

ARTICLE

The effectiveness of species distribution models in predicting local abundance depends on model grain size

Mattia Brambilla¹  | Gaia Bazzi² | Luca Ilahiane^{1,3}

¹Department of Environmental Science and Policy, Milan University, Milan, Italy

²Area Avifauna Migratrice, Istituto Superiore per la Protezione e la Ricerca Ambientale (ISPRA), Ozzano dell'Emilia, Italy

³Department of Sustainable Development and Ecological Transition, Eastern Piedmont University, Vercelli, Italy

Correspondence

Mattia Brambilla

Email: mattia.brambilla@unimi.it

Handling Editor: Nicholas J. Gotelli

Abstract

The use of species distribution models (SDMs) to predict local abundance has been often proposed and contested. We tested whether SDMs at different spatiotemporal resolutions may predict the local density of 14 bird species of open/semi-open habitats. SDMs were built at 1 ha and 1 km, and with long-term versus a mix of current and long-term climatic variables. The estimated environmental suitability was used to predict local abundance obtained by means of 275 linear transects. We tested SDM ability to predict abundance for all sampled sites versus occurrence sites, using N-mixture models to account for imperfect detection. Then, we related the R^2 of N-mixture models to SDM traits. Fine-grain SDMs appeared generally more robust than large-grain ones. Considering the all-transects models, for all species environmental suitability displayed a positive and highly significant effect at all the four combinations of spatial and temporal grains. When focusing only on occurrence transects, at the 1 km grain only one species showed a significant and positive effect. At the 1 ha grain, 62% of species models showed (over both climatic sets) a significant or nearly significant positive effect of environmental suitability on abundance. Grain was the only factor significantly affecting the model's explanatory power: 1 km grain led to lower amounts of variation explained by models. Our work re-opens the debate about predicting abundance using SDM-derived suitability, emphasizing the importance of grains and of spatiotemporal resolution more in general. The incorporation of local variables into SDMs at fine grains is key to predict local abundance. SDMs worked out at really fine grains, approaching the average size of territory or home range of target species, are needed to predict local abundance effectively. This may result from the fact that each single cell may represent a potential territory/home range, and hence a higher suitability over a given area means that more potential territories occur there.

KEYWORDS

breeding birds, density, distribution, environmental suitability, MaxEnt, N-mixture models, SDMs

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *Ecology* published by Wiley Periodicals LLC on behalf of The Ecological Society of America.

INTRODUCTION

Correlative species distribution models (SDMs) basically relate the occurrence of a species at some sites to the environmental characteristics measured at those locations to identify the environmental conditions where a species can be found and estimate its distribution over a given extent. The characteristics usually measured at occurrence (and background or absence) sites are variables that are thought to be relevant for the definition of the Hutchinsonian niche of the target species, that is, the environmental conditions under which a species survives and reproduces (Elith & Leathwick, 2009; Hutchinson, 1957; Warren, 2012). SDMs have gained primary importance for ecology, biogeography and conservation, and have become one of the dominant topics of recent literature dealing with species distribution.

Many attempts to expand their use from the “simple” prediction of distribution, to forecast past and future ranges (Maguire et al., 2015) and predict reproductive (Brambilla & Ficetola, 2012) or other population parameters (Lee-Yaw et al., 2022), including density estimates (Weber et al., 2017), have been made so far. In fact, even if they are conceived as tools to predict distribution (hence, presence/occurrence vs. absence), rather than density (abundance per space unit), SDMs could potentially provide estimates of the abundance of a species. Intuitively, the most favorable environments could support higher species abundances (Dallas & Hastings, 2018), and a positive relationship between environmental suitability computed by SDMs and local abundance can be expected because of a variety of processes acting at the population level (Lunghi et al., 2018).

However, first investigations often reported poor correlations between the two (Jiménez-Valverde, 2011 and references therein), and subsequent efforts yielded mixed evidence (see Dallas & Hastings, 2018 and references therein). In fact, it is likely that many environmental filters and other drivers may reduce the local abundance compared with the potential density, including behavior (e.g., breeding aggregation patterns, Estrada & Arroyo, 2012), position within the species’ range (Dallas & Santini, 2020; Osorio-Olvera et al., 2019) or effect of suitability on growth rates but not on carrying capacities (Dallas & Hastings, 2018), or predominant influence of other factors such as dispersal (Altermatt & Fronhofer, 2018), availability of key resources (Planillo et al., 2021), or biotic interactions (Brambilla, Scridel, et al., 2020). As a consequence, environmental suitability estimated by SDMs may predict the upper limit of abundance, rather than the actual abundance of a species (VanDerWal et al., 2009).

By evaluating the correlation between occurrence probability estimated by distribution models and species

abundance in hundreds of mammal and tree species, Dallas and Hastings (2018) suggested that in general the climatic suitability estimated by SDMs is not associated with species’ abundances, and that nonclimatic factors may be highly relevant in that sense. Recent insights based on virtual species suggested that SDMs cannot be used to predict the species’ actual abundance, and that the deriving suitability maps may instead instill a misleading perception of precision at local scales (Jiménez-Valverde et al., 2021). Nevertheless, multiple pieces of evidence for SDMs predicting local abundance exist in the literature (e.g., Bean et al., 2014; Carrascal et al., 2015), and a global meta-analysis of the correlation between environmental suitability and local abundance found a general, significantly positive, relationship (Weber et al., 2017).

Weber et al. (2017) did not find significant differences associated with range proportion, study scale, the method selected for building Environmental Niche Models (ENMs), or variable sets used for modeling. On the other side, the inclusion of local variables in combination with climatic ones led to a higher correlation between suitability and abundance (Weber et al., 2017). This suggests that variables matching the fine-scale conditions experienced by a species (or, in other words, better approaching the ultimate drivers of species’ occurrence) might increase models’ effectiveness in estimating abundance.

A related key point about SDM development, with potentially important implications also for their effectiveness in predicting local abundance, is represented by model grain size, that is, the spatial resolution at which SDMs are worked out. The choice of grain size is frequently dependent on the resolution of the available predictors (Manzoor et al., 2018). For territorial animal species, a model grain size approaching the territory size could lead to fine-scale models more likely to include the most relevant variables or, at least, better proxies than other scales (Brambilla et al., 2019). Also, the temporal resolution of predictors used in SDMs may be important, in addition to the spatial one (Pennino et al., 2019). SDM general performances and, specifically, their ability to predict density (as well as other population parameters) may depend on the spatiotemporal resolution and especially on the model’s grain. However, this potentially crucial aspect apparently had never been investigated in detail.

With this work, we tested whether SDMs based on variables at different spatiotemporal resolutions may or not predict the local density of 14 farmland bird species of conservation concerns. We focused on birds because the expertise concerning distribution modeling on this animal group is one the most prominent among taxa (Engler et al., 2017); in this sense avian SDMs can provide

valuable contributions to ongoing current debates. Here, we explored the effects of spatial and temporal resolution of environmental predictors on SDM accuracy and, especially, on their ability to predict abundance. SDMs were built with different model grain sizes (1 ha, approximately mirroring the territory/home-range size of most species, and 1 km), and with different temporal climatic variables. Then, the estimated environmental suitability was used to predict local abundance obtained by means of linear transects, in order to evaluate the effect of spatial and temporal accuracy on the abundance predictions. It can be expected that: (1) increasing spatial resolution can allow for better prediction abundance, because fine-grained models may produce estimates of habitat suitability at the home-range scale, and hence higher suitability over a given area could be associated with a

higher number of potential territories/home-ranges (see Figure 1); (2) increasing the temporal consistency between climatic predictors and occurrence data can lead to more accurate models, if bird species are (at least partly) affected by weather, and not only by indirect effects of long-term climate.

We tested the SDM ability to predict abundance for all sampled sites versus only sites with the occurrence of the target species. In fact, if the abundance–suitability relationship is based on correlations between abundance and suitability for all sites, including those where the target species is absent, the low suitability predicted for absence sites would likely inflate the rank correlation coefficients, and the apparent SDMs ability to predict density would be largely due to their ability in telling apart presence from absence (Dallas & Hastings, 2018).

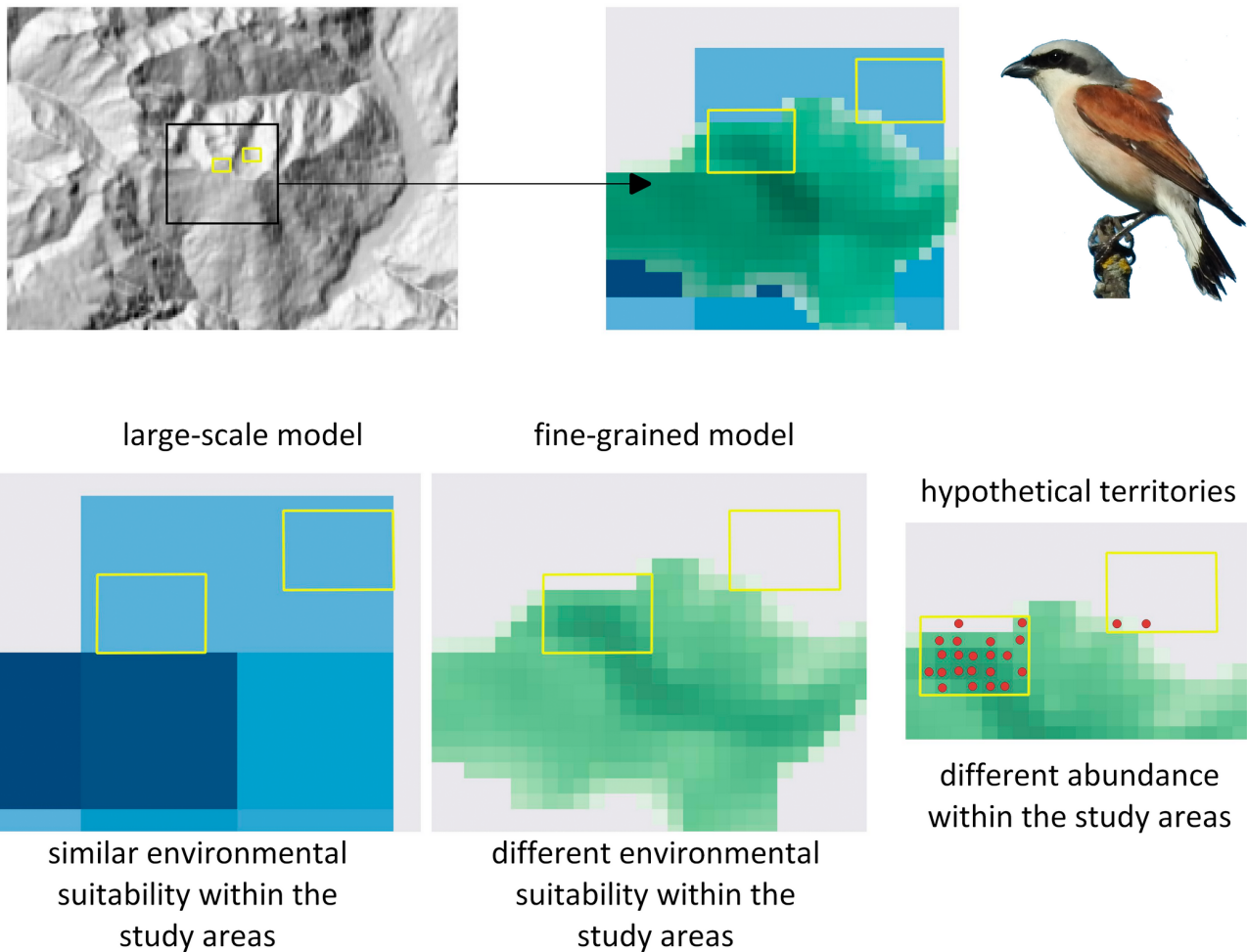


FIGURE 1 Possible differences in species distribution models (SDM) ability to predict local abundance may be due to the model’s grain, that is, spatial resolution. The figure shows a hypothetical example for two study areas (in yellow), where local abundance (red-brown dots show the possible location of different pairs/individuals) may be differently predicted by a coarse SDM (in blue color ramp) and by a fine-grained SDM (in green color ramp), respectively. Example species is red-backed shrike *Lanius collurio* (photograph: Mattia Brambilla). Example SDMs taken from Brambilla et al. (2022) and Ceresa et al. (2023a) (raster files downloaded from Brambilla, 2022; Ceresa et al., 2023b).

METHODS

Study system

The study was carried out in Lombardy, northern Italy. Lombardy region stretches over approximately 24,000 km², is densely populated (more than 400 inhabitants per square kilometer) and highly industrialized. In the northern portion, the Alps and Prealps feature the highest peaks (up to more than 4000 m above sea level) and are largely covered by forests, with high-elevation habitats toward mountain summits, and grasslands and pastures along mountainsides and in valley floors. Southward, the Po Plain covers a substantial portion of the region and is dominated by urbanized areas and intensive agriculture. In the southern extreme, the northern Apennines are found, with a hilly and low-mountain landscape largely covered by forest and farmed areas of varying intensity.

Within this area, 14 bird species related to farmland or other open or semi-open habitats were considered: one

dove (turtle dove *Streptopelia turtur*), two larks (skylark *Alauda arvensis* and woodlark *Lullula arborea*), two pipits (tawny pipit *Anthus campestris* and tree pipit *Anthus trivialis*), one chat (stonechat *Saxicola torquatus*), two warblers (whitethroat *Sylvia communis* and Moltoni's warbler *Sylvia subalpina*), one shrike (red-backed shrike *Lanius collurio*) and five bunting species (corn bunting *Emberiza calandra*, rock bunting *Emberiza cia*, ciril bunting *Emberiza cirius*, yellowhammer *Emberiza citrinella*, ortolan bunting *Emberiza hortulana*). Those species were selected as exclusively or largely tied to rural environments within the study region. Birds were surveyed by means of linear transects, which were placed in the largest and most relevant rural landscapes of northern and southern Lombardy (Figure 2), within the framework of two different projects (see <https://vignetienatura.net/> and <https://www.naturachevale.it/>) that focused (entirely or also) on avian communities within such kinds of environments. Within those areas, 275 linear transects, each one 200 m long, were placed in open and semi-open landscapes

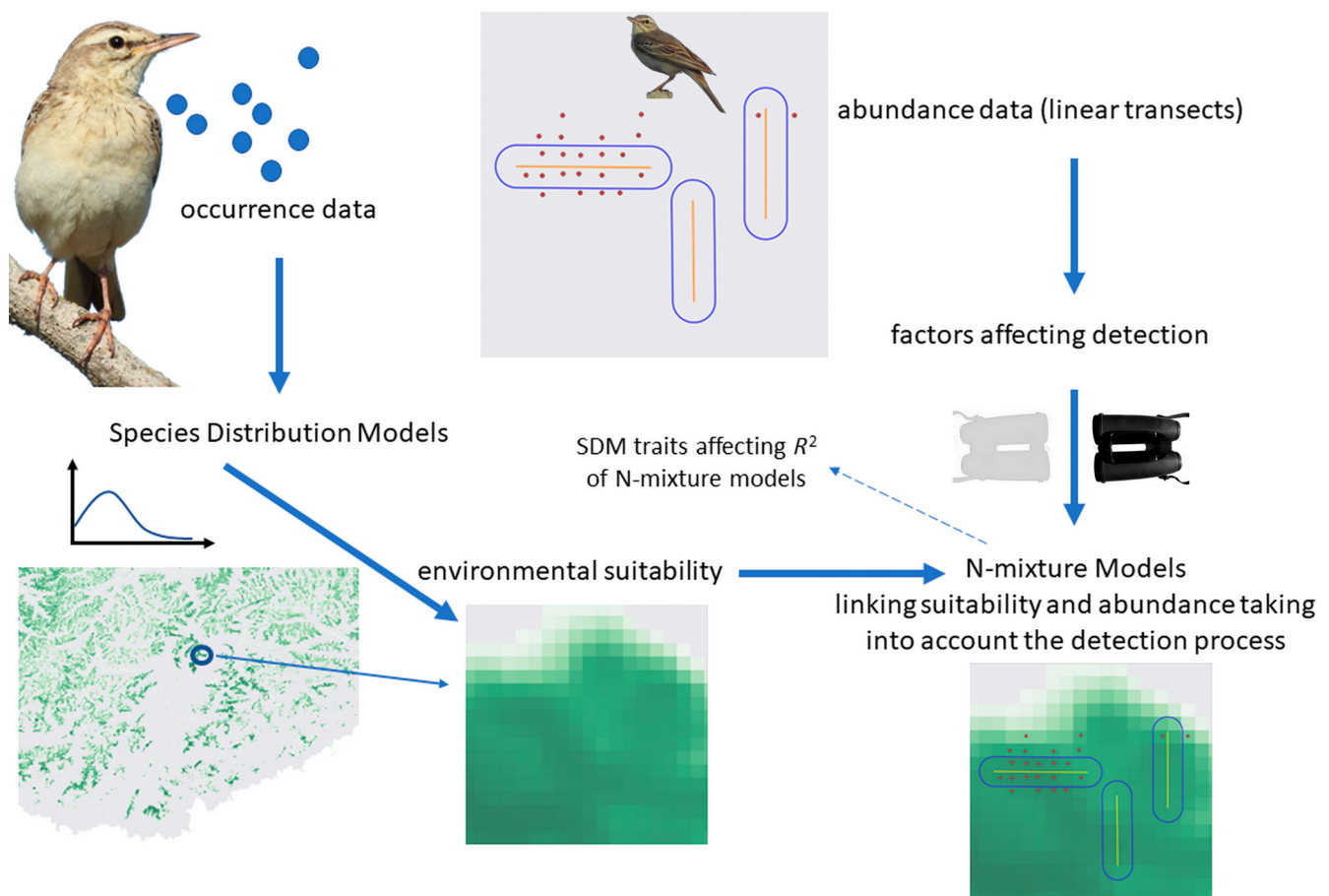


FIGURE 2 A graphical flowchart of the approach adopted in the study to relate environmental suitability estimated by species distribution models to local abundance by means of N-mixture models, taking into account the effects of factors affecting the detection process. Example species is tawny pipit *Anthus campestris* (photographs: Mattia Brambilla). Example species distribution models taken from Brambilla et al. (2022) and Ceresa et al. (2023a).

representative of the different sample contexts, in terms of elevation, slope, aspects and main vegetation types. Within such landscapes, transects were located along existing paths or unpaved roads, and were surveyed by the same authors (Mattia Brambilla, Gaia Bazzi, Luca Ilahiane) twice (May–June; 204 transects, 2017) or thrice (late April/early May–May–June; 71 transects, 2015) per year. Birds were counted within a 100-m buffer around the linear transect, and hence each sampling unit had a size of 7.14 ha (Assandri et al., 2019; Brambilla & Gatti, 2022; Rollan et al., 2019). Days with rain or wind (moderate or strong) were avoided. Contacts with the target species were accurately mapped with the best possible approximation (i.e., identifying the single tree, shrub or patch where the individual(s) had been observed), using high-resolution aerial orthophotographs of each transect area. For building SDMs, contacts were complemented with other presence data collected during the period 2013–2017 within the framework of other projects led by the authors, or by occasional observations (mostly by the authors), provided that the spatial accuracy (only data labeled as “exact location” and overlapping with potentially suitable landscapes were considered) and the period were consistent with the data collected along the transects. In total, 4800 spatially accurate and reliable occurrence records were thus used for building SDMs (skylark: 366; tawny pipit: 74; tree pipit: 453; corn bunting: 411; rock bunting: 208; ciril bunting: 505; yellowhammer: 142; ortolan bunting: 73; red-backed shrike: 704; woodlark: 447; stonechat: 331; turtle dove: 469; whitethroat: 328; and Moltoni’s warbler: 289).

Environmental variables

As environmental variables for SDMs, we combined climatic, topographic and land-use/land-cover (LULC) predictors, known to be important drivers of occurrence in these and other open-landscape species in the study region and similar contexts (Brambilla et al., 2012, 2022; Brambilla, Gustin, et al., 2020). The same datasets were used to compute environmental variables at both grains, aggregating or averaging values according to the relative grain. Climatic data consisted of long-term and short-term data. Long-term climatic data (1981–2010) were retrieved from the CHELSA database, V2.1 (Karger et al., 2017, 2021). Based on previous studies showing their relevance (Brambilla et al., 2022; Thuiller et al., 2019), the chosen bioclimatic variables were bio1 (mean annual air temperature), bio7 (annual range of air temperature), bio12 (annual precipitation amount) and bio15 (precipitation seasonality) (Karger et al., 2017, 2021). The resolution of the data was 30 arc sec; for the study

region, this corresponds to a resolution of <1 km. Short-term climatic data describing temperature and precipitations for the months of April to July (i.e., breeding period) were also retrieved at the same resolution from CHELSA V2.1 (Karger et al., 2017, 2021) for the years 2013–2017, for which mean temperature and mean precipitation during the breeding period were calculated. Two sets of climatic variables were then prepared: the first one contained the four long-term variables (annual mean temperature: bio1; temperature annual range: bio7; annual precipitation: bio12; and precipitation seasonality: bio15), while the second one contained a mix of long-term data (bio7 and bio15, i.e., those describing temperature and precipitation variations throughout the year) and short-term data (mean temperature and mean precipitation for April–July 2013–2017).

LULC and topographic predictors were derived from accurate layers made available by the regional government and freely downloadable from the regional cartographic portal (<https://www.geoportale.regione.lombardia.it/>). Topographic variables were obtained from a 20-m Digital Terrain Model (DTM) of Lombardy. From it, we computed slope and summer–spring solar radiation in GRASS (Neteler et al., 2012).

LULC variables were derived from a detailed land-cover regional map, dated 2015 (DUSAF5), with a 20 m resolution. Some LULC types were merged to reduce the number of predictors and better focus on ecologically relevant variables: different built-up surfaces were merged into “urbanized areas”; parks and gardens were merged into “urban green areas”; construction sites and rubbish dumps were merged into “degraded areas”; horticultural crops were merged into “horticulture” (see also Appendix S1: Table S1). Rare LULC types (reforestation and winter-flooded grasslands) were removed (and not merged with other variables as largely different from other categories from an ecological point of view). Then, the proportional cover of each LULC type was calculated at each resolution.

All the variables were then resampled at two different resolutions. A large-grain resolution of 1 km was intended to represent “landscape” characteristics, whereas a fine-grain resolution of 1 ha was meant to approximate “territory” characteristics for the target species (Brambilla et al., 2019), which mostly defend relatively small territories and thus should be sensitive to environmental variations at such a scale. For both grains, two different sets of predictors were considered, one including only long-term climatic variables, and one replacing mean temperature and precipitation with values calculated for the breeding season of the period 2013–2017 (see above). All the variables used for SDM training are reported in Appendix S1: Table S1.

Species distribution modeling

SDMs were built using MaxEnt (Phillips et al., 2006) in R (R Development Core Team, 2020). We selected MaxEnt as the unique approach because (1) we wanted to explore the link between spatiotemporal resolution of models and their ability to predict local abundance, rather than the potential differences between algorithms, and (2) it offers some advantages over other methods. MaxEnt is the most frequently adopted algorithm for SDMs, it reduces the dependence on “real” absences and the undesired effects of false ones (Jiménez-Valverde et al., 2008; Elith et al., 2011), and performs very well compared with other methods or ensemble modeling (Kaky et al., 2020), often providing ecologically more meaningful results (Brambilla et al., 2022).

We followed the procedure (and the relative script) described in Brambilla et al. (2022). Even if many studies used random points scattered over a calibration region, background points must be placed in areas where sampling could have potentially occurred, as they need to represent sampled environmental conditions to avoid biased evaluations of species–environment relationships due to lack of sampling in contexts characterized by variable values not found elsewhere (cf. Brambilla, Scridel, et al., 2020). For this purpose, we created two buffers, depending on the model’s grain, around the extent defined by all the occurrence points: a buffer of 7 km for the 1 km models, and of 1 km for the 1 ha ones. Background points were scattered over those buffers, and were therefore constrained to the environments sampled, or close to the sampled ones. Then, we removed duplicated records, keeping only one record for each species per grid cell, and checked correlations between environmental variables. The threshold for highly correlated variables was set at $r = |0.8|$. Occurrence data were then partitioned into training and testing datasets, adopting the “checkerboard 2” command in ENMeval (Muscarella et al., 2014). Training datasets were thus represented by databases (in species-with-data formats, SWD) including around $\frac{3}{4}$ of occurrence points, whereas testing datasets included the remaining, spatially independent data from the fourth partition.

Models were built considering only linear and quadratic relationships to avoid overfitting. An initial value of the regularization multiplier was selected testing values at 0.5 intervals between 0.5 and 5. The selected regularization multiplier was used to create a base model for each species.

Highly correlated variables were then removed based on the area under the receiver operating characteristic (ROC) curve (AUC) on the testing dataset using the “SDMTune” function (Vignali et al., 2020) to choose between the correlated variables ($r > |0.8|$), on the basis

of the model’s performance alternatively including one or another correlated variable.

All the variables for which lambda was 0, meaning that they had no effect on the species, were identified and removed. Then, the variable selection was performed by leaving out variables starting from that with the lowest permutation importance, until AIC_c (Akaike’s information criterion, corrected for small sample size) increased. The most supported model for each species was thus identified, and then it was fine-tuned and further improved. We checked if the chosen features (linear and quadratic) and the regularization multiplier were still the most appropriate for the models, and, based on the confirmed or updated value of these two features, the final model was prepared and used for predictions and model evaluation. All selection steps were based on AIC_c , unless otherwise specified.

Model evaluation was carried out using threshold-independent and threshold-based approaches. As threshold-independent methods, we computed the AUC and the True Skills Statistics (TSS) over training and testing data sets. As threshold-based methods, we calculated the omission rates over the test dataset at the 10th percentile (i.e., the suitability value that leaves out the 10% of the sites with the lowest predicted suitability) calculated on the training presence, and at the minimum training presence (i.e., the lowest suitability value at occurrence sites). The former should be close to 0.1 and the latter to 0, that is, to the theoretically expected values; higher omission rates imply model overfitting and suggest that the model is not properly working on the test dataset and thus cannot be generalized. For AUC and TSS, “valid” models should show similar values over the training and testing dataset (the larger the difference, the worse the model), whereas the absolute value is relatively poorly informative (Lobo et al., 2008).

The tuned models were used to predict environmental suitability according to Cloglog outputs. The continuous value obtained, varying between 0 and 1, was used as a predictor to test the relationship between environmental suitability and abundance.

Abundance models

The links between local density and environmental suitability estimated by SDMs were evaluated using N-mixture models. These models were chosen because they integrated both the factors affecting the detection process and those influencing the state process (i.e., abundance; see Figure 2 for a visual representation of the framework adopted) in a hierarchical modeling. Models were built in R using the package “unmarked”

(Fiske & Chandler, 2011), to obtain estimates of latent abundance, which is a measure of the “true” local abundance (and hence density, considering the fixed area of the surveys). We considered the number of individuals found during a survey as the dependent variable, SDM-derived environmental suitability as a state predictor, and some variables affecting the detection process: wind and disturbance (as two-level categorical factors), hour, date, duration (in minutes) of the survey along the transect. All continuous variables were scaled.

Two different sets of models were built. In the first one, the N-mixture models for each species were obtained considering all transects (hence including both absence and presence sites). In the second set, the species-specific models were worked out considering only presence sites. For each model, we computed the R^2 value as a measure of the predictive ability of the model and performed a validation test by using a goodness-of-fit test based on 99 simulations. From this, the p and \hat{c} values of the goodness-of-fit test were obtained using the package *AICcmodavg* (Mazerolle, 2022) and used for model validation (a significant goodness-of-fit p -value indicates a nonvalidated model). All models showing signs of severe underdispersion ($\hat{c} > 2$) also showed a significant p -value for the goodness-of-fit test, and thus we used the latter to separate between validated and nonvalidated models. Given that these models aimed to assess the possible link between environmental suitability and latent abundance, rather than to obtain a precise estimate of the latter, nonvalidated models could also provide some useful information. Therefore, they were highlighted as nonvalidated, but were not discarded. For occurrence-only sites, we re-ran models that were not validated using a negative binomial mixture; all models were thus validated (nonsignificant p -values for the goodness-of-fit test, $\hat{c} < 1.5$).

Exploring the drivers of the suitability–abundance relationship

To point out the factors shaping the relationship between environmental suitability and local abundance, we adopted a Generalized Linear Mixed Model (GLMM) approach, relating the adjusted R^2 of N-mixture models (focusing on Poisson models for consistency) to SDM features describing the performance of distribution models. Given that the latter were highly correlated and led to very high variable inflation factor (VIF) values, only a few variables could be entered into the GLMM models: model grain (1 km vs. 1 ha), climate predictors (long-term vs. current), the training sample size

for SDMs, the TSS difference between training and testing datasets, the AUC of the training dataset, and the omission rate at the minimum training presence (computed over the test dataset). In addition, we also included the number of transects at which a species was found.

GLMMs were run using the *glmmTMB* package (Brooks et al., 2017), entering the species as a random (grouping) factor. First, a full model was fitted and used for validation: using “dharma” (Hartig, 2020) validation functions and 500 simulations of residuals distribution, we checked residuals’ uniformity, the occurrence of outliers, simulated versus observed dispersion, and possible zero inflation. Then, we performed an AIC_c -based model selection using the package *MuMIn* (Bartoń, 2020) to identify the most supported models ($\Delta AIC_c < 2$), which were then fully averaged after the exclusion of uninformative parameters (Arnold, 2010). The analysis was performed on four sets of R^2 : from all-transects models, from occurrence-only models for each species, and for each of the two focusing on all models or only on those that were statistically validated (nonsignificant goodness-of-fit tests).

RESULTS

Species distribution models

A few SDMs showed poor performance, with high omission rates and/or differences in AUC and/or TSS between training and testing dataset, thus being unsuitable for extrapolation and less reliable in general than the others, while most models showed relatively robust validation statistics (Table 1). In particular, 1 km models for tawny pipit, the species with the lowest sample size, performed clearly badly on the independent test dataset, and thus should be rejected.

In general, fine-grain models appeared more robust than large-grain ones, with lower numbers of poorly performing SDMs; the same applies to models with current climates compared to those only featuring long-term climatic predictors.

N-mixture models

N-mixture models were performed for all species at both grains with the only exception of ortolan bunting, which occurred at a too low number of transects for running models on occurrence-only transects. Considering the all-transects models, for all species environmental suitability displayed a positive and highly significant ($p < 0.01$) effect at all the four combinations of spatial

TABLE 1 Summary of SDM statistics.

Species	N train	N test	TSS train	TSS test	TSS difference	AUC train	AUC test	AUC difference	MTP OR	10° perc OR	Notes
1 km, long-term climate											
Cirl bunting	148	41	0.824	0.732	0.092	0.951	0.919	0.032	0.000	0.146	
Corn bunting	69	29	0.776	0.815	−0.039	0.956	0.962	−0.007	0.000	0.069	
Moltoni's warbler	57	22	0.934	0.961	−0.027	0.987	0.986	0.001	0.000	0.227	Poor
Ortolan bunting	21	10	0.938	0.882	0.056	0.980	0.958	0.021	0.300	0.300	Poor
Red-backed shrike	162	52	0.740	0.721	0.019	0.922	0.919	0.003	0.000	0.115	
Rock bunting	53	25	0.693	0.632	0.061	0.911	0.871	0.040	0.000	0.240	Poor
Skylark	120	27	0.627	0.634	−0.007	0.887	0.865	0.021	0.037	0.111	
Stonechat	96	36	0.799	0.698	0.101	0.943	0.921	0.022	0.000	0.194	
Tawny pipit	16	5	0.863	0.529	0.334	0.957	0.771	0.186	0.600	0.600	Reject
Tree pipit	146	48	0.740	0.772	−0.032	0.926	0.921	0.005	0.000	0.125	
Turtle dove	167	66	0.732	0.751	−0.019	0.932	0.926	0.006	0.000	0.076	
Whitethroat	72	24	0.829	0.830	−0.001	0.964	0.955	0.009	0.042	0.042	
Woodlark	101	42	0.910	0.929	−0.019	0.978	0.974	0.004	0.000	0.119	
Yellowhammer	45	16	0.887	0.888	−0.001	0.972	0.960	0.012	0.000	0.313	Poor
1 ha, long-term climate											
Cirl bunting	329	114	0.720	0.732	−0.013	0.923	0.920	0.003	0.000	0.114	
Corn bunting	214	88	0.744	0.752	−0.008	0.948	0.942	0.006	0.011	0.102	
Moltoni's warbler	165	49	0.820	0.811	0.009	0.964	0.958	0.006	0.000	0.143	
Ortolan bunting	53	13	0.710	0.697	0.012	0.920	0.923	−0.003	0.000	0.077	
Red-backed shrike	434	139	0.687	0.637	0.050	0.916	0.894	0.022	0.000	0.158	
Rock bunting	129	30	0.738	0.554	0.183	0.929	0.841	0.088	0.033	0.333	Reject
Skylark	214	96	0.673	0.628	0.045	0.907	0.881	0.026	0.031	0.125	
Stonechat	205	75	0.681	0.696	−0.015	0.900	0.908	−0.008	0.000	0.080	
Tawny pipit	38	13	0.918	0.900	0.018	0.988	0.971	0.016	0.000	0.308	Poor
Tree pipit	289	94	0.751	0.748	0.003	0.935	0.929	0.006	0.000	0.096	
Turtle dove	308	118	0.628	0.604	0.024	0.879	0.865	0.014	0.017	0.136	
Whitethroat	182	71	0.747	0.710	0.038	0.933	0.915	0.018	0.014	0.183	
Woodlark	303	90	0.816	0.816	0.000	0.946	0.943	0.003	0.000	0.078	
Yellowhammer	99	24	0.867	0.905	−0.038	0.969	0.966	0.003	0.000	0.083	
1 km, current climate											
Cirl bunting	148	41	0.821	0.736	0.086	0.953	0.919	0.034	0.000	0.195	
Corn bunting	69	29	0.824	0.837	−0.014	0.968	0.968	0.000	0.000	0.069	
Moltoni's warbler	57	22	0.946	0.960	−0.013	0.988	0.987	0.002	0.000	0.091	
Ortolan bunting	21	10	0.844	0.967	−0.123	0.969	0.982	−0.014	0.000	0.000	
Red-backed shrike	162	52	0.734	0.729	0.005	0.922	0.920	0.003	0.000	0.115	
Rock bunting	53	25	0.586	0.639	−0.053	0.865	0.849	0.016	0.040	0.040	
Skylark	120	27	0.630	0.619	0.011	0.900	0.880	0.020	0.000	0.185	
Stonechat	96	36	0.811	0.724	0.087	0.947	0.913	0.034	0.028	0.194	
Tawny pipit	16	5	0.881	0.533	0.348	0.959	0.761	0.198	0.800	0.800	Reject
Tree pipit	146	48	0.725	0.773	−0.048	0.919	0.923	−0.004	0.000	0.063	
Turtle dove	167	66	0.732	0.751	−0.019	0.932	0.926	0.006	0.000	0.076	

TABLE 1 (Continued)

Species	N train	N test	TSS train	TSS test	TSS difference	AUC train	AUC test	AUC difference	MTP OR	10° perc OR	Notes
Whitethroat	72	24	0.867	0.867	0.000	0.965	0.966	-0.001	0.000	0.167	
Woodlark	101	42	0.907	0.925	-0.018	0.981	0.978	0.003	0.000	0.190	
Yellowhammer	45	16	0.893	0.873	0.020	0.974	0.959	0.015	0.125	0.313	Poor
1 ha, current climate											
Cirl bunting	329	114	0.721	0.734	-0.013	0.924	0.919	0.005	0.000	0.105	
Corn bunting	214	88	0.773	0.781	-0.008	0.954	0.949	0.005	0.011	0.102	
Moltoni's warbler	165	49	0.841	0.809	0.032	0.965	0.956	0.009	0.041	0.143	
Ortolan bunting	53	13	0.797	0.724	0.073	0.951	0.947	0.003	0.000	0.154	
Red-backed shrike	434	139	0.687	0.625	0.062	0.916	0.894	0.022	0.000	0.173	
Rock bunting	129	30	0.747	0.620	0.127	0.933	0.870	0.063	0.067	0.233	Poor
Skylark	214	96	0.660	0.646	0.014	0.906	0.887	0.019	0.042	0.104	
Stonechat	205	75	0.688	0.703	-0.015	0.898	0.909	-0.011	0.000	0.080	
Tawny pipit	38	13	0.875	0.874	0.001	0.981	0.974	0.008	0.000	0.077	
Tree pipit	289	94	0.758	0.753	0.005	0.940	0.935	0.005	0.011	0.106	
Turtle dove	308	118	0.641	0.602	0.038	0.879	0.865	0.014	0.000	0.144	
Whitethroat	182	71	0.773	0.746	0.026	0.936	0.926	0.010	0.000	0.183	
Woodlark	303	90	0.835	0.818	0.017	0.949	0.946	0.002	0.000	0.089	
Yellowhammer	99	24	0.874	0.899	-0.025	0.970	0.961	0.009	0.000	0.083	

Note: "MTP OR" stands for "omission rates at minimum training presence" and "10° perc OR" for "omission rates at the 10th percentile threshold"; both are calculated on the testing dataset. For other acronyms, see the text. Under the column "notes," "poor" indicates models with poor performance, and "reject" indicates models that should be rejected because of highly unbalanced performances over training and testing datasets. Abbreviations: AUC, Area Under the Curve; TSS, True Skills Statistics.

and temporal grains (Table 2). When focusing only on occurrence transects, at the 1 km grain (and for both climatic sets), only the model of the red-backed shrike showed a significant and positive effect; at the 1 ha grain, for 8 out of 13 species models showed (over both climatic sets) a significant ($p < 0.05$) or nearly significant ($0.05 < p < 0.1$) positive effect of environmental suitability on abundance (Table 3).

Factors affecting the suitability–abundance relationship

GLMMs showed that, at both occurrence-only and all transects, and including or excluding nonvalidated models, grain was the only factor significantly affecting the explanatory power of Poisson N-mixture models according to the full averaged models (Appendix S1: Table S2). In all cases, the 1 km grain was associated with significantly lower amounts of variation explained by the models. All models were statistically validated ($p > 0.05$ for all uniformity, outlier, dispersion and zero-inflation tests).

DISCUSSION

Whether environmental suitability generated by SDMs may be used as a proxy for local abundance, is a highly debated topic. According to a broad review of existing evidence, occurrence data can lead to models providing "a reasonable proxy for abundance" (Weber et al., 2017). This especially applies to vertebrates, and to models including local variables, and the relationship is by far not universally valid (Weber et al., 2017). These results had been largely contested by Dallas and Hastings (2018), who in particular criticized the inclusion of studies encompassing also absence sites, which can inflate the relationship between suitability and abundance because of the large differences in suitability found between occurrence (irrespective of abundance) and absence sites. Other studies indeed reported much weaker or even no linkages between suitability and abundance (Santini et al., 2019; Sporbert et al., 2020), even when including some of the studies already in the aforementioned review (Lee-Yaw et al., 2022). Recent conclusions suggested that SDMs are not a good proxy of abundance for many species (Lee-Yaw et al., 2022), likely because of the many

TABLE 2 Summary of N-mixture models over all transects, indicating number of occurrence transects for each species (“N_tr”), notes about species distribution models (“SDM notes”) (see previous paragraph and Table 1), model coefficient for environmental suitability (“Coefficient”), significance of the environmental suitability effect in the N-mixture model (first “p-value”), model’s R^2 value (“ R^2 ”), and validation probability (second “p-value”; significant values indicated nonvalidated models; see *Methods* for details).

Species	N_tr	SDM notes	Coefficient	p-value	R^2	p-value
1 km, long-term climate						
Cirl bunting	78		1.141	<0.001	0.414	<0.001
Corn bunting	40		1.442	<0.001	0.630	<0.001
Moltoni’s warbler	40	Poor	1.030	<0.001	0.353	<0.001
Ortolan bunting	8	Poor	0.660	0.005	0.187	0.091
Red-backed shrike	113		0.429	<0.001	0.424	<0.001
Rock bunting	21	Poor	0.969	<0.001	0.297	0.283
Skylark	36		1.330	<0.001	0.316	0.040
Stonechat	54		0.816	<0.001	0.297	<0.001
Tawny pipit	13	Reject	0.989	<0.001	0.149	0.414
Tree pipit	31		1.551	<0.001	0.420	0.818
Turtle dove	59		0.746	<0.001	0.203	<0.001
Whitethroat	40		1.009	<0.001	0.372	<0.001
Woodlark	91		0.960	<0.001	0.435	<0.001
Yellowhammer	28	Poor	0.944	<0.001	0.249	0.040
1 ha, long-term climate						
Cirl bunting	78		1.229	<0.001	0.572	0.162
Corn bunting	40		1.045	<0.001	0.706	0.040
Moltoni’s warbler	40		1.433	<0.001	0.617	0.384
Ortolan bunting	8		1.459	<0.001	0.292	0.202
Red-backed shrike	113		0.743	<0.001	0.554	<0.001
Rock bunting	21	Reject	1.174	<0.001	0.357	0.121
Skylark	36		1.292	<0.001	0.473	0.030
Stonechat	54		0.733	<0.001	0.283	<0.001
Tawny pipit	13	Poor	0.853	<0.001	0.272	0.929
Tree pipit	31		1.723	<0.001	0.435	0.768
Turtle dove	59		1.336	<0.001	0.343	0.242
Whitethroat	40		1.444	<0.001	0.579	0.121
Woodlark	91		1.247	<0.001	0.542	0.101
Yellowhammer	28		1.405	<0.001	0.357	0.778
1 km, current climate						
Cirl bunting	78		1.096	<0.001	0.425	<0.001
Corn bunting	40		1.235	<0.001	0.615	0.020
Moltoni’s warbler	40		1.355	<0.001	0.434	0.020
Ortolan bunting	8		1.035	<0.001	0.228	0.030
Red-backed shrike	113		0.452	<0.001	0.433	<0.001
Rock bunting	21		0.994	<0.001	0.297	0.404
Skylark	36		1.452	<0.001	0.308	0.152
Stonechat	54		0.802	<0.001	0.299	<0.001
Tawny pipit	13	Reject	0.915	<0.001	0.146	0.374
Tree pipit	31		1.645	<0.001	0.414	0.838

TABLE 2 (Continued)

Species	N_tr	SDM notes	Coefficient	p-value	R ²	p-value
Turtle dove	59		0.746	<0.001	0.203	0.051
Whitethroat	40		1.109	<0.001	0.384	<0.001
Woodlark	91		0.879	<0.001	0.414	<0.001
Yellowhammer	28	Poor	0.943	<0.001	0.259	0.061
1 ha, current climate						
Cirl bunting	78		1.204	<0.001	0.573	0.212
Corn bunting	40		1.064	<0.001	0.720	0.091
Moltoni's warbler	40		1.302	<0.001	0.588	0.152
Ortolan bunting	8		1.295	<0.001	0.295	0.596
Red-backed shrike	113		0.745	<0.001	0.555	<0.001
Rock bunting	21	Poor	1.382	<0.001	0.381	0.414
Skylark	36		1.408	<0.001	0.486	0.071
Stonechat	54		0.732	<0.001	0.285	<0.001
Tawny pipit	13		0.970	<0.001	0.287	0.889
Tree pipit	31		2.080	<0.001	0.462	0.525
Turtle dove	59		1.204	<0.001	0.318	0.152
Whitethroat	40		1.423	<0.001	0.586	0.172
Woodlark	91		1.257	<0.001	0.548	0.030
Yellowhammer	28		1.418	<0.001	0.371	0.909

other factors that affect local abundance (Brambilla, Scridel, et al., 2020; Dallas & Hastings, 2018; Holt, 2020; Osorio-Olvera et al., 2019). The awareness about the latter led to the conception of the wedge-shaped relationship between suitability and abundance (VanDerWal et al., 2009), with SDMs able to predict the upper limits of population abundance, a hypothesis later supported by further findings (Jiménez-Valverde et al., 2021 and references therein; Jiménez-Valverde, 2011). However, coarse-grained SDMs failed to find even a wedge-shaped relationship between suitability and abundance in more than 500 plant species at the European level (Sporbert et al., 2020). Taken together, the evidence led Lee-Yaw et al. (2022) to conclude that SDMs in general do not properly work as predictors of abundance, while the wedge-shaped relationship sometimes observed might make SDMs useful to identify sites where abundance could potentially be high, but care is also needed in those cases. Moving beyond occurrence-based distribution modeling looks key to improving abundance predictions (Lee-Yaw et al., 2022).

Our work focused on 14 bird species tied to open or semi-open habitats. Although based on a relatively low number of species, our results virtually re-open the debates, emphasizing the importance of grains and of spatiotemporal resolution more in general. Our study

provides strong evidence both for the occurrence of a relatively strong relationship between suitability and abundance, and for the “inflation” that including absence sites may cause when assessing such a relationship. Consistently with previous evidence (Weber et al., 2017), we found that SDMs integrating local variables may successfully predict abundance in the majority of species, even when focusing only on presence sites. Suitability obtained by SDMs worked out at the finer grain (1 ha), predicted significantly or almost so ($p < 0.1$) local latent abundance, whether including only long-term or also current climatic variables. This held true even when focusing on presence only transects (for 62% of species). Conversely, only one model at the larger grain (1 km) correctly predicted abundance. Furthermore, when focusing on all transects, that is also including those where a species was not found, all N-mixture models showed a significant effect of environmental suitability on abundance, irrespective of the spatiotemporal resolution of the SDMs. This supports the “inflation effect” of including absence sites previously highlighted (Dallas & Hastings, 2018).

Grain was also the only significant and strongly supported predictor of R^2 of N-mixture models, with fine-grained models invariably performing better than coarse-grained ones. This clearly outlines once again how

TABLE 3 Summary of N-mixture models for occurrence-only transects, with Poisson or negative binomial models; the latter were used when the Poisson model was not validated (see text).

Species	N_tr	SDM notes	Mixture family	Coefficient	p-value	R ²	p-value
1 km, long-term climate							
Cirl bunting	78		Poisson	0.073	0.374	0.260	0.283
Corn bunting	40		NB	0.175	0.214	0.388	0.172
Moltoni's warbler	40	Poor	Poisson	-0.093	0.328	0.336	0.313
Red-backed shrike	113		NB	0.093	0.150	0.478	0.626
Rock bunting	21	Poor	Poisson	0.121	0.488	0.550	0.111
Skylark	36		Poisson	0.104	0.406	0.367	0.273
Stonechat	54		Poisson	0.175	0.112	0.280	0.141
Tawny pipit	13	Reject	Poisson	0.138	0.608	0.521	0.788
Tree pipit	31		Poisson	0.064	0.643	0.173	0.667
Turtle dove	59		Poisson	-0.024	0.808	0.160	0.535
Whitethroat	40		NB	-0.030	0.810	0.201	0.141
Woodlark	91		Poisson	-0.044	0.565	0.099	0.061
Yellowhammer	28	Poor	Poisson	0.077	0.592	0.207	0.768
1 ha, long-term climate							
Cirl bunting	78		Poisson	0.179	0.027	0.299	0.424
Corn bunting	40		NB	0.378	0.001	0.493	0.091
Moltoni's warbler	40		Poisson	0.311	0.006	0.444	0.404
Red-backed shrike	113		Poisson	0.217	0.001	0.353	0.556
Rock bunting	21	Reject	Poisson	0.348	0.057	0.617	0.222
Skylark	36		Poisson	0.251	0.062	0.417	0.303
Stonechat	54		Poisson	0.199	0.039	0.302	0.202
Tawny pipit	13	Poor	Poisson	0.270	0.289	0.553	0.990
Tree pipit	31		Poisson	0.198	0.193	0.215	0.828
Turtle dove	59		Poisson	0.092	0.391	0.170	0.556
Whitethroat	40		Poisson	0.412	<0.001	0.518	0.081
Woodlark	91		Poisson	0.040	0.615	0.098	0.071
Yellowhammer	28		Poisson	0.084	0.563	0.255	0.970
1 km, current climate							
Cirl bunting	78		Poisson	0.068	0.392	0.259	0.303
Corn bunting	40		NB	0.033	0.811	0.365	0.222
Moltoni's warbler	40		Poisson	-0.026	0.786	0.321	0.242
Red-backed shrike	113		NB	0.111	0.084	0.268	0.586
Rock bunting	21		Poisson	0.116	0.505	0.549	0.081
Skylark	36		Poisson	0.124	0.361	0.370	0.333
Stonechat	54		Poisson	0.152	0.164	0.271	0.121
Tawny pipit	13	Reject	Poisson	0.167	0.520	0.527	0.788
Tree pipit	31		Poisson	0.068	0.636	0.173	0.747
Turtle dove	59		Poisson	-0.024	0.808	0.160	0.606
Whitethroat	40		NB	-0.021	0.870	0.201	0.121
Woodlark	91		Poisson	-0.047	0.544	0.100	0.061
Yellowhammer	28	Poor	Poisson	-0.021	0.893	0.246	0.949

TABLE 3 (Continued)

Species	N_tr	SDM notes	Mixture family	Coefficient	p-value	R ²	p-value
1 ha, current climate							
Cirl bunting	78		Poisson	0.081	0.018	0.306	0.444
Corn bunting	40		NB	0.418	<0.001	0.514	0.091
Moltoni's warbler	40		Poisson	0.106	0.010	0.426	0.303
Red-backed shrike	113		NB	0.229	0.001	0.323	0.667
Rock bunting	21	Poor	Poisson	0.169	0.080	0.607	0.212
Skylark	36		Poisson	0.131	0.046	0.427	0.364
Stonechat	54		Poisson	0.096	0.037	0.303	0.202
Tawny pipit	13		Poisson	0.267	0.285	0.555	0.970
Tree pipit	31		Poisson	0.173	0.115	0.240	0.848
Turtle dove	59		Poisson	0.106	0.482	0.166	0.545
Whitethroat	40		Poisson	0.093	<0.001	0.505	0.091
Woodlark	91		NB	0.078	0.330	0.106	0.091
Yellowhammer	28		Poisson	0.144	0.226	0.240	0.859

Note: For the ortolan bunting, the models were not implemented because of the too low sample size ($N = 8$ occurrence transects). See Table 2 for an explanation of the abbreviations used.

the incorporation of local variables into SDMs at fine scales is key to predicting local abundance. Interestingly, the large-scale approaches that did not find positive relationships between suitability and local abundance were often based on coarse-grain SDMs, usually much coarser than our “large-scale,” coarse-grained models (Dallas & Hastings, 2018; Santini et al., 2019; Sporbert et al., 2020), which were worked out based on a 1 km grain.

All N-mixture models at the fine grain were also statistically validated, when using negative binomial instead of Poisson error for overdispersed models. This means that suitability estimated by SDMs, coupled with factors affecting the detection process, may be useful in estimating the “true” local abundance of many species. Even more accurate results could potentially be achieved with climatic data at higher resolution: in our study, climatic predictors were only available at a resolution coarser than the 1-ha one we adopted as the fine-grain level, but the increasing availability of well performing microclimate models will promote the incorporation of more precise climatic variables.

Conclusions

While recent works pointed toward a lack of evidence in favor of a consistent relationship between environmental suitability derived from SDMs and local abundance, by working at a very fine grain we demonstrated that suitability may predict local abundance in the majority of the

farmland bird species we surveyed. Compared with most of the previous similar assessments, we worked with finer grained SDMs, to the point that even our coarse or large-scale grain is definitely finer than previously adopted grains. Our results showed that SDMs worked out at really fine scales, approaching the average size of territory or home range of target species, are needed to predict local abundance effectively. This may result from the fact that each single cell may represent a potential territory/home range, and hence having higher suitability over an area means harboring more potential territories within that area.

The increasing availability of fine-scale environmental variables, made possible by for example progress in satellite imagery (Koma et al., 2022) and microclimate modeling (Klinges et al., 2022), will increase in turn the possibility of working out ecologically representative models at the territory/home-range scale, or at even finer grains, focusing on specific resources (Alessandrini et al., 2022). Fine-grain models may increase the possibility of effectively predicting the local abundance of more species and provide better tools for conservation and habitat management.

ACKNOWLEDGMENTS

We are very grateful to R. Falco, V. Bergero, E. Vigo, M. Belardi for their helpful collaboration with the organization of data collection and project management, to L. M. Nietfeld for contributing to analyses and data management, and to colleagues and friends who provided

some additional occurrence locations of the target species. We are very grateful to the Editor, N. Gotelli, and three anonymous reviewers for helpful comments on previous versions of the manuscript.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data (Brambilla et al., 2023) are available in the Milan University Dataverse at https://doi.org/10.13130/RD_UNIMI/GUZWBX.

ORCID

Mattia Brambilla  <https://orcid.org/0000-0002-7643-4652>

REFERENCES

- Alessandrini, C., D. Scridel, L. Boitani, P. Pedrini, and M. Brambilla. 2022. "Remotely Sensed Variables Explain Microhabitat Selection and Reveal Buffering Behaviours against Warming in a Climate-Sensitive Bird Species." Edited by Nathalie Pettorelli and Tobias Kuemmerle." *Remote Sensing in Ecology and Conservation* 8(April): 615–628. <https://doi.org/10.1002/RSE2.265>.
- Altermatt, F., and E. A. Fronhofer. 2018. "Dispersal in Dendritic Networks: Ecological Consequences on the Spatial Distribution of Population Densities." *Freshwater Biology* 63(1): 22–32. <https://doi.org/10.1111/FWB.12951>.
- Arnold, T. W. 2010. "Uninformative Parameters and Model Selection Using Akaike's Information Criterion." *The Journal of Wildlife Management* 74(6): 1175–78. <https://doi.org/10.1111/j.1937-2817.2010.tb01236.x>.
- Assandri, G., G. Bogliani, P. Pedrini, and M. Brambilla. 2019. "Toward the Next Common Agricultural Policy Reform: Determinants of Avian Communities in Hay Meadows Reveal Current Policy's Inadequacy for Biodiversity Conservation in Grassland Ecosystems." *Journal of Applied Ecology* 56(January): 604–617. <https://doi.org/10.1111/1365-2664.13332>.
- Bartoń, K. 2020. "MuMIn: Multi-Model Inference." R Package Version 1.43.17. <https://cran.r-project.org/web/packages/MuMIn/index.html>.
- Bean, W. T., L. R. Prugh, H. Robert Stafford, S. Butterfield, M. Westphal, and J. S. Brashares. 2014. "Species Distribution Models of an Endangered Rodent Offer Conflicting Measures of Habitat Quality at Multiple Scales." *Journal of Applied Ecology* 51(4): 1116–25. <https://doi.org/10.1111/1365-2664.12281>.
- Brambilla, M. 2022. "Raster Files of Climate Refugia for Single Target Species and Multi-Species Refugia for High-Elevation Birds in the Alps." UNIMI Dataverse. https://doi.org/10.13130/RD_UNIMI/ESCYNG.
- Brambilla, M., G. Bazzi, and L. Ilahiane. 2023. "Replication Data for 'The Effectiveness of Species Distribution Models in Predicting Local Abundance Depends on Model Grain Size.'" UNIMI Dataverse. https://doi.org/10.13130/RD_UNIMI/GUZWBX.
- Brambilla, M., R. Falco, and I. Negri. 2012. "A Spatially Explicit Assessment of Within-Season Changes in Environmental Suitability for Farmland Birds along an Altitudinal Gradient." Edited by Steve Redpath and Darren Evans." *Animal Conservation* 15(6): 638–647. <https://doi.org/10.1111/j.1469-1795.2012.00561.x>.
- Brambilla, M., and G. F. Ficetola. 2012. "Species Distribution Models as a Tool to Estimate Reproductive Parameters: A Case Study with a Passerine Bird Species." *Journal of Animal Ecology* 81(4): 781–87. <https://doi.org/10.1111/j.1365-2656.2012.01970.x>.
- Brambilla, M., and F. Gatti. 2022. "No more Silent (and Uncoloured) Springs in Vineyards? Experimental Evidence for Positive Impact of Alternate Inter-Row Management on Birds and Butterflies." *Journal of Applied Ecology* 59: 2166–78. <https://doi.org/10.1111/1365-2664.14229>.
- Brambilla, M., M. Gustin, M. Cento, L. Ilahiane, and C. Celada. 2019. "Predicted Effects of Climate Factors on Mountain Species Are Not Uniform over Different Spatial Scales." *Journal of Avian Biology* 50(9): e02162. <https://doi.org/10.1111/jav.02162>.
- Brambilla, M., M. Gustin, M. Cento, L. Ilahiane, and C. Celada. 2020. "Habitat, Climate, Topography and Management Differently Affect Occurrence in Declining Avian Species: Implications for Conservation in Changing Environments." *Science of the Total Environment* 742: 140663. <https://doi.org/10.1016/j.scitotenv.2020.140663>.
- Brambilla, M., D. Rubolini, O. Appukuttan, G. Calvi, D. N. Karger, P. Kmecl, T. Mihelič, et al. 2022. "Identifying Climate Refugia for High-Elevation Alpine Birds under Current Climate Warming Predictions." *Global Change Biology* 28(April): 4276–91. <https://doi.org/10.1111/GCB.16187>.
- Brambilla, M., D. Scridel, G. Bazzi, L. Ilahiane, A. Iemma, P. Pedrini, E. Bassi, et al. 2020. "Species Interactions and Climate Change: How the Disruption of Species Co-Occurrence Will Impact on an Avian Forest Guild." *Global Change Biology* 26(3): 1212–24. <https://doi.org/10.1111/gcb.14953>.
- Brooks, M. E., K. Kristensen, K. J. van Benthem, A. Magnusson, C. W. Berg, A. Nielsen, H. J. Skaug, M. Mächler, and B. M. Bolker. 2017. "GlmM TMB Balances Speed and Flexibility among Packages for Zero-Inflated Generalized Linear Mixed Modeling." *The R Journal* 9(2): 378–400. <https://doi.org/10.3929/ETHZ-B-000240890>.
- Carrascal, L. M., P. Aragón, D. Palomino, and J. M. Lobo. 2015. "Predicting Regional Densities from Bird Occurrence Data: Validation and Effects of Species Traits in a Macaronesian Island." *Diversity and Distributions* 21(11): 1284–94. <https://doi.org/10.1111/ddi.12368>.
- Ceresa, F., M. Brambilla, L. Kvist, S. Vitulano, M. Pes, L. Tomasi, P. Pedrini, A. Hilpold, and P. Kranebitter. 2023a. "Matrices of Genetic Distances and Resistance Surfaces for the Paper 'Landscape Characteristics Influence Regional Dispersal in a High-Elevation Specialist Migratory Bird, the Water Pipit Anthus spinoletta'—Molecular Ecology, 2023." UNIMI Dataverse. https://doi.org/10.13130/RD_UNIMI/IFCPAX.
- Ceresa, F., M. Brambilla, L. Kvist, S. Vitulano, M. Pes, L. Tomasi, P. Pedrini, A. Hilpold, and P. Kranebitter. 2023b. "Landscape Characteristics Influence Regional Dispersal in a High-Elevation Specialist Migratory Bird, the Water Pipit Anthus spinoletta." *Molecular Ecology* 32(8): 1875–92. <https://doi.org/10.1111/MEC.16853>.

- Dallas, T. A., and A. Hastings. 2018. "Habitat Suitability Estimated by Niche Models Is Largely Unrelated to Species Abundance." *Global Ecology and Biogeography* 2018: 1448–56. <https://doi.org/10.1111/geb.12820>.
- Dallas, T. A., and L. Santini. 2020. "The Influence of Stochasticity, Landscape Structure and Species Traits on Abundant–Centre Relationships." *Ecography* 43(9): 1341–51. <https://doi.org/10.1111/ecog.05164>.
- Elith, J., and J. R. Leathwick. 2009. "Species Distribution Models: Ecological Explanation and Prediction across Space and Time." *Annual Review of Ecology, Evolution, and Systematics* 40: 677–697. <http://www.annualreviews.org/eprint/HWR4cusJrXYCSPZ9sUDj.full>.
- Elith, J., S. J. Phillips, T. Hastie, M. Dudík, Y. E. Chee, and C. J. Yates. 2011. "A Statistical Explanation of MaxEnt for Ecologists." *Diversity and Distributions* 17(1): 43–57. <https://doi.org/10.1111/j.1472-4642.2010.00725.x>.
- Engler, J. O., D. Stiels, K. Schidelko, D. Strubbe, P. Quillfeldt, and M. Brambilla. 2017. "Avian SDMs: Current State, Challenges, and Opportunities." *Journal of Avian Biology* 48: 1483–1504. <https://doi.org/10.1111/jav.01248>.
- Estrada, A., and B. Arroyo. 2012. "Occurrence Vs Abundance Models: Differences between Species with Varying Aggregation Patterns." *Biological Conservation* 152(August): 37–45. <https://doi.org/10.1016/J.BIOCON.2012.03.031>.
- Fiske, I. J., and R. B. Chandler. 2011. "Unmarked: An R Package for Fitting Hierarchical Models of Wildlife Occurrence and Abundance." *Journal of Statistical Software* 43(10): 1–23. <https://doi.org/10.18637/jss.v043.i10>.
- Hartig, F. 2020. "DHARMA: Residual Diagnostics for Hierarchical (Multi-Level/Mixed) Regression Models." R Package. <https://cran.r-project.org/web/packages/DHARMA/vignettes/DHARMA.html>.
- Holt, R. D. 2020. "Reflections on Niches and Numbers." *Ecography* 43(3): 387–390. <https://doi.org/10.1111/ecog.04828>.
- Hutchinson, G. E. 1957. "Concluding Remarks." *Cold Spring Harbor Symposia on Quantitative Biology* 22: 415–427.
- Jiménez-Valverde, A. 2011. "Relationship between Local Population Density and Environmental Suitability Estimated from Occurrence Data." *Frontiers of Biogeography* 3(2): 59–61. <https://doi.org/10.21425/F5FBG12408>.
- Jiménez-Valverde, A., P. Aragón, and J. M. Lobo. 2021. "Deconstructing the Abundance–Suitability Relationship in Species Distribution Modelling." *Global Ecology and Biogeography* 30(1): 327–338. <https://doi.org/10.1111/GEB.13204>.
- Jiménez-Valverde, A., J. M. Lobo, and J. Hortal. 2008. "Not as Good as they Seem: The Importance of Concepts in Species Distribution Modelling." *Diversity and Distributions* 14(6): 885–890. <https://doi.org/10.1111/j.1472-4642.2008.00496.x>.
- Kaky, E., V. Nolan, A. Alatawi, and F. Gilbert. 2020. "A Comparison between Ensemble and MaxEnt Species Distribution Modelling Approaches for Conservation: A Case Study with Egyptian Medicinal Plants." *Ecological Informatics* 60(November): 101150. <https://doi.org/10.1016/j.ecoinf.2020.101150>.
- Karger, D. N., O. Conrad, J. Böhner, T. Kawohl, H. Kreft, R. W. Soria-Auza, N. E. Zimmermann, H. Peter Linder, and M. Kessler. 2017. "Climatologies at High Resolution for the Earth's Land Surface Areas." *Scientific Data* 4: 170122. <https://doi.org/10.1038/sdata.2017.122>.
- Karger, D. N., O. Conrad, J. Böhner, T. Kawohl, H. Kreft, R. W. Soria-Auza, N. E. Zimmermann, H. Peter Linder, and M. Kessler. 2021. "Climatologies at High Resolution for the Earth's Land Surface Areas." *EnviDat*. <https://doi.org/10.16904/envidat.228.v2.1>.
- Klinges, D. H., J. P. Duffy, M. R. Kearney, and I. M. D. Maclean. 2022. "Mcera5: Driving Microclimate Models with ERA5 Global Gridded Climate Data." *Methods in Ecology and Evolution* 13(7): 1402–11. <https://doi.org/10.1111/2041-210X.13877>.
- Koma, Z., A. C. Seijmonsbergen, W. G. Meiert, F. Nattino, J. Groot, H. Sierdsema, P. B. F. Ruud, and D. Kissling. 2022. "Better Together? Assessing Different Remote Sensing Products for Predicting Habitat Suitability of Wetland Birds." Edited by Yoan Fourcade." *Diversity and Distributions* 00(February): 1–15. <https://doi.org/10.1111/DDI.13468>.
- Lee-Yaw, J. A., J. L. McCune, S. Pironon, and S. N. Sheth. 2022. "Species Distribution Models Rarely Predict the Biology of Real Populations." *Ecography* 2022(6): e05877. <https://doi.org/10.1111/ECOG.05877>.
- Lobo, J. M., A. Jiménez-Valverde, and R. Real. 2008. "AUC: A Misleading Measure of the Performance of Predictive Distribution Models." *Global Ecology and Biogeography* 17(2): 145–151. <https://doi.org/10.1111/j.1466-8238.2007.00358.x>.
- Lunghi, E., R. Manenti, M. Mulargia, M. Veith, C. Corti, and G. F. Ficetola. 2018. "Environmental Suitability Models Predict Population Density, Performance and Body Condition for Microendemic Salamanders." *Scientific Reports* 8(1): 7527. <https://doi.org/10.1038/s41598-018-25704-1>.
- Maguire, K. C., D. Nieto-Lugilde, M. C. Fitzpatrick, J. W. Williams, and J. L. Blois. 2015. "Modeling Species and Community Responses to Past, Present, and Future Episodes of Climatic and Ecological Change." *Annual Review of Ecology, Evolution, and Systematics* 46(1): 343–368. <https://doi.org/10.1146/annurev-ecolsys-112414-054441>.
- Manzoor, S. A., G. Griffiths, and M. Lukac. 2018. "Species Distribution Model Transferability and Model Grain Size—Finer May Not Always Be Better." *Scientific Reports* 8(1): 7168. <https://doi.org/10.1038/s41598-018-25437-1>.
- Mazerolle, M. J. 2022. "Package 'AICcmodavg'—Model Selection and Multimodel Inference Based on (Q)AIC(C)." <https://cran.r-project.org/web/packages/AICcmodavg/AICcmodavg.pdf>.
- Muscarella, R., P. J. Galante, M. Soley-Guardia, R. A. Boria, J. M. Kass, M. Uriarte, and R. P. Anderson. 2014. "ENMeval: An R Package for Conducting Spatially Independent Evaluations and Estimating Optimal Model Complexity for MAXENT Ecological Niche Models." Edited by Jana McPherson." *Methods in Ecology and Evolution* 5(11): 1198–1205. <https://doi.org/10.1111/2041-210X.12261>.
- Neteler, M., M. Hamish Bowman, M. Landa, and M. Metz. 2012. "GRASS GIS: A Multi-Purpose Open Source GIS." *Environmental Modelling & Software* 31: 124–130. <https://doi.org/10.1016/j.envsoft.2011.11.014>.
- Osorio-Olvera, L., J. Soberón, and M. Falconi. 2019. "On Population Abundance and Niche Structure." *Ecography* 42(8): 1415–25. <https://doi.org/10.1111/ecog.04442>.
- Pennino, M. G., R. Vilela, and J. M. Bellido. 2019. "Effects of Environmental Data Temporal Resolution on the Performance of Species Distribution Models." *Journal of Marine Systems*

- 189(January): 78–86. <https://doi.org/10.1016/J.JMARSYS.2018.10.001>.
- Phillips, S. J., R. P. Anderson, and R. E. Schapire. 2006. “Maximum Entropy Modeling of Species Geographic Distributions.” *Ecological Modelling* 190(3): 231–259. <https://doi.org/10.1016/j.ecolmodel.2005.03.026>.
- Planillo, A., S. Kramer-Schadt, S. Buchholz, P. Gras, M. von der Lippe, and V. Radchuk. 2021. “Arthropod Abundance Modulates Bird Community Responses to Urbanization.” *Diversity and Distributions* 27(1): 34–49. <https://doi.org/10.1111/DDI.13169>.
- R Development Core Team. 2020. *A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. <https://www.r-project.org>.
- Rollan, À., A. Hernández-Matías, and J. Real. 2019. “Organic Farming Favours Bird Communities and their Resilience to Climate Change in Mediterranean Vineyards.” *Agriculture, Ecosystems and Environment* 269(January): 107–115. <https://doi.org/10.1016/j.agee.2018.09.029>.
- Santini, L., S. Pironon, L. Maiorano, and W. Thuiller. 2019. “Addressing Common Pitfalls Does Not Provide More Support to Geographical and Ecological Abundant-Centre Hypotheses.” *Ecography* 42(4): 696–705. <https://doi.org/10.1111/ECOG.04027>.
- Sporbert, M., P. Keil, G. Seidler, H. Bruelheide, U. Jandt, S. Aćić, I. Biurrun, et al. 2020. “Testing Macroecological Abundance Patterns: The Relationship between Local Abundance and Range Size, Range Position and Climatic Suitability among European Vascular Plants.” *Journal of Biogeography* 47(10): 2210–22. <https://doi.org/10.1111/JBI.13926>.
- Thuiller, W., M. Guéguen, J. Renaud, D. N. Karger, and N. E. Zimmermann. 2019. “Uncertainty in Ensembles of Global Biodiversity Scenarios.” *Nature Communications* 10(1): 1–9. <https://doi.org/10.1038/s41467-019-09519-w>.
- VanDerWal, J., L. P. P. Shoo, C. N. N. Johnson, and S. E. E. Williams. 2009. “Abundance and the Environmental Niche: Environmental Suitability Estimated from Niche Models Predicts the Upper Limit of Local Abundance.” *American Naturalist* 174(2): 282–291. <https://doi.org/10.1086/600087>.
- Vignali, S., A. G. Barras, R. Arlettaz, and V. Braunisch. 2020. “SDMtune: An R Package to Tune and Evaluate Species Distribution Models.” *Ecology and Evolution* 10(20): 11488–506. <https://doi.org/10.1002/ece3.6786>.
- Warren, D. L. 2012. “In Defense of ‘Niche Modeling’.” *Trends in Ecology and Evolution* 27(9): 497–500. <https://doi.org/10.1016/j.tree.2012.03.010>.
- Weber, M. M., R. D. Stevens, J. A. F. Diniz-Filho, and C. E. V. Grelle. 2017. “Is there a Correlation between Abundance and Environmental Suitability Derived from Ecological Niche Modelling? A Meta-Analysis.” *Ecography* 40(7): 817–828. <https://doi.org/10.1111/ecog.02125>.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Brambilla, Mattia, Gaia Bazzi, and Luca Ilahiane. 2023. “The Effectiveness of Species Distribution Models in Predicting Local Abundance Depends on Model Grain Size.” *Ecology* e4224. <https://doi.org/10.1002/ecy.4224>