



IX Taller GBIF de Modelización de Nichos Ecológicos (sesión 1)

Introducción y herramientas

Blas M. Benito

PRESENTACIÓN DEL TALLER

EL TALLER: SESIONES

- **Sesión 1**
 - Introducción a los modelos de distribución
 - El proceso de modelado
 - Aplicaciones de los modelos de distribución
 - El entorno de trabajo
 - Introducción práctica a R
 - Preparación de presencias y variables

EL TALLER: SESIONES

- **Sesión 2**
 - Métodos de modelización
 - Ensamblado, evaluación y umbrales de corte
 - Proyección en el espacio y el tiempo
- **Sesión 3**
 - Taller de trabajo individual o en grupo
- **Sesión 4**
 - Taller de trabajo individual o en grupo
 - Presentación de resultados
 - Discusión final y cierre del taller

INTRODUCCIÓN A LOS MODELOS DE NICHO ECOLÓGICO

¿QUÉ ES UN MODELO?

Descripción simplificada de un sistema físico real, en la que aparecen algunas de sus propiedades (Joly 1988).

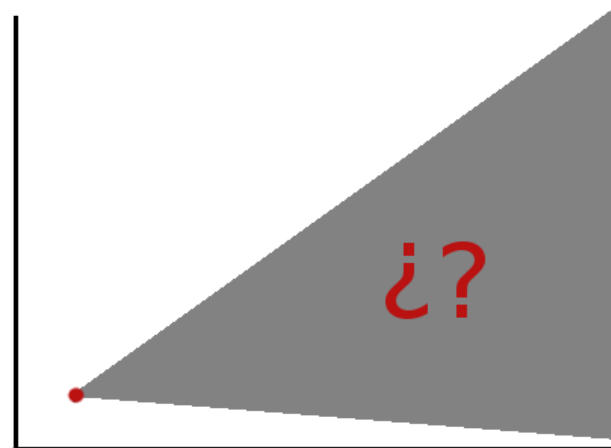
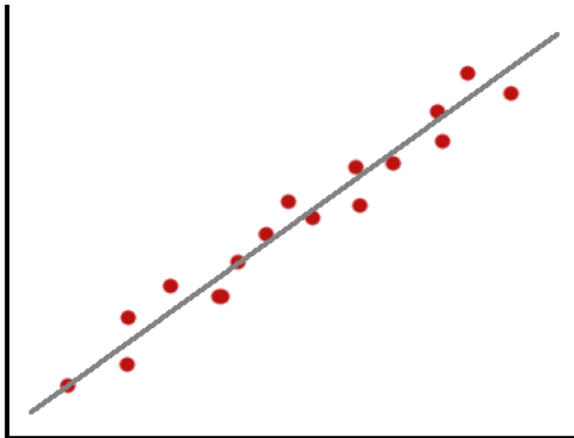
Los objetivos de un modelo son:

- Ofrecer una versión “manejable” del sistema
- Facilitar la comprensión de la estructura del sistema
- Generar y testar hipótesis
- Explorar escenarios alternativos (¿qué pasaría si...?)
- Predicción a corto plazo
- Apoyo a la toma de decisiones

HAY DISTINTOS TIPOS DE MODELOS

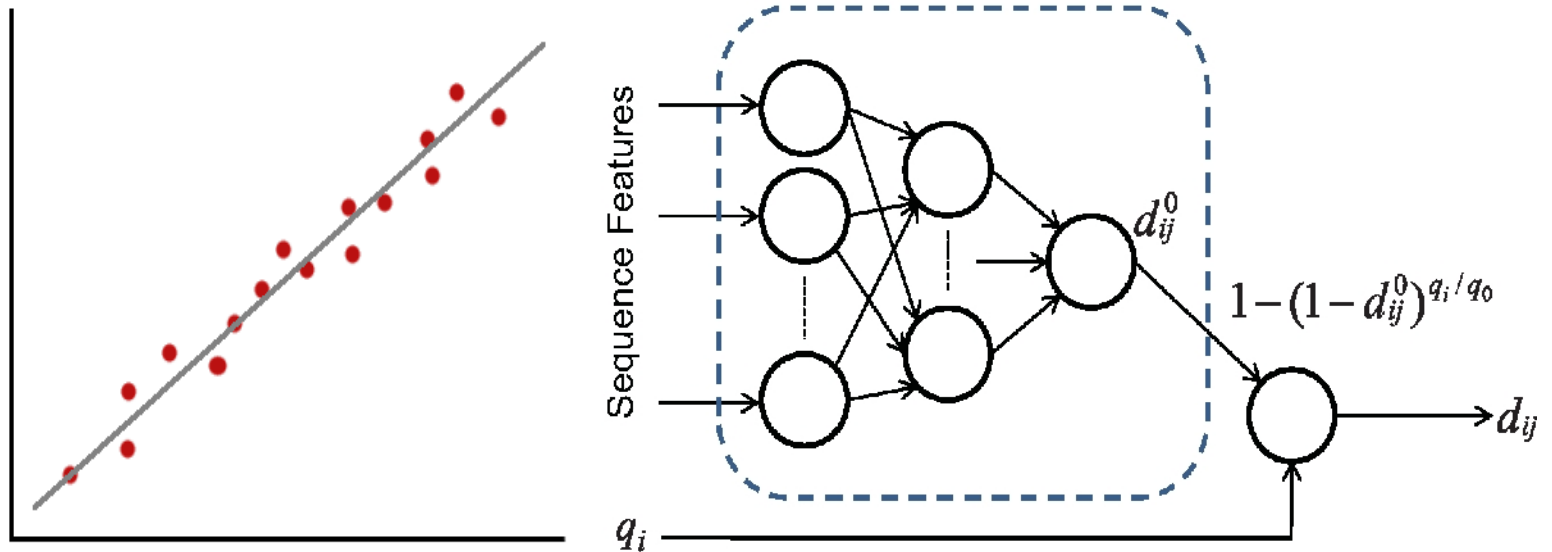
No se pueden clasificar en conjuntos cerrados, sino en gradientes.

- **Estadísticos (empíricos) vs. teóricos (mecanicistas)**



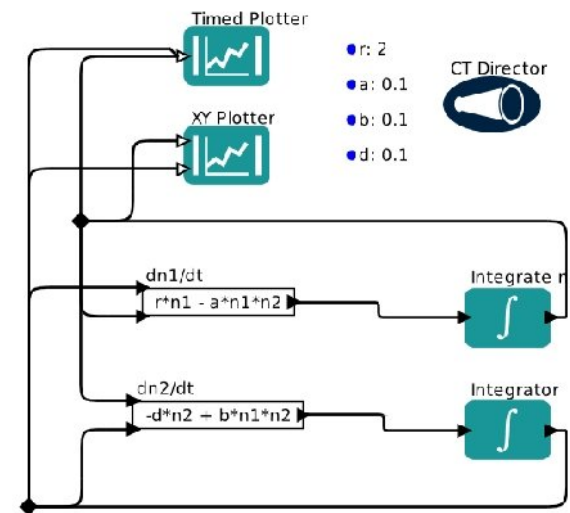
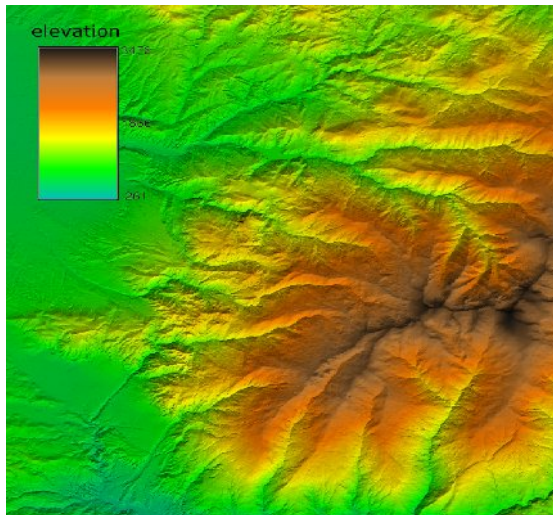
HAY DISTINTOS TIPOS DE MODELOS

- Estadísticos (empíricos) vs. teóricos (mecanicistas)
- **Simple vs. complejos**



HAY DISTINTOS TIPOS DE MODELOS

- Estadísticos (empíricos) vs. teóricos (mecanicistas)
- Simples vs. complejos
- **Estáticos vs. dinámicos**



PRESENCIA Y NICHO ECOLÓGICO

LA PRESENCIA...

presencia / ausencia



LA PRESENCIA...

presencia / ausencia



abundancia



LA PRESENCIA...

presencia / ausencia



abundancia



estructura hábitat



LA PRESENCIA...

presencia / ausencia



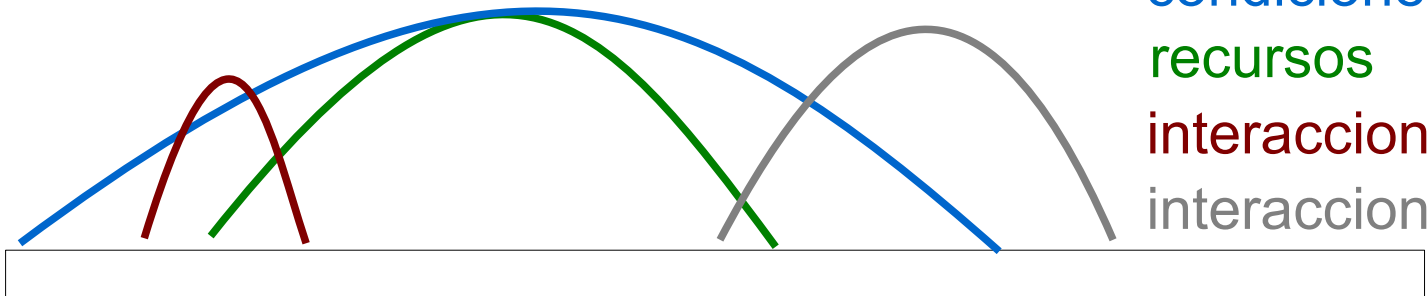
abundancia



estructura hábitat



procesos ecológicos



condiciones

recursos

interacciones -

interacciones +

LA PRESENCIA...

presencia / ausencia



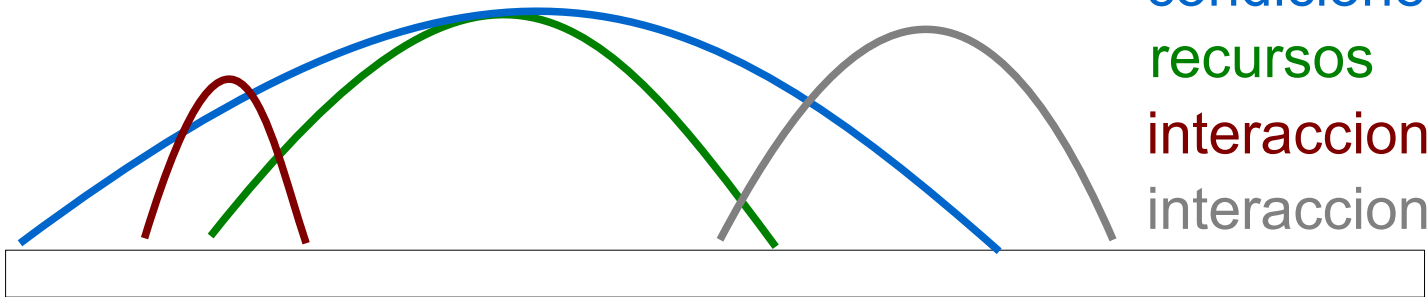
abundancia



estructura hábitat



procesos ecológicos



condiciones

recursos

interacciones -

interacciones +

tiempo - historia biogeográfica - perturbaciones

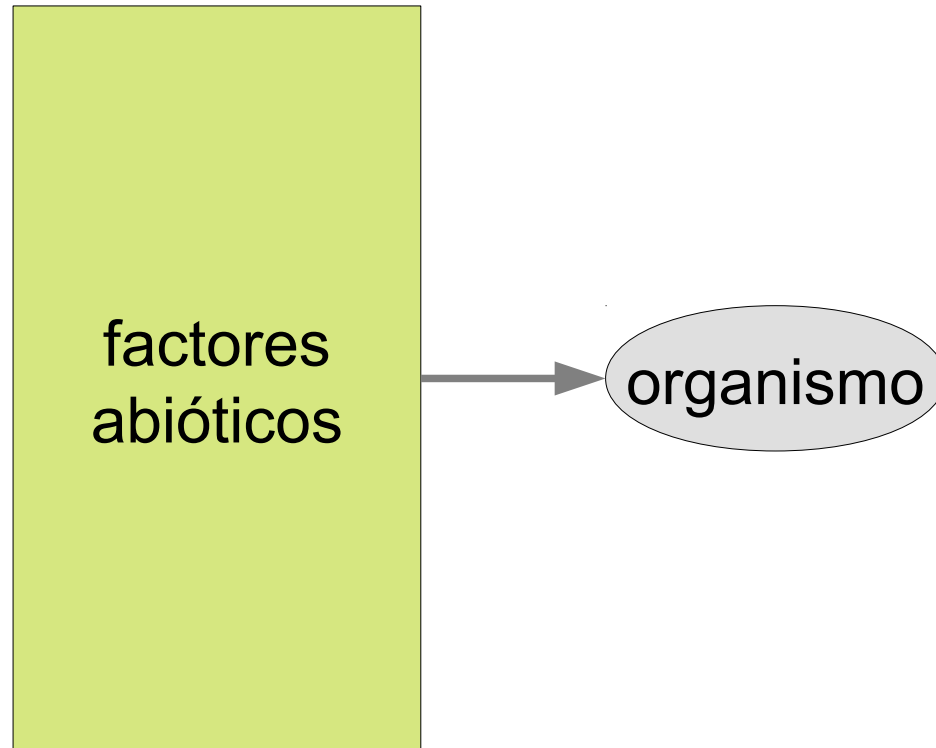
simplificación ↑

EL NICHO ECOLÓGICO

organismo

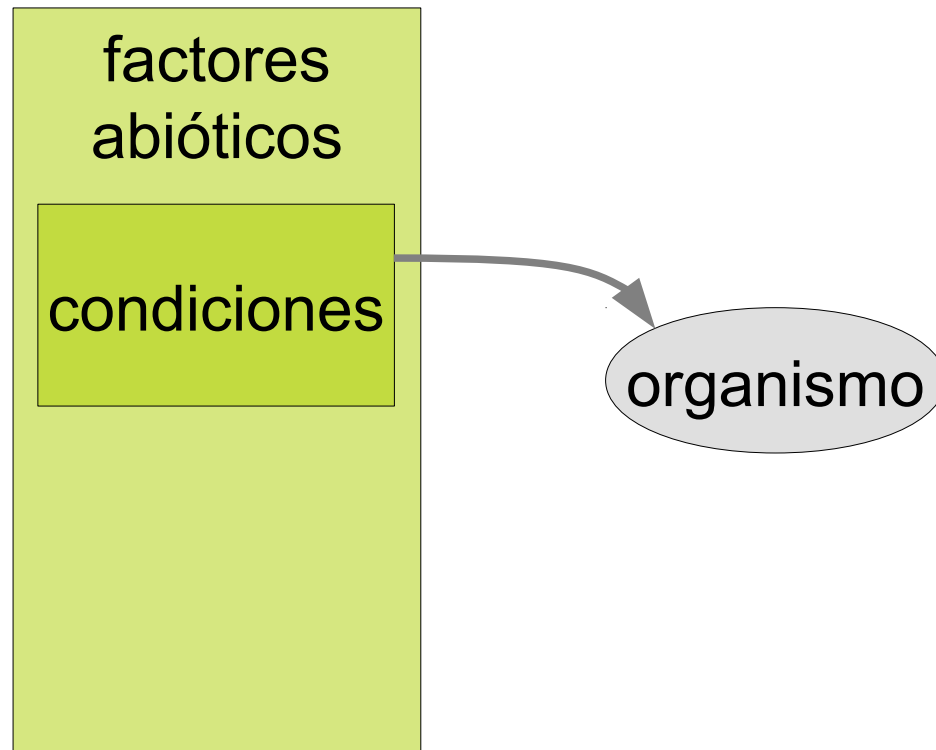
ver McInerny y Etienne 2012 J Biogeogr

EL NICHO ECOLÓGICO



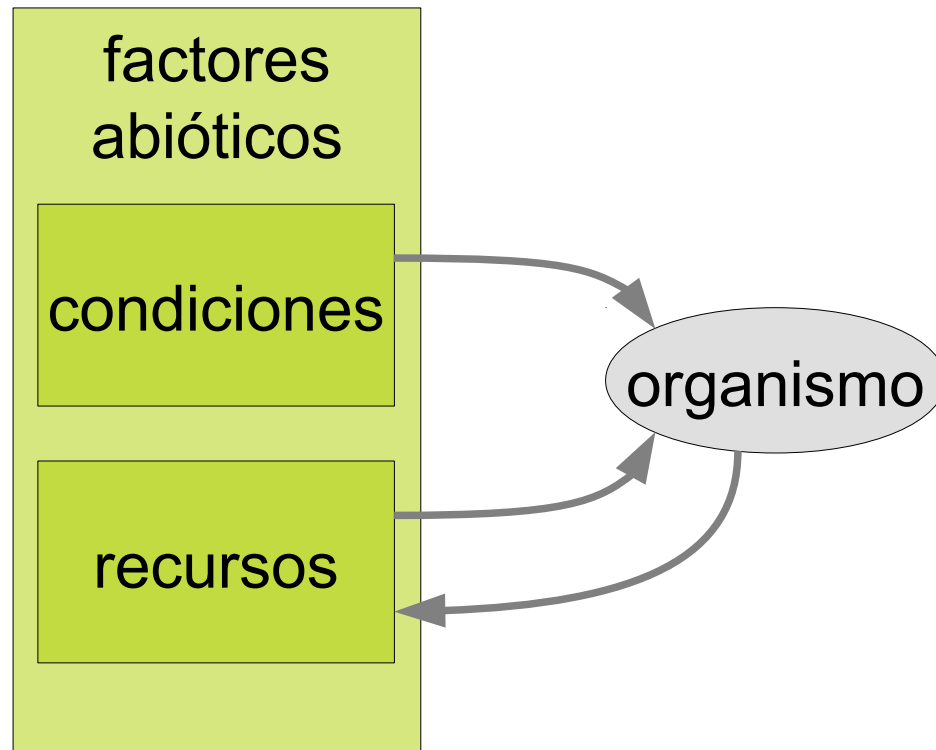
ver McInerny y Etienne 2012 J Biogeogr

EL NICHO ECOLÓGICO



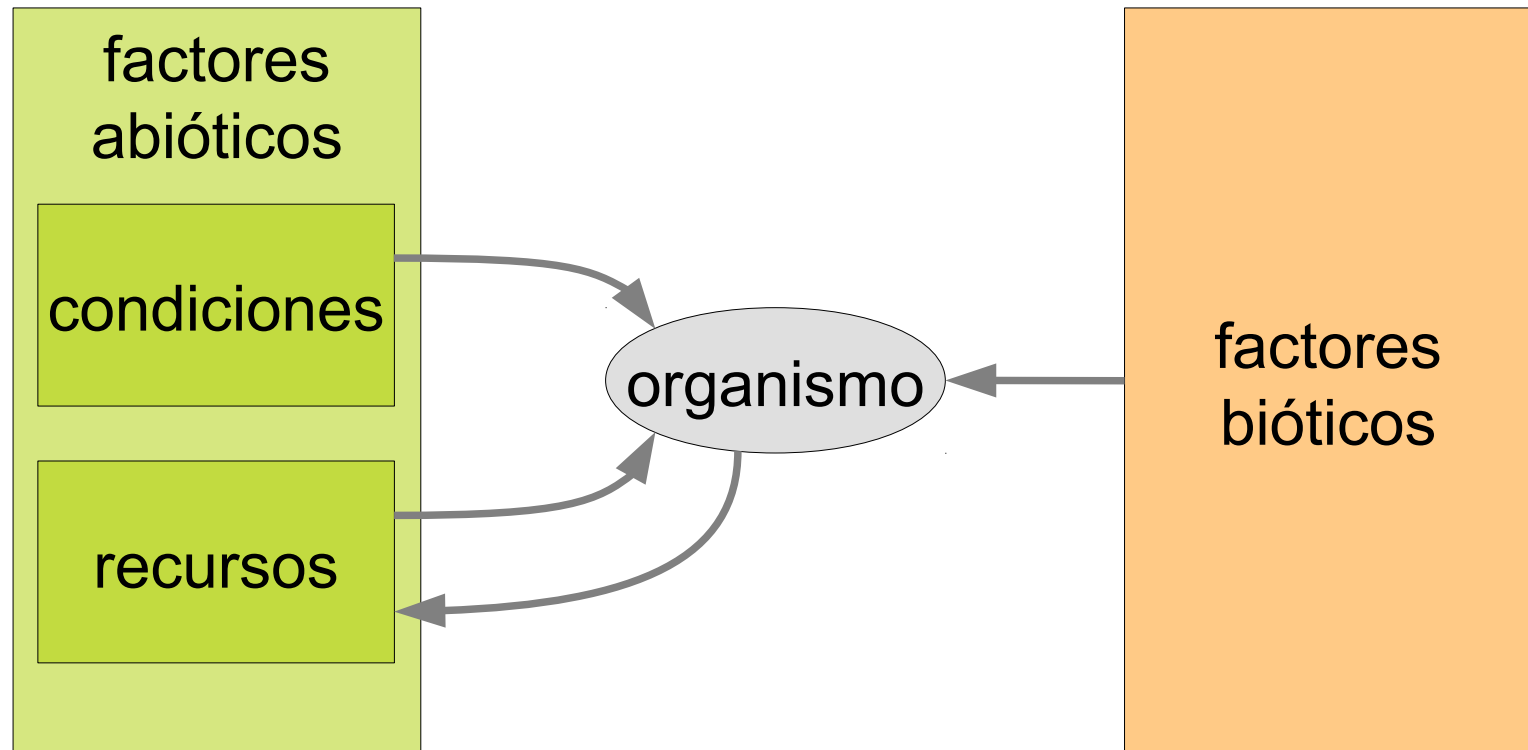
ver McInerny y Etienne 2012 J Biogeogr

EL NICHO ECOLÓGICO



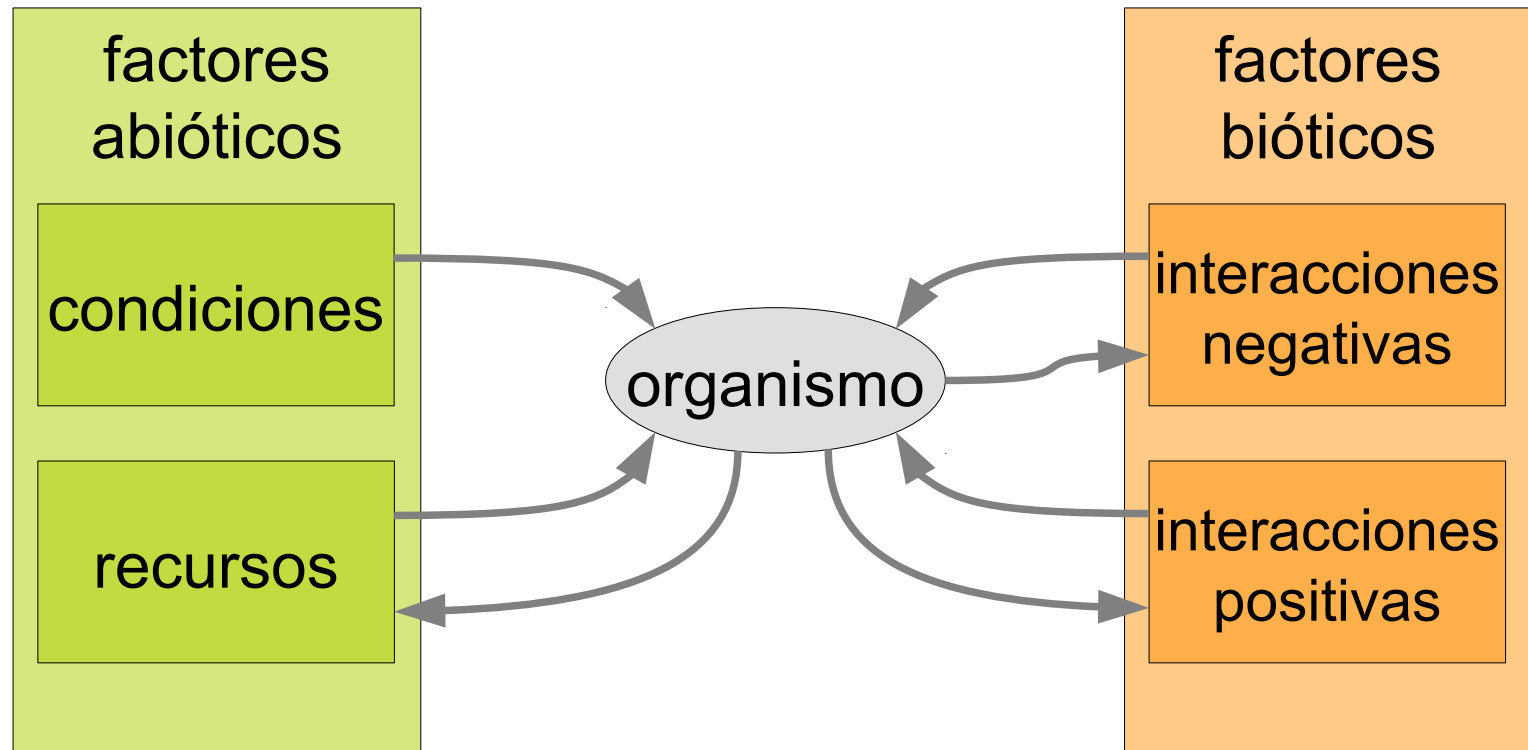
ver McInerny y Etienne 2012 J Biogeogr

EL NICHO ECOLÓGICO



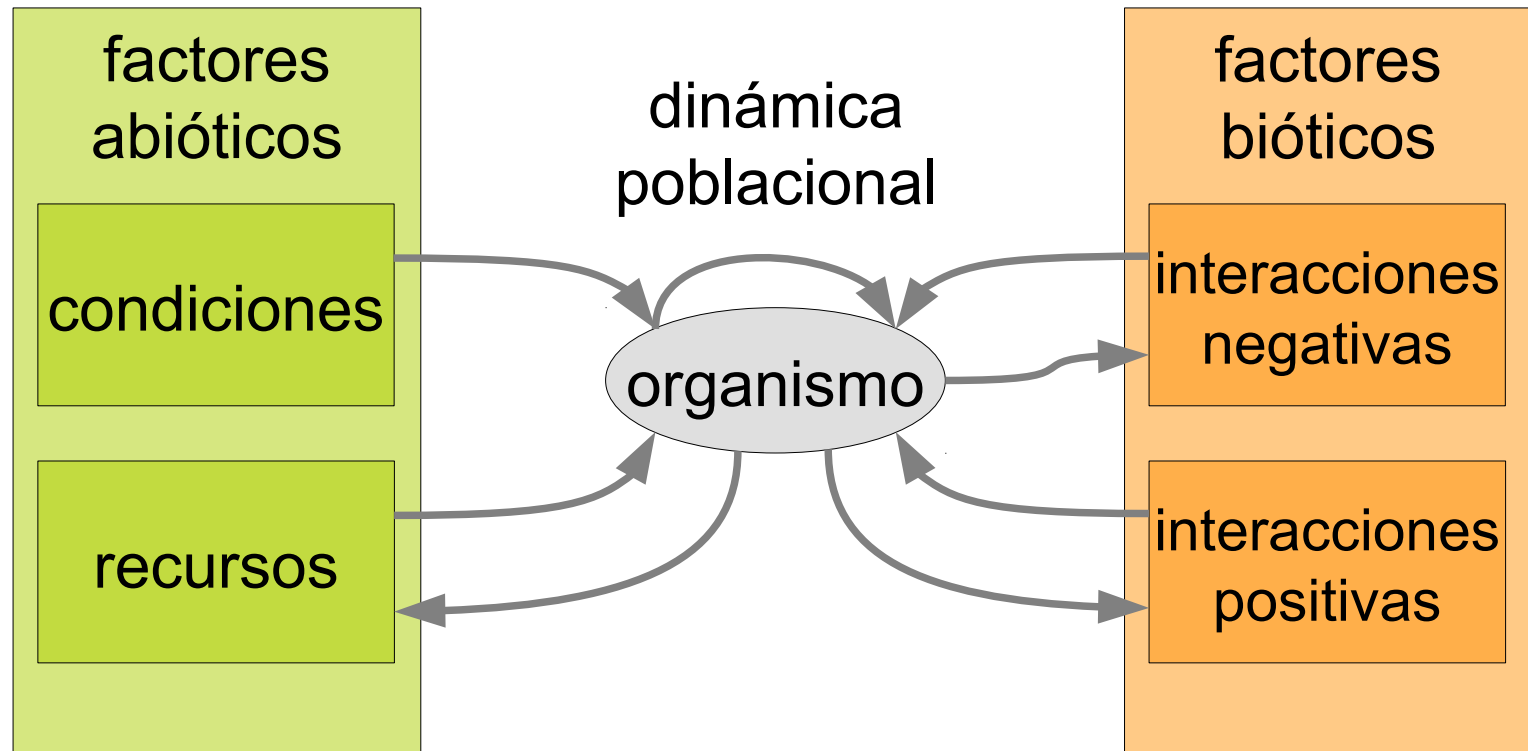
ver McInerny y Etienne 2012 J Biogeogr

EL NICHO ECOLÓGICO



ver McInerny y Etienne 2012 J Biogeogr

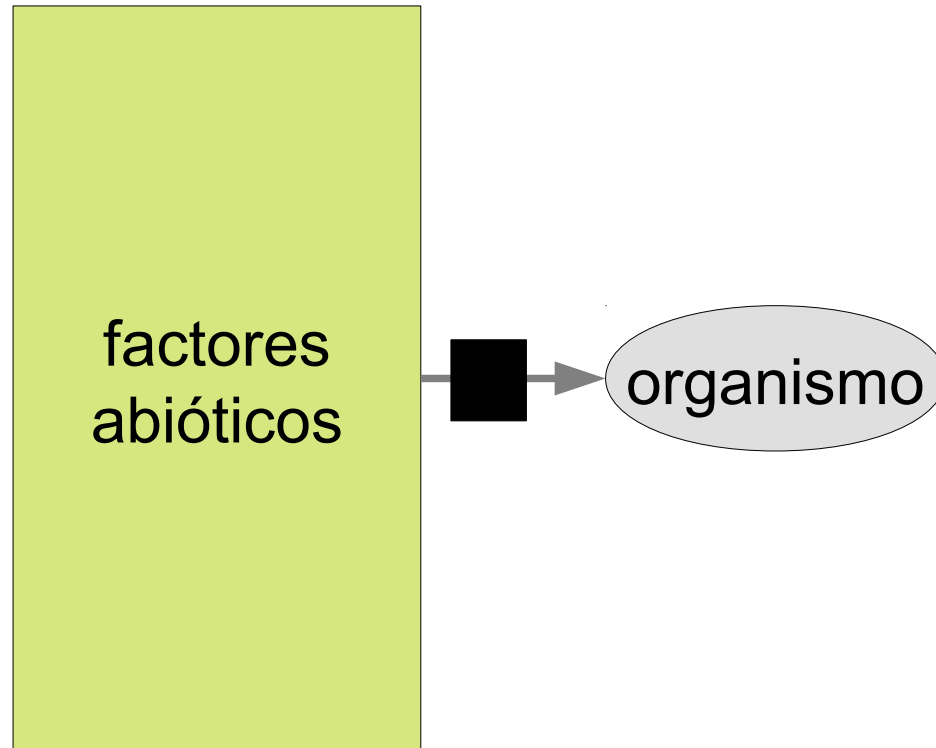
EL NICHO ECOLÓGICO



ver McInerny y Etienne 2012 J Biogeogr

NICHO ECOLÓGICO Y MODELOS DE DISTRIBUCIÓN

MODELOS CORRELATIVOS



■ = relación causal desconocida (blackbox)

MODELOS CORRELATIVOS

Frontiers in Zoology

Research

Open Access

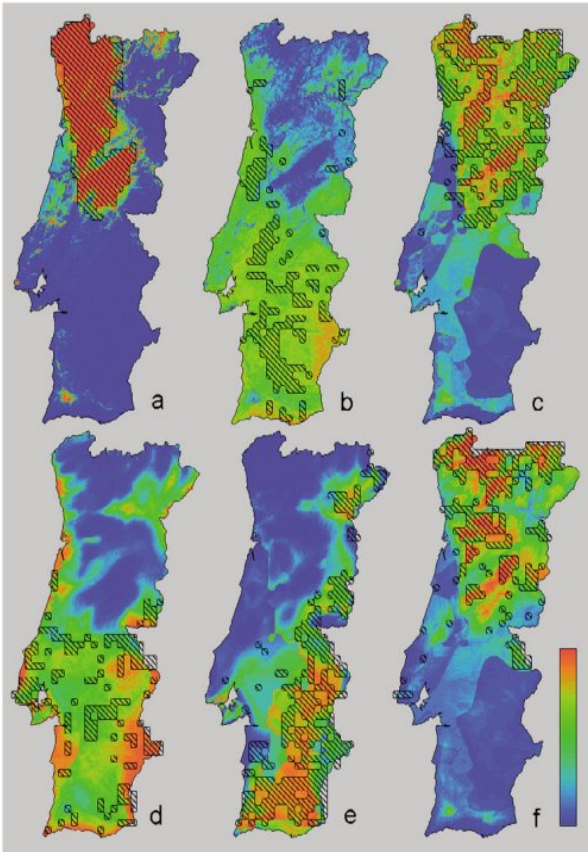
From descriptive to predictive distribution models: a working example with Iberian amphibians and reptiles

JW Arntzen*

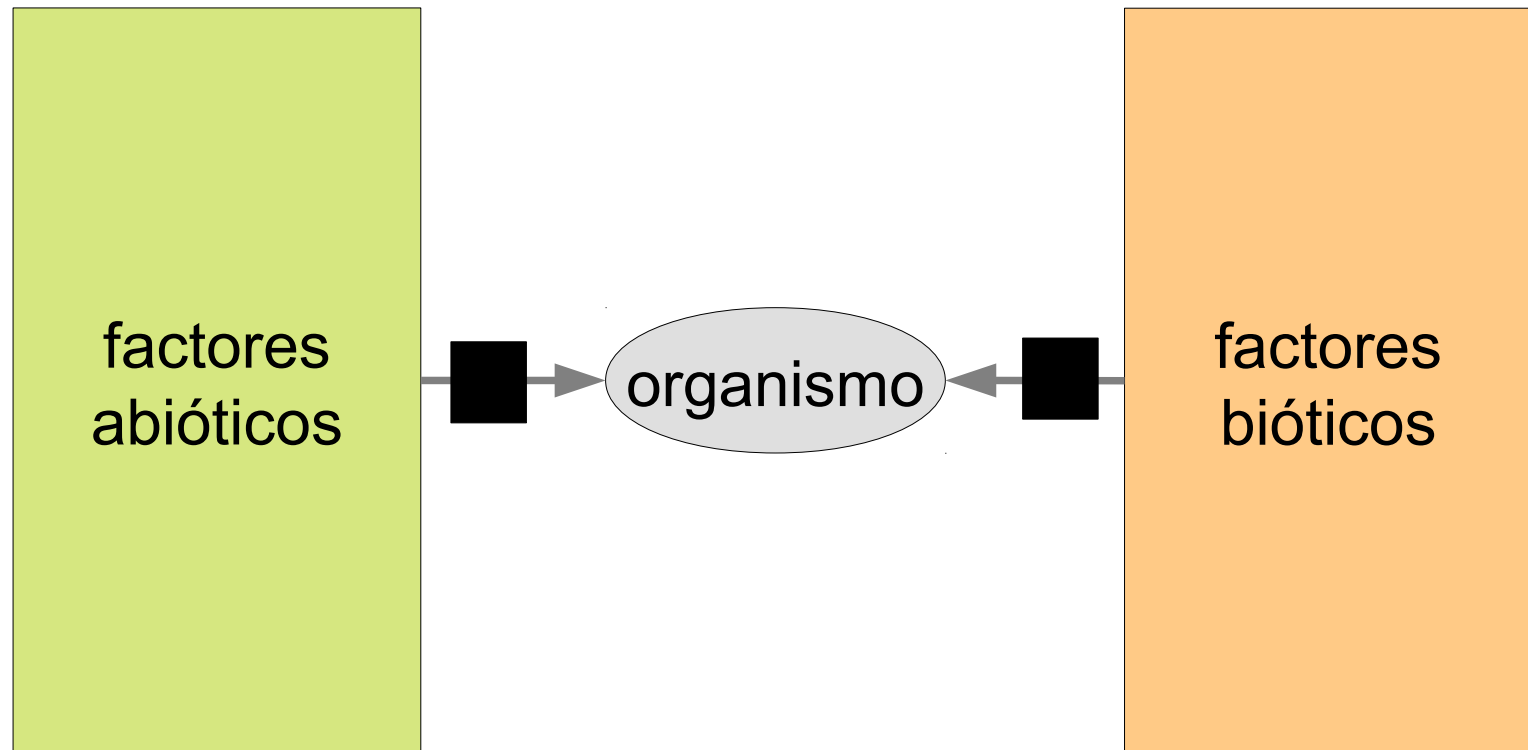
Figure 1

Descriptive distribution models for six amphibian species across Portugal.

Descriptive distribution models for six amphibian species across Portugal. Models are derived with stepwise logistic regression analysis of the dependent variable 'presence-absence of the target species' against 13 independent ecological variables (details see text and Table 1). The estimated probability of occurrence (g) ranges from 0 (blue) to 1 (red). Composite colours represent intermediate probabilities as in the colour scale bar. Species are: a) *Chioglossa lusitanica*, b) *Pleurodeles waltl*, c) *Triturus marmoratus*, d) *T. pygmaeus*, e) *Alytes cisternasii* and f) *A. obstetricans*. Recorded presences over the 10 × 10 km UTM-grid are shown by black shadings, after Godinho et al. [34].



MODELOS CORRELATIVOS CON INTERACCIONES



■ = relación causal desconocida (blackbox)

MODELOS CORRELATIVOS CON INTERACCIONES

Improving species distribution models using biotic interactions: a case study of parasites, pollinators and plants

Tereza Cristina Giannini, Daniel S. Chapman, Antonio Mauro Saraiva, Isabel Alves-dos-Santos
and Jacobus C. Biesmeijer

Ecography 36: 649–656, 2013

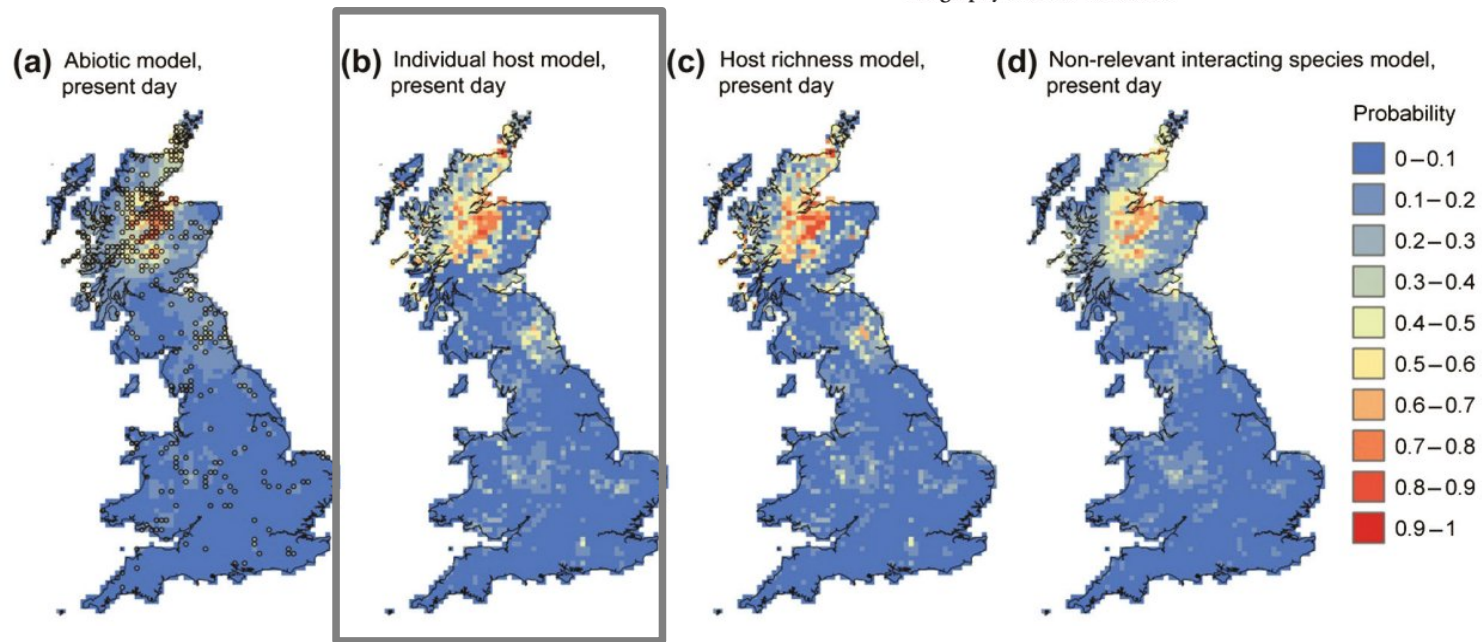
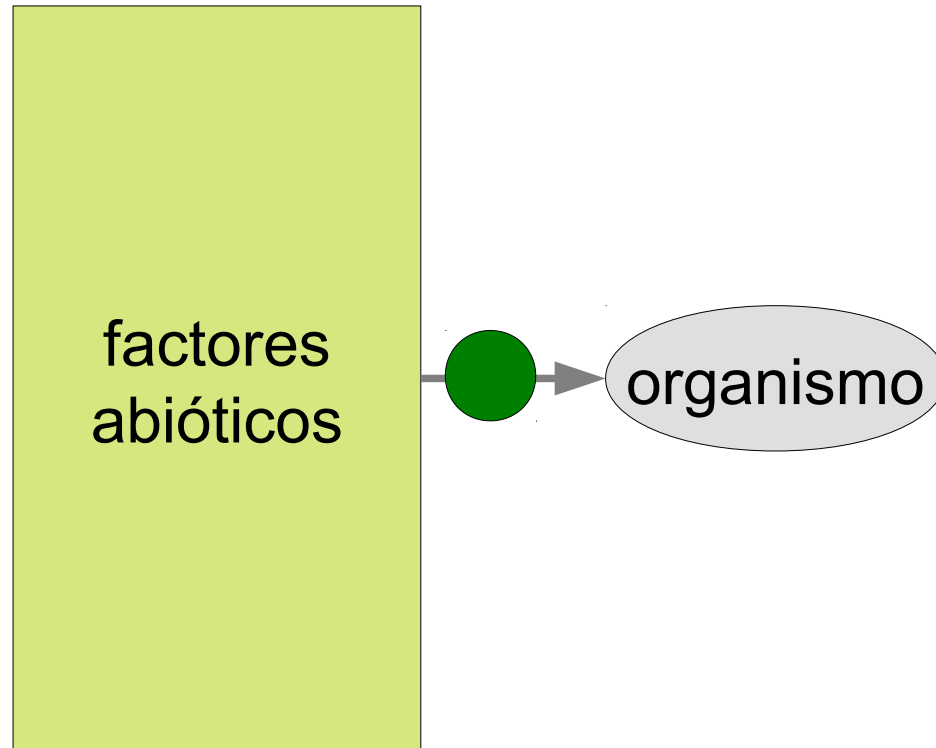


Figure 2. Projections of *Bombus bohemicus* occurrence areas. Top: present climatic conditions using (a) abiotic information only; (b) the most widespread host *B. lucorum*; (c) richness of its three host species (*B. lucorum*, *B. magnus* and *B. cryptarum*); (d) non-relevant interacting species (*B. terrestris*).

MODELOS ECO-FISIOLÓGICOS



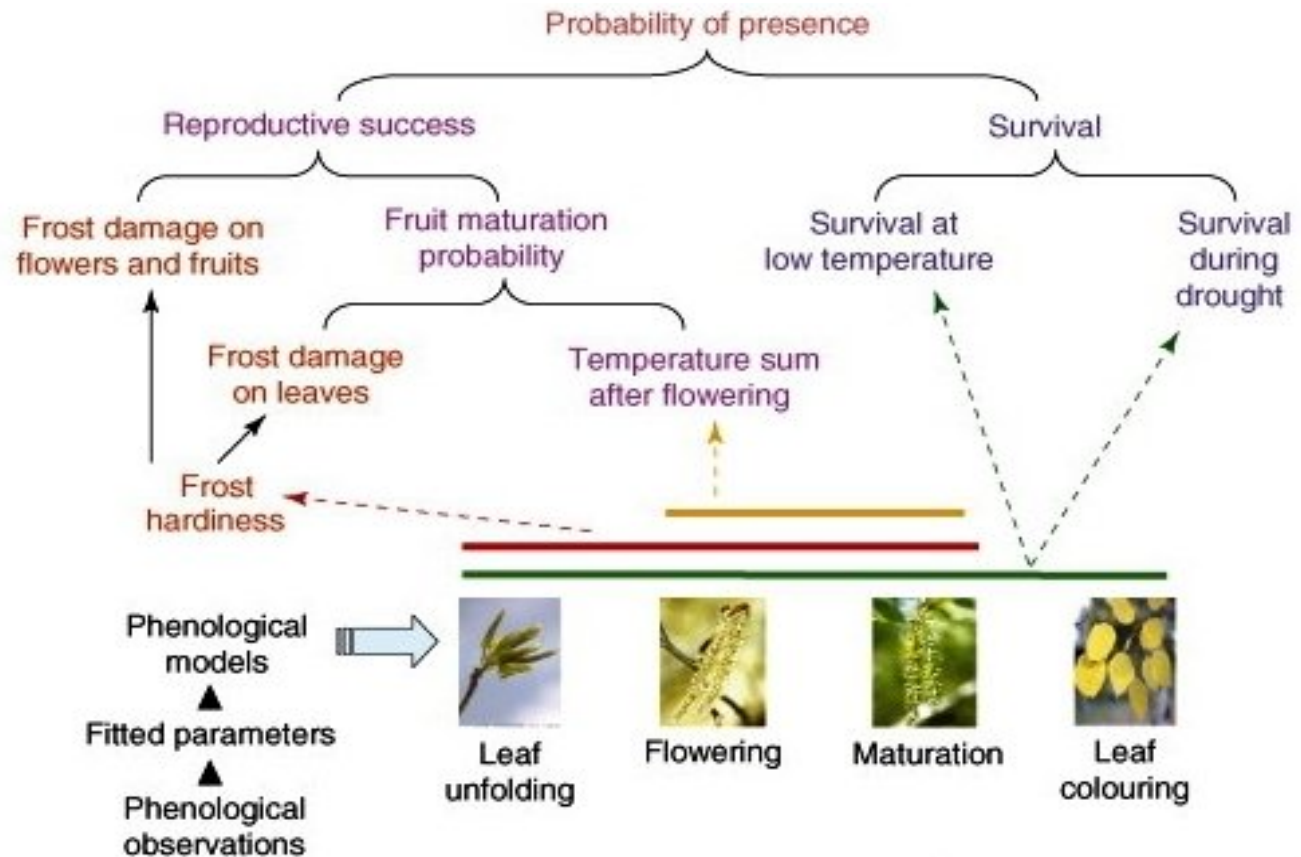
 = mecanismo ecológico explícito

MODELOS ECO-FISIOLÓGICOS

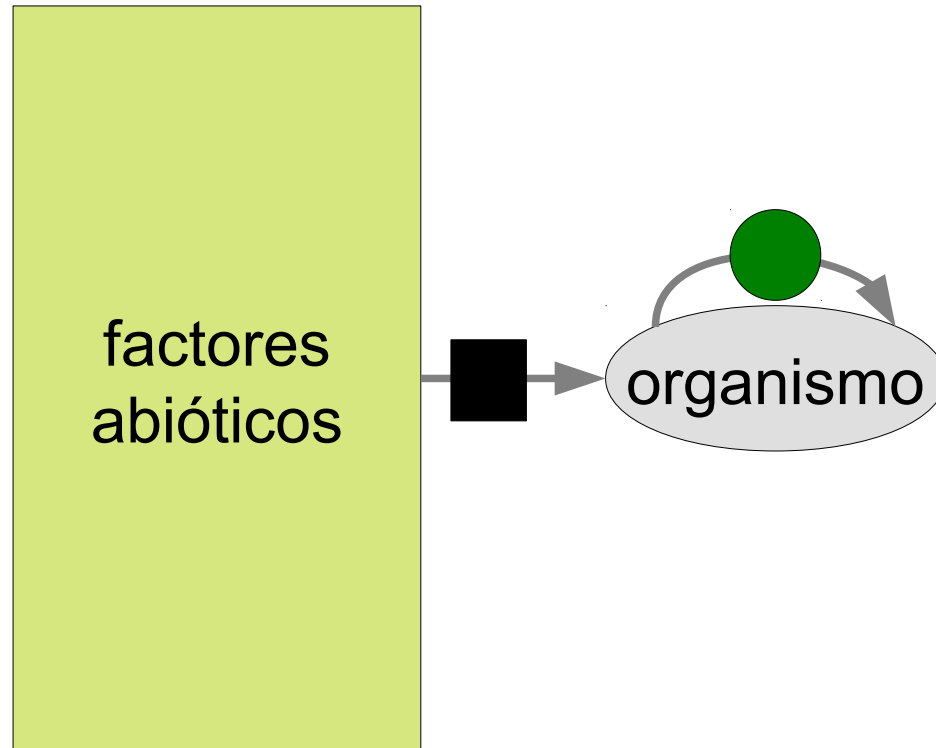


Shifting plant phenology in response to global change

Elsa E. Cleland¹, Isabelle Chuine², Annette Menzel³, Harold A. Mooney⁴ and Mark D. Schwartz⁵



MODELOS HÍBRIDOS



● = mecanismo ecológico explícito

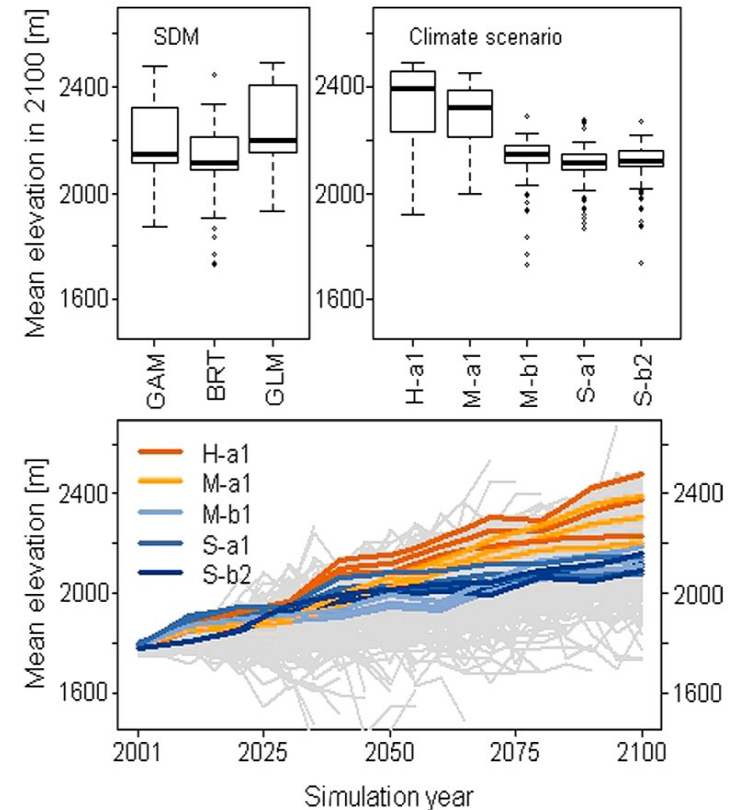
MODELOS HÍBRIDOS

Uncertainty in predictions of range dynamics: black grouse climbing the Swiss Alps

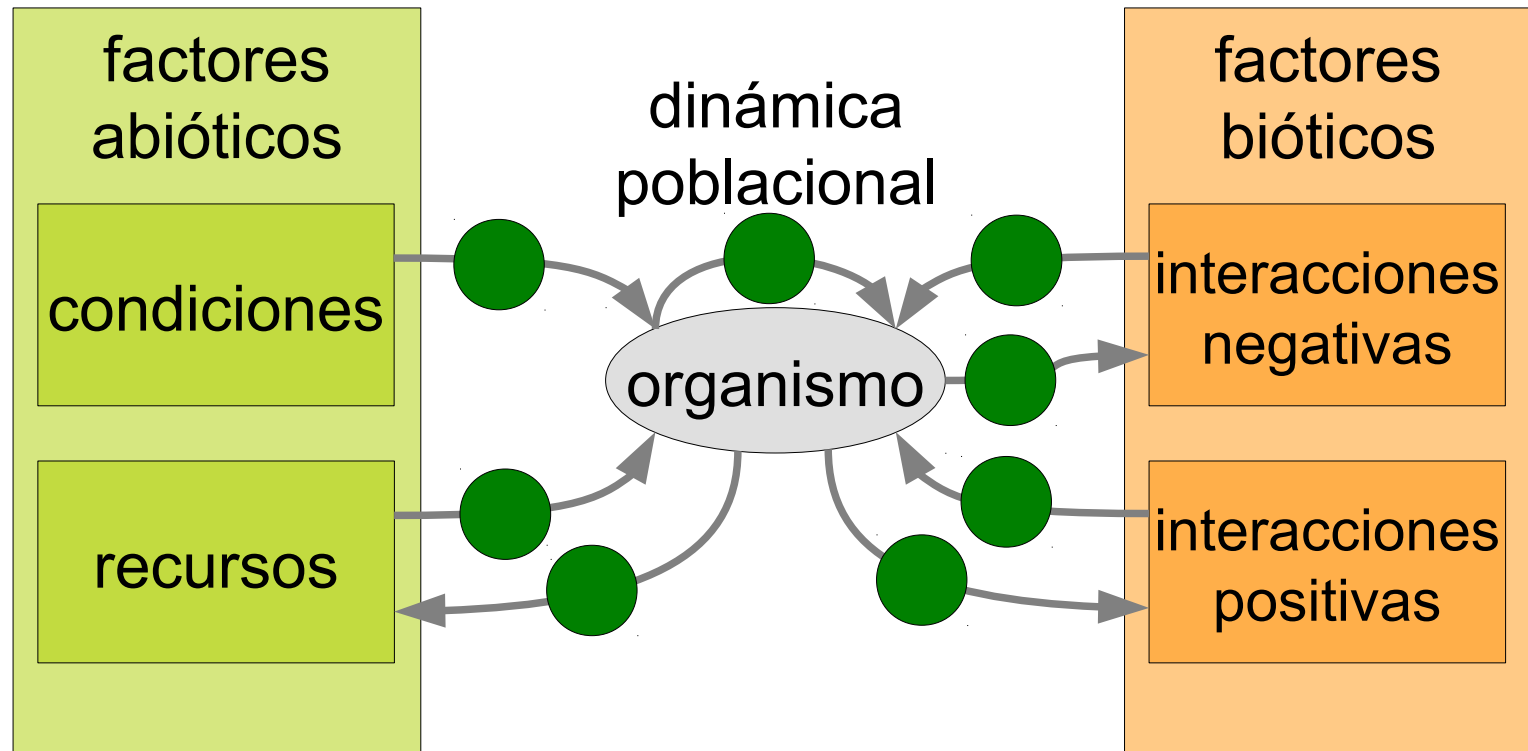
Damaris Zurell, Volker Grimm, Eva Rossmannith, Niklaus Zbinden, Niklaus E. Zimmermann and Boris Schröder

Ecography 34: 001–014, 2011

Figure 2. Mean elevation occupied by black grouse for scenarios of climate change. Bottom: grey lines show mean elevations across all simulations, coloured lines those for default IBM parameterisation (cf. Table 2) across different SDMs and climate scenarios. Top: box-plots depict variation of mean elevations predicted for the end of 21st century (2100) and for different SDMs and climate scenarios.

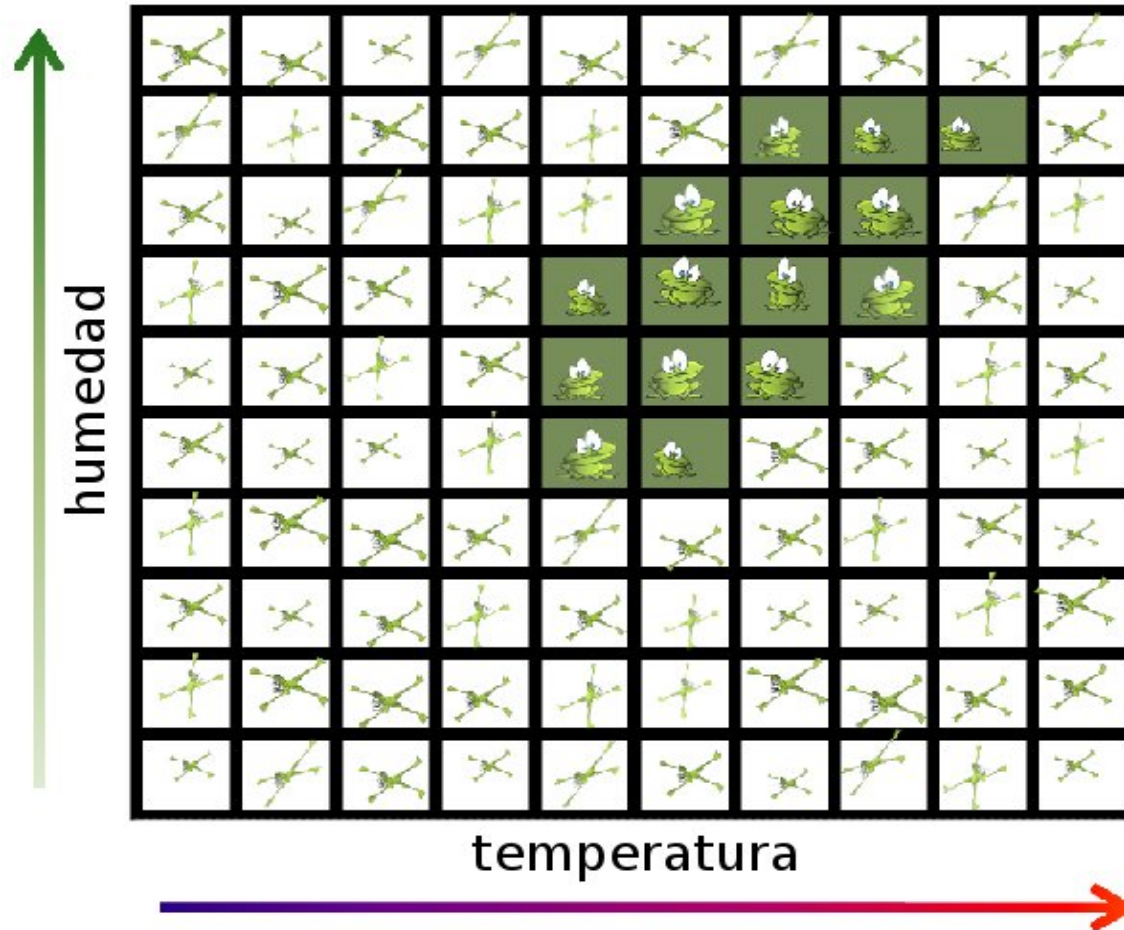


Y TAL VEZ EN EL FUTURO...



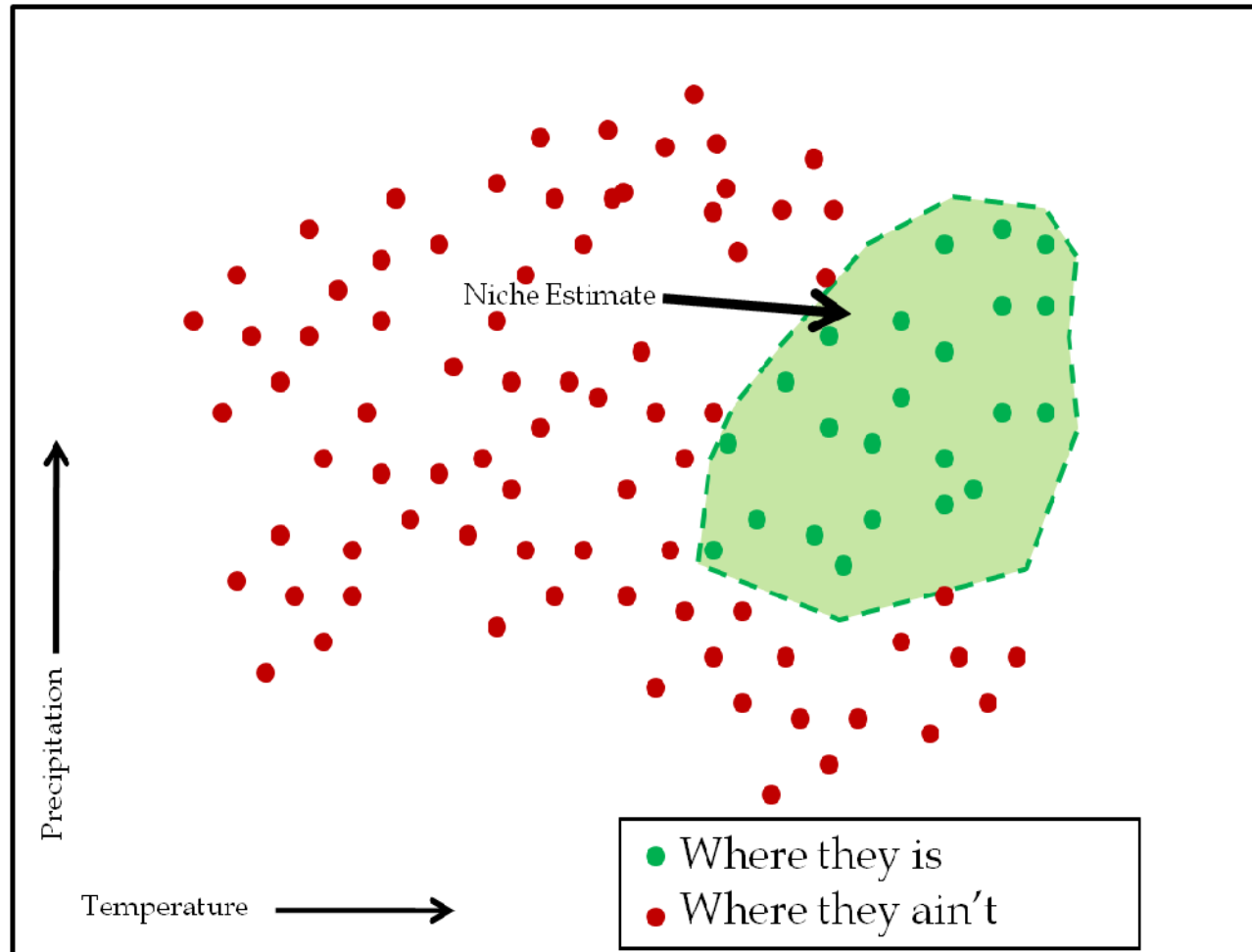
MODELOS DE DISTRIBUCIÓN DE ESPECIES

DETERMINACIÓN EXPERIMENTAL



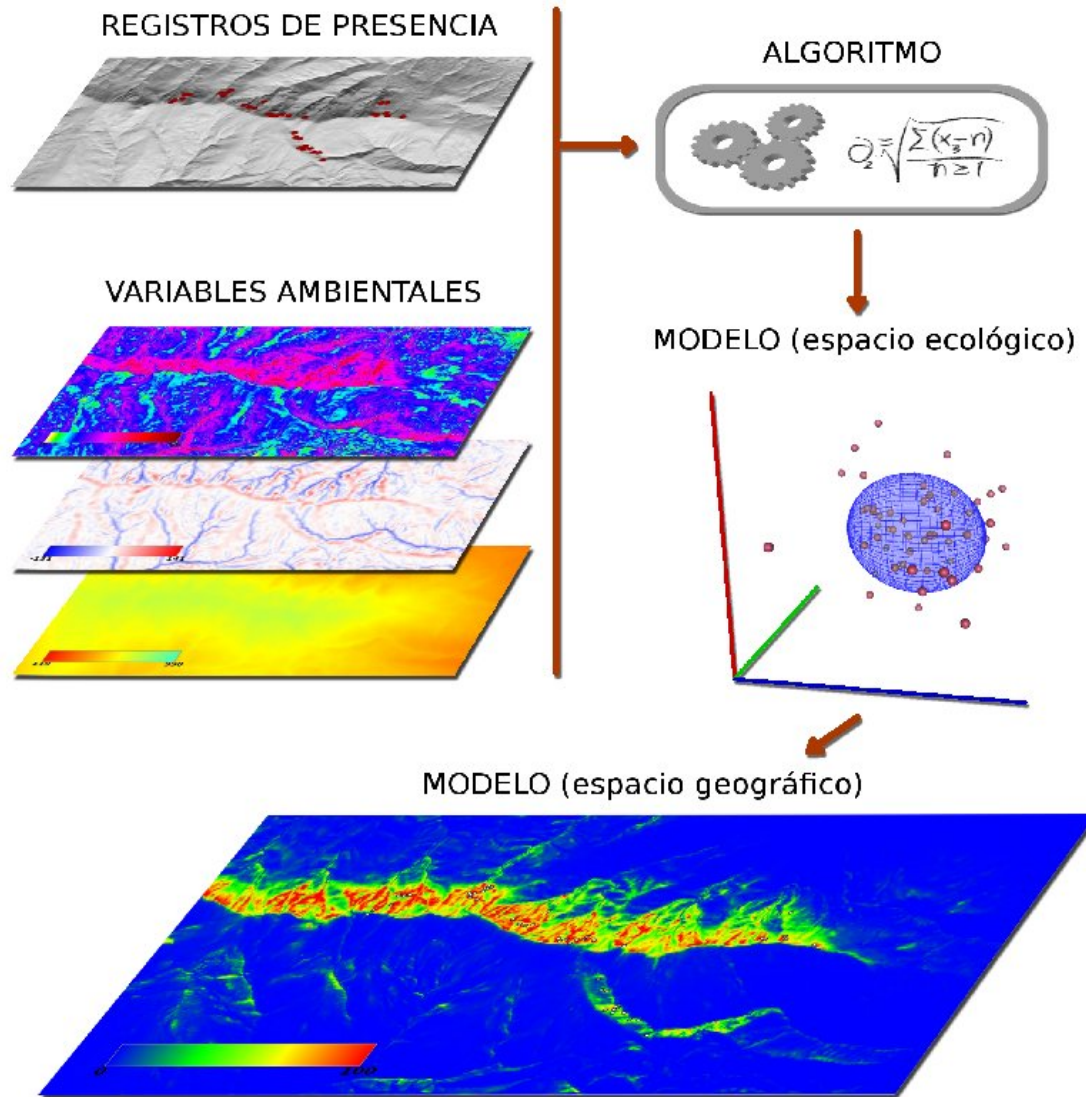
scienceasaverb.wordpress.com

DETERMINACIÓN DESCRIPTIVA

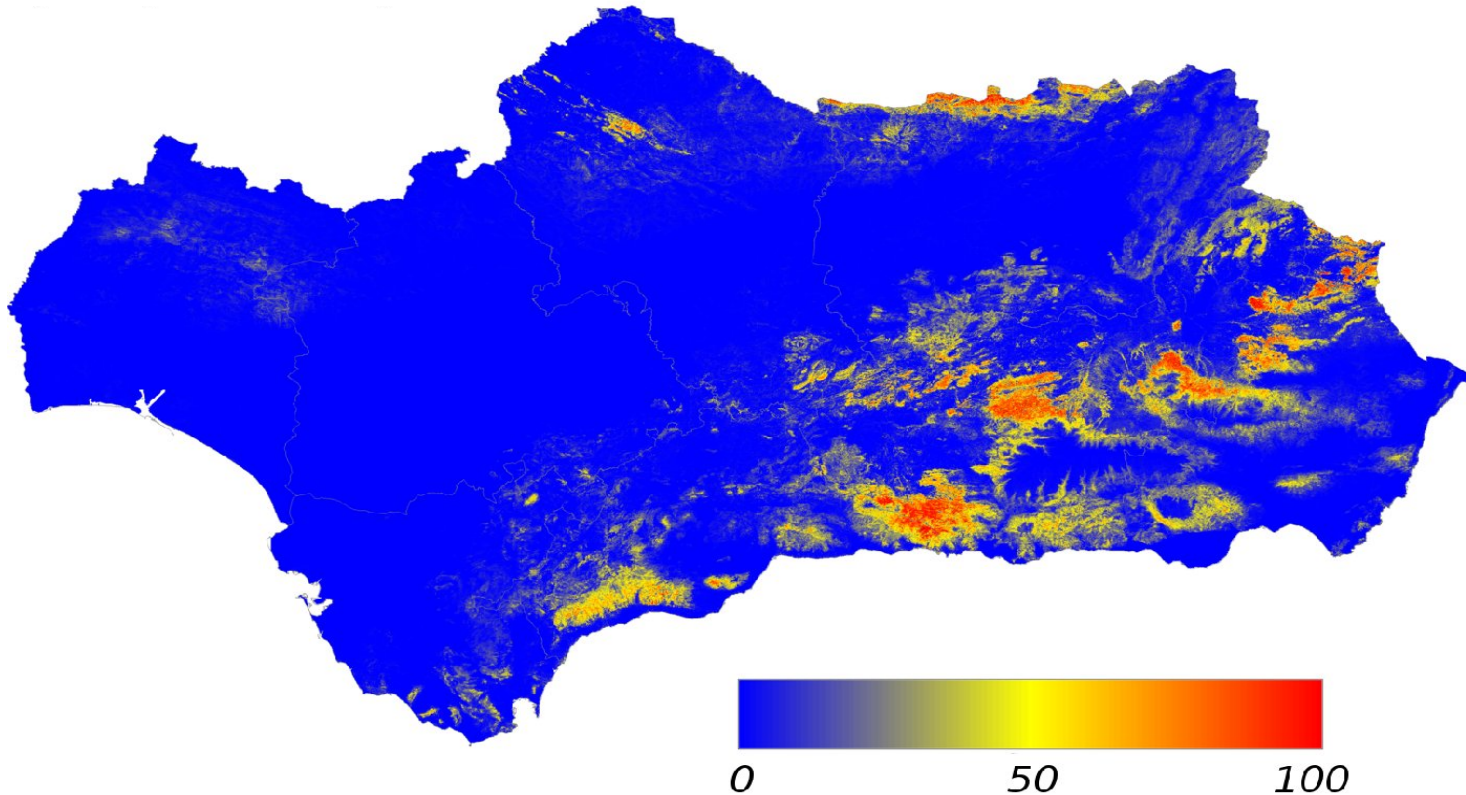


scienceasaverb.wordpress.com

MODELOS DE DISTRIBUCIÓN



MODELOS DE NICHO ECOLÓGICO



UNA ASUNCIÓN IMPORTANTE: EQUILIBRIO CON EL CLIMA

Hutchinson (1957):

“Una especie está en equilibrio con el clima si aparece en todas las áreas climáticamente apropiadas y está ausente de todas las que no lo son”

¿Están realmente las especies en equilibrio con el clima?

¿DINÁMICO O ESTÁTICO?

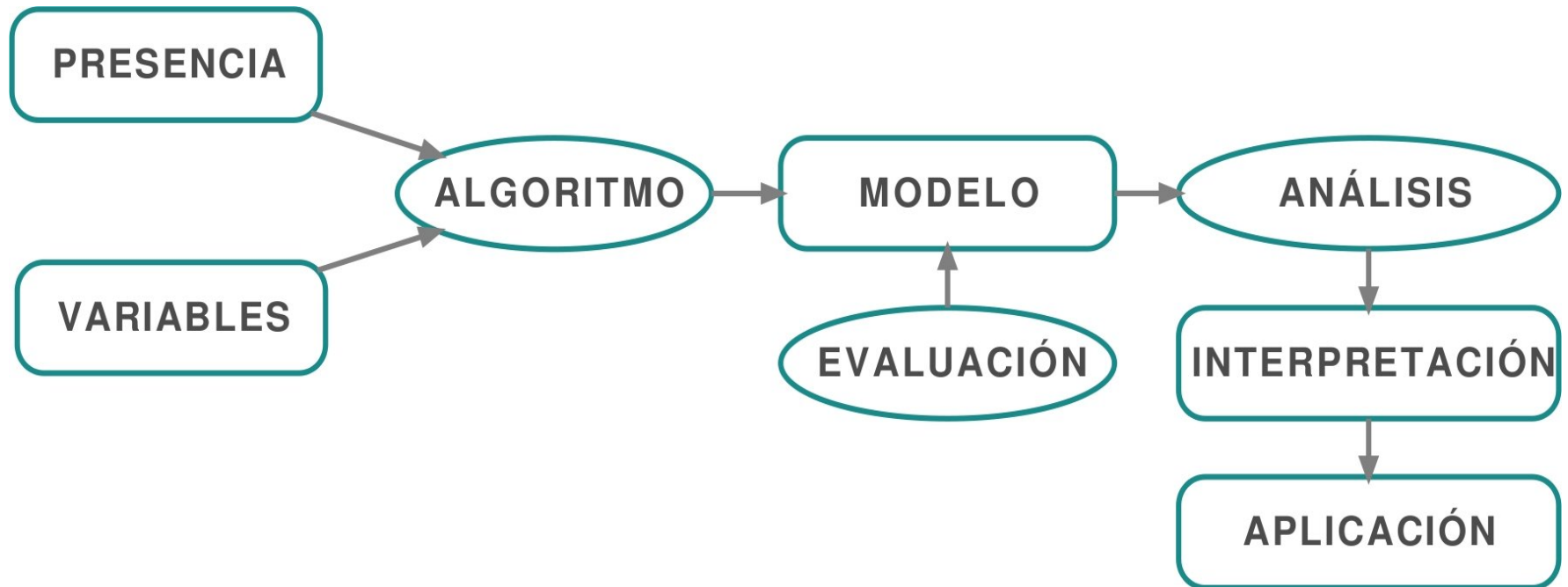
- No se sabe qué especies tienen nichos dinámicos, ni qué condiciones ecológicas que lo hacen posible
- Algunos estudios indican cambios en periodos menores a 100 años (¿especiación o adaptación?)
- Nicho dinámico posibilita la especiación, por lo que debe haber especies con nicho dinámico
- Las grandes poblaciones de especies muy extendidas presentan nichos muy conservados (flujo genético de centro a periferia previene la expansión de nicho)

EL PROCESO DE MODELADO

FORMULACIÓN TEÓRICA

- Definición de objetivos
- Conocimiento sobre la especie de trabajo
- Selección de factores potencialmente importantes
- Selección de herramientas que se van a utilizar, formatos, recursos informáticos, etc
- Diseño inicial del flujo de trabajo

FLUJO DE TRABAJO

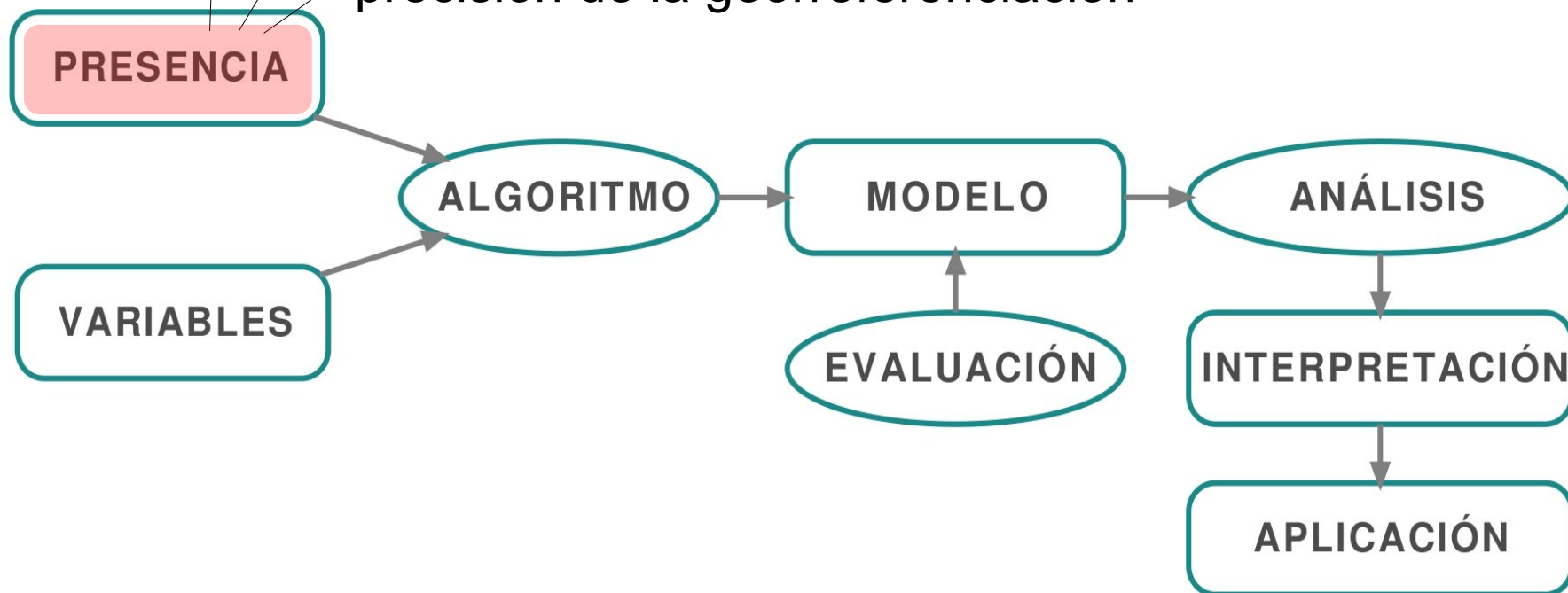


PRESENCIA

tamaño de muestra

sesgo de los datos

precisión de la georreferenciación



UN REGISTRO DE PRESENCIA...

Debería contener, al menos:

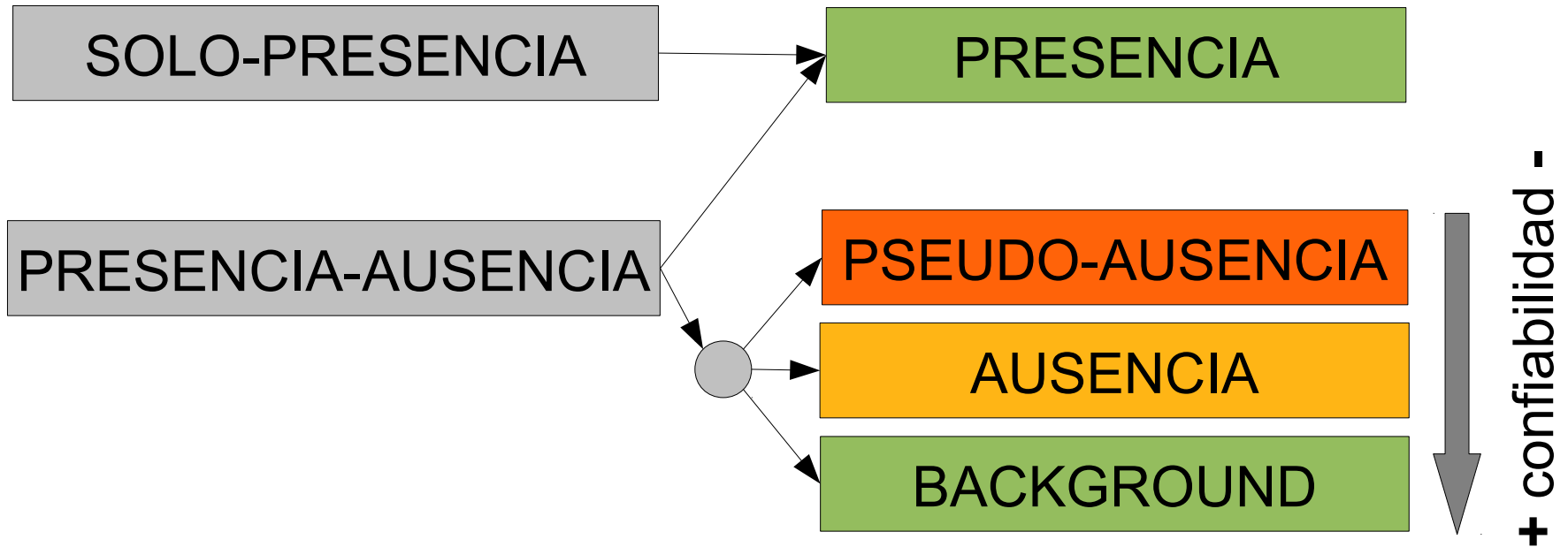
- Una coordenada X
- Una coordenada Y
- Un nombre de taxon (correctamente identificado)

Otros datos interesantes serían:

- Precisión de las coordenadas
- Fecha de recolección
- Y cualquier otro dato esencial para el trabajo en curso

¡FUNDAMENTAL!: conocer el sistema de referencia

TIPOS DE PRESENCIAS...



SOLO-PRESENCIA

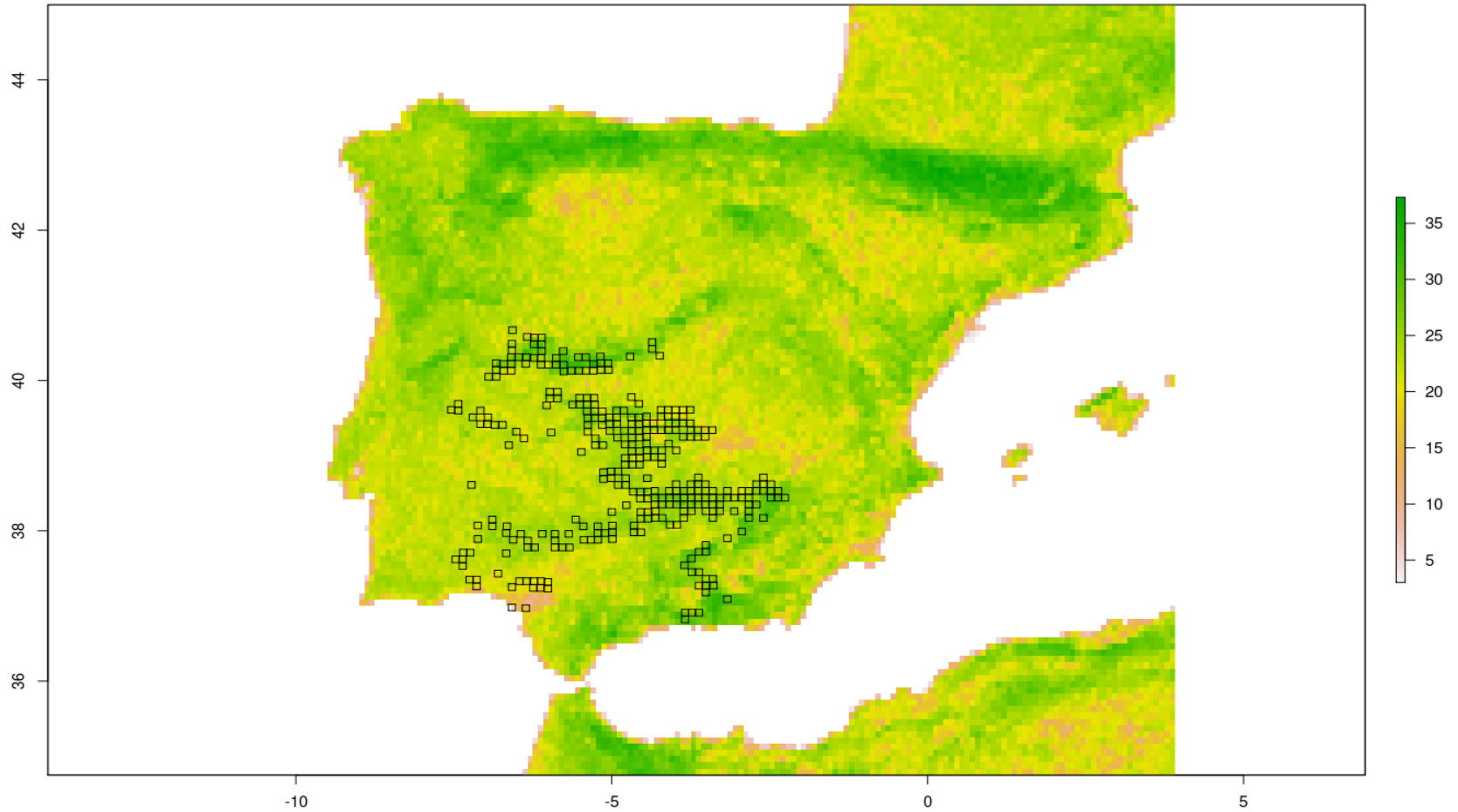
VENTAJAS

- Alta disponibilidad (colecciones → GBIF)

INCONVENIENTES

- No permiten calcular probabilidad de presencia
- No informan sobre sesgo en el muestreo
- Atención a precisión de coordenadas

SOLO-PRESENCIA



PRESENCIA-AUSENCIA

VENTAJAS

- Permiten calcular probabilidad de presencia
- Tienen en cuenta el sesgo en el muestreo

INCONVENIENTES

- ¿Cuándo y cómo una ausencia es una ausencia?
- Falsas ausencias (especies crípticas)
- Baja disponibilidad

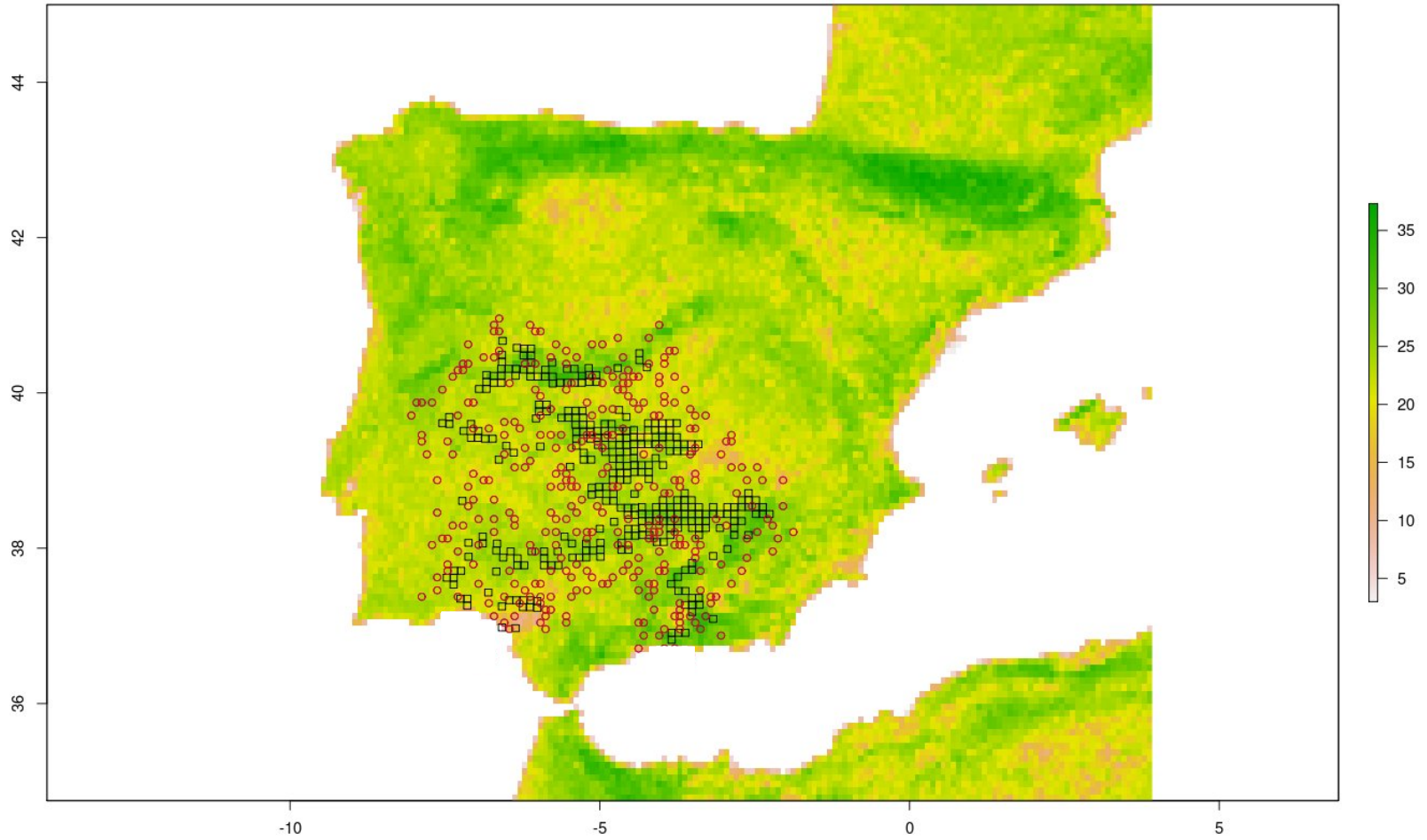
AUSENCIA

TRES TIPOS:

- **De contingencia:** en áreas idóneas, debido a restricciones dispersivas, históricas o biológicas
- **Ambientales:** por ausencia de hábitat idóneo idóneas
- **Metodológicas:** por sesgo de en el muestreo

Lobo et al. 2010

AUSENCIA



PSEUDO-AUSENCIA

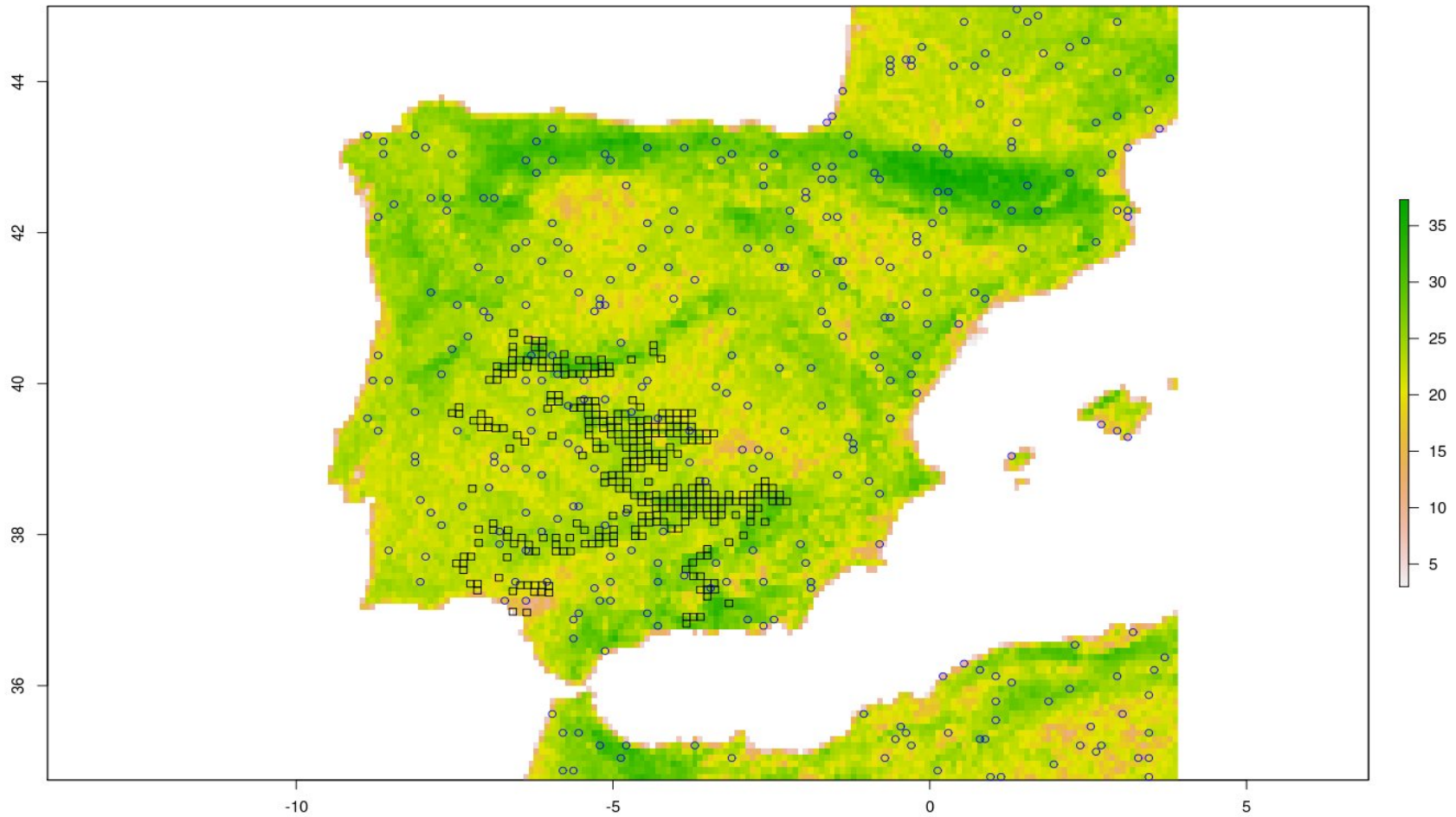
VENTAJAS

- Es fácil generarlas

INCONVENIENTES

- No son ausencias reales
- Criterio ecológico discutible
- No tienen en cuenta el sesgo del muestreo

PSEUDO-AUSENCIA



BACKGROUND:

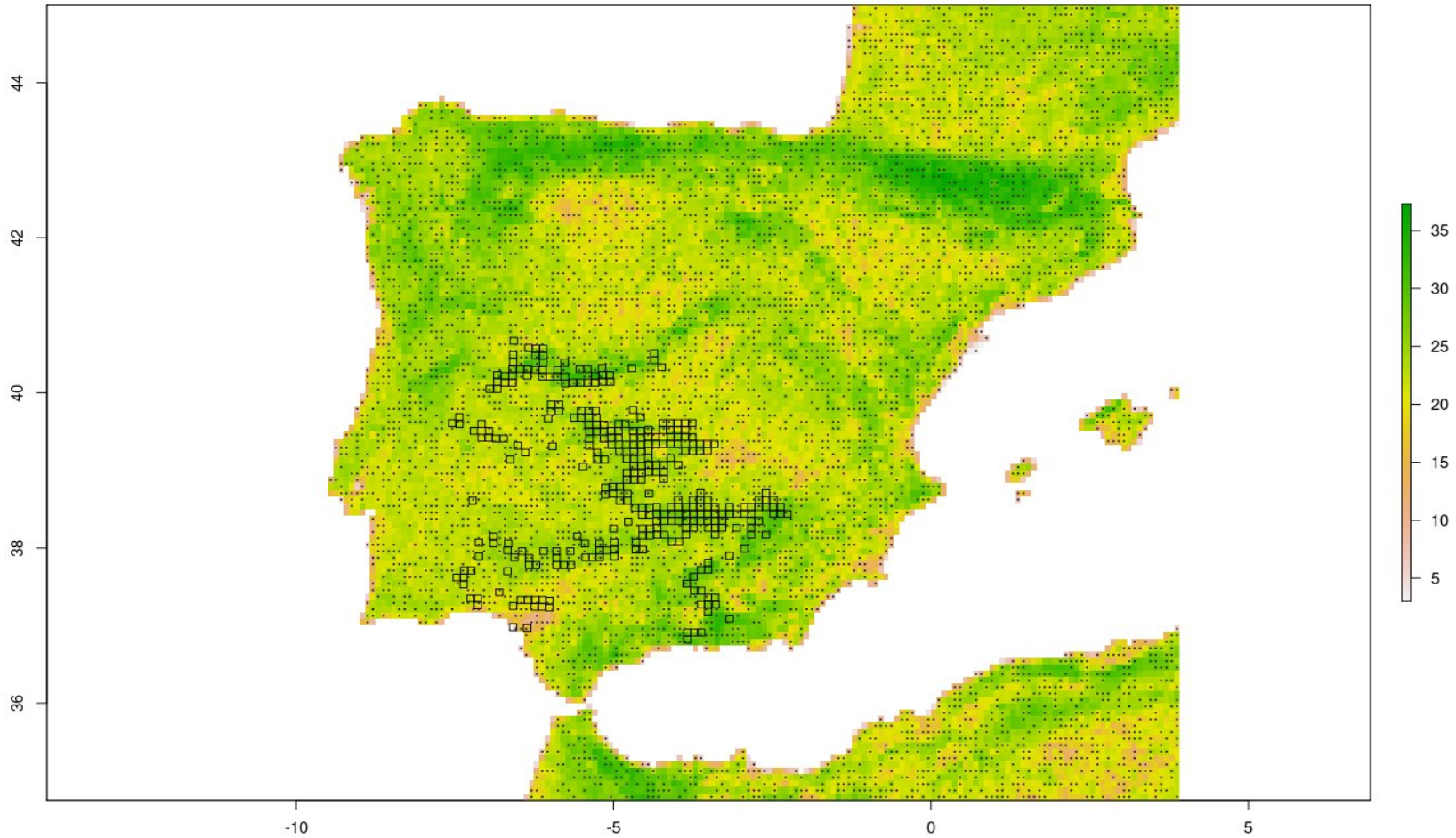
VENTAJAS

- Es fácil generarlas
- No hay problemas si solapan con las presencias.

INCONVENIENTES

- No funcionan bien para métodos de clasificación (Random Forest, Support Vector Machines)
- Tienen que ponderarse al trabajar con métodos de regresión

BACKGROUND:



SESGO

- Muestreo debe cubrir todo el rango ecológico significativo para la especie
- Un esfuerzo de muestro insuficiente o mal dirigido (solo cerca de carreteras, obviando gradientes importantes) desemboca en un SESGO
- Efecto negativo directo en MDE, porque se modelan el esfuerzo de muestreo en lugar de la distribución de la especie.

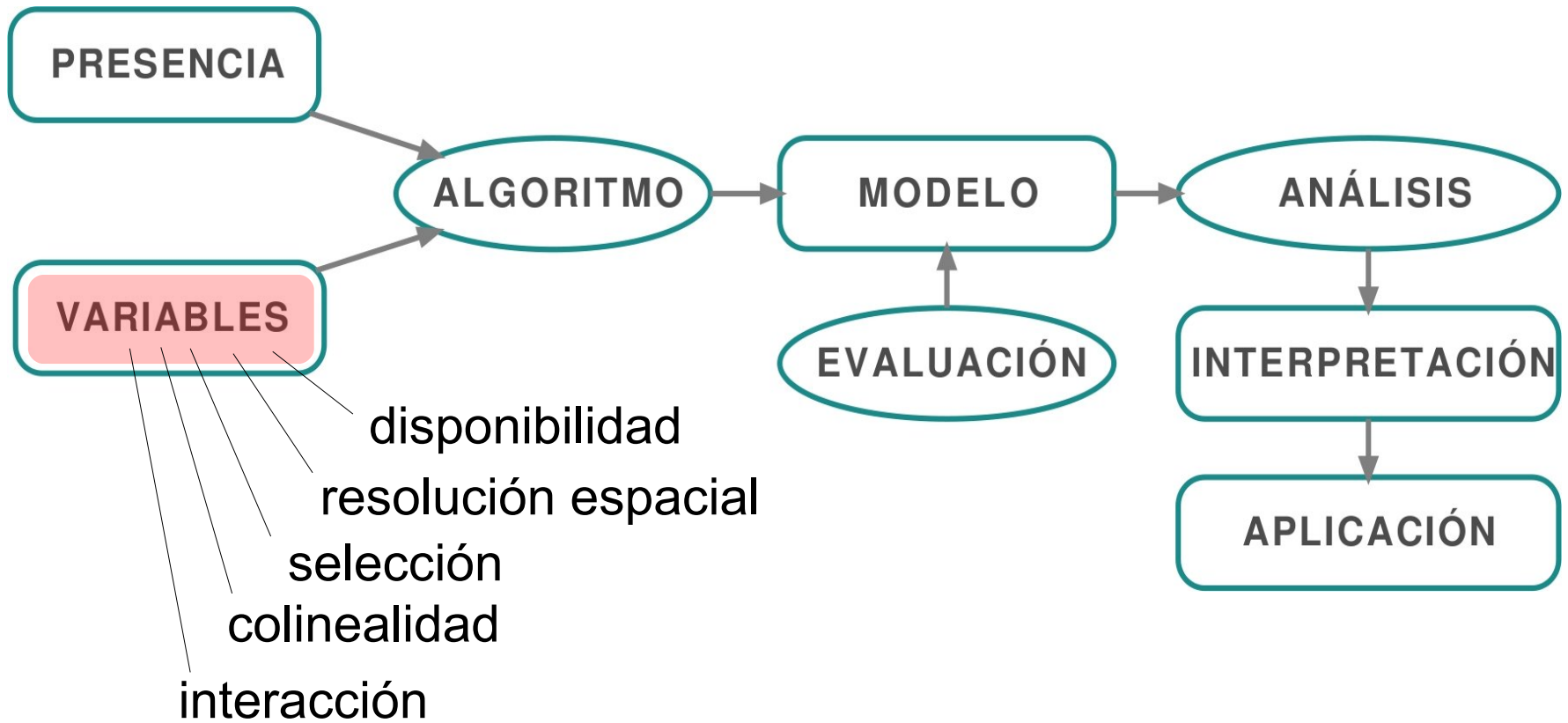
AUTOCORRELACIÓN ESPACIAL

- **Pseudorreplicación:** registros de presencia cercanos suelen tener las mismas combinaciones de valores para las variables ambientales.
- Viola la premisa de **independencia entre las observaciones**
- Para reducirla, basta con ampliar la distancia entre puntos de presencia (thinning).

PRECISIÓN DE LA GEORREFERENCIACIÓN

- Es muy variable, según la fuente (GBIF, desde metros hasta decenas de miles de metros).
- Debe ser acorde a la resolución de las variables ambientales.
- Los algoritmos de modelado toleran bien errores moderados de geolocalización (Graham et al 2008).

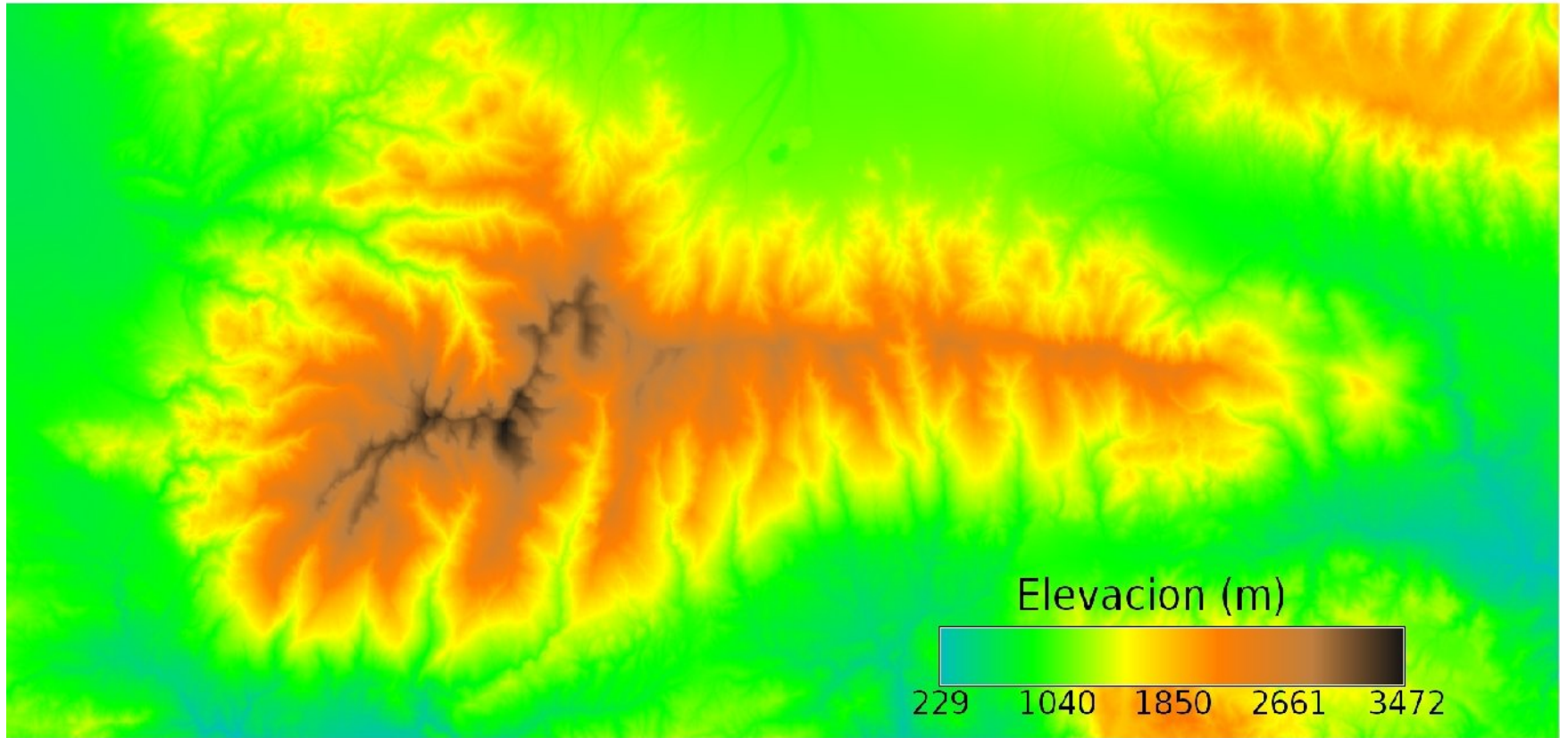
VARIABLES



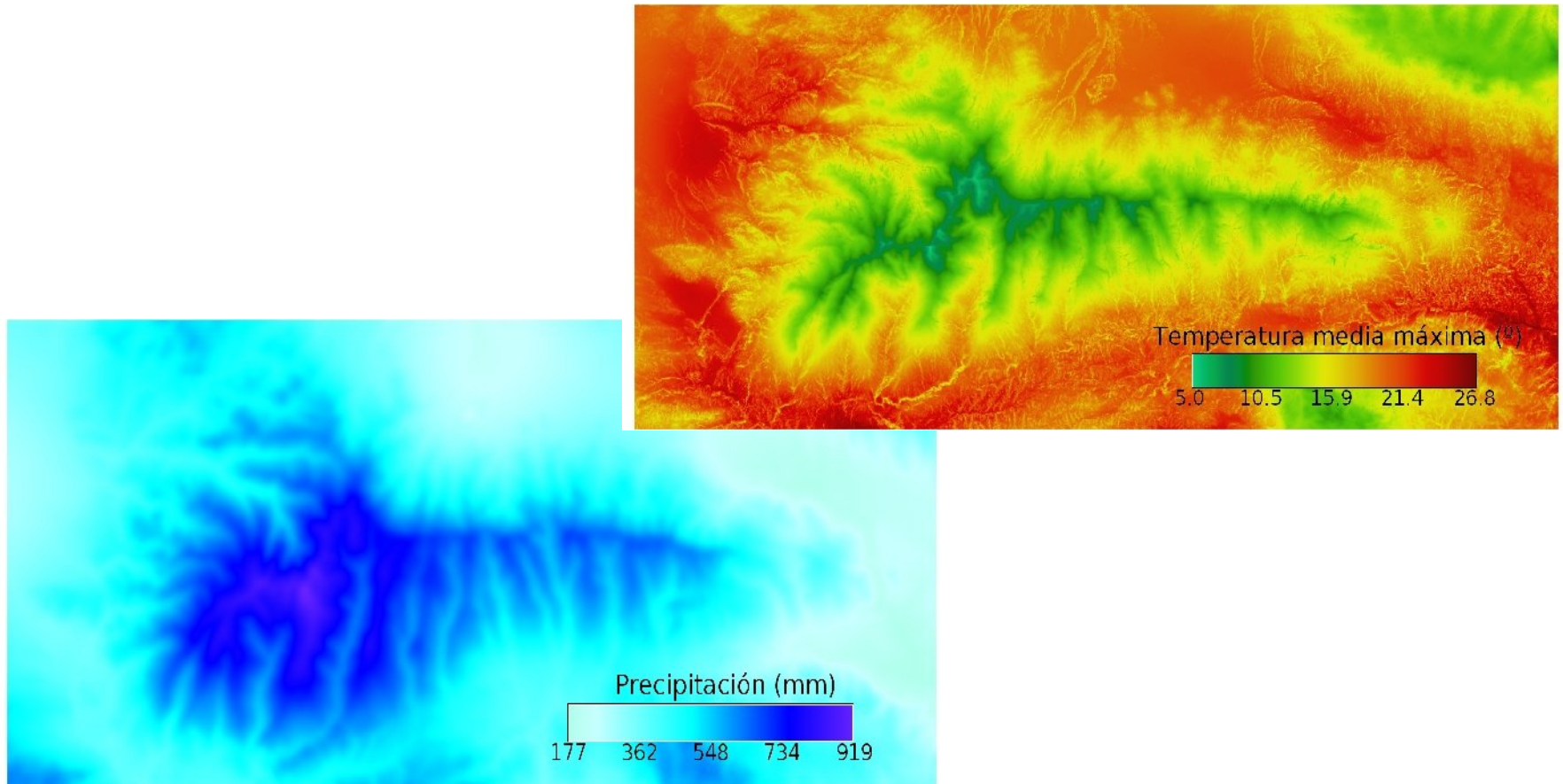
VARIABLES AMBIENTALES

- Deben tener cuatro características:
 - Afectar a la distribución de la especie
 - Mapas raster para GIS
 - Misma extensión y resolución
 - Baja colinealidad
- Resolución espacial acorde a datos de presencia
- Jerarquía de influencia (clima vs. topografía y usos del suelo según escala)
- Interacción entre variables

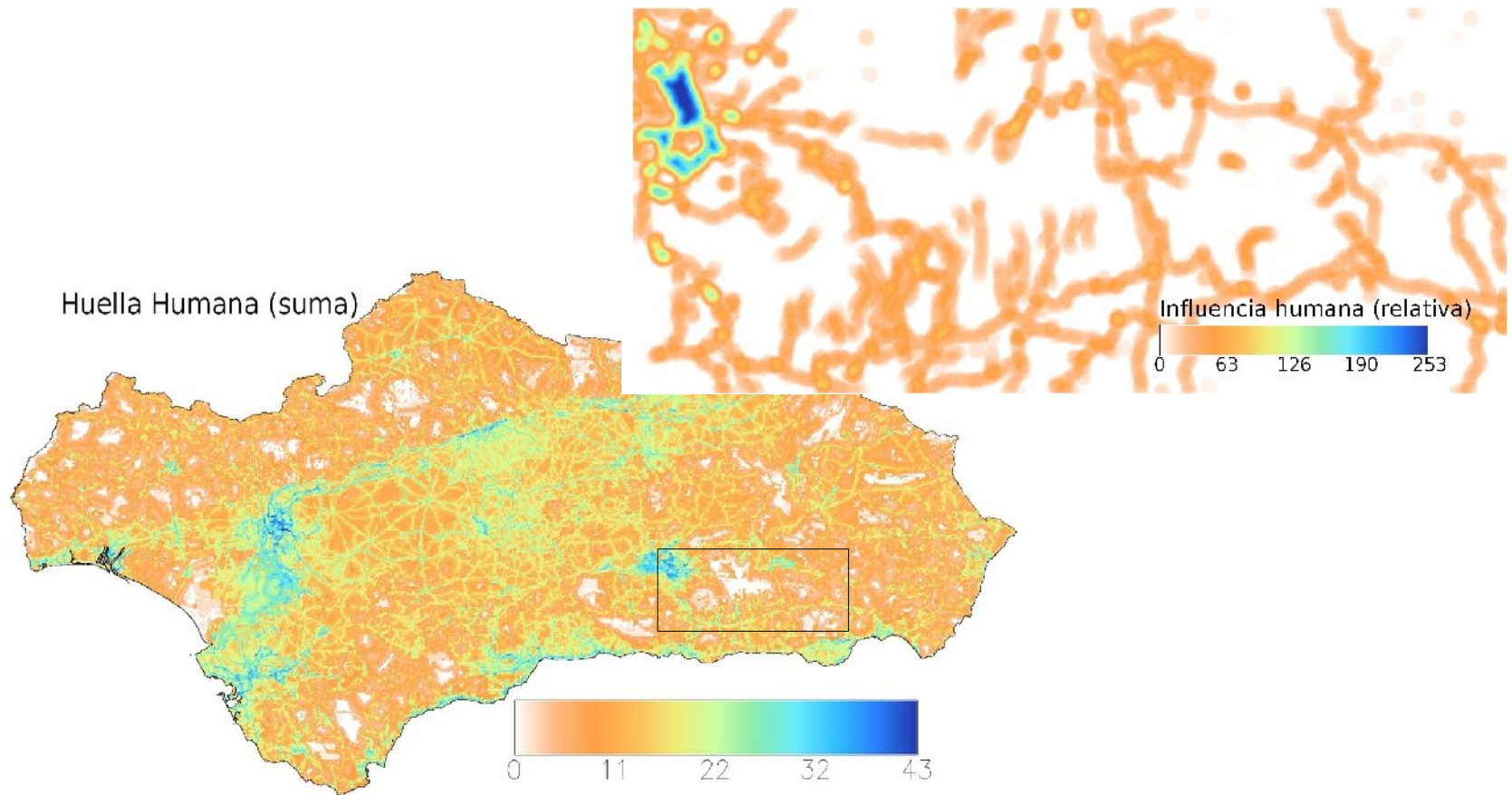
VARIABLES TOPOGRÁFICAS



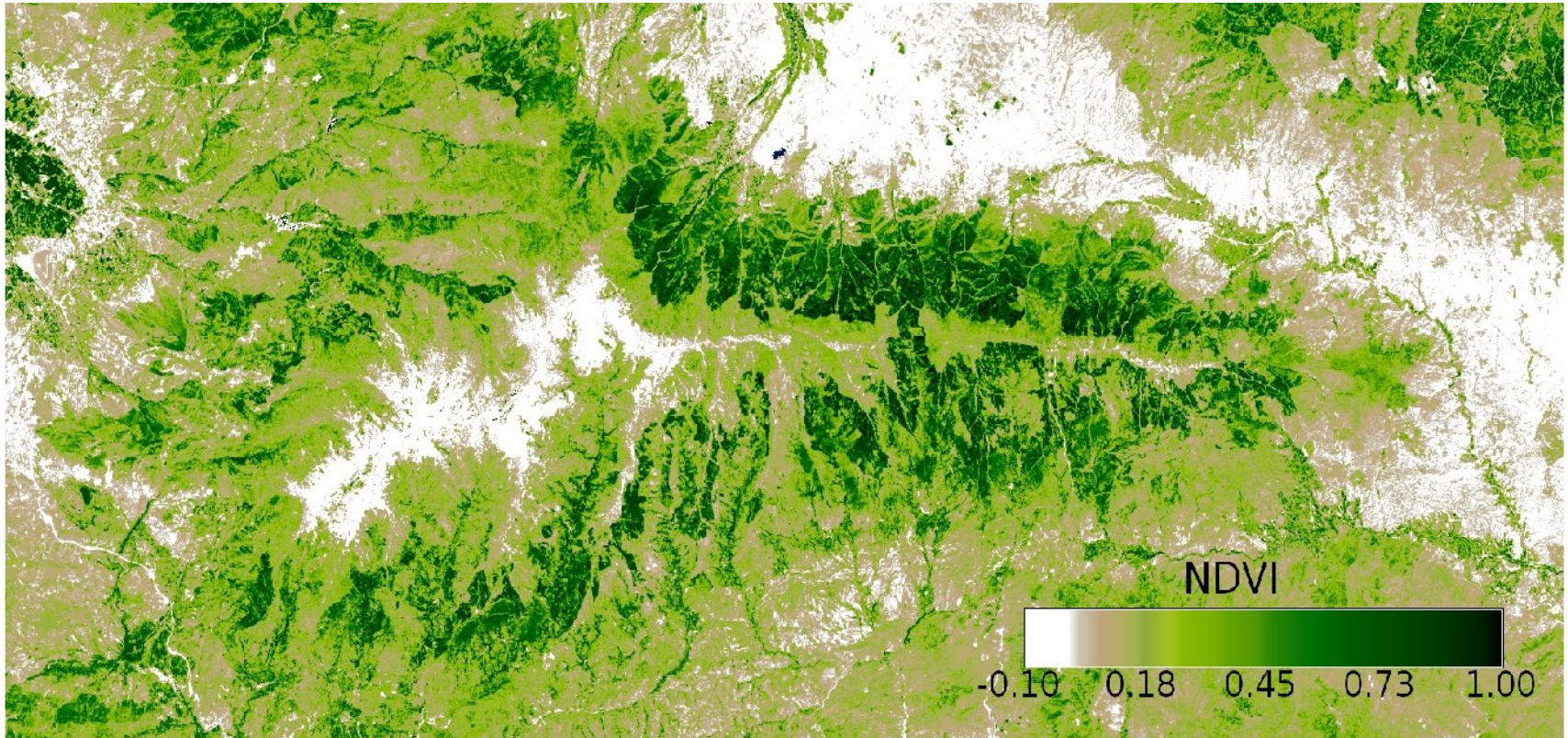
VARIABLES CLIMÁTICAS



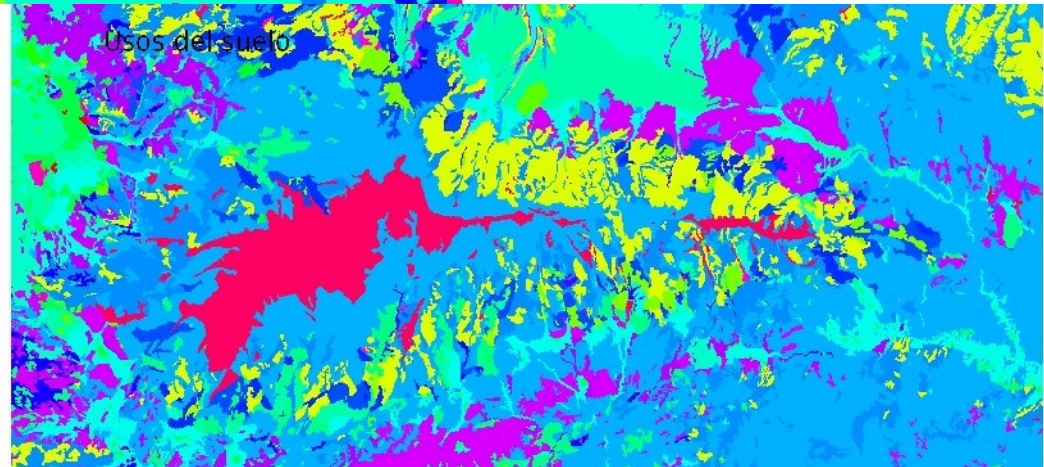
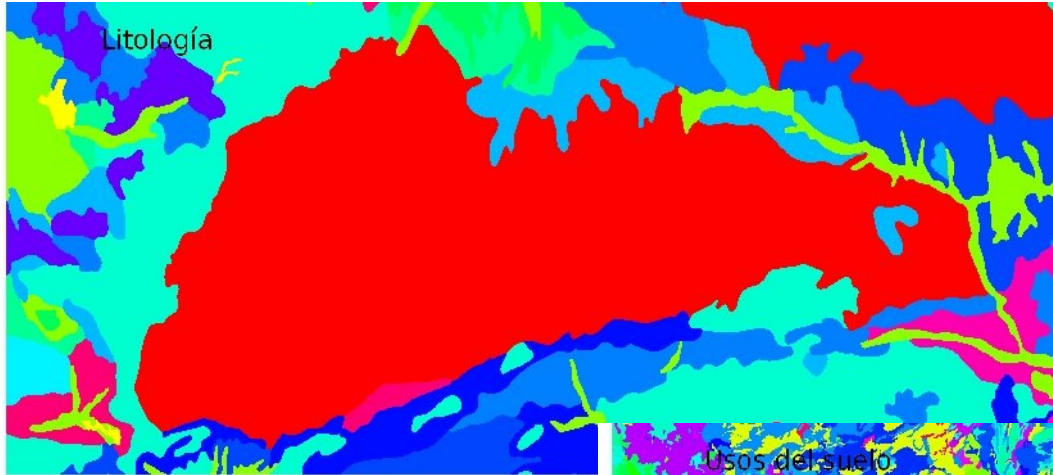
VARIABLES ANTRÓPICAS



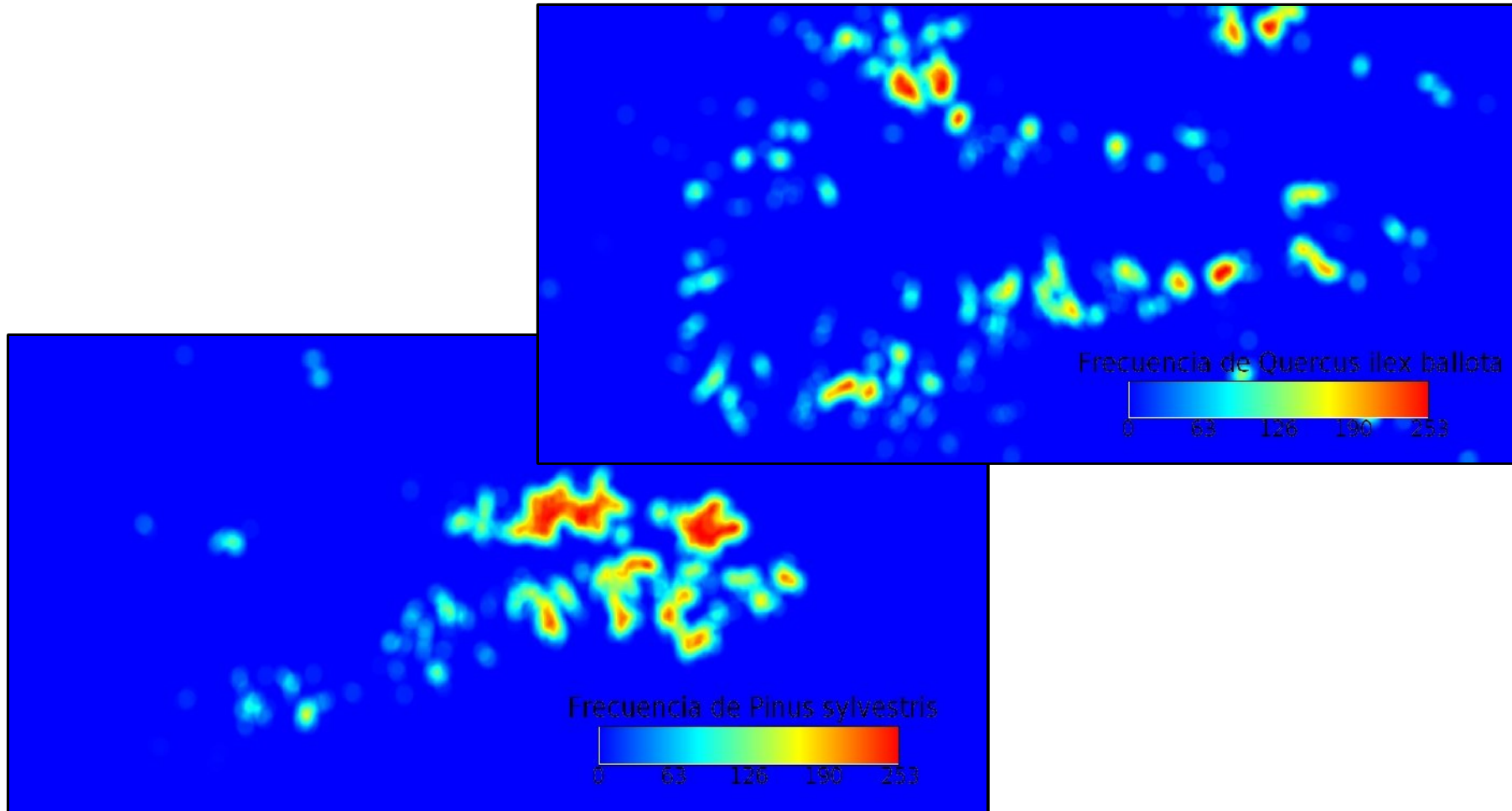
TELEDETECCIÓN



VARIABLES TEMÁTICAS



VARIABLES BIOLÓGICAS



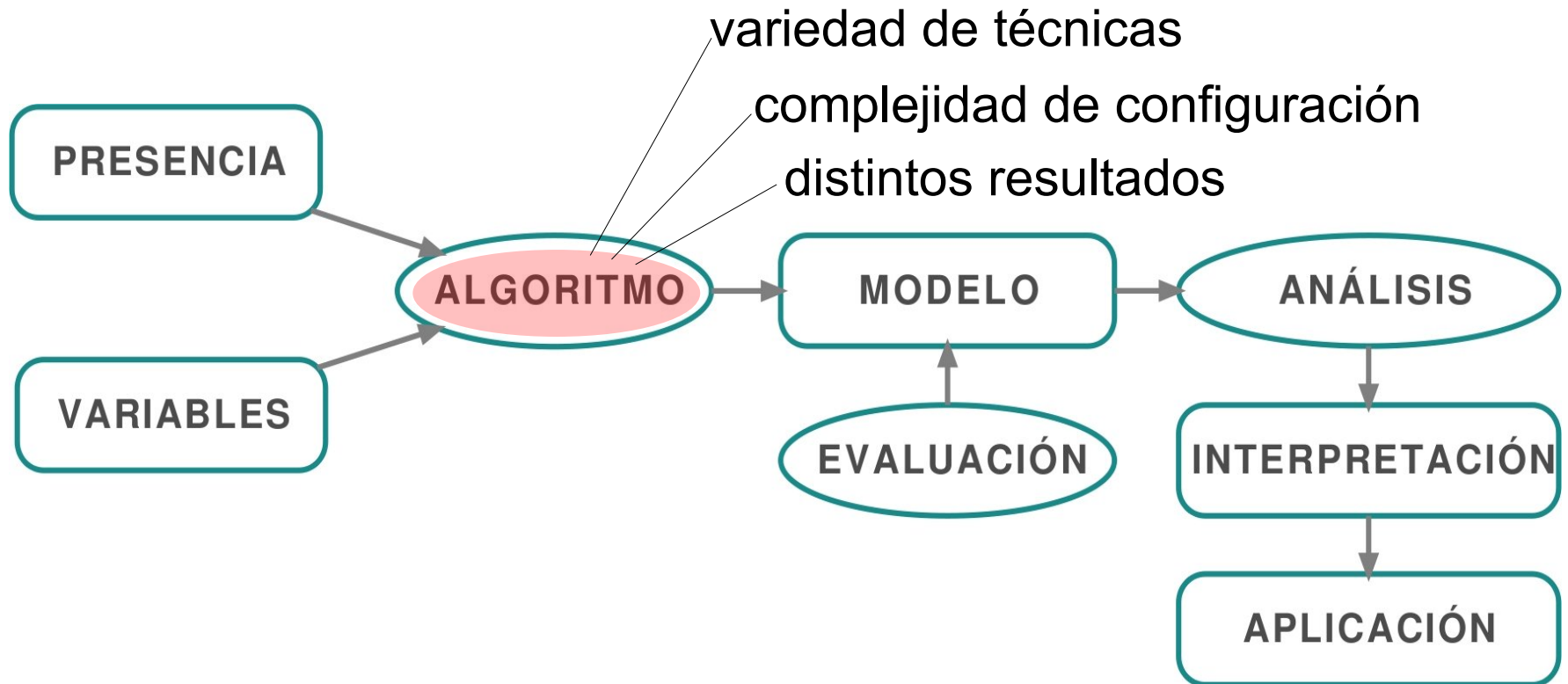
FUENTES MEDIO TERRESTRE

Tipo	Dataset	Sitio web
Clima	WORLDCLIM	http://www.worldclim.org/
Topografía	SRTM	http://www2.jpl.nasa.gov/srtm/
NDVI	GIMMS	http://glcf.umd.edu/data/gimms/
Vegetación	MODIS VCF	http://glcf.umd.edu/data/vcf/
Huella humana	Human Footprint	http://sedac.ciesin.columbia.edu
Usos del suelo	GLOBCOVER	http://www.edenextdata.com

FUENTES MEDIO MARINO

- **Bio-ORACLE** (<http://www.oracle.ugent.be/>)
 - 23 variables
 - Resolución: 9.2 km
 - Referencia: Tyberghein et al. (2012) DOI: 10.1111/j.1466-8238.2011.00656.x
- **MARSPEC** (<http://www.marspec.org/>)
 - 17 variables bioclimáticas y topográficas
 - Resolución: 1 km
 - Referencia: Sbrocco and Barber (2013) DOI: 10.1890/12-1358.1

ALGORITMO

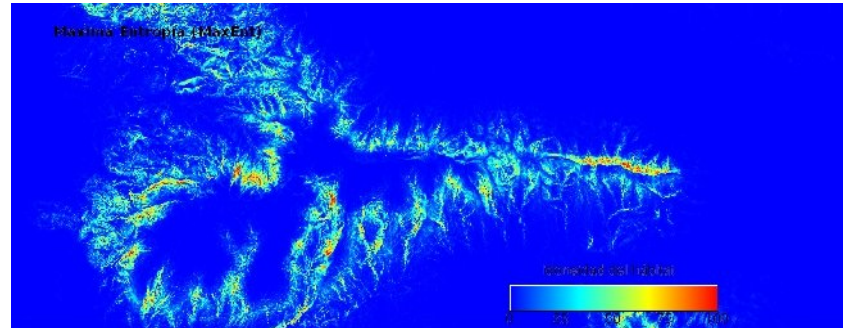
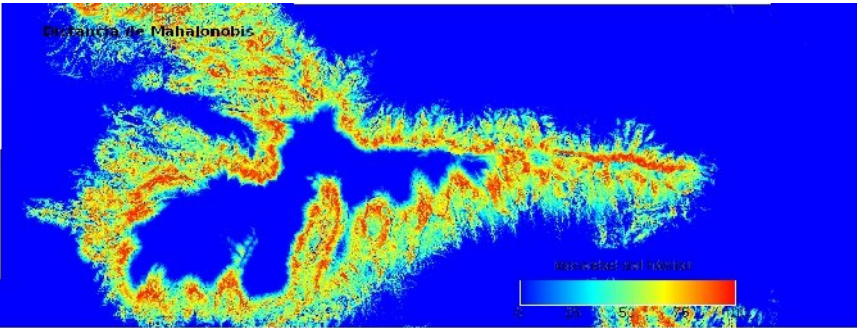
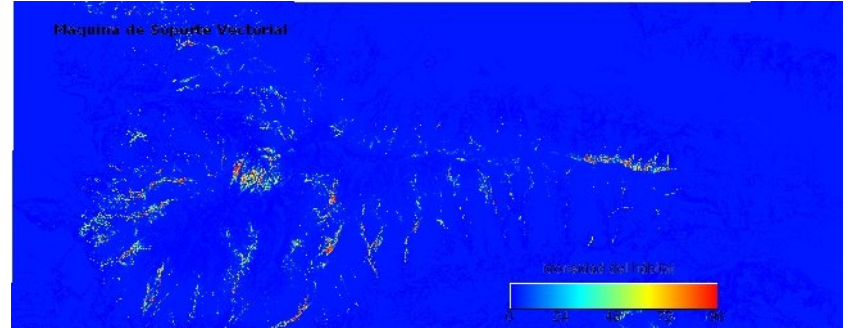
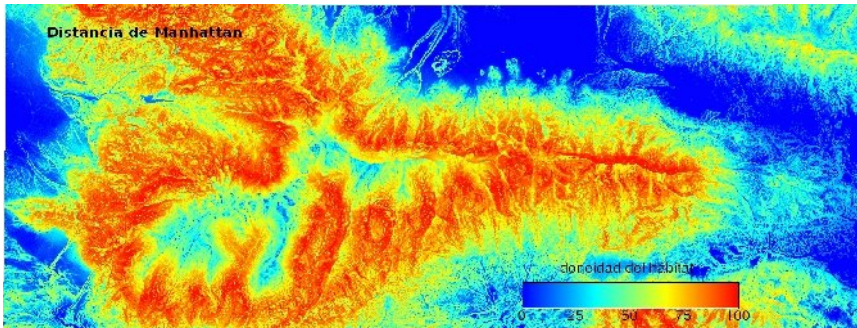
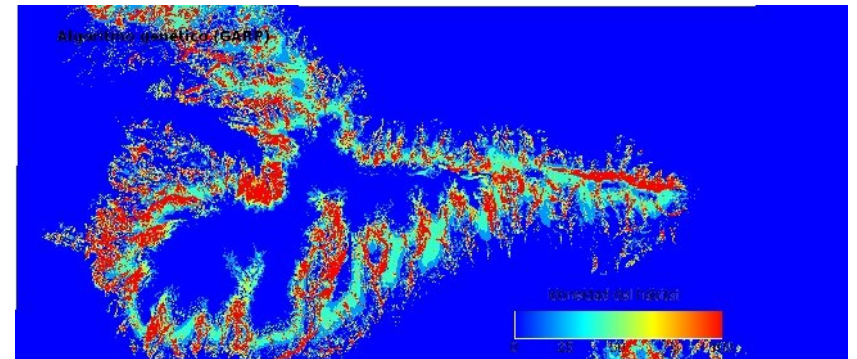
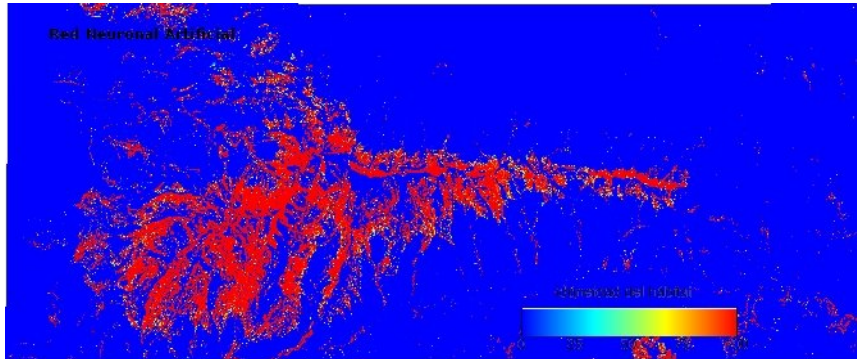


ALGORITMO

Familias de métodos (complejidad creciente):

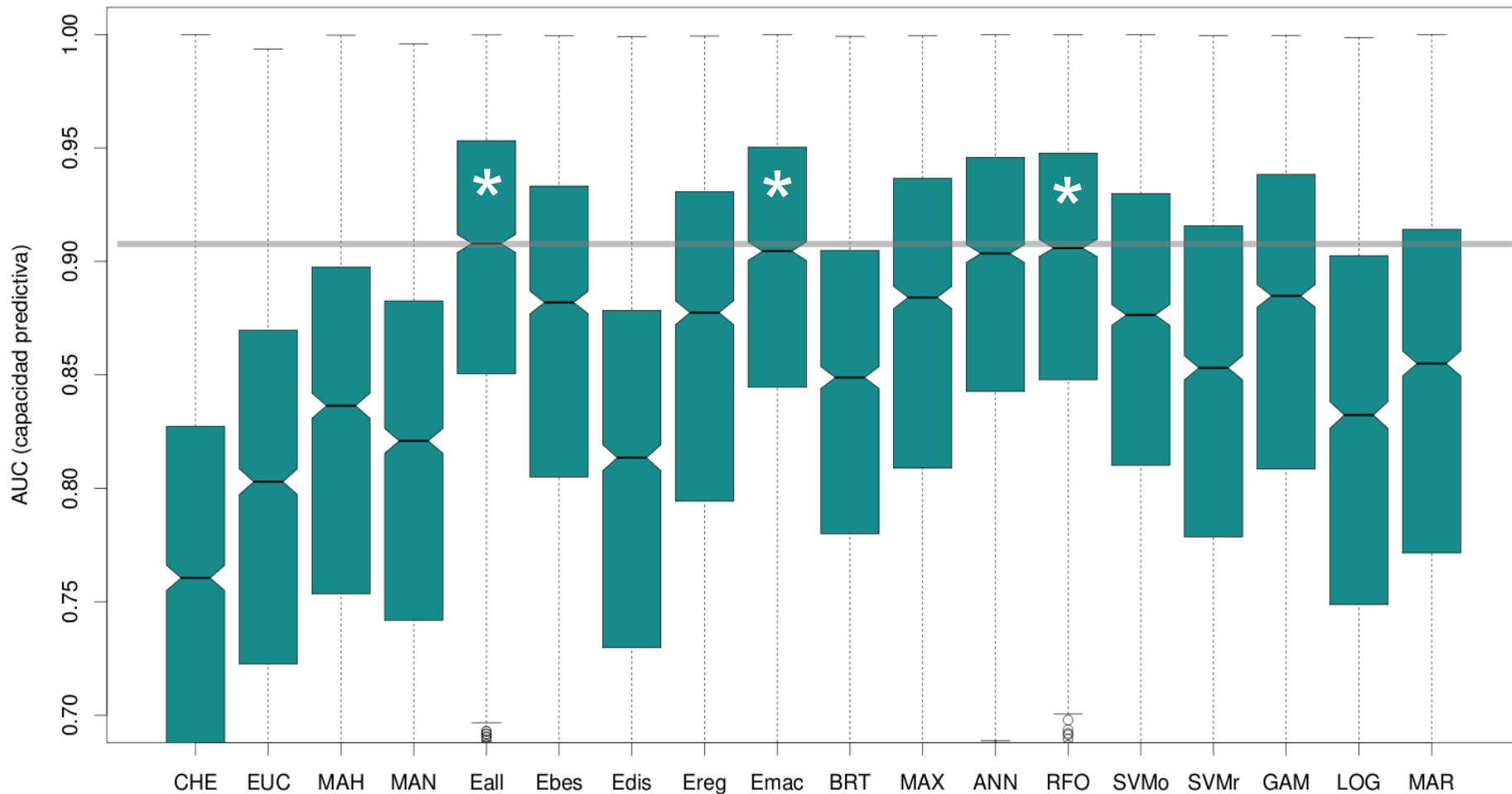
- Envueltas bioclimáticas
- Modelos de similaridad ecológica
- Métodos de regresión
- Métodos de aprendizaje artificial

DISTINTOS RESULTADOS...



...Y DISTINTO PODER PREDICTIVO

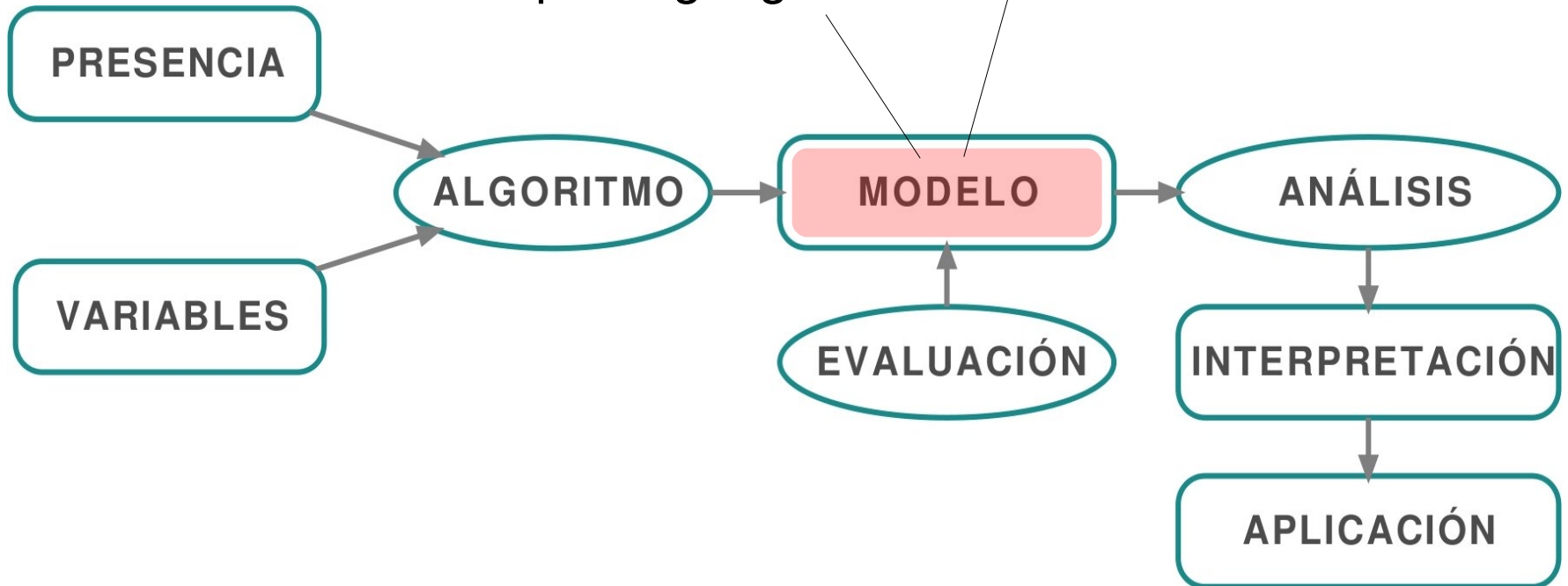
Evaluación de modelos de 1700 especies de árboles en Mesoamérica



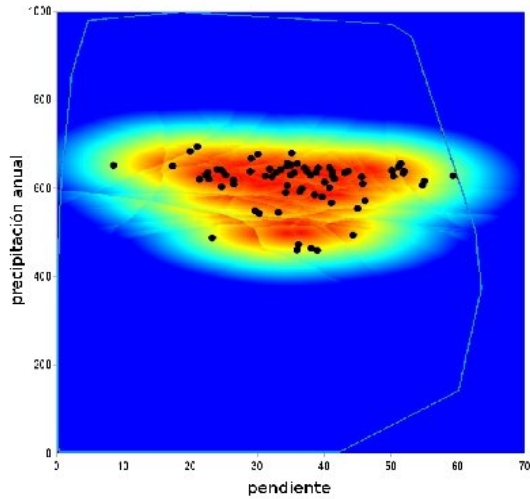
MODELO

modelo en el espacio ecológico

modelo en el espacio geográfico

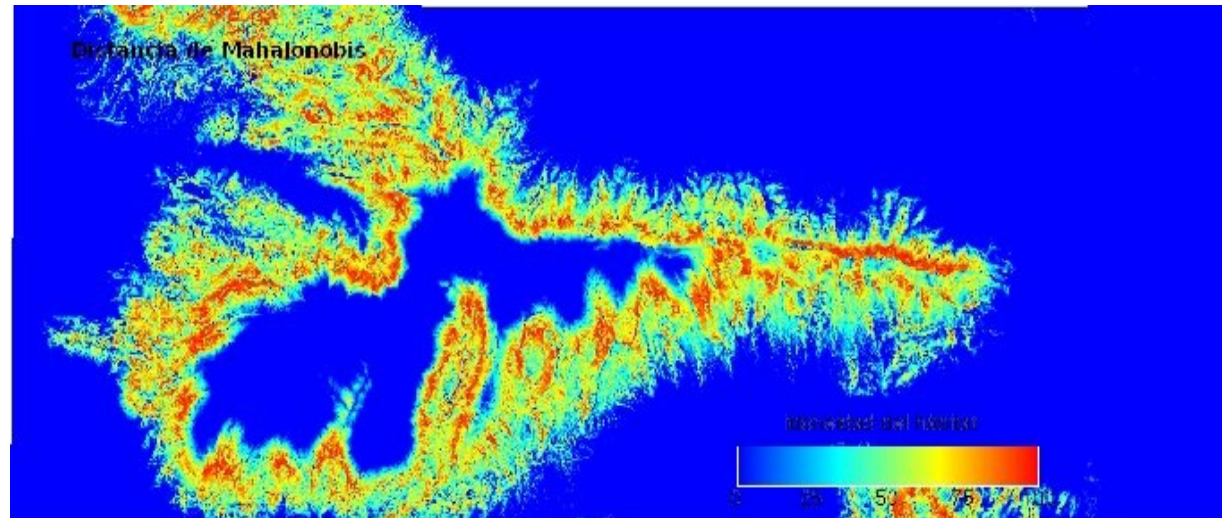


MODELO

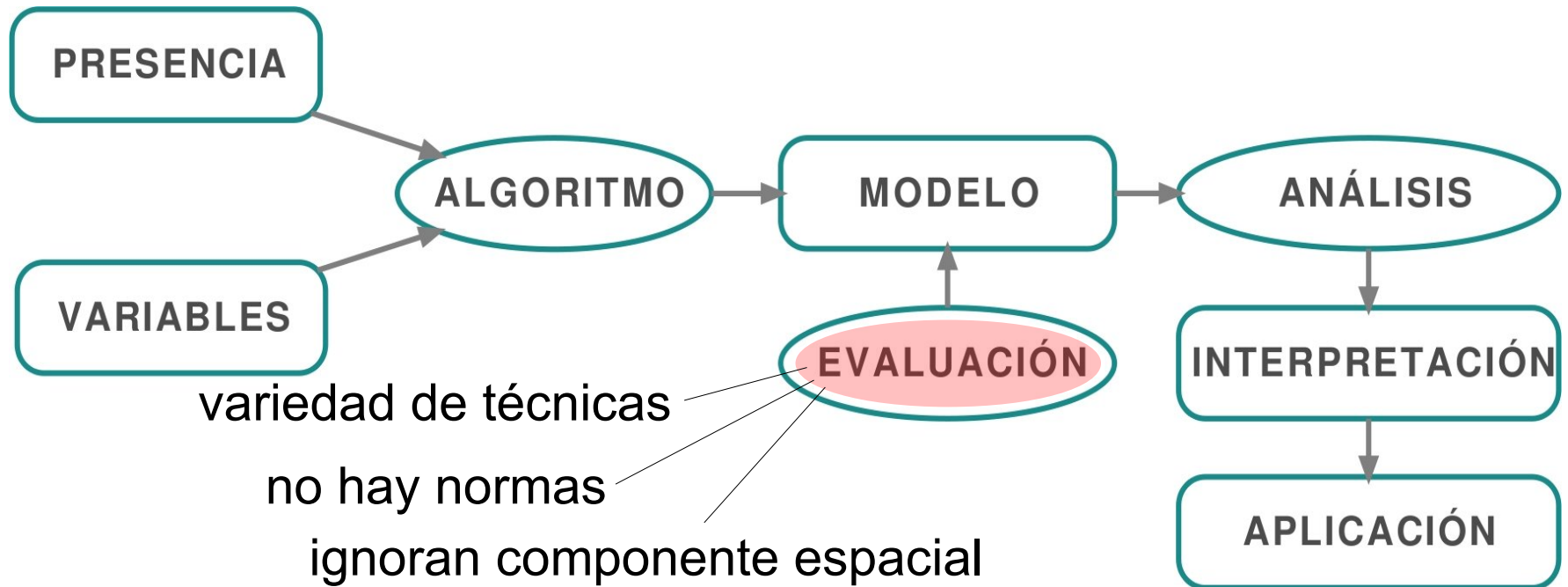


← espacio ecológico

espacio geográfico



EVALUACIÓN

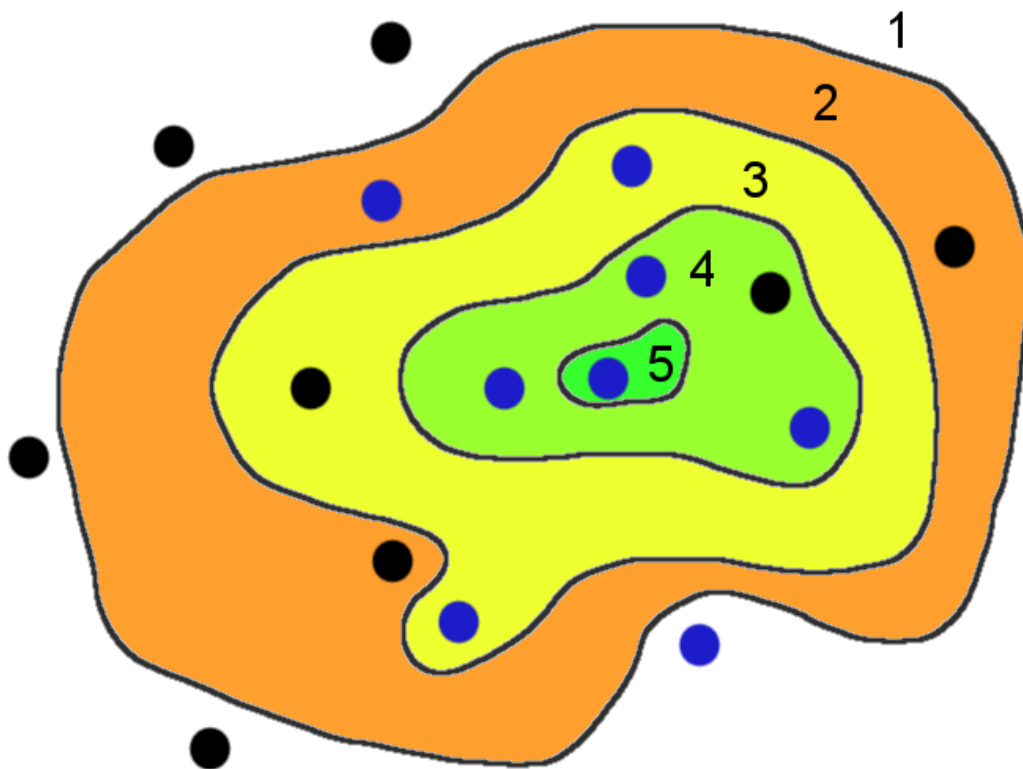


EVALUACIÓN

Artículo clave:

Fielding AH y Bell JF 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24(1), 38-49
(1739 citas en mayo de 2013)

EVALUACIÓN



1-5	Idoneidad
●	Presencia (evaluación)
●	Ausencia

MATRIZ DE CONFUSIÓN

A → presencias acertadas

D → ausencias acertadas

B → ausencias fallidas (falsos positivos o error de comisión)

C → presencias fallidas (falsos negativos o error de omisión)

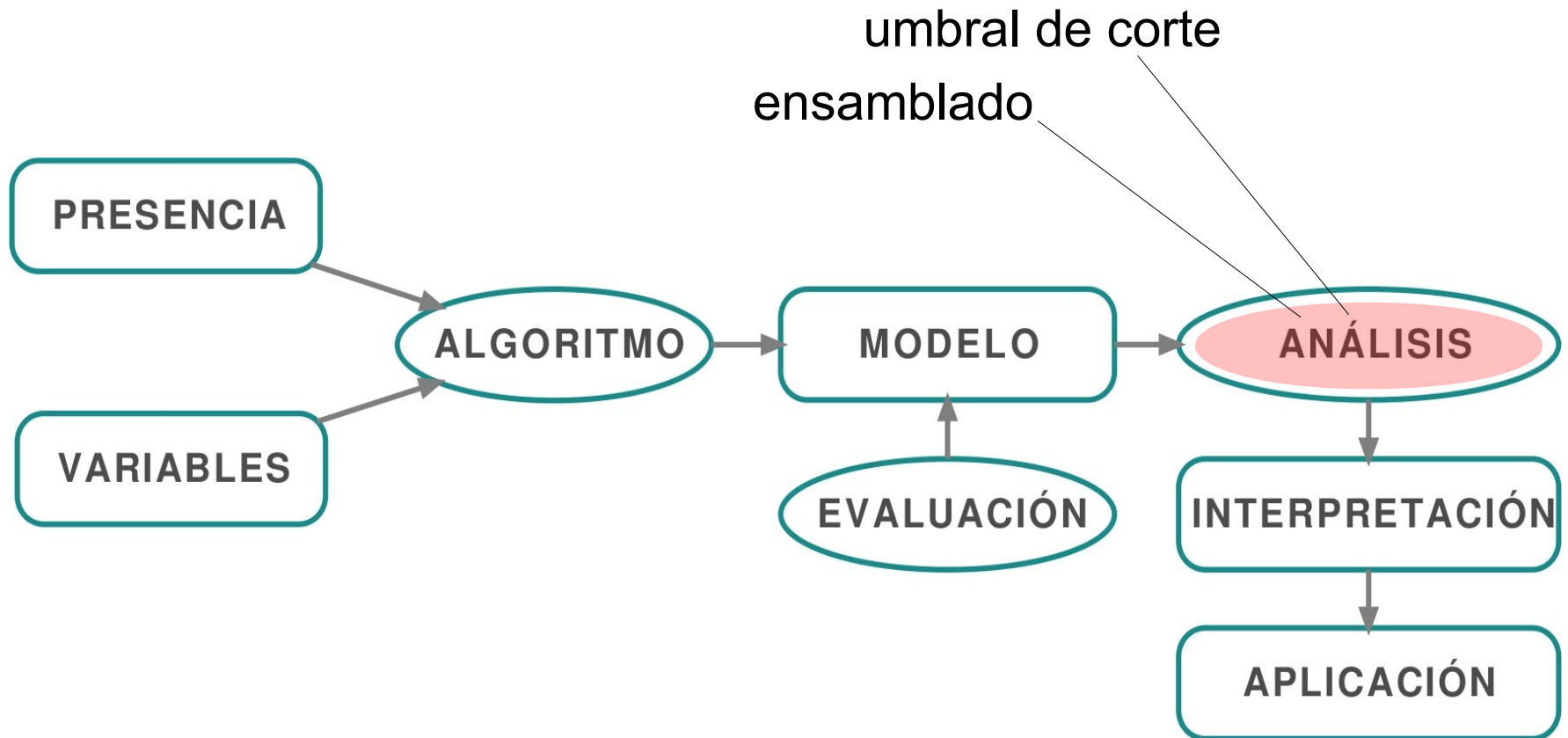
		Datos reales (registros de presencia y ausencia)	
		presencia	ausencia
Datos simulados (modelo de distribución)	presencia	A	B
	ausencia	C	D

sensibilidad → $S = A/(A+C)$

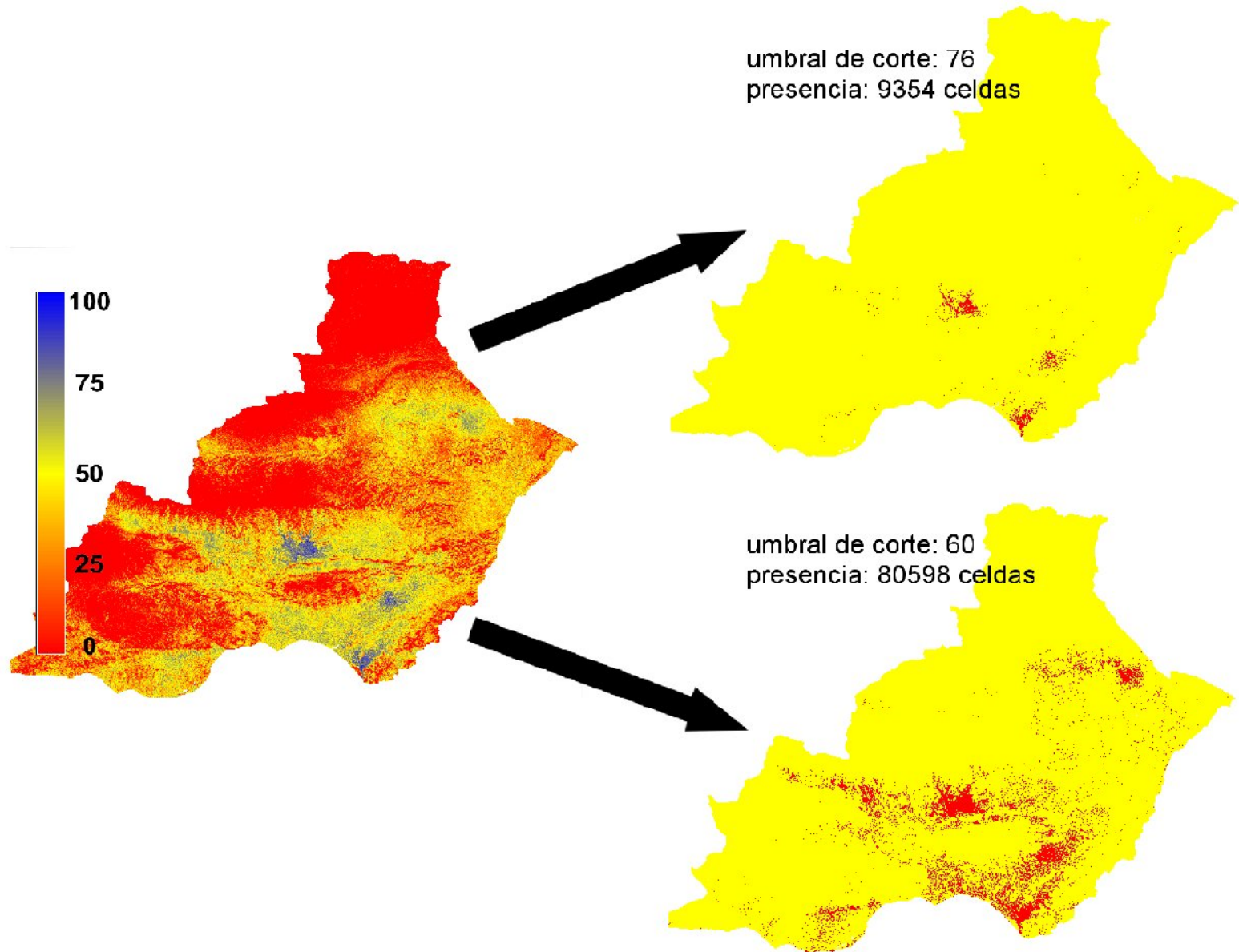
especificidad → $E = D/(B+D)$

true skill statistic → $TTS = S + E - 1$

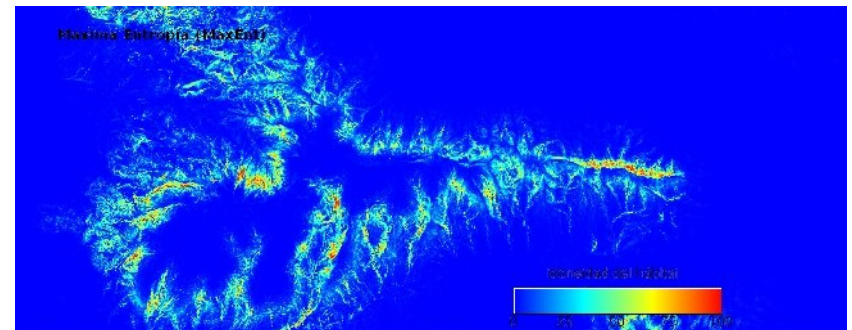
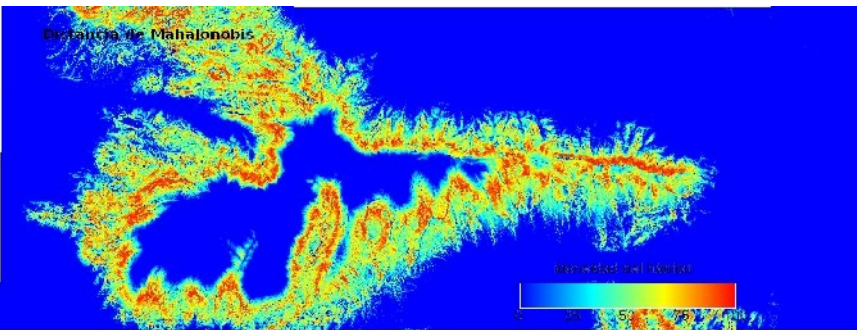
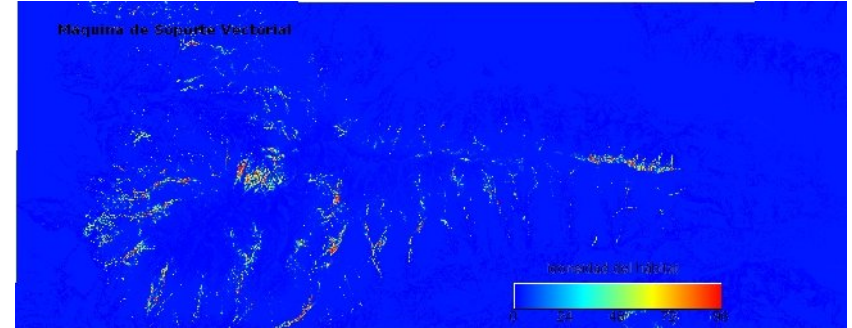
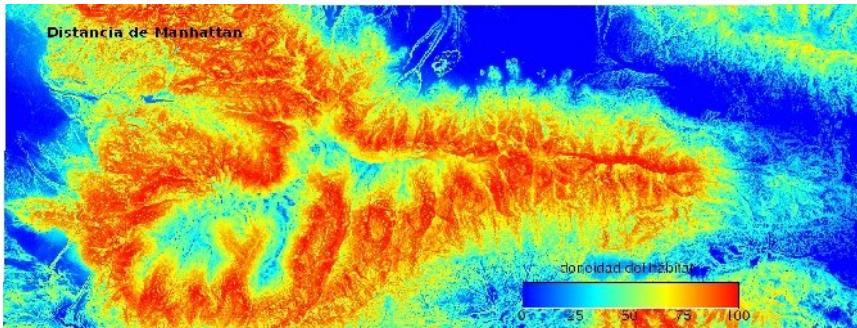
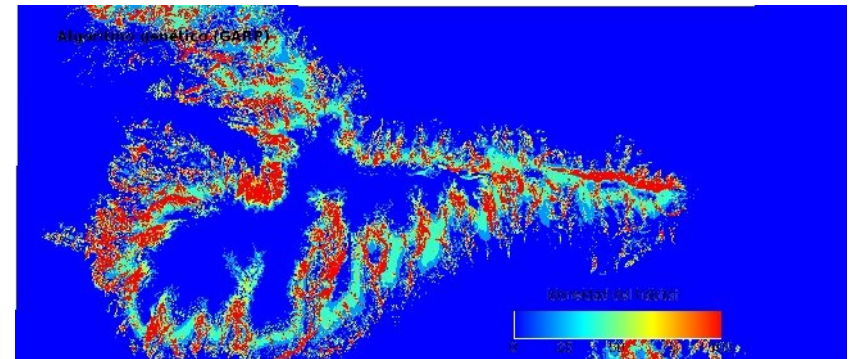
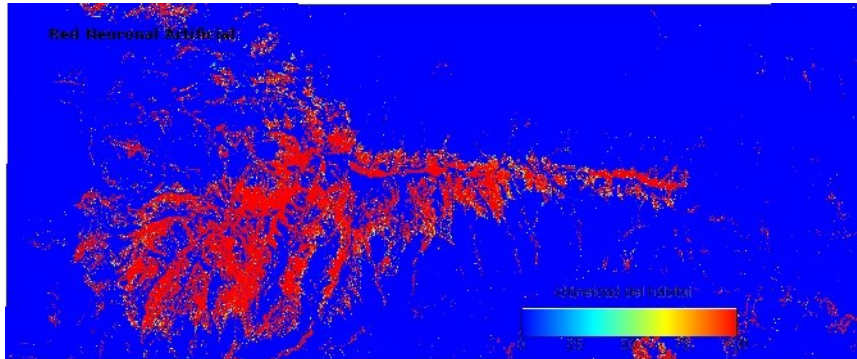
ANÁLISIS



UMBRAL DE CORTE

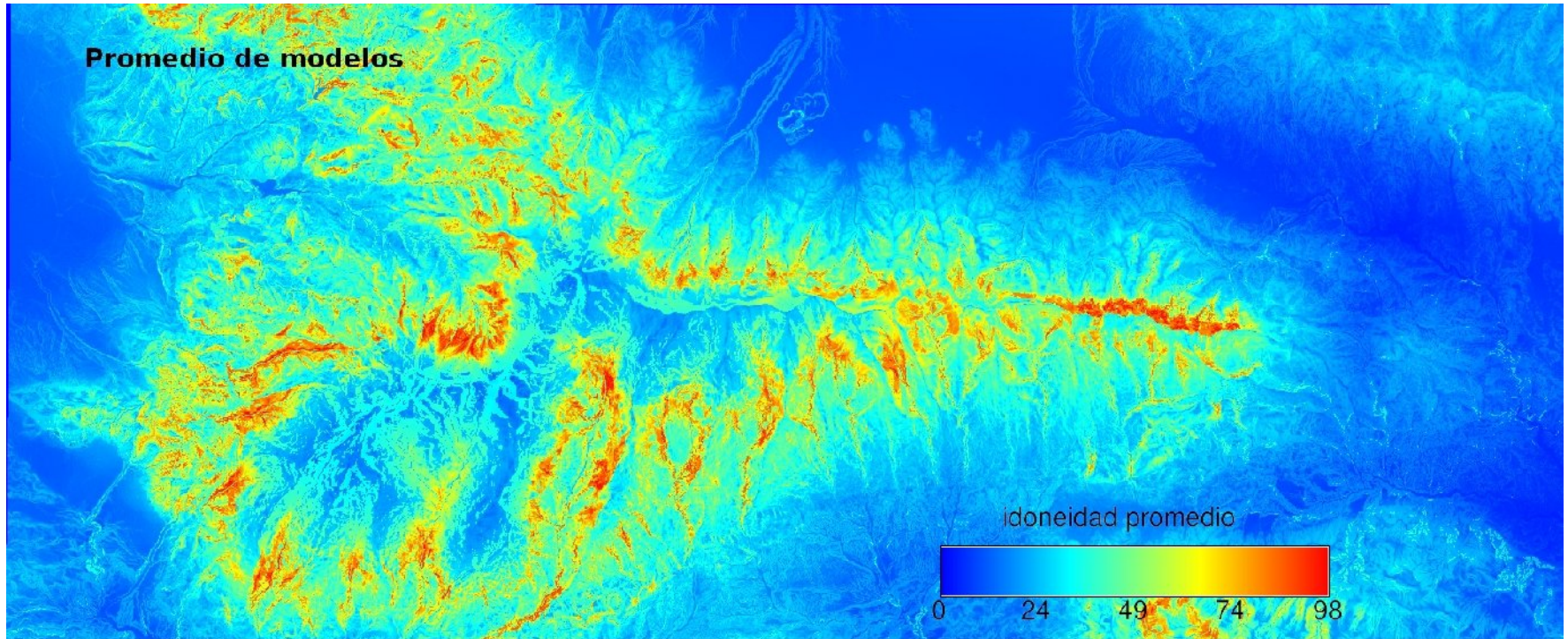


ENSAMBLADO DE MODELOS

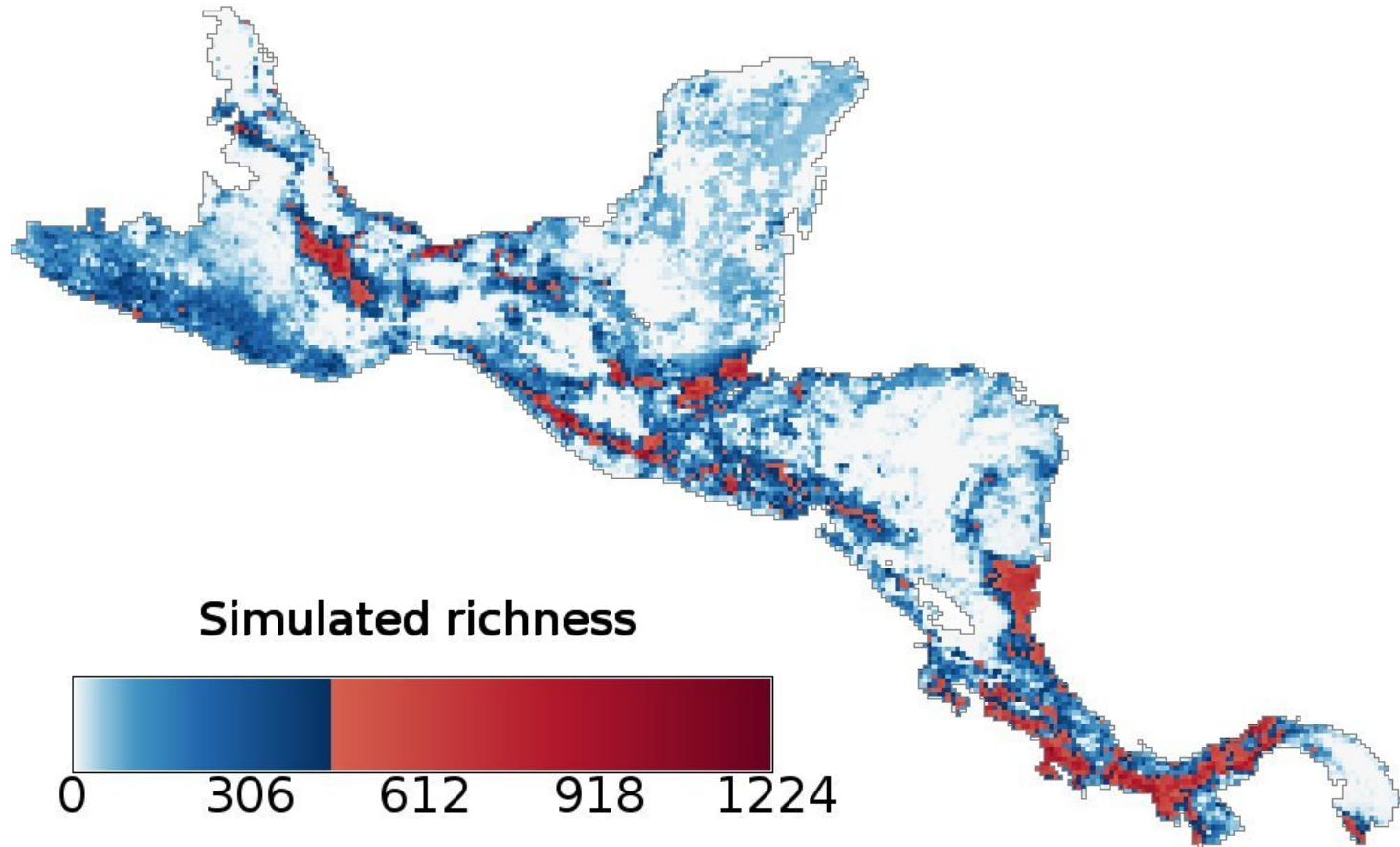


ENSAMBLADO DE MODELOS

PROMEDIO DE LOS MODELOS ANTERIORES

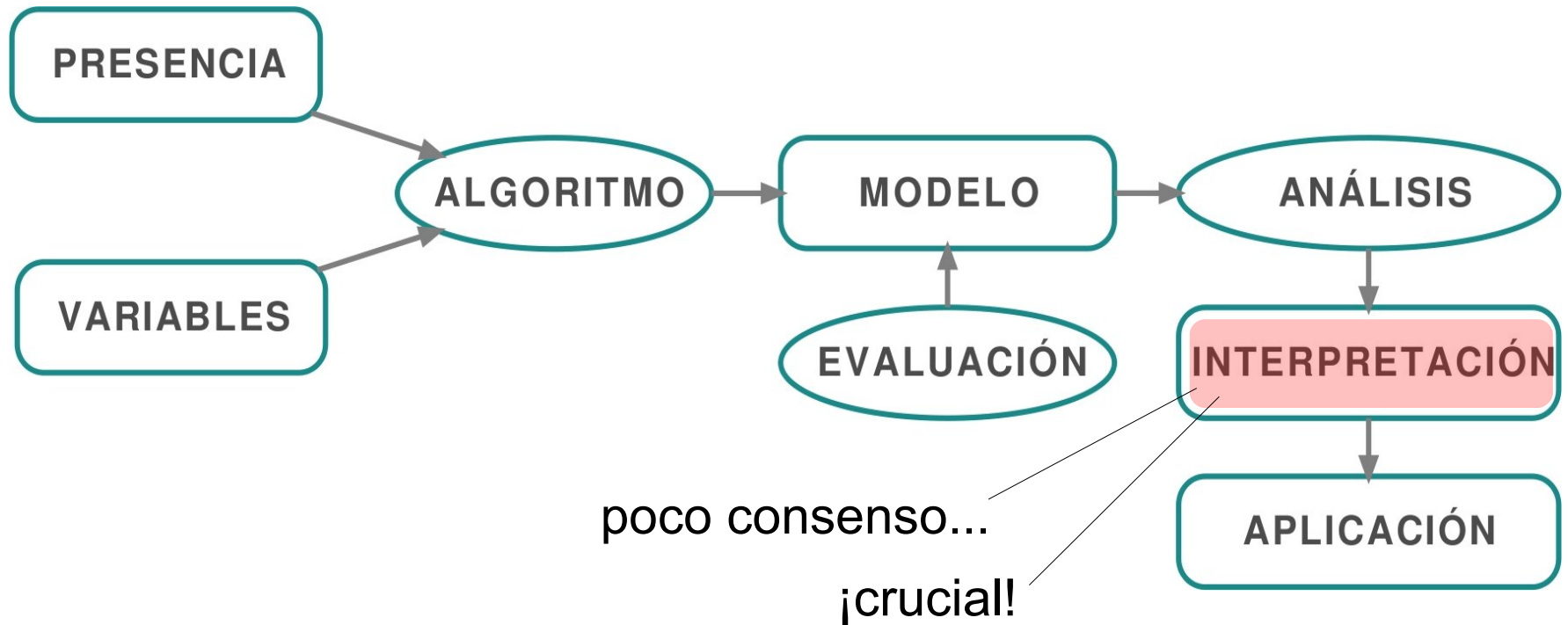


MODELOS DE RIQUEZA



Benito et al. 2013 MEE

INTERPRETACIÓN



INTERPRETACIÓN

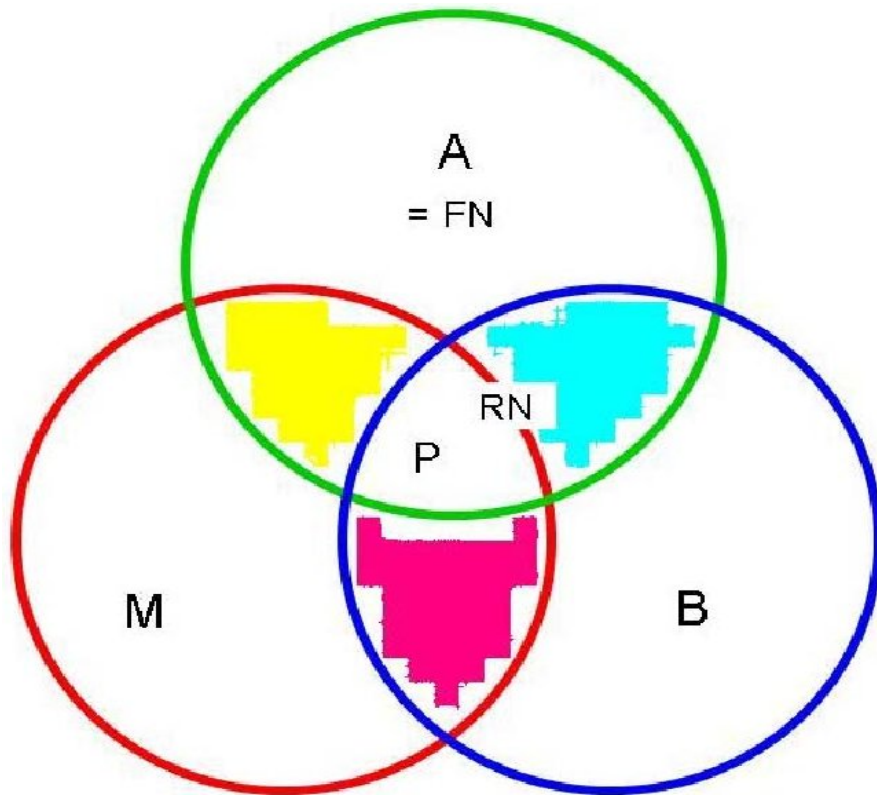


Diagrama BAM

- **A** - ambiente abiótico (*hábitat óptimo*)
- B - ambiente biótico
- M - hábitat accesible
- **RN** - *nicho realizado*
- **P** - *presencia*

Soberón 2005 Biodiversity Informatics

INTERPRETACIÓN

Dependerá del objetivo, datos de partida (presencias y variables) y diseño del modelo.

Según Soberón (2005):

*“Whether the result is interpreted as the species' distribution, the spatial extent of its fundamental niche, or some other phraseology, these algorithms **only find regions that 'resemble' those where occurrence points are located.**”*

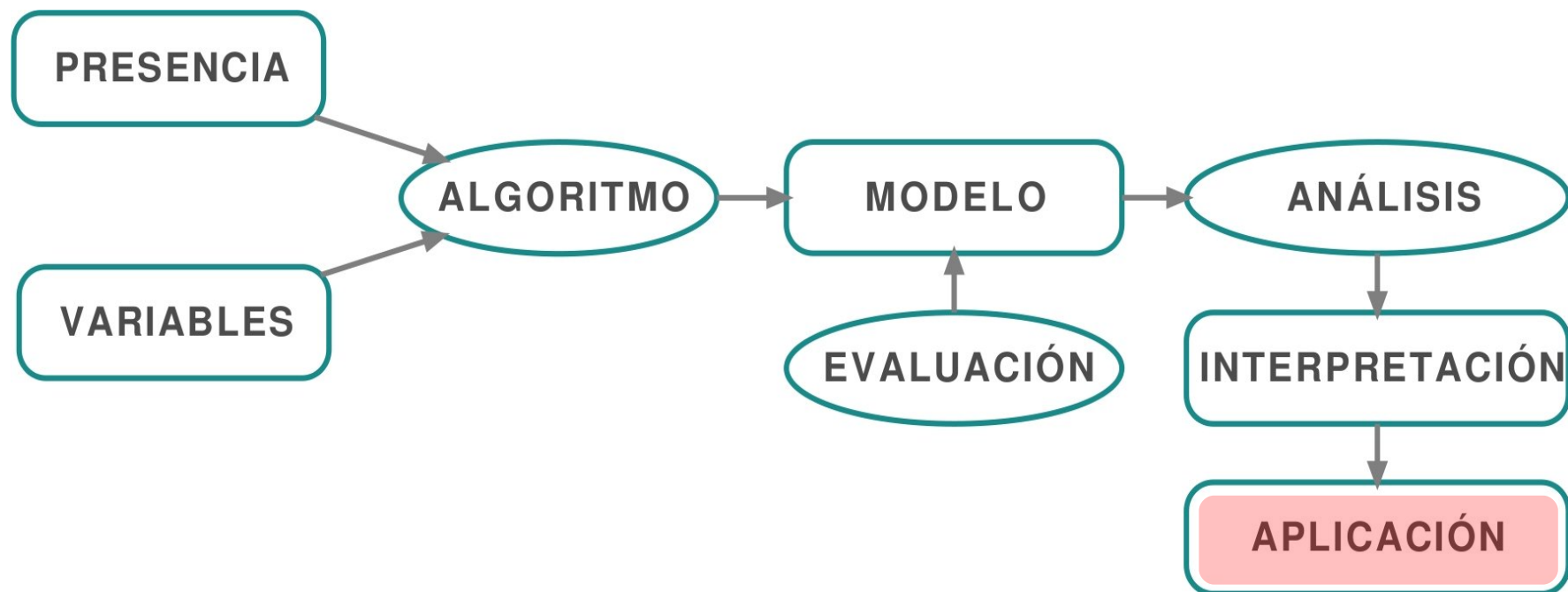
INTERPRETACIÓN

Otra cuestión importante está en la naturaleza biológica de la observación:

- ¿Se reproduce la especie en ese lugar?
- ¿Se alimenta la especie en ese lugar?
- ¿Busca la especie refugio u otros recursos en ese lugar?

Todas estas cuestiones nos ayudarán a contestar a la pregunta **¿QUÉ ESTAMOS MODELANDO?**

INTERPRETACIÓN



APLICACIONES DE LOS MODELOS DE DISTRIBUCIÓN DE ESPECIES

CARTOGRAFÍA DE ESPECIES O COMUNIDADES

Box et al. (1981) Predicting physiognomic vegetation types with climate variables.

Vegetatio 45, 127-139

- **Objetivo:** modelo terrestre de vegetación
- **Métodos:** envueltas bioclimáticas
- **Resultados:** 85 % acierto en tipos de plantas y 50 % en estructura de la vegetación
- **Conclusiones:** Un modelo muy simple y general puede predecir características de la vegetación mundial con una precisión aceptable

Box *et al.* (1981)

Modelo de envueltas bioclimáticas:

	TMAX	TMIN	DTY	PRCP	MI	PMAX	PMIN	PMTMAX
Tropical rainforest								
Max.	30.0	28.0	12.0	*****	****	*****	****	*****
Min.	20.0	18.0	0.0	1 400.	1.00	150.	5.	10.
Raingreen monsoon forest								
Max.	35.0	28.0	15.0	*****	3.00	*****	30.	*****
Min.	23.0	14.0	0.0	800.	0.90	100.	0.	20.
Summergreen broad-leaved trees								
Max.	32.0	10.0	50.0	*****	****	*****	****	*****
Min.	15.0	-20.0	8.0	300.	0.80	40.	2.	30.
Short-needed boreal/ montane trees								
Max.	22.0	3.0	60.0	*****	****	*****	****	*****
Min.	10.0	-25.0	10.0	100.	0.60	25.	0.	10.

Box *et al.* (1981)



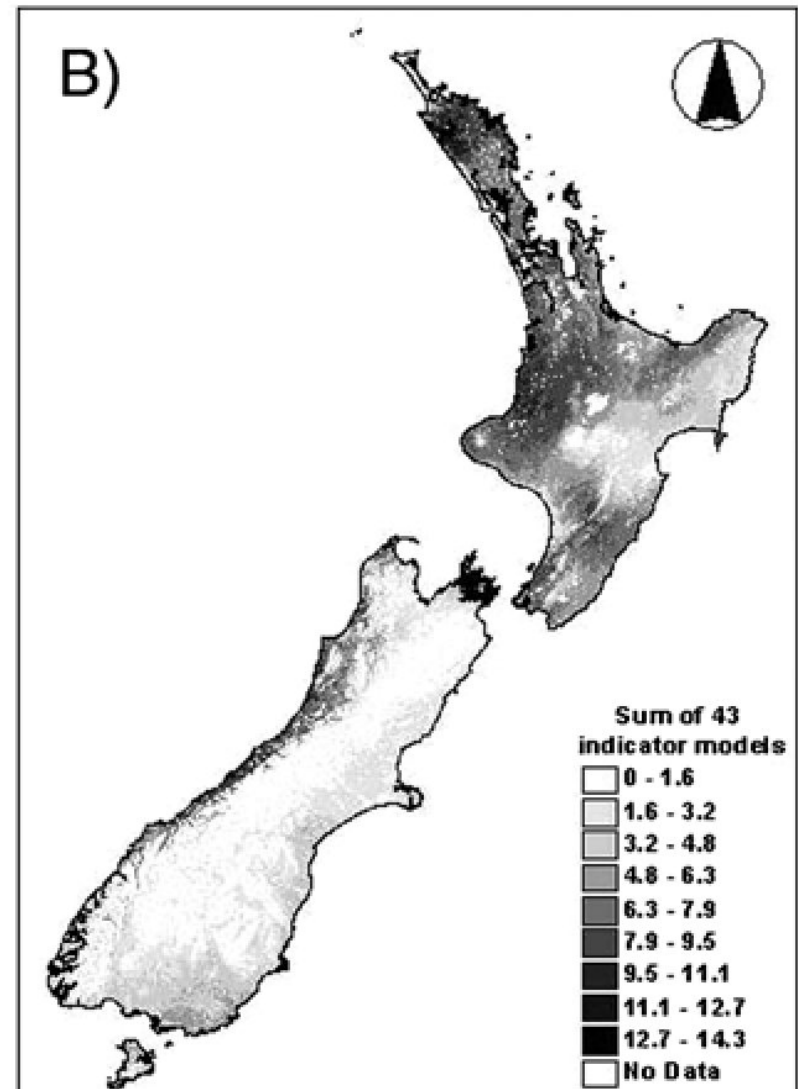
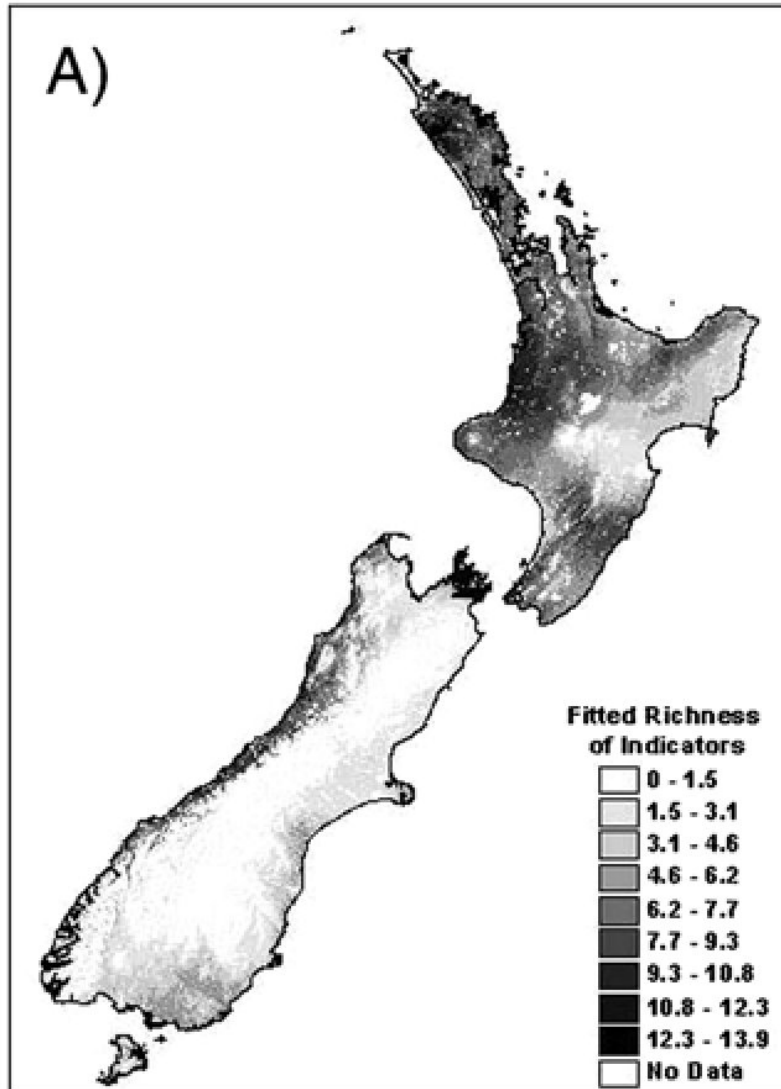
- OXAV (OXAV) = Tropical Rainforest Trees (lowland, montane, and cloud-belt subtypes)
- TX (TX) = Warm-Temperate Evergreen Trees
- Y (Y) = Mediterranean Evergreen Trees
- i°* (i°*) = Temperate Broad-Leaved Rainforest Trees
- i* (i*) = Subpolar Broad-Leaved Rainforest Trees (e.g. *Nothofagus* spp.)
- i (=) (=) = Tropical Evergreen Woodland Dwarf-Trees (e.g. 'campos cerrados')

MODELADO DE RIQUEZA

Lehmann *et al.* 2002. Assessing New Zealand fern diversity from spatial predictions of species assemblages. *Biodiversity and Conservation* 11, 2217-2238

- **Objetivo:** Modelo espacial de diversidad de arbustos
- **Métodos:** Suma de MDE comparada con predicción directa de riqueza
- **Resultados:** Alta congruencia entre modelos
- **Conclusiones:** La suma de MDE permite identificar hotspots y establecer objetivos para conservación y restauración

Lehmann *et al.* 2002

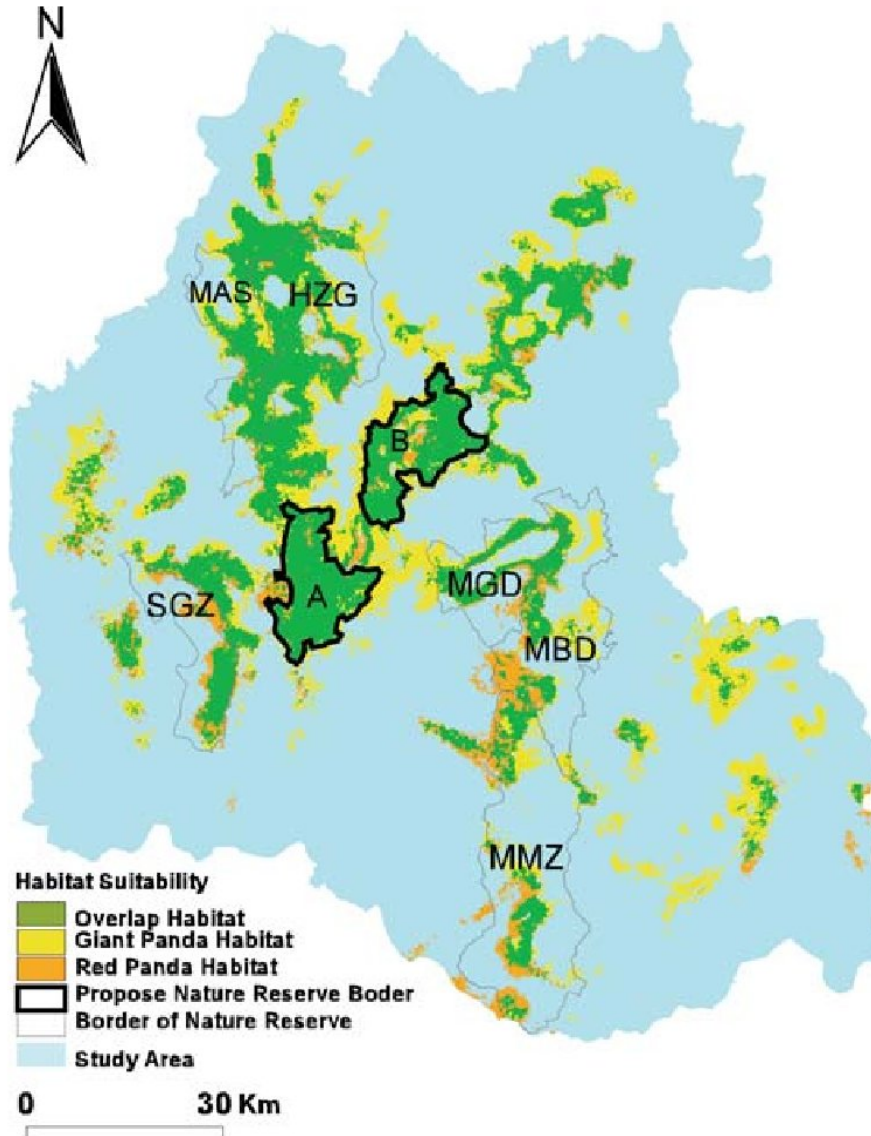


PLANIFICACIÓN DE CONSERVACIÓN

Dunwu *et al.* 2002. Ecological niche modeling of the sympatric giant and red pandas on a mountain-range scale. *Biodiversity and Conservation* 18, 2127-2141

- **Objetivo:** Recomendar áreas de reservas para pandas rojo y gigante
- **Métodos:** Solapamiento de modelos
- **Resultados:** Reservas actuales están mal diseñadas
- **Conclusiones:** Los MDE ayudan a un diseño de reservas más coherente

Dunwu et al. 2002

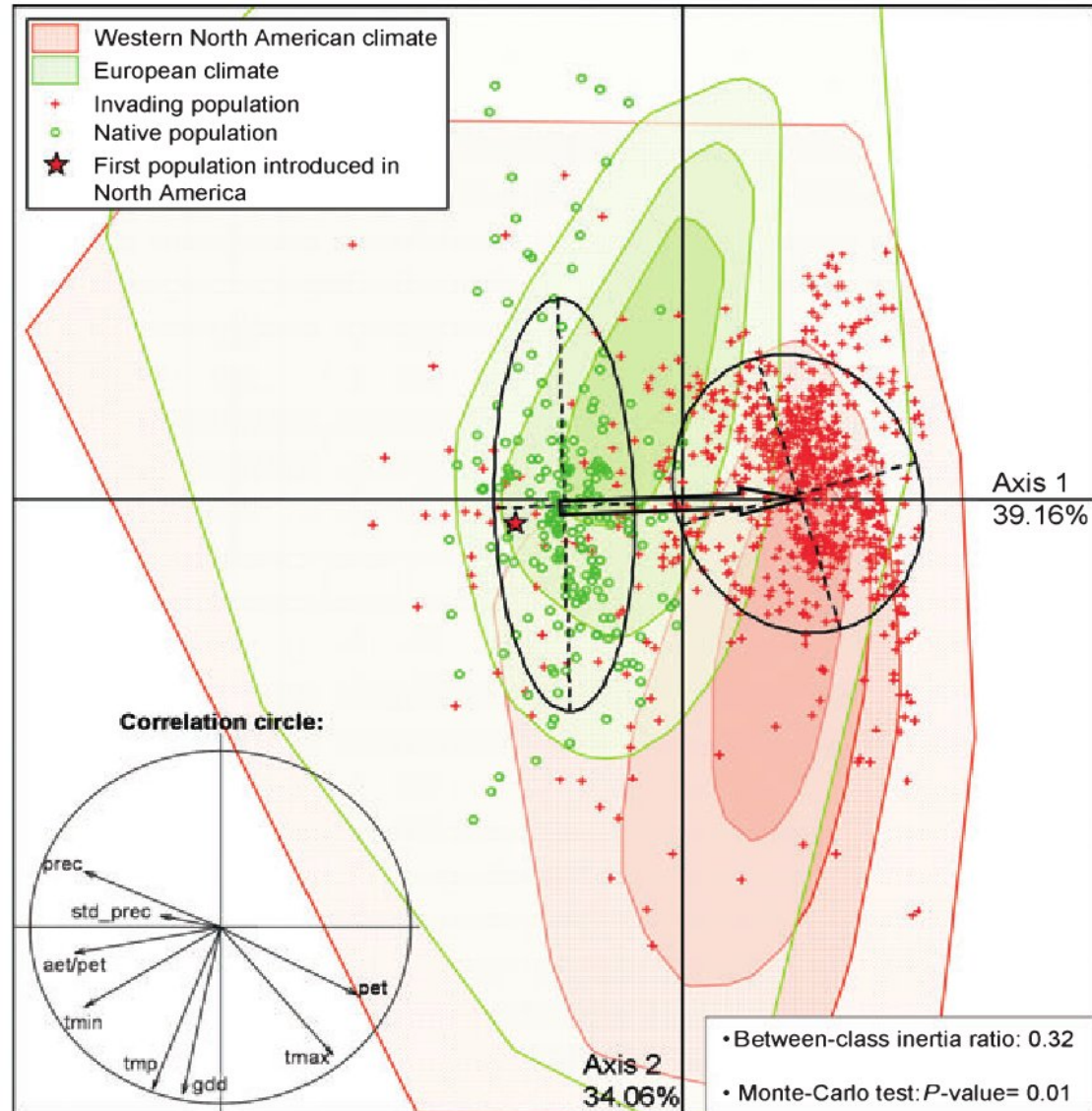


ANÁLISIS DE INVASIONES BIOLÓGICAS

Broenninmann *et al.* 2007. Evidence of climatic niche shift during biological invasions. *Ecology Letters* 10, 701-709

- **Objetivo:** Comprobar cambios en nicho ecológico durante invasión biológica
- **Métodos:** MDE de especie invasora en rango nativo en invadido
- **Resultados:** La especie cambia sus requerimientos ecológicos durante la invasión
- **Conclusiones:** Los MDE predicen áreas de introducción de especies invasoras, pero no rangos de invasión

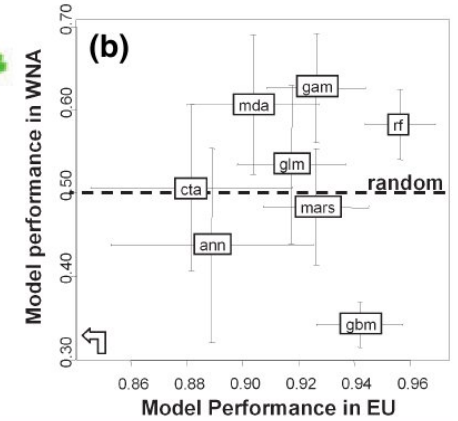
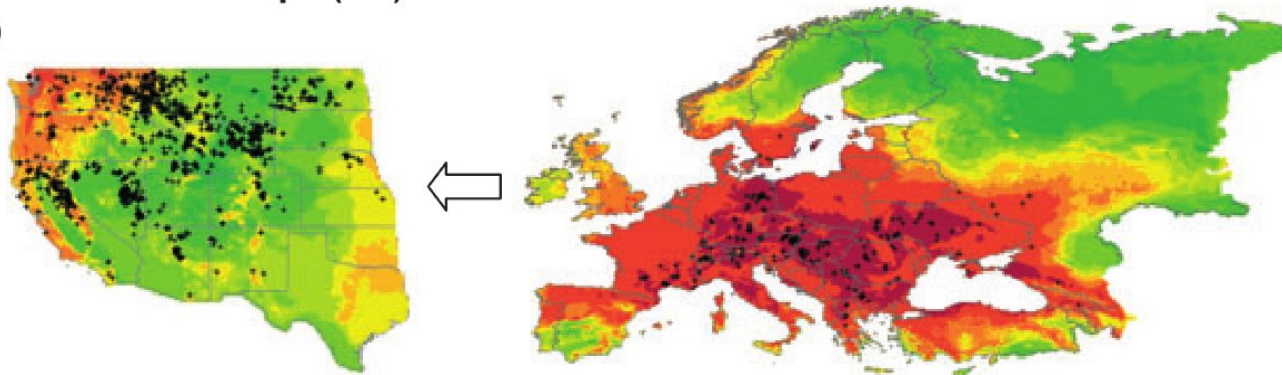
Broenninmann *et al.* 2007



Broenninmann *et al.* 2007

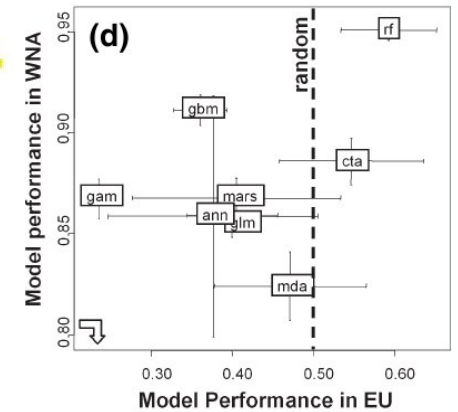
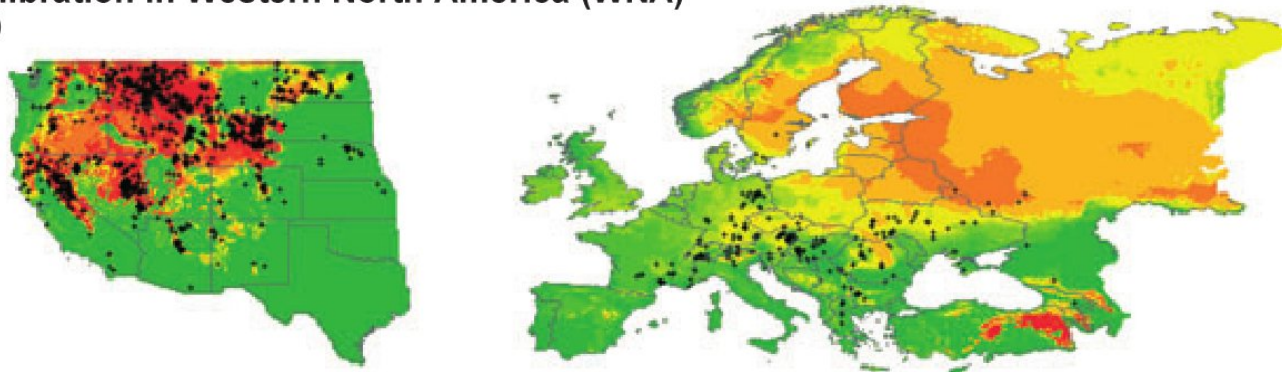
Calibration in Europe (EU)

(a)



Calibration in Western North America (WNA)

(c)

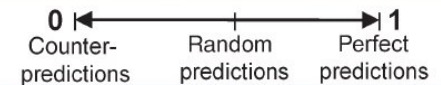


Occurrences of *Centaurea maculosa* : +

Predicted climatic suitability :



Model performance :

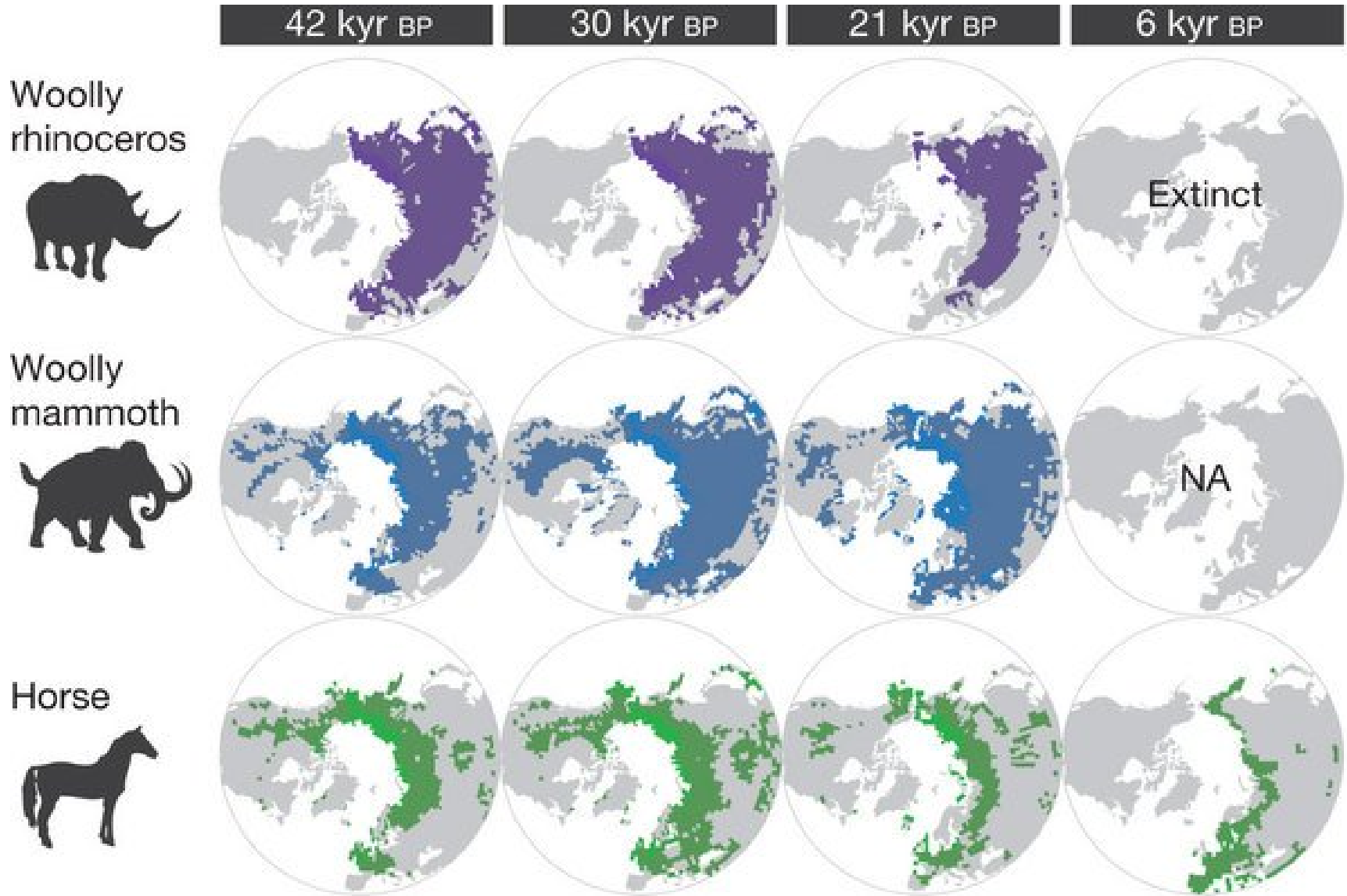


ANÁLISIS DE PALEODISTRIBUCIONES

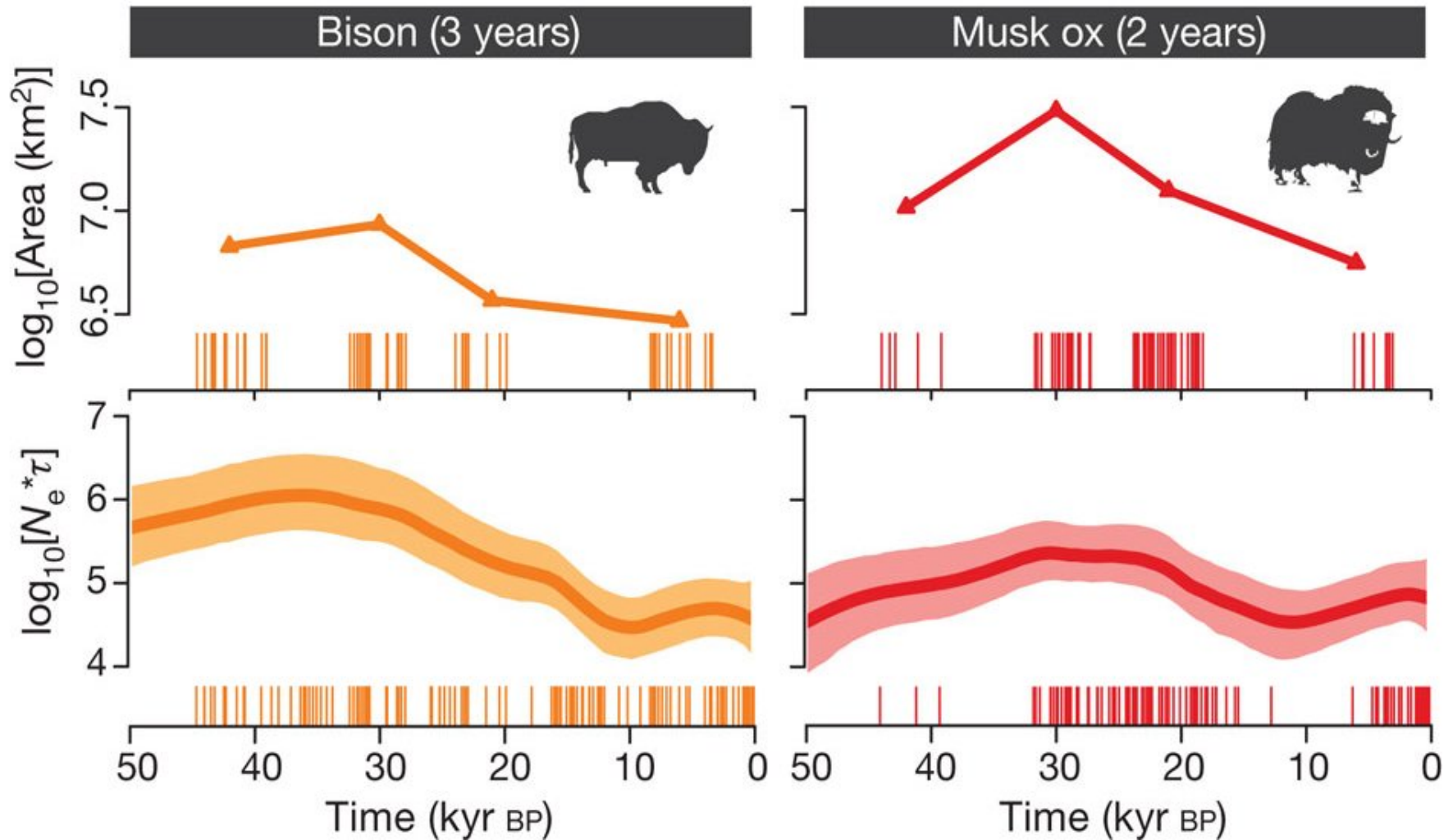
Lorenzen *et al.* 2011. Species-specific responses of Late Quaternary megafauna to climate and humans. *Nature* 479, 359-364

- **Objetivo:** Aclarar efecto del clima y el hombre en la distribución de la megafauna del Cuaternario
- **Métodos:** MDE, registro fósil, ADN
- **Resultados:** Especies presentan respuestas diferentes a cambio de clima y efecto humano
- **Conclusiones:** El cambio del clima explica algunas extinciones, pero otras no pueden entenderse sin la acción del hombre

Lorenzen et al. 2011



Lorenzen et al. 2011

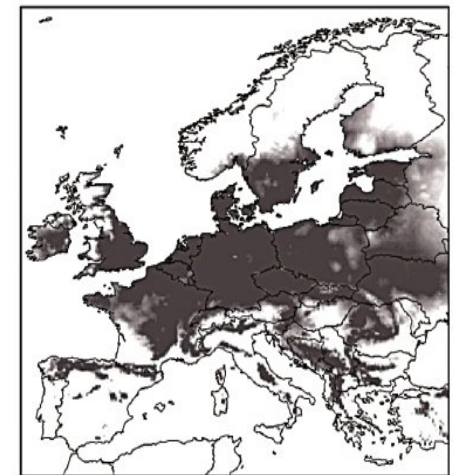
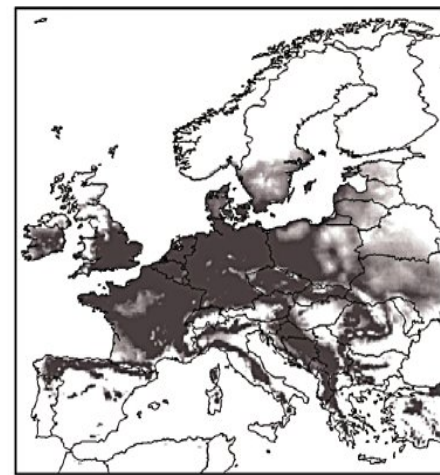
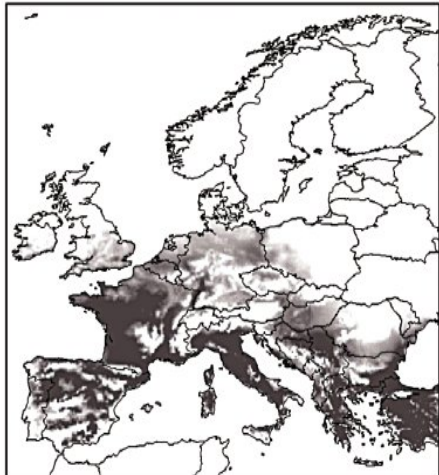
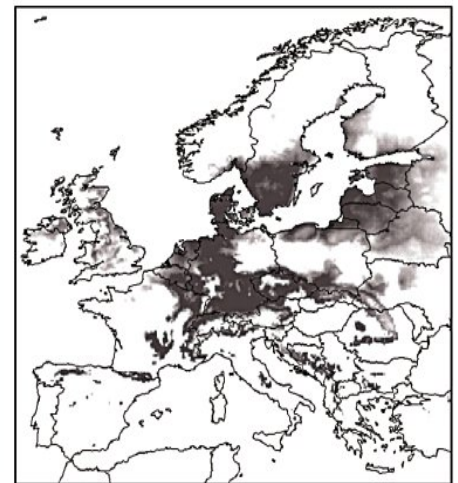
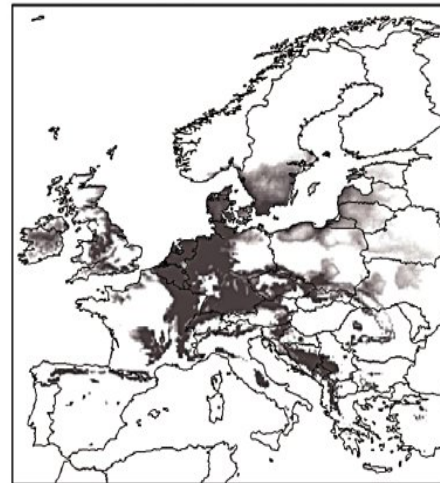
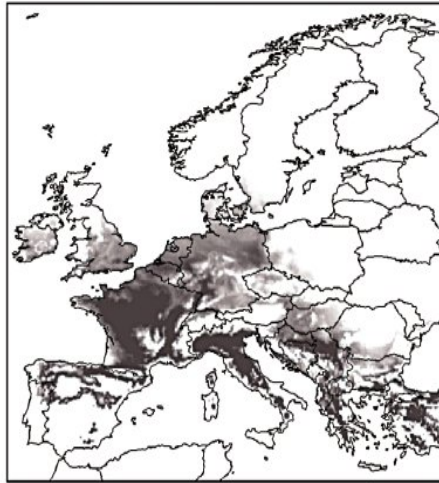
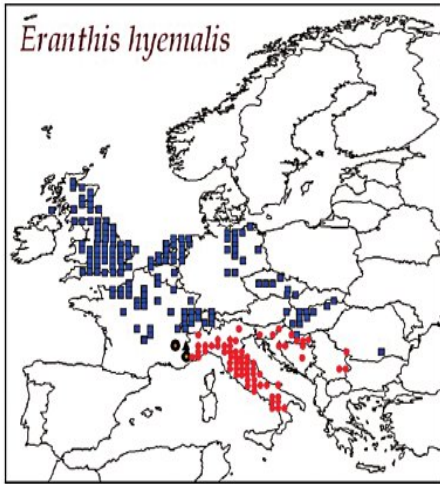


CAMBIO CLIMÁTICO Y MIGRACIÓN

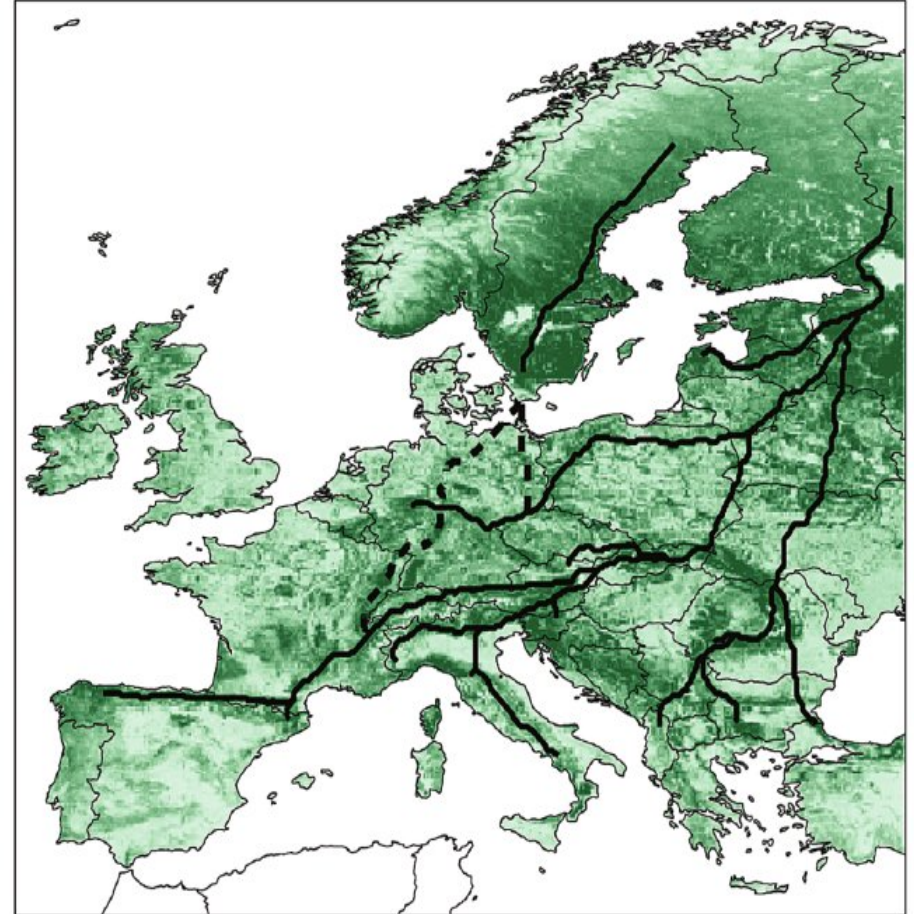
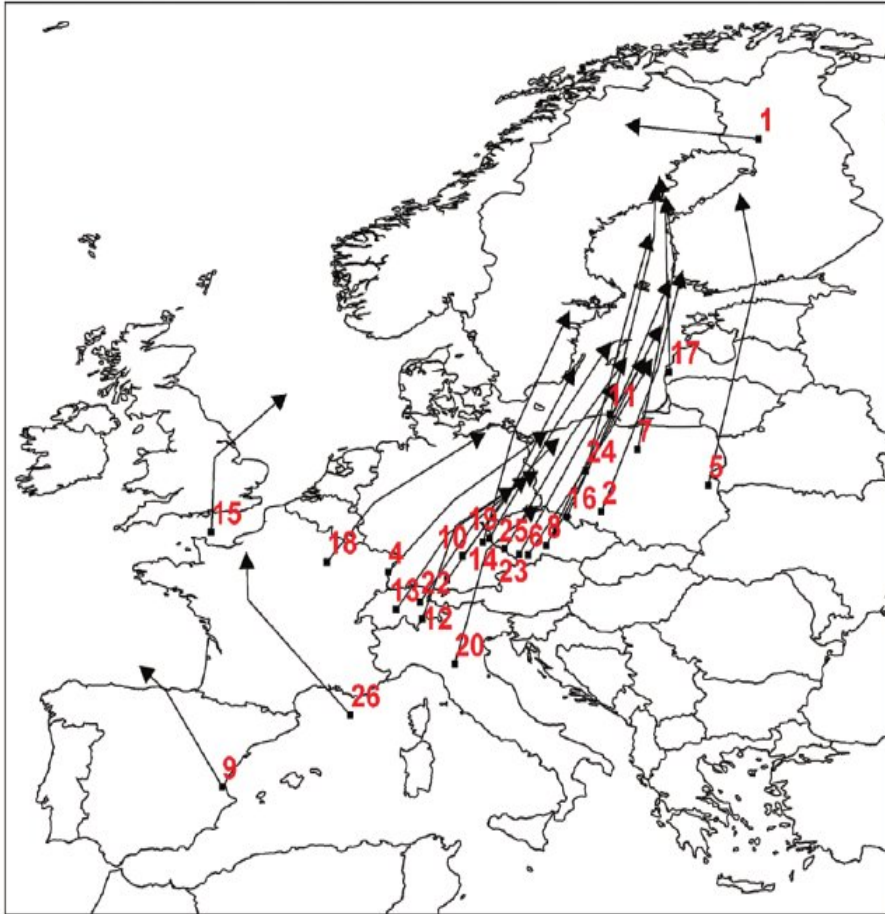
Skov y Svenning 2004. Potential impact of climatic change on the distribution of forest herbs in Europe. *Ecography* 27, 366-380

- **Objetivo:** Evaluar consecuencias del Cambio Climático en la distribución plantas europeas
- **Métodos:** Envueltas bioclimáticas difusas
- **Resultados:** Observan pérdidas moderadas de hábitat para todas las especies
- **Conclusiones:** La escasa dispersión natural de estas plantas va a impedirles migrar al ritmo exigido por el cambio climático

Skov y Svenning 2004



Skov y Svenning 2004



EN RESUMEN

- Cartografía de especies o comunidades
- Modelado de riqueza de especies
- Planificación de conservación
- Análisis de invasiones biológicas
- Análisis de paleodistribuciones
- Cambio climático y migración

EL ENTORNO DE TRABAJO

HERRAMIENTAS

PROGRAMACIÓN

- R – cran.r-project.org
- Rstudio <http://www.rstudio.com/>

MODELOS DE DISTRIBUCIÓN

- MaxEnt –
www.cs.princeton.edu/~schapire/maxent/
- Dismo - cran.r-project.org/.../dismo



- Código abierto, uso gratuito y alta potencia
- Un estándar para análisis estadístico en el ámbito académico
- Comunidad de desarrollo muy activa: evoluciona muy rápido!
- Posibilidades casi infinitas!
- Lenguaje de programación poco intuitivo



RGui

Archivo Editar Visualizar Misc Paquetes Ventanas Ayuda

R Console

```
R version 2.10.1 (2009-12-14)
Copyright (C) 2009 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R es un software libre y viene sin GARANTIA ALGUNA.
Usted puede redistribuirlo bajo ciertas circunstancias.
Escriba 'license()' o 'licence()' para detalles de distribucion.

R es un proyecto colaborativo con muchos contribuyentes.
Escriba 'contributors()' para obtener más información y
'citation()' para saber cómo citar R o paquetes de R en publicaciones.

Escriba 'demo()' para demostraciones, 'help()' para el sistema on-line de ayuda,
o 'help.start()' para abrir el sistema de ayuda HTML con su navegador.
Escriba 'q()' para salir de R.

> |
```




File Edit Code View Plots Session Project Build Tools Help

Workspace History Git

Values

Gctorture	FALSE
dir_trabajo	"/home/bias/Dropbox/DOCENCIA/CURSOS_MDE_GBIF/2014/sesiones/sesion_2/taller2"
lista_variables	character[38]
variables	RasterBrick[831972]

```
50 library(party) #ARBOLES DE CLASIFICACIÓN Y REGRESIÓN
51 library(QuantPsyc) #COEFICIENTES ESTANDARIZADOS DE MODELOS DE REGRESIÓN
52 library(HH) #VARIANCE INFLATION FACTOR
53
54
55 #####
56 #IMPORTA LAS VARIABLES PREDICTORAS
57 #####
58 #LISTADO DE VARIABLES
59 lista_variables <- list.files(path="variables",pattern="*.asc", full.names=TRUE)
60
61 #stack Y brick PREPARAN LAS VARIABLES EN UN UNICO OBJETO ESPACIAL
62 help(brick) #mapas en memoria, más rápido (pero usa mucha RAM)
63 help(stack) #mapas en disco, más lento (pero usa menos RAM)
64 variables <- brick(stack(lista_variables))
65
66 #MUESTRA LA ESTRUCTURA DEL OBJETO variables
67 str(variables)
68
69 #DIBUJA LAS VARIABLES PREDICTORAS
70 #png("resultados_R/1_variables.png", width=2000, height=2000, pointsize=20)
71 plot(variables)
72 plot(variables, maxnl=38)
73 #dev.off()
74
75
```

Console

```
empty

Attaching package: 'reshape'

The following object(s) are masked from 'package:plyr':
  rename, round_any

Loading required package: colorspace

Attaching package: 'HH'

The following object(s) are masked from 'package:boot':
  logit

> lista_variables <- list.files(path="variables",pattern="*.asc", full.names=TRUE)
> #stack Y brick PREPARAN LAS VARIABLES EN UN UNICO OBJETO ESPACIAL
> help(brick) #mapas en memoria, más rápido (pero usa mucha RAM)
> help(stack) #mapas en disco, más lento (pero usa menos RAM)
> variables <- brick(stack(lista_variables))
rgdal: version: 0.8-11, (SVN revision 479M)
Geospatial Data Abstraction Library extensions to R successfully loaded
Loaded GDAL runtime: GDAL 1.7.3, released 2010/11/10
Path to GDAL shared files: /usr/share/gdal17
GDAL does not use iconv for recoding strings.
Loaded PROJ.4 runtime: Rel. 4.8.0, 6 March 2012, [PJ_VERSION: 480]
Path to PROJ.4 shared files: (autodetected)
> plot(variables)
>
```

Files Plots Packages Help

Zoom Export Clear All

Dismo + Raster

- Dos librerías de R, una para hacer modelos de distribución y otra para procesar mapas ráster
- Desarrollada por investigadores punteros en modelos de distribución
- Combinación muy flexible de herramientas de modelización y GIS

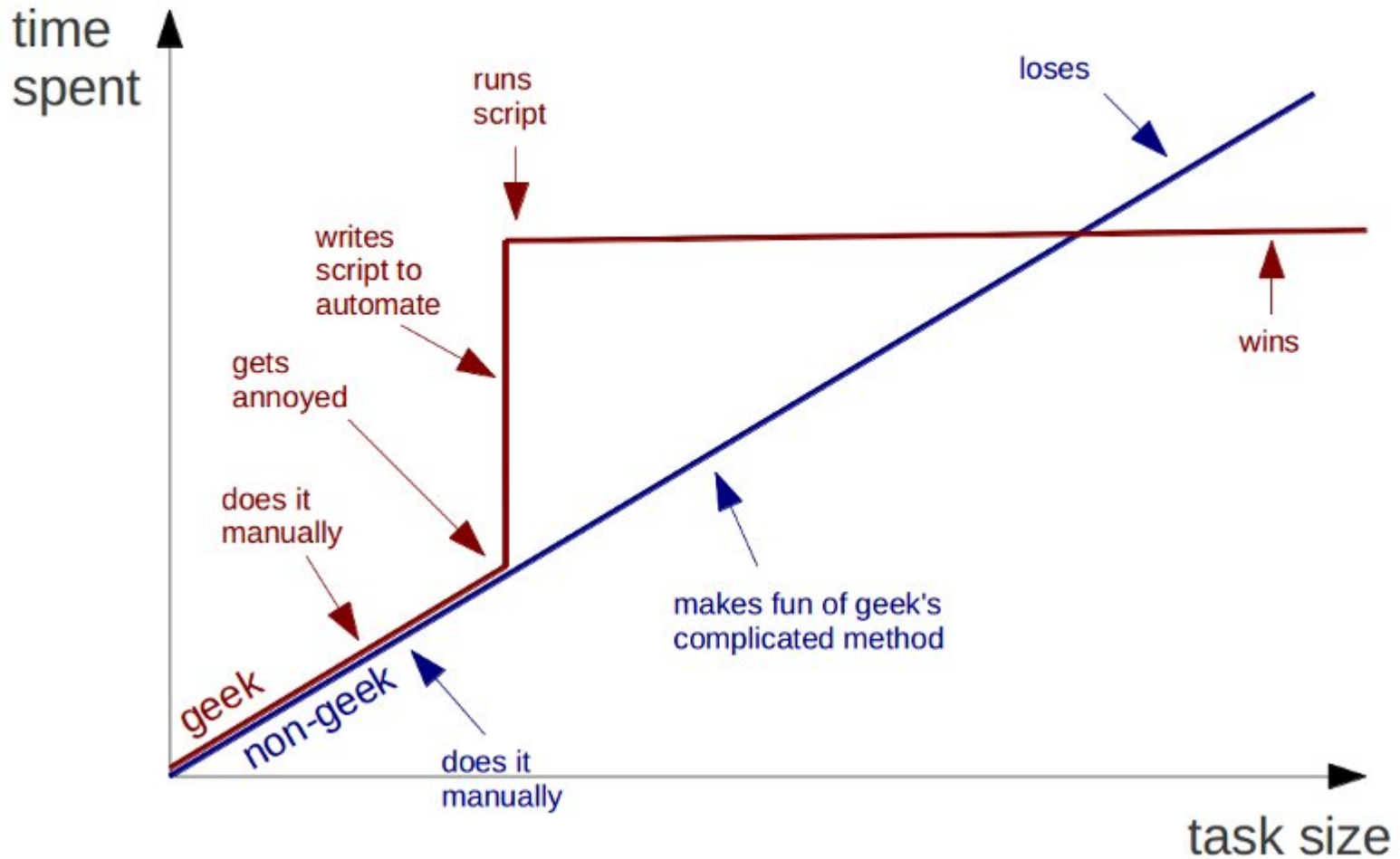
MaxEnt

- Método general para hacer predicciones a partir de información incompleta
- Permite trabajar con un número bajo de registros de presencia
- Multiplataforma (requiere Java), código cerrado pero uso gratuito
- Ofrece muy buenos resultados
- Su interfaz gráfico es muy sencillo



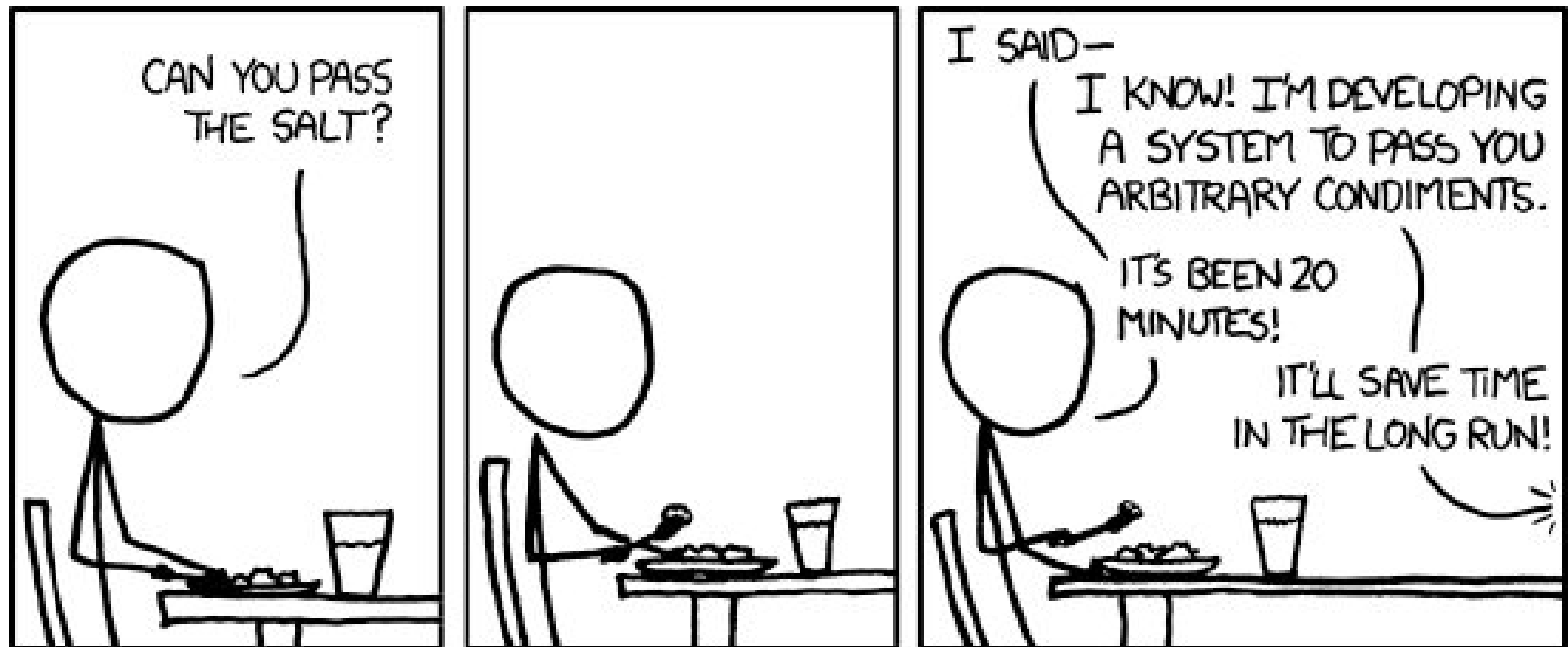
¿POR QUÉ ?

Geeks and repetitive tasks



¿POR QUÉ ?

Pero sin pasarse...



¿POR QUÉ ?

- ¿Tendrás que repetir la tarea?
 - Sí → **R**
 - No → ArgGIS + Statistica + Excel + Maxent + Access + ...
- ¿Quieres compartir tu flujo de trabajo?
- ¿Tienes que explicar paso por paso lo que has hecho?
- ¿Qué ocurre cuando haces el proceso manualmente y descubres un error al final?

NO TODO EL MONTE ES ORÉGANO

- Tendrás que LEER ayudas y manuales.
- Tendrás que ENTENDER lo que haces.
- Tendrás que APRENDER un nuevo lenguaje.
- Todo esto llevará TIEMPO.
- Al final, te convertirás en un NERD.
- Tus compañeros de trabajo te mirarán mal porque serás MUCHO MÁS PRODUCTIVO.

PRIMEROS PASOS CON R y Rstudio

- Tools>Global Options> General> desmarca “Restore .Rdata
- Tools>Global options>Pane Layout, y marcamos “Files” en el panel de arriba a la derecha
- Tools> Global options> Packages> CRAN mirror> Ireland o France o Madrid (elige uno al azar)
- Menú Session> Set Working Directory> c:/taller1/intro_R/datos
- Desde el menú 'File' abre el fichero **.../taller1/intro_R/introduccion.R**

PREPARACIÓN DE LAS VARIABLES PREDICTIVAS Y LOS REGISTROS DE PRESENCIA

EXTENSIÓN, RESOLUCIÓN Y MÁSCARA

FUENTES MEDIO TERRESTRE

Tipo	Dataset	Sitio web
Clima	WORLDCLIM	http://www.worldclim.org/
Topografía	SRTM	http://www2.jpl.nasa.gov/srtm/
NDVI	GIMMS	http://glcf.umd.edu/data/gimms/
Vegetación	MODIS VCF	http://glcf.umd.edu/data/vcf/
Huella humana	Human Footprint	http://sedac.ciesin.columbia.edu
Usos del suelo	GLOBCOVER	http://www.edenextdata.com

FUENTES MEDIO MARINO

- **Bio-ORACLE** (<http://www.oracle.ugent.be/>)
 - 23 variables
 - Resolución: 9.2 km
 - Referencia: Tyberghein et al. (2012) DOI: 10.1111/j.1466-8238.2011.00656.x
- **MARSPEC** (<http://www.marspec.org/>)
 - 17 variables bioclimáticas y topográficas
 - Resolución: 1 km
 - Referencia: Sbrocco and Barber (2013) DOI: 10.1890/12-1358.1

REQUISITOS TÉCNICOS

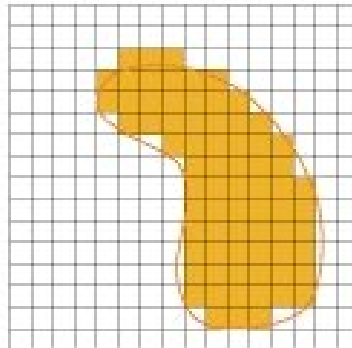
- Matrices idénticas
 - Igual resolución
 - Igual extensión
 - Iguales celdas con valores válidos
- Baja colinealidad

RESOLUCIÓN

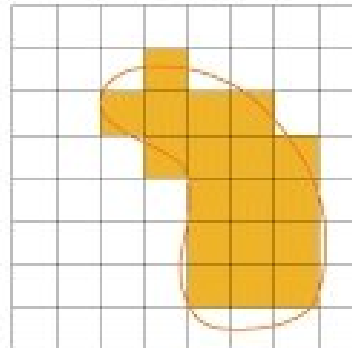
71 m²
polygon



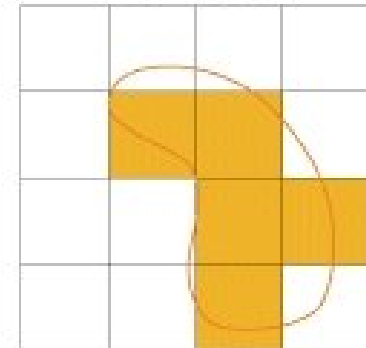
73 m²
1 m cell
16 x 16 cells



72 m²
2 m cell
8 x 8 cells



80 m²
4 m cell
4 x 4 cells

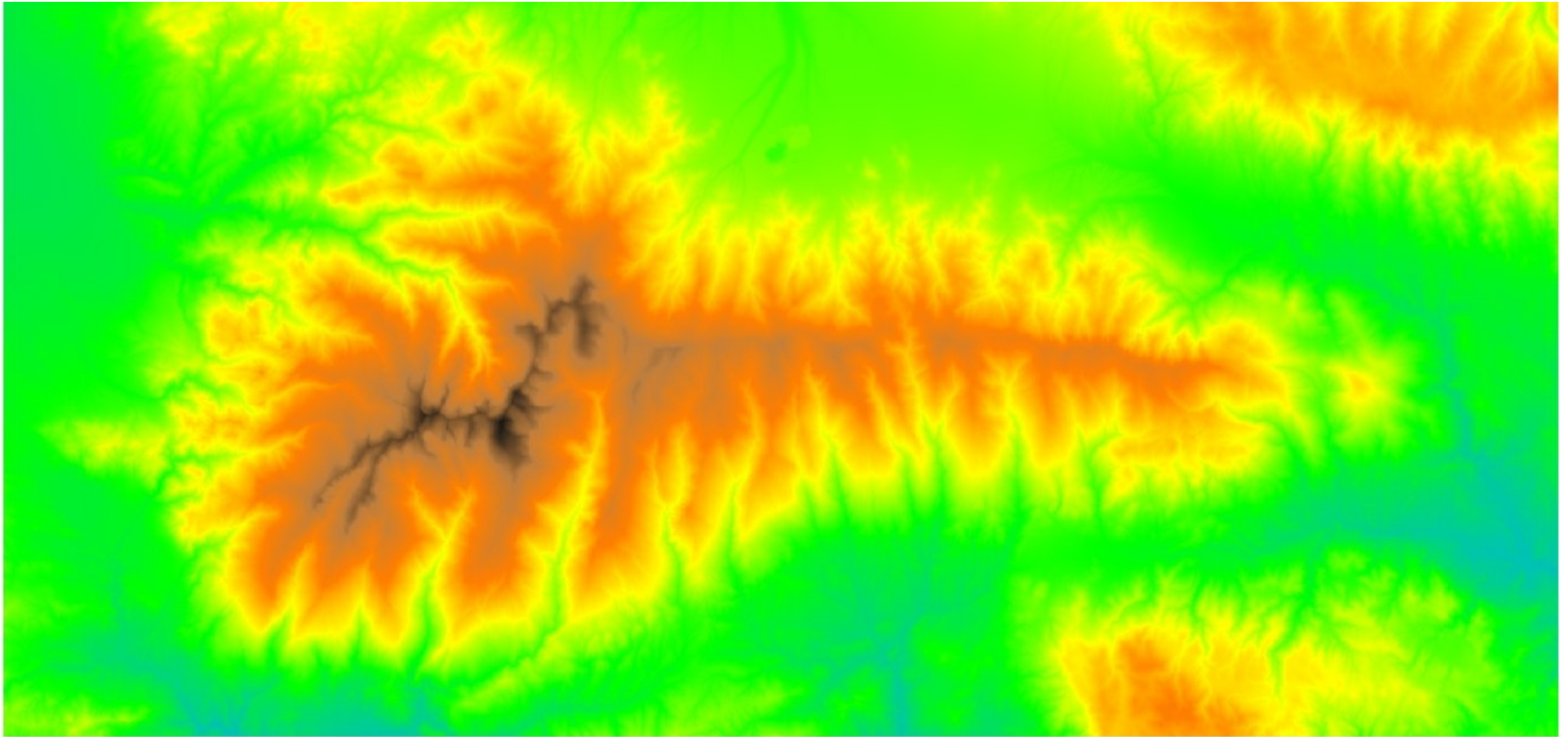


- Smaller cell size
- Higher resolution
- Higher feature spatial accuracy
- Slower display
- Slower processing
- Larger file size

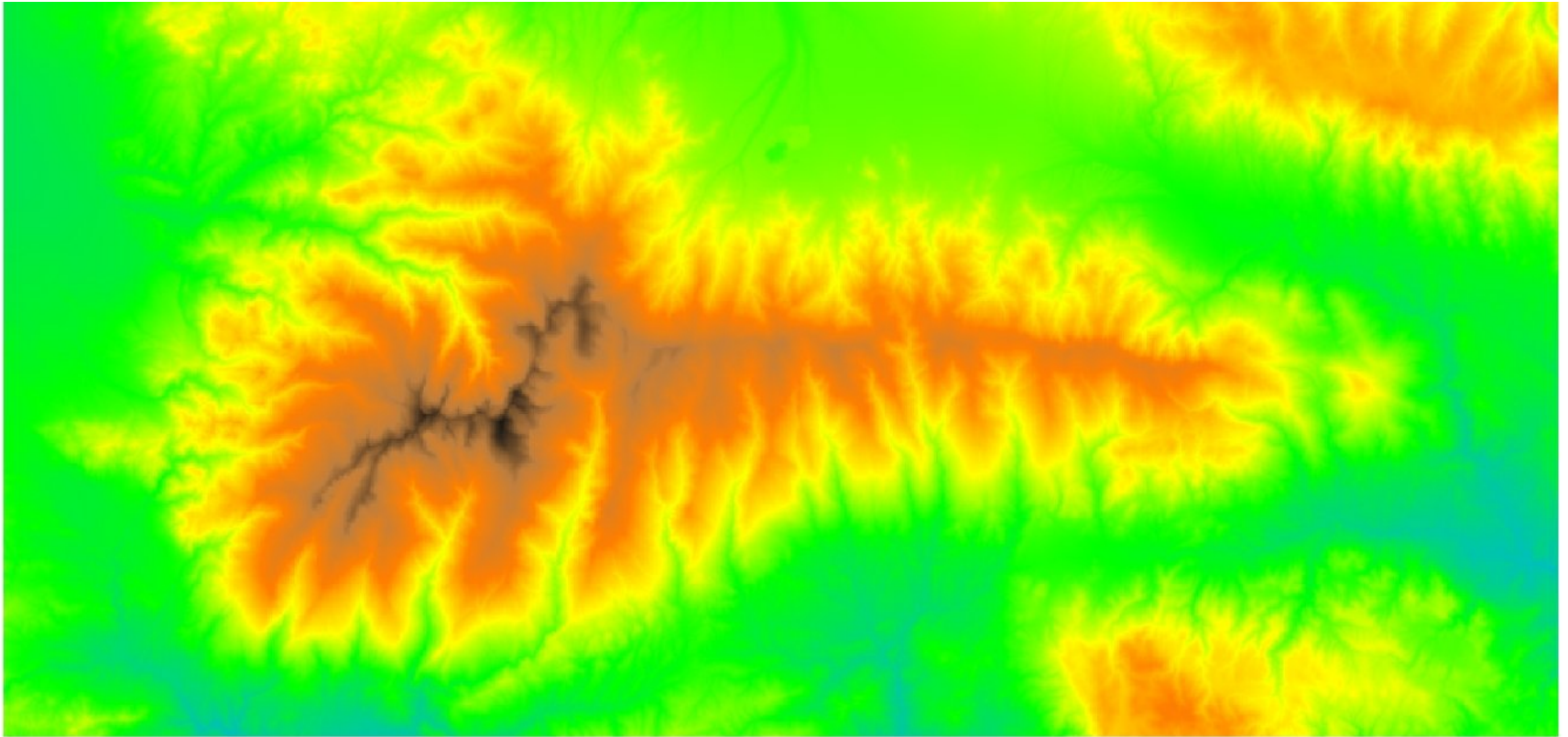
- Larger cell size
- Lower resolution
- Lower feature spatial accuracy
- Faster display
- Faster processing
- Smaller file size

Fuente: webhelp.esri.com

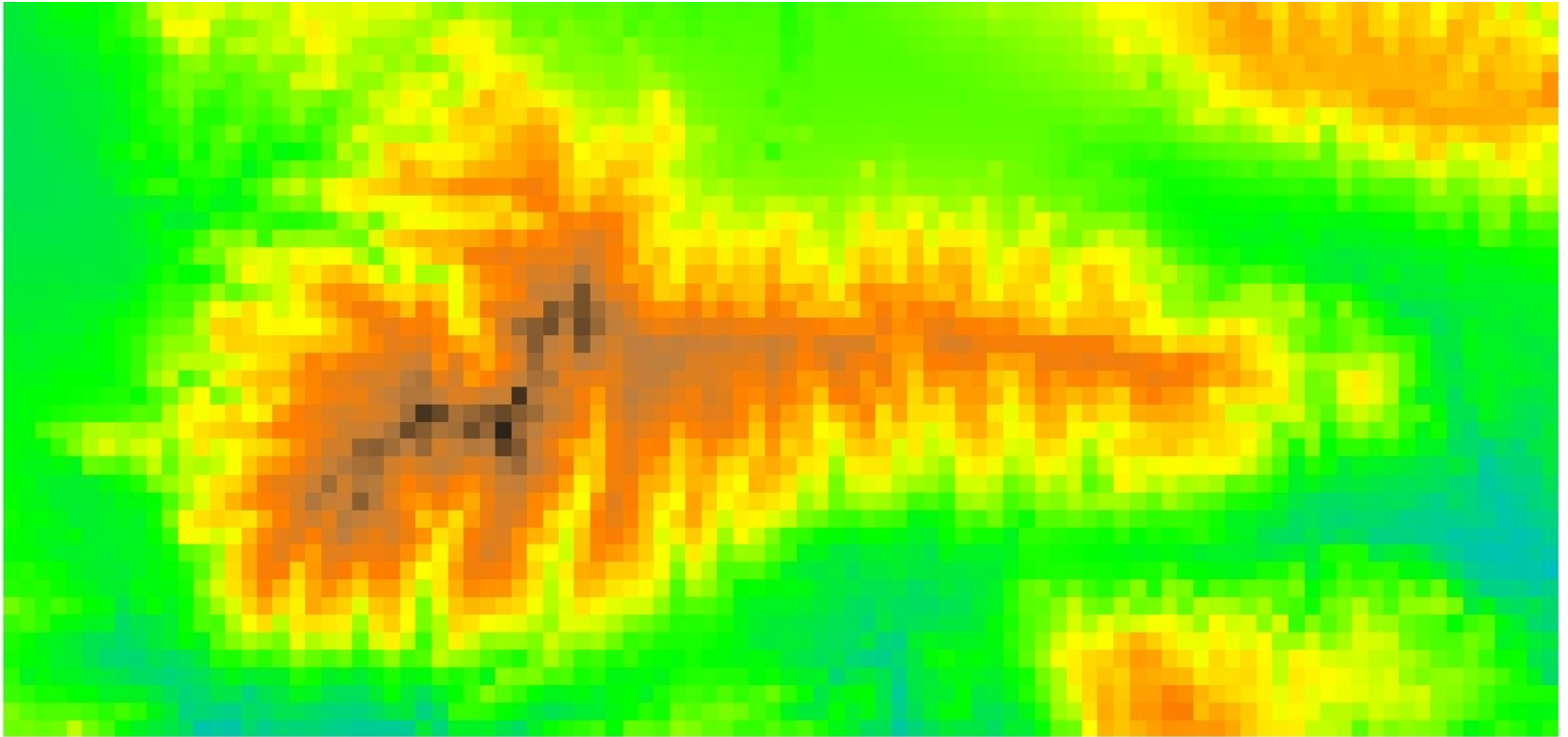
RESOLUCIÓN (10 m)



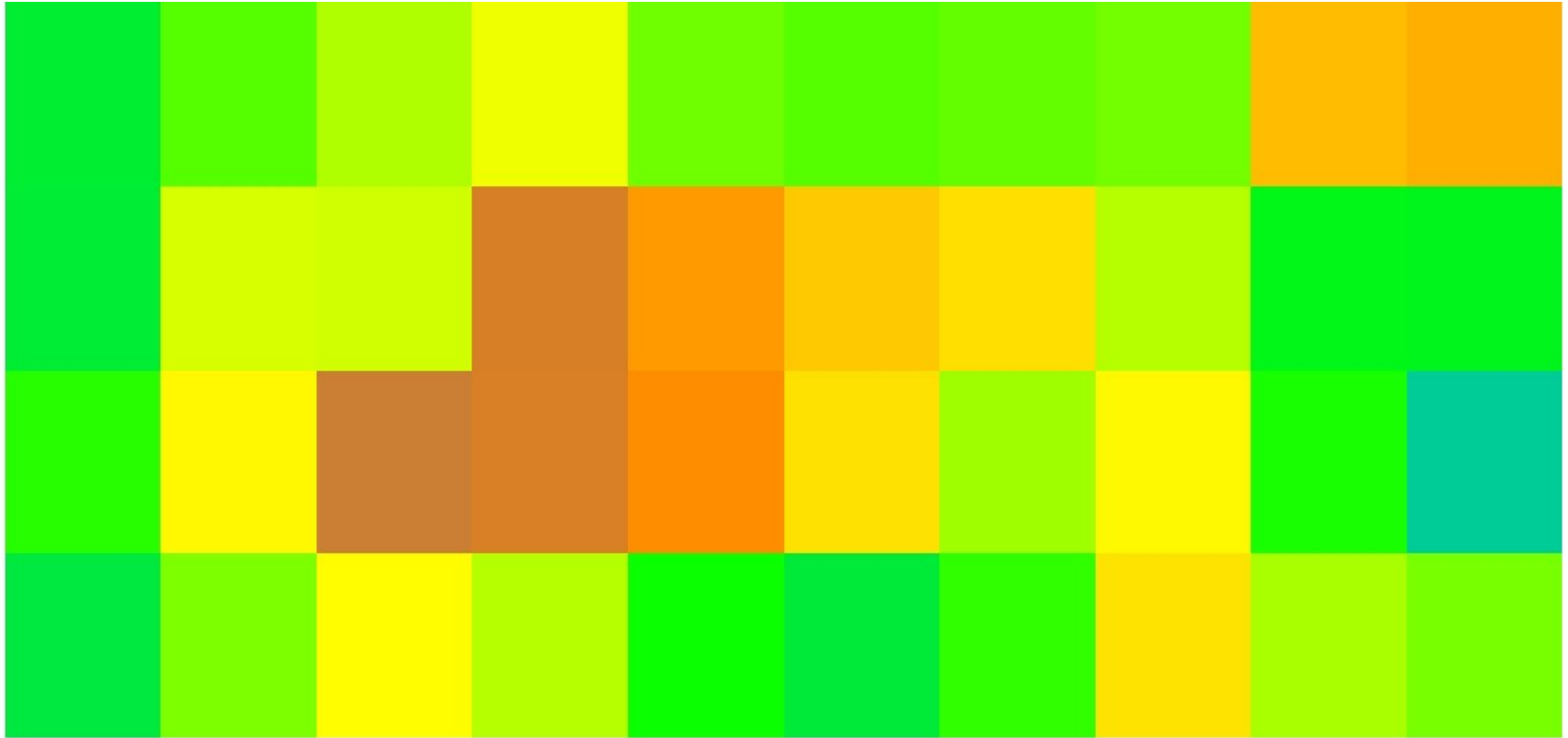
RESOLUCIÓN (100 m)



RESOLUCIÓN (1000 m)

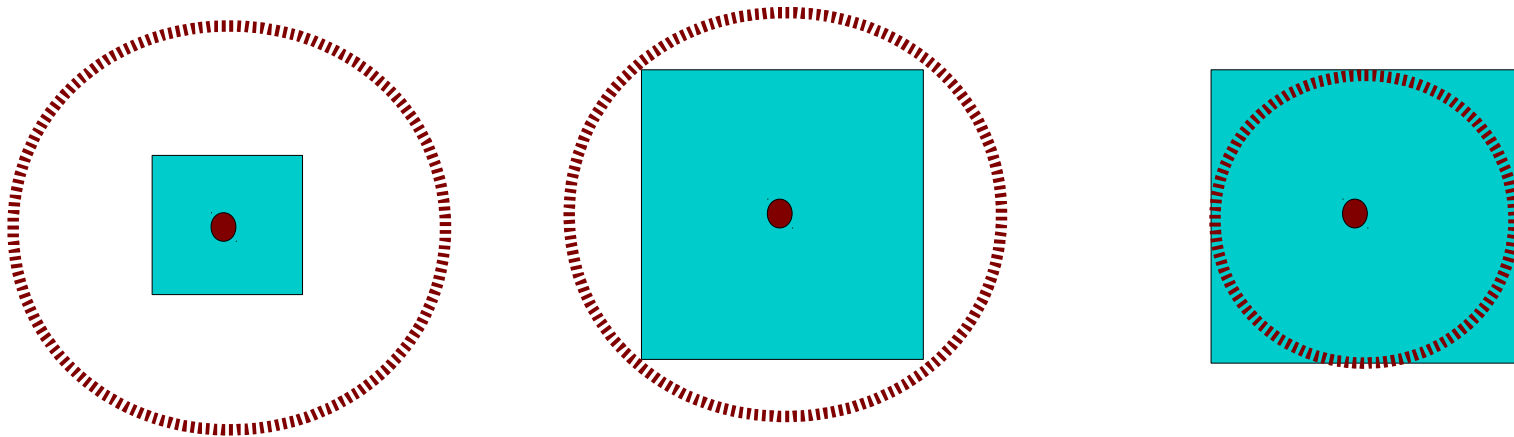


RESOLUCIÓN (10000 m)



IMPORTANTE

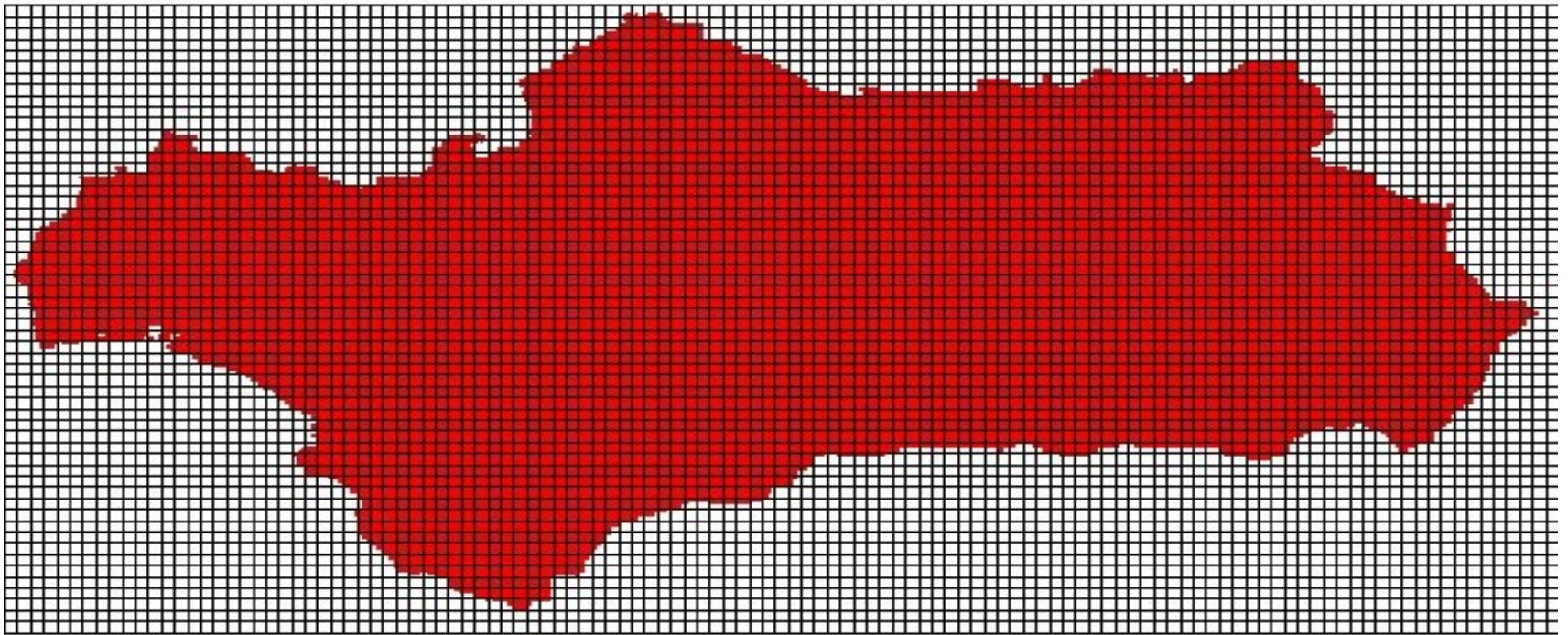
La resolución de las variables debe coincidir con la precisión de las coordenadas de las presencias



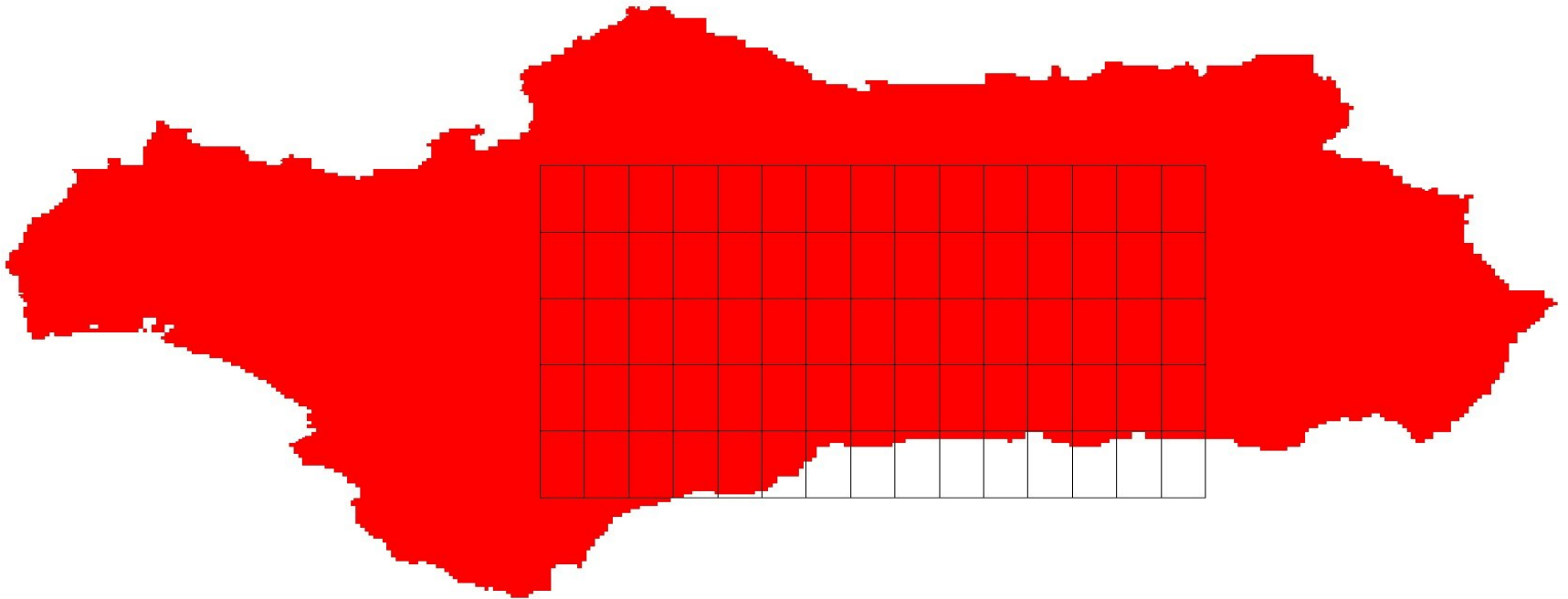
EXTENSIÓN Y RESOLUCIÓN

- **Extensión** y resolución nos permiten definir las características del área de trabajo.
- La extensión a elegir depende del área de distribución de la especie.
- La resolución depende de la calidad de datos disponibles.

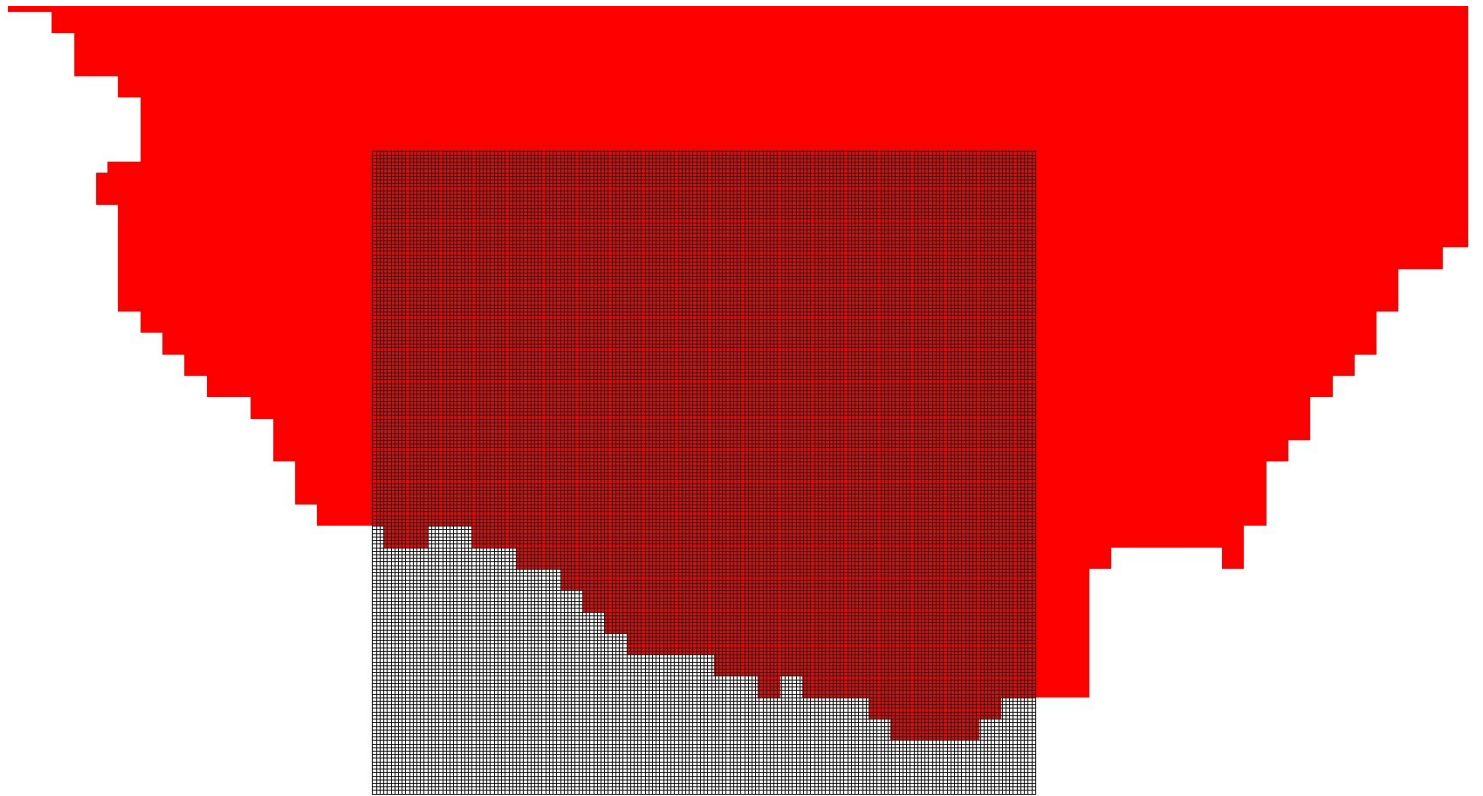
COMBINACIONES DE EXTENSIÓN Y RESOLUCIÓN



COMBINACIONES DE EXTENSIÓN Y RESOLUCIÓN



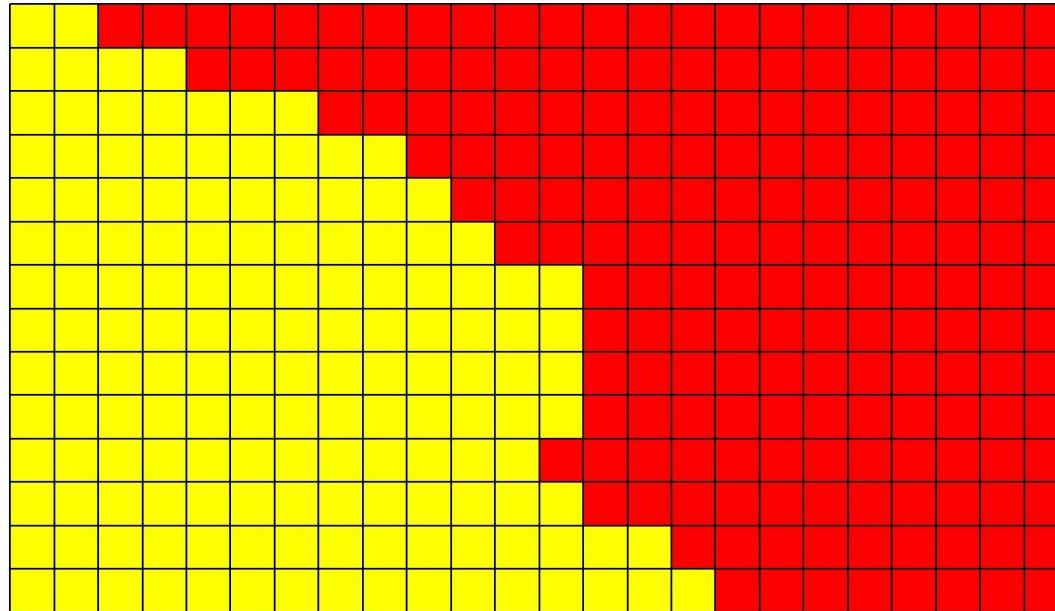
COMBINACIONES DE EXTENSIÓN Y RESOLUCIÓN



MÁSCARA

- Las celdas que no queremos usar se marcan con **valores nulos** (sin datos).
- Las celdas con valores nulos deben ser las mismas en todas las variables.
- Variables con muchas celdas nulas (teledetección) pueden dar problemas, provocando huecos sin datos en nuestro área de trabajo.

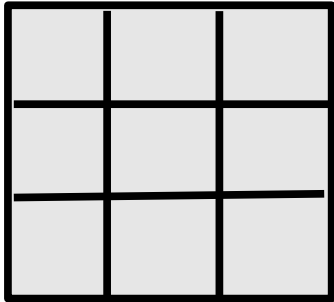
EXTENSIÓN + RESOLUCIÓN + MÁSCARA



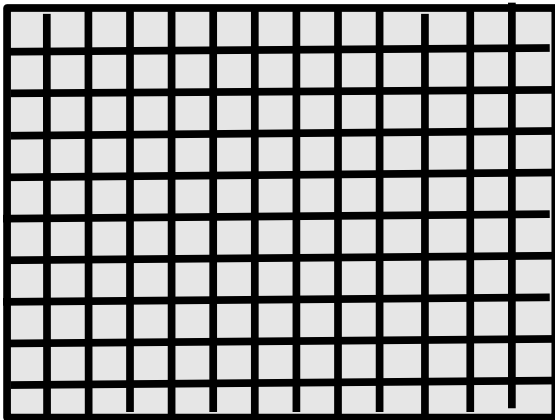
COMO LLEGAMOS HASTA AQUÍ?

1 – ELEGIR MAPA DE REFERENCIA

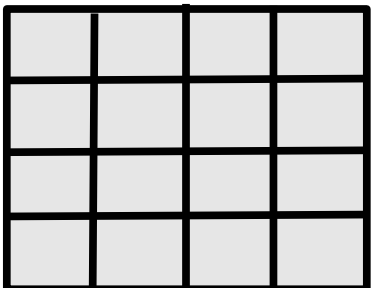
A



B



C

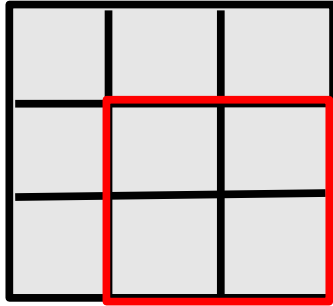


Depende de:

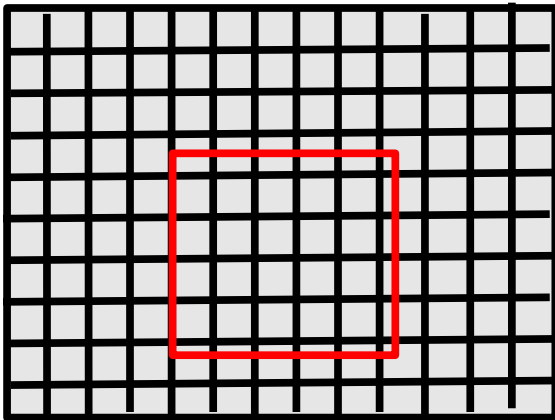
- Resolución de las presencias
- Resolución de la variable más importante para el modelo

2 – RECORTAR EXTENSIÓN

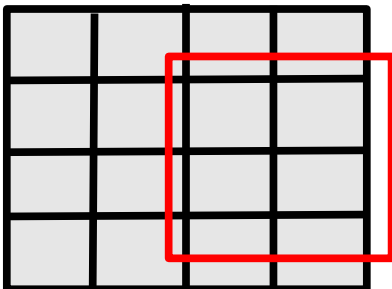
A



B

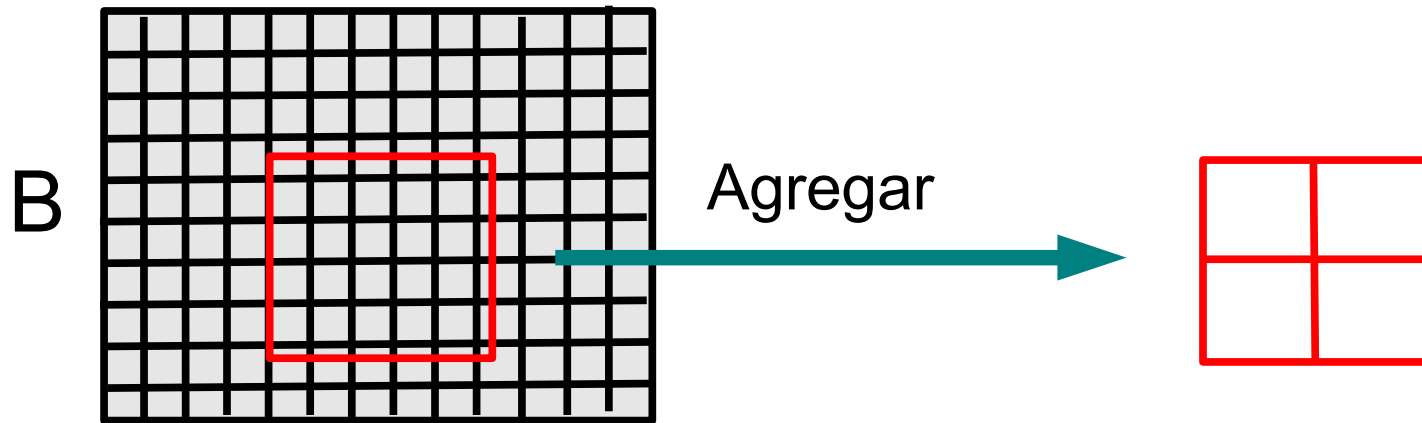


C

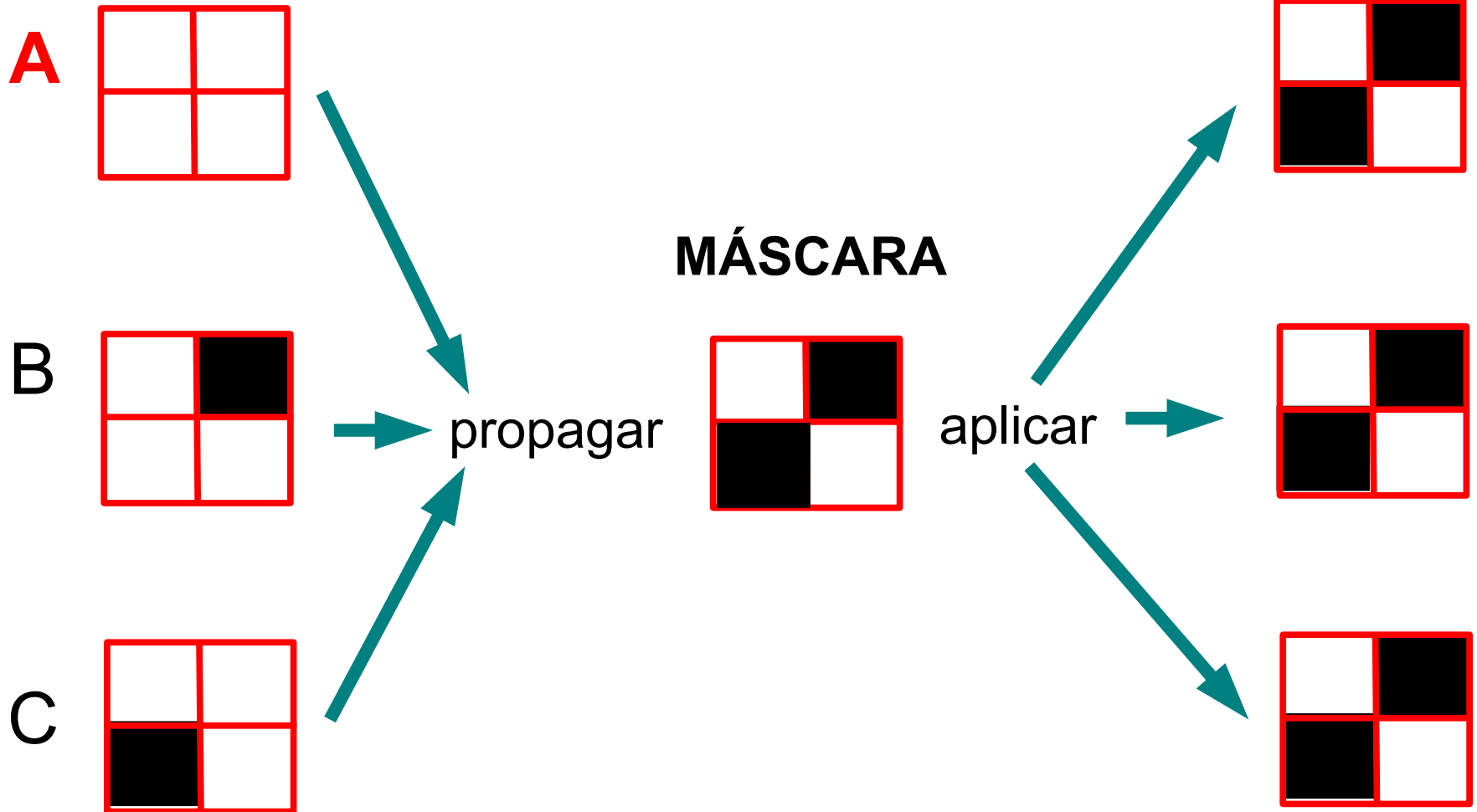


Normalmente se ajustará a la malla del mapa de referencia

3 – CAMBIAR RESOLUCIÓN



4 – IGUALAR CELDAS NULAS



SCRIPT DE R

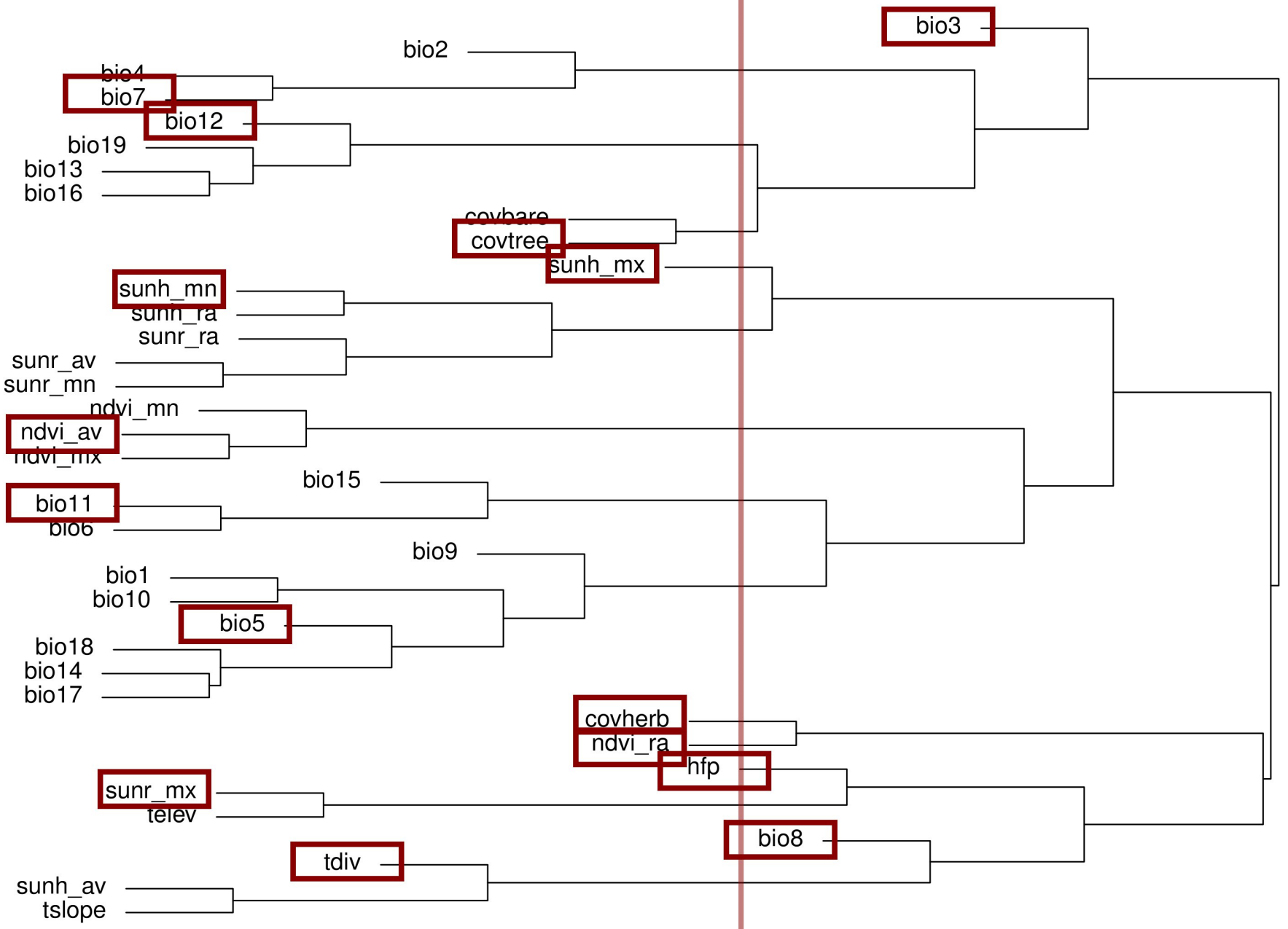
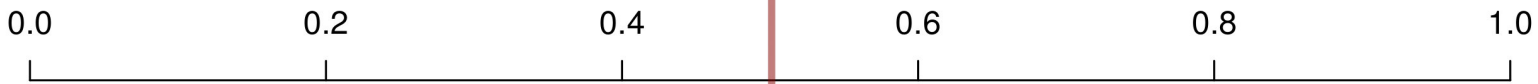
Abre

1_prepara_variables_y_presencia.R
con Rstudio

ANÁLISIS DE CORRELACIÓN DE VARIABLES PREDICTIVAS

ANÁLISIS DE CORRELACIÓN

- Permite identificar variables redundantes
- Si hay redundancia en una pareja de variables, descartamos una y retenemos la de interpretación biológica más sencilla
- R tiene todo lo necesario para un buen análisis de correlación!



VARIANCE INFLATION FACTOR

- Se calcula en R con la función **vif** del paquete **HH**
- $VIF = 1/(1-R_i^2)$
- R_i^2 es el R cuadrado de la regresión de la variable **i** frente al resto de las variables
- Si el $VIF_i > 5$, la información contenida en **i** ya está contenida en el resto de predictores, y podemos eliminar **i**

PREPARACIÓN DE REGISTROS DE PRESENCIA

PREPARACIÓN DE PRESENCIA

- Ajustar registros al área de trabajo
- Comprobar consistencia taxonómica
- Eliminar registros fósiles o muy antiguos
- Eliminar registros duplicados
- Examinar la resolución espacial de los datos
- Reducir autocorrelación espacial

PREPARACIÓN DE BACKGROUND, PSEUDO- AUSENCIA Y AUSENCIA

- **Background:** muestreo de las condiciones ambientales de todo el territorio
- **Pseudo-ausencia:** puntos al azar sin solape con las presencia
- **Ausencias:** simularemos unas ausencias

PREPARACIÓN DE PUNTOS DE EVALUACIÓN

- Seleccionaremos un porcentaje (entre el 20 y el 40%) de los puntos de presencia y las ausencias (solo de las ausencias) para evaluar todos los modelos.
- De estos puntos solo necesitamos las coordenadas.

ESTRUCTURA OBJETIVO

specie s	x	y	var_1	var_2	var_n	presen ce
Ursus	26.48	63.01	456	856	...	1
Ursus	23.60	50.20	546	452	...	1
Ursus	30.20	57.34	569	145	...	1
Ursus	32.10	62.35	412	456	...	0
Ursus	28.45	27.41	158	658	...	0
Ursus	30.24	31.22	123	456	...	0
...

TABLAS DE PRESENCIA

Con todas las presencias

- presencia
- presencia.ausencia
- presencia.pseudoausencia
- presencia.background

Misma
estructura

Sin las presencias de evaluación

- presencia.entrenamiento
- presencia.ausencia.entrenamiento
- presencia.pseudoausencia.entrenamiento
- presencia.background.entrenamiento

Tabla de evaluacion

- presencia.ausencia.evaluacion

Solo x, y y
presencia



That's all Folks!