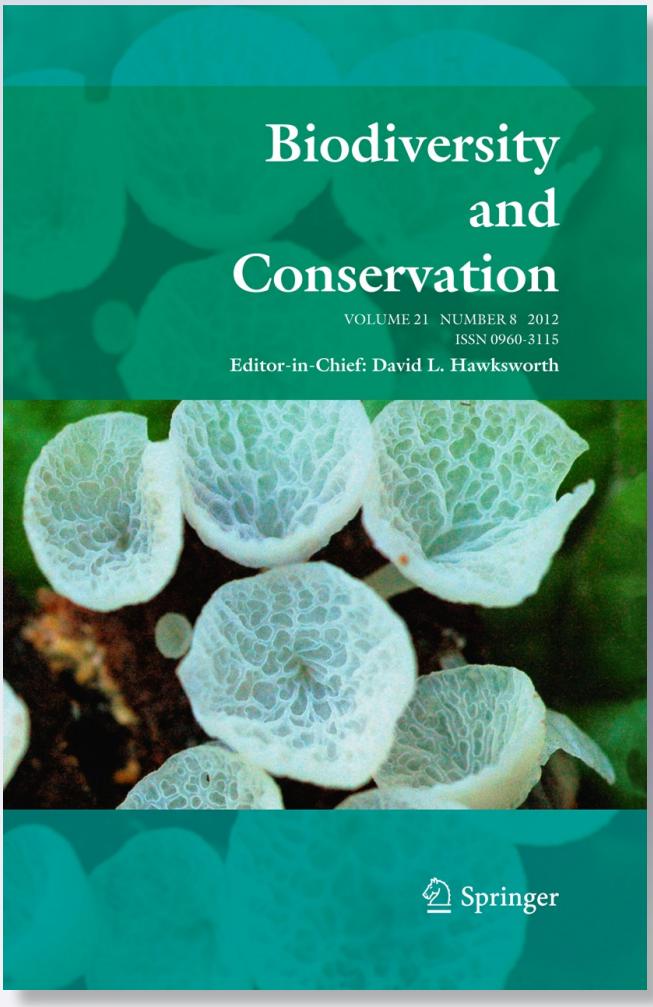
Mapping from heterogeneous biodiversity monitoring data sources

Francesc Sardà-Palomera, Lluís Brotons, Dani Villero, Henk Sierdsema, Stuart E. Newson & Frédéric Jiguet

Biodiversity and Conservation

ISSN 0960-3115
Volume 21
Number 11

Biodivers Conserv (2012) 21:2927-2948
DOI 10.1007/s10531-012-0347-6



Springer

Biodivers Conserv (2012) 21:2927–2948
DOI 10.1007/s10531-012-0347-6

ORIGINAL PAPER

Mapping from heterogeneous biodiversity monitoring data sources

Francesc Sardà-Palomera · Lluís Brotons · Dani Villero · Henk Sierdsema · Stuart E. Newson · Frédéric Jiguet

Received: 29 November 2011 / Accepted: 17 July 2012 / Published online: 9 August 2012

Abstract Field monitoring can vary from simple volunteer opportunistic observations to professional standardised monitoring surveys, leading to a trade-off between data quality and data collection costs. Such variability in data quality may result in biased predictions obtained from species distribution models (SDMs). We aimed to identify the limitations of different monitoring data sources for developing species distribution maps and to evaluate their potential for spatial data integration in a conservation context. Using Maxent, SDMs were generated from three different bird data sources in Catalonia, which differ in the degree of standardisation and available sample size. In addition, an alternative approach for modelling species distributions was applied, which combined the three data sources at a large spatial scale, but then downscaling to the required resolution. Finally, SDM predictions were

F. Sardà-Palomera (✉) · L. Brotons · D. Villero
Biodiversity and Landscape Ecology Lab., Centre Tecnològic Forestal de Catalunya,
Crta. de Sant Llorenç de Morunys, Km.2, 25280 Solsona, Spain
e-mail: francesc.sarda@ctfc.cat

F. Sardà-Palomera · L. Brotons
Institut Català d'Ornitologia (ICO), Museu de Zoologia, Passeig Picasso s/n, 08003 Barcelona, Catalonia, Spain

F. Sardà-Palomera · L. Brotons
European Bird Census Council (EBCC)
URL: http://www.ebcc.info/

L. Brotons
CREAF Center for Ecological Research and Applied Forestries, Autonomous University of Barcelona, Bellaterra, Spain

H. Sierdsema
SOVON, Antwoordnr 98189, 6500 VA Nijmegen, The Netherlands

S. E. Newson
British Trust for Ornithology, The Nunnery, Thetford IP24 2PU, UK

F. Jiguet
Centre de Recherches sur la Biologie des Populations d'Oiseaux, UMR 7204 MNHN-CNRS-UPMC,
55 Rue Buffon CP 51, 75005 Paris, France

used to identify species richness and high quality areas (hotspots) from different treatments. Models were evaluated by using high quality Atlas information. We show that both sample size and survey methodology used to collect the data are important in delivering robust information on species distributions. Models based on standardized monitoring provided higher accuracy with a lower sample size, especially when modelling common species. Accuracy of models from opportunistic observations substantially increased when modelling uncommon species, giving similar accuracy to a more standardized survey. Although downscaling data through a SDM approach appears to be a useful tool in cases of data shortage or low data quality and heterogeneity, it will tend to overestimate species distributions. In order to identify distributions of species, data with different quality may be appropriate. However, to identify biodiversity hotspots high quality information is needed.

Keywords Citizen science · Farmland Birds · Maxent · Occurrence data bias · Species distribution models

Introduction

The availability of representative spatial biodiversity data is essential for efficient species conservation, planning and management (Stem et al. 2005). However, biodiversity data are often fragmented, posing problems on how to make best use of the available information for biodiversity management and planning strategies (Soberon and Peterson 2004; Boakes et al. 2010). Species distribution models (SDMs) have been widely used to bridge the gap between fragmentary species observations and the need for continuous information at large spatial scales. SDMs rely on determining species—environmental relationships and representing these in space. However, while much has been done on technical developments of such methods (Elith et al. 2006; Araújo and New 2007; Guisan et al. 2007; Phillips and Dudik 2008), less work has focused on the effects of limitations regarding data availability, quality and spatial scale, specifically regarding the origin of the biodiversity data used for modelling (Vaughan and Ormerod 2003; Boakes et al. 2010; Feeley and Silman 2011).

Data sources of differing quality and characteristics may be available for different species in a given region, which could be used in species distribution modelling (Boakes et al. 2010). However, whilst it is assumed that models based on opportunistic observations perform less well than those based on systematic and well designed sampling (Schreuder et al. 2001), repeated standardised sampling of large areas over time is generally outside the scope of most conservation programs (Danielsen et al. 2005; Brotons et al. 2007). An alternative approach is to make use of museum specimens, and lately opportunistic records collected by volunteers through web-based tools (Roberts et al. 2007; Schmeller et al. 2009; Munson et al. 2010; Conrad and Hilchey 2011). However, this kind of opportunistic information may have considerable shortcomings, limiting its application in a spatial modelling context. For instance, observations may be biased towards areas where there are more visits, and the quality of the data is often heterogeneous (Snall et al. 2011). Little work has looked at the consequences of the method used for gathering information for species distribution modelling, but it appears important to make optimal use of resources invested in future biodiversity data collection effort. This is especially relevant in a context in which multiple data sources coexist in time and space.

Thus, due to financial constraints and lack of human resources, there is usually a trade-off between the quality of sampling, sample size and survey area covered (Danielsen et al. 2005; Snall et al. 2011). Consequently, for biodiversity conservation it is crucial to know

where the investment of resources should be prioritised, into systematic and more accurate local sampling, or into monitoring strategies based on opportunistic records. In this context, quantifying the limitations of different data characteristics and how these may be overcome is critical to guide monitoring strategies in areas where resources are limited (Braunisch and Suchant 2010). One of the issues arising when combining locally collected data across larger spatial and temporal scales relates to the need to make the best possible use of data obtained from different monitoring schemes and alternative data sources, and as a consequence, contribute to the evaluation of the efficiency of existing monitoring schemes (Brotons et al. 2007; Schmeller et al. 2009).

Here, we aim to identify the power of different data sources for mapping species distributions and the potential for map integration or compatibility of information derived from such diversity of raw information. Thus, it is important to understand the power of different types of data to produce maps from SDMs. If different data qualities and sample sizes are available for model development from different areas, one strategy may be to make the best possible use of available data in each region and integrate the maps obtained. In particular, we investigate the predictive value of models built with these different data sets of differing quality, and second we quantify the degree in spatial matching of suitable habitat maps derived from these models with high quality independent reference data already available for these species.

Our study focuses on farmland birds. This group is threatened in Europe and considerable attention is being directed towards the identification of areas where conservation effort should be invested (Green et al. 2005; Gregory et al. 2005). Getting reliable spatial information of farmland bird distributions from different data sources is important to detect farmland bird hotspots, a need for biodiversity conservation under an efficient agricultural landscape management and planning. By using a set of farmland bird species as study models, we test the impact of data quality and quantity on common and uncommon species in order to derive general guidelines relating the power of different biodiversity data sources to model species distributions. Finally, we also test the impact of data quality and quantity in identifying species richness and habitat quality hotspots, which play an important role in management and conservation programs.

Methods

Study area and species selection

We used bird data from Catalonia in north-east Spain. Catalonia is environmentally heterogeneous with a long tradition in ornithological monitoring that has allowed the gathering of several independent bird data sets of different quality.

A total of 30 farmland bird species were selected (Appendix Table 4). These species were chosen to include those in the European Farmland Bird Indicator developed by Gregory et al. (2005) that breed in Catalonia. We also included some additional farmland species in Catalonia, which are more specific to the Mediterranean farmland areas (Brotons et al. 2004a; Burfield 2004; Estrada et al. 2004). These species were classified into two groups, common and uncommon species. In order to establish a general criterion to define these two groups (common and uncommon) we used a criterion based on the frequency with which a species is found in standardized samples (Estrada et al. 2004), as a measure of their overall presence in the environment. Thus, the 50 % of the species with the lower number of occurrences were classified as uncommon, while the rest were classified as common.

Available data sets

All data used in our study originated from large scale volunteer based monitoring programs with different characteristics with respect to standardisation in data collection. We only included data from the breeding period, from 15 April to 30 June for 2009 and 2010, when all the monitoring programs were fully operational.

Occasional observational data (OBS)

These data were collected by volunteers through the web-based monitoring tool Ornitho (Kery et al. 2010) which started to be used as a tool for volunteer data collection in 2006 in Catalonia (www.ornitho.cat). These data only include opportunistic bird records, with no information on the time spent in the field or information on other species recorded on the same visit to a given site. Thus, this kind of opportunistic data is very easy to collect and increasingly gathered by ornithological societies around the world (Roberts et al. 2005; Roberts et al. 2007). However, opportunistic information is characterised by a lack on information on the area sampled and strong spatial biases. Furthermore, our expectation is that this kind of survey is more likely to sample uncommon species compared to more standardised surveys due to particular interest in these species for bird observers.

Timed bird lists data (LIST)

This kind of information was also derived from web-based monitoring tool Ornitho. In this monitoring programme, volunteers report timed species lists of variable duration in which species seen or heard at the same site during a single timed visit are recorded. Recording time in these lists ranged from 5 to 982 min, however only those lists included within 95 % confidence interval were included (5–135 min). There were occasions where the same observer sampled the same locality more than once, in these cases we selected the total number of species detected in this locality and summed the minutes expended.

Timed species lists are relatively easy to collect, and although the area sampled is not recorded, the observer provides information on species not seen (those not included in the list) and some information on time effort. Lists are expected to sample both common and uncommon species relatively well.

Breeding bird survey data (BBS)

Bird contacts were recorded along 3-km line transects divided into three 1-km sections in which bird numbers were surveyed twice every year (see Herrando et al. 2008 for detailed information). Thus, in this study we used 1-km sections as sampling units to provide information on bird presence at a 1-km resolution (Brotons et al. 2007). Data from the central 1-km section were not used to reduce spatial autocorrelation in the bird data used. Recording time for the 1-km sections ranged from 11 to 78 min. One of the main strengths of BBS is that the same observer visits the same locality over years. Thus, in each locality we selected the total number of species detected in this locality and summed the minutes expended.

BBS transects are distributed according to two different criteria: transects located in a randomly chosen 10-km square within each of five biogeographically determined strata or transects located in a 10-km square freely chosen by the observer. These two possibilities

aim at attracting volunteers to the main areas of interest without losing any possible contributor because of distant or undesired census locations.

These data contained high quality information on sampling area and effort. Common species are expected to be very well sampled since breeding bird surveys are commonly designed to track common species abundances through time. While uncommon species are also recorded, effort to count common species may lead to the under recording of uncommon species.

Coarse grained heterogeneous bird data (COARSE)

An alternative approach is to downscale coarse gridded data to the required resolution (Witte et al. 2008). Presence–absence data of species was gathered at 10-km UTM sampling squares, considering all available evidence of breeding. This represents a very comprehensive cover of the region but at the sacrifice of spatial resolution and detail. Thus, the coarse scale dataset used to train models synthesized the results of all sampling surveys (OBS, LIST and BBS) at 10-km UTM grid squares, covering 95 % of the total area (309 squares). A sub-sampling method was applied for downscaling (see McPherson et al. 2006), randomly selecting 25 % of the 1-km UTM squares as occurrences within each 10-km UTM. Only 1-km UTM squares with at least 5 % of agricultural surface were selected (information obtained from SIGPAC: Spanish Geographic Information System of Agricultural Plots). Different sub-sampling sizes (5, 20, 25 and 50 %) were previously tested, finally selecting the sample size that better performed in the final model (i.e.: 25 %).

Fine grained atlas data (ATLAS)

Reference, high quality information on species distribution was gathered through the Catalan Breeding Bird Atlas (Estrada et al. 2004; Brotons et al. 2007). In the atlas survey, 1-km UTM sampling squares were visited twice for an hour in a given year during the period 1999–2002 within the breeding period. These represent a high quality, highly standardised independent data set, with a homogenous spatial cover and large sample of 3,077 1-km squares (approximately 9 % of the total area). Sampled squares were stratified over the main habitat types within each 10-km UTM squares (Hirzel and Guisan 2002).

Atlas data was used as an independent standard reference to be compared with results obtained from the other data sets. However, atlas data can be affected by the non-detection of species that are actually present (see e.g., Kery et al. 2010; Mccarthy et al. 2012). Furthermore, the atlas data were collected in 1999–2002, while volunteer data were collected in 2009–2010. Although non-detection of presences during atlas sampling and species distribution changes over one decade may occur, we always compared monitoring data from the same period of the year with the best available and most standardized method (ATLAS). Possible non-detected presences and/or changes occurred in the last decade would be the same for all comparisons.

Data treatments

Quality treatment

Here we used data sets of increasing quality, quantified according to the amount of standardisation required for data collection. Standardisation in data collection is translated

Table 1 Information regarding data included in species distribution modelling and percentage of uncommon species detected according to each monitoring scheme

	OBS	LIST	BBS	ATLAS
Total data	5,472	7,223	5,621	18,520
Total sample units	5,472	1,337	576	3,077
% Uncommon species	21.9	13.9	11.3	12.5

according to information on sampled area and sampling effort therefore allowing modelling to account for spatial and temporal variability in bird data. In our case this gradient ranges from low quality opportunistic observational data (OBS, LISTS) and coarse scale data (COARSE) to data from standardised monitoring (BBS). Our final treatments were: species opportunistic observations (OBS, presence-only data, 1-km), opportunistic collection of species lists (LISTS, presence–absence data, low standardisation, 1-km), breeding bird survey (BBS, presence–absence data, high standardisation, 1-km), coarse combined information at 10-km (COARSE, presence data, low standardisation).

In the different monitoring methods used, information was not collected uniformly within each 1-km square. In the case of LIST and BBS the accumulated time expended in each 1-km square was used as an estimate of sampling effort. Sampling effort was therefore included as an independent variable in final models. The resulting models were projected to the whole study area using common values of 43 min for all pixels (average time expended in BBS sections). In OBS there was no information available on the sampling time and therefore this variable was not used in the model calibration.

In spite of the difficulties in comparing sampling effort between the different methods, the similar number of total records for the total number of species analysed suggest that treatments containing all available data were comparable (Table 1).

Quantity treatment

We used subsets of the overall data sets for each monitoring program to simulate limited data availability and therefore assess the impact of data shortage in mapping species distributions. In the case of BBS, sample size was selected according to the percentage of sections (5, 10, 25, 50 and 100 % of the total 1-km sections sampled). With regard to web collected bird data (OBS and LIST), the thresholds (5, 10, 25, 50 and 100 %) were randomly applied across the whole data-base including all the species. With this treatment, we expect that uncommon species would be less represented or even missing when using a small percentage of the data set.

Environmental predictors

Environmental predictors for model training included a set of 27 explanatory variables circumscribed to Catalonia including climate (6), topography (5) and land-cover (17). All variables were re-sampled from their original spatial resolution to 1-km UTM grid cells. Climate predictors included mean annual temperature, mean minimum temperature for the coldest month (January), mean maximum temperature for the hottest month (July), annual precipitation, summer precipitation and annual solar radiation, from the Digital Climatic Atlas of Catalonia (180 m resolution) (Pons 1996; Ninyerola et al. 2005) from a climatic

standardized series of 48 years (1951–1999). Topographic predictors included mean altitude, mean slope, standard deviation of slope and proportion of sunny surface, derived from a digital elevation model (10 m resolution) generated by the Cartographic Institute of Catalonia from topographic 1:50,000 maps. Land-cover predictors were derived from the Cartographic Institute of Catalonia classification of 2002 Landsat-TM remote sensing imagery (Viñas and Baulies 1995) (30 m resolution) (available from the web of the Department of the Environment and Housing of Catalonia). From the 22 original classes, we defined 17 variables as the percentage of each land-cover class within each individual 1-km UTM grid cell. Our classification comprised: coniferous, deciduous and sclerophyllous forests, shrublands, alpine grasslands, sandy areas, bare ground, wetlands, continental and marine waters, snowdrifts, dry and irrigated herbaceous and woody crops, vineyards and built-up areas (urban areas, industrial areas and infrastructures). Regarding the definition of land-cover classes, we adopted the categories given from the Land-cover map, grouping the three categories of artificial surfaces into one single category (built-up areas).

Species distribution modelling

We used Maxent to develop SDMs. Maxent is a general-purpose method for making predictions from incomplete information based on the maximum entropy principle. This approach assumes that the best approximation to an unknown probability distribution is to ensure that it satisfies any constraints that we are aware of (Phillips and Dudik 2004; Phillips et al. 2006). Generally applied to presence-only species distribution modelling, the idea is to estimate the distribution of a species as a probability distribution across a study region, subject to the constraints that each expected predictor variable has to match its empirical average over the presence sites. Of all the probability distributions that meet these constraints, Maxent chooses the most unconstrained one, i.e. the one of maximum entropy (Phillips and Dudik 2008).

Following Elith et al. (2010), we applied the method to both presence-only (OBS, COARSE) and presence–absence data (LIST, BBS and ATLAS) by using different protocols for the Maxent algorithm. In the case of presence-only data, we applied the standard Maxent protocol in which the whole study area (i.e. Catalonia) is used to sample for absences, so that the background model includes all sites with no information. On the other hand, in the case of presence–absence data, we restricted model calibration to sampled areas where there is information on presences and absences, so that the background model includes areas with absences or species not detected. Importantly 'absences' are more likely to represent true absences in the ATLAS, where there was greater survey effort and imperfect detection less likely. Imperfect detection is most likely to be an issue for LIST with short sampling sessions. To avoid bias in predictions as a result of low quality samples, when modelling from BBS and LIST, we trained models with the environmental background area restricted to 1-km UTM squares surveyed (Phillips and Dudik 2008; Phillips et al. 2009). This forces the algorithm to select background information only from sampled areas (either leading to species presence or not). In these cases, Maxent approach is roughly comparable to the one employed in logistic regression (Elith et al. 2010).

In some cases, with reduced sample sizes, some species showed less than five occurrences. In those cases models were not run. Even when modelling some species distributions with larger sample sizes, some model predictions were equivalent to random predictions (Table 2). In the case of uncommon species, a higher number of models could be run from data from low standardization programs than from more standardized

Table 2 Number of species distribution models that could not be run due to small sample size or which predictions were equal or worse than a random prediction

Sample size	OBS		LIST		BBS	
	C	U	C	U	C	U
5	2	9	1	9	2	13
10	0	2	0	7	0	9
25	0	0	0	2	0	4
50	0	0	0	0	0	2
100	0	0	0	0	0	0

Values are shown for quantity treatment (sample size), quality treatment (OBS, LIST, BBS) and for common (C) and uncommon (U) species

programmes. When modelling uncommon species from BBS some models with 5 % of the total sample size could not be run or their distributions did not overlap with the atlas distribution. In these cases results did not show omission or commission errors.

Evaluation of SDMs

We used 10 fold-cross validation strategy AUC statistics (Pearce and Ferrier 2000) on independent fine grained Atlas data as a measure to identify the predictive power of all different treatments. After testing for normality, paired sample *t* tests were applied to compare AUC results obtained for each species between treatments.

We used a second evaluation procedure to check SDMs performance in which we assessed the degree to which species distribution obtained with the developed models matched those obtained in our independent species distribution model references (ATLAS). Here we assessed the overlap of resulting species maps between treatments (OBS, LIST, BBS and COARSE) and reference models (ATLAS). Different methods can be applied to select thresholds of occurrence in the prediction of species distributions (Liu et al. 2005). In this case we applied a zoning protocol (Herrando et al. 2011) aimed at categorising species distribution in areas of increasing environmental quality based on predefined thresholds. First, we searched for an environmental quality threshold that defined the distribution of the species (hereafter species distribution), below which the absence areas for species defined corresponds to the average tenth percentile of the data used for developing the models (i.e. 10 % of the data with the lowest suitability). A second threshold was applied to identify areas above the mean environment quality of the species within its distribution (Pearson et al. 2007). This protocol allowed us to derive ecologically sound and simple representations of species models as the area within the species distribution where environment suitability was above average (hereafter high quality areas).

We then calculated the overlap between species high quality areas for each species, obtained from each monitoring method and the corresponding reference model from the ATLAS to assess the degree to which a given model differed from our reference. Thus, it allowed us to calculate the percent omission and commission errors (Feeley and Silman 2011; Sardà-Palomera and Vieites 2011).

Beyond the overall predictive statistics obtained in the first evaluation exercises, overlap measures allow a more focused and direct evaluation of how far from our best available

reference, the identification of suitable sites for the species derived from our different treatment models fell.

Species richness

Using the species distribution and high quality areas identified we calculated, for each pixel of our study area, the number of species for which species distribution and high quality area could be found. The number of species present in each square for each quality and quantity treatment was determined, thus, identifying and mapping farmland bird hotspots in Catalonia. We applied Spearman correlations to calculate the correlation between richness for each quality and quantity treatment with the actual richness, observed in the 3,077 1 km squares where the ATLAS data have been collected.

Results

Total amount of data collected by each of the three monitoring methods used in our study was roughly comparable. However, a first look at the type of data sources offers light on the type of info contained. OBS seem to have a disproportionate focus on uncommon species compared to the other data sources (Table 1). LIST and BBS data have similar ratios therefore suggesting that these methods sample species proportionally to their abundance in the area.

General patterns in model accuracy across types of data

Model performance varied across the different data treatments for model training: quality and quantity. AUC generally increased with sample size for all monitoring methods (Fig. 1). BBS tended to result in higher AUC when modelling with the complete data set. This is especially relevant when comparing results obtained when including common species only. Models with highest AUC were obtained from BBS, while data from web-based monitoring in general gave similar results in the case of both opportunistic or lists data types. COARSE always showed a lower AUC mean than BBS with the complete data set (Fig. 1). Apart from ATLAS based models, the maximum AUC values were obtained

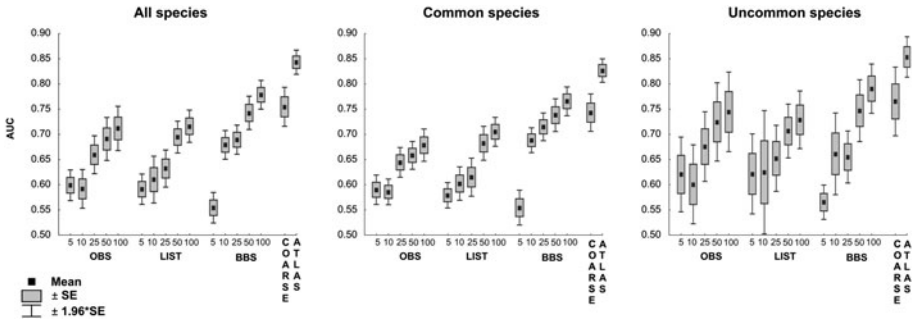


Fig. 1 Species distribution models mean AUC, boxplot, mean, standard error and confidence interval. Results are shown across different sample size percentage for each quality treatment (OBS, LIST, BBS) and for all the species, for common and uncommon species

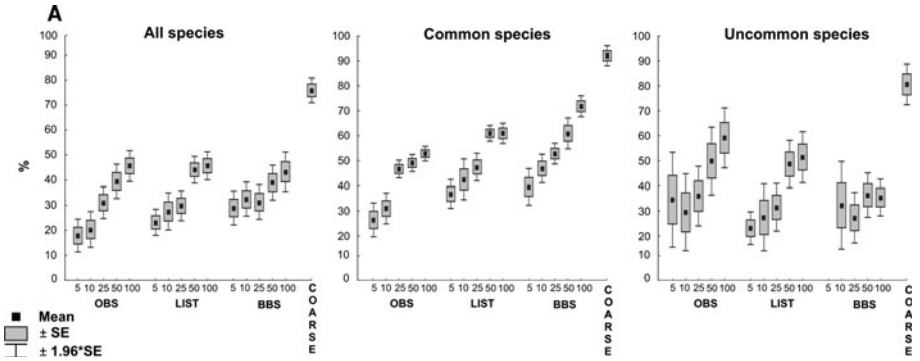


Fig. 2 Species high quality areas mean percent overlap, standard error and confidence interval. Results are shown across different sample size percentage for each quality treatment (OBS, LIST, BBS) for all the species, for common and uncommon species

from BBS when modelling uncommon species with the complete data set. However for common species, LIST using all available data (mean AUC = 0.70) performed better than OBS (mean AUC = 0.67) (t test for paired samples: $t_{14} = -3.26$, $p = 0.005$). When modelling common species models obtained from COARSE (mean AUC = 0.74) were more accurate than from LIST from the complete data set (t test for paired samples: $t_{14} = -3.26$, $p = 0.002$). When modelling uncommon species, all quality treatments showed similar results. Models for uncommon species tended to show higher AUC standard errors (Fig. 1).

The percentage of overlap with atlas was highest with high quality areas from COARSE in all cases (Fig. 2). In general, models from different monitoring programs showed similar overlapping ranges. However, BBS showed higher predictive model accuracy for common species modelling, followed by LIST, while for uncommon species OBS with the whole sample size showed the highest overlap. Models from uncommon species showed higher standard errors in all cases (Fig. 2).

An inverse pattern was shown in the percentage of omission (Fig. 3), which decreased with sample size. Models from COARSE always showed less omission errors. With regard to the different monitoring programs, when analysing all species together, there were not significant differences (t test for paired samples $p > 0.05$). In the case of common species, however, models from BBS using the complete data set showed less omission errors, on contrary models from OBS showed higher omission. However omission for uncommon species showed a different pattern, as OBS and LIST showed less omission than BBS when modelling from higher sample sizes (50 and 100 %). When modelling uncommon species from BBS, data omission did not decrease when increasing sample size (Fig. 3).

In the case of commission errors, COARSE showed the highest values when comparing with the other treatments, and especially when modelling uncommon species (Fig. 4). They also showed the highest standard errors and confidence intervals, thus they were less predictable. When taking into account all the species, models from BBS data showed lower commission when modelling from bigger sample sizes, both for common and uncommon species. Models from OBS with small sample size showed the lower commission for common species (Fig. 4). In general, models showed higher commission when modelling common than uncommon species.

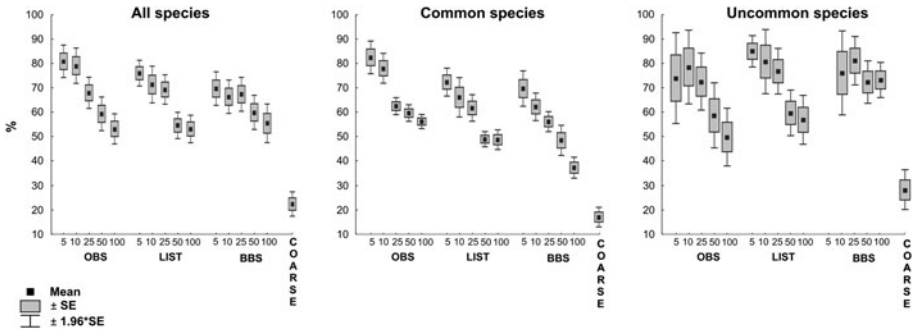


Fig. 3 Species high quality areas mean percent omission, standard error and confidence interval. Results are shown across different sample size percentage for each quality treatment (OBS, LIST, BBS) for all the species, for common and uncommon species

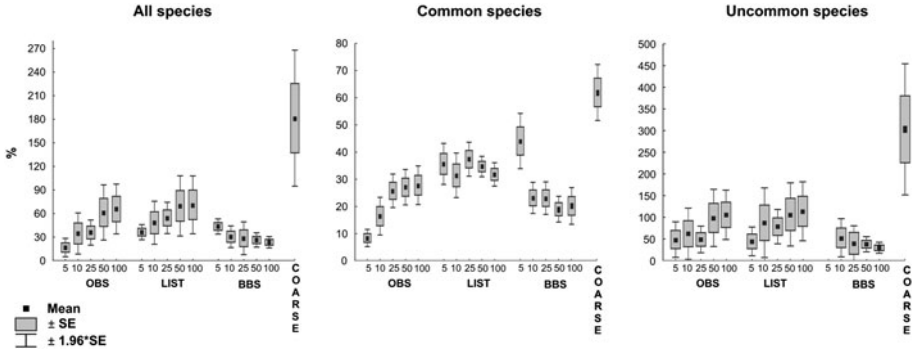


Fig. 4 Species high quality areas mean percent commission, standard error and confidence interval. Results are shown across different sample size percentage for each quality treatment (OBS, LIST, BBS) for all the species, for common and uncommon species

Species richness

Most of the species richness maps derived from different quality and quantity treatments were significantly correlated with species richness predictions from ATLAS data (Table 3). From a sample size of 10 % presences, predictions from BBS showed the highest agreements, showing the highest correlation when the complete data set was used. On the other hand, when modelling from five presences, models derived from opportunistic observations showed the highest correlation with ATLAS data (Table 3). In most of the cases, richness obtained from species distributions showed higher agreement with ATLAS data than richness obtained from high quality areas.

When predicting richness from species distributions all quality treatments overpredicted the number of species detected during ATLAS data collection. BBS was the monitoring method that showed a richness prediction closer to ATLAS richness, while predictions from OBS and LIST showed a similar pattern but with a higher overprediction than BBS (Figs. 5, 6; Figs. 8, 9 in Appendix). When predicting from high quality areas, predictions from BBS were very close to ATLAS richness with a slight overestimation in areas with few species but with a very high concordance in richest areas. OBS and LIST also followed

Table 3 *R* values for Spearman correlation when comparing species richness predictions from models obtained from different quality treatments (OBS, LIST, BBS and COARSE) with species richness obtained from atlas data

	Quantity treatment									
	5		10		25		50		100	
	SD	HQ	SD	HQ	SD	HQ	SD	HQ	SD	HQ
OBS	0.37	0.38	0.27	0.32	0.49	0.44	0.51	0.42	0.53	0.41
LIST	0.27	0.16	0.40	0.38	0.47	0.24	0.53	0.53	0.61	0.57
BBS	0.04	NS	0.56	0.40	0.65	0.56	0.66	0.62	0.67	0.66
COARSE	–	–	–	–	–	–	–	–	0.66	0.63

Results are shown for different sample size models (quantity treatment) and for species distribution (SD) and high quality (HQ) areas

N.S. non significant at $p < 0.05$

a very similar pattern, tending to increase underestimation in richer areas. Finally, model predictions from downscaled data showed the highest overestimation when predicting species richness both from species distribution and from high quality areas (Figs. 5, 6; Figs. 8, 9 in Appendix).

Accuracy patterns across species abundance

Model predictive accuracy throughout the species abundance in the region was also analysed. Only AUC values from OBS based models were negatively correlated with the number of presences per species (Spearman's $R = -0.39$, $p < 0.05$). Thus, the difference between ATLAS and OBS based models accuracy significantly increased along the uncommon to common species gradient (Spearman's $R = 0.47$, $p < 0.01$) (Fig. 7). The five most uncommon species showed the highest predictive power independent of data quality treatment used (Fig. 7).

Discussion

Data standardisation and quality

Models from monitoring breeding bird surveys showed the best results for modelling common species. This supports the results obtained by Brotons et al. (2007), which show that long term monitoring programs have the potential for being a source of good quality data for species distribution modelling. However, we have further increased our assessment of the potential of emergent monitoring programs based on the collection of opportunistic observations to derive useful distribution information. This information has the potential to deliver relevant insights into the distribution of species especially when no other information is available (Braunisch and Suchant 2010; Snall et al. 2011). They can be especially good for modelling uncommon species. In that case these two data sources may be relevant.

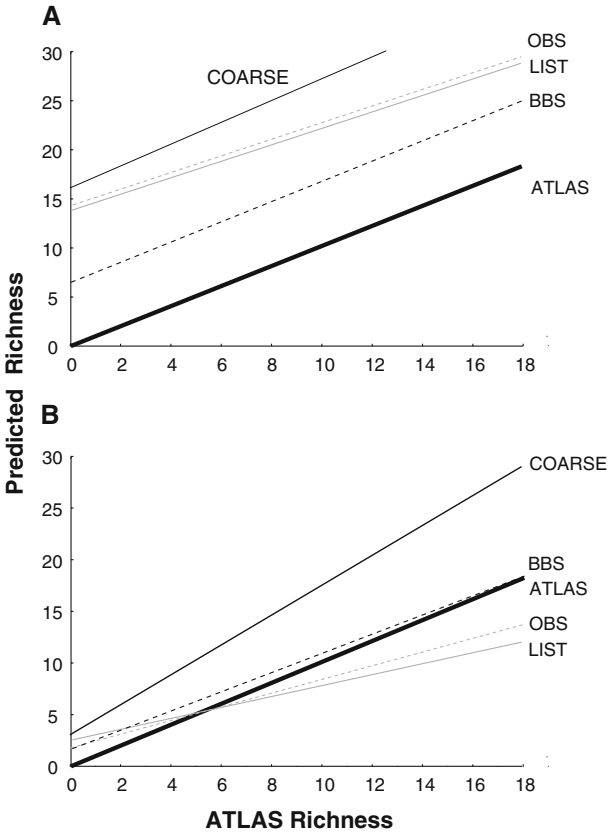


Fig. 5 Quantitative estimates of richness from each monitoring method correlated with richness from atlas data ($N = 3077$). Results are shown for ATLAS (*thick black solid line*), COARSE (*thin black solid line*), LIST (*thin grey solid line*), OBS (*grey broken line*) and BBS (*black broken line*), and for **a** species distribution area, and **b** high quality areas. Results are shown only for full sample size models

It has been previously noted that ecological characteristics of modelled species may affect model accuracy (Stockwell and Peterson 2002; Segurado and Araújo 2004) and that specialists in terms of both geographic and environmental space are easier to model than species with widespread distributions. In this study we have also shown that rarer, specialist species, may not require complex monitoring methods to obtain reliable models, providing that the sampling covers the environmental conditions used by the species. Keeping in mind that a highest predictive power (estimated through AUC) of rare species is probably related to how the AUC is calculated for presence-only models (Phillips et al. 2006), in our case, models for those species obtained highest accuracies independent of the monitoring and validation method.

How much quality data is enough?

Our results show that the standardized breeding bird survey was the most accurate method for monitoring common species and, in general, an increase in sample size will result in the best performing models (Feeley and Silman 2011). This can vary depending on species specificity (Hernández et al. 2006), as in some cases models for uncommon species can reach high accuracy even with a small sample size (Stockwell and

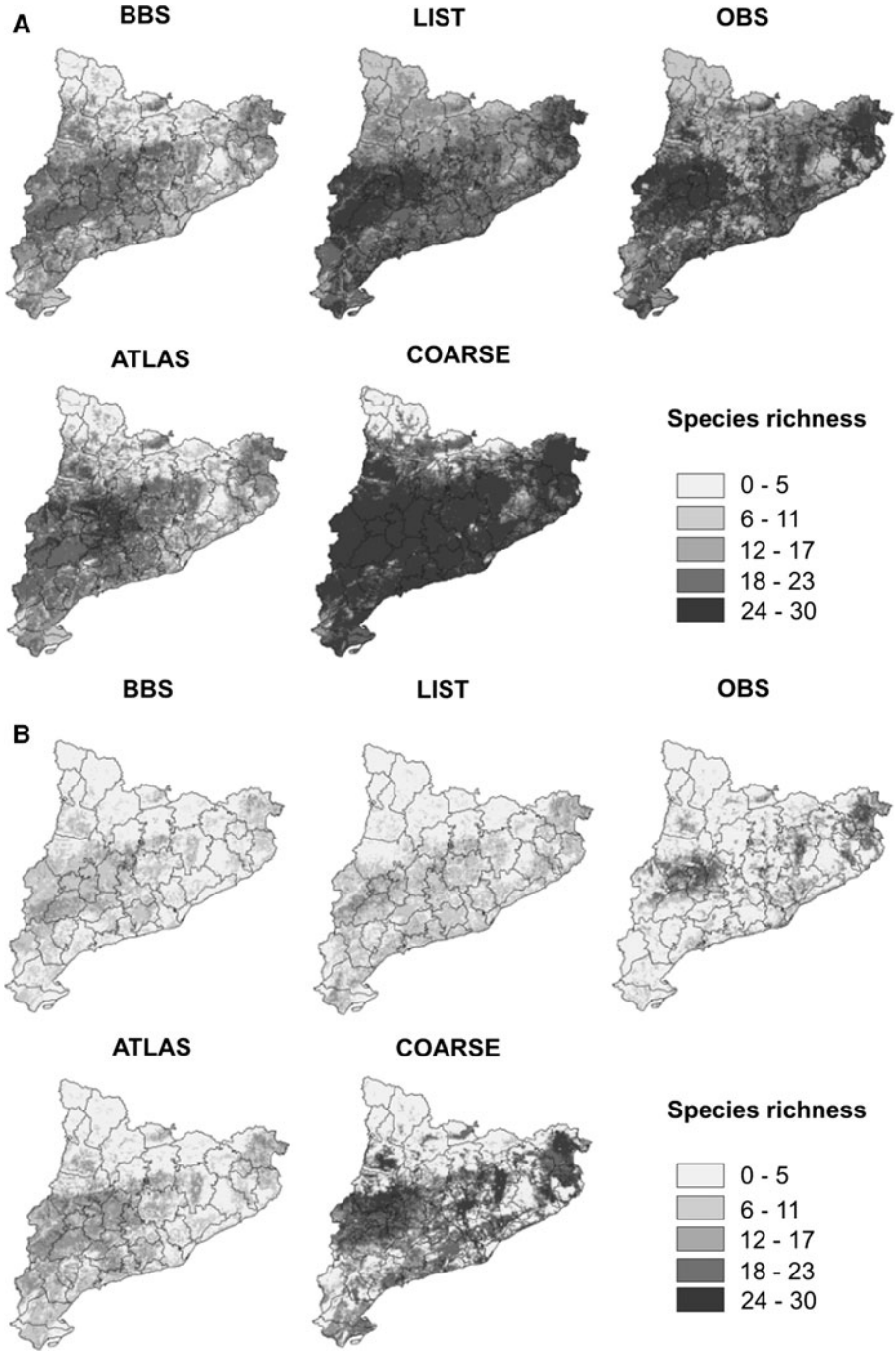


Fig. 6 Predictions of farmland bird species richness in Catalonia obtained from different quality treatments (ATLAS, BBS, LIST, OBS and COARSE) with the whole sample size. Maps are show for **a** species distribution area, and **b** high quality areas

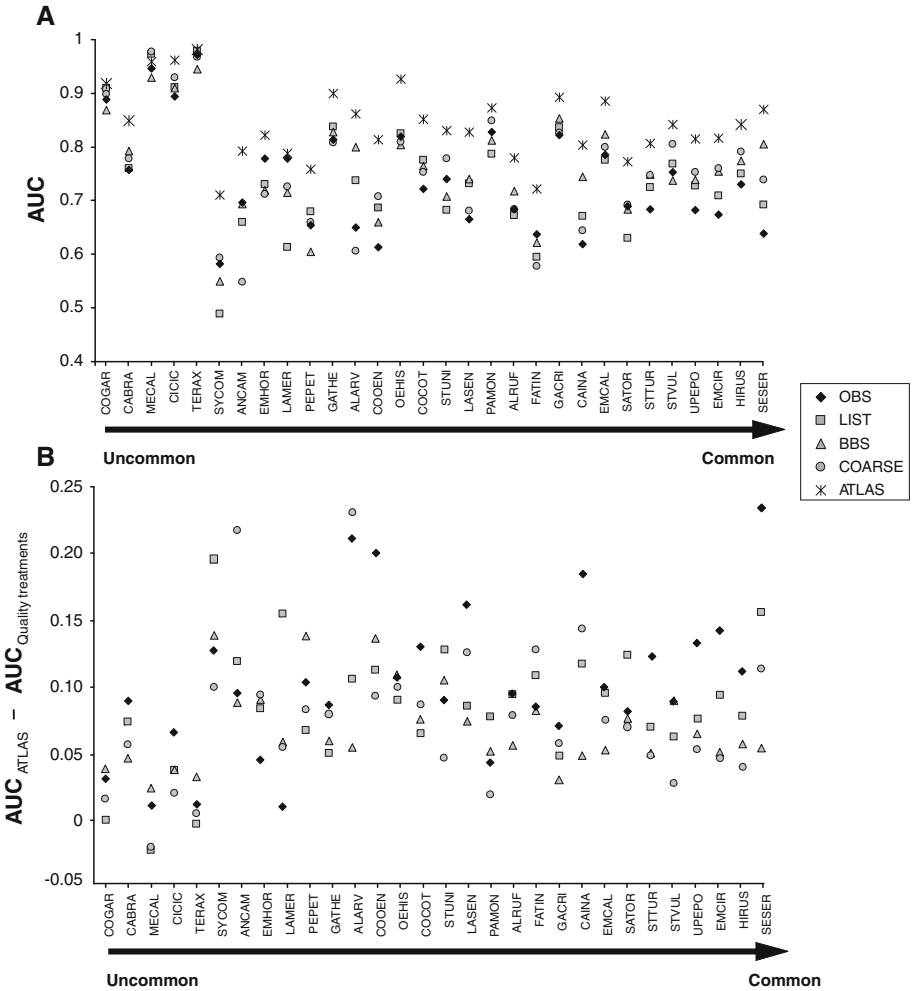


Fig. 7 **a** Species distribution models mean AUC values for each species and quality treatment and **b** mean AUC difference between quality treatments and ATLAS. Values are shown for models built with a sample size of 100 % occurrences. Farmland species (see Appendix Table 4 for codes) are sorted according the total number of occurrences in ATLAS (from uncommon to common)

Peterson 2002). In the case of commission and omission, an increase in sample size does not necessarily decrease the error, obtaining similar results from a priori less quality monitoring data. We also note that in the case of commission, an increase in sample size did not decrease the over prediction, which in some cases can even increase (Sardà-Palomera and Vieites 2011).

The fact that an asymptote was not reached for any monitoring method suggests that maximal concordance was not achieved and will likely continue to increase with sample size until a certain point. This point may depend on the species ecological characteristics (Stockwell and Peterson 2002) or when sample size reaches the optimum for an

adequate weighting of predictor variables (Feeley and Silman 2011; Sardà-Palomera and Vieites 2011). All in all, when modelling from small sample size, the predictions should be used with caution, as no algorithm will predict consistently well (Wisz et al. 2008).

More standardized methods provide the best quality data, but this comes at a cost in terms of money and man power and will normally result in a smaller sample size (Braunisch and Suchant 2010). Less standardized methods provide lower quality data but are easier to obtain, thus allowing 'a posteriori' corrections. In consequence, it will also depend on the abundance of particular species. Because uncommon species are less abundant and show smaller sample sizes, when modelling uncommon species from more standardised bird surveys, in some cases, there is not enough data to run or obtain reliable models.

Data integration across scales

Understanding environmental limitations of species distributions modelling across scales is essential to develop useful applications of SDMs that can provide reliable predictions of species distributions (Witte et al. 2008). Modelling species distributions across scales could become important in complementing information from different programs, in particular for uncommon species, so that the best use of available information is made. In this context, we have evaluated the usefulness of aggregated information at a coarse scale to subsequently downscale SDMs. While the statistics and models appeared generally good, these models tended to over predict species distribution (Araújo et al. 2005). Presence-only models show higher overprediction compared with presence–absence models (Brotons et al. 2004b). This is probably the reason that model predictions based on opportunistic observations show broader distributions than those based on lists and breeding bird survey data.

Applications for species conservation

Different treatments have different implications for predicting hotspots and detecting areas of conservation priority. Richness predictions from BBS were closer to richness predicted from the atlas data, while predictions from OBS, LIST and COARSE were less accurate. However, in all cases predictions of high quality habitat areas were less accurate than predictions of species distributions, suggesting that high quality data is needed to accurately identify hotspots.

In general, models with higher accuracy (i.e.: AUC values) are desirable. However, as estimates through AUC can be related to how the AUC is calculated (Phillips et al. 2006; Lobo et al. 2008) in some cases omission and commission errors can give more information about the suitability of a certain model than AUC values. For example, when designing nature protection areas, misclassifications of commission must be regarded as a more serious drawback than the opposite, while low omission is desirable when searching for new species or populations (see Peterson 2006). Models from data aggregated at a coarse scale resulted in maps with the highest overprediction.

In a conservation framework, overprediction can be interpreted in the light that models derived from this source of information, while adequate in some cases, may be too conservative in identifying too many good areas for species. When the aim is to protect uncommon or endangered species, overestimating areas of potentially elevated biodiversity might be preferable than underestimating their existence (Zaniewski et al. 2002).

On the other hand, overprediction of the range of species can lead to erroneously conclude that there are larger and more suitable areas for that species, and therefore no need for conservation action. Thus, downscaling of species distribution through modelling may be a useful approach for areas with scarce or heterogeneous information, and also when integrating different quality data from different regions or countries across scales. However, special attention should be paid to correct the overprediction generated from these models, and downscaled occupancies should be interpreted cautiously (Witte et al. 2008).

We must stress that any monitoring scheme involves some losses and gains of information and only a professional standardized (in terms of effort and spatial cover) project, could obtain optimal results (Engel and Voshell 2002). However, the cost of this kind of project is not always affordable whereas the acquisition of big amounts of data from simple volunteer, opportunistic monitoring schemes is relatively easy and inexpensive.

Conclusions

In general, an increase in sample size will give more accurate distribution models regardless of the monitoring source, indicating that data availability will likely constrain the predictive accuracy of species distribution modelling applications. However, in some cases models based on more standardized methods can provide the same accuracy with smaller sample size than models based on opportunistic data with bigger data sets. This leads to a trade-off between data quality and data collection costs, suggesting that in cases in which good quality data is available the use of this information for developing SDMs should be prioritised, since for comparable sample sizes information derived from carefully designed surveys appear to be more suitable for deriving species distribution maps than information derived from opportunistic data sources. We have also shown that opportunistic data sources may offer relatively good approximations especially in the case of predicting the distribution of uncommon species. In that case, opportunistic data gives similar accuracy and even a higher overlap with the real distribution than a more standardized survey. Also in that case, opportunistic data gives fewer omission errors but with higher commission. When modelling uncommon species or with very narrow ecological niches all monitoring methods show high accuracy.

Aggregation of heterogeneous data at a coarse scale and downscaling to a targeted spatial scale generated maps which tended to overpredict species distributions. Thus, although downscaling of data through a SDM approach appears useful in cases of data shortage or when data availability strongly differs in quality characteristics, it must be used with caution because will tend to overestimate species suitable habitats.

Acknowledgments This study was funded by the EU FP7 SCALES project ("Securing the Conservation of biodiversity across Administrative Levels and spatial, temporal and Ecological scales"; project #226852).

Appendix

See Appendix Table 4; Figs. 8 and 9.

Table 4 List of farmland bird species used to calibrate species distribution models from different monitoring data

Common name	Latin name	Code	Classification	Sample size
Roller	*Coracias garrulus*	COGAR	Uncommon	27
Greater short-toed lark	*Calandrella brachydactyla*	CABRA	Uncommon	35
Calandra lark	*Melanocorypha calandra*	MECAL	Uncommon	43
White stork	*Ciconia ciconia*	CICIC	Uncommon	53
Little bustard	*Tetrax tetrax*	TERAX	Uncommon	57
Common whitethroat	*Sylvia communis*	SYCOM	Uncommon	87
Tawny pipit	*Anthus campestris*	ANCAM	Uncommon	94
Ortolan bunting	*Emberiza hortulana*	EMHOR	Uncommon	112
Southern grey shrike	*Lanius meridionalis*	LAMER	Uncommon	158
Rock sparrow	*Petronia petronia*	PEPET	Uncommon	197
Thekla lark	*Galerida theklae*	GATHE	Uncommon	217
Eurasian skylark	*Alauda arvensis*	ALARV	Uncommon	264
Stock dove	*Columba oenas*	COOEN	Uncommon	284
Black-eared wheatear	*Oenanthe hispanica*	OEHIS	Uncommon	305
Common quail	*Coturnix coturnix*	COCOT	Uncommon	389
Spotless starling	*Sturnus unicolor*	STUNI	Common	487
Woodchat shrike	*Lanius senator*	LASEN	Common	503
Eurasian tree sparrow	*Passer montanus*	PAMON	Common	651
Red-legged partridge	*Alectoris rufa*	ALRUF	Common	672
Common kestrel	*Falco tinnunculus*	FATIN	Common	765
Crested lark	*Galerida cristata*	GACRI	Common	934
Linnet	*Carduelis cannabina*	CAINA	Common	962
Corn bunting	*Emberiza calandra*	EMCAL	Common	999
Common stonechat	*Saxicola torquata*	SATOR	Common	1,075
European turtle-dove	*Streptopelia turtur*	STTUR	Common	1,085
Common starling	*Sturnus vulgaris*	STVUL	Common	1,193
Eurasian hoopoe	*Upupa epops*	UPEPO	Common	1,212
Cirl bunting	*Emberiza cirlus*	EMCIR	Common	1,585
Barn swallow	*Hirundo rustica*	HIRUS	Common	1,894
European serin	*Serinus serinus*	SESER	Common	2,181

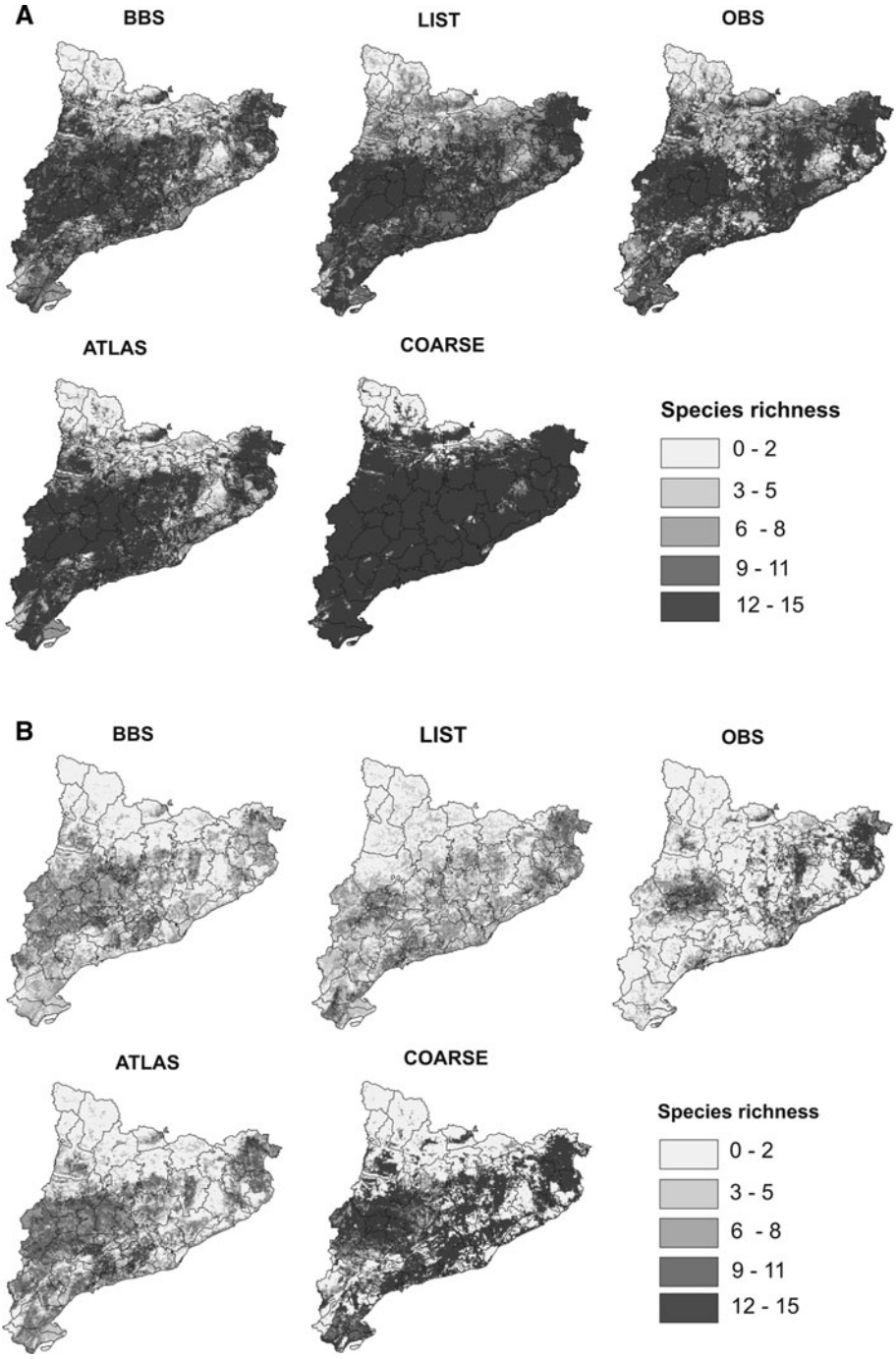


Fig. 8 Predictions of common farmland bird species richness in Catalonia obtained from different quality treatments (ATLAS, BBS, LIST, OBS and COARSE) from the whole sample size. Maps are show for **a** species distribution and **b** high quality areas

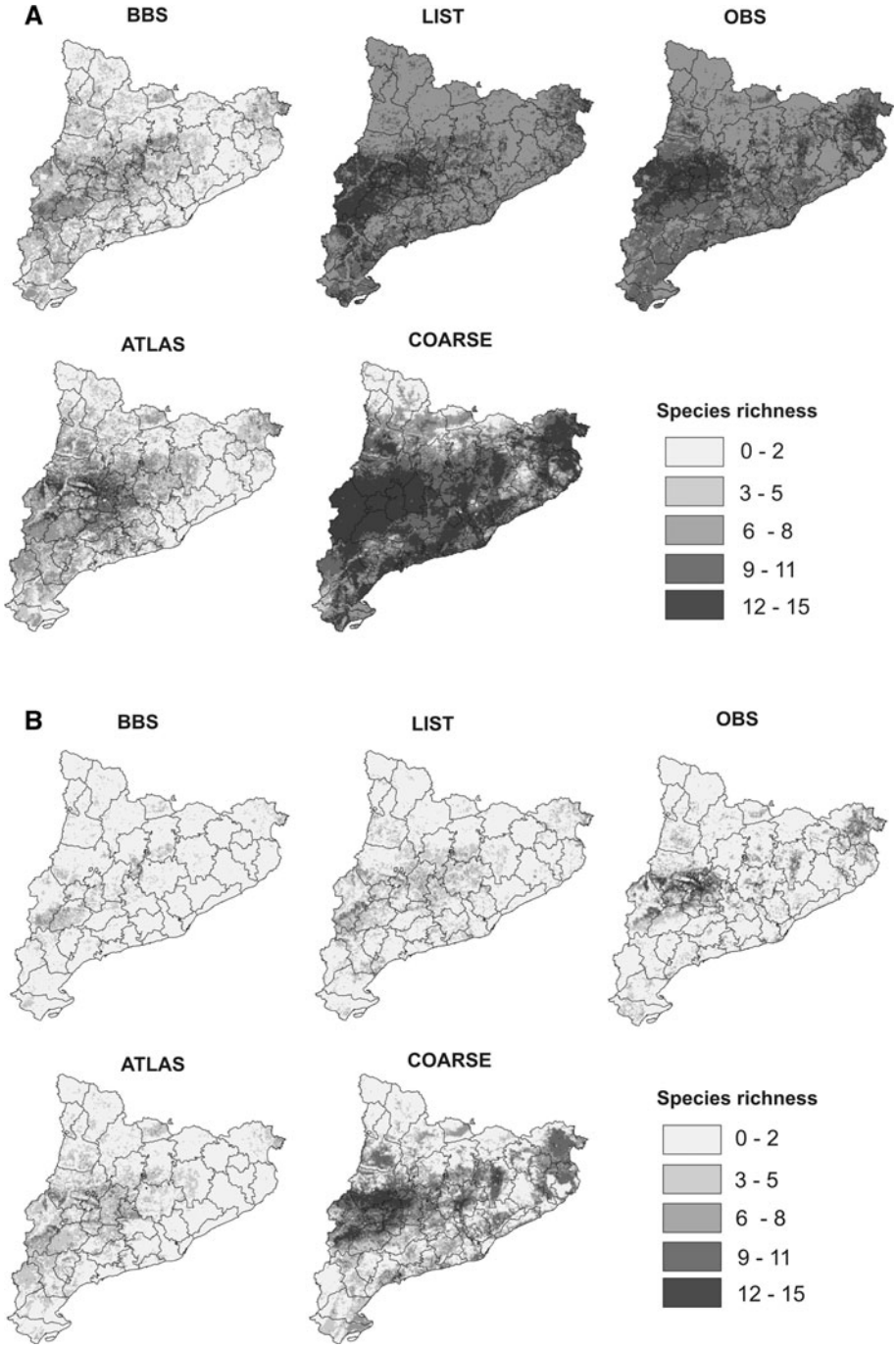


Fig. 9 Predictions of uncommon farmland bird species richness in Catalonia obtained from different quality treatments (ATLAS, BBS, LIST, OBS and COARSE) from the whole sample size. Maps are show for **a** species distribution and **b** high quality areas

References

Araújo MB, New M (2007) Ensemble forecasting of species distributions. Trends Ecol Evol 22:42–47

Araújo MB, Thuiller W, Williams PH et al (2005) Downscaling European species atlas distributions to a finer resolution: implications for conservation planning. Global Ecol Biogeogr 14:17–30

Boakes EH, McGowan PJK, Fuller RA et al (2010) Distorted views of biodiversity: spatial and temporal bias in species occurrence data. PLoS Biol 8:e1000385

Braunisch V, Suchant R (2010) Predicting species distributions based on incomplete survey data: the trade-off between precision and scale. Ecography 33:826–840

Brotons L, Mañosa S, Estrada J (2004a) Modelling the effects of irrigation schemes on the distribution of steppe birds in Mediterranean farmland. Biodivers Conserv 13:1039–1058

Brotons L, Thuiller W, Araújo MB et al (2004b) Presence–absence versus presence-only modelling methods for predicting bird habitat suitability. Ecography 27:437–448

Brotons L, Herrando S, Pla M (2007) Updating bird species distribution at large spatial scales: applications of habitat modelling to data from long-term monitoring programs. Diversity Distrib 13:276–288

Burfield I (2004) Birds in Europe. Population estimates, trends and conservation status. Birdlife International

Conrad CC, Hilchey KG (2011) A review of citizen science and community-based environmental monitoring: issues and opportunities. Environ Monit Assess 176:273–291

Danielsen F, Burgess ND, Balmford A (2005) Monitoring matters: examining the potential of locally-based approaches. Biodivers Conserv 14:2507–2542

Elith J, Graham CH, Anderson RP et al (2006) Novel methods improve prediction of species' distributions from occurrence data. Ecography 29:129–151

Elith J, Kearney M, Phillips SJ (2010) The art of modelling range-shifting species. Methods Ecol Evol 1:330–342

Engel SR, Voshell JR (2002) Volunteer biological monitoring: Can it accurately assess the ecological condition of streams? Am Entomol 48:164–177

Estrada J, Pedrocchi V, Brotons L et al. (2004) Atles dels ocells nidificants de Catalunya 1999–2002. Institut Català d'Ornitologia (ICO)/Lynx Edicions, Barcelona

Feeley KJ, Silman MR (2011) Keep collecting: accurate species distribution modelling requires more collections than previously thought. Diversity Distrib 1–9

Green RE, Cornell SJ, Scharlemann JPW et al (2005) Farming and the fate of wild nature. Science 307:550–555

Gregory RD, van Strien A, Vorisek P et al (2005) Developing indicators for European birds. Phil Trans R Soc B 360:269–288

Guisan A, Graham CH, Elith J et al (2007) Sensitivity of predictive species distribution models to change in grain size. Diversity Distrib 13:332–340

Hernández PA, Graham CH, Master LL et al (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods. Ecography 29:773–785

Herrando S, Brotons L, Estrada J et al (2008) The Catalan common bird survey (SOCC): a tool to estimate species population numbers. Revista Catalana d'Ornitologia 24:138–146

Herrando S, Brotons L, Estrada J et al (2011) Methodology. In: Herrando S, Brotons L, Estrada J et al (eds) Atles dels ocells de Catalunya a l'hivern 2006–2009. Institut Català d'Ornitologia/Lynx Edicions, Barcelona

Hirzel A, Guisan A (2002) Which is the optimal sampling strategy for habitat suitability modelling. Ecol Model 157:331–341

Kery M, Royle JA, Schmid H et al (2010) Site-occupancy distribution modeling to correct population-trend estimates derived from opportunistic observations. Conserv Biol 24:1388–1397

Liu C, Berry PM, Dawson TP et al (2005) Selecting thresholds of occurrence in the prediction of species distributions. Ecography 28:385–393

Lobo JM, Jiménez-Valverde A, Real R (2008) AUC: a misleading measure of the performance of predictive distribution models. Global Ecol Biogeogr 17:145–151

Mccarthy KP, Fletcher RJ, Rota CT, Hutto RL (2012) Predicting species distributions from samples collected along roadsides. Conserv Biol 26:68–77

McPherson JM, Jetz W, Rogers DJ (2006) Using coarse-grained occurrence data to predict species distributions at finer spatial resolutions—possibilities and limitations. Ecol Model 192:499–522

Munson MA, Caruana R, Fink D et al (2010) A method for measuring the relative information content of data from different monitoring protocols. Methods Ecol Evol 1:263–273

Ninyerola M, Pons X, Roure JM (2005) Atlas climático digital de la Península Ibérica: metodología y aplicaciones en bioclimatología y geobotánica ISBN 932860-8-7. Autonomous University of Barcelona, Bellaterra, p 44

Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. Ecol Model 133:225–245

Pearson RG, Raxworthy CJ, Nakamura M, Peterson AT (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. J Biogeogr 34:102–117

Peterson AT (2006) Uses and requirements of ecological niche models and related distributional models. Biodivers Inform 3:59–72

Phillips SJ, Dudik M (2008) Modeling of species distributions with Maxent: new extensions and a comprehensive evaluation. Ecography 31:161–175

Phillips SJ, Dudik M, Schapire RE (2004) A maximum entropy approach to species distribution modeling. Banff, Canada

Phillips SJ, Anderson RP, Schapire RE (2006) Maximum entropy modeling of species geographic distributions. Ecol Model 190:231–259

Phillips SJ, Dudík M, Elith J et al (2009) Sample selection bias and presence-only distribution models: implications for background and pseudo-absence data. Ecol Appl 19:181–197

Pons X (1996) Estimación de la radiación Solar a partir de modelos digitales de elevaciones. Propuesta metodológica. In: Juaristi J, Moro I (eds) VII coloquio de geografía cuantitativa. Sistemas de Información Geográfica y Teledetección, Vitoria

Roberts RL, Donald PF, Fisher IJ (2005) Wordbirds: developing a web-based data collection system for the global monitoring of bird distribution and abundance. Biodivers Conserv 14:2807–2820

Roberts RL, Donald PF, Green RE (2007) Using simple species lists to monitor trends in animal populations: new methods and a comparison with independent data. Anim Conserv 10:332–339

Sardà-Palomera F, Vieites DR (2011) Modelling species' climatic distributions under habitat constraints: a case study with *Coturnix coturnix*. Ann Zool Fennici 48:147–160

Schmeller DS, Henry PY, Julliard R et al (2009) Advantages of volunteer-based biodiversity monitoring in Europe. Conserv Biol 23:307–316

Schreuder HT, Gregoire TG, Weyer JP (2001) For what applications can probability and non-probability sampling be used? Environ Monit Assess 66:281–291

Segurado P, Araújo MB (2004) An evaluation of methods for modelling species distributions. J Biogeogr 13:1555–1568

Snall T, Kindvall O, Nilsson J et al (2011) Evaluating citizen-based presence data for bird monitoring. Biol Conserv 144:804–810

Soberon J, Peterson AT (2004) Biodiversity informatics: managing and applying primary biodiversity data. Philos Trans R Soc Lond Ser B 359:689–698

Stem C, Margoluis R, Salafsky N et al (2005) Monitoring and evaluation in conservation: a review of trends and approaches. Conserv Biol 19:295–309

Stockwell DRB, Peterson AT (2002) Effects of sample size on accuracy of species distribution models. Ecol Model 148:1–13

Vaughan IP, Ormerod SJ (2003) Improving the quality of distribution models for conservation by addressing shortcomings in the field collection of training data. Conserv Biol 17:1601–1611

Viñas O, Baulies X (1995) 1:250 000 Land-use map of Catalonia (32 000 km^2) using multi-temporal landsat-TM data. Int J Remote Sens 16:129–146

Wisz MS, Hijmans RJ, Peterson AT et al (2008) Effects of sample size on the performance of species distribution models. Divers Distrib 14:763–773

Witte JM, He F, Groen CLG (2008) Grid origin affects scaling of species across spatial scales. Global Ecol Biogeogr 17:448–456

Zaniewski AE, Lehmann A, Overton JMcC (2002) Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. Ecol Model 157:261–280