Presence probability and environmental favourability

Presence probability

Combines two things:

- the general probability of finding the species anywhere by chance in the study region (which depends on the species' **prevalence** there)

- the **environmental conditions** and how they make presence more likely / less likely than random expectation

If the environment were irrelevant, probability would be the same in all localities; when it is not, probability is higher in some places and (to compensate) lower in others

In any case, the sum of all probabilities equals the sum of presences, i.e. mean presence probability equals species prevalence

Presence probability

Ultra-simplified example:

There are 10 pixels recorded as presence for the lynx, and 20 pixels where the environment is perfect for this species. Presence probability can't be 100% in all those 20 pixels simultaneously, because there are only 10 presences.

At most, presence probability can be 50% in each of those 20 pixels, and this would be if all other pixels in the study region had 0% presence probability (which is unrealistic).

If you model again when more lynxes have been detected in other pixels, probability will be higher in those 20 pixels (because there are more presences), even if nothing changed environmentally.

Probability and favourability

Favourability (like suitability) can be 100% in all perfect pixels, regardless the number of presences.

Favourability (**un**like suitability) is commensurable and directly comparable across species and regions with different prevalence (proportion of presences).


Naturwissenschaften
DOI 10.1007/s00114-012-0926-0

CONCEPTS & SYNTHESIS

Favourability: concept, distinctive characteristics and potential usefulness

Pelayo Acevedo · Raimundo Real

Presence probability

Presence-absence data:

Logistic regression (GLM)

- binomial distribution

Presence probability

Presence-absence data:

Logistic regression (GLM)

- binomial distribution

Logistic regression follows the expression:

$$P = \frac{e^{y}}{1+e^{y}}$$

basis of the natural logarithm

probability of the presence of the species

linear regression:

$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n$

Presence probability

Linear predictor

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n$$

β_a $\ln(n_1/n_0)$ + β_b estimate

presences absences

$$P = \frac{e^y}{1+e^y}$$

So...

Probability values will depend on the category with the highest number of cases (presences or absences).

P = 0.5 can correspond to a very favourable area if the species is rare, or to a non-favourable area if the species is common.

Probability and favourability

$$P = \frac{e^y}{1+e^y}$$

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n$$

Favourability:

$$y' = y - \ln(n_1/n_0)$$

Probability and favourability

$$P = \frac{e^{y}}{1+e^{y}}$$

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n$$

Favourability:

$$y' = y - \ln(n_1/n_0)$$

$$F = \frac{e^{y'}}{1+e^{y'}} = \frac{e^{y}}{\frac{n_1}{n_0}+e^{y}}$$

Real *et al*. 2006. *Environmental and Ecological Statistics*, 13: 237-245.

Probability and favourability

$$P = \frac{e^{y}}{1+e^{y}}$$

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_n x_n$$

Favourability:

$$y' = y - \ln(n_1/n_0)$$

$$F = \frac{e^{y'}}{1+e^{y'}} = \frac{e^{y}}{\frac{n_1}{n_0}+e^{y}}$$

F depends only on the relationship between the presence-absence of the species and the predictors, and not on the overall prevalence.

Probability and favourability

$$F = \frac{e^{y}}{\frac{n_1}{n_0} + e^{y}}$$

F is a logistic function varying between 0 and 1.

Favourability scores greater than 0.5 correspond to localities where the probability of occurrence is higher than that expected by prevalence (or chance), while scores less than 0.5 indicate localities with a lower probability than that expected by chance.

Favourable areas do not necessarily reflect the areas where the species is predicted to be present but favourable areas for the species, even in the absence of it.

Models for all species are then levelled to the same threshold of favourability and can be directly compared and combined.

Probability and favourability

$$F = \frac{\frac{P}{(1-P)}}{\frac{n_1}{n_0} + \frac{P}{(1-P)}} = \frac{e^y}{\frac{n_1}{n_0} + e^y}$$

Favourability can be obtained from **any species distribution model** that gives **probability** of presence as an output.

Environ Ecol Stat (2006) 13:237–245
DOI 10.1007/s10651-005-0003-3

ORIGINAL ARTICLE

Obtaining environmental favourability functions from logistic regression

Raimundo Real · A. Márcia Barbosa · J. Mario Vargas

Naturwissenschaften
DOI 10.1007/s00114-012-0926-0

CONCEPTS & SYNTHESIS

Favourability: concept, distinctive characteristics and potential usefulness

Pelayo Acevedo · Raimundo Real

Probability and favourability

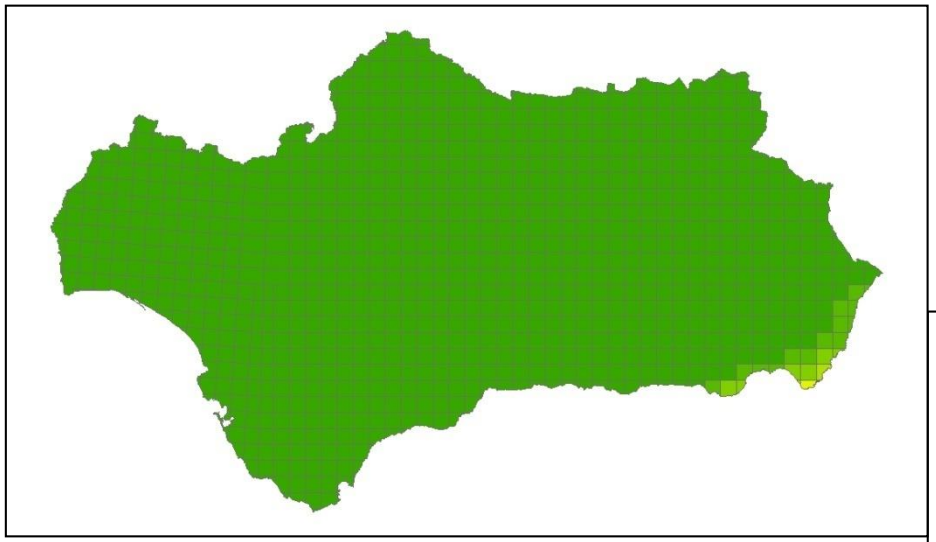
Some examples



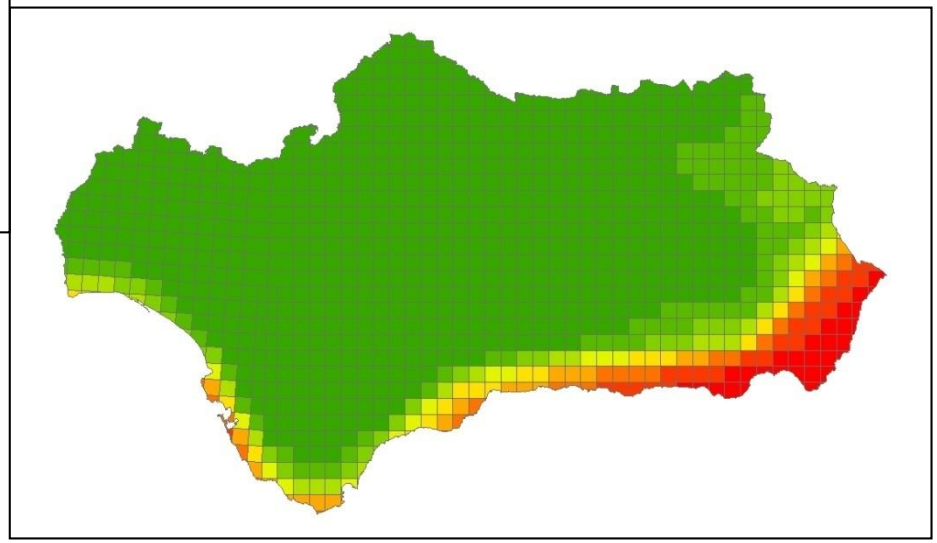


Probability and favourability

North African hedgehog (*Atelerix algirus*)



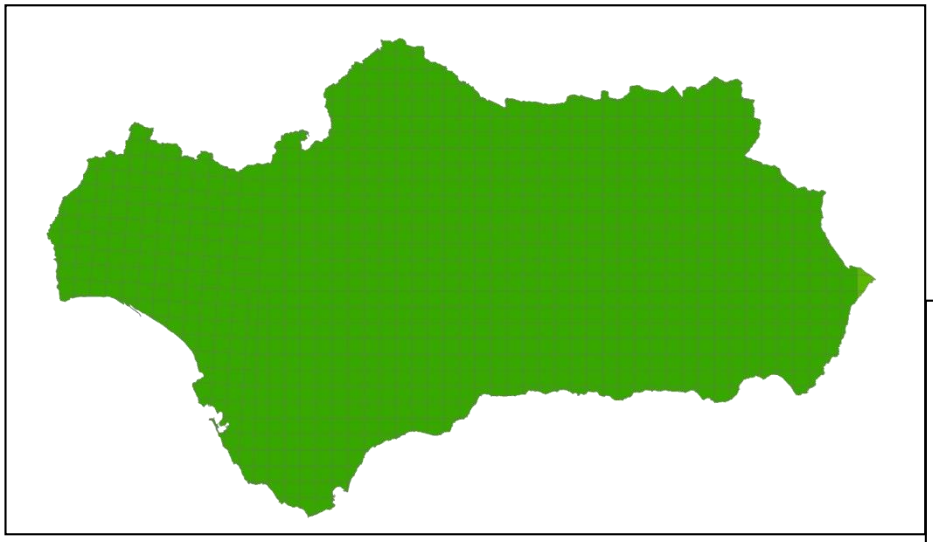
Probability



Favourability

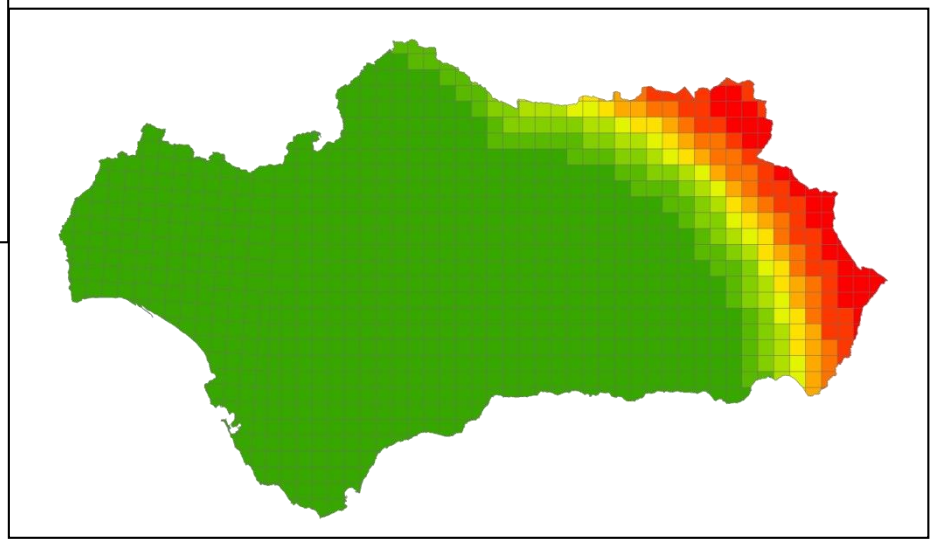
Probability and favourability

Barbary sheep (*Ammotragus lervia*)



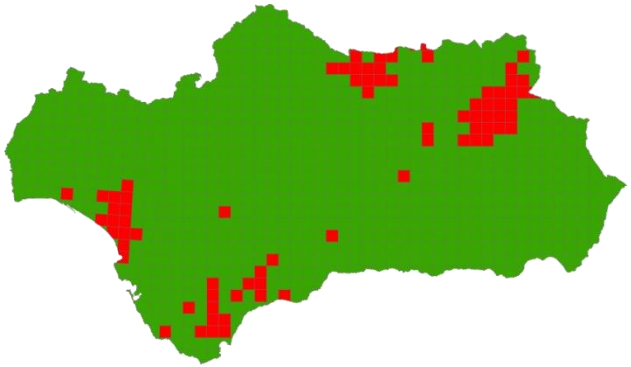
Probability

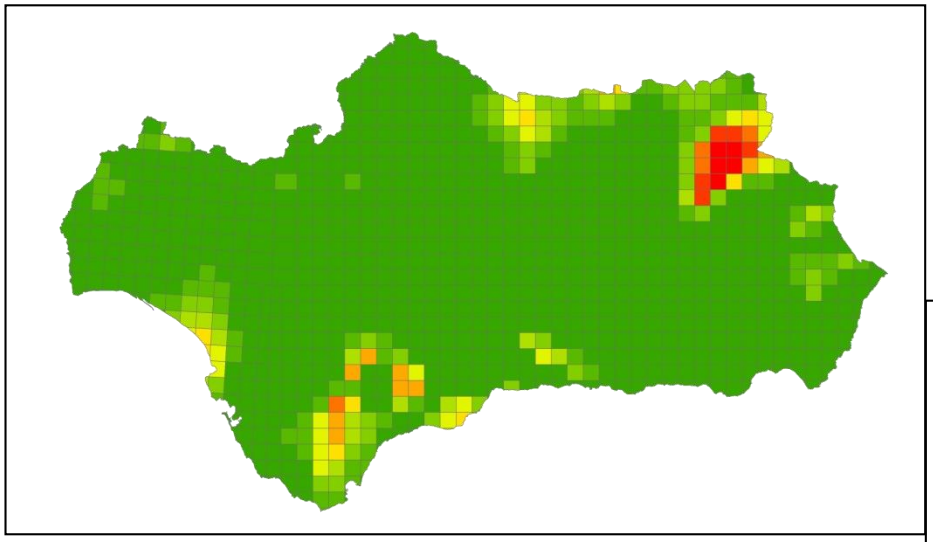
Favourability



Probability and favourability

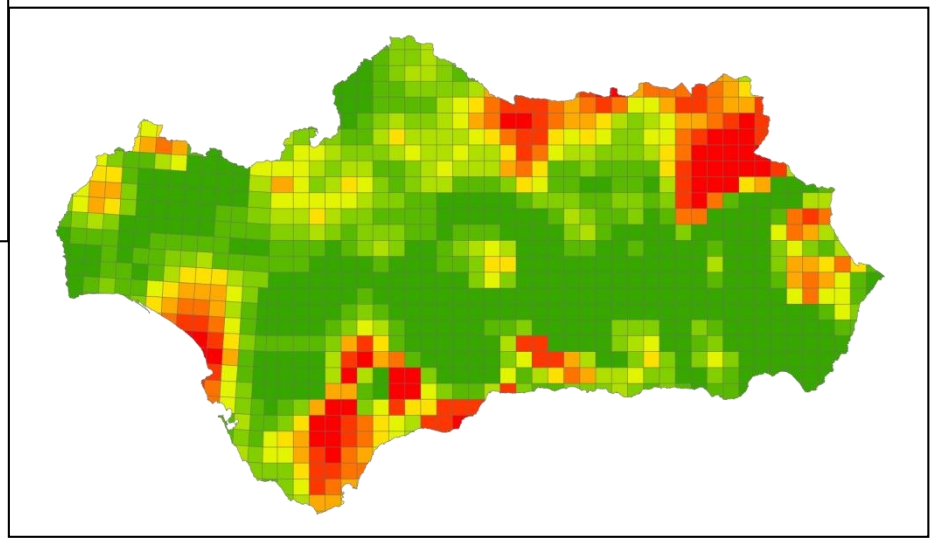
Fallow deer (*Dama dama*)





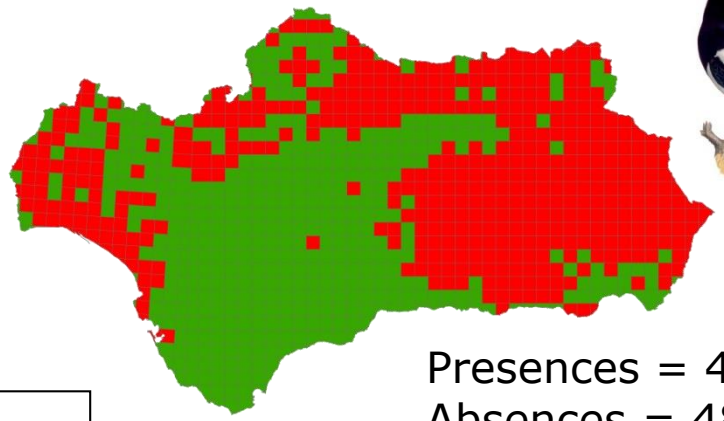
Probability

Favourability



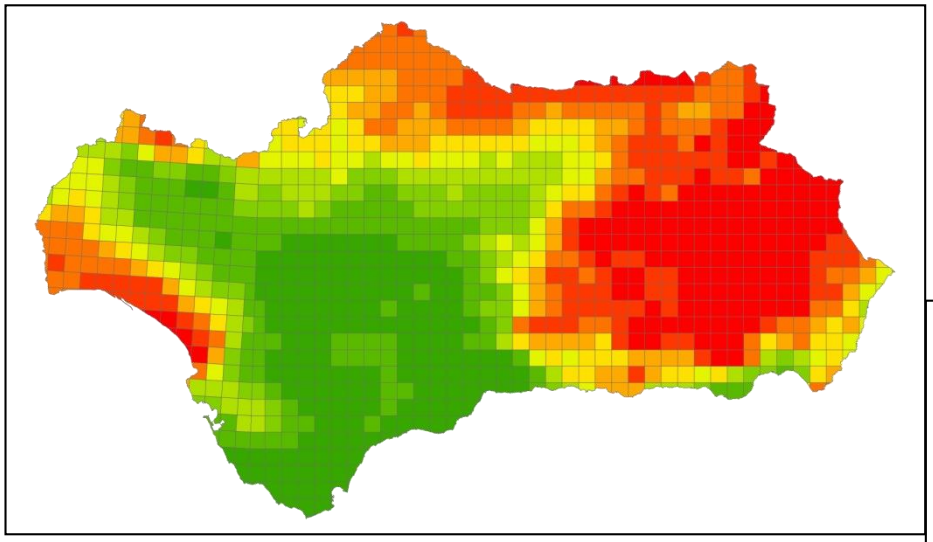
Probability and favourability

Common magpie (*Pica pica*)



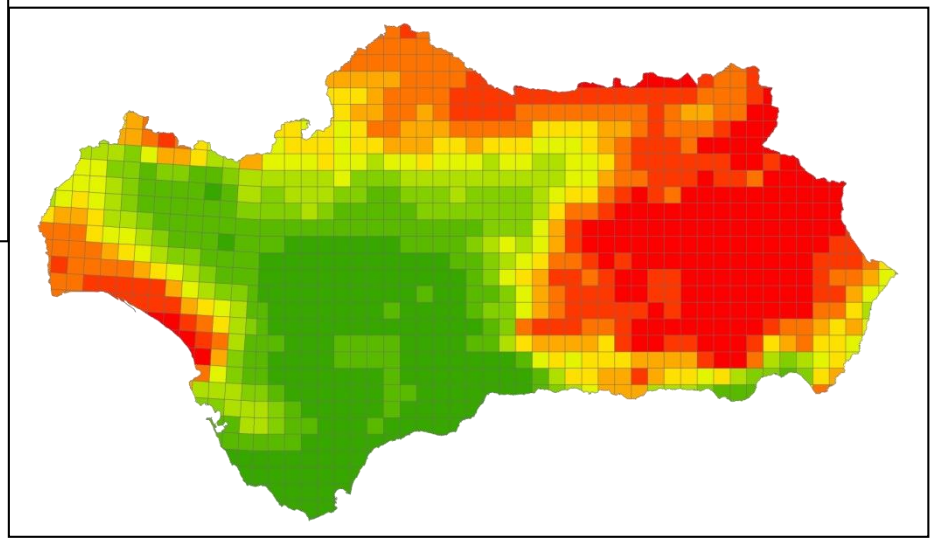
$$F = \frac{e^y}{\frac{n_1}{n_0} + e^y}$$

Presences = 478
Absences = 483



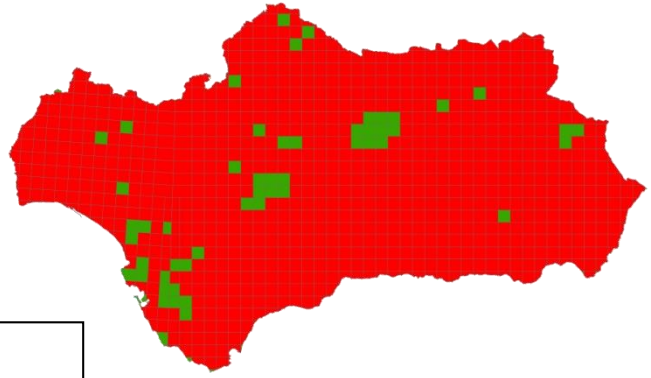
Probability

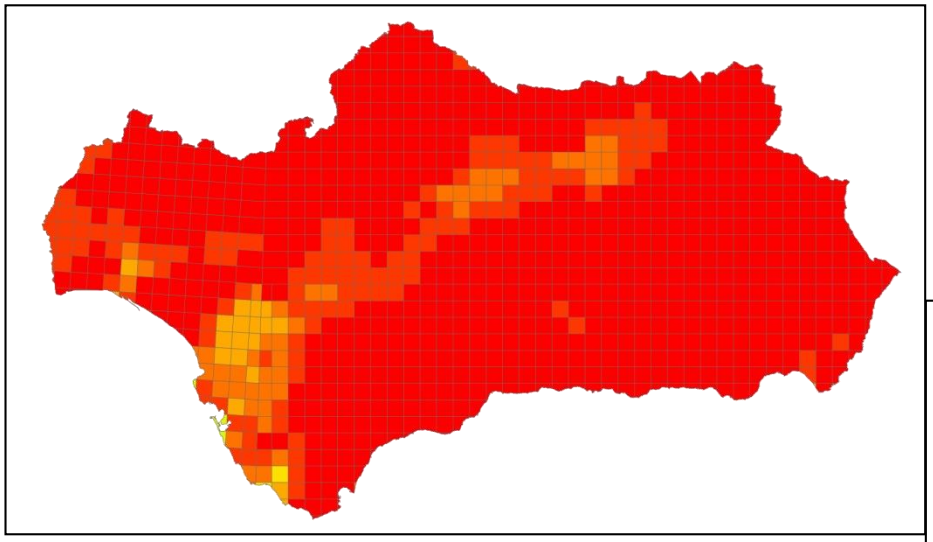
Favourability



Probability and favourability

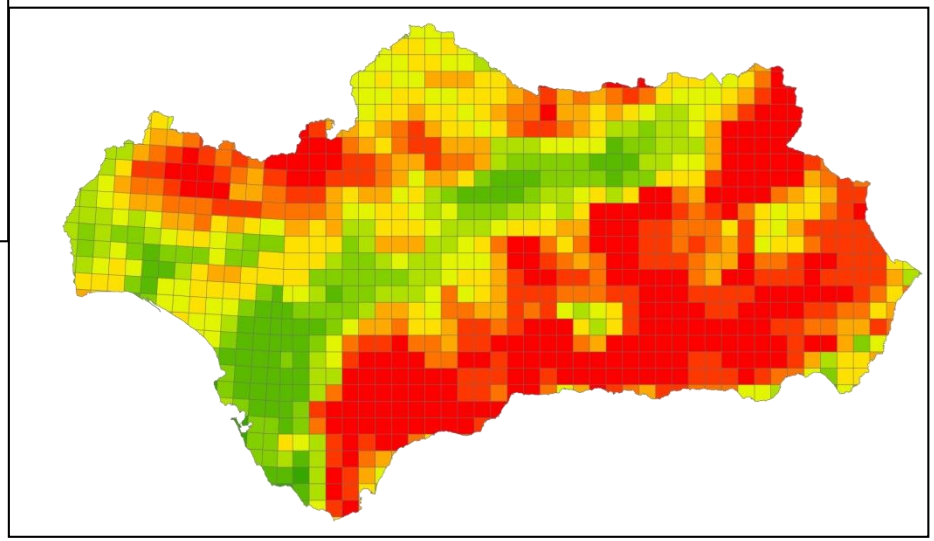
Common blackbird (*Turdus merula*)





Probability

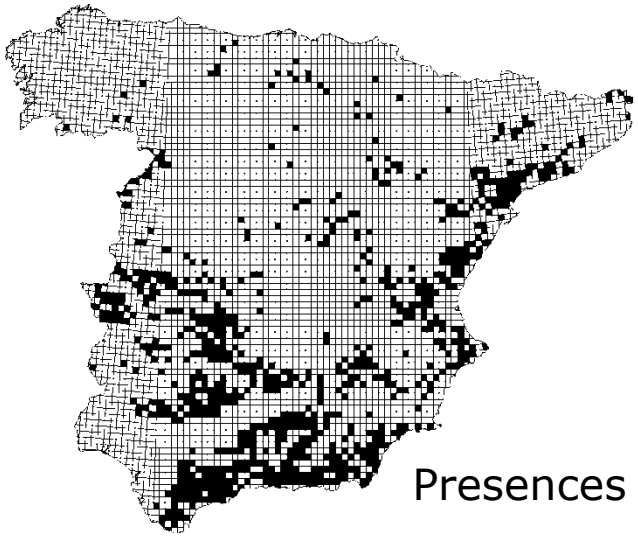
Favourability



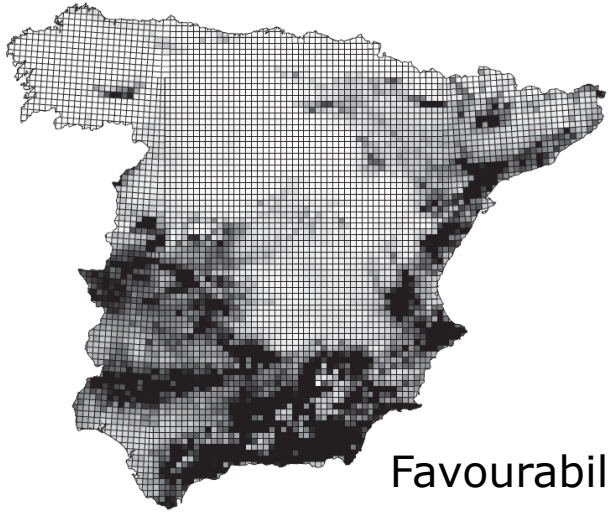
Favourability

Some applications

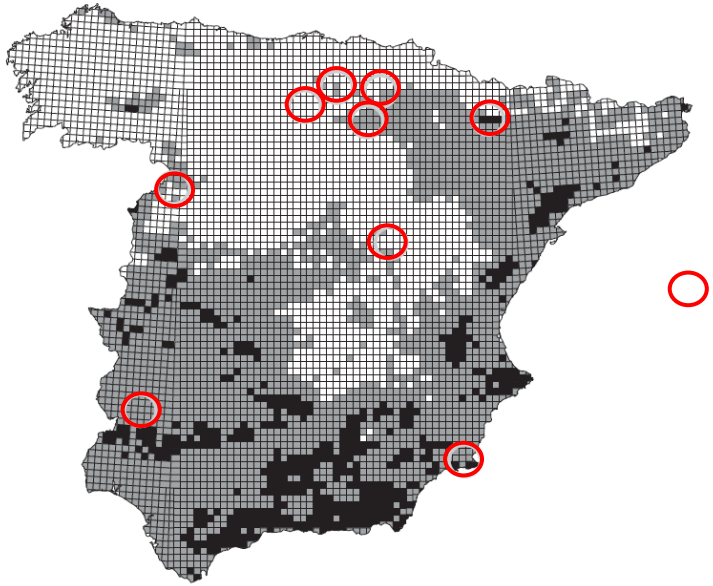
Bonelli's Eagle (*Hieraaetus fasciatus*) in Spain



Presences



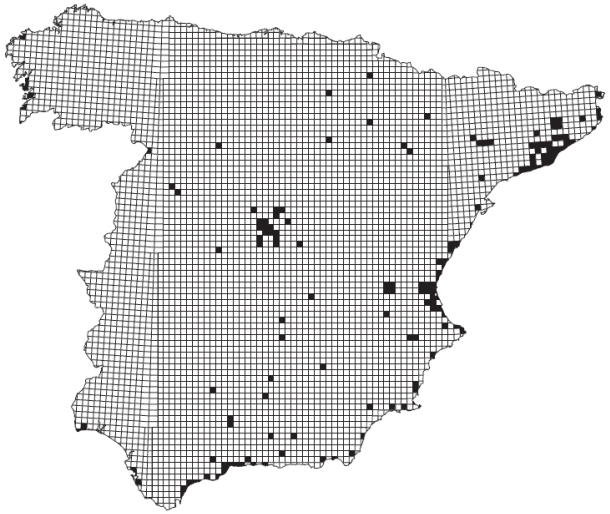
Favourability



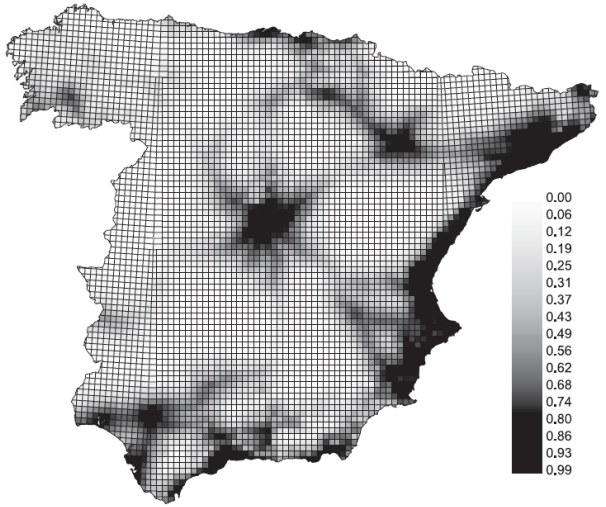
○ Conservation LIFE projects.

Muñoz *et al*. 2005. *Diversity and Distributions*, 11: 677-486.

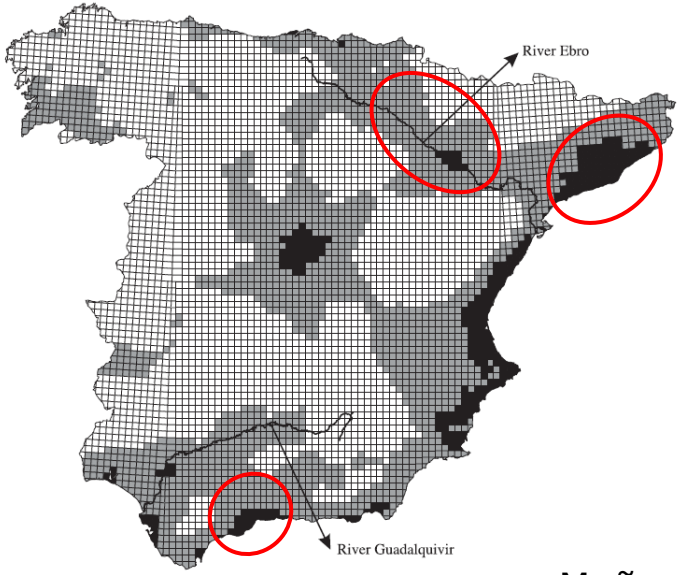
Monk parakeet (*Myiopsitta monachus*)
Invasive species in Spain

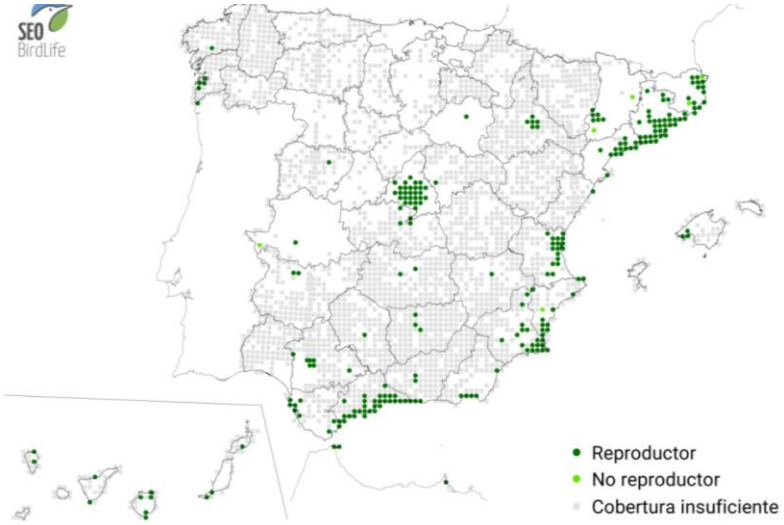


Presences

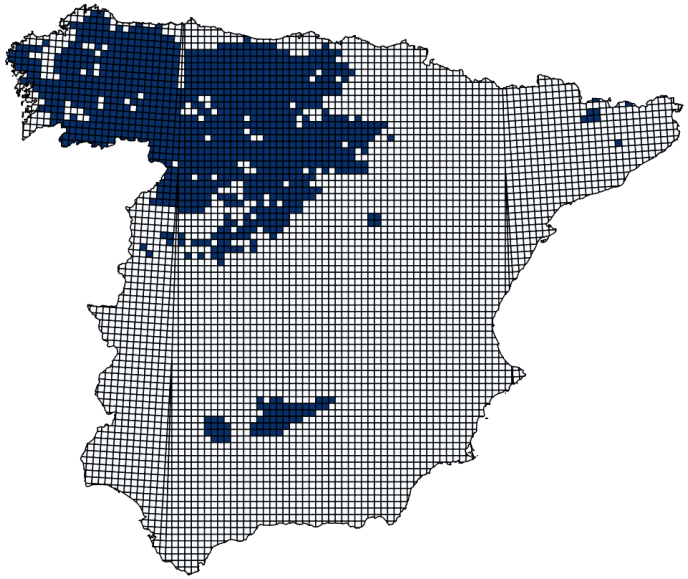


Favourability

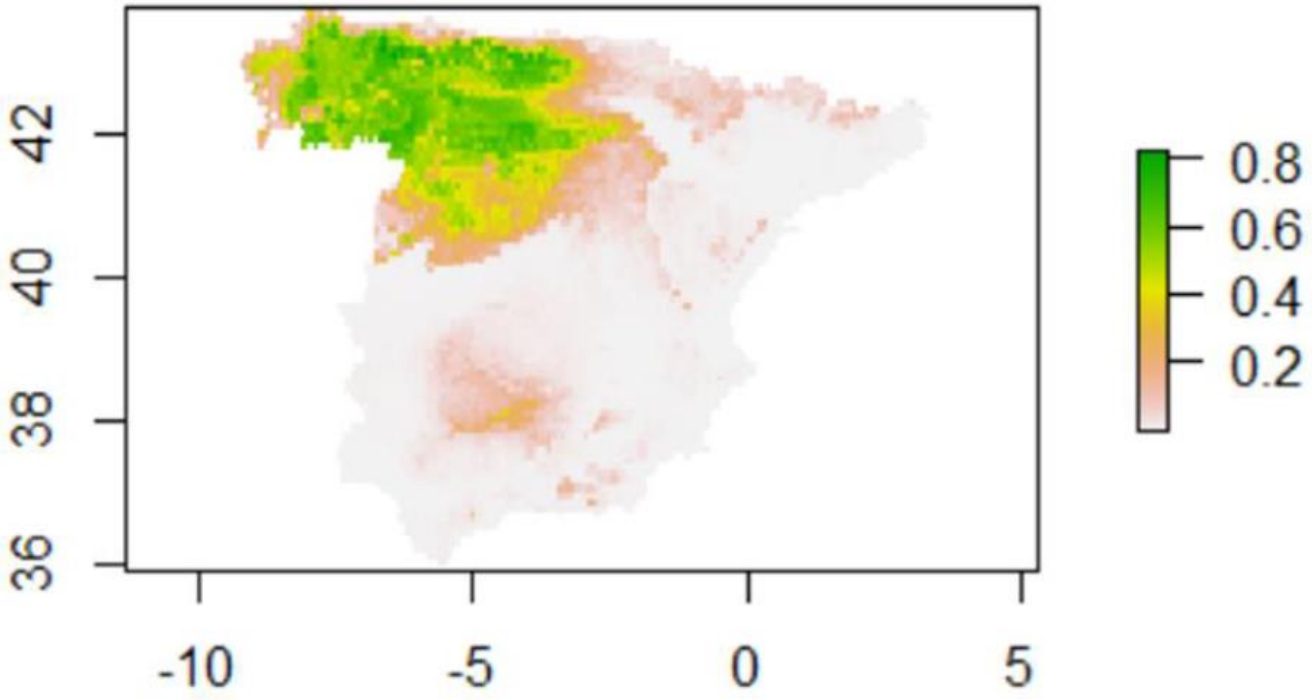




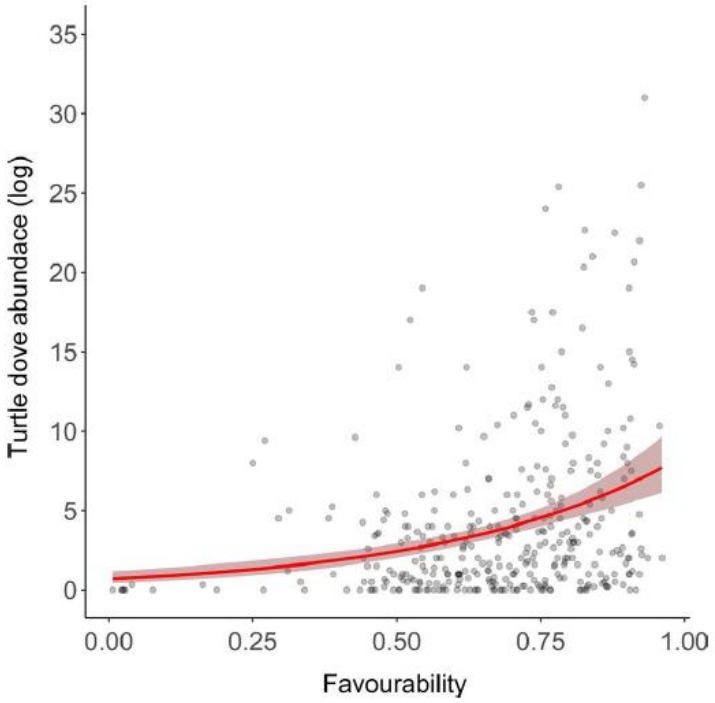
Muñoz & Real. 2006. *Diversity and Distributions*, 12: 656-665.

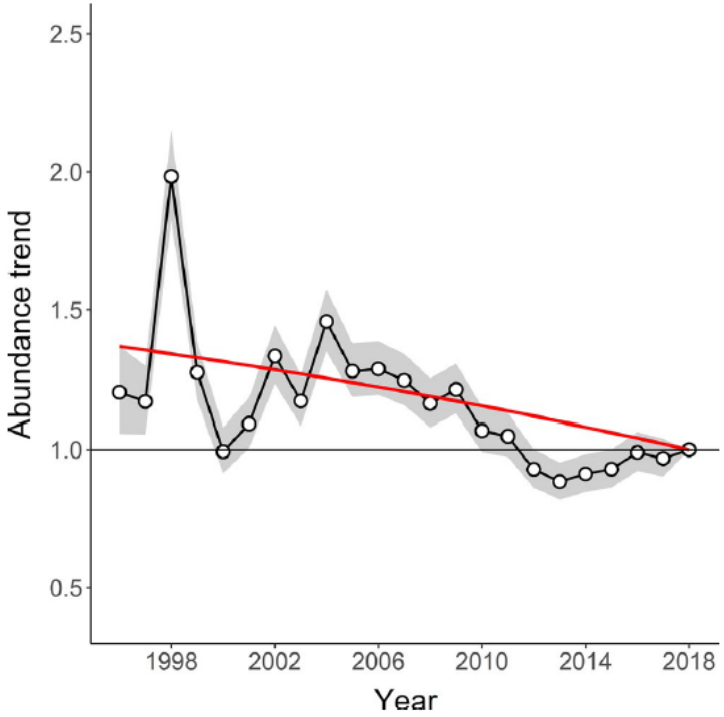


Wolf (*Canis lupus*)



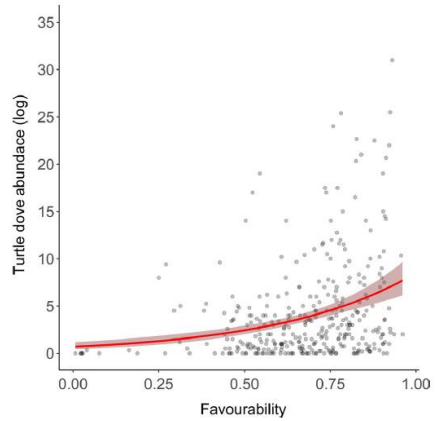
European turtle dove (*Streptopelia turtur*)

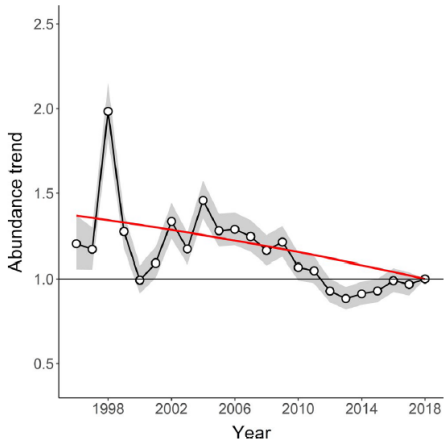


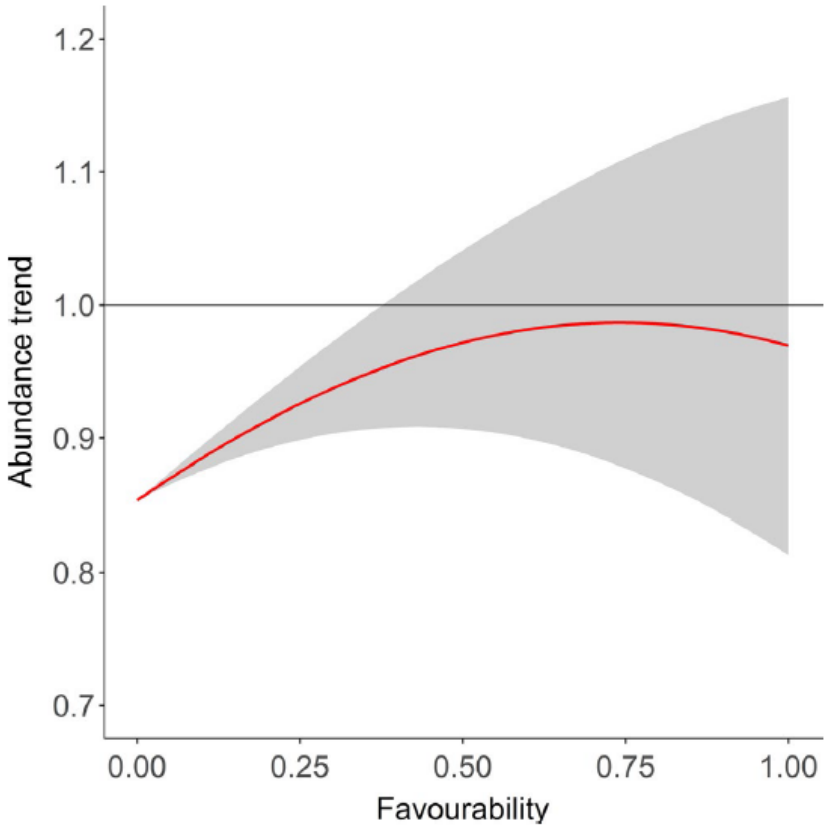


Moreno-Zarate et al. 2020. *Diversity and Distributions*, 26: 818-831

European turtle dove (*Streptopelia turtur*)



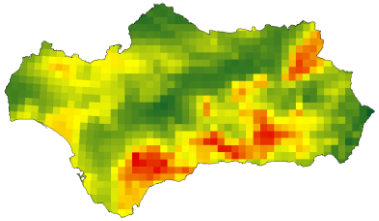




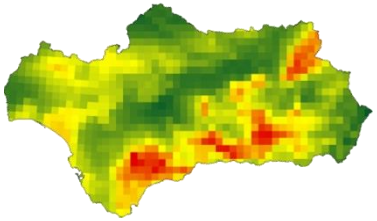
Moreno-Zarate et al. 2020. *Diversity and Distributions*, 26: 818-831

Combination of models

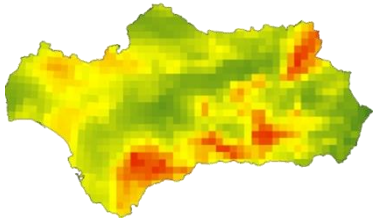
Important areas for mammals after aplying different conservation criteria and fuzzy logic



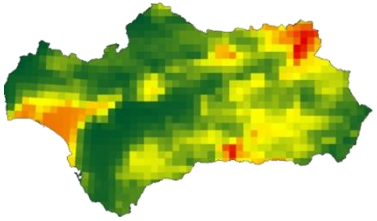
Richness



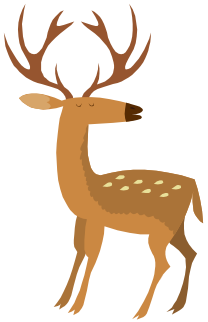
Rarity



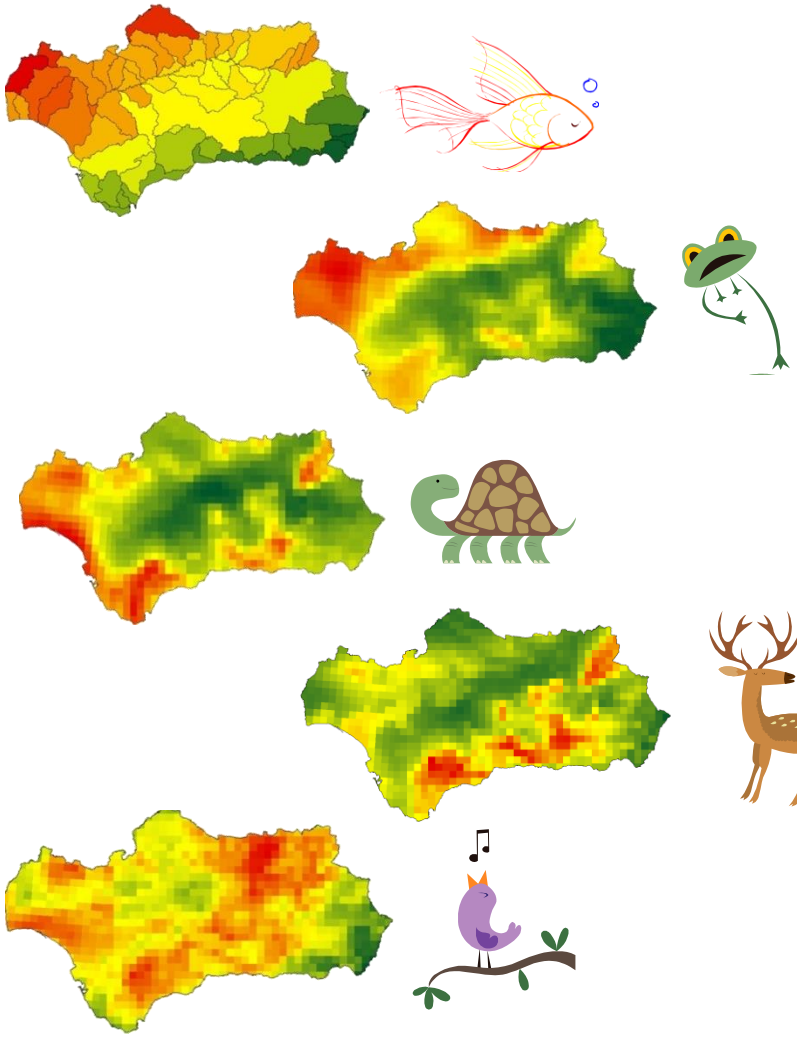
Vulnerability



Endemicity



Important areas for vertebrate groups



Estrada *et al.* 2011. *Biological Conservation*, 144: 1120-1129.

PRACTICAL

```
# convert probability to favourability:

dat$GLM_F <- Fav(pred = dat$GLM_P, sample.preval =
prevalence(dat_train[ , obs]))

dat$GAM_F <- Fav(pred = dat$GAM_P, sample.preval =
prevalence(dat_train[ , obs]))

# Maxent can't be converted to favourability because it
does not incorporate sample prevalence

# try it with your own models while we're here to help!
```