

Jakub Z. Kosicki · Krystyna Stachura
Marta Ostrowska · Eliza Rybska

Complex species distribution models of Goldcrests and Firecrests densities in Poland: are remote sensing-based predictors sufficient?

Received: 20 October 2014 / Accepted: 30 March 2015 / Published online: 17 April 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract Species distribution models should identify ecological requirements of species and predict their spatial density. However, data from remote sensing sources are often used alone as predictors in modelling distributions. Such data will only produce accurate models if features that are distinguishable by remote sensing are a good match to the environmental traits that delineate habitat requirements. Both the Goldcrest *Regulus regulus* and the Firecrest *Regulus ignicapilla* respond to complex features of habitats that are not described by simple remote sensing data. We tested the usefulness of remote sensing data as a predictor for two *Regulus* species according to data from 970 study plots sized 1 × 1 km. Predictors were aggregated using the PCAs and related to the Hayne estimator of species density using GAMs. The models based on both remote sensing data and detailed environmental data proved to be better than the model based only on remote sensing data and/or detailed forest structure data. The Goldcrest

reached the highest density in areas with a high share of old spruce-dominated forests with a substantial share of the fir, avoiding the pine, and it preferred forests with a low number of tree species. In turn, the Firecrest favoured old forests, dominated by the spruce and the beech, with an admixture of single old fir and larch trees, avoiding the pine, and preferring forests with a high number of tree species. We suggest using not only free data sources, but also more detailed data containing thorough information on forest inventory derived from ground measurements.

Keywords Firecrest · Forest stand structure · Goldcrest · Habitat selection · Species' distribution models

Electronic supplementary material The online version of this article (doi:10.1007/s11284-015-1263-5) contains supplementary material, which is available to authorized users.

J. Z. Kosicki
Laboratory and Museum of Evolutionary Ecology, Department of Ecology, Faculty of Humanities and Natural Sciences, University of Prešov, 17. novembra 1 081 16, Prešov, Slovakia

J. Z. Kosicki (✉) · M. Ostrowska
Department of Avian Biology and Ecology, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland
E-mail: kubako@amu.edu.pl
Tel.: +48618295615

K. Stachura
Laboratory of Biological Spatial Information, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland

E. Rybska
Department of Nature Education and Conservation, Faculty of Biology, Adam Mickiewicz University, Umultowska 89, 61-614 Poznań, Poland

Introduction

The alteration of forest landscapes, driven by agriculture and the needs of economy, has been identified as a primary factor affecting populations of forest-dwelling species at regional and local scales (Saunders et al. 1991; Fahrig 2002; Ewers and Didham 2006; Mazgajski et al. 2010; Bennett et al. 2014). Certain key forest birds species e.g. Chaffinch *Fringilla coelebs* and/or Tree pipit *Anthus trivialis*, (Kuczyński and Chylarecki 2012) have already suffered a long-term population decline as a result of habitat loss, fragmentation and increased isolation of woodland patches (Fahrig 2003; Mazgajski et al. 2010; Kajtoch et al. 2012; Żmihorski 2012). Further declines of other bird species are very probable, because increasing timber production is connected with the intensification of forest management (Sallabanks et al. 2000; Perry and Thill 2013). However, not all forest bird species are equally vulnerable to forest management practices and observed patterns of population changes differ across space, hampering our understanding of exact mechanisms driving the observed changes on the trans-national scale (Reif et al. 2008; Perry and Thill 2013). Therefore, studying relationships between

population density and environmental factors on large geographical scales is one of key elements enhancing conservation strategies in the European Union as a whole (Sanderson et al. 2009). The most adequate data are those obtained directly from field measurements. However, covering large areas solely by fieldwork is basically unrealistic (Spanhove et al. 2012). That is why conservationists pay a lot of attention to developing methods for collecting and analyzing data from a monitoring scheme within a large area so they can monitor population health of many bird species (Stachura-Skierczyńska et al. 2009; Spanhove et al. 2012; Brambilla et al. 2013; Buchanan et al. 2013; Linden and Roloff 2013; Walczak et al. 2013).

In the last few years many studies have focused on species distribution modelling (e.g. Franklin 2009; Brambilla et al. 2013; Guisan et al. 2013; Kosicki and Chylarecki 2013, 2014; Stachura-Skierczyńska et al. 2009) based on national and regional field monitoring programmes, e.g. the Pan-European Common Bird Monitoring Scheme (Gregory et al. 2007; Kosicki and Chylarecki 2013, 2014). Many of these studies use publicly available generalized Geographic Information Systems (GIS) datasets (i.e. Corine land cover, NDVI dataset, WordClim) as predictors (e.g., Giordano et al. 2010; Kosicki and Chylarecki 2013; Morelli and Tryjanowski 2014). However, such data are only suitable for some opportunistic species that have habitat requirements with sharp boundaries in the landscape, e.g. the Coal Tit *Periparus ater* (Da Silva et al. 2012), Common Chiffchaff *Phylloscopus collybita* (Lampila et al. 2009) and Wood Nuthatch *Sitta europea* (Kuczyński and Chylarecki 2012). In case of more specialized species, other environmental characteristics, i.e. small-scale structural elements or habitat variety might influence their habitat selection decisions. However, this kind of detailed information cannot be obtained from generalized remote sensing-based data sources (Akçakaya 2001; Giordano et al. 2010). On the large geographical scale, e.g. the country, gaining data on small-scale structural elements within habitats is only possible in case of closed schemes, e.g. forest inventory database (Rego et al. 2004). This kind of data can be very useful for ecological research and conservation purposes on the national and transnational level (Stachura-Skierczyńska et al. 2009; Skierczyński et al. 2013; Stachura-Skierczyńska and Kosiński 2014) but have not been sufficiently employed in the species distribution model (SDM) of density.

In this study we analyze spatial distribution patterns of the Goldcrest *Regulus regulus* and the Firecrest *Regulus ignicapilla*, two forest breeding bird species sharing partially an ecological niche and recently displaying a decreasing population trend (Chylarecki and Jawińska 2007; Kuczyński and Chylarecki 2012). Both species are widespread across Europe, inhabiting coniferous and mixed forests. The Goldcrest prefers the spruce *Picea abies* and the fir *Abies alba*, while the Firecrest is more opportunistic and additionally chooses habitats with deciduous trees, i.e. the beech *Fagus sylvatica* (del Hoyo

et al. 2006; Kralj et al. 2013). The effect of other tree species—mainly the pine, being one of the most abundant tree species in large parts of Europe—as well as the forest's age and structure on the density and distribution of both *Regulus* species is ambiguous. An earlier study showed that pine forests might have either negative or positive effect on both species (Purroy 1974a, b; Lebreton and Thevenot 2009). The most recent studies, (Kuczyński and Chylarecki 2012; Kralj et al. 2013) did not address this issue, as they did not examine variables related to the forest structure.

In our study, we analysed factors shaping the Goldcrest's and Firecrest's densities, however, apart from environmental variables derived from remote sensing data we also used more detailed information on the forest composition, e.g. tree species diversity, the number and age of trees as well as forest vertical structure.

In this way we were able to capture more detailed relationships between the species' densities and environmental conditions on a large spatial scale (Saveraid et al. 2001; Gibson et al. 2004; McPherson and Jetz 2007; Müller et al. 2009). A recent study also indicated that the occurrence of co-existing species might be an important predictor for SDM, but this kind of information is rarely used (Kissling et al. 2012; Morelli and Tryjanowski 2014). Therefore, we involved densities of both species as additional predictors in SDMs in order to test the usefulness of one species' density as the surrogate of other species' densities on a large geographical scale.

We used data from Poland which is particularly suitable for this kind of analyses, given a large area of forests (9163.8 thousand ha—Anon 2013) and the existing gradient of forest structure across the country (Anon 2013). We capitalize on the existing data from a countrywide survey of common breeding birds, yielding presence/absence data for >900 study plots representative for the country (Kuczyński and Chylarecki 2012). Presence/absence data recorded during planned surveys are clearly preferable to presence-only data in SDM studies due to a multitude of reasons (Franklin 2009; Royle et al. 2012).

Having access to high quality data on bird occurrences, coupled with existing environmental data from other sources, the main goals of this study are to: (1) determine the extent to which generalized remote sensing-based datasets are able to predict the distributions of the two *Regulus* species in comparison to detailed forest inventory data from ground measurements; and (2) compare the Goldcrest's and the Firecrest's densities.

Materials and methods

Bird data

The Goldcrest and Firecrest density data were derived from the common breeding birds monitoring scheme (Chylarecki and Jawińska 2007) and collected in Poland

in years 2000–2013 in 970 grid cells of 1 square km (see Appendix A, Fig. S1). Squares treated as survey plots had been chosen at random out of 311,664 squares covering whole Poland. In a particular breeding season each plot was surveyed twice. The first visit took place between 10 April and 15 May and the second between 16 May and 30 June. Each survey started between the dawn and 9 am and took about 90 min. The route of bird census consisted of two parallel one-by-one-kilometre sections (transects), along an east–west or north–south axis. Each transect was divided into five 200-meter sections, in which birds were noted in three distance categories (< 25 m, 25–100 m, > 100 m). In the analysis we only use squares which were inspected twice in breeding season, however due to the fact that surveys were carried out by volunteers, the majority of plots were not checked every year. During the 14-year span each square was inspected on average 6.5 times ($SD = 3.9$).

As the number of individuals in the given grid cell we used the highest recorded number of individuals during one of the two inspections. Finally, both species are very small and rather shy during breeding season so in the analysis we only considered birds which had been seen or heard in only two distance categories.

Environmental data

For modelling purposes, we used several environmental variables related to topography, climatic conditions, land cover (general), vegetation indices and forest stand structure (Table 1). All datasets were converted into GRASS GIS file format (Neteler and Mitasova 2008) with the grid size of 1 km² (corresponding to bird survey plots) and re-projected to coordinate system EPSG4284 projection (<http://spatialreference.org/ref/epsg/4284/>).

Altitudes above sea level (a.s.l.) were obtained from the digital elevation model dataset (GTOPO30, resolution 1 km²), originally provided by the U.S. Geological Survey's EROS Data Center (Sioux Falls, South Dakota). For modelling purposes we used both absolute altitudes and relative differences between maximal and minimal altitudes per plot.

Climate data were obtained from the WorldClim database (<http://www.worldclim.org>), which is a set of global climate layers (climate grids) with 1 km² spatial resolution. Compared to previously used large-scale climate data, this dataset has the highest spatial resolution (Hijmans et al. 2005) and has been successfully used as a predictor variable for modelling species' spatial distribution (e.g. Horn et al. 2012). We used the following six variables which determined the variability of temperature and precipitation in Poland: Annual Mean Temperature, Mean Temperature of Warmest Quarter, Mean Temperature of Coldest Quarter, Annual Precipitation, Precipitation of Warmest Quarter, and Precipitation of Coldest Quarter.

We used the CORINE Land Cover database Level 3, available at 100 meter resolution and categorized using

44 land cover classes, out of which 31 were distinguished in Poland. The database, updated in 2006, was based on Landsat TM, IRS and SPOT-4 images (EEA 2007). In this study we found on average 11 types of land cover per plot. The most frequent were non-irrigated arable land, meadows, complex cultivation patterns, agricultural areas with natural vegetation, shrub, and inland marshes. We did not use categories depicting urban areas and coniferous and deciduous forests from the original CORINE dataset, since more detailed data were available. In order to estimate the density of human settlements, we used images of lights at night. These datasets come from remote sensing (<http://svs.gsfc.nasa.gov/vis/a030000/a030000/a030028/index.html>) and the lights on the night map include cities, towns, and other sites with persistent lighting. This kind of data is highly correlated with the population level (Small et al. 2005; Doll et al. 2007). General types of the forest cover were obtained from the Joint Research Centre (<http://forest.jrc.ec.europa.eu/>) and based on remote sensing. The dataset of 25 m resolution includes two types of forest: coniferous and deciduous. Higher resolution allows detecting small patches or stripes of woodland, not distinguished in CORINE database.

We also obtained the normalized differences vegetation index (NDVI), derived from the SPOT dataset (<http://free.vgt.vito.be/>). NDVI is an index of green vegetation level, and is expressed as a mean monthly NDVI value (from March to June).

Forest stand structure information comes from the forest inventory database, consisting of stand-level numeric maps and associated attribute data. The database is owned and managed by the National Forest Holding "State Forests". Full access to data is restricted and regulated by specific project-related agreements between State Forests and particular institutions.

Stands are assumed to be homogenous spatial units. For each stand, the database contains a detailed description of the canopy (all tree species, their age and estimated percentage cover). The information was obtained from ground measurements, supplemented by high-resolution aerial photos and it was updated regularly, depending on the local schedule of management activities (usually once a year, but some records might be older). We used the following variables based on the stand description: the number of tree species, percentage cover of particular trees, age of trees and forest diversity in vertical space (number of levels) (see Appendix A, Table S1).

Spatial resolution of this data is lower than 1 km² bird survey plot. All variables are expressed as a weighted mean, where the weight is the area of spatial units.

Data processing and analysis

All statistical analyses were performed using the statistical package R (R Development Core Team 2010). The mean densities (for all years of study) of Goldcrests and

Table 1 Environmental predictors used for the model design

Variable code	Description
Predictors extracted from PCA (see Table S2)	
PREC	had the highest loadings with precipitation variables
TEMP	had the highest loadings with temperature variables
CFORFARML	represents the habitat gradient from a large area of coniferous forest to arable fields
CODEFOR	represents the habitat gradient from small area of coniferous forest to large area of deciduous forest
FARMLMFOR	represents the habitat gradient from arable fields near urban areas to small and large areas of mixed forest
DMFOR	represents the habitat gradient from a small area of deciduous forest to a large area of the forest
MARSMFOR	represents the habitat gradient from inland marshland to mixed forest
WATURB	represents the habitat gradient from water bodies to urban areas
Predictors extracted directly from remote sensing data	
LONGITUDE	geographic coordinate that specifies the east–west position
LATITUDE	geographic coordinate that specifies the south–north position
ALTITUDE	above sea level (a.s.l.)
DENIW	differences between maximal and minimal altitudes per grid cells
NDVI.MAR	mean value of green vegetation level in March
NDVI.APR	mean value of green vegetation level in April
NDVI.MAY	mean value of green vegetation level in May
NDVI.JUN	mean value of green vegetation level in Jun
Predictors extracted directly from forest inventory database. The prefix “DOM.” indicates the species’ share in grid cells, while “AGR.” indicates the mean age of tree species in grid cells	
DOM.BLAC	Black locust <i>Robinia pseudoacacia</i>
AGE.BLAC	
DOM.BEECH	Common beech <i>Fagus sylvatica</i>
AGE. BEECH	
DOM.BIRCH	Birch <i>Betula sp.</i>
AGE. BIRCH	
DOM.OAK	Oak <i>Quercus sp.</i>
AGE. OAK	
DOM.ROAK	Red oak <i>Quercus rubra</i>
AGE. ROAK	
DOM.HORN	Common hornbeam <i>Carpinus betulus</i>
AGE. HORN	
DOM.FIR	European fir <i>Abies alba</i>
AGE. FIR	
DOM.ASH	European ash <i>Fraxinus excelsior</i>
AGE. ASH	
DOM.SMAP	Sycamore maple <i>Acer pseudoplatanus</i>
AGE. SMAP	
DOM.NMAP	Norway maple <i>Acer platanoides</i>
AGE. NMAP	
DOM.LIME	Lime <i>Tilia sp.</i>
AGE. LIME	
DOM.LARCH	Larch <i>Larix decidua</i>
AGE. LARCH	
DOM.ALDE	Alder <i>Alnus sp.</i>
AGE. ALDE	
DOM.ASPE	Common aspen <i>Populus tremula</i>
AGE.ASPE	
DOM.PINE	Pine <i>Pinus sp.</i>
AGE. PINE	
DOM.ELM	Elm <i>Ulmus minor</i>
AGE. ELM	
DOM.SPRUCE	Spruce <i>Picea sp.</i>
AGE.SPRUCE	
NO.LAYER	Numbers of layers
NO.SPEC	Numbers of tree species
M.AGE.TREES	Mean age of forest

Firecrests expressed as the number of birds in each grid cell (maximum number of birds recorded during visits) were respectively 0.60 individuals/km² (95 % CL = 0.47–0.73, *n* = 970) and 0.16 individuals/km² (95 %

CL = 0.11–0.20, *n* = 970). Due to the fact that bird densities in plots depended on the transect’s length and the distance from the observer (Krebs 1999), the Hayne estimator of species density was calculated for each cell

according to the equation below (Hayne 1949; Krebs 1999):

$$\hat{D}_H = \frac{n}{2L} \left(\frac{1}{n} \sum \frac{1}{r_i} \right)$$

where, \hat{D}_H is Hayne's estimator of density, n is the number of animals seen, L is the length of transect, r_i is sighting distance to each animal i according to two distance classes, i.e.: 25 and 50 m.

In order to avoid multicollinearity among environmental variables, principal components' analysis (PCA) was performed with Varimax normalized rotation, separately for each of the two environmental datasets, i.e. climate and habitat (Quinn and Keough 2002). Principal components' axes with eigenvalues >1 were retained as predictor variables in the analyses.

PCA of climate variables produced two axes, which explained 83.7 % of the original variation in climate variables (see: Table 1 and Appendix A, Table S2A).

Habitat variables, which were derived from the corine land cover (CLC), HUMAN and the type of forest coming from the Joint Research Centre, produced seven components and explained 88.1 % of the variation (see: Table 1 and Appendix A, Table S2B). PC7 has a low value of the explained variance and is not included in the analysis.

Pearson correlation coefficients and the variance inflation factor (VIF, using HH library in R; Heiberger 2013) were used to assess relationships between predictors which came from two separate PCA analysis and NDVI variable, geographical coordination and topography (Appendix A, Table S3). VIF of 5 and above indicated a multicollinearity problem (O'Brien 2007). In the present study, VIF ranged from 1.6 to 12.9, therefore predictors whose VIF ≥ 5 were excluded from the analysis (see Appendix A, Table S3).

We employed the generalized additive model (GAM) to fit resource selection functions (Hastie and Tibshirani 1990). The Hayne estimator of species' density was used as the response variable. Forty-four variables extracted by PCAs, geographical variables (longitude, latitude, altitude and denivelation), NDVI from March to April forest stand structure information were used as predictors (Table 1). We developed three sets of GAM: (1) models based on coarse data obtained from remotely sensed images (topographic, climatic, general land cover, vegetation indices); (2) models based on detailed forest inventory data derived from ground measurements; and (3) models based on both datasets, i.e. remotely sensed images and forest structure provided by the forest inventory database. The most parsimonious model was selected using the Akaike information criterion (mgcv library in R; Wood 2013) with the lowest AIC and consequently the highest Akaike weight (Burnham and Anderson 2002). We analysed all possible models (2^n , where n = number of variables), using MuMIn library in R (Hastie and Tibshirani 1990; Bartoń 2013). Although this approach is criticized, the analysis of all

possible models is often used when there is not enough a priori information to develop a small set of models (Whittingham et al. 2005; Reino et al. 2010). The probability of including a variable in the best parsimonious model was estimated as the relative importance (RI) by summing the Akaike weights of all candidate models in which the variable was included (Burnham and Anderson 2002; Reino et al. 2010). As a measure for the best model we used evidence ratio (Burnham and Anderson 2002). In order to allow some complexity in the functions, while avoiding over-fitting the data, we defined the basis dimension, i.e. $k = 4$ (Santana et al. 2012). The Gaussian distribution of errors and the identity link function were applied. As a measure of deviance reduction, we used the D^2 coefficient, which was equivalent to R^2 , well known from least squares estimation (Weisberg 1980). Then, from the best model for each species and each of the three sets of predictors' variables we created and presented predictions of distributions. The correlation coefficient between the predicted versus the observed densities (log-transformed) was used as a measure of error prediction (Hastie et al. 2008; Kuczyński et al. 2009).

Results

Population size

Breeding populations of Goldcrests and Firecrests were recorded respectively in 36.8 and 15.4 % of grid cells. The mean density of the Goldcrest expressed as the Hayne estimator was 1.51 (95 % CL: 1.31–1.71) individuals/1 km², while the mean Hayne estimator on occupied plots was 4.12 (95 % CL: 3.7–4.5) individuals/1 km². In turn, the mean density of the Firecrest also expressed as the Hayne estimator was 0.45 (95 % CL: 0.34–0.52) individuals/1 km², while the mean Hayne estimator on occupied plots was 2.9 (95 % CL: 3.7–4.5) individuals/1 km². We found a correlation between the raw density of the Goldcrest (expressed as the number of birds in each grid cell) and the raw density of the Firecrest ($r = 0.29$, $P < 0.001$, $n = 970$), and then between the Hayne estimator of Goldcrests' and Firecrests' densities ($r = 0.58$, $P = 0.001$, $n = 970$). These correlations are not simply indications that both species are found only in the forested squares, as there was also a significant correlation even when considering data only from those squares where at least one of the species was present (raw density $r = 0.26$, $P < 0.001$ and the Hayne estimator $r = 0.61$, $P < 0.001$).

Habitat use by Goldcrest and predictive map of \hat{D}_{H-G}

The most parsimonious model of GAM (Table 2, model A2) based only on the remotely sensed coarse data and the fit to the Hayne estimator of the Goldcrest's density (\hat{D}_{H-G}) included a smooth fit to LATITUDE ($RI = 1$,

Table 2 The comparison of GAM models based on the standardized Goldcrest \hat{D}_{H-G} and Firecrest \hat{D}_{H-F} density (null and all where Akaike weights ≥ 0.001) differing in numbers of predictors, linearized or included predictors as a polynomial spline

Model	Predictor	logLink	AIC	Δ AIC	Akaike weight	D ²	R
\hat{D}_{H-G} based on the remotely sensed coarse data only							
A1	Null	-963.13	1928.31	822.08	0	0.74	0.01 n.s
A2	s(LATITUDE) + s(ALTITUDE) + s(TEMP) + s(CFORFARM) + s(CODEFOR) + s(NDVI.MAR)	-522.63	1106.23	0.00	0.67	37.21	0.38**
A3	s(LATITUDE) + s(ALTITUDE) + s(CFORFARM) + s(CODEFOR) + s(NDVI.MAR)	-523.94	1107.69	1.46	0.32	29.17	0.22**
\hat{D}_{H-G} based on the detailed data on the forest structure only							
B1	Null	-963.13	1928.31	899.34	0	0.74	0.01 n.s
B2	s(DOM.FIR) + s(AGE.FIR) + s(DOM.PINE) + s(NO.SPEC) + L(DOM.SPRUCE) + L(AGE.SPRUCE)	-475.83	1028.97	0.00	0.356	41.76	0.44**
B3	s(DOM.FIR) + s(AGE.FIR) + s(DOM.PINE) + s(NO.SPEC) + L(DOM.SPRUCE)	-475.86	1028.99	0.02	0.353	41.13	0.39**
B4	s(DOM.FIR) + s(DOM.PINE) + s(NO.SPEC) + L(DOM.SPRUCE) + L(AGE.SPRUCE)	-476.19	1029.37	0.40	0.291	39.79	0.37**
\hat{D}_{H-G} with remotely sensed coarse, forest structure and \hat{D}_{H-F} as a predictor							
C1	Null	-963.13	1928.28	915.94	0	0.7	0.01 n.s
C2	L(LATITUDE) + L(ALTITUDE) + L(TEMP) + s(PREC) + s(CFORFARM) + L(CODEFOR) + s(NDVI.MAR) + L(DOM.SPRUCE) + L(DOM.FIR) + L(AGE.SPRUCE) + s(AGE.FIR) + L(DOM.PINE) + s(NO.SPEC)	-459.45	1012.34	0	0.97	88.63	0.79**
C3	s(LATITUDE) + L(TEMP) + s(CFORFARM) + s(CODEFOR) + L(DOM.SPRUCE) + L(DOM.FIR) + L(AGE.SPRUCE) + L(DOM.PINE) + L(NO.SPEC)	-461.95	1019.59	7.25	0.02	23.74	0.21**
C4	s(CFORFARM) + s(CODEFOR) + L(DOM.SPRUCE) + L(DOM.FIR) + L(AGE.SPRUCE) + s(AGE.FIR) + s(DOM.PINE) + s(NO.SPEC)	-468.92	1026.38	14.04	0.00	11.68	0.11*
\hat{D}_{H-F} with only remotely sensed coarse data as a predictor							
D1	Null	-390.89	783.83	410.84	0	0.90	0.021 n.s
D2	s(LATITUDE) + s(ALTITUDE) + s(PREC) + s(TEMP) + s(CODEFOR) + s(FARMLMFOR) + L(DMFOR)	-149.31	372.99	0.00	0.54	41.25	0.46**
D3	s(LATITUDE) + s(ALTITUDE) + s(TEMP) + L(PREC) + L(CODEFOR) + L(FARMLMFOR) + L(DMFOR) + (NDVI.JUN)	-149.74	373.76	0.77	0.37	38.42	0.44**
D4	s(ALTITUDE) + s(TEMP) + L(PREC) + L(CODEFOR) + L(FARMLMFOR) + L(DMFOR) + s(NDVI.JUN)	-152.95	376.99	4.00	0.07	11.75	0.11*
D5	s(LATITUDE) + s(LONGITUDE) + s(TEMP) + L(PREC) + L(CODEFOR) + L(FARMLMFOR) + + s(NDVI.JUN)	-156.00	382.14	9.15	0.00	5.11	0.06 n.s
\hat{D}_{H-F} based on the detailed data on the forest structure only							
E1	Null	-390.89	783.76	445.74	0	0.90	0.021 n.s
E2	s(DOM.BEECH) + s(DOM.SPRUCE) + s(DOM.FIR) + s(AGE.FIR) + s(AGE.LARCH) + s(DOM.PINE) + s(NO.SPEC)	-125.76	338.02	0.00	0.96	44.51	0.48**
E3	s(DOM.BEECH) + s(DOM.SPRUCE) + s(DOM.FIR) + s(AGE.FIR) + s(AGE.LARCH) + s(DOM.PINE) + s(NO.SPEC)	-130.87	346.93	8.91	0.01	10.42	0.10*
E4	s(DOM.BEECH) + s(DOM.FIR) + s(AGE.FIR) + s(AGE.LARCH) + s(DOM.PINE) + s(NO.SPEC)	-132.22	347.11	9.09	0.01	10.04	0.90 n.s
E5	s(DOM.BEECH) + s(AGE.FIR) + s(AGE.LARCH) + s(NO.SPEC)	-132.24	347.43	9.41	0.00	4.73	0.02 n.s
E6	s(DOM.BEECH) + s(DOM.FIR) + s(AGE.FIR) + s(AGE.LARCH) + s(DOM.PINE)	-132.54	348.57	10.55	0.00	2.15	0.02 n.s

Table 2 continued

Model	Predictor	logLink	AIC	Δ AIC	Akaike weight	D ²	R
\hat{D}_{H-G} with remote sensing data, forest structure and \hat{D}_{H-G} as a predictors							
F1	Null	-390.89	783.83	450.4	0	0.90	0.021 n.s
F2	s(L.ALTITUDE) + s(ALTITUDE) + s(TEMP) + s(PREC) + L(CODEFOR) + s(FARMLMFOR) + s(DOM.BEECH) + s(DOM.SPRUCE) + L(DOM.PINE) + L(AGE.SPRUCE) + L(AGE.FIR) + L(AGE.LARCH) + s(D_{H-G}) + s(NO.SPEC)	-122.87	333.43	0.00	0.34	78.44	0.87**
F3	s(L.ALTITUDE) + s(ALTITUDE) + s(TEMP) + s(PREC) + s(CODEFOR) + s(FARMLMFOR) + s(DOM.BEECH) + s(DOM.SPRUCE) + L(DOM.PINE) + L(AGE.SPRUCE) + L(AGE.FIR) + L(AGE.LARCH) + s(D_{H-G}) + L(NO.SPEC)	-123.86	333.55	0.11	0.34	77.96	0.85**
F4	s(TEMP) + s(PREC) + s(CODEFOR) + s(FARMLMFOR) + s(DOM.BEECH) + s(DOM.SPRUCE) + L(DOM.PINE) + L(AGE.SPRUCE) + L(AGE.FIR) + L(AGE.LARCH) + L(D_{H-G}) + s(NO.SPEC)	-124.15	334.38	0.95	0.21	74.24	0.77**
F5	L(CODEFOR) + s(FARMLMFOR) + s(DOM.BEECH) + s(DOM.SPRUCE) + L(DOM.PINE) + L(AGE.SPRUCE) + L(AGE.FIR) + L(AGE.LARCH) + L(D_{H-G}) + s(NO.SPEC)	-125.78	338.08	4.65	0.03	18.22	0.12*
F6	s(CODEFOR) + s(FARMLMFOR) + s(DOM.BEECH) + s(DOM.SPRUCE) + s(DOM.PINE) + L(AGE.SPRUCE) + s(AGE.FIR) + L(AGE.LARCH) + L(D_{H-G}) + s(NO.SPEC)	-125.76	338.09	4.66	0.03	17.65	0.11*
F7	L(CODEFOR) + L(FARMLMFOR) + s(DOM.BEECH) + L(DOM.SPRUCE) + L(DOM.PINE) + L(AGE.SPRUCE) + L(AGE.FIR) + s(AGE.LARCH) + L(D_{H-G}) + s(NO.SPEC)	-125.61	338.14	4.71	0.03	17.62	0.10*

s reflect spline, while L linear. D² indicated measure of deviance reduction while R shows correlation coefficient between the predicted versus the observed densities. The most parsimonious model given in bold
* $P < 0.05$; ** $P < 0.01$; n.s now significant

$F = 18.57$, $P = 0.0001$), ALTITUDE ($RI = 1$, $F = 10.64$, $P = 0.0001$), CFORFARML ($RI = 1$, $F = 6.82$, $P = 0.0001$), CODEFOR ($RI = 1$, $F = 2.17$, $P = 0.0008$), NDVI.MAR ($RI = 1$, $F = 2.68$, $P = 0.0003$) and TEMP ($RI = 0.679$, $F = 2.08$, $P = 0.003$). This model was definitely better for describing the variation of \hat{D}_{H-G} than the second model from our candidate set (evidence ratio = 2.08, Table 2, model A3).

In turn the most parsimonious model of GAM (Table 2, model B2) based only on the detailed data on the forest structure and the fit to the Hayne estimator of the Goldcrest's density (\hat{D}_{H-G}) included a smooth fit to DOM.FIR ($RI = 1$, $F = 3.89$, $P = 0.002$), AGE.FIR ($RI = 0.709$, $F = 3.91$, $P < 0.001$), DOM.PINE ($RI = 1$, $F = 1.35$, $P = 0.003$) and NO.SPEC ($RI = 1$, $F = 4.67$, $P < 0.0001$) and linear fit to DOM.SPRUCE ($RI = 1$, $F = 2.01$, $P = 0.001$) and AGE.SPRUCE ($RI = 0.647$, $F = 3.11$, $P = 0.001$). This model was slightly better for describing the variation of \hat{D}_{H-G} than the second model (evidence ratio = 1.008, Table 2, model B3).

However, the most parsimonious model, which apart from the coarse data used the forest structure and the Firecrest's density (\hat{D}_{H-F}) as additional predictors, fitted to \hat{D}_{H-G} (Table 2, model C2) and revealed a smooth fit to CFORFARML, DOM.FIR, AGE.FIR, NO.SPEC, NDVI.MAR and PREC and a linear fit to ALTITUDE, LATITUDE, TEMP, DOM.SPRUCE, CODEFOR, AGE.SPRUCE and DOM.PINE. This model was definitely better than the second model (evidence ratio = 37.3, Table 2, model C3).

In this case, the most important predictors based on RI included two variables capturing the main habitat gradients, i.e. CFORFARML ($RI = 1$, $F = 13.07$, $P = 0.0001$; Fig. 1a) and CODEFOR ($RI = 1$, $F = 12.38$, $P = 0.0001$; Fig. 1b). The first of them was characterized by a non-linear shape, while the second by a linear shape of the response function and represented the habitat gradient from large (CFORFARML) and small (CODEFOR) coniferous forests to (respectively) arable fields and deciduous forests. The highest values of \hat{D}_{H-G} were found in plots where coniferous forests dominated. The next important variables described the forest stand structure. The highest \hat{D}_{H-G} was noted in plots where the spruce (DOM.SPRUCE, $RI = 1$, $F = 15.64$, $P = 0.0001$; Fig. 1c) and the fir (DOM.FIR, $RI = 1$, $F = 12.05$, $P = 0.0007$; Fig. 1d) dominated, but in both cases old trees were clearly preferred, as shown by corresponding variables AGE.SPRUCE ($RI = 1$, $F = 1.24$, $P = 0.003$; Fig. 1e); and AGE.FIR ($RI = 1$, $F = 4.84$, $P = 0.0007$; Fig. 1f). The species also avoided the pine (DOM.PINE, $RI = 1$, $F = 4.23$, $P = 0.01$; Fig. 1g) and preferred forests with a low number of tree species (NO.SPEC, $RI = 1$, $F = 1.97$, $P = 0.01$; Fig. 1h). Then, NDVI.MAR ($RI = 0.972$, $F = 2.10$, $P = 0.0001$; Fig. 1i) showed that the species' density increased in areas where the vegetation season began early. The next variables were TEMP ($RI = 0.998$, $F = 1.12$, $P = 0.003$; Fig. 1j) and

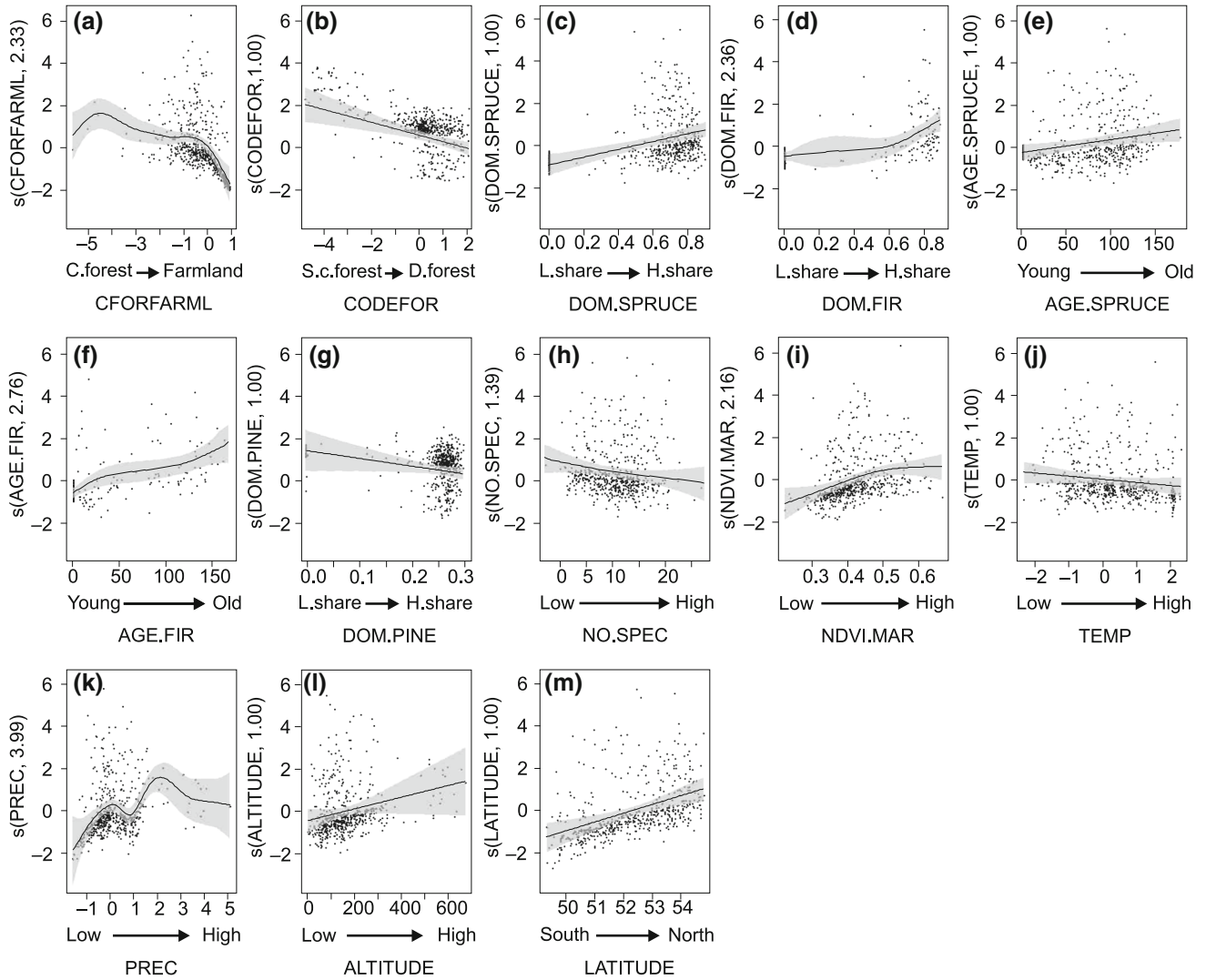


Fig. 1 GAM fit for the Goldcrest density. **a** CFORFARML represents the gradient from a large area of coniferous forest (C.forest) to arable fields (Farmland); **b** CODEFOR represented a habitat gradient from small area of coniferous forest (S.c.forest) to large area of deciduous forest (D.forest); **c** DOM.SPRUCE, **d** DOM.FIR, **e** DOM.PINE reflect forest where share of (respectively) spruce fir and pine are low (L.share) and where these species dominated (H.share); **f** AGE.SPRUCE, **g** AGE.FIR reflect forest where occur young spruce and fir (Young), while (Old)—shows forest where dominated old these species; **h** NO.SPEC represented the number of tree species (L.number)—areas with low number of tree species, while (H.number)—areas with high number of species; **i** NDVI.MAR reflect area where *green* vegetation in

March is low (*Low*) to area where vegetation in march is high (*High*); **j** TEMP represented the temperature gradient, (*Low*)—areas with low temperature, while (*High*)—areas with high temperature; **k** PREC represented the level of precipitation (*Low*)—areas with low level of precipitation, while (*High*)—areas with high level of precipitation; **l** ALTITUDE represented a geographical gradient from area of low elevation (*Low*) to area of high elevation (*High*); **m** LATITUDE represented geographical gradient from south (*South*) to north (*North*). The *y*-axis: density—(*s*) smoother function with estimate degrees of freedom in parenthesis (Hastie and Tibshirani 1990; Brown 2011). The *shaded areas* represent standard errors of the estimate curves

PREC ($RI = 0.972$, $F = 12.09$, $P = 0.0001$; Fig. 1k) that indicated that the Goldcrest favoured relatively low temperatures and high precipitation. Equally important variables were ALTITUDE ($RI = 0.998$, $F = 2.84$, $P = 0.0001$; Fig. 1l), and LATITUDE ($RI = 0.972$, $F = 1.52$, $P = 0.0001$; Fig. 1m), revealing also a non-linear shape of the response. The highest Goldcrest's density corresponded to higher elevations especially in the north of Poland. With the output of this model as the base, we created a predictive map of \hat{D}_{H-G} (Fig. 2a).

Habitat use by Firecrest and predictive map of \hat{D}_{H-F}

Based on the remotely sensed coarse data only (the same dataset as in \hat{D}_{H-G}) and the fit to the Hayne estimator of the Firecrest's density (\hat{D}_{H-F}), the model of GAM (Table 2, model D2) included a smooth fit to ALTITUDE ($RI = 0.994$, $F = 3.54$, $P = 0.0001$), and LATITUDE ($RI = 0.926$, $F = 2.22$, $P = 0.0001$); and a linear fit to CODEFOR ($RI = 1$, $F = 2.32$, $P = 0.001$), FARMLMFOR ($RI = 1$, $F = 3.11$, $P =$

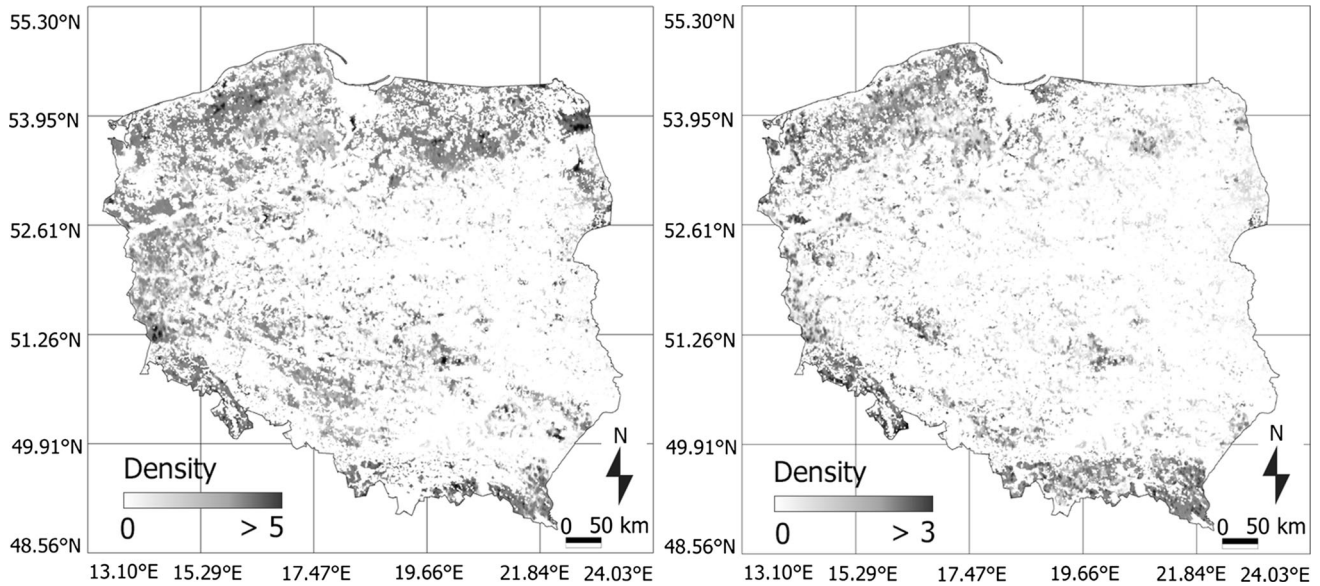


Fig. 2 Predictive maps of Hayne estimator of Goldcrest (a) and Firecrest (b) density

0.0001), TEMP ($RI = 0.999$, $F = 2.87$, $P = 0.0001$), DMFOR ($RI = 0.994$, $F = 1.97$, $P = 0.001$) and PREC ($RI = 0.627$, $F = 1.54$, $P = 0.004$). This model explained the variation of \hat{D}_{H-F} better than the second model in our candidate set (evidence ratio 1.46, Table 2, model D3).

In turn the most parsimonious model of GAM (Table 2, model E2) based only on the detailed data on the forest structure and the fit to the Hayne estimator of the Firecrest's density (\hat{D}_{H-F}) included a smooth fit to DOM.BEECH ($RI = 1$, $F = 2.07$, $P = 0.001$), DOM.SPRUCE ($RI = 0.972$, $F = 1.42$, $P = 0.003$), DOM.FIR ($RI = 0.991$, $F = 3.09$, $P < 0.001$), AGE.FIR ($RI = 1$, $F = 2.99$, $P < 0.001$), AGE.LARCH ($RI = 1$, $F = 1.66$, $P = 0.003$) and DOM.PINE ($RI = 0.991$, $F = 1.99$, $P = 0.001$) and linear fit to NO.SPEC ($RI = 0.994$, $F = 1.22$, $P = 0.001$). This model was definitely better for describing the variation of \hat{D}_{H-F} than the second model (evidence ratio = 87.36, Table 2, model E3).

However, the most parsimonious model (Table 2, model F2) which, apart from the coarse data, used the forest structure data (the same dataset as in \hat{D}_{H-G}) and the Goldcrest's density (\hat{D}_{H-G}) included a smooth fit to FARMLMFOR, \hat{D}_{H-G} , DOM.BEECH, DOM.SPRUCE, LATITUDE, ALTITUDE, TEMP, PREC and a linear fit to CODEFOR, AGE.SPRUCE, DOM.PINE, AGE.FIR, AGE.LARCH to DOM.SMAP, AGE.ELM and NO.SPEC. This model was slightly better than the second model in our candidate set (evidence ratio 1.002, Table 2, model F3). In this case, FARMLMFOR ($RI = 1$, $F = 3.22$, $P = 0.0001$; Fig. 3a) and CODEFOR ($RI = 1$, $F = 1.45$, $P = 0.000$; Fig. 3b) turned out to be the most important factors. All of them represented the habitat gradient from arable fields (FARMLMFOR) and a small area of

coniferous forest (CODEFOR) to small and large areas of mixed and deciduous forests. The third component, \hat{D}_{H-G} ($RI = 1$, $F = 16.14$, $P = 0.0001$; Fig. 3c) indicated that high densities of Firecrests were noted in areas with high densities of Goldcrests. Next important variables were DOM.BEECH ($RI = 1$, $F = 16.9$, $P = 0.00001$; Fig. 3d) and DOM.SPRUCE ($RI = 1$, $F = 1.9$, $P = 0.0002$; Fig. 3e). All of them showed that the highest \hat{D}_{H-F} corresponded to areas where the beech and spruce dominated. An old spruce tree was clearly preferred as shown by the following variable: AGE.SPRUCE ($RI = 1$, $F = 2.01$, $P = 0.0007$; Fig. 3f). At the same time the Firecrest avoided the pine (DOM.PINE, $RI = 1$, $F = 3.12$, $P = 0.0001$; Fig. 3g) and preferred a single old fir (AGE.FIR, $RI = 1$, $F = 1.98$, $P = 0.0001$; Fig. 3h) and an old larch (AGE.FIR, $RI = 1$, $F = 1.44$, $P = 0.001$; Fig. 3i). Other important factors included NO.SPEC ($RI = 1$, $F = 3.01$, $P = 0.02$; Fig. 3j), showing that Firecrests preferred a multi-species forest. The last variables included LATITUDE ($RI = 0.685$, $F = 1.29$, $P = 0.0001$; Fig. 3k), ALTITUDE ($RI = 0.685$, $F = 5.12$, $P = 0.0001$; Fig. 3l), TEMP ($RI = 0.899$, $F = 2.19$, $P = 0.0001$; Fig. 3m) and PREC ($RI = 0.899$, $F = 1.99$, $P = 0.0001$; Fig. 3n). All of them presented a non-linear shape of the response function, showing that the highest density was noted in upland, both in southern and northern parts of Poland, where the climate was relatively cold and wet. According to this model we created a predictive map of \hat{D}_{H-F} (Fig. 2b).

Discussion

Species distribution models (SDMs) are developed in order to identify ecological requirements of species and to

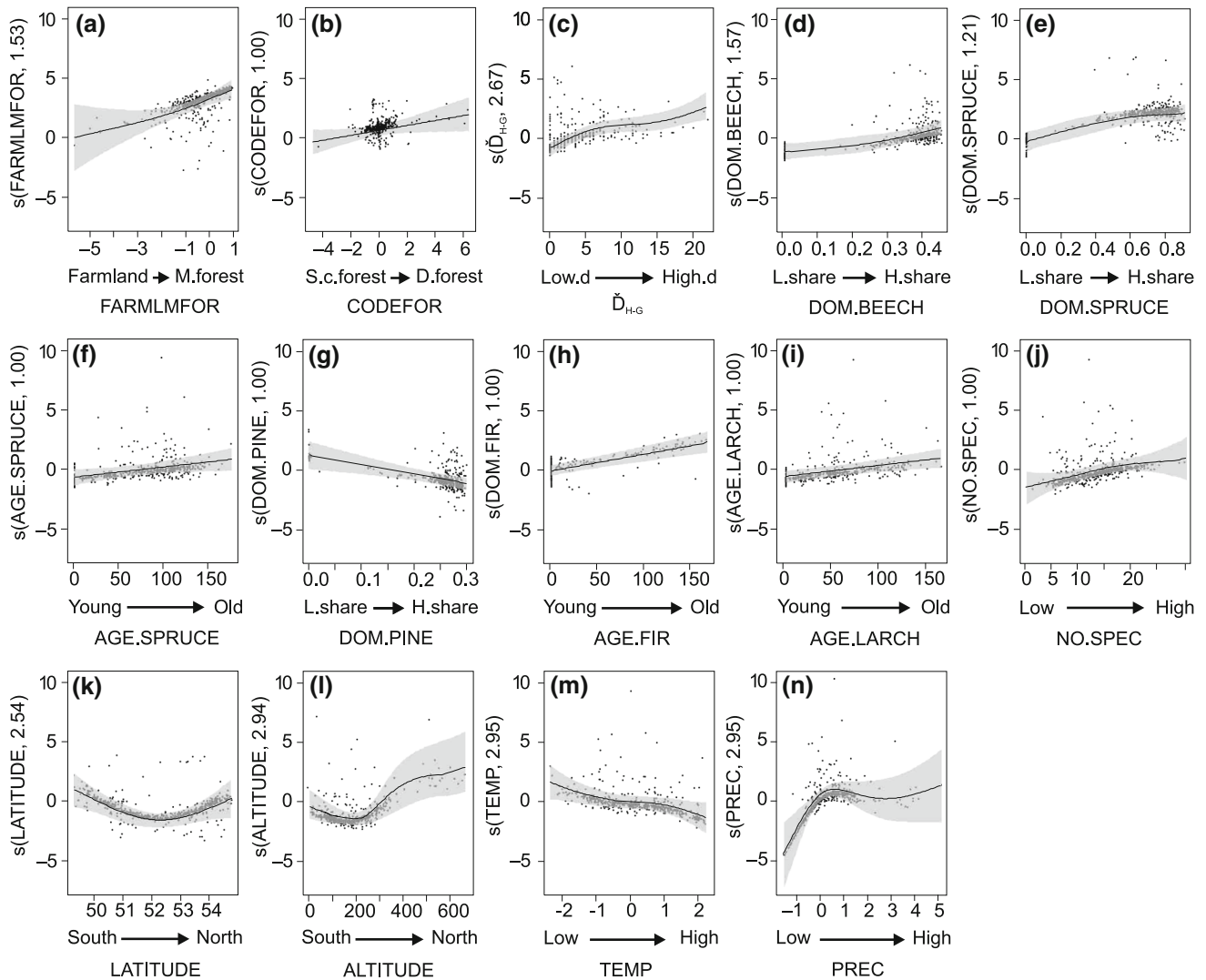


Fig. 3 GAM fit for the Firecrest density. **a** FARMLMFOR is related to the gradient from arable fields (farmland) to small and large areas of mixed forest (M. forest); **b** CODEFOR represented a habitat gradient from small area of coniferous forest (S.c. forest) to large area of deciduous forest (D. forest); **c** \hat{D}_{H-G} showing areas with low density of the Goldcrest (Low.d) and areas with high density (High.d); **d** DOM.BEECH, **e** DOM.SPRUCE, **f** DOM.PINE reflect forest where share of (respectively) beech fir and pine are low (L.share) and where these species dominated (H.share); **g** AGE.SPRUCE, **h** AGE.FIR, **i** AGE.LARCH reflect forest where occur young spruce and fir (Young), while (Old)—shows forest where dominated old these species; **j** NO.SPEC represented the number of tree species (L.number)—areas with low number of

tree species, while (H.number)—areas with high number of species; **k** LATITUDE represented geographical gradient from south (South) to north (North); **l** ALTITUDE represented a geographical gradient from area of low elevation (Low) to area of high elevation (High); **m** TEMP represented the temperature gradient, (Low)—areas with low temperature, while (High)—areas with high temperature **n** PREC represented the level of precipitation (Low)—areas with low level of precipitation, while (High)—areas with high level of precipitation. The y -axis: density—(s) smoother function with estimate degrees of freedom in parenthesis (Hastie and Tibshirani 1990, Brown 2011). The shaded areas represent standard errors of the estimate curves

predict their spatial density (Naimi et al. 2011). From the practical point of view, SDMs can provide valuable support for conservation efforts (Jetz et al. 2012) both at local, regional, and trans-boundary scales (Margules and Pressey 2000; Addison et al. 2013). SDMs can use either presence-absence or presence-only occurrence data of species of interest, but most of them use publicly available, remote sensing-based datasets as environmental predictors (Guisan et al. 2013). Such datasets, due to their generalization and/or coarse resolution, might be insufficient for the task of identifying most important environmental

factors for certain species (Morelli and Tryjanowski 2014), especially the specialized ones.

In this study, we tested the usefulness of generalized remote sensing data as predictors of two *Regulus* species. Our first step was to develop separate prediction models for Goldcrests' and Firecrests' densities using the same set of remote sensing-based predictors. The Goldcrest responded positively to the increasing area of coniferous forest, while the Firecrest preferred mixed and deciduous forests. Moreover, the predicted densities of both species increased in uplands and in cold and wet

climate of northern (especially north-east for both species as well as north-west for the Firecrest) and southern parts of Poland. For both species the relationship was obvious (Ulfstrand 1976; Tiainen et al. 1983; del Hoyo et al. 2006; Kralj et al. 2013), but it should be noted that open-access datasets like the Corine Land Cover or remote sensing-based forest cover maps do not contain information about the species' composition, only about general forest types, such as coniferous vs. deciduous. In Poland, similarly to many Central European countries, the most abundant conifer species is the pine *Pinus sylvestris*. Therefore, using only open source habitat data as a predictor might lead to a conclusion that high densities of Goldcrests and Firecrests depend on the prevalence of the pine (see also Fig. 1g and 3g). Indeed, in the Pyrenees both Goldcrests and Firecrests are recorded in pine forests (Purroy 1974a, b), but this is not the case elsewhere (Lebreton and Thevenot 2009). Therefore, models that use only open access generalized data do not provide complete information regarding the species' preferences towards specific forest conditions.

Our results also show that the Goldcrest prefers forest areas with high NDVI values at the beginning of spring while the Firecrest does not exhibit such preferences. This might be explained by the fact that the Goldcrest favours coniferous forests that are green all year round (especially in winter). The Firecrest is abundant in mixed forests that have a high NDVI rate both in early spring, when deciduous trees are still leafless, and in late spring when both deciduous and coniferous trees appear green. According to our results, both species are abundant in the uplands. These relationships reflect a general pattern of species distribution in Europe. On the scale of the continent the Goldcrest prefers middle and high altitudes (Hagemeijer and Blair 1997; del Hoyo et al. 2006), while the Firecrest is particularly abundant in only mountainous forests in southern parts of Europe (Leisler and Thaler 1982; Telleria and Santos 1994; Hagemeijer and Blair 1997).

The second group of models based on the detailed data on the forest structure proved to be definitely better than the first group of models (Goldcrest Δ AIC = 77.3, Firecrest Δ AIC = 34.9), and differences from the third group of models where both generalized and detailed environmental data were used as a predictors, were comparatively small (Goldcrest Δ AIC = 16.6, Firecrest Δ AIC = 4.6). Distributions of both Goldcrests and Firecrests were best described by a combination of features measured by remote sensing and by on-the-ground measurement of the habitat; the ground-measured habitat features played a greater role in predicting distributions of these two species. Similar patterns were found in connection to other species, e.g. the Hazel Grouse *Bonasa bonasia* (Müller et al. 2009); the Rufous Bristlebird *Dasyornis broadbenti* (Gibson et al. 2004) as well as in bird species richness in the Yellowstone National Park (Saveraid et al. 2001). However, the studies were carried out on a relatively small research area (e.g. 97 km² where it was quite easy to count the number of

small-scale habitat elements); and-or closed areas, i.e. National Parks where small-scale habitat elements were frequently recorded. In our case (study area > 300.000 km²) we used data from the forest inventory database. The primary aim of the database was to describe the main features of forests in terms of size, condition and change of growing trees, focusing mainly on productive aspects, namely wood resources (Rego et al., 2004). Nevertheless, the forest stand description can be very useful for modelling species' occurrence (Stachura-Skierczyńska et al. 2009; Stachura-Skierczyńska and Kosiński 2014), implementing monitoring schemes (Skierczyński et al. 2013), and as shown by our study on species distribution modelling of density specialized species.

The Goldcrest reached the highest density in areas with a high share of old spruce forest with fir. It generally avoided pines and preferred forests with a low number of tree species. The Firecrest chose mixed forests, dominated by beech and old spruce with an admixture of single old fir and larch trees. Firecrests also avoided pines, but favoured forests with a high number of tree species. These different habitat choices can be explained by the existing differences in the morphology and foraging behaviours of these species (Leisler and Thaler 1982; Kralj et al. 2013). Goldcrests more often forage by climbing vertically and hanging under dense spruce branches, whereas Firecrests prefer trees with less dense branches that allow them to move more rapidly and forage in a standing position (Leisler and Thaler 1982).

Furthermore, both species prefer old forests and definitely avoid pine stands. In all biomes, old-growth forests tend to have a complex structure and a variety of microhabitats characterised by different temperatures, light and humidity, which bring about a high abundance of insects, such as *Hemiptera*, *Collembola* and spiders, being major food resources for the *Regulus* and other bird species (Cramp 1992; Jokimäki and Solonen 2011; Rosenvald et al. 2011; Edenius et al. 2012). Moreover, old-growth forests are characterized by a relative spatial-temporal stability, ensuring long-term stable conditions for breeding bird communities (Wesołowski et al. 2006; Wirth et al. 2009; Rosenvald et al. 2011). Such inherent features are not demonstrated by young pine forests, which in most cases are poor in insect species and do not provide birds with a sufficient amount of food (Wirth et al. 2009; Rosenvald et al. 2011). Furthermore, longer needles of the pine do not allow birds to move efficiently through the branches, a drawback which also makes acquiring food difficult.

As for the climate and the topography, both turned out to be significant factors influencing the Goldcrest's and the Firecrest's densities. However, their importance was greater in the model based on remote sensing data only than where also detailed information about forest structure was involved. A recent study indicated the importance of climatic variables in the multiple regression model with respect to both species' abundance

(Kuczyński and Chylarecki 2012; Kralj et al. 2013). However, it should be noted that the above study was based on data from a wide longitudinal gradient (from central Spain to eastern Poland) (Kralj et al. 2013). In our study, the influence of latitude, altitude and weather conditions on Goldcrests' and Firecrests' densities is most probably indirect. The spruce and the fir prefer relatively cold and wet climate, therefore their geographical ranges in Poland are limited to the southern, mountainous and upland regions (both species) and north-eastern parts (spruce) where hemiboreal coniferous forests occur naturally. Low humidity in early spring is a limiting factor for the beech, which is present in the mountains and in the west of Poland, but does not occur further in the east where the climate is dryer (Zajac and Zajac 2001). The abundance of both studied *Regulus* species overlaps with the distribution of preferred trees, meaning that these small scale habitat variables are among the most important factors influencing their densities (all RI = 1).

We also found that the high density of the Goldcrest was likely to determine the high density of the Firecrest in areas of their coexistence, however, the opposite relationship was not established, i.e. the high abundance of the Firecrest did not always indicate the high abundance of the Goldcrest. The coexistence of species is possible only if their niches are not identical (Krebs 2009). Morelli and Tryjanowski (2014) point out that this relationship may be just a local phenomenon. Still, distribution ranges of both species overlap across a large area of sympatry in central Europe. Therefore, we suspect that this relationship reflects a general pattern within the European distribution range. Furthermore, from the behavioural point of view it is possible that the Firecrest uses the Goldcrest's presence as a clue for a suitable habitat, as is the case with the Pied Flycatcher (*Ficedula hypoleuca*) and the Titmice (*Parus* spp) (Forsman et al. 2002).

It turns out that separate models based on the open source remote sensing data and/or forest inventory data derived from ground measurements can only partially explain the distribution of the Goldcrest and the Firecrest. In our case habitat (from both sources) and climate variables as well as topography explain spatial variation of both species but do not provide a complex explanation of all factors driving the species' habitat relationships. It is highly possible that the integration different land-use types (here: small and large scale habitat data) reflects landscape complexity—one of the key elements leading to high predictive performance (McPherson and Jetz 2007). On the other hand, low predictive performance of models based only on remote sensing data can be the consequence of a non-stationary phenomenon showing different responses to the same predictor in different geographical areas (Fink et al. 2010). If Poland's territory was divided into smaller geographic sub-regions based on biogeographic zones, for example, and separate species distribution models based on remote sensing data (or only small scale data)

were constructed for these sub-regions and then averaged, predictive performance could turn out to be higher than in our approach.

Conclusion

Separate SDMs based on remote sensing data and/or detailed information on forests derived from ground measurements do not contain all important ecological factors that can affect species' distribution on a large geographical scale. From a conservationist's point of view, remote sensing data cannot replace the painstaking field study of habitat requirements for individual species (Storch et al. 2003), but it can be stated that the implementation of small scale habitat variables in SDMs can reduce overall costs and improve the efficiency of field studies since it allows to identify concentrations of high species densities with much higher accuracy.

In our opinion, habitat modelling for practical spatial planning and conservation purposes should not be based entirely on free data sources, but it also should incorporate as detailed data as possible. In case of forest areas, the best available source of data is the forest inventory database. Although designed for commercial purposes, namely inventory of wood resources, it turns out to be useful for studies on conservation.

Acknowledgments We wish to express our gratitude to numerous observers who collected data in the field. The full list of their names can be found at <http://main5.amu.edu.pl/~kubako/tsr/>. We want to thank Przemysław Chylarecki for the access to Birds' Database and analysis suggestion, Justyna Grześkowiak for her linguistic assistance and two anonymous reviewers for valuable comments. The Polish Common Breeding Bird monitoring project was funded by the Royal Society for the Protection of Birds (UK), United Nations' GEF Small Grant Programme, European Commission (The Cooperation Fund), Polish Chief Inspectorate of Environmental Protection (GIOS) and National Fund for Environmental Protection and Water Management (NFOSiGW), and coordinated by the Museum and Institute of Zoology Polish Academy of Sciences and the Polish Society for the Protection of Birds (OTOP). Forest inventory data was provided by State Forests National Forest Holding. The study has been supported by MNiSW grant N N304 025936 and National Scholarship Programme of the Slovak Republic (NSP) to JZK.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Addison PFE, Rumpff L, Bau SS, Carey JM, Chee YE, Jarrad FC, McBride MF, Burgam MA (2013) Practical solutions for making models indispensable in conservation decision-making. *Divers Distrib* 19:490–502

- Akçakaya HR (2001) Linking population-level risk assessment landscape with and habitat models. *Sci Total Environ* 274:283–291
- Anon (2013) Raport o stanie lasów w Polsce 2012. Centrum Informacyjne Lasów Państwowych, Warszawa
- Bartoń K (2013) MuMin: multi-model inference. R package version 1.9.0. <http://cran.r-project.org/web/packages/MuMin/index.html>. Accessed 28 Oct 2013
- Bennett AF, Nimmo DG, Radford JQ (2014) Riparian vegetation has disproportionate benefits for landscape-scale conservation of woodland birds in highly modified environments. *J Appl Ecol* 51:514–523
- Brambilla M, Bassi E, Bergero V, Casale F, Chemollo M, Falco R, Longoni V, Saporetti F, Viganò E, Vitulano S (2013) Modeling distribution and potential overlap between Boreal Owl *Aegolius funereus* and Black Woodpecker *Dryocopus martius*: implications for management and monitoring plans. *Bird Conserv Int* 23:502–511
- Brown GS (2011) Patterns and causes of demographic variation in a harvested moose population: evidence for the effects of climate and density-dependent drivers. *J Anim Ecol* 80:1288–1298
- Buchanan GM, Fishpool LDC, Evans MI, Butchart SHM (2013) Comparing field-based monitoring and remote-sensing, using deforestation from logging at important bird areas as a case study. *Biol Conserv* 167:334–338
- Burnham KP, Anderson DR (eds) (2002) Model selection and multimodel inference: a practical information-theoretic approach, 2nd edn. Springer-Verlag, NY
- Chylarecki P, Jawińska D (2007) Monitoring Pospolitych Ptaków Lęgowych. Raport z lat 2005–2006. OTOP Warszawa
- Cramp S (1992) Handbook of the birds of Europe, the Middle East and Africa. The Birds of the Western Palearctic Warblers, vol 4. Oxford University Press, Oxford
- Da Silva LP, Alves J, Silva AAD, Ramos JA, Fonseca C (2012) Variation in the abundance and reproductive characteristics of Great Tits *Parus major* in forest and monoculture plantations. *Acta Ornithol* 47:14–155
- del Hoyo J, Elliott A, Christie DA (eds) (2006) Handbook of the birds of the world, vol 11. Lynx Edicions, Barcelona
- Development Core Team R (2010) R: a language and environment for statistical computing. Austria, R Foundation for Statistical Computing, Vienna
- Doll CNH, Muller J-P, Morley JG (2007) Mapping regional economic activity from night-time light satellite imagery. *Ecol Econ* 57:75–92
- Edenius L, Mikusiński G, Witzell J, Bergh J (2012) Effects of repeated fertilization of young Norway spruce on foliar phenolics and arthropods: implications for insectivorous birds' food resources. *Forest Ecol Manage* 277:38–45
- EEA (2007) CLC2006 technical guidelines. EEA Technical report No 17/2007. European Environmental Agency, Copenhagen
- Ewers RM, Didham RK (2006) Confounding factors in the detection of species responses to habitat fragmentation. *Biol Rev* 81:117–142
- Fahrig L (2002) Effect of habitat fragmentation on the extinction threshold: a synthesis. *Ecol Appl* 12:346–353
- Fahrig L (2003) Effect of habitat fragmentation on biodiversity. *Ann Rev Ecol Systematics* 34:487–515
- Fink D, Hochachka WM, Zuckerberg B, Winkler DW, Shaby B, Munson MA, Hooker G, Riedewald M, Sheldon D, Kelling S (2010) Spatiotemporal exploratory models for broad-scale survey data. *Ecol Appl* 20:2131–2147
- Forsman JT, Seppänen J, Mönkkönen M (2002) Positive fitness consequences of interspecific interaction with a potential competitor. *Proc R Soc Lond B* 269:1471–2954
- Franklin J (2009) Mapping species distributions. Cambridge University Press, Spatial Inference and Prediction
- Gibson LA, Wilson BA, Cahill DM, Hill J (2004) Spatial prediction of rufous bristlebird habitat in a coastal heathland: a GIS-based approach. *J Appl Ecol* 41:213–223
- Giordano PF, Navarro JL, Martella MB (2010) Building large-scale spatially explicit models to predict the distribution of suitable habitat patches for the Greater rhea (*Rhea americana*), a near-threatened species. *Biol Conserv* 143:357–365
- Gregory DR, Vorisek P, Van Strien A, Gemelig Meyling AW, Jiguet F, Fornasari L, Reif J, Chylarecki P, Burfield IJ (2007) Population trends of widespread woodland birds in Europe. *Ibis* 149:78–97
- Guisan A, Tingley R, Baumgartner JB, Naujokaitis-Lewis I, Sutcliffe PR, Tulloch AI, Regan TJ, Brotons L, McDonald-Madden E, Mantyka-Pringle C, Martin TG, Rhodes JR, Maggini R, Setterfield SA, Elith J, Schwartz MW, Wintle BA, Broennimann O, Austin M, Ferrier S, Kearney MR, Possingham HP, Buckley YM (2013) Predicting species distributions for conservation decisions. *Ecol Lett* 16:1424–1435
- Hagemeijer WJM, Blair MJ (1997) The EBCC atlas of European breeding birds: their distribution and abundance. T & AD Poyser, London
- Hastie T, Tibshirani R (1990) Generalized additive models. Chapman and Hall, London
- Hastie T, Tibshirani RJ, Friedman J (2008) The elements of statistical learning, 2nd edn. Springer, New York
- Hayne DW (1949) An examination of the strip census method for estimation animal populations. *J Wildlife Managet* 13:145–157
- Heiberger RM (2013) Statistical Analysis and Data Display: Heiberger and Holland. R package version 2.3–37. <http://cran.r-project.org/web/packages/HH/index.html>. Accessed 1 Oct 2013
- Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* 25:1965–1978
- Horn A, Kerdelhué C, Rossi J-P (2012) Predicting the distribution of the two bark beetles *Tomicus destruens* and *Tomicus piniperda* in Europe and the Mediterranean region. *Agr Forest Entomol* 14:358–366
- Jetz W, McPherson JM, Guralnick RP (2012) Integrating biodiversity distribution knowledge: toward a global map of life. *Trends Ecol Evol* 27:151–159
- Jokimäki J, Solonen T (2011) Habitat associations of old forest bird species in managed boreal forests characterized by forest inventory data. *Ornis Fenn* 88:57–70
- Kajtoch L, Żmihorski M, Bonczar Z (2012) Hazel Grouse occurrence in fragmented forests: habitat quality and configuration is more important than quality. *Eur J For Res* 131:1783–1795
- Kissling WD, Dormann CF, Groeneveld J, Hickler T, Kühn I, McInerny GJ, Montoya JM, Römermann C, Schiffers K, Schurr FM, Singer A, Svenning J-C, Zimmermann NE, O'Hara RB (2012) Towards novel approaches to modelling biotic interactions in multispecies assemblages at large spatial extents. *J Biogeogr* 39:2163–2178
- Kosicki JZ, Chylarecki P (2013) Predictive mapping of meadow pipit density using integrated remote sensing data and atlas of vascular plants dataset. *Bird Study* 60:500–508
- Kosicki JZ, Chylarecki P (2014) The Hooded Crow *Corvus cornix* density as a predictor of wetland bird species richness on large geographical scale in Poland. *Ecol Indic* 38:50–60
- Kralj J, Flousek J, Huzak M, Ciković D, Dolenc Z (2013) Factors affecting the goldcrest/firecrest abundance ratio in their area of sympatry. *Ann Zool Fenn* 50:333–346
- Krebs CJ (1999) Ecological methodology. Addison Wesley Longman, California
- Krebs ChJ (2009) Ecology: the experimental analysis of distribution and abundance, 6th edn. Benjamin Cummings, San Francisco
- Kuczyński L, Chylarecki P (2012) Atlas of common breeding birds in Poland: distribution. Habitat Preferences and Population Trends, GIOŚ Warszawa
- Kuczyński L, Rzępała M, Goławski A, Tryjanowski P (2009) The wintering distribution of great grey shrike *Lanius excubitor* in Poland: predictions from a large-scale historical survey. *Acta Ornithol* 44:159–166

- Lampila P, Mönkkönen M, Rajasärkkä A (2009) The ability of forest reserves to maintain original fauna-Why has the Chiffchaff (*Phylloscopus collybita abietinus*) disappeared from eastern central Finland? *Ornis Fenn* 86:71–80
- Lebreton P, Thevenot M (2009) Comparaison de l'avifaune sapinières et de pinèdes méditerranéennes et méditerranéennes. *Revue D'Ecologie La Terre et La Vie* 64:171–188
- Leisler B, Thaler E (1982) Differences in morphology and foraging behaviour in the goldcrest *Regulus regulus* and firecrest *R. ignicapillus*. *Ann Zool Fenn* 19:277–284
- Linden DW, Roloff GJ (2013) Retained structures and bird communities in clearcut forests of the Pacific Northwest, USA. *Forest Ecol Manag* 310:1045–1056
- Margules CR, Pressey RL (2000) Systematic conservation planning. *Nature* 405:243–253
- Mazgajski TD, Żmihorski M, Abramowicz K (2010) Forest habitat loss and fragmentation in Central Poland during the last 100 years. *Silva Fennica* 44:715–723
- McPherson JM, Jetz W (2007) Effect of species' ecology on the accuracy of distribution models. *Ecography* 30:135–151
- Morelli F, Tryjanowski P (2014) Associations between species can influence the goodness of fit of species distribution models: The case of two passerine birds. *Ecol. Complex* (in press)
- Müller D, Schröder B, Müller J (2009) Modelling habitat selection of the cryptic Hazel Grouse *Bonasa bonasia* in a montane forest. *J Ornithol* 150:717–732
- Naimi B, Skidmore TA, Groen NAS (2011) Spatial Autocorrelation in predictors reduces the impact of positional uncertainty in occurrence data on species distribution modelling. *J Biogeogr* 38:1497–1509
- Neteler M, Mitasova H (2008) Open source GIS: A GRASS GIS approach, 3rd edn. Springer, New York
- O'Brien RM (2007) A caution regarding rules of thumb for variance inflation factors. *Qual Quant* 41:673–690
- Perry RW, Thill RE (2013) Long-term responses of disturbance-associated birds after different timber harvests. *Forest Ecol Manag* 307:274–283
- Purroy FJ (1974a) Breeding communities in the beech and fir forests of the Pyrenees. *Acta Ornithol* 14:151–157
- Purroy FJ (1974b) Contribucion al conocimiento ornitologico de los pinares Pirenaicos. *Ardeola* 20:245–261
- Quinn GP, Keough MJ (2002) Experimental design and data analysis for biologists. Cambridge University Press, Cambridge
- Rego F, Godinho-Ferreira P, Uva JS, Cunha J (2004) Combination of structural and compositional factors for describing forest types using national forest inventory data. *EFI Proc.* 51:153–162
- Reif J, Storch D, Šimová I (2008) The effect of scale-dependent habitat gradient on the structure of bird assemblages in the Czech Republic. *Acta Ornithol* 23:197–206
- Reino L, Portod M, Morgado R, Carvalhof F, Miraf A, Bejac P (2010) Does afforestation increase bird nest predation risk in surrounding farmland? *Forest Ecol Manag* 260:1359–1366
- Rosenvald R, Löhmus A, Kraut A, Remm L (2011) Bird communities in hemiboreal old-growth forests: the roles of food supply, stand structure, and site type. *Forest Ecol Manag* 262:1541–1550
- Royle JA, Chandler RB, Yackulic C, Nichols JD (2012) Likelihood analysis of species occurrence probability from presence-only data for modelling species distributions. *Methods Ecol Evol* 3:545–554
- Sallabanks R, Arnett EB, Marzluff JM (2000) An evaluation of research on the effects of timber harvest on bird populations. *Wildl Soc Bull* 28:1144–1155
- Sanderson FJ, Kloch A, Sachanowicz K, Donald PF (2009) Predicting the effect of agricultural change on farmland bird populations in Poland. *Agric Ecosys Environ* 129:37–42
- Santana J, Porto M, Gordinho L, Reino L, Beja P (2012) Long-term responses of Mediterranean birds to forest fuel management. *J Appl Ecol* 49:632–643
- Saunders DA, Hobbs RJ, Margules CR (1991) Biological consequences of ecosystem fragmentation: a review. *Conserv Biol* 5:18–32
- Saveraid EH, Debinski DM, Kindscher K, Jakubauskas ME (2001) A comparison of satellite data and landscape variables in predicting bird species occurrences in the Greater Yellowstone Ecosystem, USA. *Land Ecol* 16:71–83
- Skierczyński M, Stachura-Skierczyńska K, Kajtoch Ł, Ciach M, Strzeński P (2013) Forest inventory data in conservation planning: predicting rare woodpeckers habitat quality on regional scale. *Acta Biol Cracoviensia Ser Bot* 55:32
- Small C, Pozzi F, Elvidge CD (2005) Spatial analysis of global urban extent from DMSP-OLS nighttime lights. *Remote Sens Environ* 96:277–291
- Spanhove T, Vanden Bore J, Delalieux S, Haest B, Paelinckx D (2012) Can remote sensing estimate fine-scale quality indicators of natural habitats? *Ecol Indic* 18:403–412
- Stachura-Skierczyńska K, Kosiński Z (2014) Evaluating habitat suitability for the middle spotted woodpecker using a predictive modelling approach. *Ann Zool Fenn* 51 (in press)
- Stachura-Skierczyńska K, Tumiel T, Skierczyński M (2009) Habitat prediction model for three-toed woodpecker and its implications for the conservation of biologically valuable forest. *Forest Ecol Manag* 258:697–703
- Storch D, Konvička M, Beneš J, Martinková J, Gaston KJ (2003) Distribution patterns in butterflies and birds of the Czech Republic: separating effects of habitat and geographic position. *J Biogeogr* 30:1195–1205
- Telleria JL, Santos T (1994) Factors involved in the distribution of forest birds in the Iberian Peninsula. *Bird Study* 41:161–169
- Tiainen J, Vickholm M, Pakkala T, Piironen J, Virolainen E (1983) The habitat and spatial relations of breeding *Phylloscopus* warblers and the Goldcrest *Regulus regulus* in southern Finland. *Ann Zool Fenn* 20:1–12
- Ulfstrand S (1976) Feeding niches of some passerine birds in a South Swedish coniferous plantation in winter and summer. *Ornis Scand* 7:21–27
- Walczak L, Kosiński Z, Stachura-Skierczyńska K (2013) Factors affecting the occurrence of middle spotted woodpeckers as revealed by forest inventory data. *Baltic Forestry* 19:81–88
- Weisberg S (1980) Applied linear regression. Wiley, New York
- Wesołowski T, Rowiński P, Mitrus C, Czeszczewik D (2006) Breeding bird community of a primeval temperate forest (Białowieża National Park, Poland) at the beginning of the 21st century. *Acta Ornithol* 41:55–70
- Whittingham MJ, Swetnam RD, Wilson JD, Chamberlain DE, Freckleton RP (2005) Habitat selection by yellowhammers *Emberiza citrinella* on lowland farmland at two spatial scales: implications for conservation management. *J Appl Ecol* 42:270–280
- Wirth C, Gleixner G, Heimann M (2009) Old-growth forests: function, fate and value, ecological studies. Springer, Berlin
- Wood S (2013) mgcv: mixed gam computation vehicle with GCV/AIC/REML smoothness estimation. R package version 1.7–22. <http://cran.r-project.org/web/packages/mgcv/index.html>. Accessed 10 Oct 2013
- Zajac A, Zajac M (2001) Distribution atlas of vascular plants in Poland. Kraków
- Żmihorski M (2012) The effect of anthropogenic and natural disturbances on breeding birds of managed Scots pine forests in northern Poland. *Ornis Fenn* 89:63–73