

SEMINARIO MODELIZACIÓN DE NICHOS ECOLÓGICOS

María Ángeles Pérez
m.angeles582@gmail.com



Modelos y medidas de evaluación

Principales modelos

Modelos simples vs complejos

Medidas de adecuación dependientes de umbral

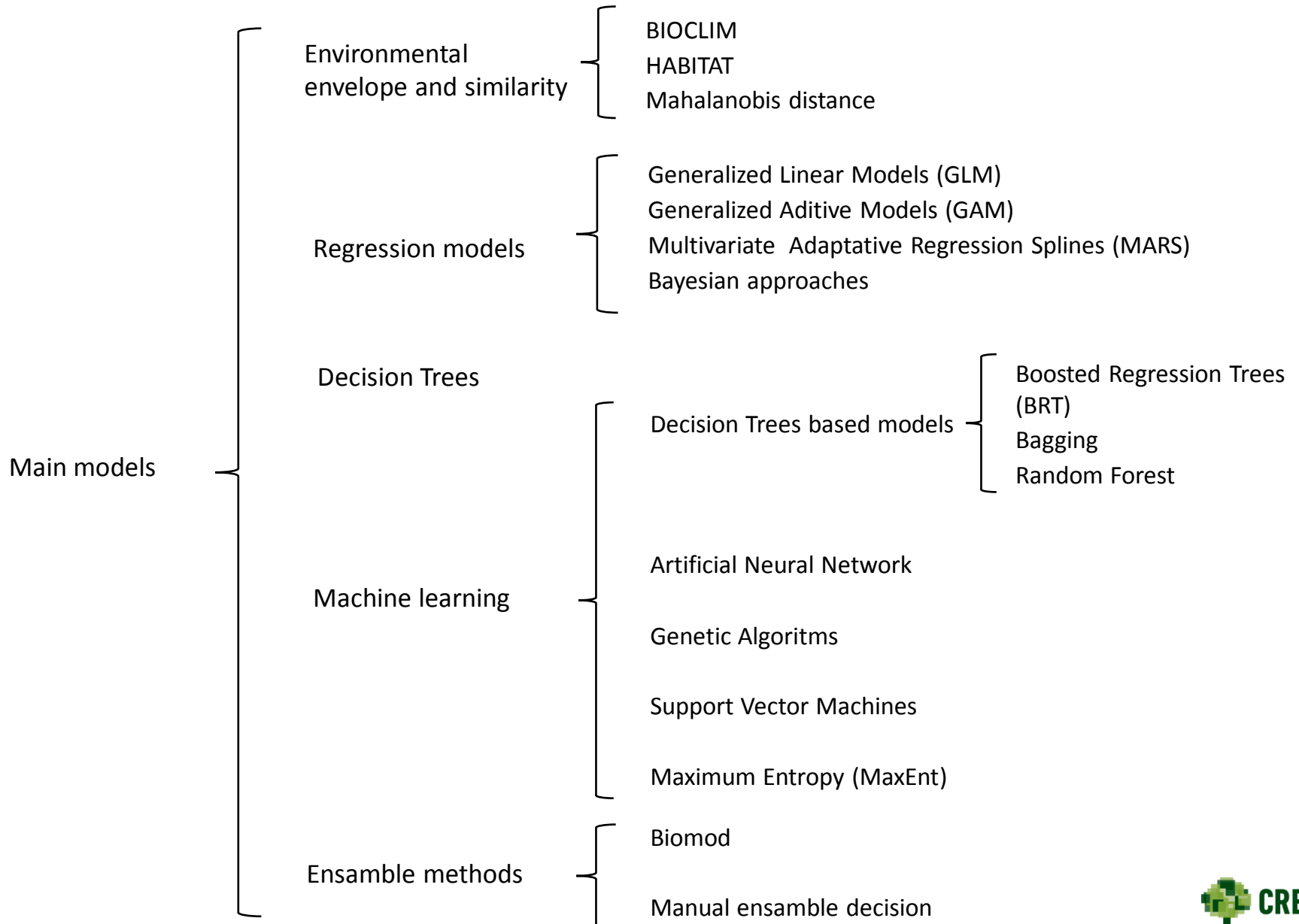
Medidas de adecuación independientes de umbral

Otros criterios de evaluación

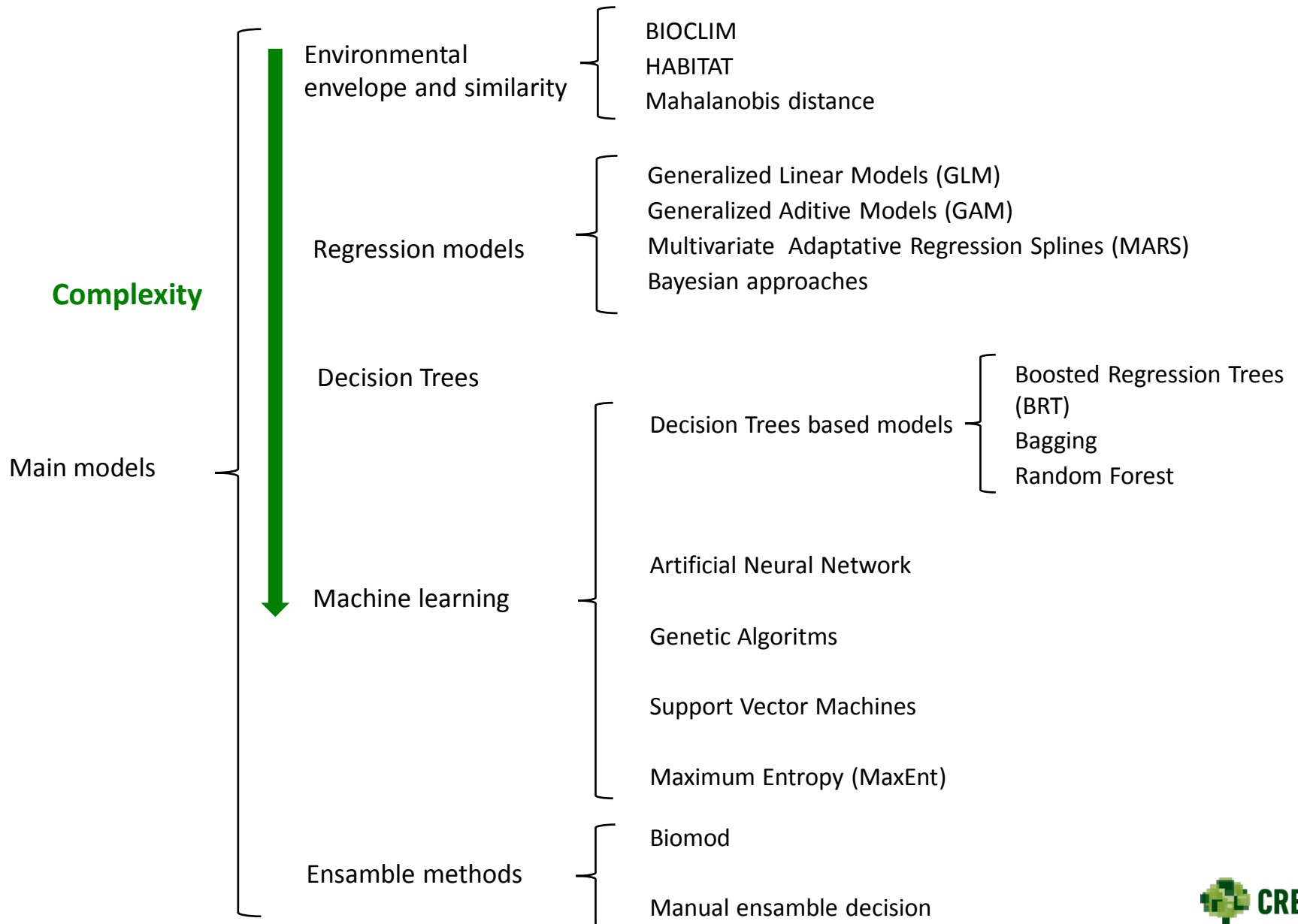
Thresholds. Selección de umbrales de probabilidad

Proyecciones en el espacio y en el tiempo. Extrapolación

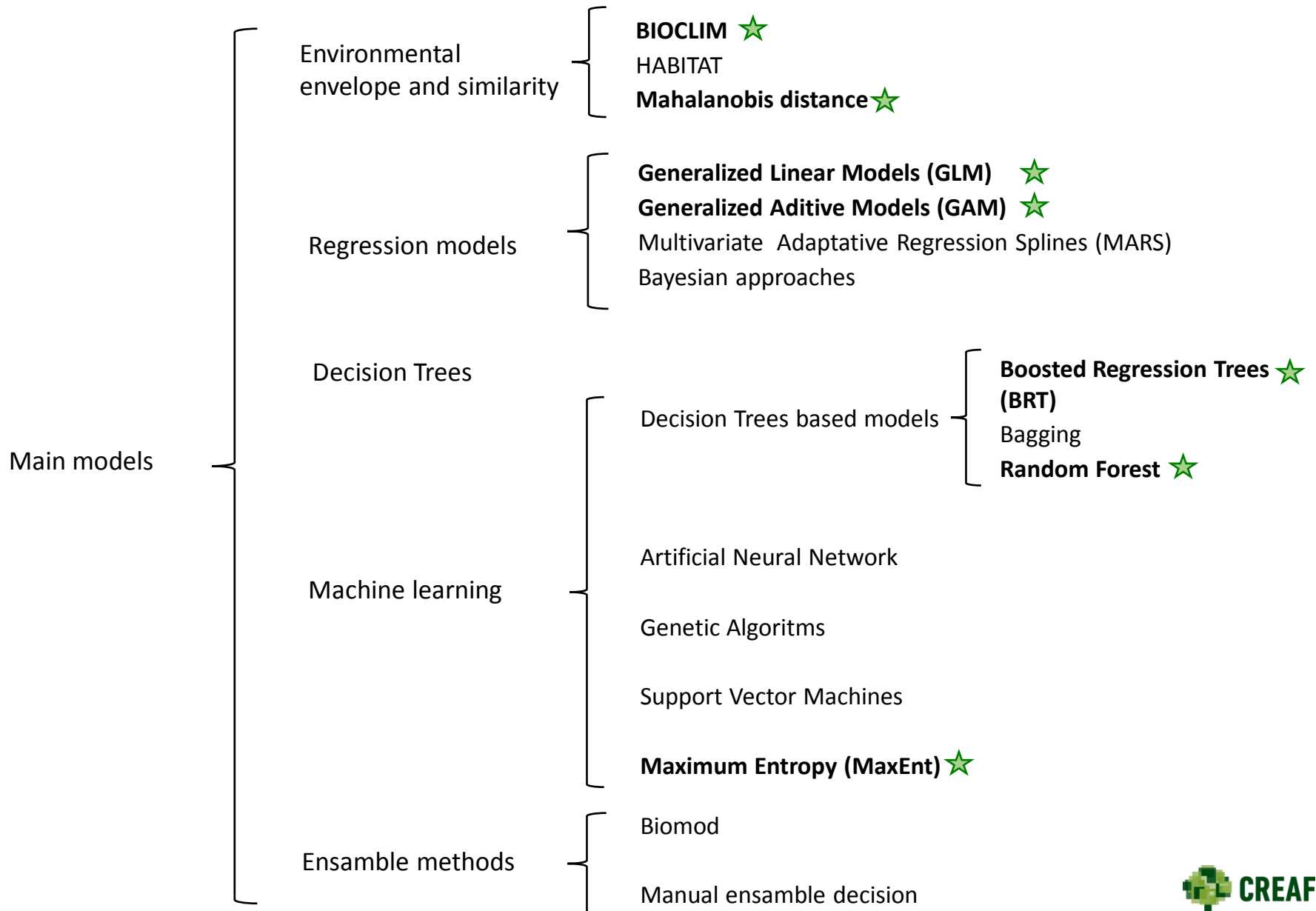
Species Distribution Models



Species Distribution Models

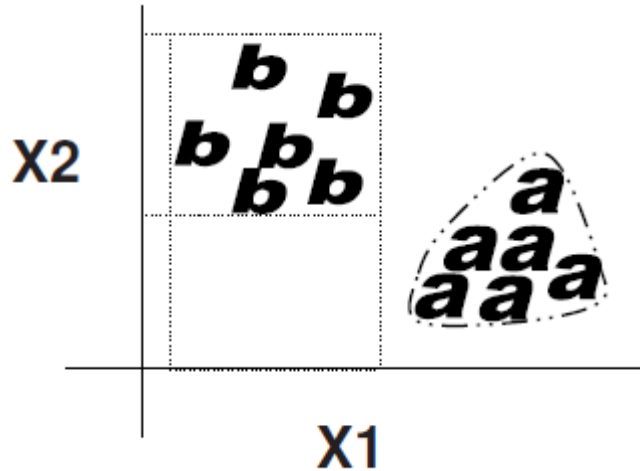


Species Distribution Models

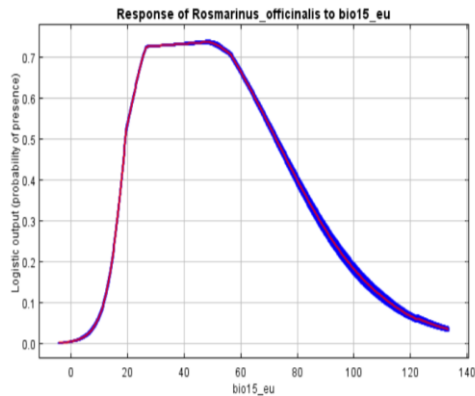
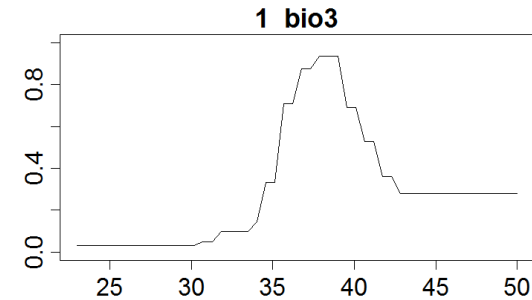


Species Distribution Models

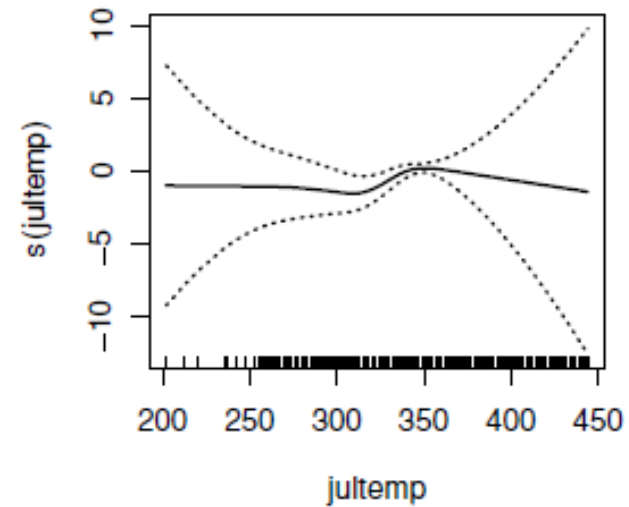
BIOCLIM



BRT



MaxEnt

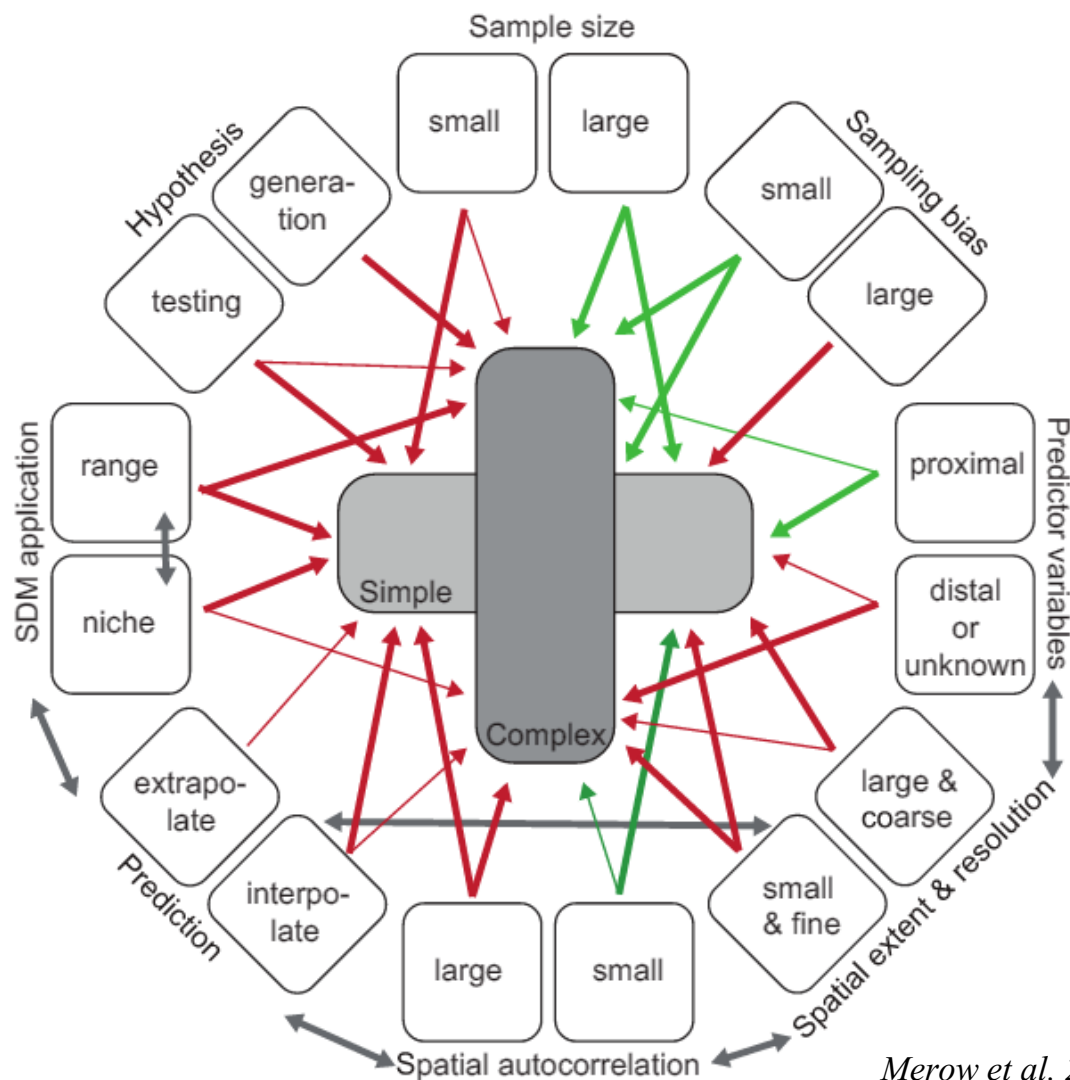


GAM

Species Distribution Models

¿Qué hace a un modelo más complejo?

- Número de variables
- Interacción entre variables
- Funciones de ajuste a los datos de ocurrencia



Merow et al. 2014 *Ecography*, 37: 1267–1281

¿Cómo medimos la adecuación de un modelo?

- **Medidas de adecuación dependientes de umbral**

Datos de evaluación categóricos + Predicciones categóricas



Matrices de confusión

		Observación	
		Presencia	Ausencia
Predicción	Presencia	TP	FP
	Ausencia	FN	TN

Sensitividad: $TP/(TP+FN)$

Especificidad: $TN/(TN+FP)$

Ratio falsos positivos: $1 - \text{Especificidad}$

Ratio falsos negativos: $1 - \text{Sensitividad}$

TP: True Positive

FP: False Positive. Error de comisión

TN: True Negative

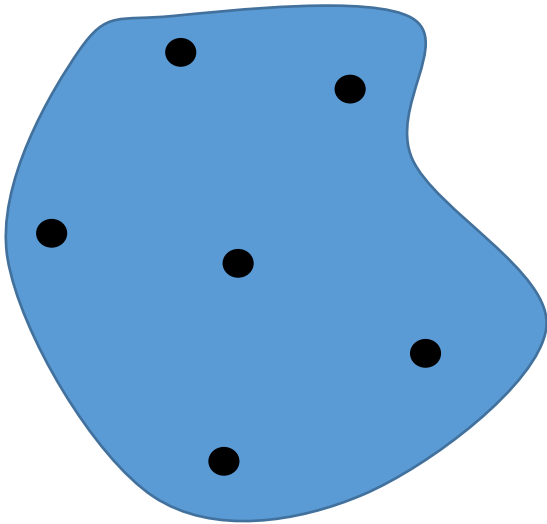
FN: False Negative. Error de omisión

Medidas dependientes de prevalencia

- Baja prevalencia → predominio errores de omisión
- Alta prevalencia → predominio errores de comisión

¿Cómo medimos la adecuación de un modelo?

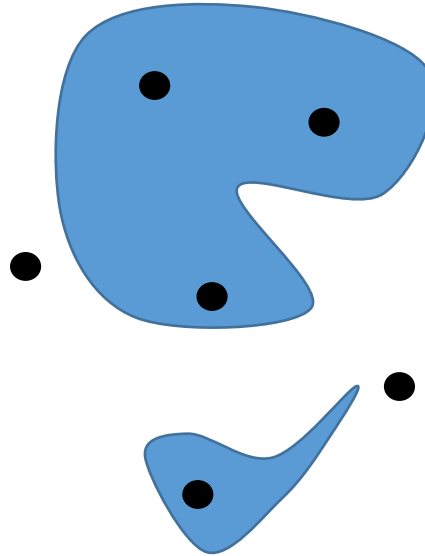
- Medidas de adecuación dependientes de umbral



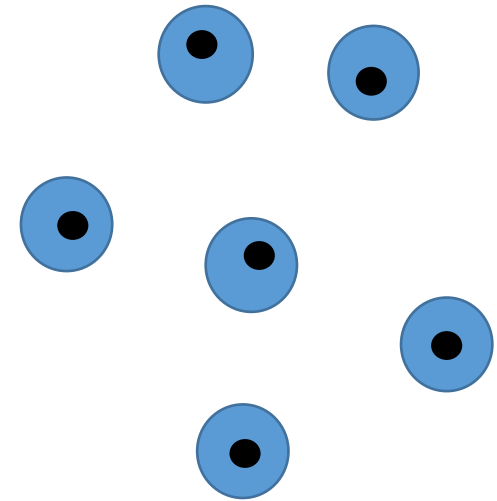
Sensibilidad= 1

Errores de comisión?

Sobreajuste?



Sensibilidad= 0.66



Sensibilidad= 1



Presencia de calibrado

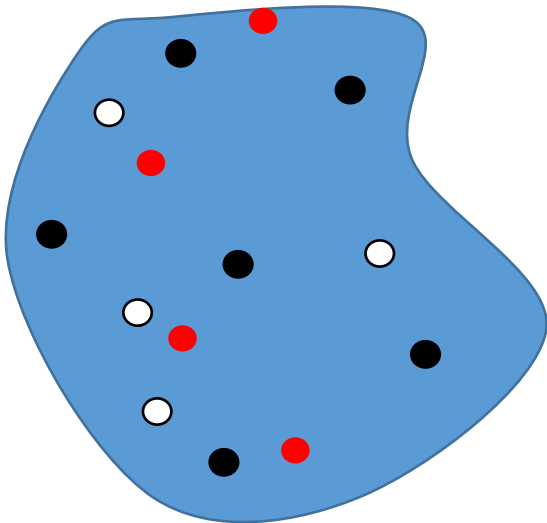


Habitat predicho como adecuado

¿Cómo medimos la adecuación de un modelo?

- **Medidas de adecuación dependientes de umbral**

Test data

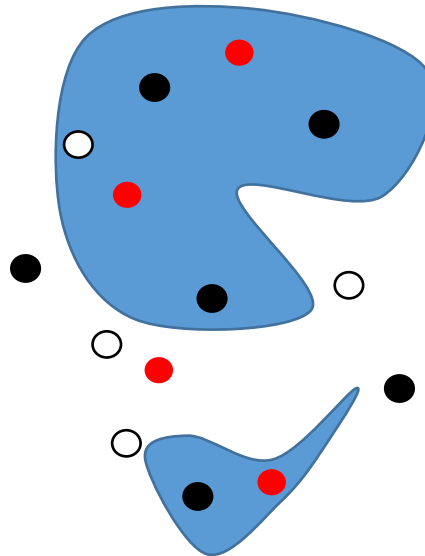


Sensibilidad= 1

Especificidad= 0

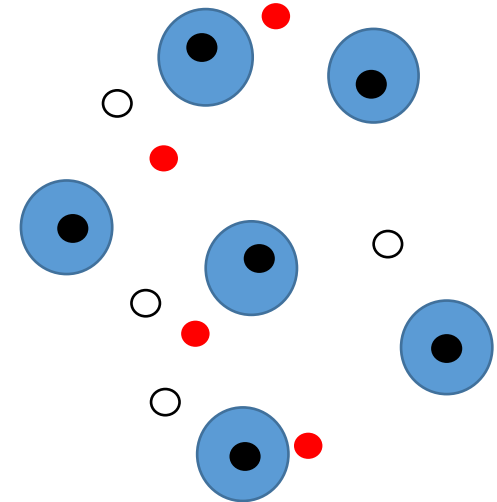
● Presencia validación (test set)

○ Ausencia validación (test set)



Sensibilidad= 0.75

Especificidad= 0.75



Sensibilidad= 0

Especificidad= 1

● Presencia de calibrado


● Habitat predicho como adecuado

¿Cómo medimos la adecuación de un modelo?

- **Medidas de adecuación dependientes de umbral**

Algunas consideraciones

Measure	Calculation
Sensitivity	$TP / (TP + FN)$
False negative rate	$1 - \text{Sensitivity}$
Specificity	$TN / (TN + FP)$
False positive rate	$1 - \text{Specificity}$
Percent correct classification	$(TP + TN) / n$
Positive predictive power	$TP / (TP + FP)$
Odds ratio	$(TP \times TN) / (FP \times FN)$
Kappa	$\frac{[(TP+TN) - (((TP+FN)(TP+FP) + (FP+TN)(FN+TN)) / n)]}{[n - (((TP+FN)(TP+FP) + (FP+TN)(FN+TN)) / n)]}$
True skill statistic	$1 - \text{maximum (Sensitivity + Specificity)}$



No tiene en cuenta la proporción de presencias y ausencias en la muestra de validación

¿Cómo medimos la adecuación de un modelo?

- **Medidas de adecuación dependientes de umbral**

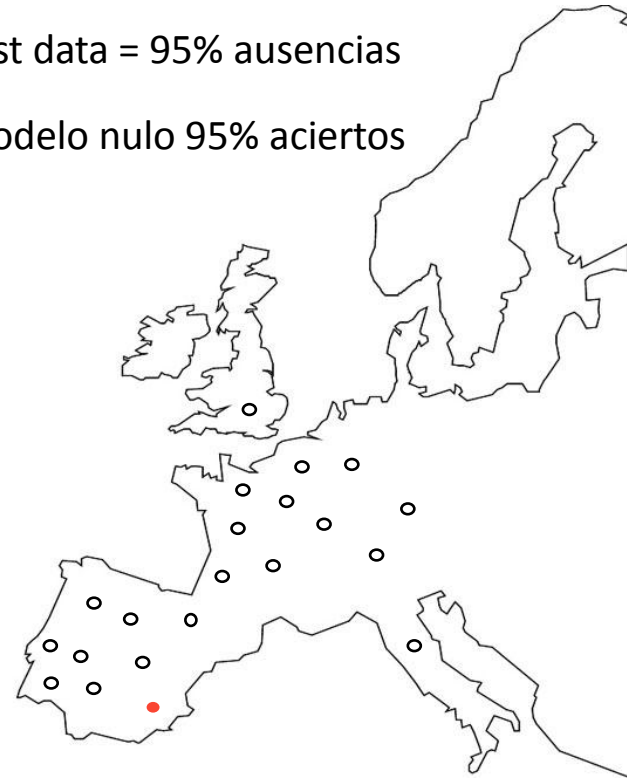
PCC → Percent of Correct Classifications



Área de distribución de la especie

Test data = 95% ausencias

Modelo nulo 95% aciertos



Test data sobre modelo nulo

¿Cómo medimos la adecuación de un modelo?

- **Medidas de adecuación dependientes de umbral**

Algunas consideraciones

Measure	Calculation
Sensitivity	$TP / (TP + FN)$
False negative rate	$1 - \text{Sensitivity}$
Specificity	$TN / (TN + FP)$
False positive rate	$1 - \text{Specificity}$
Percent correct classification	$(TP + TN) / n$
Positive predictive power	$TP / (TP + FP)$
Odds ratio	$(TP \times TN) / (FP \times FN)$
Kappa	$\frac{[(TP+TN) - (((TP+FN)(TP+FP) + (FP+TN)(FN+TN)) / n)]}{[n - (((TP+FN)(TP+FP) + (FP+TN)(FN+TN)) / n)]}$
True skill statistic	$1 - \text{maximum (Sensitivity + Specificity)}$

Busca un threshold que maximice la suma de sensibilidad y especificidad

Adecuación del modelo en relación a un modelo nulo de probabilidad de presencia

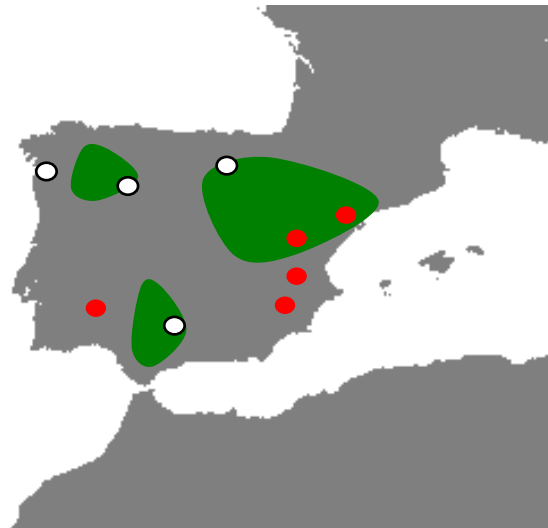
¿Cómo medimos la adecuación de un modelo?

- **Medidas de adecuación dependientes de umbral**

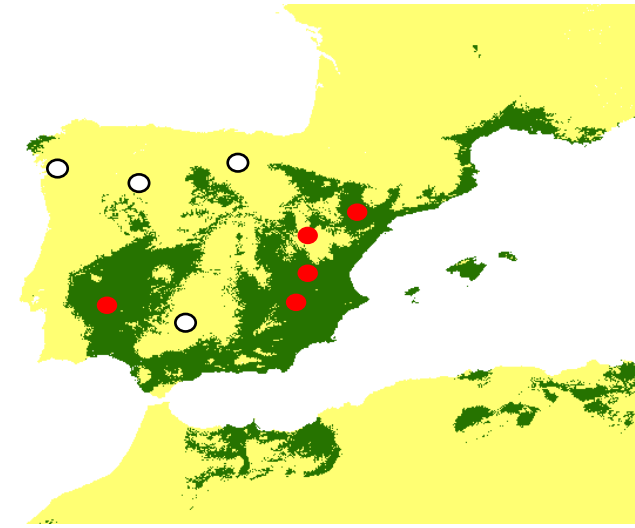
Algunas consideraciones (Kappa)



Test data

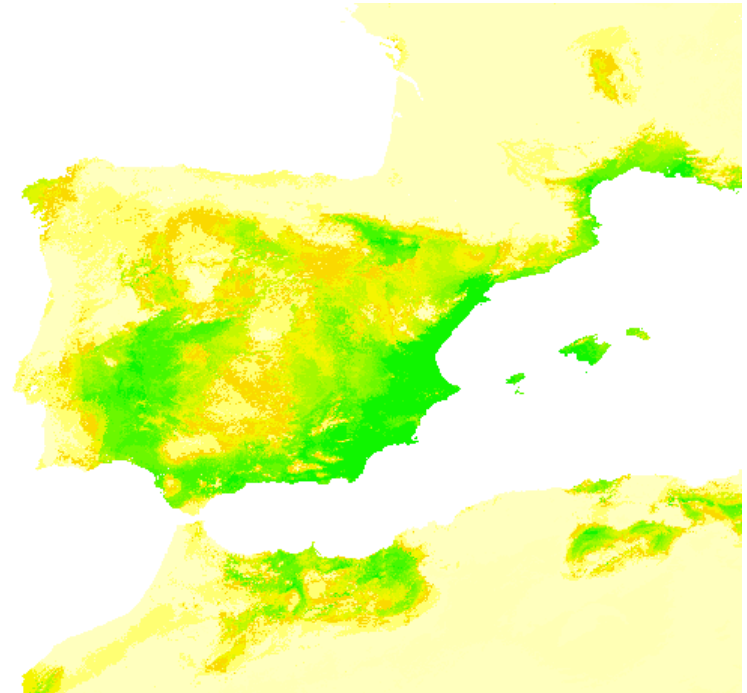
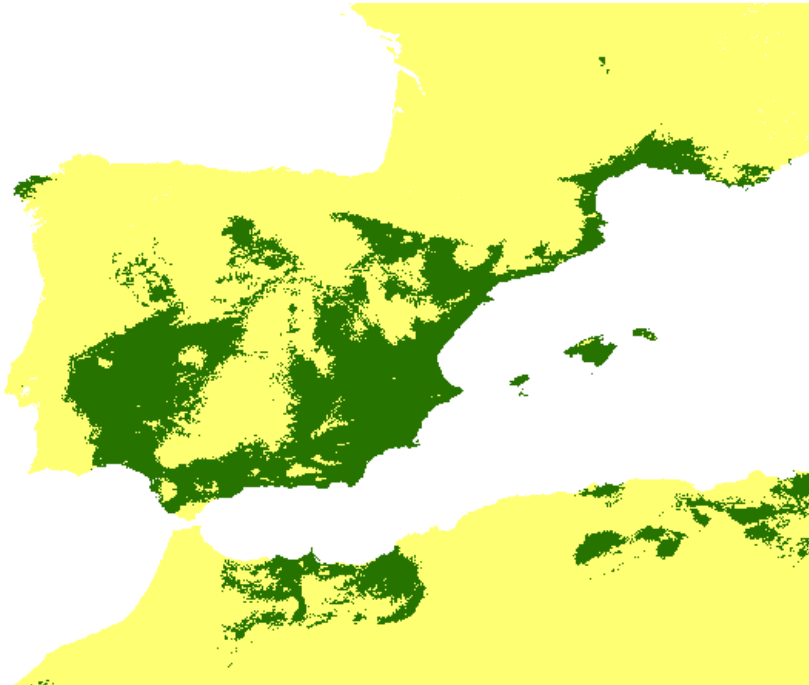


Modelo nulo

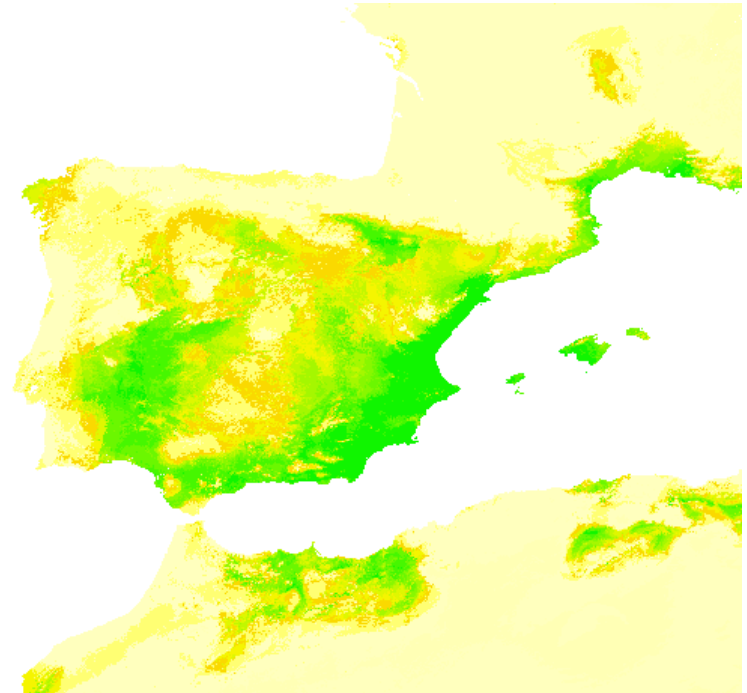
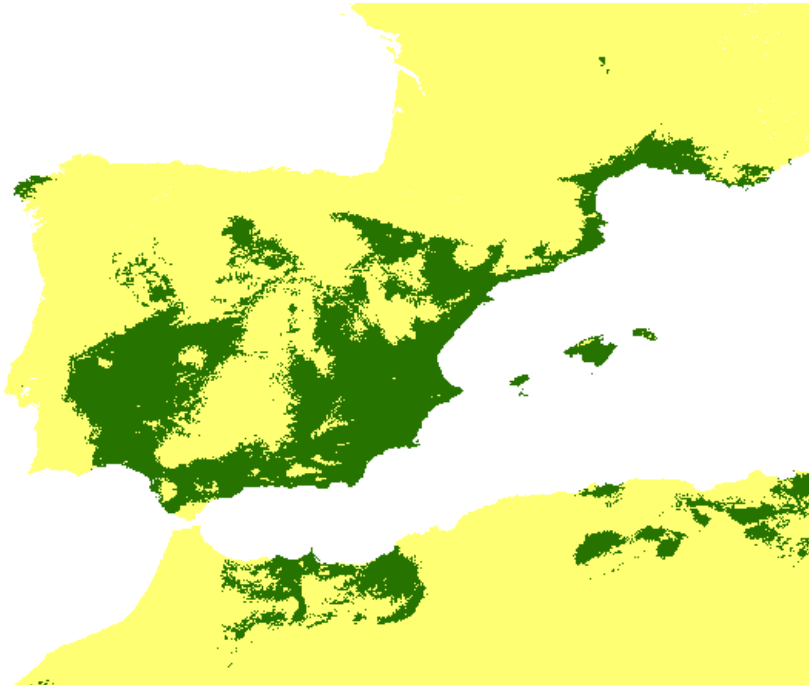


Modelo de predicción

¿Qué pasa cuando no tenemos predicciones categóricas?



¿Qué pasa cuando no tenemos predicciones categóricas?



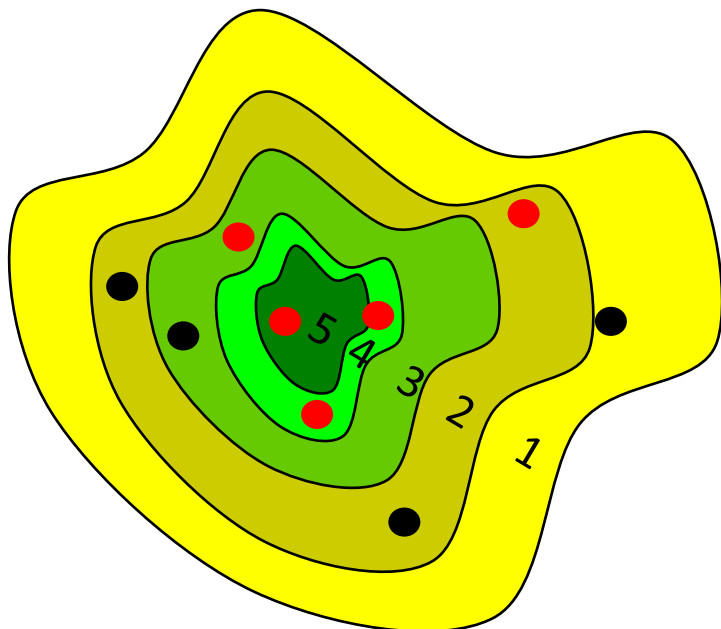
Medidas de adecuación dependientes de umbral

¿Cómo medimos la adecuación de un modelo?

- **Medidas de adecuación independientes de umbral**

Datos de evaluación categóricos + Predicciones categóricas

- Presencia test
- Ausencia test

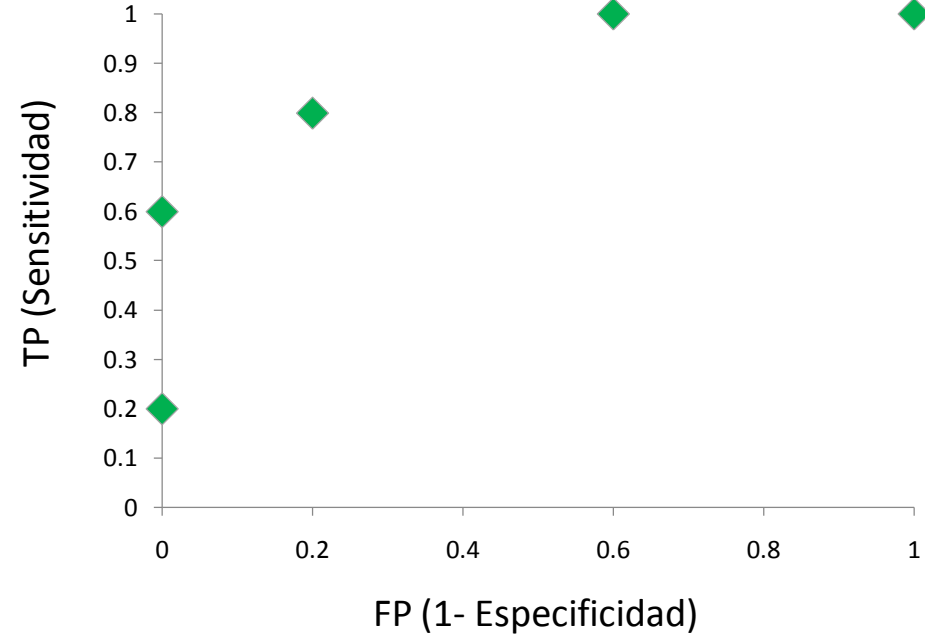
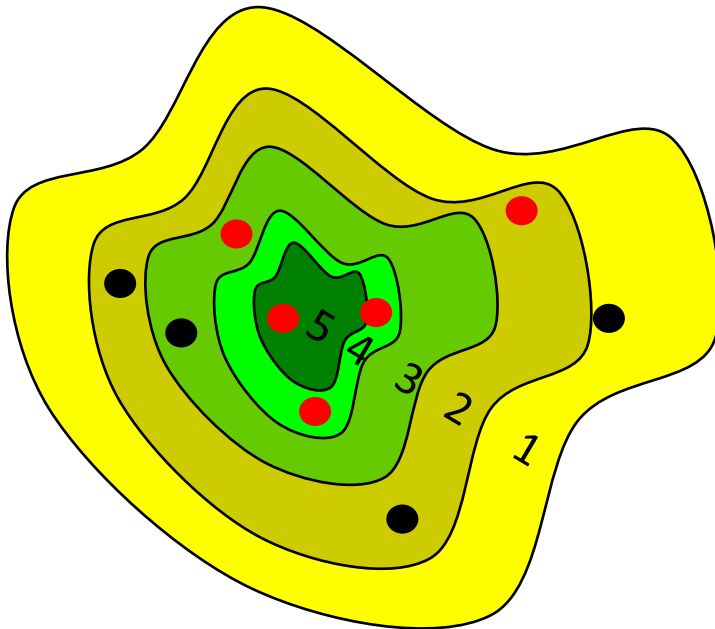


Umbral de corte	TP	FP	Sensibilidad (TP rate)	1- Especificidad (FP rate)
1	5	5	1	1
2	5	3	1	0.6
3	4	1	0.8	0.2
4	3	0	0.6	0
5	1	0	0.2	0

¿Cómo medimos la adecuación de un modelo?

- Medidas de adecuación independientes de umbral

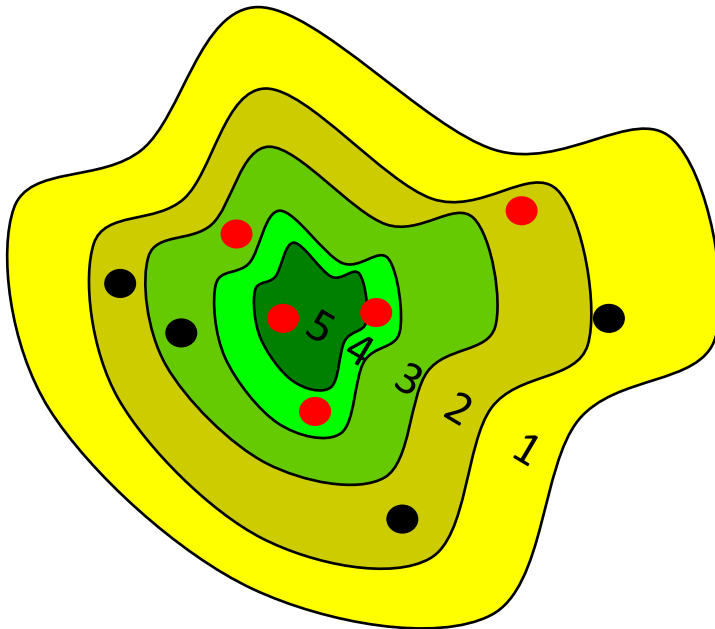
Presencia test
Ausencia test



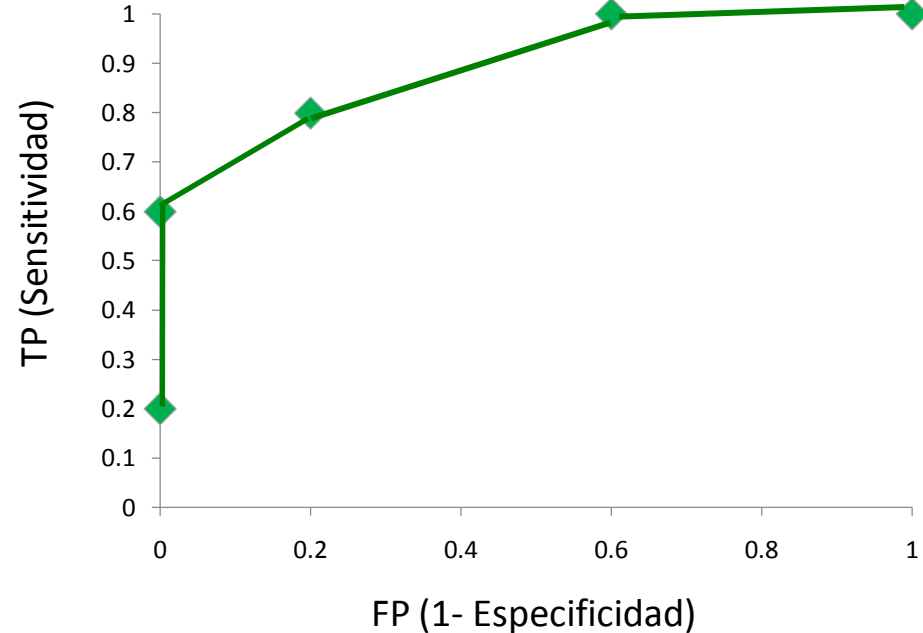
¿Cómo medimos la adecuación de un modelo?

- Medidas de adecuación independientes de umbral

Presencia test
Ausencia test

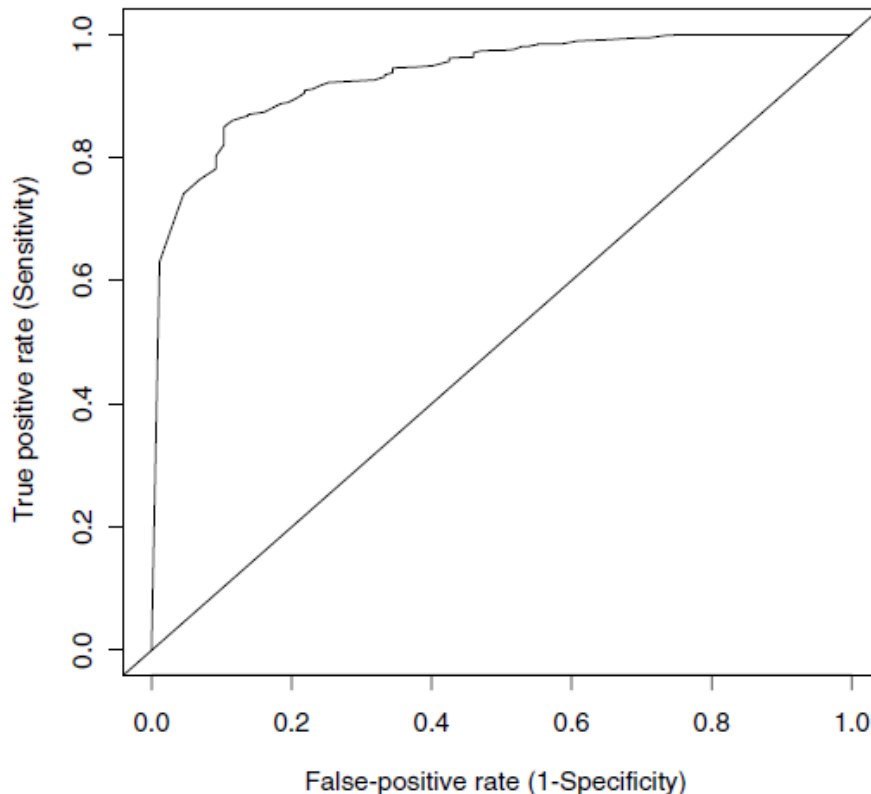


Area Under Receiver
Operating Characteristic
Curve (**AUC**)



¿Cómo medimos la adecuación de un modelo?

- Medidas de adecuación independientes de umbral

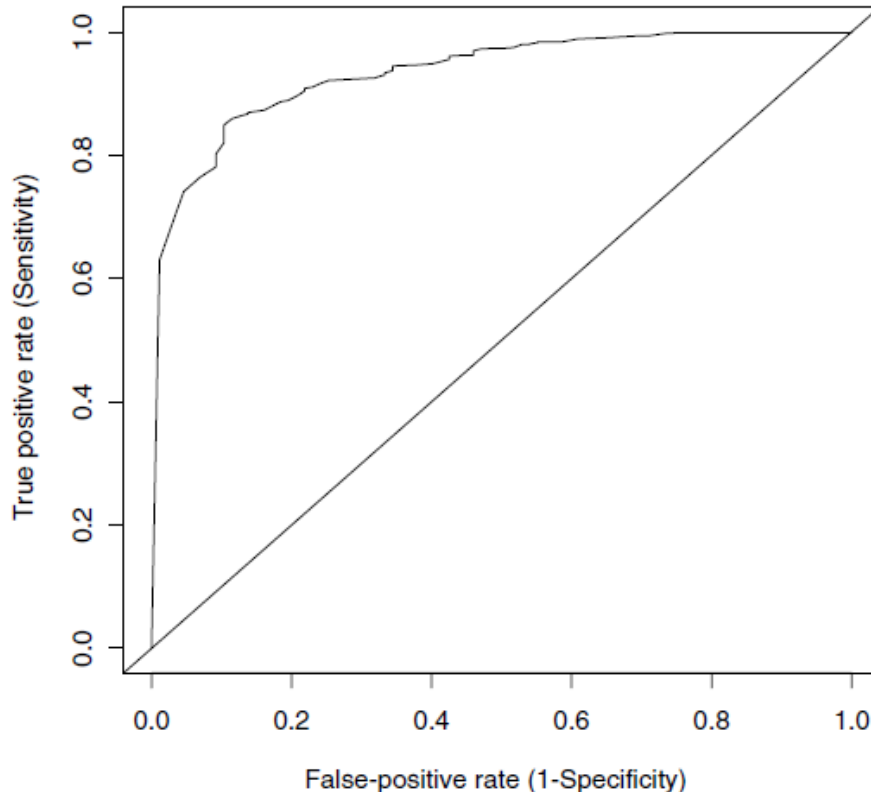


Area Under ROC Curve: A mayor pendiente, menor es el incremento de falsos positivos respecto a verdaderos positivos

La diagonal representa un modelo nulo en el que se añade un FP por cada TP. AUC=0.5

¿Cómo medimos la adecuación de un modelo?

- Medidas de adecuación independientes de umbral



Area Under ROC Curve: Probabilidad de que, entre una presencia y una ausencia seleccionadas al azar, el modelo prediga una idoneidad mayor para el dato de presencia.

AUC = 0.75 → El 75% de las ocasiones el modelo predecirá una idoneidad mayor para las presencias respecto a las ausencias

¿Cómo medimos la adecuación de un modelo?

- Medidas de adecuación con ausencias simuladas

Matriz de confusión

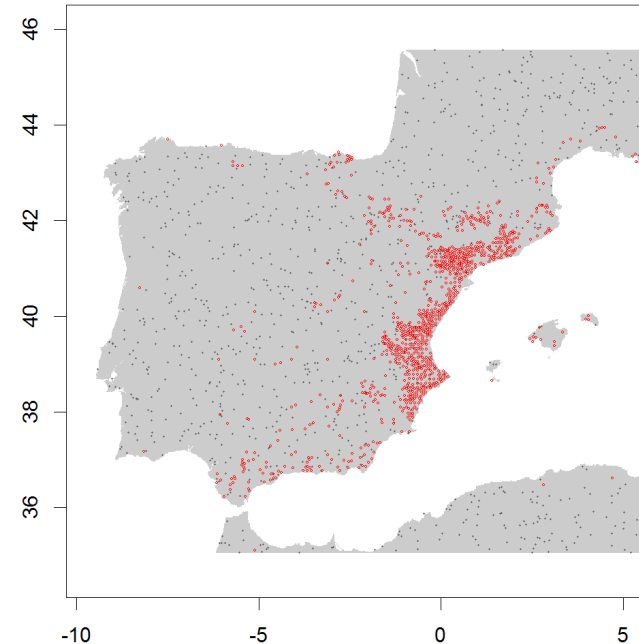
		Observación	
		Presencia	Ausencia
Predicción	Presencia	TP	FP
	Ausencia	FN	TN

TP: True Positive

FP: False Positive. Error de comisión

TN: True Negative

FN: False Negative. Error de omisión



¿Cómo medimos la adecuación de un modelo?

- Medidas de adecuación con ausencias simuladas

Matriz de confusión

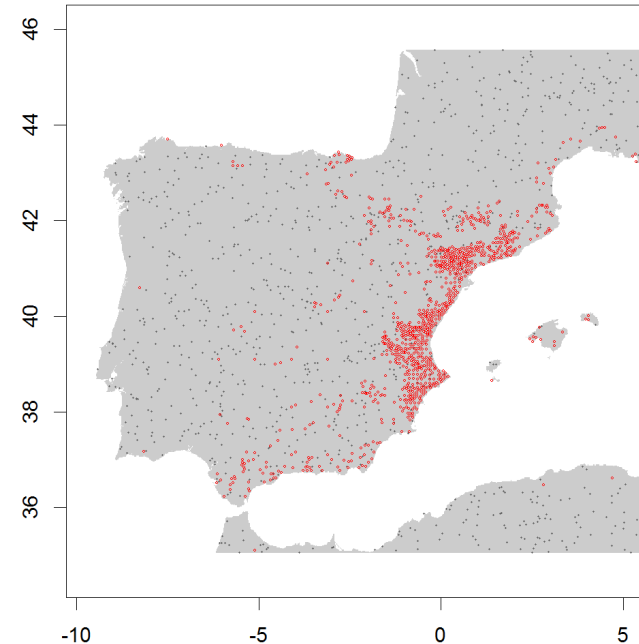
		Observación	
		Presencia	Background
Predicción	Presencia	TP	FP
	Ausencia	FN	TN

TP: True Positive

FP: False Positive

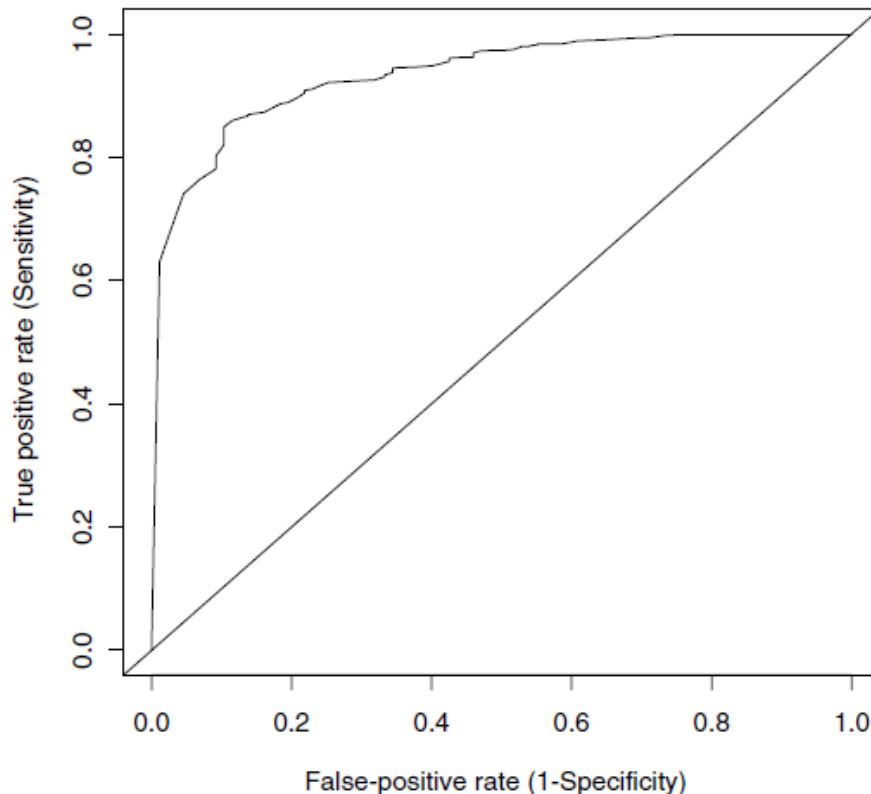
TN: True Negative

FN: False Negative. Error de omisión



¿Cómo medimos la adecuación de un modelo?

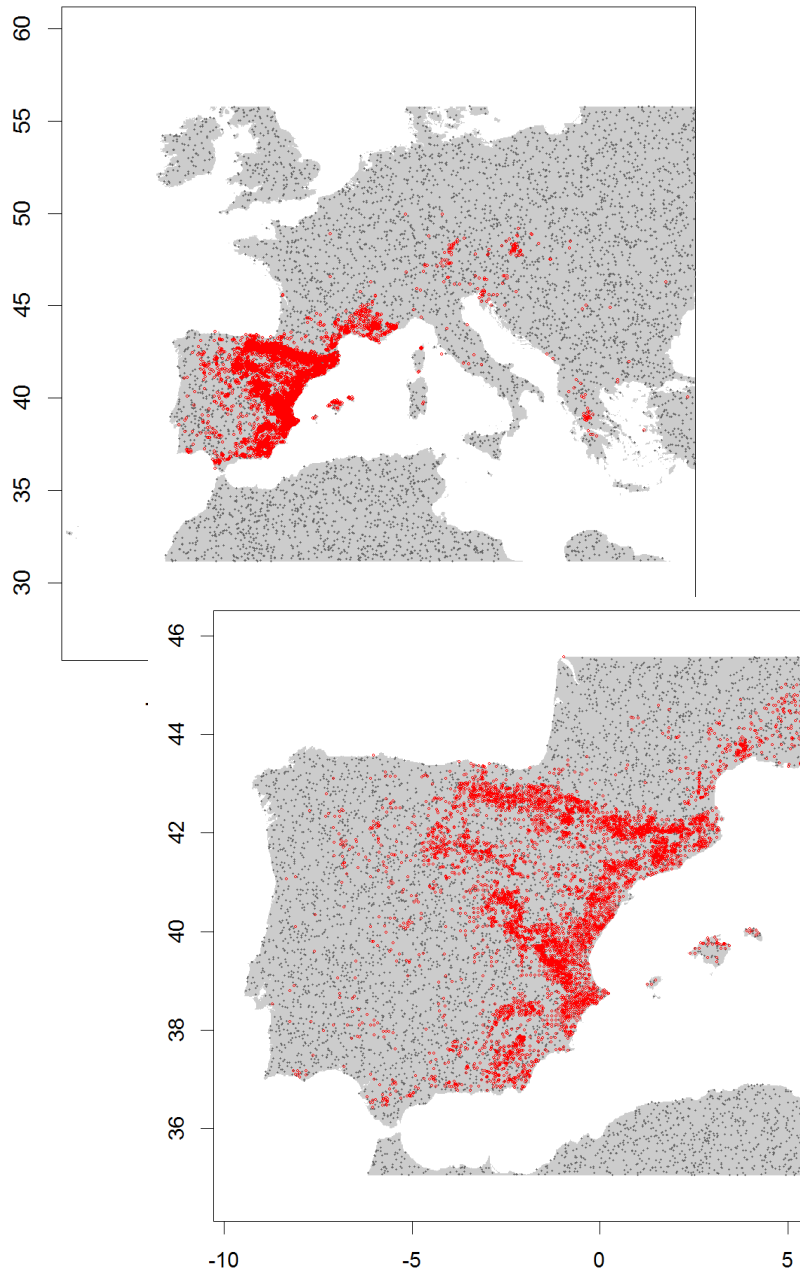
- Medidas de adecuación con ausencias simuladas



Area Under ROC Curve: Probabilidad de que, entre una presencia y un punto de background seleccionados al azar, el modelo prediga una idoneidad mayor para el dato de presencia.

$AUC = 0.75 \rightarrow$ El 75% de las ocasiones el modelo predecirá una idoneidad mayor para las presencias respecto a otro punto aleatorio del mapa

¿Cómo medimos la adecuación de un modelo?



Ventajas

- Permite calcular la adecuación con valores continuos de probabilidad de aparición
- No depende de la prevalencia

Consideraciones

- Pondera igual los errores de omisión y comisión
- Extensiones de estudio más grandes producen mayores valores de AUC
- Modelos de diferentes especies y extensiones de trabajo no son comparables

¿De dónde sacamos los datos de validación?

- Lo ideal es validar con un conjunto totalmente independiente de datos. Muy complicado en la práctica.
- Validar con datos no incluidos en el calibrado.
 - **Data splitting.** 2-fold. Ratio training:test (Huberty 1994)
50:50 → 2 variables predictoras
66:33 → 6 variables predictoras
75:25 → >10 variables predictoras
$$1/(1 + \sqrt{p - 1})$$
 - **k-fold crossvalidation:** dividir los datos k veces, obteniendo k estimas de adecuación que son promediadas. K-1 (training data) y 1 (test data)
 - **Bootstrap:** partición iterativa de los datos, para calibrar con unos, y evaluar con otros.
 - **Leave-one-out:** similar a K-fold. Tras la validación los datos son utilizados para calibrar el modelo. Sólo para muestras pequeñas.

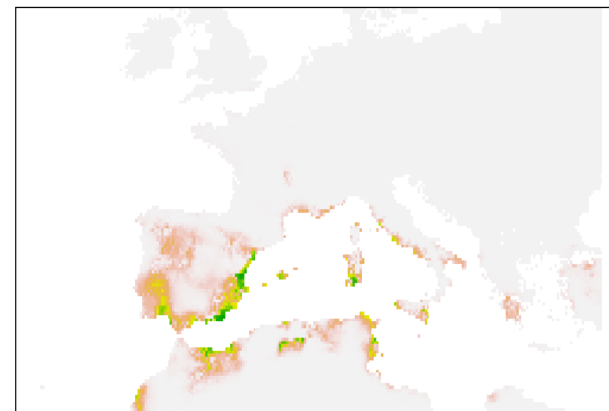
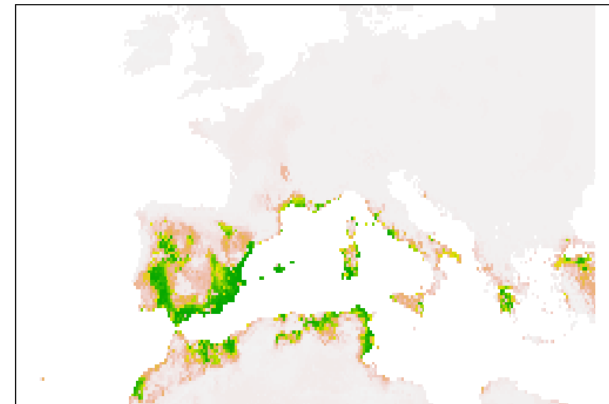
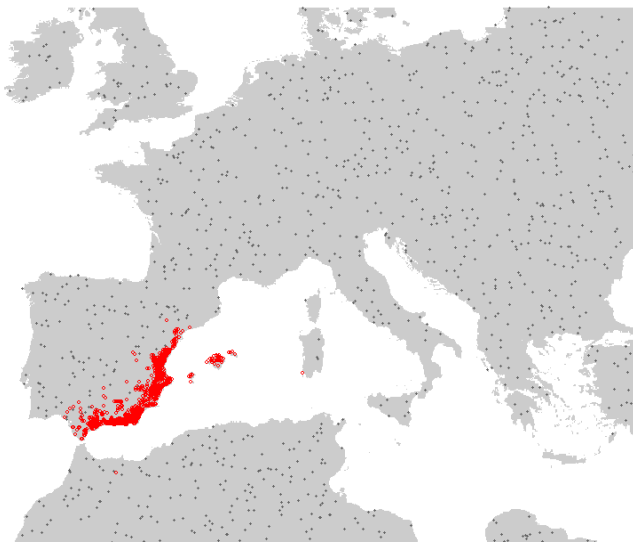
¿Cómo medimos la adecuación de un modelo?

La evaluación predictiva es sólo uno de los aspectos a contemplar para evaluar la incertidumbre de un modelo

Modeling step	Criterion	Description	Reference
Conceptual formulation	Precision	Ability to replicate system parameters	(Morrison <i>et al.</i> , 1998)
	Specification	Does the model address the problem? Does it describe the true relationship?	(Barry & Elith, 2006)
	Ecological realism	Is conceptual formulation consistent with ecological theory?	(Austin, 1980, 2002; Austin <i>et al.</i> , 2006; Barry & Elith, 2006)
Statistical formulation	Realism	Account for relevant variables and relationships?	(Morrison <i>et al.</i> , 1998)
Model calibration	Verification	Is the model logic correct?	(Rykiel, 1996)
	Calibration	Parameter estimation or model fitting and selection	(Rykiel, 1996; Chatfield, 1995)
Model evaluation	Validity, performance	Capability to produce empirically correct predictions to a degree of accuracy that is acceptable given the intended application of the model	(Morrison <i>et al.</i> , 1998; Rykiel, 1996; Barry & Elith, 2006)
	Appeal, credibility	Accepted by users, matches user intuition, sufficient degree of belief to justify use for intended application	(Morrison <i>et al.</i> , 1998; Rykiel, 1996)

¿Cómo medimos la adecuación de un modelo?

Verosimilitud de los mapas de predicción

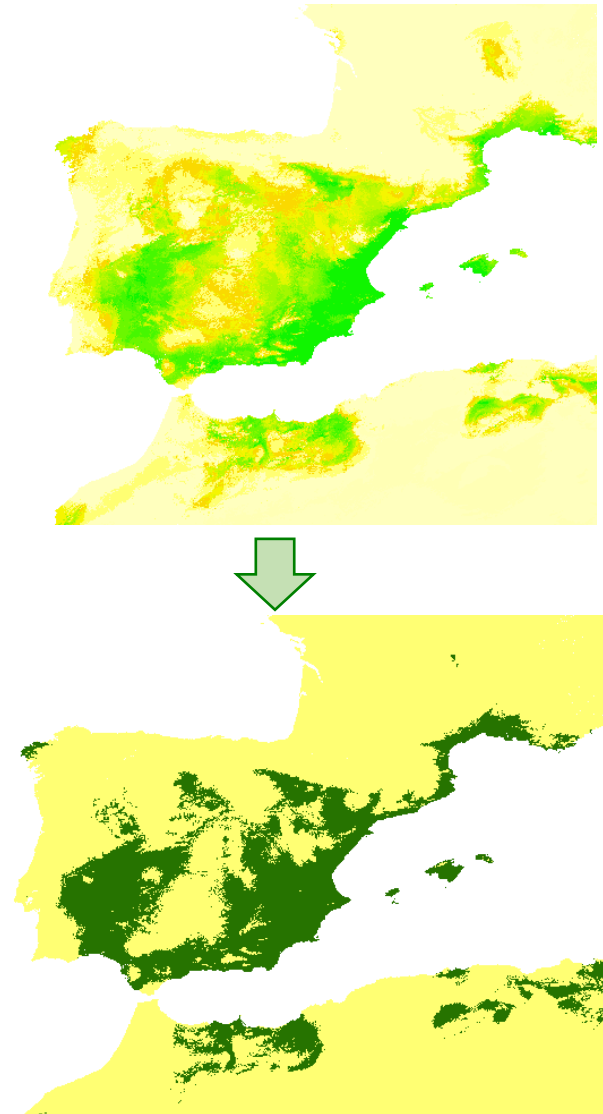


Thresholds

Supone pasar de un modelo continuo a uno binario

Se selecciona un umbral de referencia

- A todos los valores por encima del umbral se les asigna un valor de 1
- A todos los valores por debajo del umbral se les asigna un valor de 0



Thresholds

Amplio abanico de criterios

Criterio	Definición
Umbral fijos (Manel et al. 1999)	Método subjetivo. Tradicionalmente fijado en 0.5. Métodos de regresión logística y similares
Sensibilidad=especificidad (Fielding & Bell 1997)	Umbral de probabilidad donde el ratio de falsos positivos es igual al de falsos negativos
Max(Sensibilidad+Especificidad) (Manel et al. 2001)	Umbral que produce el valor total más alto de sensibilidad y especificidad. Equivalente a encontrar un punto de la curva ROC con tangente=1
Max(Kappa) (Moisen et al. 2006)	Umbral de probabilidad donde se maximiza el valor de Kappa
Max(Porcentaje Clasificaciones Correctas) (Cramer, 2003)	Umbral de probabilidad donde se maximiza el valor de PCC
Prevalencia predicha=prevalencia observada (Cramer, 2003)	Umbral de probabilidad donde la prevalencia predicha sea igual a la observada
Prevalencia observada (Cramer, 2003)	Umbral de probabilidad igual al valor de prevalencia de los datos de calibrado.
Media(valores predichos) (Cramer, 2003)	Media o mediana de los valores de todos los valores de probabilidad predichos por el modelo
Min(1-sensibilidad)^2+(especificidad -1)^2 (Cantor et al. 1999)	Umbral de probabilidad que minimiza la distancia entre la curva ROC y la esquina superior izquierda del ROC plot
Max(Sensibilidad)&Min(Especificidad) (Freeman & Moisen 2008)	Umbral de probabilidad que maximiza la sensibilidad para un mínimo de especificidad definido por el usuario.

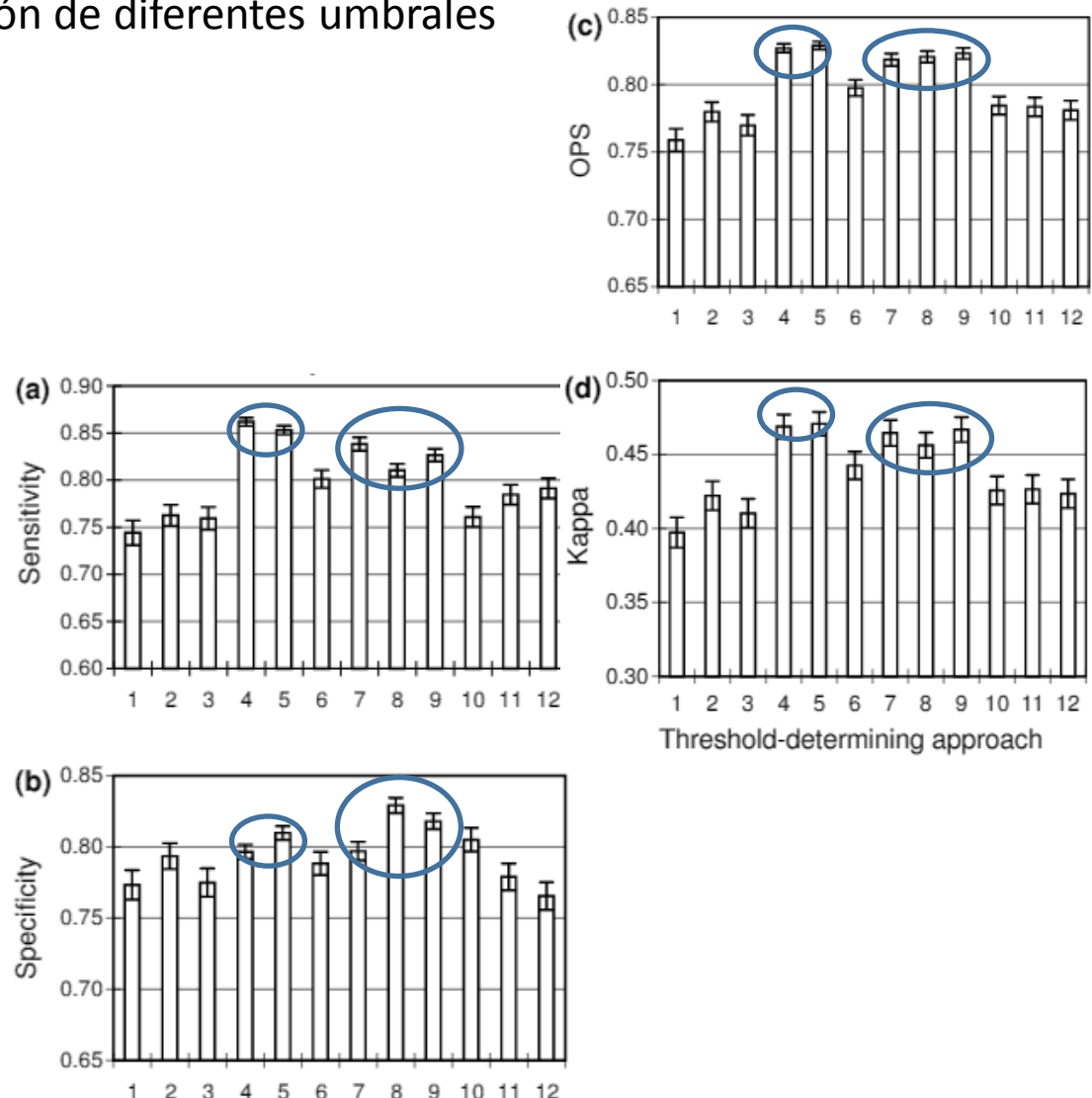
Thresholds

Adecuación predictiva en función de diferentes umbrales

Según Freman & Moisen 2008 todos los criterios tienden a converger con AUC altos y valores de prevalencia próximos a 0.5

Según Liu et al. 2005 hace una selección de los mejores umbrales

- Prevalencia observada
- Promedio probabilidad predicha
- Sensibilidad + Especificidad
- Sensibilidad = Especificidad
- Punto de la curva ROC más próximo a la esquina izquierda



Thresholds

Consideraciones

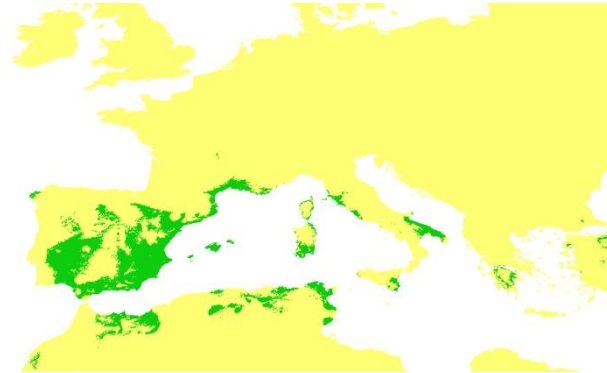
El criterio seleccionado dependerá de nuestros objetivos

Connotaciones ecológicas del concepto “presencia”

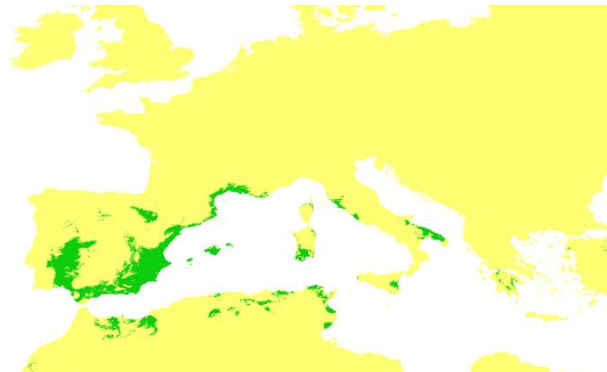
Los modelos con pocas presencias son muy sensibles a la elección de umbral

Pasar a binario siempre supone pérdida de información

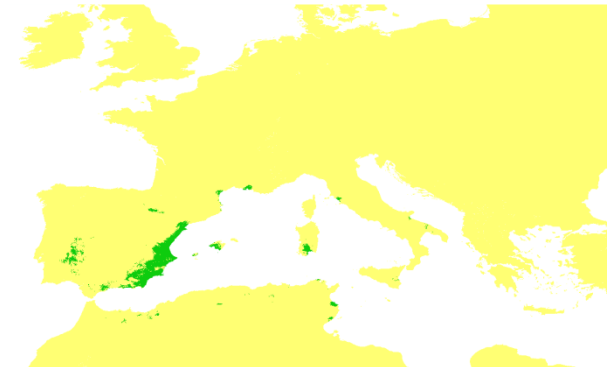
Salvo los umbrales de selección basados en la curva ROC, el resto dependen de datos de presencia/ausencia



0.5

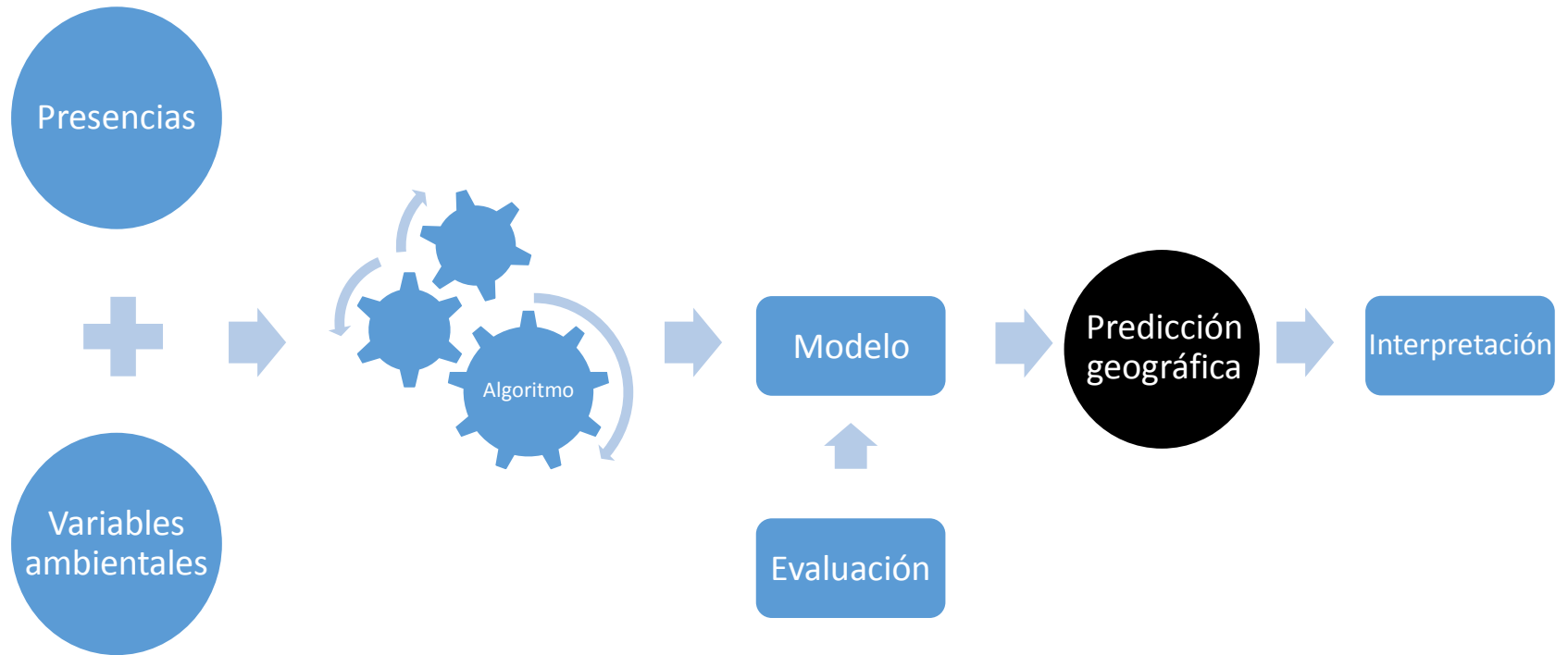


0.7

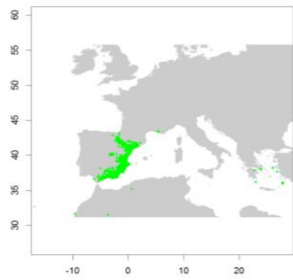


0.9

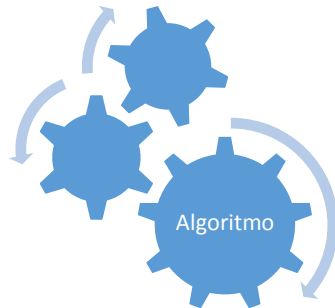
Proyecciones en el espacio y en el tiempo



Proyecciones en el espacio y en el tiempo



Presencias



Modelo



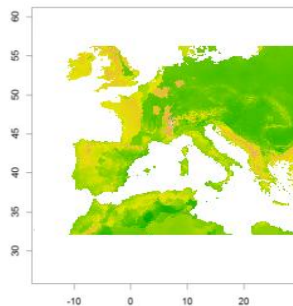
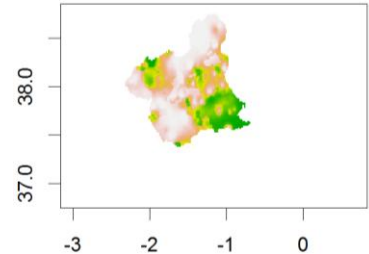
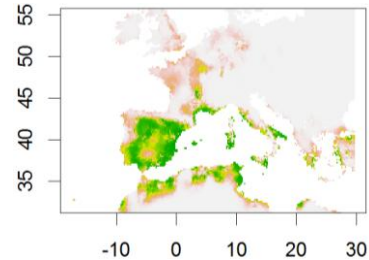
Evaluación



Condiciones de calibrado

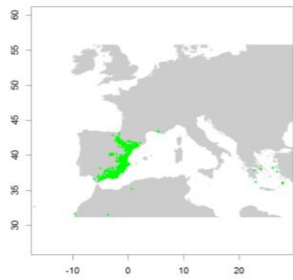


Nuevas condiciones

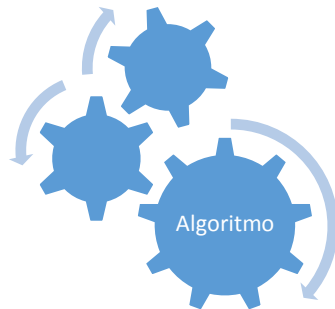


Variables ambientales

Proyecciones en el espacio y en el tiempo



Presencias



Modelo



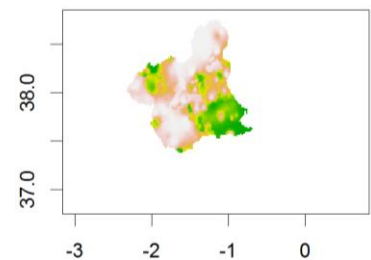
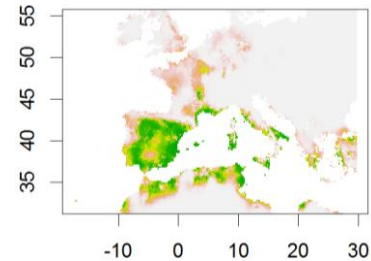
Evaluación



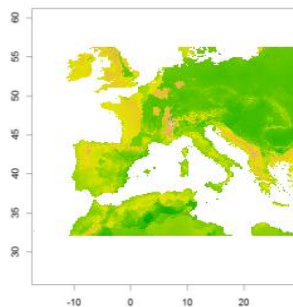
Condiciones de calibrado



Nuevas condiciones



Pueden ser muy diferentes de las condiciones de calibrado!!!!!!



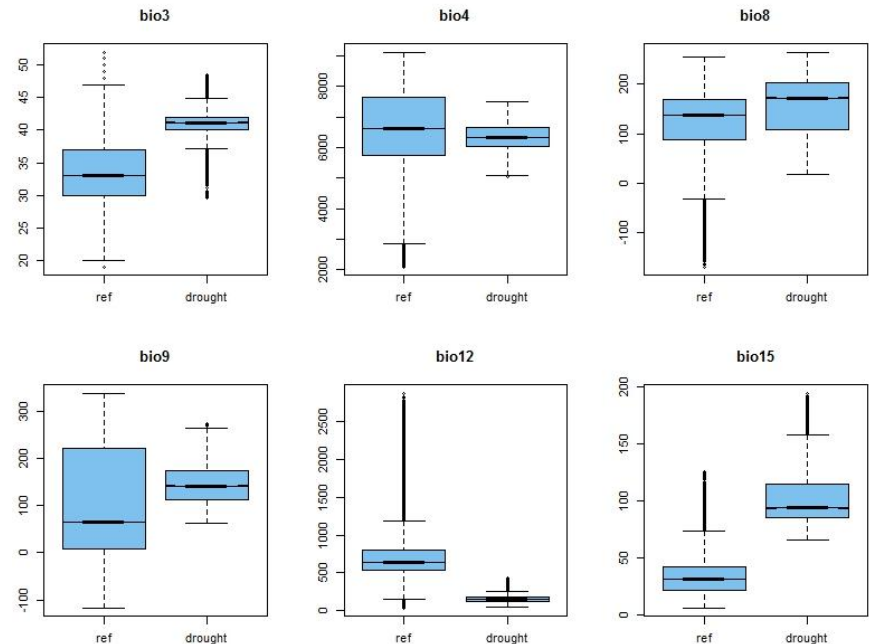
Variables ambientales

Proyecciones en el espacio y en el tiempo

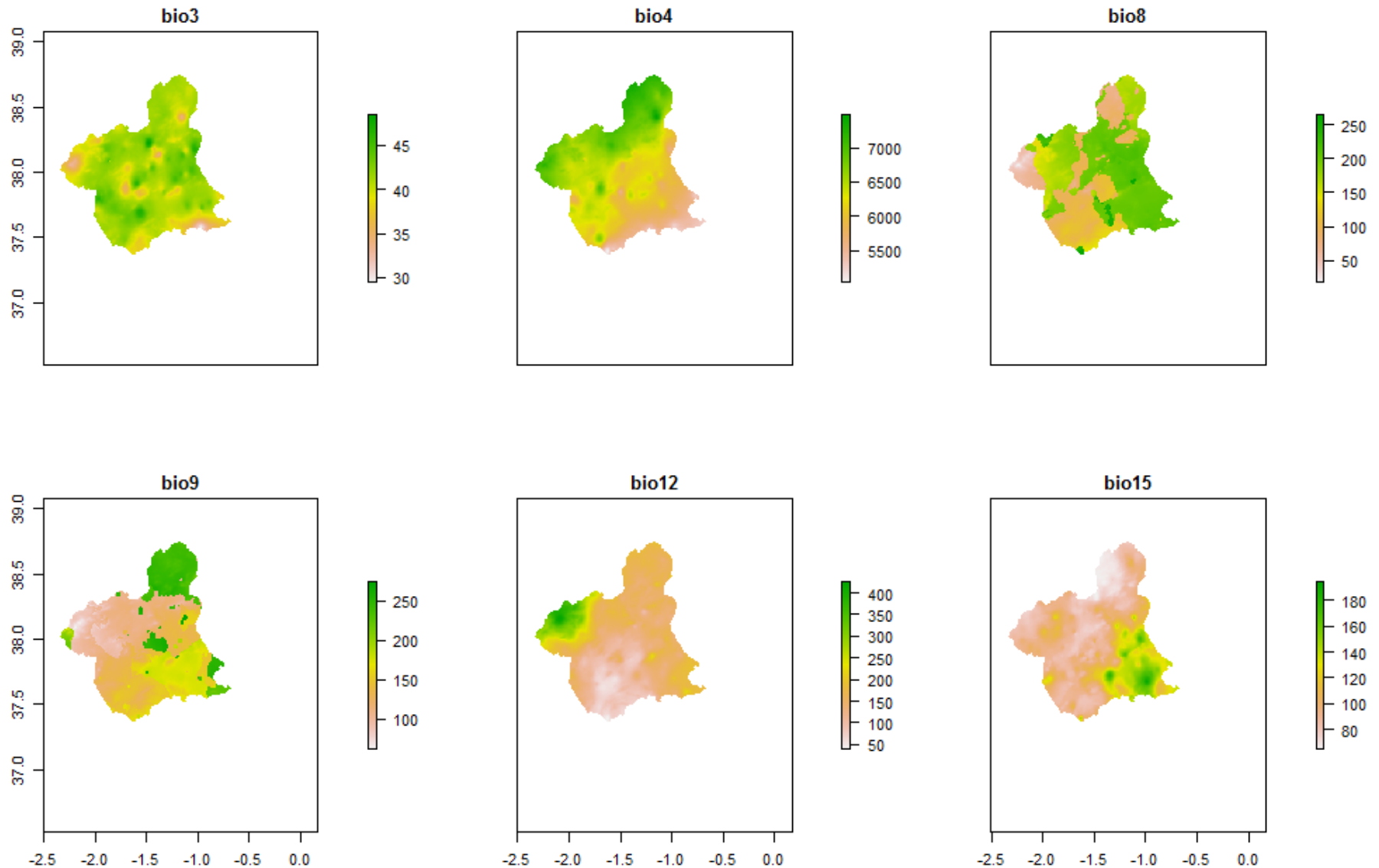
Extrapolación El modelo extrapola cuando proyecta sobre variables no incluidas en los datos de calibrado

Multivariate Environmental Similarity Surfaces (MESS) Medida de similitud entre cualquier punto dado del mapa y los puntos de calibrado, según las variables predictoras (Elith et al. 2010)

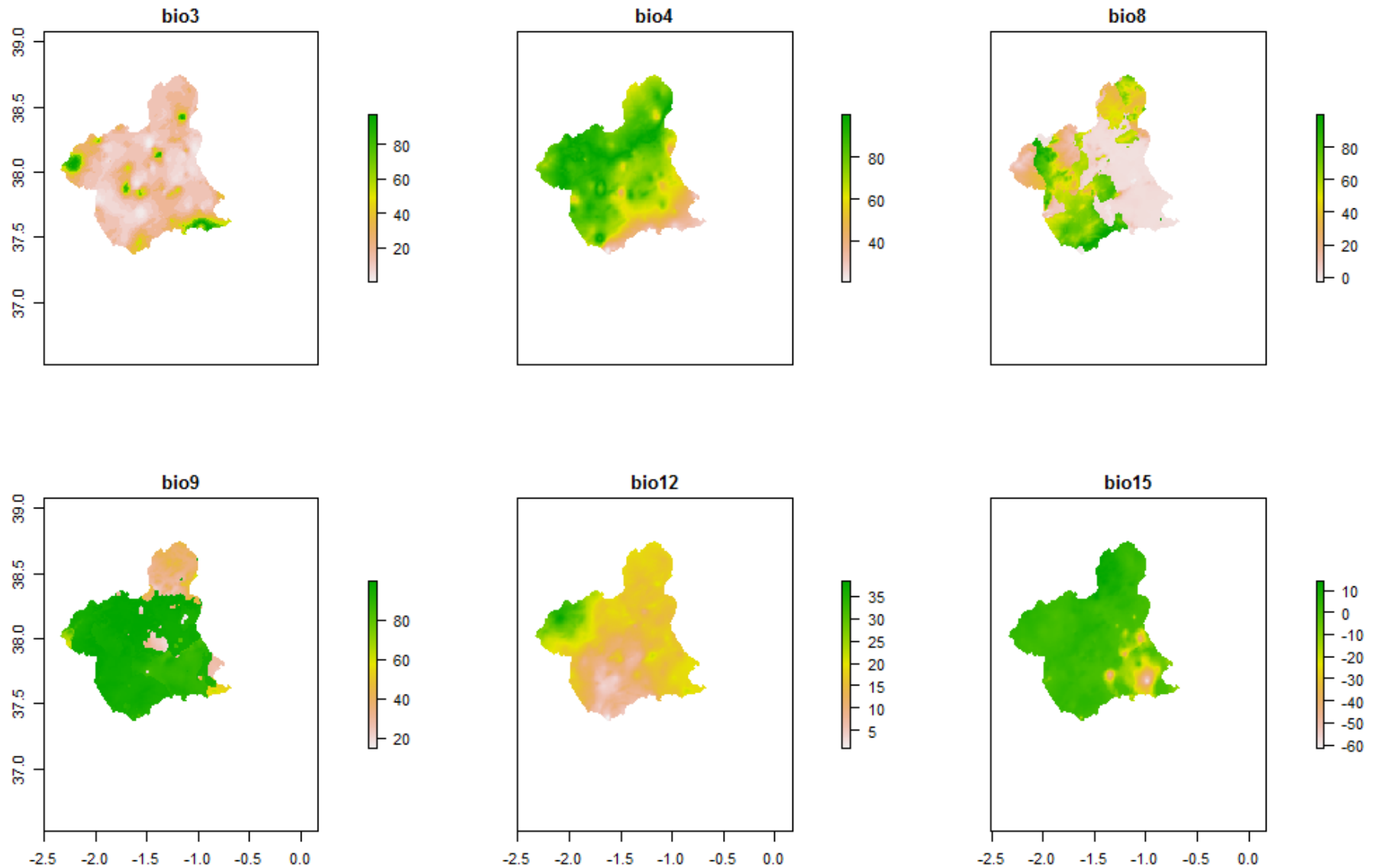
Mayor dissimilaridad → mayor extrapolación



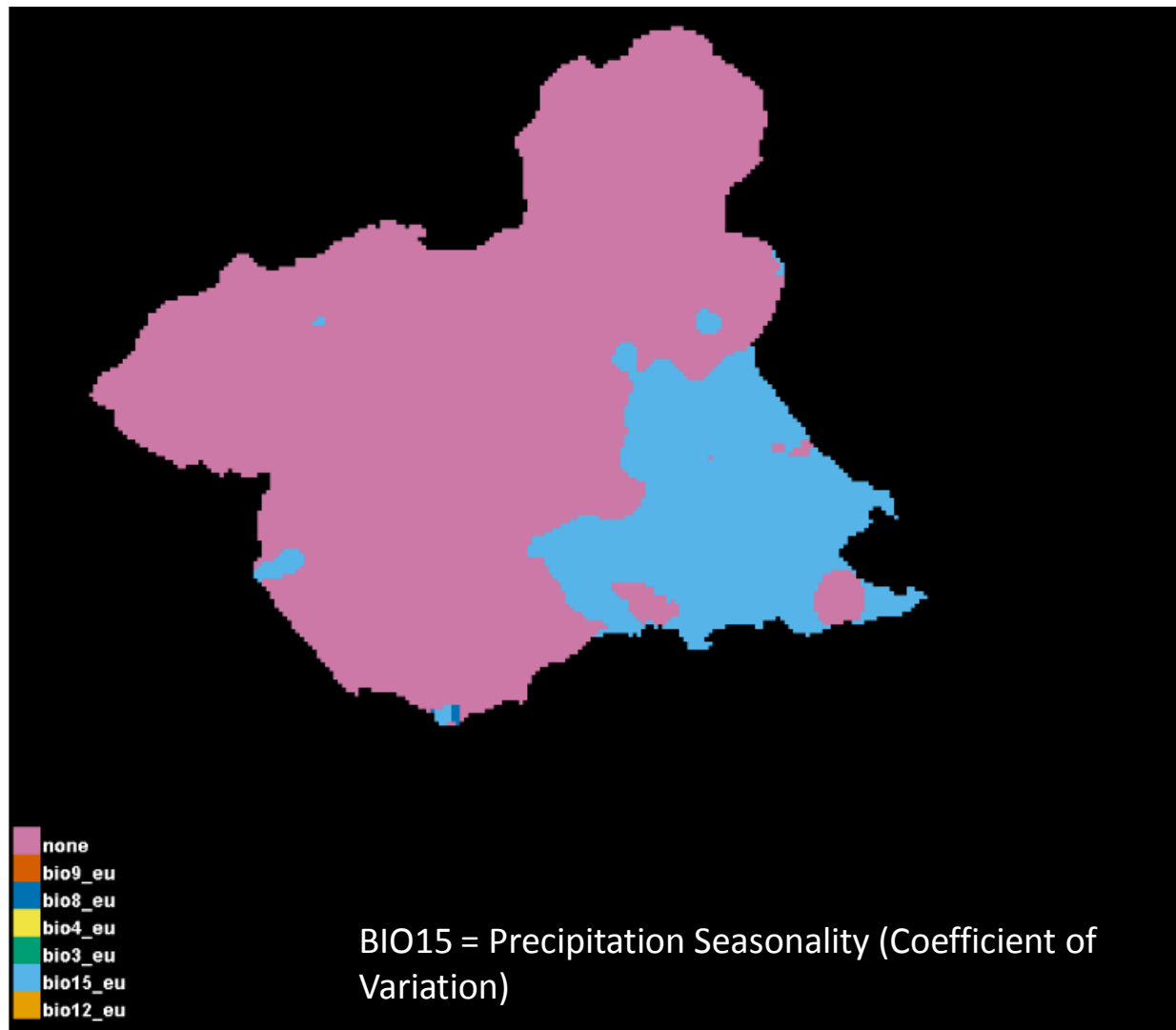
Proyecciones en el espacio y en el tiempo



Proyecciones en el espacio y en el tiempo



Proyecciones en el espacio y en el tiempo



Proyecciones en el espacio y en el tiempo



AUC mide la
adecuación para los
datos de calibrado.
NO PROYECCIONES

	Proyecciones en el espacio	Proyecciones en el tiempo
Usos	Especies invasoras	Cambio climático, paleodistribuciones
Variables	Variables área de invasión. Mismas variables, mismo nombre y resolución.	Variables futuro o pasado. Mismo nombre, mismas variables. Hay variables no disponibles: ndvi, human footprint
Calibración del modelo	Distribución de origen	Momento actual
Validación	Es posible si la especie empieza a darse en el rango de invasión	Es posible con datos de paleoregistros
Consideraciones		Incertidumbre asociada a las capas climáticas