



Grupo de investigación Ecología de Zonas Áridas

**CENTRO ANDALUZ PARA LA EVALUACIÓN Y  
SEGUIMIENTO DEL CAMBIO GLOBAL**



Modelling  
Workshops



# Evaluación y validación del modelo

**Elisa Liras**  
Dpto. Biología Vegetal y Ecología  
Universidad de Almería  
eliras@ual.es

### **Como simplificación de la realidad, ya empezamos modelizando con varios problemas:**

**La mayoría de los datos de ocurrencia disponibles de las especies son asimétricos (es decir sólo presencia), por lo que no tenemos información certera de dónde NO está la especie.**

**El esfuerzo del muestreo a través los rangos distribucionales de la mayoría de la especies es desigual, así los datos de ocurrencia están sesgados en los espacios ecológico y geográficos.**

**Las variables ambientales abarcan relativamente pocas dimensiones del nicho, y no sabemos qué variables son relevantes para cada especie**

### Además:

**Algunos de los algoritmos no manejan variables categóricas (ej. tipos de suelo; BIOCLIM, DOMAIN).**

**Varios algoritmos estocásticos presentan diversas soluciones a un problema, incluso bajo parámetros y datos de entrada idénticos (ej. GARP).**

**No conocemos la distribución real de la especie, así que no sabemos cuándo los modelos están incurriendo en equivocaciones y cuándo están llenando vacíos de información.**

*Con todo esto, podemos tener 2 tipos de errores:*

**1. OMISIÓN**

**2. COMISIÓN**

### Tipos de errores posibles:

**ERROR DE OMISIÓN:** predice la no-presencia de la especie donde realmente está (*falso negativo*)

→ *es mucho más importante, y peor: no predice lugares de presencia que pueden ser de importancia crucial para la supervivencia de la población, o contener genotipos únicos*

**ERROR DE COMISIÓN:** predice la presencia de la especie donde no está (*falso positivo*)

→ *puede ser real o aparente, ya que un “falso positivo” puede significar o una sobre-predicción del modelo o una predicción de nicho potencial de la especie*

### **Consideraciones sobre el error de OMISIÓN:**

Generalmente el error de la omisión se puede considerar un error importante.

Sin embargo, bajo algunas circunstancias, el registro de la presencia de la especie puede ser equivocado:

1. La identificación del taxón es equivocada
2. La georreferenciación de la localidad es equivocada
3. El registro representa individuos fuera de su nicho ecológico; ej., poblaciones sumidero; individuos transitorios o vagabundos, etc.

### **Consideraciones sobre el error de COMISIÓN:**

La sobrepredicción en un modelo puede o no ser un error, por ello la comisión no es un error tan importante.

La predicción en áreas donde la especie no ha sido registrada, es causado por diferentes factores:

1. El área es habitable para la especie, pero no se ha hecho ningún esfuerzo de muestreo. La especie puede estar allí.
2. El área es habitable para la especie, pero históricamente (barreras, capacidad de dispersión) o factores bióticos (competencia, depredación) han impedido que la especie la ocupe, o que se haya extinguido.
3. El área es inhabitable: verdadero error de comisión

### **Evaluación de los errores:**

Casi todas las técnicas de evaluación de modelos se centran en detectar los errores de omisión (falsos negativos).

Técnicas “jackknife”: estimar el nicho de la especie a partir de todos los puntos de aparición menos 1, tantas veces como puntos de presencia tengamos.

### Metodología para la evaluación de los errores:

Cuantificar el componente de errores con una matriz de confusión

	Presencia real	Ausencia real
Presencia predicha	<b>a</b>	<b>b</b>
Ausencia predicha	<b>c</b>	<b>d</b>

**a & d = Predicción correcta**

**b = error comisión**  
(falso positivo, sobrepredicción)

**c = error omisión**  
(falso negativo)

$[a/(a+c)]$ : SENSIBILIDAD: proporción de positivos reales

$[b/(b+d)]$ : 1- ESPECIFICIDAD: proporción del área predicha presente (error comisión)

### Metodología para la evaluación de los errores: Prueba de $\chi^2$

Mide si los valores independientes cayeron mayormente dentro de la predicción según lo esperado al azar, con base en el tamaño del área predicha y el número de puntos de verificación



Presentación original: Martínez-Meyer

### Metodología para la evaluación de los errores: Índice de Kappa de Cohen

	Presencia real	Ausencia real
Presencia predicha	<i>a</i>	<i>b</i>
Ausencia predicha	<i>c</i>	<i>d</i>

#### Kappa (ver Landis & Koch 1977 *Biometrics*):

- 0 – 0.4: pobre
- 0.4 – 0.75: buena
- 0.75 – 1.0: excelente

**Cálculo del índice de Kappa de Cohen:**

$$k = \frac{[(a + d) - (((a + c)(a + b) + (b + d)(c + d)) / n)]}{[n - (((a + c)(a + b) + (b + d)(c + d)) / n)]}$$

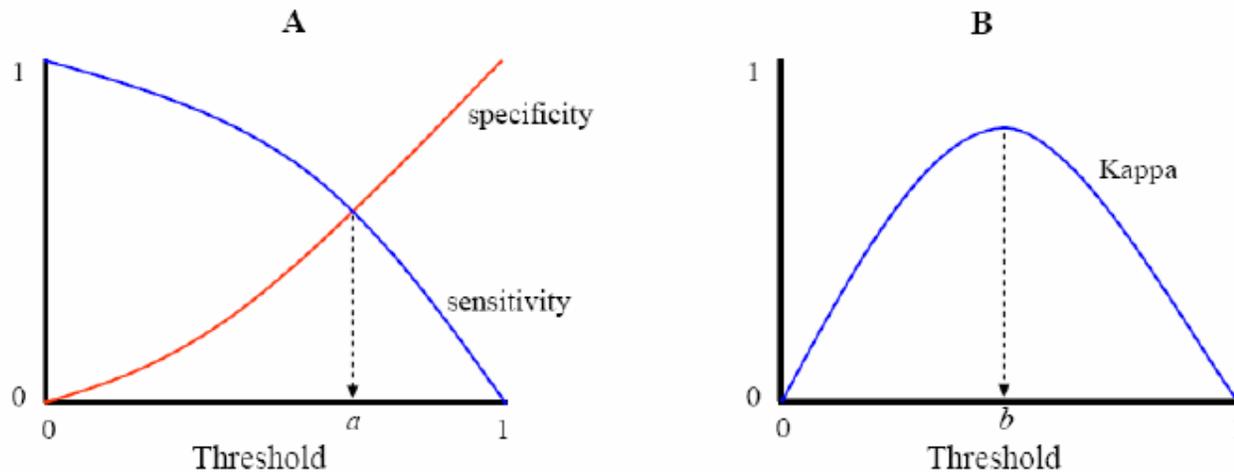
### Metodología para la evaluación de los errores: Curvas ROC (1)

**Sensibilidad: valores de presencia correctamente predichos**

**Especificidad: valores de ausencia correctamente predichos**

Hay muchos métodos para determinar el límite a partir del cual el modelo predice “presencia”, calculando estadísticas a partir de la matriz de confusión y un rango de posibles umbrales.

Por ejemplo, la sensibilidad y la especificidad pueden calcularse en límites que vayan incrementando su magnitud en 0,01, desde 0 hasta 1 (ej. 0, 0.01, 0.02, 0.03...0.99, 1).



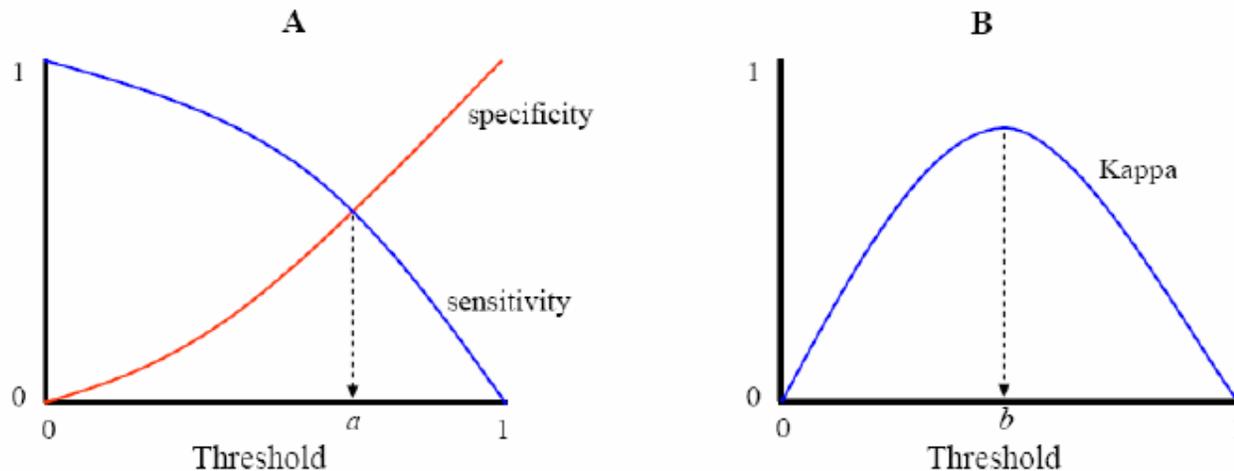
### Metodología para la evaluación de los errores: Curvas ROC (2)

Conforme aumenta el umbral, la proporción del área de estudio predicha como “favorable” para la especie, o en el que la especie estará potencialmente presente, disminuye.

Consecuentemente, el número de presencias predichas correctamente disminuye (disminuye la sensibilidad), y la proporción de ausencias observadas predichas correctamente aumenta (aumenta la especificidad).

A partir de estos datos, se puede seleccionar el límite en el cual la sensibilidad y la especificidad son iguales, o en el que su suma se maximiza.

Del mismo modo, es común calcular el coeficiente Kappa a lo largo de un gradiente de posibles umbrales y seleccionar el umbral en el que el estadístico se maximiza.



### **Metodología para la evaluación de los errores: Curvas ROC (3)**

#### **Valoración independiente del umbral:**

Cuando la salida del modelo es continua, la valoración de la predicción utilizando la estadística derivada de la matriz de confusión será sensible al método utilizado para elegir el umbral para crear una predicción binaria (favorable / no favorable).

Además, si las predicciones son binarias, la evaluación del modelo no tiene en cuenta toda la información que da el modelo (Fielding and Bell, 1997).

Por ello, normalmente es útil derivar un test estadístico que de una sola medida de evaluación del poder predictivo a través de todo el rango de posibles umbrales.

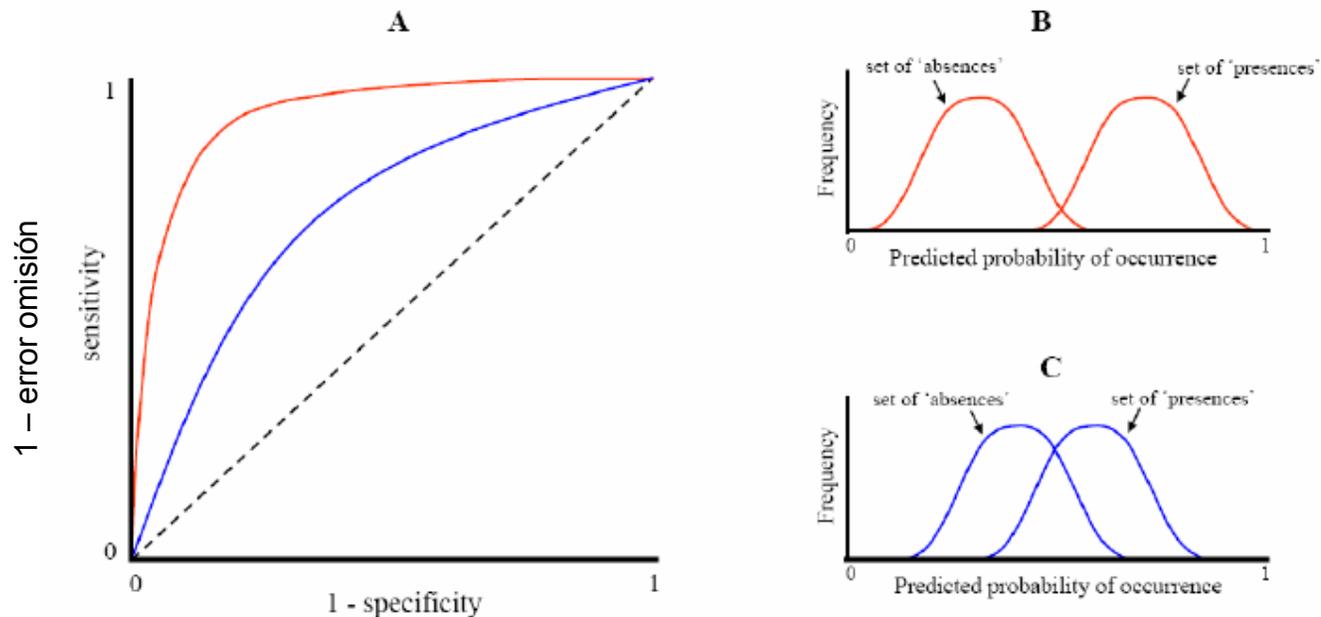
Esto se consigue con un estadístico llamado AUC: the Area Under the Receiver Operating Characteristic Curve.

### Metodología para la evaluación de los errores: Curvas ROC (4)

El AUC se deriva de la curva ROC. La curva ROC se obtiene al enfrentar la sensibilidad frente a “1-Especificidad” a los largo de todos los posibles umbrales.

Se utilizan la sensibilidad y la especificidad porque estas dos medidas tienen en cuenta los cuatro elementos de la matriz de confusión (presencias y ausencias reales y predichas).

Por conveniencia se calcula “1-Especificidad” para que la sensibilidad y la especificidad varíen en la misma dirección cuando se ajusta el umbral (Pearce and Ferrier, 2000).

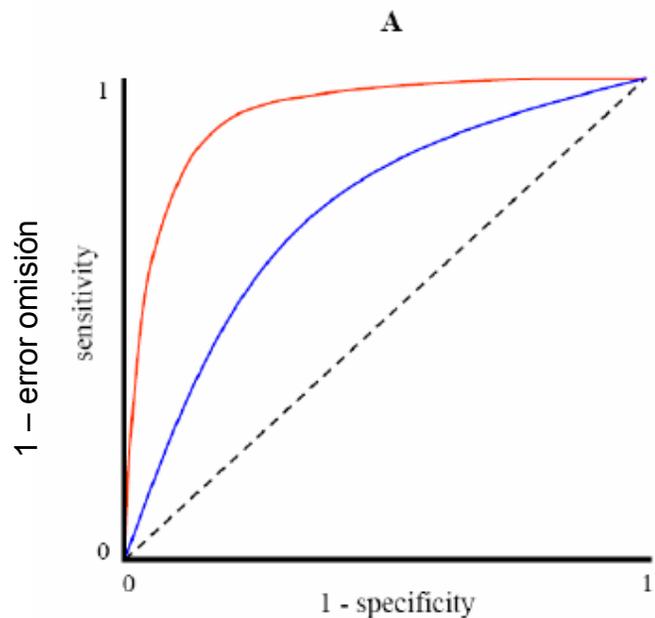


Proporción del área predicha presente

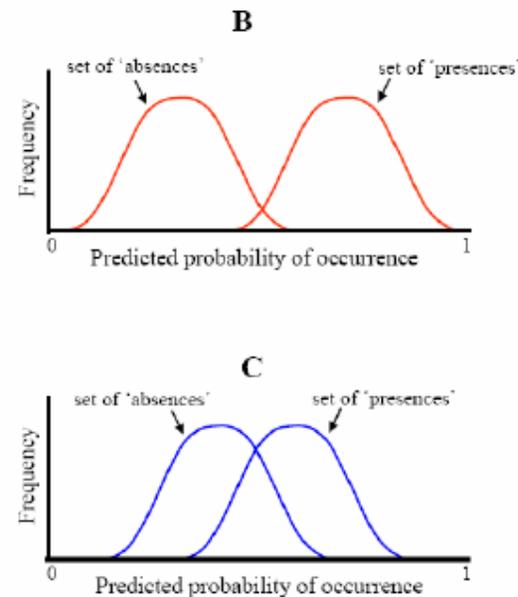
### Metodología para la evaluación de los errores: Curvas ROC (5)

La curva ROC describe la relación entre la proporción de presencia correctamente predichas (sensibilidad) y la proporción de ausencia incorrectamente predichas (1-Especificidad).

Entonces, un modelo que prediga perfectamente generará una curva ROC que siga el eje izquierdo hasta lo alto del gráfico, mientras que un modelo que prediga “no mejor que al azar” generará una curva ROC que siga la línea 1:1 (diagonal).



Proporción del área predicha presente



### **El uso de las curvas ROC: comparación de modelos**

Elith, 2006 compara diferentes algoritmos de nicho ecológico en base a los resultados de su curva ROC (y AUC)

ECOGRAPHY 29: 129–151, 2006

### **Novel methods improve prediction of species' distributions from occurrence data**

**Jane Elith\*, Catherine H. Graham\*, Robert P. Anderson, Miroslav Dudík, Simon Ferrier, Antoine Guisan, Robert J. Hijmans, Falk Huettmann, John R. Leathwick, Anthony Lehmann, Jin Li, Lucia G. Lohmann, Bette A. Loiselle, Glenn Manion, Craig Moritz, Miguel Nakamura, Yoshinori Nakazawa, Jacob McC. Overton, A. Townsend Peterson, Steven J. Phillips, Karen Richardson, Ricardo Scachetti-Pereira, Robert E. Schapire, Jorge Soberón, Stephen Williams, Mary S. Wisz and Niklaus E. Zimmermann**

Elith, J., Graham, C. H., Anderson, R. P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R. J., Huettmann, F., Leathwick, J. R., Lehmann, A., Li, J., Lohmann, L. G., Loiselle, B. A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J. McC., Peterson, A. T., Phillips, S. J., Richardson, K. S., Scachetti-Pereira, R., Schapire, R. E., Soberón, J., Williams, S., Wisz, M. S. and Zimmermann, N. E. 2006. Novel methods improve prediction of species' distributions from occurrence data. – *Ecography* 29: 129–151.

### ***El uso de las curvas ROC...y la discusión científica***

Pero...a pesar de que la curva ROC se ha utilizado para evaluar y comparar modelos entre sí, actualmente hay una discusión científica sobre su validez...y se proponen modificaciones

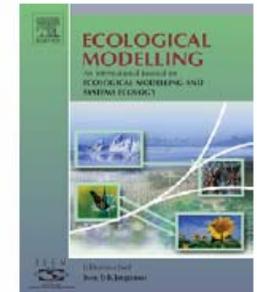
ECOLOGICAL MODELLING 213 (2008) 63–72



available at [www.sciencedirect.com](http://www.sciencedirect.com)



journal homepage: [www.elsevier.com/locate/ecolmodel](http://www.elsevier.com/locate/ecolmodel)

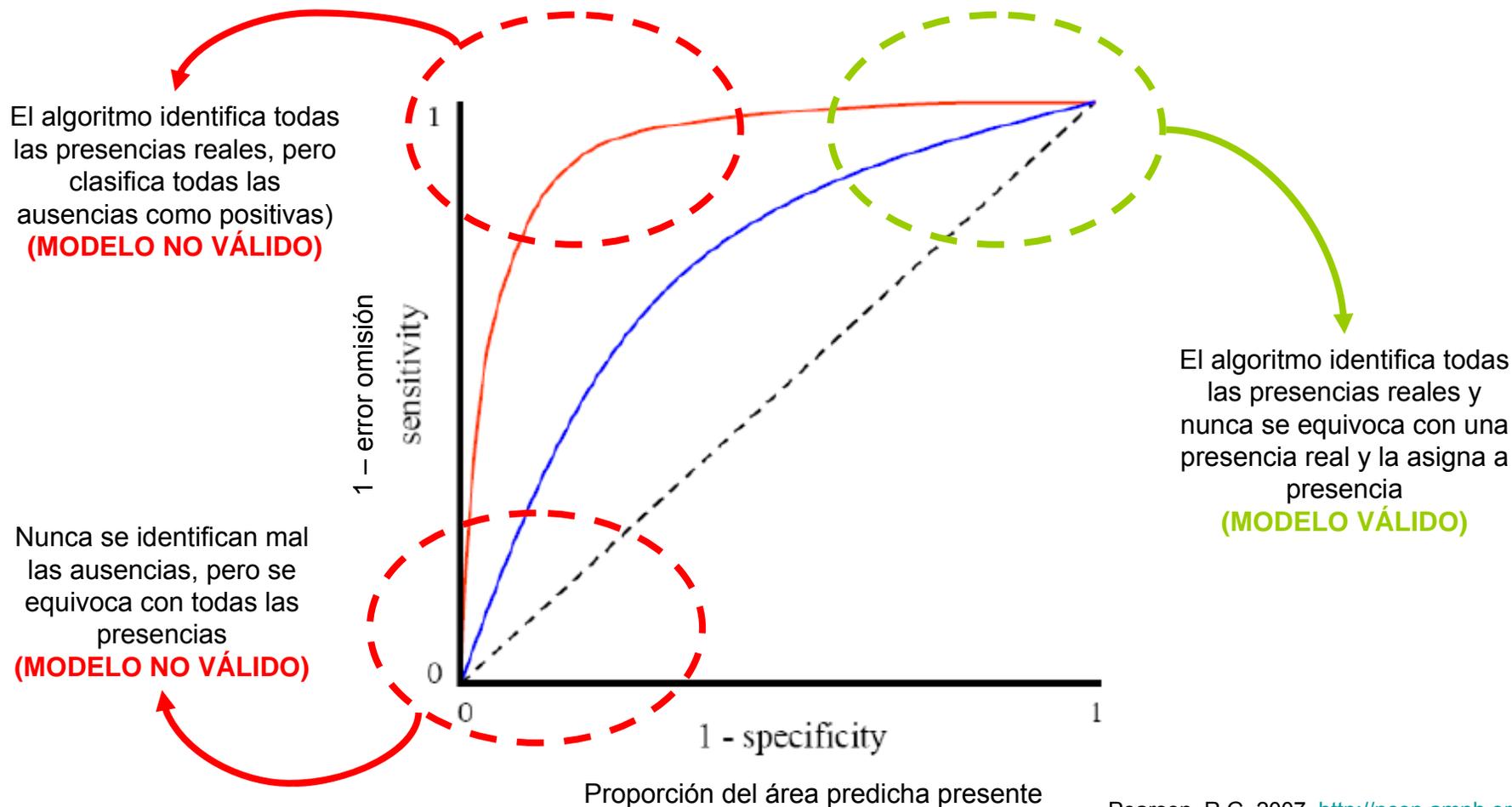


## **Rethinking receiver operating characteristic analysis applications in ecological niche modeling**

A. Townsend Peterson\*, Monica Papeş, Jorge Soberón

Natural History Museum and Biodiversity Research Center, The University of Kansas, Lawrence, KS 66045 USA

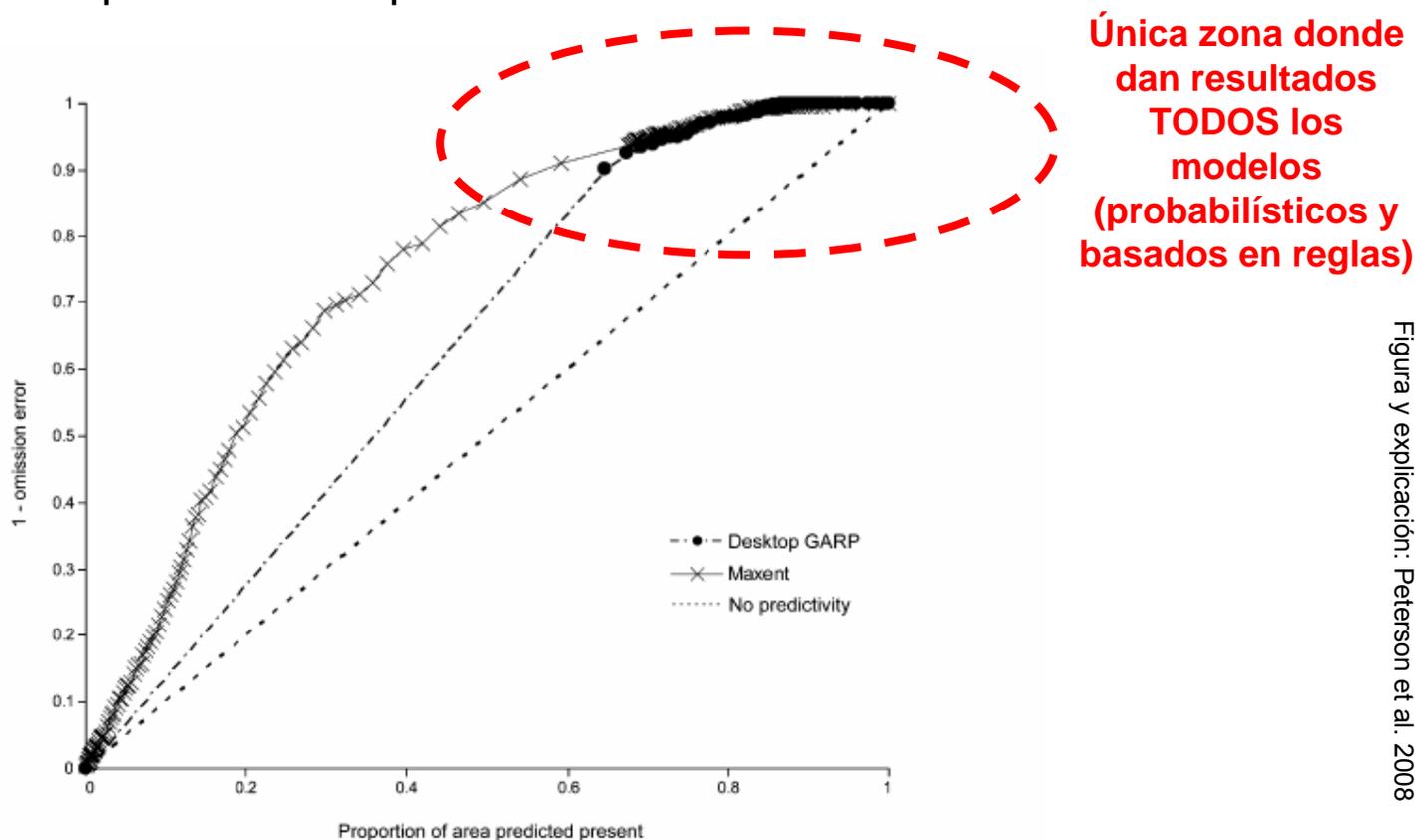
### Entonces...¿qué información da realmente la curva ROC?



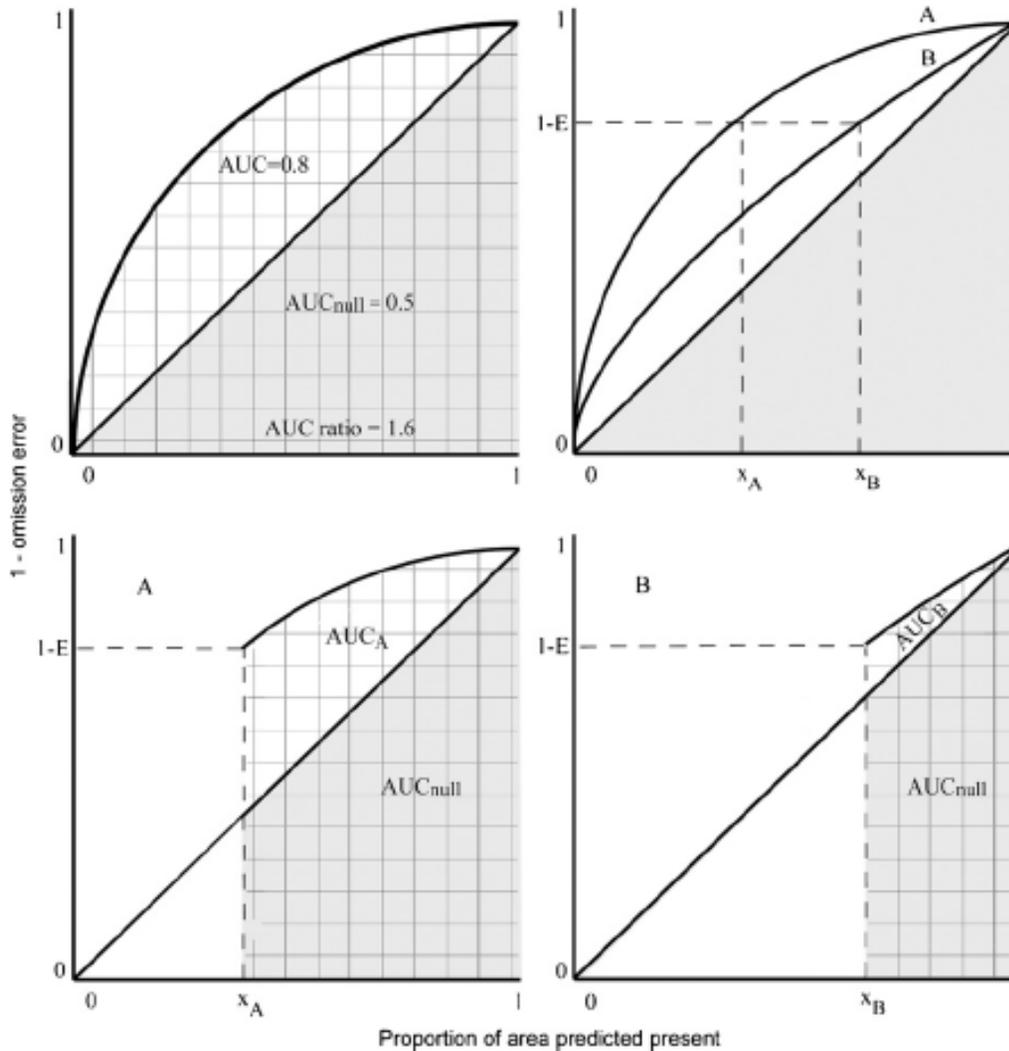
Pearson, R.G. 2007. <http://ncep.amnh.org>.  
Peterson et al. 2008

### Entonces...¿Por qué no vale la curva ROC para comparar distintos modelos?

- Los algoritmos basados en reglas sólo dan una parte de la curva ROC
- Los algoritmos probabilísticos dan toda la curva ROC
- No es válido comparar la AUC para diferentes modelos



### Solución propuesta: curva ROC modificada



Que el usuario decida:

1. Peso que le da a cada tipo de error
2. Valor umbral

# **Validación del modelo: Coincidencia del resultado con la realidad**

## 7. Evaluación y validación del modelo

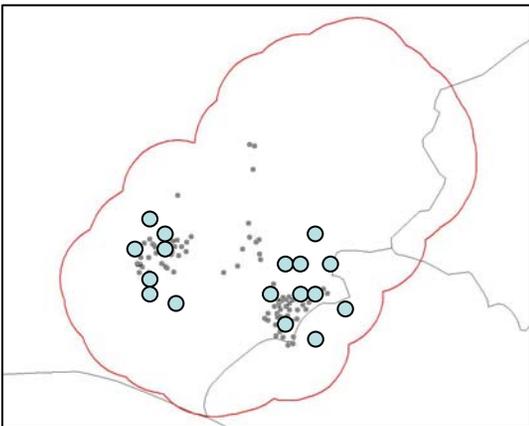
### Validación:

El modelo hay que validarlo, contrastando el resultado con la realidad.

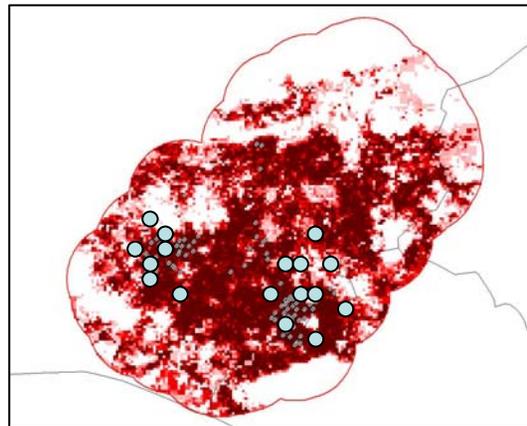
### Ejemplo:

1. Realizar el modelo con el 70% de los puntos de aparición de la especie
2. Obtener el resultado
3. Comprobar si las predicciones de alta probabilidad coinciden con el 30% de presencias restante.

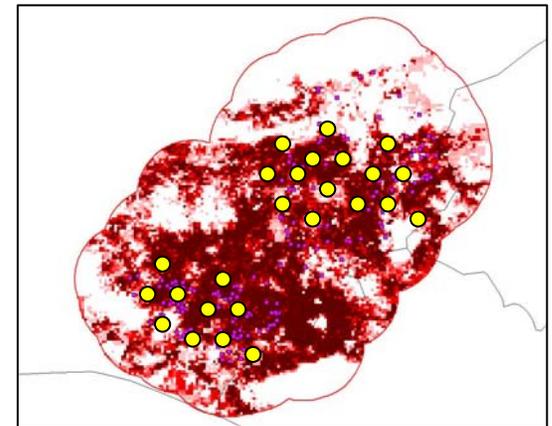
Puntos de entrenamiento



Predicción del modelo



Puntos de validación



### **Validación:**

- Un buen modelo nos estimará alta probabilidad de aparición de la especie en los lugares que coincidan con dicho 30%.
- Altamente recomendable la validación del modelo en campo.
- En casos de gestión para reintroducción de especies: realización de introducciones de la especie en los lugares considerados como aptos pero no ocupados por la misma.

**Muchas gracias**