



CONABIO
1992-2017
XXV ANIVERSARIO

Teoría y praxis en el gobierno: datos, software y ecuaciones de las áreas de distribución.

Jorge Soberon



De que voy a hablar

- La Comisión Nacional para el Conocimiento y Uso de la Biodiversidad es una institución gubernamental, con un propósito técnico y científico
- Me gustaría relatar como ciertas demandas de usuarios, relacionadas con el objetivo de la CONABIO, condujeron a avances propiamente científicos.

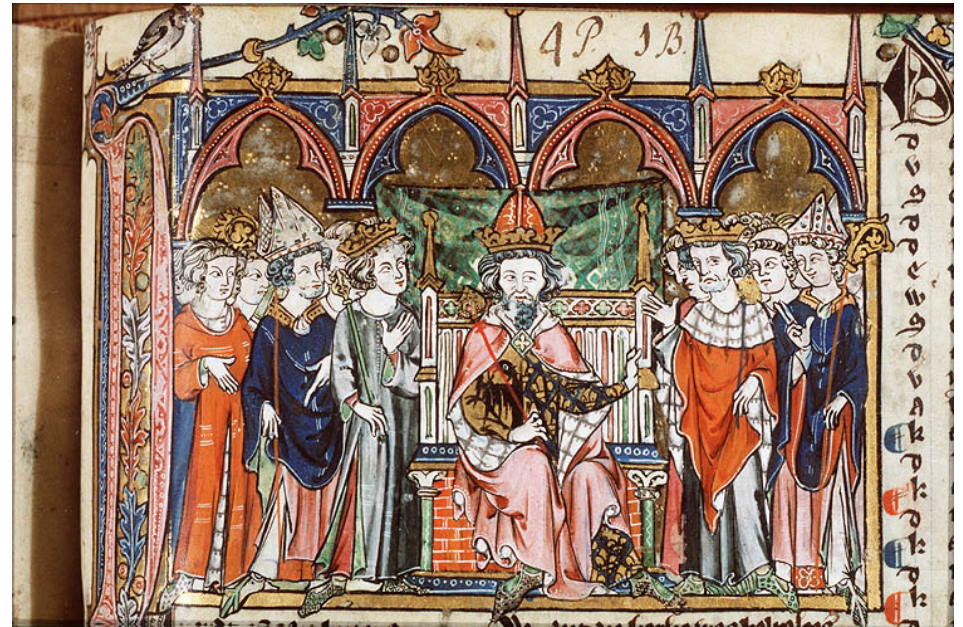
La historia completa de CONABIO es fascinante, y llena de anécdotas

- Cómo se conceptualizaron los problemas
- Cómo se convencieron a los colaboradores externos y se organizaron los proyectos externos para obtener datos
- Cómo se desarrolló la capacidad de bioinformática
- Cómo se asumió un liderazgo internacional...



En esta platica me limitaré a solamente uno de estos puntos

- Cómo es que ciertas preguntas prácticas, de usuarios no técnicos, condujeron a avances teóricos.
- Hay quien espera que un descubrimiento publicado en un artículo sea adoptado rápidamente por “los tomadores de decisiones”
- Yo voy a relatar un caso inverso. Primero la praxis, luego la teoría.



Tomadores de decisiones
Original en la *Koninklijke Bibliotheek*
Países Bajos

La CONABIO

- Comisión Intersecretarial, creada en 1992 con el objetivo de promover y coordinar acciones orientadas al conocimiento y uso sustentable de la riqueza biológica de México; especialmente aquellas dirigidas a la obtención, organización, análisis y difusión de los datos de esa riqueza.
- Está presidida *ex officio* por el Presidente de la República.
- Opera por medio de una Coordinación Nacional y una Secretaría Ejecutiva
- El Secretario Técnico es el Titular de la SEMARNAT

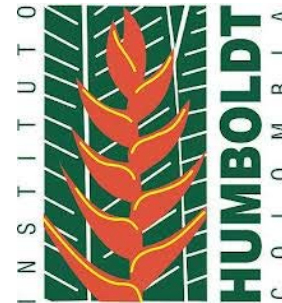


CONABIO

COMISIÓN NACIONAL PARA EL
CONOCIMIENTO Y USO DE LA BIODIVERSIDAD

La CONABIO fue creada por una necesidad política.

- En 1992, los gobiernos se iban a juntar en Rio para finalmente proteger el ambiente.
- En algunos países ya existían agencias mas o menos gubernamentales.
 - ERIN, Australia
 - InBio, Costa Rica
- La CONABIO de México fue creada con el objeto de tener algo que presentar en la cumbre de Rio



Australian Government

Pero afortunadamente la diseñaron los mejores ecólogos de México

- Jose Sarukhan
- Daniel Piñero
- Rodolfo Dirzo



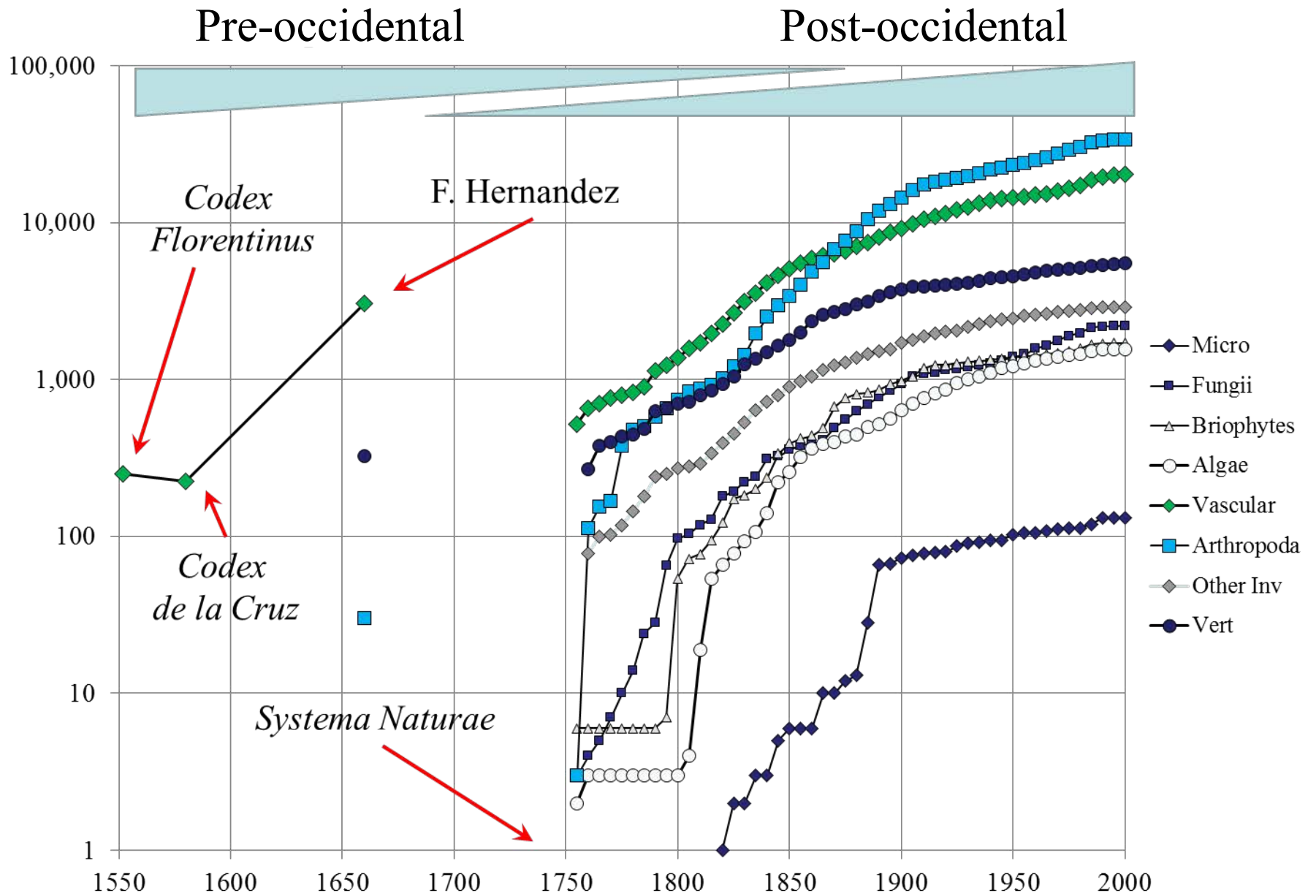
¿Y que pensaron? Bueno, México es rico en recursos naturales vivos.

En el país hay cientos de miles de especies de plantas, animales y microorganismos. Su importancia económica, cultural y científica es enorme. Por ejemplo, en México tenemos registradas a la fecha:

- Plantas medicinales y alimenticias: 4,000
- Especies forestales: 600
- Especies pesqueras: 500
- Especies invasoras (plagas potenciales): 1,000
- Especies cinegéticas y ornamentales: 200
- Especies protegidas por la legislación mexicana: 2,500
- Especies parientes silvestres de organismos transgénicos: 890
- Especies con potencial biotecnológico: decenas de miles

Ellos sabían que, si se quiere conservar y usar bien ese patrimonio, es preferible tener información que no tenerla (!)

- Pero había muy poca información *digitalizada, y organizada*.
- Había, y hay, muchísimos artículos (decenas de miles), mapas, tesis, etiquetas de colecciones científicas, *libros*, bases de datos, imágenes ... pero sin organización.
- Si bien existían usuarios (dependencias gubernamentales, NGOs, organizaciones de productores...), estos muy raramente hacían uso de la información existente, o buscaban nueva, mas que como tramite burocrático (MIAs, para obtener permisos).
- ¡No había preguntas, ni datos! O sea, si había, pero estaban desconectados.



Llorente, J y S. Ocegueda 2010 *Estado del conocimiento de la biota.*

In Soberon, G. Halffter y J Llorente. *Capital Natural de Mexico Vol I.* CONABIO, Mexico



La función principal de CONABIO:

ARTICULO SEXTO. La Comisión tendrá las siguientes funciones:

- I. [establecer] ... un programa sobre los inventarios biológicos del país ...
- II. Sintetizar la información relativa a los recursos biológicos del país, en un **banco de datos que deberá mantenerse permanentemente actualizado.**

Y la pregunta subyacente: ¿qué datos?

El examen de la experiencias, sobre todo de Australia, sugirió que muchos problemas relevantes solamente requieren respuestas sobre **localización.**

Las preguntas sobre localización se refieren a la presencia de entidades en lugares:

- ¿Qué entidades se encuentran en un lugar dado?
- ¿Dónde se puede encontrar una entidad dada?

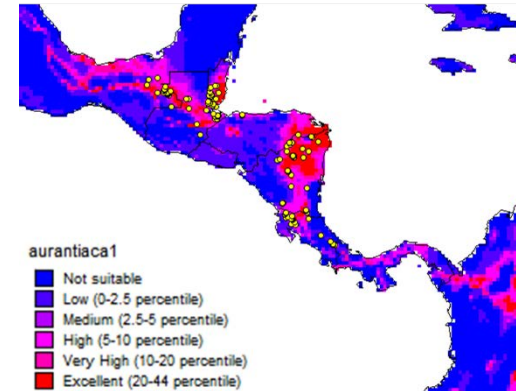
A pesar de su aparente simplicidad, estas preguntas son relevantes tanto para la ciencia básica como para la aplicada. Sobre todo, **los datos para responderlas ya existían**

Ejemplos reales de preguntas sobre localización:

- ¿Cuáles son las áreas prioritarias para la conservación de la biodiversidad en México?
- ¿Qué especies de plantas medicinales en peligro, se encuentran en las áreas Tzeltales de los altos de Chiapas?
- ¿Cuál es la distribución de los vectores del hantavirus en México?
- ¿Cuál es el área de distribución potencial de la palomilla del nopal, *Cactoblastis cactorum*?

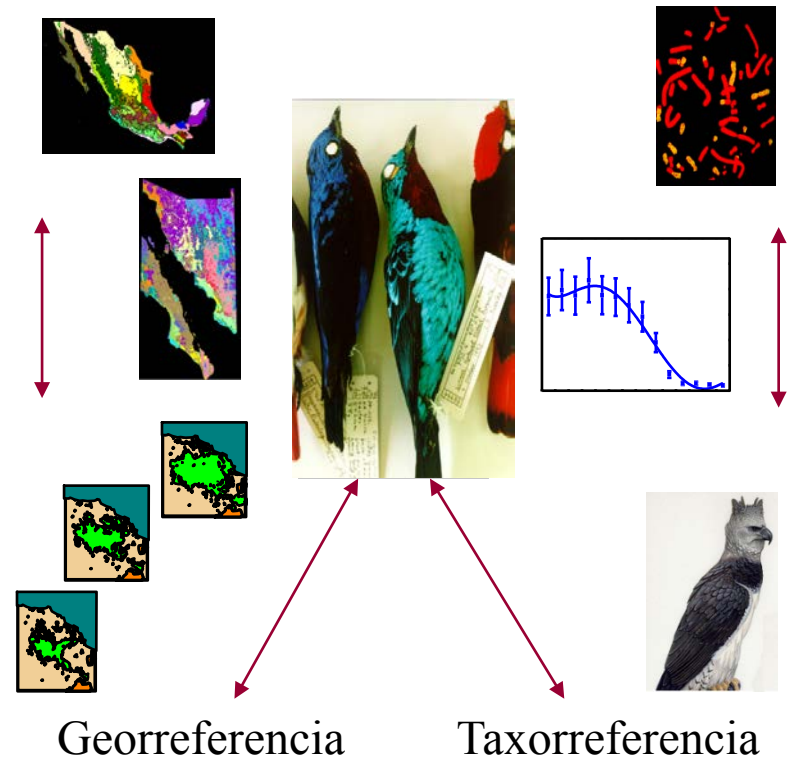
Pero igual se requería “investigación”

- No existían:
 - bases de datos,
 - ni su diseño conceptual,
 - ni sus esquemas,
 - ni sus “diccionarios de autoridades”
 - ni, obvio, su instrumentación práctica
- No existían los datos del medio ambiente físico
- No había **teoría** sobre como analizar, visualizar, interpretar ni extrapolar tales datos...



Las preguntas sobre localización y las colecciones científicas

- Las colecciones científicas del mundo, cuentan con alrededor de 3 mil millones de especímenes
- Los especímenes proveen conexiones entre la información estructurada geográficamente y la información indizada por el nombre en latín.
- Los Australianos estaban (1992) usando estos datos para contestar preguntas gubernamentales



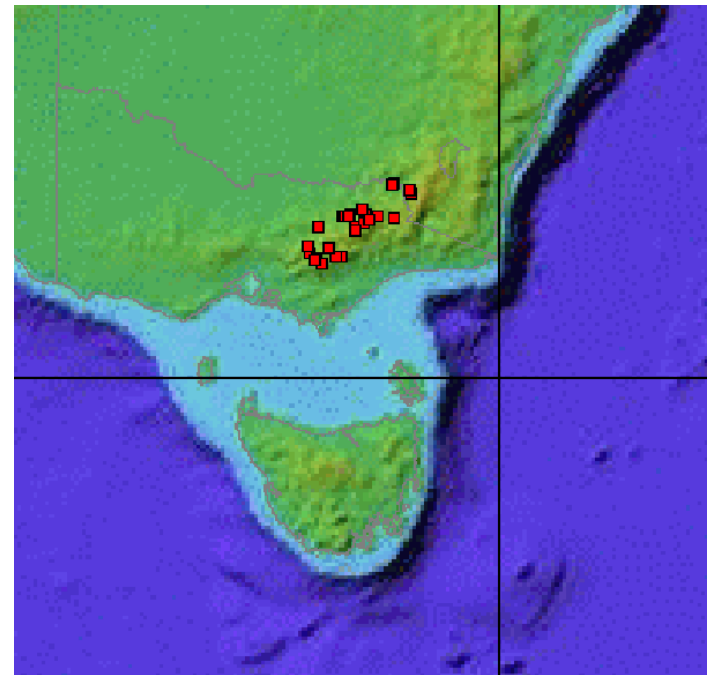
CONABIO, Mexico, 1992

- Nos hacían preguntas relacionadas con la distribución de las especies.
- El gobierno (federal) quería:
 - Identificar buenos sitios para la conservación
 - Poder defender sus decisiones
 - ¡Se necesitaban datos de las distribuciones!



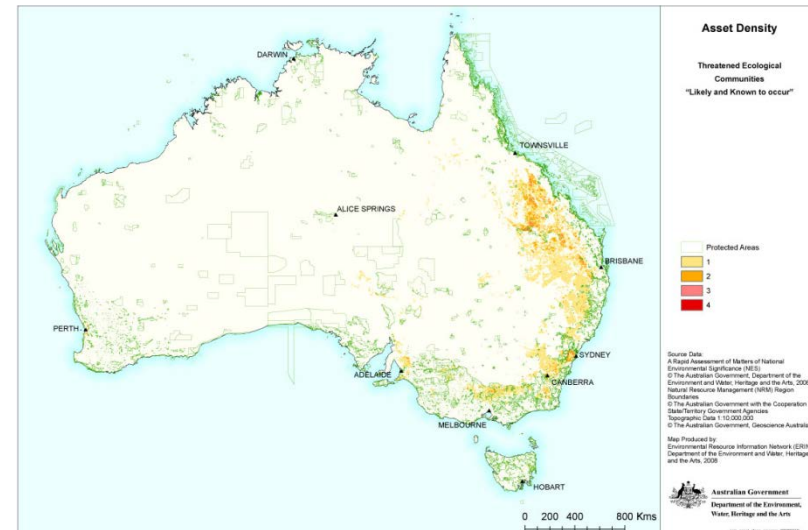
En 1992 los Australianos:

- Habían digitalizado los datos de colecciones grandes
- Habían desarrollado software para automatizar la descripción de especies
- Habían desarrollado software para **predecir la distribución de especies**
- Todo esto en CSIRO y ERIN, agencias gubernamentales



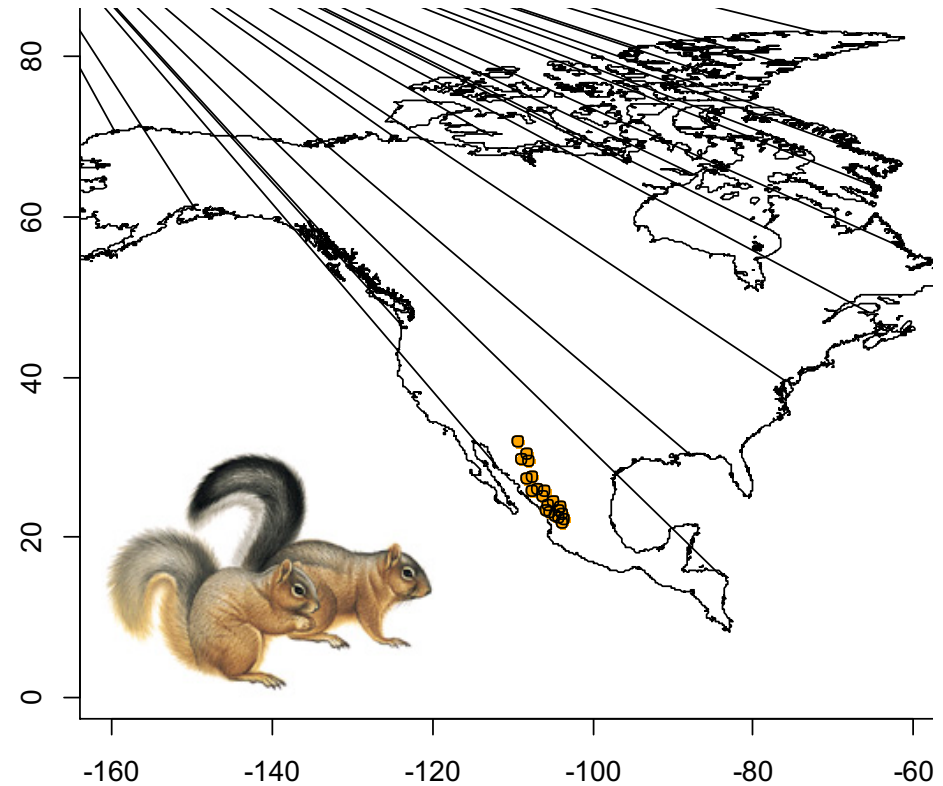
Modelacion de Distribuciones

- CONABIO mando una misión a Australia. Regresamos pasmados de admiración
- ¡Había que imitar a los Australianos! Compilar bases de datos de ejemplares de museo, obtener cartografía electrónica...
- Modelar distribuciones
- Se hacia con el software llamado GARP, que usa **inteligencia artificial** y que corría en una **supercomputadora**, operada por **Internet**. En 1994, NADA podía ser mas “cool”.



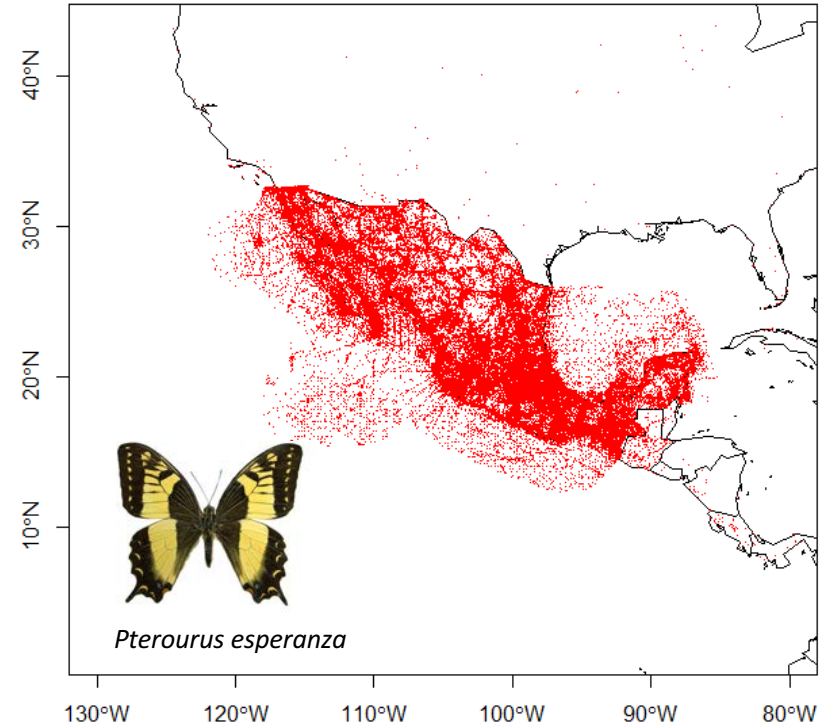
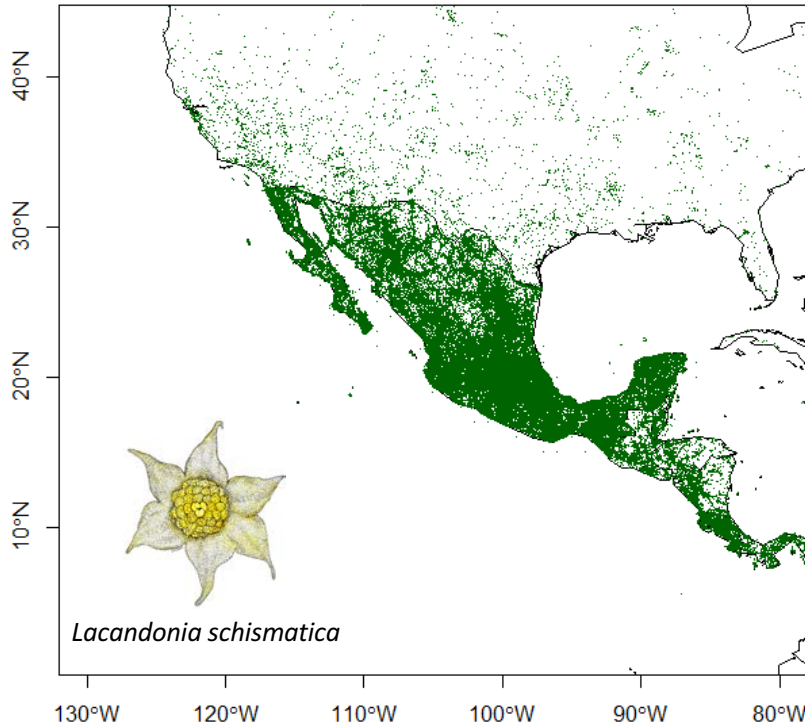
¿Que se necesita?

- Primero, puntos de “ocurrencias”
 - Hoy, hay cientos de millones, ¡disponibles públicamente! En los 90s no se pensaba que esto fuera posible
 - Cómo se crearon estas bases de datos públicas, y mundiales, es otra historia, en la cual CONABIO tuvo un papel protagónico (Raul Jimenez, Patricia Koleff, Hesiquio Benitez)



Sciurus nayaritensis

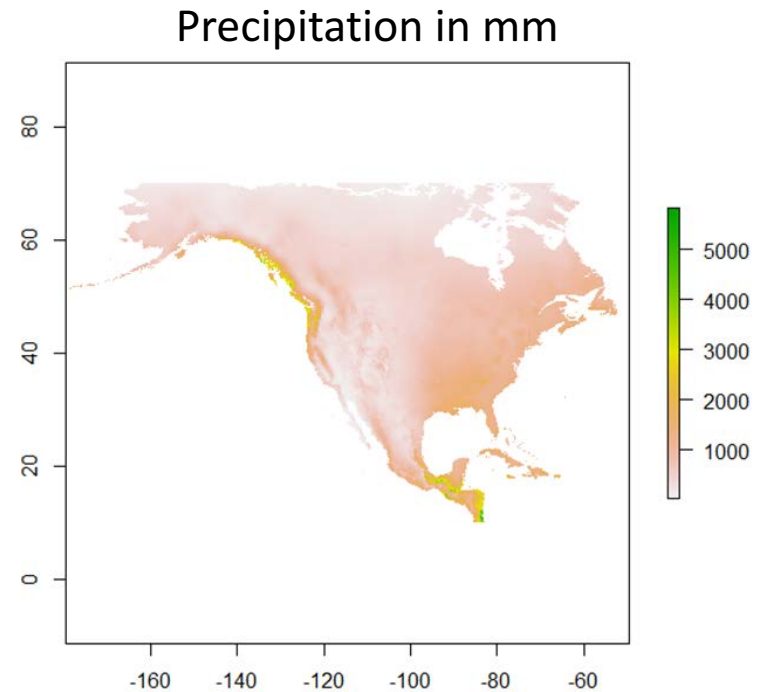
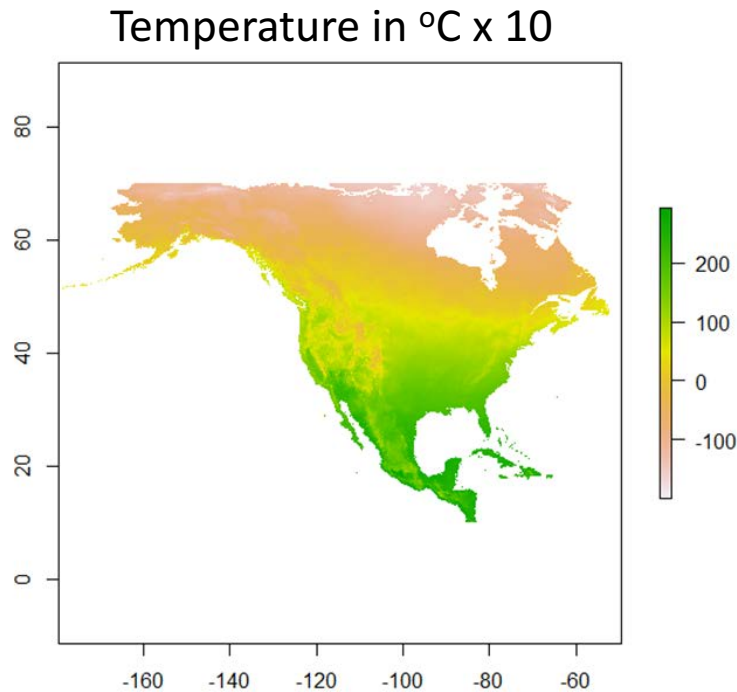
Ahora ya hay chorros de datos: *SNIB+ GBIF* (Platica de Raul Jimenez)



Plantae, 2,733,311 registros, 28,694 nombres *Animalia*, 4,492,607 registros, 48,645 nombres

Cómo obtuvimos todos estos datos es otra historia (Ana Luisa Guzman, Patricia Jorge Llorente, Patricia Koleff, personal de CONABIO, científicos mexicanos...)

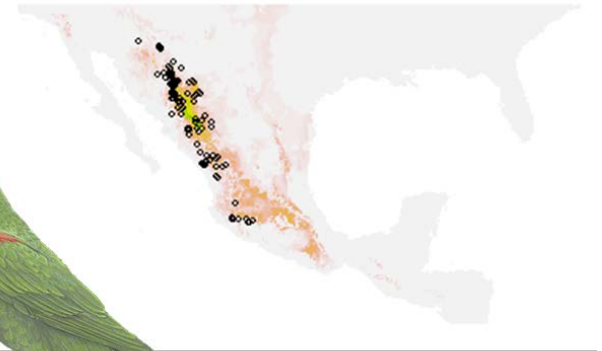
Segundo, “capas” electronicas (ahora existen Terabytes) con datos geofisicos



Cómo se obtuvieron las primeras para México, y la capacidad de usar datos satelitales, es otra historia, donde CONABIO tuvo un papel protagónico (Enriqueta Martinez, Raul Jimenez, Rainer Ressler, Michael Schmidt, Isabel Cruz...)

Y, tercero, los algoritmos y programas correctos.

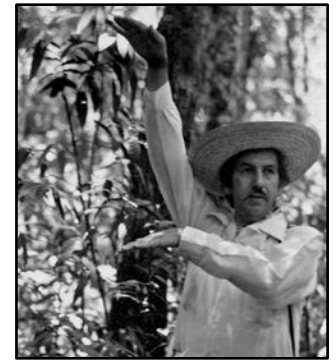
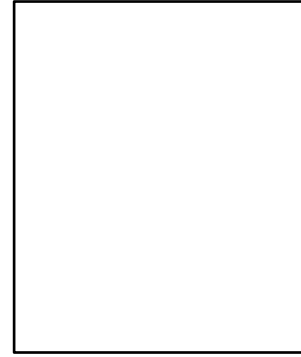
- Al principio usábamos GARP (basado en inteligencia artificial), cortesía Australiana.
- Costó tiempo y dinero, pero pronto la CONABIO estuvo lista. Usando la base de datos de aves (~200,000 registros), y datos climáticos para México...
- ¿Por qué aves? Por Town Peterson, Adolfo Navarro y Hesiquio Benitez (¡también es otra historia!)
- Y empezamos a modelar...



Rhynchopsitta pachyrhyncha
Image from CONABIO

Pequeño paréntesis

En México, desde hace medio siglo había pioneros trabajando en lo que ahora se llaman “modelos de nicho.” Solo para no olvidarnos...



Faustino Miranda

Enriqueta Martinez

Lorrain Giddings

Margarita Soto

Arturo Gomez-Pompa

Garcia, E., M. Soto & F. Miranda (1961) *Larrea* y Clima. *Ann. Inst. Biol.* 31:133-171

L. Giddings (1970) Diseño del Sistema “Bioclimas” INIREB, publicacion interna.

Gomez-Pompa A. & L. Nevling (1970). La Flora de Veracruz. *Ann. Inst. Biol.* 41:1-2
La serie “*Bioclimatologia de Veracruz en los 1990s*”

M. Soto, L. Giddings & , M. Gomez (1996) Algunos usos de Bioclimas. *Pubs. Esps.Inst. Invs. Geog.*

Empezaron a salir resultados...

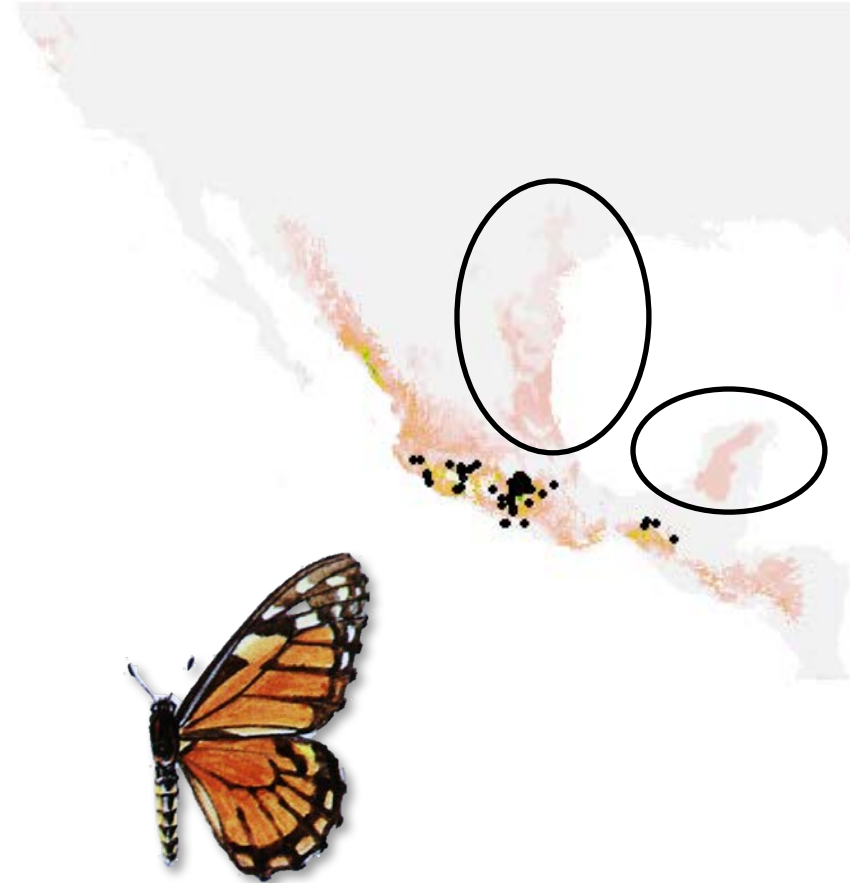
- Y a una resolución gruesa, que es lo que el Gobierno Federal necesita, ¡salían buenos resultados!
- El Supremo Gobierno estaba satisfecho con su CONABIO
- Pero había problemas.
- Cada cuarta o quinta predicción era una basura, y... **“sobrepredecía”**



Rhynchopsitta pachyrhyncha
Imagen de CONABIO

Sobrepredicciones

- Todo mundo odiaba las sobrepredicciones
- Salía (frecuentemente) una muy bonita predicción junto con algo totalmente absurdo. Áreas locas lejos de donde la especie realmente ocurría.
- Esto era muy irritante, y los investigadores en todo el mundo decidieron que había que buscar mejor **software** (GLM, GAM, RF, Maxent...) que no sobrepredijera
- O sea, una solución técnica a un problema conceptual



Baronia brevicornis.

Tomada de Llorente et al, 1997

¿Cual es el problema conceptual?

- Una “sobrepredicción”
¿de qué?
- ¿Qué, exactamente,
“predecía” GARP (or
BIOCLIM, o DOMAIN)?

Obvio (?),
áreas de distribución...



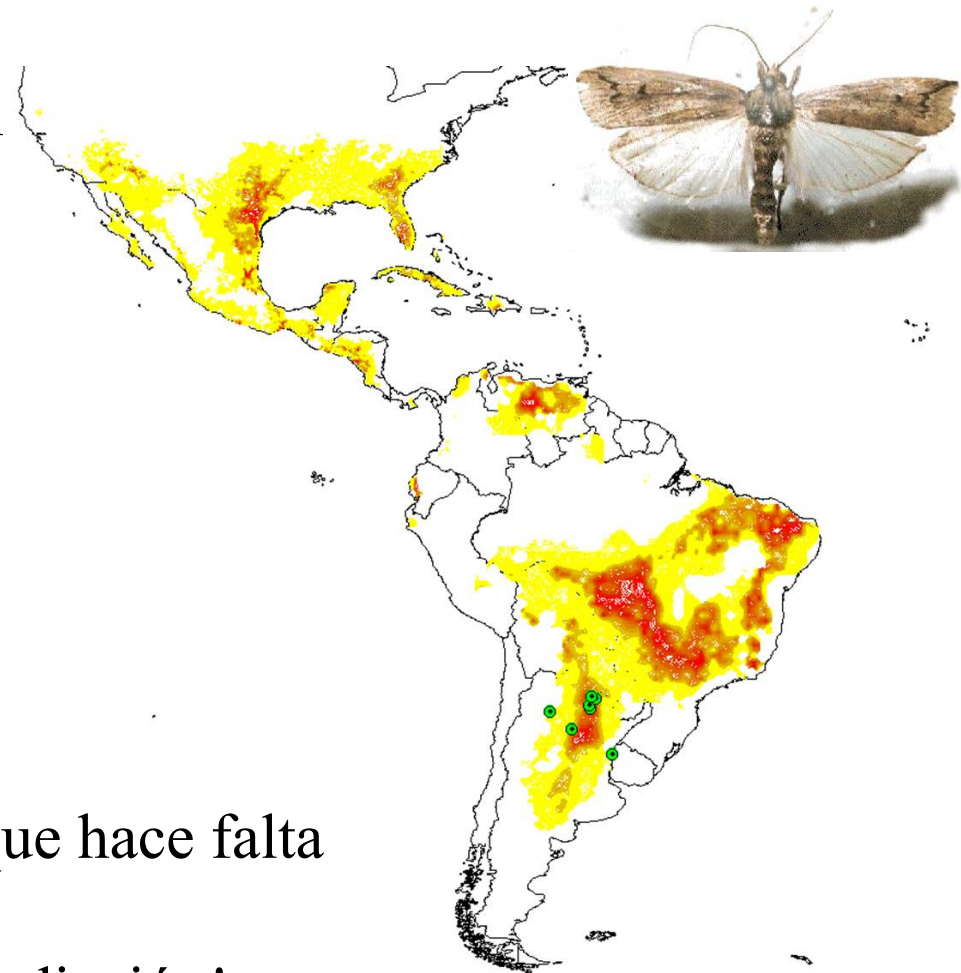
¿Seguro...? ¿ Qué es un área de distribución?

Mientras tanto, el Gobierno seguía haciendo preguntas urgentes...

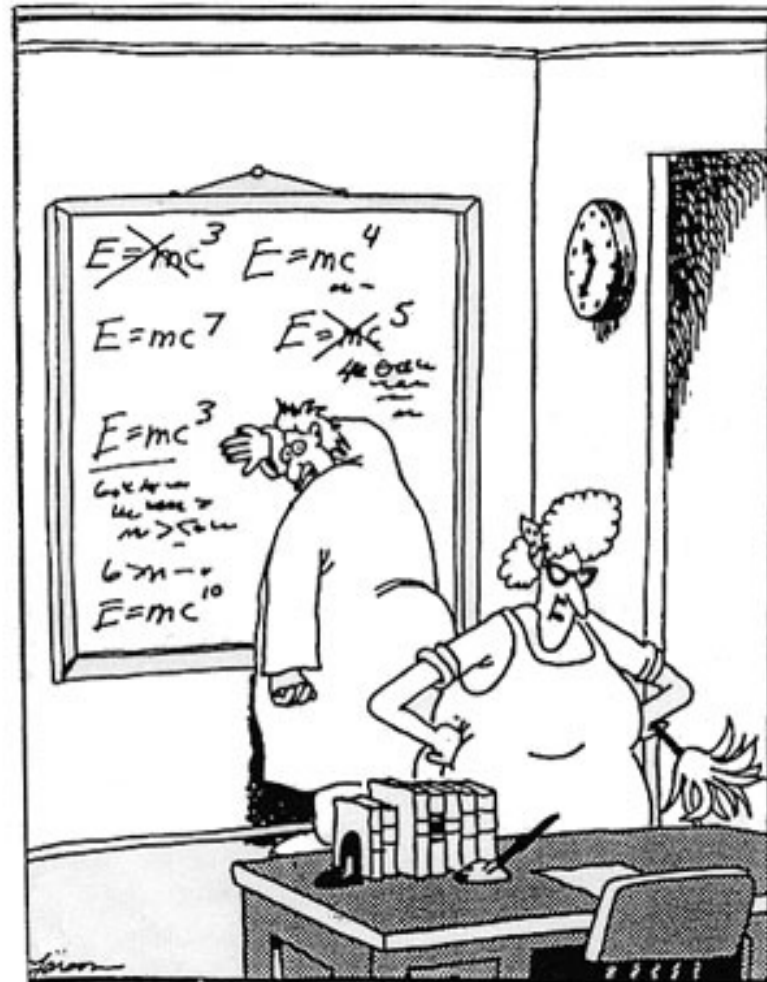
- ¿Se debía permitir la introducción al país de polinizadores (*Bombus terrestris*), o de especies milagrosas (*Cynara cardunculus* L.)?
- ¿Cual era el riesgo de posibles especies invasoras como *Cactoblastis cactorum*, *Tamarix racemosa* y otras?

Fíjense que para estas especies, lo que hace falta es la “sobrepredicción.”

¡La “sobre” es la parte útil de la predicción!



El momento de gritar ¡Eureka!



"Now that desk looks better. Everything's squared away, yessir, squaaaaared away."

¿Estariamos modelando nichos...!?

- Quiza estabamos modelando “nichos,” mas que distribuciones...
- Esta idea resulto tan sorprendente que dimos de brincos
- Solamente que... ¿qué que es un nicho?



Esto condujo a cosas muy interesantes

Necesitábamos entender y definir bien:

- que es una *distribución* y que es un *nicho*.
- Cómo se relacionan.
- Sin ignorar algoritmos y datos que sin duda capturaban una señal biológica muy real, y que ya tenían usuarios. No había que espantar a:
 - Los funcionarios, los diputados y senadores
 - Las ONGs
 - A la gran pesadilla, la Secretaria de Hacienda...

Pero ya vimos que:

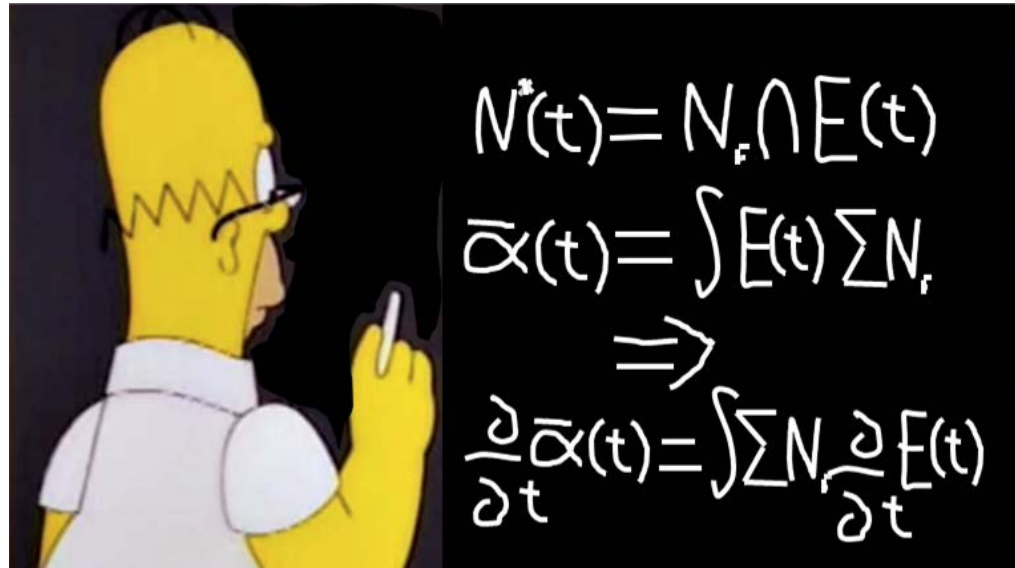
- Los nichos nadie sabe que son...
- Las distribuciones son fantasmas,
- Y al gobierno todo le urge

¿Estaba todo perdido?



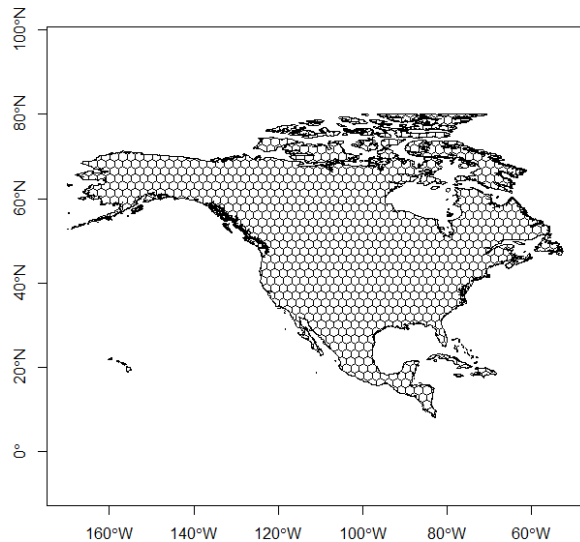
No. No estaba todo perdido

- Simplemente había que trabajarle a la teoría.
- En cuanto me salí del gobierno ya hubo tiempo para pensar con calma
- Trabajar para el gobierno tiene efectos muy perjudiciales para el pensamiento abstracto...



Mucho de la teoría se basa en la idea de la “Dualidad de Hutchinson”

Para cada punto en el espacio geográfico existe una combinación de valores ambientales única:



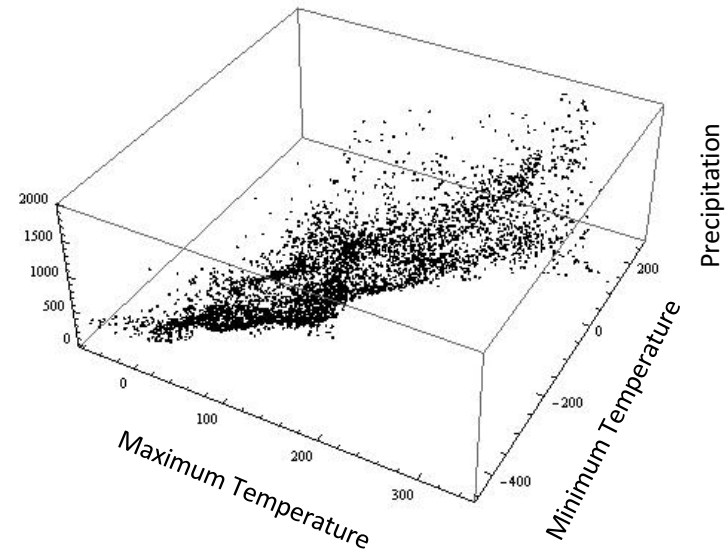
Espacio G

$$e(g) = \eta(g)$$



$$\eta^{-1}(e) = \{g\}$$

En general, $|\mathbf{G}| = |\mathbf{E}|$, pero en una versión continua de \mathbf{E} existen regions muy densas y otras muy diluidas

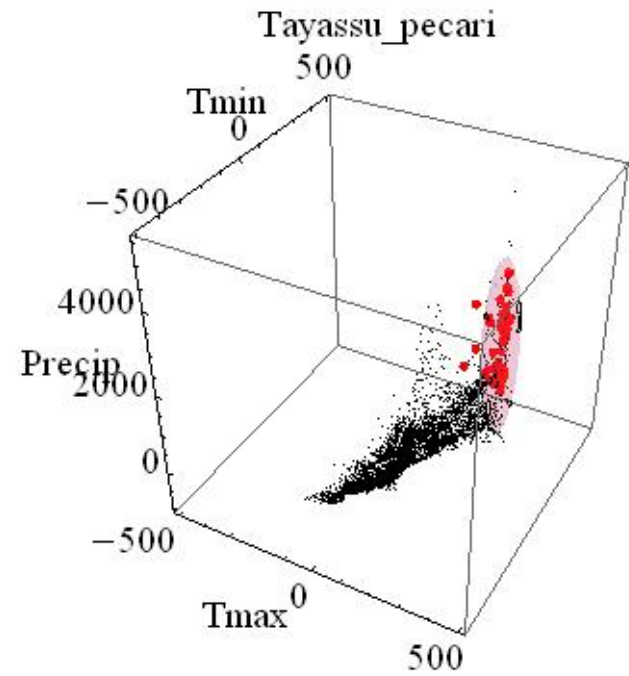
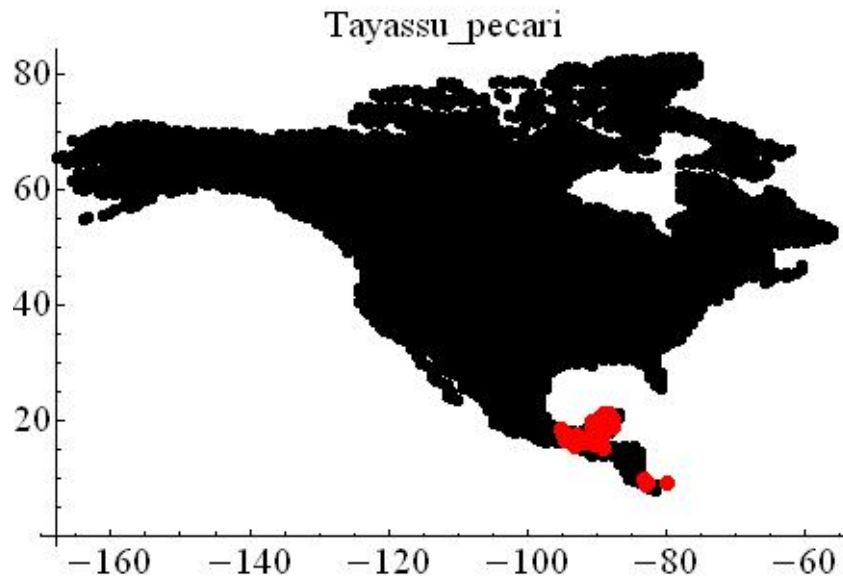


Espacio E

Por ejemplo...



Mario Antonio Pineda Malabrado/CONABIO



¿Como hago lo anterior dinámico?

- Mediante forzamiento climático

$$\frac{\partial \omega_j(t)}{\partial t} = \frac{1}{|\mathbf{G}|} \int_{\mathbf{G}} \mathbf{N}_j(\mathbf{v}^r) \frac{\partial \mathbf{E}(\mathbf{v}^r, t)}{\partial t} d\mathbf{v}^r$$

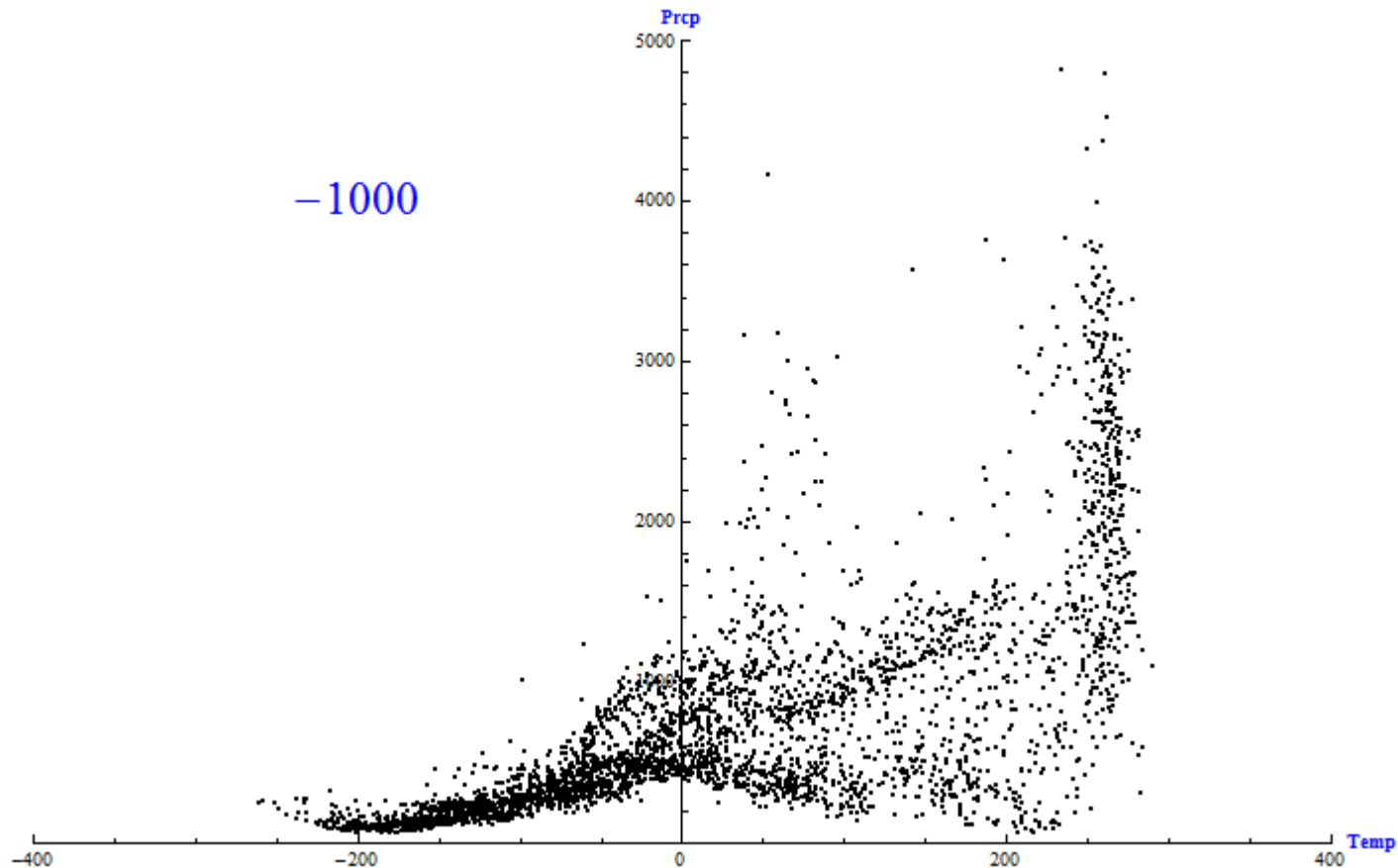
- O usando autómatas celulares de grano grueso

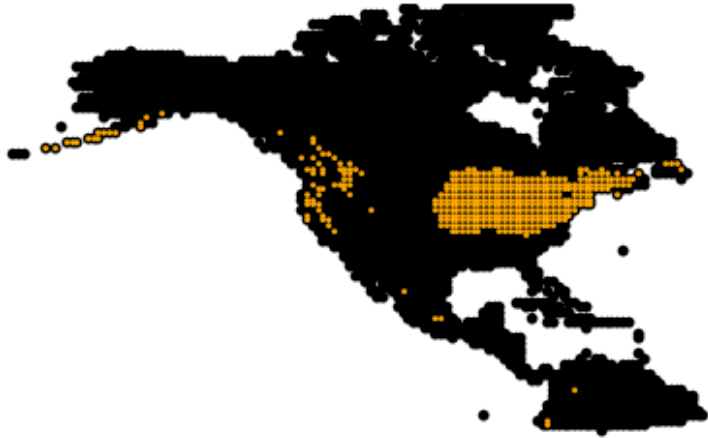
$$\mathbf{G}(t+1) \leftarrow \mathbf{S}(t) \times \mathbf{M} \times \mathbf{G}(t)$$

- O usando ecología de poblaciones básica

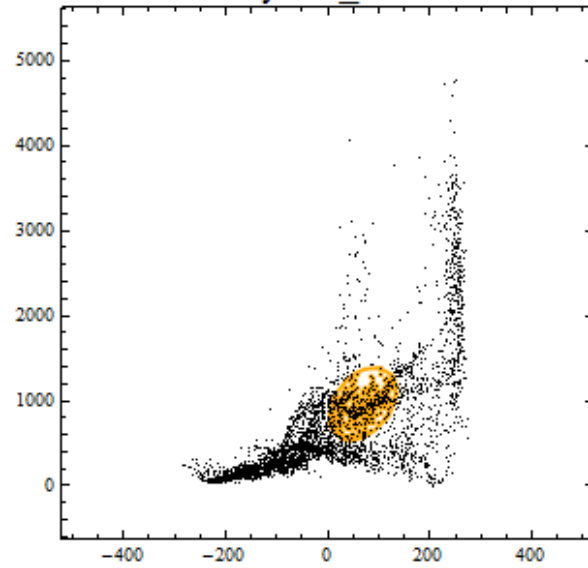
$$\frac{1}{x_{i,g}} \frac{dx_{i,g}}{dt} = r_{i,g}(\vec{e}_g; R_g^*) - \varphi_{i,g}(\vec{x}_g; \vec{e}_g, \vec{P}_{i,g}) + \psi(\vec{x}_i; \mathbf{M})$$

GCMs para Norte America, del presente a
hace 120,000 años AP (GCMs, cortesía del Centro
Hadley, UK. Dos variables.)





Condylura cristata



Condylura cristata
Smithsonian NMNH

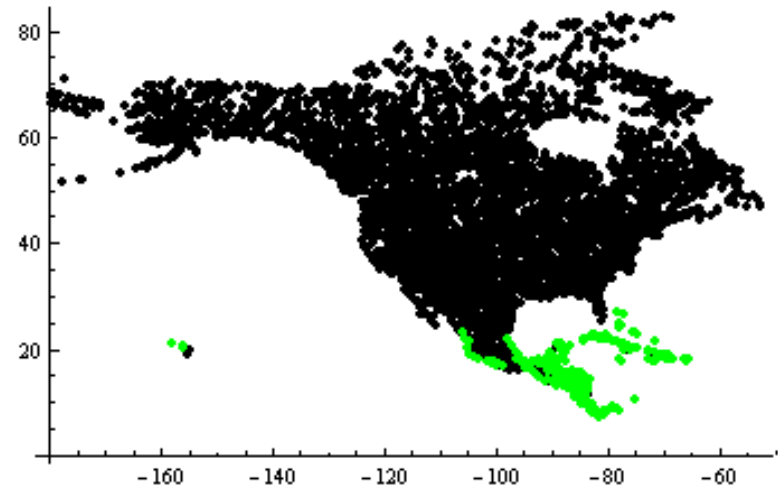
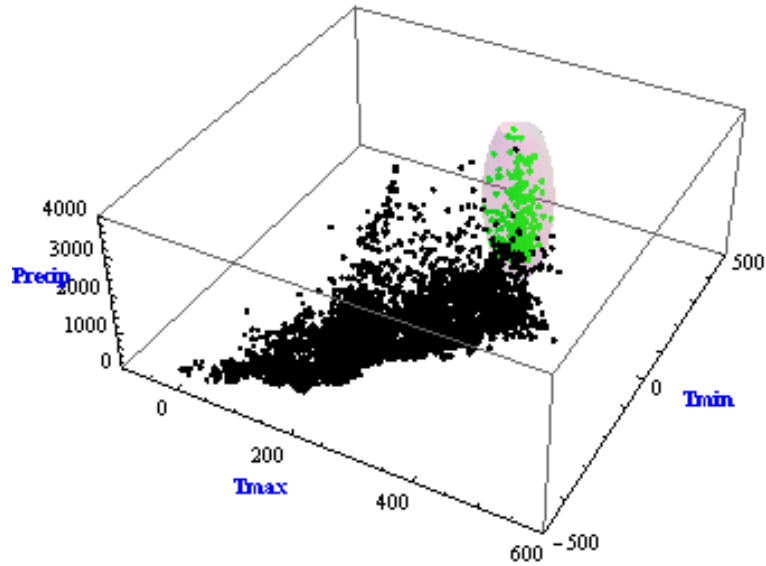


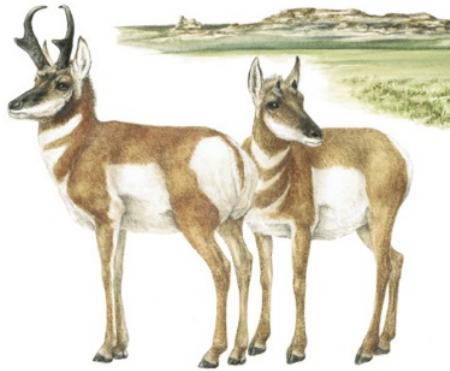


Fuente: INBio
© Derechos reservados

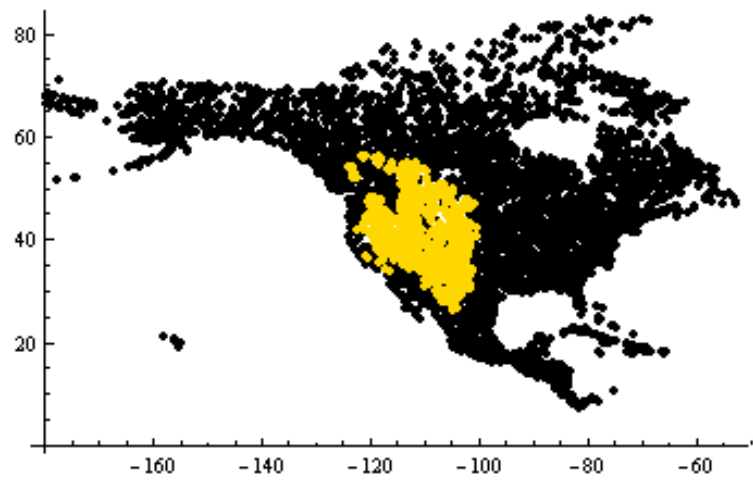
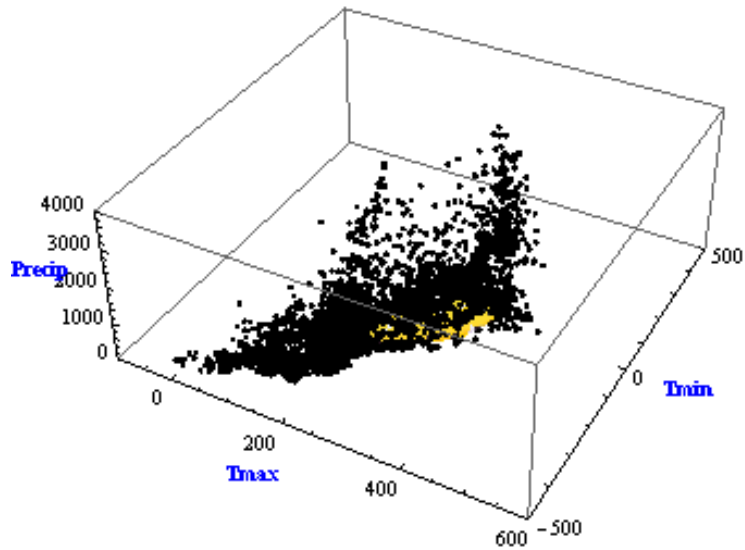
Alouatta palliata

INBio, Costa Rica






Antilocapra americana
Smithsonian NMNH



Y puede uno hacer “stacks” y calcular PAMs




 *I. gularis*



Fuente: INBio
© Derechos reservados

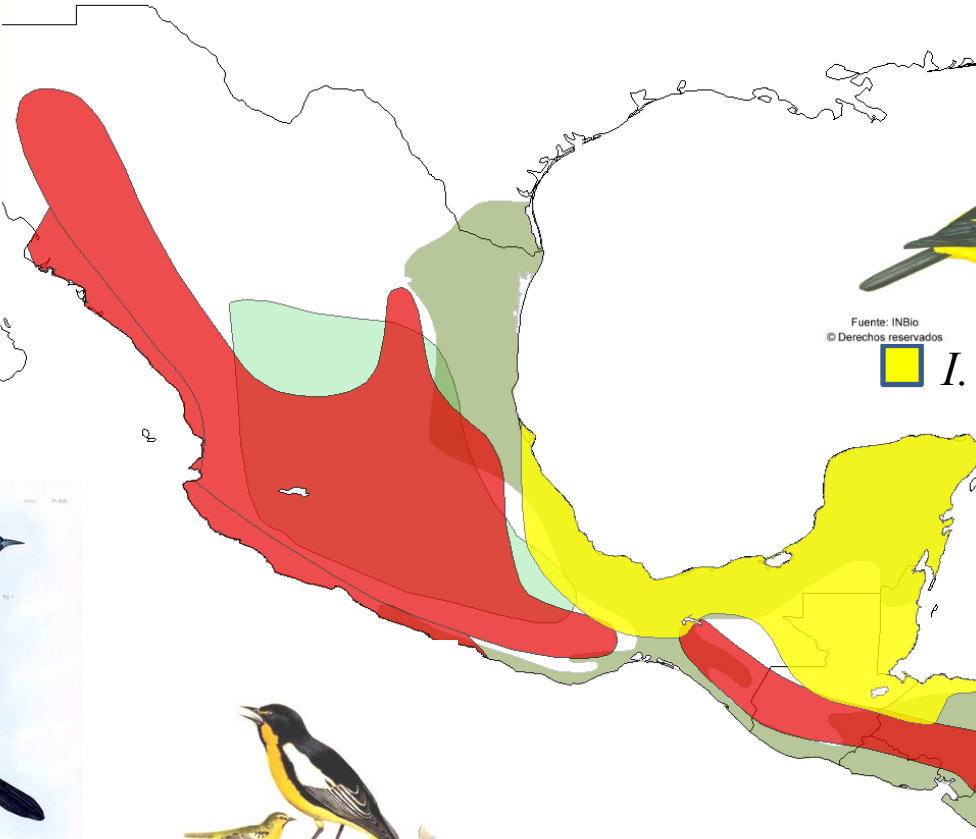
 *I. mesomelas*



 *I. abeillei*



 *I. waglerii*



PAMs = *Presence-absence matrices*

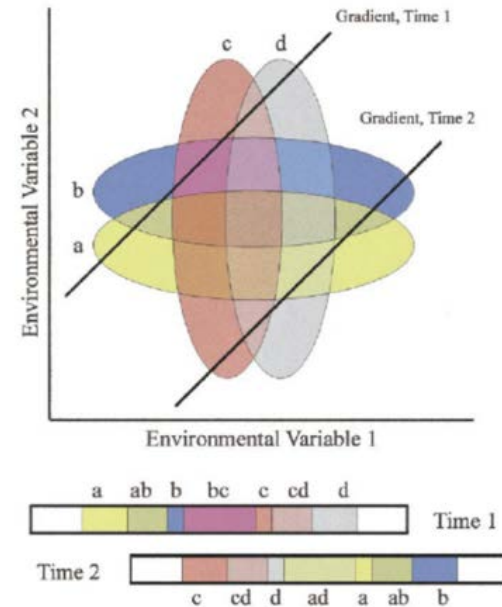
$$\begin{array}{c} \mathbf{X} \\ \left(\begin{array}{cccc} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right) \end{array} \begin{array}{c} \mathbf{X}^T \\ \left(\begin{array}{cccccccc} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \end{array} \right) \end{array} = \begin{array}{c} \left(\begin{array}{cccccccc} 4 & 2 & 2 & 3 & 3 & 3 & 1 & 1 \\ 2 & 2 & 0 & 2 & 1 & 1 & 0 & 1 \\ 2 & 0 & 2 & 1 & 2 & 2 & 1 & 0 \\ 3 & 2 & 1 & 3 & 2 & 2 & 1 & 1 \\ 3 & 1 & 2 & 2 & 3 & 2 & 1 & 1 \\ 3 & 1 & 2 & 2 & 2 & 3 & 1 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{array} \right) = \mathbf{A} \end{array}$$

$$\begin{array}{c} \mathbf{X}^T \\ \left(\begin{array}{cccccccc} 1 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 1 & 1 & 0 & 0 \end{array} \right) \end{array} \begin{array}{c} \mathbf{X} \\ \left(\begin{array}{cccc} 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{array} \right) \end{array} = \begin{array}{c} \left(\begin{array}{cccc} 4 & 3 & 3 & 2 \\ 3 & 6 & 3 & 4 \\ 3 & 3 & 5 & 2 \\ 2 & 4 & 2 & 4 \end{array} \right) = \mathbf{\Omega} \end{array}$$

“Forzamiento climático” en un stack:

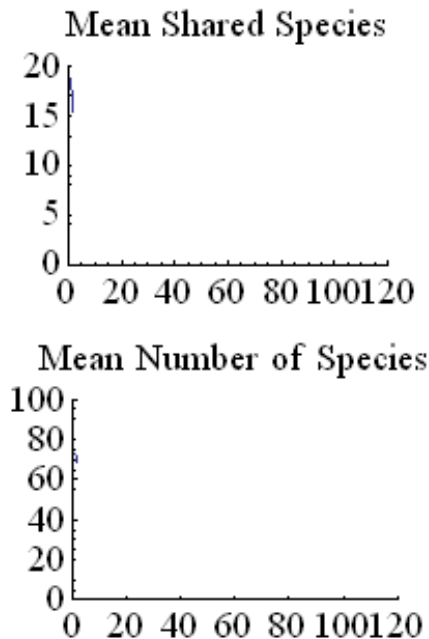
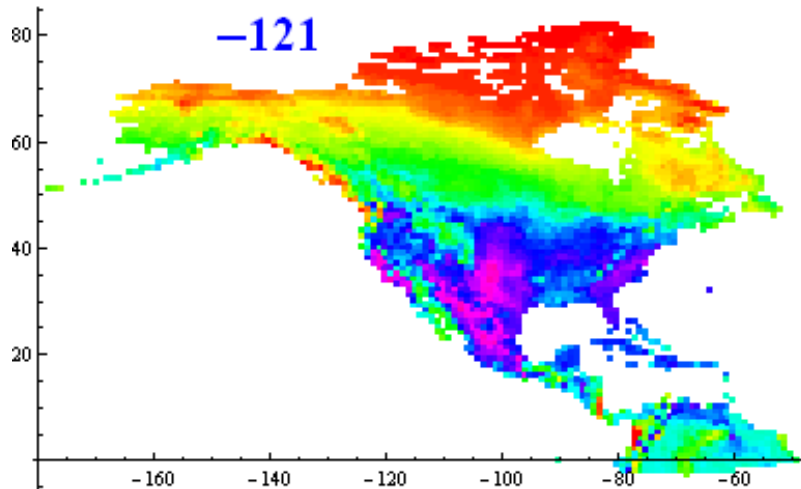
- La tasa de cambio en el número de especies en una combinación climática dada es proporcional al número de nichos fundamentales en esa combinación, multiplicada por la tasa de cambio del clima en esa combinación climática.
- Esta teoría le da al clima un papel fundamental para entender la biodiversidad

$$\frac{\partial \alpha(\mathbf{v}, t)}{\partial t} : \left(\sum_{j=1}^s N_j(\mathbf{v}) \right) \frac{\partial E(t, \mathbf{v})}{\partial t}$$

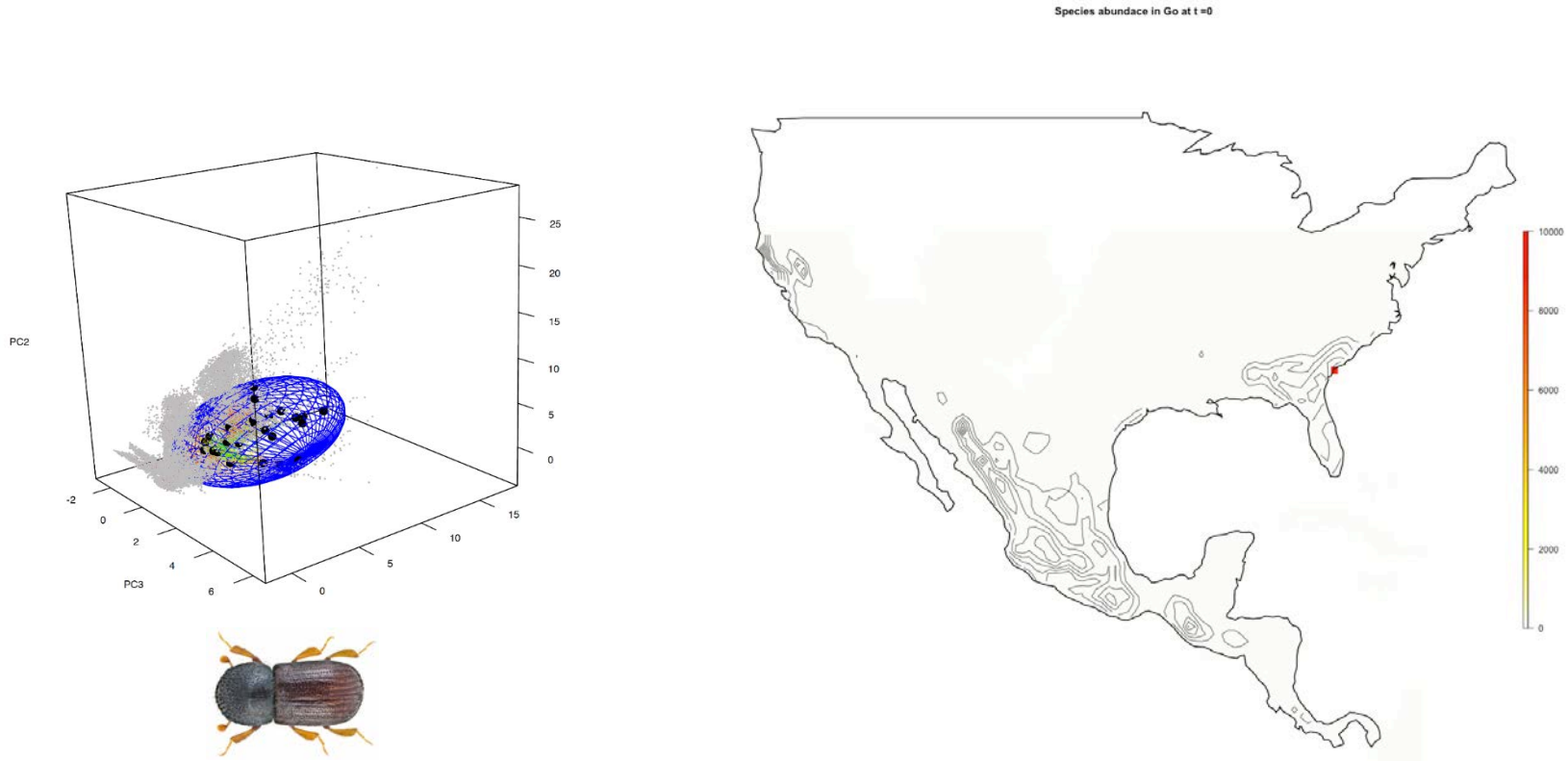


Numero potencial de especies entre el presente y el interglacial (120,000 AP)

Morado ~ 80 especies
Azul ~ 70 especies
Verde ~ 60 especies
Amarillo ~ 50 especies
Rojo ~ 40 especies



Tambien se puede hacer con ecología de poblaciones, pero eso requiere de los “otros” datos. “*Ambrosia beetle*”, tremenda amenaza para el aguacate.

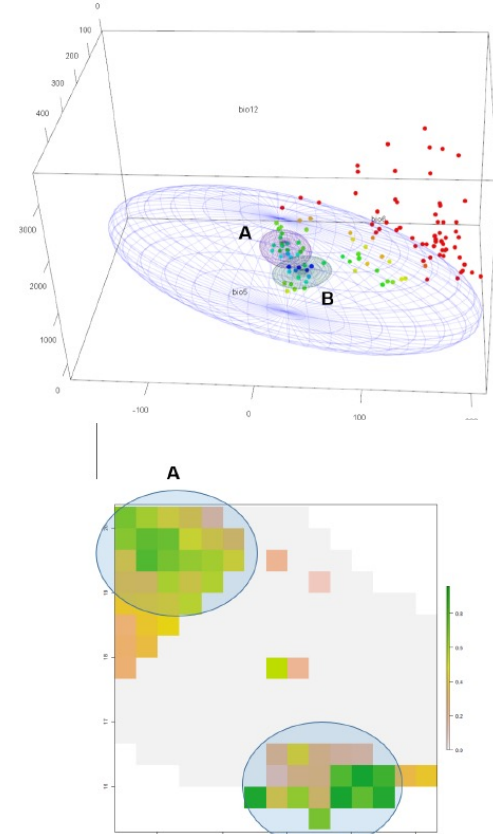


Xyleborus glabratus

<https://www.inaturalist.org/taxa/178605-Xyleborus>

Es muy difícil:

- Se necesitan datos demográficos
- Datos sobre umbrales de crecimiento (efectos Allee)
- Datos sobre distancias de dispersión
- Y sobre interacciones
- Numéricamente es endiablado
- Estamos trabajando sobre esto
- Luis Osorio, Manuel Falconi, Andres Lira (¿irá a tener la CONABIO un papel protagónico?)



On the Relationship Between Population Abundance and Niche Position

Falconi, Osorio, Soberon

January 12, 2016

Que significa todo lo anterior?

- “México” --unos mexicanos-- se plantearon la necesidad de contar con grandes cantidades de datos sobre biodiversidad, para conservarla y usarla
- Y de ahí se empezaron a responder preguntas prácticas de un usuario principal (el gobierno federal), el cual estaba satisfecho. No le importaba la teoría, pero a los académicos sí nos importa la teoría.
- El desarrollo teórico surgió como consecuencia de las preguntas de usuarios científicamente ingenuos.

Conclusiones I

- Información acumulada por siglos, en el presente aceleradamente (tecnología, ciencia ciudadana).
- Se podía organizar.
- Se puede hacer pública. Pública, pública de a de veras: toda la base de datos (hoy son 12 Gigabytes, mas la satelital, que son Terabytes)
- Se puede desarrollar bioinformática en serio

Conclusiones II

- ¿Que hizo falta? Tiempo y paciencia (nos equivocamos muchas veces).
- Respeto para platicar con gente distinta (diputados, empresarios, activistas), con prioridades muy diferentes a las nuestras.
- Compromiso por “la neta”. No lo que conviene a una causa, sino lo que dicen los mejores datos y el mejor análisis con el que cuenta uno. Bien difícil (transgénicos, conocimiento tradicional, pero eso son otras historias. Jorge Larson, Francisca Acevedo, Elleli Huerta, Jose Carlos Fernandez...)

Y finalmente

- Esta institución del gobierno, que cumple un cuarto de siglo, fue concebida por la imaginación visionaria de un científico mexicano, y dos de sus alumnos.
- Y fue construida, sobre la base de siglos de trabajo, por una comunidad de gente joven (casi todos), idealista, competente, comprometida. Es el producto de una comunidad y de una historia.
- Indispensable para México, ejemplo internacional, con gran futuro.

Mil Gracias

