

# Species distribution models as a tool to estimate reproductive parameters: a case study with a passerine bird species

Mattia Brambilla<sup>1,2</sup> and Gentile F. Ficetola<sup>3\*</sup>

<sup>1</sup>Fondazione Lombardia per l'Ambiente, Settore Biodiversità e Aree protette, Piazza Diaz 7, I-20123 Milano, Italy; <sup>2</sup>Museo delle Scienze, Sezione Zoologia dei Vertebrati, Via Calepina 14, I-38122 Trento, Italy; and <sup>3</sup>Dipartimento di Scienze dell'Ambiente e del Territorio, Università di Milano-Bicocca, Piazza della Scienza 1, I-20126 Milano, Italy

## Summary

1. Correlative species distribution models (SDMs) assess relationships between species distribution data and environmental features, to evaluate the environmental suitability (ES) of a given area for a species, by providing a measure of the probability of presence. If the output of SDMs represents the relationships between habitat features and species performance well, SDM results can be related also to other key parameters of populations, including reproductive parameters. To test this hypothesis, we evaluated whether SDM results can be used as a proxy of reproductive parameters (breeding output, territory size) in red-backed shrikes (*Lanius collurio*).

2. The distribution of 726 shrike territories in Northern Italy was obtained through multiple focused surveys; for a subset of pairs, we also measured territory area and number of fledged juveniles. We used Maximum Entropy modelling to build a SDM on the basis of territory distribution. We used generalized least squares and spatial generalized mixed models to relate territory size and number of fledged juveniles to SDM suitability, while controlling for spatial autocorrelation.

3. Species distribution models predicted shrike distribution very well. Territory size was negatively related to suitability estimated through SDM, while the number of fledglings significantly increased with the suitability of the territory. This was true also when SDM was built using only spatially and temporally independent data.

4. Results show a clear relationship between ES estimated through presence-only SDMs and two key parameters related to species' reproduction, suggesting that suitability estimated by SDM, and habitat quality determining reproduction parameters in our model system, are correlated. Our study shows the potential use of SDMs to infer important fitness parameters; this information can have great importance in management and conservation.

**Key-words:** breeding success, environmental suitability, habitat quality, *Lanius collurio*, MaxEnt, passerine birds, territory size

## Introduction

Modelling habitat suitability and species' distribution are increasingly important subjects in ecology and conservation. Correlative species distribution models (SDMs) assess relationships between species distribution data and environmental features, to evaluate the suitability of a given area for a species of interest. Models provide a measure of the probability of presence, which can be used to define species' spatial occurrence (e.g. Graham *et al.* 2004; Brambilla *et al.* 2009), inform surveys (Raxworthy *et al.* 2003; Bourg, McShea & Gill 2005), evaluate impacts of climate and habitat change (e.g. Thuiller *et al.* 2005a; Brambilla *et al.* 2010b; Fouquet

*et al.* 2010; Elith *et al.* 2011), test evolutionary hypotheses (e.g. Peterson, Soberón & Sánchez-Cordero 1999; Graham *et al.* 2004), predict species invasions (Roura-Pascual *et al.* 2004; Thuiller *et al.* 2005b; Ficetola, Thuiller & Maud 2007; Ficetola *et al.* 2010; Reshetnikov & Ficetola 2011) and inform conservation planning (e.g. Araújo & Williams 2000; Ferrier *et al.* 2002). In recent years, SDMs have become one of the most frequent tasks in conservation, and presence-only SDMs are becoming prevalent because they do not require absence data (Jiménez-Valverde, Lobo & Hortal 2008).

Collecting data on population abundance or fitness is much more complex than simply recording presence/absence data (Ficetola, Thuiller & Padoa-Schioppa 2009; VanDerWal *et al.* 2009). Presence at a given point does not always mean that location is suitable for a species or that

\*Correspondence author. E-mail: francesco.ficetola@unimib.it

habitat is sufficient to maintain a stable population, as particular conditions such as sink habitats or temporal variation in environmental features can generate occurrences in areas not highly suitable (Bateman, VanDerWal & Johnson 2012). Nevertheless, correlative SDMs implicitly assume that presence at a given location indicates that an organism has been able to settle, survive and reproduce at that point or that individuals have selected the site as potentially suitable for reproduction, and the output of SDM is considered as a measure of the suitability of environmental features [environmental suitability (ES)] for the occurrence of the target species (e.g. Kearney & Porter 2009; Soberon & Nakamura 2009; VanDerWal *et al.* 2009). If the output of SDMs represents the relationships between a species and its habitat well, it is possible that SDM results are related not only to the probability of occurrence, but also to other key parameters of populations. VanDerWal *et al.* (2009) showed that ES, obtained through presence-only SDMs, can also predict the upper limit of abundance of study species. Following the same line of reasoning, we hypothesized that ES may be related also to other critical parameters of population performance, such as those related to reproduction. To our knowledge, the potential of SDMs to predict reproductive parameters, estimating the fitness consequences of habitat selection, has not been explicitly investigated (but see Titeux *et al.* 2007; which suggested that reproductive niche is narrower than 'occurrence' niche, and Ficetola, Thuiller & Padoa-Schioppa 2009). Given that it is relatively easy to obtain well-performing SDMs, while it is often hard to get detailed data on fitness components and on habitat quality (i.e. the ability of the environment to provide conditions appropriate for individual and population persistence; Hall, Krausman & Morrison 1997) over broad spatial scales, it is important to assess whether the results of SDMs can be generalized to other population parameters.

In this study, we assessed whether suitability obtained through SDMs can be used as a proxy of two important parameters related to habitat quality for the reproduction of red-backed shrikes (*Lanius collurio*): territory size and breeding output. Individuals settled in high-quality sites often defend small territories or exploit small ranges, as they can find enough resources within relatively small areas, limiting the energetic costs of territory defence (e.g. Smith & Shugart 1987; Pasinelli 2000; Saïd *et al.* 2009). Breeding output is an important measure of fitness of breeding pairs, and it is strongly affected by habitat quality; pairs breeding in high-quality territories should experience high breeding output (Newton 1998; Sergio, Pedrini & Marchesi 2003; Förschler *et al.* 2005). Therefore, if ES provided by SDMs can be used as a proxy for parameters related to reproduction, we expect that ES should be negatively related with territory size, and positively related with breeding output.

## Materials and methods

The red-backed shrike is a passerine bird typical of 'traditional' farmed landscapes (Cramp 1993; Brambilla *et al.* 2010b and refer-

ences therein). During 2007–2010, the Lombardy region (Northern Italy; about 24 000 km<sup>2</sup>) was monitored to detect the distribution of shrike breeding territories; we obtained a total of 726 territories. Sampling was performed during a comprehensive regional project aimed at providing the most complete shrike distribution data, and covering the whole territory of the region; different localities were sampled in different years (see Brambilla *et al.* 2009, 2010b for details on field procedures). We used Maximum Entropy Modelling (MaxEnt) to build SDMs relating shrike presence to environmental features (Phillips, Anderson & Schapire 2006). MaxEnt assesses the probability of presence in a given cell on the basis of environmental features in that cell; it is considered one of the most efficient approaches to SDM using presence-only data (Elith *et al.* 2006, 2011). Presence data were represented by the centres of the shrike territories, avoiding the multiple use of the same territory if mapped during more than 1 year. As environmental variables, we used land cover variables and hedgerow length (vector lines; two different themes describing continuous and interrupted hedgerows), treated as raster cells of 20 × 20 m derived from a detailed land-use map dated to 2007 (DUSAF 2:1; Regione Lombardia & Ersaf; <http://www.cartografia.regione.lombardia.it>). Cell size corresponds to the resolution of the DUSAF data and to the accuracy at which shrikes were mapped (size of sampling unit equal to the grain size of environmental data; Elith & Leathwick 2009; Elith *et al.* 2011). For each 20 × 20 m cell, we measured land cover variables as the cover of 45 non-urban habitats in a 100-m radius from the cell centre (see Table S1 in Supporting Information). Similarly, hedgerows were treated as their respective length in the 100-m radius for each cell. We also included three topographic variables (elevation, slope and aspect) derived from a Digital Terrain Model (DTM) of the regional surface. Aspect is a circular variable and was therefore transformed into a categorical predictor before analyses (eight classes corresponding to 45° intervals). Therefore, this study integrated both raster and vector data (Elith *et al.* 2011). We did not include climatic data because in the whole study area climate is suitable for shrikes, except at the highest elevation (Casale & Brambilla 2009). Repeated surveys in the same sites revealed that land cover did not vary significantly during the study years. Background was created using 10 000 random points generated by MaxEnt. We used two approaches to assess model performance. First, we performed a 10-fold cross-validation, and calculated the area under the curve of the receiver-operator plot (AUC) as a measure of matching between the model results and presence records. Second, we evaluated omission errors. The evaluation of omission errors requires determining whether a cell is suitable; we performed this analysis using two different suitability thresholds: we assumed that a cell was suitable for shrikes if its suitability score was greater than the 10th percentile of training presence points (Pearson *et al.* 2007), or if its suitability was greater than the equal training specificity and sensitivity threshold (Bartel & Sexton 2009). Subsequently, we used a  $\chi^2$  test (1 d.f.) to compare observed frequencies of correct and incorrect predictions, and therefore to assess if models predict distribution significantly better than expected under random expectations (Roura-Pascual *et al.* 2004).

For 16 territories, we obtained an accurate measure of territory size. These were subjected to detailed mapping (multiple positions of birds recorded during several visits) in 2007 (see Brambilla *et al.* 2009). For 19 successful pairs intensively surveyed during the 2007 breeding season, we assessed the number of fledged juveniles, which was taken as a measure of breeding output. Both location and territory/breeding data come from a wide variety of survey areas, including high-quality (as pre-Alpine mown grasslands) and low-quality ones (e.g. suburban restored and residual open areas).

To test for the potential use of ES provided by SDMs as a proxy of more complex parameters related with reproduction, we assessed the relationship between (i) territory ES (measured as average ES of cells included within the territory) and territory size for the 16 territories; (ii) territory ES (taken at territory centre) and the number of fledged juveniles, for the 19 pairs. In the latter analysis, we used ES at territory centre because for some of these pairs the detailed boundary of territory was not available and the average ES could not be calculated. Notably, ES of the territory centre was strictly correlated with average ES for well-delimited territories (for the 10 well-delimited territories with fledglings recorded,  $r = 0.86$ ,  $P = 0.002$ ). In these analyses, the same data were used to build the SDM and test its ability in predicting breeding parameters. We also tested whether SDM can predict breeding parameters in territories different from those used to build the model. We split our data set into two different blocks: 2007 data (the year from which all data on reproductive parameters come), and 2008–2010 data. The 2008–2010 data set (604 territories) was used to build a second MaxEnt model: the ES calculated by that second model was related to breeding parameters of the spatially and temporally independent 2007 data.

Our data have a strong spatial structure, and spatial autocorrelation may affect the results of regression analyses (Beale *et al.* 2010). We therefore used generalized least squares models (GLS) to assess the relationship between territory size and ES, while taking into account spatial structure. GLS allows the incorporation of spatial structure into the error of the model and is considered among the techniques with the best performance for the analysis of spatial data (Dormann *et al.* 2007; Beale *et al.* 2010). We built GLS using maximum likelihood, we assessed significance using a likelihood ratio test (Beale *et al.* 2010) and calculated pseudo- $R^2$  as a measure of the proportion of explained variation (Buse 1973; Griffis & Stedinger 2007).

Our measure of breeding output (number of fledged juveniles) is a count, requiring Poisson models. We therefore used spatial Generalized Linear Mixed Models via Penalized Quasi-Likelihood (glmmPQL), assuming Poisson error distributions, for the analysis of the relationship between breeding output and ES. glmmPQL is an extension of GLS that allows building spatial models with non-normal dependent data; simulations showed that this is among the best-performing methods with non-normal spatial data (Dormann *et al.* 2007). To assess the proportion of explained variation in glmmPQL, we calculated pseudo- $R^2$  as the squared Pearson correlation coefficient between observed and predicted values (Kissling &

Carl 2008). For both GLS and glmmPQL, we used a Gaussian spatial correlation structure. In real-world analyses, the true error structure is unknown and is unlikely to exactly match a specified function (Beale *et al.* 2010); thus, we tested the robustness to the selection of error structure. Models with different structures (spherical, exponential) yielded identical results, confirming that our results are robust. We performed all analyses in R 2.12 (R Development Core Team 2010) using packages MASS and NLME (Venables & Ripley 2002; Pinheiro *et al.* 2010).

## Results

A summary of the MaxEnt model is reported in Table 1; the important variables in the model were mostly habitat types already included in habitat preference models developed in Lombardy and elsewhere (Brambilla *et al.* 2009, 2010b; and references therein) and perfectly match with previous knowledge on species ecology (Cramp 1993); the 10-fold cross-validated AUC was  $0.927 \pm 0.009$ , indicating excellent performance. Considering the 10th percentile training presence threshold, 10.7% of the study area was suitable for shrikes (Fig. 1). This threshold correctly predicted presence in 654 territories (expected correctly predicted presence: 77.7 territories), therefore the model predicted shrike distribution much better than expected by chance ( $\chi^2_1 = 4788$ ,  $P < 0.001$ ). Using the 'Equal training sensitivity and specificity threshold' yielded identical results (correctly predicted presences: 651, expected: 74.8;  $\chi^2_1 = 4954$ ,  $P < 0.001$ ).

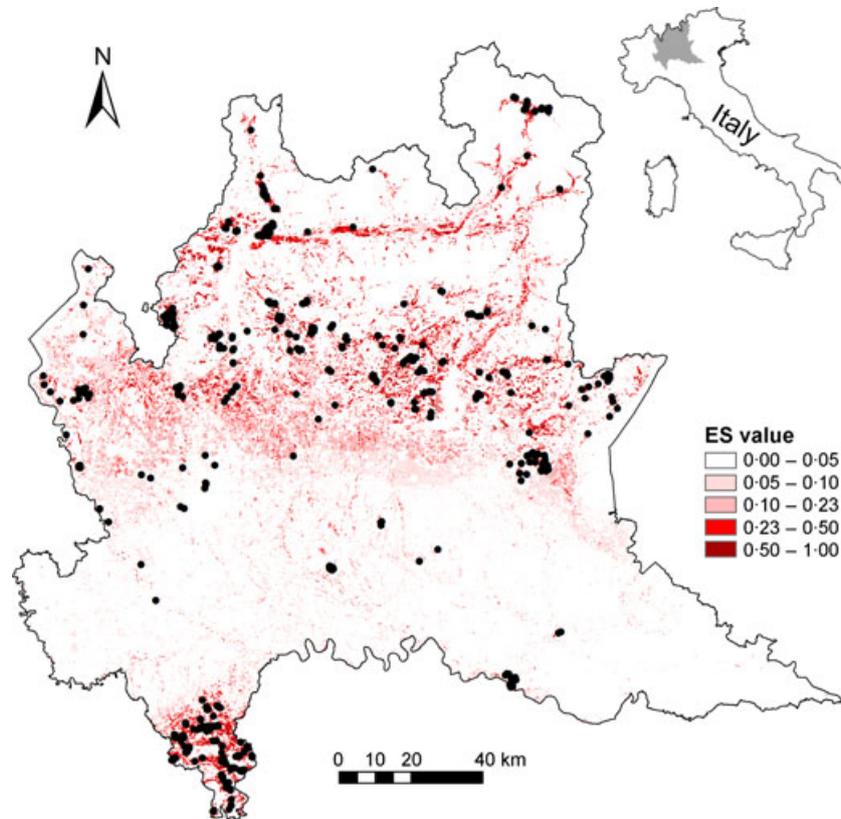
The surface of territories ranged between 2500 and 20 000 m<sup>2</sup>. Territory size was negatively related to their average ES estimated using the full-data model (GLS:  $\chi^2_1 = 4.10$ ,  $P = 0.043$ , pseudo- $R^2 = 0.23$ ; Fig. 2a). The number of fledglings ranged between one and three, and significantly increased with the ES of the territory (glmmPQL:  $t_{17} = 3.09$ ,  $P = 0.007$ , pseudo- $R^2 = 0.34$ ; Fig. 2b). Breeding shrikes successfully fledged three juveniles only in the territories with the highest ES (Fig. 2b).

If the 2008–2010 data set is used for calibration, the MaxEnt model was nearly identical (correlation between the

**Table 1.** Per cent contribution, permutation importance and type of effect for variables with a contribution to the model higher than 1%

Variable	% contribution	Permutation importance	Overall effect
Permanent grassland	34.3	13.3	Positive
Elevation	28.3	26.2	ES higher at intermediate values
Arable grassland	7.1	14.3	Positive
Broadleaved woodland (medium or high tree density; coppiced)	6.1	5.8	Negative
Permanent grassland with scattered trees and shrubs	3.7	2.4	Positive
Slope	3.4	2.1	Negative at high values
Montane grassland	3.3	2.7	Positive
Aspect (8 categories)	1.7	0.1	Slight positive effect of southern aspects
Abandoned areas with herbaceous or shrub vegetation	1.4	0.7	Positive
Shrubland with trees	1.2	1.2	Positive
Mixed woodland (medium or high tree density; coppiced)	1.0	1.7	Negative
Sum	91.5		

ES, environmental suitability.



**Fig. 1.** MaxEnt model for red-backed shrike in the Lombardy Region. The darkest colour represents the highest environmental suitability; black dots represent territories used for model calibration. 0.23 is the equal training sensitivity/specificity threshold.

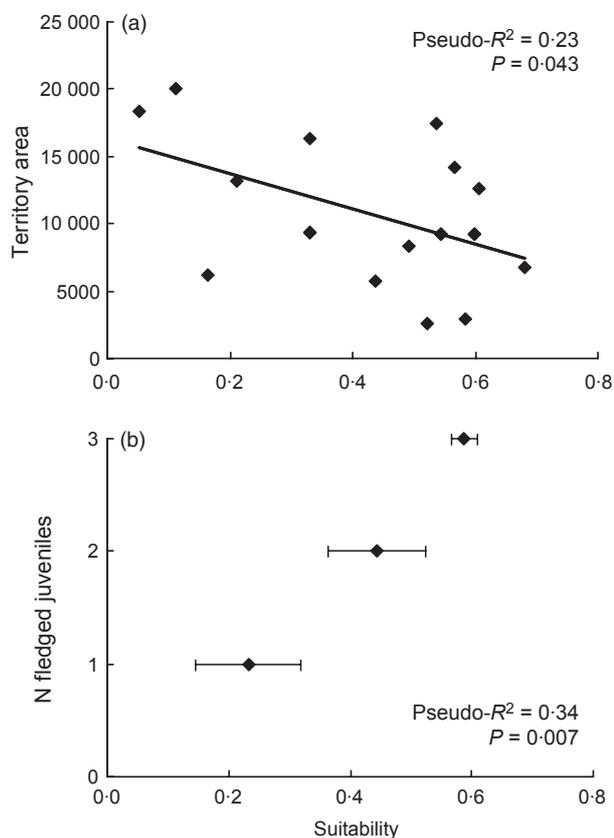
full-data and the 2008–2010 model:  $r = 0.96$ ). ES predicted by the model built with 2008–2010 data set was negatively related to territory size ( $\chi^2_1 = 4.79$ ,  $P = 0.029$ ; pseudo- $R^2 = 0.26$ ) and positively related to the number of fledglings ( $t_{17} = 2.94$ ,  $P = 0.009$ , pseudo- $R^2 = 0.32$ ) observed in 2007.

## Discussion

Our study shows a clear relationship between ES (i.e. a measure of the probability of species' occurrence according to species-habitat relationships) estimated through presence-only SDMs, and two key parameters related to species' reproduction. In high suitability sites, shrikes have higher breeding output and defend smaller territories than in low-suitability sites, suggesting that ES estimated by SDM and habitat quality determining reproduction parameters are correlated in our system. Several SDM techniques (including MaxEnt) describe species distribution very well, and their outcome has been extended to the prediction of species' abundance at local scale (VanDerWal *et al.* 2009). Our analysis shows that the results of SDMs may go beyond these applications, and can potentially be used for a first inference on important population parameters, including fitness consequences of selecting a given area, which is a relatively poorly documented topic of great relevance for conservation.

Correlative SDMs aim at evaluating the probability of presence of species on the basis of distribution data, so what is the conceptual link between predicted probability of presence and habitat quality as fitness parameters? If habitat selection is adaptive, habitat attractiveness should somehow match reproductive performance (Sergio, Pedrini & Marchesi 2003; Battin 2004). This means that the same factors affecting habitat selection (thereby determining the deriving estimates of suitability) may affect in a similar way reproduction parameters (Sergio, Pedrini & Marchesi 2003; Sergio, Marchesi & Pedrini 2004; Ortego 2007; Brambilla *et al.* 2010a).

In correlative SDMs, the true effects of biotic and abiotic conditions on species' physiology and fitness are unknown; the statistical relationships observed between species' distribution and environmental features are the indirect consequences of these unknown relationships (Kearney & Porter 2009). In fact, correlative SDMs may match species distribution well even when some variables actually influencing the species are not considered, if the variables included into the model are collinear to the influencing ones (even if the use of indirect variables would limit model transferability). This suggests that ES scores of correlative models may represent indirect but good proxies of the complex (and sometimes unknown) relationships between organisms and their environment. In the case of our study species, the known relationships between shrike fitness and habitat traits are in



**Fig. 2.** Relationship between environmental suitability predicted by species distribution models and: (a) size of shrike territories ( $m^2$ ); (b) number of fledged juveniles. Error bars are SE.

agreement with this hypothesis. For instance, some of the variables with the highest contribution to the MaxEnt model (see Table 1) are environmental factors known to affect the number of red-backed shrike fledglings also in other areas of Europe, where the abundance of grassland-like habitats (especially meadows and pastures) is the most important determinant of fledgling rate (Golawski & Meissner 2008). Recent analyses suggested that correlative SDMs can have a better performance if researchers use *a priori* knowledge on species biology, to select the environmental variables more strongly related to population performance (Rödger *et al.* 2009). Statistical relationships of these models may represent more direct biological relationships, and can therefore also have a better performance in predicting fitness components.

Nevertheless, our study has some limitations. First, the sample size used to build SDM was large (726 territories), while the sample size available for territory size or breeding output was much smaller. This is common in ecological studies, as collecting reliable data on these parameters is extremely time consuming, and is possible only over limited spatial scales (Brambilla *et al.* 2009). Second, despite we detected strong and significant relationships, ES did not explain all the variation of the considered parameters. For instance, ES explained only 23% of variation in territory size (Fig. 2a). Similarly, pairs fledging only one juvenile bred in territories with a range of ES (Fig. 2b). The large amount of

unexplained variation suggests that other unmeasured factors may be important, such as biotic interactions, presence of nest sites, dispersal limitations, weather fluctuations or extreme events (Tryjanowski, Kuzniak & Diehl 2000; Roos & Pärt 2004; Golawski 2006; VanDerWal *et al.* 2009; Bateman, VanDerWal & Johnson 2012), and that SDMs provide information useful for preliminary, large-scale assessments, but cannot replace local/regional field studies.

We used high-resolution habitat data for modelling spatial distribution, which may enable to better represent fine-scaled ecological features affecting also breeding parameters. Most modelling exercises use bioclimatic data at a much coarser resolution: climatic data could be less suited to represent habitat quality for breeding. On the other hand, recent advances in remote sensing techniques will increase the availability of high-resolution habitat data, allowing a better assessment of ecologically relevant traits also at a fine spatial scale (Mendenhall *et al.* 2011).

Our study is a first example of the potential use of SDMs to infer important fitness parameters; these information can have great importance in applied ecology and conservation, owing to the difficulty of measure these parameters over large scales. Further studies will be useful to assess for which parameters and for which taxa these relationships are strong, and under what conditions.

## Acknowledgements

We are particularly grateful to E. Bassi, V. Bergero, F. Casale, F. Cecere, M. Chemollo, V. Longoni, M. Gobbini, I. Negri, S. Ravara, F. Reginato, M. Siliprandi, S. Vitulano for help with fieldwork, to G. Bogliani, G.M. Crovetto, R. Falco, P. Lenna, F. Piccarolo for kind support. The comments of J.O. Engler and one anonymous reviewer improved an early version of this manuscript.

## References

- Araújo, M.B. & Williams, P.H. (2000) Selecting areas for species persistence using occurrence data. *Biological Conservation*, **96**, 331–345.
- Bartel, R.A. & Sexton, J.O. (2009) Monitoring habitat dynamics for rare and endangered species using satellite images and niche-based models. *Ecography*, **32**, 888–896.
- Bateman, B.L., VanDerWal, J. & Johnson, C.N. (2012) Nice weather for bettings: using weather events, not climate means, in species distribution models. *Ecography*, in press. DOI: 10.1111/j.1600-0587.2011.06871.x.
- Battin, J. (2004) When good animals love bad habitats: ecological traps and the conservation of animal populations. *Conservation Biology*, **18**, 1482–1491.
- Beale, C.M., Lennon, J.J., Yearsley, J.M., Brewer, M.J. & Elston, D.A. (2010) Regression analysis of spatial data. *Ecology Letters*, **13**, 246–264.
- Bourg, N.A., McShea, W.J. & Gill, D.E. (2005) Putting a CART before the search: successful habitat prediction for a rare forest herb. *Ecology*, **86**, 2793–2804.
- Brambilla, M., Casale, F., Bergero, V., Crovetto, M., Falco, R., Negri, I., Siccardi, P. & Bogliani, G. (2009) GIS-models work well, but are not enough: habitat preferences of *Lanius collurio* at multiple levels and conservation implications. *Biological Conservation*, **142**, 2033–2042.
- Brambilla, M., Bassi, E., Ceci, C. & Rubolini, D. (2010a) Environmental factors affecting patterns of distribution and co-occurrence of two competing raptor species. *Ibis*, **152**, 310–322.
- Brambilla, M., Casale, F., Bergero, V., Bogliani, G., Crovetto, G.M., Falco, R., Roati, M. & Negri, I. (2010b) Glorious past, uncertain present, bad future? Assessing effects of land-use changes on habitat suitability for a threatened farmland bird species. *Biological Conservation*, **143**, 2770–2778.
- Buse, A. (1973) Goodness of fit in generalized least squares estimation. *American Statistician*, **27**, 106–108.

- Casale, F. & Brambilla, M. (2009) *Averla Piccola. Ecologia e Conservazione*. Fondazione Lombardia per l'Ambiente & Regione Lombardia, Milano, Italy.
- Cramp, S. (1993) *The Birds of the Western Palearctic, Volume VII*. Oxford University Press, Oxford.
- Dormann, C.F., McPherson, J.M., Araújo, M.B., Bivand, R., Bolliger, J., Carl, G., Davies, R.G., Hirzel, A., Jetz, W., Kissling, W.D., Kühn, I., Ohlemüller, R., Peres-Neto, P.R., Reineking, B., Schröder, B., Schurr, F.M. & Wilson, R. (2007) Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, **30**, 609–628.
- Elith, J. & Leathwick, J.R. (2009) Species distribution models: ecological explanation and prediction across space and time. *Annual Review of Ecology and Systematics*, **40**, 677–697.
- Elith, J., Graham, C.H., Anderson, R.P., Dudík, M., Ferrier, S., Guisan, A., Hijmans, R.J., Huettmann, F., Leathwick, J.R., Lehmann, A., Li, J., Lohmann, L.G., Loiselle, B.A., Manion, G., Moritz, C., Nakamura, M., Nakazawa, Y., Overton, J.M., Peterson, A.T., Phillips, S.J., Richardson, K.S., Scachetti-Pereira, R., Schapire, R.E., Soberon, J., Williams, S., Wisz, M.S. & Zimmermann, N.E. (2006) Novel methods improve prediction of species' distributions from occurrence data. *Ecography*, **29**, 129–151.
- Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E. & Yates, C.J. (2011) A statistical explanation of MaxEnt for ecologists. *Diversity and Distributions*, **17**, 43–57.
- Ferrier, S., Watson, G., Pearce, J. & Drielsma, M. (2002) Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. I. Species-level modelling. *Biodiversity and Conservation*, **11**, 2275–2307.
- Ficetola, G.F., Thuiller, W. & Miaud, C. (2007) Prediction and validation of the potential global distribution of a problematic alien invasive species – the American bullfrog. *Diversity and Distributions*, **13**, 476–485.
- Ficetola, G.F., Thuiller, W. & Padoa-Schioppa, E. (2009) From introduction to the establishment of alien species: bioclimatic differences between presence and reproduction localities in the slider turtle. *Diversity and Distributions*, **15**, 108–116.
- Ficetola, G.F., Maiorano, L., Falcucci, A., Dendoncker, N., Boitani, L., Padoa-Schioppa, E., Miaud, C. & Thuiller, W. (2010) Knowing the past to predict the future: land-use change and the distribution of invasive bullfrogs. *Global Change Biology*, **16**, 528–537.
- Förschler, M.I., Borrás, A., Cabrera, J., Cabrera, T. & Senar, C.J. (2005) Inter-locality variation in reproductive success of the citril finch *Serinus citrinella*. *Journal of Ornithology*, **146**, 137–140.
- Fouquet, A., Ficetola, G.F., Haigh, A. & Gemmill, N. (2010) Using ecological niche modelling to infer past, present and future environmental suitability for *Leiopelma hochstetteri*, an endangered New Zealand native frog. *Biological Conservation*, **143**, 1375–1384.
- Golawski, A. (2006) Impact of weather on partial loss of nestlings in the red-backed shrike *Lanius collurio* in eastern Poland. *Acta Ornithologica*, **41**, 15–20.
- Golawski, A. & Meissner, W. (2008) The influence of territory characteristics and food supply on the breeding performance of the red-backed shrike (*Lanius collurio*) in an extensively farmed region of eastern Poland. *Ecological Research*, **23**, 347–353.
- Graham, C.H., Ferrier, S., Huettmann, F., Moritz, C. & Peterson, A.T. (2004) New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution*, **19**, 497–503.
- Griffis, V.W. & Stedinger, J.R. (2007) The use of GLS regression in regional hydrologic analyses. *Journal of Hydrology*, **344**, 82–95.
- Hall, L.S., Krausman, P.R. & Morrison, M.L. (1997) The habitat concept and a plea for standard terminology. *Wildlife Society Bulletin*, **25**, 173–182.
- Jiménez-Valverde, A., Lobo, J.M. & Hortal, J. (2008) Not as good as they seem: the importance of concepts in species distribution modelling. *Diversity and Distributions*, **14**, 885–890.
- Kearney, M. & Porter, W. (2009) Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, **12**, 334–350.
- Kissling, W.D. & Carl, G. (2008) Spatial autocorrelation and the selection of simultaneous autoregressive models. *Global Ecology and Biogeography*, **17**, 59–71.
- Mendenhall, C.D., Sekercioglu, C.H., Brenes, F.O., Ehrlich, P.R. & Daily, G.C. (2011) Predictive model for sustaining biodiversity in tropical countryside. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 16313–16316.
- Newton, I. (1998) *Population Limitation in Birds*. Academic Press, San Diego, CA.
- Ortego, J. (2007) Consequences of Eagle Owl nest-site habitat preference for breeding performance and territory stability. *Ornis Fennica*, **84**, 78–90.
- Pasinelli, G. (2000) Oaks (*Quercus* sp.) and only oaks? Relations between habitat structure and home range size of the middle spotted woodpecker (*Dendrocopos medius*). *Biological Conservation*, **93**, 227–235.
- Pearson, R.G., Raxworthy, C.J., Nakamura, M. & Peterson, A.T. (2007) Predicting species distributions from small numbers of occurrence records: a test case using cryptic geckos in Madagascar. *Journal of Biogeography*, **34**, 102–117.
- Peterson, A.T., Soberón, J. & Sánchez-Cordero, V. (1999) Conservatism of ecological niches in evolutionary time. *Science*, **285**, 1265–1267.
- Phillips, S.J., Anderson, R.P. & Schapire, R.E. (2006) Maximum entropy modeling of species geographic distributions. *Ecological Modelling*, **190**, 231–259.
- Pinheiro, P., Bates, D., DebRoy, S. & Sarkar, D. (2010) *Linear and Nonlinear Mixed Effects Models*. R package version 3.1-97, <http://cran.r-project.org/web/packages/nlme>.
- R Development Core Team (RDC). (2010) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna.
- Raxworthy, C.J., Martinez-Meyer, E., Horning, N., Nussbaum, R.A., Schneider, G.E., Ortega-Huerta, M.A. & Peterson, A.T. (2003) Predicting distributions of known and unknown reptile species in Madagascar. *Nature*, **426**, 837–841.
- Reshetnikov, A.N. & Ficetola, G.F. (2011) Potential range of the invasive fish rotan (*Perccottus glenii*) in the Holarctic. *Biological Invasions*, **13**, 2967–2980.
- Rödger, D., Schmidlein, S., Veith, M. & Lötters, S. (2009) Alien invasive slider turtle in unpredicted habitat: a matter of niche shift or of predictors studied? *PLoS ONE*, **4**, e7843.
- Roos, S. & Pärt, T. (2004) Nest predators affect spatial dynamics of breeding red-backed shrikes (*Lanius collurio*). *Journal of Animal Ecology*, **73**, 117–127.
- Roura-Pascual, N., Suarez, A.V., Goomez, C., Pons, P., Touyama, Y., Wild, A.L. & Peterson, A.T. (2004) Geographical potential of Argentine ants (*Linepithema humile* Mayr) in the face of global climate change. *Proceedings of the Royal Society B-Biological Sciences*, **271**, 2527–2534.
- Said, S., Gaillard, J.-M., Widmer, O., Debiat, F., Bourgoin, G., Delorme, D. & Roux, C. (2009) What shapes intra-specific variation in home range size? A case study of female roe deer *Oikos*, **118**, 1299–1306.
- Sergio, F., Marchesi, L. & Pedrini, P. (2004) Integrating individual habitat choices and regional distribution of a biodiversity indicator and top predator. *Journal of Biogeography*, **31**, 619–628.
- Sergio, F., Pedrini, P. & Marchesi, L. (2003) Spatio-temporal shifts in gradients of habitat quality for an opportunistic avian predator. *Ecography*, **26**, 243–255.
- Smith, T.M. & Shugart, H.H. (1987) Territory size variation in the ovenbird: the role of habitat structure. *Ecology*, **68**, 695–704.
- Soberon, J. & Nakamura, M. (2009) Niches and distributional areas: concepts, methods, and assumptions. *Proceedings of the National Academy of Sciences of the United States of America*, **106**, 19644–19650.
- Thuiller, W., Lavorel, S., Araújo, M.B., Sykes, M.T. & Prentice, I.C. (2005a) Climate change threats to plant diversity in Europe. *Proceedings of the National Academy of Sciences of the United States of America*, **102**, 8245–8250.
- Thuiller, W., Richardson, D.M., Pyšek, P., Midgley, G.F., Hughes, G.O. & Rouget, M. (2005b) Niche-based modelling as a tool for predicting the risk of alien plant invasions at a global scale. *Global Change Biology*, **11**, 2234–2250.
- Titeux, N., Dufrene, M., Radoux, J., Hirzel, A.H. & Defourny, P. (2007) Fitness-related parameters improve presence-only distribution modelling for conservation practice: the case of the red-backed shrike. *Biological Conservation*, **138**, 207–223.
- Tryjanowski, P., Kuzniak, S. & Diehl, B. (2000) Breeding success of the red-backed shrike (*Lanius collurio*) in relation to nest site. *Ornis Fennica*, **77**, 137–141.
- VanDerWal, J., Shoo, L.P., Johnson, C.N. & Williams, S.E. (2009) Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. *The American Naturalist*, **174**, 282–291.
- Venables, W.N. & Ripley, B.D. (2002) *Modern Applied Statistics with S*. Springer, New York.

Received: 9 September 2011; accepted 26 January 2012  
Handling Editor: Simon Griffith

## Supporting Information

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

**Table S1.** Full list of the habitat variables considered for model building under MaxEnt.

As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials 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.