entropy 2021 The Scientific Tool of the 21st Century
5–7 May 2021, ONLINE

Scientia Biodiversitatis
Biogeografía, Diversidad y Conservación

UNIVERSIDAD DE MÁLAGA

A stepwise assessment of parsimony and entropy in species distribution modelling

Raimundo Real* & Alba Estrada

*Contact: rrgimenez@uma.es

Entropy is an **intrinsic** characteristic of the geographical **distribution** of a biological **species**. A species distribution with higher entropy involves more uncertainty, i.e., is more gradually constrained by the environment.

Species distribution modelling tries to yield models with **low uncertainty**, but normally has to produce them by **increasing their complexity**, which is detrimental for another desirable property of the models, parsimony.

Methods

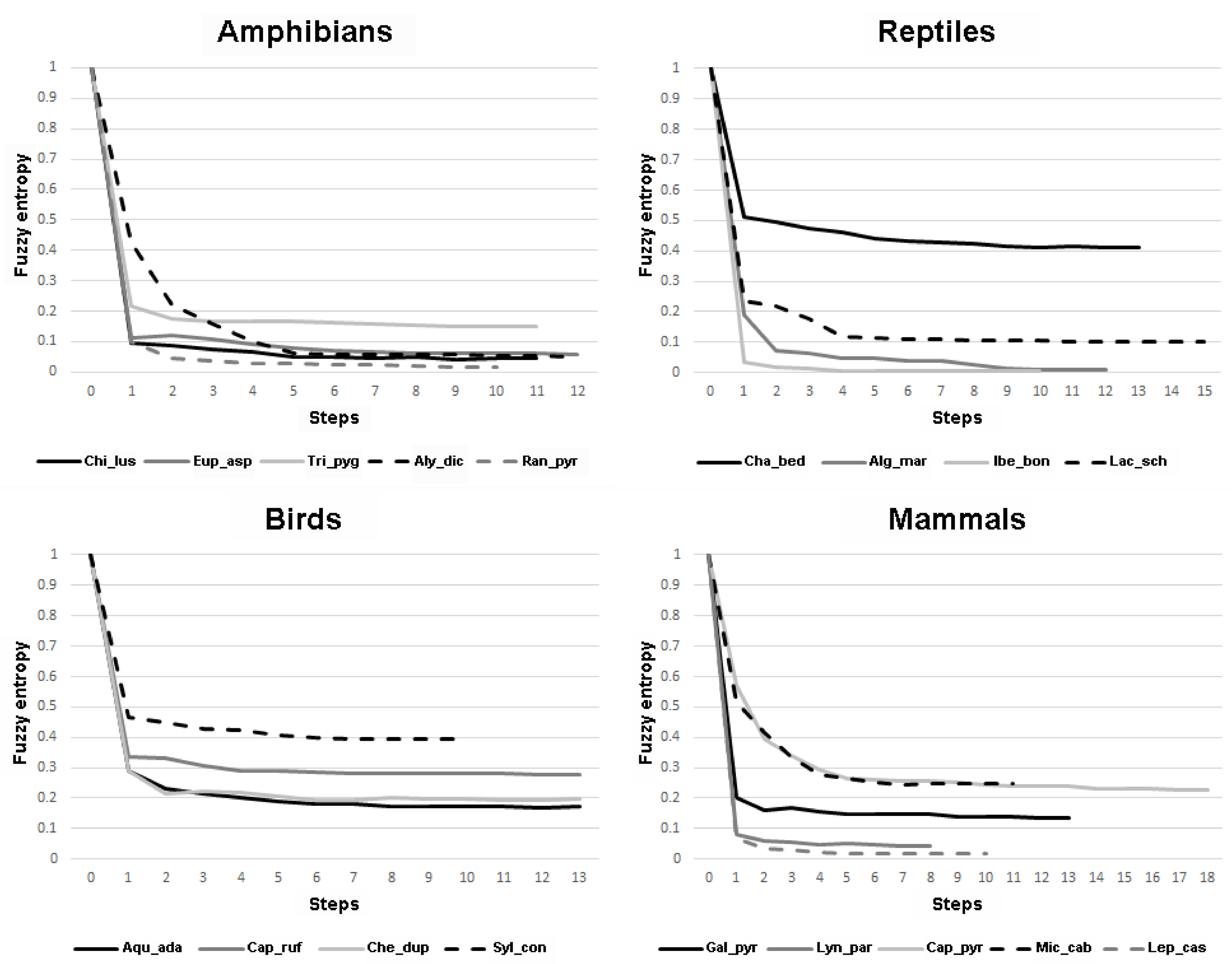
We modelled the distribution of **18 vertebrate species** in mainland Spain through Generalized Linear Models (GLMs) performed with a **forward-backward stepwise** selection of variables. We obtained favourable areas for each species in Spain after applying the **favourability** function[1]. The use of the favourability function allows the application of **fuzzy logic** to the resulting spatial analysis of the species, e.g., the calculation of the **fuzzy entropy**.

Fuzzy entropy[2] $$R = \frac{\sum_{i=1}^{n}(F_i \cap F_i^c)}{\sum_{i=1}^{n}(F_i \cup F_i^c)}$$

Fuzzy entropy has values between **zero** and **one**. If fuzzy entropy is one, the distribution of the species is completely disordered, i.e., favourability is equally distributed in all the territory with $F_i = 0.5$.

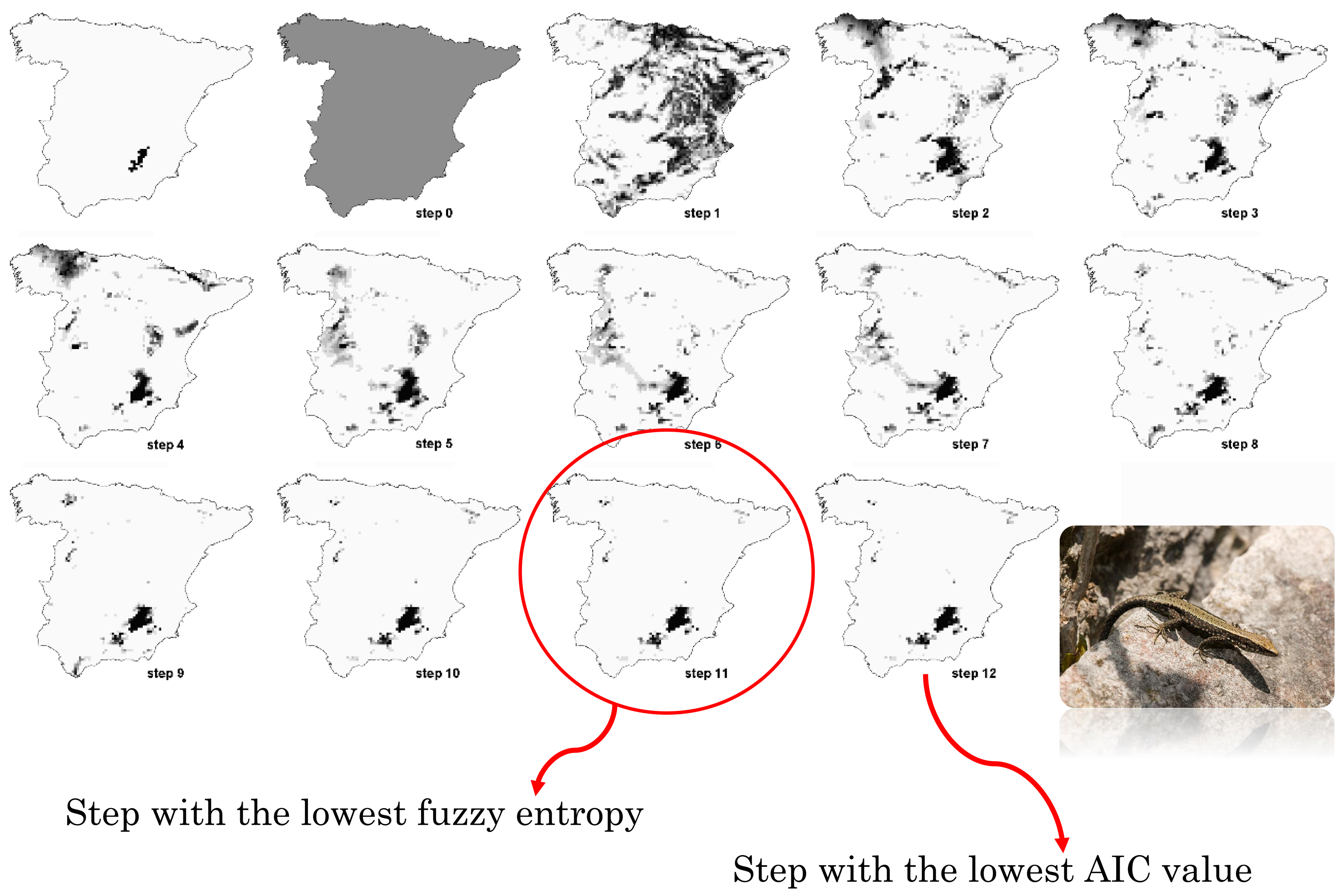
The smaller the entropy, the more orderly the distribution of the species is, i.e., the model more clearly distinguished between presences and absences. For each species, we also calculated the fuzzy entropy at each of the steps of the stepwise procedure to check whether uncertainty is reduced at each step.

Results for all species



A **reduction of entropy** was produced **asymptotically** in **each step of the model**. This asymptote could be used to identify the entropy attributable to the species distribution rather than to model misspecification.

Example species: Spanish algyroides (*Algyroides marchi*)



Favourable areas obtained in the step with lowest entropy are very similar to favourable areas obtained with all the AIC steps, as the variables of the final steps are those defining just fine-scale distribution patterns.

We suggest using the **favourability function** and the **fuzzy entropy** to obtain the **uncertainty** of the models. This approach implies a **trade-off between parsimony and uncertainty in species distribution models**. In stepwise modelling, each step yields a model with its own entropy value and, thus, **selecting the model with the lowest entropy** implies combining maximum **reduction of uncertainty and maximum parsimony**, which results in **high efficiency**.

References
1 Real, R., Barbosa, A.M. & Vargas, J.M. (2006) Obtaining environmental favourability functions from logistic regression. *Environ Ecol Stat*, 13, 237-245.
2 Kosko, B. (1986) Fuzzy entropy and conditioning. *Information Sciences*, 40, 165-174.

Address Raimundo Real and Alba Estrada:
Biogeography, Diversity and Conservation Research Team, Department of Animal Biology, Universidad de Málaga, Spain.