



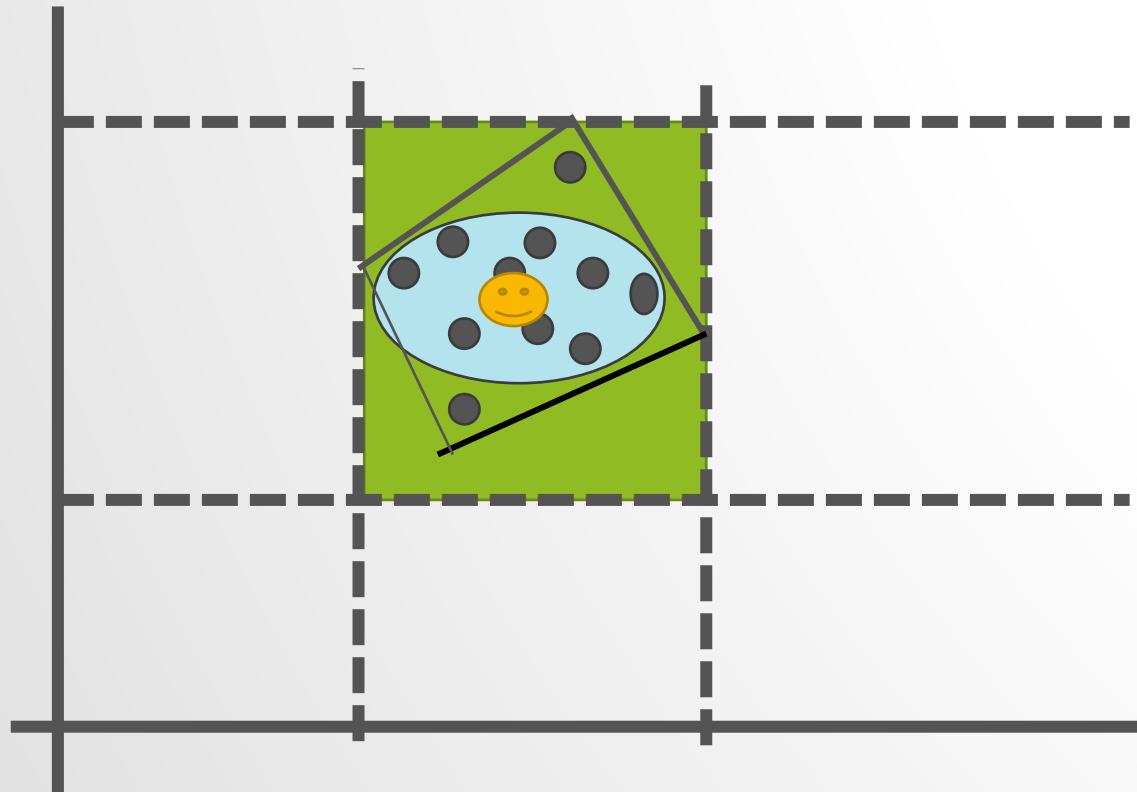
EVALUACIÓN ROC PARCIAL

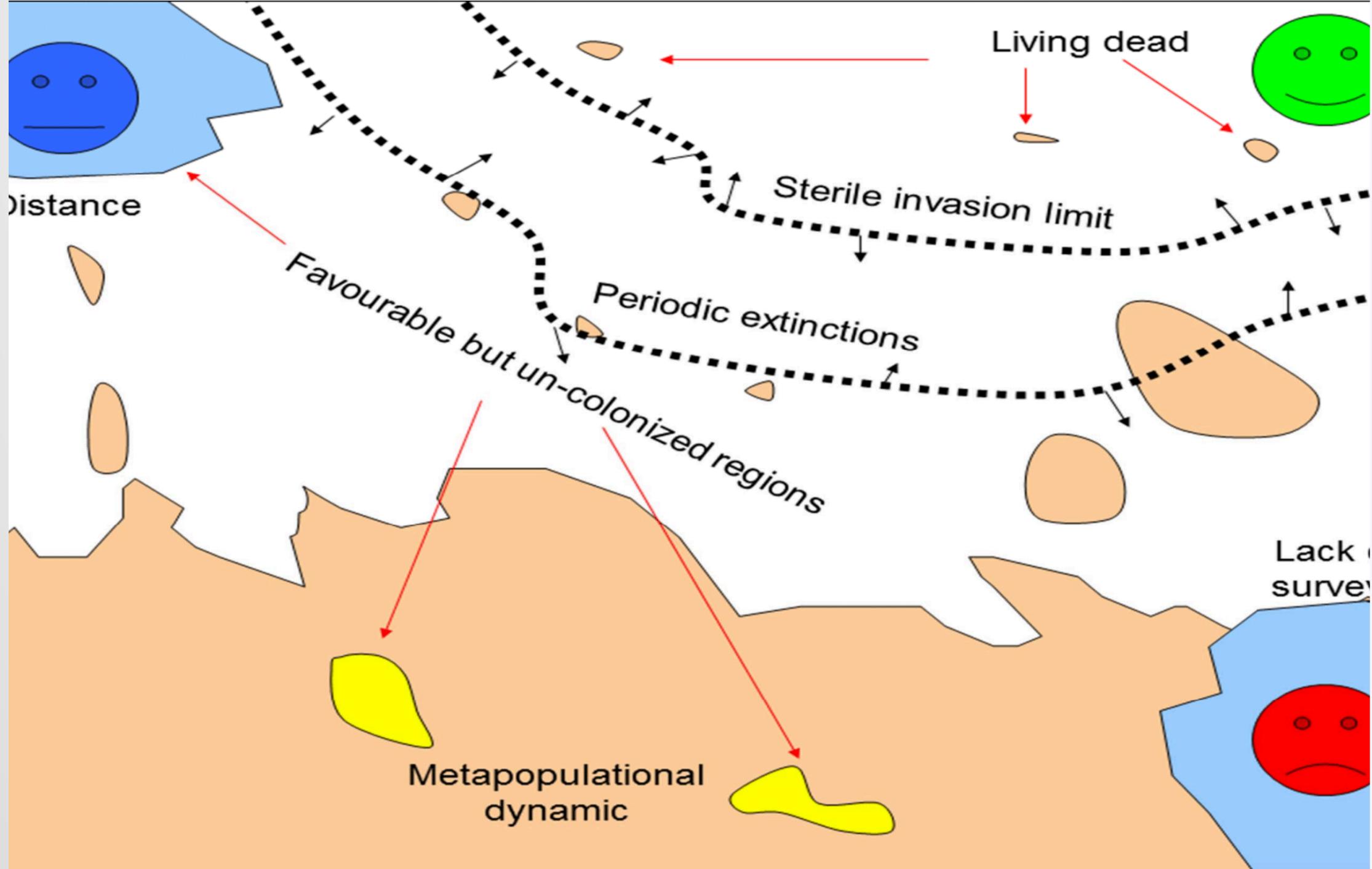
Modelaje de nicho

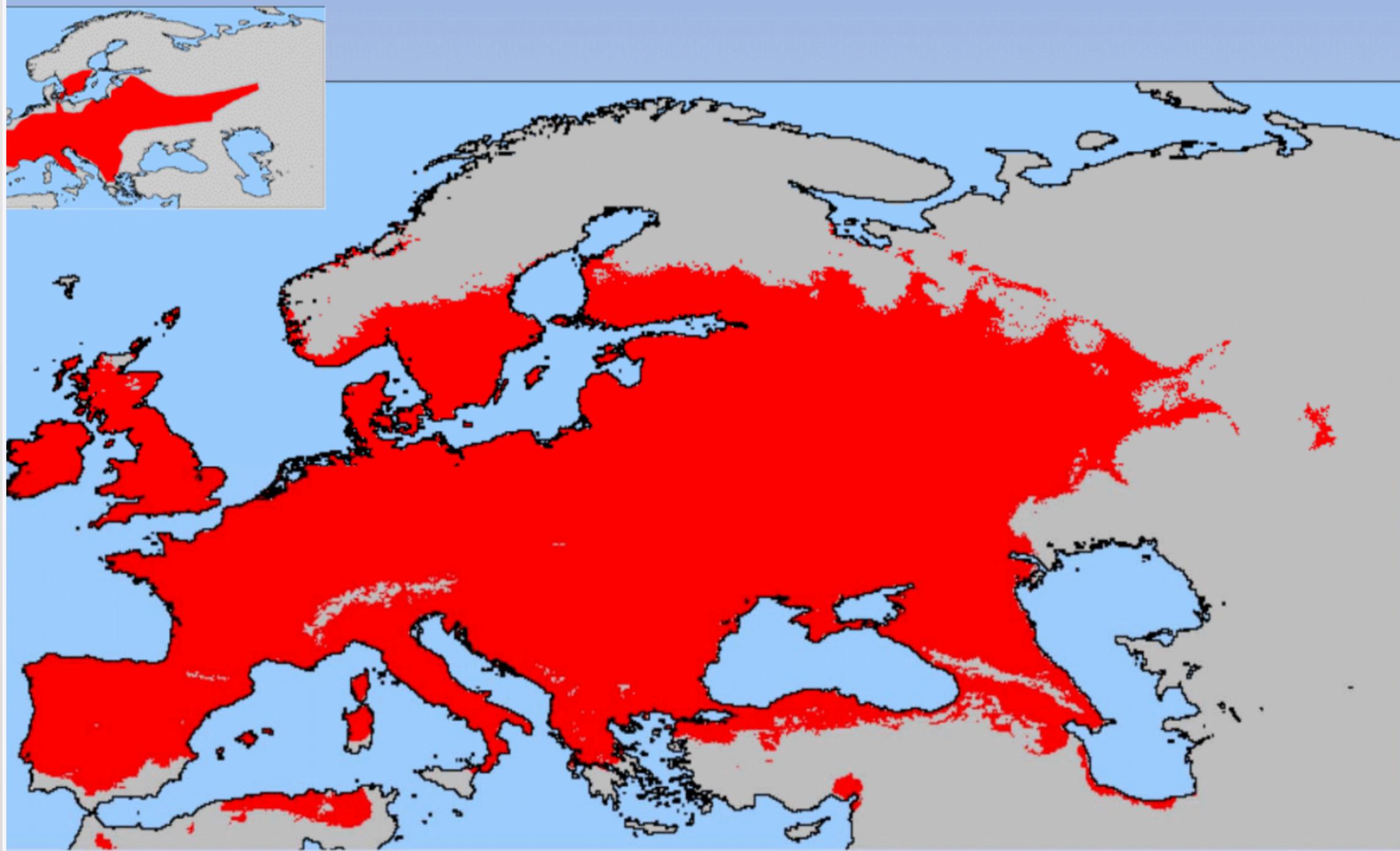
¿CÓMO VALIDAR UN MODELO?

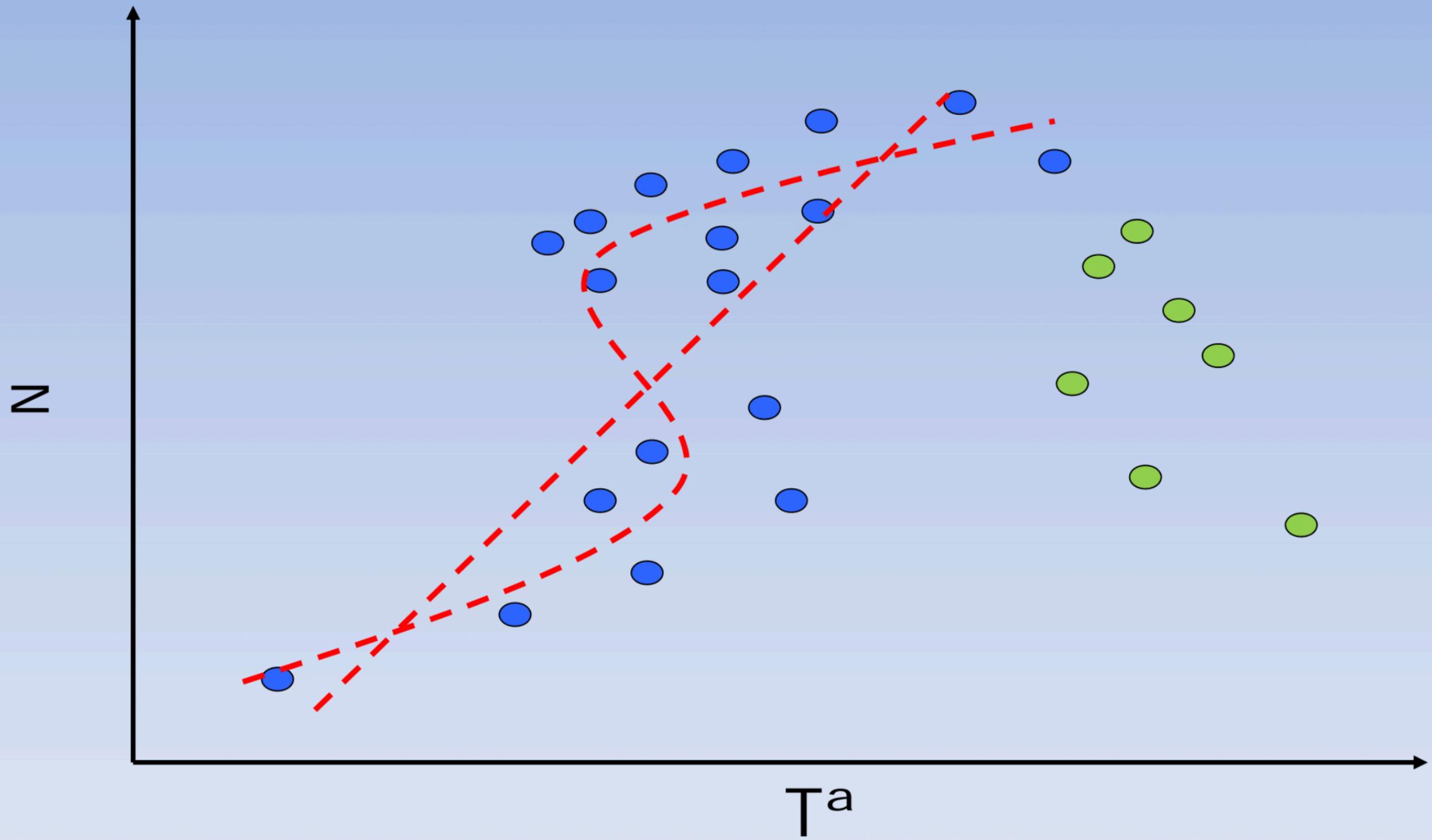
- Nicho
- Distribución potencial
- Distribución realizada

NICHO





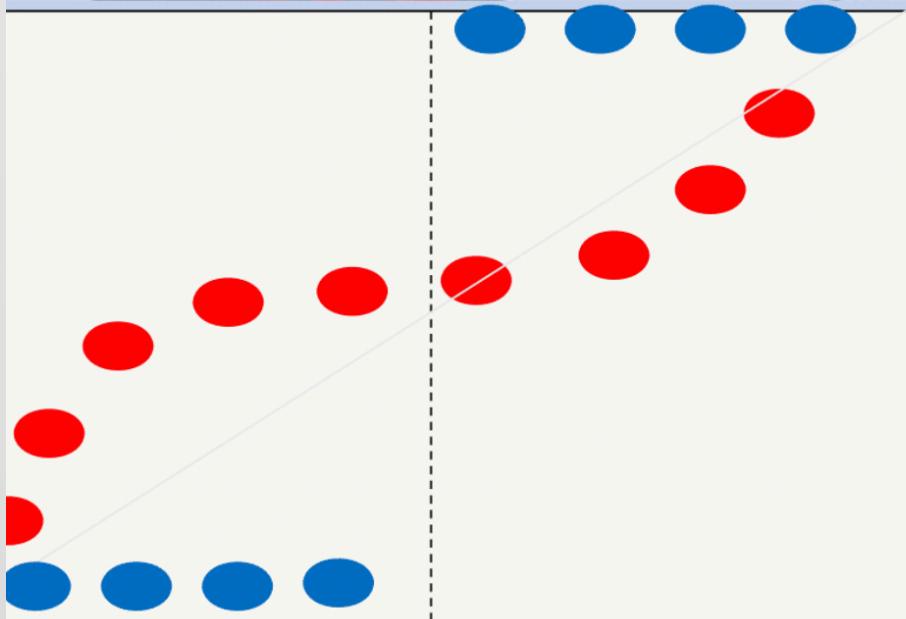




VALIDACIÓN MEDIR LA BONDAD DE UN MODELO

- Discriminación
 - Capacidad de discriminar las observaciones presencia de las de ausencia
- Calibración
 - Capacidad de que la probabilidad de presencia predicha represente la proporción observada de presencias.

Discriminación



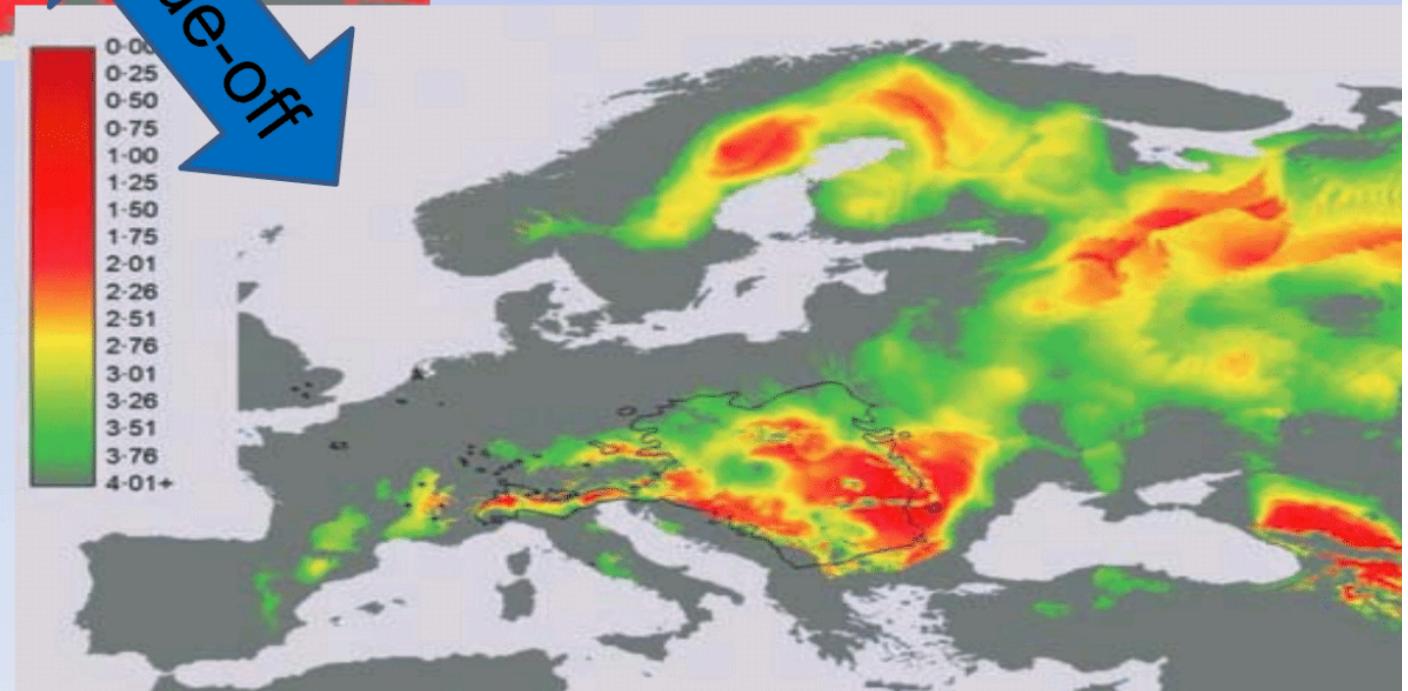
Global Ecology and Biogeography, (Global Ecol. Biogeogr.) (2013) 22, 508–516

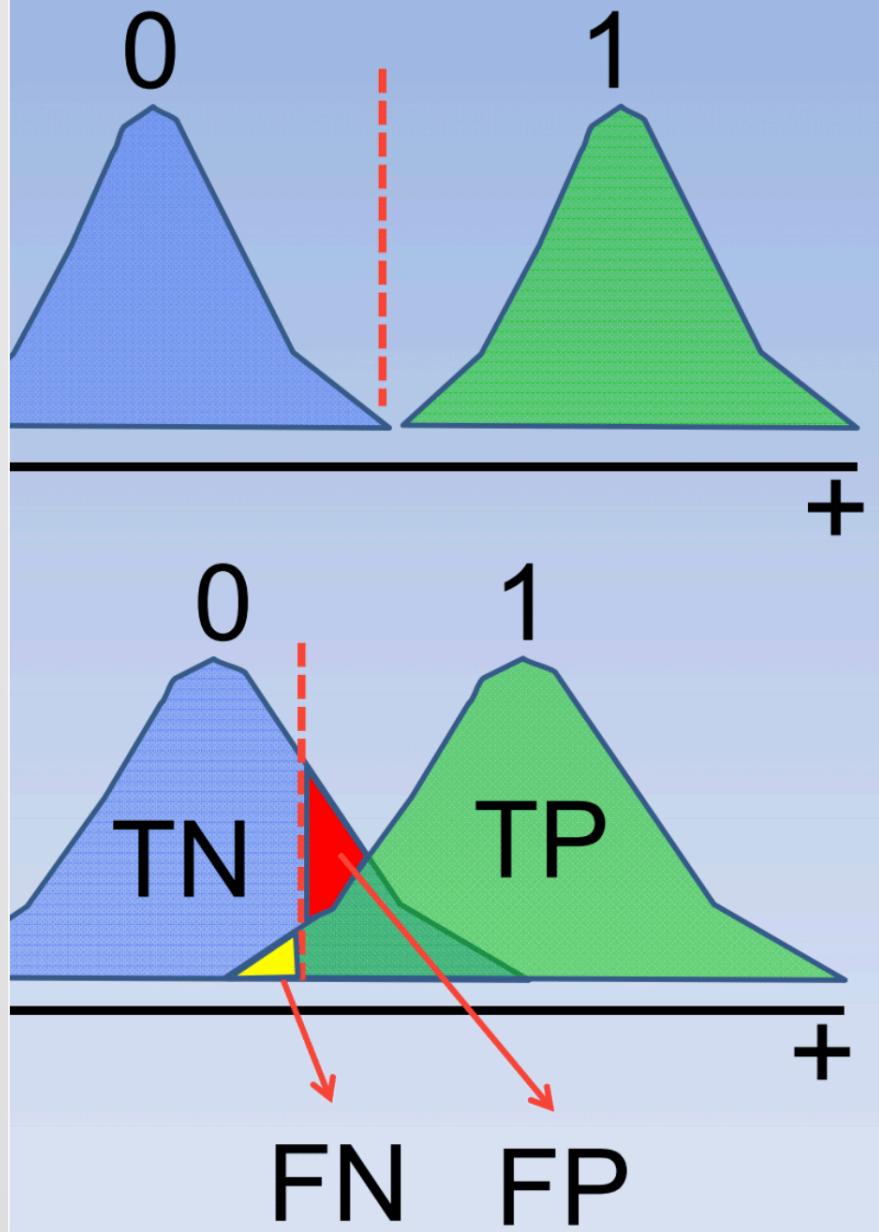


Discrimination capacity in species distribution models depends on the representativeness of the environmental domain

Alberto Jiménez-Valverde^{1*}, Pelayo Acevedo^{1,2}, A. Márcia Barbosa³
Jorge M. Lobo⁵ and Raimundo Real¹

Calibration





predicho

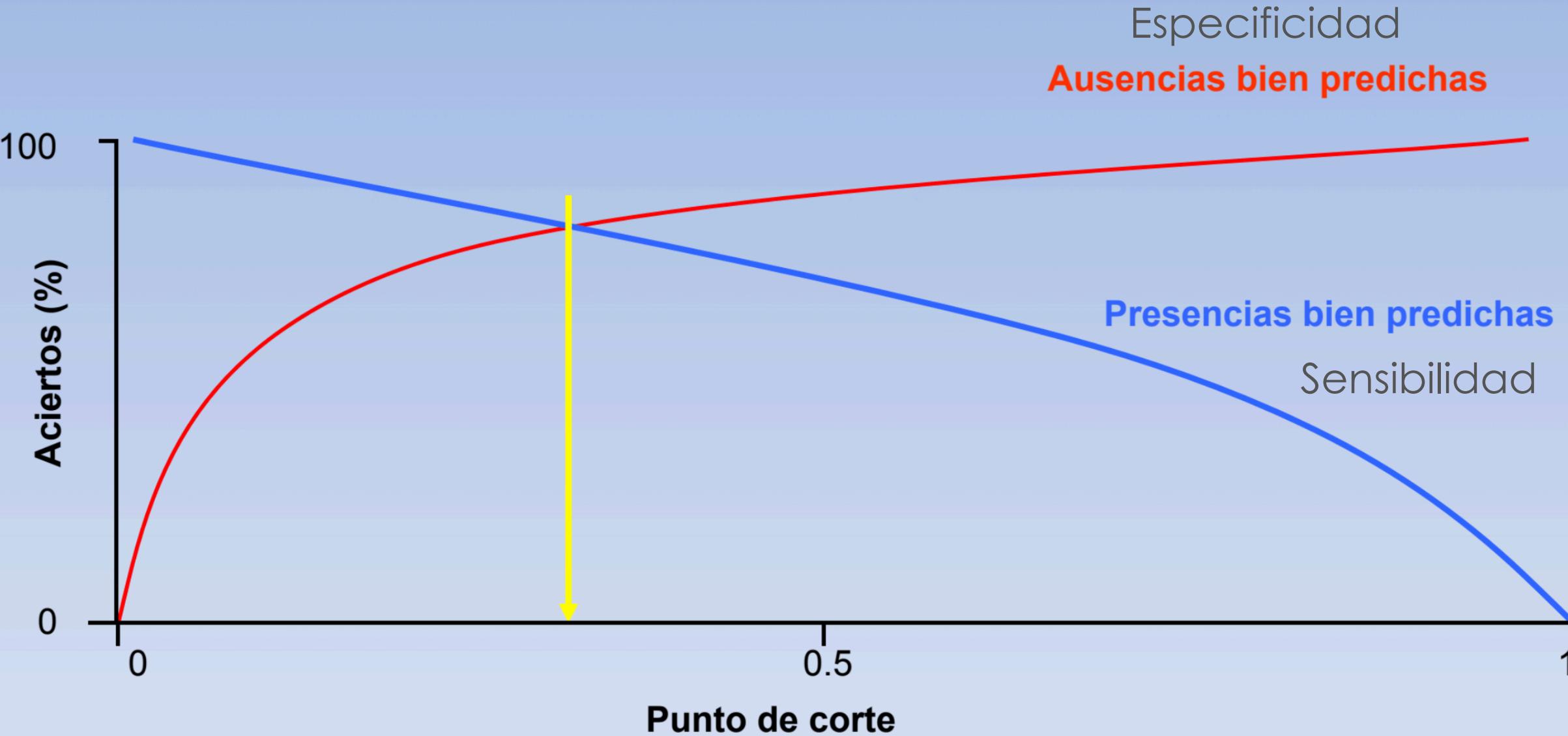
1
0

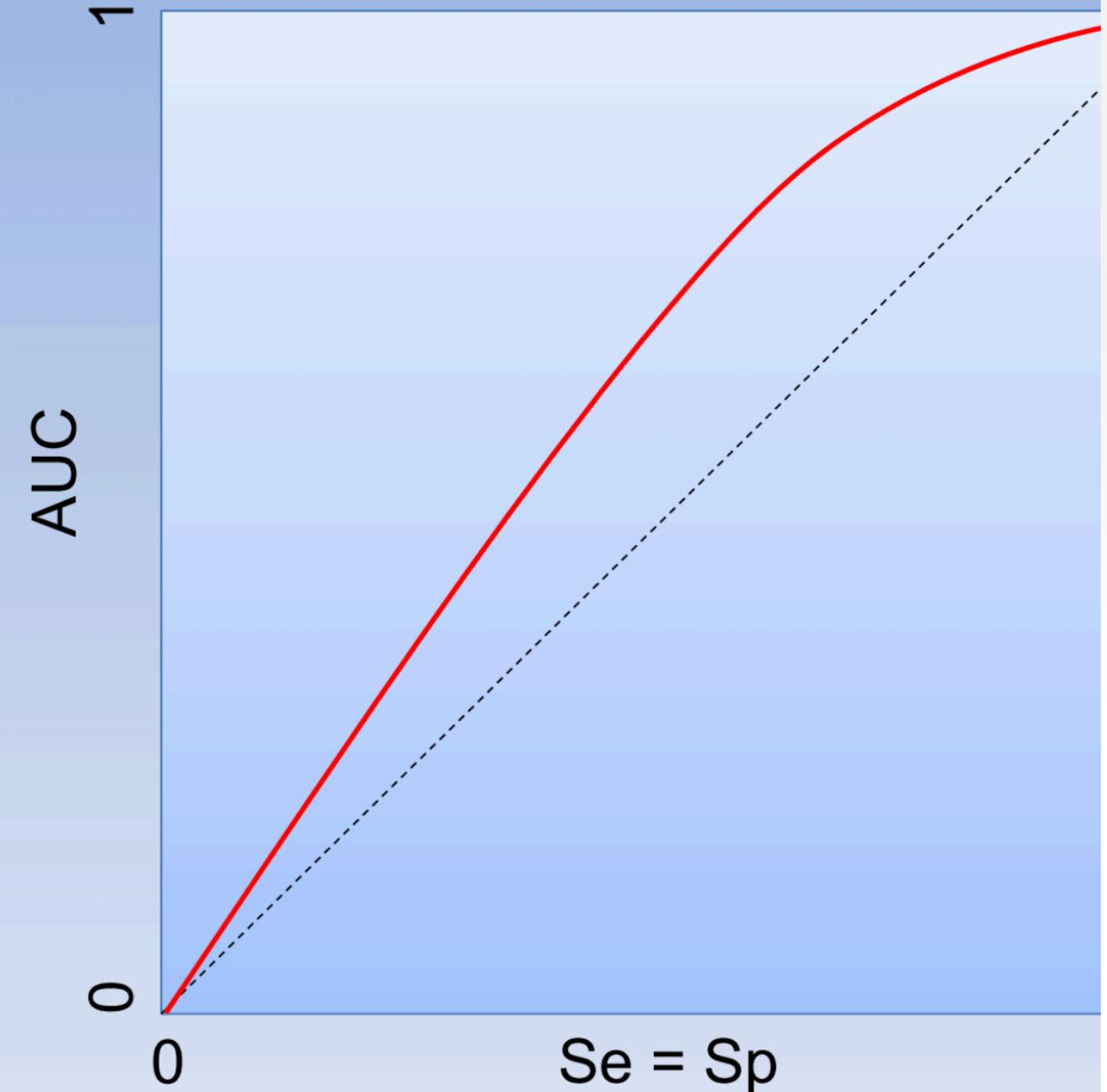
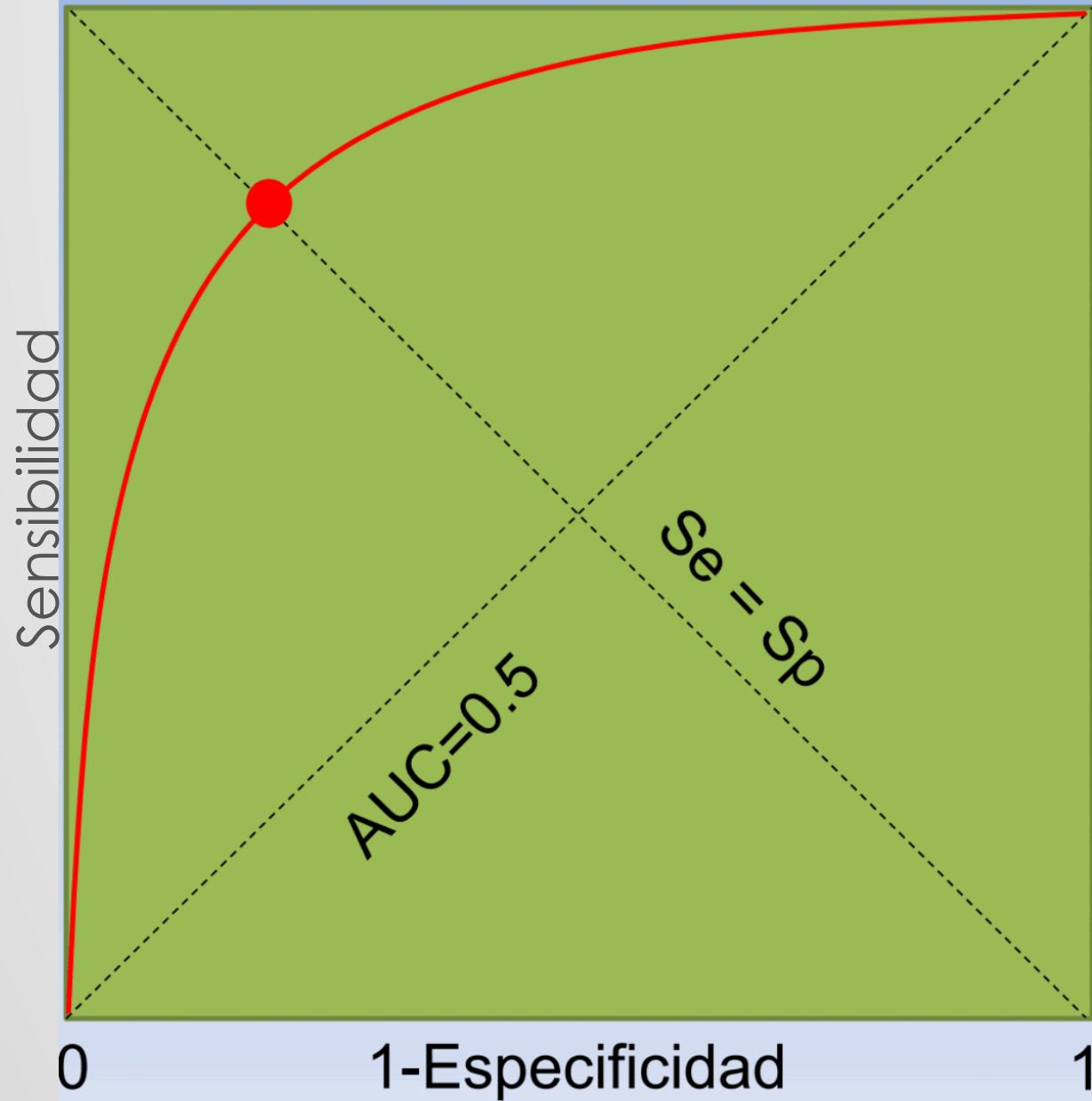
observado
1
0

observado \ predicho	1	0
1	TP	FP
0	FN	TN

Error de
omisión

Error de
comisión





Inverso de especificidad o inverso de la tasa de acierto
en las ausencias o error de comision

A review of methods for the assessment of prediction errors in conservation presence/absence models

ALAN H. FIELDING^{1*} AND JOHN F. BELL²

¹Department of Biological Sciences, the Manchester Metropolitan University, Manchester M1 5GD, UK and ²University Examinations Syndicate, University of Cambridge, Cambridge, UK

Date submitted: 18 December 1996 Date accepted: 1 March 1997

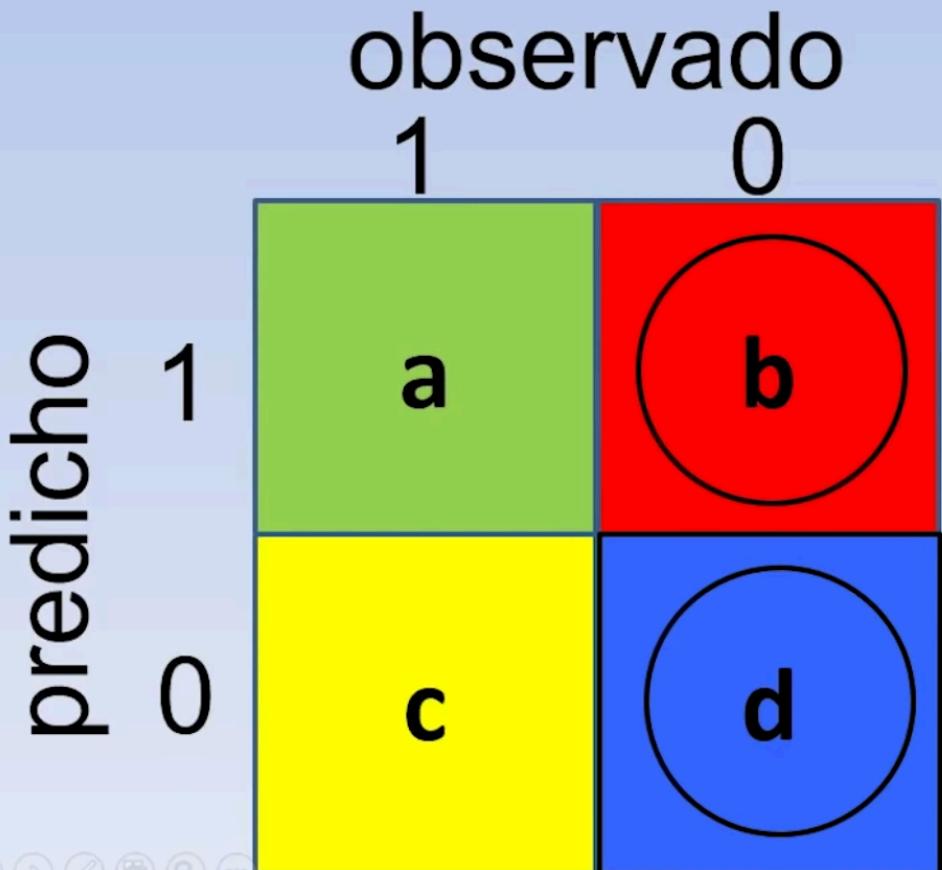


Table 2 Confusion matrix derived measures of classification accuracy.

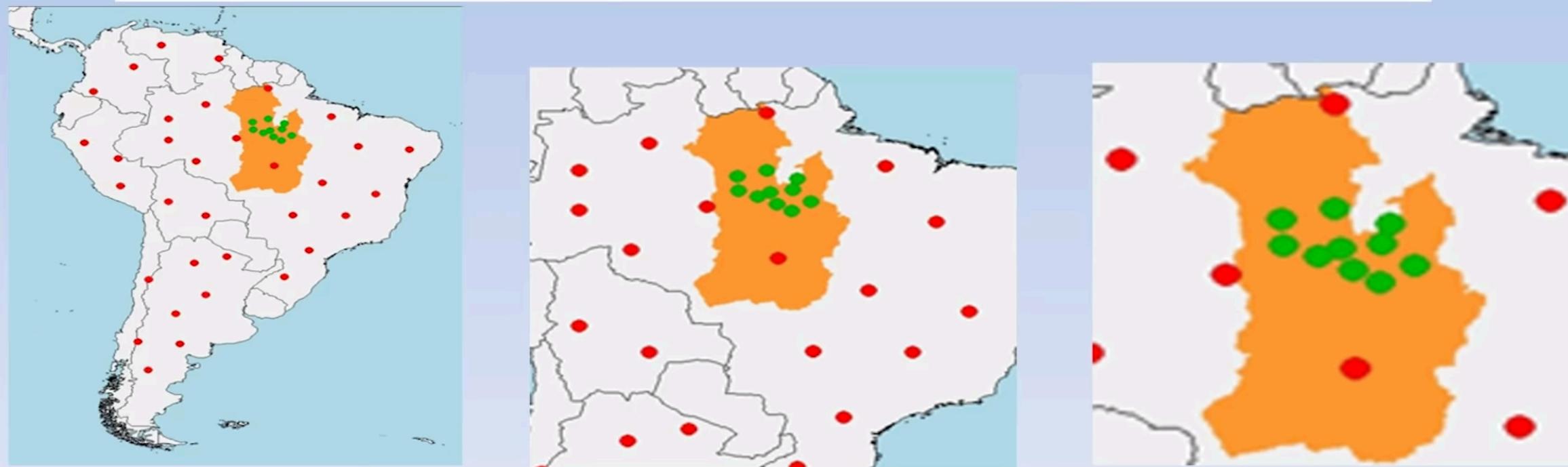
Measure	Calculation
Prevalence	$(a + c)/N$
Overall diagnostic power	$(b + d)/N$
Correct classification rate	$(a + d)/N$
Sensitivity	$a/(a + c)$
Specificity	$d/(b + d)$
False positive rate	$b/(b + d)$
False negative rate	$c/(a + c)$
Positive predictive power (PPP)	$a/(a + b)$
Negative predictive power (NPP)	$d/(c + d)$
Misclassification rate	$(b + c)/N$
Odds-ratio	$(ad)/(cb)$
Kappa	$[(a + d) - (((a + c)(a + b) + (b + d)(c + d))/N)]/[N - (((a + c)(a + b) + (b + d)(c + d))/N)]$
NMI n(s)	$[-a.\ln(a) - b.\ln(b) - c.\ln(c) - d.\ln(d) + (a+b).\ln(a+b) + (c+d).\ln(c+d)]/[N.\ln N - ((a+c).\ln(a+c) + (b+d).\ln(b+d))]$

(Youden index = Se + Sp - 1) (Youden 1950)
True Skill Statistics



AUC: a misleading measure of the performance of predictive distribution models

Jorge M. Lobo^{1*}, Alberto Jiménez-Valverde¹ and Raimundo Real²



- **Error de omisión:** es la proporción de presencias no predichas.

$$O = \frac{c}{a + c}$$

- **Sensibilidad:** mide que tan bueno es el modelo para predecir correctamente una presencia.

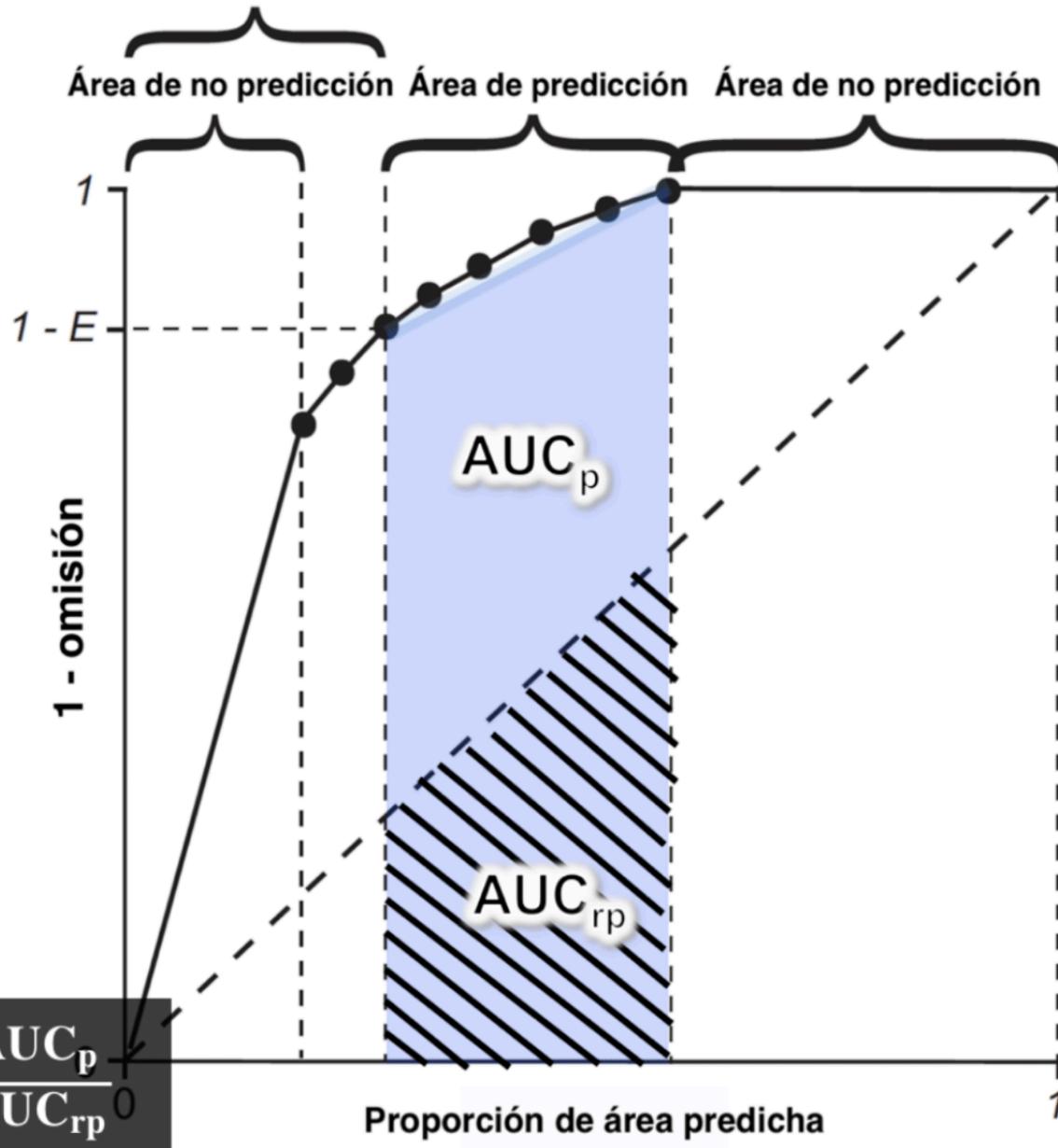
$$S = 1 - \frac{c}{a + c} = 1 - O$$

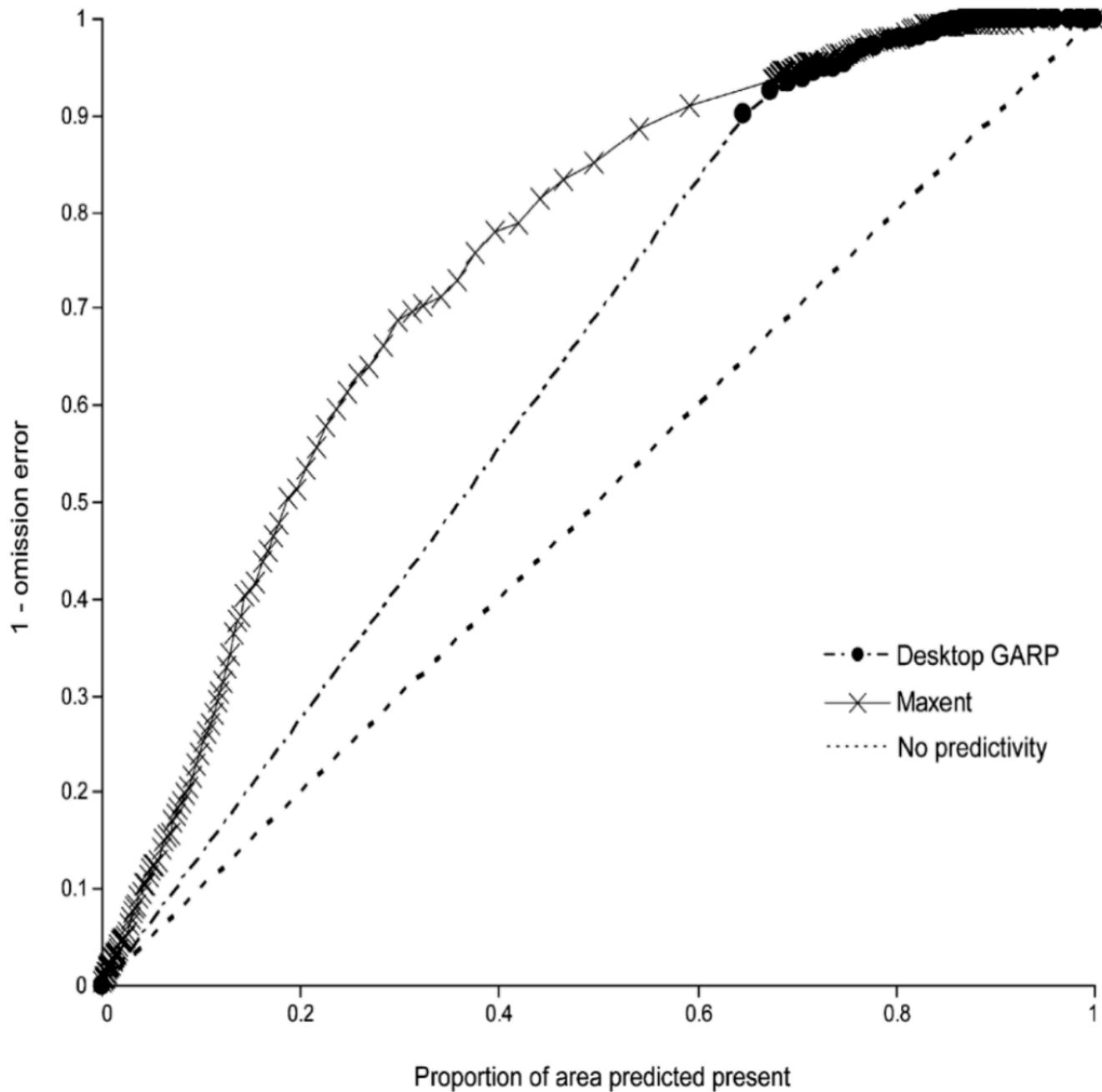
- **Especificidad:** es la proporción de ausencias no predichas.

$$Es = \frac{d}{b + d}$$

- **Error de comisión:** mide que tan bueno es el modelo para predecir correctamente una ausencia

Área de no interés dadas las
características de error de los datos





Algoritmo con bootstrapping

$$T_h = [0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1]$$

- Repetir n veces (por ejemplo, 500) los siguientes pasos para obtener la distribución de **AUC_p** de nuestros datos:
- Sacar una submuestra del 50% de los datos independientes
- Para cada umbral t_h en T_h generar un mapa binario
- Estimar la proporción del área predicha
- Estimar la sensibilidad ($1 - \text{omisión}$)
- Graficar en el eje x , la proporción de área predicha
- Graficar en el eje y , la sensibilidad
- Calcular el **AUC_{ratio}** = $\frac{\text{AUC}_p}{\text{AUC}_{p\text{rand}}}$

La significacia de ROC Parcial

- Se realiza la prueba de conteo directo (frecuencia $AUC_{ratio} > 1$)
- Los datos AUC_{ratio} no siempre están normalmente distribuídos

```
+=====
          Statistics for the difference between
          AUC random and AUC from model prediction
+=====

Ho: The difference between AUC from model prediction and AUC at random is <=0
Ha: The difference between AUC from model prediction and AUC at random is >0
+=====

      n   mean    sd median   mad  min  max range skew kurtosis
1 500 1.73 0.03   1.73 0.02 1.64 1.8  0.16 -0.02     0.13

The p-value for the difference between means (AUC random and AUC partial) is: 0 ***
Reject Ho and Accept Ha: The difference between AUC from model prediction and AUC at random is >0

$p.value
[1] 0
```

ENMGadgets

About

Gadgets for Ecological Niche Modelling

Install

Install the development version using `install_github` within Hadley's `devtools` package.

```
install.packages("devtools")
require(devtools)

install_github("narayanibarve/ENMGadgets")
require(ENMGadgets)
```

Packages `ENMGadgets` depends on

- [raster] (<http://cran.r-project.org/web/packages/raster/>)
- [maptools] (<http://cran.r-project.org/web/packages/maptools/>)
- [sqldf] (<http://cran.r-project.org/web/packages/sqldf/>)
- [fields] (<http://cran.r-project.org/web/packages/fields/>)
- [rgdal] (<http://cran.r-project.org/web/packages/rgdal/>)
- [dismo] (<http://cran.r-project.org/web/packages/dismo/>)

NicheToolBox de CONABIO

Ir a la pestaña SDM performance -> Partial ROC

The screenshot shows the NicheToolBox application interface. At the top, there is a navigation bar with tabs: NicheToolBox, Settings & Materials, Data, Niche, Bivariate correlations, Principal Component Analysis, ENM, SDM performance (which is currently selected), and Workflow. Below the navigation bar, there is a search bar labeled "Search: []". On the left side, there is a sidebar titled "NicheToolBox" containing sections for "Model prediction data" (with a file upload input) and "Presence data" (with a file upload input). The main content area has a heading "SDM performance" and a sub-section "Partial ROC" which is highlighted with a blue background. Other options in this dropdown menu include "Binary Map" and "Confusion matrix". There are also sections for "Raster" (with "Upload prediction raster" and "Presences" inputs) and "Distribution of AUC ratios". At the bottom of the page, there is a dark footer bar with the URL <http://shiny.conabio.gob.mx:3838/nichetoolb2/>.

NicheToolBox

Datos de predicción del modelo.

Suba el ráster de predicción del modelo en formato ASCII

Formato

.asc

Archivo de datos

Seleccionar archivo Lutjanus_i...mis_10.asc

Carga completa

Datos de presencia

Sube tus datos de presencia. Los datos deben estar en el siguiente formato:

Especies	largo	latitud
sp_name	-65.38	-10.38

Formato

CSV

Archivo de datos

Seleccionar archivo CSVNiche2

Carga completa

Parámetros ROC parciales

Proporción de omisión:

0.05

Porcentaje de puntos aleatorios%

Show 5 entradas

Buscar:

IteraciónNo	AUC_at_Value_0.05	AUC_at_0.5	AUC_ratio
1	0.816048223027906	0.499999994307074	1.63209646463862
2	0.814947806211201	0.499990418066359	1.6299268481242
3	0.804456913848657	0.499997730546549	1.60892113044054
4	0.840057322051344	0.499999977228298	1.68011472062083
5	0.808475473118834	0.499999977228298	1.61695101987912

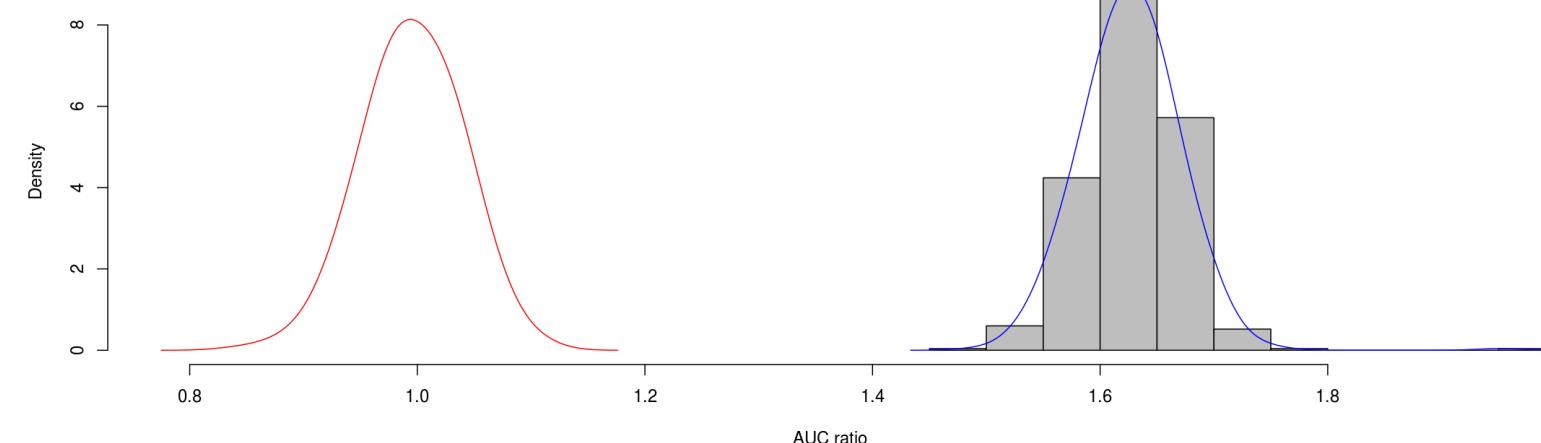
IterationNo AUC_at_Value_0.05 AUC_at_0.5 AUC_ratio

Mostrando de 1 a 5 de 500 entradas

Anterior 1 2 3 4 5 ... 100 Siguiente

Distribución de ratios de AUC

Partial AUC distribution



Estadística descriptiva para el ratio AUC

```
+++++  
Estadísticas de Partial Roc  
después de 500 simulaciones  
+++++
```

El valor medio para la relación AUC a 0.05 es: 1.626414

El valor medio para el AUC parcial en 0.05 es: 0.813201

El valor medio para el AUC parcial a 0.5 es: 0.4999963

```
+++++  
Estadísticas para la diferencia entre  
AUC aleatorias y AUC de predicción del modelo.  
+++++
```

Ho: La diferencia entre el AUC de la predicción del modelo y el AUC al azar es <= 0

Ha: La diferencia entre el AUC de la predicción del modelo y el AUC al azar es > 0

```
+++++  
n media sd mediana loca min max rango sesgo kurtosis  
1 500 1.63 0.04 1.62 0.04 1.49 1.95 0.46 0.78 5.91  
+++++
```

El valor de p para la diferencia entre medias (AUC aleatorio y AUC parcial) es: 0 ***

Rechazar Ho y Aceptar Ha: la diferencia entre el AUC de la predicción del modelo y el AUC al azar es > 0

```
$ p.valor  
[1] 0
```

