Ecological Complexity 20 (2014) 208–212

Contents lists available at ScienceDirect

Ecological Complexity

journal homepage: www.elsevier.com/locate/ecocom

Short Note

Associations between species can influence the goodness of fit of species distribution models: The case of two passerine birds

Federico Morelli [a,*], Piotr Tryjanowski [b]

[a] DiSTeVA, University of Urbino, Scientific Campus, 61029 Urbino, Italy
[b] Institute of Zoology, Poznan University of Life Sciences, Wojska Polskiego 71C, 60-625 Poznań, Poland

ARTICLE INFO

Article history:
Received 22 August 2013
Received in revised form 23 December 2013
Accepted 2 January 2014
Available online 1 February 2014

Keywords:
Species-associations
SDMs
Goodness of fit
Shrikes
Buntings
Species-specific model

ABSTRACT

Species distribution models (SDMs) are numerical tools that combine observations of species presence or abundance with environmental data, in order to develop predictive estimates about species distribution. The main variables used as regressors on SDMs are environmental parameters (climate variables, land-uses typologies, landscape metrics, etc.). However, there are several aspects that can affect the goodness of fit of species distribution models. For example, species traits, presence of associated species and anti-predator behaviors may cause differential responses to the processes that control their distribution. Considering these kinds of factors should improve the performance of models.

In this work, SDMs were performed for two bird species found associated during initial exploration of data, by mean of a correlation matrix among bird occurrence. Models were performed first on environmental variables and then on environmental variables plus the occurrence of associated species. The goodness of fit of SDMs was compared using the area under the curve (AUC) and the likelihood ratio test.

Our results showed how the associations between bird species can affect the goodness of fit of species distribution models. Specifically, we documented a significant increase in the AUC of the best model for red-backed shrike when adding the occurrence of corn bunting as a predictor. Our findings suggest how species-specific models in applied ecology can be developed to improve the predictive power of SDMs.

1. Introduction

Models are increasingly being used as key components of wildlife management programs because they provide a method to predict the outcomes of management and conservation strategies (Guisan et al., 2013). Species distribution models (SDMs) are numerical tools that combine observations of species occurrence or abundance with environmental estimates. These models can be used to provide understanding and/or to predict the species' distribution across a gradient of landscape (Elith and Leathwick, 2009). Predictive species distribution models are commonly applied as tools for the purposes of conservation planning and management of ecosystems (Guisan and Zimmermann, 2000). Furthermore, the outputs of species distribution models (spatial explicit predictions of environmental suitability for species) are considered important tools for making robust conservation decisions and to provide predictions on environmental suitability (Guisan et al., 2013).

Such models rely on the concept of the ecological niche being occupied by the detected species. However, there are several aspects that can affect the goodness of fit or performance of the species distribution models (Allouche et al., 2008; Guisan et al., 2007). The observed patterns in the spread of populations in nature are a result of the complex interplay between stochastic and deterministic factors, mainly environmental or/and demographic, as well as deterministic inter-species components (for example a predator–prey system) (Morozova et al., 2008). Environmental parameters (land-use typologies, climate variables, vegetation or landscape metrics) are often covariates (predictors) used as regressors into a model, but species traits may reflect differential responses to the processes that control their distribution and this can compromise the performance of models. Some examples are the use of biological traits (Seoane et al., 2005; Carrascal et al., 2006), but eco-ethological characteristics of the species can also have an impact. For example, is known that many Passeriformes birds usually avoid areas surrounding raptor nests (Meese and Fuller, 1989; Norrdahl and Korpimäki, 1998; Suhonen et al., 1994). In contrast, species such as woodpigeons *Columba palumbus* nest close to bolder, more aggressive birds, such as the eurasian hobby *Falco subbuteo*, which provide protection against nest predators (Bogliani et al., 1999). The presence of raven *Corvus corax* has been

* Corresponding author. Tel.: +39 3282097285..
E-mail addresses: federico.morelli@uniurb.it, morellius@libero.it (F. Morelli), piotr.tryjanowski@gmail.com (P. Tryjanowski).

1476-945X/$ – see front matter © 2014 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.ecocom.2014.01.002

shown to modify the composition of a bird community because birds nesting near raven nests can use it as an antipredator adaptation against nest-robbing by other predators (Tryjanowski, 2001).

The association may be beneficial to both species if a more timid species joins a bolder species during mobbing or if it contributes to early warning when a potential predator is approaching (Wiklund, 1979, 1982). On the other hand, numerous bird species exhibit social attraction during breeding site selection; a preference for settling near competitors that may either be conspecifics or heterospecifics – known as conspecific and heterospecific attraction (Hromada et al., 2008; Monkkonen et al., 1997; Stamps, 1988). Two species may also associate or coexist due to similar habitat selection (Kuzniak et al., 2001).

Despite our longstanding knowledge of the occurrence of associated species, this kind of biotic information is rarely used as a predictor in SDMs development (Kissling et al., 2012). Although studies are now beginning to underline how considering biotic interactions may advance the field of niche theory (Le Roux et al., 2013), few attempts have been made to develop 'species interaction distribution models' (SIDMs) that aim to incorporate multispecies interactions at large spatial extents using interaction matrices (Kissling et al., 2012).

In this study we focused on two bird species that are typical of agricultural landscapes in Central Italy (red-backed shrikes *Lanius collurio* and corn bunting *Miliaria calandra*). Red-backed shrikes are small to medium-sized passerine birds that hunt using behavioral techniques similar to raptors (Lefranc, 1993). Their food consists mainly of invertebrate prey and small vertebrates (Goławski, 2007; Tryjanowski et al., 2003) and they breed predominantly in agricultural landscapes and grasslands with scattered shrubs (Morelli, 2012). The corn bunting is a typical farmland bird in Central Italy that uses also shrublands and grasslands for breeding (Donald and Evans, 1995; Goławski and Dombrowski, 2002) and feeds mainly on invertebrates (Aebischer and Ward, 1997; Busche, 1989). Data on species association was used to test the performance of species distribution models (SDM), built using environmental variables and landscape metrics, in order to evaluate the usefulness and cost-effectiveness of modeling.

2. Methods

The study was carried out in central-eastern Italy, in the foothills of the Apennines in the Northern Marche region (43°49′ N 12°26′ E). From mid-April to end-June 2012 a mixed-farming and grassland area was surveyed by means of 600 point counts located randomly and at least 500 m apart. The sites were visited in the morning between 06:00 AM and 10:00 AM, during sunny weather conditions. Each visit lasted 10 min and during this time all birds were detected visually and acoustically recorded (Bibby et al., 1997).

Environmental data used for this study were obtained from a land-cover map of the Marche region (1:10.000) (AA.VV. 2008), the spatial scale of analysis was set at a radius 250 m (circa 20 ha) (Morelli et al., 2013). In order to quantify land-use composition and structural characteristics of the sampled sites, the area around the sampled-point was described. The percentages of land-use within the buffer was calculated using ArcGIS 10 and are summarized as follows: (a) creation of a buffer zones around each sampled point; (b) using intersect operator, overlap and clipping layers among buffers and land cover; and (c) Fragstats 3.2 and ArcGIS 9 were used to derive some landscape heterogeneity metrics (Schindler et al., 2013). The following environmental data were classified into two categories of spatial scale: landscape and land-use parameters (see Table 1).

The nature and strength of relationships between bird species occurrence and environmental parameters on the sampled sites were examined using Generalized Linear Models (GLM) (McCullagh and Nelder, 1989), with dependent variable (species occurrence) modeled specifying a binomial distribution. Explanatory variables were expressed as the arcsin root square in the case of proportions. In order to avoid the multi co-linearity of regressors, parameters with the strongest correlation between them (>0.7) were manually eliminated. A stepwise backward procedure was followed in order to select the best predictors using AIC criterion (Akaike, 1974). The best model for each bird species was selected using the lowest AIC.

The predictive performance of models was evaluated by calculating the area under curve (AUC) that considers sensitivity and specificity of model. The AUC is expressed as an index ranging from 0 to 1 (DeLong et al., 1988). An AUC equal to 0.5 is a random distribution of predictions and an AUC equal to one is a perfect prediction.

The exploration about potential relationships between bird species distribution (associated species) was checked by means of a correlation matrix built with data of occurrence of all bird species detected during the bird surveys.

In order to compare the accuracy of SDMs for each species the goodness of fit was compared with and without the addition of the occurrence of the other associated bird species as predictors. The differences were verified using the likelihood ratio test of models (that compare the output of two comparable models by mean of a Chi-square test). Model metrics were calculated using the R package "ROCR" (Sing et al., 2005). All tests were carried out using R (R Core Team, 2013).

Table 1
Environmental parameters used to perform the SDMs on red-backed shrike and corn bunting in Central Italy.

Parameter	Abbreviation	Spatial scale	Description
Altitude	alt	Landscape	Altitude of sampled-point (m/a.s.l.)
Land use number	lu	Landscape	Sum of different land use typologies
Polygons number	pn	Landscape	Sum of total polygons within the buffer
Edge density	wedg	Landscape	Sum of the perimeters of all polygons in the buffer zone per number of land use types/buffer surface (Hargis et al., 1998)
Land use diversity	ludiv	Landscape	Calculated using the Shannon–Weaver diversity index on land use types
Roads	roa	Land use	%
Urban	urb	Land use	%
Forest	for	Land use	%
Uncultivated and shrubs	unc	Land use	%
Badland	bad	Land use	%
Grassland	gra	Land use	%
Vineyard and Orchard	vin	Land use	%
Cultivated	cul	Land use	%
River	wat	Land use	%