



Calibration of hybrid species distribution models: the value of general-purpose vs. targeted monitoring data

Miquel De Cáceres^{1*} and Lluís Brotons^{1,2,3}

¹Forest Science Center of Catalonia, Ctra. St. Llorenç de Morunys km 2, E-25280 – Solsona, Catalonia, Spain, ²Centre for Ecological Research and Applied Forestries (CREAF), Autonomous University of Barcelona, Bellaterra, Catalonia, Spain, ³European Bird Census Council (EBCC) & Institut Català d'Ornitologia (ICO), Museu de Zoologia, Passeig Picasso s/n, 08003 Barcelona, Catalonia, Spain

ABSTRACT

Aim Temporally replicated observations are essential for the calibration and validation of species distribution models (SDMs) aiming at making temporal extrapolations. We study here the usefulness of a general-purpose monitoring programme for the calibration of hybrid SDMs. As a benchmark case, we take the calibration with data from a monitoring programme that specifically surveys those areas where environmental changes expected to be relevant occur.

Location Catalonia, north-east of Spain.

Methods We modelled the distribution changes of twelve open-habitat bird species in landscapes whose dynamics are driven by fire and forest regeneration. We developed hybrid SDMs combining correlative habitat suitability with mechanistic occupancy models. We used observations from two monitoring programmes to provide maximum-likelihood estimates for spread parameters: a common breeding bird survey (CBS) and a programme specifically designed to monitor bird communities within areas affected by wildfires (DINDIS).

Results Both calibration with CBS and DINDIS data yielded sound spread parameter estimates and range dynamics that suggested dispersal limitations. However, compared to calibration with DINDIS data, calibration with CBS data leads to biased estimates of spread distance for seven species and to a higher degree of uncertainty in predicted range dynamics for six species.

Main conclusions We have shown that available monitoring data can be used in the calibration of the mechanistic component of hybrid SDMs. However, if the dynamics of the target species occur within areas not well covered, general-purpose monitoring data can lead to biased and inaccurate parameter estimates. To determine the potential usefulness of a given monitoring data set for the calibration of the mechanistic component of a hybrid SDM, we recommend quantifying the number of surveyed sites that are predicted to undergo habitat suitability changes.

Keywords

Biogeography, bird atlas data, hybrid SDM, maximum-likelihood optimization, metapopulation dynamics, range shifts.

*Correspondence: Miquel De Cáceres, Centre Tecnològic Forestal de Catalunya (CTFC), Ctra. St. Llorenç de Morunys km 2, E-25280 Solsona, Catalonia, Spain.
E-mail: miquelcaceres@gmail.com

INTRODUCTION

Realizing the potential impact of contemporary changes in climate, perturbation regime and land use on species distributions has fostered our need to make reliable predictions of species range shifts (La Sorte & Jetz, 2010). To date, a major approach taken to examine the effects of environ-

mental change on species distributions has been through the development of species distribution models (SDMs), which are empirical models relating field observations to environmental predictor variables (Guisan & Zimmermann, 2000; Franklin, 2009). Once parameterized, SDMs can be used in conjunction with the expected future (or known past) environmental conditions to forecast (or hind cast) the

distribution of species (Thomas *et al.*, 2004; Thuiller, 2004; Elith & Leathwick, 2009).

When correlative SDMs are employed for temporal extrapolation, the modeller is forced to assume constant equilibrium with environmental conditions. In other words, correlative SDMs do not distinguish between potential habitat and realized species distribution. Furthermore, correlative SDMs do not account for many ecological processes, such as local population dynamics, species interactions and dispersal, which may be key to determine changes in species distributions (Pearson & Dawson, 2003; Guisan & Thuiller, 2005; Dormann, 2007; Elith & Leathwick, 2009; Zurell *et al.*, 2009; Brotons *et al.*, 2012). In the recent years, several proposals have been made that combine correlative models with mechanistic models simulating local population and dispersal processes, in an attempt to improve the accuracy of predictions related to temporal changes in species distributions (Franklin, 2010; Gallien *et al.*, 2010). The conceptual basis of such *hybrid* SDMs is that while environmental conditions determine which species could potentially colonize and persist in a given site, local and spatial biotic processes ultimately determine species presence and abundance. Hybrid SDMs can be summarized in three steps: (1) information regarding environmental (e.g. vegetation, climatic or land use) spatial dynamics is either compiled or modelled (e.g. by means of a spatially explicit landscape model); (2) a correlative habitat suitability model is built and temporally projected using the environmental data; (3) a mechanistic occupancy or population model is built and run with the constraints imposed by the predicted habitat suitability changes. For example, Keith *et al.* (2008) and Anderson *et al.* (2009) coupled bioclimatic suitability series derived from climatic model predictions with spatially explicit stochastic population models to investigate the effect of various factors on the population viability and range shifts of several species. Sometimes, environmental dynamics are lacking and the modelling exercise focuses on predicting the range or population dynamics of the target species on a static habitat suitability configuration (Cabral & Schurr, 2010; Smolik *et al.*, 2010).

The predictive performance of any model aiming to accurately predict temporal changes in species distributions should be assessed using temporally replicated observations (Araújo *et al.*, 2005; Vallecillo *et al.*, 2009; Willis *et al.*, 2009). Unfortunately, the predictive performance of SDMs is usually evaluated using observations made during the same time period as the data used for model building because high-resolution censuses covering large spatial extents are difficult to repeat (Franklin, 2009). The availability of temporally replicated surveys is not only important for model evaluation, but also to estimate parameters in the model-building phase. Although the mechanistic components of hybrid SDMs are sometimes parameterized using expert knowledge, empirical or bibliographical data (Anderson *et al.*, 2009; Willis *et al.*, 2009), it is also possible to estimate parameters for a mechanistic model using the fit of model predictions to temporal observations (Cabral & Schurr, 2010; Smolik *et al.*, 2010), as

is the usual strategy in metapopulation models (e.g. Facon & David, 2006).

Long-term monitoring programmes, such as those available for birds or butterflies, are normally conducted with the aim of estimating temporal trends for populations of common and widespread species in the surveyed area. However, they offer a great deal of temporally replicated observations that have the potential to be used to create maps showing changes in species distribution and species richness (Jiguet *et al.*, 2005; Brotons *et al.*, 2007). Like all observations, monitoring data are imprecise, subject of sampling variations, measurement errors and biases (Kéry & Schmid, 2006). When the aim is to develop hybrid SDMs, the amount of information related to the ecological process driving distribution changes will likely compromise the usefulness of monitoring data for model calibration. A monitoring programme covering large extents (i.e. regional) but with large spatial lags between sampled locations may be adequate to follow ecological processes causing large-scale distribution changes, such as those derived from climate or land-use changes (Araújo *et al.*, 2005; Gregory *et al.*, 2005). However, such observational data are likely to fail to capture the species distribution dynamics if these occur at scales smaller than the spatial lag. In contrast with general-purpose regional surveys, targeted monitoring programmes specifically sample those sites where the environmental changes that are assumed to be ecologically relevant occur (e.g. Zozaya *et al.*, 2010). By their nature, these monitoring programmes potentially contain more useful information for model calibration. However, because these targeted surveys do not cover the whole region of interest with a regular density, one cannot be certain as to whether all the areas with relevant dynamics for the target species are indeed represented. In other words, programmes that monitor distribution changes driven by specific environmental factors may overlook distribution changes caused by other environmental drivers or that are the result of population dynamics within stable habitats.

In this study, we compare the usefulness of general-purpose vs. targeted monitoring data for building hybrid SDMs. We expect that a targeted monitoring programme should yield accurate and precise parameter estimates. As general-purpose monitoring programmes are currently conducted in many regions (Jiguet *et al.*, 2005) and targeted ones are quite rare, an interesting question is 'whether a general-purpose programme yields similar parameter estimates and distribution dynamics compared to those provided by a targeted one'. To address this question, we take the distribution dynamics of twelve Mediterranean open-habitat bird species whose distribution dynamics we assume to be driven by fire and forest regeneration processes. The two monitoring programmes to be compared are a common breeding bird survey and a programme specifically analysing bird colonization patterns following large wildfires (Fig. 1). We developed hybrid species distribution models and compared parameter estimates and distribution dynamics of models calibrated with either one monitoring programme or the other.

METHODS

Study system

Catalonia (NE Spain) is a Mediterranean region whose landscapes are remarkably heterogeneous as a result of sharp climatic and geological gradients. The diversity of landscapes in Catalonia ranges from alpine habitats to coastal marshes and from evergreen forests to steppes and agricultural mosaics. Because of experiencing hot and dry summers, fire periodically affects extensive areas in Catalonia (Díaz-Delgado *et al.*, 2004; Gil-Tena *et al.*, 2009). Studies of bird communities have shown that burned areas are pioneered by bird species preferring open and low-vegetation habitat structure (Brotons *et al.*, 2005). As shrub encroachment and forest maturation progresses, however, these open-habitat species are forced to move to other areas with open structure. Species with strict open-habitat requirements can only survive regionally at the metapopulation level by dispersing far enough to allow the establishment of new populations compensating local extinctions (Keymer *et al.*, 2000; Brotons *et al.*, 2005, 2012; Vallecillo *et al.*, 2007). Among the *c.* 214 species of birds that commonly breed in Catalonia, in this study we modelled the distribution dynamics of twelve open-habitat species (Table 1). We selected these twelve species because they have been frequently detected in surveys of a monitoring programme that tracks bird colonization after wildfires. Some of these species are very selective in their preference for bare soils and burned areas, whereas others occur in a broader range of open habitats including stable shrublands, dry croplands and grasslands. Accordingly, the prevalence of the twelve species in the study area ranges from 3% to 49% (Table 1).

Overview of the modelling approach

We developed hybrid species distribution models to predict distribution changes in Catalonia for the period from 2000 to 2009, with the following components (Fig. 2): (1) a landscape dynamics model used to reproduce historical landscape changes based on the interplay between historic fires and simulated forest regeneration; (2) a first correlative model providing the initial probability of occupancy, which was also taken as the best approximation to initial habitat suitability; (3) a second correlative model that allowed us to track habitat suitability changes derived from fire and forest regeneration processes; (4) an occupancy model that updated the probability of occupancy taking into account changes in habitat suitability and the spread capability of the target species. We estimated the spread-related parameters of the occupancy model by maximizing the fit to field observations obtained from bird monitoring programmes. In what follows, we detail our modelling choices for each of these steps.

Initial forest map

We gathered and combined information from available forestry, land cover and fire data (Ministerio de Medio Ambiente 2006) to identify forested areas with different dominant species and non-forested open areas such as shrublands, croplands or grasslands. The result was a raster forest map that described the landscapes of the entire Catalan territory (32,114 km²) in year 2000 by classifying 100 × 100 m cells into eleven forest types and five land-use classes.

Table 1 Twelve open-habitat bird species included in the study: (a) Prevalence in the Catalan territory assessed using Catalan Breeding Bird Atlas observations; (b) Probability threshold for the presence of the species in the initial probability of occupancy (IPO) model; (c) Population trends assessed using the mean annual population change estimated from Common bird survey monitoring programme (period 2002–2009) and its significance at $P < 0.05$ (ns or *) (source: Servei d'Informació Ornitològica de Catalunya, <http://www.sioc.cat/>); (d) Predictive accuracy, measured using the area under the ROC curve (AUC; Fielding & Bell, 1997) of the initial probability of occurrence (IPO) model, as reported in Brotons *et al.* (2007); (e) Predictive accuracy (AUC) of the habitat suitability (HS) models, where the predicted values for the data used for model calibration were obtained by sequentially removing blocks of 20% of the data.

Species name	Code	(a) Prevalence	(b) Presence threshold	(c) Population trend	(d) IPO (AUC)	(e) HS (AUC)
<i>Anthus campestris</i>	ANCAM	3%	0.010	−11% (*)	0.822	0.727
<i>Carduelis cannabina</i>	CAINA	32%	0.101	−5% (*)	0.838	0.721
<i>Emberiza cia</i>	EMCIA	23%	0.096	2% (ns)	0.889	0.791
<i>Emberiza hortulana</i>	EMHOR	4%	0.012	0% (ns)	0.834	0.814
<i>Galerida theklae</i>	GATHE	7%	0.019	−5% (ns)	0.882	0.858
<i>Hippolais polyglotta</i>	HIPOL	27%	0.092	5% (*)	0.788	0.641
<i>Lullula arborea</i>	LUARB	30%	0.114	5% (*)	0.846	0.704
<i>Merops apiaster</i>	MEAPI	36%	0.161	6% (*)	0.848	0.763
<i>Oenanthe hispanica</i>	OEHS	10%	0.036	−3% (*)	0.925	0.855
<i>Sylvia cantillans</i>	SYCAN	29%	0.104	3% (*)	0.873	0.726
<i>Sylvia melanocephala</i>	SYMEL	49%	0.199	1% (ns)	0.927	0.716
<i>Sylvia undata</i>	SYUND	14%	0.034	1% (ns)	0.858	0.817

Landscape dynamics

We obtained fire information for the period 2000–2009 from the historical sequence of fires available in Catalonia (<http://www.gencat.cat/dmah>). As we lacked the information that would be necessary to describe vegetation regeneration following fires events, we had to simulate vegetation changes using a spatially explicit landscape dynamics model implemented in the spatially explicit landscape event simulator (SELES) modelling platform (Fall & Fall, 2001). Historical fire events were directly forced in the landscape dynamics model. Forest maturation was tracked using the years since the last fire. Post-fire changes in forest composition (i.e. changes in the dominant tree species of the forest or changes to shrubland) and the probabilities for those transitions were obtained using bibliographic sources (Rodrigo *et al.*, 2004). We induced spatial autocorrelation in post-fire forest composition changes by coupling 40% of the transitions of target cells with those of their eight surrounding neighbours. Succession was only modelled for shrublands (i.e. forests were assumed to be in a stable state). Shrublands originated from past fire events had a yearly probability of changing to forest after 20 years. In cases where a transition from shrubland to forest occurred, the resulting forest type was chosen using a multinomial distribution with probabilities obtained from the distribution of mature forest types in the surroundings cells. Details of the landscape dynamics model can be found at <http://sites.google.com/site/medfireproject/>. Preliminary results showed very low variability in simulated landscape dynamics resulting from deterministically forcing historical fires and considering a short temporal period. We therefore only used one model execution to generate landscape dynamics in this study. Running the landscape dynamics model, we obtained a sequence of forest maps that predicted yearly landscape changes between 2001 and 2009 (Fig. 2).

Observational data for the initial distribution

We used field observations collected for the Catalan Breeding Bird Atlas (Estrada *et al.*, 2004) as data to characterize the distribution of our twelve species in year 2000. This atlas was created after a large-scale survey that covered the whole of the Catalan territory between 1999 and 2002 using a grid of 10×10 km cells. This coarse sampling was complemented with a finer subsampling consisting of surveys in ten randomly chosen 1×1 km² within each 10×10 km grid cell. In this study, we used the presence–absence records corresponding to the finer sampling ($3076 \times 1 \times 1$ km²), from which we calculated the species prevalence values given in Table 1.

Observational data for distribution dynamics

We considered two distinct data sources to characterize the distribution dynamics of our twelve species for the period 2001–2009:

1. We obtained field observations describing spatial dynamics at the regional scale from the common bird monitoring survey (CBS) of Catalonia (Brotons *et al.*, 2007). Volunteer observers conduct surveys in this general-purpose programme, which has been run annually since 2002. Sampling is conducted using 3 km transects divided into six 500-m sections, in which all birds heard or seen at different distance bands are counted. Transects are distributed throughout Catalonia and located within specific 10×10 km squares (Fig. 1a). We provide the 2002–2009 population trends of the twelve species considered in this study in Table 1.
2. We obtained field observations describing spatial dynamics within burned areas from a monitoring programme that started in 2006 and is specifically oriented to investigate bird species responses to wildfires (DINDIS) (Zozaya *et al.*, 2010). Only wildfires that occurred after year 2000 and whose area burned

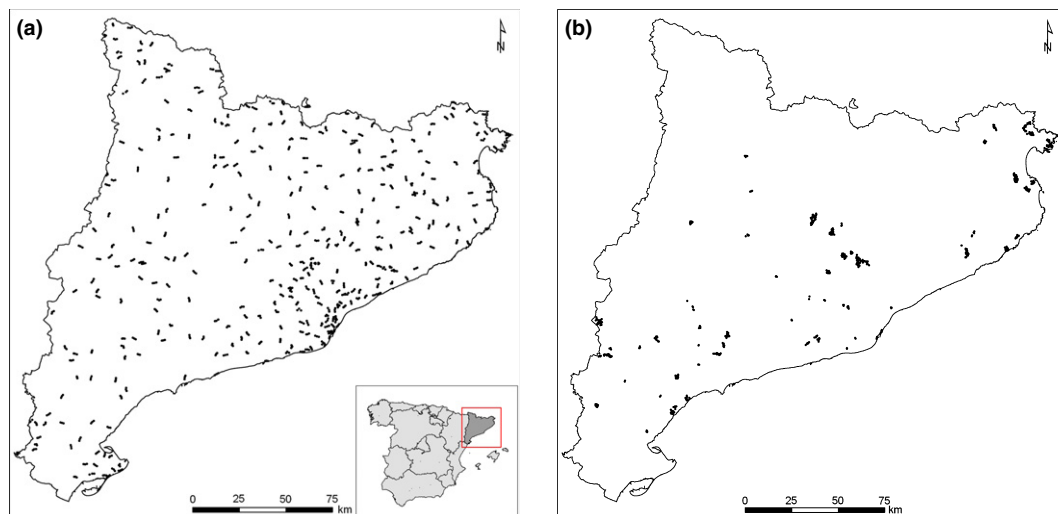


Figure 1 Geographical location of 1×1 km cells surveyed in (a) the Catalan Common Bird Survey (CBS) (937 cells) and (b) in the DINDIS fire-targeted monitoring programme (310 cells).

was greater than 50 ha are monitored (Fig. 1b). Surveys are conducted every year until ten years after the fire event. The number of transects is proportional to the logarithm of burned area, and observers have to travel along linear transects for 15–20 min (c. 500 m in length) and count all birds heard or seen within 100 m distance bands (Zozaya *et al.*, 2010).

To match the resolution of the Catalan Breeding Bird Atlas data, we transformed bird counts into species presence–absence records within 1×1 km cells. For each species, the CBS data set included 4905 observations corresponding to 937 cells, whereas the DINDIS data set included 1005 observations corresponding to 310 cells.

Initial probability of occupancy

As an estimation of the species distribution in year 2000, we used the correlative species distribution models developed at 1×1 km resolution by Brotons *et al.* (2007) from the Catalan Breeding Bird Atlas data. These are generalized linear models with binomial responses that include a broad set of environmental information (i.e. landscape, climate, topography variables) as well as the spatial structure of the response variable (i.e. using auto-covariates or other spatial predictors in the model) (McCullagh & Nelder, 1989; Augustin *et al.*, 1996; Lichstein *et al.*, 2002). We took the predicted probability values as the initial probability of occupancy (IPO) in our hybrid modelling (Fig. 2). We also calculated a threshold for species presence by first selecting the grid cells where the species had been detected and then calculating the average probability across those cells belonging to the 10% quantile of probability values (Table 1). Using these presence thresholds, we distinguished ‘initially occupied’ from ‘initially unoccupied’ sites.

Habitat suitability

To estimate changes in habitat suitability, we built a second set of correlative models where the explanatory factors were

landscape environmental variables and without any information on the spatial structure of the response. We fitted generalized linear models with binomial error (McCullagh & Nelder, 1989) using the presence–absence records of the Catalan Breeding Bird Atlas as response variable and environmental layers derived from the initial forest map (details are given in Appendix S1 in Supporting Information) (Fig. 2). Using these habitat suitability (HS) models and the sequence of maps produced by the landscape dynamics model, we predicted habitat suitability dynamics at 1×1 km resolution following changes introduced by fire and succession processes (Fig. 2). Whereas the HS models could be temporally projected, the IPO models provided a more accurate assessment of the initial distributions than the HS models [compare columns (d) and (e) in Table 1]. We therefore took the IPO models as the valid approximation for the ‘true’ habitat suitability in year 2000 and used changes in HS values to model suitability increases or decreases. This led us to the following expression for what we call here combined habitat suitability (CHS):

$$\begin{cases} \text{CHS}_{2000} = \text{IPO} \\ \text{logit}(\text{CHS}_t) = \text{logit}(\text{CHS}_{t-1}) + (\text{logit}(\text{HS}_t) - \text{logit}(\text{HS}_{t-1})) \end{cases} \quad (1)$$

The logit function was used to ensure that the arithmetic operations on probabilities produced values ranging between 0 and 1. We used the CHS probability values (and not the HS values) as estimates of habitat suitability in our hybrid SDMs (Fig. 2).

Occupancy model

We modelled changes in occupancy distribution using a simple grid-based model, implemented in R language, that updates the probability of occupancy in 1×1 km cells taking into account changes in habitat suitability and the spread capability of the target species. As input, the procedure needs initial values for the probability of occupancy (IPO), the habitat

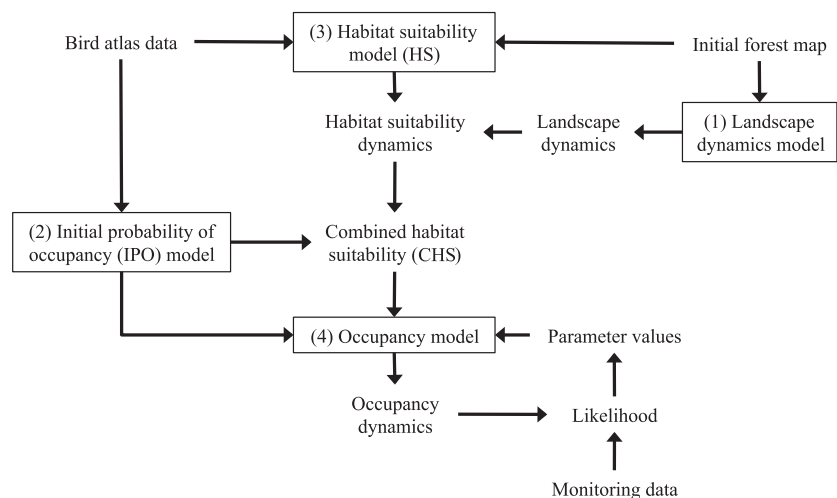


Figure 2 Schematic diagram showing the components of our approach to develop and calibrate hybrid species distribution models.

suitability (CHS) values corresponding to all time steps and the values of three parameters. Although it deals with probability values, the model is deterministic. That is, for a given probability of occupancy, the model calculates the probability of occupancy at the next time step. The probability of occupancy at cell i and time t is given by:

$$P_{i,t} = \min(\text{CHS}_{i,t}, P_{i,t-1} + (1 - P_{i,t-1}) \cdot C_{i,t}) \quad (2)$$

where $C_{i,t}$ is a coefficient describing the rate of increase in probability of occupancy derived from either local growth or colonization from neighbouring occupied cells. Decreases in habitat suitability (CHS) automatically reduce occupancy, but habitat suitability increases lead to increases in occupancy only if $C_{i,t} > 0$. The value of C is modelled much like in the incidence function model (Hanski, 1994). First, a connectivity value for the focal cell is determined using a kernel function that represents the pressure of dispersers that the focal cell receives from its occupied neighbours. For the kernel function, we used an extended negative exponential function, which is similar to the more common negative exponential function but is more flexible because it allows fat-tailed distributions (Chapman *et al.*, 2007). The formula for the connectivity value for cell i at time t is given by:

$$S_{i,t} = \sum_{j \in \Omega} P_{j,t-1} \cdot e^{-\alpha \cdot d_{ij}^\beta} \quad (3)$$

Where $P_{j,t-1}$ is the probability of occupancy at time $t-1$ of the cell j , which belongs to a set Ω of neighbours of the focal cell i , d_{ij} is the distance between cells i and j , and α is a parameter indicating the rate of exponential decay. Large values of α allow short-distance colonization only, whereas small α values allow long-distance colonization events. Parameter β further modulates the shape of the kernel function. When $\beta = 1$, the kernel is a negative exponential kernel. When $\beta < 1$ (respectively $\beta > 1$), the tail of the kernel will be longer (respectively shorter) than the one of the negative exponential kernel. For computational reasons, we used a bounding box (Ω in equation 3) to limit the number of neighbours to consider. Only the neighbours for which $e^{-\alpha \cdot d_{ij}^\beta} > 0.05$ were included in the bounding box. The rate of increase in occupancy C is determined using an increasing sigmoid function of connectivity:

$$C_{i,t} = \frac{S_{i,t}^2}{S_{i,t}^2 + \gamma^2} \quad (4)$$

where γ is a parameter that dictates how much connectivity is needed to increase occupancy. For the same connectivity value, a smaller γ value leads to a larger rate of increase in occupancy, and hence the spread rate of the species.

Calibration data sets and relevant information content

We used either CBS data (2002–2008) or DINDIS data (2006–2008) to estimate the spread parameters of the occupancy model. Although the CBS calibration data set included a much larger number of observations (4169 vs. 764) and spanned a

longer time period (7 vs. 3 years) than the DINDIS data set, we did not even out these differences. Instead, we tried to quantify the amount of relevant information that the two data sets contained for model calibration. As the three parameters to be estimated (α , β and γ) were related to the increase in probability of occupancy, we counted the number of cells that were predicted to experience habitat suitability increases, among those that were surveyed in the monitoring programme. Specifically, we first calculated the difference between CHS_{2000} and CHS_{2008} values for surveyed cells, and then, we counted the number of cells for which this change was positive. Among the surveyed cells predicted to increase in habitat suitability, we distinguished between those that were ‘initially unoccupied’ (representing new habitat) and those that were ‘initially occupied’ (representing habitat quality improvement), according to the presence thresholds of Table 1.

Maximum-likelihood estimation

We estimated the spread-related parameters of the occupancy model using a maximum-likelihood approach (Cabrál & Schurr, 2010; Smolik *et al.*, 2010). More specifically, we searched for parameter estimates whose log-likelihood was maximal, given the set of presence–absence records in the calibration data set. We used the Nelder–Mead simplex algorithm (as implemented in the R function ‘optim’) for the numerical maximization of the log-likelihood function, setting the initial estimates to $\alpha = \beta = \gamma = 1$. To assess the degree of uncertainty in parameter estimates, we ran maximum-likelihood estimation using both the complete calibration data set and bootstrap samples of it. Bootstrap samples were obtained by sampling with replacement the 1×1 km cells included in the monitoring data. Because of the slowness of likelihood optimization process in R, we only generated ten bootstrap samples for each monitoring data set. Although such a small number of bootstrap samples does not allow for an accurate assessment of the variability of parameter estimates, it provides an indication of how much the parameter estimates, and the resulting range dynamics, are affected by sampling variability in the data used for model calibration.

Range dynamics and predictive performance

We calculated the predicted species range for each year by adding the probability of occupancy over all grid cells of the study area. Those occupancy models that had been calibrated with bootstrap samples were used to assess the degree of uncertainty in predicted range dynamics.

We assessed the predictive accuracy of the occupancy models using 2009 observations coming from the two monitoring programmes pooled together. Whenever the two programmes had surveyed the same grid cell, we took the maximum value. We conducted two evaluations of predictive accuracy, one including all grid cells with observations and another restricted to ‘initially unoccupied’ cells for which habitat suitability was predicted to increase. We calculated

predictive accuracy using the coefficient of determination (R^2 ; Nagelkerke, 1991), in which we compared the likelihood of the distribution given by the occupancy model with the likelihood of the distribution given by the initial probability of occupation (IPO), taken as a static null model.

RESULTS

Suitability increases within monitoring locations

The landscape dynamics model produced an overall increase of forested areas because the rate of appearance of vegetation with open structure created by 2000–2009 wildfires did not overcome the rate of forest regeneration recovering from previous fire events (large wildfires occurred in the 1980s and 1990s). Accordingly, the CBS calibration data set included a small number of cells with predicted habitat suitability increase (Fig. 3) (between 3% and 14% of the surveyed cells, depending on the species). In other words, the quality of habitat for open-habitat birds was predicted to decrease in most of the study region. In accordance with the hypothesis that open-habitat species generally tend to benefit from the resources available in recently burned areas, the DINDIS calibration data set included many grid cells with predicted increase in habitat suitability (Fig. 3) (between 42% and 73% of the surveyed cells, depending on the species). Moreover, we found a negative relationship between the number of new habitat cells surveyed in DINDIS and the prevalence of the target species ($r = -0.769$, P -value = 0.0123, $n = 12$). Clearly, fire events created opportunities for low prevalence species to colonize new areas, whereas for widespread species, fire improved the suitability of already occupied habitats, rather than creating new habitat suitable for colonization.

Estimated spread kernels

To facilitate the comparison of the estimated kernel functions ($e^{-\alpha \cdot d_{ij}^{\beta}}$) across occupancy models, we calculated the distance that accounted for 90% of kernel values (D_{90} , the 90th quantile

of the extended negative exponential distribution) and categorized kernel functions between those leading to short ($D_{90} \leq 2$ km), moderate ($2 \text{ km} < D_{90} < 20$ km) and long ($D_{90} > 20$ km) spread distances. For five species (*Carduelis cannabina*, *Emberiza hortulana*, *Galerida theklae*, *Merops apiaster* and *Sylvia melanocephala*), calibration using DINDIS or CBS data yielded similar kernels according to this categorization (Table 2). In contrast, for five out of the remaining seven species, calibration with DINDIS data leads to longer spread distances than calibration with CBS data (*Anthus campestris*, *Lullula arborea*, *Oenanthe hispanica*, *Sylvia undata* and *Hippolais polyglotta*). Hence, spread capabilities were often underestimated when calibration was conducted using CBS data. Interestingly, those species with higher prevalence in the study area generally obtained lower estimates of spread distance, both for models calibrated with DINDIS data ($r = -0.46$, P -value = 0.136, $n = 12$) and those calibrated with CBS data ($r = -0.63$, P -value = 0.029, $n = 12$).

Predicted range dynamics

Following the predicted decrease in availability of open vegetation habitats, occupancy models predicted a decrease in range for all the species studied (Table 2 and Fig. 4). We used the range dynamics of models calibrated with bootstrap samples to determine whether models calibrated with CBS data produced similar range dynamics to models calibrated with DINDIS data. In only four species (*A. campestris*, *C. cannabina*, *Emberiza cia* and *G. theklae*), the range dynamics of models calibrated with CBS data was not distinct from the range dynamics of models calibrated with DINDIS data (Fig. 4). Moreover, for six species, the bootstrap estimates of predicted range dynamics were significantly more variable for models calibrated with CBS data than for models calibrated with DINDIS data, whereas the reverse situation occurred for only two species (Table 2). Thus, predictions of range dynamics were often more uncertain when calibrated with CBS data, despite the fact that this data set encompassed a longer time span and a larger number of surveyed locations.

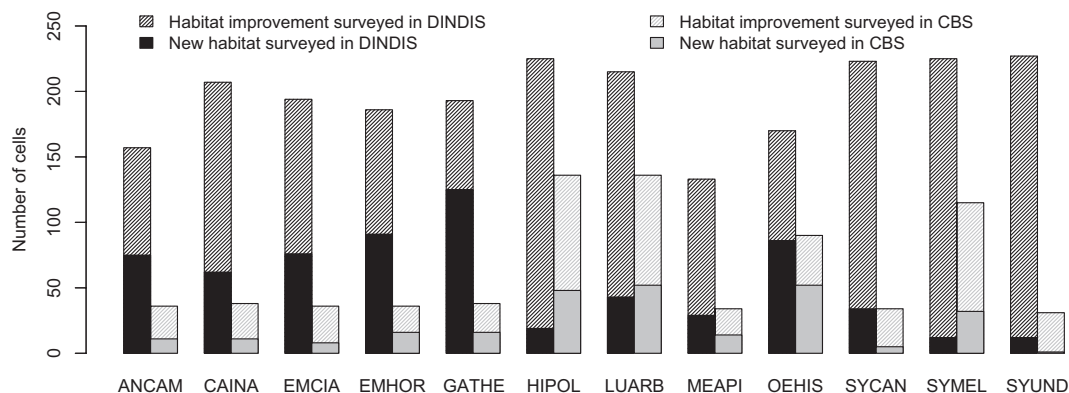


Figure 3 Number of 1×1 km cells with predicted increase in habitat suitability (CHS) for the calibration period (2000–2008), among those surveyed in the DINDIS (left bars) and CBS programmes (right bars). For each bar, we distinguish between initially occupied cells (upper parts in shaded lines) and initially unoccupied cells (lower parts). Species names are given in Table 1.

Table 2 Model calibration results (a) and predicted range dynamics (b) for each species and calibration data set (either CBS 2002–2008 or DINDIS 2006–2008). Calibration results include maximum-likelihood value (*LogLik*), maximum-likelihood estimates ($\hat{\alpha}$, $\hat{\beta}$, $\hat{\gamma}$), as well as the distance accounting for 90% of kernel values (D_{90}) and the categorization of kernel functions into those leading to *short* ($D_{90} \leq 2$ km), *moderate* ($2 \text{ km} < D_{90} < 20$ km) and *long* ($D_{90} > 20$ km) spread distances. As for range dynamics, we report the difference in the predicted range between 2000 and 2009 (*Diff*) and its estimated standard deviation (SD) that we assessed using bootstrap samples.

Species name	Calibration data set	(a) Calibration results						(b) Range dynamics	
		LogLik	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	D_{90}	Categ.	Diff	SD
<i>Anthus campestris</i>	DINDIS	-486	0.025	1.195	1.562	38.24	Long	-100	0.8*
	CBS	-508	0.609	1.080	1.087	3.24	Moderate	-130	11.2*
<i>Carduelis cannabina</i>	DINDIS	-510	14.72	7.554	2.297	0.62	Short	-726	6.8*
	CBS	-1371	9351	3557	3545	0.90	Short	-761	36.4*
<i>Emberiza cia</i>	DINDIS	-341	0.803	3.267	1.874	1.03	Short	-215	3.3*
	CBS	-1389	0.785	0.752	1.551	5.54	Moderate	-201	14.3*
<i>Emberiza hortulana</i>	DINDIS	-322	0.031	2.268	2.045	5.04	Moderate	-234	2.2*
	CBS	-500	0.728	1.226	0.865	2.25	Moderate	-250	11.5*
<i>Galerida theklae</i>	DINDIS	-430	1.169	0.401	1.166	30.06	Long	-246	59.7
	CBS	-411	1.242	0.341	1.061	65.09	Long	-189	80.0
<i>Hippolais polyglotta</i>	DINDIS	-507	3.603	0.268	3.323	7.37	Moderate	-196	0.6*
	CBS	-1977	6.433	2.604	-0.05	0.5	Short	-180	10.5*
<i>Lullula arborea</i>	DINDIS	-523	3.179	0.135	1.691	1431	Long	-419	30.6*
	CBS	-1568	0.052	3.808	8.330	2.03	Moderate	-484	10.5*
<i>Merops apiaster</i>	DINDIS	-478	6.771	16.95	0.810	0.76	Short	-346	0.5*
	CBS	-1859	9351	3557	3545	0.90	Short	-381	8.4*
<i>Oenanthe hispanica</i>	DINDIS	-535	2.412	0.187	2.506	453.5	Long	-379	55.6
	CBS	-462	1.241	0.478	2.825	11.62	Moderate	-421	25.5
<i>Sylvia cantillans</i>	DINDIS	-411	4.287	1.881	2.681	0.55	Short	-251	12.2
	CBS	-1829	0.676	0.816	16.31	5.41	Moderate	-220	9.4
<i>Sylvia melanocephala</i>	DINDIS	-548	0.0003	28.24	46.58	1.16	Short	-451	10.2*
	CBS	-1856	2.581	1.082	0.047	0.85	Short	-382	1.9*
<i>Sylvia undata</i>	DINDIS	-571	1.489	0.353	34.10	31.63	Long	-635	17.1
	CBS	-765	1.663	0.452	2.972	7.69	Moderate	-609	10.3

Asterisks (*) indicate statistical significant differences between calibration data sets in a variance ratio test.

Predictive accuracy

For several species, the distribution predicted by dynamic occupancy models was more in accordance with 2009 field observations than the distribution given by the initial probability of occupancy (IPO) model (Fig. 5a). When focusing on those unoccupied cells predicted to become suitable during the period studied, the difference in predictive accuracy was higher because of the fact that a static distribution did not allow the colonization of these locations (Fig. 5b). The best models to predict colonization of newly burned sites were obtained for low prevalence species (*A. campestris*, *E. hortulana*, *G. theklae* and *O. hispanica*). In contrast, for some widespread species, the static distribution model was more accurate than the calibrated dynamic models, and we attribute those results to misspecifications of the habitat suitability model that lead to incorrect predictions of habitat suitability increase or decrease (see AUC values for the HS model of *H. polyglotta*, *L. arborea* and *S. melanocephala* in Table 1). Although in some species, models calibrated with DINDIS data were more accurate than models calibrated with CBS data, we cannot conclude that calibration with general-purpose monitoring data generally

produced less accurate predictions (Fig 5a,b). A longer simulation period is probably necessary to find statistical differences in predictive accuracy.

DISCUSSION

General-purpose monitoring and the development of dynamic SDMs

Atlas surveys combine high-density of sampling locations with fine-resolution species information. Although these kinds of surveys provide an excellent way to analyse the distribution of species for a single time period, they are rarely repeated (Donald & Fuller, 1998; Robertson *et al.*, 2010). In comparison, monitoring programmes cover the region of interest with a lower density of sampling locations, but surveys are repeated at shorter time intervals (Brotons *et al.*, 2007). Although correlative SDMs are usually more accurate when developed from atlas data (Brotons *et al.*, 2007), modellers attempting to predict distribution changes at regional/national scales could profit from both data sources as we have performed here. Atlas data would be used to provide an accurate estimation of the species initial

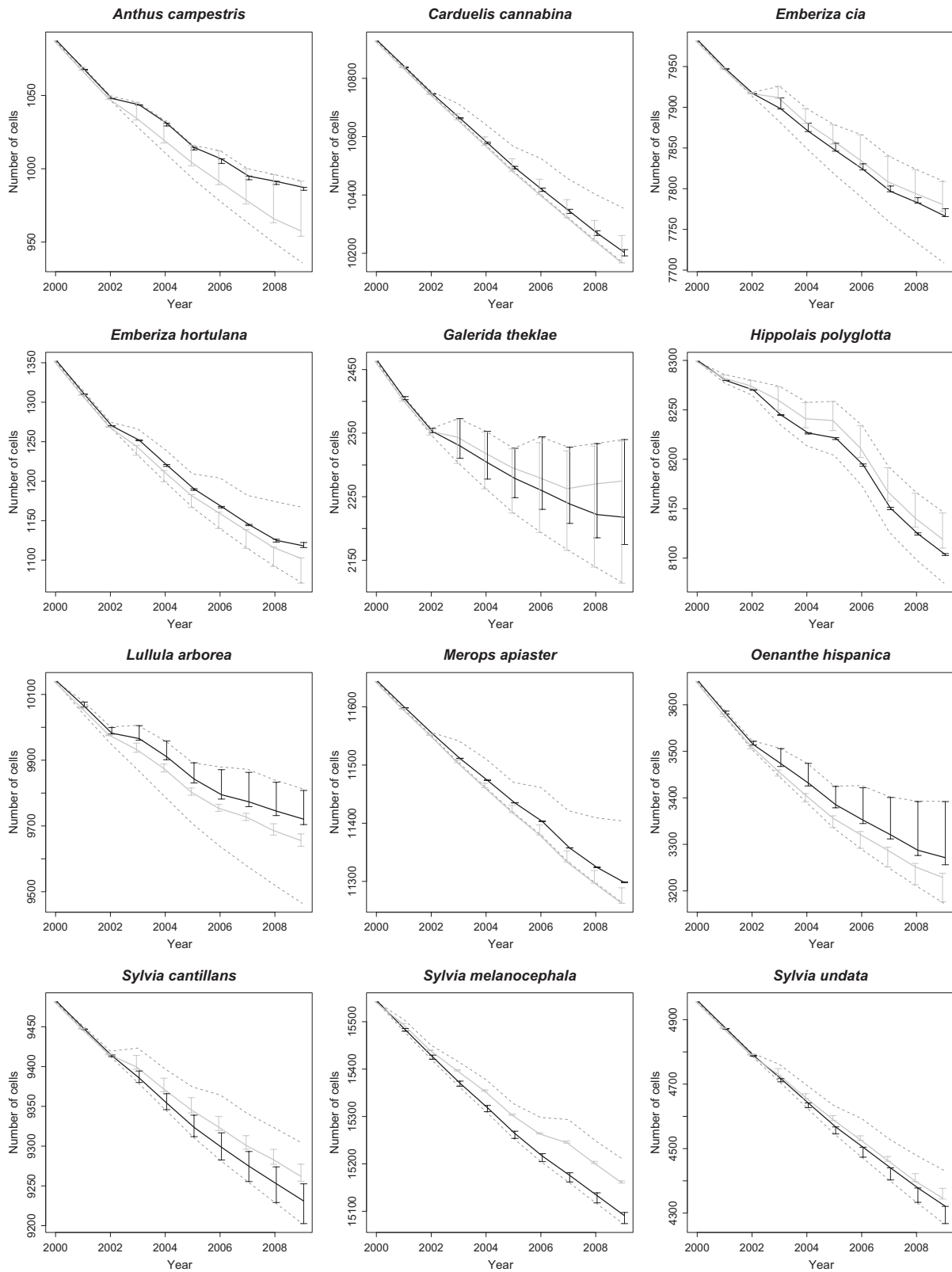


Figure 4 Predicted range dynamics according to the models calibrated using DINDIS data (in black) or CBS data (in grey). Intervals indicate minimum and maximum values among models calibrated using bootstrap samples. For comparison, we include also the range dynamics corresponding to two simple scenarios. *No spread scenario* (lower bound in dashed grey): Local populations go extinct following habitat suitability decreases, and new suitable habitat is never colonized (equivalent to setting $C_{i,t} = 0$ in the occupancy model); *Unlimited spread scenario* (upper bound in dashed grey): The target species is allowed to spread without limits (equivalent to setting $C_{i,t} = 1$ in the occupancy model).

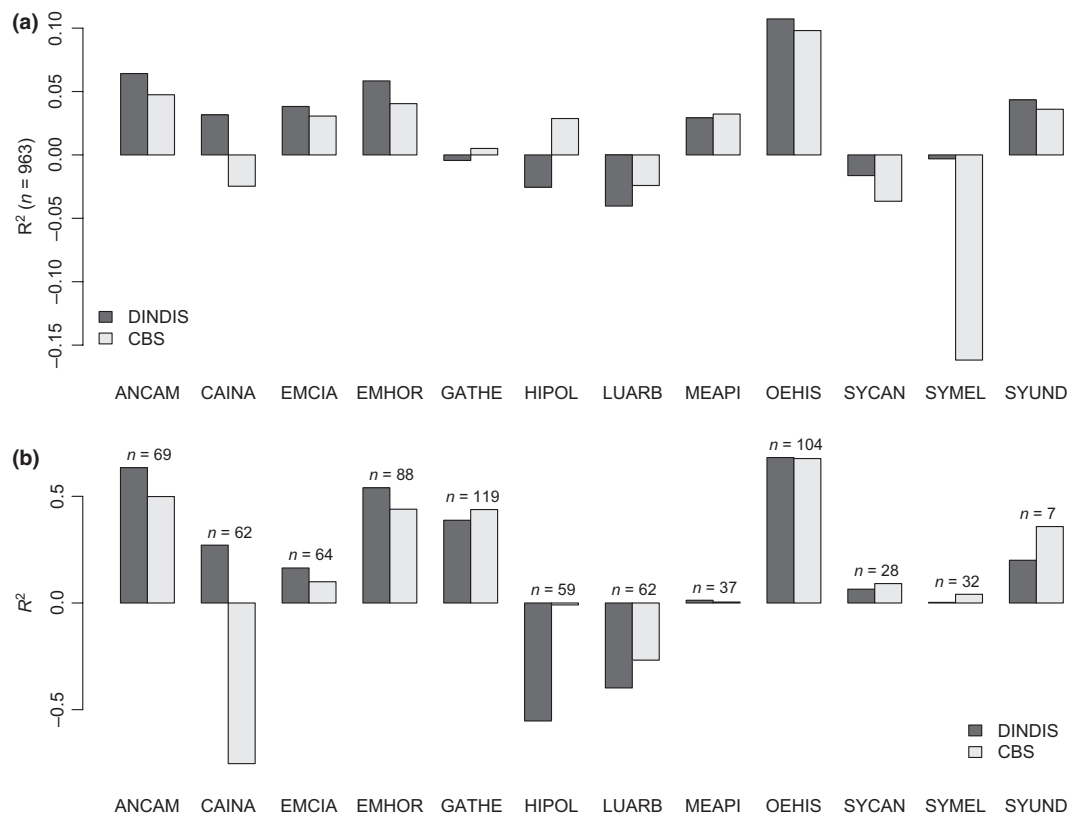


Figure 5 Accuracy of occupancy models calibrated with either DINDIS or CBS data to predict observations corresponding to year 2009. In (a) the evaluation data set included all monitored cells ($n = 963$), while in (b) the evaluation data set included only those initially unoccupied cells predicted to undergo a suitability increase (the number of cells is indicated for each species). Predictive accuracy was assessed using the coefficient of determination (R^2 , Nagelkerke, 1991). Positive R^2 values indicate a better fit to observations than the initial probability of occurrence (IPO) (the static null model), whereas negative R^2 values indicate a poorer predictive performance. Species names are given in Table 1.

distribution and habitat requirements. Complementarily, monitoring data would be used to evaluate the model dynamic outputs (Araújo *et al.*, 2005; Willis *et al.*, 2009) and to estimate parameters that modulate distribution dynamics (Cabral & Schurr, 2010; Smolik *et al.*, 2010). Instead of calibrating habitat suitability and mechanistic components of hybrid models in separate modelling steps, as we did, an alternative and promising approach is to estimate parameters for species–environment relationships and spatial dynamics in the same fitting exercise (Wikle, 2003; Pagel & Schurr, 2011). For example, Pagel & Schurr (2011) used a hierarchical Bayesian framework to calibrate a spatial population dynamics model whose population growth rate varies in response to environmental conditions and its function is calibrated from observations simultaneously to other population parameters. Moreover, hierarchical models have the additional advantage of being able to account for detectability in species distribution modelling (Kéry, 2011; Rota *et al.*, 2011). Regardless of the model structure and parameter estimation approach, it is clear that monitoring data will play a central role in the development of future dynamic SDMs.

In our study, we asked whether a general-purpose monitoring programme would be as useful for the calibration of hybrid SDMs as a monitoring programme specifically designed to survey the

areas where environmental processes expected to be relevant occurred. As the DINDIS monitoring project was specifically designed to monitoring fire patches and our target species preferred open habitats, the proportion of surveyed DINDIS sites with predicted suitability increases was much higher than in CBS. Our hope was that the larger number of CBS observations would perhaps compensate for the lower proportion of surveyed areas with relevant dynamics. Both calibration with CBS and DINDIS monitoring data yielded sound spread parameter estimates and range dynamics that were intermediate between no spread and unrestricted spread scenarios (Fig. 4), suggesting dispersal limitations in the spatial dynamics of the twelve studied species. Moreover, we did not find strong differences in predictive accuracy between occupancy models calibrated with one data set or the other. However, if we take calibration with DINDIS as a benchmark case, insufficient sampling of the areas where relevant dynamics occur resulted in calibration with CBS data to be less useful in several aspects, despite the CBS data set including a larger number of field observations. First, spread capability was underestimated or overestimated for seven of the twelve species. Second, for only in four species could the predicted range dynamics be said to be similar enough. Third, more uncertain estimates of range dynamics were often obtained.

Although we did not find general-purpose monitoring as useful as process-targeted monitoring for the calibration of hybrid SDMs, it is important to remember that in our case study, habitat suitability was predicted to decrease in almost the entire Catalan territory, whereas suitability increased in only some small areas. As the parameters used to calibrate the models were related to habitat colonization and not to extinction, it would have been surprising not to detect differences between general-purpose and targeted monitoring programmes. In other words, the CBS data set could have been very useful to estimate parameters related to extinction if the mechanistic model had allowed this extra complexity. To determine the potential usefulness of a given general-purpose monitoring data set for the calibration and evaluation of a given hybrid SDM, we recommend quantifying the number of surveyed sites that are predicted to experience habitat suitability changes as we did here.

Movement in species distribution models for animals

Unlike individual-based simulation models (e.g. Will & Tackenberg, 2008), our approach to bird movement was phenomenological (Chapman *et al.*, 2007; Smolik *et al.*, 2010). For this reason, maximum-likelihood estimates should be regarded as parameter values yielding the best phenomenological reproduction of the species distribution dynamics given the available observations (Smolik *et al.*, 2010). The fact that species prevalence correlated negatively with spread distances (D_{90}) points to a dependence of parameter estimates on the spatial context of the target species (e.g. Paradis *et al.*, 1998). A clear exception to this rule was *L. arborea*, a species that has a rather high prevalence in the territory (30%) and still obtained a very large spread distance. Excepting this case, we obtained moderate or large spread distances (using calibration with DINDIS data) with pioneer species associated with dynamic early-succession shrublands and that at the same time have low prevalence in the territory (i.e. *A. campestris*, *E. hortulana*, *G. theklae*, *O. hispanica* and *S. undata*) (Zozaya *et al.*, 2011). Dynamic occupancy models for these species were accurate when predicting the colonization of new suitable empty patches created by fire (Fig. 5b). The remaining species (*C. cannabina*, *E. cia*, *H. polyglotta*, *M. apiaster*, *S. cantillans* and *S. melanocephala*) were also predicted to experience habitat suitability changes derived from fire and forest regeneration processes. However, in the current context, those changes occurred mostly in areas that were very close to where the species was already predicted to be present (Fig. 4). In these cases, the estimates for spread distance were consequently very low. Whereas large spread distance estimates can indicate high species dispersal abilities, a low spread distance estimate is not a property of the ability of individuals of the target species to disperse, but a consequence of the species current spatial distribution. Obtaining movement-related estimates that depend on the current spatial distribution of the target species highlights a limitation of species distribution models aiming to be used for medium- or long-term predictions (Franklin, 2010), because it implies that after some time, parameter estimates may

be no-longer valid. In spatially explicit population models for animals, the dependence of parameter estimates on the spatial context of the target species can be avoided by making movement rules conditional on the status (i.e. size and carrying capacity) of the source and sink populations (Bowler & Benton, 2005), instead of using fixed dispersal kernels. Unfortunately, this seems difficult to achieve in the case of occupancy models.

Mechanistic model calibration and the accuracy of habitat suitability models

Using habitat suitability models to constrain the dynamics of mechanistic models is not without problems. Gallien *et al.* (2010) pointed at the difficulty of implementing feedbacks (for example, if seed vector animals influence the dynamics of vegetation); the potential duplication of processes accounted for in both models; and the form and strength of the relationship between habitat suitability and demographic parameters (e.g. carrying capacity). We would like to add a new issue in hybrid distribution modelling that we think should be analysed in future studies: if the mechanistic component of hybrid SDMs is calibrated from the fit to observations, the mistakes in habitat suitability predictions can be a source of bias in parameter estimates. Overestimation of spread capabilities is likely to occur if the habitat suitability model contains important omissions of good habitat that the species is using as a stepping stone. Similarly, if predicted good habitat sites are not suitable for the species, any predicted colonization in those sites will decrease the fit to observations and the spread capability will be underestimated. It is clear that this source bias arises from calibrating the mechanistic model using the predictions of the habitat suitability model.

Hierarchical Bayesian frameworks (Clark, 2005) provide an interesting way to avoid some of the problems of hybrid modelling. Because of their hierarchical random structure, these models allow for the estimation of the relationship of environmental variables on demographic parameters in the same model where these estimates operate (Pagel & Schurr, 2011), instead of having a mechanistic model that is constrained by habitat suitability estimates predicted from a previous model. However, this advantage of hierarchical Bayesian models comes at the expense of a much higher computational effort. Moreover, it is worth noting that both kinds of models (hybrid and hierarchical Bayesian) may suffer from biased parameter estimates if calibration is conducted on incorrect environmental dynamics. Given the multiple approaches and their trade-offs, design and parameterization of dynamic species distribution models still remains a challenge in the search for more robust predictions of temporal distribution changes.

ACKNOWLEDGEMENTS

This study has received financial support from the projects CGL2011-29539/BOS, CGL2008-05506-C02-01/BOS and Consolider Montes CSD2008-00040 granted by the Spanish

Ministry of Education and Science (MEC). Additional funding for MC came from a postdoctoral fellowship of the Catalan project 'Beatriu de Pinós' (2009 BP-B 00342). The study is also a contribution to the FP7-226852 EU-project SCALES and the Catalan funded programmes 2006-BE-2-52 and SGR 2009-SGR-531. The authors would like to thank Marie-Josée Fortin, Cang Hui, Andrew Fall, Núria Roura-Pascual and Elena López Zozaya and two anonymous referees for their useful comments at different stages of this research.

REFERENCES

- Anderson, B.J., Akçakaya, H.R., Araújo, M.B., Fordham, D.A., Martinez-Meyer, E., Thuiller, W. & Brook, B.W. (2009) Dynamics of range margins for metapopulations under climate change. *Proceedings of the Royal Society of London. Series B*, **276**, 1415–1420.
- Araújo, M.B., Pearson, R.G., Thuiller, W. & Erhard, M. (2005) Validation of species–climate impact models under climate change. *Global Change Biology*, **11**, 1504–1513.
- Augustin, N.H., Muggleston, M.A. & Buckland, S.T. (1996) An autologistic model for the spatial distribution of wildlife. *Journal of Applied Ecology*, **33**, 339–347.
- Bowler, D.E. & Benton, T.G. (2005) Causes and consequences of animal dispersal strategies: relating individual behaviour to spatial dynamics. *Biological Reviews*, **80**, 205–225.
- Brotons, L., Pons, P. & Herrando, S. (2005) Colonization of dynamic Mediterranean landscapes: where do birds come from after fire? *Journal of Biogeography*, **32**, 789–798.
- 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 and Distributions*, **13**, 276–288.
- Brotons, L., De Cáceres, M., Fall, A. & Fortin, M.J. (2012) Incorporating species dispersal and landscape dynamics to the modeling of bird species distribution changes in fire prone mediterranean landscapes. *Ecography*, doi: 10.1111/j.1600-0587.2011.06878.x.
- Cabral, J.S. & Schurr, F.M. (2010) Estimating demographic models for the range dynamics of plant species. *Global Ecology and Biogeography*, **19**, 85–97.
- Chapman, D.S., Dytham, C. & Oxford, G.S. (2007) Modelling population redistribution in a leaf beetle: an evaluation of alternative dispersal functions. *Journal of Animal Ecology*, **76**, 36–44.
- Clark, J.S. (2005) Why environmental scientists are becoming Bayesians. *Ecology Letters*, **8**, 2–14.
- Díaz-Delgado, R., Lloret, F. & Pons, X. (2004) Spatial patterns of fire occurrence in Catalonia, NE, Spain. *Landscape Ecology*, **19**, 731–745.
- Donald, P.F. & Fuller, R.J. (1998) Ornithological atlas data: a review of uses and limitations. *Bird Study*, **45**, 129–145.
- Dormann, C.F. (2007) Promising the future? Global change projections of species distributions. *Basic and Applied Ecology*, **8**, 387–397.
- 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.
- Estrada, J., Pedrocchi, V., Brotons, L. & Herrando, S. (2004) Atlas dels ocells nidificants de Catalunya 1999–2002. Institut Català d'Ornitologia (ICO)/Lynx Edicions, Barcelona.
- Facon, B. & David, P. (2006) Metapopulation dynamics and biological invasions: a spatially explicit model applied to a freshwater snail. *The American Naturalist*, **168**, 769–783.
- Fall, A. & Fall, J. (2001) A domain-specific language for models of landscape dynamics. *Ecological Modelling*, **141**, 1–18.
- 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.
- Franklin, J. (2009) *Mapping species distributions: spatial inference and prediction*. Cambridge University Press, Cambridge.
- Franklin, J. (2010) Moving beyond static species distribution models in support of conservation biogeography. *Diversity and Distributions*, **16**, 321–330.
- Gallien, L., Münkemüller, T., Albert, C.H., Boulangeat, I. & Thuiller, W. (2010) Predicting potential distributions of invasive species: where to go from here? *Diversity and Distributions*, **16**, 331–342.
- Gil-Tena, A., Brotons, L. & Saura, S. (2009) Mediterranean forest dynamics and forest bird distribution changes in the late 20th century. *Global Change Biology*, **15**, 474–485.
- Gregory, R.D., Strien, A., Vorisek, P., Gmelig Meyling, A.W., Noble, D.G., Foppen, R.P.B. & Gibbons, D.W. (2005) Developing indicators for European birds. *Philosophical Transactions of the Royal Society of London. Series B*, **360**, 269–288.
- Guisan, A. & Thuiller, W. (2005) Predicting species distribution: offering more than simple habitat models. *Ecology Letters*, **8**, 993–1009.
- Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
- Hanski, I. (1994) A practical model of metapopulation dynamics. *Journal of Animal Ecology*, **63**, 151–162.
- Jiguet, F., Julliard, R., Couvet, D. & Petiau, A. (2005) Modeling spatial trends in estimated species richness using breeding bird survey data: a valuable tool in biodiversity assessment. *Biodiversity and Conservation*, **14**, 3305–3324.
- Keith, D.A., Akçakaya, H.R., Thuiller, W., Midgley, G.F., Pearson, R.G., Phillips, S.J., Regan, H.M., Araújo, M.B. & Rebelo, T.G. (2008) Predicting extinction risks under climate change: coupling stochastic population models with dynamic bioclimatic habitat models. *Biology Letters*, **4**, 560–563.
- Kéry, M. (2011) Towards the modelling of true species distributions. *Journal of Biogeography*, **38**, 617–618.
- Kéry, M. & Schmid, H. (2006) Estimating species richness: calibrating a large avian monitoring programme. *Journal of Applied Ecology*, **43**, 101–110.
- Keymer, J.E., Marquet, P.A., Velasco-Hernández, J.X. & Levin, S.A. (2000) Extinction thresholds and metapopulation per-

- sistence in dynamic landscapes. *The American Naturalist*, **156**, 478–494.
- La Sorte, F.A. & Jetz, W. (2010) Avian distributions under climate change: towards improved projections. *The Journal of Experimental Biology*, **213**, 862–869.
- Lichstein, J.W., Simons, T.R., Shiner, S.A. & Franzreb, K.E. (2002) Spatial autocorrelation and autoregressive models in ecology. *Ecological Monographs*, **72**, 445–463.
- McCullagh, P. & Nelder, J.A. (1989) *Generalized linear models*. Chapman and Hall, London.
- Ministerio de Medio Ambiente (2006) *Tercer Inventario Forestal Nacional*. Dirección General de Conservación de la Naturaleza, Madrid.
- Nagelkerke, N. (1991) A note on a general definition of the coefficient of determination. *Biometrika*, **78**, 691–692.
- Pagel, J. & Schurr, F.M. (2011) Forecasting species ranges by statistical estimation of ecological niches and spatial population dynamics. *Global Ecology and Biogeography*, **21**, 293–304.
- Paradis, E., Baillie, S. & Sutherland, W. (1998) Patterns of natal and breeding dispersal in birds. *Journal of Animal Ecology*, **67**, 518–536.
- Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361–371.
- Robertson, M.P., Cumming, G.S. & Erasmus, B.F.N. (2010) Getting the most out of atlas data. *Diversity and Distributions*, **16**, 363–375.
- Rodrigo, A., Retana, J. & Picó, F.X. (2004) Direct regeneration is not the only response of Mediterranean forests to large fires. *Ecology*, **85**, 716–729.
- Rota, C.T., Fletcher, R.J., Evans, J.M. & Hutto, R.L. (2011) Does accounting for imperfect detection improve species distribution models? *Ecography*, **34**, 659–670.
- Smolik, M.G., Dullinger, S., Essl, F., Kleinbauer, I., Leitner, M., Peterseil, J., Stadler, L.M. & Vogl, G. (2010) Integrating species distribution models and interacting particle systems to predict the spread of an invasive alien plant. *Journal of Biogeography*, **37**, 411–422.
- Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N., De Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., Van Jaarsveld, A.S., Midgley, G.F., Miles, L., Ortega-Huerta, M.A., Peterson, A.T., Phillips, O.L. & Williams, S.E. (2004) Extinction risk from climate change. *Nature*, **427**, 145–148.
- Thuiller, W. (2004) Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, **10**, 2020–2027.
- Vallecillo, S., Brotons, L. & Herrando, S. (2007) Assessing the response of open-habitat bird species to landscape changes in Mediterranean mosaics. *Biodiversity and Conservation*, **17**, 103–119.
- Vallecillo, S., Brotons, L. & Thuiller, W. (2009) Dangers of predicting bird species distributions in response to land-cover changes. *Ecological Applications*, **19**, 538–549.
- Wikle, C.K. (2003) Hierarchical Bayesian models for predicting the spread of ecological processes. *Ecology*, **84**, 1382–1394.
- Will, H. & Tackenberg, O. (2008) A mechanistic simulation model of seed dispersal by animals. *Journal of Ecology*, **96**, 1011–1022.
- Willis, S.G., Thomas, C.D., Hill, J.K., Collingham, Y.C., Telfer, M.G., Fox, R. & Huntley, B. (2009) Dynamic distribution modelling: predicting the present from the past. *Ecography*, **32**, 5–12.
- Zozaya, E.L., Brotons, L., Herrando, S., Pons, P., Rost, J. & Clavero, M. (2010) Monitoring spatial and temporal dynamics of bird communities in Mediterranean landscapes affected by large wildfires. *Ardeola*, **57**, 33–50.
- Zozaya, E.L., Brotons, L. & Vallecillo, S. (2011) Bird community responses to vegetation heterogeneity following non-direct regeneration of Mediterranean forests after fire. *Ardea*, **99**, 73–84.
- Zurell, D., Jeltsch, F., Dormann, C.F. & Schröder, B. (2009) Static species distribution models in dynamically changing systems: how good can predictions really be? *Ecography*, **32**, 733–744.

SUPPORTING INFORMATION

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

Appendix S1 Details of habitat suitability modeling.

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.

BIOSKETCHES

Miquel De Cáceres is a postdoctoral researcher at the Forest Science Centre of Catalonia. He is interested in statistical methods for plant community ecologists and in the spatially explicit modelling of landscapes and species distributions at the regional/national scales.

Lluís Brotons is a full-time researcher at the Forest Technology Centre of Catalonia (CTFC). His main research aims to identify the role of spatial heterogeneity on the ecology and distribution of vertebrate species in dynamic landscapes using mediterranean and boreal regions as study models.

Author contributions: L.B. conceived the study, M.D. carried out the analyses, both authors contributed to interpreting the results and writing the manuscript.

Editor: Mark Robertson