Original article

Evaluating how species niche modelling is affected by partial distributions with an empirical case

Miguel A. Carretero [a, *], Neftalí Sillero [b]

[a] CIBIO - Research Centre in Biodiversity and Genetic Resources, InBIO, Universidade do Porto, Campus de Vairão, Rua Padre Armando Quintas, N° 7, 4485-661, Vairão, Vila do Conde, Portugal
[b] CICGE - Centro de Investigação em Ciências Geo-Espaciais, Faculdade de Ciências da Universidade do Porto (FCUP), Observatório Astronómico Prof. Manuel de Barros, Alameda do Monte da Virgem, 4430-146, Vila Nova de Gaia, Portugal

ARTICLE INFO

ABSTRACT

Ecological niche models (ENMs) will successfully identify a species' ecological niche, provided that important assumptions are fulfilled, namely environment equilibrium and niche equality across the distribution. Violations may seriously affect ENM reliability, leading to erroneous biogeographic conclusions and inappropriate conservation prioritisation. We evaluate the robustness of ENMs against incomplete knowledge of distribution with a real example, the threatened Iberian lizard *Podarcis carbonelli*, whose distribution was gradually discovered over a long time period. We used several ENM methods for presence-only data (Maxent, ENFA, Bioclim, and Domain) to infer the realised ecological niche at two spatial resolutions (1 km and 200 m). The distribution data were split into four partial datasets corresponding to separate subranges: Central System (CS); Viseu-Aveiro (VA); Atlantic coast (AC); and Doñana (DO). We then accumulated the datasets following the species discovery sequence: CS + VA, CS + VA + AC, and CS + VA + AC + DO. Niche equivalence and similarity between partial models were compared using Ecospat. ENMs were strongly affected by the violation of niche equilibrium; only the VA subrange forecasts the complete species range. ENMs were also sensitive to the violation of niche equality: only VA models were similar to the Iberian model, altitude being the most important variable followed by annual precipitation, maximum temperature in July, and annual radiation. When the ENMs were applied only to the first subrange discovered (CS), only the VA area was predicted, while the other subranges might have remained unknown, thus compromising conservation strategies. As assumptions of niche equilibrium and equality were violated, likely owing to the species' ecological multimodality, the models generated were biased and of limited predictive value. ENMs are useful tools in biogeography and conservation, but only if their basal assumptions are achieved. Partial models may be useful if they are considered as representing different suitable habitats.

© 2016 Elsevier Masson SAS. All rights reserved.

1. Introduction

Ecological niche modelling (ENM) aims to identify the species' ecological niche (Sillero, 2011; Warren, 2012, 2013). ENMs aim to provide a simple mathematical representation of how organisms' distributions are limited by the environment (Guisan and Zimmermann, 2000). Researchers have developed ENMs for almost all possible ecological processes, including range changes (speciation, expansion, invasion, extinction: Graham et al., 2004; Pyron et al., 2008; Cartens and Richards, 2007; Ficetola et al., 2009) and interactions (competition, hybridisation, morphological evolution: Rissler and Apodaca, 2007; Costa et al., 2008; Martinez-Freiria et al., 2008; Tarkhnishvili et al., 2010). In particular, ENMs have been widely applied in biogeography (Sillero et al., 2009) and conservation (Greenwald et al., 2009).

However, ENMs are based on several assumptions (Anderson, 2013): low correlation between environmental variables; reliability of presence and/or absence species records; appropriate size of study area; appropriate spatial resolution of environmental layers; uniform intensity on species' records; environment equilibrium; equality of the species' niche across all its distribution (Guisan and Zimmermann, 2000; Wiens et al., 2009; Peterson et al., 2011); and obviously correspondence between the species'

ecological niche model and study conclusions (Sillero, 2011). In fact, violations of the environment equilibrium and the equality of the species' niche may seriously affect ENM results (Wiens et al., 2009). The species is expected to be in equilibrium with the environment, i.e. it is expected to occupy all suitable habitats available and to be absent from all unsuitable ones (Araújo and Pearson, 2005; Wiens et al., 2009; Peterson et al., 2011). However, this condition may not be accomplished: if a species is still expanding its distribution (i.e. invasive species Ficetola et al., 2009), it is excluded from some areas by historical reasons (Soberon and Nakamura, 2009), or involved in large scale source-sink dynamics (Pulliam, 2000), chorological records will not fully represent its ecological niche breadth, and the ENM may not identify all potentially suitable areas (Sillero, 2011). On the other hand, the niche of the species should also remain the same across all the distribution range, i.e. all the species' populations should have the same ecological niche (Grinnell, 1917; Elton, 1927; Hutchinson, 1957). Although this condition is difficult to verify directly, it can be inferred indirectly: ENMs constructed from divided data sets from representative portions of the species' range should be similar to the ENM for the whole range (Sillero, 2010).

When ENMs violate one of these conditions (environmental equilibrium and niche equality), results may lead to erroneous conclusion, including those determining the species' conservation status (e.g. Santos et al., 2006; Santos et al., 2009). This may be crucial when species distribution ranges are expanding (Ficetola et al., 2007, 2009) or when part of their distribution is unknown, especially as a result of the underestimation of its ecological valence (Pulliam, 2000). Obviously, such underestimation may have repercussions on the species range estimations and even on the conservation priorities. In this context, the main aim of this study is to evaluate the robustness of ENMs when the knowledge of a species' distribution is incomplete in terms that records do not represent the whole ecological niche of a given species. If such cases, we intend to provide alternative interpretations for ENMs. Moreover, correlative ENMs are very sensitive to the spatial distribution of the records, as the distribution and number of presence records directly determine the result of the model (Guisan and Zimmermann, 2000; Sillero, 2011; Warren, 2012). Therefore, very different results may be expected from partial distributions when compared to the whole species' range. In order to test this hypothesis, we used a real example of a threatened lizard species, whose distribution was gradually discovered over a long period of time (several decades), in order to discuss possible conservation consequences.

2. Methods

2.1. Study species and chronological discovery

The Carbonell's wall lizard, *Podarcis carbonelli*, Pérez-Mellado (1981) is endemic to the Western Iberian Peninsula (Fig. 1). Originally described as a subspecies of *P. bocagei*, it was much later elevated to full species rank (Sá-Sousa and Harris, 2002). It ranges in western Iberia south to the Duero river, with a fragmented distribution (Sá-Sousa, 1999, 2004, 2008; Sá-Sousa and Harris, 2002, Fig. 1) across the Central Mountain Range; following the western Atlantic coast of Portugal, splitting into several isolated nuclei southwards; and finally in a highly separated range in Doñana, in the south-western Atlantic coast of Spain.

Nonetheless, in historical terms, the distribution originally recognised in 1981 was restricted to a small part of this range in the Western Central System (Table 1). During the next two decades subsequent discoveries extended that range westwards (Fig. 1) and then along the western Atlantic coast of Portugal (Table 1). The presence in the enclave of Doñana, initially reported in 1986 but later denied 1998, was finally based on mitochondrial DNA data (Table 1). Remarkably, Carretero et al. (2002) reported a coastal locality of syntopy between *P. carbonelli* and *P. bocagei* (Espinho, south to Douro river) where selection against the hybrids were later demonstrated (Pinho et al., 2009), corroborating its specific status.

2.2. Study area and distribution datasets

The Iberian Peninsula, the region were the species occurs, encompasses a large and heterogeneous area with a complex orography and a transition between Atlantic and Mediterranean domains. Anderson and Raza (2010) suggested excluding those areas where the species cannot disperse. However, the actual distribution of *P. carbonelli* is the result of reduction from a larger range during the Pleistocene and the Holocene (11,000—13,000 years ago), in a warmer post-glacial period with larger forests than present (Carretero, 2008; Sá-Sousa, 2008; Sillero and Carretero, 2013). Therefore, it might be possible to find populations of this
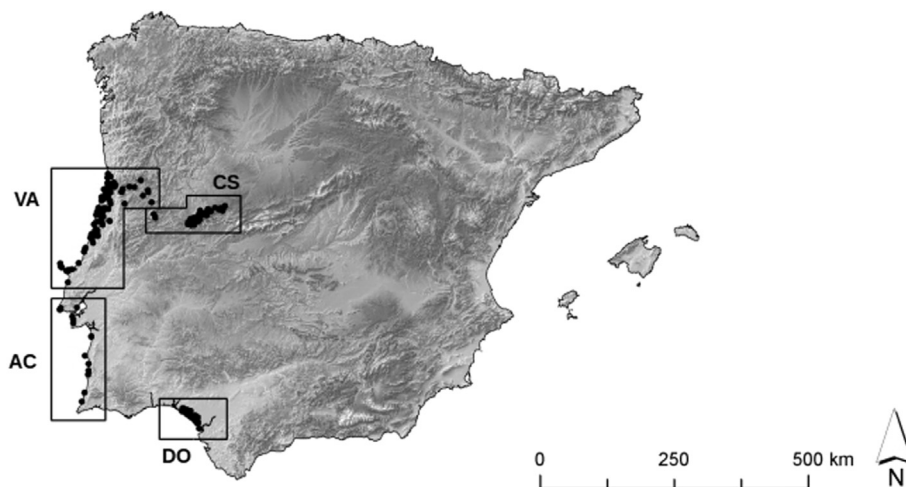
Fig. 1. Datasets used to model the distribution of *Podarcis carbonelli*. We divided species records into four partial datasets following the sequence of the discovery of the species distribution: Central System (CS), Viseu-Aveiro (VA), Atlantic coast (AC) and Doñana (DO); two merged datasets (CS + VA and CS + VA + AC); and the complete Iberian dataset. See Table 1 for number of records and Methods section for more details.