



# Incorporating spatial constraints in different periods of the annual cycle improves species distribution model performance for a highly mobile bird species

Laura Cardador<sup>1,2\*</sup>, Francesc Sardà-Palomera<sup>2,3</sup>, Martina Carrete<sup>4,5</sup> and Santi Mañosa<sup>1</sup>

<sup>1</sup>Departament de Biologia Animal, Institut de Recerca de la Biodiversitat (IRBio), Universitat de Barcelona, Facultat de Biologia, Avinguda Diagonal 643, 08028 Barcelona, Spain, <sup>2</sup>Forest Sciences Center of Catalonia (CTFC), Crta. de Sant Llorenç de Morunys Km 2, 25280 Solsona, Spain, <sup>3</sup>Catalan Ornithological Institute (ICO), Museu de Zoologia, Passeig Picasso s/n, 08003 Barcelona, Spain, <sup>4</sup>Department of Physical, Chemical and Natural Systems, Universidad Pablo de Olavide, Ctra. Utrera km 1, 41013 Sevilla, Spain, <sup>5</sup>Estación Biológica de Doñana (CSIC), Américo Vespucio s/n, E-41092 Sevilla, Spain

## ABSTRACT

**Aim** The knowledge of both potential distribution and habitat suitability is fundamental for conservation planning and management of a variety of taxa world-wide. Species distribution models (SDMs) are increasingly applied as predictive tools for these purposes. Such models are based on the concept of ecological niche and assume that species distribute themselves based on niche spaces defined by climate and habitat features. However, this assumption can be violated due to the existence of pure spatial range constraints, a factor rarely accounted for in SDMs, particularly for highly mobile species. We analyse whether pure distance effects, niche-based environmental responses or a combination of both factors can play an important role in limiting the large-scale distribution of highly mobile species.

**Location** Spain, southern Europe.

**Methods** We modelled the spatial distribution of an expanding raptor species, the marsh harrier *Circus aeruginosus*, in Spain. We implemented one conventional statistical method (generalized linear model) and one nonparametric technique (maximum entropy, Maxent) using a large dataset on marsh harrier occurrence ( $n = 1586$ ) in two different periods of the annual cycle, that is breeding and wintering seasons. We developed models that included environmental variables and that either ignored or incorporated spatial constraints using spatial eigenvector mapping (spatial filters).

**Results** By comparing model accuracy, we found evidence that the distribution of marsh harriers in Spain was spatially constrained beyond environmental variables and that the effect of spatial constraints varies depending on the period of their annual cycle.

**Main conclusions** Contrary to the equilibrium postulate, our results support the prediction that environmental specializations do not necessarily result in complete habitat matching due to dispersal limitations. Thus, ignoring spatial constraints in SDMs can lead to misunderstandings of the ecological mechanisms that explain species range limits.

## Keywords

*Circus aeruginosus*, dispersal, range expansion, raptors, spatial eigenvector mapping, species distribution model.

\*Correspondence: Laura Cardador, Forest Sciences Center of Catalonia (CTFC), Crta. de Sant Llorenç de Morunys Km 2, 25280 Solsona, Spain  
E-mail: laura.cardador@ctfc.cat;  
lcardador81@gmail.com

## INTRODUCTION

Individuals usually show non-random, aggregated spatial distributions. Species-specific environmental responses or pure distance effects can be involved in generating these patterns (Dumbrell *et al.*, 2010). Under deterministic models, species distribute themselves based on defined niche spaces, and their probabilities of occurrence are expected to change in accordance with environmental variables (Austin *et al.*, 1990; Franklin, 1994; Snyder & Chesson, 2003). In this way, distributions can be patchy at the population level if the environment is spatially autocorrelated (Legendre, 2003). Climate has often been proposed as the main range-limiting factor within environmental variables, especially at large spatial scales (Huntley *et al.*, 2007). Non-climatic environmental factors such as topography, land use or soil types are thought to increase in importance at increasingly finer scales (Mackey & Lindenmayer, 2001; Pearson & Dawson, 2003; Anadón *et al.*, 2006). Nevertheless, distribution patchiness may also arise as a result of pure spatial constraints such as dispersal limitations (e.g. physical geographical barriers or time-limited expansions from place of origin; Moore *et al.*, 2008; Blach-Overgaard *et al.*, 2010). Species can be present and persist at particular locations not because of a better habitat quality, but because individuals mostly recruit near their natal areas or close to conspecifics, regardless of the specific environmental characteristics of these locations (Serrano *et al.*, 2004; Pinto & MacDougall, 2010).

Understanding the relative roles of environmental factors and pure spatial constraints as determinants of the geographical distribution of species holds ecological and theoretical interest (Austin, 2007; Dumbrell *et al.*, 2010) but also applied relevance. Spatial constraints can limit the abilities of a species to cope with global environmental changes, either in their resilience to reductions in available suitable conditions (particularly when these involve habitat fragmentation) or in their capacity for profiting from habitat expansions (Moore *et al.*, 2008; Martínez-Morales *et al.*, 2010). Environmental change is a natural phenomenon, but the rate and magnitude of such changes have significantly increased in recent decades due to anthropogenic causes (Donald *et al.*, 2001; Carrete *et al.*, 2009). Thus, a better understanding of the importance of pure spatial factors as predictors of species distributions is needed for robust conservation planning, as these pure spatial range constraints may indicate the extent to which a given species will be able to track these rapid environmental changes.

Species distribution models (SDMs) have provided a popular analytical framework for predicting species distributions and are applied across a range of biogeographical scales (Guisan & Thuiller, 2005). SDMs represent an empirical method to draw statistical inferences about the drivers of species' ranges under different conservation, ecological and evolutionary processes (Zimmermann *et al.*, 2010). However, although empirical evidence indicates that patterns of species distribution are often constrained by pure spatial limitations

(Pinto & MacDougall, 2010), most SDMs assume equilibrium between species distribution and the environment, ignoring its potential consequences (Guisan & Thuiller, 2005). Furthermore, although for mobile organisms seasonal variation in ecological requirements and factors limiting their ability to move to distant areas may occur (Lauzon-Guay & Scheibling, 2007; Cortés-Avizanda *et al.*, 2011), intra-annual variations in the relative importance of environmental characteristics and spatial constraints on species distributions have been seldom studied (Woinarski *et al.*, 1992; Sardà-Palamera *et al.*, 2012).

Here, we evaluate whether pure spatial predictors, environmental constraints (both climatic and habitat) or combinations of both factors can play an important role in limiting the large-scale distribution of highly mobile species using the large-scale distribution of a long-lived migratory raptor species, the marsh harrier *Circus aeruginosus*, in Spain, as a model. We did so by implementing SDMs constructed with environmental variables and by excluding or incorporating spatial constraint variables obtained by eigenvector mapping from the models. The inclusion of spatial constraint variables in SDMs has recently been shown to effectively capture non-environmental spatial constraints caused by dispersal-limited, non-equilibrium range dynamics (Griffith & Peres-Neto, 2006; De Marco *et al.*, 2010). We also assessed seasonal variations in the relative importance of ecological requirements and spatial constraints as determinants of species distribution by fitting a separate model for the breeding and wintering seasons.

## MATERIALS AND METHODS

### Study model

The marsh harrier *Circus aeruginosus* is a semi-colonial, ground-nesting raptor that mainly breeds in wetlands and hunts in open habitat (Cramp & Simmons, 1994). During winter, individuals roost communally on the ground. Northern populations are migratory, while populations from the Mediterranean areas are resident. Main wintering ranges extend from the north of the Mediterranean Basin to sub-Saharan Africa (Cramp & Simmons, 1994). The European marsh harrier population underwent a dramatic population decline from 1960 to 1980 due to wetland drainage, contamination and direct persecution, but has recovered in recent decades, during which time the species has undergone a moderate increase (ca. 10%; Cramp & Simmons, 1994; Bird-Life International, 2004). This positive tendency also occurs at smaller scales, where some populations such as that in Spain (Molina & Martínez, 2008) have doubled its breeding numbers during the last decades, increasing from 481–552 breeding pairs in 1990 to 1149–1494 in 2006. At least part of this increase may represent population recovery following reductions in organochlorine use and direct persecution in recent decades (Newton, 2004), and moreover, the species has benefitted from anthropogenic changes in habitat

composition related to agriculture (Cardador *et al.*, 2011). In contrast to breeding numbers, the area occupied by this expanding population has remained nearly unchanged, with around 15% increment in those years (Fig. 1). Little is known about whether the absence of more sites with suitable habitat for the species or pure distance effects are limiting this range expansion. The expanding marsh harrier population thus provides an opportunity to identify some of the ecological factors involved in limiting range expansions of highly mobile species, such as raptors, in the current scenario of global change.

## Occurrence data

The study area covered peninsular Spain (493,770 km<sup>2</sup>). Occurrence data of marsh harriers during the breeding and wintering periods were obtained from a national census performed in 2006 (SEO/Birdlife 2006). In this census, marsh harriers were surveyed by standardized nest searching and detection of territorial pairs in potential breeding areas during the breeding period, and detection of active roosting sites (those used by at least one marsh harrier for roosting at night) during winter (Molina & Martínez, 2008). In total, 1278 nests/territories and 308 active roost sites were detected. The location of nests/territories and the central coordinates of roost sites were incorporated into a Geographic Information System (GIS) using a Universal Transverse Mercator (UTM) grid of 1-km<sup>2</sup>. Duplicate samples (i.e. two or more records within the same grid cell) were removed and handled as single observations. Resulting sample sizes were 837 for the breeding season and 284 for winter. Given the equivalent census methodology used in both seasons and the high

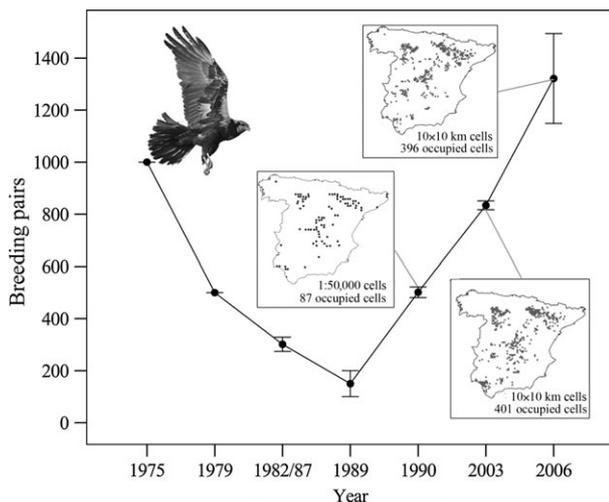
detectability of the species (Molina & Martínez, 2008), observed differences in number and distribution of breeding territories and winter roosting sites are not expected to be related with sampling bias. Nevertheless, to assess the potential influence of sample size differences between periods in our results, a subsample of 284 locations were randomly selected from the breeding period dataset to re-analyse our data (results provided in Table S2 and Fig. S3 and S5 in Supporting Information).

## Environmental predictors

We used four climatic variables commonly used in species distribution modelling to represent the climatic controls, namely mean annual temperature ( $T_{AN}$ ), minimum temperature of the coldest month ( $T_{MIN}$ ), maximum temperature of the hottest month ( $T_{MAX}$ ) and annual precipitation ( $P_{AN}$ ) (see Fig. S1 in Supporting Information). These variables were obtained with a resolution of 180 m from the 'Atlas Climático Digital de la Península Ibérica' (Ninyerola *et al.*, 2005). After performing an initial correlation analysis, we eliminated  $T_{AN}$  from modelling owing to their high correlation with  $T_{MAX}$  and  $T_{MIN}$  (Pearson correlation coefficient,  $r \geq 0.78$ ; see Table S1 in Supporting Information).

Habitat variables representing the environmental variability of the area were slope (SLO, as an indicator of landscape roughness), percentage of aquatic habitat (AQ, as an indicator of the nesting/roosting habitat of the species) and percentage of open vegetation (VEG, as an indicator of the foraging habitat). SLO was derived from the WorldClim 1.4 (Hijmans *et al.*, 2005) digital elevation model with a resolution of 1 km, while AQ and VEG were derived from remote sensing with a resolution of 250 m. AQ was obtained by pooling two original categories (wetlands and water bodies, except the subcategory sea and oceans) of the CORINE land cover 2006 (European Environment Agency, 2007). VEG was obtained by pooling seven original categories (i.e. arable land, pastures, natural grasslands, wetlands, moors and heathlands, scrub with sclerophyllous vegetation and sparsely vegetated areas) of the CORINE land cover 2006.

All layers were reprojected to the Universal Transverse Mercator projection (UTM) and resampled to a 1-km<sup>2</sup> grid size using the bilinear resampling technique to match the spatial resolution of occurrence data. However, previous information has shown that the pattern of habitat selection of marsh harriers is highly influenced by habitat characteristics at the scale of home ranges (Cardador *et al.*, 2011). Thus, we used focal statistics to transform the value of each variable in the 1 × 1 km cells to the mean value of that variable in a square of 10 × 10 km centred on the 1 × 1 km cells. We used this focal scale based on empirical data as well as analysis and results from previous studies (Cardador *et al.*, 2009; Cardador & Mañosa, 2011). As resulting variables were highly concordant to original 1 × 1 km variables (Pearson correlation coefficient range: 0.7–0.98, all  $P$ -values < 0.001; see also Fig. S2 in Supporting Information for comparison



**Figure 1** Evolution of the breeding population of the marsh harrier in Spain from 1975 to 2006. Data were compiled from various sources (revised data from 1975 to 2003: Jubete, 2003; data on 1990: Martínez *et al.*, 1996; data on 2006: Molina & Martínez, 2008). Mean, maximum and minimum estimates are shown. Maps with available historical and actual distributions of the breeding population are also shown.

on model performance), we chose the scale based on the earlier studies for subsequent analyses. All GIS operations were conducted in ArcGIS 9.3 (ESRI, Redlands, CA, USA).

### Spatial predictors

To account for non-environmental spatial constraints on the marsh harrier distribution, we used spatial variables obtained through eigenvector mapping (hereafter called *spatial filters*; Griffith & Peres-Neto, 2006; De Marco *et al.*, 2010). This method assumes that the spatial arrangement of all points of the study area (i.e. sample locations or, as in our study case, the regular grid cells of the whole study area) can be translated into a set of predictor variables (for a detailed explanation on computation, see Appendix S1 in Supporting Information), which capture spatial effects at different spatial scales (Diniz-Filho & Bini, 2005; Václavík *et al.*, 2012). The inclusion of spatial filters in the models allowed us to account for the effect of subjacent spatial structures that were not captured by the environmental factors considered (De Marco *et al.*, 2010).

### Distribution modelling

We built models to estimate the probability of occurrence of marsh harriers using generalized linear models (GLMs) and the maximum entropy modelling approach (Maxent, Phillips *et al.*, 2006; Phillips & Dudík, 2008). GLM is a parametric statistical approach that generalizes classical linear regression models, allowing for the modelling of non-normal response variables (Guisan *et al.*, 2002). We used the logistic variant (Guisan *et al.*, 2002; Václavík *et al.*, 2012) with a backward stepwise procedure. To calibrate our models, we selected a random subset of 10,000 pixels from the overall study area, in which the species is not known to reside, to use as pseudo-absences (Roura-Pascual *et al.*, 2009). GLMs were implemented in R software (R Development Core Team, 2011). In contrast, Maxent is a machine-learning method that uses a mathematical formulation to estimate the probability distribution of a species following the principle of maximum entropy, which assumes that no unfounded constraints are included in the estimation (Phillips *et al.*, 2006). Maxent does not require absence data points for the modelled distribution; instead, the probability distribution is defined only on species presences (Phillips *et al.*, 2006). For developing the model, Maxent creates random samples of background pixels (10,000) from the study area as pseudo-absences. Maxent models were implemented in MAXENT 3.1 software (Phillips *et al.*, 2006). We selected 500 iterations for model convergence and employed the regularization procedure to prevent overfitting (Phillips & Dudík, 2008).

To address our questions, we followed a hierarchical approach and ran GLM and Maxent with seven models based on different combinations of the predictors, namely (1) a climatic model that included single climatic predictors (CLIM); (2) a habitat model that included single habitat predictors

(HAB); (3) a spatial model that included single spatial filters (SPAT); (4) a full environmental model that included all the climatic and habitat predictors (CLIM + HAB); (5) a climatic and spatial model that included both climatic predictors and spatial filters (CLIM + SPAT); (6) a habitat and spatial model that included both habitat and spatial filters (HAB + SPAT); and (7) an environment and spatial model that included all environmental predictors and spatial filters (CLIM + HAB + SPAT). Both the linear and quadratic response types of the predictors were tested.

### Statistical analysis

Model performance was assessed by dividing the species occurrence data into random training (70%) and test (30%) datasets. A given model was calibrated on the training data and evaluated on the test data using the area under the receiver operating characteristics curve (AUC) as a threshold-independent assessment measure (Phillips *et al.*, 2006). We also calculated sensitivity (percentage of correctly classified presences) and specificity (percentage of correctly classified pseudo-absences) of each model. For these calculations, we used the probability value that maximized the sum of sensitivity and specificity as the threshold point (Liu *et al.*, 2005). AUC and threshold assessment for GLMs were implemented in R software by the functions 'somers2' and 'optim.thresh' from the 'Hmisc' and 'SDMTools' libraries.

Additionally, we used the Bayesian information criterion (BIC) and sample size-corrected Akaike information criterion (AICc) as measures of model fit based on the training data, to check the balance between goodness-of-fit and model complexity (Akaike, 1974; Burnham & Anderson, 2002). Functions for calculating AICc and BIC for Maxent are available in ENMTools (Warren *et al.*, 2010). AICc and BIC analyses for GLMs were implemented in R software.

To reduce uncertainty caused by sampling artefacts, we conducted 15 replicate models for each predictor set and modelling technique. We used nonparametric Mann–Whitney U-tests with a Bonferroni correction to assess how model performance of GLM and Maxent models varied between predictor sets. Further, we evaluated the independent contribution of individual predictors to model performance using a jackknife procedure. This procedure entailed the calculation of incremental improvement in performance of a model with a particular predictor compared with the equivalent model without that predictor. We also calculated model performance for each predictor when used in isolation. The estimated contributions were based on AUC. Model comparisons were performed using SPSS 15.0 (SPSS Inc, Chicago, IL, USA).

### Spatial filters and distance to previously occupied areas

We tested whether the significant effect of spatial filters on model performance found for the marsh harrier distribution

in the breeding period (see Results) could represent dispersal limitations from previously occupied areas, that is areas occupied in previous years (De Marco *et al.*, 2010). If so, it should be expected that filters reduced the probability of occurrence of marsh harriers in environmentally suitable areas (i.e. those predicted as presence by our CLIM + HAB models) most distant from areas occupied in previous years, while increasing the probability of occurrence in non-suitable areas (i.e., those predicted as absence by our CLIM + HAB models) closer to previously occupied areas. We assessed this hypothesis using published historical distribution data of the breeding marsh harrier population in peninsular Spain (distribution of probable secure breeding sites from 2003; Jubete, 2003) obtained at a resolution of 100-km<sup>2</sup> cells. To perform these analyses, we first reclassified the continuous logistic model predictions obtained in our CLIM + HAB and CLIM + HAB + SPAT models into binary maps (i.e. presence/absence) using the value that maximized the sum of sensitivity and specificity as a threshold point (Liu *et al.*, 2005) and resampled them to a 100-km<sup>2</sup> grid size using the majority resampling technique. As results obtained from GLM and Maxent models were highly in accordance (see Results), we only used Maxent outputs to conduct these analyses. Analyses were conducted by using nonparametric Mann–Whitney U-tests with SPSS 15.0 (SPSS Inc.).

## RESULTS

### Environmental models

#### *Environmental models in the breeding season*

Average values of AUC of climate (CLIM) and habitat (HAB) models in the breeding season ranged between 0.74 and 0.83, according to GLM and Maxent models, indicating overall good ability to predict the distribution of the species in the breeding season (Table 1; for comparison with results obtained according to a random subsample of the breeding period dataset, see Table S2 in Supporting Information). Among climate variables, mean annual precipitation showed the highest model performance when used in isolation, followed by minimum temperature of the coldest month (Fig. 2a). Annual precipitation also decreased the AUC the most when it was omitted from the CLIM model. This suggests that this variable shows the largest unique contribution among the studied climate variables (Fig. 2b). Probability of occurrence increased in areas with lower annual precipitation and intermediate minimum temperatures (Fig. 3). Among habitat variables, slope and percentage of open vegetation showed both the highest univariate explanatory power and the highest independent contribution to the HAB model (Fig. 2a,c). Higher probability of occurrence appeared in areas with flat terrains (lower slopes) and high percentages of open vegetation (Fig. 3).

The complete environmental model (CLIM + HAB) performed better than models including only climatic or habitat

variables according to AUC, AICc and BIC values of both GLM and Maxent models (but see some different results according to specificity and sensitivity values, Table 1). However, total unique contribution of habitat variables to these models was higher than that of climatic variables (Table 1, Fig. 2d).

#### *Environmental models in winter*

Similar to the breeding season, averaged values of AUC of climate and habitat models ranged between 0.74 and 0.81 in winter, according to GLM and Maxent models (Table 1), indicating overall good ability to predict the distribution of the species during this period. Among climate variables, mean annual precipitation showed both the highest model performance when used in isolation and the highest independent contribution to the CLIM model, followed by minimum temperature of the coldest month (Fig. 4a,b). Probability of occurrence increased in areas with lower annual precipitation and higher minimum temperatures (Fig. 3). Among habitat variables, slope followed by the percentage of aquatic vegetation had both the highest model performance when used in isolation and the highest independent contribution in the HAB model during this period (Fig. 4a,c). Probability of occurrence increased in areas with flat terrains (lower slopes), intermediate percentages of aquatic vegetation and high percentages of open habitat (Fig. 3).

The complete environmental model (CLIM + HAB) performed better than models including only climatic variables (for all measures of model accuracy and model fit, except sensitivity of Maxent models, Table 1). In contrast, additional effects of climatic variables on the habitat model were very small and statistically significant only according to AICc of both Maxent and GLMs and specificity of Maxent (Table 1, Fig. 2d).

### Spatial filters

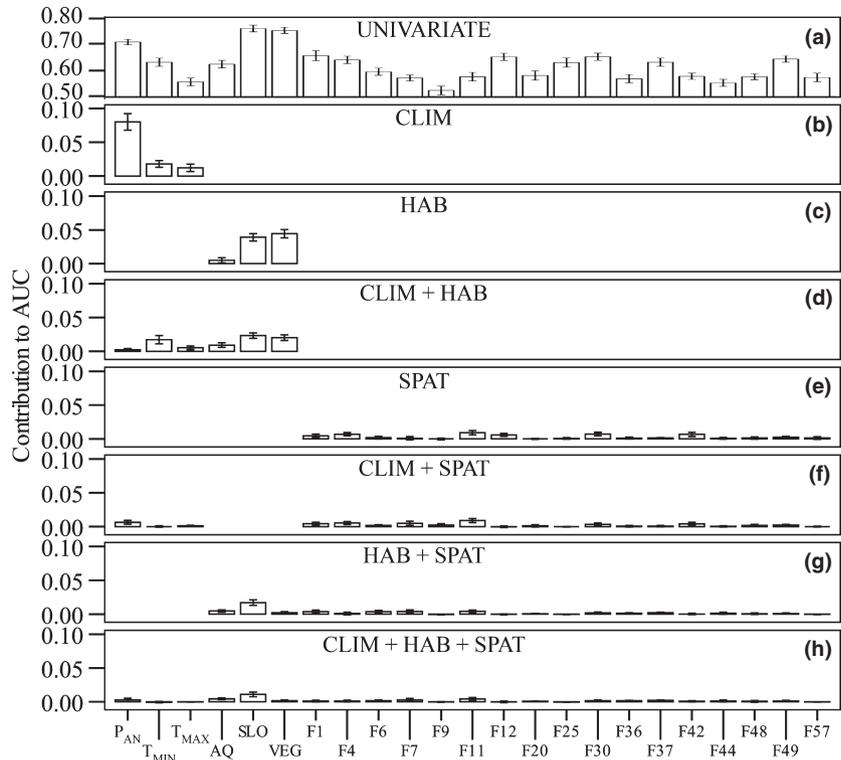
In total, 17 filters were selected for the breeding season (Fig. 5a; see results obtained according to a random subsample of the breeding period dataset in Fig. S3 in Supporting Information) and 19 for the wintering season (Fig. 5b) to describe spatial variability in marsh harrier occurrence data. The spatial complexity and variable shape of correlograms showed how selected filters reflected different spatial structures at different spatial scales (see Fig. S4 in Supporting Information for the map pattern of selected filters). Overall, filters selected for the breeding season had higher eigenvalues than those for the wintering season, meaning that aggregation patterns of the species in the breeding season occur at broader spatial scales than in winter (for more detailed explanation on selected filters see Appendix S1 in Supporting Information).

Averaged AUC values for single spatial models in the breeding and wintering season ranged between 0.80 and 0.86 (Table 1). None of the filters had a substantially greater

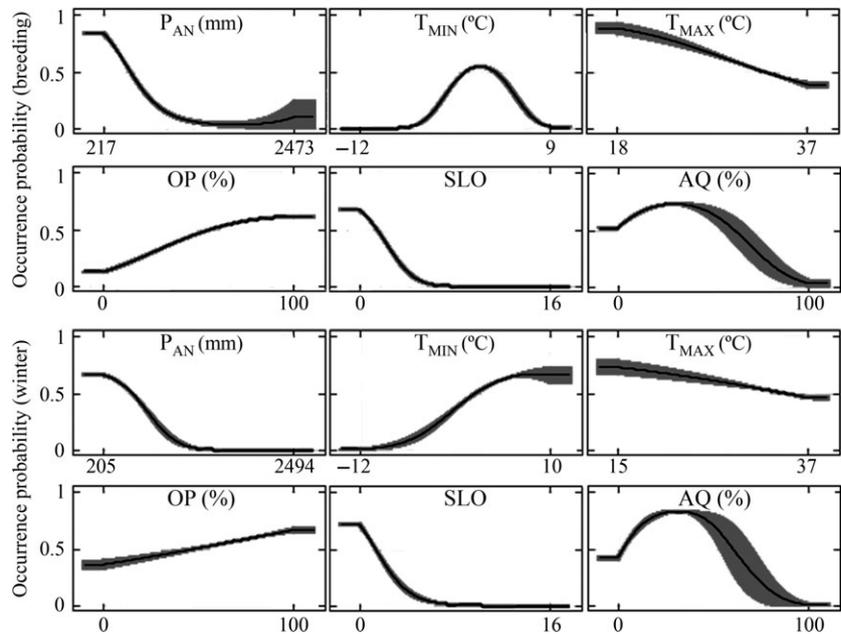
**Table 1** Differences in model performance based on different sets of predictors for the breeding and wintering periods using different measures of predictivity and model fit (AUC, AICc, BIC, sensitivity and specificity) and modelling techniques (GLM and Maxent). Note that for each set of predictors and season, 15 replicate models with different subsets of total data were conducted. The mean  $\pm$  SD of the 15 replicate models conducted for each predictor set are shown. Comparisons are based on Mann–Whitney *U*-test.

	CLIM	HAB	CLIM + HAB	SPAT	CLIM + SPAT	HAB + SPAT	CLIM + HAB + SPAT
<b>Breeding</b>							
<i>GLM</i>							
AUC	0.75 $\pm$ 0.01	0.83 $\pm$ 0.01	0.85 $\pm$ 0.01	0.86 $\pm$ 0.01	0.88 $\pm$ 0.01	<b>0.90 <math>\pm</math> 0.01<sup>a</sup></b>	<b>0.90 <math>\pm</math> 0.01<sup>a</sup></b>
AICc	3695 $\pm$ 22	3334 $\pm$ 23	3182 $\pm$ 31	3051 $\pm$ 28	2924 $\pm$ 29	2725 $\pm$ 30 <sup>a</sup>	2702 $\pm$ 29 <sup>a</sup>
BIC	3741 $\pm$ 23	3378 $\pm$ 21	3258 $\pm$ 30 <sup>a</sup>	3240 $\pm$ 24 <sup>a</sup>	3153 $\pm$ 24	2945 $\pm$ 30 <sup>b</sup>	2952 $\pm$ 29 <sup>b</sup>
Sensitivity	0.72 $\pm$ 0.04	<b>0.83 <math>\pm</math> 0.03<sup>a</sup></b>	<b>0.83 <math>\pm</math> 0.04<sup>a</sup></b>	<b>0.80 <math>\pm</math> 0.07<sup>ab</sup></b>	0.79 $\pm$ 0.04 <sup>b</sup>	<b>0.80 <math>\pm</math> 0.04<sup>ab</sup></b>	<b>0.80 <math>\pm</math> 0.04<sup>ab</sup></b>
Specificity	0.68 $\pm$ 0.02 <sup>a</sup>	0.69 $\pm$ 0.03 <sup>ab</sup>	0.74 $\pm$ 0.02 <sup>c</sup>	0.75 $\pm$ 0.06 <sup>bc</sup>	0.80 $\pm$ 0.03	<b>0.84 <math>\pm</math> 0.03<sup>d</sup></b>	<b>0.83 <math>\pm</math> 0.03<sup>d</sup></b>
<i>Maxent</i>							
AUC	0.74 $\pm$ 0.01	0.80 $\pm$ 0.01	0.83 $\pm$ 0.01 <sup>a</sup>	0.82 $\pm$ 0.01 <sup>a</sup>	0.85 $\pm$ 0.01	<b>0.87 <math>\pm</math> 0.01<sup>b</sup></b>	<b>0.87 <math>\pm</math> 0.01<sup>b</sup></b>
AICc	14,897 $\pm$ 34	14,697 $\pm$ 45	14,577 $\pm$ 69	14,361 $\pm$ 35	14,249 $\pm$ 36	14,167 $\pm$ 30	14,133 $\pm$ 34
BIC	14,922 $\pm$ 35	14,723 $\pm$ 45	14,622 $\pm$ 68	14,503 $\pm$ 34	14,410 $\pm$ 36	14,333 $\pm$ 29 <sup>b</sup>	14,317 $\pm$ 35 <sup>b</sup>
Sensitivity	0.69 $\pm$ 0.03 <sup>a</sup>	0.78 $\pm$ 0.04 <sup>bc</sup>	<b>0.82 <math>\pm</math> 0.03<sup>d</sup></b>	0.71 $\pm$ 0.08 <sup>ac,e</sup>	0.77 $\pm$ 0.06 <sup>b,e,f</sup>	<b>0.82 <math>\pm</math> 0.04<sup>b,d,f</sup></b>	<b>0.81 <math>\pm</math> 0.05<sup>d,f</sup></b>
Specificity	0.72 $\pm$ 0.01 <sup>a</sup>	0.72 $\pm$ 0.03 <sup>a</sup>	0.73 $\pm$ 0.02 <sup>a</sup>	<b>0.78 <math>\pm</math> 0.07<sup>ab</sup></b>	<b>0.80 <math>\pm</math> 0.04<sup>b</sup></b>	<b>0.80 <math>\pm</math> 0.02<sup>b</sup></b>	<b>0.81 <math>\pm</math> 0.04<sup>b</sup></b>
<b>Winter</b>							
<i>GLM</i>							
AUC	0.74 $\pm$ 0.02	0.81 $\pm$ 0.02 <sup>ab</sup>	<b>0.83 <math>\pm</math> 0.02<sup>abc</sup></b>	0.80 $\pm$ 0.02 <sup>b</sup>	0.82 $\pm$ 0.02 <sup>a</sup>	<b>0.85 <math>\pm</math> 0.02<sup>c</sup></b>	<b>0.85 <math>\pm</math> 0.02<sup>c</sup></b>
AICc	1687 $\pm$ 11	1538 $\pm$ 15 <sup>a</sup>	1506 $\pm$ 16	1536 $\pm$ 22 <sup>a</sup>	1483 $\pm$ 24	1416 $\pm$ 24 <sup>b</sup>	1403 $\pm$ 26 <sup>b</sup>
BIC	1718 $\pm$ 12 <sup>a</sup>	<b>1574 <math>\pm</math> 16<sup>b</sup></b>	<b>1574 <math>\pm</math> 18<sup>b</sup></b>	1723 $\pm$ 28 <sup>a</sup>	1669 $\pm$ 30	1609 $\pm$ 30 <sup>c</sup>	1618 $\pm$ 34 <sup>c</sup>
Sensitivity	<b>0.81 <math>\pm</math> 0.05<sup>a</sup></b>	0.75 $\pm$ 0.07 <sup>bc</sup>	0.74 $\pm$ 0.04 <sup>b</sup>	0.65 $\pm$ 0.06 <sup>d</sup>	0.69 $\pm$ 0.05 <sup>cd</sup>	<b>0.76 <math>\pm</math> 0.04<sup>ab</sup></b>	0.75 $\pm$ 0.05 <sup>b</sup>
Specificity	0.59 $\pm$ 0.02	0.74 $\pm$ 0.07 <sup>a</sup>	<b>0.79 <math>\pm</math> 0.02<sup>ab</sup></b>	<b>0.79 <math>\pm</math> 0.03<sup>ab</sup></b>	<b>0.79 <math>\pm</math> 0.03<sup>ab</sup></b>	<b>0.80 <math>\pm</math> 0.02<sup>b</sup></b>	<b>0.81 <math>\pm</math> 0.04<sup>b</sup></b>
<i>Maxent</i>							
AUC	0.74 $\pm$ 0.02	0.81 $\pm$ 0.02 <sup>a</sup>	<b>0.83 <math>\pm</math> 0.02<sup>ab</sup></b>	0.82 $\pm$ 0.02 <sup>a</sup>	0.81 $\pm$ 0.02 <sup>a</sup>	<b>0.85 <math>\pm</math> 0.02<sup>b</sup></b>	<b>0.85 <math>\pm</math> 0.02<sup>b</sup></b>
AICc	5123 $\pm$ 8	5009 $\pm$ 19	4987 $\pm$ 14 <sup>a</sup>	5047 $\pm$ 22	4985 $\pm$ 18 <sup>a</sup>	4942 $\pm$ 27 <sup>b</sup>	4938 $\pm$ 20 <sup>b</sup>
BIC	5138 $\pm$ 8 <sup>a</sup>	<b>5027 <math>\pm</math> 19<sup>b</sup></b>	<b>5021 <math>\pm</math> 13<sup>b</sup></b>	5152 $\pm$ 23 <sup>a</sup>	5098 $\pm$ 18	5057 $\pm$ 27 <sup>c</sup>	5062 $\pm$ 21 <sup>c</sup>
Sensitivity	<b>0.76 <math>\pm</math> 0.07<sup>ab</sup></b>	<b>0.74 <math>\pm</math> 0.07<sup>abc</sup></b>	<b>0.76 <math>\pm</math> 0.04<sup>a</sup></b>	0.68 $\pm$ 0.06 <sup>c</sup>	0.69 $\pm$ 0.07 <sup>bc</sup>	<b>0.73 <math>\pm</math> 0.07<sup>abc</sup></b>	<b>0.78 <math>\pm</math> 0.06<sup>a</sup></b>
Specificity	0.64 $\pm$ 0.04	0.74 $\pm$ 0.05 <sup>a</sup>	<b>0.79 <math>\pm</math> 0.01<sup>bc</sup></b>	0.78 $\pm$ 0.04 <sup>ab</sup>	<b>0.80 <math>\pm</math> 0.04<sup>bc</sup></b>	<b>0.82 <math>\pm</math> 0.03<sup>c</sup></b>	0.78 $\pm$ 0.04 <sup>ab</sup>

Letters indicate models that are not significantly different after Bonferroni corrections (i.e. at  $\alpha \leq 0.01$ ). Best models are shown in bold.



**Figure 2** Performance of environmental and climate variables in univariate models (a) and independent contribution (b–h) of individual predictors to multivariate models using different combinations of predictor variables in the breeding season. Mean predictor contributions and their standard deviations are calculated based on 15 replicate runs. The model contributions are based on AUC values from Maxent.



**Figure 3** Partial response curves illustrating the relationships between probability of occurrence of marsh harriers and our set of environmental predictors in the CLIM and HAB models for the breeding season and winter. These curves show how the shape of the response changes for a particular predictor while all other predictors are held at their mean sample value. Mean response curve of the 15 replicate Maxent runs (black) and standard deviation (grey) are shown.

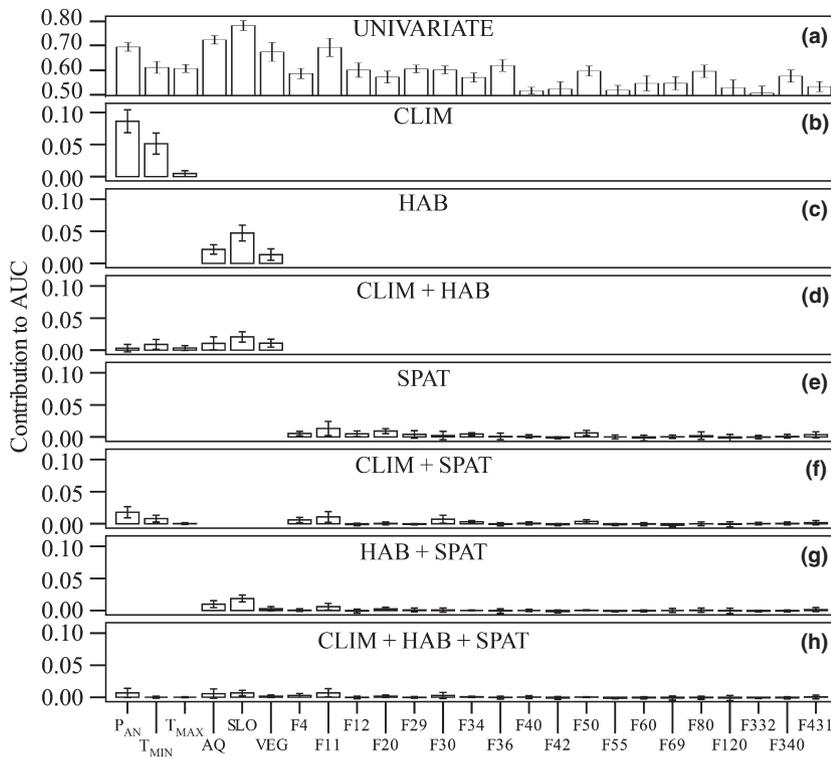
independent contribution to model performance (Figs 2e & 4e), and most of the model performance was related to their joint contributions.

**Spatial and environmental models**

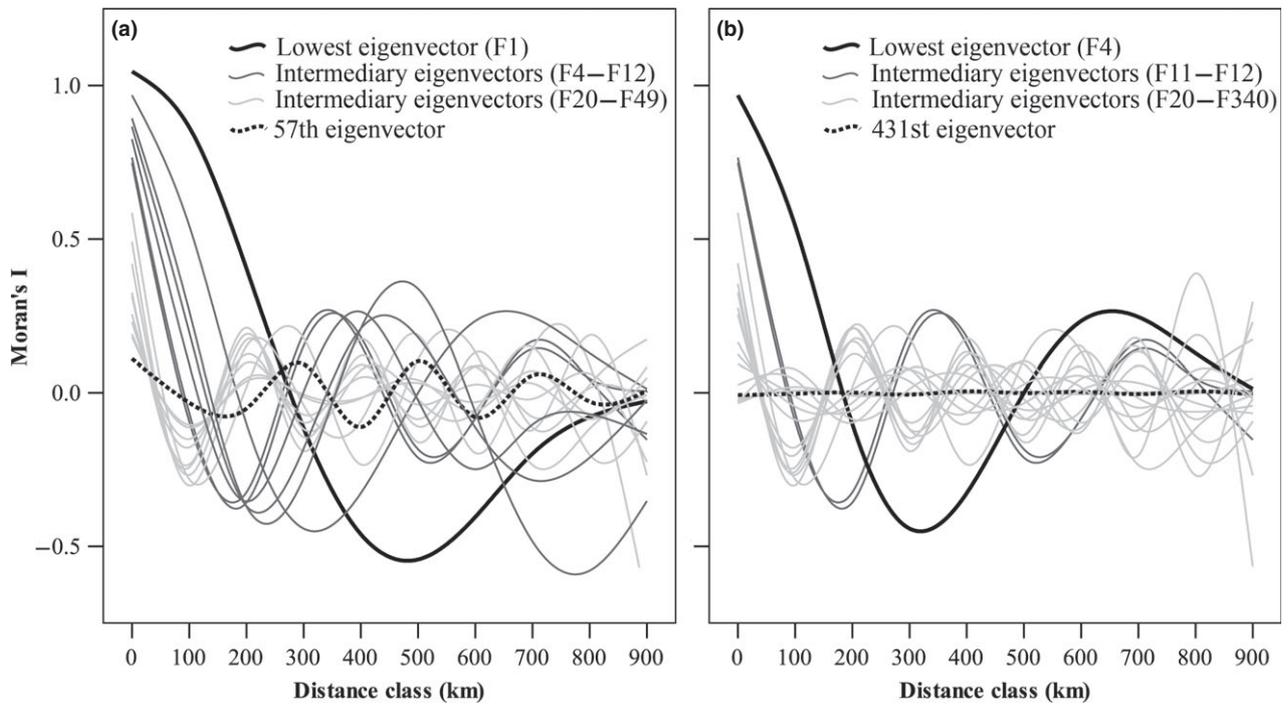
*Spatial filters and environmental models in the breeding season*

Environmental variables (i.e. climate and habitat variables) and spatial filters were both important predictors of marsh harrier

distribution during the breeding period. The CLIM + SPAT and HAB + SPAT models performed better than single CLIM and HAB models, respectively (for all measures of AUC, AICc, BIC and specificity of both GLM and Maxent models, but see some non-significant differences according to sensitivity values, Table 1; for comparison with results obtained according to a random subsample of the breeding period dataset, see Table S2 in Supporting Information). Filters independently contributed 12–15% to the CLIM + SPAT model according to AUC of Maxent and GLMs (i.e. they decreased 12–15% AUC when



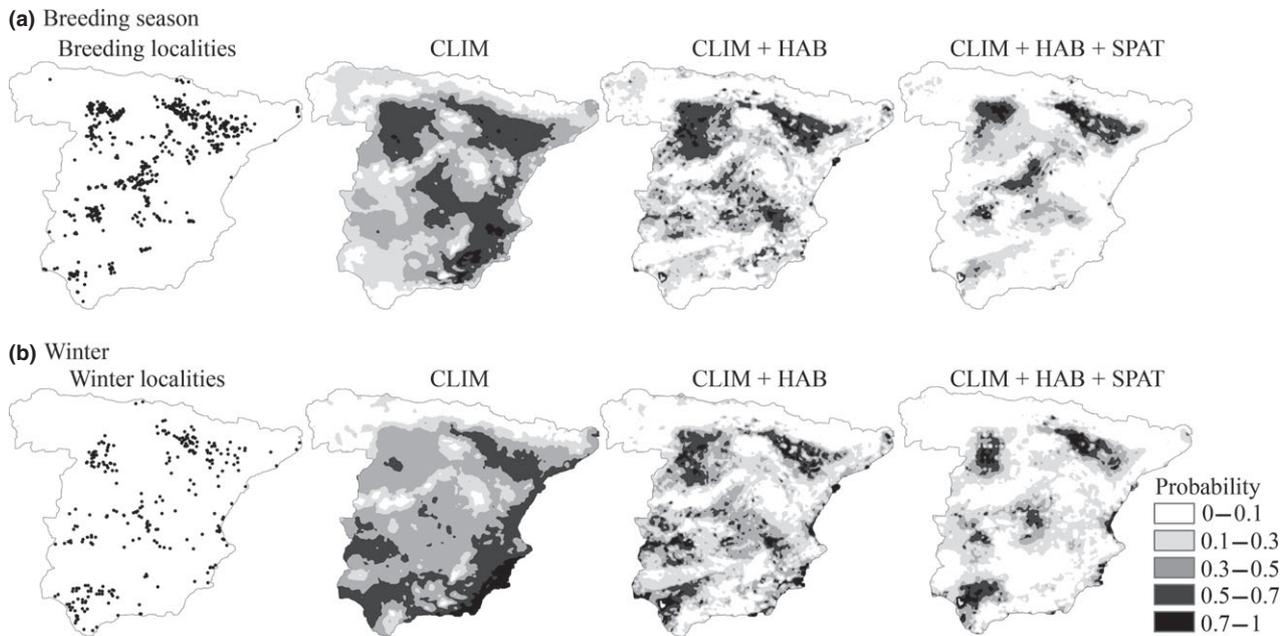
**Figure 4** Performance of environmental and climate variables in univariate models (a) and independent contribution (b–h) of individual predictors to multivariate models using different combinations of predictor variables in winter. Mean predictor contributions and their standard deviations are calculated based on 15 replicate runs. The model contributions are based on AUC values from Maxent.



**Figure 5** Spatial correlograms of spatial filters selected for the breeding (a) and winter (b) distribution of marsh harriers in peninsular Spain. Correlograms were defined by Moran's I coefficients in 10 distance classes, indicating links among points of the study area successively separated by 100 km. Filters for the breeding season are F1, F4, F6, F7, F9, F11, F12, F20, F25, F30, F36, F37, F42, F44, F48, F49, F57. Filters for the winter season are F4, F11, F12, F20, F29, F30, F34, F36, F40, F42, F50, F55, F60, F69, F80, F120, F332, F340, F431; an increased number in the name of the filters indicates subsequently lower eigenvalues.

omitted, Table 1). Filters independently contributed 8% to the HAB + SPAT model according to AUC of both GLM and Maxent. None of the filters had substantially greater independent contribution than the other filters to these models (Fig. 2f,g).

The CLIM + HAB + SPAT model also showed superior performance compared to the CLIM + HAB model (for all measures of AUC, AICc, BIC and specificity of both GLM and Maxent models, Table 1). However, non-significant differences



**Figure 6** Observed and averaged predicted distributions of the marsh harrier in Spain in the breeding (a) and wintering (b) periods. Predicted distributions are based on Maxent models using different sets of predictors: climate (CLIM model), climate and habitat predictors (CLIM + HAB model) or climate plus habitat and spatial filters (CLIM + HAB + SPAT model). Note that models developed for each set of predictors were calibrated using 15 different randomly selected subsamples of total data (averaged predictions are shown).

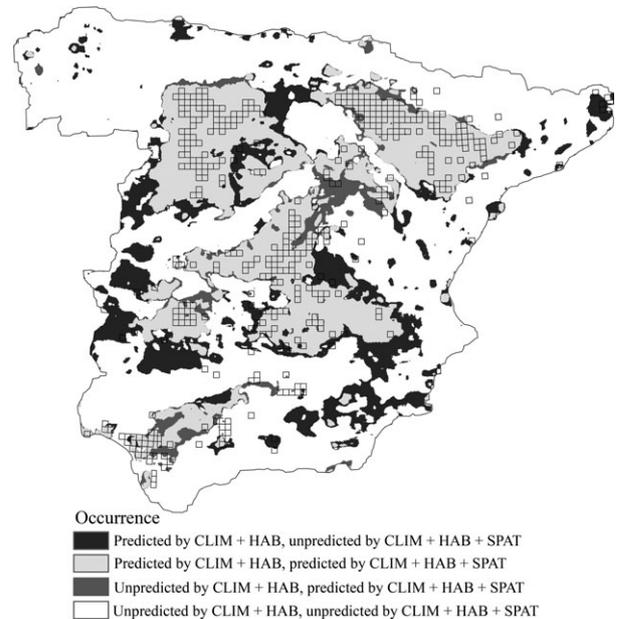
were found according to sensitivity values, meaning that increases in model performance were mainly due to the reduction in estimated probabilities of occurrence in areas where the species was not detected. Filters independently contributed 5–6% to this model according to AUC of GLM and Maxent, with none of the filters having substantially greater independent contribution than the other filters (Fig. 2h).

When compared to the previous distribution of the species, we observed that the inclusion of spatial filters in models decreased the probability of occurrence of the species in areas predicted as suitable by the CLIM + HAB model that were further away from previously occupied localities. In contrast, filters increased the probability of occurrence of the species in unsuitable areas close to its previous distribution (Figs 6–8; see mapped predicted distributions according to a random subsample of the breeding period dataset in Fig. S5 in Supporting Information).

#### *Spatial filters and environmental models in winter*

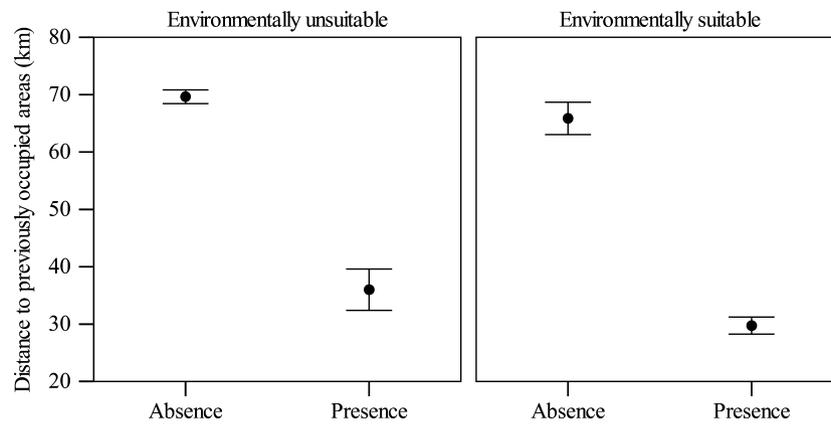
Inclusion of spatial filters also improved model performance of single climatic and habitat models in winter according to AUC, AICc, BIC and specificity of both GLM and Maxent models (but see some different results according to sensitivity values, Table 1). Filters independently contributed 13% to CLIM + SPAT according to AUC of both GLM and Maxent. Filters independently contributed 5% to HAB + SPAT according to AUC of both GLM and Maxent.

Contrary to the breeding season, spatial filters seemed to have a negligible impact on marsh harrier distribution in



**Figure 7** Predicted presence/absence of marsh harriers in peninsular Spain in 2006 during the breeding season based on Maxent models using different sets of predictors: climatic and habitat predictors (CLIM + HAB model) or climate and habitat plus spatial filters (CLIM + HAB + SPAT model). Presence of the species in 10 × 10 km UTM cells during the period 1985–2002 (Jubete. 2003) is shown.

winter when combined with both climatic and environmental predictors. AUC, sensitivity, specificity and BIC values supported models without spatial filters during this period



**Figure 8** Differences in distance to previously occupied localities between environmentally suitable UTM  $10 \times 10$  km cells predicted as absences or presences by CLIM + HAB + SPAT model (Mann–Whitney  $U$ -test,  $U = 165,394$ ,  $n = 1903$ ,  $P < 0.001$ ) and environmentally unsuitable UTM  $10 \times 10$  km cells predicted as presences or absences by these models (CLIM + HAB + SPAT model), ( $U = 126,755$ ,  $n = 3258$ ,  $P < 0.001$ ). Environmental suitability was based on presences predicted by Maxent models run with environmental predictors alone (CLIM + HAB model). Mean  $\pm$  95% IC is given.

(Table 1). Only according to AICc values did models including filters perform significantly better than the CLIM + HAB model (Table 1).

## DISCUSSION

When modelling spatial distribution of species, it is often assumed that species are in equilibrium with the environment (Guisan & Thuiller, 2005). The establishment of individuals in the most suitable areas in terms of climate and habitat heterogeneity is particularly expected for highly mobile species, such as raptors, which are expected to track the geographical position of their environmental envelope over long distances (Guisan & Thuiller, 2005; Mateo-Tomás & Olea, 2010). Indeed, the potential effect of spatial constraints and their intra-annual variations when modelling the distribution of these species at medium and large scales has been seldom studied (Carrete *et al.*, 2007; Mateo-Tomás & Olea, 2010). By analysing unique survey data of an expanding population of marsh harriers in peninsular Spain, we found evidence that distribution of at least some highly mobile species can be spatially constrained beyond environmental variables at large scales and that the relative importance of spatial constraints can vary between periods of the annual cycle.

The spatial distribution of marsh harriers across peninsular Spain was largely explained by environmental variables (climate and habitat) during both the breeding and the wintering seasons. Among these variables, habitat variables independently contributed more to environmental models (CLIM + HAB model) than climate, showing that the explanatory power of habitat variables might be important at broad scales (Luoto *et al.*, 2007). Although the good fit of a model does not necessarily imply correct inference of causation (James & McCulloch, 1990), our explanatory models

suggest that the most suitable areas for this species, in both the breeding and wintering seasons, are flat terrains (lower slopes) with high percentages of open vegetation, which is consistent with the previously described dependence of the species on these habitats (Cramp & Simmons, 1994). Surface of aquatic vegetation also seems to be important for the species, particularly during winter, probably because the presence of aquatic areas is important to allow marsh harriers to roost communally during this season (Cramp & Simmons, 1994). Unlike in other regions in Europe (Bibby & Lunn, 1982; Virkkala *et al.*, 2005), marsh harriers often breed solitarily and occupy very small aquatic areas in the study area during the breeding season, some of them are undetectable at the spatial resolution of our analyses (Cardador *et al.*, 2011), or even breed in non-aquatic habitats, such as cereal fields (Molina & Martínez, 2008).

Contrary to the equilibrium postulate (i.e. the species are in equilibrium with the environment, Guisan & Thuiller, 2005), the inclusion of spatial filters (Diniz-Filho & Bini, 2005; De Marco *et al.*, 2010) in our models provides evidence that marsh harrier distribution is spatially constrained beyond environmental variables during the breeding season at broad spatial scales. The significant effect of spatial filters during this season seems consistent with the idea of a limited dispersal from areas occupied in previous years as, in general, they reduced the probability of occurrence of the species in distant but otherwise suitable areas (i.e. those predicted by the environmental model). Limited dispersal from source areas could be related to the high philopatry and short natal dispersal distances described for the species (around 1–2 km, in a sedentary population of marsh harriers in France, Sternalski *et al.*, 2008), perhaps associated with social factors, such as conspecific attraction, which are known to affect colonial species (Serrano *et al.*, 2004).

Aggregation patterns of roosting locations of the marsh harrier in winter occurred at finer scales than that of breeding locations and seem to be mainly explained by environmental variables. Indeed, contrary to the breeding season, spatial filters seemed to have a negligible impact on marsh harrier distribution in winter when combined with both climatic and environmental predictors. It should be noted that all AUC, sensitivity, specificity and BIC values support models without spatial filters during this period, and the reduction in AICc values in the set of models that included spatial filters could be a consequence of the known tendency of this criterion to select more complex models when sample sizes are large (Burnham & Anderson, 2004). Negligible effects of spatial constraints in this season may be related to the arrival of migrant birds from central and northern Europe during winter, joining the local resident population (Molina & Martínez, 2008). Migrants can move great distances (Strandberg *et al.*, 2008) and may be able to reach suitable patches where breeders from local populations, which show high site fidelity throughout the year (Cardador *et al.*, 2009), are not present. This could be the case, for example, in suitable areas of eastern peninsular Spain (one of the most overpredicted areas by the CLIM + HAB model during the breeding season), which holds large populations of marsh harriers during winter. At these roost sites, at least some individuals are of a European breeding origin, based on ring recoveries (Molina & Martínez, 2008). In fact, migration routes described for northern populations covered the eastern (i.e. Mediterranean) region of peninsular Spain, ensuring accessibility of migrant birds to these areas (Strandberg *et al.*, 2008). On the other hand, censuses at roost sites during winter may also allow us to detect non-breeder dispersers from local populations, which may move long distances and have a different geographical distribution with respect to the breeding population, as in other birds of prey (Real & Mañosa, 2001). More specific studies on ranging behaviour of resident and migrant populations throughout the year are necessary to determine whether and to what extent spatial filters may be reflecting actual differences in spatial distribution between local and migrant birds and/or temporal movements of individuals from the local population.

Understanding the factors that limit current distributions of species and populations is crucially important in building effective management and conservation strategies. Our results show that environmental preferences may not necessarily result in complete habitat matching if species are constrained by dispersal limitation (e.g. Blach-Overgaard *et al.*, 2010; Pinto & MacDougall, 2010). Raptors are long-lived species that may move over large areas for several years before entering the breeding population (Newton, 1979). At that time, that is when settling for reproduction, individuals must optimize their fitness by increasing their foraging, mating, breeding and survival prospects. Settling near natal areas or close to conspecifics minimizes the costs of gaining appropriate knowledge of breeding sites and foraging areas while increasing mating opportunities (Arroyo *et al.*, 2001;

Serrano *et al.*, 2005). This could be occurring with marsh harriers in peninsular Spain, a demographic mechanism that impedes or delays the colonization of empty suitable patches located at great distances from previously occupied areas (Cardador *et al.*, 2012). We also propose that the effect of spatial constraints, at least in some bird species, might vary depending on the period of their annual cycle, especially in migratory species. Thus, large-scale conservation programmes should consider how events throughout the annual cycle interact to shape current species distributions.

## ACKNOWLEDGEMENTS

SEO/BirdLife provided census data of the breeding and wintering populations in 2006. F. Jubete and F. Martínez provided data on the breeding distribution in 1990. J.L. Tella, R. Heikkinen and two anonymous referees provided valuable comments on the manuscript. Original photo of marsh harrier was provided by Marcos Lacasa. S. Young revised the English. M.C. was supported by a Ramón y Cajal research contract from the Ministerio de Ciencia e Innovación (RYC-2009-04860). L.C. was supported during part of this study by the project 'Steppeahead' financed by FGCSIC and Banco Santander.

## REFERENCES

- Akaike, H. (1974) A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, **19**, 716–723.
- Anadón, J.D., Giménez, A., Martínez, M., Martínez, J., Pérez, I. & Esteve, M.A. (2006) Factors determining the distribution of the spur-thighed tortoise *Testudo graeca* in south-east Spain: a hierarchical approach. *Ecography*, **29**, 339–346.
- Arroyo, B., Mougeot, F. & Bretagnolle, V. (2001) Colonial breeding and nest defence in Montagu's Harrier (*Circus pygargus*). *Behavioral Ecology and Sociobiology*, **50**, 109–115.
- Austin, M. (2007) Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecological Modelling*, **200**, 1–19.
- Austin, M.P., Nicholls, A.O. & Margules, C.R. (1990) Measurement of the realized qualitative niche: environmental niches of five Eucalyptus Species. *Ecological Monographs*, **60**, 161–177.
- Bibby, C.J. & Lunn, J. (1982) Conservation of reed beds and their avifauna in England and Wales. *Biological Conservation*, **23**, 167–186.
- BirdLife International (2004) *Birds in Europe: population estimates, trends and conservation status*. BirdLife International (BirdLife Conservation Series No 12).
- Blach-Overgaard, A., Svenning, J.C., Dransfield, J., Greve, M. & Balslev, H. (2010) Determinants of palm species distributions across Africa: the relative roles of climate,

- non-climatic environmental factors, and spatial constraints. *Ecography*, **33**, 380–391.
- Burnham, K.P. & Anderson, D.R. (2002) *Model selection and multimodel inference: a practical information-theoretic approach*, 2nd edn. Springer-Verlag, Berlin, Germany.
- Burnham, K.P. & Anderson, D.R. (2004) Multimodel inference: understanding AIC and BIC in model selection. *Sociological Methods & Research*, **33**, 261–304.
- Cardador, L. & Mañosa, S. (2011) Foraging habitat use and selection of western Marsh-harriers (*Circus aeruginosus*) in intensive agricultural landscapes. *Journal of Raptor Research*, **45**, 168–173.
- Cardador, L., Mañosa, S., Varea, A. & Bertolero, A. (2009) Ranging behaviour of Marsh Harriers *Circus aeruginosus* in agricultural landscapes. *Ibis*, **151**, 766–770.
- Cardador, L., Carrete, M. & Mañosa, S. (2011) Can intensive agricultural landscapes favour some raptor species? The Marsh harrier in north-eastern Spain. *Animal Conservation*, **14**, 382–390.
- Cardador, L., Carrete, M. & Mañosa, S. (2012) Inter-individual variability and conspecific densities: consequences for population regulation and range expansion. *PLoS ONE*, **7**, e33375.
- Carrete, M., Grande, J.M., Tella, J.L., Sánchez-Zapata, J.A., Donazar, J.A., Díaz-Delgado, R. & Romo, A. (2007) Habitat, human pressure, and social behaviour: partialling out factors affecting territory extinction in the Egyptian vulture. *Biological Conservation*, **136**, 143–154.
- Carrete, M., Tella, J.L., Blanco, G. & Bertellotti, M. (2009) Effects of habitat degradation on the abundance, richness and diversity of raptors across Neotropical biomes. *Biological Conservation*, **142**, 2002–2011.
- Cortés-Avizanda, A., Almaraz, P., Carrete, M., Sánchez-Zapata, J.A., Delgado, A., Hiraldo, F. & Donazar, J.A. (2011) Spatial heterogeneity in resource distribution promotes facultative sociality in two trans-Saharan migratory birds. *PLoS ONE*, **6**, e21016.
- Cramp, S. & Simmons, K.E.L. (eds.) (1994) *Handbook of the birds of Europe the Middle East and North Africa. The birds of the Western Palearctic. Vol. II. Hawks to Bustards*. Oxford University Press, Oxford.
- De Marco, P. Jr, Diniz-Filho, J.A.F. & Bini, L.M. (2010) Spatial analysis improves species distribution modelling during range expansion. *Biology Letters*, **4**, 577–580.
- Diniz-Filho, J.A.F. & Bini, L.M. (2005) Modeling geographical patterns in species richness using eigenvector based spatial filters. *Global Ecology and Biogeography*, **14**, 177–185.
- Donald, P.F., Green, R.E. & Heath, M.F. (2001) Agricultural intensification and the collapse of Europe's farmland bird populations. *Proceedings of the Royal Society B: Biological Sciences*, **268**, 25–29.
- Dumbrell, A.J., Nelson, M., Helgason, T., Dytham, C. & Fitter, A.H. (2010) Relative roles of niche and neutral processes in structuring a soil microbial community. *ISME Journal*, **4**, 337–345.
- European Environment Agency (2007) *CLC2006 technical guidelines*. EEA Technical report No 17/2007. OPOCE, Luxembourg.
- Franklin, J. (1994) Predictive vegetation mapping: geographic modelling of biospatial patterns in relation to environmental gradient. *Progress in Physical Geography*, **19**, 474–499.
- Griffith, D.A. & Peres-Neto, P.R. (2006) Spatial modeling in ecology: the flexibility of eigenfunction spatial analyses. *Ecology*, **87**, 2603–2613.
- Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.
- Guisan, A., Edwards, T.C. & Hastie, T. (2002) Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling*, **157**, 89–100.
- Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G. & Jarvis, A. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.
- Huntley, B., Green, R.E., Collingham, Y.C. & Willis, S.G. (2007) *A climatic atlas of European breeding birds*. Durham University, The RSPB and Lynx Edicions, Barcelona.
- James, F.C. & McCulloch, C.E. (1990) Multivariate analysis in ecology and systematics: panacea or Pandora's box? *Annual Review in Ecology, Evolution, and Systematics*, **21**, 129–166.
- Jubete, F. (2003) Aguilucho lagunero occidental. *Atlas de las aves reproductoras de España* (ed. by R. Martí and J.C. del Moral), pp. 174–175. Dirección General de Conservación de la Naturaleza/Sociedad Española de Ornitología, Madrid.
- Lauzon-Guay, J.-S. & Scheibling, R.E. (2007) Seasonal variation in movement, aggregation and destructive grazing of the green sea urchin (*Strongylocentrotus droebachiensis*) in relation to wave action and sea temperature. *Marine Biology*, **151**, 2109–2118.
- Legendre, P. (2003) Spatial autocorrelation: trouble or new paradigm? *Ecology*, **74**, 1659–1673.
- Liu, C., Berry, P.M., Dawson, T.P. & Pearson, R.G. (2005) Selecting thresholds of occurrence in the prediction of species distributions. *Ecography*, **28**, 385–393.
- Luoto, M., Virkkala, R. & Heikkinen, R.K. (2007) The role of land cover in bioclimatic models depends on spatial resolution. *Global Ecology and Biogeography*, **16**, 34–42.
- Mackey, B.G. & Lindenmayer, D.B. (2001) Towards a hierarchical framework for modelling the spatial distribution of animals. *Journal of Biogeography*, **28**, 1147–1166.
- Martínez, F., Ortega, A. & Jubete, F. (1996) Situación actual del aguilucho lagunero (*Circus aeruginosus*) en España. Reproducción e invernada. *Biología y conservación de las rapaces mediterráneas, 1994* (ed. by J. Muntaner and J. Mayol), pp. 451–458. SEO/BirdLife, Madrid.
- Martínez-Morales, M.A., Zuria, I., Chapa-Vargas, L., MacGregor-Fors, I., Ortega-Álvarez, R., Romero-Águila, E. &

- Carbó, P. (2010) Current distribution and predicted geographic expansion of the Rufous-backed Robin in Mexico: a fading endemism? *Diversity and Distributions*, **16**, 786–797.
- Mateo-Tomás, P. & Olea, P.P. (2010) Anticipating knowledge to inform species management: predicting spatially explicit habitat suitability of a colonial vulture spreading its range. *PLoS ONE*, **5**, e12374.
- Molina, B. & Martínez, F. (2008) *El aguilucho lagunero en España. Población en 2006 y métodos de censo*. Seo/BirdLife, Madrid.
- Moore, R.P., Robinson, W.D., Lovette, I.J. & Robinson, T.R. (2008) Experimental evidence for extreme dispersal limitation in tropical forest birds. *Ecology Letters*, **11**, 960–968.
- Newton, I. (1979) *Population Ecology of Raptors*. T & AD Poyser, UK.
- Newton, I. (2004) The recent declines of farmland bird populations in Britain: an appraisal of causal factors and conservation actions. *Ibis*, **146**, 579–600.
- Ninyerola, M., Pons, X. & Roure, J.M. (2005) *Atlas climático digital de la Península Ibérica: metodología y aplicaciones en bioclimatología y geobotánica*. Universidad Autónoma de Barcelona, Barcelona.
- Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361–371.
- Phillips, S.J. & Dudík, M. (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. *Ecography*, **31**, 161–175.
- Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- Pinto, S.M. & MacDougall, A.S. (2010) Dispersal limitation and environmental structure interact to restrict the occupation of optimal habitat. *The American Naturalist*, **175**, 675–686.
- R Development Core Team (2011). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.
- Real, J. & Mañosa, S. (2001) Dispersal of juvenile and immature Bonelli's eagles in northeastern Spain. *Journal of Raptor Research*, **35**, 9–14.
- Roura-Pascual, N., Brotons, L., Peterson, A.T. & Thuiller, W. (2009) Consensual predictions of potential distributional areas for invasive species: a case study of Argentine ant in the Iberian Peninsula. *Biological Invasions*, **11**, 1017–1031.
- Sardà-Palomera, F., Puigcerver, M., Brotons, L. & Rodríguez-Teijeiro, J.D. (2012) Modelling seasonal changes in the distribution of Common quail *Coturnix coturnix* in farmland landscapes using remote sensing. *Ibis*, **154**, 703–713.
- Serrano, D., Forero, M.G., Donazar, J.A. & Tella, J.L. (2004) Dispersal and social attraction affect colony selection and dynamics of Lesser Kestrels. *Ecology*, **85**, 3438–3447.
- Serrano, D., Oro, D., Ursúa, E. & Tella, J.L. (2005) Colony size selection determines adult survival and dispersal preferences: Allee effects in a colonial bird. *The American Naturalist*, **166**, E22–E31.
- Snyder, R.E. & Chesson, P. (2003) Local dispersal can facilitate coexistence in the presence of permanent spatial heterogeneity. *Ecology Letters*, **6**, 301–309.
- Sternalski, A., Bavoux, C., Burneleau, G. & Bretagnolle, V. (2008) Philopatry and natal dispersal in a sedentary population of Western Marsh Harrier. *Journal of Zoology*, **274**, 188–197.
- Strandberg, R., Klaassen, R.H.G., Hake, M., Olofsson, P., Thorup, K. & Alerstam, T. (2008) Complex timing of Marsh Harrier *Circus aeruginosus* migration due to pre- and post-migratory movements. *Ardea*, **96**, 159–171.
- Václavík, T., Kupfer, J.A. & Meentemeyer, R.K. (2012) Accounting for multi-scale spatial autocorrelation improves performance of invasive species distribution modelling (iSDM). *Journal of Biogeography*, **39**, 42–55.
- Virkkala, R., Luoto, M., Heikkinen, R.H. & Leikola, N. (2005) Distribution patterns of boreal marshland birds: modelling the relationships to land cover and climate. *Journal of Biogeography*, **32**, 1957–1970.
- Warren, D.L., Glor, R.E. & Turelli, M. (2010) ENMTools: a toolbox for comparative studies of environmental niche models. *Ecography*, **33**, 607–611.
- Woinarski, J.C.Z., Whitehead, P.J., Bowman, D.M.J.S. & Russell-Smith, J. (1992) Conservation of mobile species in variable environment: the problem of reserve design in the Northern Territory, Australia. *Global Ecology and Biogeography Letters*, **2**, 1–10.
- Zimmermann, N.E., Thomas, C.E. Jr, Graham, C.H., Pearman, P.B. & Svenning, J. (2010) New trends in species distribution modelling. *Ecography*, **33**, 985–989.

## SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

**Appendix S1** Spatial predictors.

**Figure S1** Distribution of the six environmental variables.

**Figure S2** Comparison of 1 × 1 km and 10 × 10 km focal resolution modelling.

**Figure S3** Spatial correlograms of filters selected according to a random subsample of the breeding period dataset.

**Figure S4** Example of geographical patterns of spatial filters.

**Figure S5** Predicted distributions according to a random subsample of the breeding period dataset.

**Table S1** Correlation of the environmental variables.

**Table S2** Modelling results according to a random subsample of the breeding period dataset.

## BIOSKETCH

**Laura Cardador** is interested in ecological factors affecting spatial distribution and population dynamics of species, paying particular attention to human-induced environmental changes. Her work has focused on habitat selection models, demographic components, regulation processes, and intraspecific competition. This study is part of her PhD Thesis.

Author contributions: L.C., F.S-P., M.C. and S.M. conceived the ideas; L.C. and F.S-P. analysed the data; and L.C., F.S-P., M.C. and S.M. wrote the manuscript.

---

Editor: Risto Heikkinen