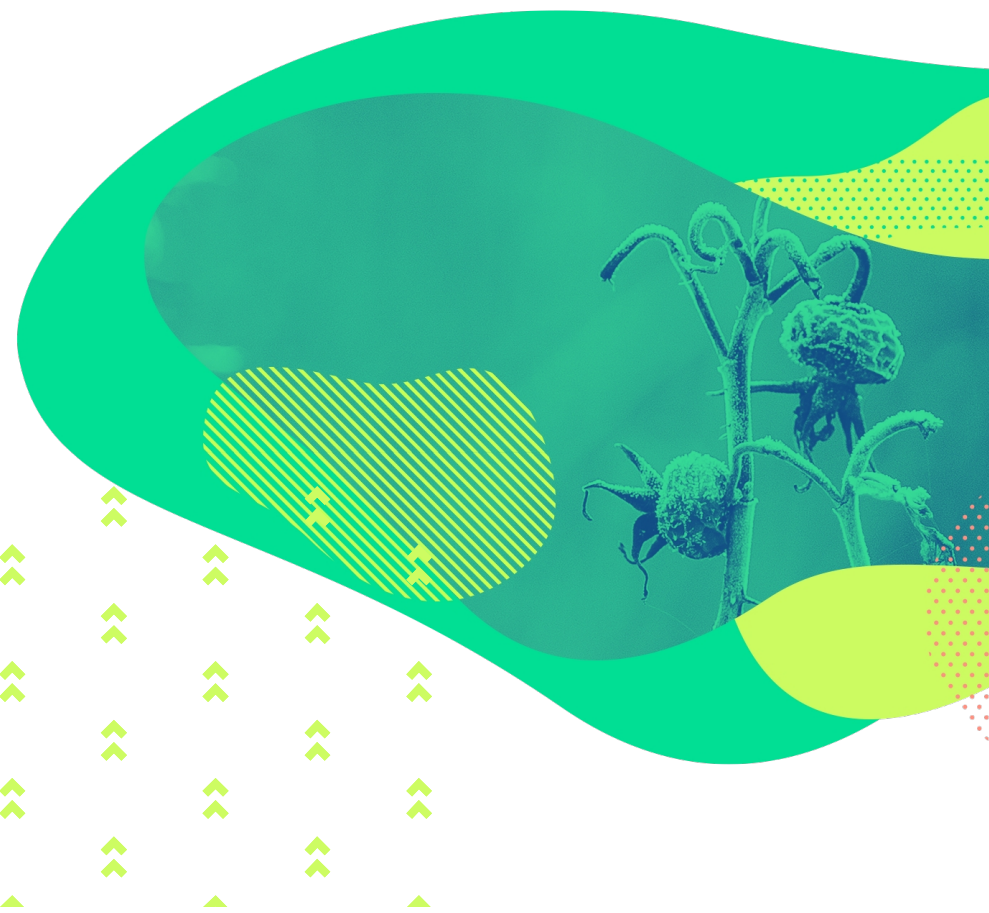


# IX Taller GBIF de Modelización de Nichos Ecológicos

Sesión 1:  
Introducción a los  
modelos de distribución  
de especies

**Blas M. Benito**



# **PRESENTACIÓN DEL TALLER**

# EL TALLER: SESIONES

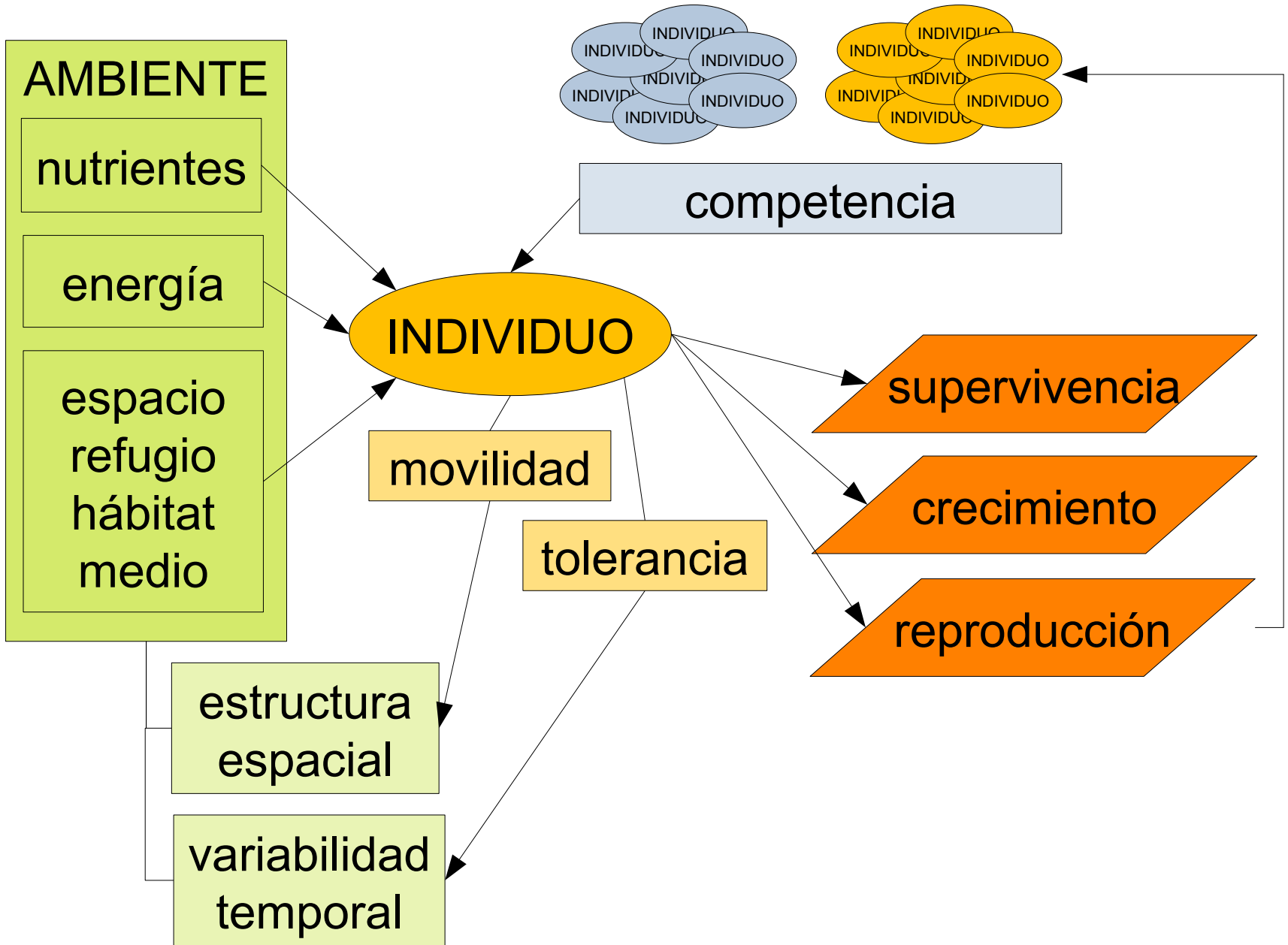
- **Sesión 1**
  - Distribución y nicho ecológico
  - Modelos de distribución: introducción y aplicaciones
  - El proceso de modelado
  - Preparación de presencias y variables

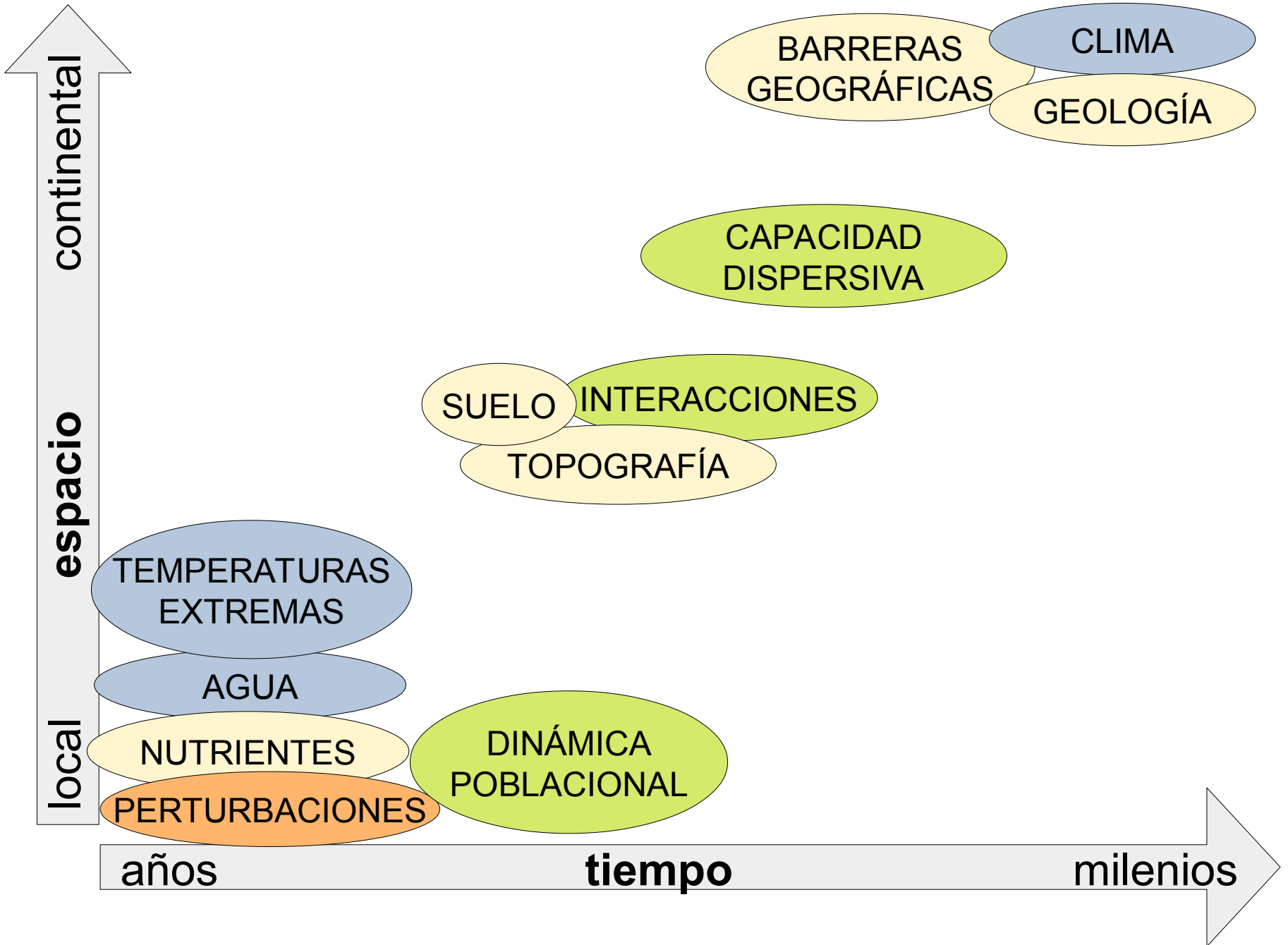
# EL TALLER: SESIONES

- **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



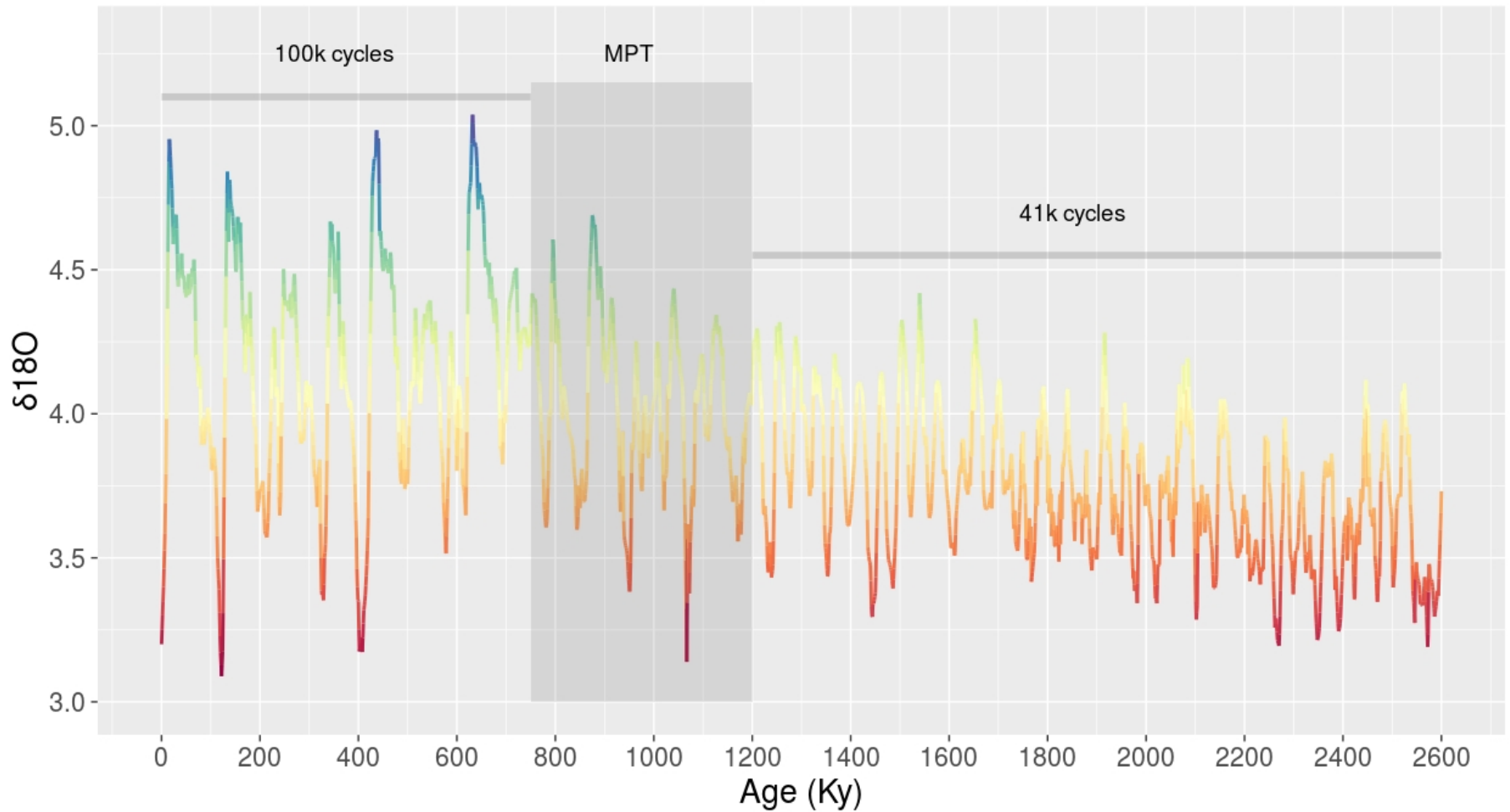
# **ESPECIES, NICHOS, Y DISTRIBUCIONES**





# TEMPERATURA

$\delta^{18}\text{O}$  during the Quaternary.



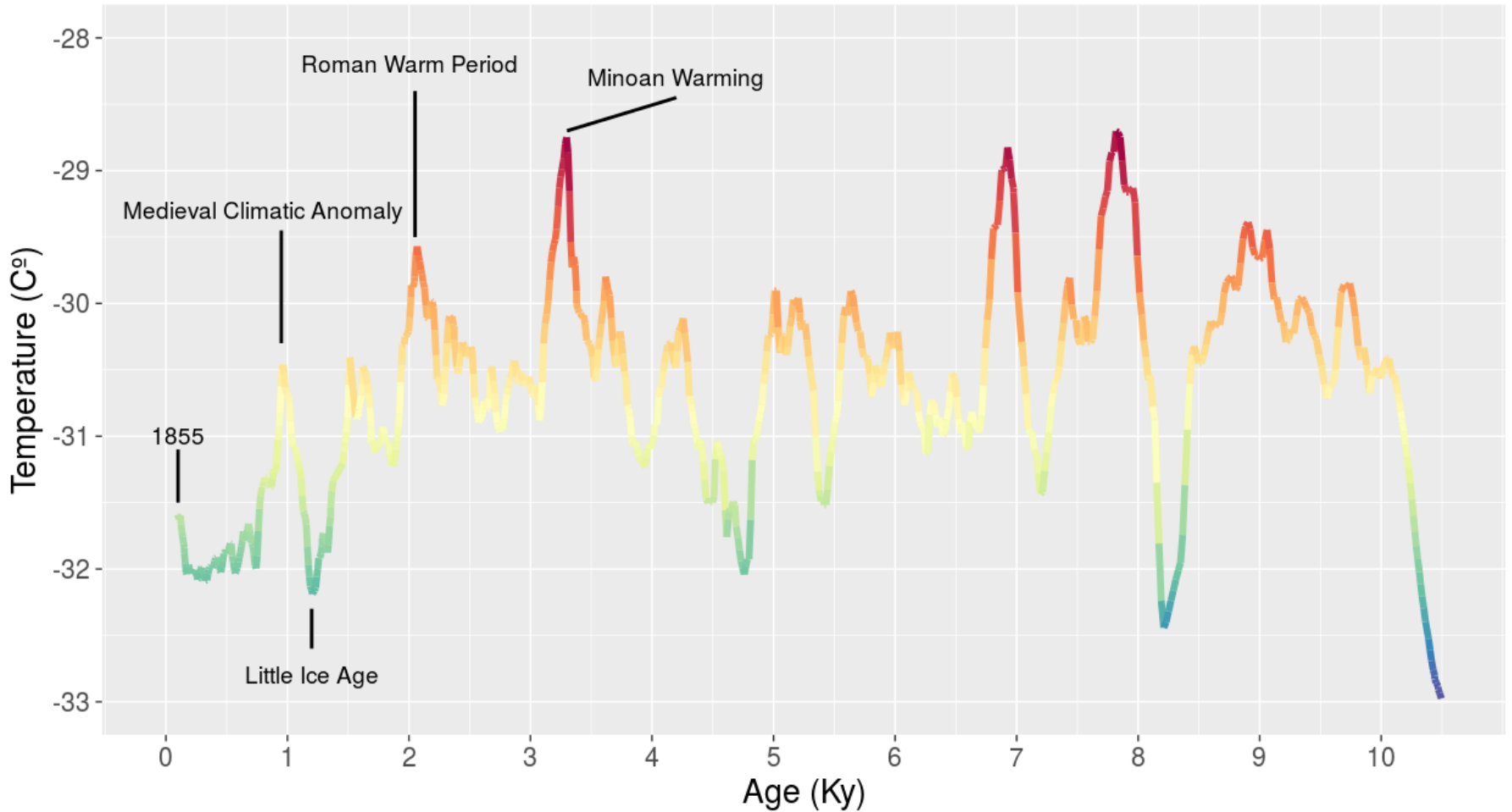
# TEMPERATURA

$\delta^{18}\text{O}$  during the last G-IG cycle.



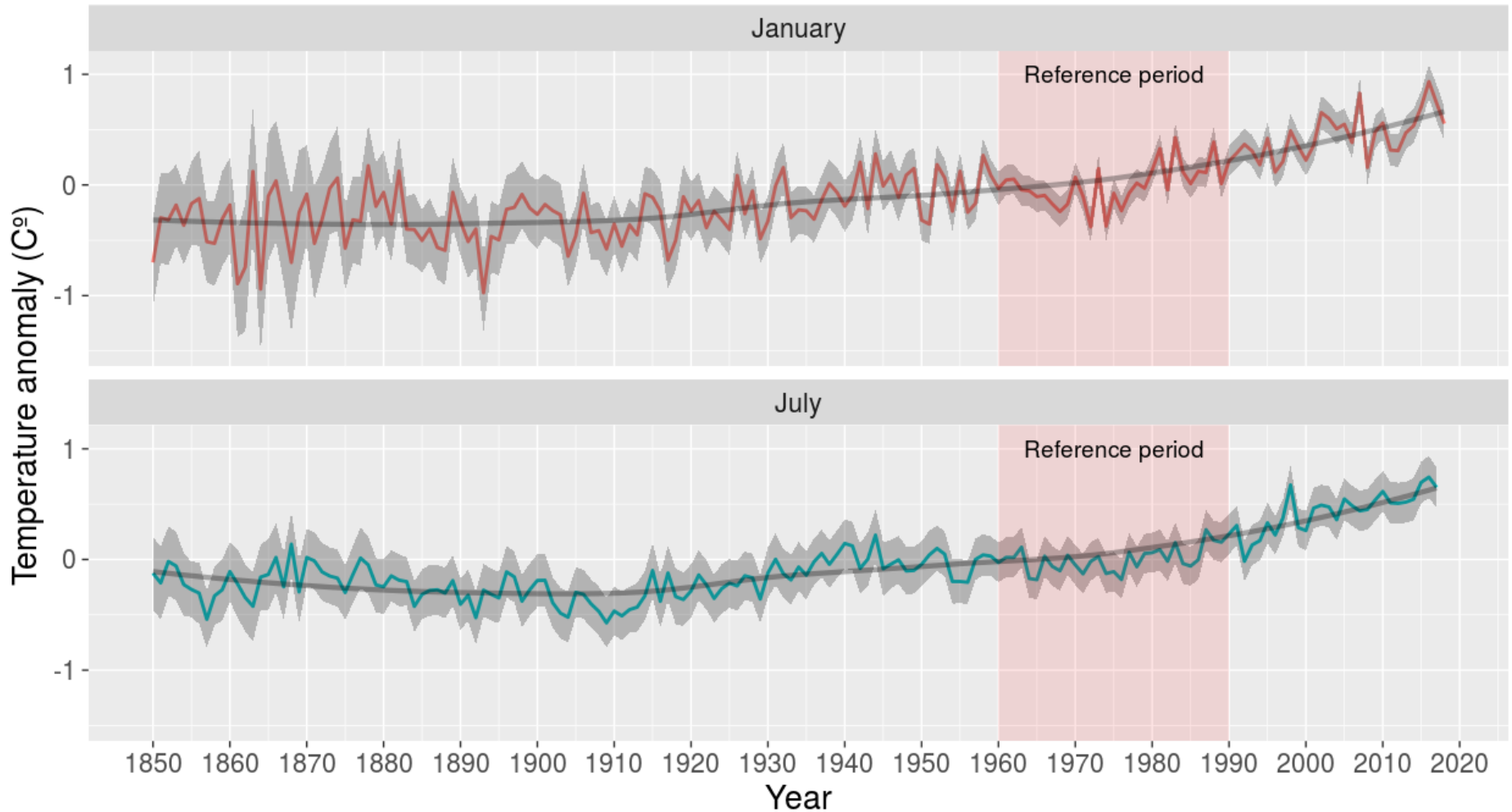
# TEMPERATURA

Temperature in central Greenland during the Holocene (GISP2 Ice Core).



# TEMPERATURA

HadCRUT (version 4) global temperature anomaly

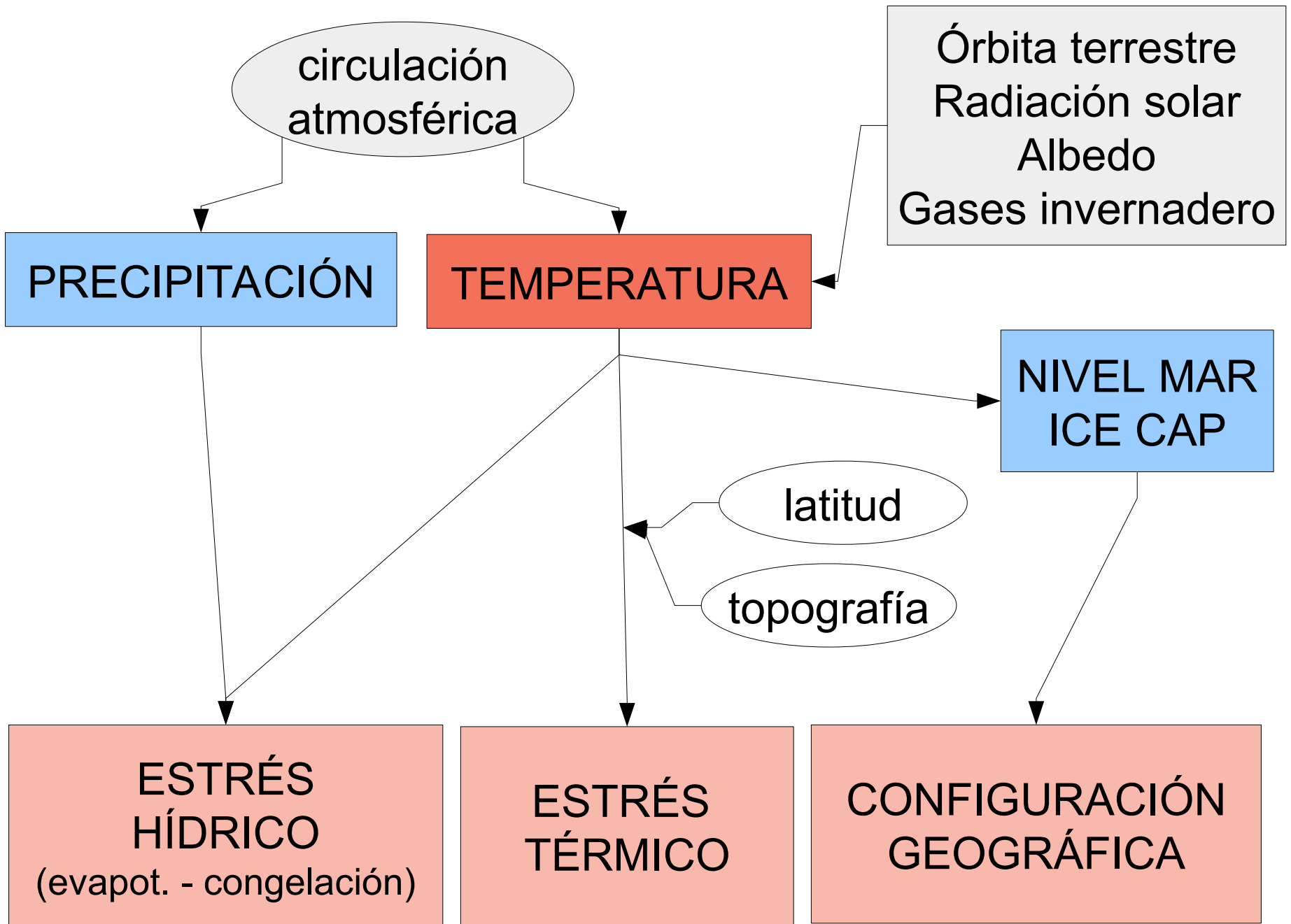




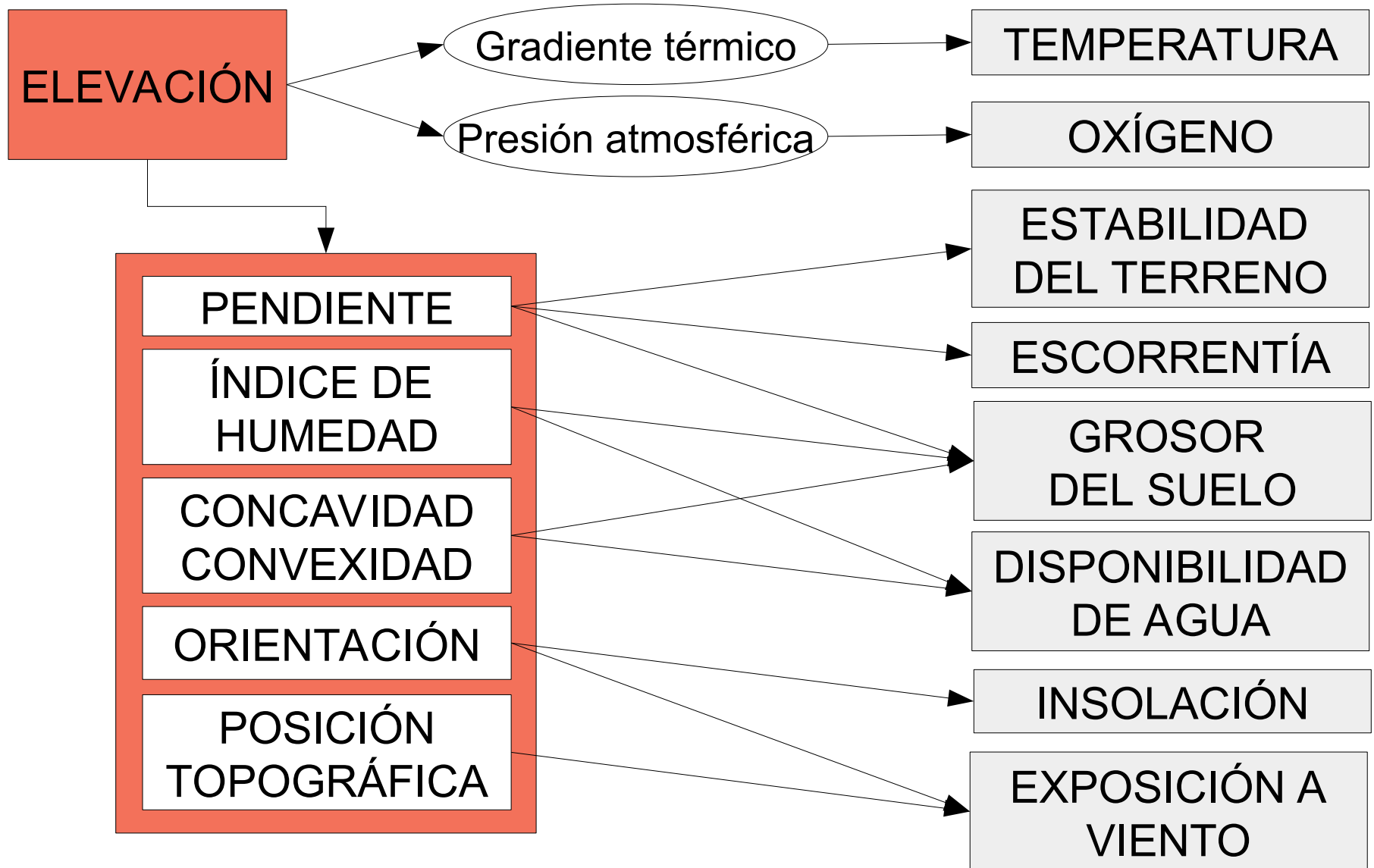


# TEMPERATURA

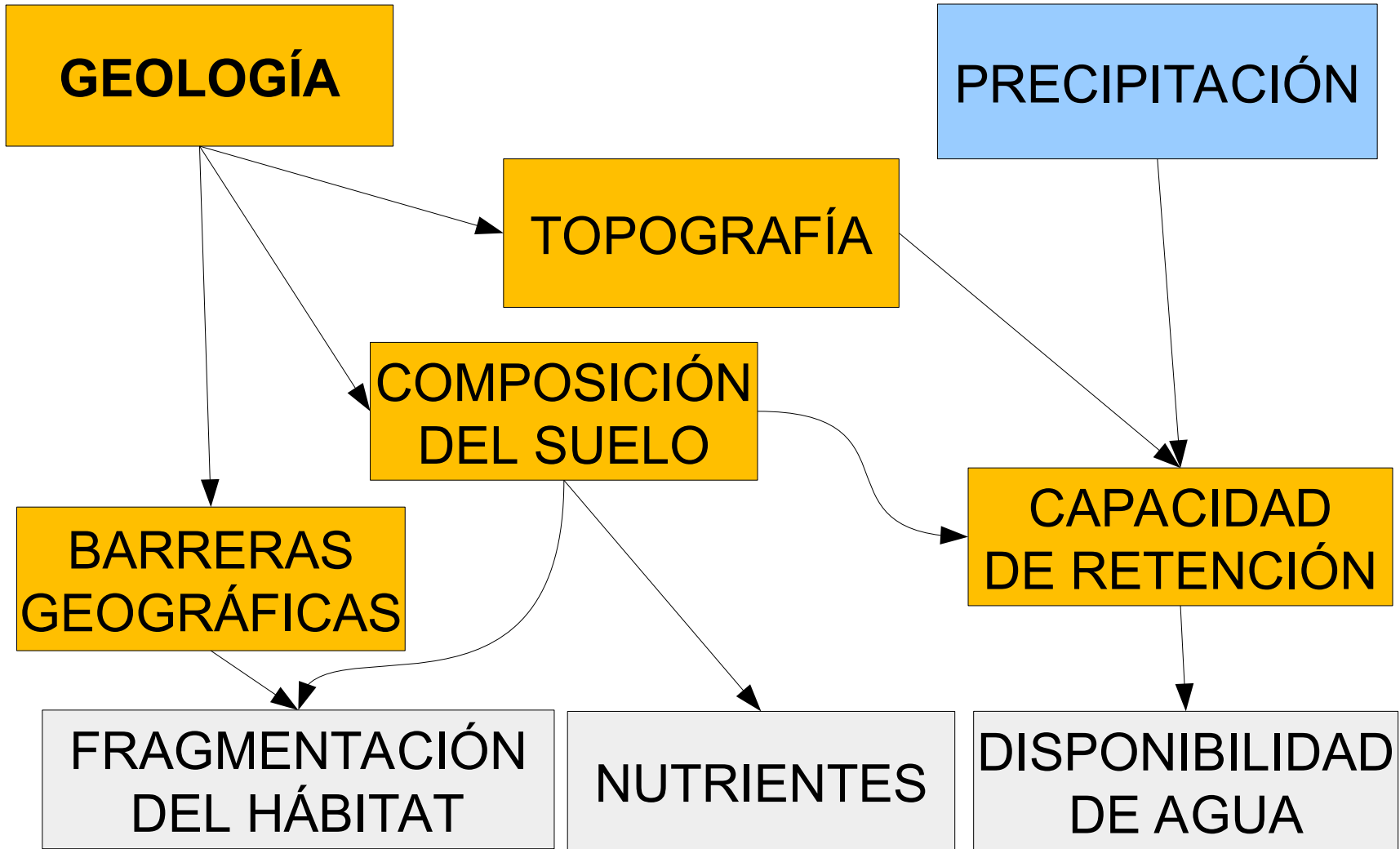
- **Efecto metabólico** → supervivencia, reproducción (fenología!), estrés térmico.
- **Temperaturas extremas** suele definir límites de distribución mejor que medias (organismos no perciben temperatura media anual).
- Se mapea con cierta precisión: estaciones meteorológicas + interpolación (*Worldclim*) o teledetección (*MODIS Land Surface Temperature*).
- Se predice con Modelos de Circulación Global



# TOPOGRAFÍA



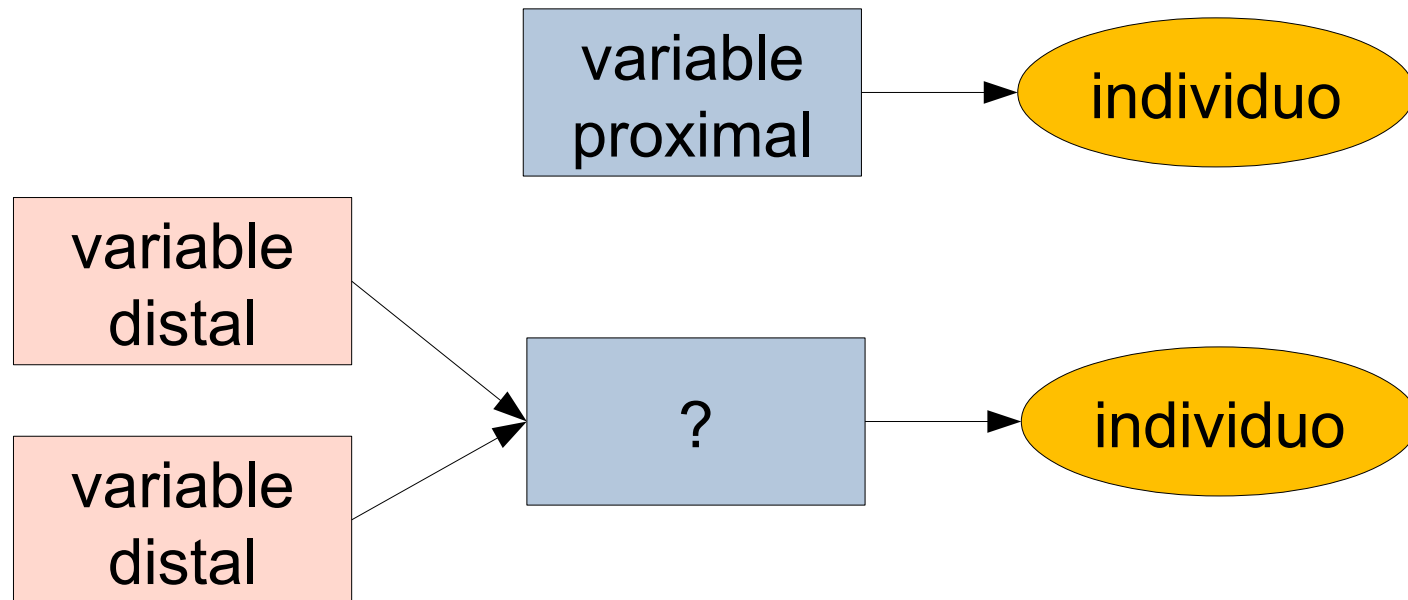
# GEOLOGÍA Y SUELOS



# VARIABLES PROXIMALES Y DISTALES

- Variable **proximal**: afecta directamente a la supervivencia/crecimiento/reproducción de los individuos.
  - Luz solar en plantas.
  - Disponibilidad de presas en carnívoros.
- Variable **distal**: está indirectamente relacionada con una o varias variables proximales.
  - Elevación
  - Topografía

# VARIABLES PROXIMALES Y DISTALES



# ACTIVIDAD HUMANA

- Fragmentación y destrucción de hábitats
- Explotación de poblaciones
- Introducción de especies invasoras
- Extracción de recursos requeridos por poblaciones silvestres (agua, suelo)
- Dispersión a larga distancia

**La actividad humana modifica la distribución natural de las especies**

The Last of the Wild (human footprint):

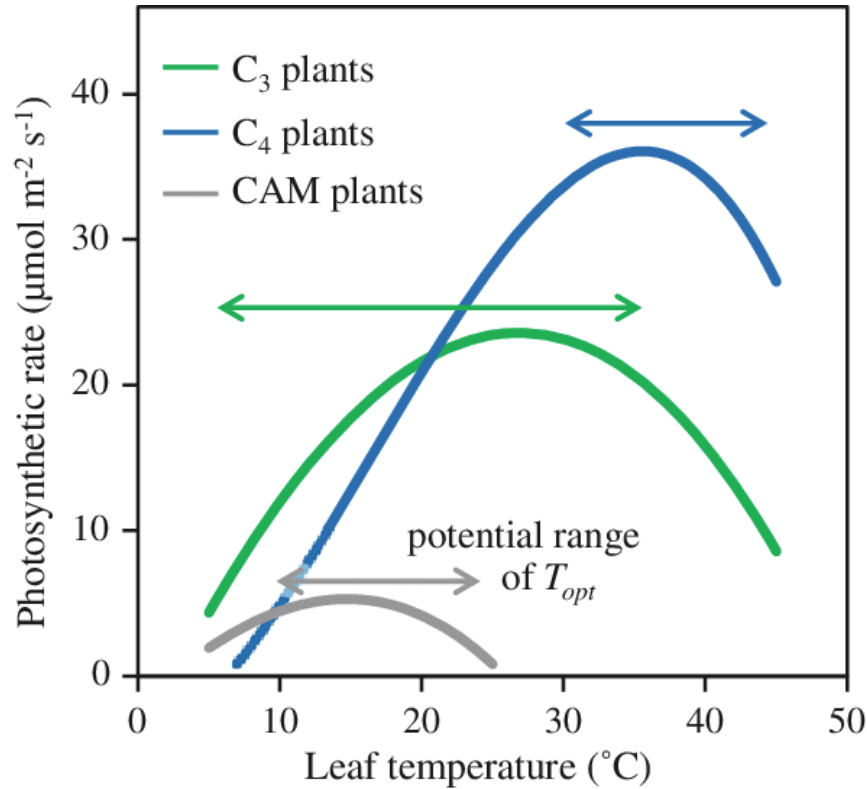
<https://sedac.ciesin.columbia.edu/data/set/wildareas-v3-2009-human-footprint>

**¿COMO RESPONDEN LOS  
INDIVIDUOS Y LAS POBLACIONES A  
ESAS VARIABLES?**



# RESPUESTAS DEL INDIVIDUO

Fotosíntesis ~ temperatura



DOI: 10.1007/s11120-013-9874-6

## fecundidad ~ temperatura

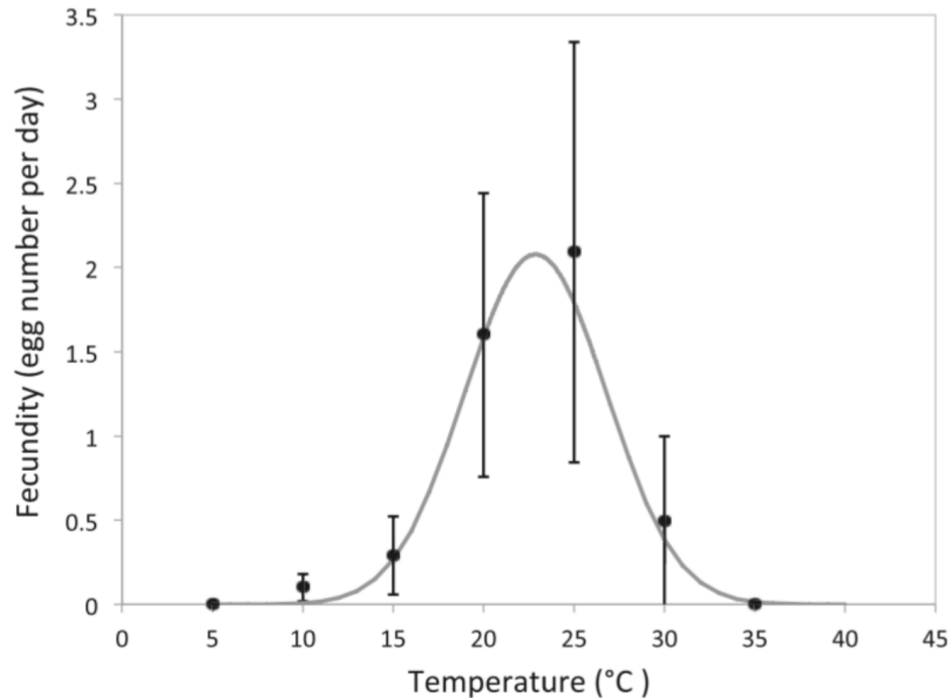


Figure 1: Temperature-dependent fecundity of *D. sukukii*. Fecundity is expressed as the number of eggs per female per day, as measured over a 10-d period following a 4-d acclimation period. Data were fitted using a polynomial function with compact support, adapted from Saryazdi and Cheriet (2007). Data points are means $\pm$ SD.

## movilidad ~ temperatura

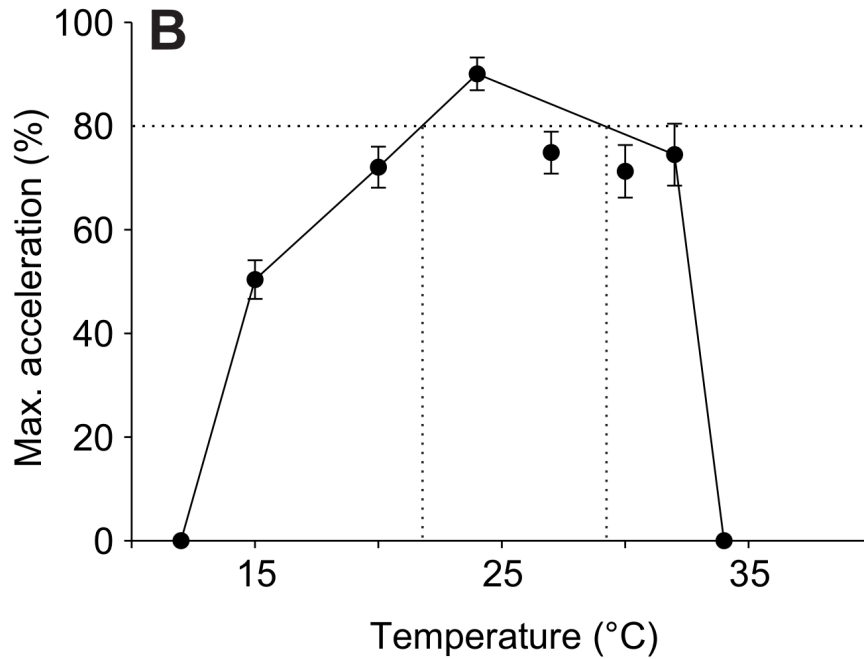
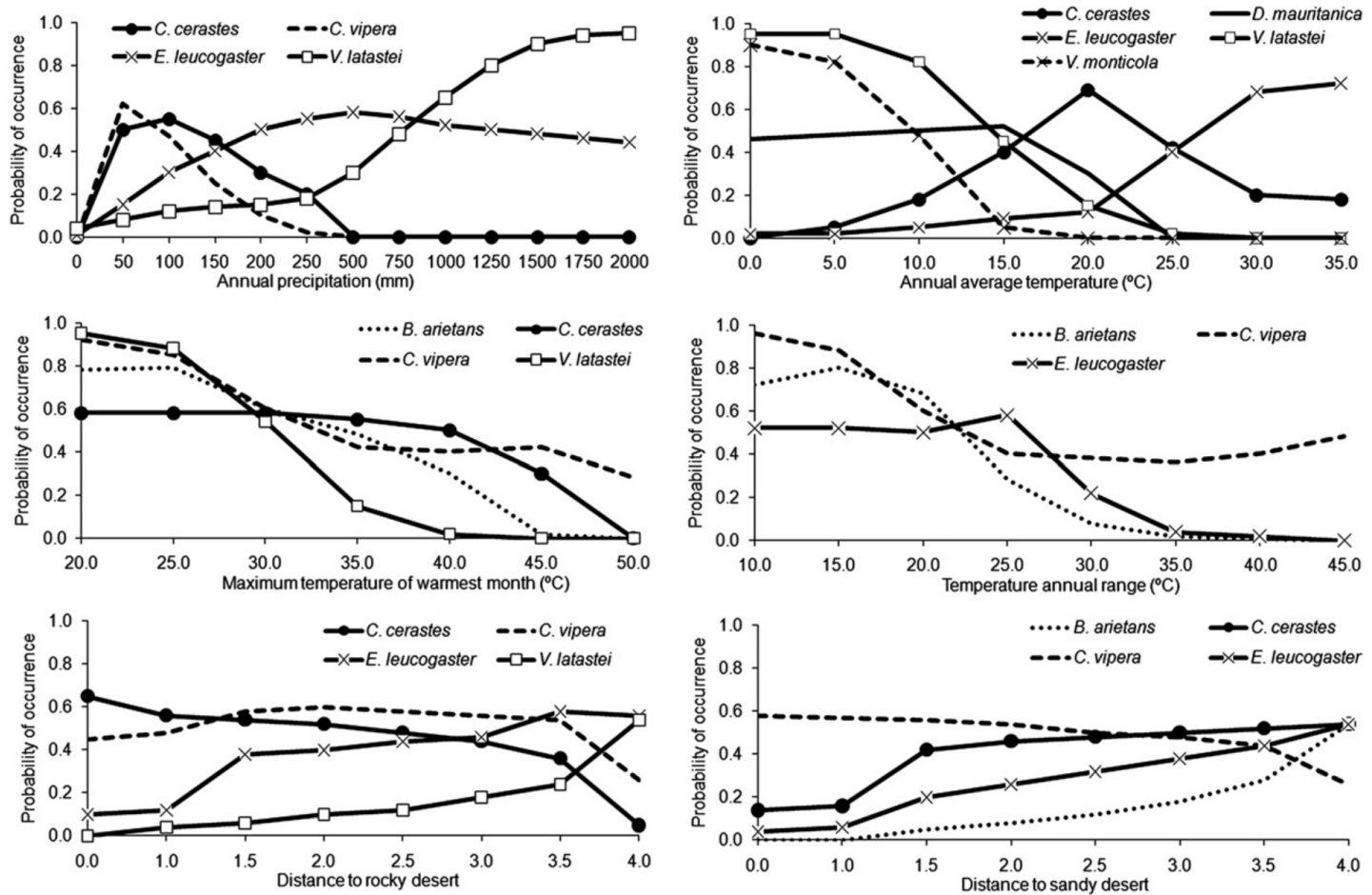


Fig. 2. Mean ( $\pm$ s.e.m.;  $N=16$  individuals) percentage of maximum performance for the burst speed measures (A) peak velocity and (B) peak acceleration. The horizontal dotted line represents the 80% performance breadth, a measure of the temperature range for which frogs can still achieve 80% of their maximal performance. Note that temperature–performance curves are bound by the critical thermal minimum and maximum temperatures (vertical dotted lines) as determined experimentally.

# RESPUESTAS DE LAS POBLACIONES

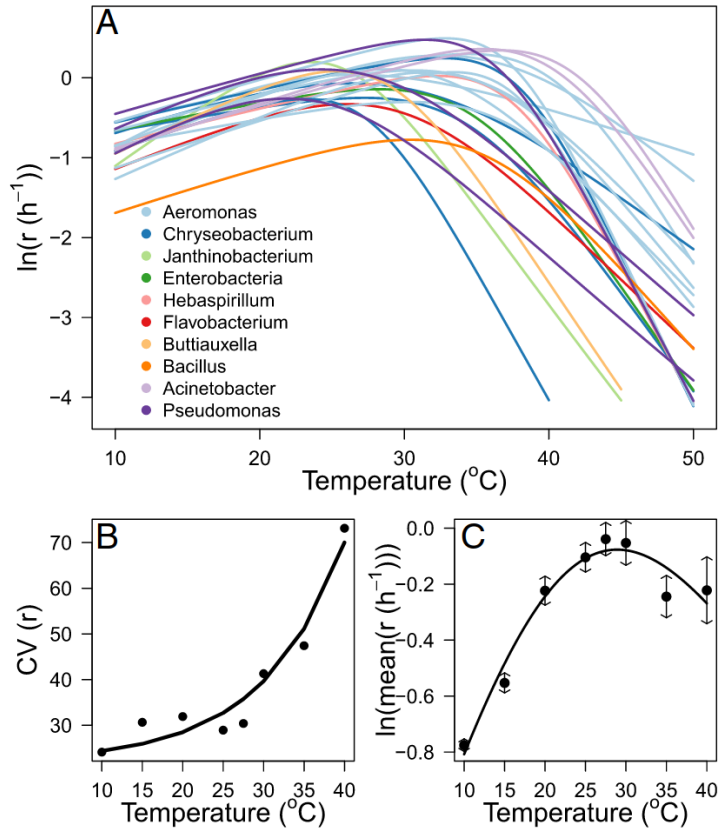
J.C. Brito et al. / Journal of Arid Environments 75 (2011) 1029–1037

1033



**Fig. 2.** Response curves for the most related environmental factors to the distribution of vipers in North-West Africa. Curves depict average probability of occurrence from 30 model replicates along the environmental gradients.

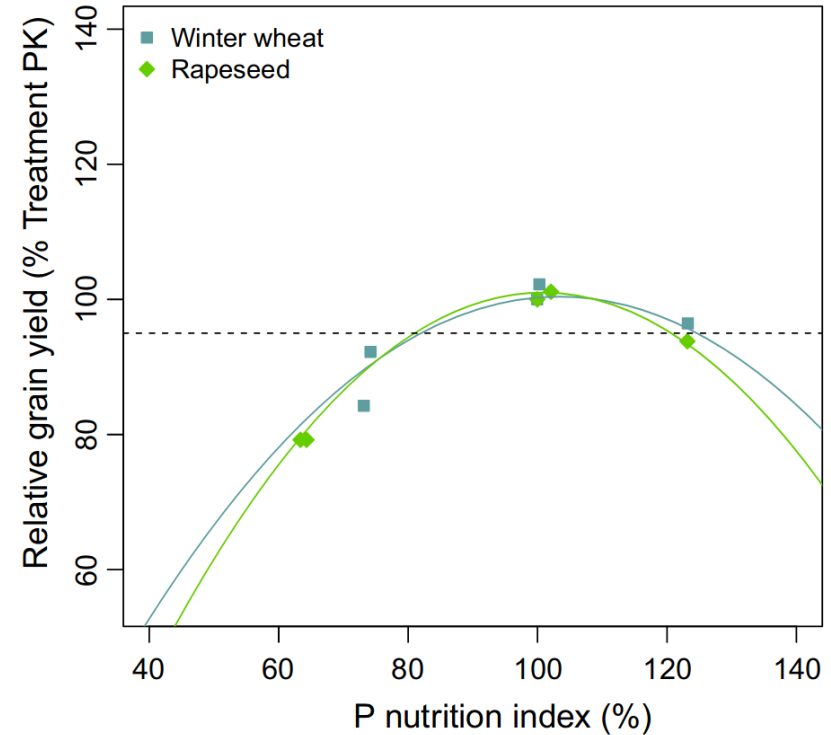
## supervivencia ~ temperatura



**Fig. 1.** Thermal tolerance curves of the 24 bacterial taxa. (A) Comparison of the fitted thermal tolerance curves of population growth rate,  $r$ , for the 24 taxa quantified using the Sharpe-Schoofield equation (*Materials and Methods*). (B) The coefficient of variance (CV) in  $r$  among the 24 taxa at each assay temperature demonstrates an exponential increase in CV with rising temperature ( $y = 21.98 e^{0.04x}$ ,  $r^2 = 0.94$ ,  $F_{1,6} = 96.22$ ,  $P < 0.001$ ). (C) Pooled thermal tolerance curve fitted using the Sharpe-Schoofield equation (*SI Appendix, Table S1*) to the mean growth rate across all 24 taxa at each assay temperature demonstrates a marked decline in average performance above 27.5  $^{\circ}\text{C}$ , which coincides with the temperature at which the CV of population growth rate rapidly increases (quasi- $r^2 = 0.92$  of the fitted model).

DOI: 10.1073/pnas.1805518115

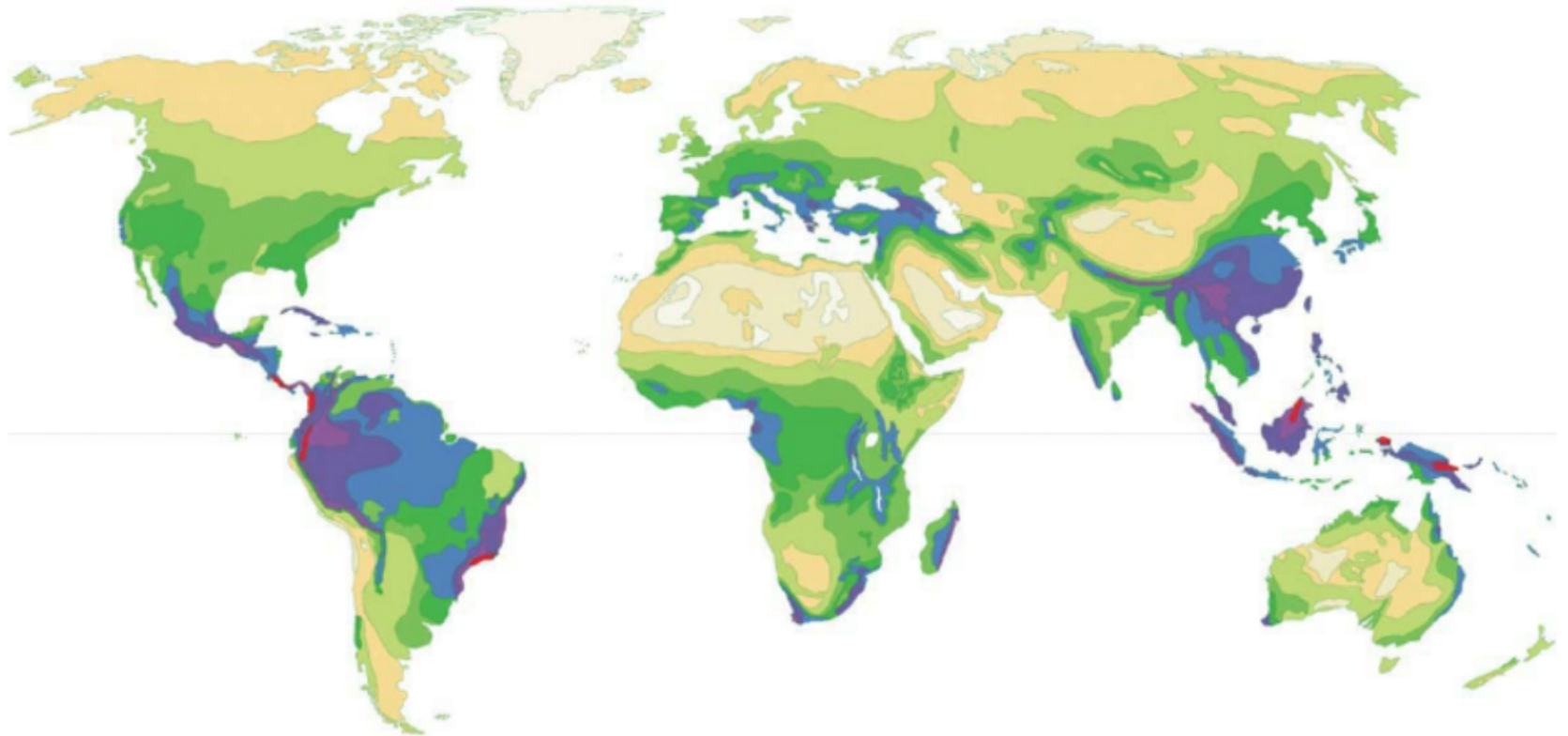
## biomasa ~ nutrientes



**Fig. 5** Relationship between relative grain yield (RGY) and the phosphorus nutrition index (PNI) (mean for all sampling dates) for winter wheat [ $\text{RGY} = -0.012\text{PNI}^2 + 2.46\text{PNI} - 26.4$ , Residual Standard Error (RSE) = 7.88,  $n = 5$ ] and rapeseed [ $\text{RGY} = -0.015\text{PNI}^2 + 3.08\text{PNI} - 54.1$ , RSE = 7.37,  $n = 5$ ]. Relative grain yields of winter wheat and rapeseed were calculated by dividing the grain yield from a given fertilization treatment by the grain yield from the PK treatment. The dashed horizontal line represents a relative grain yield of 95%

DOI: 10.1007/s10705-018-9956-0

# RESPUESTA DE LA BIODIVERSIDAD



Diversity Zones (DZ): Number of species per 10,000 km<sup>2</sup>

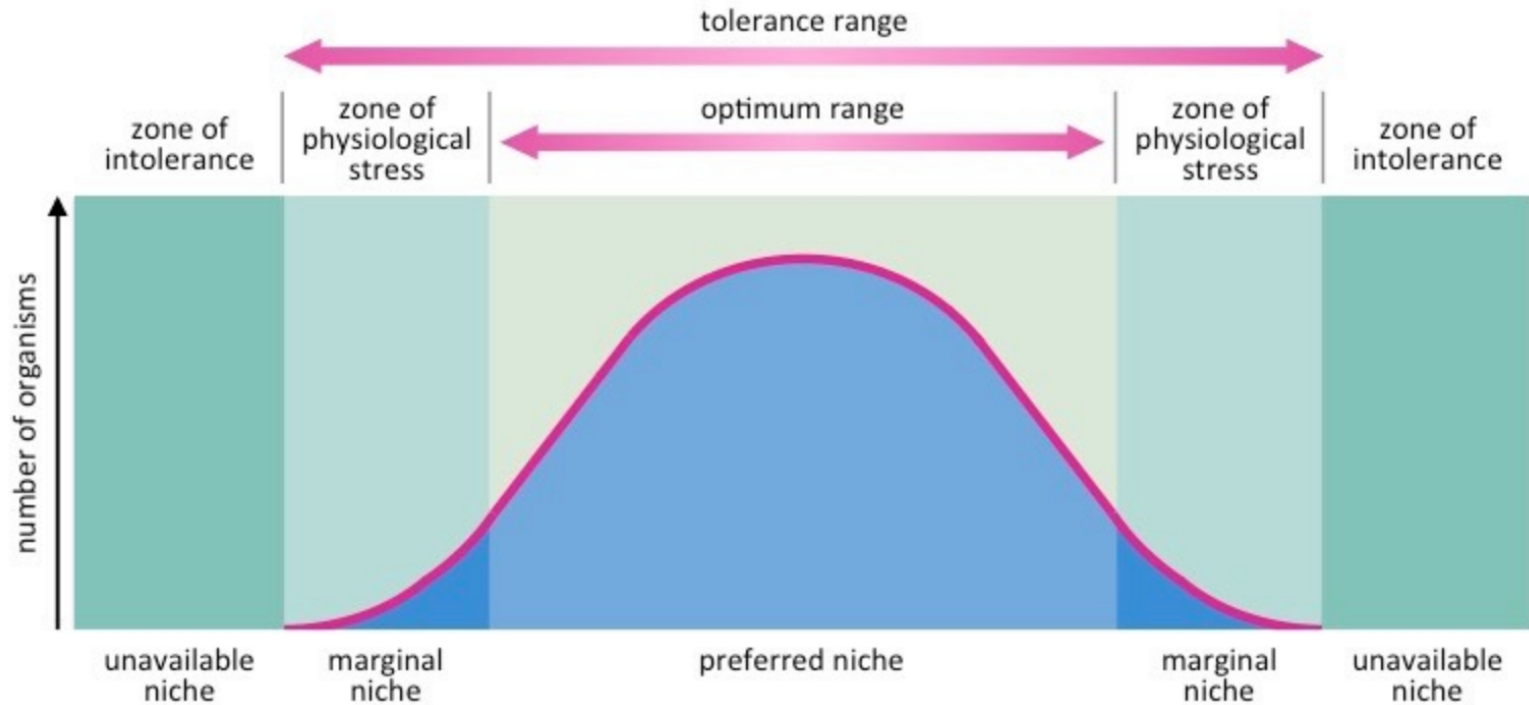
DZ 1	<20 spp.	DZ 5	1000-1500 spp.	DZ 9	4000-5000 spp.
DZ 2	20-200 spp.	DZ 6	1500-2000 spp.	DZ 10	>5000 spp.
DZ 3	200-500 spp.	DZ 7	2000-3000 spp.		
DZ 4	500-1000 spp.	DZ 8	3000-4000 spp.		

W. Barthlott, G. Kier, H. Kreft, W. Küper,  
D. Rafiqpoor & J. Mutke 2005

Nees Institute for Biodiversity of Plants  
University of Bonn

Mutke et al. 2010

# EL NICHO ECOLÓGICO

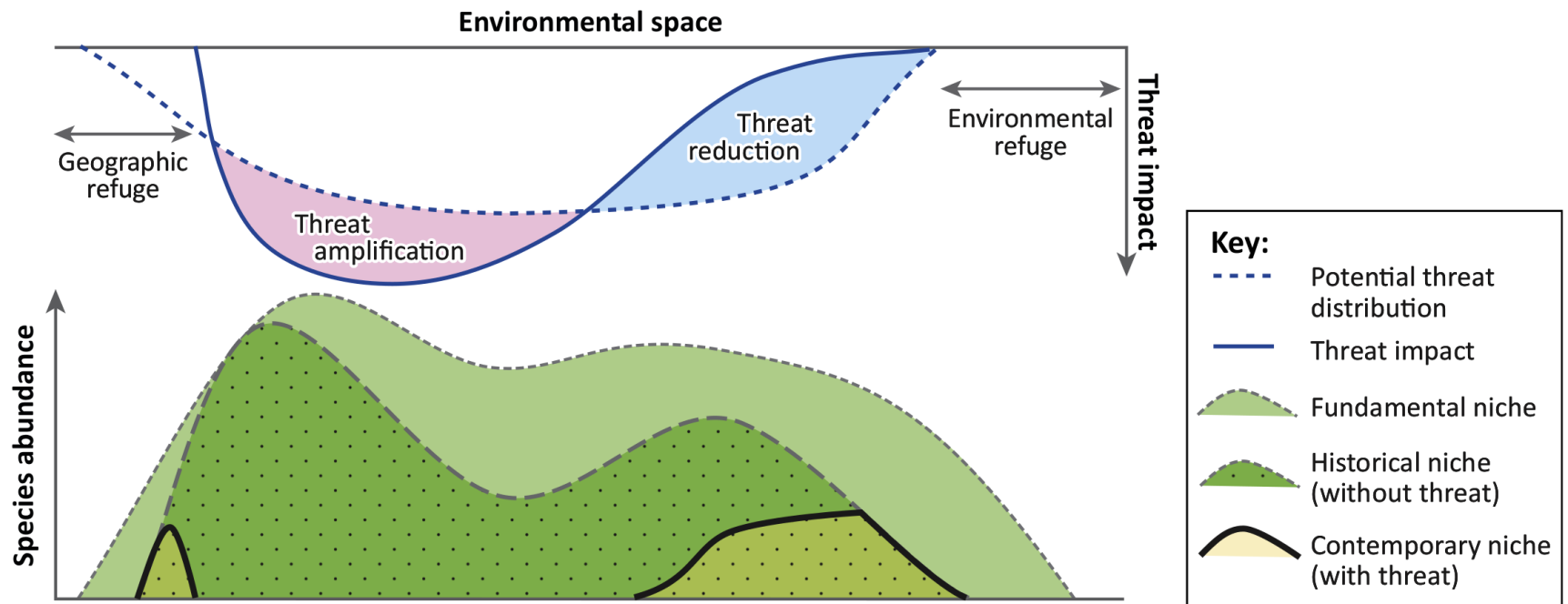


**Marginalidad:** distancia entre el óptimo del nicho y la media de la variable ambiental en el área de estudio  
**Especialización:** ancho del nicho ecológico.



# REDUCCIÓN DE NICHOS

(A) Threat distribution or impact is heterogeneous in environmental space

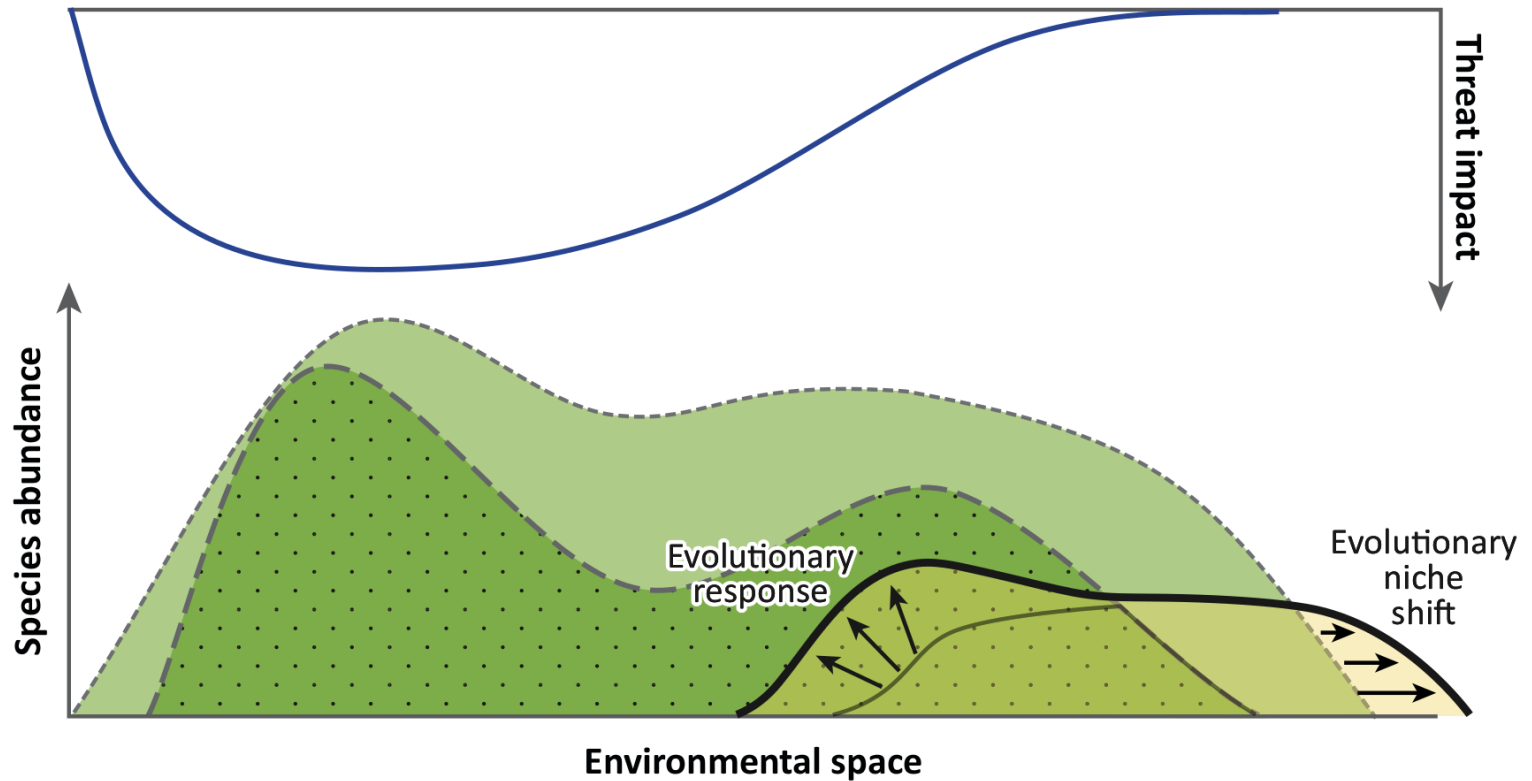


DOI: 10.1016/j.tree.2017.02.013



# EXPANSIÓN DE NICHOS

(c) The threat triggers an evolutionary response in the impacted species



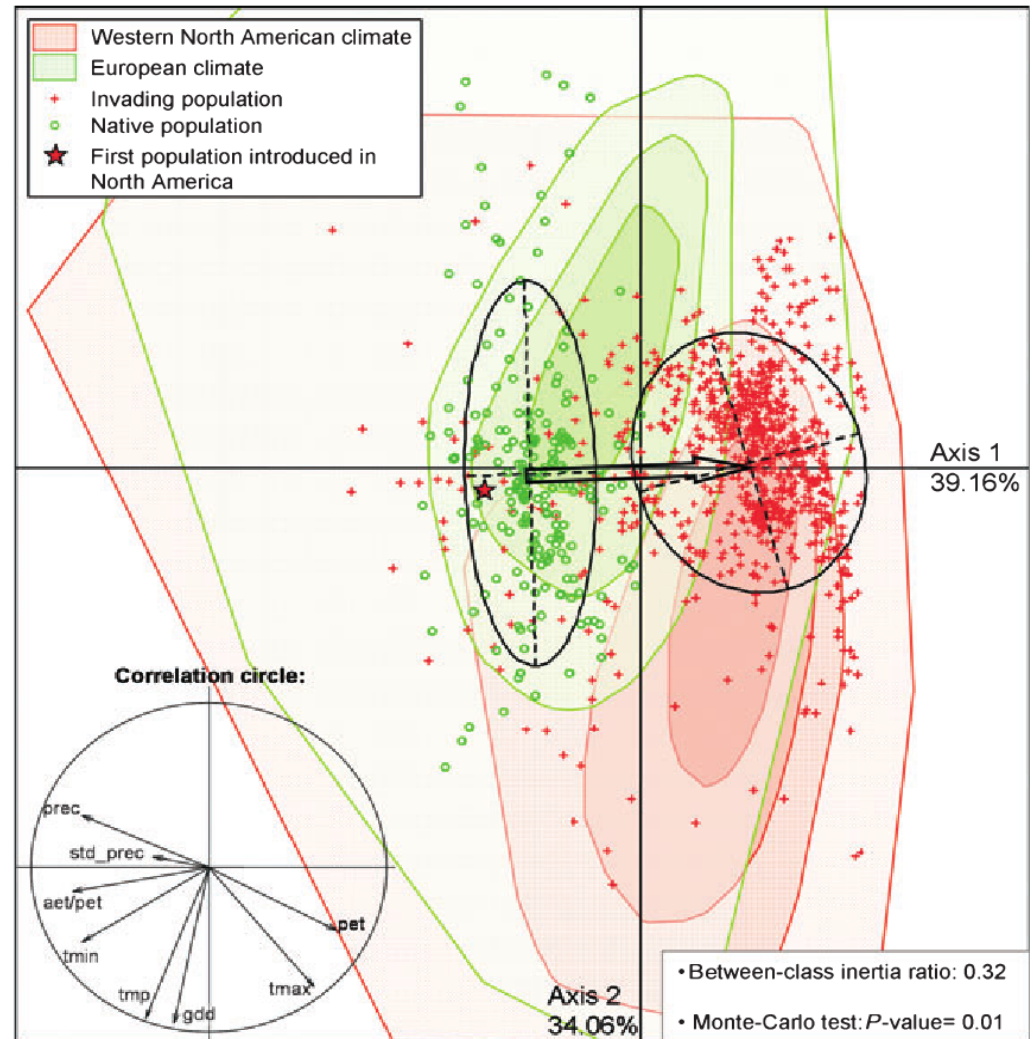
DOI: 10.1016/j.tree.2017.02.013

# EJEMPLO DE EXPANSIÓN DE NICHU

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

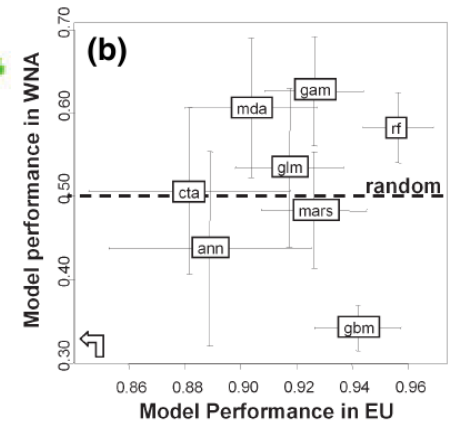
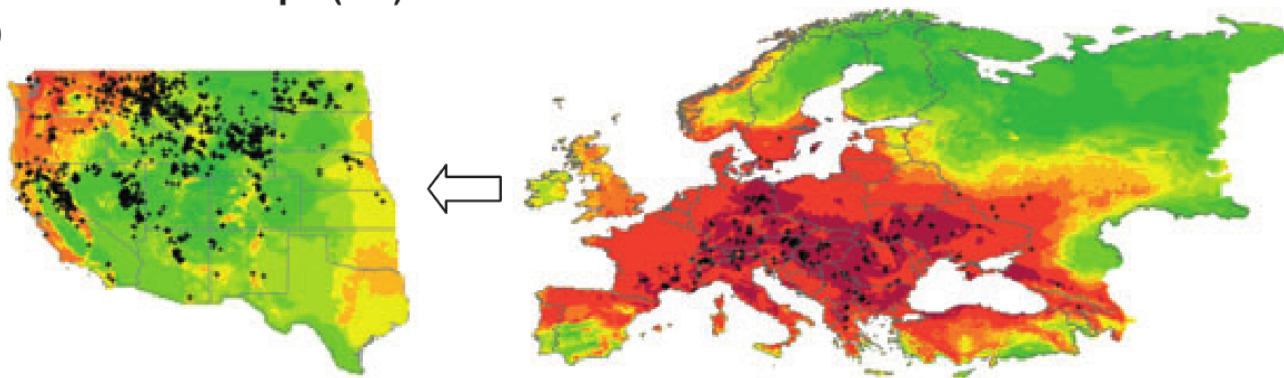


*Centaurea maculosa*



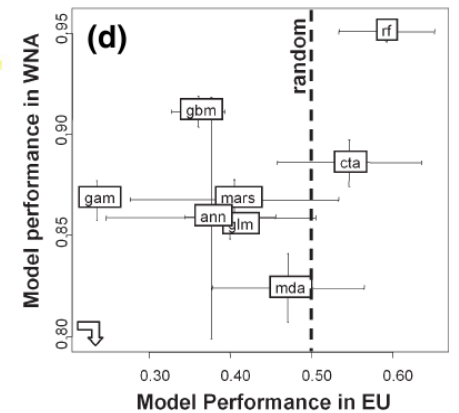
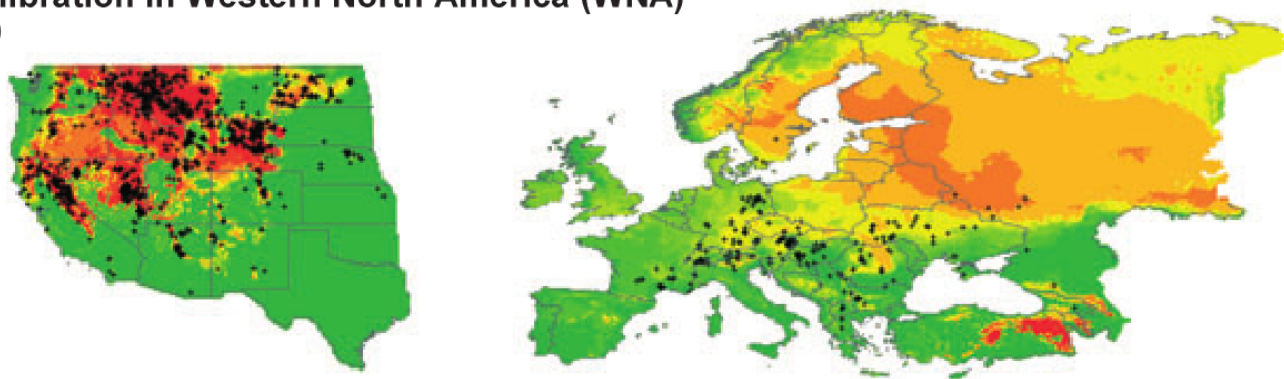
### Calibration in Europe (EU)

(a)



### Calibration in Western North America (WNA)

(c)

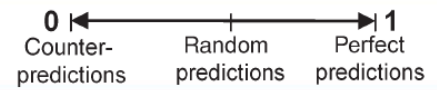


Occurrences of *Centaurea maculosa* : +

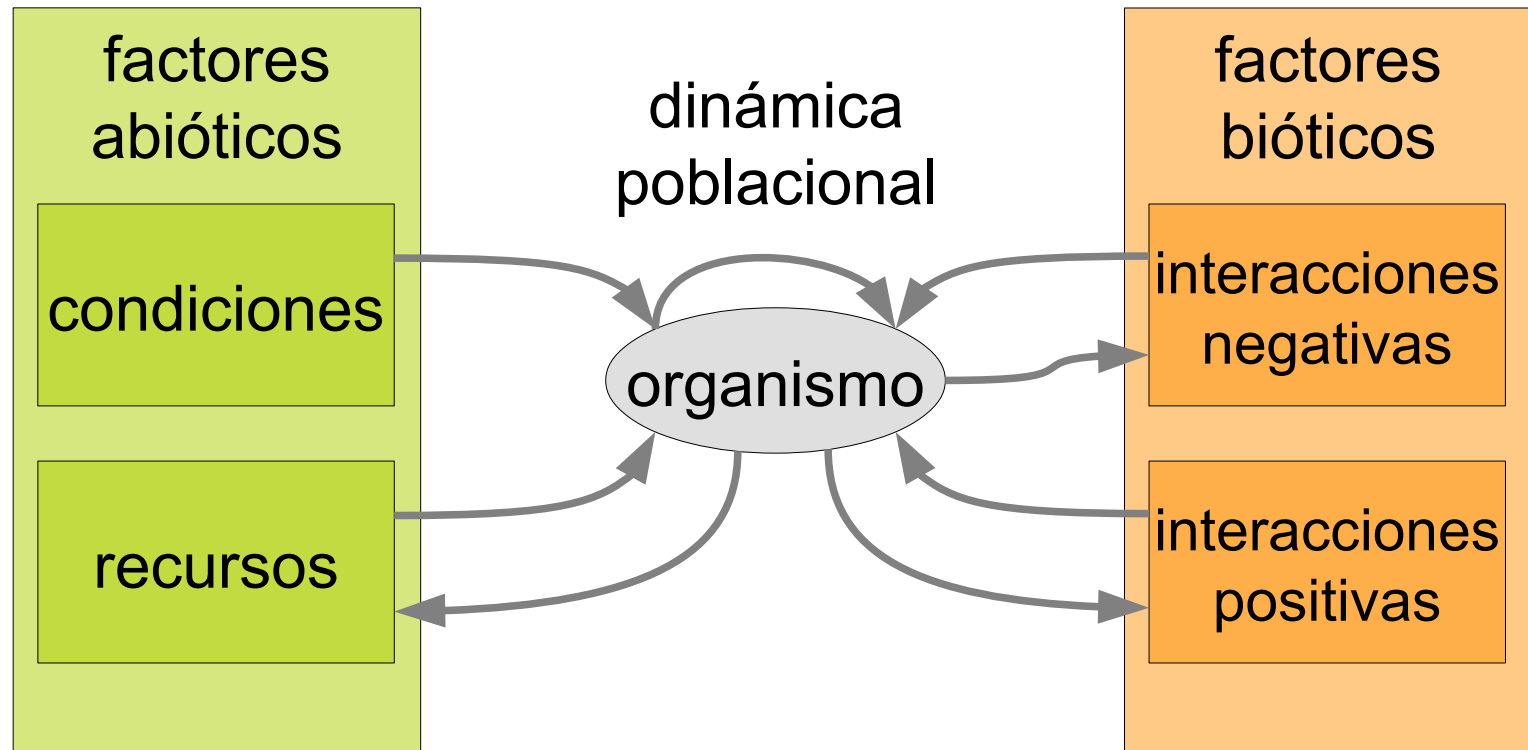
Predicted climatic suitability :



Model performance :



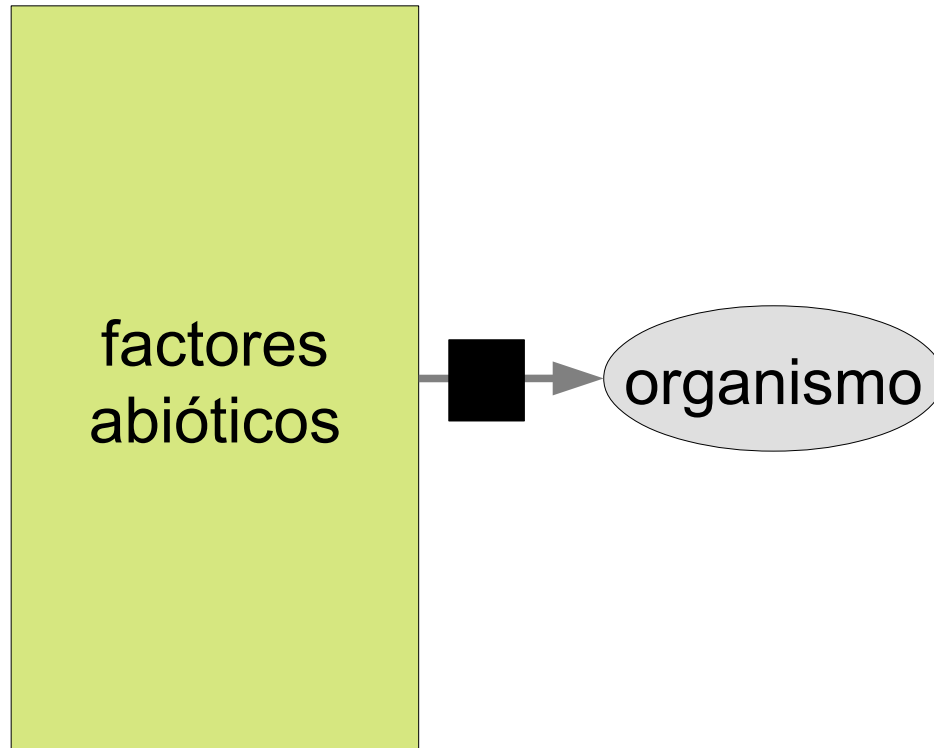
**¿QUÉ NECESITAMOS SABER  
PARA PREDECIR LA  
DISTRIBUCIÓN DE UNA ESPECIE?**



ver McInerny y Etienne 2012 J Biogeogr

# **TIPOS DE MODELOS PARA PREDECIR LA DISTRIBUCIÓN DE LAS ESPECIES**

# MODELOS CORRELATIVOS



■ = relación causal desconocida (blackbox)

# MODELOS CORRELATIVOS

## Frontiers in Zoology

Research

Open Access

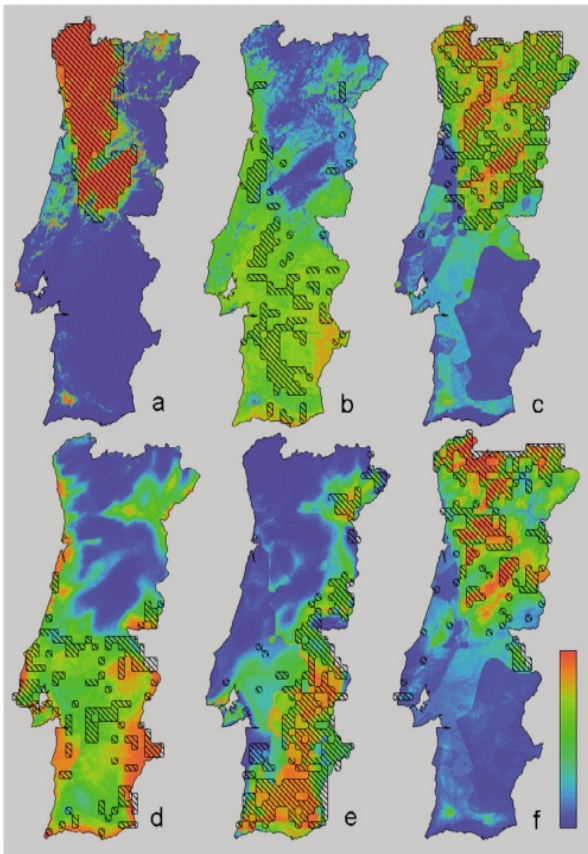
### From descriptive to predictive distribution models: a working example with Iberian amphibians and reptiles

JW Arntzen\*

#### Figure 1

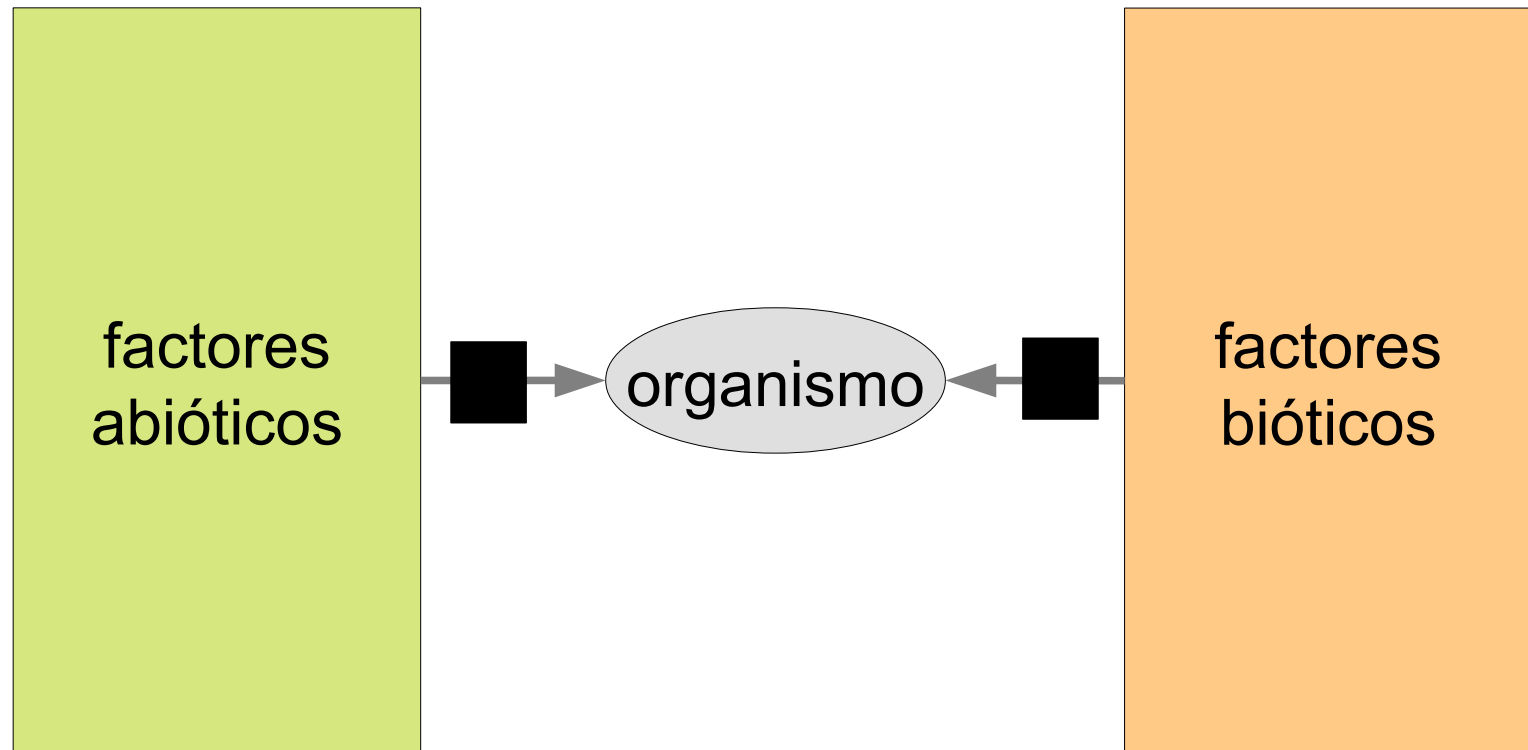
#### Descriptive distribution models for six amphibian species across Portugal.

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

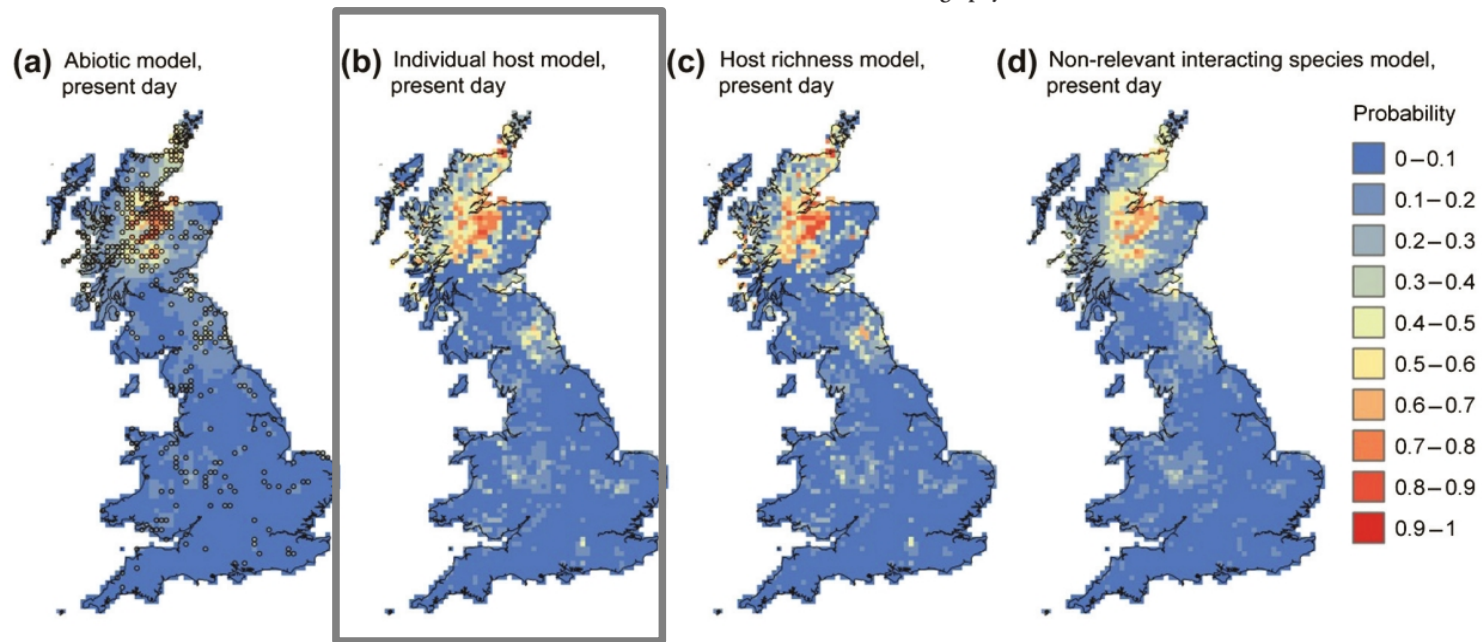
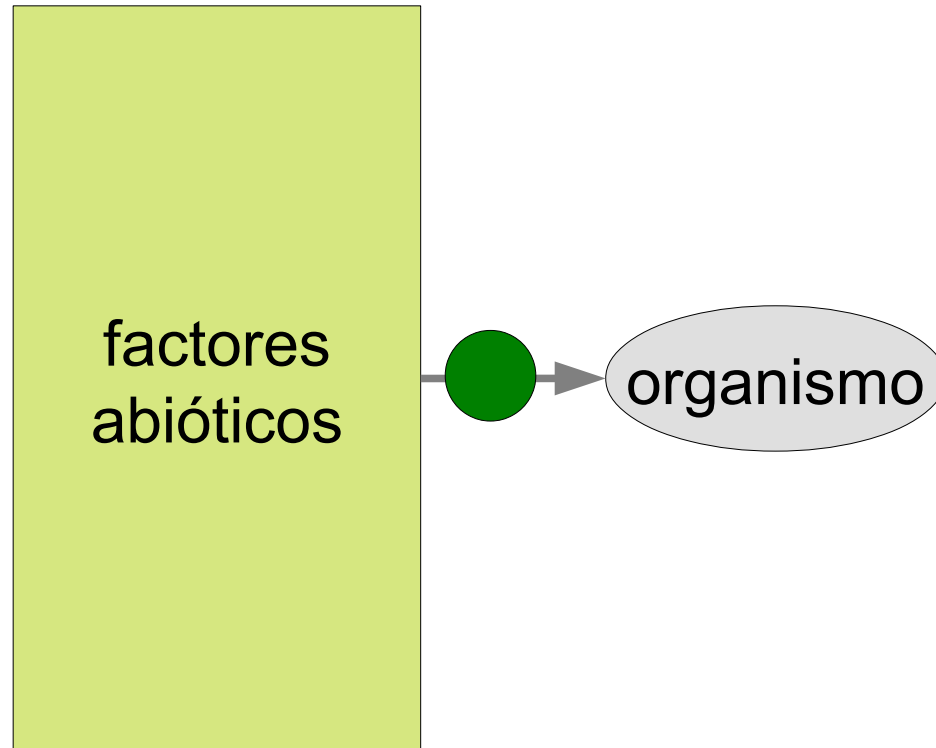


Figure 2. Projections of *Bombus bohemicus* occurrence areas. Top: present climatic conditions using (a) abiotic information only; (b) the most widespread host *B. lucorum*; (c) richness of its three host species (*B. lucorum*, *B. magnus* and *B. cryptarum*); (d) non-relevant interacting species (*B.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

# MODELOS ECO-FISIOLÓGICOS



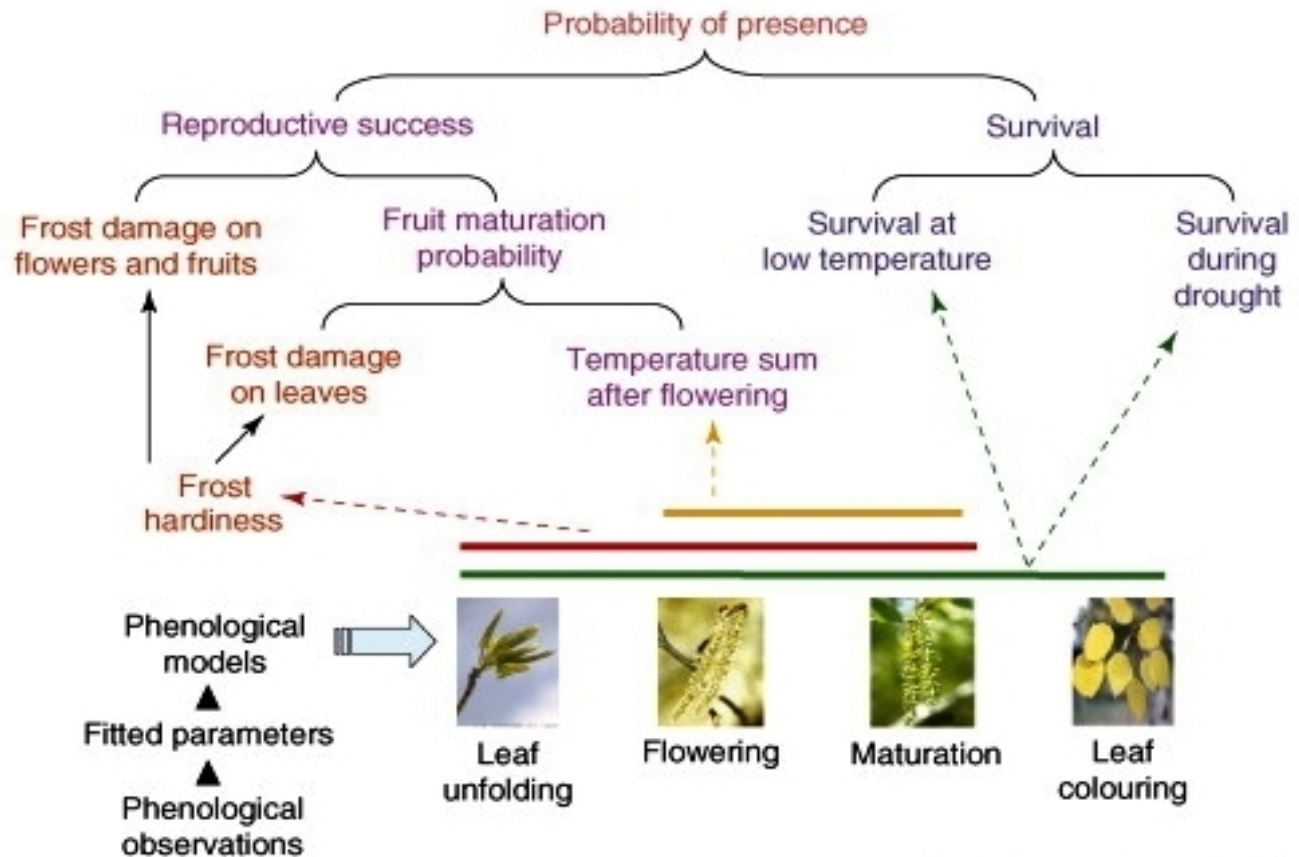
 = mecanismo ecológico explícito

# MODELOS ECO-FISIOLÓGICOS

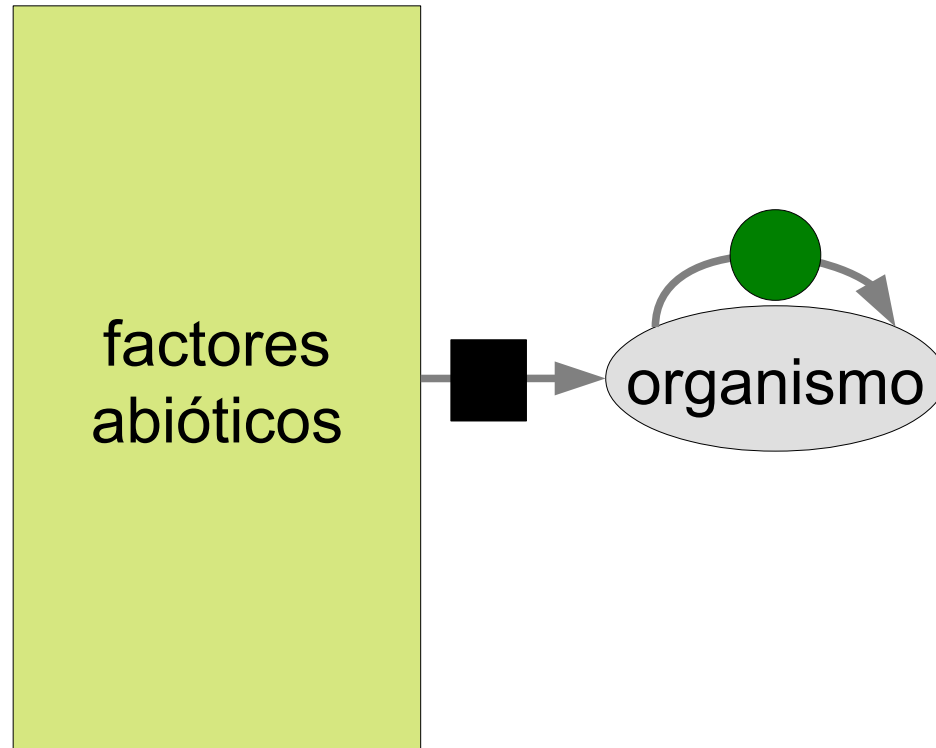


## Shifting plant phenology in response to global change

Elsa E. Cleland<sup>1</sup>, Isabelle Chuine<sup>2</sup>, Annette Menzel<sup>3</sup>, Harold A. Mooney<sup>4</sup> and Mark D. Schwartz<sup>5</sup>



# MODELOS HÍBRIDOS



● = mecanismo ecológico explícito

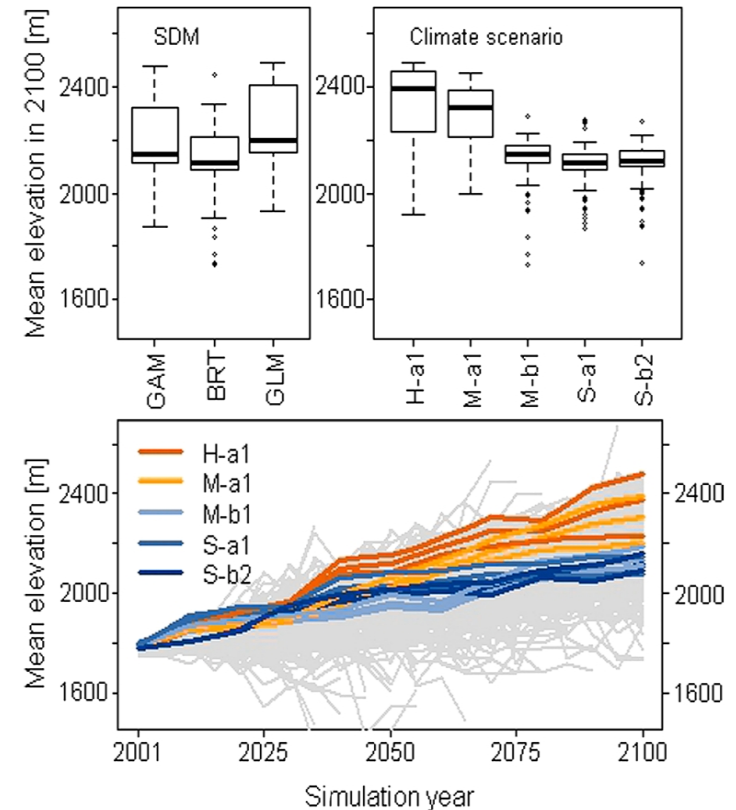
# MODELOS HÍBRIDOS

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

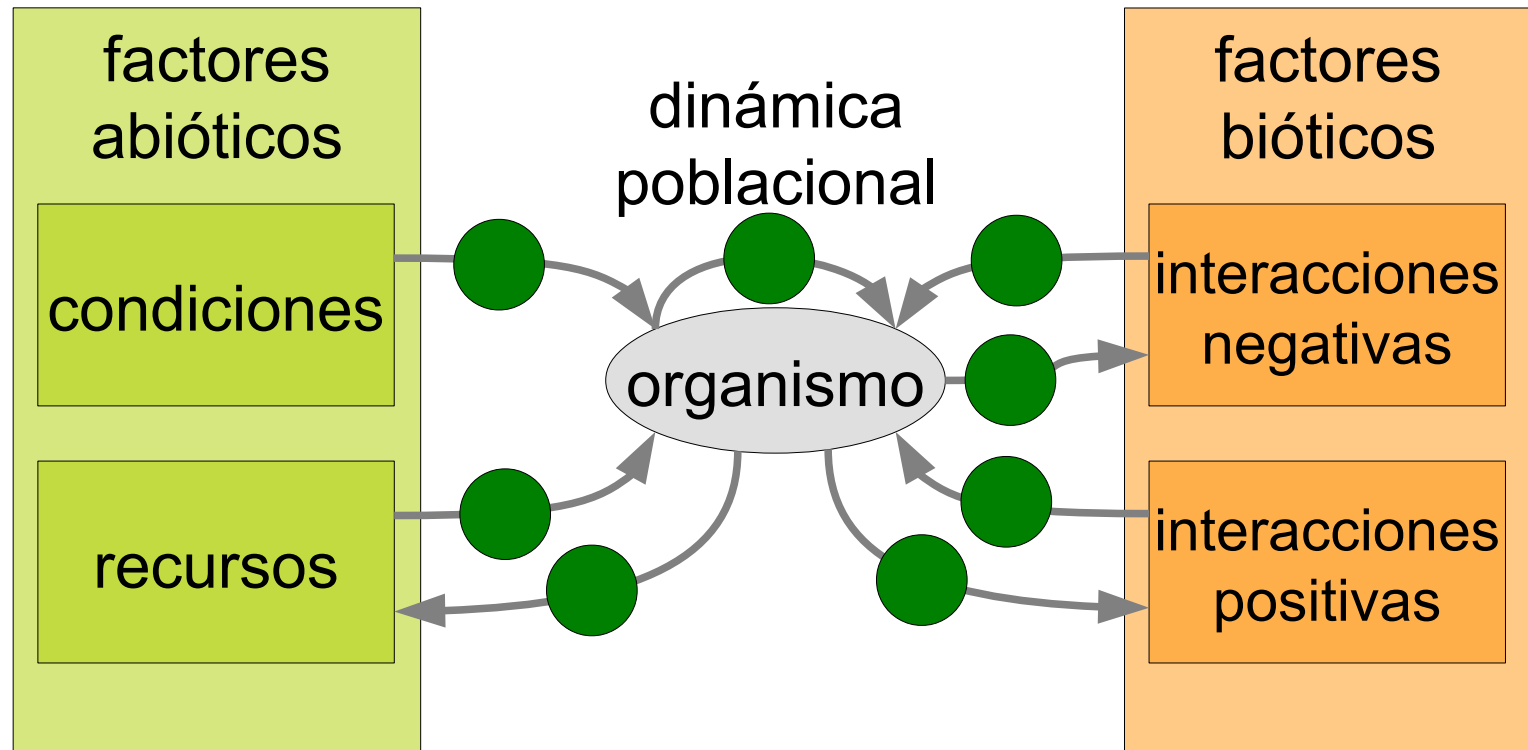
Damaris Zurell, Volker Grimm, Eva Rossmannith, Niklaus Zbinden, Niklaus E. Zimmermann and Boris Schröder

*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.



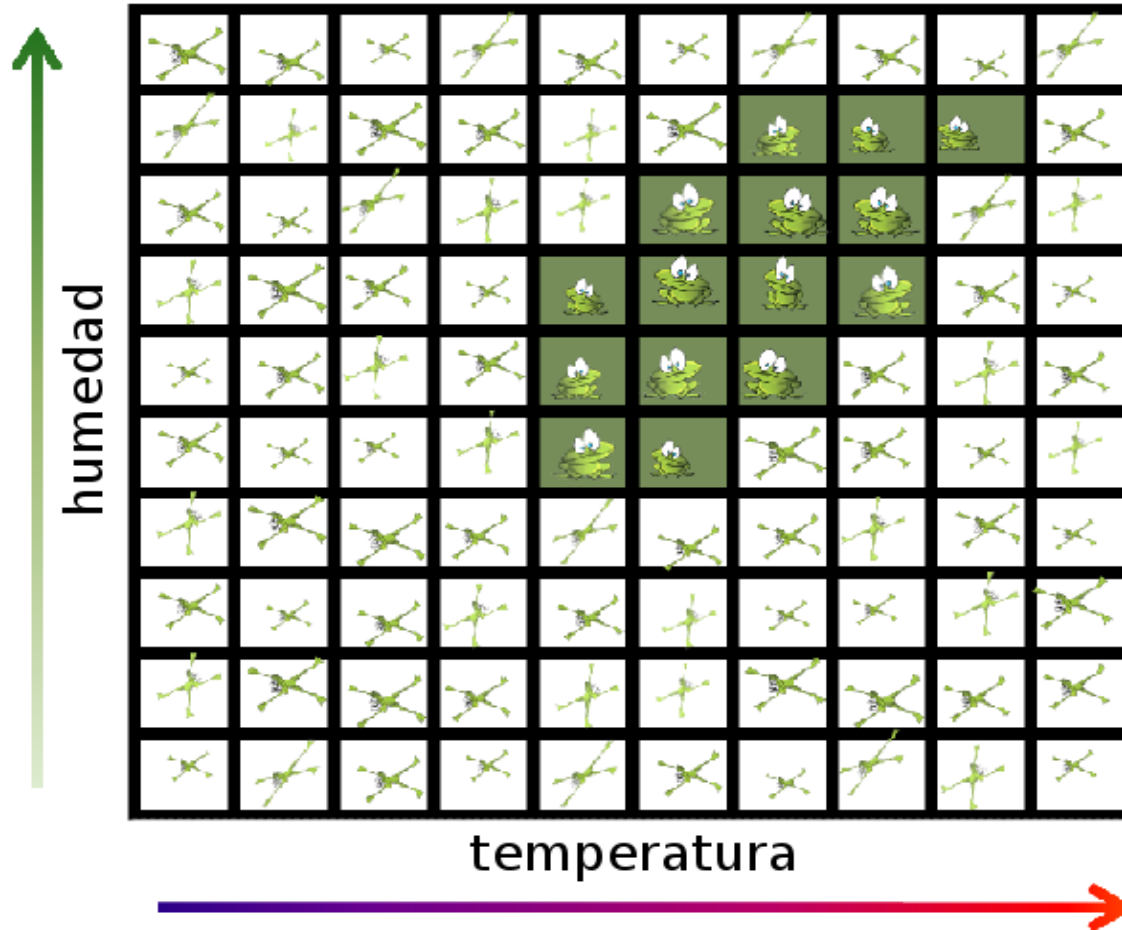
# MODELO IDEAL



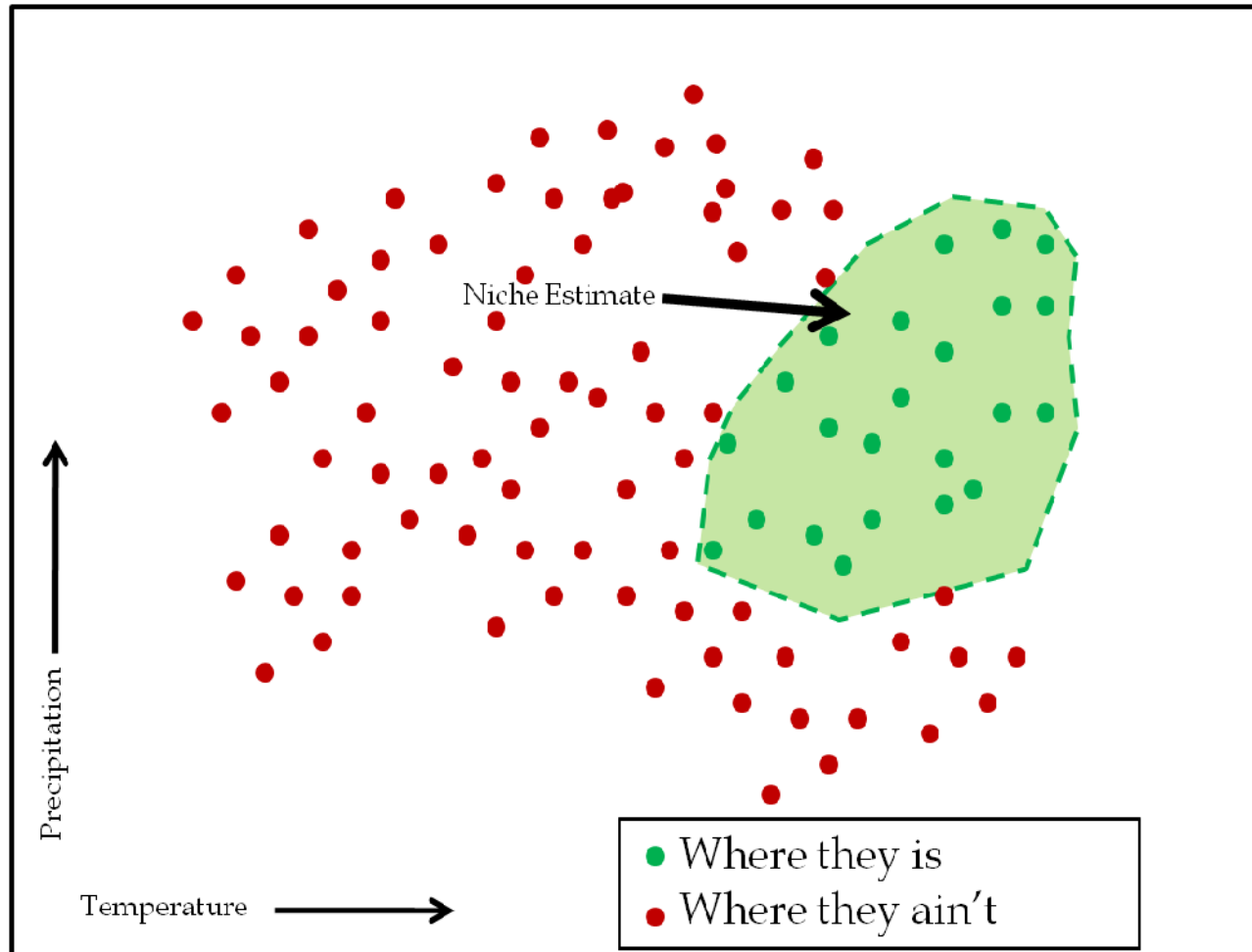
# MODELOS CORRELATIVOS



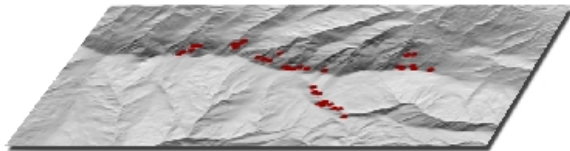
# DETERMINACIÓN DIRECTA



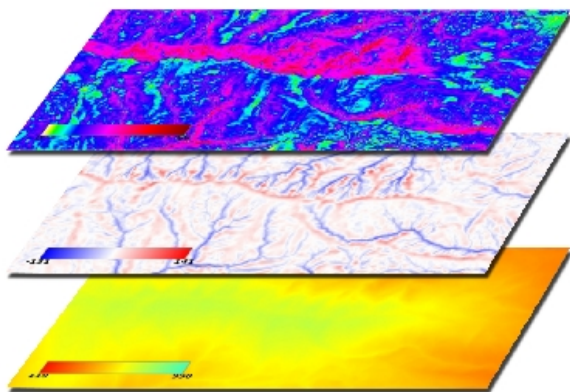
# DETERMINACIÓN INDIRECTA



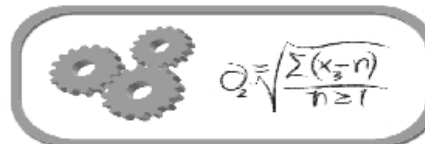
REGISTROS DE PRESENCIA



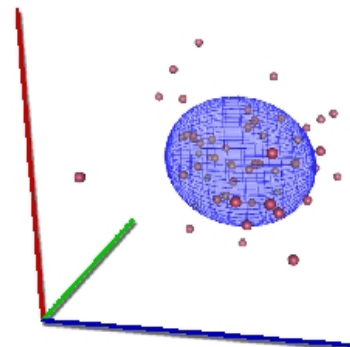
VARIABLES AMBIENTALES



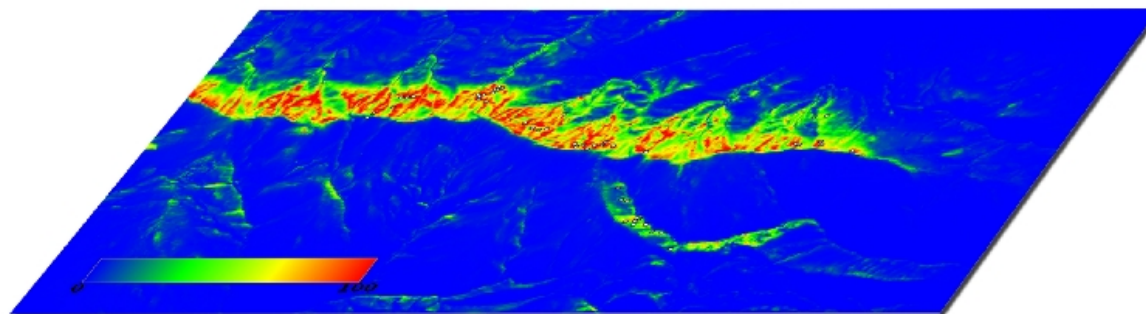
ALGORITMO



MODELO (espacio ecológico)



MODELO (espacio geográfico)



# 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?

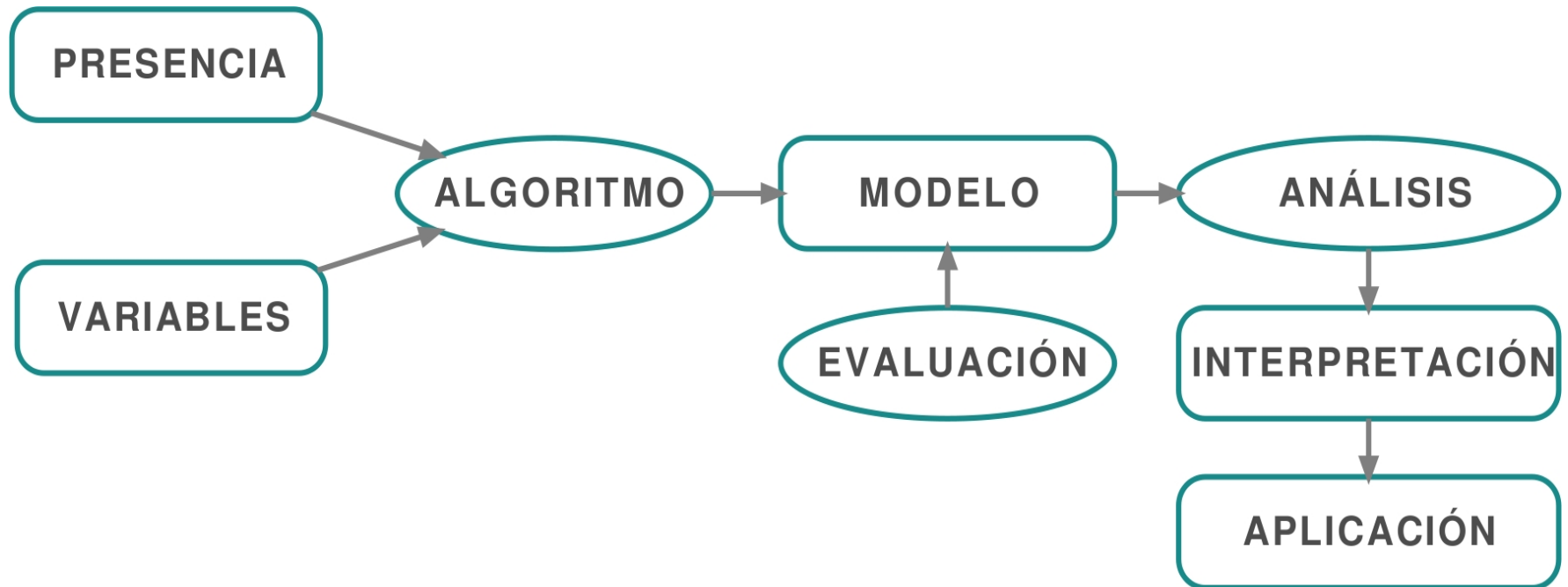
- No se sabe qué especies tienen nichos dinámicos, ni qué condiciones ecológicas que lo hacen posible
- Algunos estudios indican cambios en periodos menores a 100 años (¿especiación o adaptación?)
- Nicho dinámico posibilita la especiación, por lo que debe haber especies con nicho dinámico
- Las grandes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

# AJUSTANDO UN MODELO DE DISTRIBUCIÓN CORRELATIVO

# 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



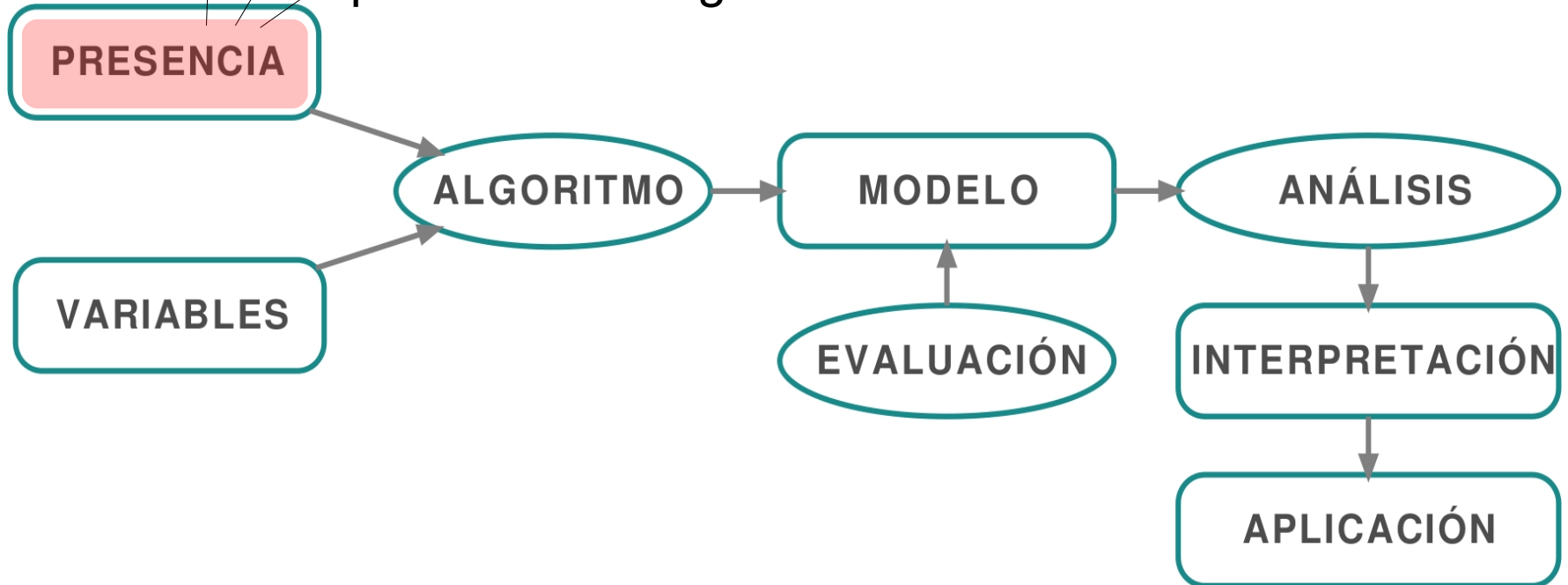


# PRESENCIA

tamaño de muestra

sesgo de los datos

precisión de la georreferenciación



# UN REGISTRO DE PRESENCIA...

Debería contener, al menos:

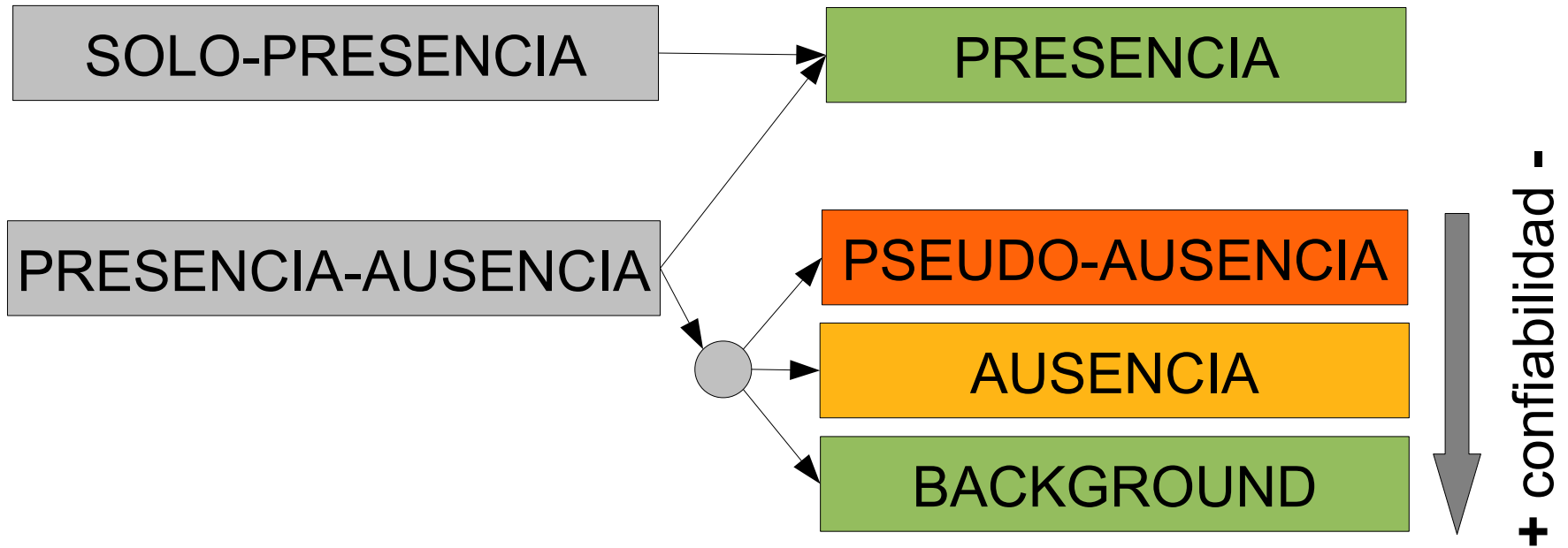
- Una coordenada X
- Una coordenada Y
- Un nombre de taxon (correctamente identificado)

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

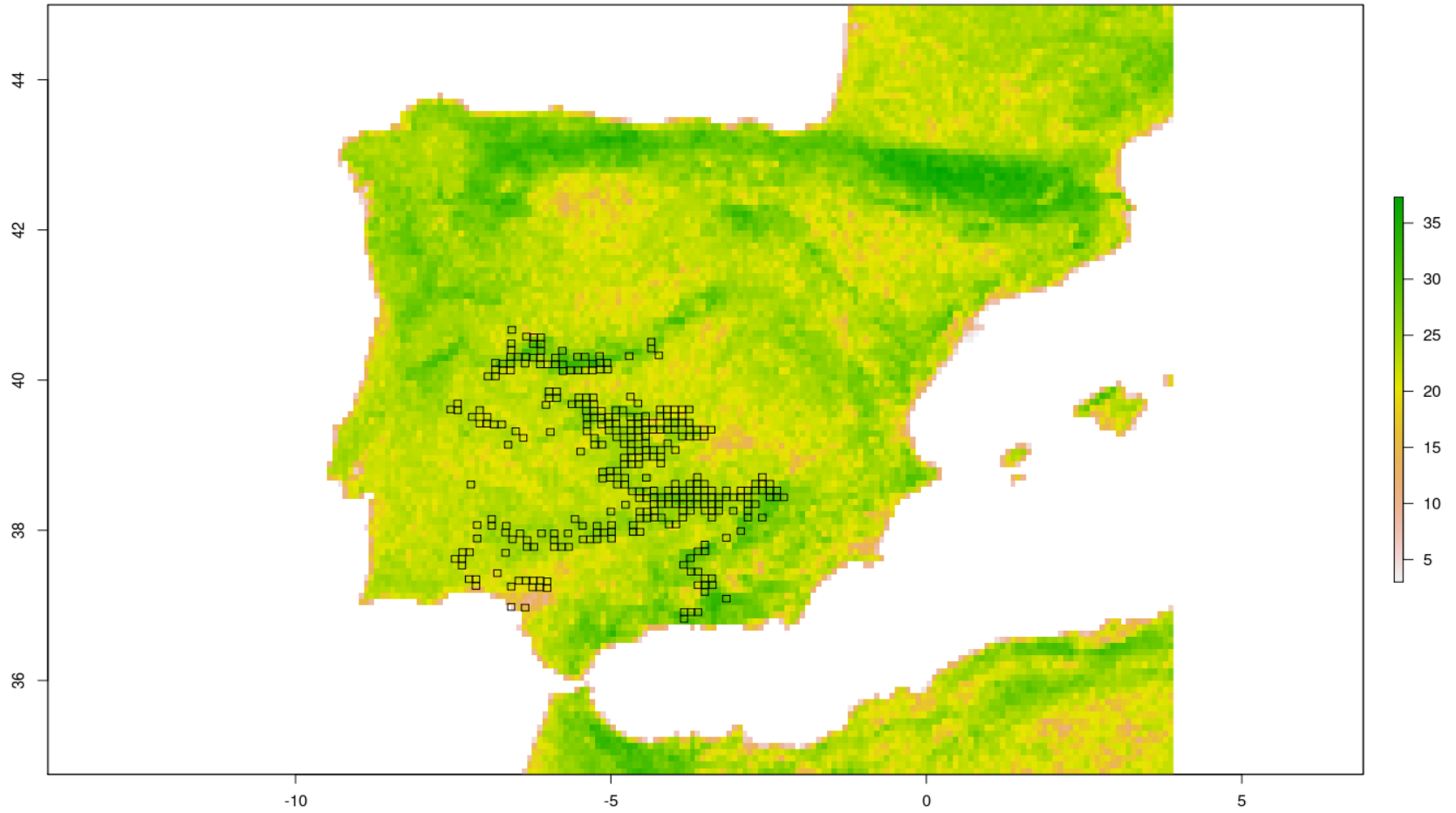
## VENTAJAS

- 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

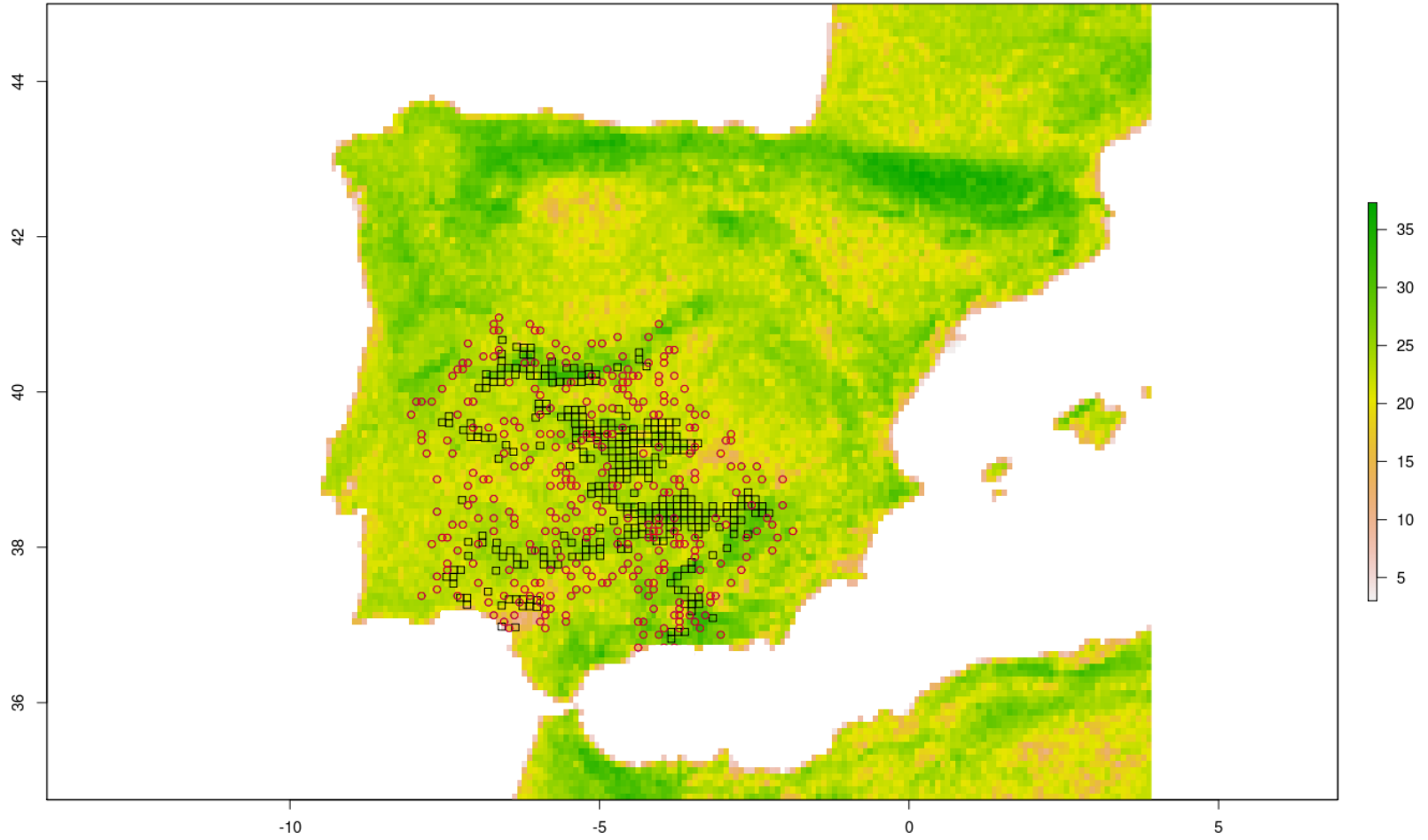
# AUSENCIA

## 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

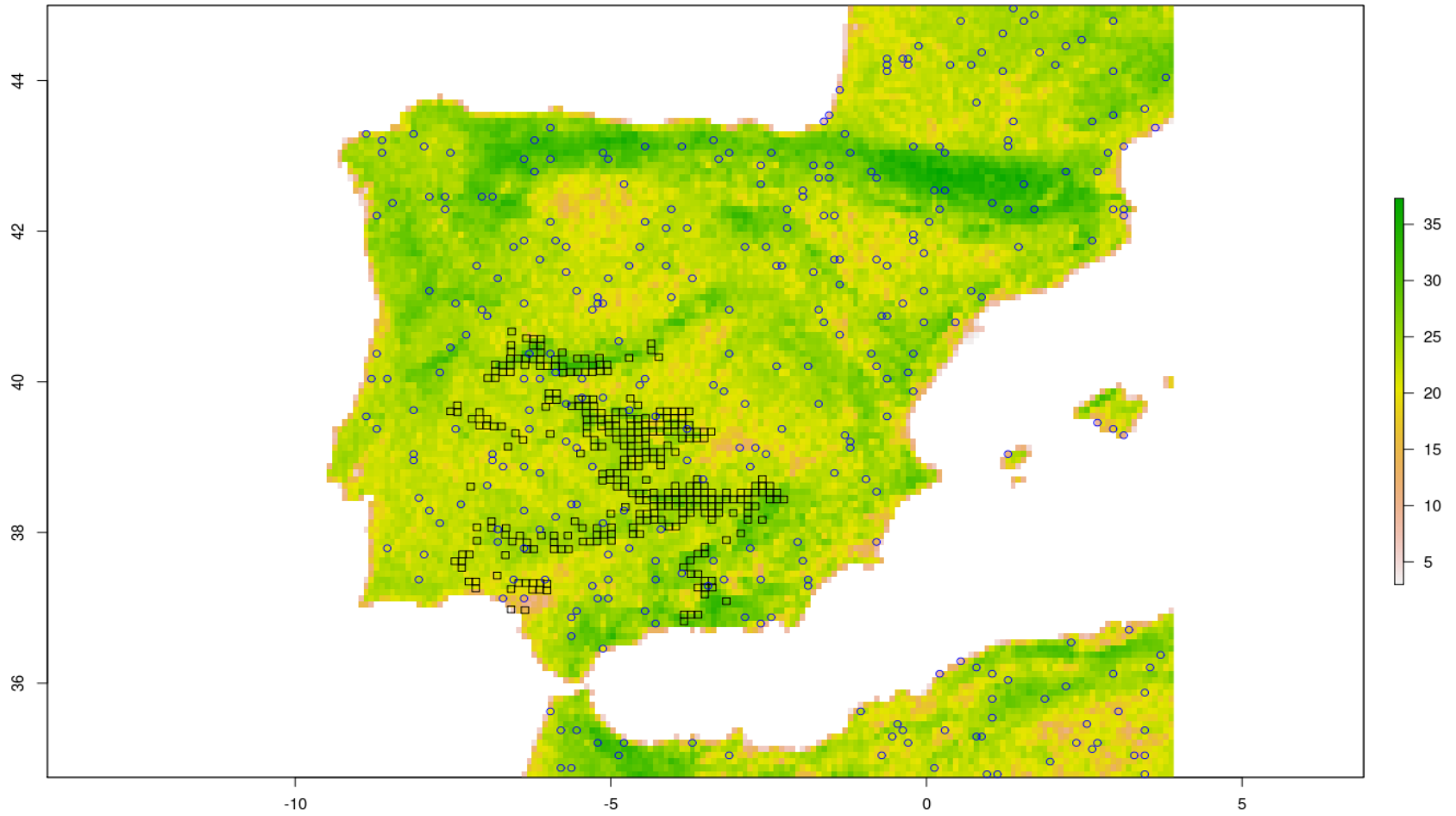
## VENTAJAS

- Es fácil generarlas

## INCONVENIENTES

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

# PSEUDO-AUSENCIA



# BACKGROUND:

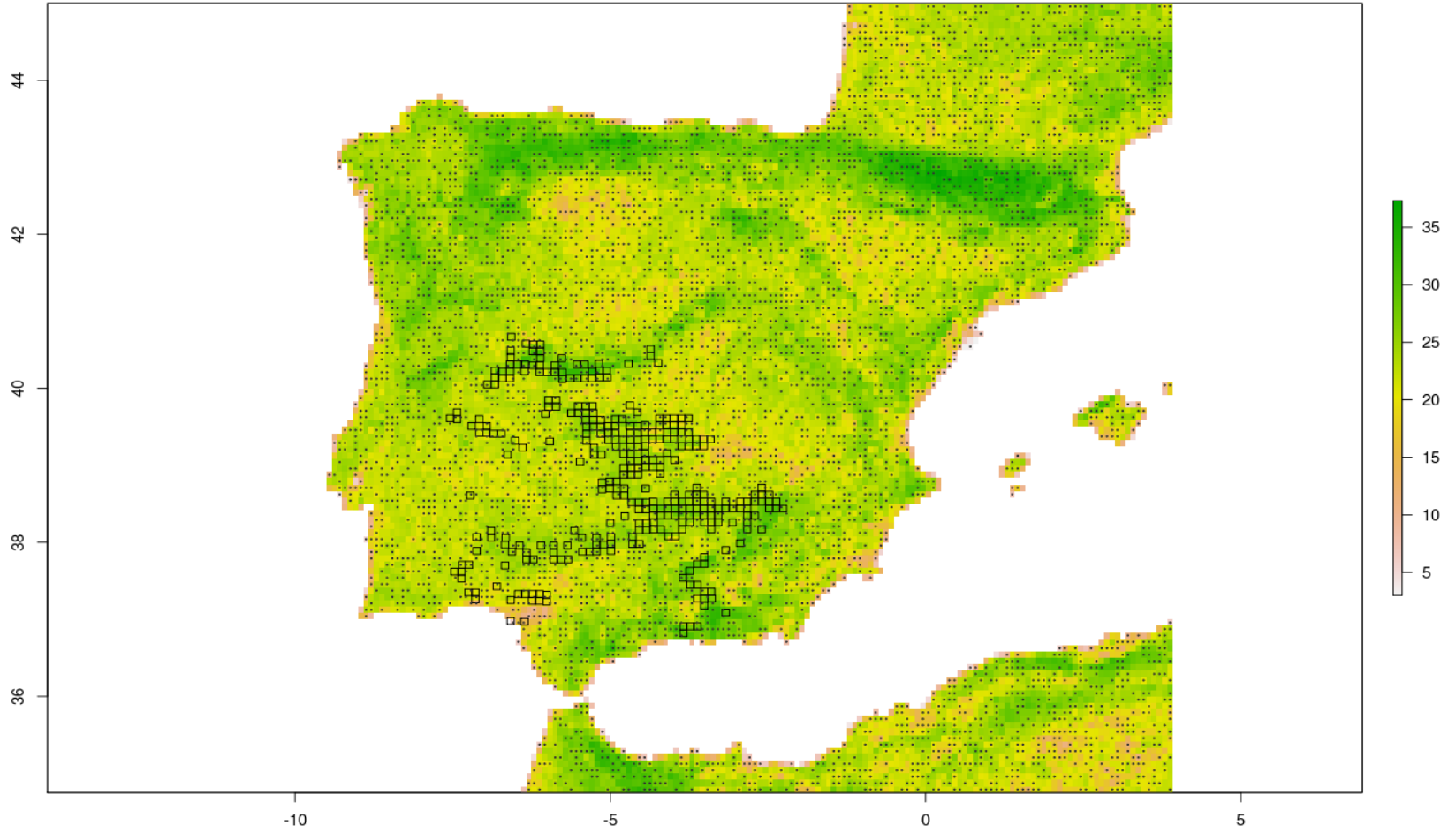
## VENTAJAS

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

## INCONVENIENTES

- No funcionan bien para métodos de que no permitan incluir pesos (weights) en el modelo.
- 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.

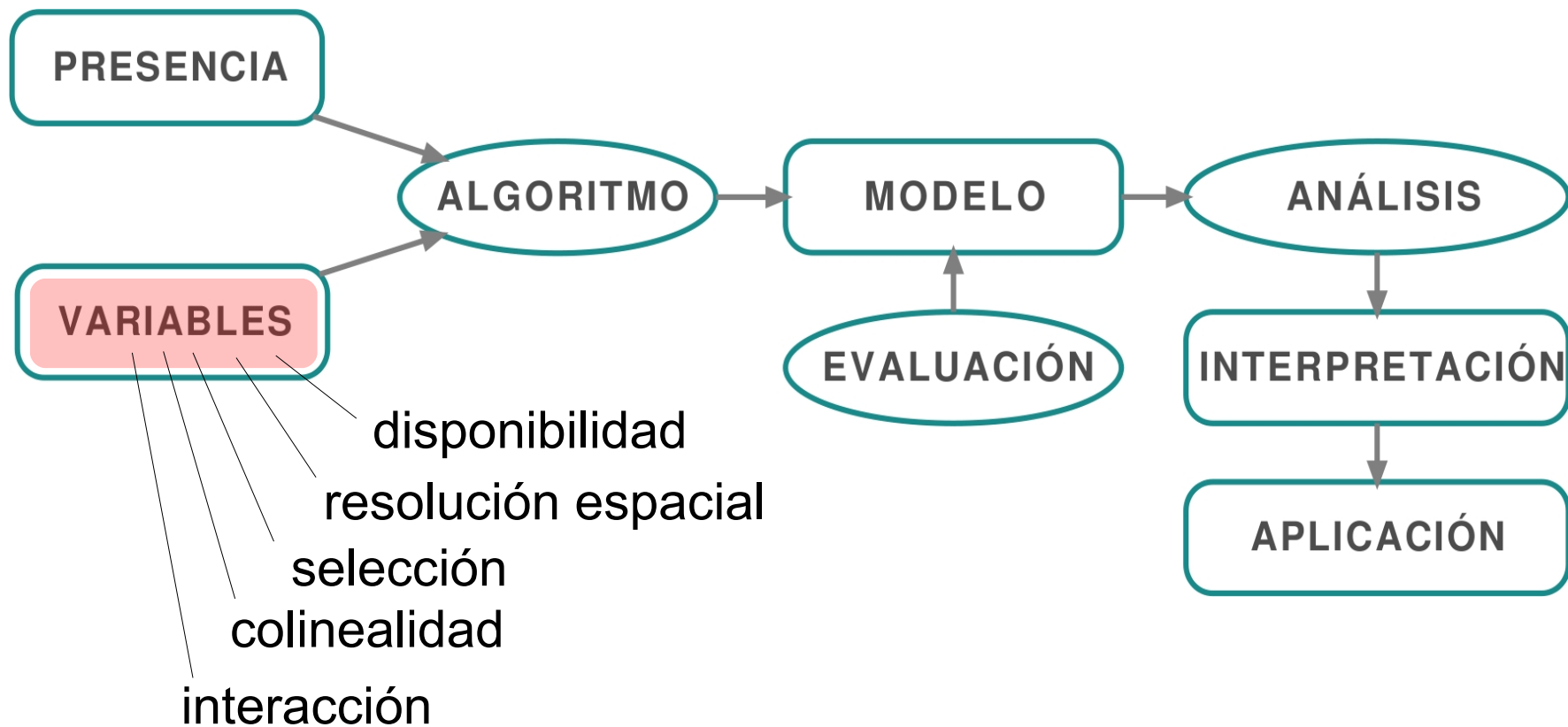
# AUTOCORRELACIÓN ESPACIAL

- **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

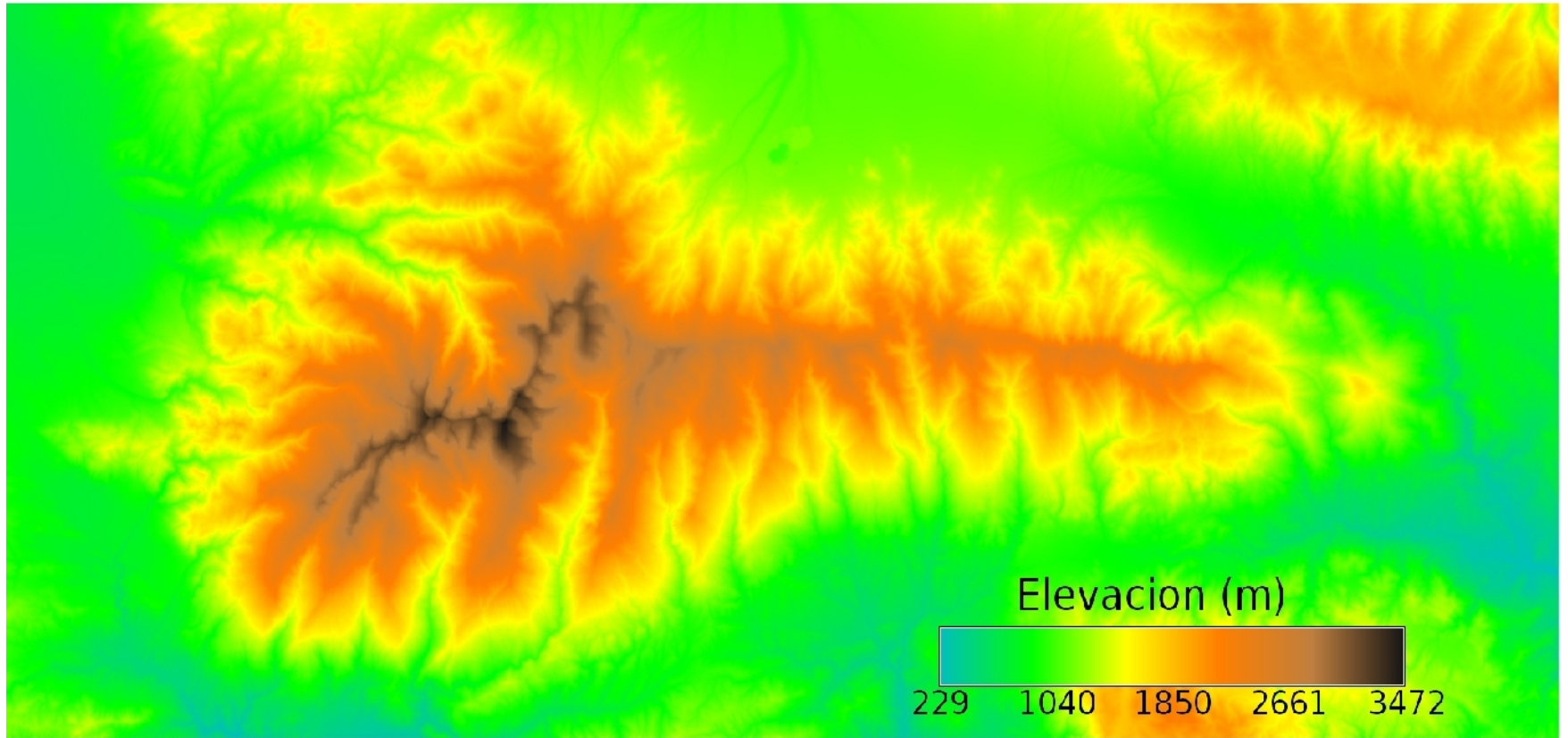




# VARIABLES AMBIENTALES

- Deben tener cuatro características:
  - Afectar directa o indirectamente a la distribución de la especie
  - Mapas raster
  - 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

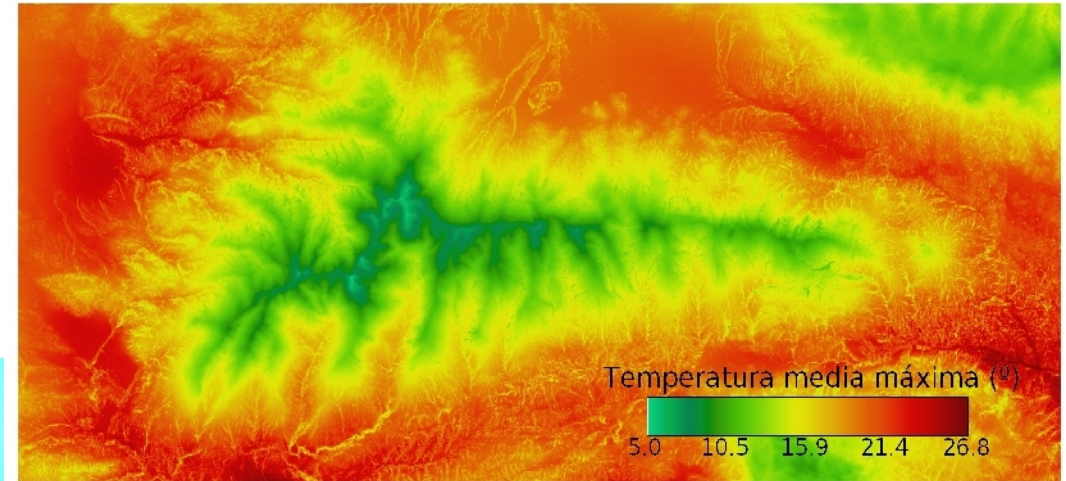
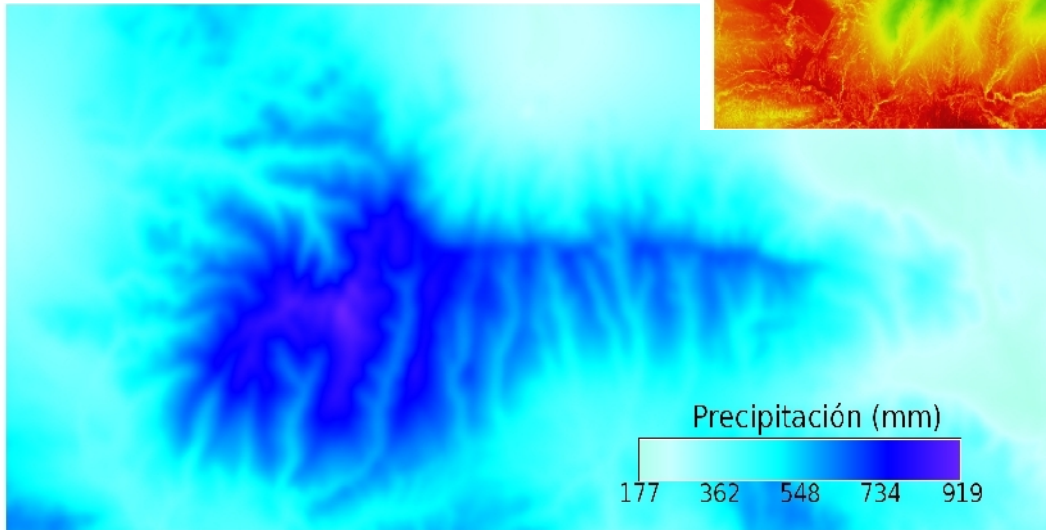
# VARIABLES TOPOGRÁFICAS



# TOPOGRAFÍA: FUENTES

- **STRM:** <http://srtm.csi.cgiar.org/>
- **ASTER-GDEM:** <https://lpdaac.usgs.gov/>
- **EU-DEM:**  
<https://land.copernicus.eu/imagery-in-situ/eu-dem>
- **Bio-Oracle** (batimetría) <http://www.bio-oracle.org/>
- **MARSPEC** (batimetría) <http://marspec.weebly.com/>
- **Centro Nacional de Información Geográfica:**  
<http://centrodedescargas.cnig.es/CentroDescargas/>

# VARIABLES CLIMÁTICAS



# CLIMA: FUENTES

• **Worldclim:** <https://www.worldclim.org/>

• **CHELSA:** <http://chelsa-climate.org/>

• **MERRAclim:**

<https://datadryad.org/resource/doi:10.5061/dryad.s2v81>

• **CLImond:** <https://www.climond.org>

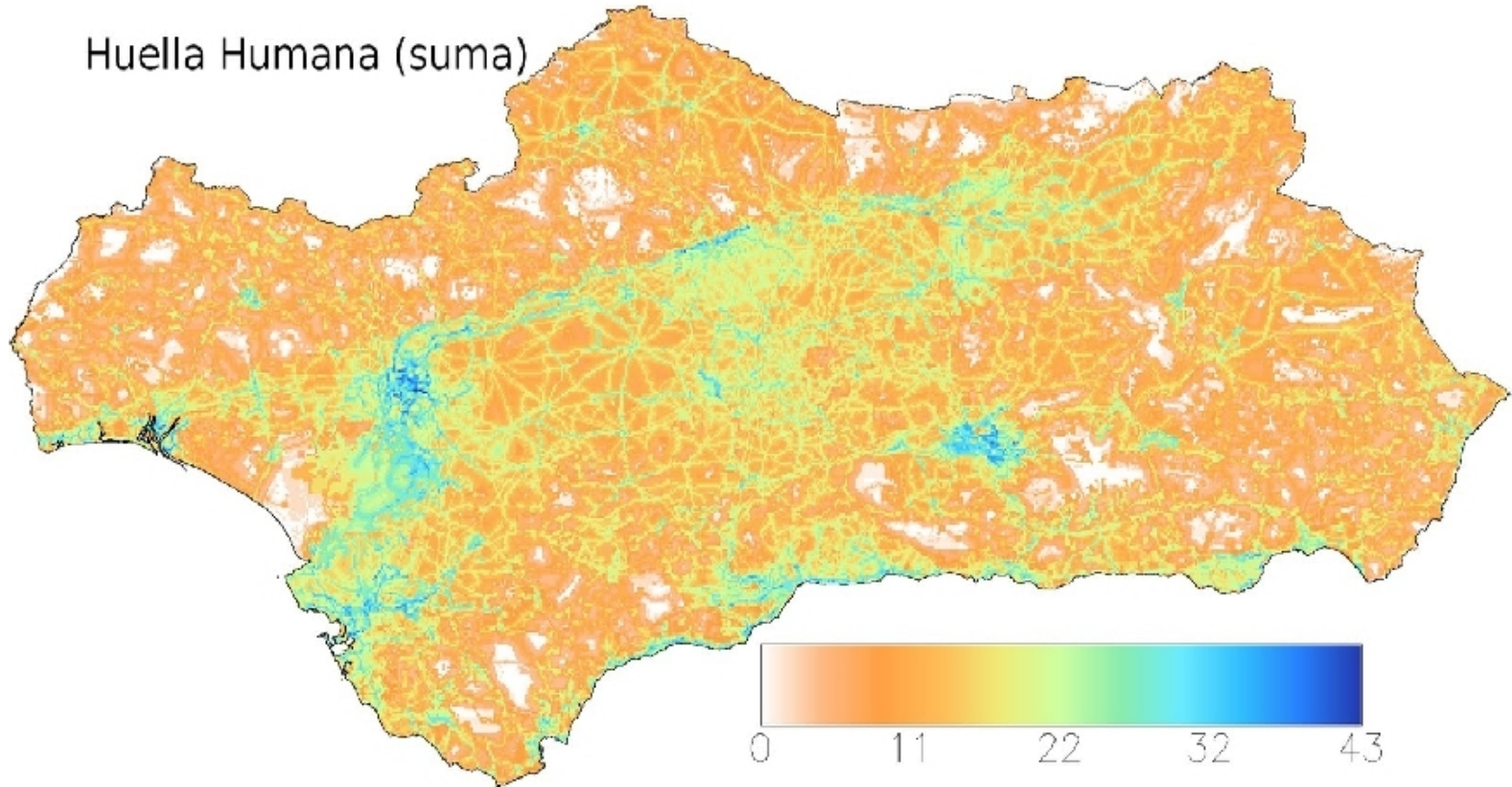
• **Bio-Oracle** (medio marino) <http://www.bio-oracle.org/> y  
paquete de R “sdmpredictors”

• **MARSPEC** (medio marino) <http://marspec.weebly.com/>



# VARIABLES ANTRÓPICAS

Huella Humana (suma)



The Last of the Wild (human footprint):

<https://sedac.ciesin.columbia.edu/data/set/wildareas-v3-2009-human-footprint>

# TELEDETECCIÓN





# MODIS (Terra and Aqua)

- Land products
  - Land surface temperature
  - Land cover
  - Vegetation index (NDVI and EVI)
  - Leaf area index
  - Evapotranspiration
  - Gross primary productivity
- Ocean products
  - Sea surface temperature
  - Reflectance
  - Chlorophyll-a concentration
  - Organic and inorganic particulate carbon.
  - Photosynthetically available radiation

Source: <https://modis.ornl.gov/data.html>



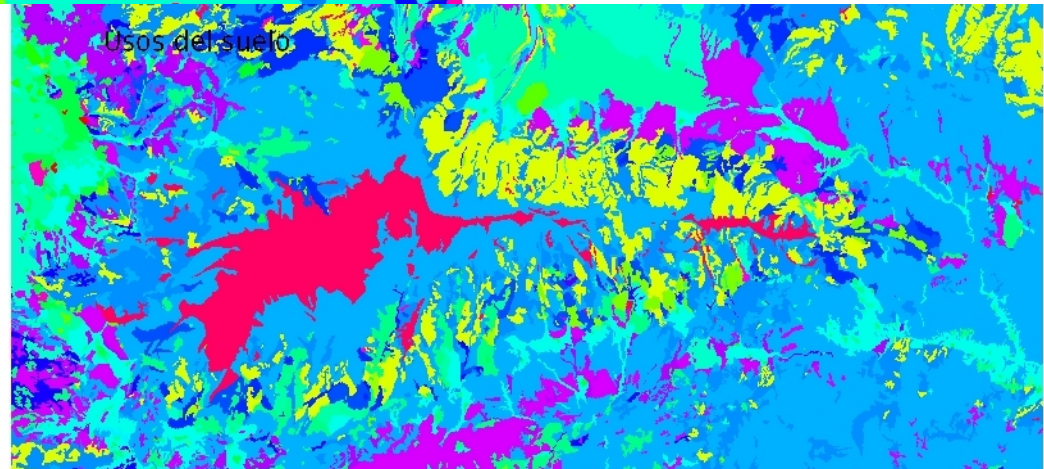
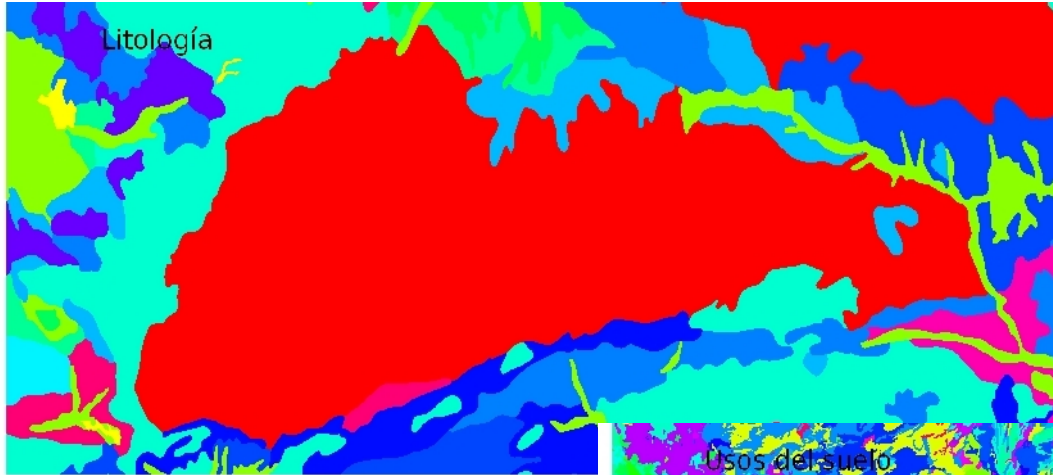
# Landsat



Source:

[https://www.usgs.gov/land-resources/nli/landsat/landsat-data-access?qt-science\\_support\\_page\\_related\\_con=0#qt-science\\_support\\_page\\_related\\_con](https://www.usgs.gov/land-resources/nli/landsat/landsat-data-access?qt-science_support_page_related_con=0#qt-science_support_page_related_con)

# VARIABLES TEMÁTICAS



# GEOLOGÍA Y SUELOS: FUENTES

- Corine European soil database version 2

<https://www.eea.europa.eu/data-and-maps/data/soil-type>

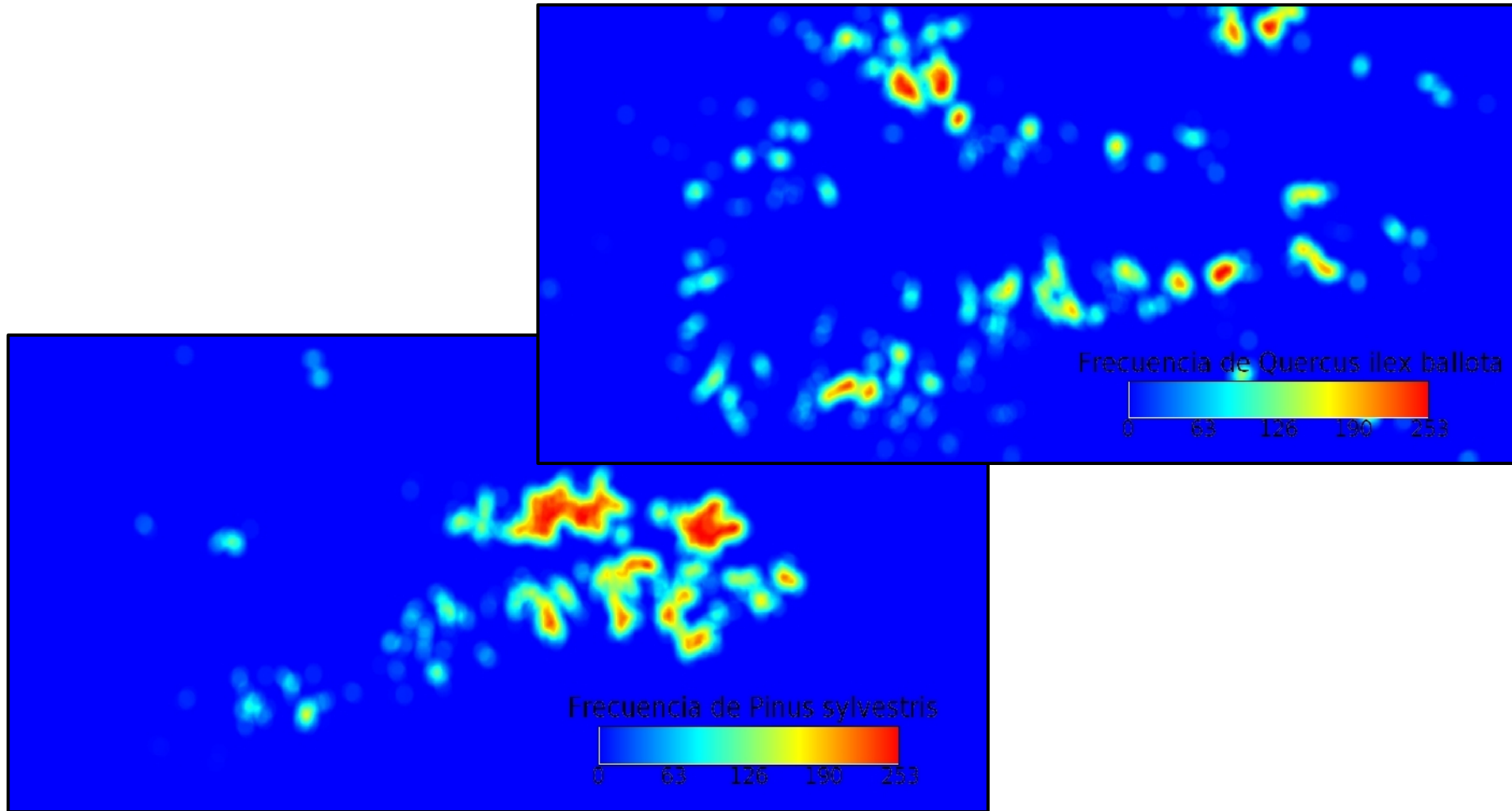
- European Soil Database & soil properties

<https://esdac.jrc.ec.europa.eu/resource-type/european-soil-database-soil-properties>

- Topsoil physical properties for Europe

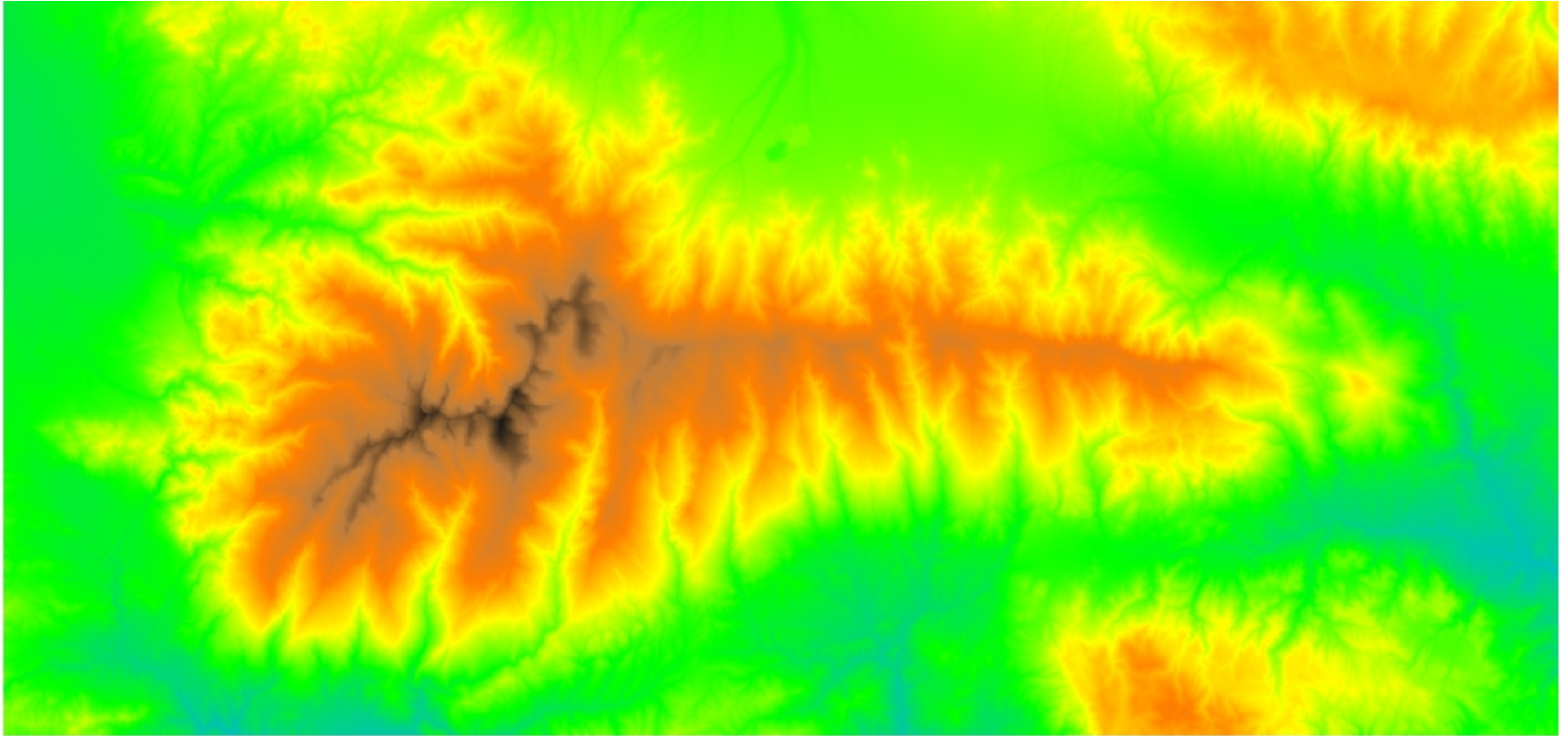
<https://esdac.jrc.ec.europa.eu/content/topsoil-physical-properties-europe-based-lucas-topsoil-data>

# VARIABLES BIOLÓGICAS

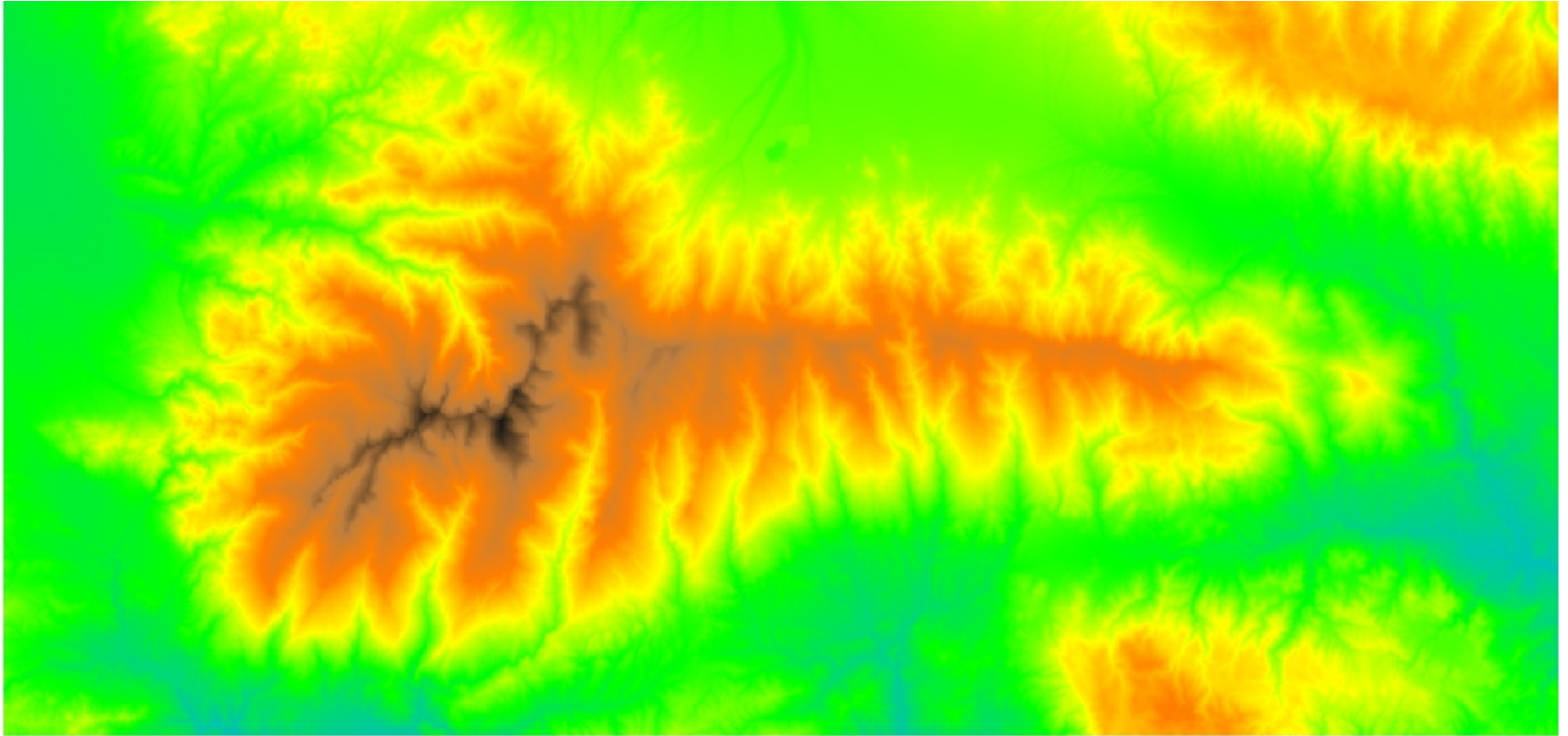




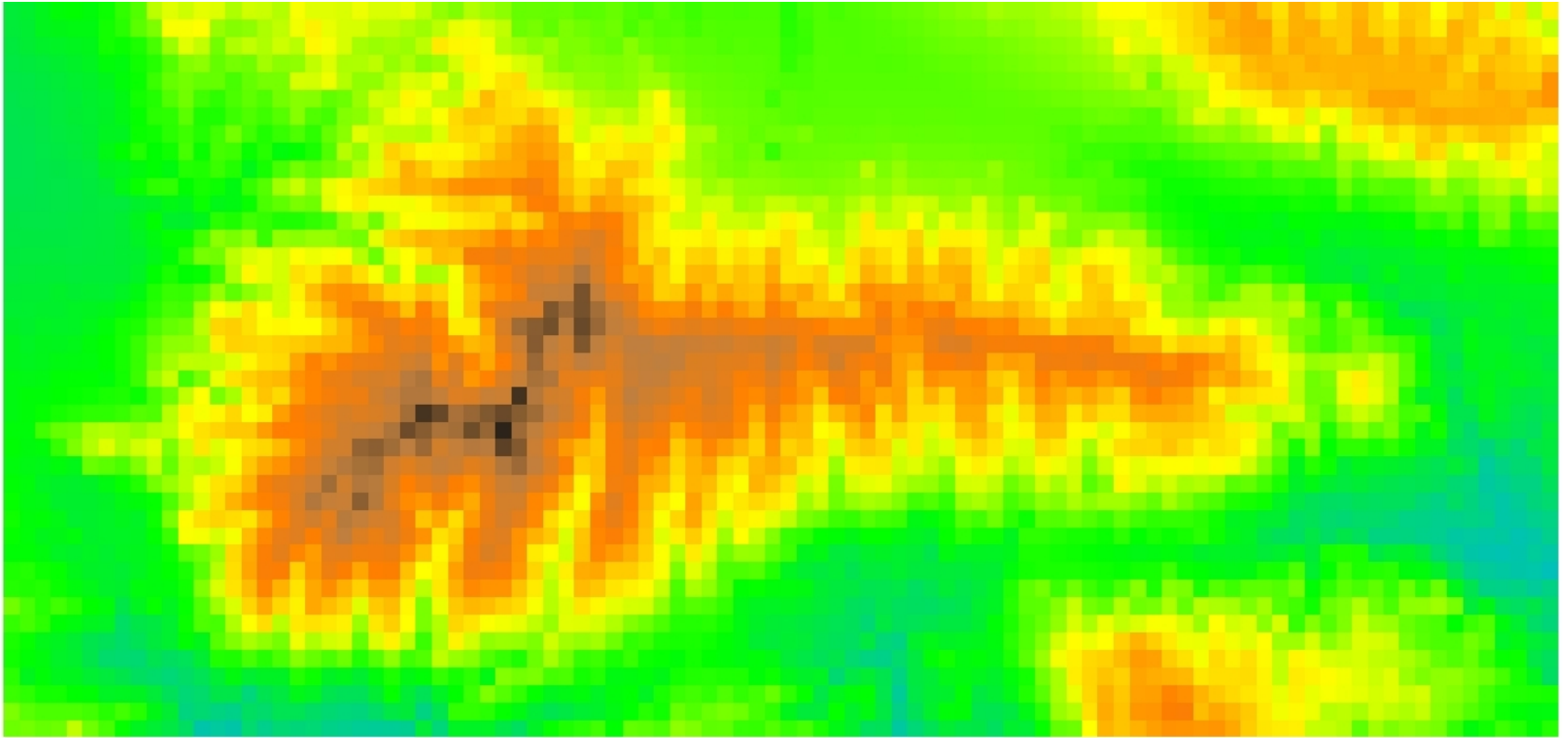
# RESOLUCIÓN (10 m)



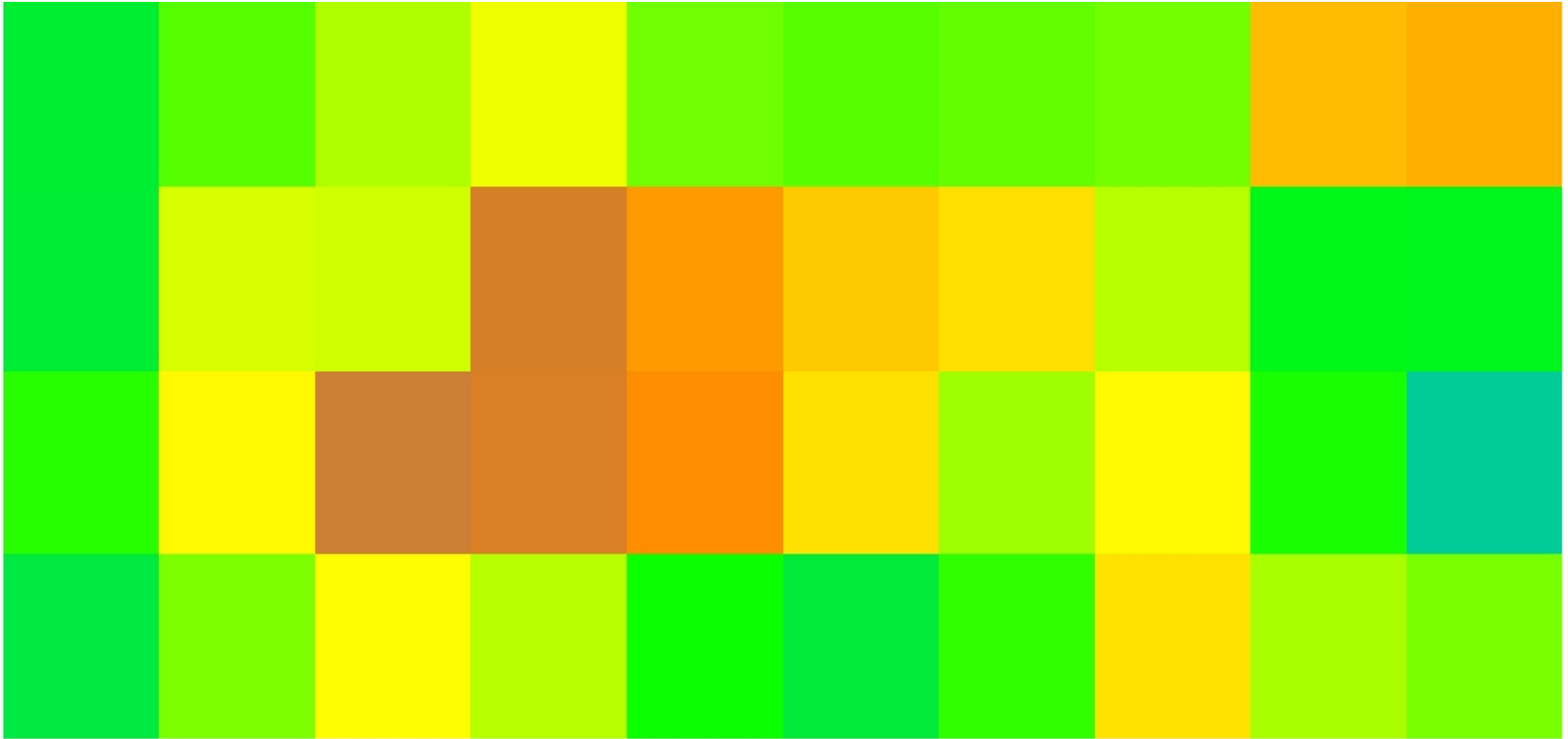
# RESOLUCIÓN (100 m)



# RESOLUCIÓN (1000 m)



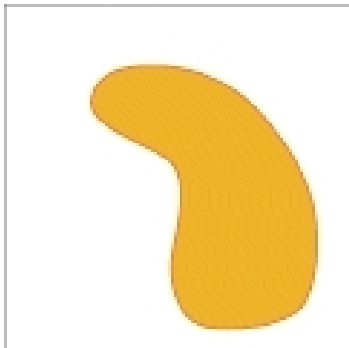
# RESOLUCIÓN (10000 m)



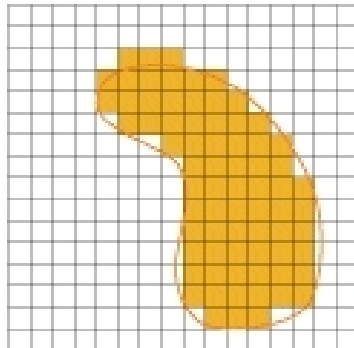


# RESOLUCIÓN

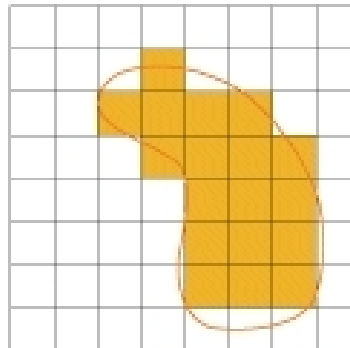
71 m<sup>2</sup>  
polygon



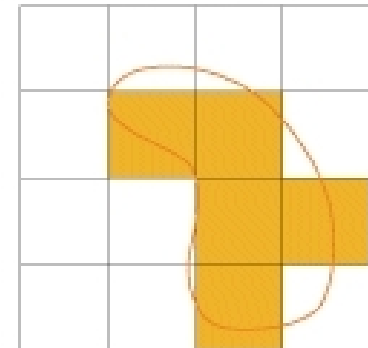
73 m<sup>2</sup>  
1 m cell  
16 x 16 cells



72 m<sup>2</sup>  
2 m cell  
8 x 8 cells



80 m<sup>2</sup>  
4 m cell  
4 x 4 cells

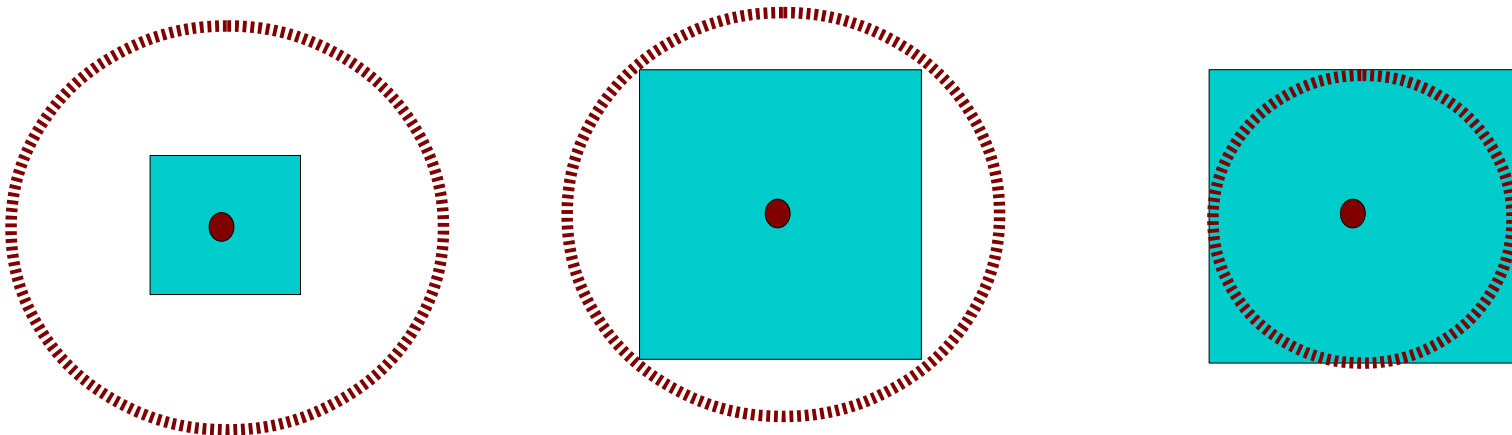


- Smaller cell size
- Higher resolution
- Higher feature spatial accuracy
- Slower display
- Slower processing
- Larger file size

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

# IMPORTANTE

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

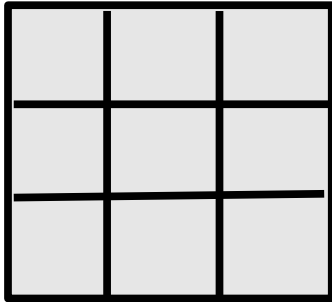


# EXTENSIÓN Y RESOLUCIÓN

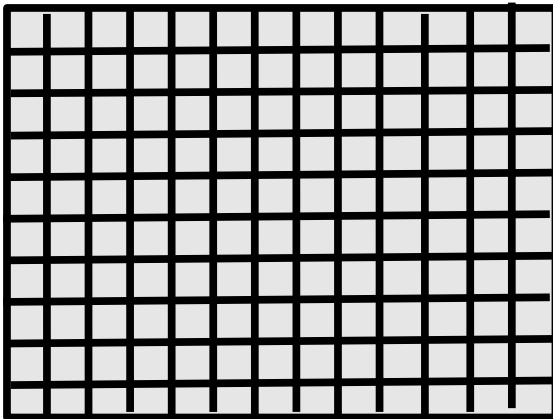
- **Extensión** y resolución nos permiten definir las características del área de trabajo.
- La extensión a elegir depende del área de distribución de la especie.
- La resolución depende de la calidad de datos disponibles.

# 1 – ELEGIR MAPA DE REFERENCIA

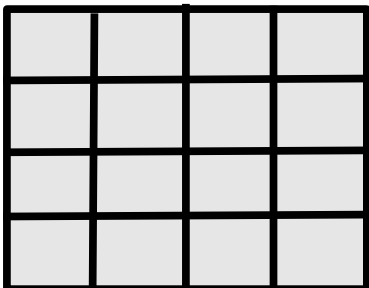
A



B



C

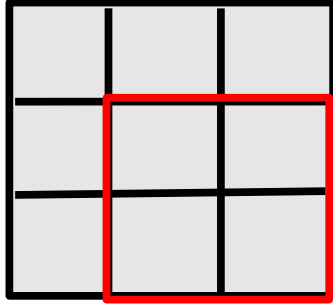


Depende de:

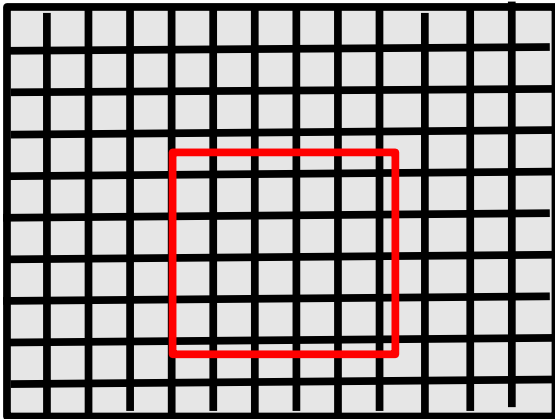
- Resolución de las presencias
- Resolución de la variable más importante para el modelo

## 2 – RECORTAR EXTENSIÓN

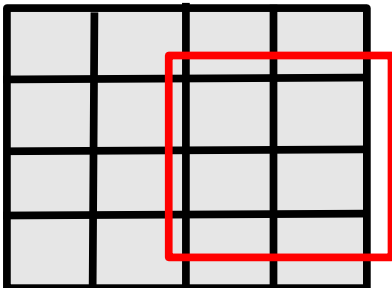
A



B

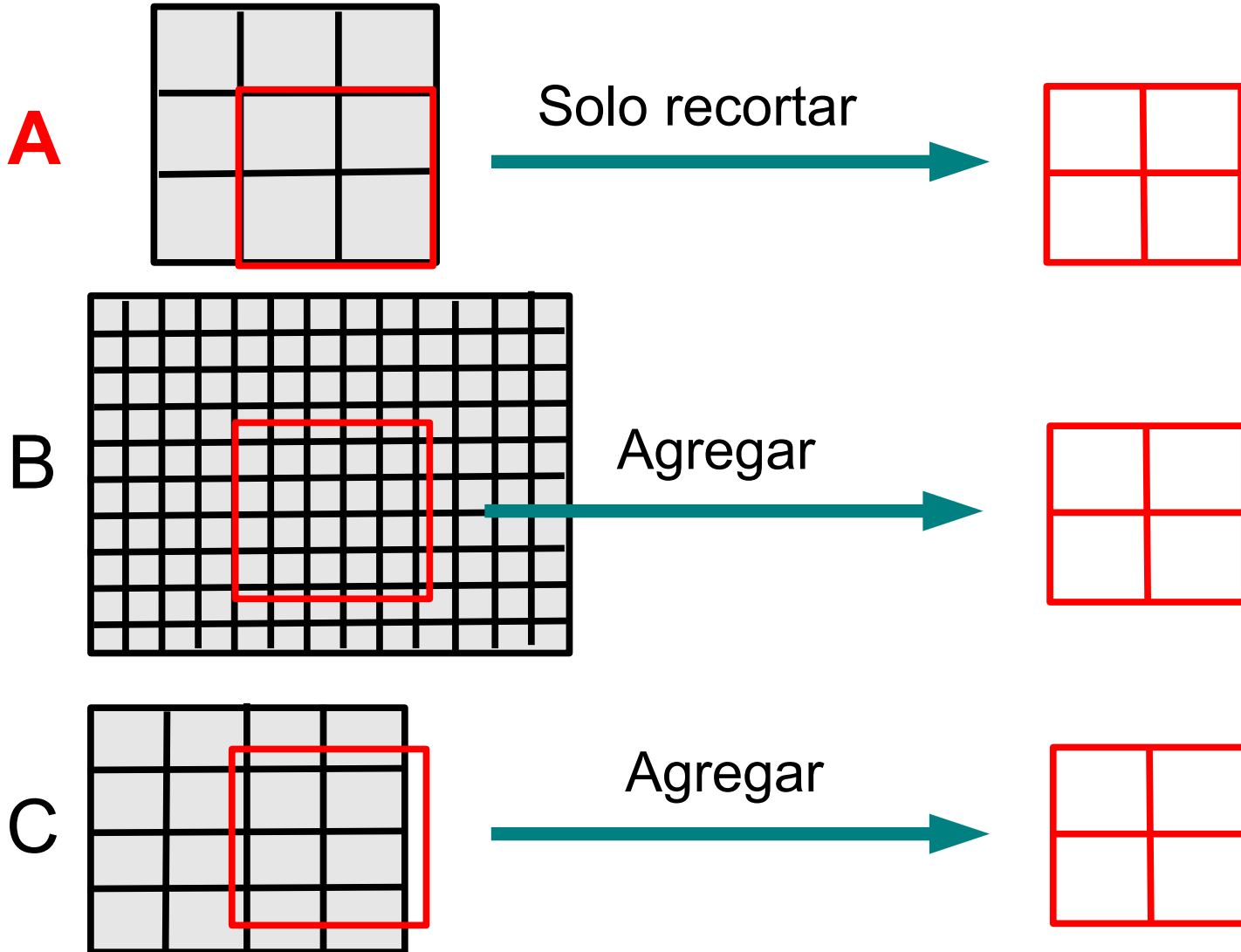


C

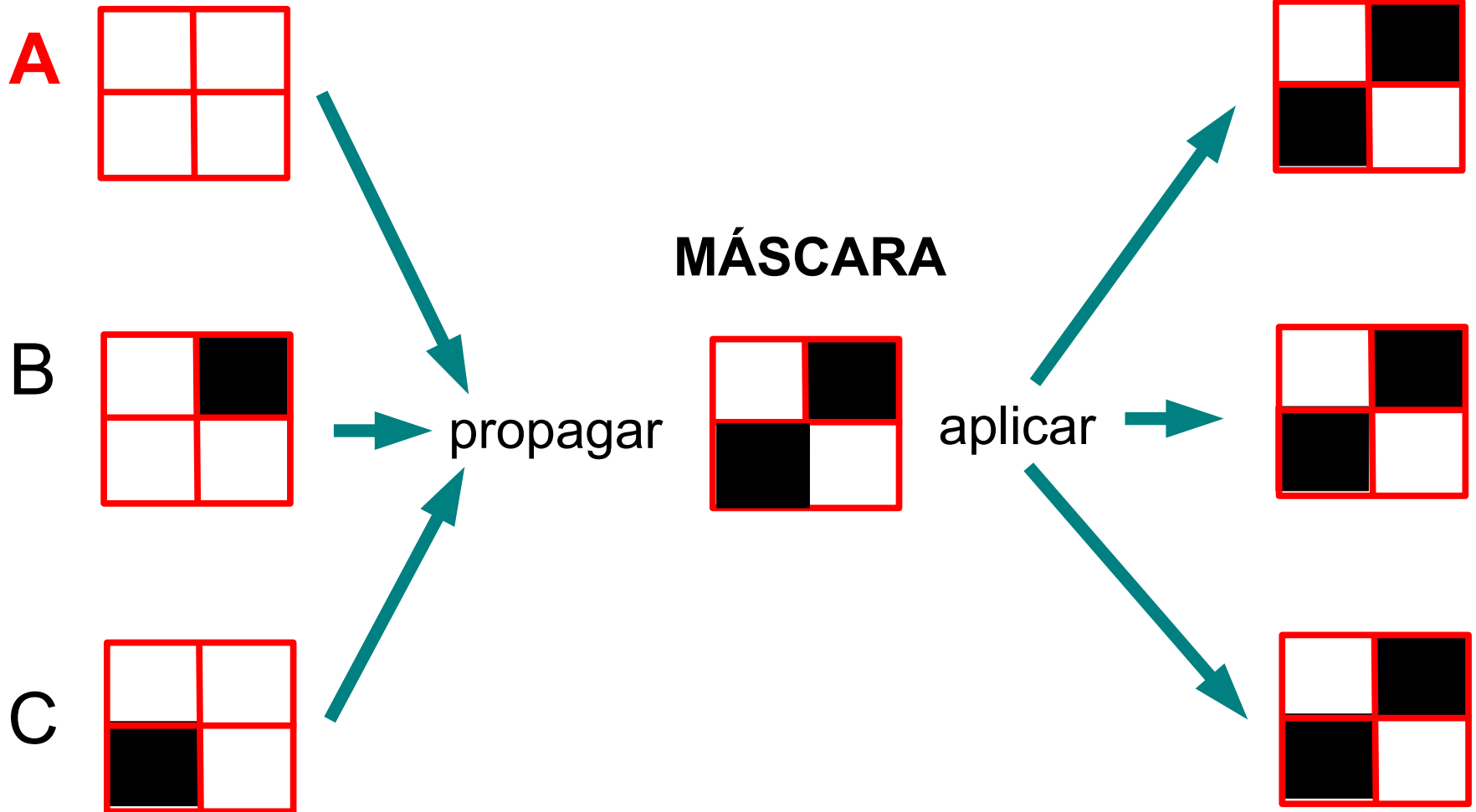


Normalmente se ajustará a la malla del mapa de referencia

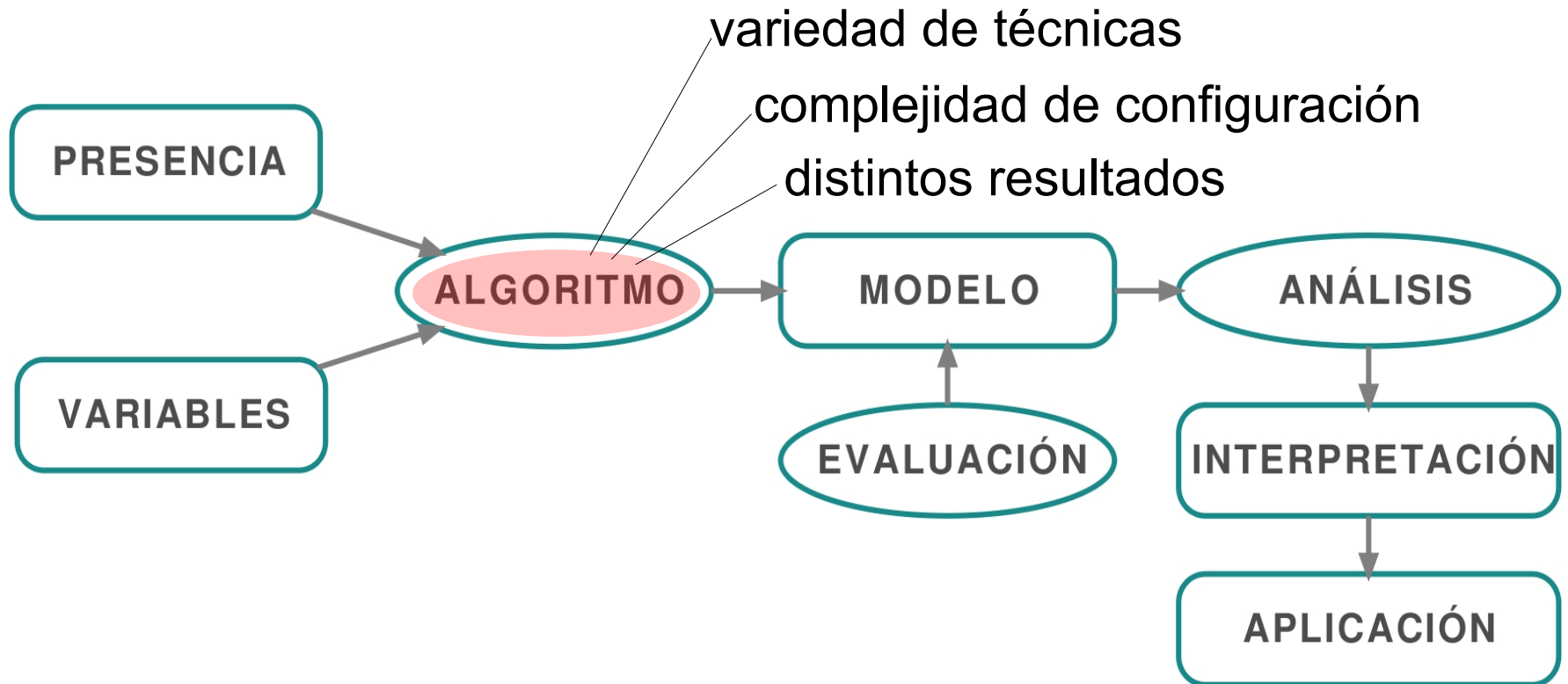
# 3 – CAMBIAR RESOLUCIÓN



# 4 – IGUALAR CELDAS NULAS



# ALGORITMO



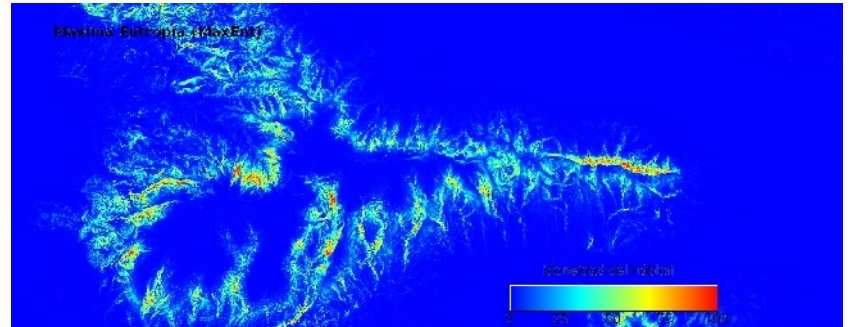
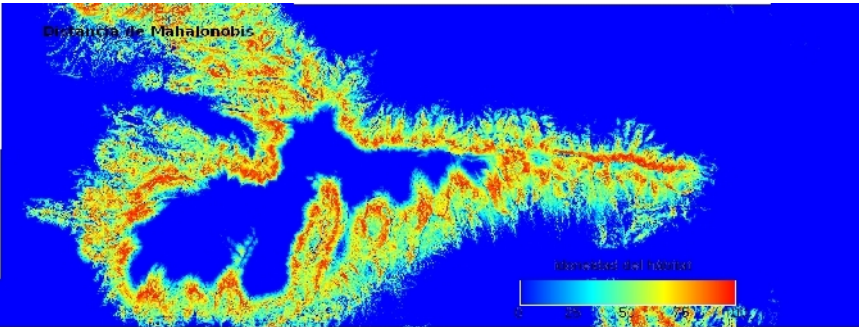
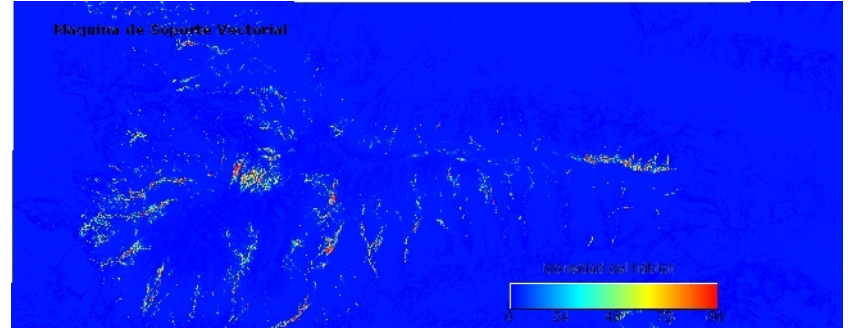
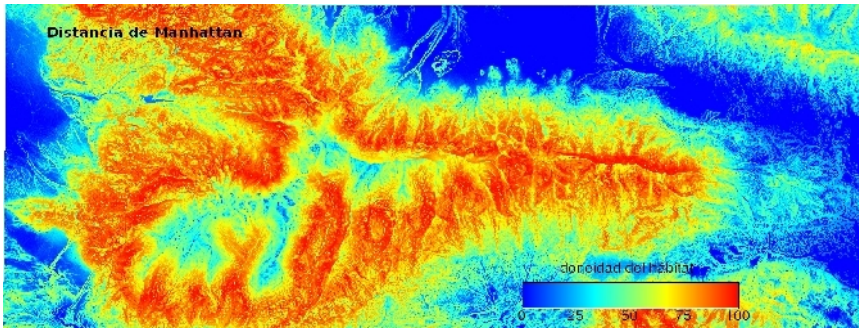
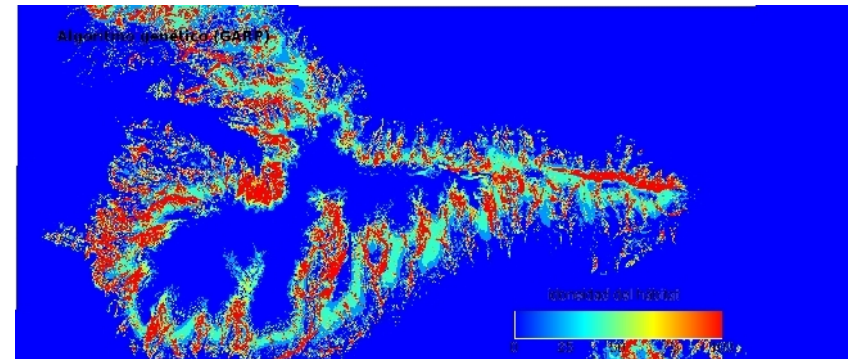
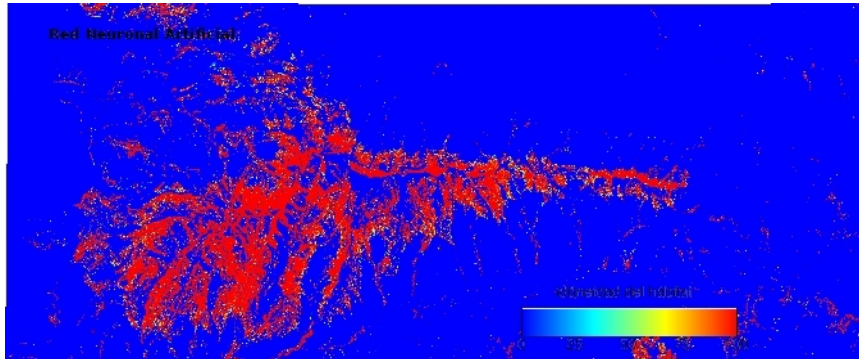


# ALGORITMO

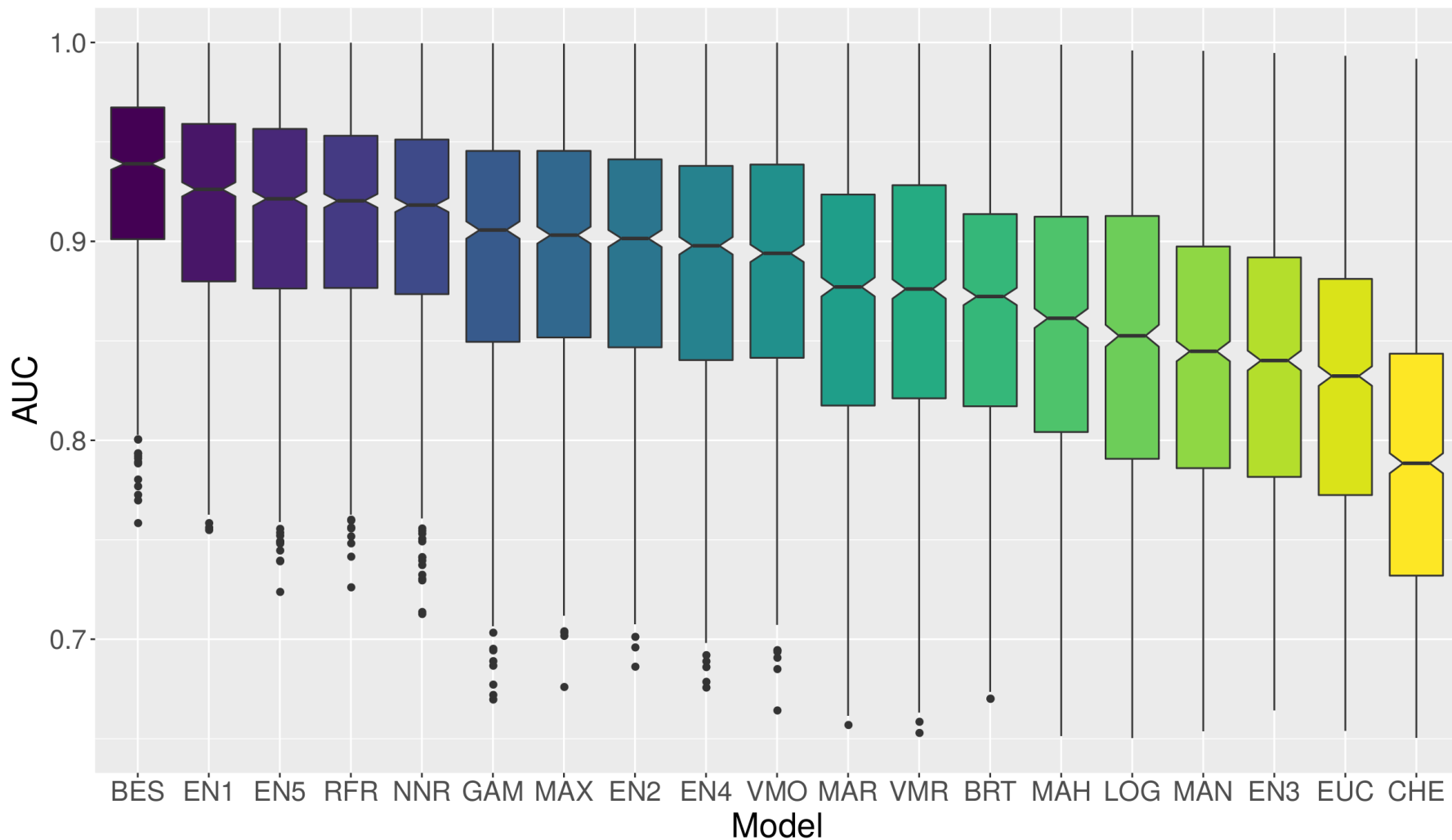
Familias de métodos (complejidad creciente):

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

# DISTINTOS RESULTADOS...



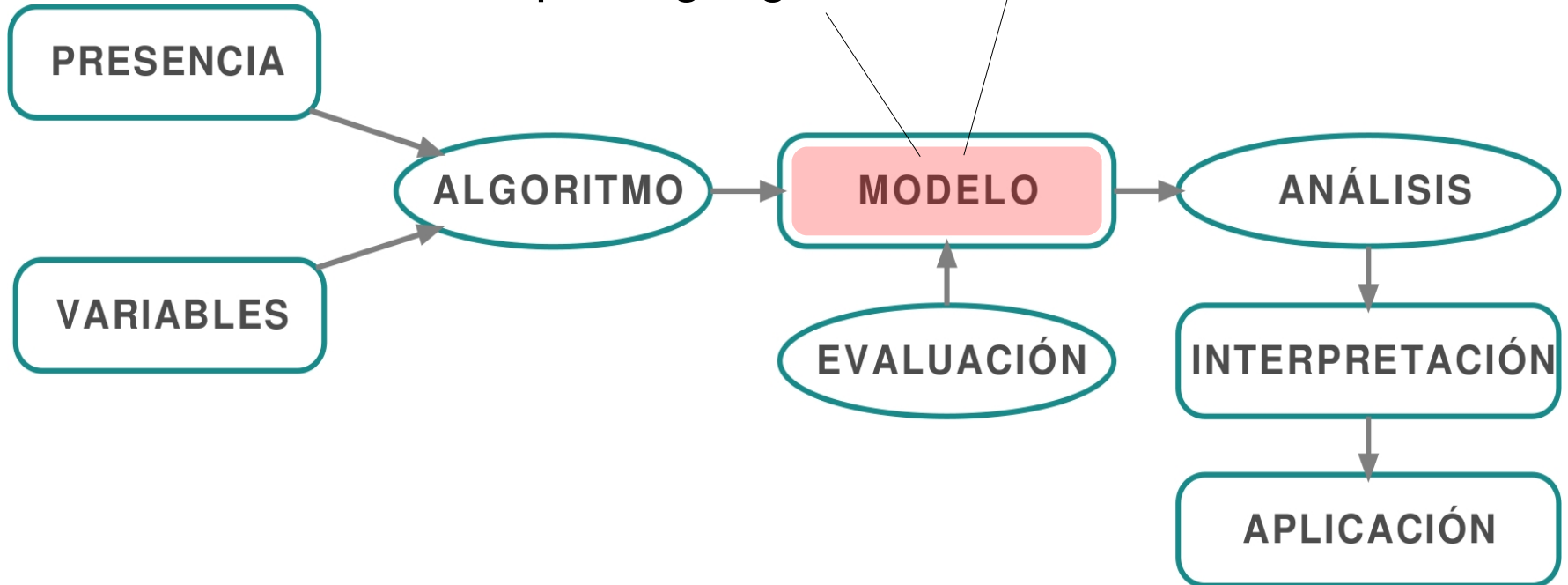
# ...Y DISTINTO PODER PREDICTIVO



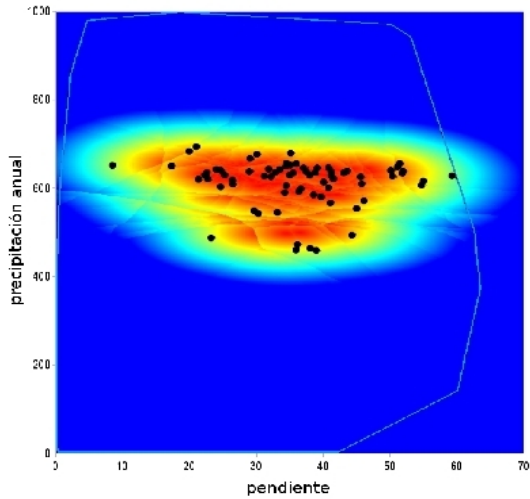
# MODELO

modelo en el espacio ecológico

modelo en el espacio geográfico

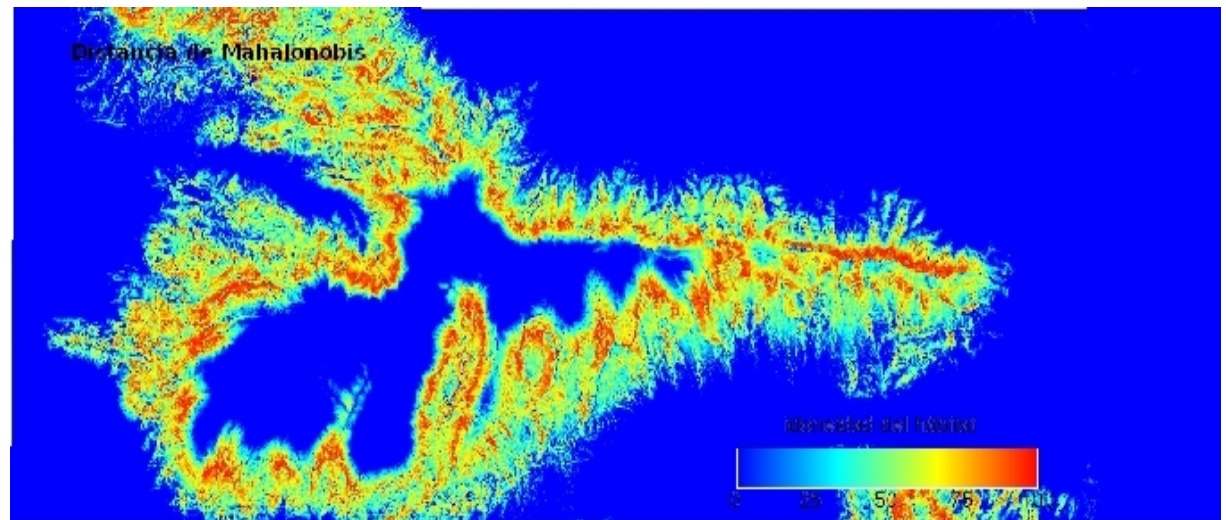


# MODELO

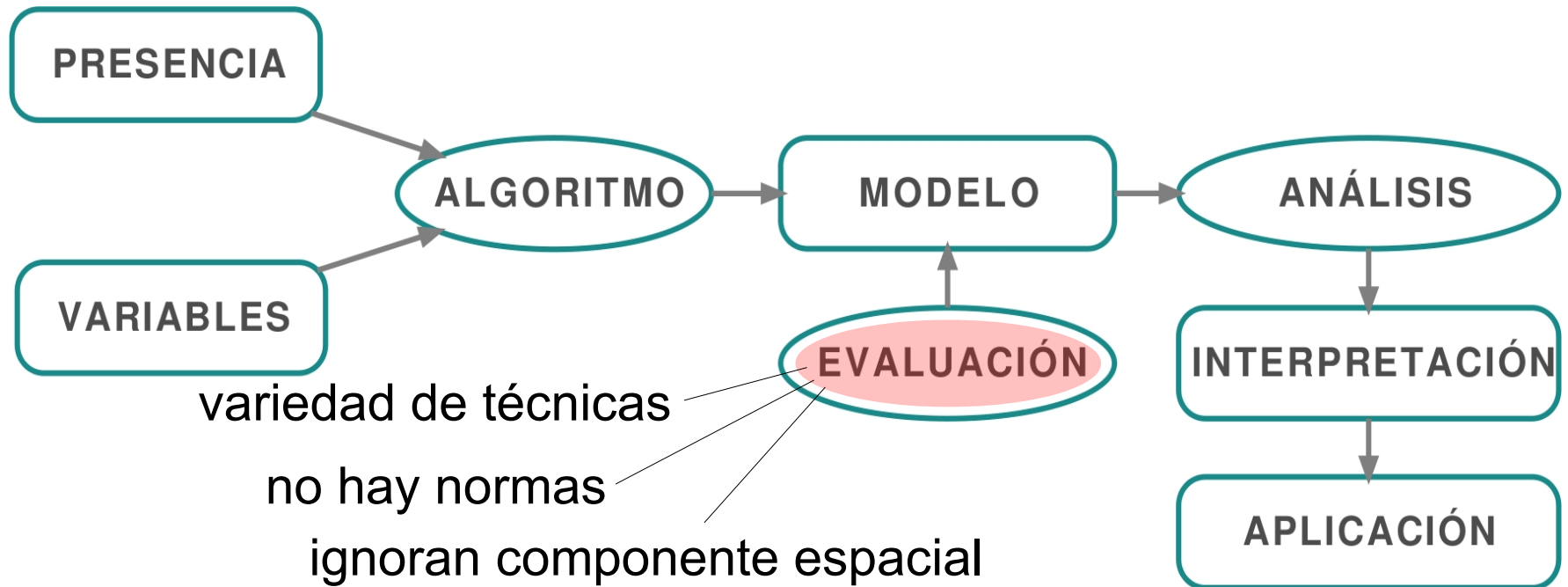


← espacio ecológico

espacio geográfico



# EVALUACIÓN



# EVALUACIÓN

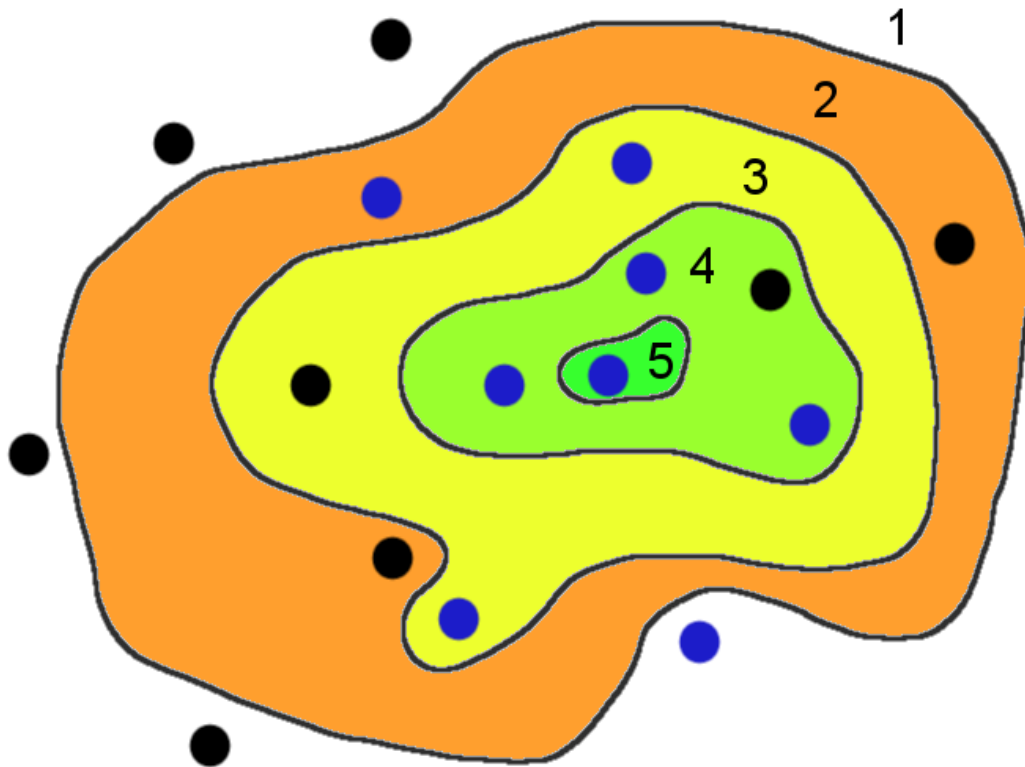
Artículo clave:

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	<b>A</b>	<b>B</b>
	ausencia	<b>C</b>	<b>D</b>

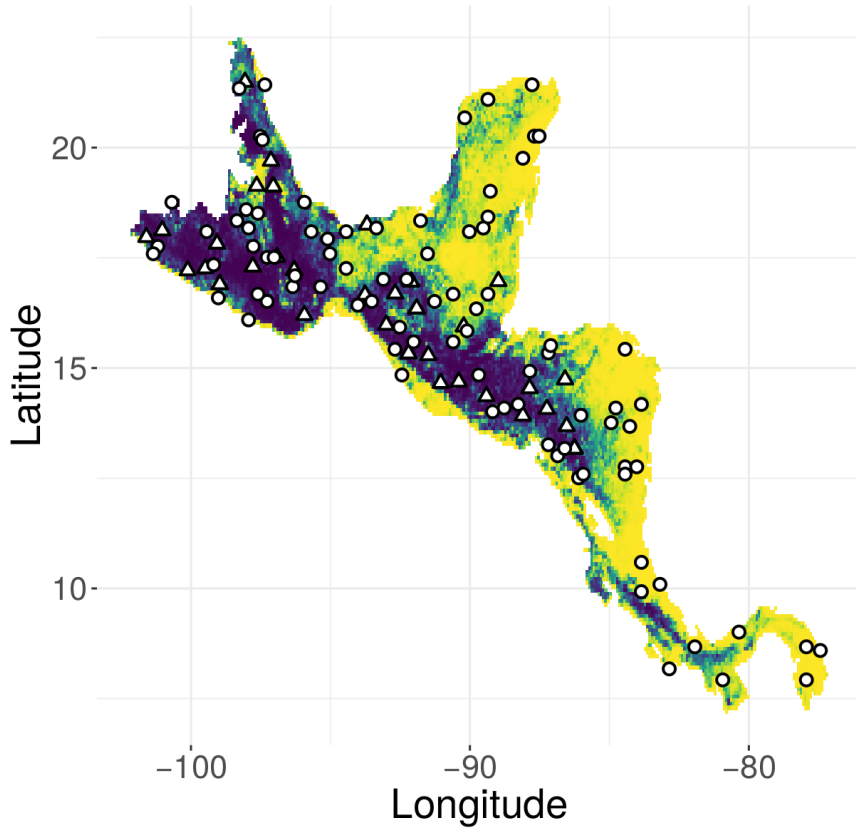
sensibilidad →  $S = A/(A+C)$

especificidad →  $E = D/(B+D)$

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

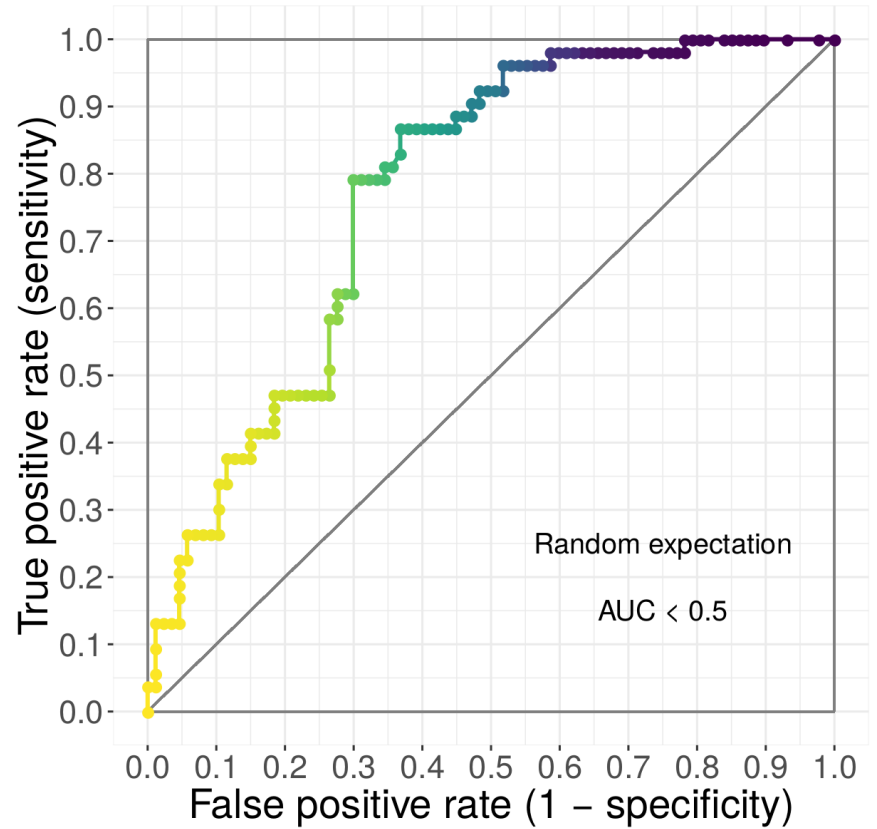
# AUC


Evaluation data



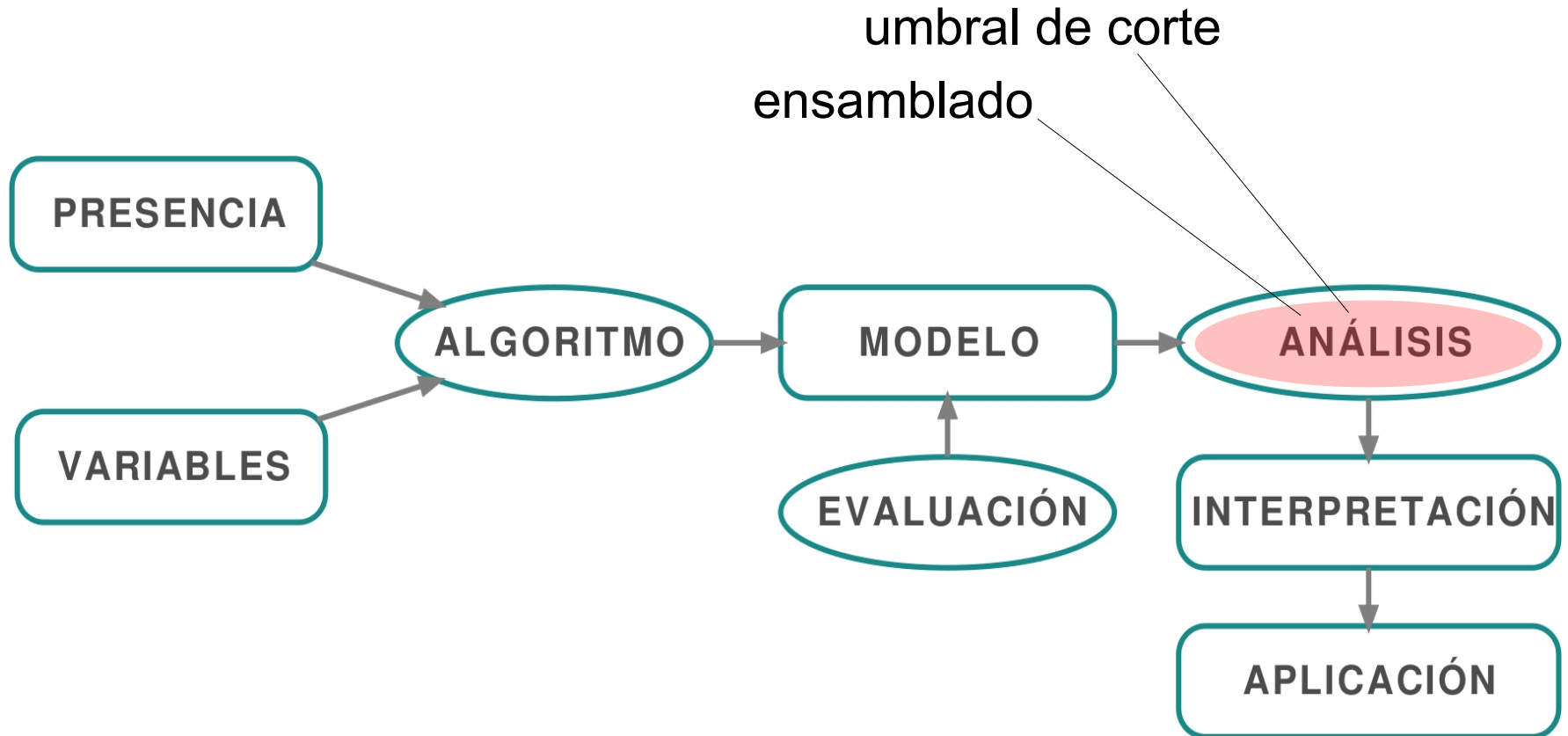
Points ○ absence ▲ presence

AUC = 0.77

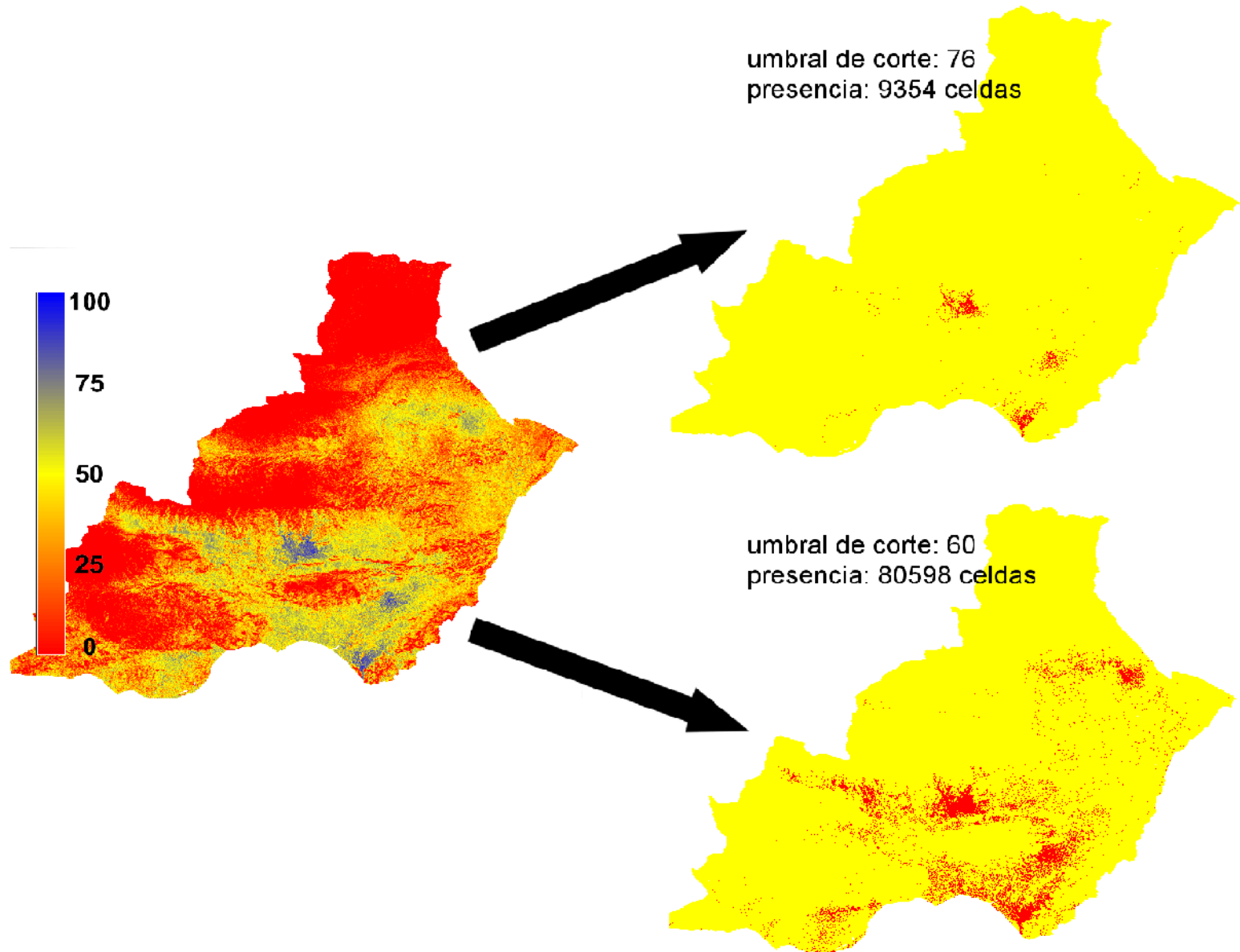


Suitability  0.25 0.50 0.75 1.00

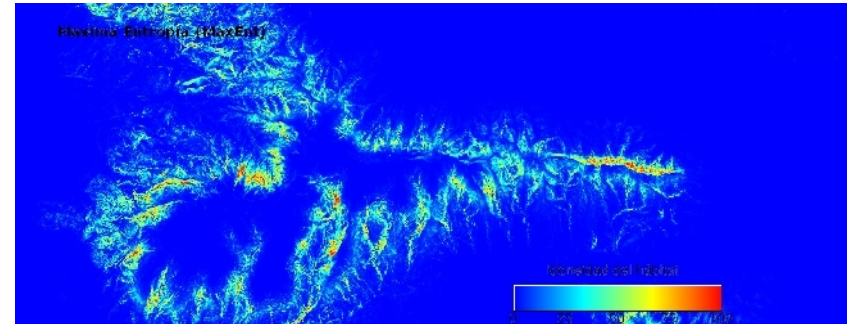
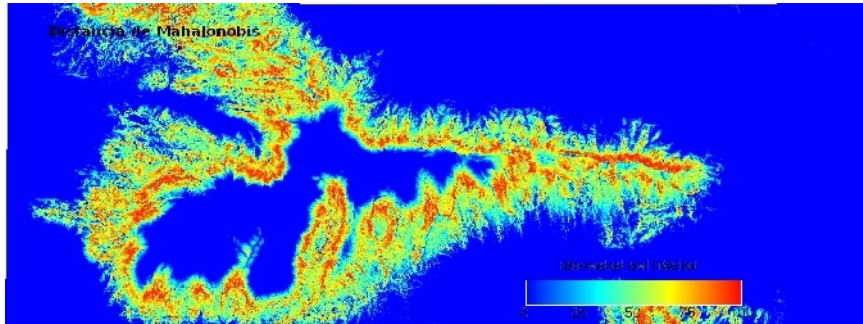
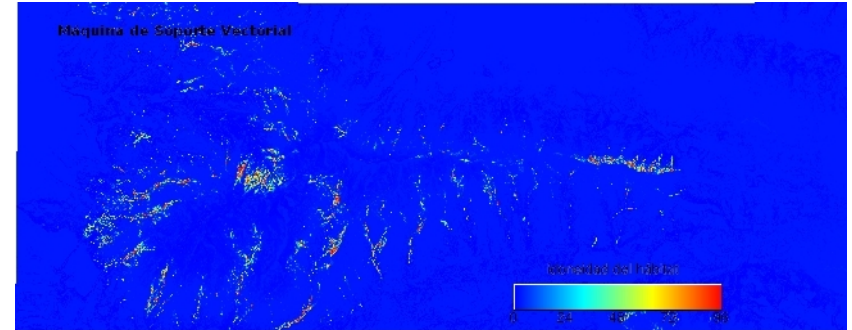
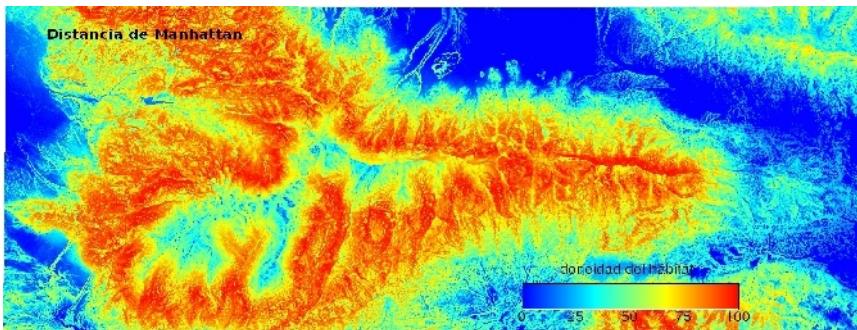
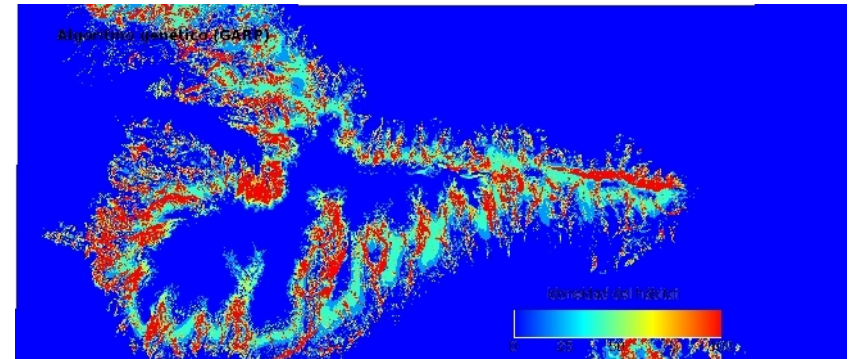
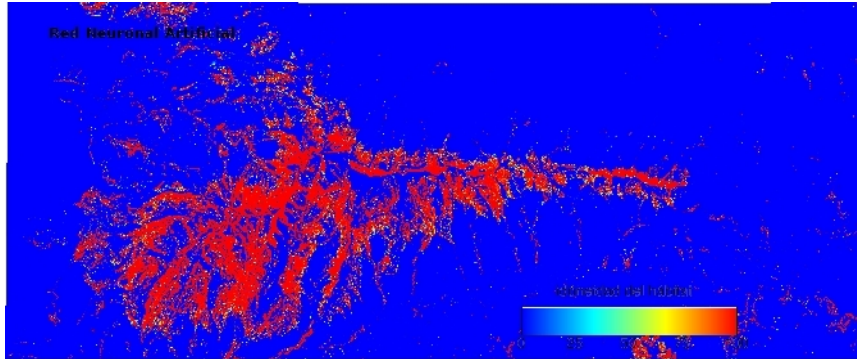
# ANÁLISIS



# UMBRAL DE CORTE



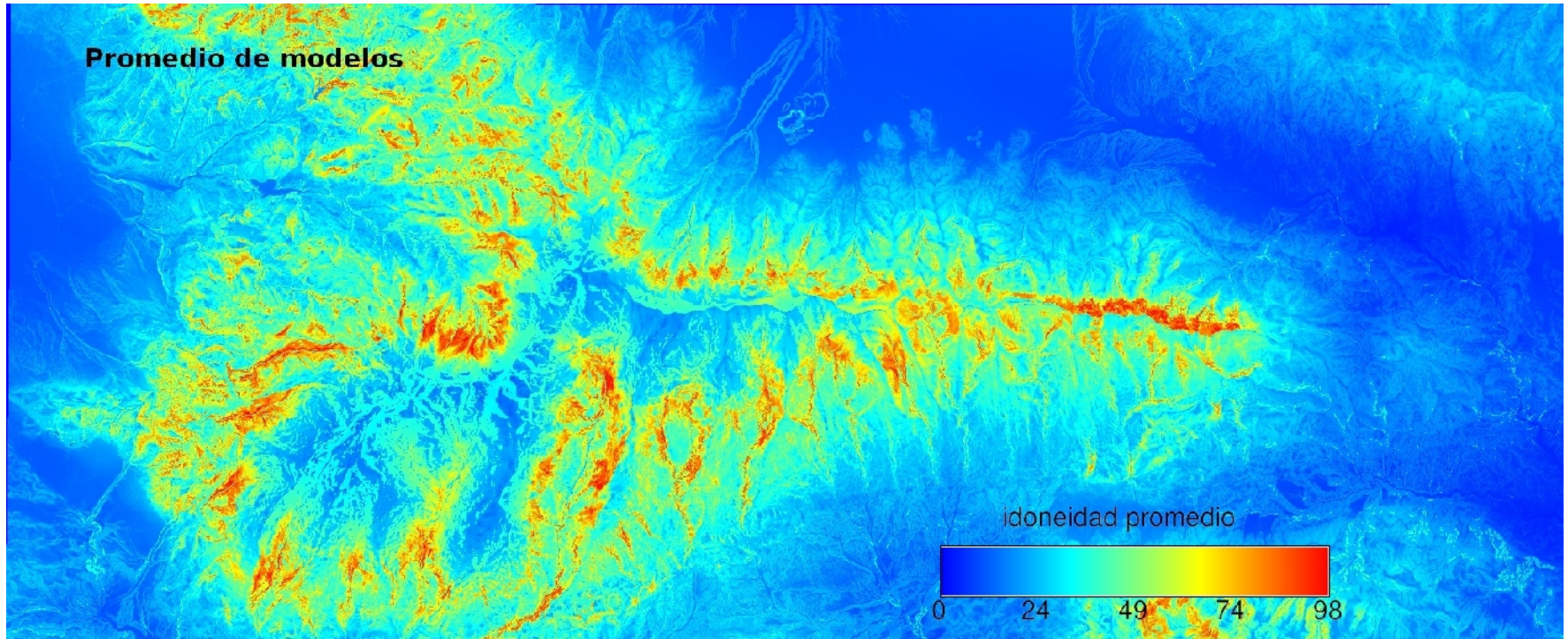
# ENSAMBLADO DE MODELOS



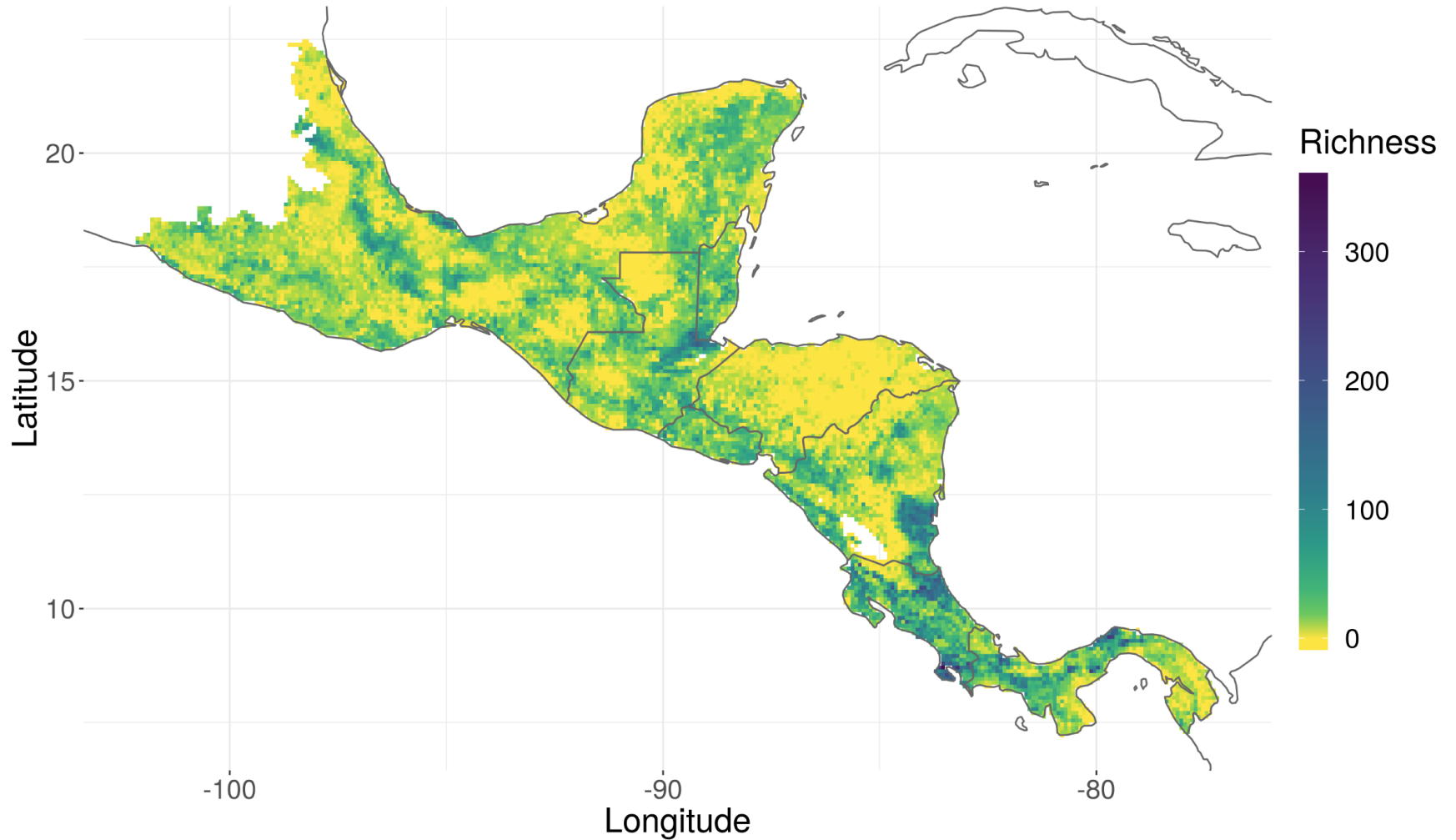


# ENSAMBLADO DE MODELOS

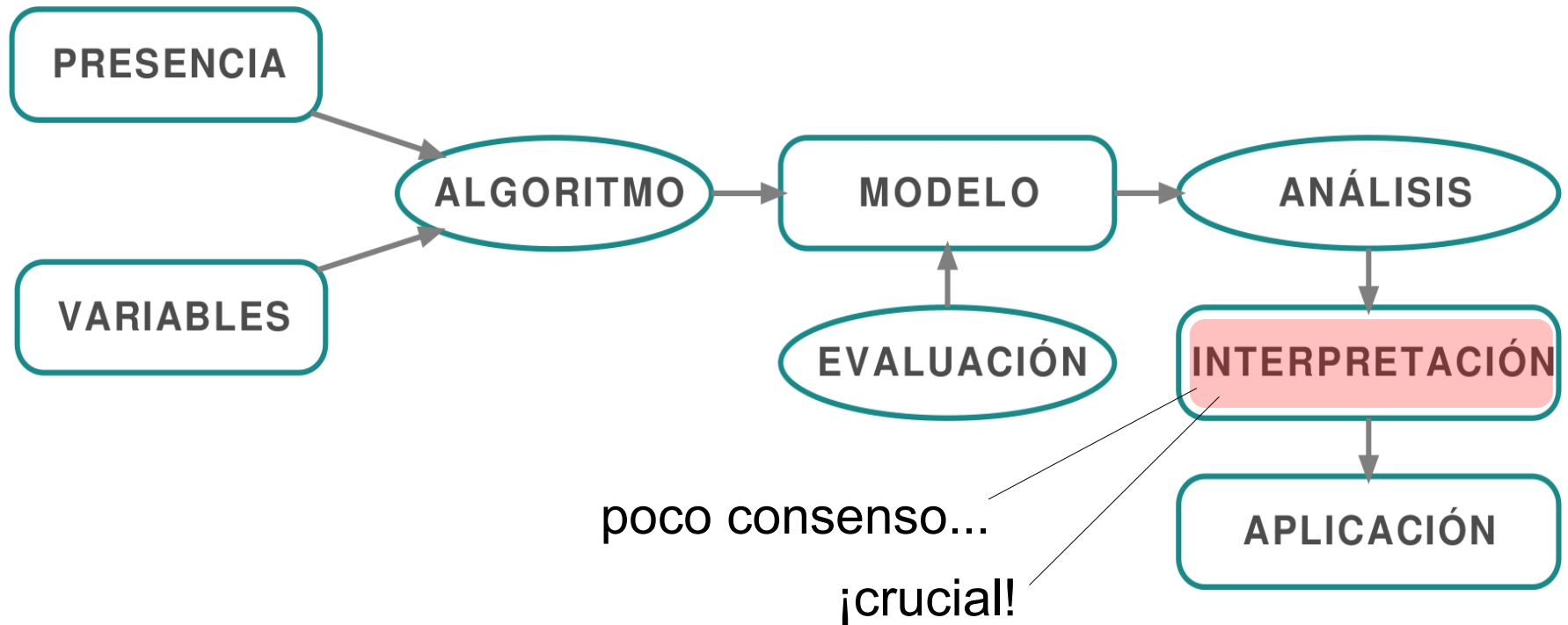
PROMEDIO DE LOS MODELOS ANTERIORES



# MODELOS DE RIQUEZA

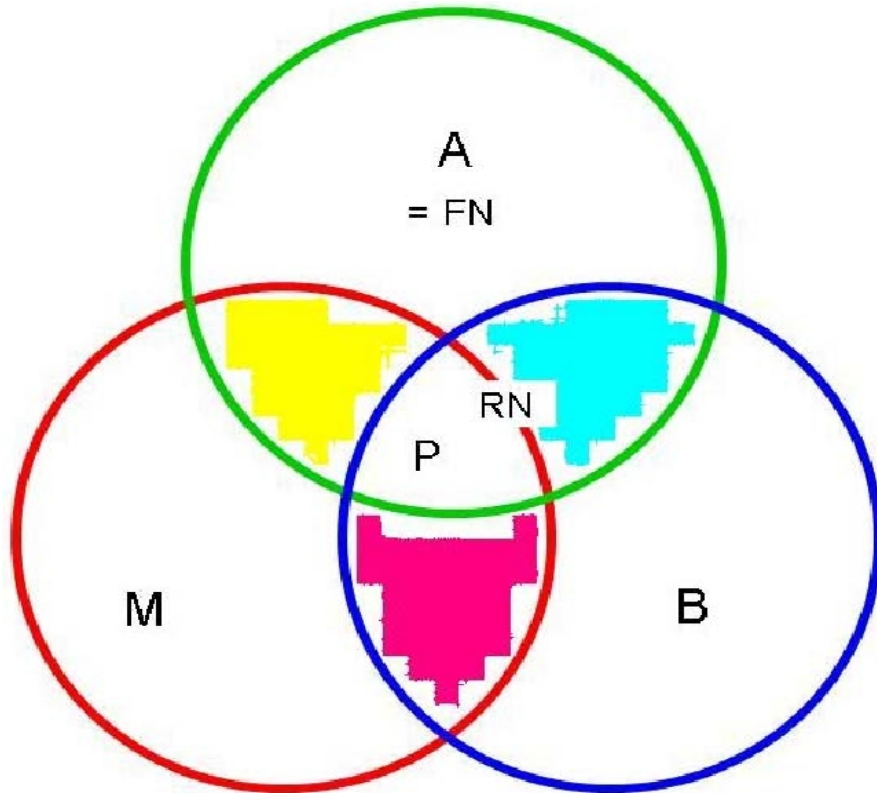


# INTERPRETACIÓN





# INTERPRETACIÓN



## Diagrama BAM

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

# 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.**”*

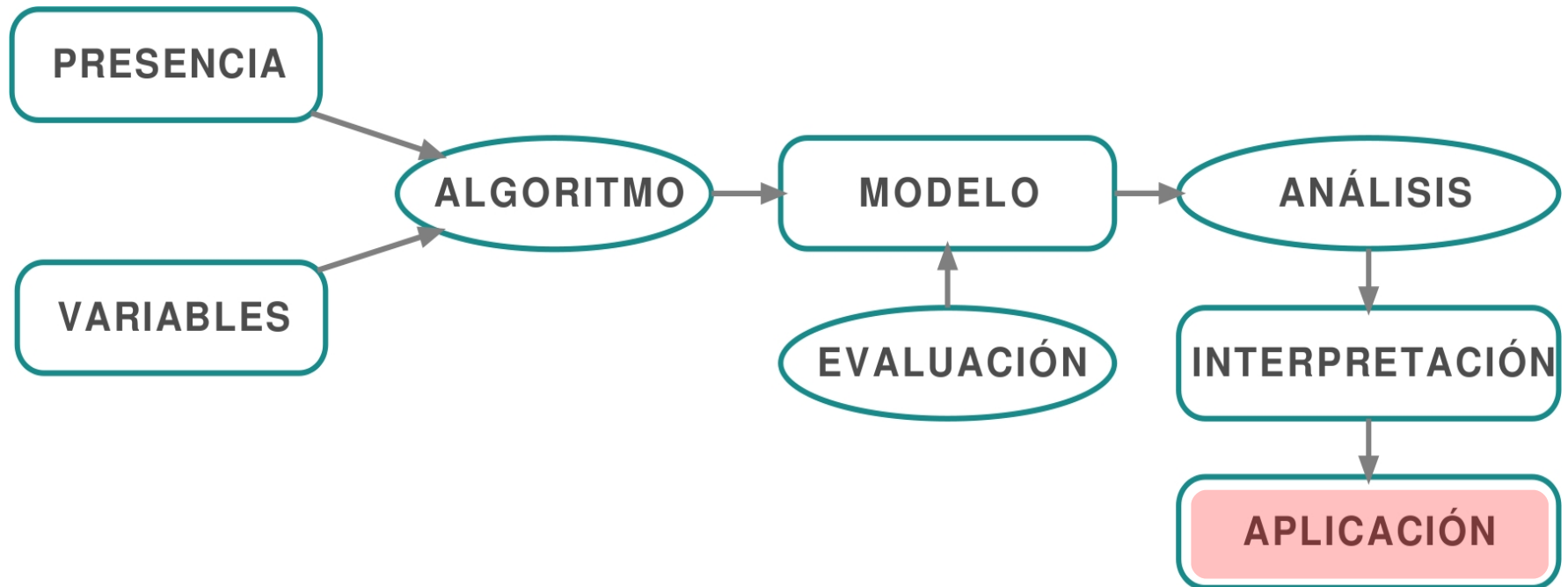
# INTERPRETACIÓN

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

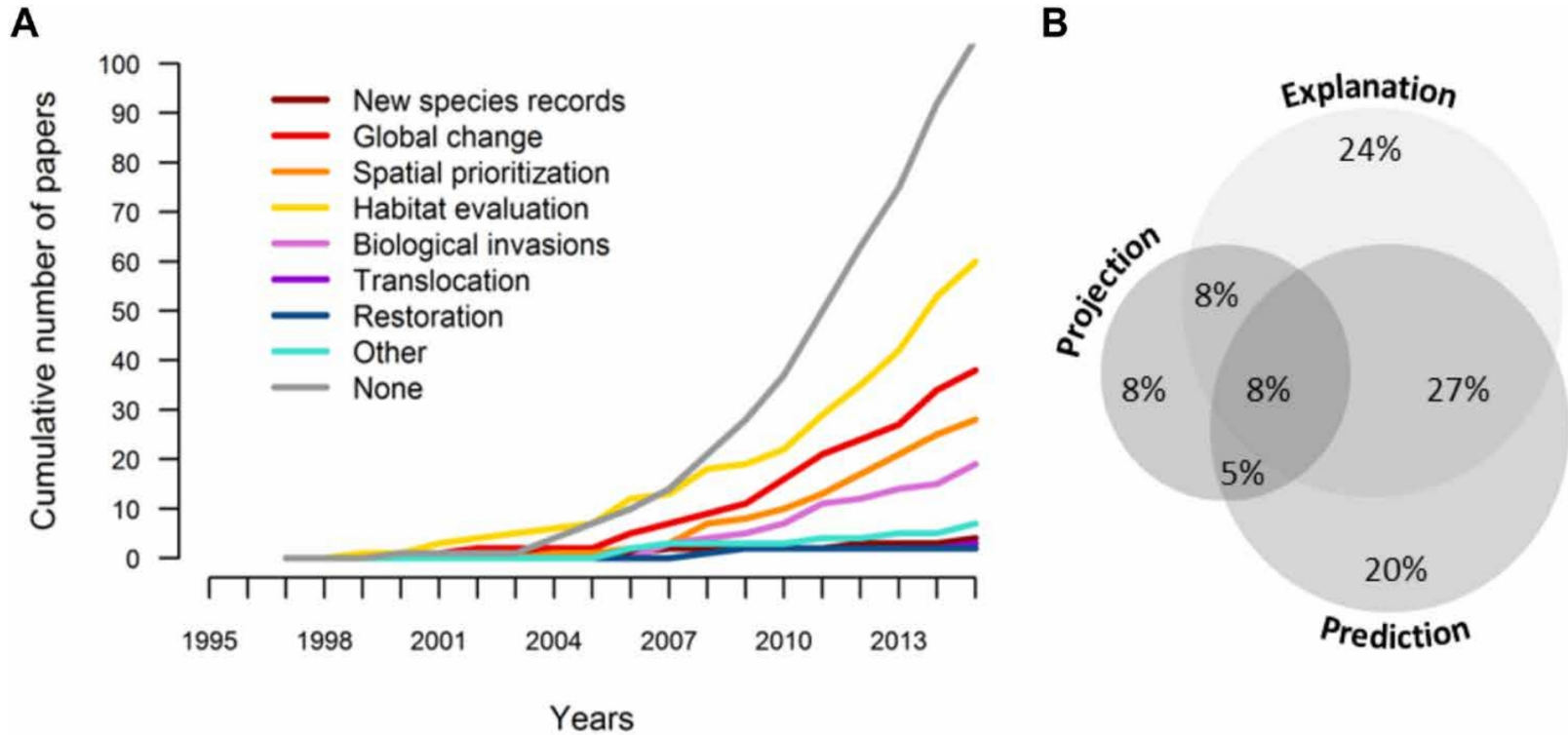
- ¿Se reproduce la especie en ese lugar?
- ¿Se alimenta la especie en ese lugar?
- ¿Busca la especie refugio u otros recursos en ese lugar?

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

# APLICACIÓN



# APLICACIONES



Araújo et al. 2019

# APLICACIONES

- **Explicación:** definir o testar hipótesis sobre los factores que influyen en la distribución de una especie.
- **Predicción:** obtener un mapa de distribución preciso de una especie.
- **Proyección:** obtener un mapa de distribución de una especie en otro tiempo o espacio.

# ESTÁNDARES DE CALIDAD DE LOS MODELOS DE DISTRIBUCIÓN

SCIENCE ADVANCES | REVIEW

---

## ECOLOGY

### Standards for distribution models in biodiversity assessments

Miguel B. Araújo<sup>1,2,3\*</sup>, Robert P. Anderson<sup>4,5,6</sup>, A. Márcia Barbosa<sup>3</sup>, Colin M. Beale<sup>7</sup>, Carsten F. Dormann<sup>8</sup>, Regan Early<sup>9</sup>, Raquel A. Garcia<sup>2,3,10,11</sup>, Antoine Guisan<sup>12,13</sup>, Luigi Maiorano<sup>14,15</sup>, Babak Naimi<sup>2</sup>, Robert B. O'Hara<sup>16,17</sup>, Niklaus E. Zimmermann<sup>18,19</sup>, Carsten Rahbek<sup>2,20</sup>

Demand for models in biodiversity assessments is rising, but which models are adequate for the task? We propose a set of best-practice standards and detailed guidelines enabling scoring of studies based on species distribution models for use in biodiversity assessments. We reviewed and scored 400 modeling studies over the past 20 years using the proposed standards and guidelines. We detected low model adequacy overall, but with a marked tendency of improvement over time in model building and, to a lesser degree, in biological data and model evaluation. We argue that implementation of agreed-upon standards for models in biodiversity assessments would promote transparency and repeatability, eventually leading to higher quality of the models and the inferences used in assessments. We encourage broad community participation toward the expansion and ongoing development of the proposed standards and guidelines.

Disponible en la carpeta de bibliografía, bajo los nombres 2019\_Araujo.pdf, y 2019\_Araujo\_ESM.pdf

# GOLD STANDARD: PRESENCIAS

- Muestreo sistemático, repetido, homogéneamente distribuido abarcando área de presencia actual e histórica.
- Identificación taxonómica mediante registros re-examinables (fotos, pliegos).
- Precisión de coordenadas más fina que resolución de las variables.



# GOLD STANDARD: VARIABLES

- Todas las variables son proximales.
- Variables distales no deben usarse para modelos explicativos o proyectados.
- Selección de variables puede estar justificada estadísticamente.
- Variables bióticas (distribución de competidores o presas) deben incluirse cuando sea posible.
- Resolución espacial ajustada a la escala de la relación entre la variable y la presencia.
- Resolución temporal ajustada a la biología de la especie (i.e. especies migratorias y medias mensuales).
- Interpolación solo aceptable cuando no hay medidas directas.
- Considera incertidumbre de las variables (i.e. escenarios de cambio climático)

# GOLD STANDARD: AJUSTE DEL MODELO

- Complejidad y sobreajuste se evalúan de distintos modos sobre datos independientes (ideal en teoría, pero imposible en la práctica).
- Corrección de sesgo en la variable respuesta, y evaluación de la corrección sobre datos independientes.
- Silver: presencias se ponderan según fiabilidad, y se comprueba resultado mediante validación cruzada.
- No hay colinealidad entre las variables.
- Evaluación de la sensibilidad a la incertidumbre en modelos y parámetros (grados de libertad del análisis).

# GOLD STANDARD: EVALUACIÓN

- Asunciones del modelo: demostrar que no se violan asunciones (o usar modelos que no requieren asunciones!)
- Evaluación del resultado:
  - Evaluación con múltiples líneas de evidencia independientes.
  - Evaluación con datos espacial o temporalmente independientes de los datos de entrenamiento (otra vez, muy difícil en la práctica).

# **PREPARACIÓN DE LAS VARIABLES PREDICTIVAS Y LOS REGISTROS DE PRESENCIA**

# Algunas cosas a aprender con R

Abre `taller1/intro_R/introduccion.R`

# PREPARACIÓN DE VARIABLES Y PRESENCIAS

Abre

**taller1/prepara\_presencias\_y\_variables/  
1\_prepara\_variables\_y\_presencia.R**

The image features a classic Looney Tunes ending screen. It consists of a series of concentric circles in shades of red and black, creating a tunnel-like effect. In the center, the text "That's all Folks!" is written in a white, elegant cursive font. The text is positioned diagonally across the center of the circles.

*That's all Folks!*