Diversity and Distributions, (Diversity Distrib.) (2011) 17, 1173–1185

A Journal of Conservation Biogeography
Diversity and Distributions
BIODIVERSITY RESEARCH

Finessing atlas data for species distribution models

Aidin Niamir[1,2]*, Andrew K. Skidmore[1,2], Albertus G. Toxopeus[1], Antonio R. Muñoz[3] and Raimundo Real[4]

[1]*Faculty of Geo-information Science and Earth Observation (ITC), University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands,* [2]*The C.T. de Wit Graduate School for Production Ecology and Resource Conservation, Wageningen University, Droevendaalsesteeg 4, 6708 PB Wageningen, The Netherlands,* [3]*Fundacion Migres, Huerta Grande, CN - 340 Km . 96 (Pelayo), 11390 Algeciras, Cadiz,* [4]*Faculty of Science, University of Malaga, E-29071 Malaga, Spain*

*Correspondence: Aidin Niamir, Faculty of Geo-information Science and Earth Observation, University of Twente, PO Box 217, 7500 AE Enschede, The Netherlands.
E-mail: niamir@itc.nl

ABSTRACT

Aim The spatial resolution of species atlases and therefore resulting model predictions are often too coarse for local applications. Collecting distribution data at a finer resolution for large numbers of species requires a comprehensive sampling effort, making it impractical and expensive. This study outlines the incorporation of existing knowledge into a conventional approach to predict the distribution of Bonelli's eagle (*Aquila fasciata*) at a resolution 100 times finer than available atlas data.

Location Malaga province, Andalusia, southern Spain.

Methods A Bayesian expert system was proposed to utilize the knowledge from distribution models to yield the probability of a species being recorded at a finer resolution (1 × 1 km) than the original atlas data (10 × 10 km). The recorded probability was then used as a weight vector to generate a sampling scheme from the species atlas to enhance the accuracy of the modelling procedure. The maximum entropy for species distribution modelling (MaxEnt) was used as the species distribution model. A comparison was made between the results of the MaxEnt using the enhanced and, the random sampling scheme, based on four groups of environmental variables: topographic, climatic, biological and anthropogenic.

Results The models with the sampling scheme enhanced by an expert system had a higher discriminative capacity than the baseline models. The downscaled (i.e. finer scale) species distribution maps using a hybrid MaxEnt/expert system approach were more specific to the nest locations and were more contrasted than those of the baseline model.

Main conclusions The proposed method is a feasible substitute for comprehensive field work. The approach developed in this study is applicable for predicting the distribution of Bonelli's eagle at a local scale from a national-level occurrence data set; however, the usefulness of this approach may be limited to well-known species.

Keywords
***Aquila fasciata*, Bayesian expert system, Bonelli's eagle, downscaling, Malaga, maximum entropy, sampling, Spain.**

INTRODUCTION

Species distribution modelling (SDM) is being widely used to predict the occurrence of species at locations where survey data are lacking (Guisan & Thuiller, 2005), which is the case for most of the earth's surface (Franklin, 2009). To define the suitability of a location for a species, SDM links data about a species distribution data to the environmental characteristics of those locations and then extrapolates the relationship over space or time (Oindo *et al.*, 2003; Elith & Leathwick, 2009). For the majority of species, data describing distributions are scarce and in the form of coarse resolution atlases (Newbold, 2010). Species atlases provide a convenient display, using near-equal areas (grids) on maps (Araujo *et al.*, 2005), of the extent of occurrence based on historical observations, museum records and often a complementary field survey. A species is marked as 'present' if there is at least one record of that species at a location within a (pre-defined) grid (Bierman *et al.*, 2010). Such grid maps are typically available in the form of systematic 10×10 km grid cells at a national level or in coarser resolutions at continental and global level, while application usually occurs on a local level of around one hundred hectares (Rouget, 2003).

Atlas data play an important role in conservation biogeography by providing species distribution data for developing new concepts, analytical approaches, and to address a range of conservation problems (Robertson *et al.*, 2010). Therefore, to generate a reliable fine-resolution distribution map, the gap between the available coarse resolution species occurrence data sets and the desired resolution has to be bridged (Hobbs, 2003). Developing an empirical approach to deal with uncertainties in the downscaling process remains a challenge (Boitani *et al.*, 1999; Barbosa *et al.*, 2003; Araujo *et al.*, 2005; Graham *et al.* 2008).

Projection to a finer resolution of statistical relationships calibrated at a coarse resolution is a conventional approach (Collingham *et al.*, 2000; Araujo *et al.*, 2005; Barbosa *et al.*, 2009, 2010). However, the extrapolation of models built for one resolution to a different resolution increases the uncertainty in the model predictions because they are extended beyond the model's original scale and error can vary among scales (Araujo *et al.*, 2005). Another downscaling approach investigated point sampling (Lloyd & Palmer, 1998; Barbosa *et al.*, 2003; Hartley *et al.*, 2003; McPherson *et al.*, 2006), where the species occurrence data set consisted of a random point within each coarse resolution grid cell (50×50 or 10×10 km) along with fine-resolution (1×1 km) environmental variables informing the SDM. A critical issue in this approach was the model's uncertainty, derived from prediction inconsistency over different random sampling iterations (Costa *et al.*, 2010).

Furthermore, the expense of an intensive field survey is often high and cannot be afforded (Skidmore & Turner, 1992; James *et al.*, 2001). Since there are insufficient field data available, expert knowledge could be an efficient source of information (Murray *et al.*, 2009). Translation of such knowledge into a rule-based script poses a challenge (Ferrier *et al.*, 2002) and has not been widely promoted (Carpenter, 2002). There have been concerns that prior knowledge (i.e. expert opinion) might drive the modelling process rather than guide it (Dennis, 1996). There are several examples that attempted to incorporate expert knowledge into SDM procedures (Pearce *et al.*, 2001; Ferrier *et al.*, 2002; Hobbs, 2003; Choy *et al.*, 2009; Bierman *et al.*, 2010), but this has not been adequately utilized in species data optimization and sampling strategies (Lehmann *et al.*, 2002). A species is recorded as present in an atlas grid if the database holds at least one record of that species from a location within that grid. The term 'recorded probability' is used to address the probability of a species being recorded at a location.

Bayesian inference is an accepted statistical tool among ecologists (Pereira & Itami, 1991; Dennis, 1996; Ellison, 2004; Clark, 2005; McCarthy, 2007). Bayes' theorem provides a clear method for estimating parameters and expressing the degree of confidence and uncertainty in those estimates. Bayesian expert systems (Lee *et al.*, 1987) have been defined to handle complex, real-world problems and attempt to solve problems by reasoning like an expert (Forsyth, 1984; Skidmore, 1989; Heikkinen & Hogmander, 1994). Expert systems used to combine diverse data sources with remotely sensed images to map soils (Skidmore, 1996), vegetation (Skidmore, 1989; Schmidt *et al.*, 2004; Wang *et al.*, 2009) and habitat characterization (Bierman *et al.*, 2010; Wang *et al.*, 2010).

This study outlines the incorporation of existing knowledge into a point-sampling approach to predict the distribution of Bonelli's eagle (*Aquila fasciata*) in Malaga province, southern Spain, at a resolution 100 times finer than atlas (modelled) data (Marti & Del Moral, 2003). An experiment was designed to evaluate the discriminative capacity of a Bayesian expert system model with an enhanced sampling scheme and four groups of environmental explanatory variables: topographic, climatic, biological and anthropogenic. This model (termed 'hybrid' model hereafter) consisted of a Bayesian expert system combined with a conventional SDM technique: maximum entropy species distribution modelling (MaxEnt) (Phillips *et al.*, 2006). The aim was to assess whether an expert system could improve the accuracy and robustness of predictions for local conservation applications by utilizing existing ecological knowledge, where empirical data were missing or difficult to obtain.

METHODS

Study area

The province of Malaga (7267 km^2) is a mountainous region situated in Andalusia, southern Spain, ranging in altitude from sea level along the Mediterranean shoreline to almost 2000 m. The climate is Mediterranean with mean annual rainfall ranging from 400 to 1200 mm and annual temperatures ranging from 12.6 to 19.2 °C (Font, 2000). The natural vegetation has been transformed in the valleys and lowlands to olive groves, cereal crops and coastal urbanization interspersed with small fragments of Mediterranean scrublands.

Species occurrence data

Bonelli's eagle is a resident species in the Malaga province, with only juvenile birds dispersing. Therefore, it was assumed that the presence of active nests indicated the presence of the species throughout the year. Occurrence data for Bonelli's eagle in 10×10 km grids for Malaga province were obtained from the Atlas of Spanish breeding birds (Marti & Del Moral, 2003) and updated with the Atlas of raptors of Malaga province (Jiménez & Muñoz, 2008). Malaga province is covered by 104 atlas grids (10×10 km), with Bonelli's eagle marked 'present' in 67 (70%) of the grids and 'absent' in the others (Fig. 1a), which indicates that the density for the species in Malaga province is the highest one known in Europe.

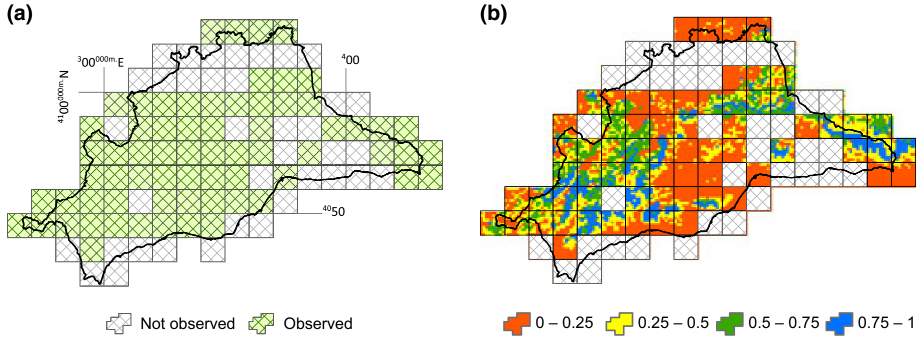

Figure 1 (a) Distribution of Bonelli's eagle (*Aquila faciata*) in Malaga, southern Spain, based on the Spanish Atlas of Breeding Birds (Marti & Del Moral, 2003) and Atlas of raptors of Malaga province (Jiménez & Muñoz, 2008). (b) Probability of a nest being recorded over presence atlas grids based on expert knowledge.

Three species sampling data sets were generated. The first data set termed the 'random points' data set consisted of 100 iterations of a random point in a coarse resolution grid (10 × 10 km), which was originally marked as 'present' in the atlas (n = 67). The second data set (termed the 'enhanced' data set) was generated based on the same procedure, except that the recorded probability (as quantified by the expert system) was used as a weighting factor. This meant that a grid cell with a higher recorded probability had a higher probability of being sampled. To test whether the developed SDMs differ significantly from what would be expected by chance, 1000 random points were generated over the study area (Malaga province) and used as inputs for the null models (Raes & ter Steege, 2007).

An independent evaluation data set was collected consisting of 80 nest locations (presences). Nest locations were provided by local ornithologists monitoring breeding territories from 2005 to 2009 in Malaga province. The absence points (n = 1000) were randomly generated using the atlas data (Marti & Del Moral, 2003) within cells where the species was not observed.

Bonelli's eagle (*Aquila fasciata*)

Bonelli's eagle is a medium to large-sized raptor with its western Palaearctic populations mainly distributed in the Mediterranean area (Del Hoyo *et al.*, 1994). The European population has suffered a marked decline and has been listed as an endangered European Species in recent decades (Rocamora, 1994; BirdLife-International 2000). Spain supports 730–800 breeding pairs (Real, 2003) or about 75% of the total European population. Malaga supports about 10% of the Spanish population of the species (Del Moral, 2006), considered one of the last strongholds of the Bonelli's eagle in Europe (Balbontín, 2005). Existing knowledge about favourable nest locations for Bonelli's eagle was elicited from literature and retrieved from the Web of Science® on 1 January 2010 through three key words: 'Bonelli's eagle', '*Hieraaetus fasciatus*' and '*Aquila fasciata*'. Table 1 lists the studies that describe the variables that are considered to be important for determining the distribution of the species. Disagreement among different sources was discussed with local experts, and their contribution has been evaluated subjectively. It was concluded that topographic attributes formed a critical factor in favourable nest locations for Bonelli's eagle. Experts consistently mentioned slope gradient and altitude as important topographic explanatory variables. Another critical factor frequently mentioned in the literature was formed by land cover attributes, e.g. distance to artificial surfaces and human disturbance.

Expert layers and rules

Using the ASTER elevation model (Abrams *et al.*, 2010), two topographic explanatory variables were compiled: elevation and slope irregularity. To generate the irregularity layer for slope (Skidmore, 1989) and elevation, the maximum values were subtracted from the mean within a 1 × 1 km grid, resulting in higher values representing higher irregularity and lower values describing homogenous topographic landscape. The CORINE [Coordination of Information on the Environment of the European Environmental Agency (EEA 2007)] land cover 1995 and 2006 maps (Bossard *et al.*, 2000) with a 250 × 250 m resolution were also downloaded and aggregated to a 1 × 1 km resolution using the Raster package (Hijmans & van Etten, 2010). Land cover and topographic irregularity were selected as environmental layers in the expert model, reflecting the fact that this species is geographically widespread in Malaga province with its fragmented bare rocks and high human disturbance. These three environmental explanatory variables (expert layers) were then used in the expert system. Since spatial correlation among (Montgomery & Peck, 1982) and within (Dormann *et al.*, 2007) predictors violates the assumption of independence of most standard statistical procedures (Legendre, 1993), the spatial structure of the expert layers was investigated using the variance inflation factor (VIF) as a measure of multicollinearity, and the Moran's I statistic as a measure of univariate spatial autocorrelation. A list of expert layers and their spatial structure is shown in Table 2.

Formulation of expert rules was based on the integration of knowledge from several sources, including the reviewed

Table 1 Details of studies that have been used for knowledge extraction on the habitat preferences of Bonelli's eagle (*Aquila fasciata*).

Species data	Region	Parameters	References
P	Cádiz, Spain	Topographic features, and human disturbance	Balbontín (2005)
P	Catalonia, Spain	Human disturbance	Bosch *et al.* (2010)
A	Spain	Climate, vegetation, and interspecific relationships	Carrascal & Seoane (2009)
P	Murcia, Spain	Prey availability	Carrete *et al.* (2002)
P	Granada, Spain	Distance to villages, topographic irregularity, and cultivation	Gil-Sánchez *et al.* (1996)
P	Granada, Spain	Human disturbance, interspecific relationships, and topographic variables	Gil-Sanchez *et al.* (2004)
P	Castellón, Spain	Cragy slopes, human disturbance, climate, and land use	Lopez-Lopez *et al.* (2006)
A	Valencian Community, Spain	Altitude, and slope	Lopez-Lopez *et al.* (2007a)
P	Castellón, Spain	Altitude	Lopez-Lopez *et al.* (2007b)
P	Alicante, Spain	Land cover	Martinez *et al.* (2008a)
P	Alicante and Murcia, Spain	Intra- and interspecific relationships	Martinez *et al.* (2008b)
SP, A	Western Europe	Prey availability, and local territorial features	Moleon *et al.* (2009)
A	Spain	Anthropogenic disturbances, temperature, and prey diversity	Moreno-Rueda *et al.* (2009)
A	Spain	Slope, temperature, and precipitation	Muñoz *et al.* (2005)
SP	Granada, Spain	Landscape features, and prey accessibility	Ontiveros & Pleguezuelos (2000)
SP	Granada, Spain	Land cover, and aspect	Ontiveros & Pleguezuelos (2003a)
SP	Western Mediterranean	Climatic constraints, and anthropogenic disturbances	Ontiveros & Pleguezuelos (2003b)
P	Andalucia, Spain	Land cover and, prey availability	Ontiveros *et al.* (2005)
P	Portugal	Prey accessibility, and landscape composition	Palma *et al.* (2006)
P	Valencia, Spain	Altitude	Rico *et al.* (1999)
SP	Spain	Anthropogenic disturbances	Soutullo *et al.* (2008)

SP, subpopulation; A, atlas/lattice; P, points.

literature, knowledge acquired through discussion with experienced local researchers as well as personal field observation. Where disagreement occurred between the different sources, a subjective decision was made based on field knowledge. The expert rules and their associated probabilities are listed in Table 2.

Explanatory variables

The environmental explanatory variable sets have a direct interaction with species and were chosen based on ecological theory (Austin, 2007). Usually, there is a wide variety of potential predictor variables to choose from, but the effects of making a choice between variables are poorly known (Synes & Osborne, 2011). In this case, the explanatory environmental variables were grouped into four categories (see Appendix S1 in Supporting Information): topographic, climatic, anthropogenic and biological. The basis for grouping was to create consistent variable sets for input to and comparison of the SDM predictions.

Topographic variables (TPG) consisted of elevation, slope and aspect along a north–south axis derived from the ASTER elevation model (Abrams *et al.*, 2010). For climatic variables (BCL), all 19 bioclimatic layers (1×1 km) (Hijmans *et al.*, 2005) were used. The approximate distance to roads, the approximate distance to urban areas and the land cover map comprised the anthropogenic variables (ANT). The most commonly used parameter for quantifying productivity and above-ground biomass of ecosystems is the Normalized Difference Vegetation Index (NDVI). For biological variables (BIO), atmospherically corrected SPOT4 and SPOT5 Vegetation Sensor were obtained from http://www.vgt.vito.be. The decade average (1998–2008) of the 10-day composite NDVI-images (S10 product) at 1 km^2 resolution was then calculated

Table 2 Variables and their corresponding rules included in the expert system to calculate the probability of recording a nest of Bonelli's eagle (*Aquila faciata*) in Malaga, southern Spain.

Variables	Moran's *I*	VIF	Evidence	$P(E\|H)$*
Elevation (m)† (Max−mean)	0.367	1.187	<100 m	0.3
			Between 100 and 200 m	0.5
			More than 200 m	0.9
Slope gradient (%)‡ (Max−mean)	0.483	1.271	<30	0.2
			30–45	0.3
			45–60	0.5
			More than 60	0.9
Land cover§	0.909	1.468	Continuous urban fabric	0.1
			Discontinuous urban fabric	0.1
			Industrial or commercial units	0.1
			Road and rail networks and associated land	0.1
			Port areas	0.1
			Airports	0.2
			Mineral extraction sites	0.2
			Construction sites	0.2
			Sport and leisure facilities	0.2
			Non-irrigated arable land	0.3
			Permanently irrigated land	0.2
			Vineyards	0.3
			Fruit trees and berry plantations	0.3
			Olive groves	0.3
			Annual crops associated with permanent crops	0.3
			Complex cultivation patterns	0.3
			Agriculture and significant natural vegetation	0.3
			Agro-forestry areas	0.4
			Broad-leaved forest	0.3
			Coniferous forest	0.4
			Mixed forest	0.4
			Natural grasslands	0.5
			Sclerophyllous vegetation	0.7
			Transitional woodland-shrub	0.3
			Bare rocks	0.9
			Burnt areas	0.5
			Water courses	0.1
			Water bodies	0.1

VIF, variance inflation factor.
*Probability that there is a piece of evidence given that a nest occurs.
†Abrams *et al.* (2010).
‡Initially computed at 30 × 30 m resolution from ASTER DEM (Abrams *et al.*, 2010).
§Aggregated from 0.25 × 0.25 to 1 × 1 km grid (EEA 2007).

(36 images). For more information on NDVI and faunal distribution, see Leyequien *et al.* (2007).

Bayesian expert system

Bayesian theory in SDM offers an alternative approach to statistical inference and differs from conventional frequentist inference in fundamental ways. Frequentist inference estimates the probability of the evidence (E) given hypothesis (H), while Bayesian inference estimates the probability that a hypothesis is true given evidence and defines it as the degree of belief in the likelihood of the evidence (Wade, 2000; Ellison, 2004). Here, a forward chaining expert system, originally developed by Skidmore (1989), was used to infer the posterior probability that a Bonelli's eagle nest occurs at a given cell based on the predictors and the expert rules. The forward chaining approach is essentially a data-driven approach (Naylor, 1984) and has been applied satisfactorily in remote sensing and image classification (Skidmore, 1996; Schmidt *et al.*, 2004; Wang *et al.*, 2010). Bayesian methods explicitly recognize and combine four components of knowledge: prior knowledge, data, model and posterior knowledge (McCarthy, 2007). In this study, prior knowledge from coarse resolution atlas data was combined with expert rules to estimate the recorded probability of a nest location at a finer resolution.

Let (N) be a nest occurring at location ($X_{i,j}$) and let (E_b) be an item of evidence (for $b = 1, ..., k$) known at location ($X_{i,j}$). Set a hypothesis (H) that a nest (N) occurs at location ($X_{i,j}$). A rule may be thus defined: given (E_b) then (H), that is, given a piece of evidence (E_b), then infer (H). However, there may be uncertainty associated with this rule (Forsyth, 1984), and the probability of the rule may not be 0 (i.e. false) or 1 (i.e. true), but rather lie between 0 and 1, depending on how 'sure' the experts are that the rule is true (Skidmore, 1989).

Bayes' theorem was used to update the probability of the rule that the nest of Bonelli's eagle (H) occurs at ($X_{i,j}$) given an environmental explanatory variable:

$$P(H|E_b) = \frac{P(E_b|H) \times P(H)}{P(E_b)} \tag{1}$$

where P ($E_b|H$) is the probability that there is a piece of evidence (E_b) (e.g. bare rocks) given that a nest occurs at location ($X_{i,j}$), also known as conditional probability that is based on expert rules (Table 2); P (H) is the probability for the hypothesis (H) that a nest occurs at location ($X_{i,j}$), initially obtained from the species atlas data (*a priori*). On iterating with further pieces of evidence, P ($E_b|H$):b = 1 replaces P (H) in equation 1. P (E_b) is the probability of the evidence alone:

$$P(E_b) = \sum_{b=1}^{n} P(E_b|H) \times P(H) \tag{2}$$

The evidence (E_b) must be spatially independent (Table 2); otherwise, P (E_b) would become larger or smaller, thereby decrementing or incrementing (H), causing the *posterior* probabilities to be incorrect.

Hybrid method

A hybrid method for species distribution was constructed, incorporating a Bayesian expert system and a machine learning model, to increase the discriminative capacity compared to methods without an expert system. Machine learning methods in contrast to deductive and knowledge-driven approaches include various kinds of algorithms implemented to learn the classification rules directly from data (Breiman, 2001). The maximum entropy model for species distribution modelling (MaxEnt) (Phillips *et al.*, 2006) has generated higher predictive accuracy than many other methods (Elith *et al.*, 2006; Hernandez *et al.*, 2006). MaxEnt also outperformed others where sampling was poor (Costa *et al.*, 2010), data were collected with sampling bias (Phillips *et al.*, 2009; Rebelo & Jones, 2010) and across different sample sizes (Wisz *et al.*, 2008). Therefore, the MaxEnt version 3.3.1 was employed here as distribution modelling core. For more information about MaxEnt and its statistical explanation, see Elith *et al.* (2011).

Evaluation

The PresenceAbsence package (Freeman & Moisen, 2008) in R 2.11.1 (R Development Core Team 2010) was used to assess the capacity of the downscaled models to discriminate between validation presence and absence points (validation data set) by analysing their receiver operation characteristic (ROC) curves. The area under this curve (AUC) provided a threshold-independent measure across all possible classification thresholds for each model (Fielding & Bell, 1997). Then, the R's wilcox.test function was used to perform a Mann–Whitney–Wilcoxon rank-sum test (termed Wilcoxon test hereafter) to check whether the hybrid model had a higher AUC value than the MaxEnt model. AUC combines specificity and sensitivity (Fielding & Bell, 1997); thus, it is not possible to explore whether the improvement in model performance has to do with reduction in commission or omission, or both errors. AUC ignores the goodness-of-fit of the predictions (Lobo *et al.*, 2008); therefore, it is possible that a poorly fitted model retains good discrimination power (Hosmer & Lameshow, 2000). To deal with this limitation, sensitivity and specificity of models were also calculated.

RESULTS

The probability of a nest being recorded (at the finest resolution of 1 × 1 km) was estimated using the Bayesian expert system. The mean of the posterior probabilities (μ = 0.24) was low. Two-thirds (76%) of all fine-resolution grids had a recorded probability of 0.5 or less. In line with expectation, the range of recorded probabilities was wide with a mean (95% CI) of 0.81 among presence coarse-resolution grids, indicating that there are few cells with a high and many cells with a low recorded probability in each presence atlas grid (Fig. 1b). Intersection of validation nest location with the recorded probability revealed that 30 (37.5%) of the nests were located in cells with a 0.75 or higher recorded probability, 47 (58.7%) were located in cells with a 0.5 or higher recorded probability and 8 (10%) in cells with a 0.1 or lower recorded probability. The average recorded probability for nest locations (presences) was 0.57 (with standard deviation 0.27).

The random, null-points and enhanced species data set along with four categories of environmental variables (TPG, BCL, ANT and BIO) were input to MaxEnt. The baseline models (built using the random points data set) were compared with the 'hybrid' models (built using the enhanced data set). The hybrid models had a higher mean AUC than the baseline models. However, not all models were equally improved by incorporating the expert system. The favourable nesting locations were accurately discriminated from unfavourable locations over the study area (Fig. 2). The mean AUC score in models with ANT variables increased the most by replacing the baseline with hybrid models: from 0.71 to 0.83 (Wilcoxon tests, $P < 0.001$); there was a 8% improvement in discrimination capacity of the models with TPG ($AUC_{Baseline}$ = 0.74 to AUC_{Hybrid} = 0.81, Wilcoxon tests, $P < 0.001$) and BIO ($AUC_{Baseline}$ = 0.75 to AUC_{Hybrid} = 0.82, Wilcoxon tests, $P < 0.01$), whereas the models with bioclimatic variables (BCL) only slightly increased the AUC score, from 0.75 to 0.79 (Wilcoxon tests, $P < 0.001$) (Table 3). The evaluation of the null models also revealed that all models are significantly different than what is expected by chance.

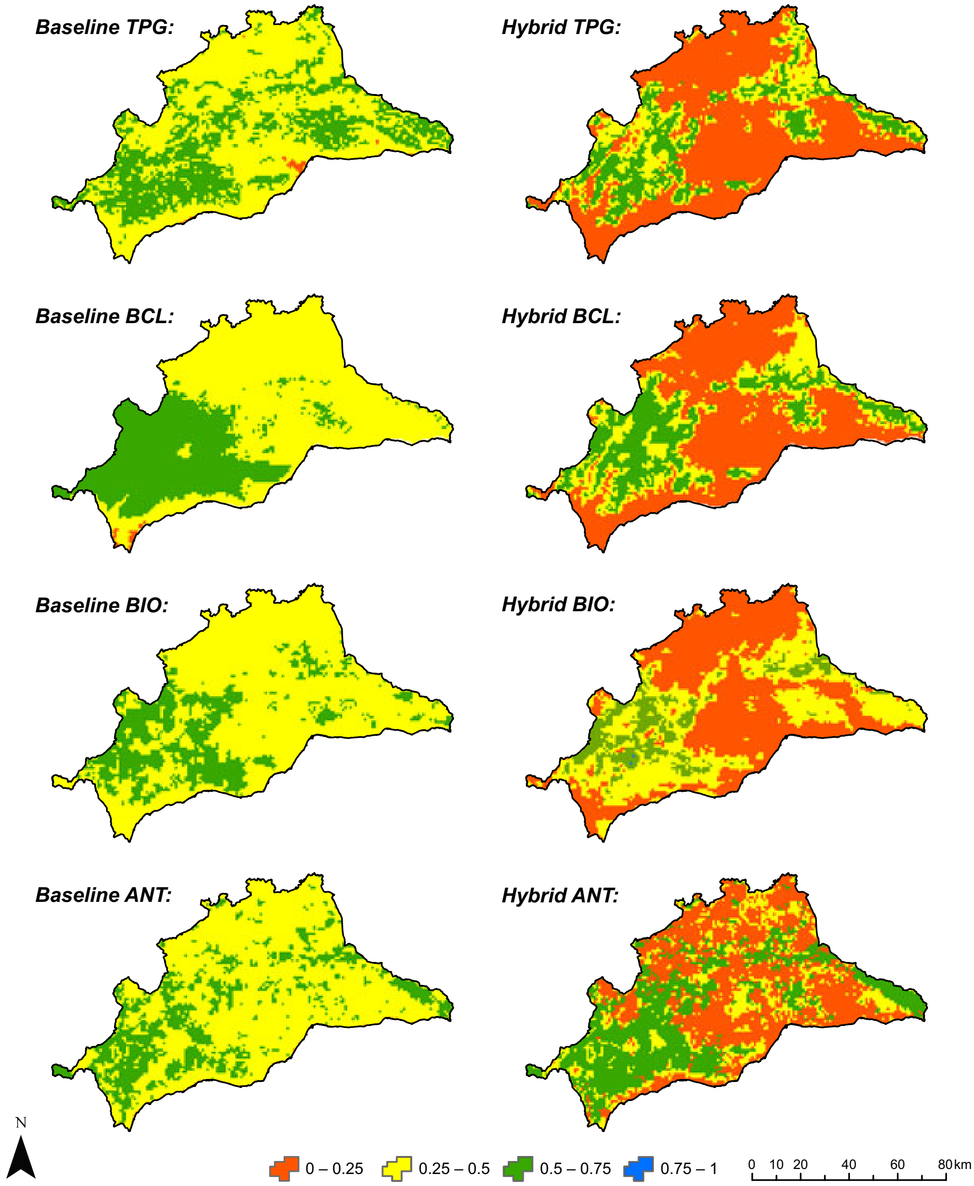

Figure 2 Comparison of fine resolution prediction maps generated by baseline (right column) and hybrid (left column) models using topographic (TPG), bioclimatic (BCL), biological (BIO) and anthropogenic (ANT) environmental explanatory variables for Bonelli's eagle (*Aquila faciata*) in Malaga, southern Spain.

Table 3 Evaluation of model performance. Mean and standard deviation of the area under the curve (AUC) scores over 100 iterations for null model (Raes & ter Steege, 2007), baseline (random sampling) and hybrid (Bayesian expert system) models. Significant levels are associated with the Wilcoxon's test.

	AUC			Standard deviation		
Model	Null model	Baseline	Hybrid	Null model	Baseline	Hybrid
Topographic	0.60	0.74	0.81*	0.078	0.114	0.027*
Bioclimatic	0.56	0.75	0.79*	0.047	0.019	0.019^{ns}
Biological	0.55	0.75	0.82†	0.039	0.045	0.015*
Anthropogenic	0.57	0.71	0.83*	0.058	0.072	0.019*

ns, not significant.
*$P < 0.001$.
†$P < 0.01$.

The mean of the probabilities (of favourable nest locations) was always significantly higher at 'presence' sites (i.e. $n = 80$ nest locations) than at 'absence' sites (Wilcoxon tests, for ANT, TPG and BIO: $P < 0.001$ and for BCL: $P < 0.01$). The proportion of presence validation points in areas with high probability (≥ 0.75) was always significantly higher than the proportion of localities available within those areas for both hybrid and baseline models. The proportion of presences in areas of low probability (≤ 0.25) was generally lower than that expected by chance (Wilcoxon tests, for ANT, TPG and BIO: $P < 0.01$ and for BCL: $P < 0.05$). The mean probability of occurrence values was also consistently higher at presence validation points and lower at absence validation points than at any of the random site samples (null points) for hybrid models (Wilcoxon tests, for presence: $P < 0.01$and for absence: $P < 0.05$).

The hybrid approach had considerably higher specificity (pseudo-absence location correctly predicted), whereas the sensitivity (nest location correctly predicted) was decreased at lower thresholds. The threshold probability where specificity is equal to sensitivity was lower in the hybrid than the baseline models. However, the proportion of the correctly predicted nest locations was higher at the above-mentioned threshold (Fig. 3). This is reflected in the predicted distribution maps. The least favourable locations (≤ 0.25) were discriminated from less favourable locations ($0.25 \leq f \leq 0.5$) in the hybrid approach (Fig. 2).

Uncertainties in the prediction (standard deviation of AUC) over 100 iterations were significantly smaller in hybrid models than in the baseline models (Table 3), revealing that the hybrid models were more robust in their discrimination.

DISCUSSION

For all environmental explanatory variables, the downscaled predictions based on the hybrid approach were significantly related to the location of survey nests and improved the discriminative capacity of the distribution models. This suggests that such a hybrid approach may be a valid way to incorporate existing knowledge into the modelling procedure in regions that are poorly known or where fine-resolution species information is too costly and time-consuming to gather.

The main aim of this paper was to introduce the Bayesian expert system as a promising alternative, incorporating existing knowledge into the analysis of species atlas data. Since the proposed approach enhances the quality of input species data, other SDMs than MaxEnt may also be employed for the prediction of species distribution. In fact, any parametric model may be suitable to build up a hybrid approach using our expert system. No *a priori* assumption was made as to the form of the relationship between the proxy variable for the expert layers and the recorded probability. However, it is possible to take into account the importance of environmental explanatory variables based on the expert knowledge by assigning weights to $P(E_b|H)$ (equation 2) and consequently updating $P(H|E_b)$ (equation 1). The methodology presented in this paper extends the methodology by Skidmore (1989) by including a species distribution model in which the recorded probability of the species is modelled as a function of environmental variables.

Higher discrimination capacity of the hybrid approach is because of its higher specificity rather than sensitivity. When

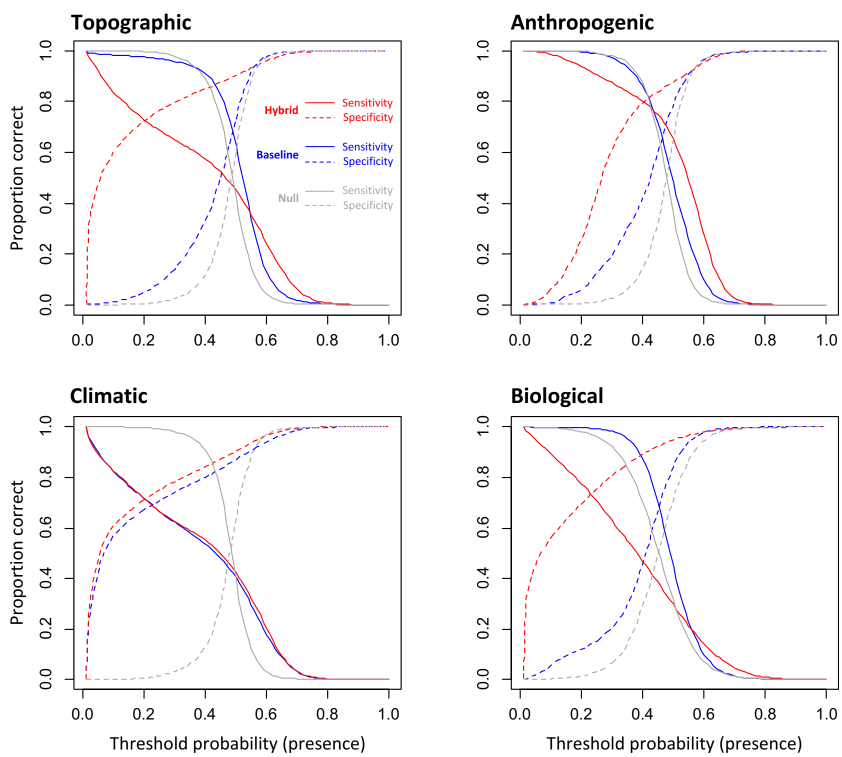

Figure 3 Variation of sensitivity (solid lines), the proportion of correctly predicted nest location versus specificity (dashed lines), the proportion of correctly predicted absences in baseline (blue) and hybrid (red) models, using four categories of environmental explanatory variables for Bonelli's eagle (*Aquila faciata*) in Malaga, southern Spain. Grey lines represent null models.

considering a 10×10 km atlas data grid, there are many favourable locations for a nest that might not be used for many reasons, e.g. the territorial behaviour of the species. Consequently, a model should be capable of discriminating unfavourable locations (even if they are biophysically favourable). A reasonable model would be expected to be not only sensitive but also specific to the nest locations. Sensitivity is equal to specificity (sensitivity and specificity lines cross each other) at a higher proportion value in the hybrid model and at a threshold closer to the prevalence of nest presence compared to the baseline model (Fig. 3).

Expert systems offer many advantages and some disadvantages over conventional statistical approaches for species distribution modelling. The major advantage is that existing knowledge about the species–environment relationship can be encapsulated into the modelling process. In this study, the inductive model (MaxEnt) produced the suitability for nest occurrence, based on random points within coarse-resolution grids. The known ecological relationship between environmental variables and the location of a nest yielded the most likely nest location (recorded probability). The expert system handled the uncertainty in these relationships through the use of probability (e.g. it is fairly certain that the Bonelli's eagle will nest on bare rock, but this may not always be the case). This is in contrast with crisp models (Estrada *et al.*, 2008), which try to describe the relationship in a binary format (i.e. does the nest of a Bonelli's eagle occur in sclerophyllous vegetation, yes or no?). Another advantage of expert systems is that expert judgment on the effect of scale can be quickly incorporated as the implicit relationship between data layers and the dependent variable being modelled becomes clear. For example, many bare rock pixels were aggregated with other dominant pixels within 1×1 km grids. Therefore, in this study, experts expressed their knowledge in the form of topographic irregularity to diminish the aggregation effects. This flexibility gives an expert the opportunity to build their own expert knowledge base, even manually drawing their ideas on a map using a graphical interface. Therefore, in such hybrid approaches, the ecological realism and acceptability to the user community may improve as well as the predictive performance.

An obvious disadvantage of expert systems for species distribution modelling is that experts may not agree among themselves, causing inconsistencies in the existing knowledge on the ecological factors affecting distribution of the target species. The 'robustness' of a recorded probability map generated by an expert system has to be gauged against the criteria defined by the expert or other source of knowledge, as well as against a validation data set. Another disadvantage of expert systems, as with all other modelling techniques, is that they do not respond well to incomplete knowledge or extrapolation into an area beyond the region of expertise (Murray *et al.*, 2009). In such a situation, the probabilities associated with the rules may have to be adjusted to better reflect the gap in knowledge, or the addition of explanatory variables may be required (Skidmore, 1989). As the results revealed, existing knowledge was more accurate in describing unfavourable locations for nests than favourable locations. Therefore, it may be beneficial to model the unsuitability, using absence points instead of presence points and then taking the inverse of the predicted distribution map (Lobo *et al.*, 2010). In other words, the presented hybrid model may be employed to incorporate expert knowledge on either suitability or unsuitability for a certain species into distribution models.

A number of studies have investigated the performance of inductive (data-driven) SDM at predicting distribution and species ecological characteristics (Segurado & Araujo, 2004; Elith *et al.*, 2006; Guisan *et al.*, 2007; Evangelista *et al.*, 2008), and it appeared that generalist species with a wide geographical range yielded models with a lower discriminative capacity than species with strict geographical boundaries (Buisson *et al.*, 2010). This is the stage that our hybrid approach will be most usefully applied. It is easier to discriminate favourable from unfavourable habitat for generalist species, by incorporating a deductive (knowledge-driven) method, thereby describing the ecological niche. The novel technique described here may not be successful for species that are not well known or minimally studied. However, all SDMs suffer the same restrictions.

ACKNOWLEDGEMENTS

This research was funded by the European Union Erasmus Mundus program External Cooperation Windows 2008/2441/001 MUN ECW 15-7-2008, and by the Spanish Ministry of Education and Science and FEDER (project CGL2009-11316, BOS subprogram). The authors thank the local ornithologists who kindly shared their knowledge and observations of the Bonelli's eagle in Andalucia, as well as the many volunteer fieldworkers who contributed to the Spanish atlas of breeding birds. Babak Naimi provided crucial tips on the use of R. The language editing was done by Eva Skidmore. Many thanks to the anonymous reviewers for their comments on a previous version of this paper.

REFERENCES

Abrams, M., Bailey, B., Tsu, H. & Hato, M. (2010) The ASTER Global DEM. *Photogrammetric Engineering and Remote Sensing*, **76**, 344–348.

Araujo, M.B., Thuiller, W., Williams, P.H. & Reginster, I. (2005) Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. *Global Ecology and Biogeography*, **14**, 17–30.

Austin, M. (2007) Species distribution models and ecological theory: a critical assessment and some possible new approaches. *Ecological Modelling*, **200**, 1–19.

Balbontín, J. (2005) Identifying suitable habitat for dispersal in Bonelli's eagle: an important issue in halting its decline in Europe. *Biological Conservation*, **126**, 74–83.

Barbosa, A.M., Real, R., Olivero, J. & Mario Vargas, J. (2003) Otter (*Lutra lutra*) distribution modeling at two resolution scales suited to conservation planning in the Iberian Peninsula. *Biological Conservation*, **114**, 377–387.

Barbosa, A.M., Real, R. & Mario Vargas, J. (2009) Transferability of environmental favourability models in geographic space: the case of the Iberian desman (*Galemys pyrenaicus*) in Portugal and Spain. *Ecological Modelling*, **220**, 747–754.

Barbosa, A.M., Real, R. & Vargas, J.M. (2010) Use of Coarse-Resolution Models of Species' Distributions to Guide Local Conservation Inferences. *Conservation Biology*, **24**, 1378–1387.

Bierman, S.M., Butler, A., Marion, G. & Kühn, I. (2010) Bayesian image restoration models for combining expert knowledge on recording activity with species distribution data. *Ecography*, **33**, 451–460.

BirdLife-International (2000) *European bird populations: estimates and trends*. BirdLife International (BirdLife Conservation Series no. 10), Cambridge, UK.

Boitani, L., Corsi, F., Biase, A.D., Carranza, I.D.I., Ravagli, M., Reggiani, G., Sinibaldi, I. & Trapanese, P. (1999) *A databank for the conservation and management of the African mammals*. Institute of Applied Ecology (IEA), Roma.

Bosch, R., Real, J., Tinto, A., Zozaya, E.L. & Castell, C. (2010) Home-ranges and patterns of spatial use in territorial Bonelli's Eagles *Aquila fasciata*. *Ibis*, **152**, 105–117.

Bossard, M., Feranec, J. & Otahel, J. (2000) *CORINE land cover technical guide – Addendum 2000*. European Environment Agency, Copenhagen.

Breiman, L. (2001) Statistical modeling: the two cultures. *Statistical Science*, **16**, 199–231.

Buisson, L., Thuiller, W., Casajus, N., Lek, S. & Grenouillet, G. (2010) Uncertainty in ensemble forecasting of species distribution. *Global Change Biology*, **16**, 1145–1157.

Carpenter, S.R. (2002) Ecological futures: building an ecology of the long now. *Ecology*, **83**, 2069–2083.

Carrascal, L.M. & Seoane, J. (2009) Factors affecting large-scale distribution of the Bonelli's eagle *Aquila fasciata* in Spain. *Ecological Research*, **24**, 565–573.

Carrete, M., Sanchez-Zapata, J.A., Martinez, J.E., Sanchez, M.A. & Calvo, J.F. (2002) Factors influencing the decline of a Bonelli's eagle *Hieraaetus fasciatus* population in southeastern Spain: demography, habitat or competition? *Biodiversity and Conservation*, **11**, 975–985.

Choy, S.L., O'Leary, R. & Mengersen, K. (2009) Elicitation by design in ecology: using expert opinion to inform priors for Bayesian statistical models. *Ecology*, **90**, 265–277.

Clark, J.S. (2005) Why environmental scientists are becoming Bayesians. *Ecology Letters*, **8**, 2–14.

Collingham, Y.C., Wadsworth, R.A., Huntley, B. & Hulme, P.E. (2000) Predicting the spatial distribution of non-indigenous riparian weeds: issues of spatial scale and extent. *Journal of Applied Ecology*, **37**, 13–27.

Costa, G.C., Nogueira, C., Machado, R.B. & Colli, G.R. (2010) Sampling bias and the use of ecological niche modeling in conservation planning: a field evaluation in a biodiversity hotspot. *Biodiversity and Conservation*, **19**, 883–899.

Del Hoyo, J., Elliot, A. & Sargatal, J. (1994) *Handbook of the birds of the world, vol. II. New world vultures to guineafowl*. Lynx Edicions, Barcelona.

Del Moral, J.C. (2006) *El águila perdicera en España. Población en 2005 y método de censo*. SEO/BirdLife, Madrid.

Dennis, B. (1996) Discussion: should ecologists become Bayesians? *Ecological Applications*, **6**, 1095–1103.

Dormann, C.F., McPherson, J.M., Araujo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemuller, R., Peres-Neto, P.R., Reineking, B., Schroder, B., Schurr, F.M. & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609–628.

EEA (2007) *European Environmental Agency data service*. European Environment Agency, Copenhagen.

Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology Evolution and Systematics*, **40**, 677–697.

Elith, J., Graham, C.H., Anderson, R.P. *et al.* (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. & Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.

Ellison, A.M. (2004) Bayesian inference in ecology. *Ecology Letters*, **7**, 509–520.

Estrada, A., Real, R. & Vargas, J.M. (2008) Using crisp and fuzzy modelling to identify favourability hotspots useful to perform gap analysis. *Biodiversity and Conservation*, **17**, 857–871.

Evangelista, P.H., Kumar, S., Stohlgren, T.J., Jarnevich, C.S., Crall, A.W., Norman Iii, J.B. & Barnett, D.T. (2008) Modelling invasion for a habitat generalist and a specialist plant species. *Diversity and Distributions*, **14**, 808–817.

Ferrier, S., Watson, G., Pearce, J. & Drielsma, M. (2002) Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling. *Biodiversity and Conservation*, **11**, 2275–2307.

Fielding, A.H. & Bell, J.F. (1997) A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, **24**, 38–49.

Font, I. (2000) *Climatología de España y Portugal*. Universidad de Salamanca, Salamanca.

Forsyth, R. (1984) *Expert systems: principles and case studies*. Chapman and Hall, London.

Franklin, J. (2009) *Mapping species distributions: spatial inference and prediction*. Cambridge University, Cambridge.

Freeman, E.A. & Moisen, G. (2008) PresenceAbsence: an R package for presence absence analysis. *Journal of Statistical Software*, **23**, 1–31.

Gil-Sanchez, J.M., Moleon, M., Otero, M. & Bautista, J. (2004) A nine-year study of successful breeding in a Bonelli's eagle population in southeast Spain: a basis for conservation. *Biological Conservation*, **118**, 685–694.

Gil-Sánchez, J.M., Molino-Garrido, F. & Valenzuela, G. (1996) Selección de hábitat de nidificación por el Águila Perdicera (*Hieraaetus fasciatus*) en Granada (SE de España). *Ardeola*, **43**, 189–197.

Graham, C.H., Elith, J., Hijmans, R.J., Guisan, A., Peterson, A.T., Loiselle, B.A. & The Nceas Predicting Species Distributions Working Group (2008) The influence of spatial errors in species occurrence data used in distribution models. *Journal of Applied Ecology*, **45**, 239–247.

Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.

Guisan, A., Zimmermann, N.E., Elith, J., Graham, C.H., Phillips, S. & Peterson, A.T. (2007) What matters for predicting the occurrences of trees: techniques, data, or species' characteristics? *Ecological Monographs*, **77**, 615–630.

Hartley, S., Kunin, W.E., Lennon, J.J. & Pocock, M.J.O. (2003) Coherence and discontinuity in the scaling of species' distribution patterns. *Proceedings of the Royal Society B: Biological Sciences*, **271**, 81–88.

Heikkinen, J. & Hogmander, H. (1994) Fully Bayesian Approach to Image Restoration with an Application in Biogeography. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, **43**, 569–582.

Hernandez, P.A., Graham, C.H., Master, L.L. & Albert, D.L. (2006) The effect of sample size and species characteristics on performance of different species distribution modelling methods. *Ecography*, **29**, 773–785.

Hijmans, R.J. & van Etten, J. (2010) raster: Package for reading, writing, and manipulating raster (grid) type geographic (spatial) data. Available at: http://cran.r-project.org/web/packages/raster/index.html.

Hijmans, R.J., Susan, E., Cameron, J.L., Parra Peter, G. & Jones, A.J. (2005) Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology*, **25**, 1965–1978.

Hobbs, N.T. (2003) Challenges and opportunities in integrating ecological knowledge across scales. *Forest Ecology and Management*, **181**, 223–238.

Hosmer, D.W. & Lameshow, S. (2000) *Applied logistic regression*. John Wiley & Sons, New York.

I.G.N. (1999) *Mapa de carreteras. Península Ibérica, Baleares y Canarias*. Instituto Geográfico Nacional/Ministerio de Fomento, Madrid.

James, A., Gaston, K.J. & Balmford, A. (2001) Can we afford to conserve biodiversity? *BioScience*, **51**, 43–52.

Jiménez, J.J. & Muñoz, A.R. (2008) *Atlas de las rapaces diurnas de la provincia de Málaga (Reproducción, invernada y migración)*. Spain, Centro de Ediciones de la Diputación de Málaga.

Lee, T., Richards, J.A. & Swain, P.H. (1987) Probabilistic and evidential approaches for multisource data analysis. *Geoscience and Remote Sensing, IEEE Transactions on GE*, **25**, 283–293.

Legendre, P. (1993) Spatial autocorrelation: trouble or new paradigm. *Ecology*, **74**, 1659–1673.

Lehmann, A., Overton, J.M. & Leathwick, J.R. (2002) GRASP: generalized regression analysis and spatial prediction. *Ecological Modelling*, **157**, 189–207.

Leyequien, E., Verrelst, J., Slot, M., Schaepman-Strub, G., Heitkönig, I.M.A. & Skidmore, A. (2007) Capturing the fugitive: applying remote sensing to terrestrial animal distribution and diversity. *International Journal of Applied Earth Observation and Geoinformation*, **9**, 1–20.

Lloyd, P. & Palmer, A.R. (1998) Abiotic factors as predictors of distribution in southern African Bulbuls. *Auk*, **115**, 404–411.

Lobo, J.M., Jimenez-Valverde, A. & Real, R. (2008) AUC: a misleading measure of the performance of predictive distribution models. *Global Ecology and Biogeography*, **17**, 145–151.

Lobo, J.M., Jimenez-Valverde, A. & Hortal, J. (2010) The uncertain nature of absences and their importance in species distribution modelling. *Ecography*, **33**, 103–114.

Lopez-Lopez, P., Garcia-Ripolles, C., Aguilar, J., Garcia-Lopez, F. & Verdejo, J. (2006) Modelling breeding habitat preferences of Bonelli's eagle (*Hieraaetus fasciatus*) in relation to topography, disturbance, climate and land use at different spatial scales. *Journal of Ornithology*, **147**, 97–106.

Lopez-Lopez, P., Garcia-Ripolles, C., Soutullo, A., Cadahia, L. & Urios, V. (2007a) Are important bird areas and special protected areas enough for conservation?: the case of Bonelli's eagle in a Mediterranean area. *Biodiversity and Conservation*, **16**, 3755–3780.

Lopez-Lopez, P., Garcia-Ripolles, C. & Urios, V. (2007b) Population size, breeding performance and territory quality of Bonelli's Eagle *Hieraaetus fasciatus* in eastern Spain. *Bird Study*, **54**, 335–342.

Marti, R. & Del Moral, J.C. (2003) *Atlas de las aves reproductoras de España*. Sociedad Española de Ornitologia, Madrid.

Martinez, J.A., Calvo, J.F., Martinez, J.E., Zuberogoitia, I., Zabala, J. & Redpath, S.M. (2008a) Breeding performance, age effects and territory occupancy in a Bonelli's Eagle *Hieraaetus fasciatus* population. *Ibis*, **150**, 223–233.

Martinez, J.E., Martinez, J.A., Zuberogoitia, I., Zabala, J., Redpath, S.M. & Calvo, J.F. (2008b) The effect of intra- and interspecific interactions on the large-scale distribution of cliff-nesting raptors. *Ornis Fennica*, **85**, 13–21.

McCarthy, M.A. (2007) *Bayesian methods for ecology*. Cambridge University Press, New York.

McPherson, J.M., Jetz, W. & Rogers, D.J. (2006) Using coarse-grained occurrence data to predict species distributions at finer spatial resolutions-possibilities and limitations. *Ecological Modelling*, **192**, 499–522.

Moleon, M., Sanchez-Zapata, J.A., Real, J., Garcia-Charton, J.A., Gil-Sanchez, J.M., Palma, L., Bautista, J. & Bayle, P. (2009) Large-scale spatio-temporal shifts in the diet of a predator mediated by an emerging infectious disease of its main prey. *Journal of Biogeography*, **36**, 1502–1515.

Montgomery, D.C. & Peck, E.A. (1982) *Introduction to linear regression analysis*. John Wiley and Sons, New York.

Moreno-Rueda, G., Pizarro, M., Ontiveros, D. & Pleguezuelos, J.M. (2009) The coexistence of the eagles Aquila chrysaetos and Hieraaetus fasciatus increases with low human population density, intermediate temperature, and high prey diversity. *Annales Zoologici Fennici*, **46**, 283–290.

Muñoz, A.R., Real, R., Barbosa, A.M. & Vargas, J.M. (2005) Modelling the distribution of Bonelli's eagle in Spain: implications for conservation planning. *Diversity and Distributions*, **11**, 477–486.

Murray, J.V., Goldizen, A.W., O'Leary, R.A., McAlpine, C.A., Possingham, H.P. & Choy, S.L. (2009) How useful is expert opinion for predicting the distribution of a species within and beyond the region of expertise? A case study using brush-tailed rock-wallabies *Petrogale penicillata*. *Journal of Applied Ecology*, **46**, 842–851.

Naylor, C. (1984) How to build an inferencing engine. *Expert systems: principals and case studies* (ed. by R. Forsyth), pp. 84. Chapman and Hall, London.

Newbold, T. (2010) Applications and limitations of museum data for conservation and ecology, with particular attention to species distribution models. *Progress in Physical Geography*, **34**, 3–22.

Oindo, B.O., Skidmore, A.K. & De Salvo, P. (2003) Mapping habitat and biological diversity in the Maasai Mara ecosystem. *International journal of remote sensing*, **24**, 1053–1069.

Ontiveros, D. & Pleguezuelos, J.M. (2000) Influence of prey densities in the distribution and breeding success of Bonelli's eagle (*Hieraaetus fasciatus*): management implications. *Biological Conservation*, **93**, 19–25.

Ontiveros, D. & Pleguezuelos, J.M. (2003a) Physical, environmental and human factors influencing productivity in Bonelli's eagle *Hieraaetus fasciatus* in Granada (SE Spain). *Biodiversity and Conservation*, **12**, 1193–1203.

Ontiveros, D. & Pleguezuelos, J.M. (2003b) Influence of climate on Bonelli's eagle's (*Hieraaetus fasciatus* V, 1822) breeding success through the Western Mediterranean. *Journal of Biogeography*, **30**, 755–760.

Ontiveros, D., Pleguezuelos, J.M. & Caro, J. (2005) Prey density, prey detectability and food habits: the case of Bonelli's eagle and the conservation measures. *Biological Conservation*, **123**, 19–25.

Palma, L., Beja, P., Pais, M. & Da Fonseca, L.C. (2006) Why do raptors take domestic prey? The case of Bonelli's eagles and pigeons. *Journal of Applied Ecology*, **43**, 1075–1086.

Pearce, J.L., Cherry, K., Drielsma, M., Ferrier, S. & Whish, G. (2001) Incorporating expert opinion and fine-scale vegetation mapping into statistical models of faunal distribution. *Journal of Applied Ecology*, **38**, 412–424.

Pereira, J.M.C. & Itami, R.M. (1991) GIS-based habitat modeling using logistic multiple regression: a study of the Mt Graham Red Squirrel. *Photogrammetric Engineering and Remote Sensing*, **57**, 1475–1486.

Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modelling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.

Phillips, S.J., Dudik, M., Elith, J., Graham, C.H., Lehmann, A., Leathwick, J. & Ferrier, S. (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. *Ecological Applications*, **19**, 181–197.

R Development Core Team (2010) *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.

Raes, N. & ter Steege, H. (2007) A null-model for significance testing of presence-only species distribution models. *Ecography*, **30**, 727–736.

Real, J. (2003) Águila-azor Perdicera, *Hieraaetus fasciatus*. *Atlas de las Aves Reproductoras de España* (ed. by R. Marti and J.C. del Moral), pp. 192–193. Ministerio de Medio Ambiente-Sociedad Española de Ornitología, Madrid.

Rebelo, H. & Jones, G. (2010) Ground validation of presence-only modelling with rare species: a case study on barbastelles *Barbastella barbastellus* (Chiroptera: Vespertilionidae). *Journal of Applied Ecology*, **47**, 410–420.

Rico, L., Sánchez-Zapata, J.A., Izquierdo, A., García, J.R., Morán, S. & Rico, D. (1999) Tendencias recientes en las poblaciones del Águila Real Aquila chrysaetos y el Águila-azor Perdicera *Hieraaetus fasciatus* en la provincia de Valencia. *Ardeola*, **46**, 235–238.

Robertson, M.P., Cumming, G.S. & Erasmus, B.F.N. (2010) Getting the most out of atlas data. *Diversity and Distributions*, **16**, 363–375.

Rocamora, G. (1994) Bonelli's eagle *Hieraaetus fasciatus*. *Birds in Europe, their conservation status* (ed. by G.M. Tucker and M.F. Heath), pp. 184–185. BirdLife International (BirdLife Conservation Series no. 3), Cambridge, UK.

Rouget, M. (2003) Measuring conservation value at fine and broad scales: implications for a diverse and fragmented region, the Agulhas Plain. *Biological Conservation*, **112**, 217–232.

Schmidt, K.S., Skidmore, A.K., Kloosterman, E.H., van Oosten, H.H., Kumar, L. & Janssen, J.A.M. (2004) Mapping coastal vegetation using an expert system and hyperspectral imagery. *Photogrammetric Engineering and Remote Sensing*, **70**, 703–715.

Segurado, P. & Araujo, M.B. (2004) An evaluation of methods for modelling species distributions. *Journal of Biogeography*, **31**, 1555–1568.

Skidmore, A.K. (1989) An expert system classifies Eucalypt forest types using thematic mapper data and digital terrain model. *Photogrammetric Engineering and Remote Sensing*, **55**, 1449–1464.

Skidmore, A.K. (1996) Operational GIS expert system for mapping forest soils. *Photogrammetric Engineering and Remote Sensing*, **62**, 501–511.

Skidmore, A.K. & Turner, B.J. (1992) Map accuracy assessment using line intersect sampling. *Photogrammetric Engineering and Remote Sensing*, **58**, 1453–1457.

Soutullo, A., Lopez-Lopez, P. & Urios, V. (2008) Incorporating spatial structure and stochasticity in endangered Bonelli's eagle's population models: implications for conservation and management. *Biological Conservation*, **141**, 1013–1020.

Synes, N.W. & Osborne, P.E. (2011) Choice of predictor variables as a source of uncertainty in continental-scale species distribution modelling under climate change. *Global Ecology and Biogeography*, In press.

Wade, P.R. (2000) Bayesian methods in conservation biology. *Conservation Biology*, **14**, 1308–1316.

Wang, T.J., Skidmore, A.K. & Toxopeus, A.G. (2009) Improved understorey bamboo cover mapping using a novel hybrid neural network and expert system. *International journal of remote sensing*, **30**, 965–981.

Wang, T., Ye, X., Skidmore, A.K. & Toxopeus, A.G. (2010) Characterizing the spatial distribution of giant pandas (*Ailuropoda melanoleuca*) in fragmented forest landscapes. *Journal of Biogeography*, **37**, 865–878.

Wisz, M.S., Hijmans, R.J., Li, J., Peterson, A.T., Graham, C.H. & Guisan, A. (2008) Effects of sample size on the performance of species distribution models. *Diversity and Distributions*, **14**, 763–773.

SUPPORTING INFORMATION

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

Appendix S1 Environmental explanatory variables used in SDMs.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

BIOSKETCH

Aidin Niamir is a PhD candidate at the Biodiversity in Fragmenting Landscapes (BIOFRAG) laboratory, Faculty for Geo-Information Science and Earth Observation (ITC), University of Twente, the Netherlands. His primary interest is in methods for implementing and evaluating species distribution models with focus on the incorporation of existing knowledge into the modelling procedure.

Author contributions: A.N. and A.K.S. conceived the ideas and developed the methods; A.N. implemented the study and analysed the results; A.G.T. and A.R.M. provided the case study data and expertise; A.K.S., R.R. and A.G.T. supervised the research. All authors contributed to the writing.

Editor: Mark Robertson