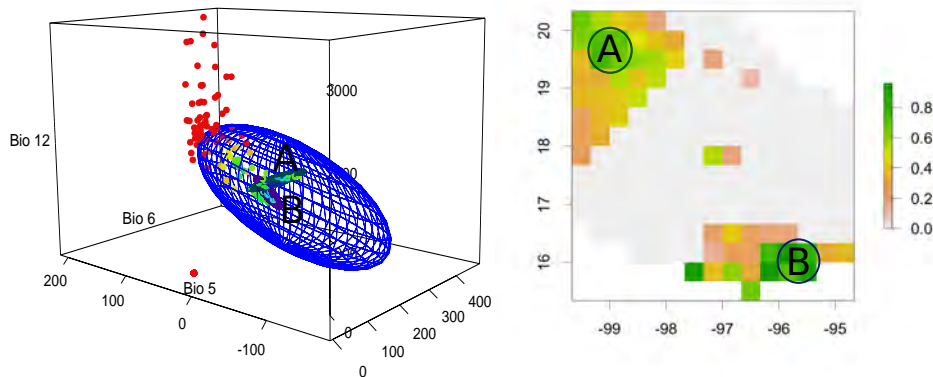


Figure 3.1: El nicho fundamental y el mapa de idoneidad de la especie virtual. Panel izquierdo: el elipsoide azul represental el nicho fundamental de la especie (N_e); los puntos rojos son las combinaciones ambientales del espacio geográfico (G_e) que caen fuera del nicho; y en naranja, verde y azul los puntos que caen dentro de éste. Panel izquierdo: Valores de idoneidad de los parches en el espacio geográfico; la intensidad del verde corresponde a sitios con mayor idoneidad (más cercanos al centroide). A y B denotan el mismo conjunto de parches tanto en el espacio de nicho como en el geográfico. Observando los dos paneles se puede notar que A y B están muy cercanos en el espacio de nicho, pero que a su vez están separados por regiones no idóneas (correspondientes a los puntos rojos del N_e) en G_e .

§3.2.2. Kernel de dispersión

Para modelar la dispersión entre celdas usamos un Kernel exponencial [109] dado por

$$\delta_{ih}(w_{ih}) = \begin{cases} k \exp[-w_{ih}^b] & w_{ih} \leq D_{max} \\ 0 & w_{ih} > D_{max} \end{cases} \quad (3.8a)$$

$$(3.8b)$$

donde D_{max} es la distancia máxima que un individuo puede viajar y $\delta_{ih}(w)$ es la tasa de movimiento de los individuos que viajan de la celda i a la celda h . k define la tasa máxima de movimiento. b es una constante que define la velocidad en que la tasa de movimiento disminuye.

Para medir la distancia entre parches se usó a la distancia Euclidiana, de donde se obtuvo una matriz de distancias \mathbf{W} (ecuación 3.9). Cada elemento w_{ih} de la matriz \mathbf{W} representa la distancia del parche i al parche h .

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & w_{13} & \cdot & \cdot & \cdot & w_{1n} \\ w_{21} & w_{22} & w_{32} & & & & \cdot \\ w_{31} & w_{32} & w_{33} & & & & \cdot \\ \cdot & & & \cdot & & & \cdot \\ \cdot & & & & \cdot & & \cdot \\ \cdot & & & & & \cdot & \cdot \\ w_{n1} & & & & & & w_{nn} \end{bmatrix} \quad (3.9)$$

Aplicando el Kernel de dispersión de las ecuaciones 3.8a y 3.8b a cada entrada w_{ih} de la matriz \mathbf{W} se obtiene la matriz de dispersión $\mathbf{\Delta}$ en la que cada entrada δ_{ih} representa la tasa de dispersión del parche i al parche h .

$$\mathbf{\Delta} = \begin{bmatrix} \delta_{11} & \delta_{12} & \delta_{13} & \cdot & \cdot & \cdot & \delta_{1n} \\ \delta_{21} & \delta_{22} & \delta_{32} & & & & \cdot \\ \delta_{31} & \delta_{32} & \delta_{33} & & & & \cdot \\ \cdot & & & \cdot & & & \cdot \\ \cdot & & & & \cdot & & \cdot \\ \cdot & & & & & \cdot & \cdot \\ \delta_{n1} & & & & & & \delta_{nn} \end{bmatrix} \quad (3.10)$$

§3.3. Método

En esta sección se describe las herramientas estadísticas para analizar cada uno de los casos de interés. En general las herramientas estadísticas que se utilizaron en todos los casos fueron la regresiones lineales y el coeficiente de determinación como medida de la relación entre abundancia poblacional y posición en el espacio de nicho.

§3.3.1. Método para determinar los efectos de la métrica en la relación y del tiempo que llevan las poblaciones establecidas en las localidades

Para estudiar los efectos del tiempo de establecimiento de una población en los parches y el de la métrica de distancia en la relación entre abundancia poblacional y posición en el espacio de nicho:

- Primero se determina la distancia Euclidiana y de Mahalanobis entre cada punto en el espacio de nicho y el centroide.
- Se hacen regresiones lineales entre las distancias (Euclidianas y Mahalanobis) y las abundancias poblacionales para diferentes tiempos:
 - 10, 50 y 2000 años de establecimiento.
- Posteriormente, se determina el coeficiente de determinación de las métricas para cada escenario de tiempo.

§3.3.2. Método para estudiar el efecto de la estructura espacial de las idoneidades en la relación

Para estudiar el efecto de la estructura espacial de las idoneidades en relación a los sitios donde comienza el proceso de dispersión:

- Primero se modela el crecimiento de la población sin (ecuación 3.2) y con efecto Allee (ecuación 3.3).
- Se inicializa una población en un parche determinado (100 individuos) y se determina la correlación entre abundancia y posición en el espacio de nicho (este proceso se hace para todos los parches en el espacio geográfico).

§3.4. Resultados

Se usó una retícula de 180 celdas con combinaciones ambientales de tres variables (bio5, bio6 y bio12). El nicho fundamental de la especie virtual se modeló como un elipsoide (Fig. 3.1) y la idoneidad del hábitat de cada celda se definió como una función de la distancia al centroide, como se describe en la sección §3.2.1. Los resultados de cada simulación dependen del tiempo de inicio de la dispersión, del tipo de métrica que se utilizó para medir la distancia al centroide y de la condición inicial en relación a la celda donde comenzó la dispersión; sin embargo como patrón general, se observó una mejor correlación de la métrica de Mahalanobis para todas las condiciones iniciales en los modelos con y sin efecto Allee. También se encontró una disminución del coeficiente de determinación r^2 para las simulaciones donde se modeló el efecto Allee. Para cada tiempo se reporta el valor de r^2 obtenida para cada métrica y para el parche donde se inicia la dispersión. Las simulaciones se realizaron con y sin efecto Allee.

§3.4.1. Efecto de la métrica y del tiempo de inicio de la dispersión.

Se encontró una mejor correlación entre abundancia y posición en el espacio de nicho para la distancia de Mahalanobis en todos los escenarios de tiempo; en este sentido incluso para los datos ambientales transformados con componentes principales, la métrica que obtuvo mejores correlaciones fue la de Mahalanobis, tanto para las abundancias modeladas sin efecto Allee (figura 3.2a) como para las modelas con efecto Allee (figura 3.2b). En la figura 3.2 se muestra la distribución de densidad del valor de r^2 para la distancia de Mahalanobis (A); la distancia Euclidiana (B); y para la distancia Euclidiana de los datos ambientales transformados utilizando componentes principales

(C). Los escenarios de tiempo que se muestran corresponde a 10 (rosa), 50 (verde) y 2000 (morado) años de inicio del proceso de dispersión. Queda claro que la elección de la métrica de distancia al centroide es un punto relevante en el estudio de la relación entre abundancia y posición en el espacio de nicho.

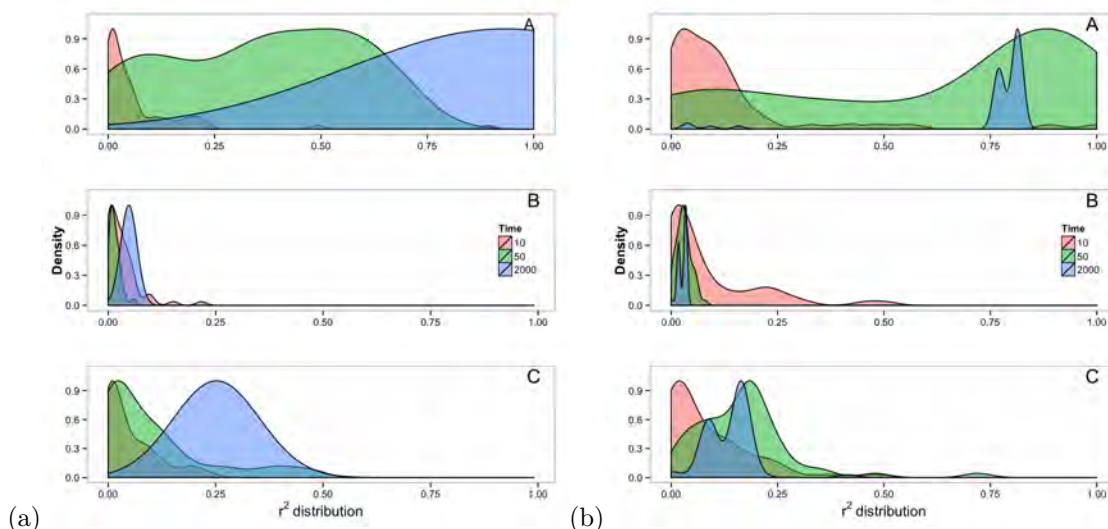


Figura 3.2: Distribución de las correlaciones entre abundancia y posición en el espacio de nicho para las abundancias modeladas sin (a) y con efecto Allee (b). En el primer panel (A), la distribución de r^2 para la distancia de Mahalanobis; en (B) la distancia Euclidiana; y en (C) la distancia Euclidiana para los datos transformados utilizando componentes principales. Los colores representan los escenarios de tiempo después del comienzo del proceso de dispersión: 10 (rojo), 50 (verde) y (2000) años .

El tiempo tiene un efecto importante en la correlación entre abundancia y posición en espacio de nicho; en términos generales, sin importar su posición en el espacio de nicho, al inicio las poblaciones que han sido recientemente fundadas presentan densidades más bajas que las de sus fundadoras. Como consecuencia de lo anterior, los valores de r^2 para estos tiempos son bajos. Por otra parte, conforme pasa el tiempo y las poblaciones de los parches invadidos van alcanzando su estado estacionario, el valor de r^2 va incrementando (animación 3.3).

§3.4.2. Efecto de la estructura espacial de las idoneidades

Al inicio del proceso de dispersión, casi cualquier condición inicial (valor de idoneidad) presenta un valor bajo de r^2 , tanto para las poblaciones modeladas sin efecto Allee (figura 3.4a), como las que lo presentan (figura 3.4b). Por otro lado, el valor de r^2 va incrementando conforme las poblaciones van alcanzando su estado estacionario.

Figura 3.3: Animación de relación entre abundancia y distancia al centroide del nicho para las poblaciones modeladas sin efecto Allee utilizando como métrica a la distancia de Mahalanobis para $t = 0$ hasta $t = 1001$. A la derecha se muestra el mapa dinámico de la expansión de la especie virtual

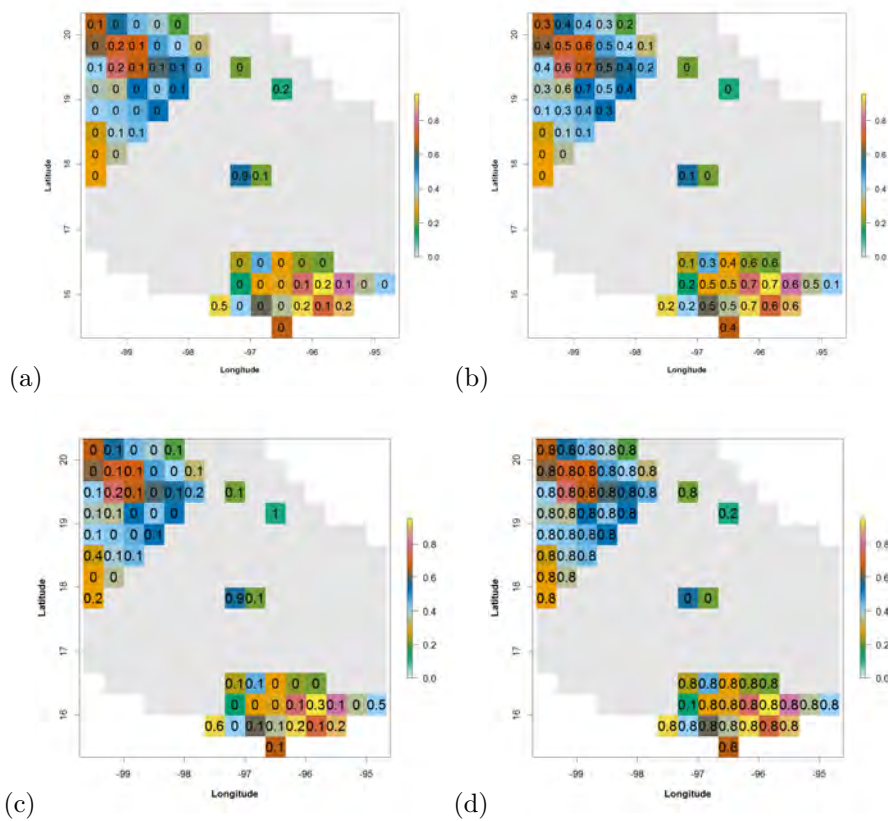


Figura 3.4: Coeficientes de determinación de la relación abundancia y posición en espacio de nicho como función del parche de inicio de la dispersión. (a) y (b) correlaciones para el modelo que no tiene efecto Allee en los dos escenarios de tiempo: después de 10 años de inicio de la dispersión (a) y para 2000 años (b). (c) y (d) correlaciones para el modelo con efecto Allee para 10 años (c) y 2000 años (d) de inicio del proceso de dispersión. En general se encontró una mejor correlación para el modelo sin efecto Allee. Los colores representan el valor de idoneidad del parche (ver leyenda).

En la figura 3.5 se muestra la dispersión de los valores de r^2 para diferentes escenarios de tiempo (10, 50 y 2000 años). Las poblaciones de los parches que presentan crecimiento sin efecto Allee muestran menor variabilidad en los valores de r^2 (figura 3.5a) que aquellas con efecto Allee (figura 3.5b).

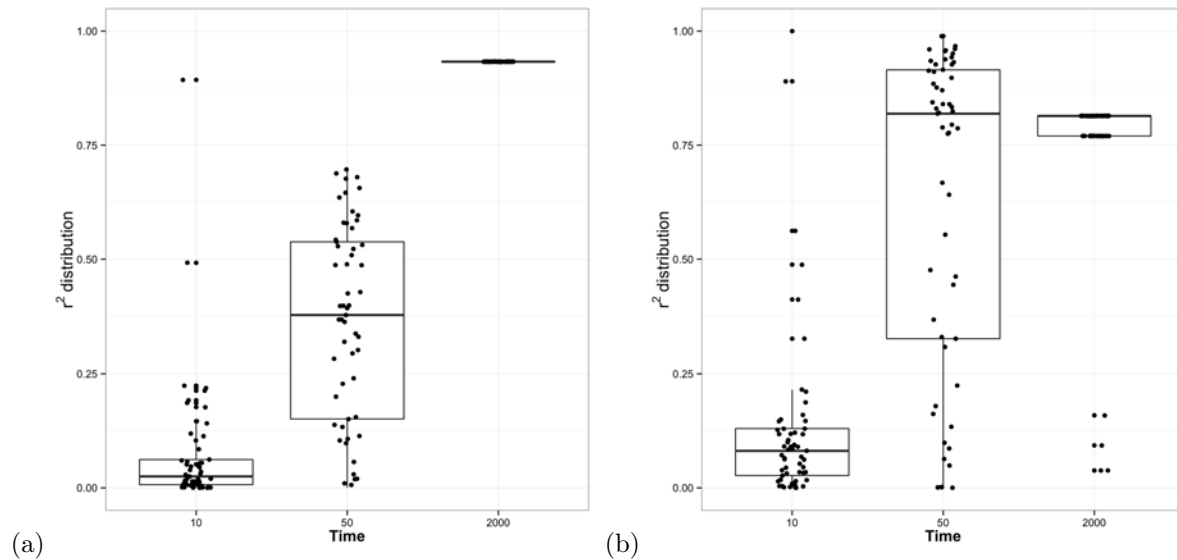


Figura 3.5: Diagrama de cajas y bigotes de la distribución de las r^2 para diferentes escenarios de tiempo (10, 50 y 2000 años). (a) Dispersión de las r^2 para las poblaciones modeladas sin efecto Allee. (b) Dispersión de las r^2 para el modelo con efecto Allee.

§3.4.3. Barrera biológica a la dispersión

El efecto Allee disminuyó de manera drástica la distribución geográfica de la especie virtual, de modo tal que se observó que solo una fracción de los sitios idóneos, fueron colonizados exitosamente por las poblaciones modeladas con este efecto (ver animación 3.6). De manera análoga, el efecto Allee también disminuyó considerablemente los valores de r^2 en todos los escenarios de tiempo estudiados. En general se encontró que tanto la distribución geográfica como los valores de las correlaciones disminuyen, entre más fuerte es este efecto (ver animación 3.6b).

(a)

(b)

Figura 3.6: Animación de relación entre abundancia y distancia al centroide del nicho para las poblaciones modeladas con efecto Allee utilizando como métrica a la distancia de Mahalanobis para $t = 0$ hasta $t = 1001$. (a) A la derecha se muestra el mapa dinámico de la expansión geográfica de la especie virtual para una especie que presenta un efecto Allee débil ($M = 2$); b) Mapa dinámico y correlaciones para una especie con efecto Allee fuerte ($M = 5$)

§3.5. Discusión

El propósito de este trabajo es mostrar la forma en que algunos factores ecológicos pueden modificar la relación esperada entre abundancia y posición en el espacio de nicho, de manera particular

se estudiaron los efectos de la dispersión, los de la métrica para medir la distancia al centroide del nicho y el de una barrera biológica a la dispersión. Se definió un modelo matemático relativamente sencillo para estudiar la dinámica espacio-temporal de una especie virtual, esto se realizó bajo el supuesto de que: 1) hay una relación directa entre la tasa intrínseca de crecimiento y la posición en el espacio de nicho (*ver* [96, 114]); 2) la especie crece de manera logística; 3) no hay evolución del nicho; 4) no hay competidores interespecíficos y 5) la dispersión se da a primeros vecinos (entre celdas contiguas). Si se asume que las poblaciones en estas localidades han llegado a su estado de equilibrio y que todas las localidades en el espacio geográfico están disponibles, es decir no hay barreras a la dispersión (a lo que se conoce como el mundo de Hutchinson, *ver* [144]), la posición en el espacio de nicho predice muy bien la abundancia poblacional, tal como se muestra en la figura 3.2a. En este sentido, el crecimiento logístico y la forma del nicho (ecuación 3.5), producen densidades de equilibrio cuya relación es inversa a la distancia de Mahalanobis al centroide del nicho. Lo anterior se deriva del supuesto de que la adecuación es más alta en los sitios que se encuentran en el centro de un nicho fundamental convexo, lo cual ha sido mencionado en la literatura casi desde los trabajos iniciales de Hutchinson [76]. Esta idea es exactamente lo que está detrás de los razonamientos de trabajos recientes como los de Yañez-Arenas et al. (2012)[185], Martínez-Meyer et al. (2013)[101] y Ureña et al. (2015) [169].

Por otra parte se espera que los datos se alejen de esta situación cuando:

- La métrica para medir la distancia al centroide del nicho no sea la adecuada.
- Todas las poblaciones en los parches no hayan alcanzado su estado de equilibrio.
- Existan barreras a la dispersión que provoquen que haya una proporción mayor de poblaciones que no han alcanzado su estado estacionario respecto a las que lo han alcanzado.
- Hay interacciones con otras especies.

§3.5.1. Métrica de distancia

A pesar de que solo es una cuestión metodológica, la elección de una métrica adecuada para medir la distancia de un punto en el espacio ambiental al centroide del nicho es muy importante, ya que ésta puede dar resultados contrastantes en la correlación entre abundancia y posición en el espacio de nicho como se muestra en la figura 3.2. En este trabajo se comparó el desempeño de la métrica de distancia Euclidiana (con datos ambientales no transformados y transformados con componentes principales) y la de Mahalanobis encontrándose que la segunda en general correlaciona mucho mejor con la abundancia poblacional para todos los escenarios de tiempo estudiados. Esto se debe a que se asumió que la distribución del nicho fundamental es normal y por ello tiene una forma de elipsoide en el espacio de n dimensiones con parámetros μ y Σ^{-1} . Dicho de otra manera, la idoneidad del hábitat decrece con la distancia de Mahalanobis entre el centroide del nicho μ y los ambientes presentes en un parche dado: $(\mathbf{x} - \mu)^T \Sigma^{-1} (\mathbf{x} - \mu)$. Si bien la forma en como se definió el nicho tiene mucho que ver en el desempeño de la métrica de Mahalanobis, también se han reportado en la literatura mejores correlaciones cuando se usa esta métrica que cuando se utiliza a la distancia Euclidiana como medida de la distancia al centroide; por ejemplo, Martínez-Meyer et al. [101] utilizando distancias Euclidianas encontró valores de correlación muy bajos, por otro lado, Yañez-Arenas et al. [184] encontraron $r^2 > 0.9$ al usar distancias de Mahalanobis, esto se puede relacionar al hecho de que en estas últimas se considera la covarianza que hay entre las variables de nicho, a diferencia de ello cuando se utiliza a la distancia Euclidiana se asume que no

hay correlación entre las variables, por lo tanto un punto situado en una combinación ambiental cuya distancia Euclidiana sea la misma que otro con una combinación ambiental diferente, tiene la misma idoneidad (para ejemplificar lo anterior, imagínese que una especie encuentra igualmente idóneo un ambiente cuya temperatura es de 1° y un pH de 7 que otro 7° de temperatura y 1 pH).

§3.5.1.1. Relación entre la salida cruda (raw output) de Maxent y la distancia al centroide del nicho

Una pregunta de interés derivada de la teoría que relaciona a las abundancias con el nicho es cómo es que las salidas directas de algoritmos de modelación de nichos, como Maxent por ejemplo, tienen una menor correlación con las abundancias que las medidas directas de la posición de un punto en el espacio de nicho, como ha sido observado en el contexto de relacionar abundancias con modelación de nicho [170, 165, 185]. Esto puede deberse a que la salida “cruda” de Maxent es proporcional a la probabilidad de que una combinación ambiental dada esté en la colección de ambientes donde una especie ha sido observada, lo cual es un número que puede no tener relación directa con la idoneidad del ambiente en particular [39, 107]. Mas aún, en Maxent el cálculo de esta probabilidad se realiza minimizando la entropía de una distribución relativa al “background” (los ambientes de referencia), por lo que su valor es dependiente del background particular que se haya escogido al realizar los cálculos. Por estas dos razones no resulta nada obvio que la salida directa de Maxent pueda tener relación con la idoneidad de una combinación climática. Por el contrario, la distancia al centroide del conjunto de puntos es una medida al menos ordinal de idoneidad. Los puntos más cercanos al centroide, independientemente de su valor numérico, están probablemente en la región de más alta idoneidad. Esta es una posible explicación de que no se haya encontrado una buena relación entre abundancia poblacional y las salidas de Maxent, como tales. En cambio algoritmos de envoltentes climáticos como BIOCLIM calculan directamente un valor proporcional a la distancia al centroide de un rectángulo que se usa como modelo del nicho [65]. No es de extrañar que un algoritmo así correlacione mucho mejor con la abundancia de una especie que un algoritmo como Maxent [165, 185].

§3.5.2. Estructura de las idoneidades en el espacio geográfico.

La dualidad de Hutchinson [30] establece que hay una correspondencia entre el espacio ambiental y el geográfico, por tanto la área de distribución está afectada por la combinación de ambos componentes. En este trabajo se muestra como es que la capacidad de dispersión además de determinar la forma de la distribución, también afecta la abundancia poblacional de formas potencialmente complejas. Esto se debe en gran parte al hecho de que a pesar de que haya una correspondencia entre las localidades geográficas y los puntos en el espacio de nicho [126], la estructura topológica de los dos espacios es radicalmente diferente, en el sentido de que puntos muy cercanos en un espacio pueden estar muy alejados en el otro (ver figura 3.1). La condición inicial para el comienzo de la dispersión, las barreras o el grado de avance de la onda de expansión geográfica [158], puede causar que áreas con idóneas altas tengan abundancias bajas respecto a los sitios con idoneidades bajas e incluso menores que sitios sumidero [130].

§3.5.3. Barrera biológica a la dispersión (efecto Allee)

El efecto Allee considera el hecho de que para que una población se pueda establecer de manera exitosa en un sitio determinado, es necesario superar un umbral mínimo de densidad por debajo del

cual no hay crecimiento. En función de la estructura de las idoneidades en el espacio geográfico, el efecto Allee puede provocar que sitios más cercanos al centroide tengan menores abundancias poblacionales que aquellos más alejados, debido a que en la geografía los primeros pueden encontrarse más alejados que los otros con idoneidades bajas. La intensidad del efecto Allee también tiene un impacto importante en la forma del área de distribución [82], en general se espera que las especies que tengan un efecto fuerte presenten área de distribución más restringidas que aquellos donde es menor (ver animación 3.6).

§3.6. Artículo: A Theoretical Explanation for the Relationship Between Population Abundance and Niche Structure, and When to Expect Exceptions.

The American Naturalist

A Theoretical Explanation for the Relationship Between Population Abundance and Niche Structure, and When to Expect Exceptions

--Manuscript Draft--

Manuscript Number:	58033
Full Title:	A Theoretical Explanation for the Relationship Between Population Abundance and Niche Structure, and When to Expect Exceptions
Short Title:	Population Density and Niche Structure
Article Type:	Major Article
Manuscript Classifications:	Ecology: population; Ecology: spatial; Ecology: theoretical
Additional Information:	
Question	Response
Please provide the word count for the main text (excluding the abstract, the literature cited, tables, or figure legends) [original submission]	6570

A Theoretical Explanation for the Relationship Between Population Abundance and Niche Structure, and When to Expect Exceptions

Manuscript elements: Figure 1, figure 2, figure 3, figure 4, figure 5, online appendices A and B (including figure A1, figure B1 and figure B2). Figure 2 is to print in color.

Keywords: Population abundance, Grinnellian niches, spatial structure, position in niche space.

Manuscript type: Article.

Prepared using the suggested L^AT_EX template for *Am. Nat.*

Abstract

Recently published evidence has indicated a negative correlation between population abundance and its distance from a suitably defined "niche centroid". This is an empirical observation without substantial theoretical ground. We provide a theoretical underpinning for the empirical relationship between population density and position in niche space, and use this framework to understand the circumstances under which the relationship will fail. We propose a population model for the area of distribution, as a system of ordinary differential equations coupled with a dispersal kernel. We present an analytical approximation to the solution of the system as well as an R package to solve the full model numerically. We use this tool to analyze various scenarios and assumptions. General and realistic demographic assumptions imply a good correlation between position in niche space and population abundance. Factors that modify this correlation are: transitory states, a heterogeneous spatial structure of suitability, and Allee effects. We also explain why the raw output of the niche modeling algorithm MaxEnt is not a good predictor of population density. Our results elucidate the empirical results for spatial patterns of population size in niche terms, and provide a theoretical basis for a structured theory of the niche.

1 **Introduction**

2 Population density was for some time thought to be related to the position of populations in
3 a geographic range, with larger populations close to the center of the range (Rapoport, 1975;
4 Udvardy, 1969). This so called “abundant centre” hypothesis has been criticized on empirical
5 grounds (Sagarin and Gaines, 2002; Sagarin et al., 2006; Tuya et al., 2008). Another attempt to
6 establish relationships between population density and distributional range used the outputs of
7 ecological niche models (ENMs) as predictors of density (Nielsen et al., 2005; Thuiller et al., 2014;
8 Tôrres et al., 2012; Van der Wal et al., 2009). However, these authors have found inconsistent or
9 weak relationships between density and the output of ENMs.

10 Finally, on the basis of ideas suggested by Maguire (1973) and Brown (1984), and on growing
11 empirical evidence, it has been proposed that abundance in a site has been proposed to be related
12 to the position of the environment of the site in niche space, with the highest population densities
13 in those sites with environments closest to the center of the fundamental niche (Martínez-Meyer
14 et al., 2013; Ureña-Aranda et al., 2015; Yañez-Arenas et al., 2012). The notion of a relationship
15 between population variables and position in niche space, as estimated using ENM methods, has
16 also been applied to the genetic structure of populations (Lira-Noriega and Manthey, 2014), with
17 positive results.

18 Although there is indeed empirical evidence of a relationship between position in niche space
19 and population density (Martínez-Meyer et al., 2013; Yañez-Arenas et al., 2012), a very recent re-
20 view (Dallas et al., 2017) reported consistent near zero correlation values between population
21 density and distance to a niche centroid (calculated using convex-hulls) for several thousand
22 species. The careful explanation that Dallas et al. (2017) give for the lack of a consistent rela-
23 tionship include several possibilities that can be actually studied more appropriately by the use
24 of virtual species and an explicit model of the relationship. It is difficult to understand any
25 body of empirical results without a theoretical formulation of the problem. To provide one such
26 framework is the objective of this paper.

27 There are three major suppositions underlying the idea that position in niche space and
28 population density may be related: (1) there is a function relating the position in niche space to
29 the value of the fitness of a population experiencing such niche conditions, (2) the multivariate
30 fundamental niche (\mathbf{N}_F) can be defined in such a way that it is meaningful to talk about its center
31 and periphery (for instance, that fundamental niches are convex shapes), and (3) regions closest
32 to the center have the highest values of the fitness function. That multivariate niches may have
33 such a structure was hinted by Hutchinson (1957) and elaborated in greater detail by Maguire
34 (1973). If true, this is an insight with many important implications.

35 However, although the idea of a relationship between niche position and population prop-
36 erties has clear intuitive appeal, a comprehensive treatment of the problem requires rigorously
37 defining the variables used (interacting, non-interacting), the niche being modeled (fundamen-
38 tal, existing, realized), and what shape it is assumed to have (for instance, a convex hull, or
39 an ellipsoid). When comparing with the outputs of different ENMs (sometimes referred to in
40 the literature as Species Distribution Models, or SDMs), one needs to understand exactly what
41 is the mathematical meaning of these outputs, often referred to vaguely as "probabilities" or
42 "suitability indexes". In the present work we show that the problem can be analyzed in great de-
43 tail, and using realistic data, as long as the niche definition is restricted to using non-interactive
44 "scenopoetic" variables (Hutchinson, 1978; Peterson et al., 2011; Soberón, 2010), that define ellip-
45 soidal fundamental niches, and if a reasonably realistic model of the metapopulation structure is
46 provided.

47 Two caveats about ENM methods are also needed. First, ENMs calculate combinations of en-
48 vironmental variables that are likely related to the existing niche, rather than to the fundamental
49 niche (Peterson et al., 2011). Second, the output of presence-only ENM algorithms are not fitness
50 measures directly. In presence-only methods these measures are not even "probability of occur-
51 rence" measures (Pearce and Boyce, 2006). They are measures of how similar an environment is
52 to environments where the species has been observed.

53 In light of the above, the aim of our paper is to examine, using theoretical tools, the hypothesis

54 that subpopulation abundance in a metapopulation is related to the positions of subpopulations
55 in the niche space of scenopoetic variables, and to search for complicating factors and exceptions.

56 To study the problem we propose a process-based model of a metapopulation in a discrete
57 grid of cells (Nenzén et al., 2012; Osorio-Olvera et al., 2016; Schurr et al., 2012; Smolik et al., 2010).
58 We assume an explicit ellipsoidal structure for the fundamental niche. Such a model will generate
59 population abundances over a landscape with known environmental features and explicit inter-
60 cell movements. This establishes a direct relationship between the niche position of a site, and
61 the abundance of the species in the site. The major complicating factor is movement, which is
62 modeled explicitly, and thus relaxes the steady-state assumption of much species-distribution's
63 modeling (Svenning and Skov, 2004).

64 Process-based models include both structure of niche space and spatially-explicit dynamics in
65 geographic space. We follow previous work by Keith et al. (2008); Kirkpatrick and Barton (1997);
66 Vandermeer (1972) and others, who proposed different approaches to the problem. We use con-
67 tinuous logistic population growth (with and without Allee effects) in a spatially discretized land-
68 scape connected by dispersal. This approach has been used to address a number of problems,
69 for instance, by Pulliam (2000), who performed numerical simulations of an spatially-explicit
70 model with two species, to explore the consequences of competition on species distributions, but
71 without the emphasis on abundance, or Holt et al. (1997), who focused on the spatial distribution
72 of abundance but mostly ignored niche considerations; or Hanski (1999), who concentrated on
73 the extinction-colonization dynamics. A spatially-explicit population dynamics model provides
74 a natural first-principles model for species distributions in terms of numbers (Carter and Prince,
75 1981), because the distribution of a species is simply the collection of localities in which the prob-
76 ability of detecting a viable population is greater than some threshold. Therefore, fundamentally,
77 the geographic distribution of a species is a manifestation of its population dynamics in space.

78 Although simplified in the sense that we assume a deterministic single-species ecology with
79 no age-structure and no evolution, our model is, on the one hand, general and robust, and
80 on the other, quite challenging numerically and analytically because the population growth is

81 modeled by non-linear equations, over a grid of potentially many thousands of cells, coupled by
82 dispersal. We present the results of: (i) analytic approximations, and a general theorem on the
83 existence of a solution, as well as (ii) numerical results of the relationship between the position
84 of a population in niche space and its size, under a number of scenarios. Although a relationship
85 between population density and distance to a niche centroid is built into the structure of our
86 model, several complications can upset the basic relationship. We wish to explore four major
87 questions:

- 88 1. The importance of choosing the appropriate metric to express distances in environmental
89 space.
- 90 2. The effect of the spatial structure of suitability in relation to initial conditions and Allee
91 effects.
- 92 3. The effect of transient time.
- 93 4. Why direct outputs of ENMs often fail to correlate with abundance

94 **Methods**

95 We assume a region divided by a regular grid with $i = 1, 2, \dots, 638$ cells or patches. The extent
96 of the grid is Mesoamerica, from the eastern side of the Isthmus of Tehuantepec to the border of
97 Panama with Colombia; and the resolution of the grid's cells is half a degree. This region was
98 selected only for purposes of illustration. Each cell is characterized by environmental features.
99 We use the first three principal components on the 19 bioclimatic variables of the WorldClim
100 database (Hijmans et al., 2005). We assume that climatic variables affect the growth rate of
101 the population located in a particular cell. Populations in the cells grow in a sigmoid manner,
102 with and without an Allee effect (Keitt et al., 2001). The intrinsic growth rate depends on the
103 environments of the different cells, a constant (among cells) density-dependent parameter, and
104 the rate of dispersal to other cells, as follows:

$$\dot{x}_i = (r_i x_i - a x_i^2) (x_i - M) + \sum_h \delta_{hi} x_h - \sum_h \delta_{ih} x_i, \quad (1)$$

105 where \dot{x} is the rate of change of a population, growing in the cell i at time t . r_i is the intrinsic
 106 growth rate in cell i , a is a constant density-dependent parameter, M is the threshold of the Allee
 107 effect, and $\delta_{k,l}$ is the migration rate from cell k to cell l . Without migration and Allee effects,
 108 equations (1) are simple logistics with steady-state values of $\hat{x}_i = \frac{r_i}{a}$. Notice that although the
 109 density-dependent parameter a is constant over cells, the steady-state population size changes as
 110 a function of the environment, via the intrinsic growth-rate.

111 We define the growth rates in each cell as direct functions of the Mahalanobis distance be-
 112 tween the centroid of a hypothetical fundamental niche (\mathbf{N}_F) and the environmental values in
 113 the cell. If μ is the centroid of \mathbf{N}_F , Σ its shape matrix, \vec{e}_i is the vector of environmental variables
 114 in cell i , and r_m is the maximum possible value of the intrinsic growth rate, then

$$r_i(\mu, \vec{e}_i) = r_m \exp\left(-\frac{1}{2}(\mu - \vec{e}_i)^\top \Sigma (\mu - \vec{e}_i)\right). \quad (2)$$

115 This is, we assume an ellipsoidal structure for a fundamental niche. Other structures have
 116 been proposed, for instance convex hulls (Dallas et al., 2017), but we prefer the ellipsoid model
 117 because every data point is used to define the ellipsoid (Farber and Kadmon, 2003), whereas a
 118 convex hull uses only points in the periphery of a data set. This means that the position of the
 119 centroid in a convex hull may change rather substantially due to a few outliers. If the distance
 120 between the actual environment in a cell and the centroid of the niche is smaller than a given
 121 threshold, the local intrinsic growth rate would be positive. The above is a simple model of
 122 population growth expressed as a function of niche distance, and by design creates the effect of
 123 correlating suitability with population density.

124 The maximum value of the intrinsic growth rate corresponds to the value of a species com-
 125 plex, *Cryptotis nigrescens*, which has a distributional range in Mesoamerica (Ceballos and Oliva,
 126 2005). A value for its growth rate ($r_m = 0.115 y^{-1}$) was obtained from the database of life-history

127 parameters COMADRE (Salguero-Gómez et al., 2016). This value is modulated by the Maha-
 128 lanobis distance, in such a way that when the distance of environmental values to the centroid
 129 of the niche is zero, the intrinsic growth rate is $0.115 y^{-1}$. The value of the density-dependence
 130 parameter was set to $a = .001$, to yield a population densities (in units of individuals per km^2 of
 131 about 100 in the most suitable cells). This value was selected to produce reasonable simulation
 132 times.

133 To define the fundamental niche, we used the first three principal components performed on
 134 the 19 variables of WorldClim, at $10'$ resolution, for the entire world. These explain 85% of the
 135 total variance. The centroid of our species was $\mu = (5.258301, 3.135807, 1.820532)$ and the shape
 136 matrix is

$$\Sigma = \begin{pmatrix} 2.7366 & -1.43800 & -1.50494 \\ -1.4380 & 1.05709 & 0.75770 \\ -1.5049 & 0.75770 & 2.02778 \end{pmatrix}$$

137 The above correspond to climates commonly found in Mesoamerica, and defines a minimum-
 138 volume ellipsoid that contains all of the GBIF occurrence points for *C. nigrescens*. This ellipsoid
 139 is our hypothetical model for the \mathbf{N}_F of the species.

140 Dispersal kernel

141 To model the dispersal between cells we used an exponential kernel given by

$$\delta_{ih}(w_{ih}) = \begin{cases} k \exp[-w_{ih}^b] & w \leq D_{max} \\ 0 & w > D_{max} \end{cases} \quad (3a)$$

142 where D_{max} is the maximum distance that an individual can travel, set to 2 cells per year, w
 143 is the distance between patches, $\delta_{i,h}(w)$ is the migration rate of individuals going from cell i
 144 to cell h , k is the maximum capacity of migration, that we set to 0.00575/day, b is a positive
 145 constant such that the greater it is the less migration will occur to distant patches. The distance

146 between cells is simply the Euclidean distance using the coordinates of cells i and h . Note that
147 barriers that prevent the migration from patch i to patch h can be simulated by taking $\delta_{i,h} = 0$.
148 A dispersal matrix that can be partitioned in subsets inaccessible from other subsets is called
149 reducible. Whether subsets of cells can be reached or not from another subset depends on the
150 dispersal capabilities of the species in question, and on whether there are Allee effects. Using the
151 above dispersal kernel, and ignoring Allee effects, every element of the grid can be reached from
152 every other element, therefore the dispersal matrix is irreducible.

153 **Distance metrics**

154 The hypothesis we wish to explore is that there is a correlation between population abundance
155 and the "environmental distance" to the center of the niche. Each point in niche space is described
156 by m variables. This is a multivariate space in which variables tend to be highly correlated and
157 of very different units (for instance, mean temperature in °C and precipitation in mm). Therefore
158 it is important to use a distance metric that will not be affected by those complications. In
159 the past, authors have used simple Euclidean distances (Martínez-Meyer et al., 2013); in this
160 work we will also try a Mahalanobis distance (Legendre and Legendre, 1998). Other authors
161 have directly used the outputs of a popular ENM algorithm, Maxent (Phillips et al., 2006) as
162 predictors of population size (Tórres et al., 2012; Van der Wal et al., 2009; Weber and Grelle,
163 2012). We compared the performance of these three metrics by obtaining correlations between
164 the value of the metric and population abundance at different time scenarios: 10, 500 and 2500
165 years after invasion of an initial cell.

166 **Euclidean distance**

167 To compute the Euclidean distance to the centroid, we first extracted the values of the occurrence
168 points for each niche layer (PC1, PC2 and PC3) for every occurrence point; this is a matrix \mathbf{O}
169 with entries for the niche values of the occurrence points. The niche centroid was defined as the
170 vector $\vec{\mu} = (\mu_1, \dots, \mu_m)$, where μ_j is the mean value of the column j of \mathbf{O} . The Euclidean distance

171 to the centroid for each cell i was calculated as follows:

$$D_E(\vec{e}_i) = \sqrt{(\vec{e}_i - \vec{\mu}) \cdot (\vec{e}_i - \vec{\mu})} \quad (4)$$

172 Mahalanobis distance

173 As in the Euclidean distance, we computed the niche centroid of the occurrence points. Second,
174 we estimated the shape matrix Σ as the inverse of the covariance matrix $(\text{cov}(\mathbf{O}))^{-1}$ of the oc-
175 currence points in the niche space. For each environmental combination in cell i the squared
176 Mahalanobis distance is:

$$D_M(\vec{e}_i) = \sqrt{(\vec{\mu} - \vec{e}_i)^\top \Sigma (\vec{\mu} - \vec{e}_i)} \quad (5)$$

177 Niche modeling algorithm outputs

178 We use a single ENM method, MaxEnt (Phillips et al., 2006), which has become the *de facto*
179 standard ENM algorithm. The outputs of MaxEnt have been used to correlate with abundance
180 data. Therefore, three MaxEnt models were used for comparison purposes. We only used the
181 Linear, Quadratic, and Product features of MaxEnt. This avoids over-fitted models with dozens
182 of parameters. We disabled clamping and left the other parameters with their default values.
183 Because MaxEnt is sensitive to different backgrounds, we fitted models for three different back-
184 grounds (*see* figure A1 in Supplementary information); the first was located in Mesoamerica; the
185 second one corresponds to a region between Eastern Mexico and Northern Colombia; finally,
186 the third extends from the United States to Brazil. (*see* Merow et al. (2013) for a discussion on
187 parameterizing a Maxent analysis).

188 Allee effects

189 We considered two cases, no Allee effect, and a strong Allee effect (20 individuals are required
190 to start a new population, see below). When considering Allee effects, initial conditions can

191 be critical to results. We used as initial conditions all of the single-cell possibilities, for two
192 situations: We initiated a population in every possible cell, and report the abundances at three
193 different times: 10, 500 and 2,500 years. After 2,500 years, changes in abundance are very minor,
194 and a steady state has been achieved. We know by the theoretical result proved in the appendix
195 that without an Allee effect there is a single global equilibrium point for the system (and all our
196 simulations reach one).

197 There are dozens of possible ways to model an Allee effect (Boukal and Berec, 2002); we
198 selected, as a preliminary choice, a cubic equation used by Keitt et al. (2001). Notably, this choice
199 still requires further theoretical analysis. This is model (1).

200 **Analytic approximation**

201 Because an analytic solution for equations (1) is generally impossible, we developed a simple
202 analytic approximation in order to enable preliminary explorations into the behavior of equations
203 (1). First, in the appendix it is demonstrated that equations (1) have a unique positive equilibrium
204 point; however, analytical expressions for this point are intractable. To approximate the solution
205 of the system in the steady state state, we first consider a global emigration rate for cell i . This is
206 simply the sum

$$\sum_h \delta_{i,h} = \delta_i \quad (6)$$

207 Now we make two assumptions. First, we assume that there is an average, constant total
208 number of migrants approximated by the sum of the migration-independent equilibrium (in
209 every cell) with no Allee effects:

$$R = \frac{\sum_h \frac{r_i}{a}}{n} \quad (7)$$

210 Therefore, noticing that the dispersal rates are symmetric, the term describing immigration
211 reduces to:

$$R \sum_h \delta_{i,h} = R\delta_i \quad (8)$$

212 Now the model becomes:

$$\frac{dx_i}{dt} = (r_i - \delta_i) x_i - ax_i^2 + \delta_i R \quad (9)$$

213 which has an analytic solution for the steady state (denoted with an asterisk) in every cell:

$$\begin{bmatrix} x_1^* \\ x_2^* \\ \vdots \\ x_n^* \end{bmatrix} = \begin{bmatrix} \frac{r_1 - \delta_1 + \sqrt{(r_1 - \delta_1)^2 + 4\delta_1 a R}}{2a} \\ \frac{r_2 - \delta_2 + \sqrt{(r_2 - \delta_2)^2 + 4\delta_2 a R}}{2a} \\ \vdots \\ \frac{r_n - \delta_n + \sqrt{(r_n - \delta_n)^2 + 4\delta_n a R}}{2a} \end{bmatrix} \quad (10)$$

214 Equations (10) are expressions of the steady-state equilibrium point for every cell in the grid
 215 under the assumption that there was a suitable initial population (this is, we ignore initial con-
 216 ditions and Allee effects).

217 Numerical solutions

218 Our analytic approximation only estimates the steady-state and ignores the transitory period
 219 and Allee effects; to obtain solutions over time intervals we resort to numerical simulations
 220 (developed as an R package deposited in the GitHub archive). This package permits the user to
 221 create and parametrize the set of ordinary differential equations of the model (1), and then to
 222 solve them numerically. A compiled C++ routine for ordinary differential equations (from the
 223 *odeint* library (Ahnert and Mulansky, 2011)) was used. The computation time grows as a power
 224 of the number of cells; thus, despite using compiled and highly optimized numerical routines, if
 225 the grid contains more than a few hundred cells, computation time becomes unmanageable.

226 We are now solving to find trajectories over a time interval. Thus, the initial conditions
 227 (the actual spot where the spread of a species begins) may determine the steady-state spatial
 228 distribution. If there are barriers to dispersal, or regions that are unsuitable for the species,

229 different initial conditions for the spread may create very contrasting steady-state distributions
230 (Soberón et al., 2017). Moreover, in simulating trajectories from initial conditions, the transitory
231 phase needs to be determined. To illustrate different aspects of the time dynamics of the system,
232 we take three time slices: 10 years, 500 years and 2,500 years.

233 Our main results are presented as correlations between observed population density (at each
234 time slice), and the predictors (suitability, Mahalanobis distance to centroid, Euclidean distance
235 to centroid, and the three MaxEnt raw outputs). The relationships between abundance and the
236 predictors tends to be non-linear; thus, we use the Spearman rank correlation, which is not
237 affected by monotonous nonlinearity. In figure 1, we take two starting points of the simulation to
238 illustrate how the steady-state abundance is related to various predictors, for every point in the
239 grid.

240 **Results**

241 **Analytic results**

242 Without Allee effects, as long as the environments in cell i are inside the fundamental niche, and
243 thus $r_i > 0$, a cell can be invaded. With no migration, the steady state per cell would be the
244 steady-state abundance would be simply the steady-state of a logistic equation.

$$\hat{x}_i = \frac{r_i}{a}. \quad (11)$$

245 Because of the way the model is built, with the intrinsic growth rate as a function of the
246 environmental parameters of each cell $[r(\mu, \vec{e}_i)]$, at the steady-state and ignoring movements
247 there will be a direct relationship between suitability and abundance. There will therefore also
248 be an inverse relationship between the Mahalanobis distance of the environments in cell i and
249 the centroid of N_F , as shown in equation (2). However, the analytical approximation reveals
250 an immediate complication relative to the non-movements case. As shown by equations (10),
251 the steady-state abundance is also a function of the amount of dispersing individuals, in ways

252 unrelated to local suitability (unless the parameter determining the immigration rate d_i would be
253 directly correlated with $r(\mu, \vec{e}_i)$). Therefore, we expect that in general, immigration may interfere
254 with the direct correlation between abundance and suitability. In figure 2, we show how the
255 analytic prediction tends to over-predict the results of the simulation, as well as the simple set of
256 logistic models without movements, thereby under-predicting abundance.

257 Thus, we have a null result: without the complicating effects of dispersal, barriers, transi-
258 tory states, and Allee effects, equilibrium population density should be inversely related to the
259 distance between the local environments and the centroid of the niche, at least in an ordinal
260 sense.

261 Numerical results

262 To display the effect of the time dynamics we show box plots of the correlation between abun-
263 dance and the predictor variables at three different points in time. In figure 3 we show boxplots
264 of the Spearman rank correlation (ρ), for times 10 (A), 500 (B) and 2500 (C) years, as well as be-
265 tween abundance and (i) suitability, (ii) Mahalanobis distance to the centroid and (iii) Euclidean
266 distance to the centroid; additionally, Pearson correlation for the three MaxEnt (raw data) pre-
267 dictions were obtained using different backgrounds (*see* methods). We show the results for pop-
268 ulation sizes reached after initiating growth in each one of the 638 cells, against the suitability
269 value, without Allee effects.

270 The results clearly show that during the transitory period correlations between population
271 density and suitability of habitat can be bad or mediocre. However, towards the steady state
272 (row C), the ρ value becomes very high when correlating abundance and suitability (which is
273 to be expected due to the way the model was constructed). This value is also as high for a
274 Mahalanobis distance to the niche centroid (because we are using a Spearman rank correlation)
275 and substantially lower for the Euclidean distance. The linear Pearson correlation is very poor
276 for any of the MaxEnt raw outputs.

277 The inclusion of Allee effects greatly complicates this. In figure 4 we show boxplots of ρ

278 values between population density at the steady-state (2,500 years), and suitability, Mahalanobis
279 distance and Euclidean distance for every patch in the grid, for three times along the trajectory
280 (10, 500 and 2,500 years).The inclusion of an Allee effect greatly affects the result, primarily
281 reducing the median value of the rank correlation ρ , and also increasing its variance, because
282 there are now many localities from which a starting population cannot reach suitable regions.
283 The strong Allee effect makes this effect very noticeable.

284 **Discussion**

285 Following the classic work of Maguire (1973), ecologists have struggled to define niche structure
286 in terms of fitness, which (all things being equal), would imply that niche structure would corre-
287 late with population density. As long as the fitness function defining the niche has some regular
288 structure, for instance an ellipsoid, with a centroid located in the environmental combinations
289 of maximum fitness, as Maguire (1973) suggested, an appropriate measure of distance to the
290 centroid should correlate with abundance.

291 However, niche space and geographic space are tied in inseparable ways, forming what has
292 been called a “Hutchinson’s Duality” (Colwell and Rangel, 2009), and the ecological and evolu-
293 tionary dynamics of a species depends on the combination of processes occurring in both spaces
294 (Colwell and Rangel, 2009; Soberón et al., 2017).

295 Therefore, a fuller understanding of the problem of the relationship between niche and pop-
296 ulation density requires to state the metapopulation structure, which we model by assuming a
297 grid in which, without complications due to Allee effects, every cell is reachable from every other
298 cell (an irreducible dispersal matrix). Ellipsoidal niche structure and comprehensive dispersal ca-
299 pacities, in essence constitute a null model. However, we report on a number of complicating
300 factors.

301 **Effects due to choice of distance metric**

302 Since we defined fitness as a function of a Mahalanobis distance to a niche centroid, it is to be
303 expected that using another distance will yield poorer correlations, which is what we observe.
304 However, the lesson from our results is more than an artifact of an ellipsoidal choice for the
305 fundamental niche. Environmental space is intrinsically non Euclidean, and the variables have
306 non-zero correlations and also different units. Therefore using an Euclidean distance probably
307 is not appropriate in general, as we saw in figure 4. Even if the space has been transformed
308 by principal components, as it is the case in the study of Dallas et al. (2017), the specific set of
309 occurrence points may be not well described by the Princial Componnets transformation. The
310 Mahalanobis distance automatically corrects for different units and for non-orthogonality of axes
311 for the occurrence points, and for these reasons we would expect that in general it would be a
312 better metric to calculate distance to a niche centroid.

313 **Effects due to transient phases**

314 Unsurprisingly, when the system is far from its steady-state, the correlation between abundance
315 and any of its predictors is not very good. Although this is an intuitive result that has been
316 reported in the context of distribution modeling (Svenning and Skov, 2004), a more general
317 conclusion can be extracted: systems suspected to remain far from a steady-state (for instance,
318 an invasive species at an early phase of its invasion) should not be expected to have a good
319 correlation between **presence** and measures of niche suitability is not to be expected (Vaclavik
320 and Meentemeyer, 2012); thus, a good correlation with **abundance** is much less likely.

321 **Movements and Allee effects**

322 The main complication that we discuss is the effect, shown in equations (10), and confirmed by
323 the simulations, that dispersal for a population with a spatial structure (i.e., a set of metapopula-
324 tions), can substantially interfere with the expected theoretical correlation between local environ-

325 mental suitability and population density. In our model, metapopulation dynamics can interfere
326 with the relationship between abundance and suitability even when such relationship is there by
327 design! Therefore, the complicating factor introduced by movements would be worse when the
328 predictor of abundance is some distance to the centroid of a hypothetical niche function, which
329 we assume correlates with abundance only in an ordinal manner.

330 Equations (10) suggest that the confusing effect of immigration is mediated by the rates of
331 arrival to patches. If there is no covariance between suitability and patch attractiveness, as deter-
332 mined by δ_i , or the covariance is negative, then the expected null relationship between abundance
333 and suitability (distance to the centroid) may be seriously upset (Soberón, 1986).

334 However, another subtler effect of dispersal is expressed via the Allee effect. Having an Allee
335 effect implies that there may be barriers due to regions of relatively low suitability. This is an in-
336 teraction between the process of actually moving, and the process of establishing subpopulations
337 that may act as stepping-stones. The Allee effects amplify the barrier effects of low suitability
338 regions to act as barriers, and make information about the initial conditions crucial to determine
339 the time trajectory of the spread of an invasion, or the expansion of a population. This effect has
340 been studied theoretically by Kirkpatrick and Barton (1997) and Holt (2009a), but our method
341 and software permits very realistic simulations. The main lesson we extract, in the context of
342 the problem under study, is that if a geographic region contains regions of low suitability, even
343 if movements are allowed among all patches (which in mathematical terms means that an ad-
344 jacency matrix is irreducible; there are paths from every patch to every other patch), then low
345 suitability and Allee effects may create empty but perfectly suitable subregions, depending on the
346 initial conditions. Clearly, such suitable but empty localities would upset the correlation between
347 abundance and distance to the centroid. This is one of the mechanisms that Dallas *et al.* (Dallas
348 *et al.*, 2017) use to explain the low correlations they found in their review of the relationship
349 between population density and distance to niche centroid.

350 MaxEnt

351 As stated previously, empirical studies have sought a correlation between the mean abundance
352 and the direct output of presence-only ENM algorithms such as MaxEnt (Nielsen et al., 2005;
353 Torres et al., 2012; Van der Wal et al., 2009; Weber et al., 2017). The results have been inconsistent
354 or weak. However, the same outputs organized as niche distances can be predictive of abundance
355 (Martínez-Meyer et al., 2013). We hypothesize that this is essentially because the raw output of
356 the MaxEnt program is not a measure of fitness or suitability, but simply a measure of how
357 similar the environments in the background cells are to the environments of those cells in which
358 the species has been found (Merow et al., 2013). MaxEnt finds a solution for the probability
359 $P(\vec{e}_i|Y = 1)$, i.e., the probability of having the environmental variables \vec{e}_i given the fact that the
360 species has been observed ($Y = 1$) in cell i . This probability is an index of similarity between
361 environments (relative to a background) and not an index of suitability.

362 To estimate a probability of presence from presence-only data, further assumptions are re-
363 quired. Statistically, random sampling and a constant probability of detection, and an appro-
364 priate probability model (Royle et al., 2012) need to be true. Ecologically, the assumptions are:
365 no interference from other species, and a dispersal steady-state (Peterson, 2011). Therefore, in
366 general, ENM using presence only data does not estimate a probability of presence. This is an im-
367 portant and almost universally disregarded point: the outputs of presence-only ENM algorithms
368 may be used to classify cells in a spatial grid in terms of **similarity to a subset where a species**
369 **has been detected**; and whether this would coincide with classifying cell using a –generally
370 unavailable– **fitness function** is an open question. Simply assuming that the coincidence would
371 exist is unwarranted.

372 Now, it may initially appear that if an occurrence of a species is reported in a locality, then
373 fitness there should be high; and viceversa, (that a non-presence should imply negative fitness);
374 however, these expectations may fail for a variety of reasons. For instance, a population may be a
375 sink, with negative fitness (Pulliam, 2000) but a non-zero population maintained by immigration.
376 In contrast, a perfectly suitable set of cells, with potential positive fitness, may be unoccupied

377 due to dispersal disequilibrium, or to the presence of barriers, as shown in our numerical results.

378 However, the assumption that the fitness function defines a regular shape in niche space
379 allows points to be ordered in relation to such a shape. In their empirical explorations, Martínez-
380 Meyer et al. (2013) use not the value of the raw output of MaxEnt, but first selected the cells with
381 observed presences, ordered them per value of the raw output, retained the upper 90%, and used
382 these to estimate a mean value per variable in the model. This set of mean values constitutes
383 their approximation to the centroid of a niche with an implicitly assumed structure expressed in
384 fitness terms; and it is from this centroid that distances (Euclidean, in their case) can be calculated
385 to each point in the region. If it is true that the fitness function indeed describes an ellipsoidal
386 structure in niche space, then the distance between the environments \vec{e}_i and the hypothesized
387 niche centroid μ is at least an ordinal measure of suitability. The empirical association between
388 the distance to the centroid and abundance found by Martínez-Meyer et al. (2013) is therefore
389 consistent with the hypothesis of a structure of the niche in terms of fitness.

390 In their recent review of this problem, Dallas et al. (2017) found almost no evidence for
391 a relationship between distance to the centroid of niches and population density. Dallas et al.
392 (2017) choose convex hulls as models for niches. Our work provides formal basis for several of the
393 reasons Dallas et al. (2017) cite to explain why they did not find consistently good relationships
394 between niche and population density. But besides the many complications that we reviewed
395 and that may obscure such relationship, the issue of whether a convex-hull is a good model for
396 a fundamental niche (as opposed to a realized niche) needs to be addressed. We suspect that
397 the use of convex hulls, known to be sensitive to outliers and to sample size (Siväranta et al.,
398 2013), are not as good models for a hypothetical fundamental niche as ellipsoids, since hulls
399 are more affected by outliers and thus by errors in the data used to calculate them. Moreover,
400 convex hulls are determined only by the periphery of points in a cloud. Rather different clouds
401 of points (with very different centroids) can be modeled by exactly the same convex hull, with
402 the same centroid. Perhaps the use of convex hulls as models of fundamental niches introduces
403 yet another complicating factor to the analysis of the relationship between niche structure and

404 population density.

405 Finally, when true absence data is available, it may be possible to estimate a true probability
406 of occurrence. Thuiller et al. (2014) used presence-absence data to study the correlation between
407 probability of occurrence of species in plots (using a GLM regression), and estimations of three
408 demographic parameters (the intrinsic growth rate r , a "carrying capacity" K , and the initial
409 population size N). These were obtained from locally-defined Ricker models for population
410 growth, without metapopulation structure. Their results show that there is some tendency for
411 probability of occurrence to predict the value of N ; however, as they remark, their results are
412 inconclusive, among other issues, because without dispersal effects, simple species distribution
413 models ignore critical information required to understand range dynamics, which is precisely
414 the main result of our work: spatial dynamics is essential to understand the relationship between
415 niche and population abundance.

416 **Conclusion**

417 We propose that the relationship between population abundances in a locality, and the distance
418 to some "niche centroid" estimated using ENM, should be negative, for theoretical reasons related
419 to basic population dynamics: fitness, and hence abundance (under general assumptions) should
420 be related to environmental suitability. This is in some sense the null hypothesis. However, by
421 analytically and numerically examining a spatially-explicit population model, we show how the
422 use of an inappropriate measure of distance, transitory (disequilibrium) dynamics (Svenning
423 and Skov, 2004), and metapopulation dynamics (Pulliam, 2000) coupled with an Allee effect
424 (Holt, 2009b) can interfere with the simple pattern expected by a population dynamics argument.
425 Keeping in mind these complications will help to interpret correctly the growing body of data
426 on the relationship between population dynamics and the structure of niche space.

427 **Acknowledgments**

Figure 1:

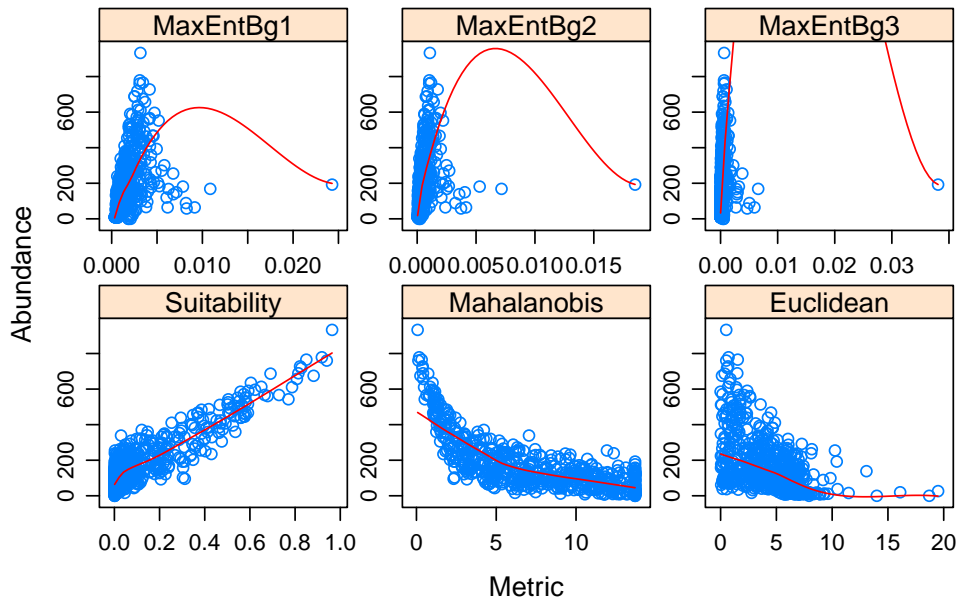
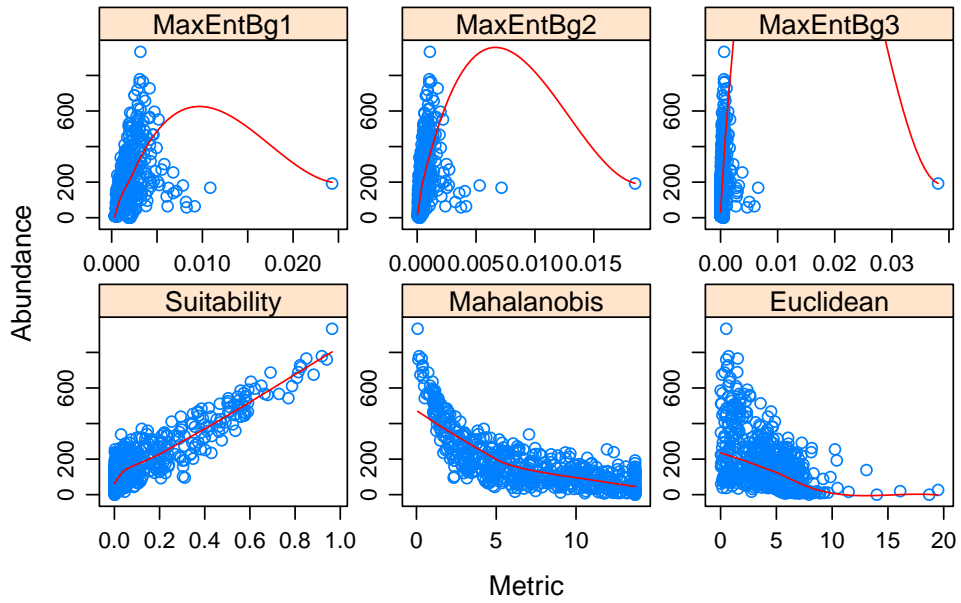


Figure 2:

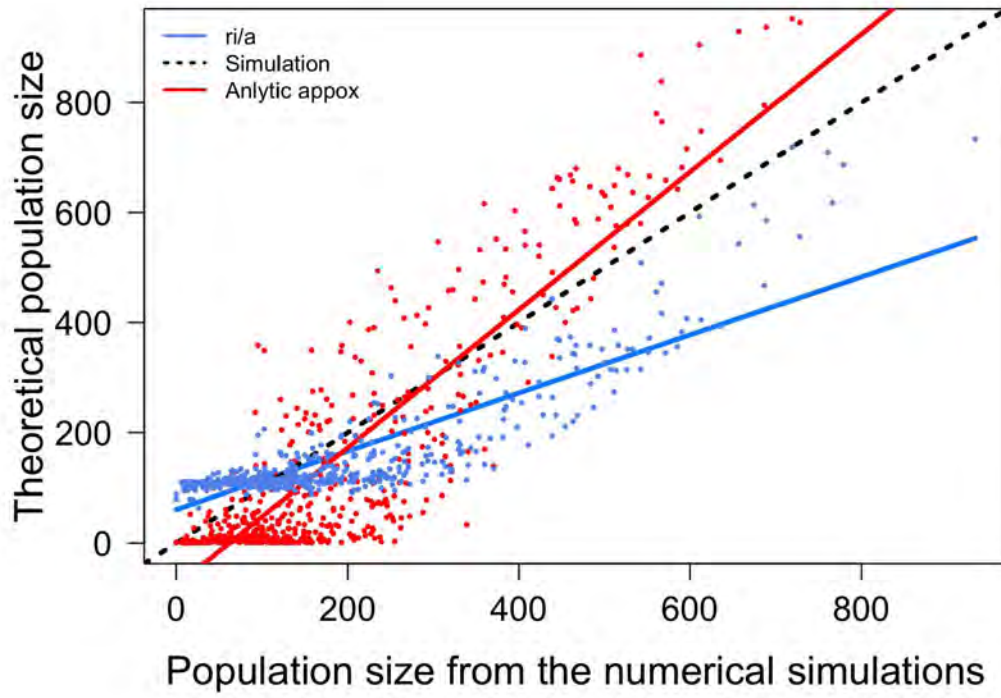


Figure 3:

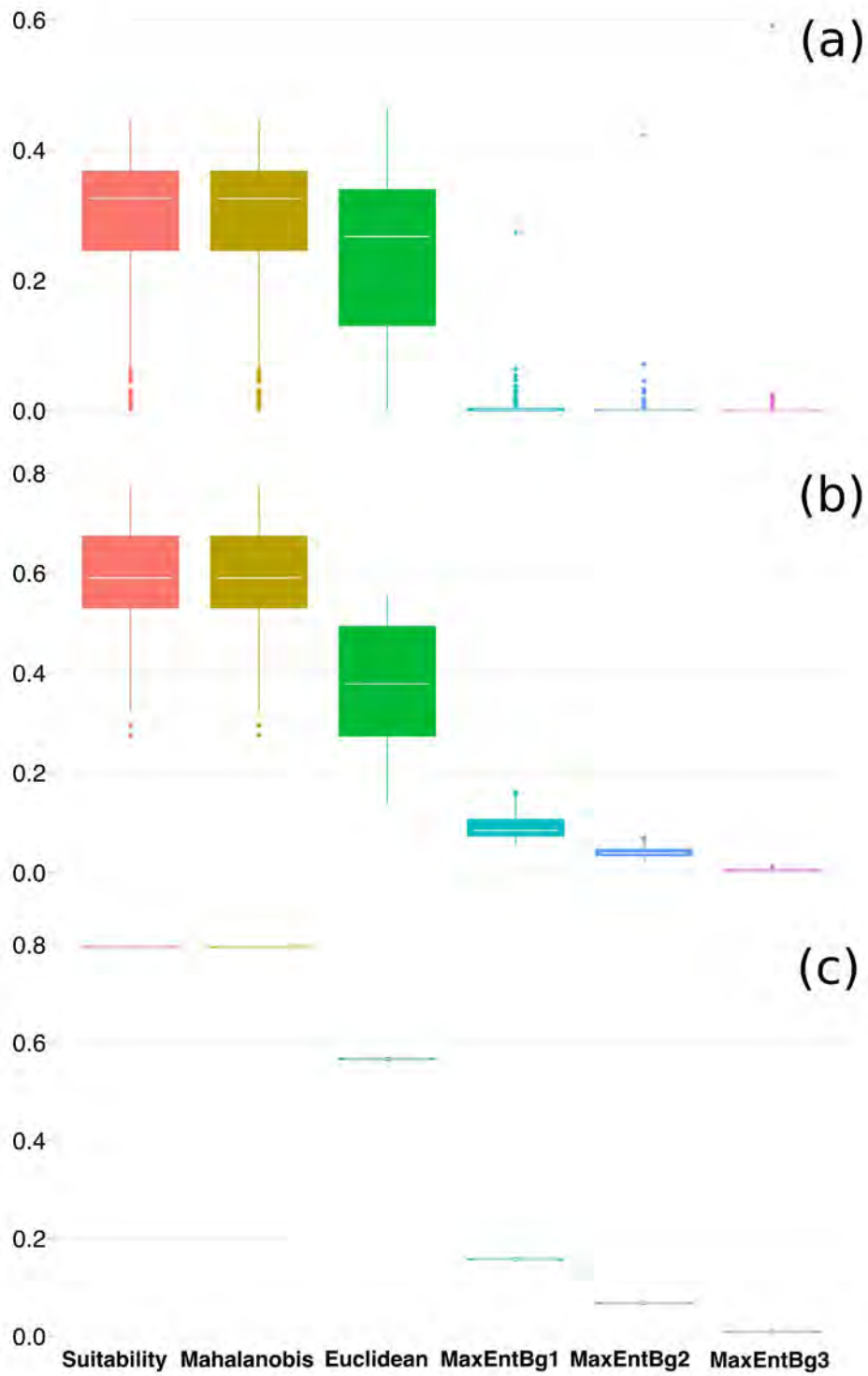


Figure 4:

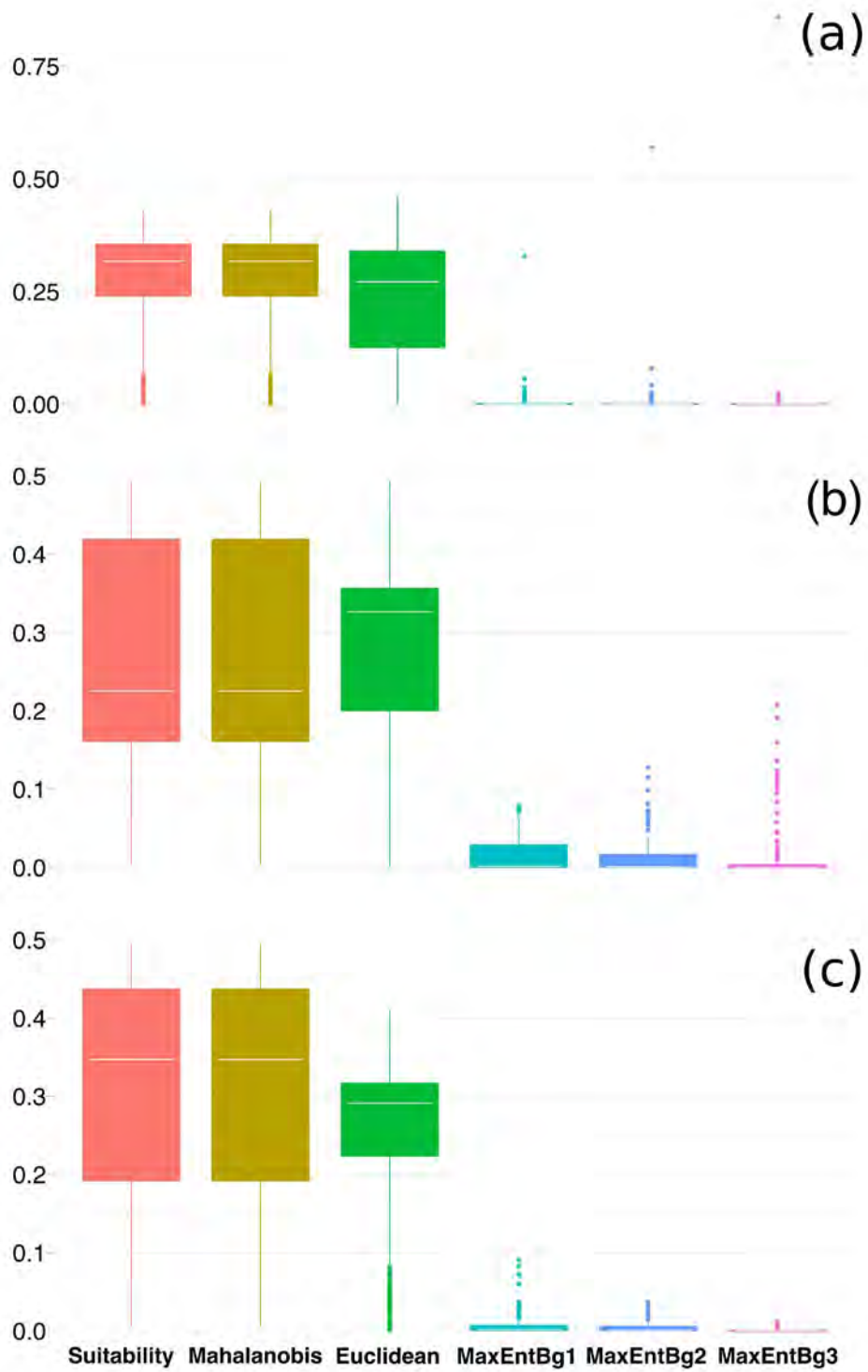
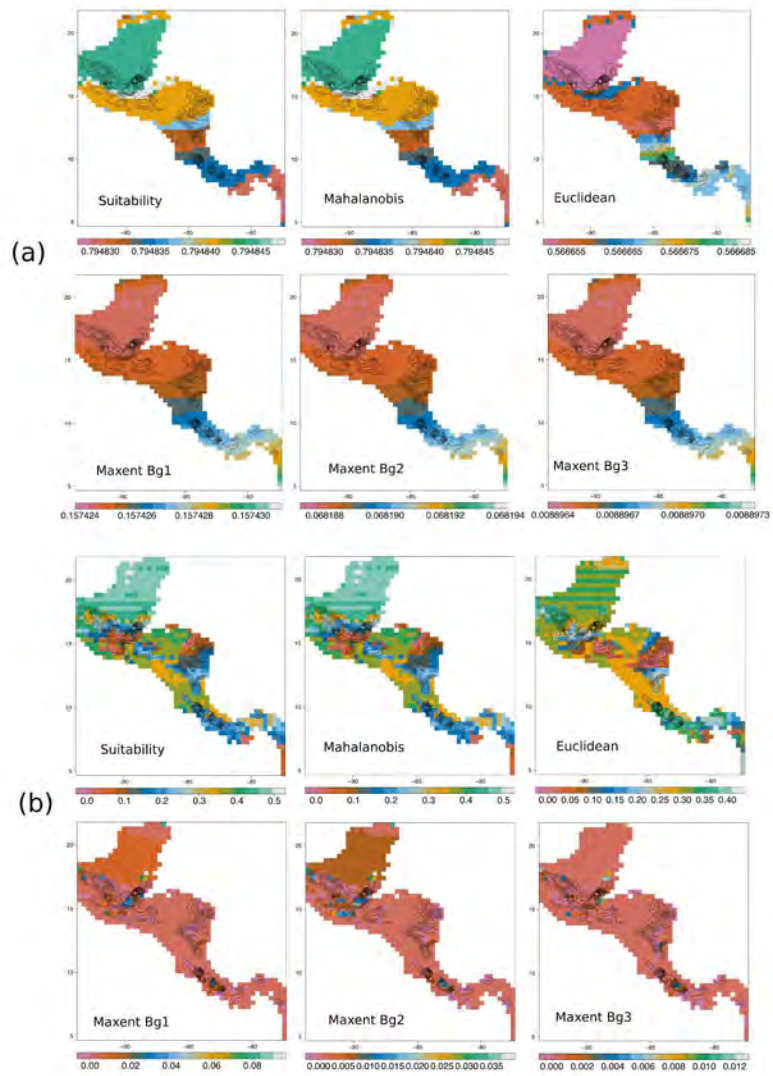


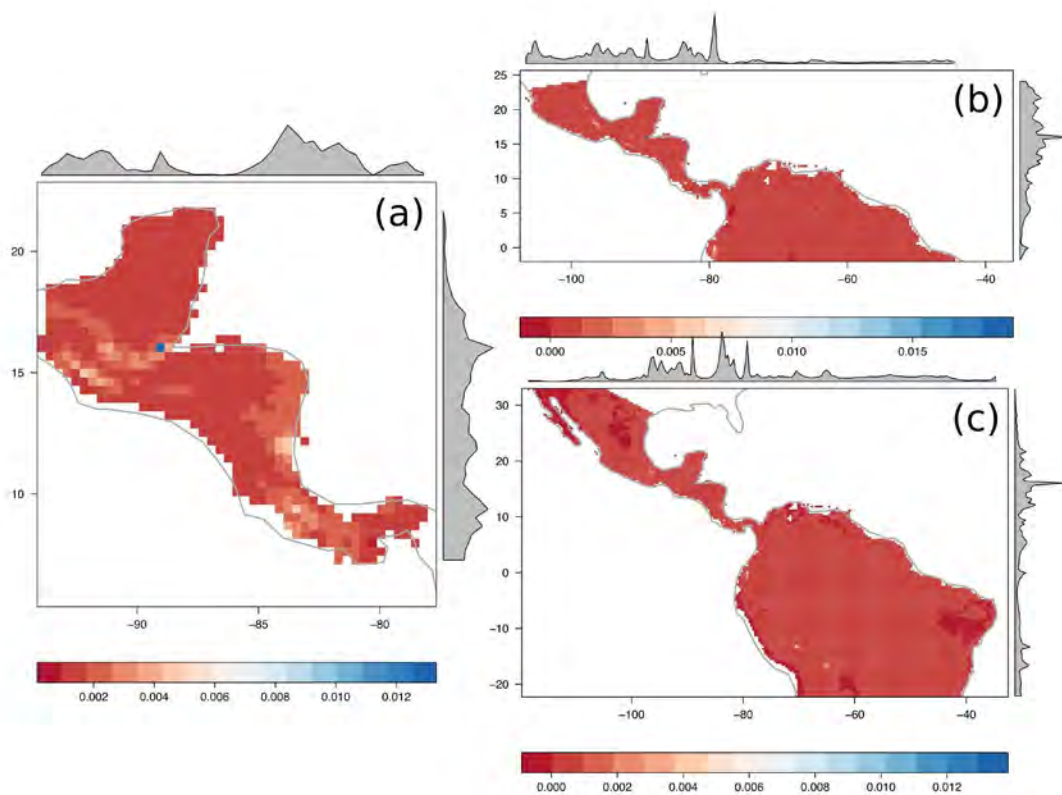
Figure 5:



430 **A Supplementary Figures**

431 **MaxEnt backgrounds**

Figure A1:



432 **B The uniqueness of the positive equilibrium point of the system 1**

433 The uniqueness and existence of the equilibrium point of the system 1 without Allee effect has
 434 been proved in Takeuchi (1989) by applying some theoretical results on cooperative systems.
 435 Here we sketch a geometrical proof of the uniqueness which seems to be more intuitive than
 436 the former. Indeed, we consider the case for three cells as a guide for the general n-dimensional
 437 model.

438 The three cells system is given by

$$\begin{aligned}
 \dot{x} &= r_1x - x^2 + \delta_{12}y + \delta_{13}z - \delta_{21}x - \delta_{31}x = F_1(x, y, z), \\
 \dot{y} &= r_2y - y^2 + \delta_{21}x + \delta_{23}z - \delta_{12}y - \delta_{32}y = F_2(x, y, z), \\
 \dot{z} &= r_3z - z^2 + \delta_{31}x + \delta_{32}y - \delta_{13}z - \delta_{23}z = F_3(x, y, z).
 \end{aligned}
 \tag{B1}$$

439 The equilibrium points are the solutions of the system of equations $F_i = 0$, $i = 1, 2, 3$ in
 440 the region $\mathcal{O}^+ = \{x \geq 0, y \geq 0, z \geq 0\}$. The isoclines of the System (B1) are the surfaces
 441 $F_i(x, y, z) = 0$, $i = 1, 2, 3$. To determine the number of equilibrium points of the system B1,
 442 we analyze the geometry of the isoclines (These are denoted by *Eqx*, *Eqy* and *Eqz* according to
 443 $i = 1, 2, 3$). From the equation $F_1(x, y, z) = 0$ it follows that the isocline *Eqx* is contained in the
 444 surface given by

$$z = -\frac{\delta_{12}}{\delta_{13}}y + \frac{x^2 - (r_1 - \delta_{21} - \delta_{31})x}{\delta_{13}}.
 \tag{B2}$$

445 The surface (B2) is a parabolic cylinder; for each point in the parabola $P(x) = (x, 0, \frac{x^2 - (r_1 - \delta_{21} - \delta_{31})x}{\delta_{13}})$,
 446 there is a straight line $L_x(y) = (x, 0, \frac{x^2 - (r_1 - \delta_{21} - \delta_{31})x}{\delta_{13}}) + y(0, 1, -\frac{\delta_{12}}{\delta_{13}})$. The vertex of the parabola
 447 $P(x)$ is reached at $x^* = \frac{1}{2}(r_1 - \delta_{21} - \delta_{31})$; The line $L_{x^*}(y)$ corresponding to x^* will be called *Vertex*
 448 *Line*. The cylinder is symmetric with respect to the plane which contains the Vertex Line and is
 449 orthogonal to $x - y$ plane. Notice that the Vertex Line is located under the plane $z = 0$, for all
 450 $x, y \geq 0$, therefore, if $r_1 - \delta_{21} - \delta_{31} > 0$ the isocline *Eqx* is the cylindric region corresponding to

451 $x \geq r_1 - \delta_{21} - \delta_{31}, y \geq 0$. Analogously, if $r_1 - \delta_{21} - \delta_{31} < 0$, Eqx is the region corresponding to
 452 $x \geq 0, y \geq 0$ (see Figure B. B1).

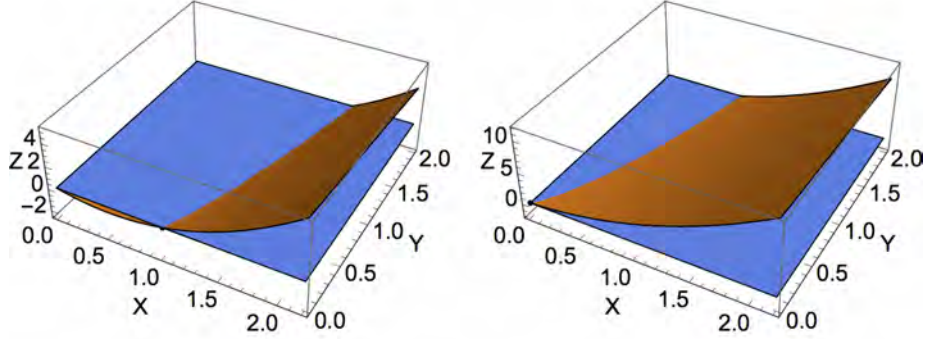


Figure B1: Isocline $\dot{x} = 0$. a) $r_1 - \delta_{21} - \delta_{31} > 0$; b) $r_1 - \delta_{21} - \delta_{31} \leq 0$

453 The surfaces Eqy and Eqz are similar to Eqx , since the equations $F_i(x, y, z) = 0$ are anal-
 454 ogous. The straight lines that conform the surface $F_2(x, y, z) = 0$ are defined by $L_y(x) =$
 455 $(0, y, \frac{y^2 - (r_2 \delta_{12} - \delta_{32})y}{\delta_{23}}) + x(1, 0, -\frac{\delta_{21}}{\delta_{23}})$; the Vertex Line L_{y^*} corresponds to $y^* = \frac{1}{2}(r_2 - \delta_{12} - \delta_{32})$.
 456 Notice that this line is contained in a plane which is transversal to the Vertex Line $L_{x^*}(y)$. Hence,
 457 the intersection γ of Eqx and Eqy in the positive octant is a curve whose intersection with any
 458 parallel plane to $z = 0$ consists of a just one point.

459 Finally, The surface $F_3(x, y, z) = 0$ is contained in the cylinder determined by the lines $L_z(y) =$
 460 $(\frac{z^2 - (r_3 - \delta_{13} - \delta_{23})z}{\delta_{31}}, 0, z) + y(-\frac{\delta_{32}}{\delta_{31}}, 1, 0)$; The Vertex Line L_{z^*} corresponds to $z^* = \frac{1}{2}(r_3 - \delta_{13} - \delta_{23})$.
 461 Thus, Eqz cuts to γ in a single point (see Figure (B. B2)). This proves the existence and uniqueness
 462 of the coexistence equilibrium point for three cells.

463 It is hard and very cumbersome to prove the existence and uniqueness of the equilibrium point
 464 for the general case of n cells. Shortly, the method is based in a reduction of n cells to $n - 1$ cells.
 465 We show how to reduce 4 cells to just 3. Let x_i be the population in the patch $i = 1, \dots, 4$, then

$$\dot{x}_i = r_i x_i - x_i^2 + \sum_{j \neq i, j=1}^4 \delta_{i,j} x_j - \left(\sum_{j \neq i, j=1}^4 \delta_{j,i} \right) x_i. \quad (\text{B3})$$

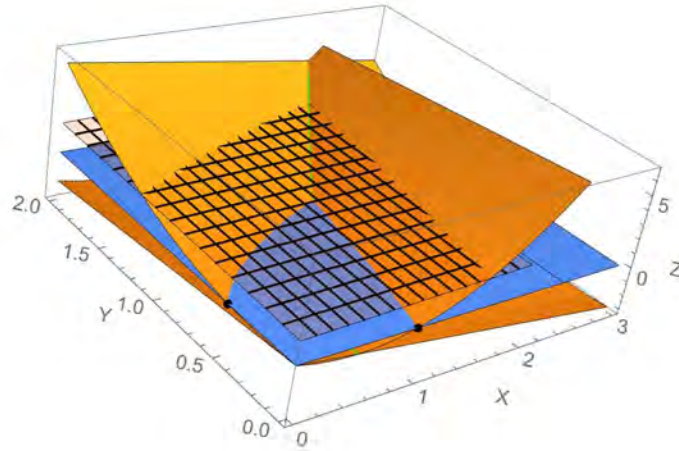


Figure B2:

466 Notice that for each positive value of x_3 , the isoclines

$$\dot{x}_1 = 0, \quad \dot{x}_2 = 0, \quad \dot{x}_3 = 0,$$

467 are analogous to the case of three cells, hence their intersection consists of a single point
 468 which depends on x_3 . Thus, the intersection of these three isoclines generates a curve whose x_4
 469 coordinate grows monotonically, when x_3 varies continuously. Therefore, the intersection of this
 470 curve with the isocline $\dot{x}_4 = 0$ has a unique point.

471 Literature Cited

- 472 Ahnert, K., and M. Mulansky. 2011. Odeint - Solving ordinary differential equations in C++. AIP
473 Conference Proceedings 1389:1586–1589.
- 474 Boukal, D. S., and L. Berec. 2002. Single-species Models of the Allee Effect: Extinction Boundaries,
475 Sex Ratios and Mate Encounters. *J. theor. Biol* 218:375–394.
- 476 Brown, J. H. 1984. On the relationship between abundance and distribution of a species. *The*
477 *American Naturalist* 124:255–279.
- 478 Carter, R. N., and S. D. Prince. 1981. Epidemic models used to explain biogeographical distribu-
479 tion limits. *Nature* 293:644–645.
- 480 Ceballos, G., and G. Oliva. 2005. *Los Mamíferos Silvestres de México* .
- 481 Colwell, R. K., and T. F. Rangel. 2009. Hutchinson’s duality: The once and future niche. *Proceed-*
482 *ings of the National Academy of Sciences USA* 106:19651–19658.
- 483 Dallas, T., R. R. Decker, and A. Hastings. 2017. Species are not most abundant in the centre of
484 their geographic range or climatic niche. *Ecology Letters* pages n/a–n/a.
- 485 Farber, O., and R. Kadmon. 2003. Assessment of alternative approaches for bioclimatic modeling
486 with special emphasis on the Mahalanobis distance. *Ecological Modelling* 160:115–130.
- 487 Hanski, I. 1999. *Metapopulation Ecology*. *Metapopulation Ecology*. OUP Oxford.
- 488 Hijmans, R. J., S. Cameron, J. Parra, P. G. Jones, and A. Jarvis. 2005. Very high resolution interpo-
489 lated climate surfaces for global land areas. *International Journal of Climatology* 25:1965–1978.
- 490 Holt, R. 2009*a*. Predation and Community Organization. Chap. 3.3, pages 274–281 *in* S. A. Levin,
491 ed. *The Princeton Guide to Ecology*. Princeton University Press, Princeton, NJ.
- 492 Holt, R., J. H. Lawton, J. K. Gaston, and T. M. Blackburn. 1997. On the relationship between
493 range size and local abundances: back to the basics. *Oikos* 78:183–190.

494 Holt, R. D. 2009b. Bringing the Hutchinsonian niche into the 21st century: Ecological and evolu-
495 tionary perspectives. *Proceedings of the National Academy of Sciences USA* 106:19659–19665.

496 Hutchinson, G. E. 1957. Concluding Remarks. *Cold Spring Harbor Symposia on Quantitative*
497 *Biology* 22:415–427.

498 ———. 1978. *An Introduction to Population Ecology*. Yale University Press, New Haven.

499 Keith, D. A., H. R. Akçakaya, W. Thuiller, G. F. Midgley, R. G. Pearson, S. J. Phillips, H. M. Regan,
500 M. B. Araújo, and T. G. Rebelo. 2008. Predicting extinction risks under climate change: coupling
501 stochastic population models with dynamic bioclimatic habitat models. *Biology Letters* 4:560–
502 563.

503 Keitt, T. H., M. a. Lewis, and R. D. Holt. 2001. Allee effects, invasion pinning, and species’
504 borders. *American Naturalist* 157:203–216.

505 Kirkpatrick, M., and N. H. Barton. 1997. Evolution of a species range. *American Naturalist*
506 150:1–23.

507 Legendre, P., and L. Legendre. 1998. *Numerical Ecology*, vol. 20. Second edi ed. Elsevier Scientific
508 Publishing Company, Amsterdam.

509 Lira-Noriega, A., and J. D. Manthey. 2014. Relationship of genetic diversity and niche centrality:
510 A survey and analysis. *Evolution* 68:1082–1093.

511 Maguire, B. 1973. Niche response structure and the analytical potentials of its relationship to the
512 habitat. *American Naturalist* 107:213–246.

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

515 Merow, C., M. J. Smith, and J. a. Silander. 2013. A practical guide to MaxEnt for modeling species’
516 distributions: What it does, and why inputs and settings matter. *Ecography* 36:1058–1069.

517 Nenzén, H. K., R. M. Swab, D. a. Keith, and M. B. Araújo. 2012. *demoniche* - an R-package for
518 simulating spatially-explicit population dynamics. *Ecography* 35:577–580.

519 Nielsen, S. E., C. J. Johnson, D. C. Heard, and M. S. Boyce. 2005. Can models of presence-absence
520 be used to scale abundance? Two case studies considering extremes in life history. *Ecography*
521 28:197–208.

522 Osorio-Olvera, L. A., M. Falconi, and J. Soberón. 2016. Sobre la relación entre idoneidad del
523 hábitat y la abundancia poblacional bajo diferentes escenarios de dispersión. *Revista Mexicana*
524 *de Biodiversidad* 87:1080–1088.

525 Pearce, J., and M. S. Boyce. 2006. Modelling distribution and abundance with presence-only data.
526 *Journal of Applied Ecology* 43:405–412.

527 Peterson, A. T. 2011. Ecological niche conservatism: A time-structured review of evidence. *Jour-*
528 *nal of Biogeography* 38:817–827.

529 Peterson, A. T., J. Soberón, R. G. Pearson, R. P. Anderson, E. Martínez-Meyer, M. Nakamura, and
530 M. Bastos Araujo. 2011. Ecological niches and geographic distributions.

531 Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. Maximum entropy modeling of species
532 geographic distributions. *Ecological Modelling* 190:231–259.

533 Pulliam, H. 2000. On the relationship between niche and distribution. *Ecology Letters* 3:349–361.

534 Rapoport, E. H. 1975. *Areografía: Estrategias Geográficas de las Especies*. Fondo de Cultura
535 Económica, Mexico City.

536 Royle, J. A., R. B. Chandler, C. Yackulic, and J. D. Nichols. 2012. Likelihood analysis of species
537 occurrence probability from presence-only data for modelling species distributions. *Methods*
538 *in Ecology and Evolution* 3:545–554.

539 Sagarin, R. D., and S. D. Gaines. 2002. Geographical abundance distributions of coastal inverte-
540 brates: using one-dimensional ranges to test biogeographic hypotheses. *Journal of Biogeogra-*
541 *phy* 29:985–997.

542 Sagarin, R. D., S. D. Gaines, and B. Gaylord. 2006. Moving beyond assumptions to understand
543 abundance distributions across the ranges of species. *Trends in Ecology and Evolution* 21:524–
544 530.

545 Salguero-Gómez, R., O. R. Jones, C. R. Archer, C. Bein, H. de Buhr, C. Farack, F. Gottschalk,
546 A. Hartmann, A. Henning, G. Hoppe, G. Römer, T. Ruoff, V. Sommer, J. Wille, J. Voigt, S. Zeh,
547 D. Vieregg, Y. M. Buckley, J. Che-Castaldo, D. Hodgson, A. Scheuerlein, H. Caswell, and J. W.
548 Vaupel. 2016. COMADRE: a global data base of animal demography. *Journal of Animal*
549 *Ecology* 85:371–384.

550 Schurr, F. M., J. Pagel, J. S. Cabral, J. Groeneveld, O. Bykova, R. B. O'Hara, F. Hartig, W. D.
551 Kissling, H. P. Linder, G. F. Midgley, B. Schroder, A. Singer, and N. E. Zimmermann. 2012.
552 How to understand species' niches and range dynamics: a demographic research agenda for
553 biogeography. *Journal of Biogeography* 39:2146–2162.

554 Smolik, M., S. Dullinger, F. Essl, I. Kleinbauer, M. Leitner, J. Peterseil, L.-M. Stadler, and G. Vogl.
555 2010. Integrating species distribution models and interacting particle systems to predict the
556 spread of an invasive alien plant. *Journal of Biogeography* 37:411–422.

557 Soberón, J. 1986. The relationship between use and suitability of resources and its consequence
558 to insect population dynamics. *American Naturalist* 127:338–357.

559 Soberón, J., L. Osorio-Olvera, and A. T. Peterson. 2017. Diferencias conceptuales entre modelación
560 de nichos y modelación de áreas de distribución. *Revista Mexicana de Biodiversidad* 88:437–
561 441.

562 Soberón, J. M. 2010. Niche and area of distribution modeling: a population ecology perspective.
563 *Ecography* 33:159–167.

- 564 Svenning, J.-C., and F. Skov. 2004. Limited filling of the potential range in European tree species.
565 Ecology Letters 7:565–573.
- 566 Syväranta, J., A. Lensu, T. J. Marjomäki, S. Oksanen, and R. I. Jones. 2013. An Empirical Evalua-
567 tion of the Utility of Convex Hull and Standard Ellipse Areas for Assessing Population Niche
568 Widths from Stable Isotope Data. PLoS ONE 8.
- 569 Takeuchi, Y. 1989. Cooperative systems theory and global stability of diffusion models. Acta
570 Applicandae Mathematicae 14:49–57.
- 571 Thuiller, W., T. Münkemüller, K. H. Schiffers, D. Georges, S. Dullinger, V. M. Eckhart, T. C.
572 Edwards, D. Gravel, G. Kunstler, C. Merow, K. Moore, C. Piedallu, S. Vissault, N. E. Zimmer-
573 mann, D. Zurell, and F. M. Schurr. 2014. Does probability of occurrence relate to population
574 dynamics? Ecography 37:1155–1166.
- 575 Torres, N., P. d. M. Júnior, T. Santos, L. Silveira, A. T. Jácomo de Almeida, and A. F. Diniz-Filho.
576 2012. Can species distribution modelling provide estimates of population densities? A case
577 study with jaguars in the Neotropics. Diversity and Distributions 18:615–627.
- 578 Tôrres, N. M., P. De Marco, T. Santos, L. Silveira, A. T. de Almeida Jácomo, and J. A. F. Diniz-
579 Filho. 2012. Can species distribution modelling provide estimates of population densities? A
580 case study with jaguars in the Neotropics. Diversity and Distributions 18:615–627.
- 581 Tuya, F., T. Wernberg, and M. S. Thomsen. 2008. The spatial arrangement of reefs alters the
582 ecological patterns of fauna between interspersed algal habitats. Estuarine, Coastal and Shelf
583 Science 78:774–782.
- 584 Udvardy, M. 1969. Dynamic Zoogeography. With Special Reference to Land Animals. van Nos-
585 trand Reinhold, New York.
- 586 Ureña-Aranda, C., O. Rojas-Soto, E. Martínez-Meyer, C. Yañez-Arenas, R. Landgrave Ramírez,
587 and A. Espinosa de los Monteros. 2015. Using range-wide abundance modeling to identify key

588 conservation areas for the micro-endemic Bolson Tortoise (*Gopherus flavomarginatus*). PLoS
589 ONE 10:1–14.

590 Vaclavik, T., and R. K. Meentemeyer. 2012. Equilibrium or not? Modelling potential distribution
591 of invasive species in different stages of invasion. *Diversity and Distributions* 18:73–83.

592 Van der Wal, J., L. Shoo, C. Johnson, and S. Williams. 2009. Abundance and the environmental
593 niche: environmental suitability estimated from niche models predicts the upper limit of local
594 abundance. *American Naturalist* 174:282–291.

595 Vandermeer, J. H. 1972. Niche theory. *Annual Review of Ecology and Systematics* 3:107–132.

596 Weber, M. d. M., and C. E. V. Grelle. 2012. Does environmental suitability explain the relative
597 abundance of the tailed tailless bat, *Anoura caudifer*? *Natureza & Conservação* 10:221–227.

598 Weber, M. M., R. D. Stevens, J. A. F. Diniz-Filho, and C. E. V. Grelle. 2017. Is there a correlation
599 between abundance and environmental suitability derived from ecological niche modelling? A
600 meta-analysis. *Ecography* 40:817–828.

601 Yañez-Arenas, C., E. Martínez-Meyer, S. Mandujano, and O. Rojas-Soto. 2012. Modelling geo-
602 graphic patterns of population density of the white-tailed deer in central Mexico by imple-
603 menting ecological niche theory. *Oikos* 121:2081–2089.

604 **Figure legends**

Figure 1: Scatterplots of the relationship between population size, in every grid-cell, as a function of the different predictors, for two randomly-selected initial points in the simulations.

Figure 2: Relationship between population size, calculated numerically, from the 638 initial conditions, without Allee effect, and: the analytical approximation (red); the identical function (dashed line in black); and the simple uncoupled logistic steady-state values (blue).

Figure 3: Distribution of the correlation value between niche metric and population abundance for different times in the model without Allee effect. A) Ten years after dispersal; B) 500 years period; C) 2500 year period. A Spearman rank correlation was used for variables with a monotonic relationship with abundance (Suitability, Mahalanobis and Euclidean distances), and a Pearson correlation for the MaxEnt outputs.

Figure 4: ρ distribution of the relationship between niche metric and population abundance for different times in the model with Allee effect. A) Ten years after dispersal; B) 500 years period; C) 2500 year period. A Spearman rank correlation was used for variables with a monotonic relationship with abundance (Suitability, Mahalanobis and Euclidean distances), and a Pearson correlation for the MaxEnt outputs.

Figure 5: ρ maps of the relationship between niche metric and population abundance after 2500 years of invasion. (a) ρ map for each initial condition (grid-cell) in the model without Allee effect. (b) ρ map for each initial condition (pixel) in the model with an Allee effect.

605 **Online figure legends**

Figure A1: Raw outputs for MaxEnt with three different backgrounds. The three geographic regions (a), (b) and (c) were used as background for the MaxEnt calculations in the text, with the details described in the Methods section.

Capítulo 4

Herramientas informáticas para la modelación de la distribución de las especies

Resumen

En este capítulo se muestran algunas de las herramientas computacionales que he desarrollado para los análisis de esta tesis. Estas herramientas consisten una aplicación web llamada “Niche-ToolBox” y un paquete de R [133] llamado “ntb”. NicheToolbox es una plataforma web basada en el framework “shiny” [22] de R [133], la cual sirve para la obtención, limpieza y visualización de datos de presencia, así como para la modelación y validación de los modelos de distribución de especies (SDMs por sus siglas en ingles). ntb es un paquete de R [133] para simular la dinámica de crecimiento espacio-temporal de una o varias poblaciones que crecen en ciertas regiones del espacio geográfico y se expanden a través de éste según un kernel de dispersión y la idoneidad del hábitat; el paquete utiliza los modelos de distribución basados en proceso presentados en el capítulo 2 (ecuaciones 3.1 y 3.2). Ambas herramientas son software.

§4.1. Introducción

Desde los intereses teóricos hasta los aplicados, entender el por qué una especie está presente o no en un sitio determinado es una pregunta de relevancia biológica, de tal modo que durante varios siglos los ecólogos y biogeógrafos constantemente han documentado y tratado de comprender las relaciones existentes entre la distribución de las especies y su ambiente. Para ayudar a entender estas relaciones se han propuesto una serie de modelos cualitativos y cuantitativos que tienen como objetivo estimar el área de distribución de una o varias especies [38]. A tales modelos se les conoce como Modelos de Distribución de Especies (Species Distribution Models -SDMs-) [108]. Estos modelos utilizan un conjunto de herramientas matemáticas y estadísticas para predecir la distribución geográfica de una especie [126].

Los SDMs han mostrado tener varias aplicaciones, por ejemplo; se han aplicado para elaborar

planes conservación [99, 55], localizar sitios de prioridad ecológica (por ejemplo, *hotspots de biodiversidad*) [6], estudiar relaciones evolutivas [72], predecir sitios donde se darán brotes epidémicos [58] y para estudiar los efectos del cambio climático en la riqueza y distribución de las especies [179]. Dado el gran campo de aplicaciones de estos modelos, actualmente se han desarrollado una variedad de paquetes de software y aplicaciones web que facilitan el proceso de modelado de nichos¹ y de la distribución².

A continuación se enlistan los principales pasos implicados en tal proceso y algunos de los paquetes más frecuentemente utilizados para realizar cada etapa:

1. Obtención de datos de presencia georreferenciados (por ejemplo, el sitio web de GBIF [161], spocc de R [21], dismo de R [67],).
2. Limpieza y filtrado de las bases de datos, las cuales frecuentemente incluyen información duplicada, redundante y mal georreferenciada (Acces, Excel, MySQL, sphin de R [1], etc).
3. Extracción de la información ambiental de los puntos donde se encuentra la especie (sistemas de información geográfica como GRASS [110], QGIS [132]; ArcGIS [42], Raster de R [66] etc).
4. Modelación del nicho ecológico y el área de distribución (MaxEnt [127], dismo de R [67], BIOMOD de R [163], GARP [157], openModeller [154], etc).
5. Evaluación del desempeño de los SDMs utilizando métricas como ROC [45], ROC parcial [125], Matriz de confusión y métricas asociadas a ésta [46] (MaxEnt [127], ENMGadgets de R [7]).

Cuadro 4.1: Proceso de modelación y algunos paquetes de software conocidos

A pesar de que existe una gran variedad de paquetes para realizar cada etapa del proceso de modelado, su uso puede ser problemático ya que para poder llevar a cabo este proceso completo el usuario tiene que: a) aprender a programar en un lenguaje como R [133], en el cual existen varias librerías que contienen las funciones necesarias para realizar los 5 pasos mencionados en el cuadro 4.1; o bien b) si se desea trabajar con una interfaz gráfica de usuario, es necesario aprender a utilizar varios programas, ya que en ninguno de estos se pueden realizar todas las tareas implicadas en el proceso de modelación.

NicheToolBox

NicheToolBox es una aplicación de web basada en el framework «shiny»[22] de R la cual tiene como objetivo proveer al usuario las herramientas necesarias para realizar el proceso de modelación de nichos y distribuciones geográficas en una interfaz gráfica de usuario (GUI) amigable. La aplicación tiene 4 funciones principales: I) Limpieza y obtención de datos, II) Análisis exploratorio de

¹Con esto se pretende estimar las combinaciones ambientales en las que la especie puede subsistir.

²Se refiere a la estimación del área de distribución de una especie, e decir los sitios en la geografía en los que se puede encontrar a la especie

datos ambientales, III) Modelación de nichos y áreas de distribución; IV) Evaluación de los modelos de distribución.

La aplicación cuenta con un generador de flujos de trabajo (workflow) el cual pretende hacer el proceso de modelado transparente; en workflow se encuentra almacenado el código ejecutado para realizar los análisis que el usuario hizo en la interfaz gráfica de usuario. Actualmente la aplicación está alojada en un servidor de la CONABIO :

<http://shiny.conabio.gob.mx:3838/nichetoolb2/>

Dentro de las utilidades que se pueden encontrar en NicheToolBox están:

1. Obtención de datos de presencia (usa la API de GBIF mediante el paquete “spocc”[21]).
 - a) visualización del historial de los reportes de presencias por año.
2. Limpieza de los datos de presencia.
 - a) Eliminación de duplicados.
 - b) Filtrado de los datos que caen en el área de estudio.
 - c) Mapas dinámicos que permiten la exploración y eliminación de datos en una interfaz gráfica.
3. Visualización del nicho n-dimensional utilizando las 19 bios de WorldClim [64].
4. Análisis de Cluster de los puntos de presencia en el espacio ambiental.
 - a) k-means clustering.
 - 1) Visualizador de los Clusters en el espacio geográfico **G** y ecológico **E**.
 - b) Cluster dendrogram para explorar asociaciones entre las variables de nicho.
5. Análisis de correlaciones entre los valores de las variables ambientales.
 - a) Correlograma.
 - b) Tabla de correlaciones.
 - c) Algoritmo para filtrado de las variables correlacionadas .
6. Modelación de Nichos Ecológicos.
 - a) Elipsoides.
 - b) Bioclim.
 - c) MaxEnt (en versión de prueba).
 - d) ENM proyección de los modelos al espacio **G**.
7. Roc Parcial [125].
 - a) prueba de hipótesis

8. Conversor del mapa continuo de idoneidad a binario
 - a) Estimación del área de distribución basada en el mapa binario.
9. Algunas métricas de evaluación de los modelos de distribución:
 - Matriz de confusión, TSS, Kappa, Prevalencia, Sensibilidad, Especificidad. [46]

§4.1.1. Manual de NicheToolBox (en inglés).

Nichetoolbox

Luis Osorio-Olvera

Contents

NicheToolBox	1
Installation	1
Launching the app	2
The AppSettings section	3
The Data section	5
Geographic explorations using Dynamic Maps	10
Saving the workflow	15
Niche space (steps to make a cluster analysis)	17
1. Extracting niche values from raster layers	17
2. Niche explorations	18
3. Niche clustering	20
4. Niche correlations	23
6. Ecological niche modeling	24
7. Species distribution model performance	39
References	47

NicheToolBox

In this tutorial I show how the application works. The application was developed for GSOC 2016

Installation

Go to the github repo of the project nichetoolbox repo.

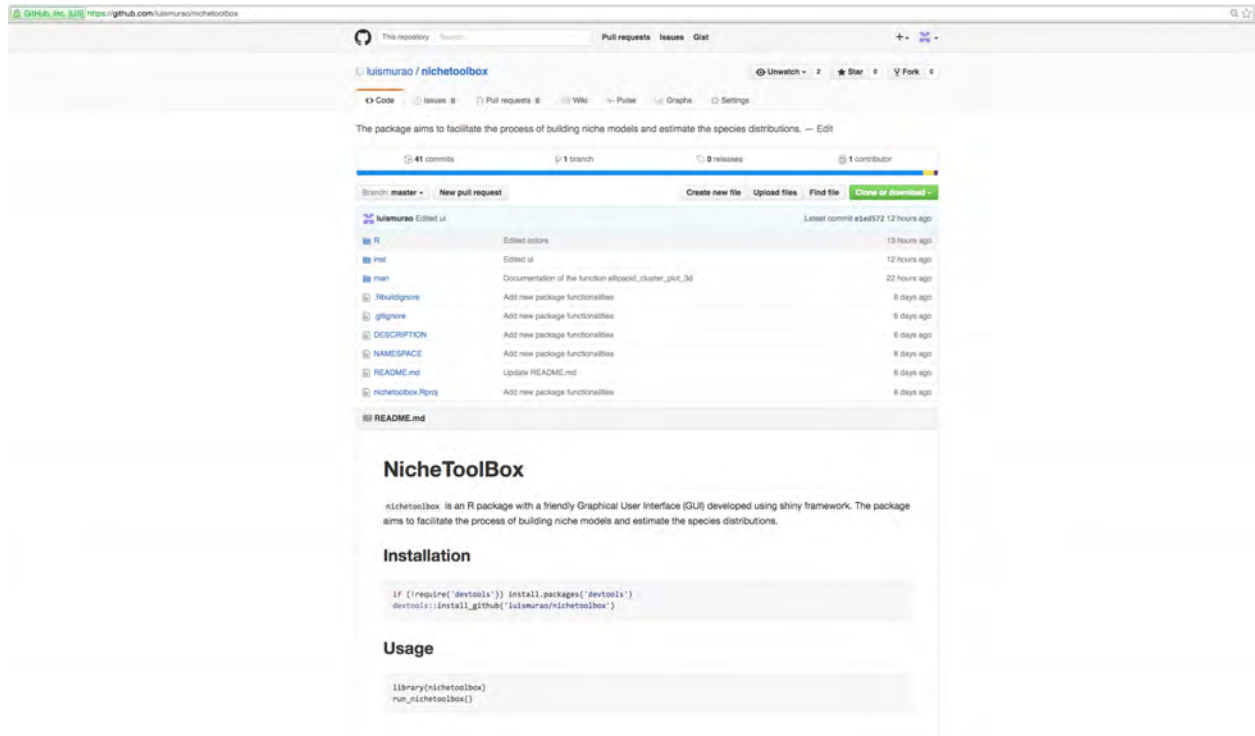


Figure 1:

Then copy and run the installation instructions in R:

```
if (!require('devtools')) install.packages('devtools')
devtools::install_github('luismurao/nichetoolbox')
```

Launching the app

```
library(nichetoolbox)
run_nichetoolbox()
```

First look

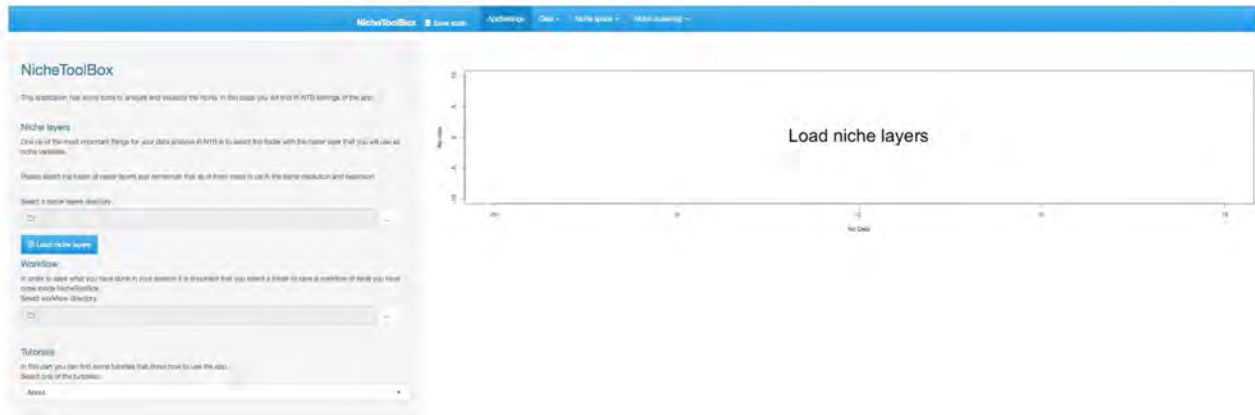


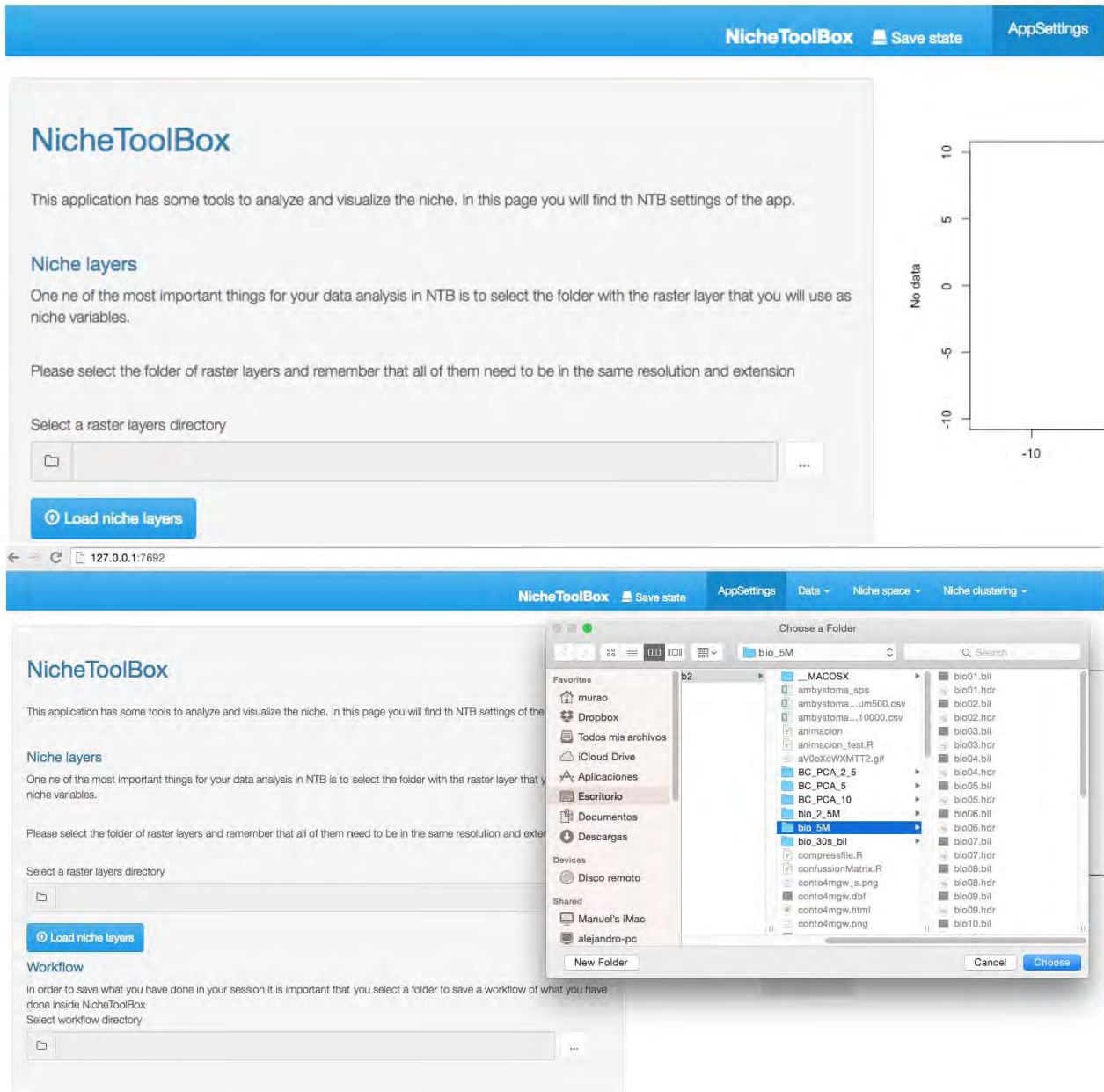
Figure 2:

The AppSettings section

In this section you need to specify the folder that contains the niche layers that you will use for the modeling process, as well as the folder where you will save your workflow.

Loading niche layers

On the left panel go to **Niche layers** section and select the folder where your niche raster layers are. Remember that they need to have the same spatial extent and resolution (**raster** formats accepted: .asc, .bil, .sdat, .rst, .nc, .tif, .envi, .img).



Press the **Load niche layers** button and wait. In a few seconds appears a plot showing one of the layers contained in the folder



Figure 3:

Workflow

To get track of your work in NicheToolBoox, you need to specify the folder where you want to save your analysis, data, maps etc. Go to **Workflow** section and select the folder.

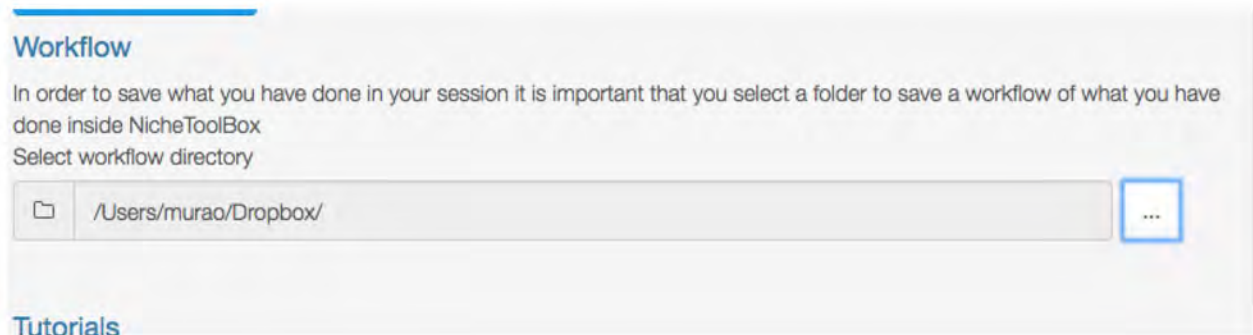


Figure 4:

The Data section

Now, we are ready to work with NicheToolBoox. First, we need some georeferenced records of the species we want to model. NicheToolBoox can work with two source of longitude/ latitude data: a) GBIF records, which you can search, download and clean GBIF records, b) you can upload and clean your own occurrence data from a local file.

Seaching GBIF records

Go to Data -> GBIF data. Enter species **genus**, **species name** where corresponds and optionally specify the number of records that you want to search (**occ search limit**). Press **Search GBIF** button and wait. If the species is in the GBIF portal a data table will be displayed, if the species is not in GBIF, it will display the following message: No occurrences found

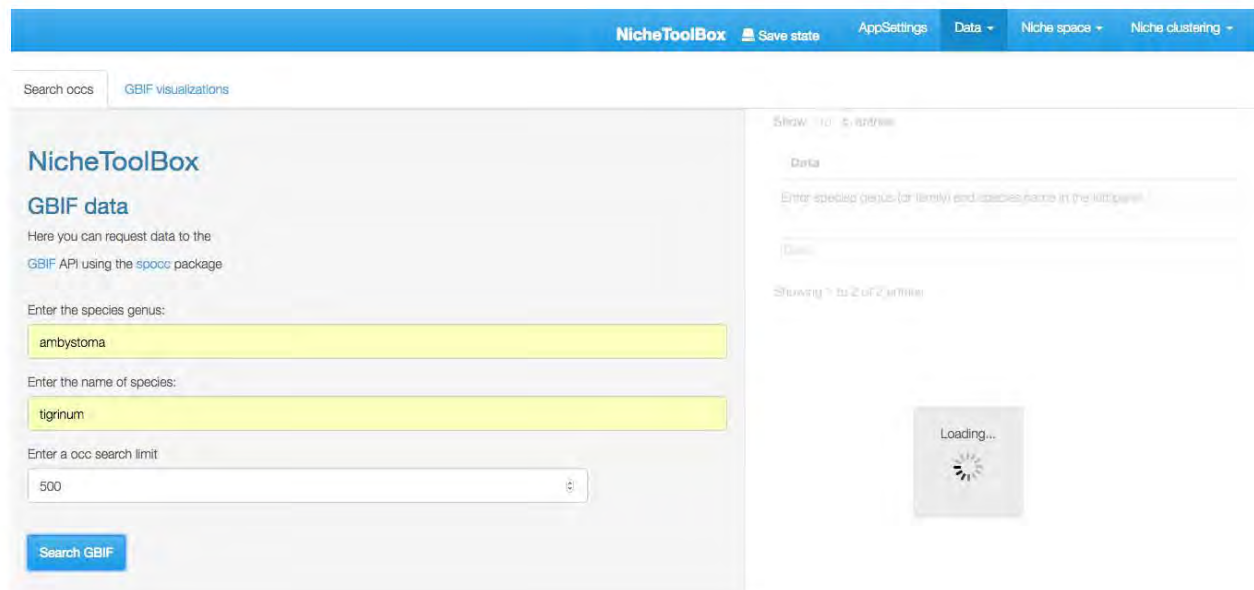


Figure 5:

In the example we searched for the species *Ambystoma tigrinum* which generated 480 records.

GBIF data cleaning

You can remove duplicate records using a separation distance in decimal degrees (default is 0). For *Ambystoma tigrinum* I had **480 records** before cleaning, and after clicking **Clean duplicates** with a $\delta = 0$ distance it remained **154**, so there were **326** duplicate records!!!

NicheToolBox

GBIF data

Here you can request data to the GBIF API using the `spocc` package

Enter the species genus:

Enter the name of species:

Enter a occ search limit:

[Search GBIF](#)

Cleaning duplicates

In this section you can clean your longitude and latitude data using a threshold distance between points to be considered as duplicates.

Select longitude

Select latitude

The data set has

154 rows and 106 columns

Threshold distance for duplicates (δ)

[Clean duplicates](#)

Clean duplicates by group

Sometimes you may want to clean your data on a specific location (country, continent), this option will clean you data given a grouping variable)

Show 10 entries

name	longitude	latitude	issues	prov	key
Ambystoma tigrinum	-88.78191	37.20232	cdround_gas84	gbif	1262379614
Ambystoma tigrinum	-88.92973	37.37986	cdround_gas84	gbif	1265906433
Ambystoma tigrinum	-95.91780	31.98853	cdround_gas84	gbif	1269551639
Ambystoma tigrinum	-96.77353	49.13301	cdround_gas84	gbif	1269570108
Ambystoma tigrinum	-95.89340	31.88686	cdround_gas84	gbif	1269551674
Ambystoma tigrinum	-106.42736	35.84869	cdround_gas84	gbif	1272084151
Ambystoma tigrinum	-92.48258	34.46026	cdround_gas84	gbif	1265535984
Ambystoma tigrinum	-95.88532	31.93524	cdround_gas84	gbif	1088947578
Ambystoma tigrinum	-93.27306	45.21076	cdround_gas84	gbif	1088946504
Ambystoma tigrinum	-88.93064	37.37630	cdround_gas84	gbif	1088971059

Figure 6:

Clean duplicates by group

Suppose that your species has a huge geographic range and you want to work only with the records that match certain criteria, for example records that lie within Canada. You can curate duplicate records using a grouping variable; in this example the grouping variable must be country. Go to **Clean duplicates by group** section and select the grouping variable in this case **country**, then select the country (Canada) and click **Clean duplicates by group**.

Clean duplicates by group

Sometimes you may want to clean your data on a specific location (country, continent), this option will clean you data given a grouping variable)

Select the grouping variable

country

Select the ids you want to clean

Canada

Clean duplicates by group

Download your clean data

Download Data

Figure 7:

From 154 records only 2 are in Canada.

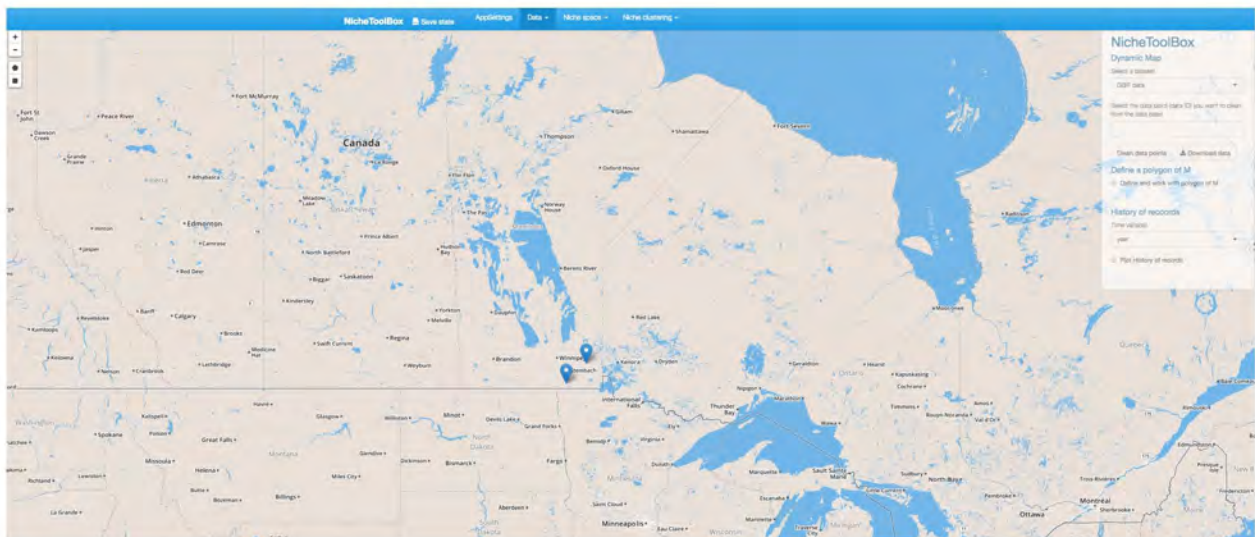
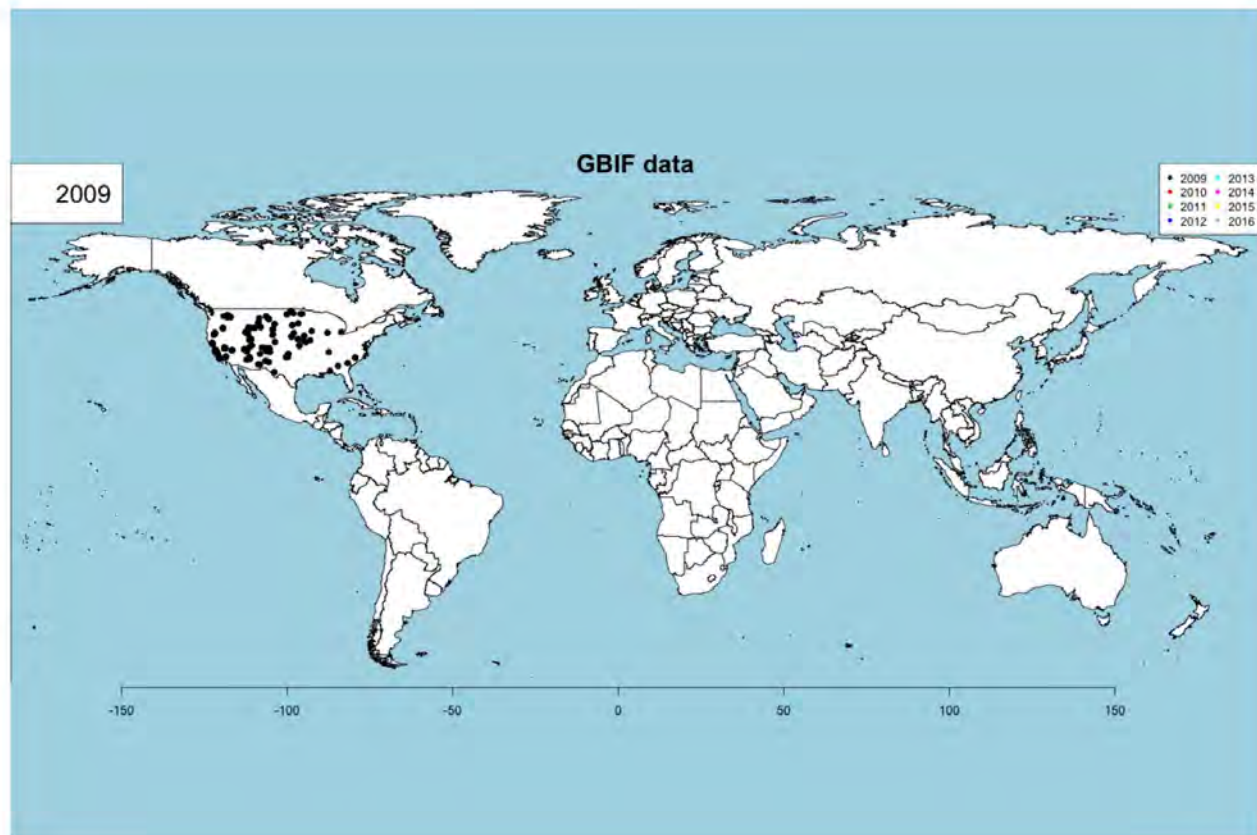


Figure 8:

GBIF visualizations

The GBIF dataset has some fields that can be used to get some exciting visualizations, particularly fields related to observation date (year, month, day) and country. In Data -> GBIF data -> GBIF visualizations tab you can play with interactive plots, create animated visualizations and display a calendar of the reported records by year.



User data

You can use and clean your own latitude and longitude data for the modeling process. Go to Data -> User data and upload your data. The data cleaning process is exactly the same as the GBIF data.

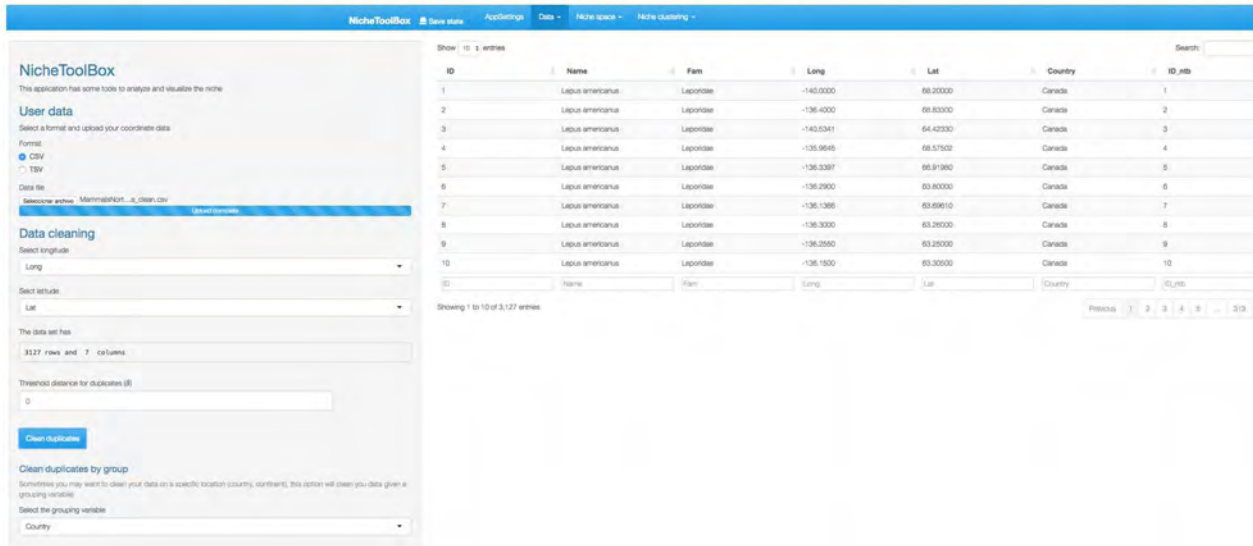


Figure 9:

Geographic explorations using Dynamic Maps

We have seen how to curate data using threshold distances and grouping variables in NicheToolBox. Now let's see how to use leaflet maps: 1) to display longitude/latitude data, 2) to clean data and 3) to define our study area (M data referring to the M concept, which in the niche modeling world is the accessible area where the species has been able to reach even if has not established), 4) Clean data using the M polygon. The above can be done for either the GBIF dataset or the User dataset.

Display longitude and latitude data

Go to Data -> Dynamic Map and on the right panel **Select a dataset** that you want to work with, in this case I will show the work with GBIF data.

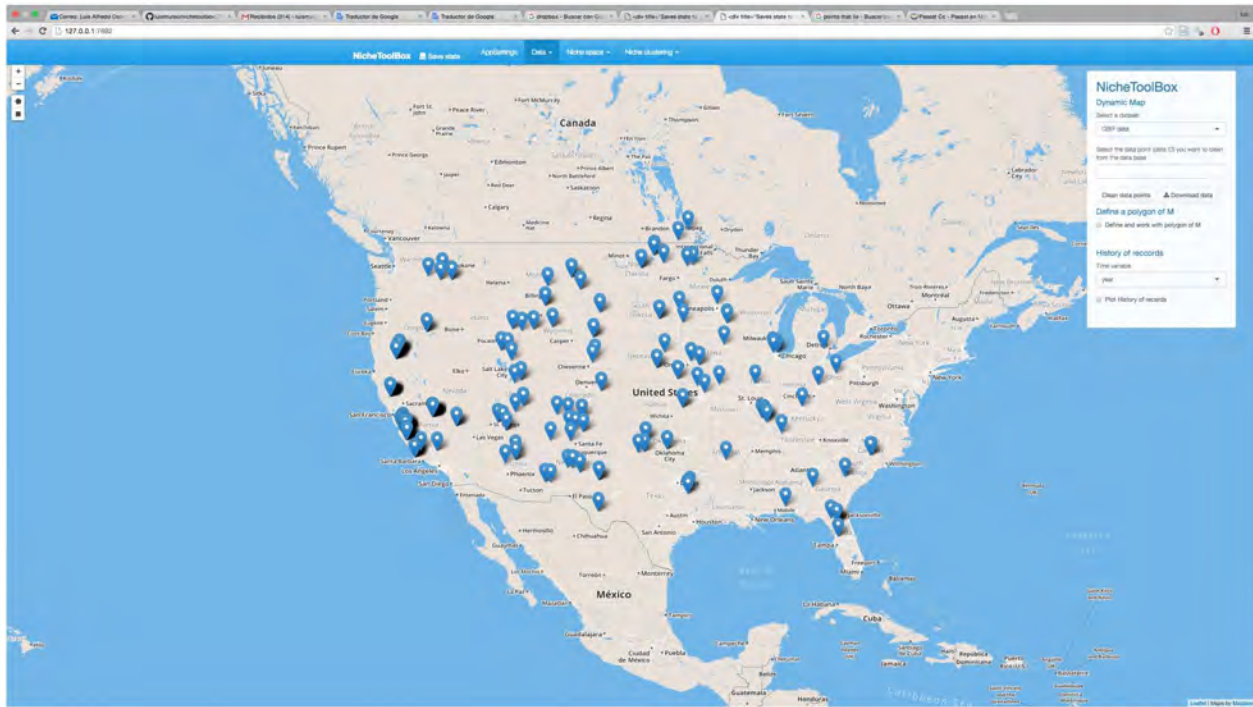


Figure 10:

Data curation using dynamic map

On the right side panel there is an option where you can specify the data point id to clean it from the dataset. Click on pop-up to see the point id, select it in the select input form from the right panel and press **Clean data points** button to clean.

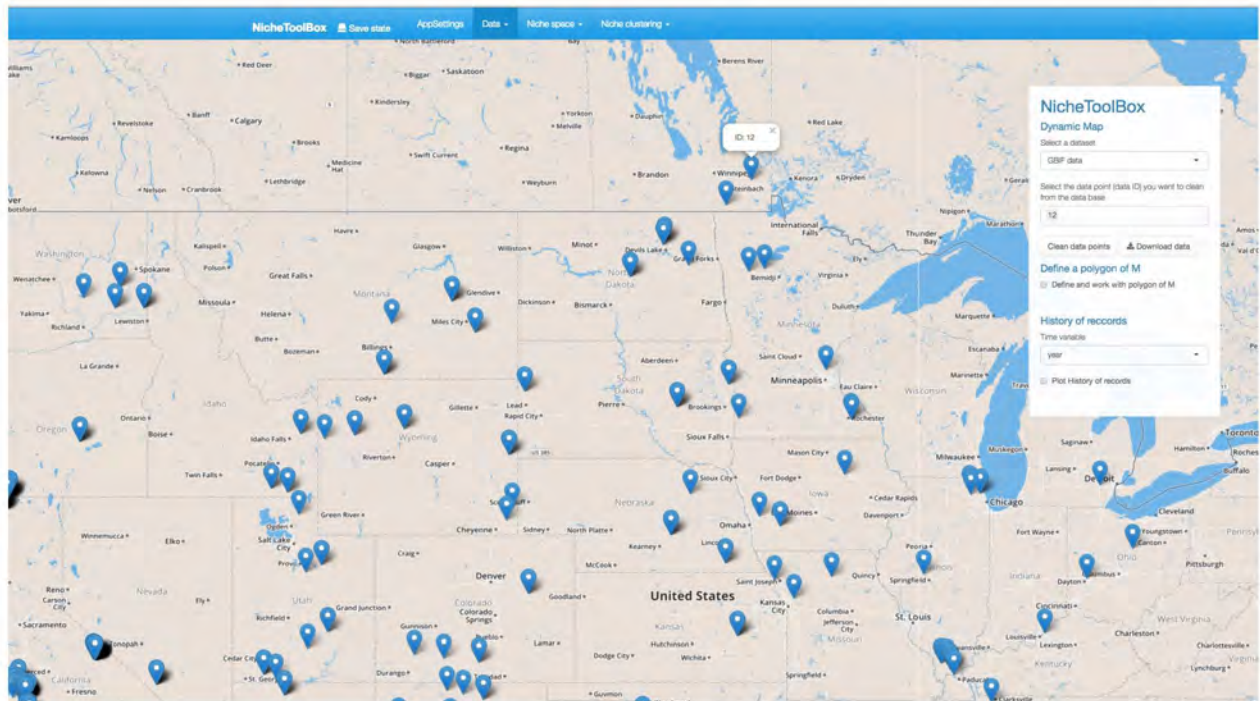


Figure 11:

Define an 'M' map

You can use NicheToolBox to define your study area. Go to Data -> Dynamic Map and in the right-side panel turn-on the button **Define and work with polygon M**, when activated you can either draw a polygon using the drawing tools (topright corner) from NicheToolBox or select a local shapefile. If you prefer to define the M polygon using NicheToolBox press the polygon tool and draw it:

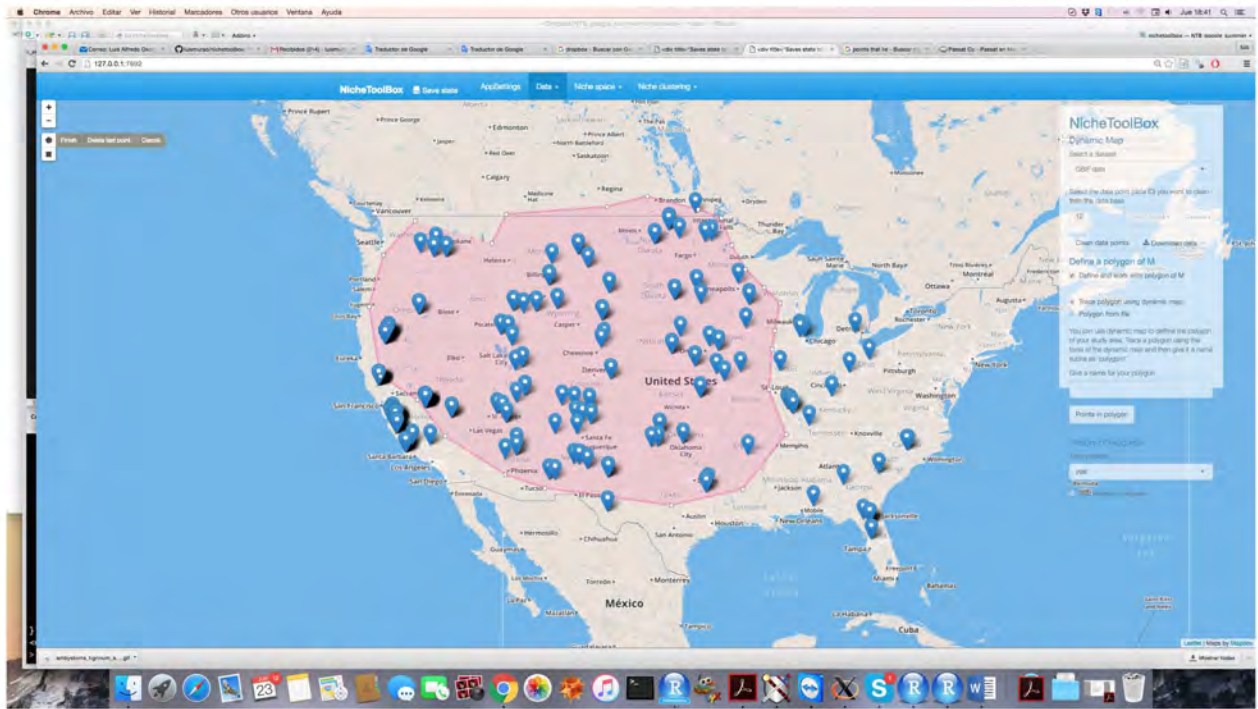


Figure 12:

Once defined, the polygon can be saved. In the right panel there is a form where you can give a name for your polygon.

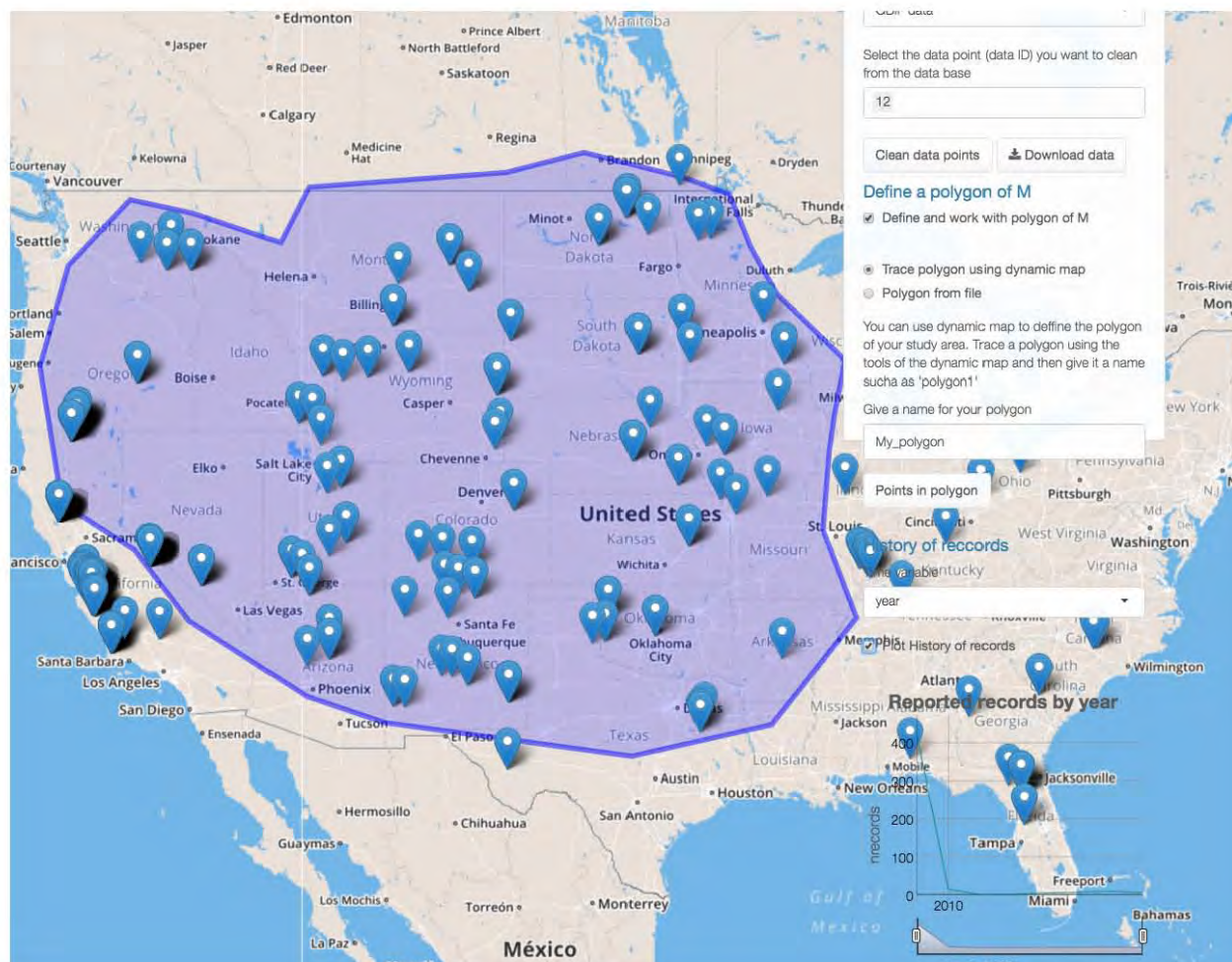


Figure 13:

Data curation using the M polygon

We can filter the data points that lie inside the polygon. In the right panel just press the button **Points in polygon**

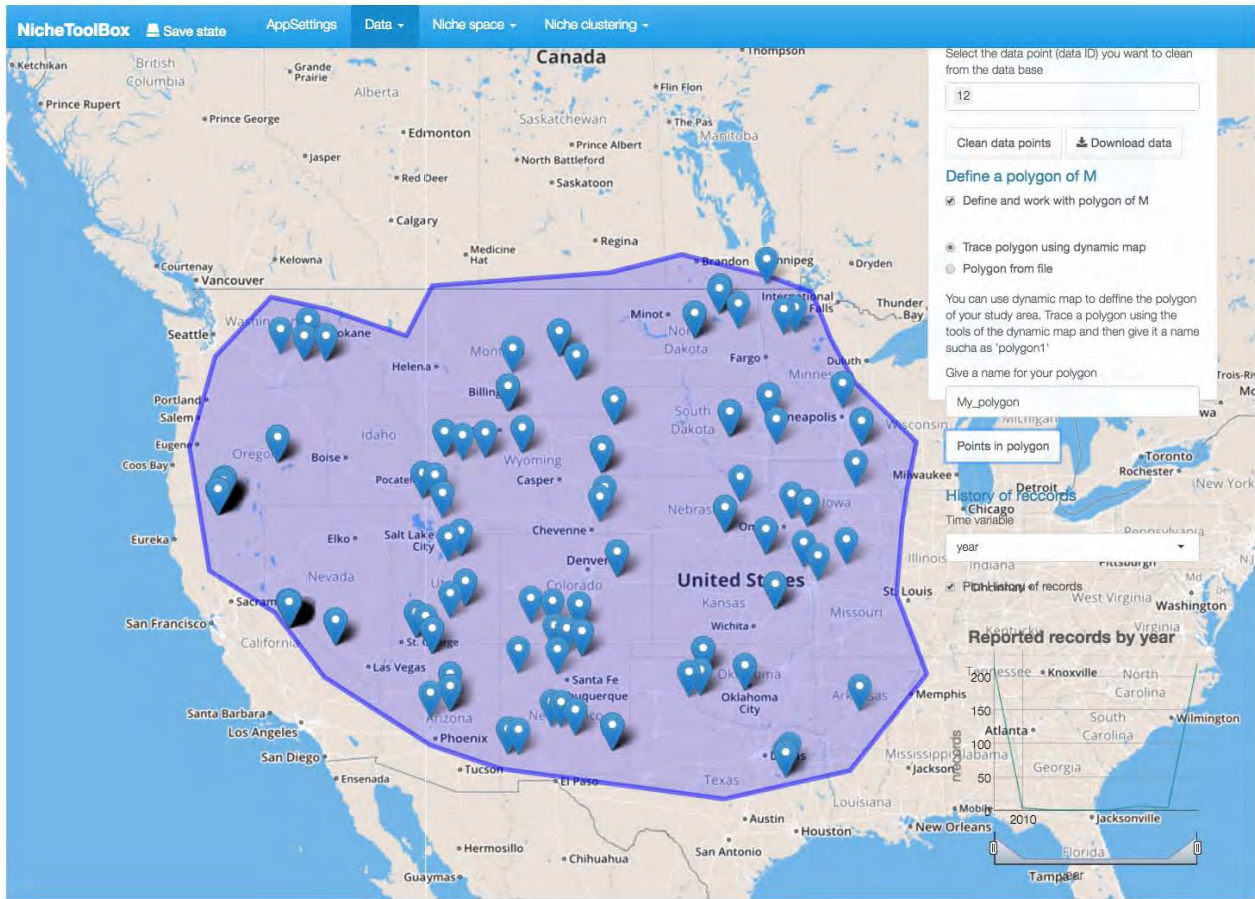


Figure 14:

Saving the workflow

Once specified the workflow directory (AppSettings “go to the first section of the tutorial”) which is the directory where all the information generated in the app is stored, we just need to press the **Save state** button in order to save everything!! (in this stage only the geographic data related work).

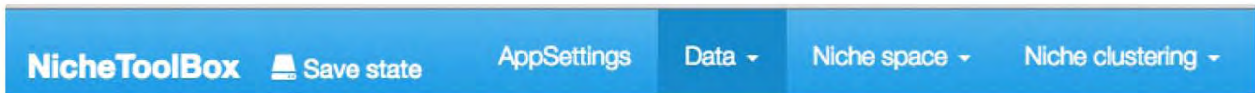


Figure 15:

The workflow report

One of the files generated when you press **Save state** is the **data_report.html** which is an html file with a summary of the geographic data-related information that you have done with NicheToolBox.

NicheToolBox data report

- GBIF data
 - GBIF visualizations
- User data
- Dynamic Map

This is **NicheToolBox data report** here you can find a quick view of the things that you have done in the software. Remember that this is free software so it comes with no warranty, please report any bugs you find at luismurao@comunidad.unam.mx

GBIF data

The total number of occurrences points were: **480**

```
occ_data <- data_gbif()
kable(head(occ_data[,c(input$LongitudeGBIF,
                      input$LatitudeGBIF,
                      "country", "countryCode")]), digits=4)
```

	longitude	latitude	country	countryCode
2	-88.7819	37.2023	United States	US
4	-88.9297	37.3799	United States	US
6	-95.9178	31.9885	United States	US
7	-96.7735	49.1330	Canada	CA
8	-95.8934	31.8869	United States	US
9	-106.4274	35.8489	United States	US

GBIF visualizations

Pie chart of occs by country (all years)

```
plot(GBIF_vis())$pie)
```



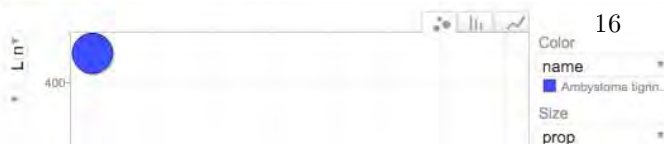
Pie chart of occs by country (year 2009)

```
plot(calendar())
```



Motion chart of records reported by year

```
plot(GBIF_vis())$motion)
```



Niche space (steps to make a cluster analysis)

To work in **Niche space** we need to have loaded our **niche raster layers** (AppSettings “go to the first section of the tutorial”) and have a **longitude/latitude dataset** (GBIF data or User data).

1. Extracting niche values from raster layers

Go to Niche space -> Niche data extraction and select a longitude and latitude dataset. In the example I selected the **GBIF dataset**. If the dataset is not empty and we have loaded the **raster layers** the app will not show any message:

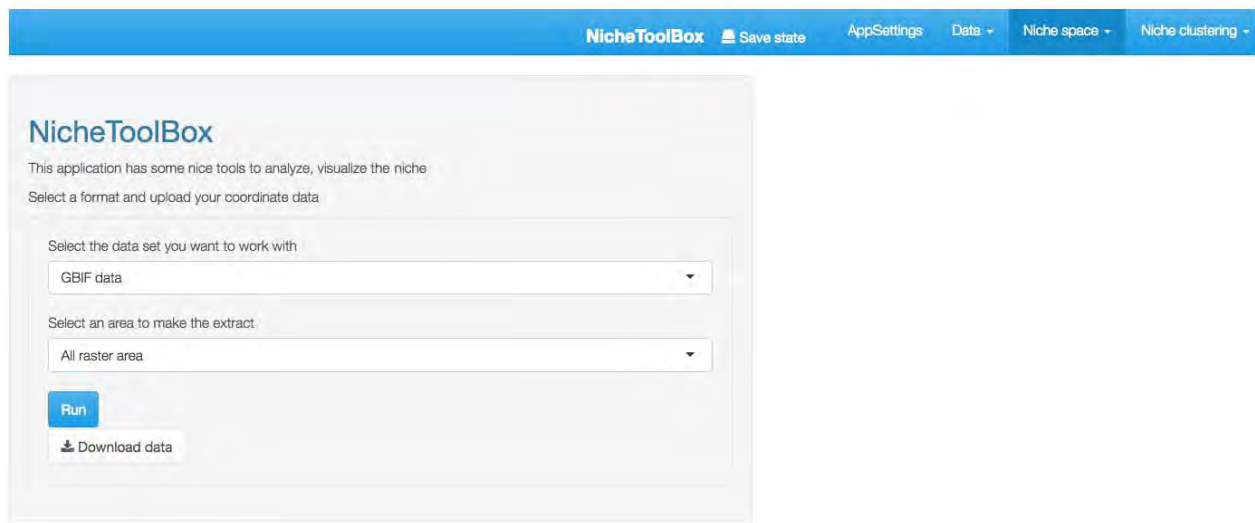


Figure 16:

On the contrary, if we have not loaded either the **raster layers** or the **longitude/latitude data** a messages indicating what to do will be displayed.

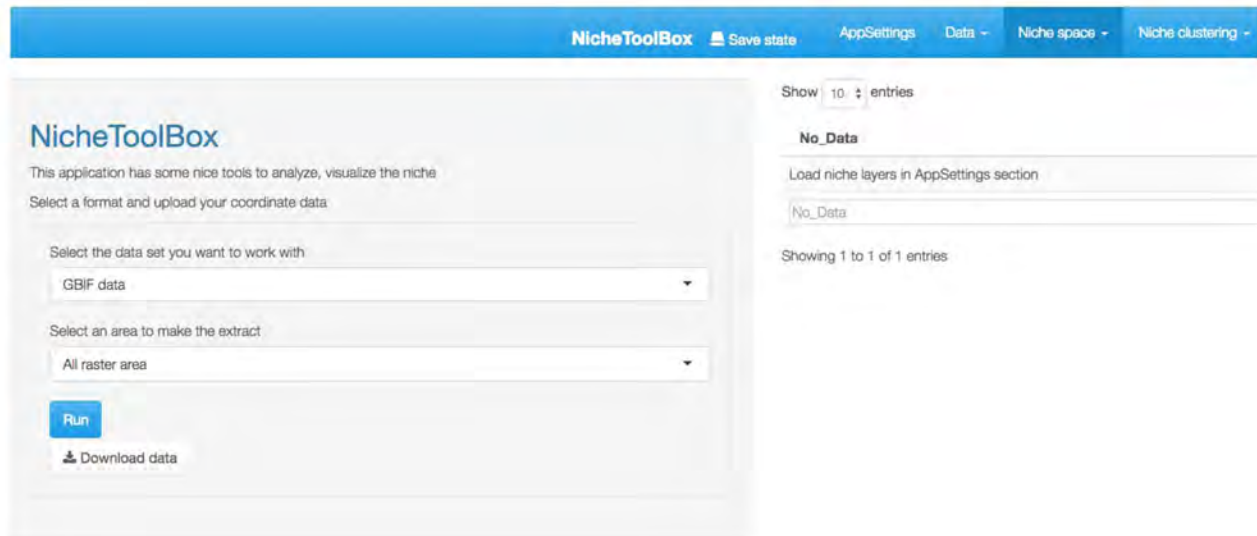


Figure 17:

When the dataset and the layers are in the App memory we can proceed to the next step. Here you just need to press the **Run** button and then a data table with the niche values of our longitude and latitude data will be displayed.



2. Niche explorations

We can explore our niche data using some exciting 3-Dimensional plots. Go to Niche space -> Known niche and play with x , y and z variables of the ellipsoid plot.

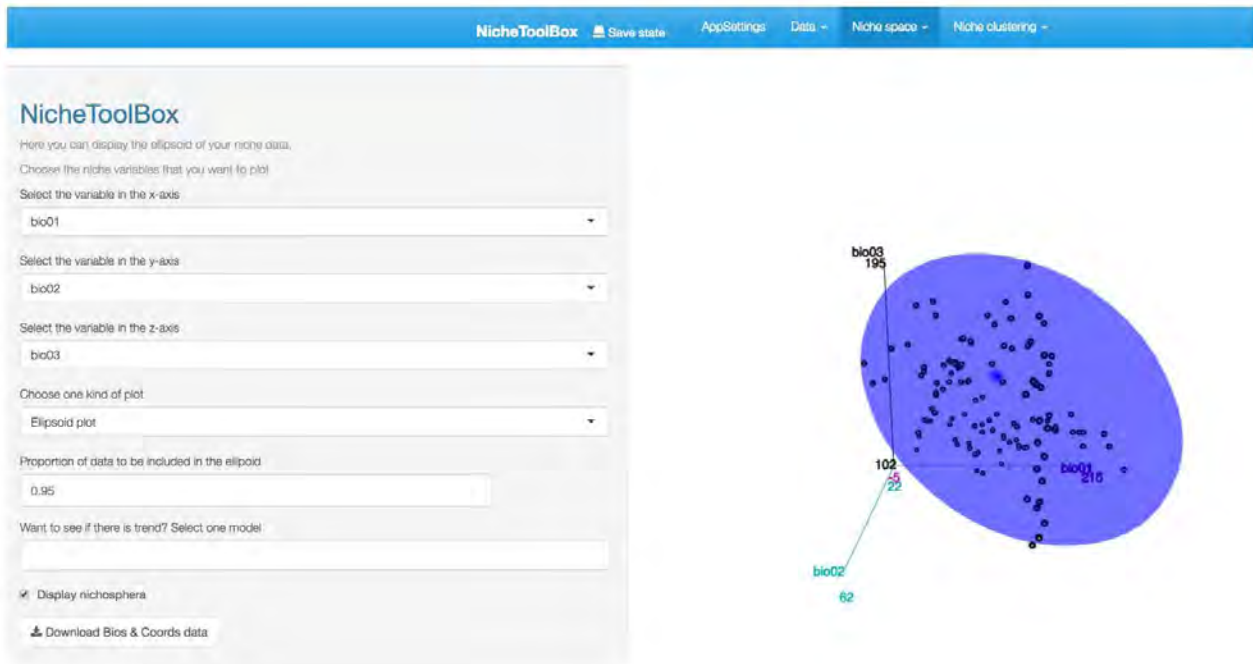


Figure 19:

2.1 Niche trends

You can fit a (linear, quadratic, additive, smooth) model to see if your niche data have a trend.

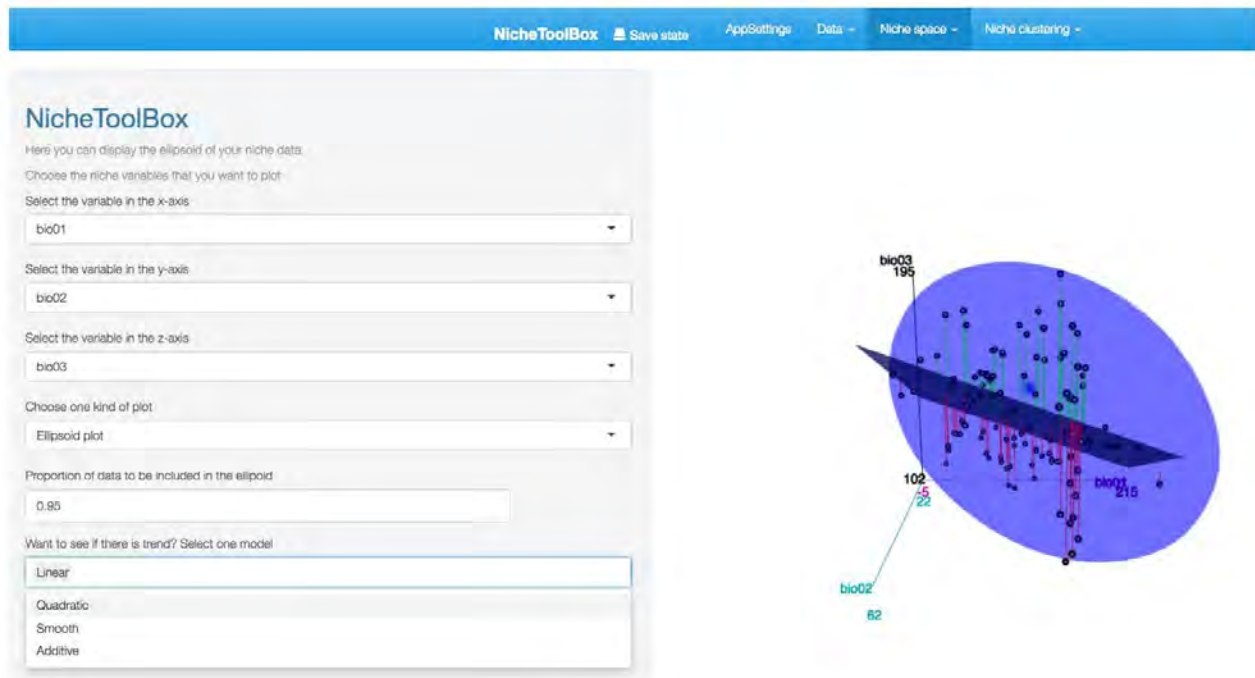


Figure 20:

3. Niche clustering

When studying species niches and distributions, one of the biggest questions that comes to my mind is whether or not the species are adapting to different niche conditions. One way to explore this question is using clustering algorithms (a statistical tool which aims to observe if a multivariate data have a cluster structure in such a way that the data belonging to the same cluster are highly similar between them and different with respect to other groups). If the clusters are different, we can think that populations of the same species are responding in different ways to the same set of niche variables (they are adapting to local conditions). This is just an exploratory tool.

Go to Niche clustering -> K-means section and select at least 3 niche variables to make the cluster analysis. In my case, as I selected the bios of the WorldClim database as my niche layers, I used 19 niche variables, but if you want to work with few variables just delate some of them (**Select at least 3 niche variables** section).

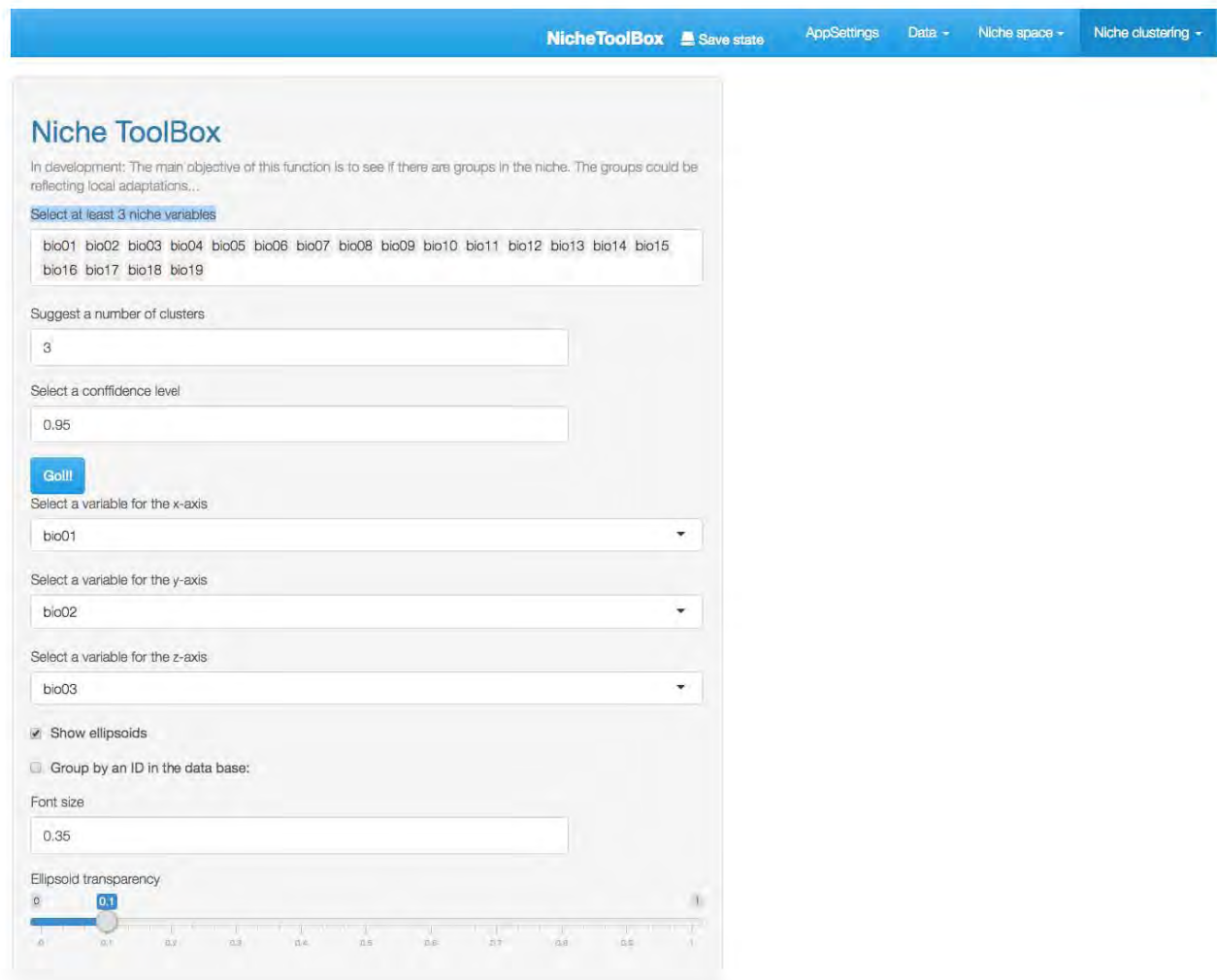


Figure 21:

Here it is necessary to indicate a number of clusters, the default value is 3 (in the future the app will have algorithms to help you to make this decision). Press the **Go!!!** button and you will see a 3 dimensional plot with ellipsoids representing the number of clusters you suggested. Below this plot you will see a **leaflet** map with the geographic projection of the points that fall inside each ellipsoid (colors help to identify to which cluster each data point belongs).

Niche ToolBox

In development. The main objective of this function is to see if there are groups in the niche. The groups could be reflecting local adaptations...

Select at least 3 niche variables

bx07 bx02 bx03 bx04 bx05 bx06 bx07 bx08 bx09 bx10 bx11 bx12 bx13 bx14 bx15
bx16 bx17 bx18 bx19

Suggest a number of clusters

3

Select a confidence level

0.95

Go!!

Select a variable for the x-axis

bx07

Select a variable for the y-axis

bx02

Select a variable for the z-axis

bx03

Show ellipsoids

Group by an ID in the data base

Font size

0.35

Ellipsoid transparency

0.7

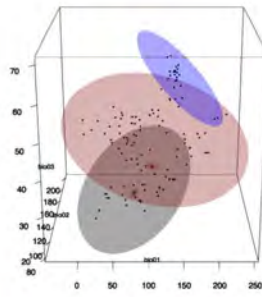


Figure 22:

Let's play with the number of cluster (now 4) and see how the results change...

Selecting local adaptations...

Select at least 3 niche variables

bx07 bx02 bx03 bx04 bx05 bx06 bx07 bx08 bx09 bx10 bx11 bx12 bx13 bx14 bx15
bx16 bx17 bx18 bx19

Suggest a number of clusters

4

Select a confidence level

0.95

Go!!

Select a variable for the x-axis

bx01

Select a variable for the y-axis

bx02

Select a variable for the z-axis

bx03

Show ellipsoids

Group by an ID in the data base

Font size

0.35

Ellipsoid transparency

0.7

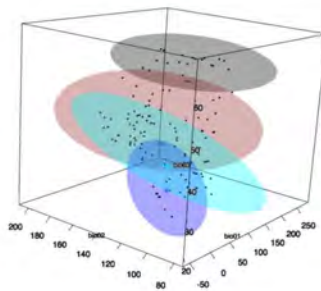


Figure 23:

4. Niche correlations

One popular method to select the niche variables for modeling species niches and distributions is to study correlations among niche variables and filter those variables that are highly correlated. In *NicheToolbox* you can filter the variables that summarize the environmental information of your presences data according to a correlation threshold; this algorithm suggest which variables to use for the modeling part.



Figure 24:

Also, you can explore the correlation matrix and download it in .csv format

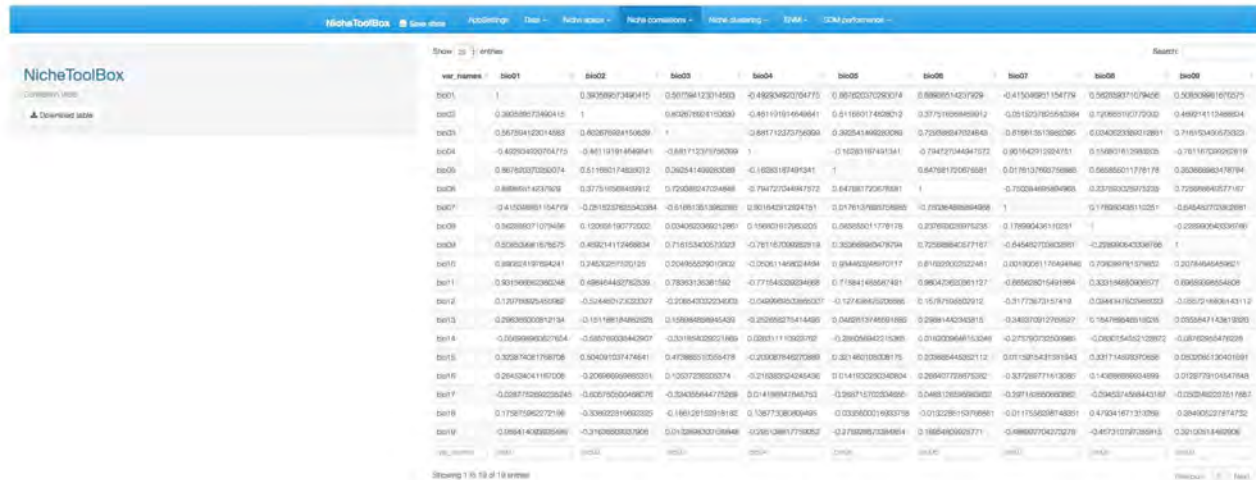


Figure 25:

Another thing that the user can do is to plot a correlogram.

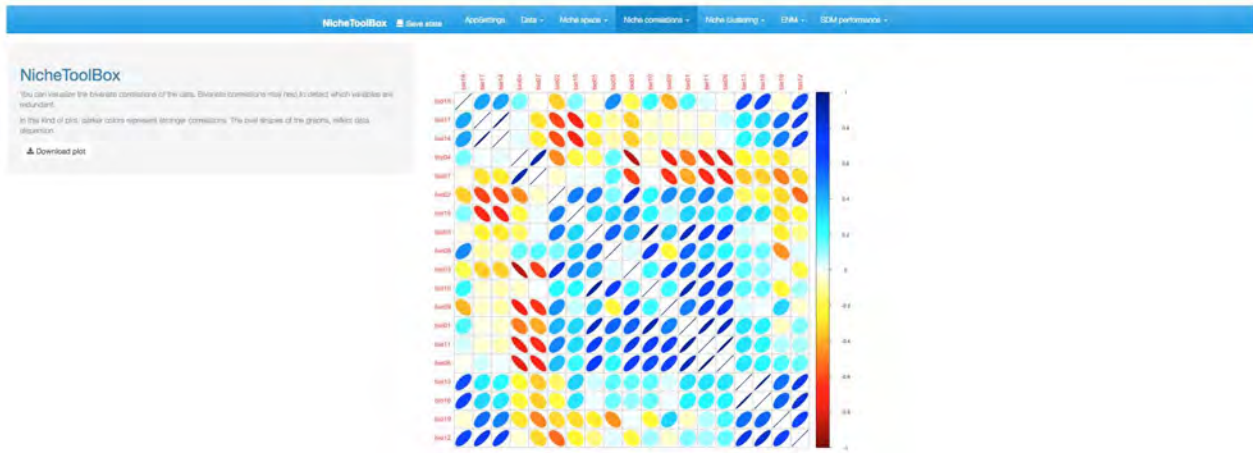


Figure 26:

6. Ecological niche modeling

Ecological Niche Modeling (ENM), is a growing field of ecology and biogeography which aims to reconstruct the multidimensional ecological niche of species and from here to approximate its geographical distribution. ENM uses a set of mathematical and statistical tools to study the relationship between some environmental variables and species occurrences to estimate species niches and predict potential areas where the species can survive. These models have proved useful in ecology and conservation biology because they have been used to identify geographic localities that can be used to relocate endangered species, to study the impacts of climate change in biodiversity, to find biodiversity hotspots, vulnerability to invasive species and pathogens, among other applications (Peterson & Vieglais 2001; Peterson et al 2011).

In `Nichetoolbox` you can model ecological niches by using one of following modeling algorithms:

- 1) Minimum volume ellipsoid
- 2) Bioclim
- 3) MaxEnt

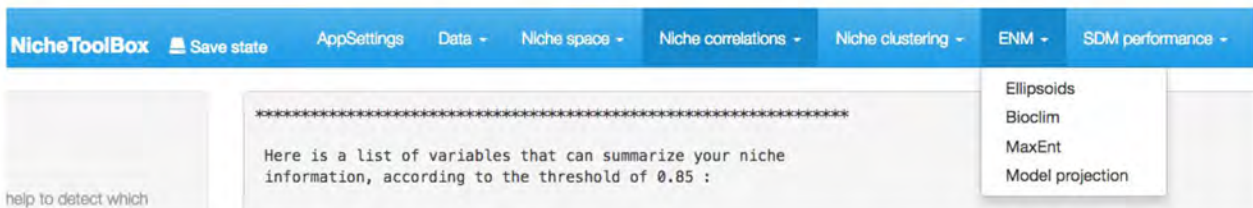


Figure 27:

6.1 Minimum volume ellipsoid model

Ellipsoid models use the multinormal probability density function (PDF; equation 1) to compute the niche suitability index; the PDF is rescaled in order to have a suitability index defined in the interval $[0, 1]$.

$$f(x_1, x_2, x_3, \dots, x_k) = \frac{1}{(2\pi)^k |\Sigma|} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})\right) \quad (1)$$

$$f(x_1, x_2, x_3, \dots, x_k) = \frac{1}{|\Sigma|} \exp\left(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})\right)$$

where \mathbf{x} is the vector of environmental variables such that each x_i represents an observation of the environmental variable i . Σ is the covariance matrix of the occurrence data. $\boldsymbol{\mu}$ is the vector of means (centroids).

The $(\mathbf{x} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu})$ is the square of the Mahalanobis distance.

In `NicheToolbox`, to make an ellipsoid model you just need the environmental information of your occurrence points and select which layers you want to model the niche.

The model can be trained either with all occurrence data or with the occurrence points that lie inside your polygon of M.

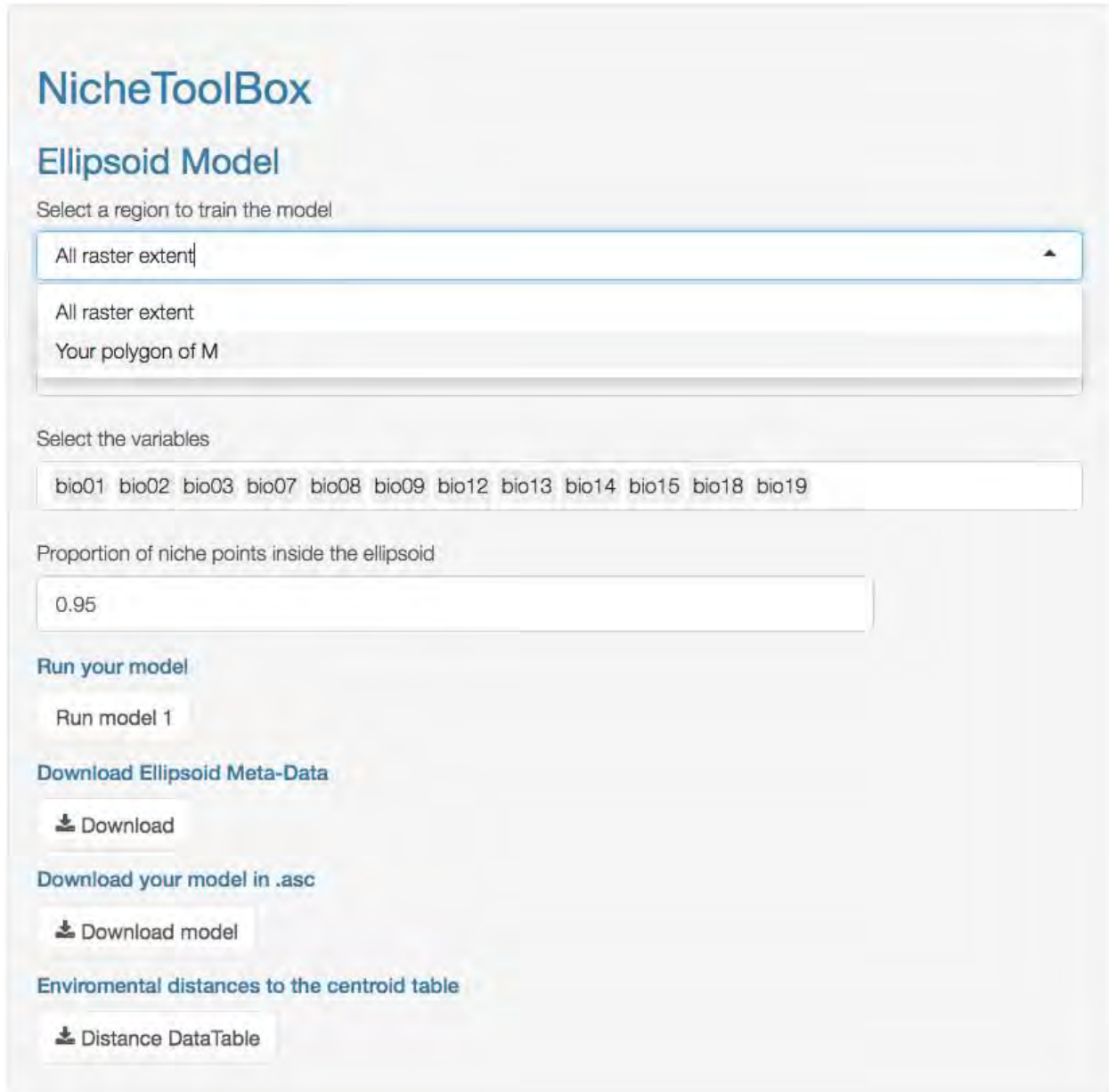


Figure 28:

Similarly, you can project the model to the geography by using either the full extent of rasters or the extent of the 'M' polygon.

Using full extent

NicheToolBox

Ellipsoid Model

Select a region to train the model

All raster extent

Select a region to project the ellipsoid

All raster extent

All raster extent

Your polygon of M

Proportion of niche points inside the ellipsoid

0.95

Run your model

Run model 1

Download Ellipsoid Meta-Data

Download

Download your model in .asc

Download model

Enviromental distances to the centroid table

Distance DataTable

Figure 29:

Select the niche variables and run your model.

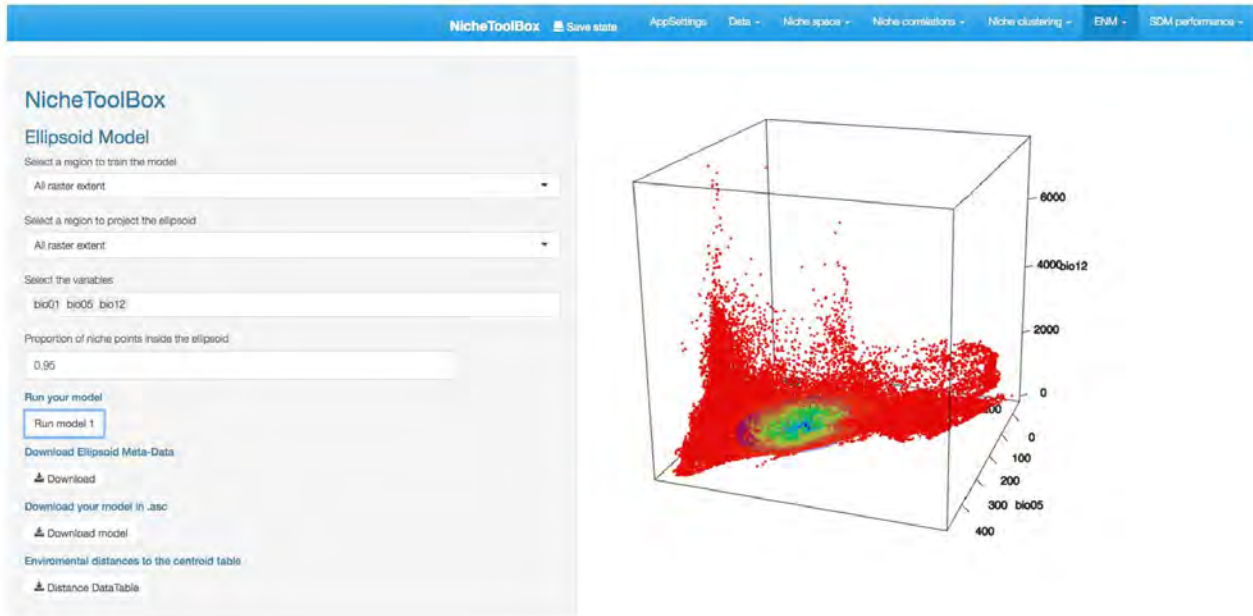


Figure 30:

Using the extent of the 'M' polygon

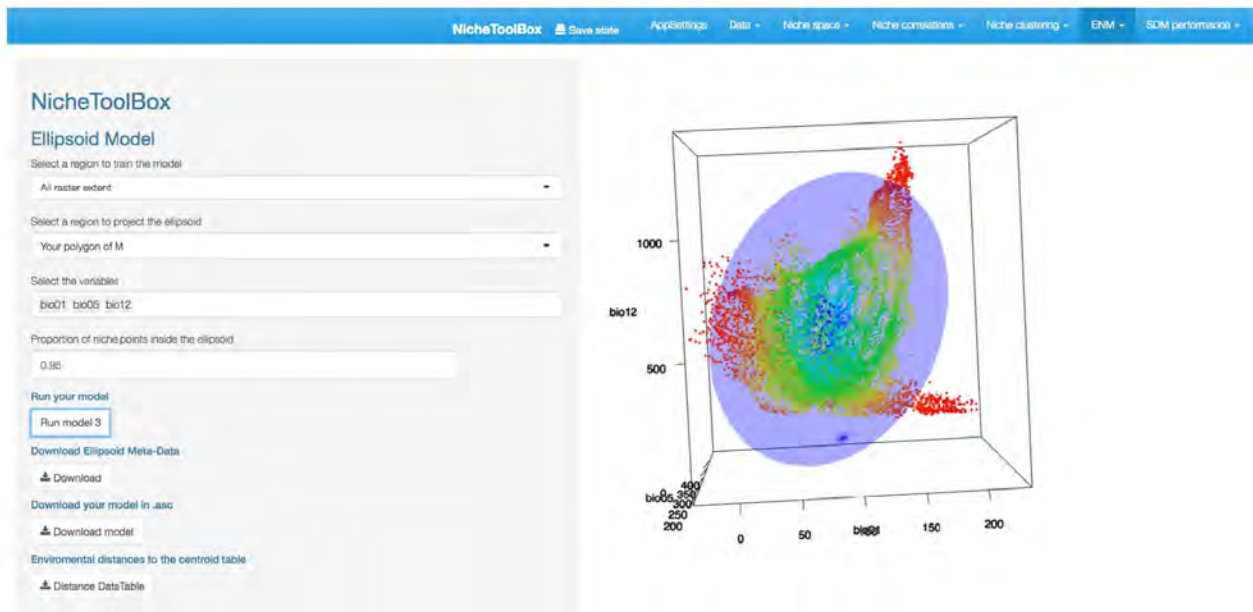


Figure 31:

Download ellipsoid metadata

```
EllipsoidMetaData_wWorld_var_matrix_wWorld (2).txt - Modificatio
$centroid
  bio01    bio05    bio12
89.02299 293.98851 606.36782

$covariance
      bio01    bio05    bio12
bio01 2066.697 1581.535 2076.759
bio05 1581.535 1563.593 -1286.984
bio12 2076.759 -1286.984 75748.072

$data
  bio01 bio05 bio12
1      28   259   338
2      82   318   399
3      52   291   443
4      68   260   806
5      51   274   793
6     117   311   964
7     170   318   479
8      23   225   464
9     105   335   177
10     48   277   276
11     39   226   759

90     63   286   709

$niche_volume
[1] 6418938

$axis_length
      a      b      c
550.73529 116.77468 23.82775
```

Figure 32:

Download ellipsoid raster model

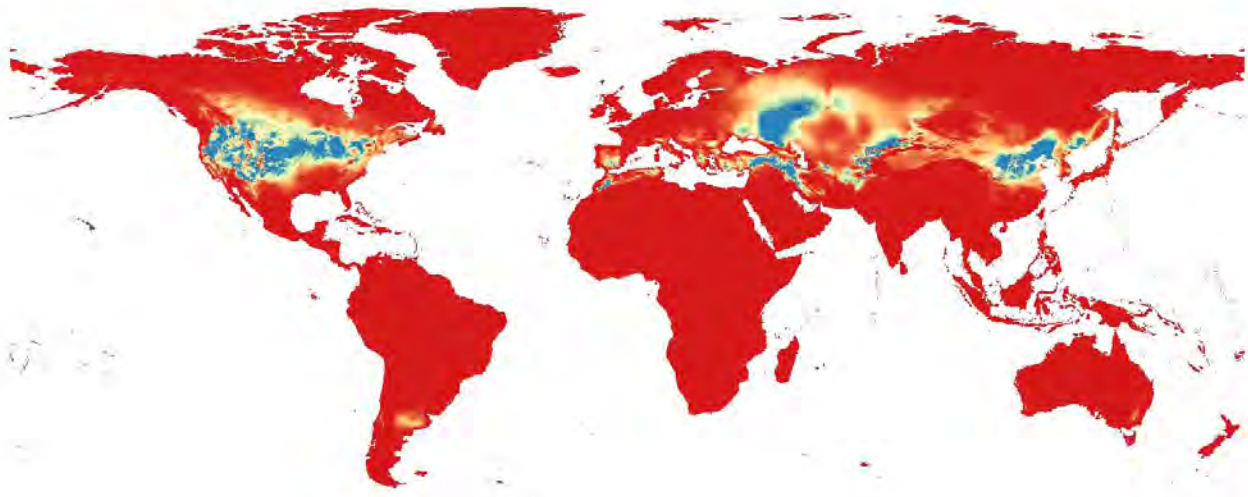


Figure 33:

Download distance to the centroid table

x	y	suitability	bio01	bio05	bio12	ncel	mahalanobis	ecucleidean
-37.916667	83.583333	0	-174	37	144	1	50.707492	982.37931
-37.75	83.583333	0	-174	37	143	2	50.733446	983.37931
-36.916667	83.583333	0	-172	39	136	3	50.214163	986.37931
-36.75	83.583333	0	-173	39	136	4	50.041259	987.37931
-36.583333	83.583333	0	-173	39	136	5	50.041259	987.37931
-36.416667	83.583333	0	-173	38	137	6	50.538349	987.37931
-36.25	83.583333	0	-173	38	137	7	50.538349	987.37931
-36.083333	83.583333	0	-173	38	137	8	50.538349	987.37931
-35.916667	83.583333	0	-173	38	137	9	50.538349	987.37931
-35.75	83.583333	0	-173	38	137	10	50.538349	987.37931
-35.583333	83.583333	0	-173	38	138	11	50.512192	986.37931
-35.416667	83.583333	0	-173	38	138	12	50.512192	986.37931
-35.25	83.583333	0	-173	38	138	13	50.512192	986.37931
-35.083333	83.583333	0	-174	38	138	14	50.339554	987.37931
-34.916667	83.583333	0	-174	37	138	15	50.863815	988.37931
-34.75	83.583333	0	-176	35	144	16	51.414243	986.37931
-34.583333	83.583333	0	-178	34	150	17	51.443625	983.37931

Figure 34:

6.2 Bioclim model

The way that bioclim model is implemented in `Nichetoolbox` is the same as the ellipsoid model:

- The model can be trained either with all occurrence data or with occurrence points that lie inside your 'M' polygon.

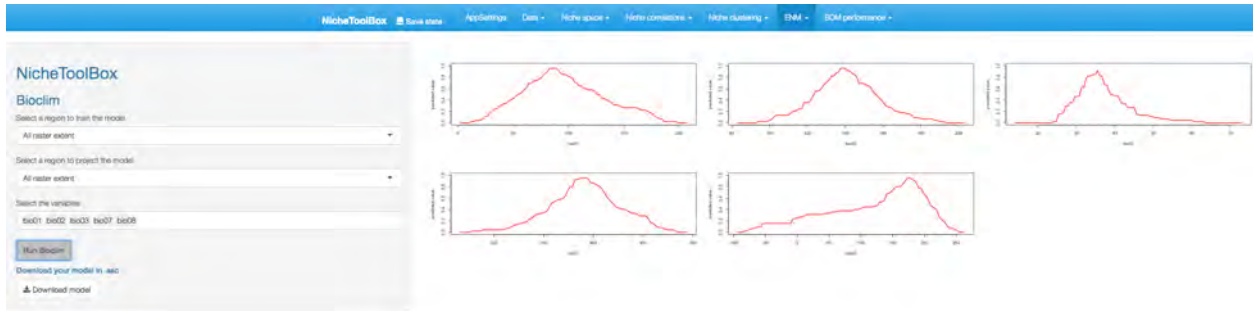


Figure 35:

- Similarly, you can project the model to the geography using either the full extent of rasters or the extent of the 'M' polygon.



Figure 36:

6.3 MaxEnt model

You can run MaxEnt within Nichetoolbox. Nichetoolbox call the Maxent function from dismo package. In order to use MaxEnt within Nichetoolbox you need to install rJava and paste the .jar file of maxent in the java folder of dismo. To test if maxent is aviable run the following comand:

```
jar <- paste(system.file(package="dismo"), "/java/maxent.jar", sep='')
# Ask if necessary files are in java folder of dismo
file.exists(jar)
```

```
## [1] TRUE
```

```
# test if rJava is installed  
"rJava" %in% installed.packages()
```

```
## [1] TRUE
```

If everything is fine, you can make Maxent models within `NicheToolbox` with your own data or the data that you have downloaded from GBIF and by choosing between full raster extent or the 'M' extent.

Most of MaxEnt features and settings are implemented in the app

Main features

MaxEnt

Main features

Select a dataset to train the model

User data

Select a region to project the model

M raster extent

Select the variables

bio01 bio02 bio03 bio07 bio08 bio09 bio12 bio13 bio14 bio15 bio18 bio19

Features

- Linear features
- Quadratic features
- Product features
- Threshold features
- Hinge features
- Logscale pictures
- Do Jackknife
- Create response curves

Output format

Logistic

Run

Model

Figure 37:

Basic settings

Settings

Basic

Remove duplicate presence records

Random test percentage

0

Regularization multiplier

1

Max number of background points

10000

Replicates

1

Replicated run type

Crossvalidate

Figure 38:

Advanced settings

Advanced

Add samples to background

Add all samples to background

Extrapolate

Number of iterations

Convergence threshold

Prevalence

Figure 39:

Experimental settings

Experimental

Lq to Lqp threshold

Linear to lc threshold

Hinge threshold

Beta threshold

Beta categorical

Beta lqp

Beta hinge

Figure 40:

Runing maxent

Once you have configured your Maxent settings press the run button. A window with the basic statistics of maxent will be displayed

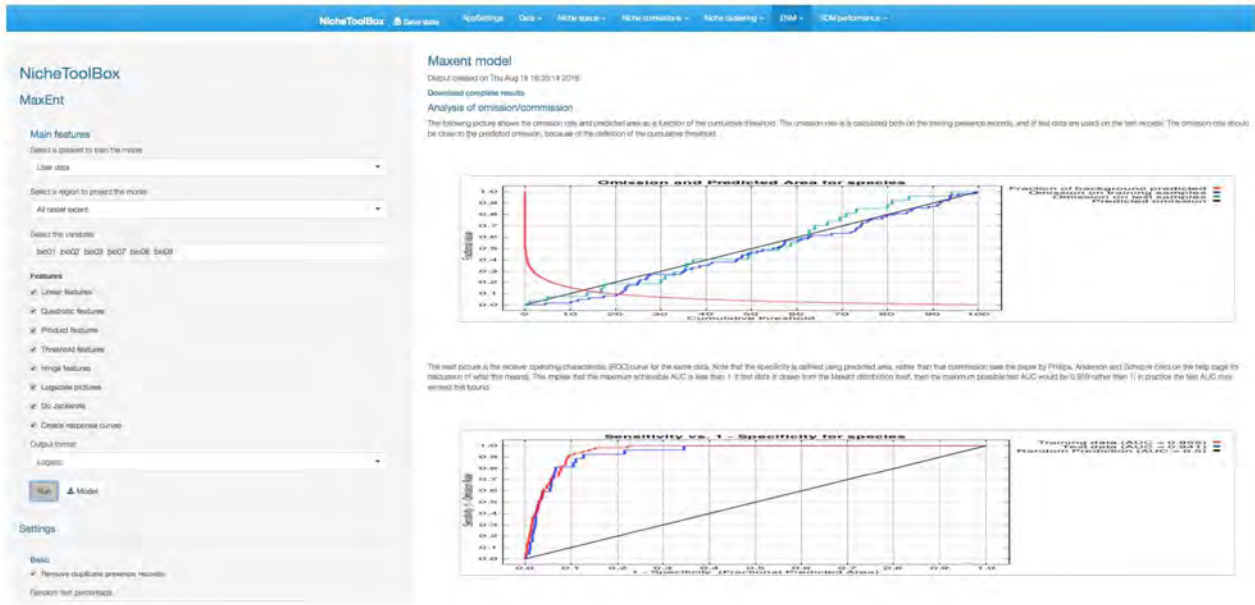


Figure 41:

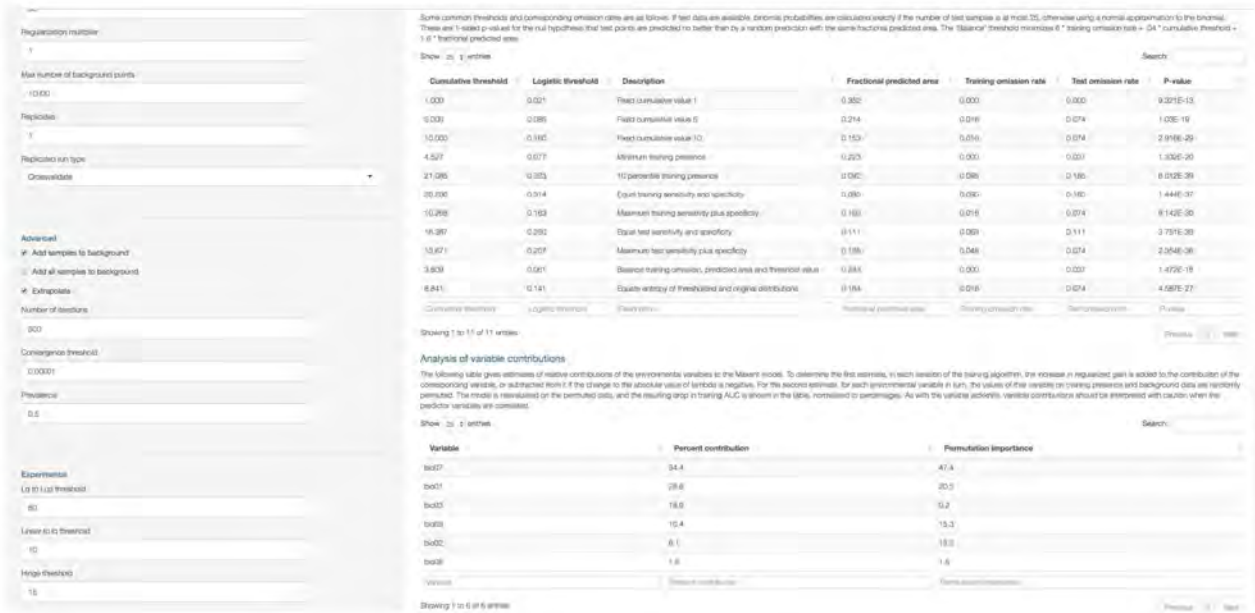


Figure 42:

Download Maxent results

To download maxent results click on Download complete results link



Figure 43:

Download Maxent raster model

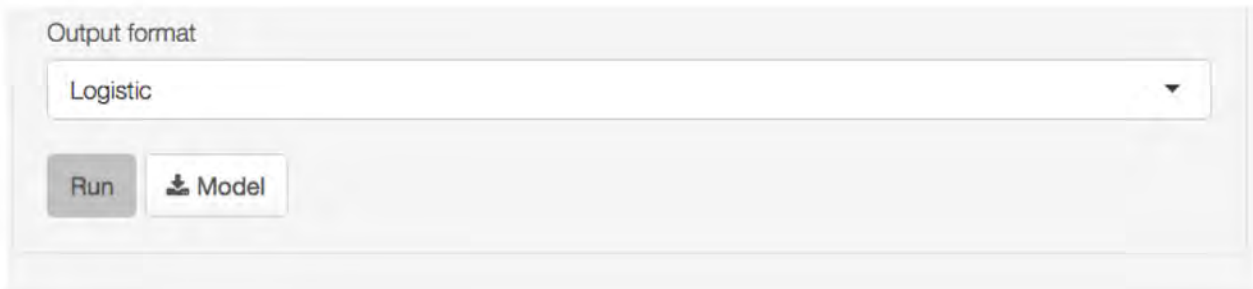


Figure 44:

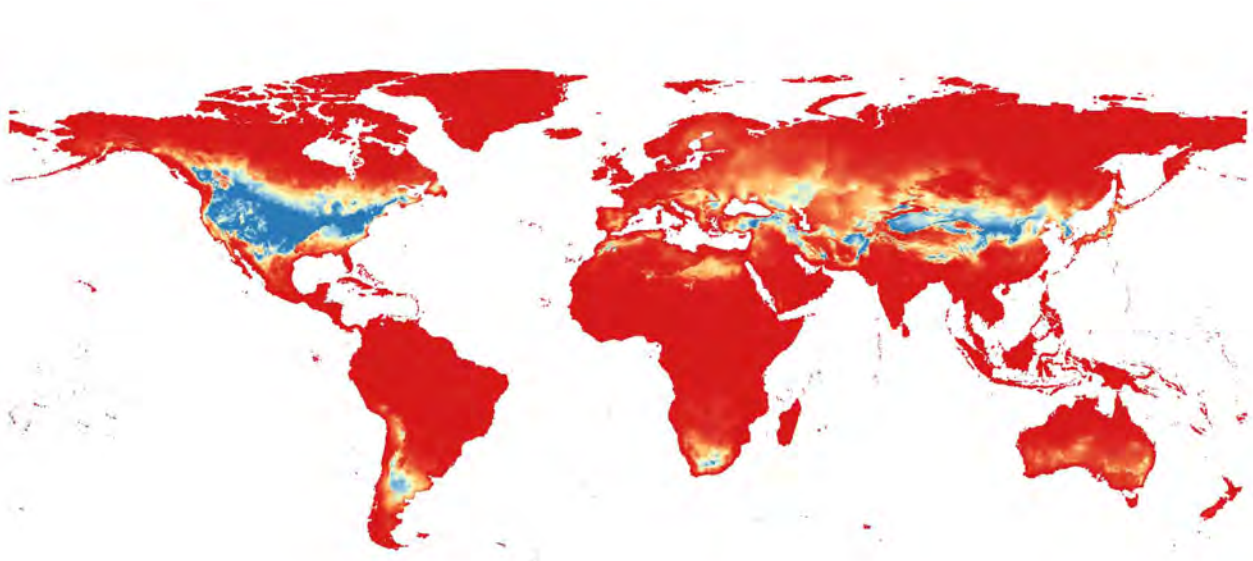


Figure 45:

ENM projection in geographic space

Once you have modeled your species' niche using one or all modeling algorithms, you can explore them in geographic space by using the model visualizer. The visualizer is interactive (you can zoom in/out a map) and uses the `leaflet` library.

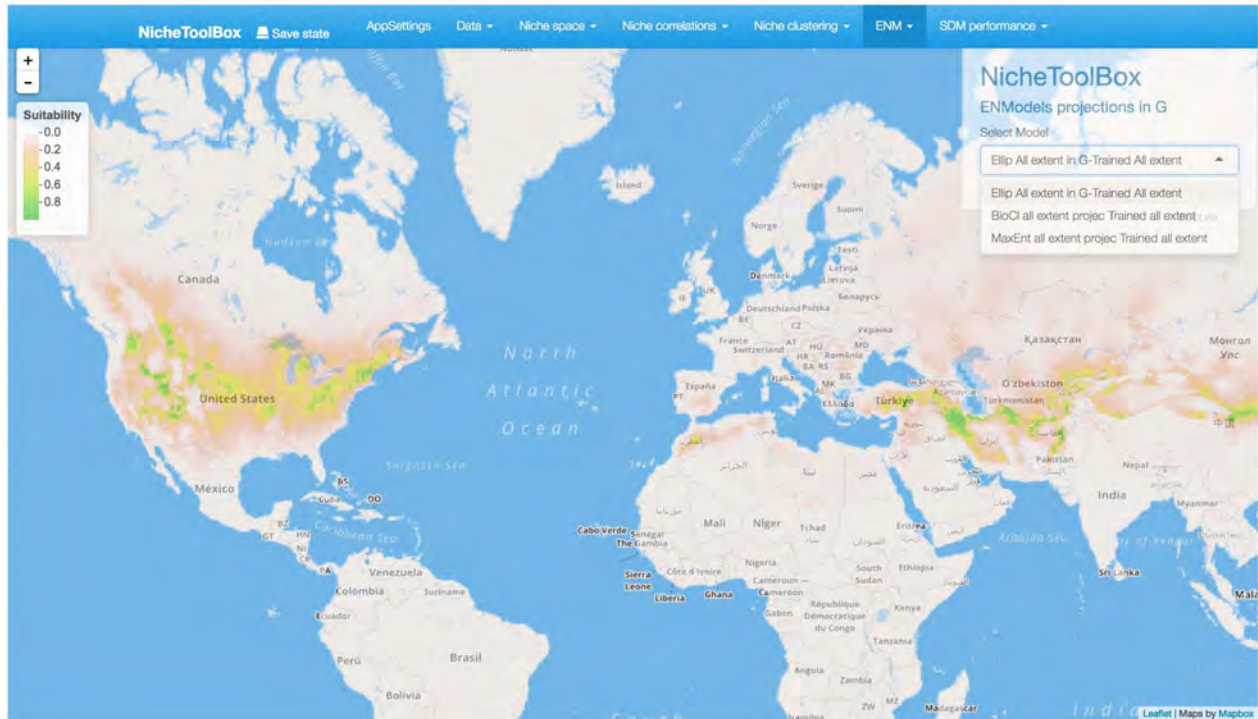


Figure 46:

7. Species distribution model performance

The last part of the project deals with species distribution model evaluation and performance. `Nichetoolbox` has two ways to evaluate models:

- 1) Partial Roc: uses the function implemented on `ENMGadgets` package that does Partial Roc (Peterson et al. 2008).
- 2) Confusion matrix metrics: You can compute prevalence, specificity, sensitivity, TSS, Kappa, correct classification rate, misclassification rate, negative predictive power, positive predictive power, omission error fraction, commission error fraction, false negative rate, and false positive rate from the confusion metrics (Fielding and Bell, 1997).

7.1 Partial ROC

To do Partial ROC analysis in `Nichetoolbox` upload your continuous niche model aoutput map (e.g., Maxent) and your validation dataset.

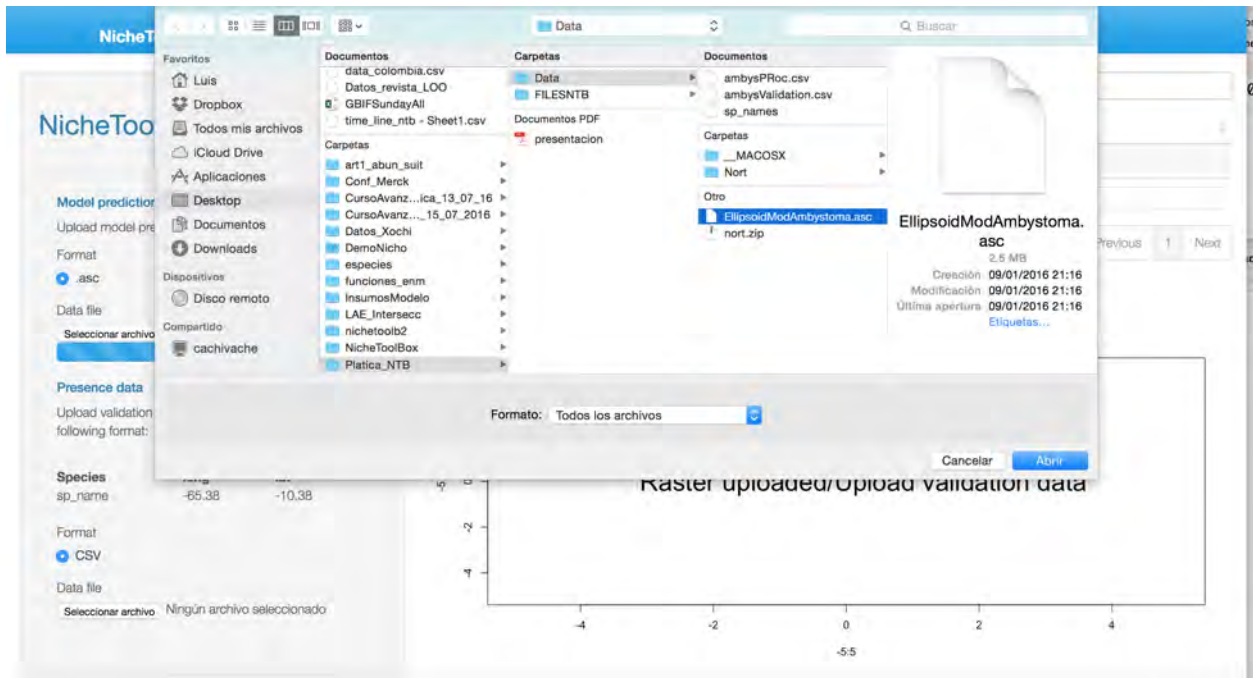


Figure 47:

Validation data must be in the following format:

sp_name	longitude	latitude
Ambystoma tigrinum	-107.08333	51.08333
Ambystoma tigrinum	-102.41667	44.41667
Ambystoma tigrinum	-99.75000	45.91667
Ambystoma tigrinum	-85.75000	45.25000
Ambystoma tigrinum	-91.75000	45.75000
Ambystoma tigrinum	-91.41667	39.75000

Partial ROC output

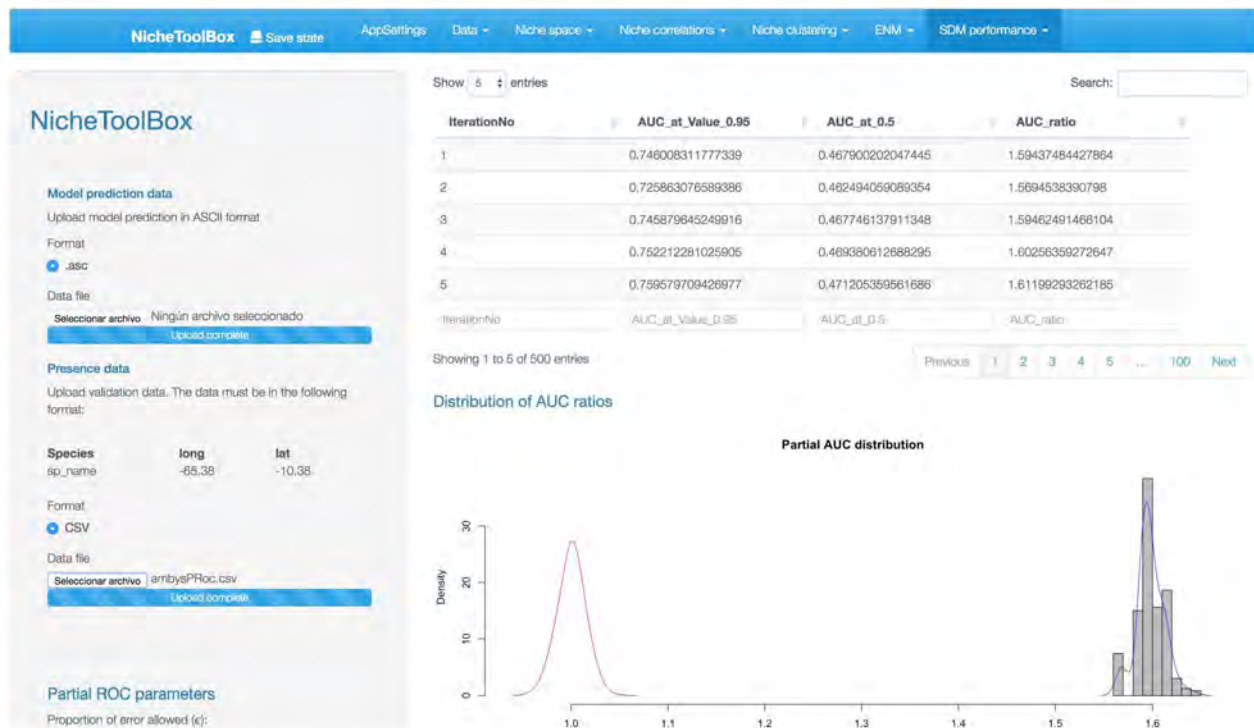
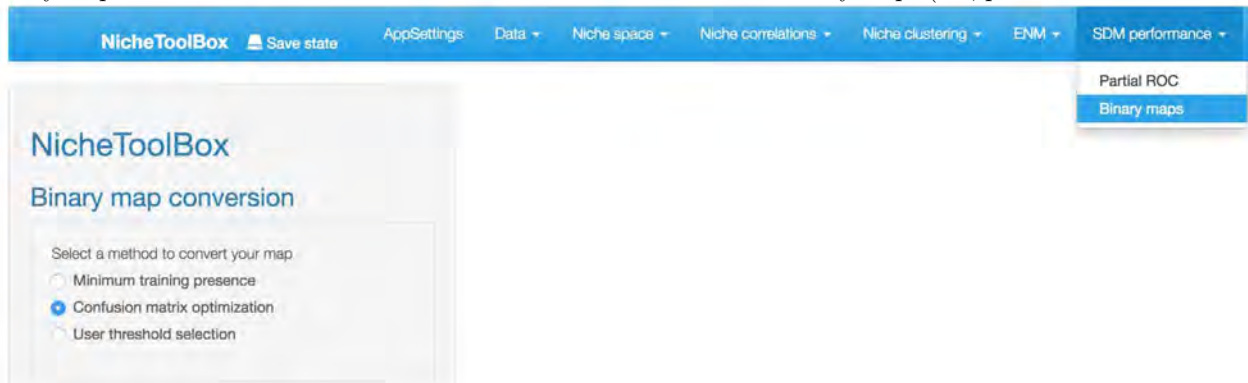


Figure 48:

7.2 Binary maps

The 'Binary maps' section has functions to transform continuous models into binary maps (i.e., presence and ab-



sence).

The conversion can be done by using one of following methods: 1) Confusion matrix optimization: By using true presences and absences the algorithm search for the cut-off threshold that optimizes the value of Kappa and/or TSS statistic.

- 2) Minimum training presence: Uses the lowest suitability value where a presence occurs as cut-off threshold.
- 3) User defined threshold: The user specifies the cut-off threshold.

7.2.1 Confusion matrix optimization

The user uploads both the continuous map (.asc) and the presences/absences data file (.csv). The presences/absences data have to be in the following format

sp_name	longitude	latitude
Ambystoma tigrinum	-107.08333	51.08333
Ambystoma tigrinum	-102.41667	44.41667
Ambystoma tigrinum	-99.75000	45.91667
Ambystoma tigrinum	-85.75000	45.25000
Ambystoma tigrinum	-91.75000	45.75000
Ambystoma tigrinum	-91.41667	39.75000

Once uploaded press specify the range of thresholds to look for and press **Search threshold** button

Real Presences and Absences

Upload your validation data. The data must be in the following format:

long	lat	PrecenseAbsence
-65.38	-10.38	1
-62.11	-11.22	0

Format

CSV

Data file

Seleccionar archivo ambysValidation.csv

Upload complete

Select a threshold range to look for the best cut off



Select a partition step for your range

0.00982112348079681

Optimization by

Kappa

TSS

Search threshold

The output

NicheToolBox

Binary map conversion

Select a method to convert your map

- Minimum training presence
- Confusion matrix optimization
- User threshold selection

Upload the model prediction raster in ASCII format

Format

- .asc

Data file

Seleccionar archivo EllipsoidMod...bystoma.asc

Real Presences and Absences

Upload your validation data. The data must be in the following format:

Show 10 entries Search:

threshold	a	b	c	d	kappa	tss	prevalence	specificity	sensibility	correct_classification
0.127674605250359	57	0	13	60	0.8019	0.8143	0.4615	1	0.8143	0.9
0.117853481769562	59	2	11	58	0.8009	0.8095	0.4615	0.9667	0.8429	0.9
0.137495728731155	55	0	15	60	0.7719	0.7857	0.4615	1	0.7857	0.8848
0.147316852211952	52	0	18	60	0.7273	0.7429	0.4615	1	0.7429	0.8615
0.157137975692749	52	0	18	60	0.7273	0.7429	0.4615	1	0.7429	0.8615
0.166959099173546	52	0	18	60	0.7273	0.7429	0.4615	1	0.7429	0.8615
0.176780222654343	52	0	18	60	0.7273	0.7429	0.4615	1	0.7429	0.8615
0.186601346135139	52	0	18	60	0.7273	0.7429	0.4615	1	0.7429	0.8615
0.108032359288765	59	8	11	52	0.707	0.7095	0.4615	0.8657	0.8429	0.8538
0.196422469615936	49	0	21	60	0.6829	0.7	0.4615	1	0.7	0.8385

Showing 1 to 10 of 51 entries

Previous 1 2 3 4 5 6 Next

looks like this

long	lat	PresenceAbsence
-65.38	-10.38	1
-62.11	-11.22	0

Format

- CSV

Data file

Seleccionar archivo ambysValidation.csv

Select a threshold range to look for the best cut off

Select a partition step for your range

Optimization by

- Kappa
- TSS

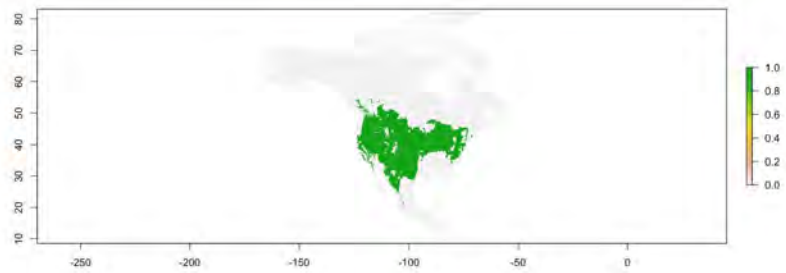
Search threshold

Confusion matrix

Enter values for: a, b, c y d

Correctly predicted presences (a)

Species Distribution Area 5793655.9635 in km² (at 0.127674605250359 , threshold)



Confusion matrix

	Observed presence	Observed absence
Predicted presence	a = 57	b = 0
Predicted absence	c = 13	d = 60

Predicted as present but actually absent(b)

Predicted as absent but actually present (c)

Correctly predicted absences (d)

[Download binary map](#)

[Download Binary](#)

[Download binary metadata](#)

[Download metadata](#)

Select a metric

Show all metrics

Prevalence

Specificity

Sensitivity

TSS

Kappa

Correct classification rate

Misclassification rate

Negative predictive power

Positive predictive power

Omission error fraction

Commission error fraction

False negative rate

False positive rate

Evaluation metrics for SDMs

```

*****
                          Symbols
*****
a = Correctly predicted presences
b = Predicted present but actually absent
c = Predicted absent but actually present
d = Correctly predicted absences

*****
                          Confusion Matrix
*****
                          Observed presence Observed absence
Predicted presence "a = 57"      "b = 0"
Predicted absence "c = 13"      "d = 60"

*****
                          Evaluation metrics for SDMs
*****
Kappa: 0.8018757
TSS: 0.8142857
Prevalence: 0.4615385
Correct classification rate: 0.9
Misclassification rate: 0.1
Negative predictive power: 0.8219178
Positive predictive power: 1
False negative rate: 0.1857143
False positive rate: 0
Specificity: 1
Sensitivity: 0.8142857
Omission error (fraction): 0.1857143
Commission error (fraction): 0
*****

```

7.2.2 Minimum training presence

Just upload your continuous model (.asc) and your training data file (.csv).

Validation data must be in the following format:

	sp_name	longitude	latitude
85	Ambystoma tigrinum	-100.58333	31.91667
86	Ambystoma tigrinum	-91.08333	38.91667
87	Ambystoma tigrinum	-113.41667	42.75000
88	Ambystoma tigrinum	-121.41667	39.75000
89	Ambystoma tigrinum	-114.58333	42.91667
90	Ambystoma tigrinum	-94.41667	45.41667

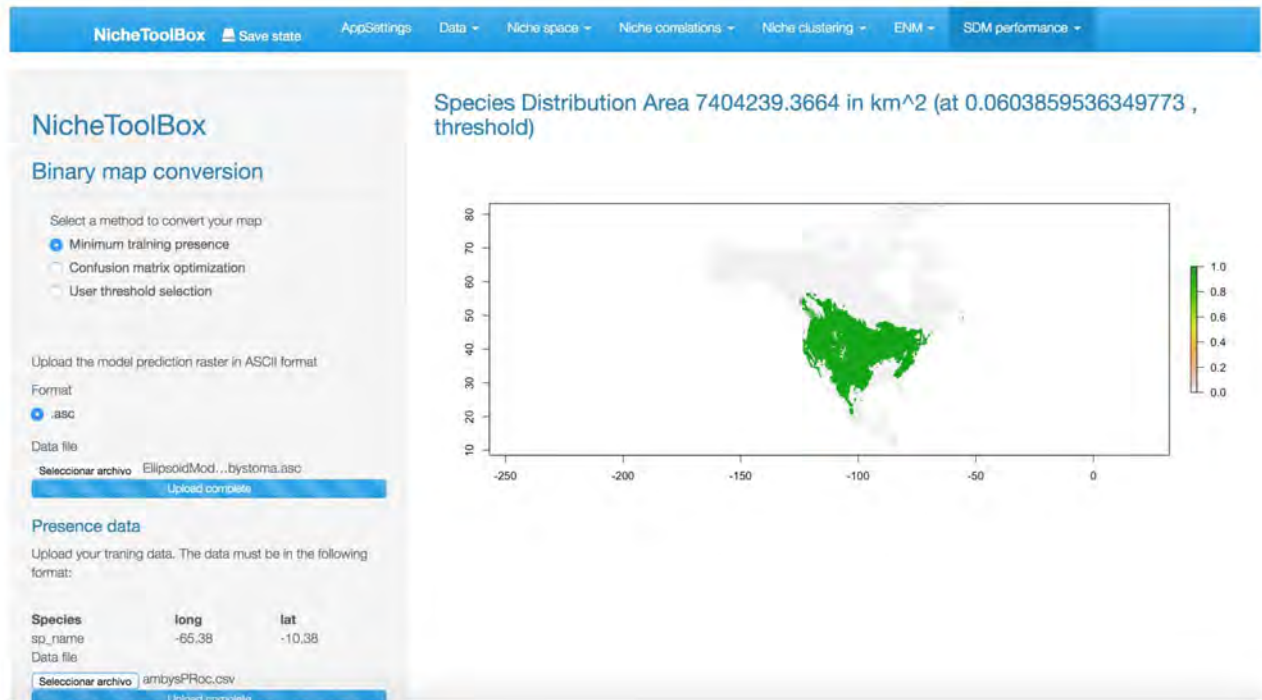


Figure 49:

7.2.3 User defined threshold

Specify a cut-off threshold

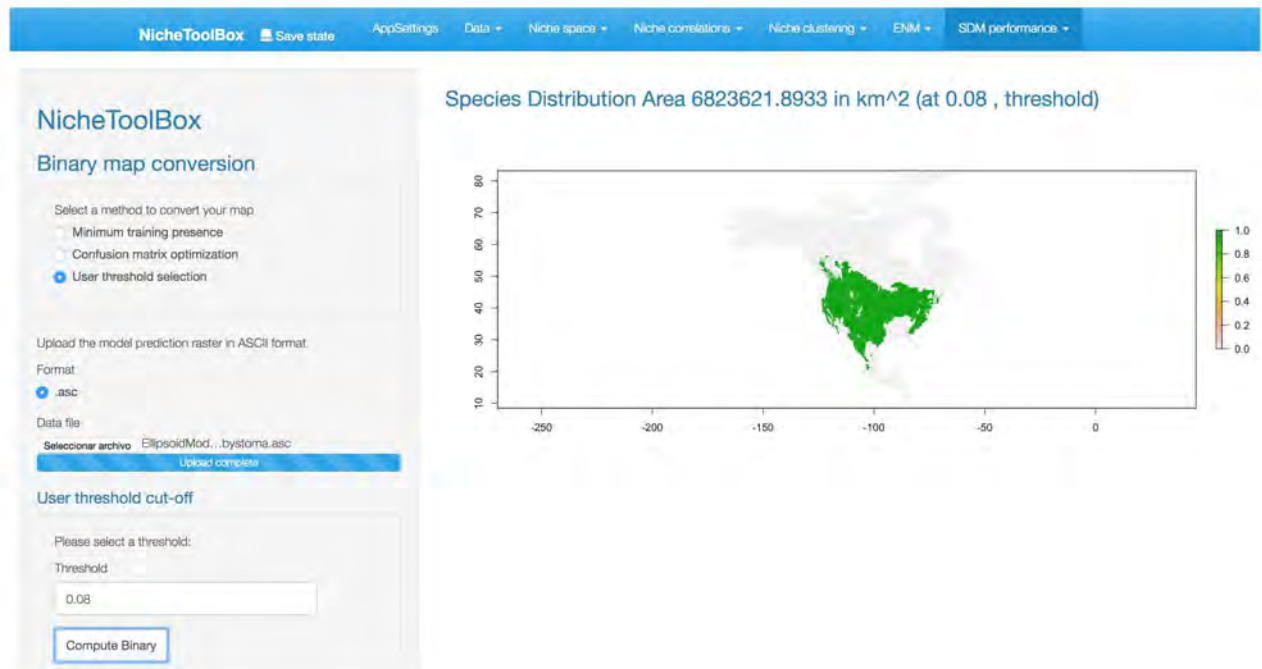


Figure 50:

References

- Fielding AH and Bell F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24(1):38–49
- Peterson AT, Vieglais DA (2001) Predicting species invasions using ecological niche modeling: New approaches from bioinformatics attack a pressing problem. *Bioscience* 51:363–371
- Peterson AT (2003) Predicting the geography of species' invasions via ecological niche modeling. *Quarterly Review of Biology* 78:419–433
- Peterson AT, Papes M., Soberon J. (2008) Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological modeling* 213:63–72
- Peterson AT, Soberón J., Pearson R., Anderson R., Martínez-Meyer E., Nakamura M. & Araújo M. (2011) *Ecological Niches and Geographic Distributions*. Princeton University Press

§4.1.2. Artículo: Resumen del seminario de NicheToolbox.

INTRODUCCION A LOS ANÁLISIS ESPACIALES CON ÉNFASIS EN MODELOS DE NICHO ECOLÓGICO

ANGELA P. CUERVO-ROBAYO^{1,2}, LUIS E. ESCOBAR^{3*}, LUIS A. OSORIO-OLVERA^{1,4}, JAVIER NORI⁵, SARA VARELA⁶, ENRIQUE MARTÍNEZ-MEYER¹, JORGE VELÁSQUEZ-TIBATÁ⁷, CLARITA RODRÍGUEZ-SOTO⁸, MARIANA MUNGUÍA², NORA P. CASTAÑEDA-ÁLVAREZ⁹, ANDRÉS LIRA-NORIEGA¹⁰, MARIANO SOLEY-GUARDIA¹¹, JOSEP M. SERRA-DÍAZ¹², A. TOWNSEND PETERSON¹³

¹*Instituto de Biología, Universidad Nacional Autónoma de México, México Distrito Federal 04510, México.* ²*Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), Insurgentes Sur-Periférico 4903, Tlalpan, México Distrito Federal 14010, México.* ³*Department of Fisheries, Wildlife and Conservation Biology, University of Minnesota, Estados Unidos*¹. ⁴*Departamento de Matemáticas de la Facultad de Ciencias, Universidad Nacional Autónoma de México, Ciudad de México, México.* ⁵*Instituto de Diversidad y Ecología Animal (IDEA-CONICET), Centro de Zoología Aplicada, Universidad Nacional de Córdoba, Rondeau 798, Córdoba, Argentina.* ⁶*Museum für Naturkunde, Leibniz Institute for Evolution and Biodiversity Science, Berlin, Germany.* ⁷*Laboratorio de Biogeografía Aplicada, Instituto Alexander von Humboldt, Calle 28ª # 15-09, Bogotá D.C., Colombia.* ⁸*Centro de Estudios e Investigación en Desarrollo Sustentable, Universidad Autónoma del Estado de México. Toluca, Estado de México.* ⁹*Global Crop Diversity Trust, Platz der Vereinten Nationen 7, 53113, Bonn, Germany.* ¹⁰*Catedrático CONACyT, Red de Estudios Moleculares Avanzados, Instituto de Ecología A.C., Xalapa, Veracruz.* ¹¹*Escuela de Biología, Universidad de Costa Rica, Ciudad Universitaria, 11501-2060, San Pedro, Costa Rica.* ¹²*Ecoinformatics and Biodiversity, Department of Bioscience, Aarhus University, Ny Munkegade 114, Aarhus C DK-8000, Denmark.* ¹³*Biodiversity Institute, University of Kansas, Lawrence, KS 66045, Estados Unidos.*

Resumen.—En 2016 implementamos un sistema de seminarios de enseñanza, en formato de videos libres y accesibles desde internet, con la finalidad de dar a conocer de forma sencilla y en castellano, las bases conceptuales y aplicaciones de los modelos de nicho ecológico en estudios de ecología, conservación biológica, epidemiología y agrobiodiversidad, así como su implementación para el diseño de políticas públicas de los recursos naturales. Cada seminario fue desarrollado por uno o varios expertos discutiendo conceptos, métodos y diferentes herramientas disponibles para elaborar modelos de distribución de especies. Este manuscrito reúne los resúmenes de cada uno de los seminarios en línea, dando referencias clave para cada tema y el enlace al video correspondiente. Los videos están disponibles de forma libre en YouTube o en formato .mp4 bajo solicitud.

Palabras clave.—Ecología espacial; biogeografía; conservación; distribución de especies; en línea; modelos de nicho ecológico; seminarios; video; enseñanza; capacitación en línea.

¹Enviar correspondencia a: E-mail: lescobar@umn.edu.

Abstract.—In 2016, we implemented a system of seminars for teaching, in open access video format available via the Internet, aiming to show in Spanish and in simple form the conceptual bases and applications of ecological niche modeling in studies of ecology, biological conservation, epidemiology, and agro-biodiversity, as well as its implementation for designing public policies for natural resources. Each seminar was developed by one or several experts discussing concepts, methods, and tools available to construct species distribution models. This manuscript assembles the abstracts for each of the seminars online, providing key references for each topic and links to the corresponding videos. The videos are available freely via YouTube or in .mp4 format by request.

Key words.—Spatial analysis; biogeography; conservation; species distributions; online; ecological niche modeling; seminar; video; teaching, online training.

El modelado de nicho ecológico permite estudiar la distribución geográfica de las especies e identificar aquellos factores ambientales que la limitan (Peterson et al. 2011). En general, los modelos de nicho ecológico relacionan datos de presencia (y algunas veces también de ausencia) de las especies con una serie de parámetros ambientales para generar una aproximación de las condiciones que favorecen la presencia de las poblaciones de la especie (el nicho ecológico). Este modelo se calcula en un espacio ambiental multidimensional para luego ser proyectado al espacio geográfico para generar un mapa que representa una distribución potencial (Peterson et al. 2011). Junto con otra serie de análisis espaciales, los modelos de nicho ecológico han permitido ampliar las preguntas que se abordan desde el campo de la biogeografía (Peterson 2008). El modelado de nicho ecológico es un campo en constante evolución y adaptación a nuevas preguntas y métodos. Por esto, los ecólogos, biólogos, epidemiólogos y biogeógrafos necesitan estar en constante actualización.

En este sentido, las redes sociales son un medio eficiente para conectar y actualizar a la comunidad científica, diluyendo las barreras en espacio y tiempo, ya que permiten la comunicación en tiempo real en casi todo el mundo (Kaplan & Haenlein 2010); por ello, actualmente son un medio altamente utilizado entre jóvenes investigadores para el uso compartido de la información (Tachibana 2014). Este trabajo tiene como objetivo dar a conocer a científicos de habla hispana interesados en el modelado de nichos ecológicos y otros análisis espaciales una primera serie de seminarios en línea. Se presentaron 13 seminarios por parte de 15 especialistas en el campo radicados en 6 países.

Según las estadísticas de la página de YouTube (www.youtube.com; a enero 2017) se registraron al menos ~8500 visualizaciones a los videos, con un rango de edad de la audiencia de 25–34 años, seguido por un rango de 35–44 años. Las visualizaciones se efectuaron desde 19 países, con mayor audiencia en México, Colombia, Perú, Argentina y Ecuador (Cuadro 1), pero la presente publicación puede ampliar la audiencia de los seminarios (expandiendo el número de usuarios y la lista de países desde los cuales se efectúan las visualizaciones).

Los seminarios están organizados en tres grupos: (i) bases conceptuales de los modelos de nicho ecológico; (ii) aplicaciones de los modelos; y (iii) tutoriales de las herramientas disponibles. Este artículo pretende estimular el interés en este campo de investigación presentando una síntesis de los seminarios ofrecidos y los enlaces a sus videos correspondientes en los que se explican las bases conceptuales y métodos modernos para generar modelos de nicho ecológico.

BASES CONCEPTUALES DE LOS MODELOS DE NICHOS ECOLÓGICOS *[El Diagrama BAM](#)*

El diagrama **BAM** (Soberón & Peterson 2005) es un marco conceptual de referencia para pensar en torno a los factores bióticos (**B**), abióticos (**A**) y de accesibilidad o movilidad (**M**) que son determinantes para explicar las áreas de distribución de las especies. Cuando los conjuntos de áreas **B**, **A** y **M** de una especie coinciden geográficamente, se podrán encontrar poblaciones de dicha especie, ya que esas localidades serían accesibles, **M**, y presentarían las condiciones ambientales, **A**, y las interacciones interespecíficas suficientes, **B**, para mantener una tasa

demográfica neta positiva. Estas localidades constituyen el área ocupada ($\mathbf{B} \cap \mathbf{A} \cap \mathbf{M}$).

El diagrama **BAM** puede usarse para entender por separado la influencia de cada uno de los tres conjuntos de factores, pero también para deducir cómo su variación espacial y temporal puede estar afectando la distribución de una especie (Hutchinson 1978, Soberón 2007). En este sentido **B** está compuesto por variables de tipo *biónómico* (relacionadas dinámicamente con la especie), que caracterizan interacciones interespecíficas (positivas o negativas), incluyendo recursos de los cuales depende la especie de interés. **A** corresponde a variables *escenopoéticas* o no dinámicas (no son modificadas, en un sentido amplio, por la especie), y que típicamente son más estables en el tiempo, tales como el clima de una región. Finalmente, **M** corresponde a una hipótesis sobre el área sobre la cual la especie tiene, o ha tenido, acceso para dispersarse (Barve et al. 2011; ver también Anderson & Raza 2010).

Cada uno de estos tres factores se cuantifica y está disponible de manera diferente. La información de **B** es, por lo general, muy escasa, ya que generalmente desconocemos cuáles son las interacciones con otras especies y cómo varían éstas a través del espacio y tiempo. Gran parte de la información de **A** está disponible a diversas resoluciones espaciales y temporales, por ejemplo, en forma de coberturas bioclimáticas; y **M** debería calcularse a partir de la capacidad de dispersión de la especie a través del tiempo. En la práctica, debido al amplio desconocimiento de **B**, se suele operar con los factores **A** y **M**, aunque **B** pueda ser un factor determinante a escalas pequeñas (Soberón & Nakamura 2009) e incluso se ha sugerido un efecto importante de **B** a escalas continentales (e.g., Gutiérrez et al. 2014; Wisz et al. 2012). Quizás lo más importante es que el diagrama **BAM** es una referencia para pensar formalmente sobre distintas combinaciones de los tres conjuntos de factores (Peterson et al. 2011) para preguntarse: ¿Qué zonas son potencialmente habitables para una especie (área invadible)? ¿Qué determina una localidad de presencia o ausencia de una especie? ¿Cómo operan los algoritmos para hacer modelos de nicho ecológico dependiendo de las configuraciones del **BAM** (por ej., **BAM** clásico, mundos de Wallace y Hutchinson; Saupe et al. 2012)? ¿Los algoritmos analizan la periferia ambiental y riesgos de extrapolación (Owens et al. 2013)? Este enfoque también permite explorar una teoría más general

de la biodiversidad que incorpore la demografía (Soberón 2010). La sugerencia es que un buen ejercicio de modelado de nichos ecológicos, así como la edición posterior del modelo para aproximarlos al área de distribución deseada, debe ir acompañado de su propia configuración del diagrama **BAM**.²

Cuadro 1. Lista de países en los que se ha reportado por lo menos una hora de visualización.

País	Horas de visualización
México	711
Colombia	295
Perú	102
Argentina	78
Ecuador	60
Estados Unidos	59
Brasil	42
Chile	29
España	27
Guatemala	15
Costa Rica	14
Venezuela, Reino Unido	8 c/u
Bolivia	5
Francia	3
Nueva Zelanda, Canadá, Uruguay, Nicaragua	2 c/u

APLICACIONES DE LOS MODELOS DE NICHOS Y MODELOS DE DISTRIBUCIÓN

[Biogeografía de la Conservación, Anfibios y Reptiles](#)

En la actualidad existe una crisis global de pérdida de biodiversidad. Históricamente muchas decisiones para la conservación de la biodiversidad se han basado en disposiciones políticas sin dar suficiente importancia a la preservación de los ambientes naturales, su biodiversidad y los servicios ecosistémicos que proveen (Margules & Pressey 2000). No obstante, este tipo de decisiones son críticas en la conservación, y requieren un entendimiento de la situación al menos a escalas espaciales gruesas. En los últimos años, se ha consolidado la *biogeografía de la conservación*, sub-disciplina de la biología de la conservación enfocada en los procesos y patrones que operan a escalas espacio-temporales amplias y resoluciones gruesas (Ladle & Whittaker 2011). La biogeografía de la conservación utiliza principios, teorías y análisis de la biogeografía relacionados a la dinámica de las distribuciones de diversos taxa, para abordar

²https://youtu.be/Px_bGt-neh8?list=PLU_3tLNPCPDZd0VvXVCt9XHcNx7L9E_e

problemas relacionados a la conservación de la biodiversidad (Ladle & Whittaker 2011).

En este seminario se abordan estudios de biogeografía de la conservación que buscan generar información científica para la toma de decisiones para la conservación de los anfibios y reptiles de Sudamérica. Entre los ejes temáticos se destacan: (i) la determinación de zonas vulnerables a invasiones de la rana toro norteamericana (*Lithobates catesbianus*) en Argentina y, luego, a lo largo de Sudamérica, considerando escenarios de cambio climático (Nori et al. 2011) y zonas vulnerable a la invasión de las tortugas acuáticas más comercializadas en Argentina (Nori et al. 2016); (ii) el estudio de la eficiencia del sistema global de áreas protegidas para representar a los anfibios (Nori et al. 2015; Nori & Loyola 2015); (iii) el estudio de la exposición de los reptiles de Argentina al cambio climático (Nori et al. 2016); y (iv) la determinación de áreas prioritarias para conservación en el Gran Chaco Sudamericano (Nori et al. 2016) y la provincia de Córdoba en Argentina (Nori et al. 2013). Estos ejemplos representan una muestra de los artículos científicos realizados exclusivamente para aportar información útil en la toma de decisiones para la conservación de la biodiversidad. Algunos estudios están siendo considerados por los tomadores de decisiones o incluso han sido generados específicamente por requerimiento de éstos. Sin embargo, muchas veces la valiosa información, luego de ser generada, no es utilizada o siquiera detectada por los tomadores de decisiones. En ese sentido, la interacción y articulación entre los científicos del área y las esferas políticas encargadas de la toma de decisiones resulta indispensable.³

[Territorios de Oportunidad para Mejorar la Conservación Considerando Factores Socio-ambientales](#)

La identificación y planeación de áreas prioritarias para la conservación, así como el uso eficiente de los recursos son aspectos fundamentales en el éxito de la conservación biológica (Valenzuela & Vázquez 2007). Por ello, se ha desarrollado un enfoque de investigación cuyo objetivo es identificar las áreas que deben priorizarse para la distribución de los escasos recursos dedicados al manejo de la biodiversidad y desvincular estas áreas de los factores que amenazan su persistencia, por ejemplo, la

planificación sistemática para la conservación (PSC; Margules & Sarkar 2009).

La PSC debe tener en cuenta muchos criterios para garantizar un equilibrio entre la conservación de la biodiversidad y otros tipos de uso de la tierra (Faleiro & Loyola 2013). Su eficiencia puede ser mayor si se tienen en cuenta las dimensiones sociales y humanas, incluyendo la gobernabilidad y la voluntad política (p.ej., Faleiro & Loyola 2013). Debido a que la biodiversidad es difícil de estimar y conservar en su totalidad, se propone usar medidas parciales o subrogados que pueden ser determinadas más fácilmente (Margules & Sarkar 2009). En este sentido, existen especies de mayor interés para la conservación, como las especies bandera y las especies con ámbitos hogareños amplios (Valenzuela & Vázquez 2007).

En este seminario se presenta la aplicación de la PSC en la identificación de escenarios de priorización para la conservación de vertebrados terrestres en México, considerando factores socio-ambientales que pueden influir en la conservación. En este ejercicio se usó el software Zonation (Moilanen et al. 2011) para desarrollar diferentes escenarios de priorización a través de un análisis de costo-beneficio que reduce las limitaciones socioeconómicas y permite aprovechar las oportunidades políticas para la conservación (para un enfoque similar ver Faleiro & Loyola 2013). El algoritmo de priorización Zonation calcula la contribución relativa de cada celda para lograr el objetivo de conservación, utilizando la regla de eliminación de superficie original “área núcleo” (Moilanen et al. 2009 para más detalles). Zonation ofrece la posibilidad de penalizar las zonas de acuerdo a la importancia de los factores, lo que permite un equilibrio entre los beneficios y costos para las acciones de conservación (Moilanen et al. 2011). En este trabajo, a cada especie y variable se le asignó un valor de importancia diferente (ver Faleiro & Loyola 2013). Los resultados mostraron que la mayor concentración de biodiversidad converge en regiones con gran persistencia de la cobertura natural y alta gobernabilidad local, definida como la voluntad de las personas de participar en acciones de conservación. Los resultados del estudio de caso presentado en este seminario resaltan la relevancia de las variables socioeconómicas para futuros modelos de nicho ecológico, diseño de políticas ambientales e implicaciones para el cambio climático.⁴

³ https://youtu.be/semcgnitLd8?list=PLu_3tLNPCPDZd0VvVXVC9XHcNx7L9E

⁴ https://youtu.be/Dmykn5votC0?list=PLu_3tLNPCPDZd0VvVXVC9XHcNx7L9E_e

Gradiente de Impacto Humano y Biodiversidad

Las actividades humanas han causado cambios drásticos en los paisajes naturales; estos cambios han provocado la extinción local de algunas especies y por lo tanto pérdida de la biodiversidad. La extinción no es azarosa, si no depende de una serie de factores biológicos y ambientales. En particular, la sensibilidad de las especies a la degradación puede estar asociada a caracteres especie-específicos. Por lo tanto, asociar los caracteres de las especies (p.ej., grupos tróficos, masa corporal) con atributos ambientales, permite detectar el potencial impacto ecológico de la pérdida de una especie en los hábitats, debido a la función que desempeña la especie en los ecosistemas estudiados (p.ej., los frugívoros son potenciales dispersores de semillas y los nectarívoros polinizadores). Adicionalmente, esta relación permite la detección de sitios con riesgo latente para la biodiversidad mediante la identificación de los impactos que más afectan al grupo taxonómico de interés.

En este seminario se presenta una propuesta para diseñar un gradiente de impacto humano de biodiversidad (GIHB) basado en el análisis de ordenación RLQ (Dolédec et al. 1996) y "fourth corner" (Dray & Legendre 2008), el cual tiene como objetivo identificar especies y caracteres asociados a sitios con diferentes grados de degradación humana (Munguía et al. 2016). Se utilizaron datos de localidades de especies de mamíferos (CONABIO 2012) para seleccionar sitios de 50 km de radio con alta completitud de especies en las comunidades analizadas. Se incluyeron 211 especies de mamíferos terrestres y nueve variables ambientales biofísicas, geofísicas y de impacto humano, las cuales fueron procesadas en un sistema de información geográfico (ESRI 2014) y analizadas con la librería ade4 (Dray & Dufor 2007) en el paquete estadístico R (R Core Team 2014). Los caracteres de las especies evaluados fueron de tres tipos: modo de locomoción, hábito trófico y masa corporal. El GIHB detectado para los mamíferos fue conformado principalmente por el porcentaje de cobertura de plántulas, la riqueza de plantas, el porcentaje de cobertura vegetal bajo dosel (CONAFOR 2009) y el índice de densidad humana basado en imágenes espaciales de luces nocturnas (NOAA/NESDIS/NCEI 2011). Los resultados muestran que los caracteres asociados con sitios menos impactados por el hombre fueron los grupos tróficos carnívoros y los frugívoros,

así como los mamíferos con tamaño corporal mayor a 17.8 kg, lo que sugiere que estos son caracteres sensibles a la degradación. Por otro lado, los mamíferos con hábito trófico granívoro, y tipo de locomoción fosorial y semi-fosorial están asociados a los sitios más impactados. Este marco puede ser replicado en otras áreas con menos información, en donde se necesite identificar especies en riesgo al cambio en el uso del suelo. Las variables más importantes asociadas al impacto humano del GIHB fueron mapeados en conjunto para resaltar los sitios con mayor riesgo para los mamíferos en el país. Dada la crisis de biodiversidad y la acelerada tasa de extinción sin precedentes que actualmente enfrentamos (Levin 2005; Barnosky et al. 2011) existe la necesidad de continuar detectando diversos indicadores y especies sensibles a la degradación para el monitoreo de los ecosistemas.⁵

Identificación de Prioridades de Conservación de los Parientes Silvestres de Cultivos

Los sistemas de producción de alimentos enfrentan nuevos retos que impulsan el desarrollo de alternativas para la producción de más alimentos para una población humana en crecimiento, utilizando recursos naturales de manera eficiente, y reduciendo los impactos ambientales negativos asociados a la agricultura. Con el mejoramiento genético de cultivos es posible obtener variedades para incrementar la producción por hectárea cultivada, reducir gases de efecto invernadero (Subbarao et al. 2007), asimismo variedades más nutritivas y tolerantes a plagas y enfermedades (Borlaug 1983). El mejoramiento genético de cultivos requiere como materia prima los recursos fitogenéticos, ya que estos ofrecen la variabilidad genética para obtener nuevas variedades (Gepts 2006). Los recursos fitogenéticos se pueden organizar de acuerdo a la variabilidad genética que ofrecen: variedades obtenidas a través del mejoramiento genético, variedades tradicionales y especies silvestres emparentadas con las especies cultivadas (también conocidas como parientes silvestres de cultivos). Estas últimas son consideradas como portadoras de una alta diversidad genética ya que no han sido expuestas a procesos de selección propios de la domesticación (Mariat et al. 2006). Es así, como los parientes silvestres son utilizados para contribuir a que la agricultura sea una actividad más sostenible, debido a que son fuentes

⁵ https://youtu.be/vxuyvGkM9e4?list=PLu_3tLNPcPDZa0VvXVCt9XHcNx7L9E_e

de genes en mejoramiento con el potencial para ayudar a adaptar cultivos a las condiciones ambientales extremas asociadas al cambio climático y para obtener variedades que usen el agua, la tierra y los fertilizantes de forma eficiente (Guarino & Lobell 2011; Dempewolf et al. 2013).

En este seminario se presenta una introducción sobre parientes silvestres, la cual incluye su importancia, usos en la agricultura, algunas de las amenazas que enfrentan en sus hábitats, y los objetivos globales de conservación y desarrollo que reconocen su importancia para la seguridad alimentaria global. Adicionalmente, en el video se presenta en detalle la metodología utilizada por Castañeda-Álvarez et al. (2016) para estimar la representatividad de estos recursos fitogenéticos en bancos de germoplasma, y cómo los análisis obtenidos a través de esta metodología están siendo utilizados para establecer prioridades de conservación de las especies analizadas. El caso presentado en este video es un ejemplo práctico de cómo pueden utilizarse herramientas de modelación como MaxEnt, y datos de acceso público, como los facilitados a través de la Infraestructura Mundial de Información en Biodiversidad (GBIF⁶), para establecer prioridades de elementos de la naturaleza importantes para la alimentación.⁷

[Cambio Global, Plantas y Modelos de Distribución de Especies](#)

El nicho ecológico de las especies vegetales ha sido ampliamente interpretado en términos de clima, puesto que el clima condiciona en gran medida las respuestas fisiológicas y la ecología de las especies. En consecuencia, los modelos de distribución de especies o modelos de nicho ecológico (Franklin 2010, Peterson et al. 2011) utilizan principalmente variables climáticas, ya que son generalmente accesibles, no obstante existen otras variables que también se relacionan con el desempeño fisiológico. En este seminario discutimos diferentes dimensiones de la distribución de especies comparando modelos eco-fisiológicos, modelos de nicho ecológico y modelos de interacciones. Asimismo, analizamos las diferencias en las proyecciones de cambio climático bajo esos modelos. En una primera comparación entre modelos de nicho ecológico y modelos ecofisiológicos, observamos que diversas especies de árboles de la Península Ibérica podrían crecer en espacios climáticos que no ocupan actualmente (Serra-Diaz et al. 2013).

Menos frecuente fue el caso de encontrar espacios climáticos ocupados por la especie que no tuvieran un bajo rendimiento fisiológico (p.ej., crecimiento). En la comparación entre proyecciones de cambio climático realizadas con modelos ecofisiológicos y modelos de nicho correlativos, observamos que el efecto del CO₂ puede modificar la relación entre crecimiento y clima debido a un aumento en la eficiencia del uso de agua. Por ello, incluso en el caso de que la distribución actual de especies estuviera en equilibrio climático, el no considerar las relaciones biogeoquímicas que alteran el uso de variables climáticas y su respuesta en plantas (p.ej., producción, esfuerzo reproductivo) puede influir las proyecciones realizadas a partir de los modelos correlativos (Keenan et al. 2011).

Utilizando un modelo basado en individuos (LANDIS-ii, Scheller et al. 2007), proyectamos el cambio en la distribución de un conjunto de especies debido a un desplazamiento espacial del nicho. Este experimento de modelado identificó potenciales expansiones y contracciones del área de distribución de muchas especies arbóreas debido a la interacción de características funcionales, competición-facilitación, heterogeneidad espacial y distribución de microrefugios. Incluso, para especies cuyo nicho potencial se expandía con el cambio climático, el área de distribución se reducía debido a la combinación de perturbaciones y heterogeneidad espacial (Serra-Diaz et al. 2016). Como resultado de tales comparaciones, podemos entender que los modelos correlativos de nicho ecológico son en realidad una buena técnica para captar la exposición al cambio climático. Así, se analizaron los cambios de distribución de la exposición al cambio climático para varias especies de árboles de California y se desarrollaron varias métricas para comparar la velocidad del cambio en la exposición de las especies. Se observó que, para varias especies con distribución espacial relativamente similar, la exposición al cambio climático varía fuertemente entre ellas y entre mitad y final del siglo 21 (Serra-Diaz et al. 2014). En la actualidad existen varias aproximaciones y modelos que caracterizan bien diferentes dimensiones de la distribución de la especie (BAM *sensu* Soberón 2007). En el caso de plantas y cambio climático los modelos correlativos pueden caracterizar algunas facetas, pero deben ser analizados e interpretados conjuntamente con otras aproximaciones para poder proyectar

⁶ <http://www.gbif.org/>

⁷ https://youtu.be/wiFm11_wQ8Q?list=PLu_3iLNPcPDZ40VvXVCi9XHcN7L9E_e

cambios de distribución de plantas que sean plausibles en el curso de este siglo (Franklin et al. 2016).⁸

Mapeo de Riesgo de Transmisión de Enfermedades

Los sistemas de transmisión de enfermedades representan un conjunto de especies interactuando: patógenos, vectores y hospederos. Como tal, el modelado de las dimensiones geográficas del riesgo de transmisión requiere, en efecto, la integración de la ecología de la distribución de múltiples especies en un solo modelo compuesto. Una aplicación útil de los modelos de nicho ecológico es la de evaluar la geografía del riesgo de transmisión de enfermedades; las diversas y complejas consideraciones involucradas en estas aplicaciones han sido revisadas en un reciente compendio que abarca un libro (Peterson 2014).

En este seminario, se revisan los diversos conceptos y consideraciones prácticas involucradas en el mapeo del riesgo de transmisión de enfermedades usando modelos de nicho ecológico. Una bifurcación clave aparece entre los esfuerzos que buscan reconstruir el nicho ecológico y la distribución potencial de cada componente del ciclo de transmisión de la enfermedad, versus situaciones en las cuales solo está disponible la información sobre los casos de la enfermedad, lo que se define como aplicaciones de “caja negra”. Se presenta una serie de ejemplos y casos de estudio para ilustrar diferentes tipos de aplicaciones, así como algunos inconvenientes clave y problemas que se encuentran durante su desarrollo.⁹

HERRAMIENTAS DISPONIBLES

NicheToolBox: De la Obtención de Datos de Biodiversidad a la Validación de los SDMs

El modelado de nicho es un campo de la ecología que ha permitido estimar partes del nicho ecológico y la distribución geográfica de las especies (Elith & Leathwick 2009). Se utiliza un conjunto de herramientas estadísticas, matemáticas y computacionales para estimar la relación entre variables ambientales y la presencia de las especies (Franklin 2010). El proceso de modelación de la distribución involucra por lo menos cuatro etapas: (i) obtención de datos georreferenciados de presencia (y algunas veces de ausencia) de especies, (ii) depuración de las bases de datos,

(iii) ajuste de un modelo de distribución utilizando variables de información ambiental y los registros depurados y (iv) validación del modelo. Si bien existe un gran desarrollo de herramientas computacionales disponibles en R para modelar el nicho ecológico de las especies (p.ej., dismo, sdm), estas herramientas se encuentran dispersas y no cuentan con un flujo de trabajo que permita construir modelos de distribución de forma estandarizada. Más aún, algunos de los programas de uso libre en cierto sentido son considerados cajas negras debido a que su código es cerrado y por lo tanto no se sabe con certeza lo que hacen y, por otro lado, aprender a usar programas de código abierto representa un gran reto para aquellas personas que no están familiarizadas con la programación.

En este sentido, este seminario presenta el programa NicheToolBox (Osorio-Olvera et al. 2016), una plataforma y paquete de R con una interfaz de usuario amigable desarrollada en Shiny (Chang et al. 2016). El objetivo de NicheToolBox es facilitar el proceso de modelado de nicho ecológico y de las distribuciones de las especies. La plataforma incorpora funciones propias y otras disponibles en diferentes paquetes de R (p.ej., dismo, ENMGadgets, spocc) para buscar datos de presencia, limpiar duplicados, seleccionar variables ambientales, calibrar algoritmos de nicho ecológico (p.ej., Bioclim, MaxEnt y modelos basados en elipsoides) y evaluar los modelos de distribución con las métricas dependientes de umbral, como la sensibilidad, o independientes, como el AUC (*area under the curve*) de la curva ROC (*receiver operating characteristic*) y la ROC parcial (Peterson et al. 2008). Una de las características notables de NicheToolBox es que cuenta con funciones para descargar el flujo de trabajo (*workflow*) de lo que el usuario ha hecho dentro de la aplicación. Este flujo de trabajo, además de contener los archivos de los análisis realizados en la sesión, guarda en un documento el código de R con el que los produjo. Con lo anterior se pretende hacer que el proceso de modelado sea transparente, además de que los usuarios aprendan a programar en R mientras realizan su investigación.¹⁰

NicheA

Los modelos de nicho ecológico buscan una aproximación de la distribución de las especies a través de la estimación sus nichos ecológicos (ver

⁸ https://youtu.be/bTby0saYUck?list=PLu_3lNlNCPDZd0VvXVC9XHcNx7L9E_e

⁹ https://youtu.be/yBzLlSuuFY?list=PLu_3lNlNCPDZd0VvXVC9XHcNx7L9E_e
¹⁰ https://youtu.be/cc03KfWzAL4?list=PLu_3lNlNCPDZd0VvXVC9XHcNx7L9E_e

el diagrama de **BAM**). Así, los modelos de nicho ecológico son análisis generados en espacios ambientales que pueden ser proyectados al espacio geográfico (Warren 2012). No obstante, muchas veces los estudios son diseñados, evaluados, e interpretados considerando únicamente la geografía (p.ej., Peterson et al. 2008; Fischer et al. 2014; Radosavljevic & Anderson 2014). Por otro lado, los avances en los métodos y las variables disponibles para la calibración de los modelos de nicho ecológico han permitido avanzar considerablemente en este campo (Escobar & Craft 2016). Por ejemplo, cada vez se desarrollan y proponen nuevos algoritmos para el modelado de los nichos ecológicos (p.ej., Blonder et al. 2014; Drake 2015; Qiao et al. 2015). Estos nuevos algoritmos requieren de rigurosas evaluaciones para validar su aplicación en la práctica (Peterson et al. 2008), y las evaluaciones a su vez requieren datos robustos y confiables para determinar si los modelos pueden o no reconstruir los nichos ecológicos y la distribución de las especies. Sin embargo, los datos frecuentemente sufren de sesgo en el muestreo o no representan correctamente el nicho ecológico de las especies, lo que limita la correcta evaluación de los modelos (Kadmon et al. 2004). En este sentido, han surgido las especies virtuales como alternativa que permite controlar la forma, posición y amplitud de un nicho ecológico. Además, permiten manejar la respuesta de las especies a las variables ambientales y controlar la distribución espacial de las muestras para obtener muestras con o sin sesgo en el muestreo.

En este seminario se presenta el programa NicheA, que permite diseñar especies virtuales en espacios ambientales multidimensionales (Qiao et al. 2016). NicheA está construido en un ambiente Java enlazado a R, permitiendo desarrollar nichos virtuales en una plataforma amigable (Figura 1). Una vez creada la especie virtual, NicheA permite generar desde análisis descriptivos básicos hasta análisis predictivos multivariados complejos. La versatilidad y facilidad de uso de NicheA hacen de este programa una herramienta ideal en docencia en cursos de biología básica o capacitaciones avanzadas para el modelado de nichos ecológicos. En este video, se explican las bases conceptuales y teorías ecológicas que respaldan el uso de NicheA para crear especies virtuales. Además, se muestran aspectos introductorios del funcionamiento de NicheA, el

manejo de los datos, aplicaciones de las especies virtuales creadas y material de ayuda en línea para el proceso de instalación y utilización del programa. Finalmente, se muestra cómo NicheA puede ayudar a interpretar y comparar modelos en espacios ambientales, convirtiendo a NicheA en una poderosa herramienta de última generación para el modelado e interpretación de nichos ecológicos en ecología moderna.¹¹

[Modelamiento de Distribución de Especies con MaxEnt en R: Tutorial y Fundamentos](#)

El modelaje de distribuciones de especies es un campo de rápido crecimiento en el que nuevos programas, métodos y recomendaciones de buenas prácticas son publicados continuamente. Un elemento frecuente en los avances recientes en el campo es el uso del lenguaje estadístico R (R Core Team 2016), el cual permite el desarrollo de flujos completos de trabajo, facilita la replicación y permite el desarrollo de experimentos virtuales de alta complejidad.

En este seminario se presenta una introducción al modelamiento de distribuciones de especies en R usando MaxEnt (Phillips et al. 2016), a través de dismo (Hijmans et al. 2016) y un conjunto de datos para *Bradypus variegatus* (Phillips et al. 2016, disponible en¹²). Inicialmente se facilita una “traducción” de las opciones de ingreso de datos y desarrollo de modelos con la configuración predeterminada de MaxEnt. Posteriormente se explica cómo (i) evaluar modelos y seleccionar umbrales usando validación cruzada, AUC y TSS; (ii) transferir modelos a otras épocas/regiones y (iii) configurar distintos argumentos para el desarrollo de modelos y su predicción en el espacio geográfico en MaxEnt a través de R. Finalmente se presentan dos implementaciones de buenas prácticas recomendadas en la literatura: (1) la optimización del multiplicador de regularización (Warren & Seifert 2011) y (2) el muestreo de datos de entorno del área accesible, **M** (Anderson & Raza 2010; Barve et al. 2011). Este tutorial está dirigido a usuarios con conocimientos mínimos del uso de MaxEnt y de la sintaxis de R^{13,14}. El código empleado en las demostraciones se encuentra disponible en: ¹⁵.

[ENMeval: La Teoría detrás del Paquete](#)

En el modelaje existe un clásico balance entre generalidad y sobreajuste, en el cual la complejidad óptima se encuentra entre ambos

¹¹ https://youtu.be/cSbod4vQdTg?list=PLu_3tLNPCPDZd0VvXVCi9XHcNx7L9E_e

¹² <https://www.cs.princeton.edu/~schapire/maxent/datasets/samples.zip>

¹³ Parte I: https://youtu.be/2pKC-xFmBo8?list=PLu_3tLNPCPDZd0VvXVCi9XHcNx7L9E_e

¹⁴ Parte II: https://youtu.be/q-x3VXppq5c?list=PLu_3tLNPCPDZd0VvXVCi9XHcNx7L9E_e

¹⁵ <https://github.com/jivelasquez/courses>

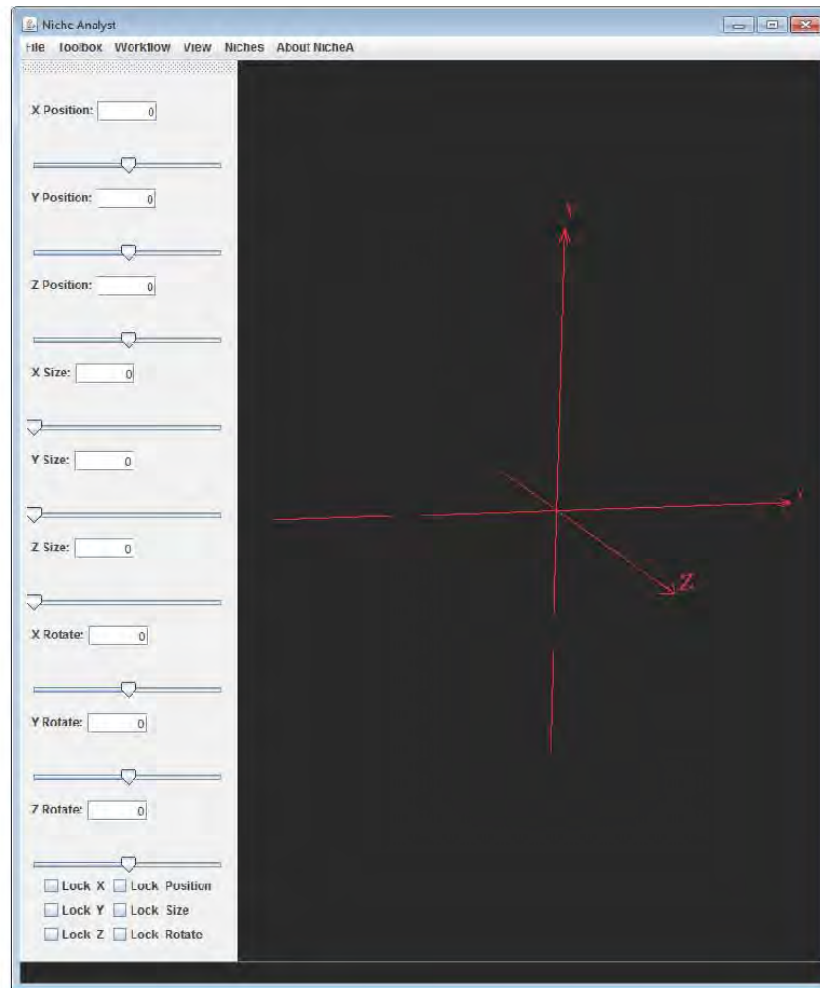


Figura 1. Interface de usuario de NicheA. Los botones superiores (File, Toolbox, Workflow, View, Niches, About NicheA) despliegan herramientas para selección, análisis de las variables ambientales y nichos e información general sobre el programa. Las herramientas laterales (izquierda) permiten diseñar el nicho ecológico de una especie virtual de forma manual (barras) o a través de valores de la dimensión y posición del nicho (cuadros blancos). Derecha: las condiciones ambientales del área de estudio y los modelos de nicho son proyectados en un espacio ambiental de una (X), dos (Y), o tres dimensiones (Z).

extremos y puede ser determinada mediante evaluaciones rigurosas (Peterson et al. 2011). En este seminario se discuten los conceptos básicos detrás de la elaboración y evaluación de un modelo correlativo de nicho ecológico haciendo énfasis en el programa de modelación MaxEnt (Phillips et al. 2006). Seguidamente, se exponen y explican las limitantes importantes que presentan las evaluaciones automatizadas de dicho programa, y cómo éstas pueden ser solucionadas mediante el paquete complementario de R (R Core Team 2016) ENMeval (Muscarella et al. 2014).

En particular, MaxEnt permite extraer relaciones altamente complejas entre la especie y el ambiente, las cuales pueden ser más representativas de lo que realmente ocurre en sistemas naturales, otorgando así una alta capacidad predictiva (Olden et al. 2008; Evans et al. 2013). No obstante, la complejidad potencial de MaxEnt también puede resultar en un serio sobreajuste a la información existente. En este caso, los modelos pueden convertirse más bien en una representación de los problemas típicamente asociados a datos (p.ej., muestreo incompleto o sesgado) reduciendo su capacidad predictiva e incluso invalidando su utilidad en general (Anderson 2012; Merow et al. 2013). Mientras MaxEnt permite al usuario variar la complejidad potencial del modelo mediante el uso de distintas características de los argumentos (*feature classes*) y niveles de regularización (Phillips & Dudík 2008), las evaluaciones implementadas por este software no toman en cuenta la autocorrelación espacial presente entre los datos de calibración y evaluación. Esto resulta en valores artificialmente inflados para las métricas de evaluación (Velo 2009), lo que no solo dificulta la determinación de una complejidad óptima, sino que conlleva a que los usuarios ni siquiera intenten calibrar modelos alternativos.

Para solventar este problema, ENMeval permite al usuario calibrar múltiples modelos en MaxEnt desde la plataforma de R y simultáneamente evaluarlos mediante diversas métricas y particiones de los datos (Muscarella et al. 2014). Estas particiones incluyen esquemas que reducen o eliminan la autocorrelación espacial (Radosavljevic & Anderson 2014) y se ajustan a distintos requerimientos según el tamaño de muestra y la configuración geográfica del sistema. Por su lado, las diversas métricas permiten evaluar los modelos de acuerdo a distintos criterios, como lo son la tasa de omisión (*omission rate*: OR),

discriminación del entorno (AUC) y complejidad en general (AICc). De esta manera, ENMeval permite al usuario escoger los modelos que mejor se ajusten a su determinado sistema y objetivos en general.¹⁶

CONSIDERACIONES FINALES

Los videos en vivo permitieron la interacción de los expositores con la audiencia en tiempo real, mientras otros usuarios continúan interactuando con los expositores a través del sitio web de los videos a partir de comentarios o correos electrónicos. En conclusión, la constante evolución de modelado de nichos y distribución de especies hace necesario un constante desarrollo de plataformas de interacción y actualización para la comunidad científica. Más esfuerzos son necesarios para que las nuevas generaciones de investigadores y revistas científicas tomen ventaja de las nuevas tecnológicas para la transferencia de información. El uso de redes sociales y seminarios en líneas son una prometedora y sustentable plataforma para la interacción entre expertos y estudiantes.

AGRADECIMIENTOS

Los seminarios en línea se realizaron con el apoyo de la Dirección de Análisis y Prioridades de la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), y el Laboratorio de Análisis Espaciales del Instituto de Biología de la Universidad Nacional Autónoma de México (UNAM). LEE fue apoyado por el Minnesota Environment and Natural Resources Trust Fund, the Minnesota Aquatic Invasive Species Research Center, and the Clean Water Land and Legacy, y agradece a Huijie Qiao por su incansable esfuerzo para el desarrollo de NicheA. ALN agradece a Jorge Lobo, Octavio Rojas y Juan L. Parra por los comentarios. LOO agradece al Posgrado en Ciencias Biológicas de la UNAM, al PAPIIT-UNAM IN112715 (2015) y a GSoC 2016 por el apoyo parcial brindado para la realización de su investigación. Agradecemos a Jorge Soberón y Eliécier Gutiérrez por sus comentarios, los cuales permitieron mejorar la versión final de este artículo.

REFERENCIAS

- Anderson, R. P. 2012. Harnessing the world's biodiversity data: Promise and peril in ecological niche modeling of species distributions. *Ann. N. Y. Acad. Sci.* 1260:66–80.

¹⁶https://youtu.be/07Ved2088U?list=PLu_3tLNPCPDZd0VvVXVC9XHcNx7L9E_e

- Anderson, R. P., & A. Raza. 2010. The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *J. Biogeogr.* 37:1378–1393.
- Barnosky, A. D., N. Matzke, S. Tomiya, G. O. U. Wogan, B. Swartz, T. B. Quental, C. Marshall, J. L. McGuire, E. L. Lindsey, K. C. Maguire, B. Mersey, & E. A. Ferrer. 2011. Has the Earth's sixth mass extinction already arrived? *Nature* 471:51–57.
- Barve, N., V. Barve, A. Jiménez-Valverde, A. Lira-Noriega, S. P. Maher, A. T. Peterson, J. Soberón, & F. Villalobos. 2011. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecol. Modell.* 222:1810–1819.
- Blonder, B., C. Lamanna, C. Violle, & B. J. Enquist. 2014. The n-dimensional hypervolume. *Glob. Ecol. Biogeogr.* 23:595–609.
- Borlaug, N. E. 1983. Contributions of conventional plant breeding to food production. *Sci.* 219:689–693.
- Castañeda-Álvarez, N. P., C. K. Khoury, H. A. Achicanoy, V. Bernau, H. Dempewolf, R. J. Eastwood, L. Guarino, R. H. Harker, A. Jarvis, N. Maxted, J. V. Müller, J. Ramirez-Villegas, C. C. Sotelo, P. C. Struik, H. Vincent, J. Toll. 2016. Global conservation priorities for crop wild relatives. *Nat. Plants* 2:16022.
- Chang W., J. Cheng, J. J. Allaire, Y. Xie, & J. McPherson. 2016. shiny: Web application framework for R. R package version 0.13.2. <https://CRAN.R-project.org/package=shiny>
- CONABIO, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. 2012. Base de datos de localidades de especies de mamíferos terrestres. Sistema Nacional de Información sobre Biodiversidad de México (SNIB). Consultado en línea en el 2012: <http://www.conabio.gob.mx/institucion/snib/doctos/acerca.html>
- CONAFOR, Comisión Nacional Forestal. 2009. Inventario Nacional Forestal y de Suelos, México 2004–2009.
- Dempewolf, H., R. J. Eastwood, L. Guarino, C. K. Khoury, J. V Müller, & J. Toll. 2013. Adapting agriculture to climate change: A global initiative to collect, conserve, and use Crop Wild Relatives. *Agroecol. and Sust. Food.* 38:369–77.
- Dolédec, S., D. Chessel, F. C. J. Ter Braak, & S. Champely. 1996. Matching species traits to environmental variables: a new three-table ordination method. *Environ. Ecol. Stat.* 3:143–146.
- Drake, J. M. 2015. Range bagging: A new method for ecological niche modelling from presence-only data. *J. R. Soc. Interface.* 12:1–9.
- Dray, S., & A. B. Dufor. 2007. The ade4 package: Implementing the duality diagram for ecologists. *J. Stat. Softw.* 22:1–20.
- Dray, S., & P. Legendre. 2008. Testing the species traits-environment relationships: The fourth-corner problem revisited. *Ecology* 89:3400–3412.
- Elith, J., & J. Leathwick. 2009. Species distribution models: Ecological explanation and prediction across space and time. *Annu. Rev. Ecol. Evol. Syst.* 40:677–697.
- Escobar, L. E., & M. E. Craft. 2016. Advances and limitations of disease biogeography using ecological niche modeling. *Front. Microbiol.* 7:1–21.
- Gutiérrez, E. E., Boria R. A. & R. P. Anderson. 2014. Can biotic interactions cause allopatry? Niche models, competition, and distributions of South American mouse opossums. *Ecography* 37: 741–753.
- Evans, M. R., V. Grimm, K. Johst, T. Knuuttila, R. de Langhe, C. M. Lessells, M. Merz, M. A. O'Malley, S. H. Orzack, M. Weisberg, D. J. Wilkinson, O. Wolkenhauer, & T. G. Benton. 2013. Do simple models lead to generality in ecology? *Trends Ecol. Evol.* 28:578–583.
- Faleiro, F. V., & R. D. Loyola. 2013. Socioeconomic and political trade-offs in biodiversity conservation: A case study of the Cerrado Biodiversity Hotspot, Brazil. *Divers. Distrib.* 19:977–987.
- Fischer, D., S. M. Thomas, M. Neteler, N. B. Tjaden, & C. Beierkuhnlein. 2014. Climatic suitability of *Aedes albopictus* in Europe referring to climate change projections: Comparison of mechanistic and correlative niche modelling approaches. *Euro Surveill.* 19:1–13.
- Franklin, J. 2010. Mapping species distributions: Spatial inference and prediction. Cambridge: Cambridge University Press.
- Franklin, J., J. M. Serra-Diaz, A. D. Syphard, & H. M. Regan. 2016. Global change and terrestrial plant community dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 113:3725–3734.
- Gepts, P. 2006. Plant genetic resources conservation and utilization: The accomplishments and future of a societal insurance policy. *Crop Sci.* 46: 2278–2292.
- Guarino, L. & D. B. Lobell. 2011. A walk on the wild Side. *Nat. Clim. Change.* 1:374–375.
- Hijmans, R. J., S. Phillips, J. Leathwick, & J. Elith. 2016. dismo: Species distribution modeling. R package version 1.0–15. <https://CRAN.R-project.org/package=dismo>
- Hutchinson, G. E. 1957. Concluding remarks. *Cold Spring Harb. Symp. Quant. Biol.* 22:415–427.
- Hutchinson, G. E. 1978. An introduction to population ecology. New Heaven: Yale University Press.
- Kadmon, R., O. Farber, & A. Danin. 2004. Effect of roadside bias on the accuracy of predictive maps produced by bioclimatic models. *Ecol Appl.* 14:401–413.

- Kaplan, A. M., & Haenlein, M. 2010. Users of the world, unite! The challenges and opportunities of social media. *Bus Horiz.* 53:59–68.
- Keenan, T., J. M. Serra, F. Lloret, M. Ninyerola, & S. Sabate. 2011. Predicting the future of forests in the Mediterranean under climate change, with niche- and process-based models: CO₂ matters! *Glob Chang Biol.* 17:565–579.
- Ladle, R., & R. J. Whittaker. Editors. 2011. *Conservation biogeography*. John Wiley & Sons. Chichester, UK: Wiley-Blackwell.
- Margules, C. R., & R. L. Pressey. 2000. Systematic conservation planning. *Nat.* 405: 243–53.
- Margules, C. R., & S. Sarkar. 2009. Planeación sistemática de la conservación. (Trad. V. Sánchez-Cordero & F. Figueroa). Universidad Nacional Autónoma de México, Comisión Nacional de Áreas Naturales Protegidas y Comisión Nacional para el Conocimiento y Uso de la Biodiversidad. 304 pp. México D.F. (Original en inglés, 2007).
- Mariac, C., V. Luong, I. Kapran, A. Mamadou, F. Sagnard, M. Deu, J. Chantreau, B. Gerard, J. Ndjounga, G. Bezançon, J. L. Pham, & Y. Vigouroux. 2006. Diversity of wild and cultivated pearl millet accessions (*Pennisetum glaucum* [L.] R. Br.) in Niger assessed by microsatellite markers. *Theor. Appl. Genet.* 114: 49–58.
- Merow C., M. J. Smirh, & J. A. Silander. 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. *Ecography.* 36:1058–1069.
- Moilanen, A., H. P. Possingham, & K. A. Wilson. 2009. Spatial conservation prioritization: past, present and future. En: A. Moilanen, K. A. Wilson, & H. P. Possingham (Eds.) *Spatial conservation prioritization: Quantitative methods and computational tools*. Oxford University Press, Oxford, pp. 260–268.
- Moilanen, A., B. J. Anderson, F. Eigenbrod, A. Heinemeyer, D. B. Roy, S. Gillings, P. R. Armsworth, K. J. Gaston, & C. D. Thomas. 2011. Balancing alternative land uses in conservation prioritization. *Ecol Appl.* 21:1419–1426.
- Munguía, M., I. Trejo, C. González-Salazar, & O. Pérez-Maqueo. 2016. Human impact gradient on mammalian biodiversity. *Glob. Ecol. Conserv.* 6:79–92.
- Muscarella, R., P. J. Galante, M. Soley-Guardia, R. A. Boria, J. M. Kass, M. Uriarte, & R. P. Anderson. 2014. ENMeval: An R package for conducting spatially independent evaluations and estimating optimal model complexity for MaxEnt ecological niche models. *Methods. Ecol. Evol.* 5:1198–1205.
- NOAA/NESDIS/NCEI. 2011. DMSP-OLS Nighttime light data. <http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html>
- Nori, J., M. S. Akmentins, R. Ghirardi, N. Frutos, & G. C. Leynaud. 2011. American bullfrog invasion in Argentina: Where should we take urgent measures? *Biodivers. Conserv.* 20:1125–1132.
- Nori, J. D. L. M. Azócar, F. B. Cruz, M. F. Bonino, & G. C. Leynaud. 2016. Translating niche features: Modelling differential exposure of Argentine reptiles to global climate change. *Austral. Ecol.* 41:373–381.
- Nori, J., P. Lemes, N. Urbina-Cardona, D. Baldo, J. Lescano, & R. Loyola. 2015. Amphibian conservation, land-use changes and protected areas: A global overview. *Biol. Conserv.* 191:367–374.
- Nori, J., J. N. Lescano, P. Illoldi-Rangel, N. Frutos, M. R. Cabrera, & G. C. Leynaud. 2013. The conflict between agricultural expansion and priority conservation areas: Making the right decisions before it is too late. *Biol. Conserv.* 159:507–513.
- Nori, J., & R. Loyola. 2015. On the worrying fate of data deficient amphibians. *PLoS ONE* 10:1–8.
- Nori, J. G. Tassarolo, G. F. Ficetola, V. Di Cola, R. Loyola, & G. C. Leynaud. 2016. Buying environmental problems: The invasive potential of imported turtles in Argentina. *Aquat. Conserv.* In press.
- Nori, J., R. Torres, J. N. Lescano, J. M. Cordier, M. E. Periago, & D. Baldo. 2016. Protected areas and spatial conservation priorities for endemic vertebrates of the Gran Chaco, one of the most threatened ecoregions of the world. *Divers. Distrib.* 22:1212–1219.
- Nori, J., J. N. Urbina-Cardona, R. D. Loyola, J. N. Lescano, & G. C. Leynaud. 2011. Climate change and American bullfrog invasion: What could we expect in South America? *PLoS ONE* 6: e25718.
- Olden, J. D., J. J. Lawler, & N. L. Poff. 2008. Machine learning methods without tears: A primer for ecologists. *Q. Rev. Biol.* 83:171–193.
- Osorio-Olvera L., V. Barve, N. Barve, & J. Soberón. 2016. nichetoolbox: From getting biodiversity data to evaluating species distribution models in a friendly GUI environment. R package version 0.1.6.0. <https://github.com/luismurao/nichetoolbox>
- Owens, H. L., L. P. Campbell, L. L. Dornak, E. E. Saupe, N. Barve, J. Soberón, K. Ingenloff, A. Lira-Noriega, C. M. Hensz, & C. E. Myers. 2013. Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecol. Modell.* 263:10–18.
- Peterson, A. T. 2014. *Mapping disease transmission risk in geographic and ecological contexts*. Baltimore: Johns Hopkins University Press.
- Peterson, A. T., M. Papeş, & J. Soberón. 2008. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecol. Modell.* 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*. Princeton: Princeton University Press.

- Phillips, S. J., & M. Dudík. 2008. Modeling of species distributions with MaxEnt: New extensions and a comprehensive evaluation. *Ecography* 31:161–175.
- Phillips, S. J., R. P. Anderson, & R. E. Shapire. 2006. Maximum entropy modeling of species geographic distributions. *Ecol. Modell.* 190:231–259.
- Qiao, H., C. Lin, Z. Jiang, & L. Ji. 2015. Marble algorithm: A solution to estimating ecological niches from Presence-only records. *Sci. Rep.* 5:14232.
- Qiao, H., A. T. Peterson, L. P. Campbell, J. Soberón, L. Ji, & L. E. Escobar. 2016. NicheA: Creating virtual species and ecological niches in multivariate environmental scenarios. *Ecography* 39:805–813.
- R Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Radosavljevic, A., & R. P. Anderson. 2014. Making better MaxEnt models of species distributions: Complexity, overfitting and evaluation. *J. Biogeogr.* 41:629–43.
- Saupe, E. E., V. Barve, C. E. Myers, J. Soberón, N. Barve, C. M. Hensz, A. T. Peterson, H. L. Owens, & A. Lira-Noriega. 2012. Variation in niche and distribution model performance: The need for a priori assessment of key causal factors. *Ecol. Modell.* 237:11–22.
- Scheller, R. M., J. B. Domingo, Brian R. Sturtevant, J. S. Williams, A.D. Rudy, E.J. Gustafson, & D. J. Mladenoff. 2007. Design, development, and application of LANDIS-II, a spatial landscape simulation model with flexible temporal and spatial resolution. *Ecol. Modell.* 201:409–419.
- Serra-Diaz, J. M., J. Franklin, Miquel Ninyerola, F.W. Davis, A. D. Syphard, H. M. Regan, & M. Ikegami. 2014. Bioclimatic velocity: The pace of species exposure to climate change. *Divers. Distrib.* 20:169–180.
- Serra-Diaz, J. M., T. F. Keenan, M. Ninyerola, S. Sabaté, C. Gracia, & F. Lloret. 2013. Geographical patterns of congruence and incongruence between correlative species distribution models and a Process-based ecophysiological growth model. *J. Biogeogr.* 40:1928–1938.
- Serra-Diaz, J. M., R. M. Scheller, A. D. Syphard, & J. Franklin. 2015. Disturbance and climate microrefugia mediate tree range shifts during climate change. *Landsc. Ecol.* 30:1039–53.
- Soberón, J. 2007. Grinnellian and Eltonian niches and geographic distributions of species. *Ecol. Lett.* 10:1115–1123.
- Soberón, J. 2010. Niche and area of distribution modeling: A population ecology perspective. *Ecography* 33:159–167
- Soberón, J., & M. Nakamura. 2009. Niches and distributional areas: Concepts, methods, and assumptions. *Proc. Natl. Acad. Sci. U.S.A.* 106:19644–19650.
- Soberón, J., & A. T. Peterson. 2005. Interpretation of models of fundamental ecological niches and species' distributional areas. *Biodivers. Infor.* 2:1–10.
- Subbarao, G. V., Ban Tomohiro, Kishii Masahiro, Ito Osamu, H. Samejima, H. Y. Wang, S. J. Pearse, S. Gopalakrishnan, K. Nakahara, A. K. M. Zakir Hossain, H. Tsujimoto & W.L. Berry. 2007. Can biological nitrification inhibition (BNI) genes from perennial *Leymus racemosus* (Triticeae) combat nitrification in wheat farming? *Plant Soil.* 299:55–64.
- Tachibana, C. 2014. A scientist's guide to social media. *Science.* 343:1032–1035.
- Valenzuela, D. G., & L. B. Vázquez. 2007. Consideraciones para priorizar la conservación de carnívoros mexicanos. En: G. Sánchez-Rojas, & A. Rojas-Martínez (Eds) *Tópicos en sistemática, biogeografía ecología y conservación de mamíferos.* Universidad Autónoma del Estado de Hidalgo, México. Pp. 197–214.
- Veloz, S. D. 2009. Spatially autocorrelated sampling falsely inflates measures of accuracy for presence only niche models. *J. Biogeogr.* 36:2290–2299.
- Warren, D. L., & S. N. Seifert. 2011. Ecological niche modeling in MaxEnt: The importance of model complexity and the performance of model selection criteria. *Ecol. Appl.* 21:335–342.
- Warren, D. L. 2012. In defense of 'niche modeling'. *Trends Ecol. Evol.* 27:497–500.
- Wisz, M. S. J. Pottier, W. D. Kissling, L. Pellissier, J. Lenoir, C.F. Damgaard, M. F. Forchhammer, J-A. Grytnes, A. Guisan, R. K. Heikkinen, T. T. Høye, I. Kuhn, M. Luoto, L. Maiorano, M-C. Nilsson, S. Normand, E. Ockinger, N. M. Schmidt, M. Termansen, A. Timmermann, D. A. Wardle, P. Aastrup, J-C. Svenning. 2012. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biol. Rev.* 88: 15–30.

§4.2. Paquete de R para modelar la distribución usando modelos basados en proceso.

Los modelos basados en proceso son aquellos modelos espacio-temporales que combinan medidas del nicho con propiedades poblacionales como la dispersión y el crecimiento poblacional. Este tipo de modelos pueden dar estimaciones de las abundancias poblacionales en un espacio geográficamente explícito [115, 109, 20]. Los modelos basados en preceso se han utilizado frecuentemente para: 1) estimar la distribución de las especies en escenarios de cambio climático; y 2) estudiar la relación entre demografía y la dinámica de rango de algunas especies invasoras [162, 80, 48, 50, 100, 115].

El paquete que desarrollé para modelar la distribución utilizando modelos basados en proceso se llama «ntb» y en este tiene como funciones principales:

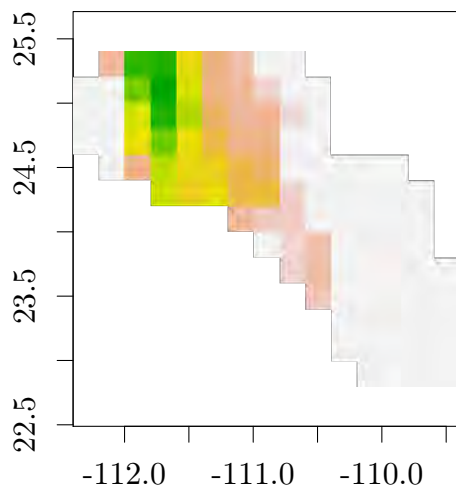
1. Una función para estimar la matriz de covarianzas usada para generar el elipsoide de volumen mínimo; esto lo realiza utilizando los valores de los puntos de presencia en el espacio de nicho.
2. Una función para ajustar el modelo de elipsoide, los outputs de esta función son:
 - a) Un raster de las idoneidades en el espacio geográfico.
 - b) Un data.frame con las idoneidades de cada pixel.
 - c) Un data.frame con las distancias Euclidianas y de Mahalanobis de las variables de nicho de cada pixel al centroide.
3. Una función para generar las tasas de dispersión entre parches del raster de idoneidad; esta función está basada en el kernel de dispersión exponencial de la ecuación 3.10.
4. Una función para correr el modelo basado en procesos (es un modelo metapoblacional). Se puede correr el modelo con y sin efecto Allee (el modelo se explica en el capítulo 2 y su ecuación es la 3.3).

A continuación se muestra el código para ajustar el modelo basado en proceso.

```
# Cargar la libreria
library(ntb)
# Primero se obtiene el mapa de idoneidades:
# Leer los datos de presencia
d_cardon <- read.csv(system.file("extdata", "cardon_virtual.csv", package = "ntb"))
# Calcular la matriz de covarianza de volumen mínimo
covar_centroid <- cov_center(d_cardon,mve=TRUE,level=0.99,vars=c(3,4,5))
# Leer los raster de las variables que se usaran para modelar el nicho
nicheStack <- stack(list.files(system.file("extdata",
                                         package = "ntb"),pattern = ".asc$",full.names = TRUE))
# Ajustar el modelo de elipsoide
ellipsoidMod <- ellipsoidfit(nicheStack,
                             covar_centroid$centroid,
                             covar_centroid$covariance,
                             level=0.95,threshold=0.05,plot=TRUE,size=3)

## NOTE: rgdal::checkCRSArgs: no proj_defs.dat in PROJ.4 shared files
```

```
# Graficar el raster de suitability  
suitabilityRaster <- ellipsoidMod$suitRaster  
plot(suitabilityRaster)  
  
## NOTE: rgdal::checkCRSArgs: no proj_defs.dat in PROJ.4 shared files  
## NOTE: rgdal::checkCRSArgs: no proj_defs.dat in PROJ.4 shared files
```



Lo que se mostró solo es el código para correr el modelo basado en proceso. El modelo metapoblacional se explica en el capítulo 3 de esta tesis y los detalles del modelo de distribución utilizando elipsoides de volumen mínimo se encuentra en el apéndice 1.

Capítulo 5

Discusión general y conclusiones generales

A lo largo de esta tesis se ha mostrado como el nicho ecológico *sensu* Grinnell [62] junto con procesos poblacionales como la dispersión, el crecimiento demográfico y las barreras bióticas pueden determinar los patrones de abundancia geográfica de una especie. En esta dirección se puso especial atención en el estudio de la relación entre abundancia y posición en el espacio de nicho bajo diferentes escenarios de dispersión (*ver* artículo del Capítulo 2) considerando al efecto Allee como barrera biológica que dificulta la colonización (Capítulo 3).

La propuesta de que existe una relación entre posición en el espacio de nicho y la abundancia poblacional, se deriva a partir de los trabajos teóricos realizados por Hutchinson [76] y de las conjeturas propuestas por Maguire [96]; este último propuso de manera explícita que diferentes regiones del espacio de nicho corresponden a distintos valores de la tasa intrínseca de crecimiento y de la abundancia poblacional. Una idea que surgió de lo anterior, es la hipótesis del «centroide del nicho», la cual menciona que las mayores abundancias se alcanzan en el centro del nicho y que conforme una población se encuentre en un sitio más alejado de este, tenderá a presentar menores abundancias [101]; de manera reciente se empezó a acumular evidencia de carácter empírico sobre este tópico [170, 185, 101, 41, 176] y en esta tesis se da una explicación plausible en términos de procesos poblacionales de las correlaciones que se han reportado entre abundancia y distancia al centroide del nicho. En este sentido con base en el análisis del modelo matemático estudiado en el artículo de investigación del capítulo §3.1, se sugieren algunas hipótesis acerca de las condiciones en las que los procesos demográficos y de dispersión pueden hacer que la correlación entre abundancia y distancia al centroide del nicho sea baja, por ejemplo, se espera que las especies cuyas tasas de migración son altas, muestren una correlación baja entre su abundancia y su posición en el espacio de nicho. Lo anterior ya ha sido puesto a prueba con datos de abundancia y modelos de nicho ecológico de 443 especies de aves de Norte América (el artículo ha sido enviado a PNAS y es colaboración con Carlos Yáñez-Arenas, Townsend Peterson y Enrique Martínez Meyer), encontrándose que efectivamente las especies migratorias tienen correlaciones más débiles que las que no lo son (figura 5.1). Este resultado concuerda con lo señalado por Pulliam [131], quien afirmó que las especies móviles frecuentemente presentan dinámicas fuente-sumidero, es decir, es posible encontrar poblaciones con abundancias altas en sitios donde la tasa de crecimiento es negativa (sumideros) [129, 131, 115, 146]; este resultado fue analizado mediante un modelo matemático a tiempo y a espacio discreto en el

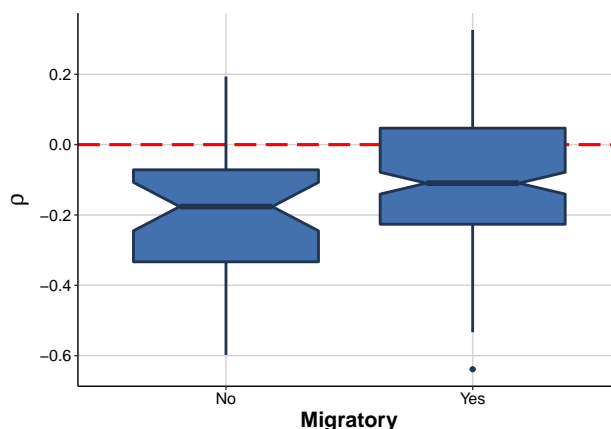


Figura 5.1: Correlaciones entre abundancia y distancia al centroide del nicho para 443 especies de aves de Norte América. 360 de estas son migratorias y 83 no lo son.

artículo que corresponde al capítulo 2 de esta tesis.

Es importante tener en cuenta que si bien la dispersión es uno de los procesos que regulan el valor de la correlación entre distancia al centroide del nicho y abundancia, existen otros factores que también pueden modificar este valor; en el capítulo 3 se analizaron los siguientes:

1. El efecto del tiempo transcurrido antes de alcanzar el estado de equilibrio poblacional (duración de los estados transitorios);
2. La estructura espacial de las idoneidades;
3. Barreras biológicas a la colonización en términos del efecto Allee.

Existe mucha controversia sobre la existencia de la relación entre abundancia y nicho; en efecto, en la literatura hay trabajos como que muestran que no clara esta relación. Por ejemplo, un trabajo reciente Dallas y colaboradores [35] analizando 1400 especies pertenecientes a 4 grupos taxonómicos distintos (1109 aves, 63 peces, 81 mamíferos y 166 especies de árboles), concluyeron de manera estadística que están relacionadas débilmente, sin embargo, Soberón y colegas (bajo revisión, Anexo 6) dan una explicación del por qué los autores obtuvieron tal resultado; una de las razones que dan Soberón et al., es que la métrica (distancia Euclidiana) que usaron en el trabajo de Dallas et al [35] no es apropiada para medir la relación debido a que no considera la estructura de la covarianza entre las variables que definen el nicho. En este sentido, en el capítulo 3 se comparó el desempeño de las métricas de Mahalanobis, Euclidiana y la salida cruda del algoritmo de modelación de nichos MaxEnt, concluyéndose que la de Mahalanobis es la que tiene mejor ajuste a los datos. Asimismo, el autor de esta tesis en colaboración con otros, compararon el desempeño de la métrica Euclidiana y la de Mahalanobis en 443 especies de Aves de Norte América y llegaron a la misma conclusión que Soberón y colaboradores (figura 5.2).

Diferentes estudios han hecho uso de las salidas directas de los algoritmos de modelación de nichos para predecir la abundancia, de manera específica varios de ellos usan la salida de MaxEnt [166, 170, 111, 176]; en tales trabajos se han encontrado tanto correlaciones altas como bajas para distintas especies. En el artículo correspondiente al capítulo 3 se sugiere que esto se debe a que

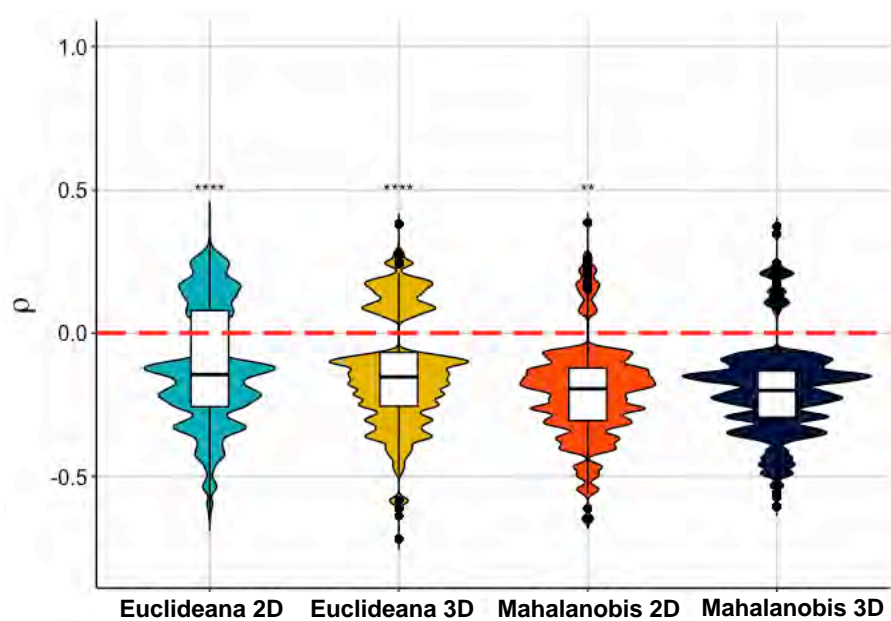


Figura 5.2: Correlaciones de Spearman (ρ) significativas de para el método Convex hull (usan distancias Euclidiana) y Elipsoides (usan distancias de Mahalanobis) para medir la relación entre distancia al centroide del nicho y la abundancia de 443 especies de Aves de Norte América. Para definir el nicho se usaron los primeros 2 y 3 componentes principales de las variables de WorldClim a una resolución de 10 minutos de arco [64].

la salida cruda de MaxEnt no es una medida de adecuación o idoneidad, sino una medida de que tan similares son los ambientes del entorno «background» con el que se ajustó MaxEnt respecto a aquellos donde se ha encontrado a la especie [107]; lo anterior sugiere que el valor de la correlación pueda cambiar en función de los datos entorno: en una de las secciones del artículo que corresponde al capítulo 3, se estudió la correlación entre abundancia y el índice idoneidad del hábitat estimado por MaxEnt usando 3 diferentes datos de entorno; observamos que efectivamente el valor de la correlación cambia para cada uno de los 3 entornos.

Sobre el efecto de los estados transitorios, el análisis del modelo que corresponde al capítulo 3 muestra que la correlación entre abundancia y nicho se hace más fuerte conforme las poblaciones alcanzan su estado de equilibrio (*ver* figura 3 y 4 del artículo del capítulo 3); este resultado nos permite conjeturar sobre la posibilidad de que algunas de las correlaciones de las poblaciones de las especies analizadas por Dallas y colaboradores [35], se encuentren alejadas de su estado de equilibrio. Lo anterior es hasta cierto punto lo que ha sido reportado también por Svenning y colaboradores [158]; de manera general se espera que las poblaciones que se encuentren lejos de su estado de equilibrio, presenten poca o ninguna relación entre abundancia y su distancia al centroide del nicho.

La dualidad de Hutchinson [30] establece que hay una correspondencia entre el espacio ambiental y el geográfico, por tanto la área de distribución está afectada por la combinación de ambos componentes. En este trabajo se muestra como es que la capacidad de dispersión además de determinar la forma de la distribución, también afecta la abundancia poblacional de formas potencialmente complejas. Esto se debe en gran parte al hecho de que la estructura topológica del espacio geográfico y el de nicho, es radicalmente diferente: puntos muy cercanos en un espacio pueden estar muy alejados en el otro (*ver* figura 3.1). La condición inicial para el comienzo de la dispersión, las barreras al movimiento o el grado de avance de la onda de expansión geográfica [158], puede causar que áreas muy idóneas tengan abundancias bajas respecto a los sitios con idoneidades relativamente bajas e incluso menores que sitios sumidero [130].

La presencia de efectos Allee implica la aparición de barreras a la colonización debido a que las poblaciones que se encuentran en sitios con un nivel bajo de idoneidad (aquellos que se encuentran lejos del óptimo de su nicho), tenderán a presentar poblaciones no viables y por lo tanto estas se extingan. Esta interacción entre movimiento, idoneidad del hábitat y efectos Allee genera patrones de distribución en forma de parches, donde hay zonas con sitios cuya idoneidad es alta pero sin poblaciones, debido a que la especie no pudo cruzar las barreras asociadas a los sitios de baja idoneidad donde es muy probable que el efecto Allee sea muy fuerte; esto concuerda con los resultados obtenidos por Keitt y colaboradores [83] quienes analizaron por medio de simulaciones numéricas la colonización de 5 sitios hipotéticos como función de la dispersión y de que tan susceptibles eran de ser ocupados (idóneos).

Finalmente, a pesar de que el enfoque de modelación utilizado en esta tesis (modelos híbridos que consideran aspectos demográficos y de nicho) es prometedor, para trabajos posteriores será necesario estudiar el papel de las interacciones bióticas -i.e. competencia, mutualismo, depredación, etc- en la relación abundancia y posición en el espacio de nicho; de esta manera la ecuación de abundancia de una especie en un espacio geográficamente explícito debe ser una función de su nicho fundamental, su capacidad de dispersión y las interacciones que tenga con las demás especies [153].

Conclusiones generales

Este trabajo busca contribuir a los fundamentos teóricos sobre la cuales se pueda avanzar en el entendimiento de la compleja red de factores que determinan la geometría y la estructura interna del área de distribución de una especie; para ello se utilizaron modelos matemáticos (a tiempo discreto y continuo) en donde se incorporaron mecanismos de dispersión y crecimiento poblacional en un espacio discreto. Las conclusiones generales de este proyecto de investigación son las siguientes:

- Existe una correlación negativa entre abundancia poblacional y distancia al centroide del nicho.
- La distancia de Mahalanobis en la mayoría de las ocasiones, es mejor predictor de la abundancia que la distancia Euclidiana.
- El análisis de los modelos muestra que la correlación entre distancia al centroide del nicho y la abundancia está regulada por los procesos de dispersión, la estructura espacial de las idoneidades y las interacciones denso-dependientes como el efecto Allee.
- La dispersión tiende a disminuir la correlación entre abundancia y posición en el espacio de nicho; de manera general, se espera que las especies con mayores tasas de dispersión presentes correlaciones de abundancia y nicho bajas.
- La población inicial, la idoneidad del sitio donde comienza la dispersión, las barreras bióticas y/o geográficas, el grado de avance de la onda de expansión puede causar que áreas muy idóneas tengan abundancias bajas respecto a los sitios con idoneidades relativamente bajas e incluso menores que sitios sumidero [130].

Bibliografía

- [1] Matthew E. Aiello-Lammens, Robert A. Boria, Aleksandar Radosavljevic, Bruno Vilela, and Robert P. Anderson. spThin: An R package for spatial thinning of species occurrence records for use in ecological niche models. *Ecography*, 38(5):541–545, 2015.
- [2] P. Anderson. *Dispersal in rodents : a resident fitness hypothesis*. The American Society of Mammalogist, 1989.
- [3] S. A. Aniadir Cain. *Foundations of Plant Geography*. Harper & Brothers, New York, 1994.
- [4] Alexandre Antonelli and Isabel Sanmartín. Mass Extinction, gradual cooling, or rapid radiation? reconstructing the spatiotemporal evolution of the ancient angiosperm genus *hedysmum* (Chloranthaceae) using empirical and simulated approaches. *Systematic Biology*, 60(5):596–615, 2011.
- [5] Lars A. Bach, René Thomsen, Cino Pertoldi, and Volker Loeschcke. Kin competition and the evolution of dispersal in an individual-based model. *Ecological Modelling*, 192(3-4):658–666, 2006.
- [6] Andrew P. Baltensperger and Falk Huettmann. Predictive spatial niche and biodiversity hotspot models for small mammal communities in Alaska: applying machine-learning to conservation planning. *Landscape Ecology*, 30(4):681–697, 2015.
- [7] Narayani Barve. ENMGadgets: Gadgets for Ecological Niche Modelling, 2016.
- [8] Narayani Barve, Vijay Barve, Alberto Jiménez-Valverde, Andrés Lira-Noriega, Sean P. Maher, a. Townsend Peterson, Jorge Soberón, and Fabricio Villalobos. The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecological Modelling*, 222(11):1810–1819, jun 2011.
- [9] Marc Bekoff. Social Play Behavior. *BioScience*, 34(4):228–233, 1984.
- [10] Graham Bell. *The Masterpiece of Nature: the Evolution and Genetics of Sexuality*. Croom Helm, London, 1982.
- [11] Comisión Nacional De Biodiversidad. Interpretation of models of fundamental ecological niches and species distributional areas. pages 1–10, 2005.
- [12] Daniel T Blumstein, Tina W Wey, and Karisa Tang. A test of the social cohesion hypothesis: interactive female marmots remain at home. *Proceedings of the Royal Society B: Biological Sciences*, 276(1669):3007–3012, 2009.

-
- [13] S. Bondrup-Nielsen and F. Karlsson. Movements and spatial patterns in populations of *Clethrionomys* species: A review. *Ann. Zool. Fennici*, 22:385–392, 1985.
- [14] Patrick John Brenchley. Extinction : Late Ordovician Mass Extinction. *Life Sciences*, pages 1–5, 2002.
- [15] Olivier Broennimann, Matthew C Fitzpatrick, Peter B Pearman, Blaise Petitpierre, Loïc Pellissier, Nigel G Yoccoz, Wilfried Thuiller, Marie-Josée Fortin, Christophe Randin, Niklaus E Zimmermann, Catherine H Graham, and Antoine Guisan. Measuring ecological niche overlap from occurrence and spatial environmental data. *Global Ecology and Biogeography*, 21(4):481–497, 2012.
- [16] T. Broquet, C. A. Johnson, E. Petit, I. Thompson, F. Burel, and J. M. Fryxell. Dispersal and genetic structure in the American marten, *Martes americana*. *Molecular Ecology*, 15(6):1689–1697, 2006.
- [17] James H. Brown. On the relationship between abundance and distribution of a species. *The American Naturalist*, 124(2):255–279, 1984.
- [18] Ryan Calsbeek. Sex-specific adult dispersal and its selective consequences in the brown anole, *Anolis sagrei*. *Journal of Animal Ecology*, 78(3):617–624, 2009.
- [19] S. P. Carroll, A. P. Hendry, D. N. Reznick, and C. W. Fox. Evolution on ecological time-scales, 2007.
- [20] Stephen Catterall, Alex R. Cook, Glenn Marion, Adam Butler, and Philip E. Hulme. Accounting for uncertainty in colonisation times: a novel approach to modelling the spatio-temporal dynamics of alien invasions using distribution data. *Ecography*, 35(10):901–911, 2012.
- [21] S. Chamberlain, K. Ram, and T. Hart. spocc: Interface to Species Occurrence Data Sources, 2015.
- [22] Wiston Chang, Joe Cheng, JJ Allaire, Yihui Xie, and Jonathan McPherson. shiny: Web Application Framework for R. 2015.
- [23] Pierre-Olivier Cheptou and Antoine Dornier. Urban metapopulation dynamics, and evolution of dispersal traits in the weed *Crepis sancta*. In Jean Clobert, Michel Baguette, Tim G. Benton, and James M. Bullock, editors, *Dispersal Ecology and Evolution*, pages 304–312. Oxford University Press, 2012.
- [24] D Chitty. *The Natural Selection of Self-regulatory Behaviour in Animal Populations*. 1967.
- [25] J. Christian. Social subordination, population density, and mammalian evolution. *Science*, 168:84–90, 1970.
- [26] E a Cittadino, M Busch, and F O Kravetz. Population abundance and dispersal in *Akodon azarae* (pampean grassland mouse) in Argentina. *Canadian Journal of Zoology-Revue Canadienne De Zoologie*, 76(6):1011–1018, 1998.
- [27] Alice L Clarke, Bernt-Erik Sæther, and Eivin Røskaft. Sex Biases in Avian Dispersal: A Reappraisal. *Oikos*, 79(3):429–438, 1997.

- [28] A. Cockburn. Habitat heterogeneity and dispersal: environmental and genetic patchiness. In N. Stenseth and W. Lindicker, editors, *Animal Dispersal*, chapter 4, pages 65–95. Springer, 1992.
- [29] B Colas, I Olivieri, and M Riba. *Centaurea corymbosa*, a cliff-dwelling species tottering on the brink of extinction: a demographic and genetic study. *Proceedings of the National Academy of Sciences of the United States of America*, 94(7):3471–3476, 1997.
- [30] Robert K. Colwell and Thiago F. Rangel. Hutchinson’s duality: The once and future niche. *Proceedings of the National Academy of Sciences USA*, 106(Supplement 2):19651–19658, 2009.
- [31] E. Conti, T. Eriksson, J. Schonemberger, K. J. Sysma, and D.A. Baum. Early Tertiary out of India dispersal of Crypteroniaceae: Evidence from phylogeny and molecular dating. *Evolution*, 56:1931–1942, 2002.
- [32] J. Cote and J. Clobert. Risky dispersal: Avoiding kin competition despite uncertainty. *Ecology*, 91(5):1485–1493, 2010.
- [33] Keith Cressman. Monitoring Desert Locusts in the Middle East: An Overview. *Yale F&ES Bulletin*, 103(Figure 1):125–140, 1996.
- [34] Beata Czyz, Marta Borowiec, Aleksandra Wasinska, Robert Pawliszko, and K. Mazaur. Breeding-season dispersal of male and female Penduline Tits (*Remiz pendulinus*) in south-western Poland. *Ornis Fennica*, 89:216–221, 2012.
- [35] Tad Dallas, Robin R Decker, and Alan Hastings. Species are not most abundant in the centre of their geographic range or climatic niche. *Ecology Letters*, pages n/a—n/a, 2017.
- [36] Marcelo de Moraes Weber and Carlos Eduardo Viveiros Grelle. Does environmental suitability explain the relative abundance of the tailed tailless bat, *Anoura caudifer*? *Natureza & Conservação*, 10(2):221–227, 2012.
- [37] Stephanie Dolrenry, Jennifer Stenglein, Leela Hazzah, R. Scott Lutz, and Laurence Frank. A metapopulation approach to African lion (*Panthera leo*) conservation. *PLoS ONE*, 9(2), 2014.
- [38] Jane Elith and John R. Leathwick. Species Distribution Models: Ecological Explanation and Prediction Across Space and Time. *Annual Review of Ecology, Evolution, and Systematics*, 40(1):677–697, 2009.
- [39] Jane Elith, Steven J Phillips, Trevor Hastie, Miroslav Dudik, Yung En Chee, and Colin J Yates. A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, 17(1):43–57, 2011.
- [40] Charles Elton. *Animal Ecology*. Sedgwick and Jackson, London, 1927.
- [41] Tania Escalante and Enrique Martínez-Meyer. Ecological niche modeling and wildlife management units (UMAS): An application to deer in campeche, México. *Tropical and Subtropical Agroecosystems*, 16(2):183–191, 2013.
- [42] ESRI. ArcGIS Desktop: Release 10.2, 2013.

- [43] Manuel Falconi Magaña. Una introducción a los modelos de competencia. In Manuel Esteva Peralta, Lourdes, Falconi Magaña, editor, *Biología Matemática. Un Enfoque Desde los Sistemas Dinámicos*, chapter 7, pages 165–187. UNAM. Facultad de Ciencias, México, DF., 2 edition, 2012.
- [44] Oren Farber and Ronen Kadmon. Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecological Modelling*, 160(1-2):115–130, 2003.
- [45] Tom Fawcett. An introduction to ROC analysis. *Pattern Recognition Letters*, 27(8):861–874, 2006.
- [46] Alan H. FIELDING and JOHN F. BELL. A review of methods for the assessment of prediction errors in conservation presence/absence models, 1997.
- [47] Jacqueline L. Frair, Evelyn H. Merrill, Darcy R. Visscher, Daniel Fortin, Hawthorne L. Beyer, and Juan M. Morales. Scales of movement by elk (*Cervus elaphus*) in response to heterogeneity in forage resources and predation risk. In *Landscape Ecology*, volume 20, pages 273–287, 2005.
- [48] Janet Franklin. Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions*, 16(3):321–330, 2010.
- [49] J.M. Fryxell. Aggregation and migration by grazing ungulates in relation to resources and predators. In P. Sinclair and A.R.E. Arcese, editors, *Serengeti II, Dynamics, Management, and Conservation of an Ecosystem*, pages 257–273. University of Chicago Press., Chicago, IL, USA., 1995.
- [50] Laure Gallien, Tamara Münkemüller, Cécile H. Albert, Isabelle Boulangeat, and Wilfried Thuiller. Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions*, 16(3):331–342, 2010.
- [51] K J Gaston. *The Structure and Dynamics of Geographic Ranges*. Oxford series in ecology and evolution. Oxford University Press, 2003.
- [52] a G Gatehouse. Behavior and ecological genetics of wind-borne migration by insects. *Annual review of entomology*, 42:475–502, 1997.
- [53] F. Gause. The Struggle for Existence. *The Yale journal of biology and medicine*, 7(6):609, 1935.
- [54] Sergey Gavrillets and Jonathan B Losos. Adaptive radiation: contrasting theory with data. *Science*, 323(5915):732–737, 2009.
- [55] Sandra M. Gelviz-Gelvez, Numa P. Pavón, Patricia Illoldi-Rangel, and Claudia Ballesteros-Barrera. Ecological niche modeling under climate change to select shrubs for ecological restoration in Central Mexico. *Ecological Engineering*, 74:302–309, 2015.
- [56] Melanie Gibbs, Marjo Saastamoinen, Aurélie Coulon, and Virginie M Stevens. Organisms on the move: ecology and evolution of dispersal. *Biology Letters*, 6(2):146–8, 2010.
- [57] J Gliwicz. Dispersal in bank voles - benefits to emigrants or to residents? *Acta Theriologica*, 38(1):31–38, 1993.

- [58] Constantino González-Salazar, Christopher R. Stephens, and Pablo A. Marquet. Comparing the relative contributions of biotic and abiotic factors as mediators of species' distributions. *Ecological Modelling*, 248:57–70, 2013.
- [59] Digpal Singh Gour, Jyotsna Bhagavatula, Maradani Bhavanishankar, Patlolla Anuradha Reddy, Jaya A. Gupta, Mriganka Shekhar Sarkar, Shaik Mohammed Hussain, Segu Harika, Ravinder Gulia, and Sisinthy Shivaji. Philopatry and Dispersal Patterns in Tiger (*Panthera tigris*). *PLoS ONE*, 8(7), 2013.
- [60] P J Greenwood and P H Harvey. The Natal and Breeding Dispersal of Birds. *Annual Review of Ecology and Systematics*, 13(1):1–21, 1982.
- [61] Paul J. Greenwood. Mating systems, philopatry and dispersal in birds and mammals. *Animal Behaviour*, 28(4):1140–1162, 1980.
- [62] Joseph Grinnell. The Niche-Relationships of the California Thrasher. *The Auk*, 34:427–433, 1917.
- [63] J E Herbert-Read, S Krause, L J Morrell, T M Schaerf, J Krause, and A J W Ward. The role of individuality in collective group movement. *Proceedings. Biological sciences / The Royal Society*, 280(1752):20122564, 2013.
- [64] R. J.; Hijmans, S. E. Cameron, J. L.; Parra, P. G.; Jones, and A. Jarvis. The WorldClim interpolated global terrestrial climate surfaces., 2005.
- [65] R J Hijmans and J Elith. Species distribution modeling with R. *R CRAN Project*, 2015.
- [66] Robert J Hijmans and Jacob van Etten. raster: Geographic analysis and modeling with raster data, 2010.
- [67] J. Hijmans, R.J, Phillips, S., Leathwick, J. and Elith. Package 'dismo'. <http://cran.r-project.org/web/packages/dismo/index.html>.
- [68] Tadao Hirota. The evolution of sex-biased dispersal by pre-dispersal copulation and fluctuating environment. *Journal of Animal Ecology*, 73(Roff 1994):1115–1120, 2004.
- [69] D. J. Hoare, I. D. Couzin, J. G J Godin, and J. Krause. Context-dependent group size choice in fish. *Animal Behaviour*, 67(1):155–164, 2004.
- [70] R. Holt and M. Gaines. Analysis of adaptation in heterogenous landscapes: implications for the evolution of fundamental niches. *Evolutionary ecology*, 6:433–447, 1992.
- [71] R. D. Holt, J H Lawton, Kevin J Gaston, and Carol A Blanchette. On the Relationship between Range Size and Local Abundance: Back to Basics. *Oikos*, 78(1):183–190, 1997.
- [72] Robert D. Holt and Michael S. Gaines. Analysis of adaptation in heterogeneous landscapes: Implications for the evolution of fundamental niches. *Evolutionary Ecology*, 6(5):433–447, 1992.
- [73] R. J. Huggett. *Fundamentals of Biogeography*. Taylor & Francis e-Library, 2002.

- [74] Colin Hughes and Ruth Eastwood. Island radiation on a continental scale: exceptional rates of plant diversification after uplift of the Andes. *Proceedings of the National Academy of Sciences of the United States of America*, 103(27):10334–10339, 2006.
- [75] Jeffrey a Hutchings and Leah Gerber. Sex-biased dispersal in a salmonid fish. *Proceedings. Biological sciences / The Royal Society*, 269(1508):2487–2493, 2002.
- [76] G. E. Hutchinson. Concluding Remarks. *Cold Spring Harbor Symposia on Quantitative Biology*, 22(0):415–427, 1957.
- [77] Duncan J. Irschick, Christopher C. Austin, Ken Petren, Robert N. Fisher, Jonathan B. Losos, and Olaf Ellers. A comparative analysis of clinging ability among pad-bearing lizards. *Biological Journal of the Linnean Society*, 59(1):21–35, 1996.
- [78] David G. Jenkins, Camille R. Brescacin, Craig V. Duxbury, Jennifer A. Elliott, Jennafer A. Evans, Katherine R. Grablow, Melissa Hillegass, Boyd N. Lyon, Genevieve A. Metzger, Michelle L. Olandese, Diana Pepe, Greg A. Silvers, Hiliary N. Suresch, Tracy N. Thompson, Christina M. Trexler, Graham E. Williams, Natalie C. Williams, and Sara E. Williams. Does size matter for dispersal distance? *Global Ecology and Biogeography*, 16(4):415–425, 2007.
- [79] M L Johnson and M S Gaines. Evolution of Dispersal: Theoretical Models and Empirical Tests Using Birds and Mammals. *Annual Review of Ecology and Systematics*, 21(1):449–480, 1990.
- [80] Eelke Jongejans, Katriona Shea, Olav Skarpaas, Dave Kelly, Andy W Sheppard, and Tim L Woodburn. Dispersal and demography contributions to population spread of *Carduus nutans* in its native and invaded ranges. *Journal of Ecology*, 96(4):687–697, 2008.
- [81] Ian D. Jonsen, Ransom A. Myers, and Joanna Mills Flemming. Meta-analysis of animal movement using state-space models, 2003.
- [82] Andrew R. Kanarek, Colleen T. Webb, Michael Barfield, and Robert D. Holt. Allee effects, aggregation, and invasion success. *Theoretical Ecology*, 6(2):153–164, 2013.
- [83] T H Keitt, M a Lewis, and R D Holt. Allee effects, invasion pinning, and species' borders. *American Naturalist*, 157(2):203–216, 2001.
- [84] Carole Kerdelhué, Lorenzo Zane, Mauro Simonato, Paola Salvato, Jérôme Rousset, Alain Roques, and Andrea Battisti. Quaternary history and contemporary patterns in a currently expanding species. *BMC evolutionary biology*, 9:220, 2009.
- [85] Ullasa Kodandaramaiah. Vagility: The neglected component in historical biogeography. *Evolutionary Biology*, 36(3):327–335, 2009.
- [86] C Krebs. A review of the Chitty hypothesis of population regulation. *Canadian J Zool*, 56:2463–2480, 1978.
- [87] X. Lambin, J. Aars, and S. Piertnet. Dispersal, intraspecific competition, kin competition and kin facilitation: a review of the empirical evidence. In J. Clobert, E. Danchin, A. Dhondt, and J. Nchols, editors, *Dispersal*. Oxford University Press, 2001.

- [88] Russell Lande. Risks of Population Extinction from Demographic and Environmental Stochasticity and Random Catastrophes. *The American Naturalist*, 142(6):911–927, 1993.
- [89] Pierre-Alexandre Landry and François-Joseph Lapointe. The genetic heterogeneity of deer mouse populations (*Peromyscus maniculatus*) in an insular landscape. *Population Ecology*, 41(3):263–268, 1999.
- [90] Allan Larson and Jonathan B Losos. Phylogenetic systematics of adaptation, 1996.
- [91] W. Lidicker. The role of dispersal in demography of small mammal populations. In F. Golley, K. Pretusewicz, and K. Ryszowski, editors, *Small mammals: their productivity and population dynamics*, pages 105–128. Cambridge University Press, 1975.
- [92] Eric B. Liebgold, Edmund D. Brodie, and Paul R. Cabe. Female philopatry and male-biased dispersal in a direct-developing salamander, *Plethodon cinereus*. *Molecular Ecology*, 20(2):249–257, 2011.
- [93] Jonathan B. Losos. Ecomorphology, Performance Capability, and Scaling of West Indian Anolis Lizards: An Evolutionary Analysis. *Ecological Monographs*, 60:369–388, 1990.
- [94] Jonathan B. Losos. *Lizards in an Evolutionary Tree: Ecology and Adaptive Radiation of Anoles*. University of California Press, 2009.
- [95] Jonathan B. Losos, Duncan J. Irschick, and T. W. Schoener. Adaptation and Constraint in the Evolution of Specialization of Bahamian Anolis Lizards. *Evolution*, 48(6):1786–1798, 1994.
- [96] Basset Maguire. Niche response structure and the analytical potentials of its relationship to the habitat. *American Naturalist*, 107(954):213–246, 1973.
- [97] A. Majumder, S. Basu, K. Sankar, Q. Qureshi, Y. V. Jhala, P. Nigam, and R. Gopal. Home ranges of the radio-collared Bengal tigers (*Panthera tigris tigris* L.) in Pench Tiger Reserve, Madhya Pradesh, Central India. *Wildlife Biology in Practice*, 8(1):36–49, 2012.
- [98] M Mangel and C Tier. Dynamics of Metapopulations with Demographic Stochasticity and Environmental Catastrophes. *Theoretical Population Biology*, 44(1):1–31, 1993.
- [99] Miguel Ângelo Marini, Morgane Barbet-Massin, Jaime Martinez, Nêmora P. Prestes, and Frédéric Jiguet. Applying ecological niche modelling to plan conservation actions for the Red-spectacled Amazon (*Amazona pretrei*). *Biological Conservation*, 143(1):102–112, 2010.
- [100] Glenn Marion, Greg J. McInerny, Jörn Pagel, Stephen Catterall, Alex R. Cook, Florian Hartig, and Robert B. O’Hara. Parameter and uncertainty estimation for process-oriented population and distribution models: data, statistics and the niche. *Journal of Biogeography*, 39(12):2225–2239, dec 2012.
- [101] Enrique Martínez-Meyer, Daniel Díaz-Porrás, Peterson Townsend, and Carlos Yáñez-Arenas. Ecological niche structure and rangewide abundance patterns of species Ecological niche structure and rangewide abundance patterns of species. *Biol. Lett.*, 9(November 2012), 2013.

- [102] P. Matos-Maraví, R. Núñez Águila, C. Peña, J. Miller, A. Sourakov, and N. Wahlberg. Causes of endemic radiation in the Caribbean: evidence from the historical biogeography and diversification of the butterfly genus *Calisto* (Nymphalidae: Satyrinae: Satyrini). *BMC evolutionary biology*, 14(199):1–18, 2014.
- [103] Erik Matthysen. Multicausality of dispersal : a review. In *Dispersal ecology and evolution*, pages 3–17. 2012.
- [104] E Mayr. *Systematics and the Origin of Species*, volume 102 Suppl. 1942.
- [105] R M Mcdowall. What biogeography is : a place for process. *Journal of Biogeography*, 31:345–351, 2004.
- [106] Brett A Melbourne and Alan Hastings. Extinction risk depends strongly on factors contributing to stochasticity. *Nature*, 454(7200):100–103, 2008.
- [107] Cory Merow, Matthew J. Smith, and John a. Silander. A practical guide to MaxEnt for modeling species’ distributions: What it does, and why inputs and settings matter. *Ecography*, 36(10):1058–1069, 2013.
- [108] Babak Naimi and Miguel B. Araújo. sdm: a reproducible and extensible R platform for species distribution modelling. *Ecography*, 39:001–008, 2016.
- [109] Hedvig K. Nenzén, Rebecca M. Swab, David a. Keith, and Miguel B. Araújo. demoniche - an R-package for simulating spatially-explicit population dynamics. *Ecography*, 35(7):577–580, jul 2012.
- [110] Markus Neteler and Helena Mitasova. *Open source GIS: A GRASS GIS approach*. Springer US, 2008.
- [111] Scott E. Nielsen, Chris J. Johnson, Douglas C. Heard, and Mark S. Boyce. Can models of presence-absence be used to scale abundance? Two case studies considering extremes in life history. *Ecography*, 28(2):197–208, 2005.
- [112] E P Odum and G W Barrett. *Fundamentals of ecology*. Thomson Brooks/Cole, 2005.
- [113] L. A. Osorio. *Determinación de la Respuesta Funcional del Ajolote (*Ambystoma mexicanum*) y modelación de su dinámica de crecimiento en presencia de Tilapia (*Oreochromis niloticus*)*. Tesis de maestría, UNAM, México D.F., 2013.
- [114] Luis Alfredo Osorio-Olvera, Manuel Falconi, and Jorge Soberón. Sobre la relación entre idoneidad del hábitat y la abundancia poblacional bajo diferentes escenarios de dispersión. *Revista Mexicana de Biodiversidad*, 87(3):1080–1088, 2016.
- [115] Jörn Pagel and Frank M. Schurr. Forecasting species ranges by statistical estimation of ecological niches and spatial population dynamics. *Global Ecology and Biogeography*, 21(2):293–304, feb 2012.
- [116] R T Paine. A note on trophic complexity and community stability. *The American Naturalist*, 103(929):91–93, 1969.

- [117] R. T. Paine. Intertidal community structure: experimental studies on the relationship between a dominant competitor and its principal predator. *Oecologia*, 15:93–120, 1974.
- [118] Robert T Paine. Food Web Complexity and Species Diversity. *The American Naturalist*, 100(910):65–75, 1966.
- [119] Robert T Paine. The Pisaster-Tegula interaction: prey patches, predator food preference, and intertidal community structure. *Ecology*, 50(6):950–961, 1969.
- [120] F. D. Panetta and N. D. Mitchell. Bioclimatic prediction of the potential distributions of some weed species prohibited entry to New Zealand, 1991.
- [121] E Paradis, S R Baillie, W J Sutherland, and R D Gregory. Patterns of natal and breeding dispersal in birds. *Journal of Animal Ecology*, 67(4):518–536, 1998.
- [122] Jennie Pearce and Simon Ferrier. The practical value of modelling relative abundance of species for regional conservation planning: A case study. *Biological Conservation*, 98(1):33–43, 2001.
- [123] N. Perrin. Dispersal. In Simon Levin, editor, *The Princeton guide to ecology*, pages 45–50. 2009.
- [124] A Townsend Peterson. Predicting the geography of species’ invasions via ecological niche modeling. *The Quarterly review of biology*, 78(4):419–433, 2003.
- [125] A. Townsend Peterson, Monica Papes, and Jorge Soberon. Rethinking receiver operating characteristic analysis applications in ecological niche modeling. *Ecological Modelling*, 213(1):63–72, 2008.
- [126] A. Townsend Peterson, Jorge Soberón, Richard G Pearson, Robert P Anderson, E Martínez-Meyer, M Nakamura, and M Bastos Araujo. *Ecological niches and geographic distributions*. 2011.
- [127] Steven J. Phillips and Miroslav Dudík. Modeling of species distributions with Maxent: New extensions and a comprehensive evaluation. *Ecography*, 31(2):161–175, 2008.
- [128] Jitka Polechová and David Storch. Ecological niche. In S. E. Jorgensen and B. D. Fath, editors, *Encyclopedia of Ecology*, pages 1088–1097. Oxford: Elsevier, 2 edition, 2008.
- [129] H Ronald Pulliam. Sources, Sinks, and Population Regulation. *The American Naturalist*, 132(5):652–661, 1988.
- [130] H Ronald Pulliam and Nickolas M Waser. Ecological Invariance and the Search for Generality in Ecology.
- [131] H.R. Pulliam. On the relationship between niche and distribution. *Ecology Letters*, 3(4):349–361, jul 2000.
- [132] QGIS Development Team. QGIS Geographic Information System, 2015.
- [133] R Core team. R: a language environment for statistical computing, 2015.

- [134] EH. Rapoport. *Aerography: geographical strategies of species*. Pergamon, Oxford, UK., 1982.
- [135] S Reddy, A Driskell, D L Rabosky, S J Hackett, and T S Schulenberg. Diversification and the adaptive radiation of the vangas of Madagascar. *Proceedings of the Royal Society B: Biological Sciences*, 279(1735):2062–2071, 2012.
- [136] Susanne S Renner. Multiple Miocene Melastomataceae dispersal between Madagascar, Africa and India. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences*, 359(1450):1485–1494, 2004.
- [137] J Ridley and W J Sutherland. Kin competition within groups: the offspring depreciation hypothesis. *Proceedings. Biological sciences / The Royal Society*, 269(1509):2559–64, 2002.
- [138] Gordon H Rodda and Julie a Savidge. Biology and Impacts of Pacific Island Invasive Species. 2. *Boiga irregularis*, the Brown Tree Snake (Reptilia: Colubridae)1. *Pacific Science*, 61(3):307–324, 2007.
- [139] Daniel Ruf, Silvia Dorn, and Dominique Mazzi. Females leave home for sex: Natal dispersal in a parasitoid with complementary sex determination. *Animal Behaviour*, 81(5):1083–1089, 2011.
- [140] Raphael D. Sagarin and Steven D. Gaines. Geographical abundance distributions of coastal invertebrates: using one-dimensional ranges to test biogeographic hypotheses. *Journal of Biogeography*, 29(8):985–997, 2002.
- [141] Raphael D. Sagarin, Steven D. Gaines, and Brian Gaylord. Moving beyond assumptions to understand abundance distributions across the ranges of species. *Trends in Ecology and Evolution*, 21(9):524–530, 2006.
- [142] Isabel Sanmartín, Paul Van Der Mark, and Fredrik Ronquist. Inferring dispersal: A Bayesian approach to phylogeny-based island biogeography, with special reference to the Canary Islands. In *Journal of Biogeography*, volume 35, pages 428–449, 2008.
- [143] Luca Santini, Moreno Di Marco, Piero Visconti, Daniele Baisero, Luigi Boitani, and Carlo Rondinini. Ecological correlates of dispersal distance in terrestrial mammals. *Hystrix, the Italian Journal of Mammalogy*, 24(2):181–186, 2013.
- [144] E. E. Saupe, Vijay Barve, C. E. Myers, Jorge Soberón, Narayani Barve, C. M. Hensz, A. T. Peterson, H. L. Owens, and Andrés Lira-Noriega. Variation in niche and distribution model performance: The need for a priori assessment of key causal factors. *Ecological Modelling*, 237-238:11–22, 2012.
- [145] Frank M. Schurr, Guy F. Midgley, Anthony G. Rebelo, Gail Reeves, Peter Poschlod, and Steven I. Higgins. Colonization and persistence ability explain the extent to which plant species fill their potential range. *Global Ecology and Biogeography*, 16(4):449–459, jul 2007.
- [146] Frank M. Schurr, Jorn Pagel, Juliano Sarmiento Cabral, Jurgen Groeneveld, Olga Bykova, Robert B. O’Hara, Florian Hartig, W. Daniel Kissling, H. Peter Linder, Guy F. Midgley, Boris Schroder, Alexander Singer, and Niklaus E. Zimmermann. How to understand species’ niches and range dynamics: a demographic research agenda for biogeography. *Journal of Biogeography*, 39(12):2146–2162, dec 2012.

- [147] Tadeu Siqueira, Luis Mauricio Bini, Marcus Vinicius Cianciaruso, Fabio Oliveira Roque, and Susana Trivinho-Strixino. The role of niche measures in explaining the abundance-distribution relationship in tropical lotic chironomids. *Hydrobiologia*, pages 1–10, 2009.
- [148] M Slatkin. Gene flow and the geographic structure of natural populations. *Science*, 236(4803):787–792, 1987.
- [149] J Soberón and A T Peterson. Interpretation of models of fundamental ecological niches and species’ distributional areas. *Biodiversity Informatics*, 2:1–10, 2005.
- [150] Jorge Soberón. Grinnellian and Eltonian niches and geographic distributions of species, 2007.
- [151] Jorge Soberón and Miguel Nakamura. Niches and distributional areas: concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences of the United States of America*, 106 Suppl:19644–19650, 2009.
- [152] Jorge Soberón and A Townsend Peterson. INTERPRETATION OF MODELS OF FUNDAMENTAL ECOLOGICAL NICHES AND SPECIES DISTRIBUTIONAL AREAS. *Biodiversity Informatics*, 2:1–10, 2005.
- [153] Jorge M. Soberón. Niche and area of distribution modeling: a population ecology perspective. *Ecography*, 33(1):159–167, feb 2010.
- [154] Mauro Enrique Souza Muñoz, Renato Giovanni, Marinez Ferreira de Siqueira, Tim Sutton, Peter Brewer, Ricardo Scachetti Pereira, Dora Ann Lange Canhos, and Vanderlei Perez Canhos. openModeller: a generic approach to species’ potential distribution m. *GeoInformatica*, 15(1):111–135, 2011.
- [155] N. Stenseth. Causes and consequences of dispersal in small mammals. In I. Swingland and P. J. Greenwood, editors, *The Ecology of Animal Movement*, pages 63–101. Clarendon Press, Oxford, 1983.
- [156] F. Stephen Dobson. Competition for mates and predominant juvenile male dispersal in mammals. *Animal Behaviour*, 30(4):1183–1192, 1982.
- [157] David Stockwell. The GARP modelling system: problems and solutions to automated spatial prediction. *International Journal of Geographical Information Science*, 13(2):143–158, 1999.
- [158] Jens-Christian Svenning and Flemming Skov. Limited filling of the potential range in European tree species. *Ecology Letters*, 7(7):565–573, 2004.
- [159] Linda L. Sweanor, Kenneth A. Logan, and Maurice G. Hornocker. Cougar dispersal patterns, metapopulation dynamics, and conservation. *Conservation Biology*, 14(3):798–808, 2000.
- [160] Z. Tablado, E. Revilla, D. Dubray, S. Said, D. Maillard, and A. Loison. From steps to home range formation: species-specific movement upscaling among sympatric ungulates. *Functional Ecology*, pages 1–13, 2015.
- [161] Anders Telenius. Biodiversity information goes public: GBIF at your service. *Nordic Journal of Botany*, 29(3):378–381, 2011.

- [162] W Thuiller, C Albert, M B Araujo, P M Berry, M Cabeza, A Guisan, T Hickler, G F Midgely, J Paterson, F M Schurr, M T Sykes, and N E Zimmermann. Predicting global change impacts on plant species' distributions: Future challenges. *Perspectives in Plant Ecology Evolution and Systematics*, 9(3-4):137–152, 2008.
- [163] Wilfried Thuiller, Bruno Lafourcade, Robin Engler, and Miguel B. Araújo. BIOMOD - A platform for ensemble forecasting of species distributions. *Ecography*, 32(3):369–373, 2009.
- [164] P C Tobin, S L Whitmire, D M Johnson, O N Bjornstad, and A M Liebhold. Invasion speed is affected by geographical variation in the strength of Allee effects. *Ecology Letters*, 10(1):36–43, 2007.
- [165] Natália Torres, Paulo de Marco Júnior, Thiago Santos, Leandro Silveira, Anah T. Jácomo de Almeida, and A. F. Diniz-Filho. Can species distribution modelling provide estimates of population densities? A case study with jaguars in the Neotropics. *Diversity and Distributions*, 18(6):615–627, 2012.
- [166] Natália M Tôrres, Paulo De Marco, Thiago Santos, Leandro Silveira, Anah T de Almeida Jácomo, and José A F Diniz-Filho. Can species distribution modelling provide estimates of population densities? A case study with jaguars in the Neotropics. *Diversity and Distributions*, 18(6):615–627, 2012.
- [167] F. Tuya, T. Wernberg, and M. S. Thomsen. The spatial arrangement of reefs alters the ecological patterns of fauna between interspersed algal habitats. *Estuarine, Coastal and Shelf Science*, 78(4):774–782, 2008.
- [168] M Udvardy. *Dynamic Zoogeography. With Special Reference to Land Animals*. van Nostrand Reinhold, New York, 1969.
- [169] Cinthya Ureña-Aranda, Octavio Rojas-Soto, Enrique Martínez-Meyer, Carlos Yañez-Arenas, Rosario Landgrave Ramírez, and Alejandro Espinosa de los Monteros. Using range-wide abundance modeling to identify key conservation areas for the micro-endemic Bolson Tortoise (*Gopherus flavomarginatus*). *PLoS ONE*, 10(6):1–14, 2015.
- [170] J Van der Wal, LP Shoo, CN Johnson, and SE Williams. Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. *American Naturalist*, 174(2):282–291, 2009.
- [171] Bram Van Moorter, Darcy Visscher, Simon Benhamou, Luca B??rger, Mark S. Boyce, and Jean Michel Gaillard. Memory keeps you at home: A mechanistic model for home range emergence. *Oikos*, 118(5):641–652, 2009.
- [172] J H Vandermeer. Niche theory. *Annual Review of Ecology and Systematics*, 3(1):107–132, 1972.
- [173] Mei Hui Wang and Mark Kot. Speeds of invasion in a model with strong or weak Allee effects. *Mathematical Biosciences*, 171(1):83–97, 2001.
- [174] Ben H. Warren, Eldredge Bermingham, Robert P. Prys-Jones, and Christophe Thebaud. Tracking island colonization history and phenotypic shifts in Indian Ocean bulbuls (Hypsipetes: Pycnonotidae). *Biological Journal of the Linnean Society*, 85(3):271–287, 2005.

- [175] Sd Webb. Mammalian faunal dynamics of the Great American interchange. *Paleobiology*, 2(3):220–234, 1976.
- [176] Marcelo M. Weber, Richard D. Stevens, José Alexandre F. Diniz-Filho, and Carlos Eduardo V. Grelle. Is there a correlation between abundance and environmental suitability derived from ecological niche modelling? A meta-analysis. *Ecography*, 40(7):817–828, jul 2017.
- [177] Jason T. Weir and Momoko Price. Andean uplift promotes lowland speciation through vicariance and dispersal in *Dendrocincla* woodcreepers. *Molecular Ecology*, 20(21):4550–4563, 2011.
- [178] Piran C. L. White and Stephen Harris. Encounters Between Red Foxes (*Vulpes-Vulpes*) : Implications for Territory Maintenance, Social Cohesion and Dispersal. *Journal of Animal Ecology*, 63(2):315–327, 1994.
- [179] John A. Wiens, Diana Stralberg, Dennis Jongsomjit, Christine A. Howell, and Mark A. Snyder. Niches, models, and climate change: Assessing the assumptions and uncertainties. *Proceedings of the National Academy of Sciences*, 106(Supplement 2):19729–19736, 2009.
- [180] E. O. Wiley and B. S. Liberman. *Phylogenics theory and practice*. Wiley-Blackwell, 2 edition, 2011.
- [181] Mary Susanne Wisz, Julien Pottier, W Daniel Kissling, Loïc Pellissier, Jonathan Lenoir, Christian F Damgaard, Carsten F Dormann, Mads C Forchhammer, John-Arvid Grytnes, Antoine Guisan, Risto K Heikkinen, Toke T Høye, Ingolf Kühn, Miska Luoto, Luigi Maiorano, Marie-Charlotte Nilsson, Signe Normand, Erik Öckinger, Niels M Schmidt, Mette Termansen, Allan Timmermann, David a Wardle, Peter Aastrup, and Jens-Christian Svenning. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological reviews of the Cambridge Philosophical Society*, 88(1):15–30, mar 2013.
- [182] Mary Susanne Wisz, Julien Pottier, W Daniel Kissling, Loïc Pellissier, Jonathan Lenoir, Christian F Damgaard, Carsten F Dormann, Mads C Forchhammer, John-Arvid Grytnes, Antoine Guisan, Risto K Heikkinen, Toke T Høye, Ingolf Kühn, Miska Luoto, Luigi Maiorano, Marie-Charlotte Nilsson, Signe Normand, Erik Öckinger, Niels M Schmidt, Mette Termansen, Allan Timmermann, David a Wardle, Peter Aastrup, and Jens-Christian Svenning. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. *Biological reviews of the Cambridge Philosophical Society*, 88(1):15–30, mar 2013.
- [183] Sewall Wright. Isolation by Distance. *Genetics*, 28(2):114–138, 1943.
- [184] Carlos Yañez-Arenas, Roger Guevara, Enrique Martínez-Meyer, Salvador Mandujano, and Jorge M Lobo. Predicting species abundances from occurrence data: Effects of sample size and bias. *Ecological Modelling*, 294:36–41, 2014.
- [185] Carlos Yañez-Arenas, Enrique Martínez-Meyer, Salvador Mandujano, and Octavio Rojas-Soto. Modelling geographic patterns of population density of the white-tailed deer in central Mexico by implementing ecological niche theory. *Oikos*, 121(March):2081–2089, 2012.

-
- [186] Yutaka Yawata, Otto X Cordero, Filippo Menolascina, Jan-Hendrik Hehemann, Martin F Polz, and Roman Stocker. Competition-dispersal tradeoff ecologically differentiates recently speciated marine bacterioplankton populations. *Proceedings of the National Academy of Sciences of the United States of America*, 111(15):5622–7, 2014.

Capítulo 6

Apéndices

§6.1. Apéndice 1

Artículo en colaboración publicado: Predictible invasion dynamics in North American populations of the Euroasian collared dove *Streptopelia decaocto*

Research



Cite this article: Ingenloff K *et al.* 2017 Predictable invasion dynamics in North American populations of the Eurasian collared dove *Streptopelia decaocto*. *Proc. R. Soc. B* **284**: 20171157.
<http://dx.doi.org/10.1098/rspb.2017.1157>

Received: 25 May 2017

Accepted: 28 July 2017

Subject Category:

Ecology

Subject Areas:

ecology, environmental science

Keywords:

invasion dynamics, ecological niche modelling, MAXENT, minimum volume ellipsoid, niche centroid, abundant centre

Author for correspondence:

A. Townsend Peterson

e-mail: town@ku.edu

[†]Present address: Florida Museum of Natural History, 1659 Museum Road, Gainesville, FL 32611, USA.

[‡]Present address: Committee on Evolutionary Biology, University of Chicago, Culver Hall 402, Chicago, IL 60637, USA.

[§]Present address: Posgrado en Ciencias Biológicas, Facultad de Ciencias, Universidad Nacional Autónoma de México, Circuito Exterior s/n, Ciudad Universitaria, 04510, Ciudad de México, México.

^{||}Present address: Entomology Department, Faculty of Science, Ain Shams University, Abbassia, Cairo, 11566, Egypt.

Electronic supplementary material is available online at <https://dx.doi.org/10.6084/m9.figshare.c.3852121>.

Predictable invasion dynamics in North American populations of the Eurasian collared dove *Streptopelia decaocto*

Kathryn Ingenloff, Christopher M. Hensz, Tashitso Anamza, Vijay Barve[†], Lindsay P. Campbell, Jacob C. Cooper[‡], Ed Komp, Laura Jimenez, Karen V. Olson, Luis Osorio-Olvera[§], Hannah L. Owens[†], A. Townsend Peterson, Abdallah M. Samy^{||}, Marianna Simões and Jorge Soberón

Biodiversity Institute, University of Kansas, Lawrence, KS 66045 USA

JCC, 0000-0003-2182-3236; ATP, 0000-0003-0243-2379

Species invasions represent a significant dimension of global change yet the dynamics of invasions remain poorly understood and are considered rather unpredictable. We explored interannual dynamics of the invasion process in the Eurasian collared dove (*Streptopelia decaocto*) and tested whether the advance of the invasion front of the species in North America relates to centrality (versus peripherality) within its estimated fundamental ecological niche. We used ecological niche modelling approaches to estimate the dimensions of the fundamental ecological niche on the Old World distribution of the species, and then transferred that model to the New World as measures of centrality versus peripherality within the niche for the species. Although our hypothesis was that the invasion front would advance faster over more favourable (i.e. more central) conditions, the reverse was the case: the invasion expanded faster in areas presenting less favourable (i.e. more peripheral) conditions for the species as it advanced across North America. This result offers a first view of a predictive approach to the dynamics of species' invasions, and thereby has relevant implications for the management of invasive species, as such a predictive understanding would allow better anticipation of coming steps and advances in the progress of invasions, important to designing and guiding effective remediation and mitigation efforts.

1. Introduction

Invasive species represent a significant agent of global change [1] with potentially serious economic and environmental implications in the form of negative effects on agriculture, public health and transportation [2,3]. Indeed, as invasive species spread across novel landscapes, they can affect broad areas, such that the ability to predict and anticipate their geographical potential is much desired [2]. Given the importance of anticipating and understanding the complexities of invasion dynamics in real-world applications, new methodological frameworks are needed for predicting species' invasions [2,4,5].

A predictive framework that has been applied to questions of species' invasions is that of ecological niche modelling (ENM; [6]). The idea is that the geographical occurrence of the species in question is related to digital maps summarizing relevant environmental parameters to estimate coarse-resolution, non-interactive dimensions of the fundamental ecological niche. Such a 'niche model' can be transferred to other regions to anticipate the invasive potential of the species on a given landscape [6], although this potential is clearly dependent on context and many contingencies. ENM transfers are not without significant complications and challenges [4,7–9]; however, the most significant shortcoming is that the models provide only a snapshot of the invasive potential of the species without consideration of the dynamics of how the species may or may not be able to access

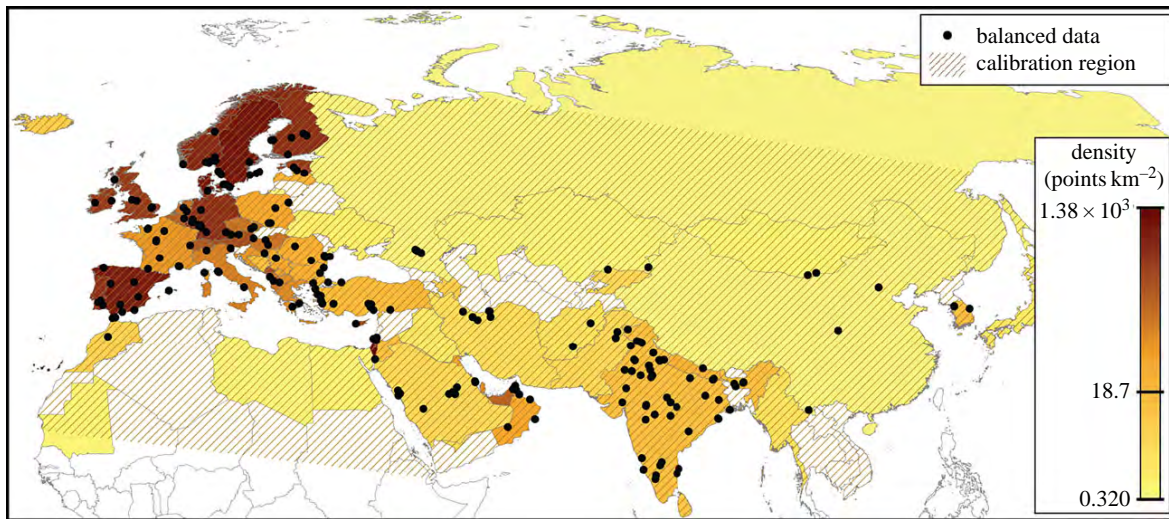


Figure 1. Balanced training data (black circles). Countries are filled on a shading ramp denoting relative density of records of Eurasian collared doves, such that lightest shading represents low density, and darkest shading denotes high density. Hatched lines show the model calibration region. Projection: Eckert III. (Online version in colour.)

different sectors of its potential distribution [10,11]. Still, ENM approaches have now been applied to many invasive species, providing useful information about their invasive potential (e.g. [12,13]).

Therefore, an important goal is to progress beyond simply estimating potential geographical distributions of species towards understanding elements of the dynamics of invasions [4]. Progress in this area has been made to incorporate dispersal to enrich model projections [14,15], yet models remain rather assumption-laden, lacking detailed information on how dispersal relates to environments across complex landscapes [16]. A crucial and related challenge is understanding how a given locality's position in fundamental niche space determines population features, such as abundance or the direction and strength of movement rates [17–19].

In this contribution, we address the latter challenge, by testing how the speed of the invasion relates to environmental centrality within the ecological niche, in the spread of Eurasian collared doves (*Streptopelia decaocto*) across North America. Our approach centres on ideas regarding relationships between the centrality of conditions within the fundamental ecological niche of species and population density or abundance [19]. We estimate coarse-resolution aspects of the fundamental ecological niche of the species in its distributional areas across Eurasia, calculate distances in environmental space from all existing conditions to the centroid of the ecological niche of the species, and classify North American environments by niche centroid distances relative to conditions estimated in Eurasia. We aim more to increase understanding of how niche structure affects invasion dynamics, rather than to obtain a detailed and specific prediction for the future distribution of Eurasian collared doves since such predictions have already been presented [20]. Overall, then, we set out to test the idea that the species would expand its range faster across more favourable conditions, a relationship that would ideally have predictive power regarding the interannual dynamics of range expansion in an invasive species.

2. Methods

As we explore and test a number of rather novel methodologies in this analysis, we provide a flowchart summarizing our methods in

the electronic supplementary material. We refer to labelled parts of this flowchart in the methods descriptions below.

(a) Study species and occurrence data

We sought a model species that: (i) is broadly invasive on continental extents (i.e. a species with distributional limits not set principally by dispersal considerations [21]), and (ii) has extensive occurrence data available across both native and invaded range areas. A clear choice was the Eurasian collared dove, which is native to Asia and limited areas of Europe, but invaded across Europe in the 1930s. In the early 1970s, it arrived in the Americas on New Providence in the Bahamas [22]. The species probably arrived in Florida during the late 1970s, with the first definitive nesting record on mainland North America coming from Homestead (Florida) in 1982 [22]. It has since spread across much of North America. We obtained occurrence data from eBird [23] and the Global Biodiversity Information Facility (GBIF; www.gbif.org, 9 February 2015); institutions providing data used in these analyses via the GBIF portal are available at <http://hdl.handle.net/1808/23373>.

We downloaded an initial set of 639 623 occurrence records, and subjected them to a variety of quality- and bias-control steps (step 'a', in the methods flowchart in the electronic supplementary material). We removed records that were extralimital for the reported observation date (particularly as regards North American data), as individuals of this species have appeared in many far-flung sites that do not result in population establishment in succeeding years, particularly in view of local releases that may have occurred. Next, we filtered travelling counts (reports of birds sighted along a path or road) in eBird data to ≤ 20 km; we also excluded all records from prior to 1970. We reduced the data by removing records lacking geographical coordinates or with latitude and longitude of 0, 0; we also discarded nine points from Australia. For Eurasia, then, the dataset reduced to 64 078 records; owing to coarse spatial resolution and other problems, an additional 7895 records were also removed, leaving 54 125 records from 63 countries (figure 1; data available at <http://hdl.handle.net/1808/23356>).

We termed this dataset the 'thinned' occurrence dataset. Because models can be affected by biases in sampling (e.g. between western Europe and central Asia; [24]), in addition to the niche-related environmental biases shown by species' occurrences, it is necessary to remove these biases to every extent possible, which we have accomplished by several means. First, we generated a 'bias layer' by extracting from the same data

sources a geographical summary of the number of records of any bird species (step 'b', in the methods flowchart in the electronic supplementary material). We derived this summary using customized code developed in R v. 3.2.5 [25], and consisted of a raster grid data layer at 5' spatial resolution (approx. 9 km at the Equator), which had pixel values ranging 0–783 376; because of the massive differences in sampling intensity among regions (i.e. five orders of magnitude), we also explored a natural-logarithm-transformed version ('ln bias') that rescaled the bias layer values to a more compact range.

In light of diverse results regarding use of bias layers in ENM calibration in previous studies (e.g. [26]), we explored an alternative approach: we further refined Eurasian collared dove occurrence data by balancing the density of occurrences on a country-by-country basis (step 'c', in the methods flowchart in the electronic supplementary material), given that much of reporting variation worldwide depends on national boundaries [27]. We then used the thinned dataset to calculate the occurrence record density per country, dividing the number of records by the land area of the country. As a reasonable intermediate-density reference point, we arbitrarily chose the record density in India (18.7 records 10^{-6} km $^{-2}$) as a reference point. Hence, for all countries with record densities of more than 25 records 10^{-6} km $^{-2}$, we reduced record density to 25 records 10^{-6} km $^{-2}$ by random sampling. In countries for which record density was less than 25 records 10^{-6} km $^{-2}$, we retained all points. The final 'balanced' occurrence dataset for Eurasia thus comprised 1170 occurrence records (figure 1). Both the thinned and balanced datasets were divided into five replicate pairs of datasets (step 'd', in the methods flowchart in the electronic supplementary material), using the 'sample' function in R [25]. Each thinned data subset consisted of 10 725 records for model calibration and 100 records for model evaluation per sub-sample; each balanced dataset consisted of 134 records for model calibration and 100 records for model evaluation per sub-sample.

We drew records from the invaded distributional area in North America only from eBird to maximize consistency and retain as much metadata as possible, and processed them as detailed in step 'a' in the methods flowchart in the electronic supplementary material. We reduced the eBird dataset to 147 350 points that had implicit or explicit spatial accuracies of ≤ 20 km and automatically removed all points with dates prior to 1970. We restricted the data to points within the United States and Canada to avoid: (i) colonizations across the Caribbean Basin, and (ii) undersampled regions in northern Mexico and Middle America. In R, we identified localities with persistent populations (i.e. localities at which the species had been recorded within 10 km in multiple years). Conclusions about persistence were not conditional on detections in sequential years, as some localities are not visited yearly nor are all established populations detected during every visit. This refined dataset resulted in yearly subsets covering 1986–2014, with three localities in 1986 and 84 444 localities in 2014; however, because expansion during 1986–1989 was minimal, we restrict our analyses to the period 1989–2014.

(b) Environmental variables

We used climate data from the WorldClim climate data archive (v. 1.4; [28]) at 5' spatial resolution. We used 15 bioclimatic variables (annual mean temperature, mean diurnal range, isothermality, temperature seasonality, maximum temperature of warmest month, minimum temperature of coldest month, temperature annual range, mean temperature of warmest quarter, mean temperature of coldest quarter, annual precipitation, precipitation of wettest month, precipitation of driest month, precipitation seasonality, precipitation of wettest quarter, and precipitation of driest quarter), eliminating four variables in light of known spatial artefacts in those layers (i.e. artificial discontinuities in the midst of gentle climate gradients). To reduce

dimensionality and variable intercorrelation, we subjected the 15 data layers to principal components analysis (PCA) based on the correlation matrix via the default *princomp* function in R [25]. We used the first five principal components, which summarized 97.1% of the overall variation in the environmental data, as summaries of environmental variation across landscapes for model calibration. In general, the first component contrasted annual means with seasonal variation, and the second contrasted temperature and precipitation variables; as such, these components provide a rich summary of the dimensions of climate across relevant regions (see the electronic supplementary material for a full summary of PCA results).

(c) Model calibration

A crucial step in calibrating ENMs appropriately is determining the region across which models are to be calibrated [29]. This region should correspond to areas that have been accessible to the species over relevant time periods [30]. We outlined an area corresponding to the limits of Eurasia, plus part of North Africa, removing distant parts of northern Asia from the calibration region (figure 1). We chose this area in view of the Atlantic, Indian and Pacific oceans acting as major barriers to dispersal; we interpreted the Sahara Desert as acting as a barrier in Africa. As regards barriers to the northeast of the species' Asian distribution, we detected no major barriers, and so relied instead on distance as a barrier.

Our model calibration efforts followed the general idea of the 'no silver bullet' approaches outlined by Qiao *et al.* [31]. That is, rather than trusting that certain modelling algorithms will perform best consistently [32], we recognize that an algorithm which may not be the 'best' on average may be the best in a given situation. As such, we submitted a variety of candidate models to detailed significance and performance testing to identify which modelling approach offered the most appropriate predictions for this particular species, landscape, and set of study goals. We developed niche models using two niche modelling methodologies: MAXENT [33] and minimum volume ellipsoids (MVEs) [34], a variation on Mahalanobis distance methods for ENM [35], each under diverse settings and conditions, varying the bias layer in MAXENT, and the *E* parameter that sets amount of omission permitted in MVEs. These two approaches are in some senses at opposite poles of the niche modelling universe: MAXENT fits highly complex response surfaces that will fit well to complex environmental landscapes [32], whereas the MVEs are simple and convex shapes that may correspond better to the idea of a fundamental ecological niche [34]. Our goal was to compare predictive performance across the five subsamples described above as an analogue to expected performance in model transfer. We calibrated models using the two occurrence datasets (thinned versus balanced), and MAXENT with one, the other, or neither of the two bias layers described above; MVEs are much less sensitive to density of points within the occurrence data cloud, and no background or pseudoabsence points are sampled in this method, so we did not consider bias for those models. We projected final models to North America for exploration of model implications.

In MAXENT, we used 50% of calibration points for training, with 10 bootstrapped replicates chosen with a random seed. In all, we explored six sets of MAXENT models: thinned versus balanced occurrence data and no bias layer, and raw and log-transformed bias layers. As the purpose of the ENMs was to characterize the niche centroid, final MAXENT runs were developed with no clamping and no extrapolation, to avoid the extrapolation that can characterize MAXENT models in model transfers [8]. We used the median of the raw outputs across 10 replicate modelling runs as a final model output.

MVEs can be used as niche models, particularly when one is interested in fitting a niche model that is only minimally affected

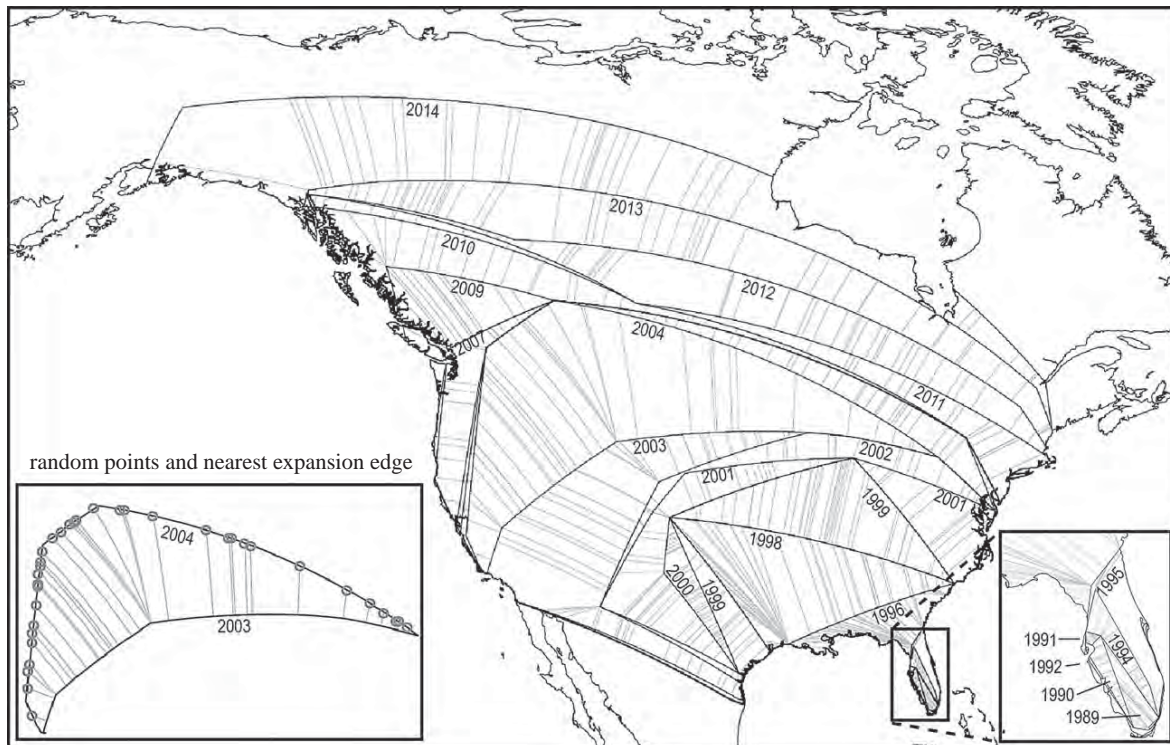


Figure 2. Range expansion map (convex hulls by year, 1989–2014; early years shown in detail in lower-right inset) and schematic illustrating what is being measured (lower-left inset; circles indicate the random points). Only selected years are labelled. Map is not projected for simplicity.

by biases among the input occurrence data [36]. MVEs can be calculated in many dimensions, and are characterized by a centroid and a symmetric positive-definite matrix that describes the directions of the axes and their lengths. We used the first five principal component axes of the environmental data to calibrate MVEs under two thresholds of allowable omission error ($E = 1\%$ and $E = 5\%$). We implemented MVEs in R v. 3.3.2 [25] using code available at <http://hdl.handle.net/1808/23357>.

(d) Model evaluation

We evaluated niche models using partial receiver operating characteristic (ROC) approaches via the `PARTIALROC` function available in the R package ‘`ENMGADGETS`’ [37]. We chose partial ROC (pROC) as a significance test in light of critiques of the appropriateness of traditional ROC approaches [38,39]. We chose an acceptable omission rate of $E = 5\%$, ran 100 iterations of pROC calculations based on random subsamples of 50% of the test data for each of the five testing subsamples, and determined probability values by direct count of iterations for which the area under the curve (AUC) ratio was ≤ 1 . In view of the need for models to anticipate the full geographical (and environmental) distribution of the species, we tested models using random subsets of the balanced occurrence dataset. MVE outputs in the form of distances to the niche centroid in environmental space were converted to similarity values by standardizing and subtracting from unity prior to pROC analysis.

(e) Invasion process in North America

Dove occurrence data in North America (described above), were partitioned by year from 1989–2014 (initial years of the invasion process, in which the species was restricted to the Florida Keys, were removed); we also removed areas outside of the continental USA owing to the need for detailed and dense occurrence data on an interannual basis. For each year, in geographical space,

we generated a convex hull around occurrences using the R package ‘`rgeos`’ [40]. We transformed all convex hulls to use the shortest geodesic path between vertex points using the package ‘`geosphere`’ [41] in R, and defined the ‘expansion edge’ as only those hull edges falling outside the hull from the previous year. We cast 50 random points along the expansion edge for each year and calculated the shortest distance to edge of the previous year’s hull (figure 2) to characterize the expansion distance associated with each point; we then used the mean of those distances to characterize the amount of expansion between each consecutive pair of years. We used these distances as measures of the amount of range expansion from one year to the next.

To summarize suitability of the landscapes across which the species was expanding its range, we used raster data layers summarizing distances in environmental space to the centroids of the ecological niches estimated across the Eurasian range of the species. This niche-centroid-distance approach is founded on previous results indicating that population abundances are a negative function of distance to the centroid in environmental space, in essence a measure of centrality versus peripherality within the niche [19,42].

To explore relationships between these niche centroid distances and population expansions, we extracted distance values for the species from: (i) the hull edge points representing the sources from which expanding populations were derived, and (ii) along the transect line connecting the hull edge and the expansion edge. For the latter, we focused on mean and maximum niche centroid distance values, as they summarize the overall suitability and the lowest suitability confronting the species; we calculated these values along the transect line using the ‘`raster`’ package [43] in R. In view of the triangular relationships that are common in such assessments (i.e. suitability defines the maximum possible abundance or density, but many points have values lower than this maximum; e.g. [44]), we used quantile regressions to assess tendencies in the top percentiles of the relationship of expansion distance to environmental suitability; as a consequence, we considered only the top 5% of the expansion distances (i.e. $\tau = 0.95$).

Table 1. Summary of partial receiver operating characteristic tests for MAXENT and minimum volume ellipsoid (MVE) models, in terms of number of non-significant ($\alpha = 0.05$) iterations, range of partial ROC AUC ratio scores, and mean AUC ratio for each sub-sample; AUC ratios above 1.0 are indicative of better-than-random model predictions. (Partial ROC (pROC) tests evaluate the statistical significance of niche model predictions in comparison with a random classifier, subject to a maximum proportional amount of omission error (E ; [39]): no. NS indicates the number out of 100 random replicate analyses for which the AUC ratio was greater than 1, pROC range and mean indicate the general magnitude of the AUC ratios across the 100 replicate analyses, and ln bias indicates results from the bias surface that was natural-log transformed.)

		thinned			balanced			
		subsample	no. NS	pROC range	mean pROC	no. NS	pROC range	mean pROC
MVE	$E = 1\%$	1	100	0.963–0.998	0.979	5	0.986–1.327	1.125
		2	81	0.963–1.063	0.989	19	0.977–1.925	1.177
		3	96	0.964–1.030	0.987	3 ^a	0.979–1.334	1.157
		4	93	0.964–1.021	0.986	2 ^a	0.983–1.303	1.133
		5	96	0.963–1.038	0.979	21	0.962–1.344	1.129
	$E = 5\%$	1	95	0.964–1.038	0.983	10	0.984–1.327	1.125
		2	60	0.963–1.087	0.996	5	0.981–1.300	1.197
		3	87	0.964–1.066	0.990	4 ^a	0.974–1.342	1.177
		4	80	0.965–1.050	0.993	1 ^a	0.980–1.326	1.210
		5	96	0.963–1.046	0.980	33	0.962–1.320	1.094
MAXENT	no bias	1	0 ^a	1.018–1.377	1.088	0 ^a	1.196–1.467	1.353
		2	0 ^a	1.032–1.393	1.109	0 ^a	1.236–1.455	1.351
		3	0 ^a	1.020–1.394	1.077	0 ^a	1.153–1.499	1.347
		4	0 ^a	1.013–1.171	1.056	0 ^a	1.320–1.494	1.367
		5	0 ^a	1.029–1.391	1.134	0 ^a	1.140–1.532	1.334
	bias	1	15	0.994–1.092	1.028	0 ^a	1.022–1.276	1.110
		2	2 ^a	0.997–1.130	1.048	0 ^a	1.010–1.207	1.054
		3	1 ^a	0.998–1.224	1.078	0 ^a	1.016–1.216	1.105
		4	8	0.997–1.153	1.047	0 ^a	1.005–1.191	1.060
		5	1 ^a	0.996–1.560	1.045	0 ^a	1.017–1.262	1.105
	ln bias	1	13	0.996–1.099	1.029	0 ^a	1.111–1.292	1.190
		2	4 ^a	0.997–1.130	1.043	0 ^a	1.076–1.245	1.153
		3	0 ^a	1.006–1.205	1.074	0 ^a	1.163–1.319	1.209
		4	4 ^a	0.999–1.259	1.046	0 ^a	1.108–1.237	1.188
		5	7 ^a	0.996–1.141	1.036	0 ^a	1.063–1.297	1.196

^aDenotes significant sub-sample evaluation.

3. Results

(a) Model calibration

We found stark imbalances in density of occurrence data across the Old World range of the species (figure 1); models based on occurrence data without a sufficient mechanism for compensation of this phenomenon showed corresponding biases towards western Europe at the expense of range areas farther to the east. MAXENT models based on thinned data without a bias layer performed particularly poorly (table 1). The ln-transformed bias layer performed better than the raw bias layer, probably owing to scale differences of several orders of magnitude in terms of numbers of occurrence records between western Europe and the rest of Eurasia. MVE models were less sensitive to the differences between the two occurrence datasets, such that differences among MVE models were smaller.

We performed a simple ANOVA on the five random subsamples, first for algorithm, finding no significant effect

($F = 4.968$, d.f. residuals = 48, $p = 0.305$). Therefore, we split the dataset into MAXENT and MVE analyses, and performed two ANOVAs using data type (thinned versus balanced) and type of bias as factors. For MAXENT, both data type and bias type were significant ($F_1 = 336.64$, $p = 1.21 \times 10^{-15}$ and $F_2 = 130.64$, $p = 1.26 \times 10^{-13}$, respectively); for MVE, data type was significant ($F_1 = 21194.37$, $p = 4.72 \times 10^{-5}$) but bias type had a rather minor effect ($F_8 = 155.79$, $p = 0.0064$). Because overall best performance was from the MAXENT model based on balanced data with no bias layer, we present those results here, but provide results from the MVE model based on balanced data in the electronic supplementary material.

(b) Model transfers to North America

Transferring the best MAXENT model (balanced occurrence data, no bias layer) indicated that the southeastern United States (except Florida) and high-latitude regions are relatively

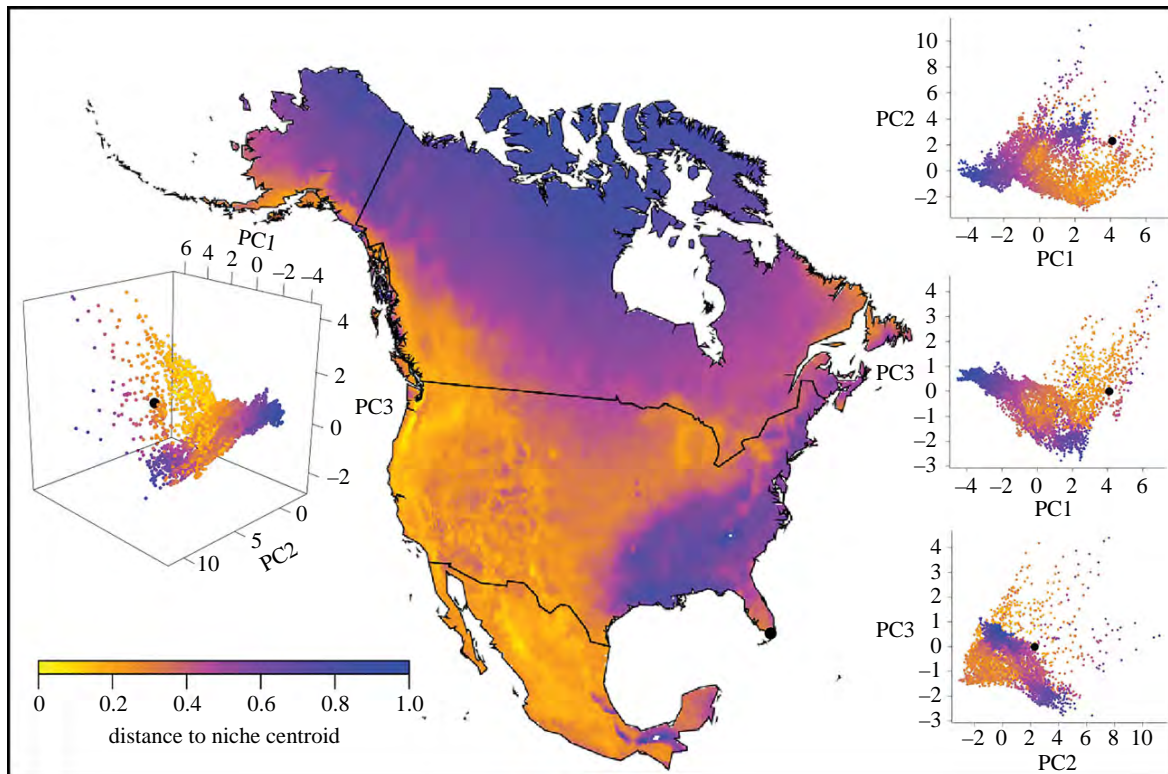


Figure 3. Median projection and visualizations in North America of the MAXENT model (balanced occurrence data, no bias layer) based on an adjusted least training presence threshold ($E = 5\%$). Insets: Projected environmental suitability. Environmental distance from calculated niche centroid is displayed on a continuous gradient scale. The niche-centroid distances are shown over geography in the middle part of the figure, and in various combinations of three-dimensional and two-dimensional plots to show variation in environmental dimensions (these dimensions represent the principal components derived from analysis of the WorldClim bioclimatic dimensions). Projection: North America Lambert Conformal Conic, for purposes of visualization.

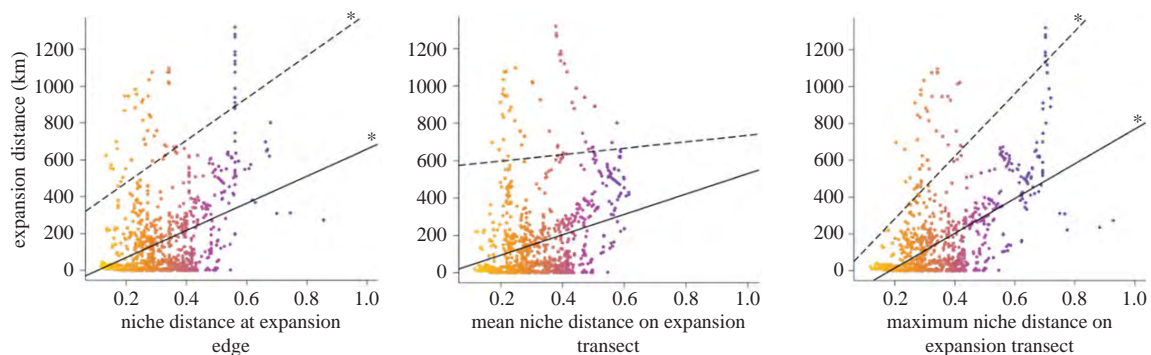


Figure 4. Plots of simple linear (solid lines) and 5% quantile (dashed lines) regressions of the final MAXENT model (balanced occurrence data, no bias layer) for distance from niche centroid to expansion edge, mean niche distance on expansion transect, and maximum niche distance on expansion transect. Asterisks (*) denote significance ($\alpha = 0.05$).

less suitable; much of the western United States was only moderately suitable (figure 3). MVE models showed high suitability across most of North America, except for a less suitable area in the southeastern United States (excluding Florida), as well as for highest-latitude regions of Alaska and Canada (electronic supplementary material, figure S1).

(c) Expansion distance and suitability

We assessed the relationship between expansion distance (figure 2) and niche centroid distance (figure 3). Although initial expectations were that the species would invade more rapidly across more suitable conditions (i.e. lower

niche centroid distances), the results were the opposite: we found positive relationships between (i) expansion distance and niche centroid distance at the hull edge in the previous year (quantile regression $R^2 = 0.0427$, $p = 0.0178$, although, again, these values may be inflated by spatial covariances), and (ii) expansion distance and maximum niche centroid distance along the expansion lines (quantile regression $R^2 = 0.1448$, $p < 0.0001$, although these values may be inflated by spatial covariances; figure 4). Relationships between expansion distance and mean niche centroid distance along the expansion lines were not significant (quantile regression $R^2 = 0.0027$, $p = 0.51$). Results for MVE-based relationships were similar (not reported).

4. Discussion

(a) Basic outcomes

Literature on species invasions emphasizes the chance nature of many aspects of invasion events [2]. Our analyses, however, point out a regularity in the process: at least in the case examined here, invasion proceeds more rapidly in areas presenting relatively unsuitable conditions. The Grinnellian ecological niche sets limits on the geographical potential of the invader as has been noted in many retrospective studies of invasive species and their spread across suitable distributional areas [6]. Clearly, the tendency for non-native species to become ‘invasive’—in the sense of expanding populations and affecting other native species—depends to some degree on the characteristics of the species [45]. The dynamics of how a species expands (once it has become invasive) across real-world landscapes to fill the potential distributional area, however, have remained rather opaque to modelling and prediction efforts.

In this study, we have appealed to recent advances in ecological niche modelling, which have connected the topology of environmental conditions within the limits of modelled ecological niches to the population biology of the species [19]. Species maintain higher abundances and densities under conditions near the centroid of the ecological niche [42], as well as higher genetic diversity [46] and more positive long-term population trends [18]. At least for Eurasian collared doves, we now note a further tie to population ecology: a species invasion proceeds more rapidly over environmental conditions less able to hold high densities of the species.

Our results offer an intriguing view of invasion dynamics. In Eurasian collared doves, the invasion front advances faster when the source areas and intervening areas are unsuitable for the species—it appears that the species has sufficient dispersal ability such that it can ‘jump’ longer distances when closer areas are not particularly attractive. When, however, no suitable areas are located beyond the unsuitable areas, of course, no long-distance colonization is possible, as can be observed in figure 2 in the northeastern United States and eastern Canada. In many ways, this suggestion runs counter to other ideas, such as the ‘Ideal Free Distribution’ of Fretwell and Calver ([47]; i.e. the best areas would fill up first, before other, less suitable areas are colonized), although this proposition depends on full information availability, which is unlikely to be fulfilled at broad extents.

(b) Caveats and limitations

We suspect that the result for Eurasian collared doves will not prove general. Rather, our guess is that further analyses of this sort will reveal dependence on species-level traits, such as reproductive rates, carrying capacities and dispersal abilities. Clearly, the geometry of suitable and unsuitable areas across landscapes will also affect results, as can be seen in this paper in the contrasting expansion patterns westwards and northwards from the southeastern United States. Certainly, though, interesting insights into invasion dynamics will derive from exploration of parallel effects and patterns in other invasive species.

This study required a number of methodological explorations and tests. Initially, we assessed diverse means of modelling niches, which took considerable time and effort in fine-tuning models and obtaining logical fits of models and

data, even on the native range of the species. Our relatively large occurrence datasets, such as for the North American range of the species, required considerable cleaning and quality assessment, such as in the careful removal of single, isolated records that did not result in future, established populations, created either by isolated dispersers or by local (unsuccessful) introductions. We also had to evaluate many approaches to testing and interpreting the distance-suitability relationships that were the focus on this proposal—an outstanding challenge remains that of distinguishing true relationships from the effects of spatial autocorrelation in assessing the significance of our quantile regressions, as we have not succeeded in developing a means of taking such relationships into account in quantile regression analyses.

We adhere closely to the idea of ‘no silver bullets’ in ecological niche modelling—the idea that no single algorithm is likely to prove to be ‘best’ across diverse modelling challenges [31]. Hence, we explored two contrasting niche modelling algorithms, four ways of managing sampling bias (thinning, balancing, bias layer, log-bias layer), and different amounts of allowable omission error in thresholding niche models. Lessons learned include the observation that bias surfaces may require transformation in order not to neglect minimally sampled areas, but also more fundamentally that a lot of experimentation is necessary to create the best possible models for these situations.

(c) Implications

In one species at least, and based on what is in effect a *post hoc* analysis, a predictable aspect can now be perceived for the rate and direction of advance of a species invasion. Clearly, this is a first result, and will require testing and exploration in other species and other regions, so as to clarify its generality. Nonetheless, many previous studies have concluded that such predictability is lacking (see review in [2]), save for expected effects of introduction (propagule) pressure [48] and natural history traits of species [49]. This study opens a new dimension of questions: how much can one anticipate the geographical course and rate of invasion across a novel region?

If such predictability exists more generally than just in this first species analysed, then this result would have interesting and important management implications. Although the usual wisdom is that once an invasive species is established, the invasion is unstoppable and eradication is impossible, attempts at mitigation and eradication are still made, and are sometimes even at least partly successful [50,51]. The potential signalled by this study would be instrumental in guiding investment of resources in such efforts.

Ethics. No ethical concerns were involved in the development of this contribution.

Data accessibility. Datasets supporting this article can be found in KU Scholarworks: <http://hdl.handle.net/1808/23356>.

Authors’ contributions. This study was conceived in the course of group discussions including all of the authors. K.I., C.H., A.M.S., M.S., T.A., V.B., L.P.C., L.J. and K.V.O. developed the statistical analyses. K.I. and C.H. executed the final analyses and created the figures. E.K. and L.O.-O. assisted with particular challenges in coding and analysis. H.L.O. and J.C.C. prepared the occurrence data and participated in the analysis stage. A.T.P. and J.S. provided conceptual frameworks. All authors edited and commented on the manuscript.

Competing interests. We have no competing interests.

Funding. J.S. and J.C.C. were supported partially by a grant from the National Science Foundation (DEB-1208472). A.M.S. was supported

by the Graduate Fulbright Egyptian Mission Program (EFMP). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements. We thank the eBird initiative and the diverse biodiversity information sources listed in <http://hdl.handle.net/1808/>

23373 for providing free and open access to the valuable data resources that they have assembled and curated. We thank the broader University of Kansas Ecological Niche Modeling group, which was the context within which we developed this project, for encouragement and input.

References

- Vitousek PM, Antonio CM, Loope LL, Westbrooks R. 1996 Biological invasions as global environmental change. *Am. Sci.* **84**, 468.
- NAS. 2002 *Predicting invasions of nonindigenous plants and plant pests*. Washington, DC: National Academy of Sciences. (<http://www.nap.edu/openbook.php?isbn=0309082641>).
- Pimentel D, Lach L, Zuniga R, Morrison D. 2000 Environmental and economic costs of nonindigenous species in the United States. *Bioscience* **50**, 53–65. (doi:10.1641/0006-3568(2000)050[0053:EAECON]2.3.CO;2)
- Gallien L, Münkemüller T, Albert CH, Boulangeat I, Thuiller W. 2010 Predicting potential distributions of invasive species: where to go from here? *Divers. Distrib.* **16**, 331–342. (doi:10.1111/j.1472-4642.2010.00652.x)
- Sakai AK *et al.* 2001 The population biology of invasive species. *Annu. Rev. Ecol. Syst.* **32**, 305–332. (doi:10.1146/annurev.ecolsys.32.081501.114037)
- Peterson AT. 2003 Predicting the geography of species' invasions via ecological niche modeling. *Q. Rev. Biol.* **78**, 419–433. (doi:10.1086/378926)
- Jiménez-Valverde A, Peterson AT, Soberón J, Overton J, Aragón P, Lobo JM. 2011 Use of niche models in invasive species risk assessments. *Biol. Invasions* **13**, 2785–2797. (doi:10.1007/s10530-011-9963-4)
- Owens HL *et al.* 2013 Constraints on interpretation of ecological niche models by limited environmental ranges on calibration areas. *Ecol. Modell.* **263**, 10–18. (doi:10.1016/j.ecolmodel.2013.04.011)
- Warren DL, Glor RE, Turelli M. 2008 Environmental niche equivalency versus conservatism: quantitative approaches to niche evolution. *Evolution* **62**, 2868–2883. (doi:10.1111/j.1558-5646.2008.00482.x)
- Arim M, Abades SR, Neill PE, Lima M, Marquet PA. 2006 Spread dynamics of invasive species. *Proc. Natl Acad. Sci. USA* **103**, 374–378. (doi:10.1073/pnas.0504272102)
- Pearson RG, Dawson TP. 2003 Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Glob. Ecol. Biogeogr.* **12**, 361–371. (doi:10.1046/j.1466-822X.2003.00042.x)
- Kulhanek SA, Leung B, Ricciardi A. 2011 Using ecological niche models to predict the abundance and impact of invasive species: application to the common carp. *Ecol. Appl.* **21**, 203–213. (doi:10.1890/09-1639.1)
- Zhu G-P, Peterson AT. In press. Do consensus models outperform individual models? Transferability evaluations of diverse modeling approaches for an invasive moth. *Biol. Invasions*. (doi:10.1007/s10530-017-1460-y)
- Merow C, Allen JM, Aiello-Lammens M, Silander JA. 2016 Improving niche and range estimates with MAXENT and point process models by integrating spatially explicit information. *Glob. Ecol. Biogeogr.* **25**, 1022–1036. (doi:10.1111/geb.12453)
- Osorio-Olvera LA, Falconi M, Soberón J. 2016 Sobre la relación entre idoneidad del hábitat y la abundancia poblacional bajo diferentes escenarios de dispersión. *Rev. Mex. Biodivers.* **87**, 1080–1088. (doi:10.1016/j.rmb.2016.07.001)
- Muthukrishnan R, West NM, Davis AS, Jordan NR, Forester JD. 2015 Evaluating the role of landscape in the spread of invasive species: the case of the biomass crop *Miscanthus × giganteus*. *Ecol. Modell.* **317**, 6–15. (doi:10.1016/j.ecolmodel.2015.08.022)
- Lira-Noriega A, Soberón J, Miller CP. 2013 Process-based and correlative modeling of desert mistletoe distribution: a multiscalar approach. *Ecosphere* **4**, 1–23. (doi:10.1890/ES13-00155.1)
- Manthey JD *et al.* 2014 A test of niche centrality as a determinant of population trends and conservation status in threatened and endangered North American birds. *Endanger. Species Res.* **26**, 201–208. (doi:10.3354/esr00646)
- Martínez-Meyer E, Díaz-Porras D, Peterson AT, Yáñez-Arenas C. 2012 Ecological niche structure determines rangewide abundance patterns of species. *Biol. Lett.* **9**, 20120637. (doi:10.1098/rsbl.2012.0637)
- Hooten MB, Wikle CK. 2008 A hierarchical Bayesian non-linear spatio-temporal model for the spread of invasive species with application to the Eurasian collared-dove. *Environ. Ecol. Stat.* **15**, 59–70. (doi:10.1007/s10651-007-0040-1)
- Saupe EE, Barve V, Myers CE, Soberón J, Barve N, Hensz CM, Peterson AT, Owens HL, Lira-Noriega A. 2012 Variation in niche and distribution model performance: the need for *a priori* assessment of key causal factors. *Ecol. Modell.* **237**, 11–22. (doi:10.1016/j.ecolmodel.2012.04.001)
- Smith PW. 1987 The Eurasian collared-dove arrives in the Americas. *Am. Birds* **41**, 1371–1379.
- eBird. 2014 *eBird Basic Dataset, version: EBD_relNov-2014*. Ithaca, NY: Cornell University.
- Peterson AT, Anamza T. 2015 Ecological niches and present and historical geographic distributions of species: a 15-year review of frameworks, results, pitfalls, and promises. *Folia Zool.* **64**, 207–217.
- R Development Core Team. 2016 *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. (<http://www.r-project.org/>).
- Campbell LP, Luther C, Moo-Llanes D, Ramsey JM, Danis-Lozano R, Peterson AT. 2015 Climate change influences on global distributions of dengue and chikungunya virus vectors. *Phil. Trans. R. Soc. B* **370**, 20140135. (doi:10.1098/rstb.2014.0135)
- Jiménez-Valverde A, Lira-Noriega A, Peterson AT, Soberón J. 2010 Marshalling existing biodiversity data to evaluate biodiversity status and trends in planning exercises. *Ecol. Res.* **25**, 947–957. (doi:10.1007/s11284-010-0753-8)
- Hijmans R, Cameron S, Parra J, Jones P, Jarvis A. 2005 Very high resolution interpolated climate surfaces for global land areas. *Int. J. Climatol.* **25**, 1965–1978. (doi:10.1002/joc.1276)
- Anderson RP, Raza A. 2010 The effect of the extent of the study region on GIS models of species geographic distributions and estimates of niche evolution: preliminary tests with montane rodents (genus *Nephelomys*) in Venezuela. *J. Biogeogr.* **37**, 1378–1393. (doi:10.1111/j.1365-2699.2010.02290.x)
- Barve N, Barve V, Jimenez-Valverde A, Lira-Noriega A, Maher SP, Peterson AT, Soberón J, Villalobos F. 2011 The crucial role of the accessible area in ecological niche modeling and species distribution modeling. *Ecol. Modell.* **222**, 1810–1819. (doi:10.1016/j.ecolmodel.2011.02.011)
- Qiao H, Soberón J, Peterson AT. 2015 No silver bullets in correlative ecological niche modeling: insights from testing among many potential algorithms for niche estimation. *Methods Ecol. Evol.* **6**, 1126–1136. (doi:10.1111/2041-210X.12397)
- Elith J *et al.* 2006 Novel methods improve prediction of species' distributions from occurrence data. *Ecography* **29**, 129–151. (doi:10.1111/j.2006.0906-7590.04596.x)
- Phillips SJ, Anderson RP, Schapire RE. 2006 Maximum entropy modeling of species geographic distributions. *Ecol. Modell.* **190**, 231–259. (doi:10.1016/j.ecolmodel.2005.03.026)
- Qiao H, Peterson AT, Campbell LP, Soberón J, Ji L, Escobar LE. 2016 NicheA: creating virtual species and ecological niches in multivariate environmental scenarios. *Ecography* **39**, 805–813. (doi:10.1111/ecog.01961)
- Farber O, Kadmon R. 2003 Assessment of alternative approaches for bioclimatic modeling with special emphasis on the Mahalanobis distance. *Ecol. Modell.*

- 160, 115–130. (doi:10.1016/S0304-3800(02)00327-7)
36. Van Aelst S, Rousseeuw P. 2009 Minimum volume ellipsoid. *Comput. Stat.* **1**, 71–82. (doi:10.1002/wics.19)
37. Barve N, Barve V. 2013 ENMGadgets: tools for pre and post processing in ENM workflows. See <https://github.com/vijaybarve/ENMGadgets>.
38. Lobo JM, Jiménez-Valverde A, Real R. 2008 AUC: a misleading measure of the performance of predictive distribution models. *Glob. Ecol. Biogeogr.* **17**, 145–151. (doi:10.1111/j.1466-8238.2007.00358.x)
39. Peterson AT, Papeş M, Soberón J. 2008 Rethinking receiver operating characteristic analysis applications in ecological niche modelling. *Ecol. Modell.* **213**, 63–72. (doi:10.1016/j.ecolmodel.2007.11.008)
40. Bivand R, Rundel C. 2015 rgeos: interface to geometry engine—open source (GEOS) R package, version 0.3-21. See <https://cran.r-project.org/package=rgeos>.
41. Hijmans RJ. 2016 Geosphere: spherical trigonometry R package, version 1.5-5. See <https://cran.r-project.org/package=geosphere>.
42. 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. (doi:10.1111/j.1600-0706.2012.20350.x)
43. Hijmans RJ, van Etten J. 2012 raster: geographic analysis and modeling with raster data, R package version 2.0-12. See <http://cran.r-project.org/package=raster>.
44. Warren PH, Gaston KJ. 1997 Interspecific abundance-occupancy relationships: a test of mechanisms using microcosms. *J. Anim. Ecol.* **66**, 730–742. (doi:10.2307/5924)
45. Moore RP, Robinson WD, Lovette IJ, Robinson TR. 2008 Experimental evidence for extreme dispersal limitation in tropical forest birds. *Ecol. Lett.* **11**, 960–968. (doi:10.1111/j.1461-0248.2008.01196.x)
46. Lira-Noriega A, Manthey JD. 2014 Relationship of genetic diversity and niche centrality: a survey and analysis. *Evolution* **68**, 1082–1093. (doi:10.1111/evo.12343)
47. Fretwell SD, Calver JS. 1969 On territorial behavior and other factors influencing habitat distribution in birds. *Acta Biotheor.* **19**, 37–44. (doi:10.1007/BF01601954)
48. Lockwood JL, Cassey P, Blackburn T. 2005 The role of propagule pressure in explaining species invasions. *Trends Ecol. Evol.* **20**, 223–228. (doi:10.1016/j.tree.2005.02.004)
49. Thuiller W, Richardson DM, Rouget M, Procheş Ş, Wilson JRU. 2006 Interactions between environment, species traits, and human uses describe patterns of plant invasions. *Ecology* **87**, 1755–1769. (doi:10.1890/0012-9658(2006)87[1755:IBESTA]2.0.CO;2)
50. Killeen GF. 2003 Following in Soper's footsteps: Northeast Brazil 63 years after eradication of *Anopheles gambiae*. *Lancet Infect. Dis.* **3**, 663–666. (doi:10.1016/S1473-3099(03)00776-X)
51. Rinderer TE, Wright JE, Shimanuki H, Parker F, Erickson E, Wilson WT. 1987 The proposed honey-bee regulated zone in Mexico. *Am. Bee J.* **127**, 160–164.

§6.2. Artículo en colaboración publicado: Complex groundwater flow systems as traveling agent models

Complex groundwater flow systems as traveling agent models

Oliver López Corona^{1,7}, Pablo Padilla², Oscar Escolero³,
Tomas González⁴, Eric Morales-Casique³ and Luis Osorio-Olvera^{5,6}

¹ Posgrado en Ciencias de la Tierra, Instituto de Geología, Universidad Nacional Autónoma de México, México D.F., Mexico

² IIMAS, Universidad Nacional Autónoma de México, México D.F., Mexico

³ Instituto de Geología, Universidad Nacional Autónoma de México, México D.F., Mexico

⁴ Instituto de Geofísica, Universidad Nacional Autónoma de México, México D.F., Mexico

⁵ Posgrado en Ciencias Biológicas, Facultad de Ciencias, Universidad Nacional Autónoma de México, México D.F., Mexico

⁶ Departamento de Matemáticas, Universidad Nacional Autónoma de México, México D.F., Mexico

⁷ Current affiliation: Theoretical Astrophysics, Instituto de Astronomía, Universidad Nacional Autónoma de México, México D.F., Mexico

ABSTRACT

Analyzing field data from pumping tests, we show that as with many other natural phenomena, groundwater flow exhibits complex dynamics described by $1/f$ power spectrum. This result is theoretically studied within an agent perspective. Using a traveling agent model, we prove that this statistical behavior emerges when the medium is complex. Some heuristic reasoning is provided to justify both spatial and dynamic complexity, as the result of the superposition of an infinite number of stochastic processes. Even more, we show that this implies that non-Kolmogorovian probability is needed for its study, and provide a set of new partial differential equations for groundwater flow.

Subjects Environmental Sciences, Computational Science, Coupled Natural and Human Systems

Keywords Hydrogeology, $1/f$ noise, Quantum game theory, Complex systems, Spatially extended games

Submitted 8 September 2013

Accepted 12 August 2014

Published 16 October 2014

Corresponding author

Oliver López Corona,
olopez@astro.unam.mx

Academic editor

Budiman Minasny

Additional Information and
Declarations can be found on
page 10

DOI 10.7717/peerj.557

© Copyright

2014 López Corona et al.

Distributed under

Creative Commons CC-BY 3.0

OPEN ACCESS

INTRODUCTION

Pink or $1/f$ noise (sometimes also called Flicker noise) is a signal or process with a frequency spectrum such that the power spectral density is inversely proportional to the frequency (*Montroll & Shlesinger, 1982; Downey, 2012*). This statistical behavior appears in such diverse phenomena as Quantum Mechanics (*Bohigas & Schmit, 1984; Faleiro et al., 2006; Haq, Pandey & Bohigas, 1982; Relanyo et al., 2002*), Biology (*Cavagna et al., 2009; Buhl et al., 2006; Boyer & López-Corona, 2009*), Medicine (*Goldberger, 2002*), Astronomy and many other fields (*Press, 1978*). Recently the universality of $1/f$ noise has been related with the manifestation of weak ergodicity breaking (*Niemann, Szendro & Kantz, 2013*) and with statistical phase transition (*López-Corona et al., 2013*).

In Geosciences the idea of self-organized criticality (SOC) associated with $1/f$ power spectrum showed to be important for example in modeling seismicity (*Bak, Tang & Weisenfeld, 1987; Bak & Tang, 1989; Bak & Chen, 1991; Sornette & Sornette, 1989*). The basic idea of SOC is that large (spatially extended) interactive systems evolve towards a state

in which a minor new event can have dramatic consequences. In seismicity this means that earthquakes contribute to organize the lithosphere both in space and time (*Sornette, Davy & Sornette, 1990*). In this context, the lithosphere may be understood as an unstable and non-linear system made of hierarchy of interacting blocks and in which dynamics has a characteristic $1/f$ signal (*Keilis-Botok, 1990*).

A particular active research field in Geoscience is the study of groundwater, which may be considered as a complex dynamic system characterized by non-stationary input (recharge), output (discharge), and response (groundwater levels). For example, groundwater levels in unconfined aquifers never reach steady state and may vary over multiple spatial and temporal scales showing fractal scaling characterized by inverse power law spectra (*Zhang & Schilling, 2004; Jianting, Young & Osterberg, 2012*). For a review on groundwater transport see *Dentz et al. (2012)*.

Spectral analysis has proven to be a powerful analytical tool for the study of variations in hydrologic processes. Ever since *Gelhar (1974)* studied temporal variations of groundwater levels for the first time with spectral analysis, it has been widely used. Spectral densities have been used in the study of: the earth tides effect on water level fluctuations (*Shih et al., 2000; Maréchal et al., 2002*); temporal scaling in discharge (*Tessier et al., 1996; Sauquet et al., 2008*); scaling in hydraulic head and river base flow (*Zhang & Schilling, 2004; Zhang & Li, 2005; Zhang & Li, 2006*); water quality variations in space and time domains (*Duffy & Gelhar, 1985; Duffy & Al-Hassan, 1988; Kirchner, Feng & Neal, 2000; Schilling et al., 2009*). Using the Detrended Fluctuation Analysis (DFA) method, *Zhongwei & You-Kuan (2007)* have proved that groundwater levels exhibit $1/f$ behavior for large time scales.

The groundwater flow process may be considered as the motion of agents (water particles) in a heterogeneous medium (*Tranouez, Bertelle & Olivier, 2001; Cortis & Knudby, 2006; Park et al., 2008*). This problem is analogous to the model of traveling agents presented in *Boyer & López-Corona (2009)*. In that model, the agents are frugivorous animals who feed on randomly located vegetation patches, in a similar way to anomalously diffusing particles in a physical context. The displacement patterns of a variety of animals as albatrosses, bumblebees, primates, gastropods, jackals, seals and sharks, among others (*Viswanathan et al., 1999; Ramos-Fernández et al., 2004; Seuront, Duponchel & Chaperon, 2007; Atkinson et al., 2002; Austin, Bowen & McMillan, 2004; Sims et al., 2008*) involve many spatio-temporal scales and are sometimes well described by Lévy walks. This is the case of the traveling agents of the model referred to *Boyer & López-Corona (2009)*. A good review on biological aspects of the subject may be found in *Miramontes, Boyer & Bartumeus (2012)* and for Lévy process see *Shlesinger, Klafter & Wong (1982)*, *Klafter, Blumen & Shlesinger (1987)* and *Lomholt et al. (2008)*.

The frequent occurrence of pink noise in a seemingly unrelated set of physical systems has prompted an extensive search for common underlying physical principles (*Miller, Miller & McWhorter, 1993*). In this paper we present a heuristic reasoning for the emergence of $1/f$ noise in groundwater and propose a new set of groundwater equations for flow in complex media (see [Supplemental Information 1](#)).

The traveling agent model

Let us consider a two-dimensional square domain of unit area with N fixed, point-like food patches randomly and uniformly distributed. Each patch contains an amount of food k .

Initially, an agent is located on a patch chosen at random. Then the following deterministic foraging rules are iteratively applied at every time step:

- (i) The agent located at patch i feeds on that patch, the fruit content decreasing by one unit: $k_i \rightarrow k_i - 1$.
- (ii) If k_i has reached the value 0, the agent chooses another patch, j , such that k_j/d_{ij} is maximal over all the allowed patches $j \neq i$ in the system, where k_j is the food content of patch j and d_{ij} the Euclidean distance between patches i and j . With this rule, the next visited patch (the “best” patch) has large food content and/or is at a short distance from i . It was assumed that the travel from i to j takes one time unit.
- (iii) The agent does not revisit previously visited patches.

This model produces complex trajectories that have been studied in detail in [Boyer et al. \(2006\)](#) and [Boyer, Miramontes & Larralde \(2009\)](#) and discussed in connection with spider monkeys foraging patterns observed in the field ([Ramos-Fernández et al., 2004](#)). Most interesting is the fact that when this model is combined with a forest one, the coupled model exhibits self-organized criticality and $1/f$ power spectrum for biomass time series ([Boyer & López-Corona, 2009](#)).

We propose that it is possible to use an equivalent model to study groundwater flow, conceptualizing it as the motion of water particles (agents) in a hydrogeological medium.

Assume the existence of a scale of support w where porous media properties can be measured. This scale of support is kept constant and is small enough such that, at the scale of the flow domain, w can be represented as a point-like quantity. Let us consider a two-dimensional square domain of unit area with N fixed, point-like Hydrogeological Units (HU) randomly and uniformly distributed. Each HU is characterized by its hydraulic flow potential, defined as $K_i = H_i/R_i$, where H_i and R_i are hydraulic head and hydraulic resistivity at point i , respectively. Thus K_i has units of time.

Initially, an agent (water particle) is located on an HU chosen at random. Then the following deterministic motion rules are iteratively applied at every time step:

- (i') An agent located in an HU stays there for a dimensionless time T proportional to $K_{max}/(K + a)$, where K_{max} is the maximum hydraulic flow potential in the domain and a is an arbitrary normalization constant such that $K_{max} \gg a$. For $K \rightarrow 0$ then the waiting time is the maximum possible; for $K \rightarrow K_{max}$ then the waiting time is the minimum possible.
- (ii') When an agent has spent T time in the HU_i , it chooses another HU_j , such that $\Delta K_{ij}/d_{ij}$ is maximal over all the allowed HU ($i \neq j$) in the domain, where $\Delta K_{ij}/d_{ij}$ is the hydraulic flow potential difference between HU_i and j , and d_{ij} is the Euclidean distance between points i and j . With this rule, the next visited HU has the largest

Table 1 Relation between media homogeneity coefficient β , type of medium, agent motion, and the noise type observed.

Homogeneity	Medium type	Agent motion type	Displacement noise type
$\beta = 2$	Inhomogenous Disordered	Random confined	White Uncorrelated
$\beta = 3$	Complex Transition point	Lévy Fractal	Pink ($1/f$) Transition point
$\beta = 5$	Homogeneous Ordered	Brownian	Brown Highly correlated

hydraulic flow potential gradient. It is assumed that the travel from i to j takes one time unit.

(iii') For a particular set of initial and boundary conditions, the agent does not revisit previously visited HU.

With this set of rules, both models, biological and groundwater flow, have the same statistical properties despite representing very different systems and then a direct analogy may be considered.

This traveling agent model exhibits some remarkable properties. Let us define the displacement of an agent $R(t) = |R(t + t_0) - R(t_0)|$ with $R(t)$ is the location of the agent at time t . For analysis, averages were taken over different and independent realizations. If the hydraulic flow potential K follows an inverse power-law distribution $P(K) = cK^{-\beta}$, where c is an arbitrary constant and β is a coefficient that represents the medium homogeneity. When β is large ($\beta \gg 1$) the medium is very homogeneous, meaning that all HU have similar values of hydraulic flow potential. On the contrary when β is small ($\beta \sim 1$) the medium is very heterogeneous, meaning that HUs with high hydraulic flow potential are numerous. The intermediate case is when $\beta = 3$ and corresponds to a complex medium where HUs with high hydraulic flow potential are present but they are not so numerous.

LÉVY WALKS AND $1/f$ DYNAMICS

In recent works (Eliazar & Klafter, 2009a; Eliazar & Klafter, 2009b; Eliazar & Klafter, 2010) proved that Lévy walks and $1/f$ are the result of systems which superimpose the transmission of many independent stochastic signals.

With this in mind, we proceeded to investigate if the power spectrum of the agent's motion follows a $1/f$ dynamic. We found a non-trivial relationship between the homogeneity coefficient β , the motion of the traveling agent and the type of noise observed. These results (summarized in Table 1) are new and differ from previous work since now the motion of the agents is explicitly analyzed.

Fifty time series for $R(t)$ were generated using the implemented traveling agents model in Boyer & López-Corona (2009) which we propose is analogous to groundwater flow. Three values of $\beta = \{2, 3, 5\}$ were considered, corresponding to disordered, complex and ordered media. Then all the 50 power spectra were averaged and fitted by an inverse power law $S(f) \sim f^{-\lambda}$. White noise correspond to a $\lambda \approx 0$, pink to a $\lambda \approx 1$, and brown to a $\lambda \approx 2$.

These results show that the emergence of pink noise for a traveling agent in a heterogeneous medium depends on its degree of heterogeneity. Thus this dynamical behavior may naturally arise from the motion of agents in a complex medium. The agents may be frugivorous monkeys, and the complex medium a rain forest; or the agents may be water particles and the medium an aquifer with a complex geology. Our results suggest that $1/f$ noise may be a fingerprint of a statistical phase transition from randomness (low correlation associated with white noise), to predictability (high correlation associated to brown noise) an idea suggested to us by A Frank (pers. comm., 2011) and discussed in *Fossion et al. (2010)*.

STUDY CASE

As part of an academic collaboration between German Karlsruhe Institute of Technology (KIT) and Mexico's National University (UNAM), pumping tests were performed on a set of urban wells in the metropolitan zone of San Luis Potosi City in Mexico (ZMSLP), which hydrogeology is described in *Martinez, Escolero & Kralish (2010)* and *Martinez, Escolero & Wolf (2011)*.

The metropolitan area is located approximately 400 km northwest of Mexico City. It lies in the San Luis Potosi Valley in the west-centre of the state of the same name at an altitude between 1,850 and 1,900 m above sea level. The area is flanked by the hills of Sierra San Miguelito to the west and Cerro San Pedro to the east; the hills have an altitude of more than 2300 m. The climate is semi-arid with an average rainfall of 356 mm between 1989 and 2006, an average annual temperature of 17.68 °C, and average annual potential evaporation of approximately 2,000 mm. The San Luis Potosi aquifer system underlies much of the surface endorheic basin. It consists of a shallow aquifer and a deep one, separated by a lens of fine material that permits very little interaction. The shallow aquifer is recharged by rainfall in the valley and the Sierra San Miguelito foothills, as well as by leaks from the urban water system. The deep aquifer is recharged in the Sierra San Miguelito and beyond. The 300 km² of shallow aquifer underlies the urban zone and its periphery. The thickness of the aquifer is within a range estimated at four to 60 m, while the depth of the phreatic level has been reported in general terms as between five and 30 m. The less deep levels are to be found within the urban zone and they deepen towards the east and northeast in the area of peripheral farmland, following the direction of the flow. The deep aquifer covers about 1,980 km² and underlies the municipalities of San Luis Potosi and Soledad de G. Sanchez, as well as part of Cerro San Pedro, Mexquitic and Zaragoza. The aquifer consists of granular material and fractured volcanic rock, and is confined over most of the flat part of the basin. Usually, wells tapping this aquifer terminate at a depth of 350–450 m and exceptionally at 700 m.

The time series from three pumping well tests, in the shallow aquifer, were analyzed. A pumping test is conducted to evaluate an aquifer by “stimulating” the aquifer through constant pumping, and observing the aquifer's response (drawdown) in observation wells. The power spectrum from all tests shows that there are two statistical regimes (*Figs. 1 and 2*). The first regime is characterized by time periods from 101 s to 103 s and $1/f$ noise

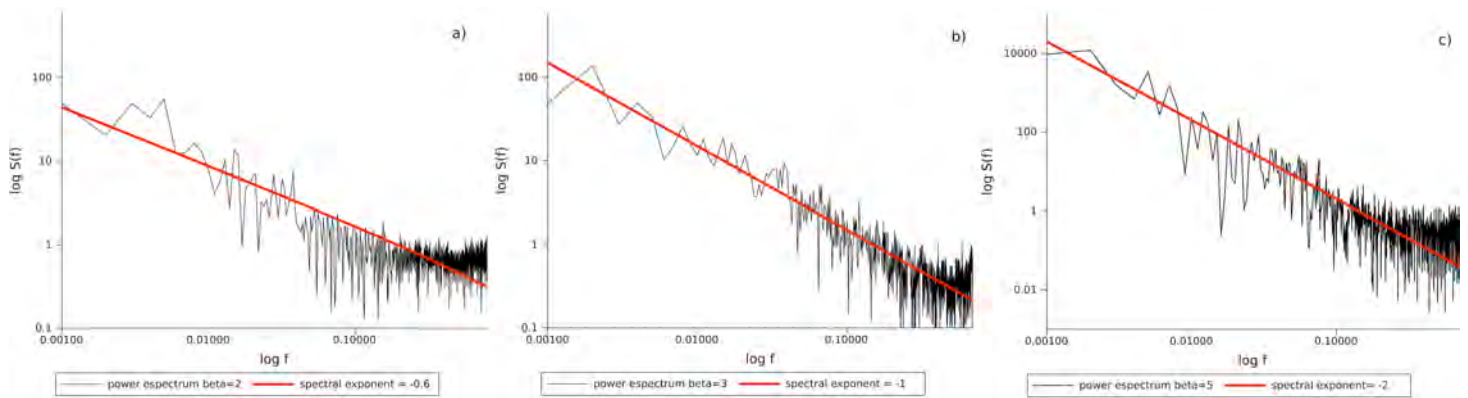


Figure 1 Power spectra for traveling agents with three values of homogeneity. First column $\beta = 2$, the medium is very inhomogeneous (disordered) and the signal is a white noise. Second column $\beta = 3$, the medium is complex and the signal is a pink noise. Third column 5, the medium is very homogeneous (ordered) and the signal is a brown noise. Power Spectrum is taken as $S(f) \equiv \hat{R}(f)\hat{R}(-f)$, where $\hat{R}(f)$ is the Fourier transformation of the displacement calculated by a Fast Fourier Transformation technique.

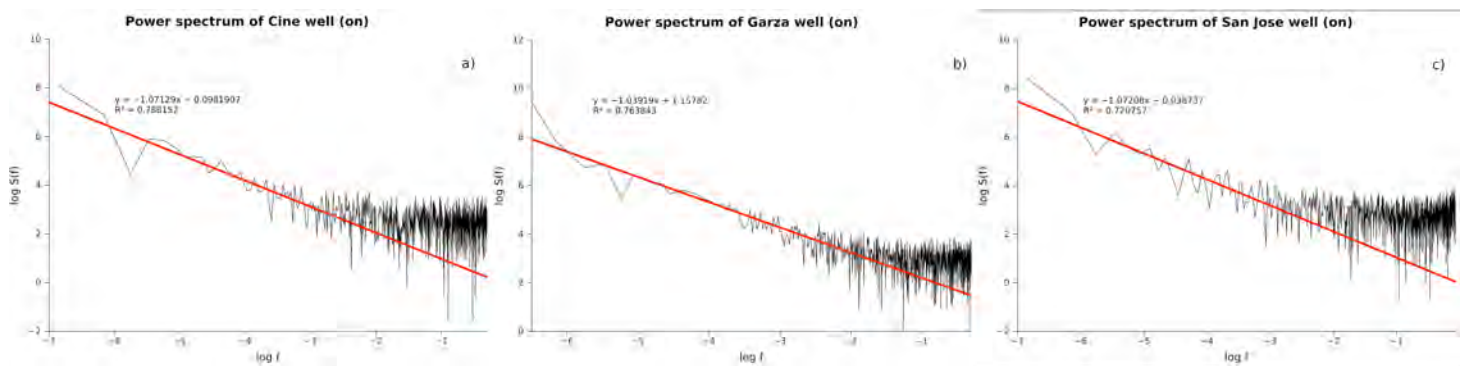


Figure 2 Power spectra for three pumping tests in the aquifer of San Luis Potosi City in Mexico. Drawdown data was acquired in 3 s intervals basis, with a total of 1800 measurements. There are two statistical regimes 101 s to 103 s with $1/f$ noise statistical behavior, and the second one with periods of seconds or less and a white noise type of signal.

statistical behavior, and the second one with periods of seconds or less and a white noise type of signal.

DISCUSSION AND CONCLUSIONS

Major sources of uncertainty have been identified in groundwater modeling. Model parameters are uncertain because they are usually measured at a few locations which are not enough to fully characterize the high degree of spatial variability at all length scales; thus, it is impossible to find a unique set of parameters to represent reality correctly. Stresses and boundary conditions are also uncertain; the extraction of water through wells and vertical recharge due to rain are not known exactly and they must be provided to the model; lateral boundaries are often virtual boundaries and water exchange through them is usually uncertain. Even model structure can be uncertain because a mathematical model is an approximation of reality and thus some physical processes are not completely known or partially represented (Neuman, 2003). In fact, the problem of characterizing subsurface

heterogeneity has been one of the biggest obstacles in constructing realistic models of groundwater flow (Fleckenstein, Niswonger & Fogg, 2006). Koltermann & Gorelick (1996) and De Marsily et al. (1998) present a good review on the subject.

Prediction with classical deterministic process models is constrained by several mathematical limitations. For one side, there is measurement error, non-linearity and sensitivity to boundary conditions (chaos) and on the other side we most face model error and inaccessible or uncertain parameters and variables (Little & Bloomfield, 2010). For these reasons, systematic oversimplifications in groundwater problems have been commonly made, under the assumption that if the most important processes are identified, groundwater flow may be sufficiently characterized.

On the other hand, Kirchner, Feng & Neal (2000) found that long-term, time series measurements of chloride, a natural passive tracer, in runoff in catchments exhibits a $1/f$ dynamic and later (Scher, Cortis & Berkowitz, 2002) gave a physical model to explain these findings in terms of CTRW.

Significant deviations from standard solutions have been observed in pumping tests (Raghavan, 2004). Moreover, it has been reported that $1/f$ dynamics are observed in time series of pumping test (Zhongwei & You-Kuan, 2007) and we showed evidence that support their findings. One approach to deal with this anomalous behavior has been to formulate the groundwater flow problem in the continuous time random walk (CTRW) framework (Cortis & Knudby, 2006). Alternatively we propose a traveling agent model for groundwater flow. The model proposed is an analogy of a previous one presented by Boyer & López-Corona (2009) which was used to construct time series for agent's mean-displacement. In agreement with field results, the model generates $1/f$ dynamics when the ambient where the agent's move is complex. For this type of medium, the step length follows a power law distribution $P(l) \sim l^{-a}$ with $a \approx 2$; the waiting time distribution follows a power law $y(t) \sim t^{-d}$ with $d = 2$ and the mean displacement a power law $\langle R^2 \rangle \sim Tg$ with $g \approx 1.2$ (Boyer & López-Corona, 2009; López-Corona, 2007). If the process was a CTRW then the following relationship should hold $g = 2 + d - a$ and a value of $g = 2$ would be expected (Klafter, Zumofen & Shlesinger, 1995). This suggests that groundwater flow is even more complex than a CTRW, which, in fact, also occurs in spider monkeys' foraging process for which $g = 1.7$ (Ramos-Fernández et al., 2004). In this sense, the model proposed could be a forward step in the study of groundwater flow complexity.

Another advantage of the traveling agent model for explaining the emergence of $1/f$ is that we may identify in which type of hydrogeological medium this kind of dynamic behavior is observed. We proved that pink noise is present when the environment heterogeneities in which the agents are moving are distributed as a power law with a scaling exponent of $\beta = 3$, corresponding to a complex medium. Labat et al. (2011) has pointed out that the complex characteristics of karst aquifers make their exploitation more complicated than other porous or fractured aquifers. These types of aquifers are spatially complex (as our $\beta = 3$ medium) groundwater systems characterized by an inherent temporal non-stationarity and nonlinearity of their hydrological response.

Eliazar & Klafter (2009a) and *Eliazar & Klafter (2009b)* have proven that the $1/f$ statistical dynamic is originated by the superposition of an infinite number of stochastic processes. This suggests that for complex media (as karstic or rock fractured aquifer) no groundwater modeling simplification is valid. This ambient induces $1/f$ noise and an infinite number of stochastic processes are in play. Therefore, the assumption that groundwater flow may be sufficiently characterized if the most important processes are identified is no longer valid.

Even more, the results may be interpreted also from a physical standpoint; the observable macroscopic behavior of a hydrogeological system at a given location is the result of the superposition of different physical processes at different scales, such as: diurnal barometric variations that affect groundwater levels, temporal fluctuations in recharge rates, moon's gravitational effects over the aquifer, tide variations in coastal aquifers, variations in the income flow from rivers and discharge through base flow, temporal increase on total stress due to trains, the effect of extraordinary recharge events provoked by an hurricane presence, and the regime of operation of wells in the area. In *Labat et al. (2011)* it has been proved, using DFA analysis, that in karstic stream flow fluctuations there are three distinct temporal scale ranges: from 1 h to around 100 h, from around 100 h up to 1 year and scales larger to 1 year. Fluctuations in flow show a clearly anti-correlated behavior on time scales above 1 year, with a slope around 0.3 corresponding to white noise. In the intermediate regime from a few days up to 1 year, a positive Hurst effect is observed, with a slope around 0.8 (almost a $1/f$ noise) as expected. On time scales below the crossover at a few days, the scaling behavior is highly non-stationary and corresponds to a random walk with positively correlated steps (with a slope around 1.75, near a Brown noise type). The authors explain these findings from a hydrogeological point of view. The first temporal scale, 1 to 100 h, is interpreted as the rapid response of the aquifer (associated with the main drain in the karstic system) to the rainfall; the second temporal scale, 100 h to 1 year, is the global response of the aquifer to rainfall input including the temporal structure of the peak flow; the third temporal scale, larger than 1 year, corresponds to the annual response of rainfall input including the regulation of the discharge via annex systems in the saturated zone. It has also been suggested that an explanation for the scale invariance of groundwater levels involve the system response to constantly changing driving inputs and boundary conditions, including boundaries imposed by management regimes, (*Little & Bloomfield, 2010*). In this way, the $1/f$ power spectrum observed in groundwater time series may be originated by both complexity of the geological medium and the presence of complex external factors (as time dependent boundary conditions).

Given this, either we accept that these types of complex groundwater systems are not suited to being modeled or we learn to deal with this infinite superposition of stochastic processes. Once groundwater flow is modeled on a traveling agent framework, we propose to describe it as a spatially extended game. Using this approach we have been able to deduce a set of partial differential equations starting from the discrete description of the model (the details of the derivation are presented as [Supplemental Information 1](#)). The

probability of finding an agent (water particle) in the position (x, y) at the time t obeys

$$\partial_t P(x, y, t) = \text{div}[e(x, y, t) \nabla P], \quad (1)$$

where $e(x, y, t)$ is the strategy (micro-physics of the flow process) that the agent in (x, y) plays at time t . The strategy in turn obeys the equation

$$\partial_t e(x, y, t) = -\text{div}[D_1(x, y, t)e(x, y, t)] + \nabla^2[D_2(x, y, t)e(x, y, t)]. \quad (2)$$

While in continuous time random walk approaches few parameters suffice to describe a complex system, Eqs. (1) and (2) introduce field (x, y) dependent diffusion and drift coefficients, and thus represent a quite complex approach. *Codec & Metzler (2013)* has provided an exact expression for the diffusion coefficient in anomalous diffusion process modeled by Lévy walks under linear response regime.

If you take the simple case when $e(x, y, t)$ is a constant (assuming that the porous medium is relatively constant in the observation time scale, and it is sufficiently homogeneous and isotropic, all of which are common assumptions in hydrogeology), then you recover the classical groundwater flow equation $S_s \partial h / \partial t = k \nabla^2 h$. Our equations then satisfy the correspondence principle since they recover classical formulation and establish the ground for new insights of groundwater flow process, other porous media transport phenomena and even in Game Theory.

Typically, a system is considered complex when it is constituted from a large number of subsystems that interact strongly enough, but there is another source of complexity that has been widely ignored. A system is also complex when the system itself changes over time in the same scale of its dynamics, which is the case in some karstic aquifers. This second source of complexity is taken into account directly in our equations making a contribution in this respect and might have some important interpretation in Game Theory.

Finally, most interesting, using the traveling agent model described in the method section, we proposed (*López-Corona et al., 2013*) that $1/f$ noise is a fingerprint of a statistical phase transition, from randomness (disorder—white noise) to predictability (order—brown noise). In this context, one may interpret *Labat et al.'s (2011)* results as follows: first temporal scale (from 1 to 100 h) represents a rapid response of the aquifer and should be dominated by random processes (white noise); the second (100 h to 1 year) is the global response of the aquifer to rainfall input including the temporal structure of the peak flow one may be interpreted as a complex (with multiple spatio and temporal scales included) process ($1/f$ noise); and as the third correspond to mean (1 year or more) response is a more predictable process (brown noise). We have then a transition from randomness to predictability consistent with power spectra exponent values. In this way, the results of *Labat et al. (2011)* is only one example of a universal statistical kind of phase transition.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was supported by a CONACYT fellowship within the Earth Science Graduate School at UNAM. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Grant Disclosures

The following grant information was disclosed by the authors:
CONACYT.

Competing Interests

The authors declares there are no competing interests.

Author Contributions

- Oliver López Corona conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper, revision of the manuscript.
- Pablo Padilla conceived and designed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper, revision of the manuscript.
- Oscar Escolero conceived and designed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.
- Tomas González analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper, revision of the manuscript.
- Eric Morales-Casique conceived and designed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, reviewed drafts of the paper.
- Luis Osorio-Olvera analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.557#supplemental-information>.

REFERENCES

Atkinson R, Rhodes C, MacDonald D, Anderson R. 2002. Scale-free dynamics in the movement patterns of jackals. *Oikos* 98:134–140 DOI 10.1034/j.1600-0706.2002.980114.x.

- Austin D, Bowen W, McMillan J. 2004.** Intraspecific variation in movement patterns: modeling individual behaviour in a large marine predator. *Oikos* **105**:15–30 DOI [10.1111/j.0030-1299.1999.12730.x](https://doi.org/10.1111/j.0030-1299.1999.12730.x).
- Bak PC, Chen K. 1991.** Self-organized criticality, scientific American, self-organized criticality. *Scientific American* **264**(1):46–54 DOI [10.1038/scientificamerican0191-46](https://doi.org/10.1038/scientificamerican0191-46).
- Bak PC, Tang K. 1989.** Earthquakes as a self-organized critical phenomenon. *Journal of Geophysical Research* **94**(B11):15635–15637 DOI [10.1029/JB094iB11p15635](https://doi.org/10.1029/JB094iB11p15635).
- Bak PC, Tang K, Weisenfeld K. 1987.** Self-organized criticality: an explanation of $1/f$ noise. *Physical Review Letters* **59**(4):381–384 DOI [10.1103/PhysRevLett.59.381](https://doi.org/10.1103/PhysRevLett.59.381).
- Bohigas OMG, Schmit C. 1984.** Characterization of chaotic quantum spectra and universality of level fluctuation laws. *Physical Review Letters* **52**(1):1–4 DOI [10.1103/PhysRevLett.52.1](https://doi.org/10.1103/PhysRevLett.52.1).
- Boyer D, López-Corona O. 2009.** Self-organization scaling and collapse in a coupled automaton model of foragers. *Journal of Physics A: Mathematical and Theoretical* **42**:434014 DOI [10.1088/1751-8113/42/43/434014](https://doi.org/10.1088/1751-8113/42/43/434014).
- Boyer D, Miramontes O, Larralde H. 2009.** Lévy-like behaviour in deterministic models of intelligent agents exploring heterogeneous environments. *Journal of Physics A: Mathematical and Theoretical* **42**:434015.
- Boyer D, Ramos-Ramos-Fernández G, Miramontes O, Mateos JL, Cocho G, Larralde H, Ramos H, Rojas F. 2006.** Scale-free foraging by primates emerges from their interaction with a complex environment. *Proceedings of the Royal Society B: Biological Sciences* **273**:1743–1750 DOI [10.1098/rspb.2005.3462](https://doi.org/10.1098/rspb.2005.3462).
- Buhl J, Sumpter D, Couzin D, Hale JJ, Despland E, Miller ER, Simpson SJ. 2006.** From disorder to order in marching locusts. *Science* **312**(5778):1402–1406 DOI [10.1126/science.1125142](https://doi.org/10.1126/science.1125142).
- Cavagna A, Cimarelli A, Giardina I, Parisi G, Santagati R, Stefanini F, Viale M. 2009.** Scale-free correlations in bird flocks. ArXiv preprint. [arXiv:0911.4393](https://arxiv.org/abs/0911.4393).
- Cortis A, Knudby C. 2006.** A continuous time random walk approach to transient flow in heterogeneous porous media. *Water Resources Research* **42**:W10201 DOI [10.1029/2006WR005227](https://doi.org/10.1029/2006WR005227).
- De Marsily G, Delay F, Teles V, Schafmeister MT. 1998.** Some current methods to represent the heterogeneity of natural media in hydrogeology. *Hydrogeology Journal* **6**:115–130 DOI [10.1007/s100400050138](https://doi.org/10.1007/s100400050138).
- Dentz M, Guze P, Russian A, Dweik J, Delay F. 2012.** Diffusion and trapping in heterogeneous media: an inhomogeneous continuous time random walk approach. *Advances in Water Resources* **49**:13–22 DOI [10.1016/j.advwatres.2012.07.015](https://doi.org/10.1016/j.advwatres.2012.07.015).
- Downey A. 2012.** *Think complexity: complexity science and computational modeling*. O'Reilly Media, 79.
- Duffy CJ, Al-Hassan S. 1988.** The time and frequency response of tracer experiments. *Journal of Hydrology* **97**(1–2):59–73 DOI [10.1016/0022-1694\(88\)90066-2](https://doi.org/10.1016/0022-1694(88)90066-2).
- Duffy CJ, Gelhar LW. 1985.** A frequency domain approach to water quality modeling in groundwater: theory. *Water Resources Research* **21**(8):1175–1184 DOI [10.1029/WR021i008p01175](https://doi.org/10.1029/WR021i008p01175).
- Eliazar I, Klafter J. 2009a.** A unified and universal explanation for Lévy laws and $1/f$ noises. *Proceedings of the National Academy of Sciences of the United States of America* **106**(30):12251–12254 DOI [10.1073/pnas.0900299106](https://doi.org/10.1073/pnas.0900299106).
- Eliazar I, Klafter J. 2009b.** Universal generation of statistical self-similarity: a randomized central limit theorem. *Physical Review Letters* **103** (4): 040602 DOI [10.1103/physrevlett.103.040602](https://doi.org/10.1103/physrevlett.103.040602).

- Eliazar I, Klafter J. 2010.** Universal generation of $1/f$ noises. *Physical Review E* 82(2):021109 DOI 10.1103/PhysRevE.82.021109.
- Faleiro E, Kuhl U, Molina R, Muñoz L, Relayo A, Retamosa J. 2006.** Power spectrum analysis of experimental Sinai quantum billiards. *PLA* 358:251–255 DOI 10.1016/j.physleta.2006.05.029.
- Fleckenstein J, Niswonger R, Fogg G. 2006.** River-aquifer interactions, geologic heterogeneity, and low-flow management. *Groundwater* 44(6):837–852 DOI 10.1111/j.1745-6584.2006.00190.x.
- Fossion R, Landa E, Stransky P, Velazquez V, Lopez Vieyra JC, Garduno I, Garcia D, Frank A. 2010.** Scale invariance as a symmetry in physical and biological systems: listening to photons, bubbles and heartbeats. *AIP Conference Proceedings* 1323(74):76–90 DOI 10.1063/1.3537868.
- Gelhar LW. 1974.** Stochastic analysis of phreatic aquifer. *Water Resources Research* 10(3):539–545 DOI 10.1029/WR010i003p00539.
- Godec A, Metzler R. 2013.** Linear response, fluctuation–dissipation, and finite-system-size effects in superdiffusion. *Physical Review E* 88:012116 DOI 10.1103/PhysRevE.88.012116.
- Goldberger A. 2002.** Fractal dynamics in physiology: alterations with disease and aging. *Proceedings of the National Academy of Sciences of the United States of America* 99(1):2466–2472 DOI 10.1073/pnas.012579499.
- Haq R, Pandey A, Bohigas O. 1982.** Fluctuation properties of nuclear energy levels do theory and experiment, agree. *Physical Review Letters* 48(16):1086–1089 DOI 10.1103/PhysRevLett.48.1086.
- Jianting Z, Young M, Osterberg J. 2012.** Impacts of riparian zone plant water use on temporal scaling of groundwater systems. *Hydrological Processes* 26:1352–1360 DOI 10.1002/hyp.8241.
- Keilis-Botok V. 1990.** The lithosphere of the Earth as a non-linear system with implications for earthquake prediction. *Reviews of Geophysics* 28(1):19–34 DOI 10.1029/RG028i001p00019.
- Kirchner JW, Feng X, Neal C. 2000.** Fractal stream chemistry and its implications for contaminant transport in catchments. *Nature* 403:524–527 DOI 10.1038/35000537.
- Klafter J, Blumen A, Shlesinger M F. 1987.** Stochastic pathway to anomalous diffusion. *Physical Review A* 35(7):3081–3085 DOI 10.1103/PhysRevA.35.3081.
- Klafter J, Zumofen G, Shlesinger MF. 1995.** Lévy description of anomalous diffusion in dynamical systems. In: Shlesinger MF, Zaslavsky GM, Frisch U, eds. *Lévy flights in physical sciences and related phenomena*. Springer, 196–215.
- Koltermann D, Gorelick S. 1996.** Heterogeneity in sedimentary deposits: a review of structure-imitating, process-imitating, and descriptive approaches. *Water Resources Research* 32(9):2617–2658 DOI 10.1029/96WR00025.
- Labat D, Masbou J, Beaulieu E, Mangin A. 2011.** Scaling behavior of the fluctuations in stream flow at the outlet of karstic watersheds, France. *Journal of Hydrology* 410:162–168 DOI 10.1016/j.jhydrol.2011.09.010.
- Little M, Bloomfield J. 2010.** Robust evidence for random fractal scaling of groundwater levels in unconfined aquifers. *Journal of Hydrology* 393:362–369 DOI 10.1016/j.jhydrol.2010.08.031.
- Lomholt MA, Tal K, Metzler R, Joseph K. 2008.** Lévy strategies in intermittent search processes are advantageous. *Proceedings of the National Academy of Sciences of the United States of America* 105(32):11055–11059 DOI 10.1073/pnas.0803117105.
- López-Corona O. 2007.** Modelos estadísticos de difusión en ambientes desordenados. Aplicación al forrajeo y dispersión de semillas en selvas. Bachelor degree dissertation.
- López-Corona O, Padilla P, Escolero O, Frank A, Fossion R. 2013.** Lévy flights, $1/f$ noise and self organized criticality in a traveling agent model. *Journal of Modern Physics* 4(3):337–343 DOI 10.4236/jmp.2013.43046.

- Maréchal JC, Sarma MP, Ahmed S, Lachassagne P. 2002.** Establishment of earth tide effect on water-level fluctuations in an unconfined hard rock aquifer using spectral analysis. *Current Science* **83**(1):61–64.
- Martínez S, Escolero O, Kralish S. 2010.** Water management in San Luis Potosi Metropolitan area, México. *International Journal of Water Resources Development* **26**(3):459–475
DOI [10.1080/07900627.2010.489292](https://doi.org/10.1080/07900627.2010.489292).
- Martínez S, Escolero O, Wolf L. 2011.** Total urban water cycle models in semi-arid environments — Quantitative scenario analysis at the area of San Luis Potosi, Mexico. *Water Resources Management* **25**(1):239–263 DOI [10.1007/s11269-010-9697-6](https://doi.org/10.1007/s11269-010-9697-6).
- Miller SL, Miller WM, McWhorter PJ. 1993.** Extremal dynamics: a unifying and activated processes physical explanation of fractals, *i/f* noise. *Journal of Applied Physics* **73**(6):2617–2628
DOI [10.1063/1.353079](https://doi.org/10.1063/1.353079).
- Miramontes O, Boyer D, Bartumeus F. 2012.** The effects of spatially heterogeneous prey distributions on detection patterns in foraging seabirds. *PLoS ONE* **7**(4):e34317
DOI [10.1371/journal.pone.0034317](https://doi.org/10.1371/journal.pone.0034317).
- Montroll EW, Shlesinger MF. 1982.** On $1/f$ noise and other distributions with long tails. *Proceedings of the National Academy of Sciences of the United States of America* **72**(10):3380–3383 DOI [10.1073/pnas.79.10.3380](https://doi.org/10.1073/pnas.79.10.3380).
- Neuman SP. 2003.** Maximum likelihood Bayesian averaging of uncertain model predictions. *Stochastic Environmental Research and Risk Assessment* **17**(5):291–305
DOI [10.1007/s00477-003-0151-7](https://doi.org/10.1007/s00477-003-0151-7).
- Niemann M, Szendro I, Kantz H. 2013.** $1/f^\beta$ noise in a model for weak ergodicity breaking. ArXiv preprint. [arXiv:1002.0993](https://arxiv.org/abs/1002.0993).
- Park CH, Beyer C, Bauer S, Kolditz O. 2008.** An efficient method of random walk particle tracking: accuracy and resolution. *Geosciences Journal* **12**(3):285–297
DOI [10.1007/s12303-008-0029-2](https://doi.org/10.1007/s12303-008-0029-2).
- Press WH. 1978.** Flicker noises in astronomy and elsewhere. *Comments on Modern Physics, Part C—Comments on Astrophysics* **7**(4):103–119.
- Raghavan R. 2004.** A review of applications to constrain pumping test responses to improve on geological description and uncertainty. *Reviews of Geophysics* **42**:RG4001 DOI [10.1029/2003RG000142](https://doi.org/10.1029/2003RG000142).
- Ramos-Fernández G, Mateos JL, Miramontes O, Cocho G, Larralde H, Ayala-Orozco B. 2004.** Lévy walk patterns in the foraging movements of spider monkeys (*Ateles geoffroyi*). *Behavioral Ecology and Sociobiology* **55**:223–230 DOI [10.1007/s00265-003-0700-6](https://doi.org/10.1007/s00265-003-0700-6).
- Relanyo A, Gómez J, Molina R, Retamosa J, Faleiro E. 2002.** Quantum chaos and 1 over f noise. *Physical Review Letters* **89**:244102 DOI [10.1103/PhysRevLett.89.244102](https://doi.org/10.1103/PhysRevLett.89.244102).
- Sauquet E, Ramos MH, Chapel L, Bernardara P. 2008.** Streamflow scaling properties: investigating characteristic scale from different statistical approaches. *Hydrological Processes* **22**(17):3462–3475 DOI [10.1002/hyp.6952](https://doi.org/10.1002/hyp.6952).
- Seuront L, Duponchel A, Chapperon C. 2007.** Heavy-tailed distributions in the intermittent motion behaviour of the intertidal gastropod *Littorina littorea*. *Physica A* **385**:573–582
DOI [10.1016/j.physa.2007.07.029](https://doi.org/10.1016/j.physa.2007.07.029).
- Scher H, Cortis A, Berkowitz B. 2002.** A unified framework for anomalous transport in geological formations. In: *AGU fall meeting abstracts*, vol. 1. 4.

- Schilling KE, Palmer JA, Bettis III EA, Jacobson P, Schultz RC, Isenhardt TM. 2009. Vertical distribution of total carbon, nitrogen and phosphorus in riparian soils of Walnut Creek, southern Iowa (USA). *Catena* 77(3):266–273 DOI 10.1016/j.catena.2009.02.006.
- Shlesinger MF, Klafter J, Wong YM. 1982. Random walks with infinite spatial and temporal moments. *Journal of Statistical Physics* 27(3):499–512 DOI 10.1007/BF01011089.
- Shih D, Lee C, Chiou K, Tsai S. 2000. Spectral analysis of tidal fluctuations in ground water level. *Journal of the American Water Resources Association* 36(5):1087–1099 DOI 10.1111/j.1752-1688.2000.tb05712.x.
- Sims DW, Southal EJ, Humphries NE, Hays GC, Bradshaw CJA, Pitchford JW, James A, Ahmed MZ, Brierley AS, Hindel MA, Morritt D, Musyl MK, Righton D, Shepard ELC, Wearmouth VJ, Wilson RP, Witt MJ, Metcalfe JD. 2008. Scaling laws of marine predator search behaviour. *Nature* 451:1098–1102 DOI 10.1038/nature06518.
- Sornette A, Sornette D. 1989. Self-organized criticality and earthquakes. *Europhysics Letters* 9:197–202 DOI 10.1209/0295-5075/9/3/002.
- Sornette D, Davy P, Sornette A. 1990. Structuration of the lithosphere in plate tectonics as a self-organized criticality phenomenon. *Journal of Geophysical Research* 95(B11):17353–17361 DOI 10.1029/JB095iB11p17353.
- Tessier Y, Lovejoy S, Hubert P, Schertzer D, Pecknold S. 1996. Multifractal analysis and modeling of rainfall and NGWA.org river flows and scaling, causal transfer functions. *Journal of Geophysical Research* 101(D21):26427–26440 DOI 10.1029/96JD01799.
- Tranouez P, Bertelle C, Olivier D. 2001. Changing the level of description of a fluid flow in a agent-based simulation. In: *ESS 2001 conference*. France: Marseilles.
- Viswanathan GM, Buldyrev SV, Havlin S, da Luz M, Raposo E, Stanley H. 1999. Optimizing the success of random searches. *Nature* 401:911–914 DOI 10.1038/44831.
- Zhongwei L, You-Kuan Z. 2007. Quantifying fractal dynamics of groundwater systems with detrended fluctuation analysis. *Journal of Hydrology* 336:139–146 DOI 10.1016/j.jhydrol.2006.12.017.
- Zhang Y-K, Li Z. 2005. Temporal scaling of hydraulic head fluctuations: nonstationary spectral analyses and numerical simulations. *Water Resources Research* 41(7):W07031 DOI 10.1029/2004WR003797.
- Zhang Y-K, Li Z. 2006. Effect of temporally correlated recharge on fluctuations of groundwater levels. *Water Resources Research* 42(10):W10412 DOI 10.1029/2005WR004828.
- Zhang YK, Schilling K. 2004. Temporal scaling of hydraulic head and river base flow and its implication for groundwater recharge. *Water Resources Research* 40(3):W03504 DOI 10.1029/2003WR002094.

§6.3. Artículo en colaboración (enviado): Assessing sustainability in North America's ecosystems using criticality and information theory

PLOS ONE

Assessing sustainability in North America's ecosystems using criticality and information theory.

--Manuscript Draft--

Manuscript Number:	PONE-D-18-04888
Article Type:	Research Article
Full Title:	Assessing sustainability in North America's ecosystems using criticality and information theory.
Short Title:	Sustainability in North America's ecosystems
Corresponding Author:	Oliver López-Corona, PhD Comision Nacional para el Conocimiento y Uso de la Biodiversidad CDMX, CDMX MEXICO
Keywords:	Criticality; Entropy; Fisher Information; Ecosystem Health; Sustainability
Abstract:	Sustainability is a key issue in economical and political discussions, and for good reasons. The question of whether current patterns of human demography and socioeconomic activity can continue in the face of great environmental impact has led to calls for "sustainability". However, the word sustainability is often used in a qualitative sense, and thus there is a great need for a quantitative measure of sustainability through a system-level index. In this work, we use an informational approach to quantify some aspects of ecosystem sustainability, particularly health and stability
Order of Authors:	Elvia Ramírez-Carrillo Oliver López-Corona, PhD Juan C. Toledo-Roy Jon Lovett Fernando de León-González Luis Osorio-Olvera Julian Equihua Everardo Robredo Alejandro Frank Rodolfo Dirzo Vanessa Pérez-Cirera
Opposed Reviewers:	
Additional Information:	
Question	Response
Financial Disclosure	Catedras CONACyT fellowship program (project number 30) and Sistema Nacional de Investigadores (62929)
Please describe all sources of funding that have supported your work. This information is required for submission and will be published with your article, should it be accepted. A complete funding statement should do the following: Include grant numbers and the URLs of any funder's website. Use the full name,	

Assessing sustainability in North America's ecosystems using criticality and information theory

Elvia Ramírez-Carrillo¹, Oliver López-Corona^{2,3*}, Juan C. Toledo-Roy^{3,4}, Jon Lovett⁵, Fernando de León-González⁶, Luis Osorio-Olvera⁷, Julian Equihua⁸, Everardo Robredo⁸, Alejandro Frank^{3,4,9}, Rodolfo Dirzo¹⁰, and Vanessa Pérez-Cirera¹¹

¹Doctorado en Ciencias Agropecuarias, Universidad Autónoma Metropolitana-Xochimilco, Ciudad de México, México

²Cátedra CONACyT, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, México

³Centro de Ciencias de la Complejidad (C3), Universidad Nacional Autónoma de México, Ciudad de México, México.

⁴Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, Ciudad de México, México.

⁵Chair of Global Challenges, University of Leeds, UK

⁶Departamento de Producción Agrícola y Animal, Universidad Autónoma Metropolitana-Xochimilco, Ciudad de México, México

⁷Posgrado en Ciencias Biológicas, Facultad de Ciencias, Universidad Nacional Autónoma de México

⁸Comisión Nacional para el Conocimiento y Uso de la Biodiversidad, México

⁹Miembro del Colegio Nacional

¹⁰Biology Department, Stanford University, US

¹¹Instituto de Investigación para el Desarrollo con Equidad (EQUIDE), Universidad Iberoamericana, Ciudad de México, México.

Corresponding author:

Oliver López-Corona¹

Email address: lopezoliverx@otrasenda.org

ABSTRACT

Sustainability is a key issue in economics and politics today. The question of whether current patterns of human demography and socioeconomic activity can continue in the face of great environmental impact has led to calls for "sustainability". However, the word sustainability is often used in a qualitative sense. There is a great need for a quantitative, system-wide measure of sustainability. In this work, we use an informational approach to quantify some aspects of ecosystem sustainability, in particular health and stability.

INTRODUCTION

Sustainability has been defined in many ways, but the most frequently-quoted definition is from *Our Common Future*, also known as the Brundtland Report: "Sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs". This description, which might be considered the standard definition of sustainability, is politically correct but difficult to implement. What exactly do we mean by future? 5, 10 or 50 years? What does the word "necessities" mean in this context? What would a New Yorker say about his water needs in comparison with a peasant?

Furthermore, Cabezas and coworkers (Cabezas et al., 2000; Mayer et al., 2007) have pointed out that

45 while there is considerable interest in the core idea of sustainability, it remains a loosely-defined concept
46 with a limited mathematical definition, and even then it has only been defined within the restricted scope
47 of a few disciplines. There is no widely-accepted, precise, and testable multidisciplinary definition of
48 sustainability, and there is no general theory of the subject. This is a problem, because without precise
49 mathematical definitions, a general theory, and a testable hypothesis, it is virtually impossible to apply
50 the scientific method to make progress in any area of study. For example, Burger et al. (2012) criticized
51 that the macroecology of sustainability is woefully underrepresented in sustainability science. While
52 ecological principles are central to sustainability, their use has been largely concerned with socioenviron-
53 mental interactions, and they barely consider variables such as physical limits on resource use. Burger
54 et al. describe some calculations that are consistent with analyses reporting peak oil, fresh water, and
55 phosphate, to examine how global stocks of these important resources affect the patterns of global con-
56 sumption decline and the likelihood of global depletion (Gleick and Palaniappan, 2010; Nel and van Zyl,
57 2010).

58 From a basic ecological perspective, sustainability encompasses "the ability of one or more entities,
59 either individually or collectively, to exist and thrive (either unchanged or in evolved forms) for lengthy
60 timeframes, in such a manner that the existence and flourishing of other collectivities of entities is permit-
61 ted at related levels and in related systems" (Starik and Rands, 1995). From the multiplicity of elements
62 considered in this definition, we identify two core aspects relevant to ecological sustainability: ecosys-
63 tem stability and health (Chen et al., 2008). The concept of sustainability faces important challenges if
64 it is to consolidate as a solid scientific discipline. In this paper we only focus on the ecological angle of
65 sustainability, particularly ecological stability and health, using a thermodynamic-informational frame-
66 work. We consider that this approach provides precise mathematical concepts and takes into account
67 constraints in ecology and system dynamics.

68 Ecosystem Health

69 Ecosystem health is a diffuse concept that has been defined several times since the late eighties (Rapport et al.,
70 1998). This conceptual diversity has given rise to different measurement methods, which in turn has gener-
71 ated a wide range of narratives related to ecosystem health (O'Brien et al., 2016). Ultimately it has
72 become an ongoing priority for governments, scientists and managers around the world (Burger et al.,
73 2012).

74 Originally, ecosystem health was conceived within a control-optimization perspective in which health
75 is defined as a desired management target or reference condition (Rapport, 1989; Schaeffer et al., 1988).

76 Figure 1 lays out a schematic relationship between the ontologies that result from different disci-
77 plinary biases. The horizontal axis represents anthropocentric disciplinary bias, and the vertical axis
78 represents biological bias. Anthropocentric bias understands the issue from a managerial perspective,
79 focusing on optimization and control. On the other hand, as seen through a biological bias, ecosys-
80 tem health is usually defined through structure assessment and ecological functions. This usually in-
81 volves the measurement of certain indicators, such as the lack of algal blooms in rivers (Di Battista et al.,
82 2016; Pont et al., 2006); food web performance (Thompson et al., 2012); nutrient recycling and main-
83 tenance of biodiversity (US National Research Council, 2005), and resilience to external perturbations
84 (Mageau et al., 1995).

85 As is the case with science in general, the definition of ecosystem health has entered the realm of
86 multi- and interdisciplinarity (diagonal vector in Fig. 1). At the level of multidiscipline, ecosystem health
87 can be characterized by ecosystem services such as the provision of clean drinking water (Keeler et al.,
88 2012). As the idea of sustainability has permeated society, ecosystem health has become increasingly
89 associated with the integration of environmental, economic and human domains (Rapport et al., 1998;
90 Costanza et al., 1992).

Figure 1. The narrative surrounding ecosystem health has shifted from a very disciplinary framework
all the way through multi, inter and transdisciplinarity, before finally being defined in terms of complex
systems.

91 Finally, the shift towards interdisciplinarity in science has led to a complexity-based approach in
92 which ecosystem health is conceived as a property of a complex system. A dynamic system is com-
93 pletely characterized by a set of (state) variables. When assigned a particular set of numeric values,

94 these variables define the state of the system. There are also evolution rules that describe the way in
95 which the system transitions from any one state to any other. Although the very definition of a com-
96 plex system is still under active discussion (Prokopenko et al., 2009; Gershenson and Fernández, 2012;
97 Koorehdavoudi and Bogdan, 2016), in general a complex system emerges from a sufficiently large num-
98 ber of elements that have strong enough (usually non-linear) interactions, or when the state space changes
99 fast enough in terms of scale-space observer. These qualities make it impossible to describe the behavior
100 of the system in terms of the simpler behavior of its components.

101 Within this narrative of system dynamics, complexity is usually studied by analyzing the time series
102 of the fluctuations in state variables that have been identified as central to the dynamics of the system
103 (Van Kampen, 1965). Ultimately, they are at the center of the modern description of out-of-equilibrium
104 dynamics (Bertini et al., 2015).

105 A typical analytical method is studying the time series through spectral and fractal analysis, in
106 particular through the Power Spectral Density (PSD) or Detrended Fluctuation Analysis (DFA). It is
107 often the case that these fluctuations exhibit scale invariance (e.g. when the power spectrum follows
108 a power-law ($S \sim f^{-\beta}$), in which case it is customary to compare and classify fluctuation dynamics
109 according to their similarity to three archetypal classes of noises: white ($\beta \sim 0$), pink ($\beta \sim 1$) and
110 Brownian ($\beta \sim 2$) (Bak et al., 1988; Landa et al., 2011; Kleinen et al., 2003; Peng et al., 1994).

111 It has also been reported in the literature that several complex systems display behaviors related to
112 dynamic criticality, usually associated with some kind of scale invariance, and in many cases with pink
113 noise (Solé et al., 1996; Bak et al., 1988; Landa et al., 2011; Fossion et al., 2010). In this sense, it has
114 been suggested that systems in dynamic criticality require some kind of trade off between robustness and
115 adaptability (López-Corona et al., 2013). If the system is in too strict an order, it tends to be too rigid to
116 react promptly to the external perturbations, and in contrast, if the system behaves too randomly, then it
117 is too fragile, and any external perturbation will impair fundamental properties of the system. Thus, by
118 reaching criticality, complex systems are capable of being adaptive to changes while maintaining their
119 robustness (Morales et al., 2015).

120 In this line of thought, several authors have found empirical evidence of this dynamic criticality in
121 physiological process such as heart activity, and have speculated that it may be a key feature of a healthy
122 state. (Kiyono et al., 2005; Ivanov et al., 1996; Rivera et al., 2016)

123 Our proposal in this paper for measuring ecosystem health is based on this idea of dynamic criticality
124 as the combination of scale invariance and balance between adaptability and robustness.

Figure 2. Our proposal for measuring the dynamic dimension of ecosystemic health is based on the
idea of criticality as the combination of scale invariance and balance between adaptability and
robustness (pink noise). By combining a scale invariance index based on BIC values with the value of
the scalar coefficients (beta) in power spectra, we propose an Ecosystemic Health Index, whose
maximum for beta values equals 1, and that is associated with a balance between adaptability and
robustness. In this way, an ecosystem may lose health by losing robustness and exhibiting white-noise
dynamics, or by losing adaptability leading to Brownian-noise dynamics.

125 But which environmental signal could be used as an analog of a systemic physiological variable such
126 as heart rate, which is one of the most widely used experimental variables?

127 We consider that potential candidates should be related to soil because it is a complex system (Janzen,
128 2015) that integrates several scales as well as main ecosystemic processes. On the one hand, its dy-
129 namics is defined by the interaction of different subsystems such as the biosphere, atmosphere, geo-
130 sphere and hydrosphere and all their components (Chesworth and others, 2008), which in turn inter-
131 act in the geographic space, generating different pedogenetic processes related to climate or geoforms
132 (James and Danoff-Burg, 2008; Ehrenfeld et al., 2005; Yaalon, 1983). On the other hand, according to
133 Janzen (2015), carbon flows also connect ecosystems in the temporal dimension, since carbon persists in
134 the soil as a kind of biogeochemical memory (Targulian and Goryachkin, 2004).

135 Considering the above, and since soil plays a central role in the flux of CO₂, we decided to analyze
136 the Ameriflux database that measures energy and matter (mainly CO₂) fluxes in America (mostly North
137 America).

138 **Ecosystem Stability (Out-of-equilibrium Thermodynamics)**

139 Following Michaelian (2005), let's consider an abstract resource space in which we can define a vector \bar{r}_i
140 that represents a particular socio-ecological entity (a species, an environmental entity such as a wetland,
141 or a human community). In this way, the projection of the entity's vector \bar{r}_i over a resource axis represents
142 the root mean square value of the resource gradient that the entity i requires to subsist.

143 Michaelian (2005), proposes that the movement of an entity in this space implies changing the re-
144 source gradient requirements as well as the strength of the interactions with the other entities. Thus, an
145 entity vector \bar{r}_i may exhibit length changes in the form of vibrations (i.e. changes in a species popu-
146 lation), small direction changes in the form of rotations (i.e. niche plasticity) and larger changes in its
147 direction (i.e. niche evolution).

148 A well known fact discovered by Helmholtz is that, when left alone, all systems tend to more stable
149 states of greater longevity by reducing their free energy F as defined by

$$F = U - TS \quad (1)$$

150 where U is the internal energy of the system given by interactions between the internal constituents,
151 T is its temperature (associated with randomness) and S its entropy. Then, given the restrictions of the
152 resource space, this optimization problem guides the dynamics of the exploration of this space and hence
153 the evolution of the strengths of the interactions between entities and between entities and environment.
154 Stability is reached by minimizing Helmholtz free energy, and one way to achieve this is to maximize
155 the entropy of the whole system (not just a single component) Michaelian (2000).

156 **Ecosystem Stability (Fisher Information)**

157 Mayer and co-workers (Mayer et al., 2007) have proposed that Fisher information offers a robust method
158 to assess the stability of a system over time, being essentially able to aggregate multiple variables, each
159 one capturing different aspects of a system, and outputting a global indicator of stability.

160 Following Frieden and Soffer (1995) and Mayer et al. (2007) let us consider the basic problem of
161 estimating the real value of a state variable θ . The estimation comes from an inference process from
162 imperfect observation $y = \theta + x$ in the presence of some random noise x .

163 This kind of measurement-inference process will hence be called "smart measurement" of θ whose
164 result is an estimator $\hat{\theta}$ that is function of imperfect observation $\hat{\theta}(y)$.

This is a closed system, meaning that it's well described by $\{y, \hat{\theta}, x\}$ without the need to consider ad-
ditional sources of noise. Consider also that the estimator is unbiased in terms of being a good estimator
on average $\langle \hat{\theta}(y) \rangle = \theta$. In this case, the mean-square error obeys the Cramer-Rao inequality

$$e^2 I \geq 1, \quad (2)$$

165 where I is the Fisher Information of the system, calculated as

$$I = \int \frac{dy}{P_0(y|\theta)} \left[\frac{dP_0(y|\theta)}{d\theta} \right]^2, \quad (3)$$

166 in which $P_0(y|\theta)$ is the probability density function of measuring a particular value of y given the true
167 value θ of the state variable in question. Then, since the error decreases as information increases, Fisher
168 information may be understood as the quality of the estimation θ from a smart measurement.

169 Then if the system is characterized by a phase space with m state variables x_i that define the phase
170 vector $s = (x_1, \dots, x_i, \dots, x_m)$ associated with a smart measurement y , then we can prove that

$$I(s) = \frac{1}{T} \int_0^T \frac{s'^2}{s'^4} dt \quad (4)$$

171 where T is the time period required for one cycle of the system; $s'(t)$ is the tangential speed and $s''(t)$ is
172 scalar acceleration tangential to the system path in phase space. Both are calculated in terms of the state
173 variables x_i as

$$s'(t) = \sqrt{\sum_i^m \left(\frac{dx_i}{dt}\right)^2}, \quad (5)$$

$$s''(t) = \frac{1}{s'(t)} \sum_i^m \left(\frac{dx_i}{dt} \frac{d^2x_i}{dt^2}\right). \quad (6)$$

174 A simple and robust approach to calculating tangential velocity and acceleration uses the three-point
175 difference scheme

$$\frac{dx_i}{dt} = \frac{\alpha x_i(t + \Delta t_a) - (\alpha^2 - 1)x_i(t) - x_i(t - \alpha \Delta t_a)}{\alpha(\alpha + 1)\Delta t_a} \quad (7)$$

$$\frac{d^2x_i}{dt^2} = \frac{\alpha x_i(t + \Delta t_a) - (\alpha + 1)x_i(t) - x_i(t - \alpha \Delta t_a)}{\alpha(\alpha + 1)\Delta t_a^2/2} \quad (8)$$

176 where $x_i(t)$ is a central data point, $x_i(t - \Delta t_a)$ is the later point to the central x_i and $x_i(t - \Delta t_p)$ is the
177 previous point. For evenly-spaced points $\Delta t_a = \Delta t_p$ and $\alpha = \Delta t_p/\Delta t_a$ is the ratio of the previous and
178 following time space.

179 The thesis proposed by Frieden and Gatenby (2007) is that a change in Fisher information can signal
180 a regime change in a dynamic system, and is based on the following premises: (1) if a change in the
181 dynamic regime is observable then there must be a corresponding change in the measurable variables of
182 the system; (2) an observable change in the measurable variables implies a corresponding change in the
183 distribution of system states; and (3) a change in the distribution of system states implies a change in the
184 system's Fisher information. An interesting feature of this proposal is that it gives us a way to measure
185 order, since very little information would be inferred from a disordered (non-correlated) system with no
186 observable patterns. This would translate to a Fisher information that approaches zero. On the other
187 hand, one would obtain the highest values of information from ordered (highly-correlated) systems that
188 exhibit patterns in behavior.

189 Following these ideas, Cabezas and Fath (2002) has proposed that:

- 190 • Fisher information is a function of the variability of the observations. Low variability leads to high
191 Fisher information and high variability leads to low Fisher information.
- 192 • Systems in stable dynamic states have constant Fisher information. Systems losing organization
193 migrate toward higher variability and lose Fisher information.
- 194 • Self-organizing systems decrease their variability and acquire Fisher information.

195 These considerations led them to propose a sustainability hypothesis: "sustainable systems do not
196 lose or gain Fisher information over time." From equations (5) and (6) this means that the system is in a
197 state of constant tangential velocity and acceleration in the phase space, and therefore, in a stable state.

198 METHODS

199 The data were taken from the AmeriFlux researcher-driven network of sites in North, Central and South
200 America, measuring ecosystem respiration, water, and energy fluxes. The network was established to
201 provide compatible data from a large number of sites representing major climate and ecological biomes,
202 including tundra, grasslands, savannah, farmland, and coniferous, deciduous, and tropical forests. Each
203 site has instruments tailored to suit each ecosystem. The network grew from about 15 sites in 1997 to
204 more than 110 active sites registered today. Sixty-one other sites, now inactive, have flux data stored in
205 the network's database. In 2012, the U.S. DOE established the AmeriFlux Management Project (AMP)
206 at Lawrence Berkeley National Laboratory (LBNL) to support the broad AmeriFlux community and the
207 AmeriFlux sites. The data is publically available from the Ameriflux database (<http://ameriflux.lbl.gov>).
208 The time series measures CO₂ flux fluctuations every half hour.

209 First, we performed an Analytical Hierarchical Process (AHP) using the following criteria: data re-
210 quirements for analysis, ecological relevance and quantity of data available. From this analysis, ecosys-
211 tem respiration was identified as the best measure of ecological processes.

212 We did not use data from intensively managed farmland sites, because they are subject to high external
213 inputs of nutrients, pesticides and herbicides and are artificially maintained in a temporary ecological
214 condition. In the absence of external control, they would markedly change their physiological signals
215 and so their ecological state should be considered fragile.

216 In order to derive the Criticality Index and a Scale Invariance Index from annual Ameriflux site data
217 we used the following steps.

218 Firstly, we extracted the desired variable from the raw file and identified any missing or invalid values
219 in the data series. We then scanned for gaps in the time series, as the computation of the power spectrum
220 demands that the time series have constant time intervals between data values. When the gaps are small
221 (a few data values) it is possible to perform a simple interpolation to fill in the gap. However, to ensure
222 that this interpolation does not alter the data, we only analyzed time series with no gaps at all.

223 Secondly, we filtered out long-term trends and obvious periodicities from the time series. Although
224 the Ameriflux records are not expected to exhibit long-term trends, we still computed and subtracted a
225 linear trend for the whole time series for each site. If no trend was present the data was left essentially
226 unchanged. Daily periodicities are to be expected for many of the variables. While there are many techniques
227 to extract periodicities for time series, we found that applying a digital infinite impulse response
228 filter worked well for the Ameriflux records. In particular, we used *scipy*'s notch filter, which is a band-
229 stop filter that rejects a narrow bandwidth around a chosen frequency and leaves the rest of the spectrum
230 unchanged. We did this three times: once around a frequency of one cycle per day using a quality factor
231 $Q=12$, and then twice with a narrower band with ($Q=30$) for the first two harmonics of this frequency.
232 After this step the time series was considered to be trend-free and most high-energy periodicities are
233 removed. We termed this filtered time series the 'fluctuations'.

234 The third step was to apply a traditional spectral analysis using a Fast Fourier Transformation of
235 the time series and to compute the spectral index by fitting power-laws to the spectrum. Two fits were
236 obtained. The first is a direct single power-law fit obtained by fitting a straight line to the logarithmic
237 spectrum through least-squares linear regression. The negative of the slope of this line is the first spectral
238 index, β_0 , which is a measure of criticality. We used this value to define the Criticality Indicator (*Icrit*).

239 The second fit model is a piecewise-defined and double power-law function, composed of a low-
240 frequency power-law with a spectral index β_1 , followed by a high-frequency power-law with a spectral
241 index β_2 , with a crossover frequency to be determined. We employed the *scipy* curve-fit routine (a
242 nonlinear least-squares optimizer that uses the Trust Region Reflective algorithm) to obtain the best-
243 fit double power-law model for each of the fluctuation time series. In order to obtain a measure of
244 scale invariance we then compare the two power-law models by computing their Bayesian Information
245 Criterion (BIC) using the residual sum of squares of the models and the relevant number of parameters
246 for each one, including the data variance (i.e. 3 for the single power-law model, 5 for the double-power
247 model). The BIC provides a model comparison that penalizes a model for having more parameters. The
248 BIC values of the models then yields the Scale Invariance Index (*Iscale*).

249 We defined the Criticality Index (*Icrit*) as a function of the "distance" to a $1/f$ type of signal ($\beta \sim -1$)
250 in such a way that it equals 1 when $\beta = -1$ and zero when $\beta \leq -0.5$ and $\beta \geq 1.5$. Between those
251 values *Icrit* grows and decreases linearly as $Icrit = 2\beta - 1$ and $Icrit = -2\beta + 3$.

252 In the same manner, we defined the Scale Invariance Index (*Iscale*) in terms of model selection
253 between a one linear model or a two lines model for fit the PSD using the BIC. Taking the BIC model
254 difference $dBIC = BIC(model1) - BIC(model2)$, *Iscale* is zero for $dBIC \leq 2$ and 1 for $dBIC > 10$.
255 For intermediate values it increases linearly as $Iscale = (1/8) * dBIC - 1.75$.

256 Then, we define the Ecosystemic Health Index (*Ih*) following the functional form of the Human
257 Development Index (HDI) as the square root of the criticality index times the scale index.

$$Ih = \sqrt{Icrit * Iscale} \quad (9)$$

258 To derive the value of ecosystem stability under the Michaelian framework based on entropy (Michaelian,
259 2005), we used the Statcomp library in R (Sippel et al., 2016; R Core Team, 2017), which calculates sta-
260 tistical complexity and other information measures for the Time Series. Statcomp (Sippel et al., 2016)
261 is based on Bandt and Pompe (2002) and calculates simple complexity measures using the concept of
262 permutation entropy defined by the order relations among values of a time series. Permutation entropy
263 assumes that patterns may not have the same probability of occurrence, and that this probability may

264 unveil relevant knowledge about the underlying system. In the same way, we use the *fis* function in the
265 Statcomp library to calculate Fisher information over the annual fluctuation time series.

266 RESULTS

267 In Figure 3 we show the land condition variable used as validation in terms of noise type. The colors
268 represent ecosystem types in the IGBP nomenclature. We expected that if criticality is a good proxy for
269 ecosystem health, then sites with $1/f$ dynamics should have the highest land condition values.

270 By assessing an ANOVA test using land condition as a proxy variable and grouping data by noise type
271 (color) we obtain statistically significant differences ($F = 28.16$; $Pr(> F) = 2.2 \times 10^{-16}$ ***) between
272 noise types, although it is clear that this difference between pink and white might be marginal, perhaps
273 because more data is needed.

Figure 3. In this figure we show the Land Condition variable used as validation in terms of noise type.
Color scale corresponds to ecosystem types in the IGBP nomenclature
gaia.agraria.unitus.it/IGBPdesignations.pdf

274 Since land condition is not a dynamic measurement, it is not expected to have strong correlation
275 with our measurement of ecosystemic health, but it is consistent in statistical terms. Sites with pink
276 noise ($\beta \sim -1$) behavior are statistically in better land condition than those sites with white ($\beta \sim 0$) or
277 brown ($\beta \sim -2$) noise type. So we may have cases of sites with an external non-health condition that
278 nevertheless exist in a systemically healthy state. An analogy would be a person that has a broken arm:
279 this state of non-health means nothing in terms of systemic processes such as heart activity. Conversely,
280 we also have sites that are externally healthy, but systemically non-healthy. A great analogy would be
281 the cases of Sudden Cardiac Death in Young Athletes (Chandra et al., 2013).

282 In Fig. 4 we show a boxplot for beta values for each ecosystem type in IGBP nomenclature again.
283 As expected, most ecosystems fall into pink noise behavior. And in general, we found that ecosystems
284 out of criticality are older forests or have been altered by human activity or events such as wildfires. One
285 example is the UA-Me1 (Metolius-Eyerly burn) site, which is an intermediate aged ponderosa pine forest
286 in Oregon-USA that was severely burned in 2002 by the famous Eyerly wildfire, a stand-replacing event
287 in which all trees were killed.

Figure 4. As expected, most ecosystems fall into pink noise. In general, we found that ecosystems out
of criticality are older forests or have been altered by human activity or events such as wildfires.

288 Figure 5 shows the values of our Ecosystemic Health Index for all ecosystem types (IGBP). For this
289 data set, it seems that Ecosystemic Health is basically driven by the value of beta.

Figure 5. Values of our Ecosystemic Health Index for all ecosystem types (IGBP). The color of each
data point corresponds to the type of noise (white, pink and brown)

290 Our results (see figure 6) are also consistent with Michaelian (Michaelian, 2005) ideas about ecosys-
291 tem stability and entropy, where stable ecosystems (higher values of entropy) correspond to healthier
292 states (criticality - pink noise).

293 Again we have significant difference between noise types, with pink noise ($\beta \sim -1$) behavior being
294 statistically more stable than white ($\beta \sim 0$) or brown ($\beta \sim -2$) noise.

Figure 6. Permutation entropy boxplot in terms of noise color

295 We see in Figure 7 a very interesting behavior for permutation entropy as a function of beta. We see
296 that entropy reaches its highest values around the range of beta values for which pink noise is defined,
297 meaning that the most stable behavior corresponds to a criticality dynamic.

Figure 7. Permutation entropy scatter plot in terms of β value. Entropy reaches highest values around the beta values range for which pink noise is defined.

298 Interestingly enough, as proposed by Gershenson and Fernández (2012) we see in Figure 8 a quadratic
299 relation between information (i.e. entropy) and complexity:

$$C = aI(1 - I) \tag{10}$$

300 We think that this might be the first time this relation is found directly in the data and not from
301 modeling.

Figure 8. Complexity as a quadratic function of Permutation Entropy.

302 And finally, our results are consistent with Cabezas and Fath (2002): (a) Fisher information is a
303 function of the variability of the observations such that low variability leads to high Fisher information
304 and high variability leads to low Fisher information; (b) Systems in stable dynamic states have constant
305 Fisher information. Systems losing organization migrate toward higher variability and lose Fisher in-
306 formation; (c) Self-organizing systems decrease their variability and acquire Fisher information. The
307 authors under these considerations propose a sustainability hypothesis: “sustainable systems do not lose
308 or gain Fisher information over time.”

309 In figure 9 we show a time series of Fisher information for the Harvard Forest site (US-Ha1). We
310 can see that from 1991 to 2003 the ecosystem was in a stable state of low health (combination of white
311 and pink noise) with a low Fisher information value (around 0.025). After that, it enters a process of
312 self-organization gaining Fisher information and starts to stabilize around a higher Fisher information
313 value (around 0.15), dominated by healthier pink noise dynamics, and therefore, according to Cabezas’s
314 hypothesis, a more sustainable state.

Figure 9. Since criticality (pink noise) appears to be the most healthy and stable (sustainable) type of dynamics, we use it as a leaf variable in a classification tree using the C4.5 algorithm in WEKA.

315 Since criticality (pink noise) appears to be the most healthy and stable (sustainable) type of dynamics,
316 we use it as a leaf variable in a classification tree using the C4.5 algorithm in WEKA. Results are
317 consistent with what was previously described: sites with an entropy value lower than 0.85 are out of
318 criticality, that is, in non-healthy and non-sustainable states. Sites with a land condition value under 30
319 are also out of criticality and hence in non-healthy and non-sustainable states. Pink noise in this branch
320 of the tree is a spurious result because, just like with intensive crop lands, this Sherman Island site (US-
321 Snd) is known to be very degraded but under intensive control by the California Department of Water
322 Resources, so this result is a false positive. Specific ecosystem types with a combination of entropy
323 values higher than 0.85 and Land Condition value higher than 30 are in healthy, stable and therefore
324 sustainable states. The rest of the tree is harder to interpret, and individual sites stories might play an
325 important role (events of wildfires, site management, and so on.)

326 In Figure 10 we present the corresponding maps using a combination of circle size and color to
327 encode one or two variables of interest.

Figure 10. In this map we represent the type of noise in a color scale, and land condition as size of the circles.

328 Finally we show the complete Sustainability Index as the square root of the Health Index times the
329 Stability Index 11.

Figure 11. Here, Ecosystem Stability (Permutation Entropy) is shown in a color scale and Ecosystem Health (as the square root of scale invariance indicator times the criticality indicator) as the size of the circles.

DISCUSSION AND CONCLUSIONS

A complexity perspective based on information theory seems to be a promising starting point to develop a general framework for the measurement of sustainability.

We showed that the use of criticality, defined by scale invariance and pink noise behavior, may be one way to measure systemic ecosystem health. Nevertheless, we consider it necessary to use a non-dynamic variable, such as land condition, to complement the analysis.

In general, we found that sites with pink noise ($\beta \sim -1$) behavior are statistically in better land condition than those sites with white ($\beta \sim 0$) or brown ($\beta \sim -2$) noise type. Interestingly enough, one may find systems with good land condition but in low systemic health and vice versa.

We understand the first case, where we have high values of land condition but low values of Ecosystem Health, in terms of an analogy with human health considering phenomena such as Sudden Cardiac Death in young athletes (Chandra et al., 2013). Consider an Olympic athlete in their twenties: it would be difficult to think of someone with better external health qualities, and yet this athlete could drop dead on the track due to this syndrome, which is related to systemic health. On the other hand, consider a person with a broken arm (a clear external signal of non-health) but perfectly healthy in terms of heart, brain and general system functioning. Additionally, we identified a third case that, in this analogy, matches a patient in intensive care who is maintained in some form of artificial health using life-support devices.

We recognize that as usually happens with complex systems: the basis of the configuration space of the problem is not known a priori. Whether ecosystem respiration is the correct or only “physiological” signal for Ecosystemic Health measurements remains an open question.

Our results are consistent with stability ideas developed by Michaleian, where high entropic sites are also in criticality. We refer to a case study where our results are also in good agreement with the Fisher information framework for system stability developed by Mayer and co-workers (Mayer et al., 2007). In this study, stable (more sustainable) systems do not lose or gain Fisher information over time.

Finally, we acknowledge that more data and controlled experiments would be necessary to understand complicated patterns that emerge from current analysis.

ACKNOWLEDGMENTS

OLC thanks to Cátedras CONACyT fellowship program (project number 30) and Sistema Nacional de Investigadores (62929); LOO acknowledges partial support from CONACyT 245648 scholarship and PAPIIT IN116018 for partial support in postgraduate studies.

REFERENCES

- Bak, P., Tang, C., and Wiesenfeld, K. (1988). Self-organized criticality. *Physical Review A*, 38(1):364–374.
- Bandt, C. and Pompe, B. (2002). Permutation Entropy: A Natural Complexity Measure for Time Series. *Physical Review Letters*, 88(17):174102.
- Bertini, L., De Sole, A., Gabrielli, D., Jona-Lasinio, G., and Landim, C. (2015). Macroscopic fluctuation theory. *Reviews of Modern Physics*, 87(2):593.
- Burger, J. R., Allen, C. D., Brown, J. H., Burnside, W. R., Davidson, A. D., Fristoe, T. S., Hamilton, M. J., Mercado-Silva, N., Nekola, J. C., Okie, J. G., and Zuo, W. (2012). The Macroecology of Sustainability. *PLoS Biology*, 10(6):e1001345.
- Cabezas, H. and Fath, B. D. (2002). Towards a theory of sustainable systems. *Fluid Phase Equilibria*, 194-197:3–14.
- Cabezas, H., Harten, H. P., and Green, M. (2000). Designing greener solvents. *Chemical Engineering*, 107(3):107–109.
- Chandra, N., Bastiaenen, R., Papadakis, M., and Sharma, S. (2013). Sudden Cardiac Death in Young Athletes. *Journal of the American College of Cardiology*, 61(10):1027–1040.

- 376 Chen, A. J., Boudreau, M., and Watson, R. T. (2008). Information systems and ecological sustainability.
377 *Journal of Systems and Information Technology*, 10(3):186–201.
- 378 Chesworth, W. and others (2008). Encyclopedia of soil science.
- 379 Costanza, R., Norton, B. G., and Haskell, B. D. (1992). *Ecosystem health : new goals for environmental*
380 *management*. Island Press.
- 381 Di Battista, T., Fortuna, F., and Mauro, F. (2016). Environmental monitoring through functional biodi-
382 versity tools. *Ecological Indicators*, 60:237–247.
- 383 Ehrenfeld, J. G., Ravit, B., and Elgersma, K. (2005). Feedback in the plant-soil system. *Annu. Rev.*
384 *Environ. Resour.*, 30:75–115.
- 385 Fossion, R., Landa, E., Stránský, P., Velázquez, V., Vieyra, J. C. L., Garduño, I., García, D., Frank, A.,
386 Wolf, K. B., Benet, L., Torres, J. M., and Hess, P. O. (2010). Scale invariance as a symmetry in
387 physical and biological systems: listening to photons, bubbles and heartbeats. pages 74–90.
- 388 Frieden, B. R. and Gatenby, R. A. (2007). *Exploratory data analysis using Fisher information*. Springer.
- 389 Frieden, B. R. and Soffer, B. H. (1995). Lagrangians of physics and the game of Fisher-information
390 transfer. *Physical Review E*, 52(3):2274–2286.
- 391 Gershenson, C. and Fernández, N. (2012). Complexity and information: Measuring emergence, self-
392 organization, and homeostasis at multiple scales. *Complexity*, 18(2):29–44.
- 393 Gleick, P. H. and Palaniappan, M. (2010). Peak water limits to freshwater withdrawal and use. *Proceed-*
394 *ings of the National Academy of Sciences of the United States of America*, 107(25):11155–62.
- 395 Ivanov, P. C., Rosenblum, M. G., Peng, C.-K., Mietus, J., and others (1996). Scaling behaviour of
396 heartbeat intervals obtained by wavelet-based time-series analysis. *Nature*, 383(6598):323.
- 397 James, A. and Danoff-Burg, A. (2008). The Terrestrial Influence: Geology and Soils. *Columbia Univer-*
398 *sity. pp46*.
- 399 Janzen, H. H. (2015). Beyond carbon sequestration: soil as conduit of solar energy. *European Journal*
400 *of Soil Science*, 66(1):19–32.
- 401 Keeler, B. L., Polasky, S., Brauman, K. A., Johnson, K. A., Finlay, J. C., O’Neill, A., Kovacs, K., and
402 Dalzell, B. (2012). Linking water quality and well-being for improved assessment and valuation of
403 ecosystem services. *Proceedings of the National Academy of Sciences of the United States of America*,
404 109(45):18619–24.
- 405 Kiyono, K., Struzik, Z. R., Aoyagi, N., Togo, F., and Yamamoto, Y. (2005). Phase transition in a healthy
406 human heart rate. *Physical review letters*, 95(5):58101.
- 407 Kleinen, T., Held, H., and Petschel-Held, G. (2003). The potential role of spectral properties in detect-
408 ing thresholds in the Earth system: application to the thermohaline circulation. *Ocean Dynamics*,
409 53(2):53–63.
- 410 Koorehdavoudi, H. and Bogdan, P. (2016). A Statistical Physics Characterization of the Complex Sys-
411 tems Dynamics: Quantifying Complexity from Spatio-Temporal Interactions. *Scientific reports*, 6.
- 412 Landa, E., Morales, I. O., Fossion, R., Stránský, P., Velázquez, V., López Vieyra, J. C., and Frank,
413 A. (2011). Criticality and long-range correlations in time series in classical and quantum systems.
414 *Physical Review E*, 84(1):016224.
- 415 López-Corona, O., Padilla, P., Escolero, O., Frank, A., and Fossion, R. (2013). Lévy Flights, 1/f Noise
416 and Self Organized Criticality. *Journal of Modern Physics*, 4(3).
- 417 Mageau, M., Costanza, R., and Ulanowicz, R. (1995). The development and initial testing of a quantita-
418 tive assessment of ecosystem health. *Ecosyst. Health*, 1:201–213.
- 419 Mayer, A. L., Pawlowski, C., Fath, B. D., and Cabezas, H. (2007). Applications of Fisher Information to
420 the Management of Sustainable Environmental Systems. In *Exploratory Data Analysis Using Fisher*
421 *Information*, pages 217–244. Springer London, London.
- 422 Michaelian, K. (2000). A Physical Basis of Evolution and Speculation on an Evolutionary Basis of
423 Physics. In Heras, J., editor, *Topics in Contemporary Physics*, pages 195–210.
- 424 Michaelian, K. (2005). Thermodynamic stability of ecosystems. *Journal of Theoretical Biology*,
425 237(3):323–335.
- 426 Morales, I. O., Landa, E., Angeles, C. C., Toledo, J. C., Rivera, A. L., Temis, J. M., and Frank, A. (2015).
427 Behavior of early warnings near the critical temperature in the two-dimensional Ising model. *PLoS*
428 *one*, 10(6):e0130751.
- 429 Nel, W. P. and van Zyl, G. (2010). Defining limits: Energy constrained economic growth. *Applied*
430 *Energy*, 87(1):168–177.

- 431 O'Brien, A., Townsend, K., Hale, R., Sharley, D., and Pettigrove, V. (2016). How is ecosystem health
432 defined and measured? A critical review of freshwater and estuarine studies. *Ecological Indicators*,
433 69:722–729.
- 434 Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., and Goldberger, A. L. (1994).
435 Mosaic organization of DNA nucleotides. *Physical Review E*, 49(2):1685–1689.
- 436 Pont, D., Hugueny, B., Beier, U., Goffaux, D., Melcher, A., Noble, R., Rogers, C., Roset, N., and
437 Schmutz, S. (2006). Assessing river biotic condition at a continental scale: a European approach
438 using functional metrics and fish assemblages. *Journal of Applied Ecology*, 43(1):70–80.
- 439 Prokopenko, M., Boschetti, F., and Ryan, A. J. (2009). An information-theoretic primer on complexity,
440 self-organization, and emergence. *Complexity*, 15(1):11–28.
- 441 R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statis-
442 tical Computing, Vienna, Austria.
- 443 Rapport, D. J. (1989). What Constitutes Ecosystem Health? *Perspectives in Biology and Medicine*,
444 33(1):120–132.
- 445 Rapport, D. J., Costanza, R., and McMichael, A. J. (1998). Assessing ecosystem health.
- 446 Rivera, A. L., Estañol, B., Sentís-Madrid, H., Fossion, R., Toledo-Roy, J. C., Mendoza-Temis, J.,
447 Morales, I. O., Landa, E., Robles-Cabrera, A., Moreno, R., and others (2016). Heart rate and sys-
448 tolic blood pressure variability in the time domain in patients with recent and long-standing diabetes
449 mellitus. *PLoS one*, 11(2):e0148378.
- 450 Schaeffer, D. J., Herricks, E. E., and Kerster, H. W. (1988). Ecosystem health: I. Measuring ecosystem
451 health. *Environmental Management*, 12(4):445–455.
- 452 Sippel, S., Lange, H., and Gans, F. (2016). statcomp: Statistical Complexity and Information Measures
453 for Time Series Analysis. R package version 0.0.1.1000.
- 454 Solé, R. V., Manrubia, S. C., Luque, B., Delgado, J., and Bascompte, J. (1996). Phase transitions and
455 complex systems: Simple, nonlinear models capture complex systems at the edge of chaos. *Complexity*,
456 1(4):13–26.
- 457 Starik, M. and Rands, G. P. (1995). Weaving an Integrated Web: Multilevel and Multisystem Perspectives
458 of Ecologically Sustainable Organizations. *Academy of Management Review*, 20(4):908–935.
- 459 Targulian, V. O. and Goryachkin, S. V. (2004). Soil memory: Types of record, carriers, hierarchy and
460 diversity. *Revista Mexicana de Ciencias Geológicas*, 21(1):1–8.
- 461 Thompson, R. M., Brose, U., Dunne, J. A., Hall, R. O., Hladysz, S., Kitching, R. L., Martinez, N. D.,
462 Rantala, H., Romanuk, T. N., Stouffer, D. B., and Tylianakis, J. M. (2012). Food webs: reconciling
463 the structure and function of biodiversity. *Trends in ecology & evolution*, 27(12):689–97.
- 464 Van Kampen, N. G. (1965). Fluctuations in nonlinear systems. *Fluctuation Phenomena in Solids*, Aca-
465 demic Press, New York.
- 466 Yaalon, D. H. (1983). Climate, time and soil development. *Developments in Soil Science*, 11:233–251.

Figure 1. The narrative surrounding ecosystem health has shifted from a very disciplinary framework all the way through multi, inter and transdisciplinarity, before

[Click here to download Figure diagrama.tif](#)

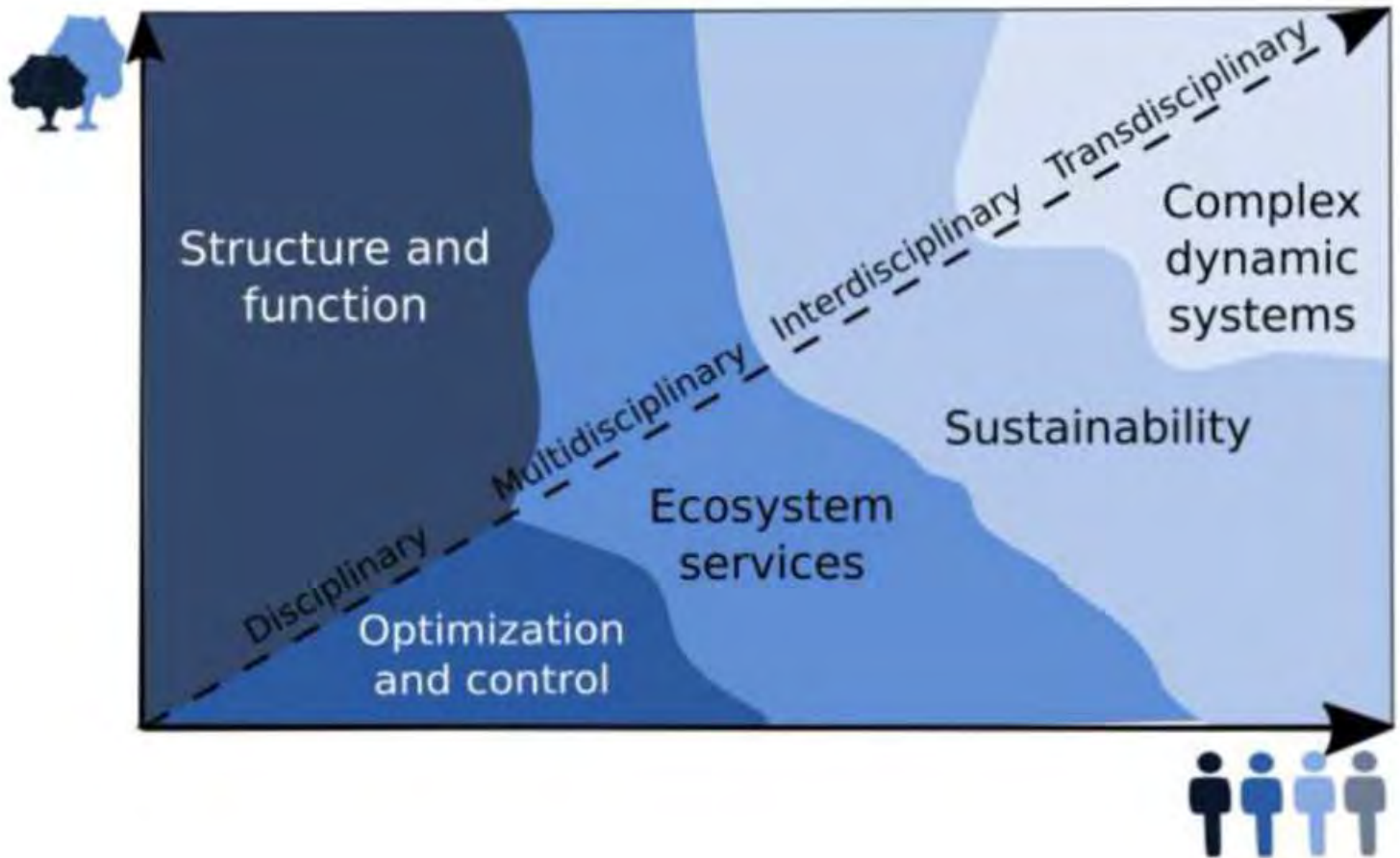


Figure 2. Our proposal for measuring the dynamic dimension of ecosystemic health is based on the idea of criticality as the combination of scale invariance and a balance

[Click here to download Figure esquema.tiff](#)

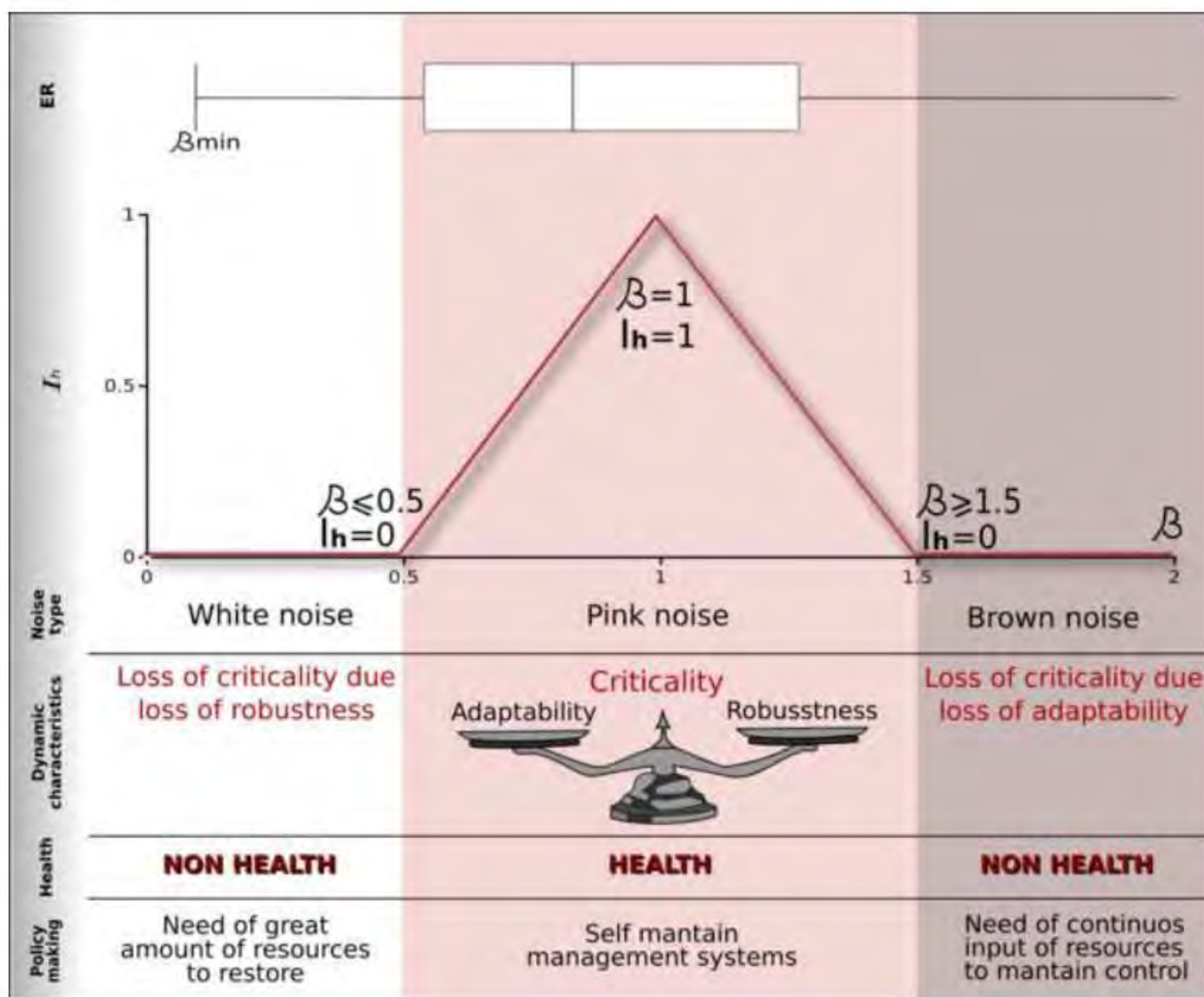


Figure 3. In this figure we show the Land Condition variable used as validation in terms of noise type. Color scale corresponds to ecosystem types in the IGBP nomenclature [Click here to download Figure img01.eps](#)

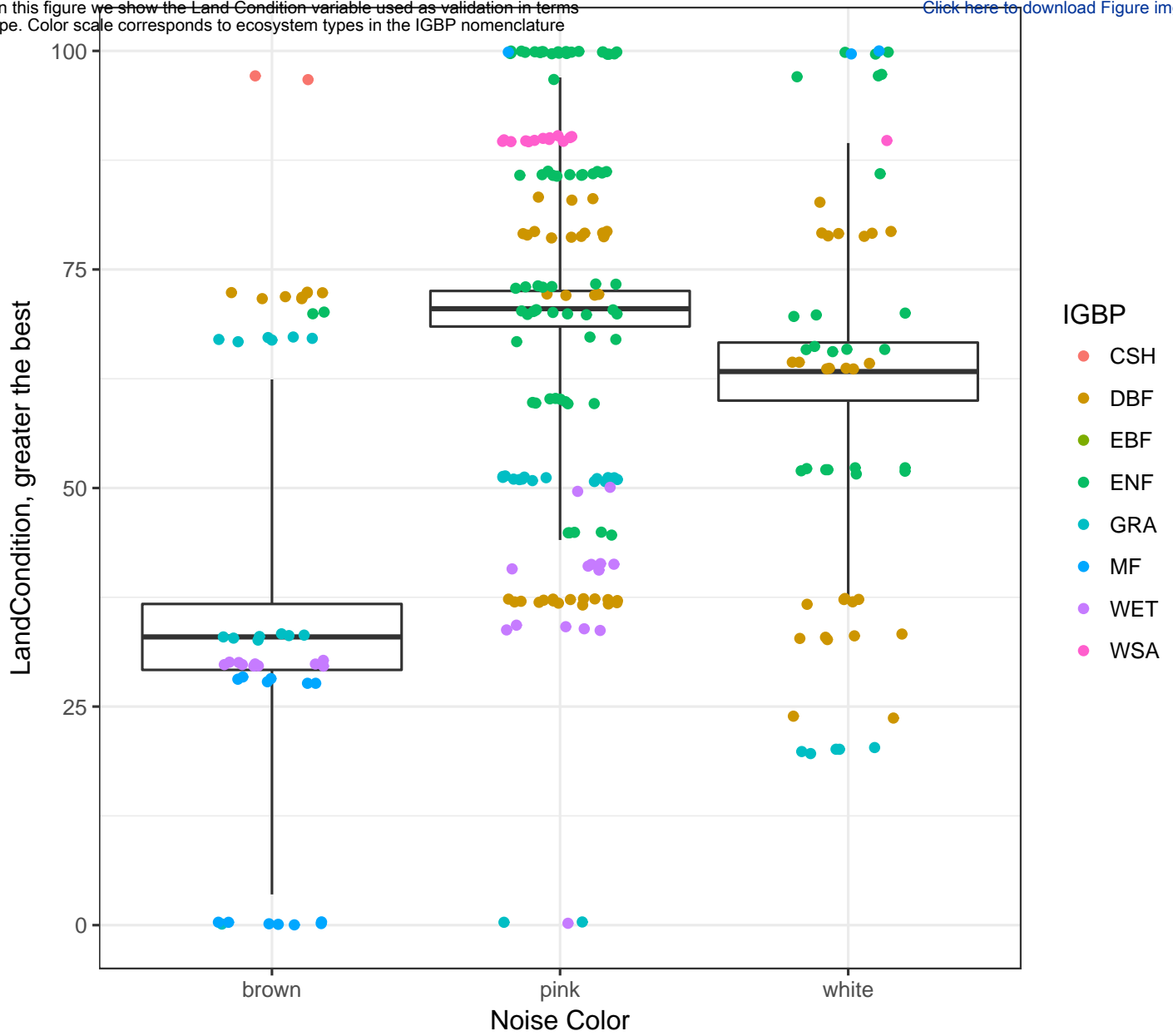


Figure 4. As expected, most ecosystems fall into pink noise. In general, we found that ecosystems out of criticality are older forests or have been altered by human activity or

[Click here to download Figure img02.eps](#)

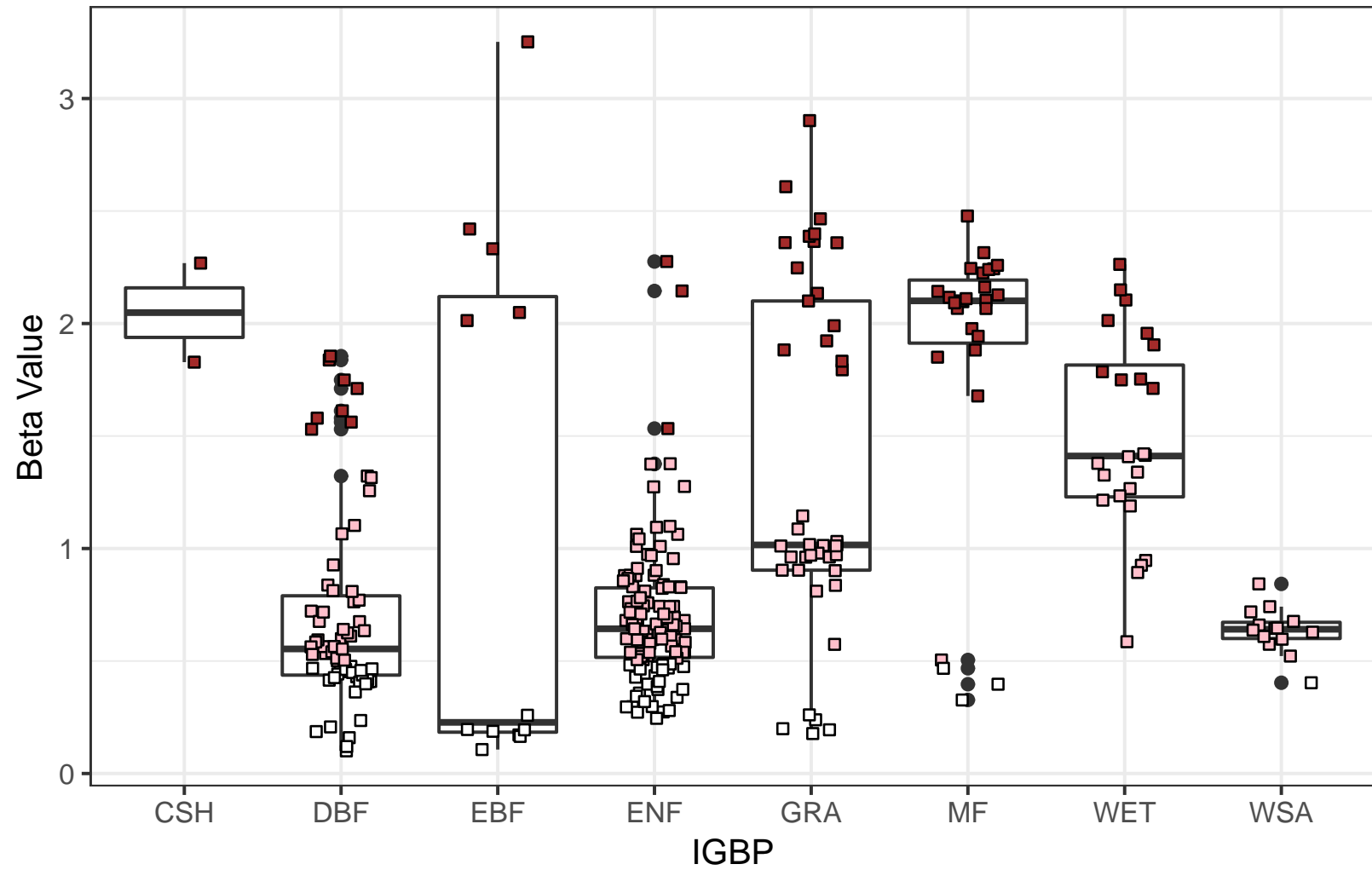


Figure 5. Values of our Ecosystemic Health Index for all ecosystem types (IGBP). The color of each data point corresponds to the type of noise (white, pink and brown)

[Click here to download Figure img03.eps](#)

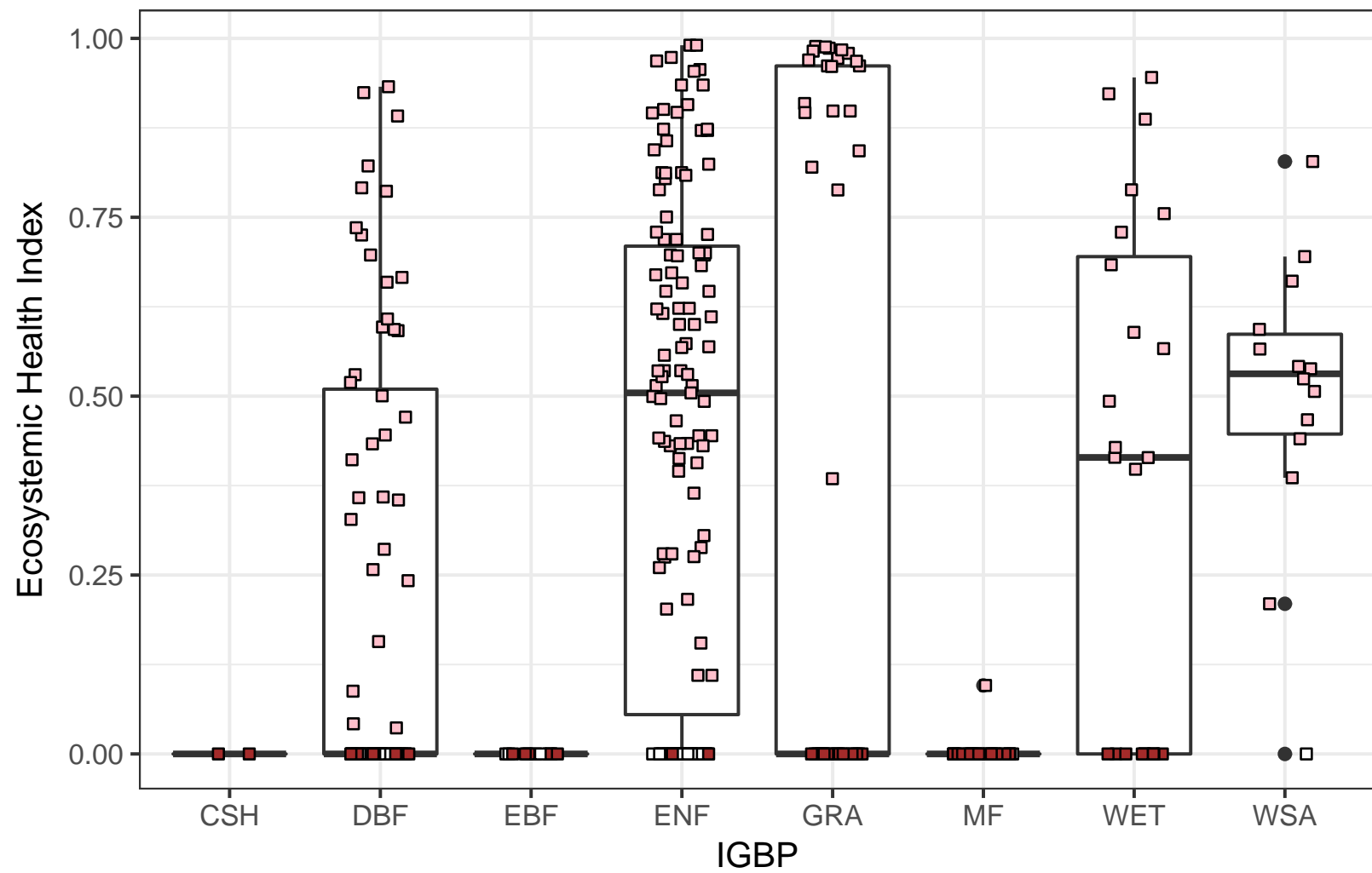


Figure 6. Permutation entropy boxplot in terms of noise color

[Click here to download Figure img04.eps](#)

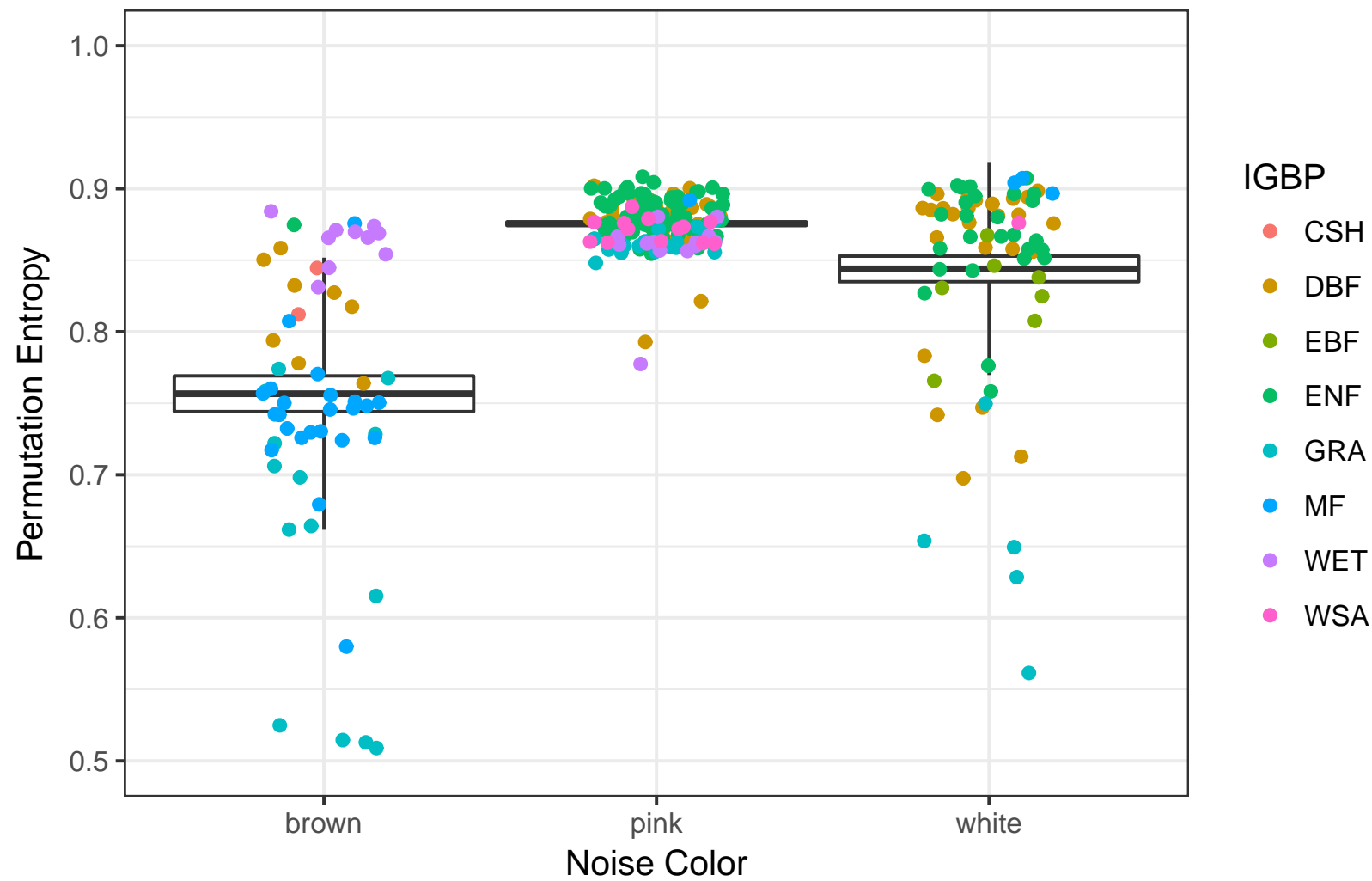


Figure 7. Permutation entropy scatter plot in terms of β value. Entropy reaches highest values around the beta values range for which pink noise is defined

[Click here to download Figure img05.eps](#)

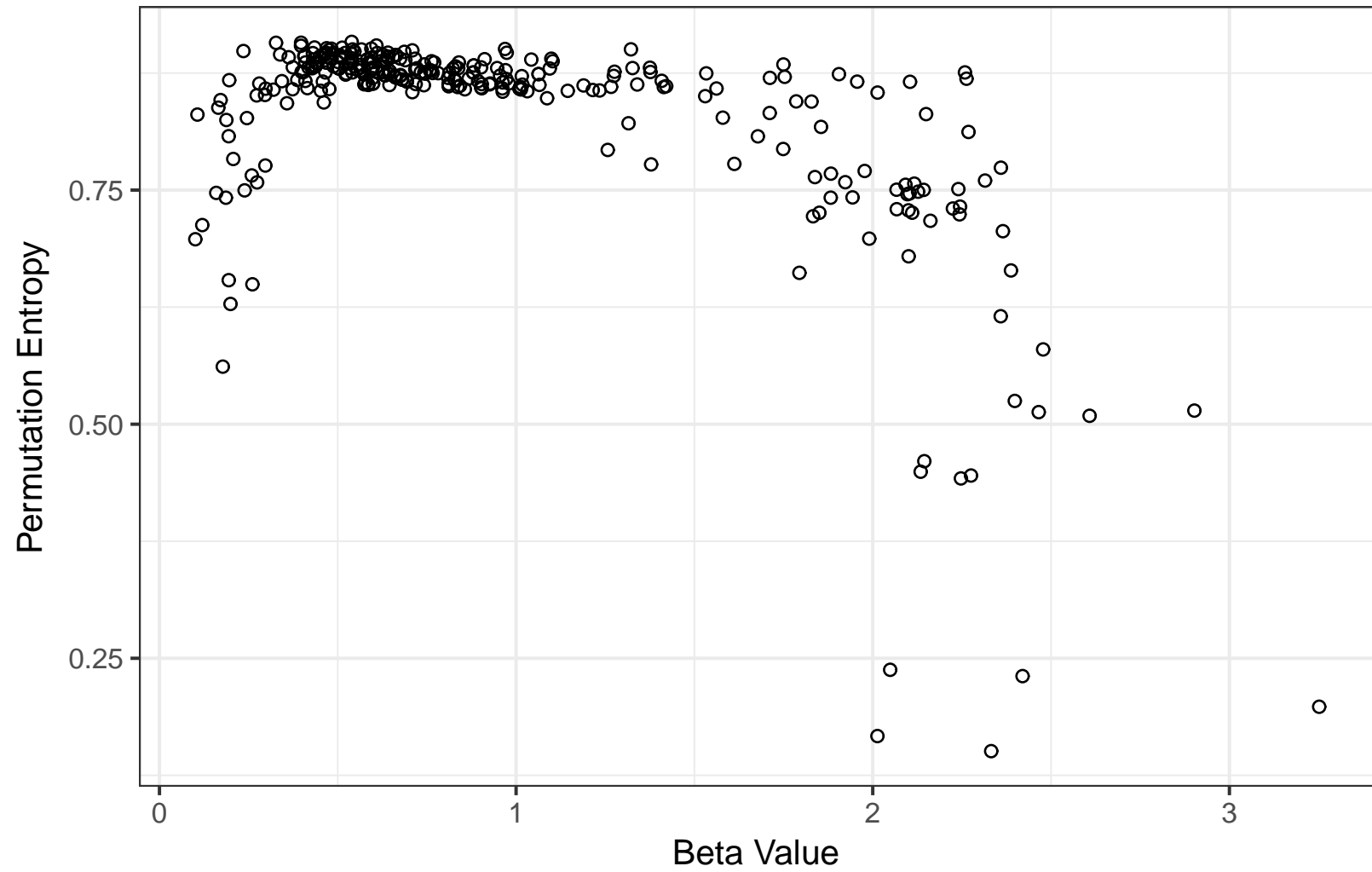


Figure 8. Complexity as a quadratic function of Permutation Entropy

[Click here to download Figure img06.eps](#)

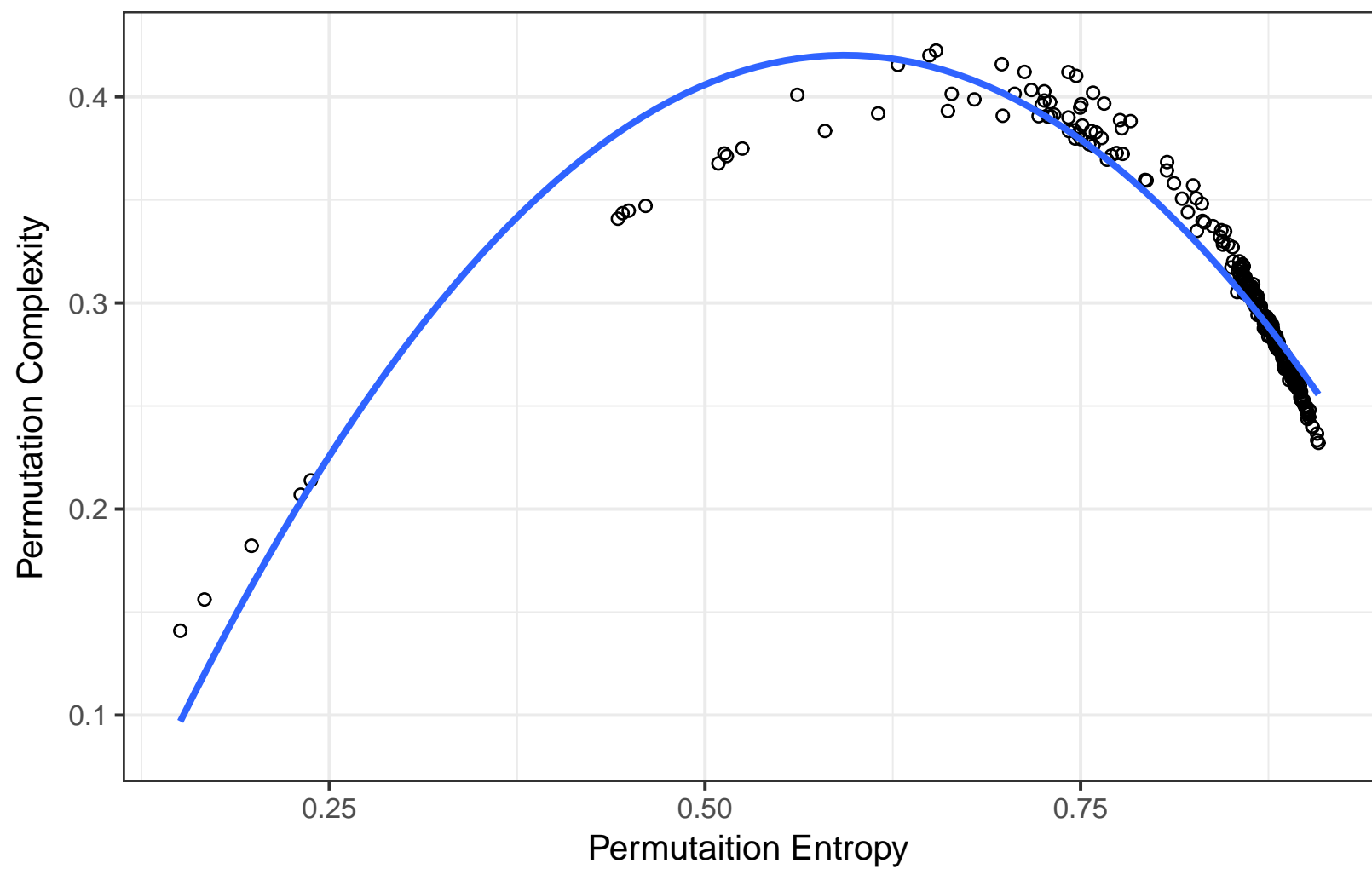


Figure 9. Since criticality (pink noise) appears to be the most healthy and stable (sustainable) type of dynamics, we use it as a leaf variable in a classification tree using

[Click here to download Figure arboIRE.tiff](#)

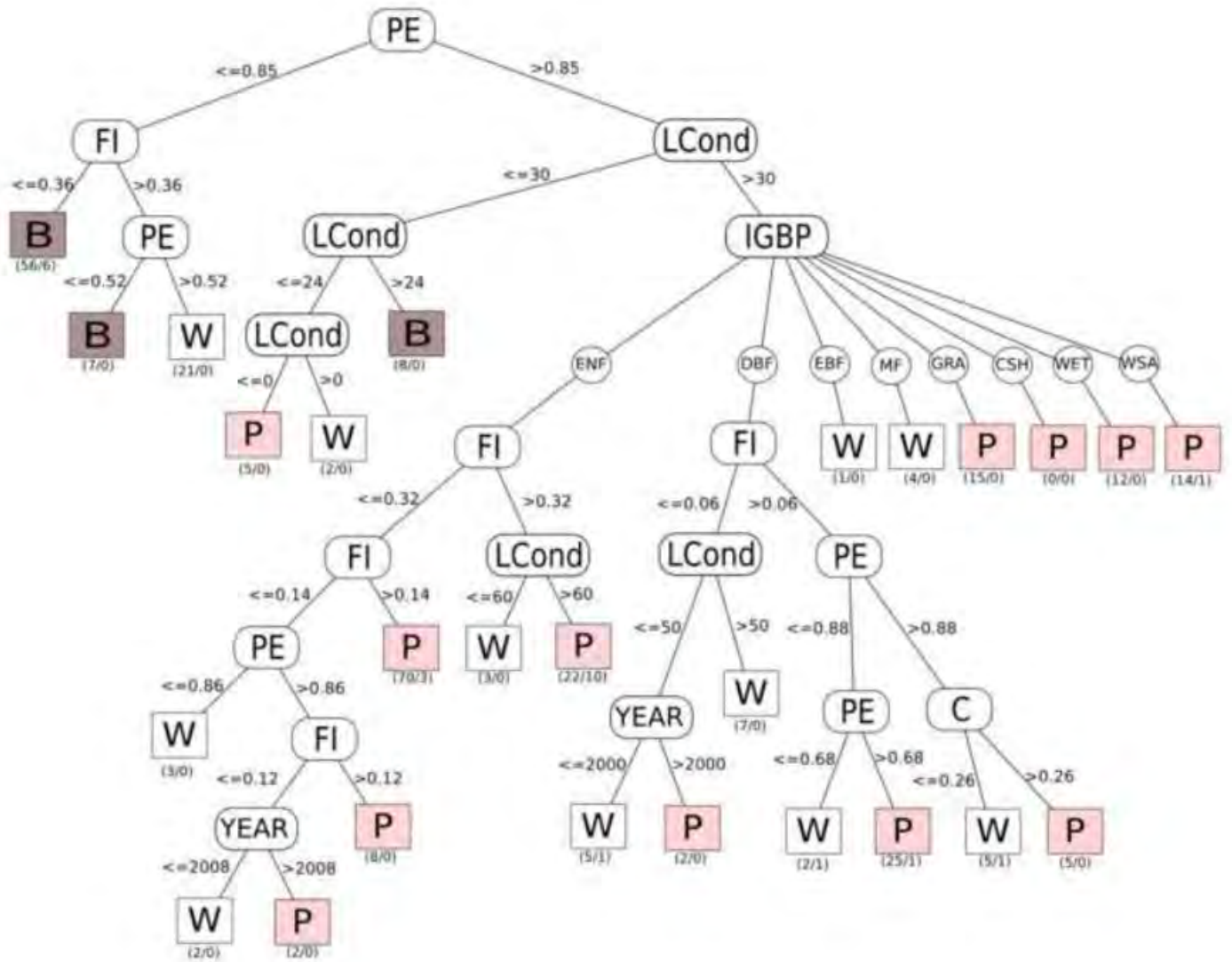


Figure 10. In this map we represent the type of noise in a color scale, and Land Condition as size of the circles

[Click here to download Figure ncolor_vs_lcond.tif](#)

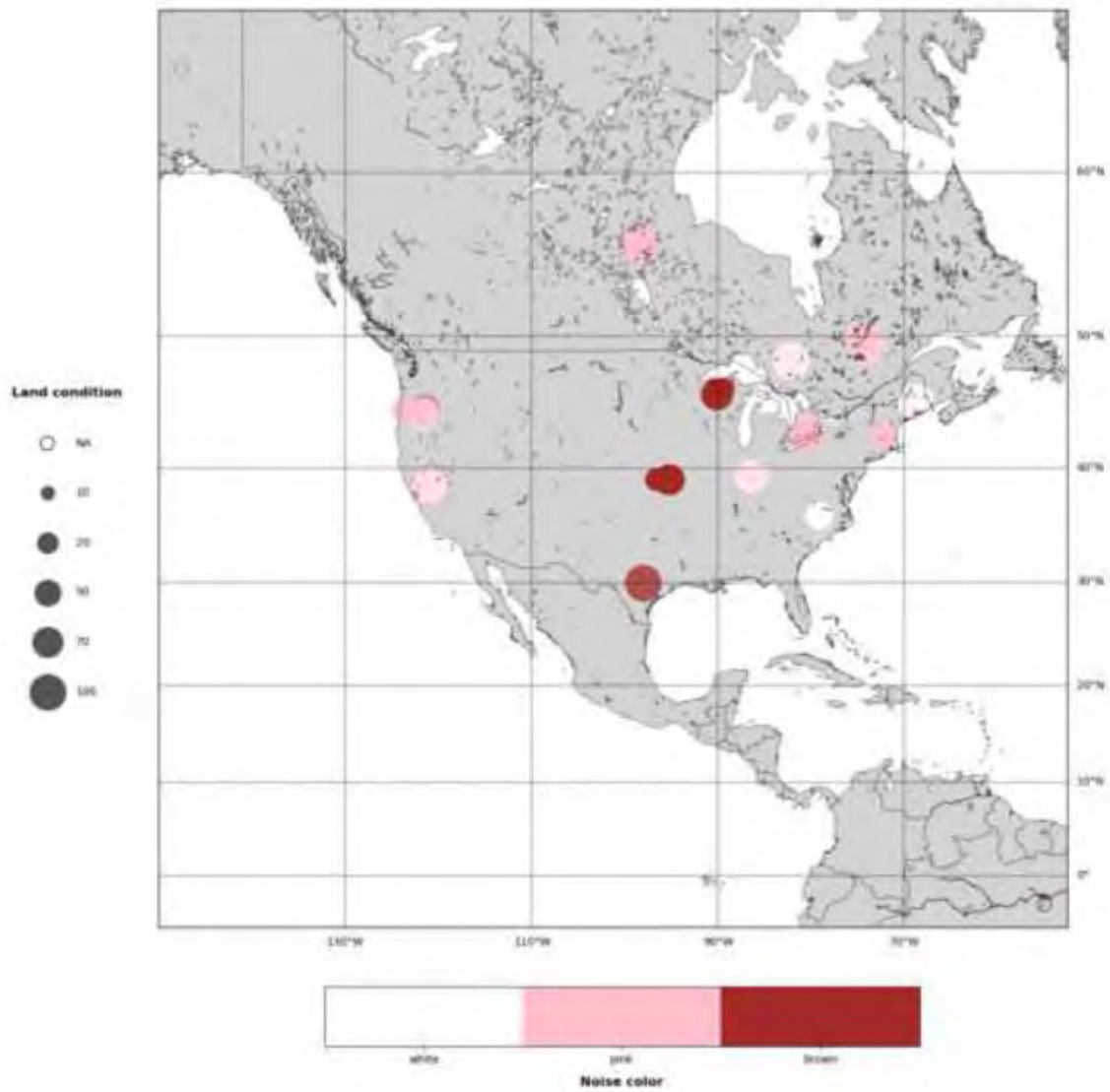
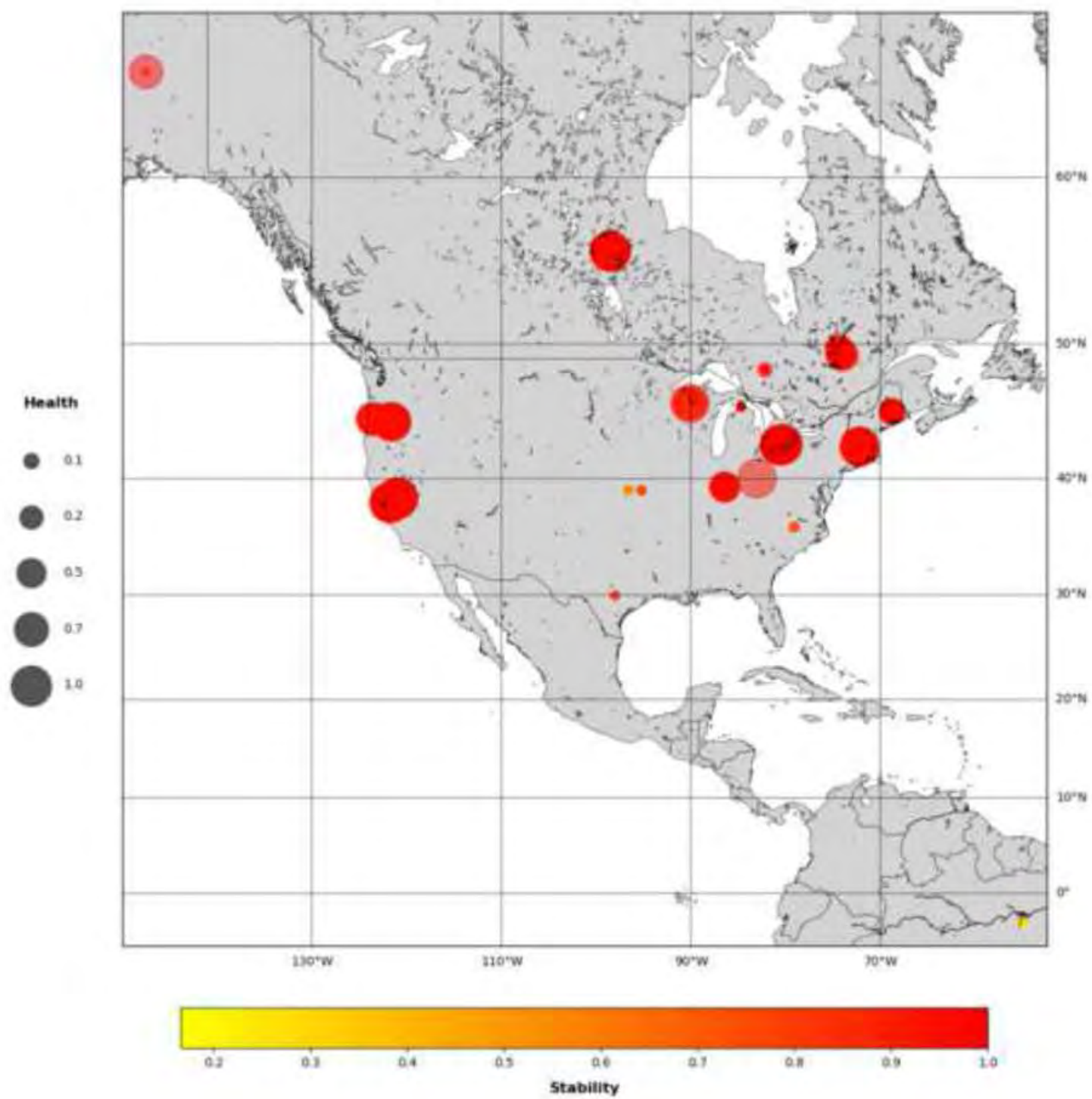


Figure 11. In here we show Ecosystem Stability (Permutation Entropy) in a color scale and Ecosystem Health (as the square

[Click here to download Figure stab_vs_health.tiff](#)



§6.4. Nota de respuesta al artículo de Dallas et al. (2017) en bioRxiv.

1 A comment on “*Species are not most abundant in the centre of their geographic range or climatic niche*”

2

3 Jorge Soberón ⁽¹⁾⁽²⁾

4 A. Townsend Peterson ⁽¹⁾⁽³⁾

5 Luis Osorio-Olvera ⁽³⁾

6 (1) Department of Ecology and Evolutionary Biology and Biodiversity Institute, Lawrence, KS 66049,

7 USA.

8 (2) Corresponding author. jsoberon@ku.edu.

9 (3) town@ku.edu

10 (4) Facultad de Ciencias, Universidad Nacional Autónoma de México, Cd. Universitaria, 04510 Ciudad

11 de México, México. luismurao@comunidad.unam.mx

12

13 **Running title:** Comment on niche-abundance relationships

14 **Keywords:** population density, niche centrality, niche distances

15 **Technical Comment**, 2 figures, 11 references, 789 words

16

17 **Abstract**

18 A study recently published argued against a relationship between population density and position in

19 geographic and environmental spaces. We found a number of methodological problems underlying the

20 analysis. We discuss the main issues and conclude that these problems hinder a robust conclusion about

21 the original question.

22

23 **Introduction**

24 The question of whether population density is related to position in geographic (Sagarin 2002)
25 or ecological niche space (Yañez-Arenas et al. 2012, Martínez-Meyer et al. 2013) is important and still
26 unresolved. In a recent paper in *Ecology Letters*, Dallas et al. (2017) examined the problem using a large
27 dataset of 118,000 sampled populations of >1400 species birds, mammals, and trees. Dallas et al. (2017)
28 failed to detect consistent and significant correlations between population density and distance to the
29 centroids of species' distributions in geographic or environmental spaces, and concluded against the
30 generality of such distance-density relationships. However, the authors' failure to detect significant
31 relationships may result from methodological artifacts, rather than to non-existence of such
32 relationships. We focus on five problems inherent in their analysis.

33

34 **Results**

35 1) The largest dataset analyzed by Dallas et al. (2017) was eBird observations (Sullivan et al. 2009),
36 which are collected without any sampling protocol or plan (there are alternative and better
37 databases, like the Breeding Bird Survey). eBird has biases frequent among observational data, like
38 more observers near cities, and more reporting where a species is rare. Therefore, confounding
39 effects between effort and observer bias may be present, at least for the birds.

40

41 2) Dallas et al. (2017) caution about maximum abundances falling at the periphery of sampled ranges
42 for two of the datasets that they analyzed, but we still worry that true niche centroids will not be
43 represented appropriately. Dallas et al. (2017) largely disregarded parts of species' distributions
44 falling outside the regions for which they had abundance data available. We illustrate this point
45 using the rodent *Dipodomys merriami* Dallas et al. (2017). Figure 1 shows the spatial minimum
46 convex hull (CH) for occurrences in the United States (region in gold on map below). This is

47 considerably less extensive than the range outline for this species from IUCN (Patterson et al. 2003).

48 The geographic centroids based on the two range outlines are markedly distinct.

49

50 A similar problem exists in environmental space. We downloaded the 2-dimensional principal

51 components (PC) used by Dallas et al. (2017). For 1799 localities (debugged and thinned to 0.1°, out

52 of 40,000 available via GBIF), we extracted the PC values for each of the points. Figure 1 shows that

53 the range of environmental space in the full distributional area of *Dipodomys merriami* extends into

54 environmental space not represented in the CH used by Dallas et al. (2017).

55

56 3) Dallas et al. (2017) used CHs to characterize ecological niches of species. CHs are sensitive to outliers

57 (Syväranta et al. 2013), and their centroids may be quite distinct from those obtained using robust

58 estimators (Van Aelst and Rousseeuw 2009). In Figure 2, based on the *D. merriami* example, the CH

59 and a minimum volume ellipsoid (MVE) centroids around the same US data are located in very

60 different positions in niche space.

61

62 4) Dallas et al. (2017) use Euclidean distances as measures of distance to niche centroids, which trace

63 equidistant circles around the centroid. A Mahalanobis distance, estimated using the covariance

64 matrix of the observations of the species in question, would be preferable. Figure 2, shows the US

65 distribution of *D. merriami* occurrences, with centroids and outlines of the CH and MVE. Ignoring the

66 covariance in the realized niche of the species contributes extra bias to characterizations of

67 distances in niche space. Dallas et al. (pers. comm.) explored Mahalanobis distances, without finding

68 major differences in the results, an observation deserving further exploration.

69

70 **5)** The data provided in the Supplementary Materials of Dallas et al. (2017) reveal that the population-
71 density data points have coordinates with precision of 100 m or finer. However, the climate data
72 used in the paper have a resolution of $\sim 0.042^\circ$, or squares of ~ 4600 m on a side. This means that
73 multiple abundance data points may fall within a single climate pixel, introducing a further problem
74 in the analysis, as shown in lines 365-370 of the code provided by the authors. Correcting this
75 methodological problem leaves 40 instead of 81 species for mammals, 49 instead of 63 for fishes
76 with ≥ 10 different abundance/climate points; all of the birds and 165 of the 166 tree species.

77
78

79 **Conclusions**

80 Dallas et al. (2017) provide the largest-scale analysis available to date of relationships between
81 population density and positions in geographic and environmental spaces. Their negative results
82 contrast with previous empirical work (Yáñez-Arenas et al. 2012, Martínez-Meyer et al. 2013) and with
83 theoretical arguments supporting such a relationship (Maguire 1973, Osorio-Olvera et al. 2016).

84 In this communication, we identify a series of methodological problems underlying the results of
85 Dallas et al. (2017). Although some of them may be of minor importance (e.g., Euclidean vs.
86 Mahalanobis distances), others may be more significant. We suggest that this important and interesting
87 problem remains far from settled.

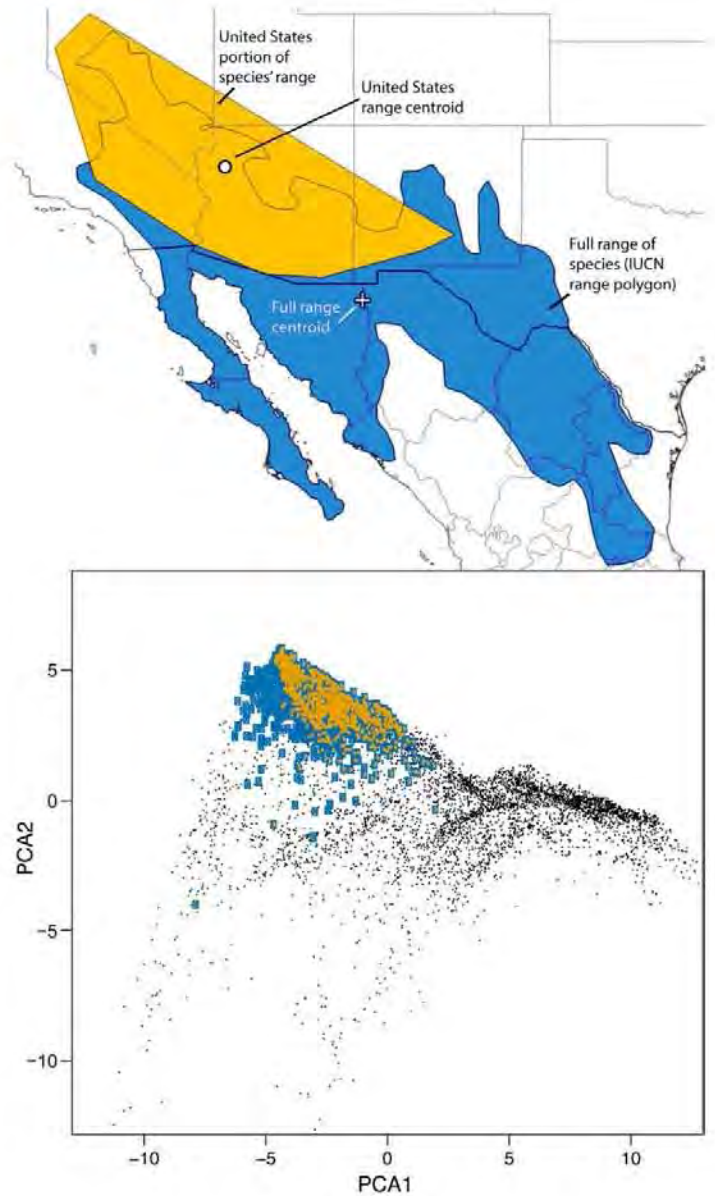
88

89 **Literature Cited**

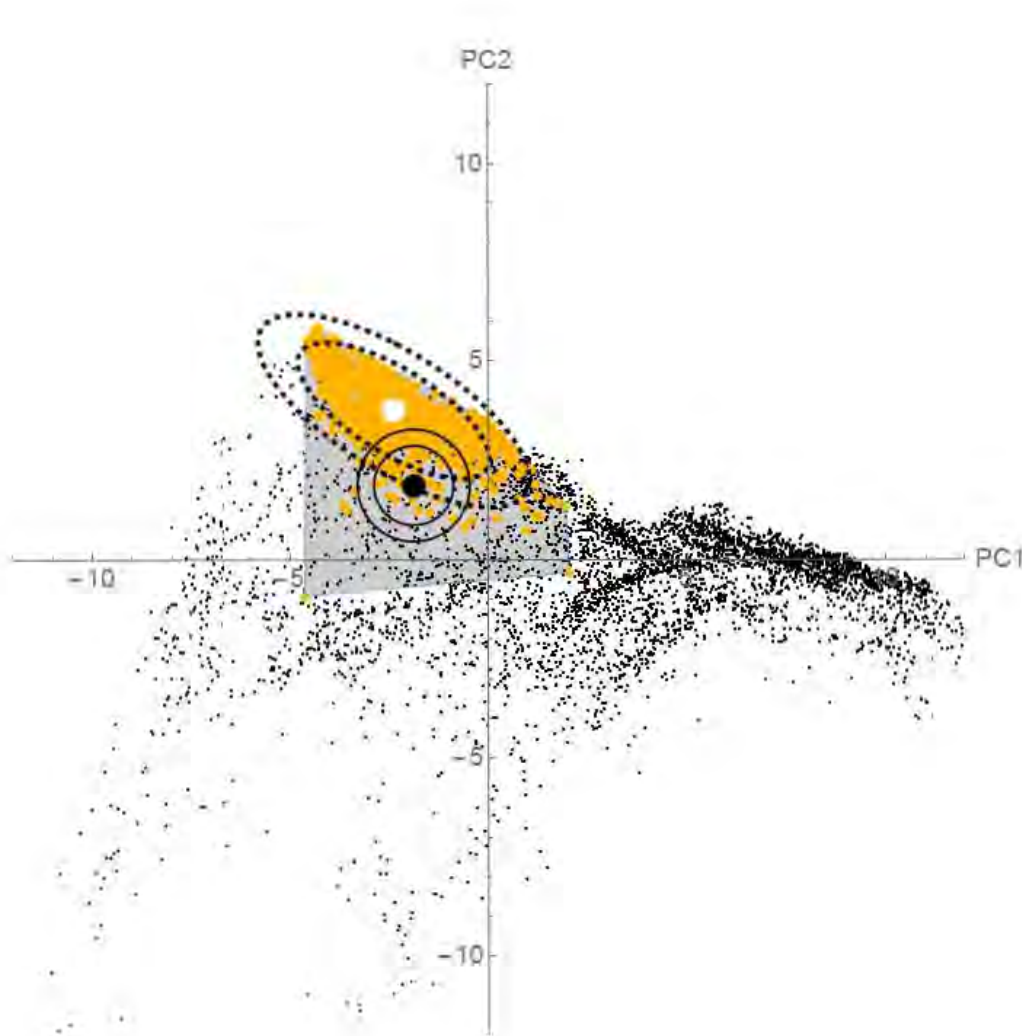
90 Dallas, T., R. R. Decker, and A. Hastings. 2017. Species are not most abundant in the centre of their
91 geographic range or climatic niche. *Ecology Letters* **20**:1526-1533.
92 Maguire, B. 1973. Niche response structure and the analytical potentials of its relationship to the
93 habitat. *American Naturalist* **107**:213-246.
94 Martínez-Meyer, E., D. Díaz-Porras, A. T. Peterson, and C. Yáñez-Arenas. 2013. Ecological niche structure
95 and rangewide abundance patterns of species. *Biology Letters* **9**:20120637.
96 Osorio-Olvera, L., M. Falconi, and J. Soberón. 2016. Sobre la relación entre idoneidad del hábitat y la
97 abundancia poblacional bajo diferentes escenarios de dispersión. *Revista Mexicana de*
98 *Biodiversidad* **87**:1080-1088.

- 99 Patterson, B. D., G. Ceballos, W. Sechrest, M. F. Tognelli, T. Brooks, L. Luna, . . . B. E. Young. 2003. Digital
100 distribution maps of the mammals of the Western Hemisphere. CD-ROM compiled by
101 NatureServe.
- 102 Sagarin, R. 2002. The 'abundant centre' distribution: to what extent is it a biogeographical rule? *Ecology*
103 *Letters* **5**:137-147.
- 104 Sullivan, B. L., C. L. Wood, M. J. Iliff, R. E. Bonney, D. Fink, and S. Kelling. 2009. eBird: A citizen-based bird
105 observation network in the biological sciences. *Biological Conservation* **142**:2282-2292.
- 106 Syväranta, J., A. Lensu, T. J. Marjomäki, S. Oksanen, and R. I. Jones. 2013. An empirical evaluation of the
107 utility of convex hull and standard ellipse areas for assessing population niche widths from
108 stable isotope data. *PLoS ONE* **8**:e56094.
- 109 Van Aelst, S. and P. Rousseeuw. 2009. Minimum volume ellipsoid. *Wiley Interdisciplinary Reviews:*
110 *Computational Statistics* **1**:71-82.
- 111 Yañez-Arenas, C., E. Martínez-Meyer, S. Mandujano, and O. Rojas-Soto. 2012. Modelling geographic
112 patterns of population density of the white-tailed deer in central Mexico by implementing
113 ecological niche theory. *Oikos* **121**:2081-2089.
- 114

115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134



135 Figure 1. Geographic and environmental spaces for *Dipodomys merriami*. Top: extent of occurrence
136 polygon for the distribution of *D. merriami* in the United States (gold, centroid shown by black and white
137 circle), representing the range area analyzed by Dallas et al. (2017), and the full range of the species
138 (blue, centroid show as a cross; polygon from IUCN). Note that the true range centroid falls outside of
139 the convex hull analyzed by Dallas et al. (2017). Bottom: 1799 data points from GBIF (see text) in a space
140 of the first two climatic principal components used by Dallas. et al. (2017; see text). Points in gold are
141 the reduced portion (United States) of the species range analyzed by Dallas et al. (2017); points in in
142 blue cover the entire range of the species.



143

144 Figure 2. GBIF points for *Dipodomys merriami* in environmental space, showing differences between
145 methods for delimiting niches and calculating niche-centroid distances. The black circle is the centroid of
146 the convex hull (gray-shaded polygon), showing the strong effect of one outlier point. The white circle is
147 the centroid of a 95% minimum volume ellipsoid that is able to ignore the outlier. Circles are Euclidean
148 distances of radii 1 and 2, for the convex hull centroid; the dashed ellipsoids are the equivalent distances
149 (Mahalanobis distances) taking into account the covariance shown by the points in gold (see text). Note
150 the striking differences between the two methodologies in both shape of the niche estimated and the
151 distances that result; in particular, note that the centroid estimated via convex hulls falls at the
152 periphery of the cloud of points for the species' occurrence.

153

154