

Réplica del III taller sobre modelización de nichos ecológicos de GBIF

Madrid, España. 14-16 marzo 2007



Hilo conductor de los diferentes métodos que veremos.

- Describiremos una serie de modelos diferentes y aplicaremos cada uno de ellos a un único caso práctico real.

¡¡Esto no es muy sensato!!

**Es bueno probar varios modelos, pero
hay que hacerlo con sentido común**

**Lo haremos para aprender, sólo
estamos jugando ...**

Introducción a los métodos de modelización

Hilo conductor de los diferentes métodos que veremos.

Trabajaremos con *Salix hastata* subsp. *sierrae-nevadae*

• **CR** SALICACEAE
Salix hastata subsp. *sierrae-nevadae* Rchb. f.



Sauce de Sierra Nevada

Taxón geográficamente aislado y relictivo, acantonado en lugares inaccesibles a los herbívoros, que constituyen la principal amenaza. Existen dos poblaciones con bajo número de individuos. Está protegido por la legislación autonómica desde 1994.

M. Ruiz



Hilo conductor de los diferentes métodos que veremos.

- Trabajaremos con *Salix hastata* subsp. *sierrae-nevadae*
 - Nuestro objetivo es buscar sitios para llevar a cabo reintroducciones que mejoren el estado de conservación del taxón
 - El principal problema que tenemos es que hay sólo dos poblaciones identificadas: **la distribución actual no es representativa de su nicho fundamental**
 - La principal amenaza es el sobre pastoreo: **variables humanas no contempladas en los modelos ...**



Introducción a los métodos de modelización

Hilo conductor de los diferentes métodos que veremos.

- Variables seleccionada para modelizar el nicho de *S. hastata*
 - **Altitud:** las poblaciones están por encima de los 2300 metros
 - **Índice de humedad (CTI):** necesita vivir en suelos permanentemente húmedos.
 - **Precipitación:** También requiere humedad “macroclimática”
 - **Insolación:** Suponemos que “preferirá” vivir en lugares umbríos
 - **Zona de estudio:** Hemos “recortado” Sierra Nevada para que los modelos se ejecuten más rápido. Sólo usaremos una porción del macizo.



Modelos “probabilísticos”

Objetivos

- Mostrar de forma muy general las bases teóricas de los modelos probabilísticos (sobre todo para enfrentarlos con los algorítmicos y comprenderlos mejor)
- Analizar con la regresión logística como ejemplo de modelo probabilístico



Generalidades sobre modelización: Culturas de modelización

Cultura “probabilística”



- Hay un modelo de datos probabilístico dentro de la caja negra. Esto implica que la caja no es tan negra ...
- Como resultado se obtiene una función probabilística que muestra la probabilidad de ocurrencia de la especie en función de las variables independientes.
- La validación se realiza evaluando si los resultados satisfacen las condiciones de aleatoriedad de partida. Surge el concepto de bondad de ajuste del modelo

Características generales de la regresión logística multivariante

- Es un método estadístico que permite establecer una relación entre una variable dependiente y una o varias variables independientes. En nuestro caso, entre la presencia/ausencia de una especie y una serie de variables ambientales predictoras.
- El modelo matemático toma la siguiente forma:

$$\text{logit}(p_i) = \ln \left(\frac{p_i}{1 - p_i} \right) = \alpha + \beta_1 x_{1,i} + \dots + \beta_k x_{k,i},$$



- El resultado de valores continuos entre 0 y 1
- Las variables dependientes pueden ser categóricas

■ Características generales de la regresión logística multivariante

■ Ventajas de la regresión logística

- Fácil de implementar y de interpretar
- Acepta datos categóricos
- Al obtenerse una función, podemos analizar la importancia de cada variable
- Ha sido profusamente estudiado y hay muchas herramientas en las que está implementada
- Es posible evaluar la bondad del modelo, ya que se trata de un ajuste estadístico

■ Desventajas de la regresión logística

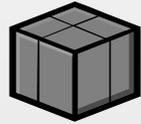
- No funciona bien cuando hay pocos puntos de ocurrencia de la especie
- A veces es difícil determinar la aproximación no lineal que explica la distribución.
- Es muy sensible a la relación entre ausencias y presencias de la especie.

Modelos “probabilísticos”

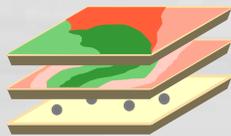
Caso práctico de regresión logística

$$Y = 1 - 0*[Solar_ra] + 0*[Dtm] - 5.82077e-011*[Cti] + 2e-12*[Rain]$$

“No le desagradan” los suelos poco saturados y los lugares donde llueve mucho



Regresión logística
(grid and theme regression
www.jennessent.com)



Datos de partida

