



Multivariate analysis of a fine-scale breeding bird atlas using a geographical information system and partial canonical correspondence analysis: environmental and spatial effects

Nicolas Titeux^{1,2}, Marc Dufrêne³, Jean-Paul Jacob³, Marc Paquay⁴ and Pierre Defourny^{1*}

¹Department of Environmental Sciences and Land Use Planning and ²Biodiversity Research Centre, Université catholique de Louvain, Louvain-la-Neuve, ³Observatoire de la Faune, de la Flore et des Habitats, Gembloux and ⁴Rue des Marmozets 1, Ciergnon, Belgium

ABSTRACT

Aim To assess the relative roles of environment and space in driving bird species distribution and to identify relevant drivers of bird assemblage composition, in the case of a fine-scale bird atlas data set.

Location The study was carried out in southern Belgium using grid cells of 1 × 1 km, based on the distribution maps of the *Oiseaux nicheurs de Famenne: Atlas de Lesse et Lomme* which contains abundance for 103 bird species.

Methods Species found in < 10% or > 90% of the atlas cells were omitted from the bird data set for the analysis. Each cell was characterized by 59 landscape metrics, quantifying its composition and spatial patterns, using a Geographical Information System. Partial canonical correspondence analysis was used to partition the variance of bird species matrix into independent components: (a) 'pure' environmental variation, (b) spatially-structured environmental variation, (c) 'pure' spatial variation and (d) unexplained, non-spatial variation.

Results The variance partitioning method shows that the selected landscape metrics explain 27.5% of the variation, whilst 'pure' spatial and spatially-structured environmental variables explain only a weak percentage of the variation in the bird species matrix (2.5% and 4%, respectively). Avian community composition is primarily related to the degree of urbanization and the amount and composition of forested and open areas. These variables explain more than half of the variation for three species and over one-third of the variation for 12 species.

Main conclusions The results seem to indicate that the majority of explained variation in species assemblages is attributable to local environmental factors. At such a fine spatial resolution, however, the method does not seem to be appropriated for detecting and extracting the spatial variation of assemblages. Consequently, the large amount of unexplained variation is probably because of missing spatial structures and 'noise' in species abundance data. Furthermore, it is possible that other relevant environmental factors, that were not taken into account in this study and which may operate at different spatial scales, can drive bird assemblage structure. As a large proportion of ecological variation can be shared by environment and space, the applied partitioning method was found to be useful when analysing multispecific atlas data, but it needs improvement to factor out all-scale spatial components of this variation (the source of 'false correlation') and to bring out the 'pure' environmental variation for ecological interpretation.

*Correspondence: Pierre Defourny, Department of Environmental Sciences and Land Use Planning, Université catholique de Louvain, Croix du Sud 2/16, B-1348 Louvain-la-Neuve, Belgium. E-mail: defourny@enge.ucl.ac.be

Keywords

Atlas, birds, partial Canonical Correspondence Analysis, variation partitioning, spatial distribution, landscape, Geographical Information Systems.

INTRODUCTION

The mapping of bird distributions using grid-based atlases at different spatial scales is one of the most frequently used form of ornithological survey (e.g. Jacob & Paquay, 1992; Hagemeyer & Blair, 1997; Schmid *et al.*, 1998). Atlases have numerous potential uses including bird–environment relationship assessment (Donald & Fuller, 1998) and the prediction of species distribution (Bibby *et al.*, 1992; Tobalske & Tobalske, 1999; Pearson & Dawson, 2003). A constraint to such analyse is the availability of relevant environmental data at the same spatial scale as the atlas data (Donald & Fuller, 1998). The development of Geographical Information Systems (GIS) and the increasing availability of digital maps permit the manipulation of spatially-distributed data (Burrough & McDonnell, 1998) to extract habitat data, i.e. various landscape metrics, at selected scales (Berry & McGarigal, 1998), and thus to avoid this technical problem.

The use of atlas data and GIS to develop habitat models (Gates *et al.*, 1993; Tobalske & Tobalske, 1999; Siriwardena *et al.*, 2000; Vanhinsbergh & Chamberlain, 2001), to predict bird species distributions (e.g. Osborne & Tigar, 1992), and to assess and predict relative distribution changes (Böhning-Gaese & Bauer, 1996; Chamberlain & Fuller, 2000; Chamberlain *et al.*, 2001; Telfer *et al.*, 2002) has already been explored in a species-specific way, often by means of Generalized Linear Models (GLM). Other taxonomic groups have also been investigated in the same way (e.g. Dennis & Hardy, 1999; Guisan *et al.*, 1999; Luoto *et al.*, 2002). In the context of multispecies analysis, atlases and GIS have already been used to provide environmental interpretation of avifaunal zonation (Pasinelli *et al.*, 2001), to model species richness distribution (Lobo *et al.*, 2002; Lobo & Martin-Piera, 2002; Maes *et al.*, 2003) and to relate avian (Natuhara & Imai, 1996; Storch *et al.*, 2003) or floristic (Guisan *et al.*, 1999) assemblages to environmental conditions. In order to explain relationships between species assemblages and environmental variables, ordination methods are often used, especially direct gradient analyses in which species occurrence or abundance are directly related to environmental variables (ter Braak, 1986).

Spatial structuring is a fundamental and functional component of ecosystems: the observed distribution of organisms (response variables) result from the spatial dependence in these response variables (spatial autocorrelation) – because of biotic processes such as growth, mortality, dispersal or predation – or are caused by the dependence of response variables on several

explanatory variables which are themselves spatially-structured, or both (Legendre & Legendre, 1998). Spatial heterogeneity is not the result of some random, noise-generating processes, and it is thus necessary to take it into account. However, incorporating the spatial structure of the data (response and explanatory variables) into the modelling process is rarely done in the case of multispecific atlas data analysis. Appropriate methods dealing with this very general property of ecological variables exist (Legendre, 1993). Borcard *et al.* (1992) have proposed a method based on partial Canonical Correspondence Analysis (CCA) to factor out the spatial component of the ecological variation of species assemblages and to partition this variation as follows: (a) non-spatial environmental variation, (b) spatially-structured environmental variation, (c) spatial variation that is not shared by environmental variables and (d) unexplained, non-spatial variation. This method has been used for several purposes, notably in the assessment of the relative roles of environment and space in driving fauna or vegetation distributions (i.e. Borcard & Legendre, 1994; Pinelalloul *et al.*, 1995; Monti *et al.*, 1996; Hobson *et al.*, 2000). To our knowledge it has never been applied to atlas data sets in spite of these data being a special case of spatially-structured ecological data. Therefore, such an assessment should be performed. The study objective is to evaluate the applicability of this variance partitioning technique for a fine-scale breeding bird atlas for a region of 279 km² encompassing various land use types, in order to quantify the relative roles of environment and space in driving bird species assemblages, to display substantial ecological or spatial gradients and to highlight relevant environmental variables for bird assemblage composition.

MATERIALS AND METHODS

Study area

The region is located in the southern part of Belgium and forms a vast depression (altitude: 140–300 m) characterized by schistous and calcareous rocks. The study area corresponds to the Central Famenne and covers 279 km². It is centred on the confluence of two rivers: the Lesse and the Lomme (Fig. 1). This area has poorly drained clay soils derived from schist, which are unfavourable to agriculture. Consequently, land use is dominated by grassland with scattered hedgerows (39% of the total surface) and forests of varying extents and types (37%), while cultivated fields (16.5%) and urban areas (4%) are less frequent.

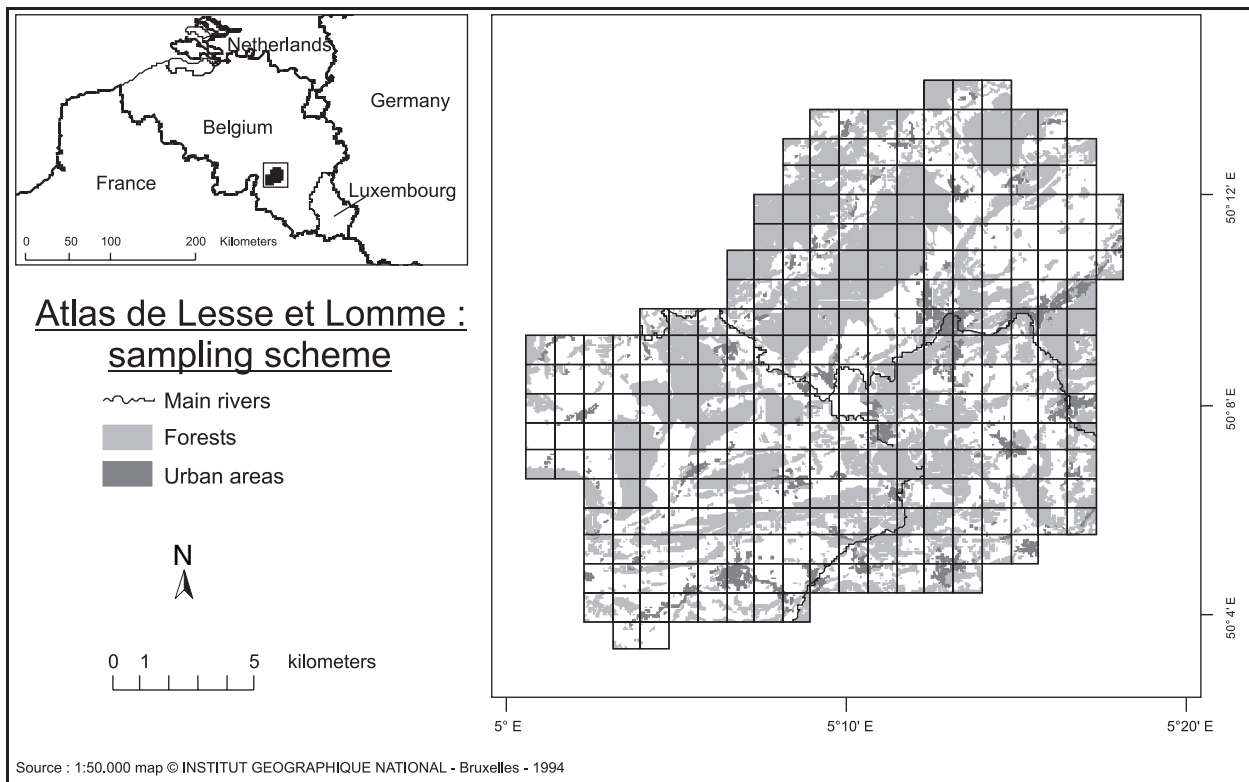


Figure 1 Study area covered by the *Atlas de Lesse et Lomme* (Jacob & Paquay, 1992) showing the 279 atlas squares.

Avifauna and landscape data

Avifauna

The distribution maps presented in the *Oiseaux nicheurs de Famenne: Atlas de Lesse et Lomme* (Jacob & Paquay, 1992) are composed of 279 1 km² grid cells, hereafter called 'squares' (Fig. 1). Bird censuses were conducted from 1985 to 1989 during spring and summer (breeding season) by 70 members of the Centrale Ornithologique Aves (COA, Belgium).

Each year between 1985 and 1988, a different set of squares was surveyed three times during the breeding season (March–April, May and June–July). Surveys began 15–20 min before sunrise and ended within 4 h after sunrise. Each square was totally surveyed during one season. The last year (1989) was devoted to verification to ensure that each square was surveyed with the same intensity and to limit the observer bias that could arise from *a priori* knowledge of the avifauna of certain squares or from differential census intensity (J.-P. Jacob and M. Paquay, pers. comm.). The survey time by square ranges from 5 to 25 h depending on the habitats complexity within the square.

For each square, the observer recorded the presence–absence and evaluated the number of breeding pairs (abundance) for each species. Abundance was aggregated to six classes as described in Table 1. For this study, the middle of each interval was retained except for the last one (see Table 1). Only records of probable or confirmed breeding – as defined by the

Table 1 Classes of abundance as counted at the square level and used by the *Atlas de Lesse et Lomme* (Jacob & Paquay, 1992)

Class	Number of breeding pairs	Number of breeding pairs used for analysis
0	0	0
1	1–5	2.5
2	6–10	8
3	11–20	15
4	21–40	30
5	>40	50

international EOAC code (Timothy & Sharrock, 1974) – were considered, while possible breeding was excluded from the species set (for more details, see Jacob & Paquay, 1992).

GIS database

The GIS software ArcInfo 7.0.4. (ESRI, Redlands, CA, USA) was used to combine several habitat data layers from different sources.

A 30-m resolution digital elevation model (Institut Géographique National, edited in 1995) was used to represent the topography of the region. Digital 1 : 50,000 vector topographic maps (Institut Géographique National, edited in 1994) were used as a planimetric reference and land-cover information. In addition, a set of elements that could potentially have a role in

explaining the birds distribution was retained for analysis. In particular, several types of linear and surface features were digitized from aerial 1 : 10,000 colour orthophotographs (acquired between 1985 and 1990) and digital 1 : 10,000 raster topographic maps (Institut Géographique National, edited in 1981 or 1988–89). Map digitizing based on visual interpretation also allows to ‘downdate’ the landscape description to the bird atlas period (1985–89).

Linear elements are hedgerows, railways, roads and the hydrographic network. Hedges were allocated to four classes by visual interpretation of the aerial photographs: (1) broken hedges formed by scattered and isolated shrubs, (2) continuous, but narrow hedges, (3) continuous and wide hedges and (4) tree rows. Roads and railways were allocated to four classes: (1) motorways, major and secondary roads, and railways, (2) smaller roads, (3) pathways and (4) footpaths. The hydrographical network comprised rivers (1) narrower than 3 m, (2) from 3 to 15 m, and (3) from 15 to 30 m. No rivers exceed 30 m.

All surface features were classified into one of 10 land-cover types: pond, marsh, urban area, hardwood forest, coniferous forest, mixed forest, orchard, grassland, crop field and heathland.

Landscape metrics computation

The focus of this study is on the relationships between landscape structure, i.e. composition and spatial patterns, and bird assemblages. Consequently it was necessary to characterize the composition and the spatial configuration of the landscape for the 279 squares of the atlas.

For each atlas square, the total length of the 11 linear elements and the surface occupied by each of the 10 land-cover types were calculated. These 21 metrics document the landscape composition in the 279 atlas squares (Table 2).

Landscape configuration analysis was performed using the standalone ArcInfo application FRAGSTATS*ARC version 2 (Berry & McGarigal, 1998). Numerous metrics of landscape pattern are available, but many are highly correlated or redundant (Coppedge *et al.*, 2001) and no consensus exists as to which metric best describes any particular ecological phenomenon (Gustafson, 1998). These metrics quantify different aspects of landscape configuration, and 34 were selected to quantify the landscape configuration of the 279 atlas squares (Table 2). Two groups of indices were retained: class level metrics (for each land cover or linear element type) and landscape level metrics (for the whole atlas square).

Several core area metrics were computed. Core area is defined as the area within a patch beyond some specified buffer width. Core area metrics were arbitrarily based, for the purpose of this study, on a 50-m wide buffer along the perimeter of each patch. Core area metrics reflect both landscape composition and configuration, and are related to the ‘edge effect’ (Woodward *et al.*, 2001). To avoid the effect of patch truncation at atlas square boundaries on the calculation of core areas, the full size of patches that overlap two squares was

used, i.e. the feature area extending outside of the squares was included for the calculation of core areas (see Trzcinski *et al.*, 1999).

In each atlas square, several statistics representing the total edge length of a particular patch type (class level) were computed. Several interface metrics were also calculated, representing the total edge length between two particular patch types or between a patch type and a linear element.

Landscape level metrics that quantify the square mosaic as a whole were also computed. The area-weighted mean shape index measures the average perimeter-to-area ratio for all patches in the square, by weighting patches according to their size, so that the larger the patch the more strongly it contributes to the value of the metric. Shannon’s diversity index and Shannon’s evenness index refer to the number of patch types and the distribution of area among these different types.

Four metrics were computed to describe the topographic characteristics of each square: minimum elevation, maximum elevation, mean elevation and standard deviation of elevation. The latter metric is a measure of the altitude variability in a square.

The nature and number of explanatory variables (59) resulted from a trade-off between concerns to maintain a high ratio of observations-to-variables for subsequent multivariate analysis (CCA) and a relevant and coherent description of the landscape structure of the squares within the framework of this study. All the landscape metrics, hereafter called ‘environmental variables’, are detailed in Table 2.

Data analysis

Canonical correspondence analysis

The avian assemblages were related to landscape composition and pattern using CCA with the program CANOCO 4.0 for Windows (ter Braak & Smilauer, 1998). CCA is an eigenvector ordination technique for multivariate direct gradient analysis (ter Braak, 1986) and represents a special case of multivariate regression (Palmer, 1993). The ordination axes are constrained to be linear combinations of the explanatory variables (ter Braak, 1986; Natuhara & Imai, 1996; Guisan *et al.*, 1999). CCA explains community variation by detecting patterns of variation in species abundance that can best be explained by a set of environmental variables (ter Braak, 1986), with no *a priori* knowledge about potential predictor variables (Boren *et al.*, 1999).

This technique is appropriated when the responses of the dependent variables are expected to be unimodal along environmental gradients (ter Braak & Smilauer, 1998). Redundancy Analysis is another ordination technique where the underlying response model is a monotonic distribution of species along environmental gradients, which limits its use when large gradients are analysed (Guisan & Zimmermann, 2000). CCA is also appropriate when dealing with occurrence data (ter Braak, 1986).

Table 2 The 59 metrics used to describe the landscape pattern of the 1-km atlas squares

Acronym	Units	Description
Hedge1	m	Total length of scattered and isolated shrubs
Hedge2	m	Total length of continuous but narrow hedges
Hedge3	m	Total length of continuous and wide hedges
Trees	m	Total length of tree rows
Riv1	m	Total length of rivers from 0 to 3 m wide
Riv2	m	Total length of rivers from 3 to 15 m wide
Riv3	m	Total length of rivers from 15 to 30 m wide
RoA	m	Total length of motorways, major and secondary roads and railways
RoB	m	Total length of smaller roads
Path	m	Total length of pathways
Foot	m	Total length of footpaths
Pond	ha	Total area of ponds
Marsh	ha	Total area of marshes
Urban	ha	Total area of urban zones
HFor	ha	Total area of hardwood forests
CFor	ha	Total area of coniferous forests
MFor	ha	Total area of mixed forests
Orch	ha	Total area of orchards
Grass	ha	Total area of grasslands
Field	ha	Total area of fields
Heath	ha	Total area of heathlands
HForCore	ha	Total amount of hardwood forests core area
CForCore	ha	Total amount of coniferous forests core area
MForCore	ha	Total amount of mixed forests core area
GrassCore	ha	Total amount of grasslands core area
FieldCore	ha	Total amount of fields core area
RivFor	m	Total length of edges between rivers and forested areas
RivOpen	m	Total length of edges between rivers and open areas
PondFor	m	Total length of edges between ponds and forested areas
PondOpen	m	Total length of edges between ponds and open areas
RoAFor	m	Total length of edges between roads A and forested areas
RoAOpen	m	Total length of edges between roads A and open areas
RoBFor	m	Total length of edges between roads B and forested areas
RoBOpen	m	Total length of edges between roads B and open areas
HForCFor	m	Total length of edges between hardwood and coniferous forests
HForMFor	m	Total length of edges between hardwood and mixed forests
CForMFor	m	Total length of edges between coniferous and mixed forests
HForField	m	Total length of edges between hardwood forests and fields
HForGrass	m	Total length of edges between hardwood forests and grasslands
CForField	m	Total length of edges between coniferous forests and fields
CForGrass	m	Total length of edges between coniferous forests and grasslands
MForField	m	Total length of edges between mixed forests and fields
MForGrass	m	Total length of edges between mixed forests and grasslands
FieldGrass	m	Total length of edges between fields and grasslands
HForTE	m	Total length of edges involving hardwood forests
CForTE	m	Total length of edges involving coniferous forests
MForTE	m	Total length of edges involving mixed forests
GrassTE	m	Total length of edges involving grasslands
FieldTE	m	Total length of edges involving fields
NP	none	Number of patches (all patch types together)
PR	none	Patch richness (number of patch types)
Awmsi	none	Area-weighted mean shape index = $\sum_{i=1}^m \sum_{j=1}^n [(p_{ij}/2\sqrt{\pi a_{ij}})(a_{ij}/A)]$
TE	m	Total length of edges
Shdi	none	Shannon's diversity index = $-\sum_{i=1}^m (P_i \ln P_i)$
Shei	none	Shannon's evenness index = $[-\sum_{i=1}^m (P_i \ln P_i)] / \ln m$
Altmin	m	Minimal elevation
Altmax	m	Maximal elevation

Acronym	Units	Description
Altmean	m	Mean elevation
Altstd	m	Elevation standard deviation

Table 2 *continued*

p_{ij} , perimeter of patch j for type i ; a_{ij} , surface of patch j for type i ; A , total surface of the square (i.e. 1 km²); P_i = proportion of the total surface of the square occupied by patch type i ; m , number of patch types in the square.

Ideally, rare species would also be examined with gradient analysis. Unfortunately these species tend to obscure community patterns because their weak occurrence introduces a large number of zeros, i.e. absence, in the bird data set. These species may have an unduly large influence on the analysis (ter Braak & Smlauer, 1998), by creating an increase in the total inertia of the species data set or a distortion of the ordination. Furthermore, rare species are generally associated with specific habitat conditions not summarized by the 59 landscape metrics. For these reasons, species found in < 10% of the atlas squares were omitted from the bird data set for the subsequent analysis. Likewise, ubiquitous species will not exhibit a relationship with environmental gradients and may skew or obscure community patterns. Consequently, species found in more than 90% of the atlas squares were also omitted.

Distinction between environmental and spatial effects

Spatial autocorrelation can be defined as the property of random variables taking values, at pairs of locations a certain distance apart, that are more similar (positive autocorrelation) or less similar (negative autocorrelation) than expected for randomly associated pairs of observations (Legendre, 1993). In the case of atlas squares, the observed bird species assemblage in one square might be influenced by the species assemblages in the surrounding squares because of contagious biotic processes. Similarly, environmental variables used to describe the squares are also neither randomly nor uniformly spatially-distributed, but structured by physical processes causing gradients and/or patchy structures. One consequence of this general property of ecological variables is that the assumption of independence of the observations is not respected (Legendre, 1993). Therefore, it is necessary to incorporate the spatial structure of the data within the modelling process. In the present case of multivariate dependent data set, the statistical methods to use to handle such data form a family called *partial (constrained) ordination analyses* (Legendre, 1993; Jongman *et al.*, 1995): the computation are made after removing, by multiple linear regression, the effects of known or undesirable variables, called covariables. Borcard *et al.* (1992) proposed the use of a polynomial function of the geographical coordinates of the sampling locations, i.e. centres of atlas squares, as covariables in the partial CCA. Therefore, a matrix of x and y geographical coordinates of the centres of the squares was compiled by adding all terms for a cubic trend surface regression of the form:

$$f(x, y) = b_1x + b_2y + b_3x^2 + b_4xy + b_5y^2 + b_6x^3 + b_7x^2y + b_8xy^2 + b_9y^3$$

This ensures the detection of more complex spatial features in the species data set than by linear gradient patterns only (Borcard *et al.*, 1992), because third order polynomial regressions simulate either skewed or bimodal responses, or even combinations of both of these (Guisan & Zimmermann, 2000).

In total, three matrices were constructed: species assemblages, environmental variables and polynomial function of geographical coordinates of the squares. By making two canonical ordinations, each of them constrained by one set of explanatory variables, one obtained a measure of the effects of the environmental conditions and the spatial structure that are not independent because of spatially-structured environmental descriptors. Consequently, the two runs are partially redundant (Borcard *et al.*, 1992) and one should be able to partition the variation of the bird species matrix as follows (Legendre, 1993): (a) non-spatial environmental variation, (b) spatially-structured environmental variation, (c) spatial variation that is not shared by environmental variables and (d) unexplained, non-spatial variation. This partition is possible by using the sum of all canonical eigenvalues of two canonical ordinations, each of them constrained by one set of explanatory variables, and of two partial canonical ordinations, each of them constrained by one set of explanatory variables while controlling for the effect of the others (covariables) (Borcard *et al.*, 1992).

Environmental and spatial variables selection

In the canonical ordination constrained by the polynomial function of geographical coordinates of the squares (hereafter called 'spatial' CCA), the geographical terms were submitted to forward selection and Monte Carlo permutation tests (999 permutations; $P < 0.05$) in CANOCO to retain only the most important ones and to avoid the artificial increase of explained variation by chance. These terms will only be incorporated into subsequent partial canonical ordination constrained by spatial effects while controlling for the environmental effects (hereafter called 'spatial' partial CCA). In the partial canonical ordination constrained by environmental variables while controlling for the selected geographical terms (hereafter called 'environmental' partial CCA), variables were submitted to a forward procedure and selected manually on the basis of the additional variation explained. Their significance was

tested by Monte Carlo permutations tests (999 permutations), to retain the 10 ones that best explain the variation in the breeding bird abundance. Furthermore, the interactions of these 10 environmental variables were computed in CANOCO and added as explanatory variables. An additional 'environmental' partial CCA was run using the 10 selected environmental variables and their interactions, and the best 10 variables or interactions were again selected manually during a forward selection procedure. These selection procedures are not species specific: selections are made of the variables which explain successively the highest proportion of variation in the species data set as a whole (Guisan *et al.*, 1999). Forward selection includes collinear variables if they make important contributions to the variation in species composition, but will eliminate completely redundant variables (Palmer, 1993). Finally, only the 10 retained environmental (interactions included) and the selected geographical variables were used in all 'environmental' or 'spatial' (partial) canonical ordinations.

CCA and partial CCA fit and interpretation measures

Canonical eigenvalues measure the amount of variation in the species data set that is explained by the explanatory variables. The total variation in the species matrix (called 'Total Inertia') is measured by the chi-square statistic of the sample-by-species table divided by the table's total (ter Braak & Smilauer, 1998). This is equal to the sum of all eigenvalues of a CA on species assemblages. Dividing the sum of all canonical eigenvalues by the total inertia gives the percentage of total variance in the species data set that is explained by the explanatory variables, i.e. an overall measure of the CCA fit (ter Braak & Smilauer, 1998). Similarly, the proportion of the total inertia in the species data that is explained by each canonical axis was computed. Unrestricted Monte Carlo permutations were used to test the statistical significance of the first canonical axis and of all canonical axes. Tests of significance in CCA do not depend on parametric distributional assumptions (Palmer, 1993). Therefore, species data were not transformed, and environmental and spatial variables were simply standardized.

Canonical coefficient and intraset correlations were examined to evaluate the relative contributions of environmental variables to the prediction of the community composition (Boren *et al.*, 1999). Canonical coefficients are the coefficients of a weighted multiple regression of the sample scores on the environmental variables and thus, define the ordination axes as linear combinations of the environmental variables (ter Braak, 1986). Intraset correlations provide the correlation between an environmental variable and a canonical axis, adjusted for the covariables (partial correlations) (ter Braak, 1986). In order to interpret the ecological signification of the canonical axes, this latter coefficient must be used instead of the canonical coefficient in the case of multicollinearity.

The marginal effect of an environmental variable is equal to the eigenvalue of a partial CCA if the corresponding variable

was the only environmental variable (additionally to the variance explained by covariables). The conditional effect of an environmental variable is equal to the additional amount of variance in species assemblages explained by the corresponding variable at the time it was included into the model during a selection procedure (additionally to the variance explained by covariables) (ter Braak & Smilauer, 1998). Such effects were also examined to assess the relative contributions of environmental variables for predicting the community composition.

Species–environment biplots

The CCA allows a visual interpretation of multiple habitat dimensions by plotting species and environmental variables in the ordination diagram formed by two canonical axes obtained by the 'environmental' partial CCA. The type-2 scaling of CANOCO was used, where the species are represented at the centroids of the sampling sites where they occur. On such a biplot, environmental variables are depicted by lines. The length of each line represents the relative importance of the environmental variable to the community distribution (in the two-dimensional plane under consideration) (ter Braak, 1986). The further the species points are from the origin, the more informative the ordination: a cluster of species near the origin is difficult to interpret and indicates a weak species–environment relationship (MacFaden & Capen, 2002). A perpendicular line can be drawn from each species position to each environmental variable. Endpoints of the perpendicular lines indicate relative positions of species distribution centres along this environmental variable (ter Braak & Smilauer, 1998).

RESULTS

The bird census program detected 103 breeding bird species during 1985–89. The species richness of the squares is 45.7 ± 8.9 (range: 9–68 breeding species by km²). The number of atlas squares occupied by each species is highly variable. Only 80 species were found consistently enough (more than 10% of atlas squares) to allow habitat analysis. Among them, 17 species were detected in more than 90% of atlas squares and were therefore omitted (Table 3).

Distinction between environmental and spatial effects

The following geographical terms were retained by the forward selection procedure ($P < 0.05$) during the 'spatial' CCA:

$$f(x, y) = b_1x + b_2y + b_3x^2 + b_8xy^2$$

The 10 environmental variables or interactions selected during 'environmental' partial CCA, are summarized in Table 4. The number of explanatory variables was arbitrarily fixed at 10. The use of additional variables in the analysis does not substantially improve the results because these explain only a negligible part of the variation. Species abundance transformation, taking Napierian logarithms [$z' = \ln(z + 1)$], does not considerably improve the results either.

Table 3 List of all recorded species with English and scientific names (taxonomic order), code (based on scientific names), total number of atlas squares occupied (*N*) and proportion of variance explained by environmental variables and by space as a result of the partial Canonical Correspondence Analysis (only for species present in > 10% and < 90% of atlas squares)

Species					
English name	Scientific name	Code	<i>N</i>	% of variance explained by environment	% of variance explained by space
Little grebe	<i>Tachybaptus ruficollis</i> *	Taru	13	–	–
Grey heron	<i>Ardea cinerea</i> *	Arci	2	–	–
Common mallard	<i>Anas platyrhynchos</i>	Anpl	91	28	< 1
European honey-buzzard	<i>Pernis apivorus</i>	Peap	34	7	3
Red kite	<i>Milvus milvus</i> *	Mimi	1	–	–
Eurasian sparrowhawk	<i>Accipiter nisus</i>	Acni	86	19	2
Common buzzard	<i>Buteo buteo</i>	Bubu	181	37	5
Common kestrel	<i>Falco tinnunculus</i>	Fati	50	10	4
Hazel grouse	<i>Bonasa bonasia</i> *	Bobo	4	–	–
Common quail	<i>Coturnix coturnix</i> *	Coct	24	–	–
Common pheasant	<i>Phasianus colchicus</i>	Phco	138	18	2
Water rail	<i>Rallus aquaticus</i> *	Raaq	3	–	–
Corn crake	<i>Crex crex</i> *	Crcr	5	–	–
Common moorhen	<i>Gallinula chloropus</i> *	Gach	22	–	–
Eurasian coot	<i>Fulica atra</i> *	Fuat	3	–	–
Northern lapwing	<i>Vanellus vanellus</i>	Vava	67	11	2
Little ringed plover	<i>Charadrius dubius</i> *	Chdu	3	–	–
Eurasian woodcock	<i>Scolopax rusticola</i>	Scru	44	8	5
Stock pigeon	<i>Columba oenas</i>	Cooe	29	3	2
Common wood-pigeon	<i>Columba palumbus</i> †	Copa	266	–	–
Eurasian collared-dove	<i>Streptopelia decaocto</i>	Stde	102	6	3
European turtle-dove	<i>Streptopelia turtur</i>	Sttu	177	8	2
Eurasian cuckoo	<i>Cuculus canorus</i>	Cuca	177	13	2
Barn owl	<i>Tyto alba</i>	Tyal	27	19	1
Eurasian eagle-owl	<i>Bubo bubo</i> *	Bubb	3	–	–
Little owl	<i>Athene noctua</i>	Atno	62	14	2
Tawny owl	<i>Strix aluco</i>	Stal	80	16	3
Long-eared owl	<i>Asio otus</i>	Asot	71	12	4
Eurasian nightjar	<i>Caprimulgus europaeus</i> *	Caeu	2	–	–
Common swift	<i>Apus apus</i>	Apap	49	37	3
Common kingfisher	<i>Alcedo atthis</i> *	Alat	10	–	–
Eurasian green woodpecker	<i>Picus viridis</i>	Pivi	104	9	6
Black woodpecker	<i>Dryocopus martius</i>	Drma	44	14	3
Great Spotted woodpecker	<i>Dendrocopos major</i>	Dema	227	44	< 1
Middle spotted woodpecker	<i>Dendrocopos medius</i> *	Deme	24	–	–
Lesser spotted woodpecker	<i>Dendrocopos minor</i>	Demi	68	16	< 1
Wood lark	<i>Lullula arborea</i> *	Luar	12	–	–
Skylark	<i>Alauda arvensis</i>	Alar	216	45	4
Sand martin	<i>Riparia riparia</i> *	Riri	11	–	–
Barn swallow	<i>Hirundo rustica</i>	Hiru	130	40	1
Northern house-martin	<i>Delichon urbica</i>	Deur	78	31	3
Tree pipit	<i>Anthus trivialis</i> †	Antr	261	–	–
Meadow pipit	<i>Anthus pratensis</i>	Anpr	122	20	2
Grey wagtail	<i>Motacilla cinerea</i>	Moci	59	25	2
White wagtail	<i>Motacilla alba</i>	Moal	223	20	1
White-throated dipper	<i>Cinclus cinclus</i>	Cici	36	17	4
Winter wren	<i>Troglodytes troglodytes</i> †	Trtr	271	–	–
Hedge accentor	<i>Prunella modularis</i> †	Prmo	268	–	–
European robin	<i>Erithacus rubecula</i> †	Erru	271	–	–
Common nightingale	<i>Luscinia megarhynchos</i>	Lume	116	15	2
Black redstart	<i>Phoenicurus ochruros</i>	Phoc	137	25	1
Common redstart	<i>Phoenicurus phoenicurus</i>	Phph	73	21	1
Common stonechat	<i>Saxicola torquata</i>	Sato	34	8	2

Table 3 continued

Species					% of variance explained by environment	% of variance explained by space
English name	Scientific name	Code	N			
Common blackbird	<i>Turdus merula</i> †	Tume	277	–	–	–
Fieldfare	<i>Turdus pilaris</i>	Tupi	80	13	–	5
Song thrush	<i>Turdus philomelos</i> †	Tuph	270	–	–	–
Mistle thrush	<i>Turdus viscivorus</i>	Tuvi	183	22	–	1
Common grasshopper-warbler	<i>Locustella naevia</i>	Lona	95	11	–	1
Marsh warbler	<i>Acrocephalus palustris</i>	Acpa	91	14	–	3
Eurasian reed-warbler	<i>Acrocephalus scirpaceus</i> *	Acsc	21	–	–	–
Melodious warbler	<i>Hippolais polyglotta</i>	Hipo	59	7	–	1
Lesser whitethroat	<i>Sylvia curruca</i>	Sycu	221	18	–	3
Greater whitethroat	<i>Sylvia communis</i>	Syco	227	28	–	1
Garden warbler	<i>Sylvia borin</i> †	Sybo	274	–	–	–
Blackcap	<i>Sylvia atricapilla</i> †	Syat	276	–	–	–
Wood warbler	<i>Phylloscopus sibilatrix</i>	Phsi	204	57	–	< 1
Common chiffchaff	<i>Phylloscopus collybita</i> †	Phcl	276	–	–	–
Willow warbler	<i>Phylloscopus trochilus</i> †	Phtr	269	–	–	–
Goldcrest	<i>Regulus regulus</i>	Rere	247	36	–	4
Firecrest	<i>Regulus ignicapillus</i>	Reig	214	27	–	4
Spotted flycatcher	<i>Muscicapa striata</i>	Must	107	10	–	2
Long-tailed tit	<i>Aegithalos caudatus</i>	Aeca	215	21	–	3
Willow tit	<i>Parus montanus</i> †	Pamo	265	–	–	–
Marsh tit	<i>Parus palustris</i>	Papa	202	32	–	1
Crested tit	<i>Parus cristatus</i>	Pacr	162	37	–	5
Coal tit	<i>Parus ater</i>	Paat	154	26	–	10
Blue tit	<i>Parus caeruleus</i> †	Paca	268	–	–	–
Great tit	<i>Parus major</i> †	Pama	277	–	–	–
Eurasian nuthatch	<i>Sitta europaea</i>	Sieu	199	52	–	2
Eurasian tree-creeper	<i>Certhia familiaris</i> *	Cefa	17	–	–	–
Short-toed tree-creeper	<i>Certhia brachydactyla</i>	Cebr	243	38	–	1
Eurasian golden-oriole	<i>Oriolus oriolus</i> *	Oror	22	–	–	–
Red-backed shrike	<i>Lanius collurio</i>	Laco	55	8	–	3
Great grey shrike	<i>Lanius excubitor</i>	Laex	45	22	–	1
Eurasian jay	<i>Garrulus glandarius</i>	Gagl	244	35	–	1
Black-billed magpie	<i>Pica pica</i>	Pipi	158	18	–	2
Eurasian jackdaw	<i>Corvus monedula</i>	Como	40	22	–	3
Rook	<i>Corvus frugilegus</i> *	Cofr	13	–	–	–
Carrion crow	<i>Corvus corone</i> †	Coco	258	–	–	–
Common starling	<i>Sturnus vulgaris</i>	Stvu	187	27	–	3
House sparrow	<i>Passer domesticus</i>	Pado	146	54	–	1
Eurasian tree sparrow	<i>Passer montanus</i>	Pamn	105	17	–	5
Chaffinch	<i>Fringilla coelebs</i> †	Frcs	274	–	–	–
European serin	<i>Serinus serinus</i> *	Sese	24	–	–	–
European greenfinch	<i>Carduelis chloris</i>	Cach	178	25	–	1
European goldfinch	<i>Carduelis carduelis</i>	Caca	69	6	–	6
Eurasian linnet	<i>Carduelis cannabina</i>	Cacn	222	27	–	< 1
Red crossbill	<i>Loxia curvirostra</i> *	Locu	26	–	–	–
Eurasian bullfinch	<i>Pyrrhula pyrrhula</i>	Pypy	165	11	–	1
Hawfinch	<i>Coccothraustes coccothraustes</i>	Cocc	220	29	–	3
Yellowhammer	<i>Emberiza citrinella</i> †	Emci	263	–	–	–
Reed bunting	<i>Emberiza schoeniclus</i>	Emsc	47	12	–	2
Corn bunting	<i>Miliaria calandra</i> *	Mica	5	–	–	–

*Species present in < 10% of atlas squares.

†Species present in > 90% of atlas squares.

Table 4 Environmental variables selected by the manual forward procedure, in order of their inclusion in the model during the ‘environmental’ partial Canonical Correspondence Analysis of the species assemblages. The canonical coefficients and intraset correlations, the additional variance each variable explains at the time of inclusion (i.e. conditional effect, $\lambda - A$), the marginal effect of each variable ($\lambda - 1$), the statistics of the Monte Carlo significance test for the forward procedure (F) and the associated probability (P) are reported for each variable

Variable	Canonical coefficients				Intraset correlations				$\lambda - A$	$\lambda - 1$	F	P -value
	Axis 1	Axis 2	Axis 3	Axis 4	Axis 1	Axis 2	Axis 3	Axis 4				
HFor	0.6326	-0.6497	-0.5956	-0.0438	0.7493	-0.3406	-0.4401	0.0255	0.16	0.16	31.31	0.001
Urban \times GrassTE	0.1141	-0.1032	-0.4268	1.1805	-0.6710	-0.5577	-0.0806	0.3362	0.12	0.15	24.27	0.001
MFor	0.3241	-0.3263	0.1560	0.0908	0.3926	-0.2087	0.2887	0.0734	0.05	0.05	10.35	0.001
CFor	0.2580	-0.2229	0.5720	-0.1212	0.4617	-0.1522	0.6095	-0.0814	0.03	0.07	7.79	0.001
FieldCore	0.0753	0.2927	-0.2427	0.3732	0.0166	0.6975	-0.1468	0.3775	0.03	0.06	6.85	0.001
RivOpen	-0.0197	-0.2328	0.4262	-0.4125	-0.4211	-0.1794	0.3461	-0.1572	0.02	0.06	3.61	0.001
Urban ²	0.5204	0.6586	0.6065	0.7342	-0.3991	-0.3806	0.0359	0.5376	0.01	0.07	3.67	0.001
Urban	-0.9459	-1.2093	-0.3641	-1.1449	-0.6305	-0.5518	-0.0478	0.3945	0.02	0.14	2.94	0.001
Altmin \times GrassTE	-0.0292	-0.1148	-0.2231	-0.9806	-0.3766	0.0133	-0.2141	-0.5799	0.01	0.05	3.40	0.001
Pond ²	-0.0569	-0.1075	0.2894	0.2588	-0.1172	-0.1171	0.3409	0.2360	0.01	0.02	3.15	0.006

The four (partial) CCA runs give the following results:

1. ‘environmental’ CCA of the species assemblages: sum of all canonical eigenvalues = 0.532;
2. ‘spatial’ CCA of the species assemblages: sum of all canonical eigenvalues = 0.109;
3. ‘environmental’ partial CCA of the species assemblages: sum of all canonical eigenvalues = 0.464;
4. ‘spatial’ partial CCA of the species assemblages: sum of all canonical eigenvalues = 0.042.

The total variation in the species assemblages (total inertia) is 1.69. The total percentage of variance explained by the environmental variables and by the polynomial function of geographical coordinates is equal to the sum of all canonical eigenvalues of a CCA (step 1 or 2) plus the sum of all canonical eigenvalues of a partial CCA (step 4 or 3, respectively), divided by total inertia, namely $(0.532 + 0.042) \times 100/1.69 = (0.109 + 0.464) \times 100/1.69 = 33.96\%$.

The variation of species assemblages can be partitioned as follows (Borcard *et al.*, 1992): (a) non-spatial environmental variation ($0.464 \times 100/1.69 = 27.46\%$), (b) spatially-structured environmental variation [$(0.532 - 0.464) \times 100/1.69 = (0.109 - 0.042) \times 100/1.69 = 4.02\%$], (c) spatial variation of the species assemblages that are not shared by environmental variables ($0.042 \times 100/1.69 = 2.49\%$) and (d) unexplained, non-spatial variation ($100 - 33.96 = 66.04\%$). Figure 2 illustrates the four partitions of the variation of species assemblages.

For the ‘environmental’ partial CCA of the species assemblages, the Monte Carlo test of the relationships found that the first partial CCA axis ($F = 50.465$; $P < 0.001$) and the overall analysis ($F = 10.98$; $P < 0.001$) were highly significant, indicating that axes of partial CCA significantly explain species assemblages after accounting for ‘pure’ spatial variations. On the contrary, for the ‘spatial’ partial CCA of the species assemblages, the Monte Carlo test of the relationships found that the first partial CCA axis ($F = 3.619$; $P < 0.001$) and the

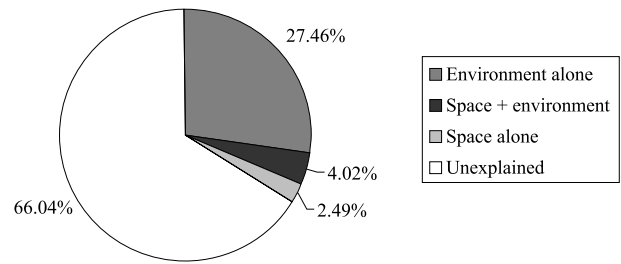


Figure 2 Variation partitioning of the bird species assemblages.

overall analysis ($F = 2.468$; $P < 0.001$) were highly significant, indicating that axes of partial CCA significantly explain species assemblages after accounting for ‘pure’ environmental variations.

Environmental drivers of bird species assemblages

All subsequent results are relative to ‘environmental’ partial CCA.

The first two partial CCA axes accounted respectively for 54.6% (eigenvalue = 0.254) and 25.4% (eigenvalue = 0.118) of the extracted variance in the species–environment relationship, and for 16% and 7.5% of the inertia of species data, after removing the effects due to ‘pure’ spatial variation of the species assemblages. Therefore, the first two canonical axes explain about 80% of the bird species – non-spatial environment relationships and 23.5% of the non-spatially-structured variation of species assemblages. The third (eigenvalue = 0.03) and fourth axes (eigenvalue = 0.02) taken together explain about 10% of the bird species – non-spatial environment relationships and 3.2% of the non-spatially-structured variation of species assemblages. Eigenvalues calculated for each ordination axis indicate the degree of separation of species along this axis and serve as a measure of the

importance of the axis (ter Braak & Smilauer, 1998). Consequently, the first axis is about twice as important as the second, and so on.

The percentage of the variance explained varies markedly from species to species (Table 3) and ranges from 3% for stock pigeon (*Columba oenas*) to about 57% for wood warbler (*Phylloscopus sibilatrix*). About half of the species have a percentage of the variance explained higher than 20% while more than half of the variation is explained for three species.

No multicollinearity was detected in the 10 selected environmental variables as Variance Inflation Factors were all <20 (ter Braak & Smilauer, 1998). Therefore, both intraset correlations and canonical coefficients may be used to assess the relative contributions of environmental variables to the prediction of community composition, as well as marginal and conditional effects of selected variables, respectively $\lambda - 1$ and $\lambda - A$ (Table 4).

Several species–environment biplots were reported (Figs 3 & 4), illustrating species assemblages in the three-dimensional ordination space of the first three canonical axes. These ordination diagrams show well-spread distributions of species and environmental variables, indicating continuous changes in avifauna along ecological gradients. Therefore, the relationships between avian assemblages and environmental variables

in terms of the main axes of variation is easy to interpret. The positions of samples on the diagrams are not shown here, but do not demonstrate substantial discontinuities. In this unimodal ordination analysis, the species score, depicted on the biplots, is a weighted average of the sample scores, i.e. the centre of the species bell-shaped distribution along the ordination axis, namely an approximation of the species optimum.

The examination of intraset correlations in Table 4 and of biplots on Figs 3 and 4 indicates that the first canonical axis is positively correlated with forested areas and negatively correlated with the amount of urban–grassland areas. Therefore, the first axis separates urban and grassland species, such as house and eurasian tree sparrows (*Passer domesticus* and *P. montanus*), little owl (*Athene noctua*) or eurasian jackdaw (*Corvus monedula*), from forest species such as wood warbler (*Phylloscopus sibilatrix*), goldcrest (*Regulus regulus*), great spotted woodpecker (*Dendrocopos major*) or eurasian nut-hatch (*Sitta europaea*). The second axis is negatively correlated with urban and forested areas but positively correlated with the variable *FieldCore* while the variable *Altmin × GrassTE* obtains an intermediate canonical coefficient. Thus, within open habitat species, the second axis separates those which are more associated to urban areas, such as common swift (*Apus apus*) or barn owl (*Tyto alba*), from those that are

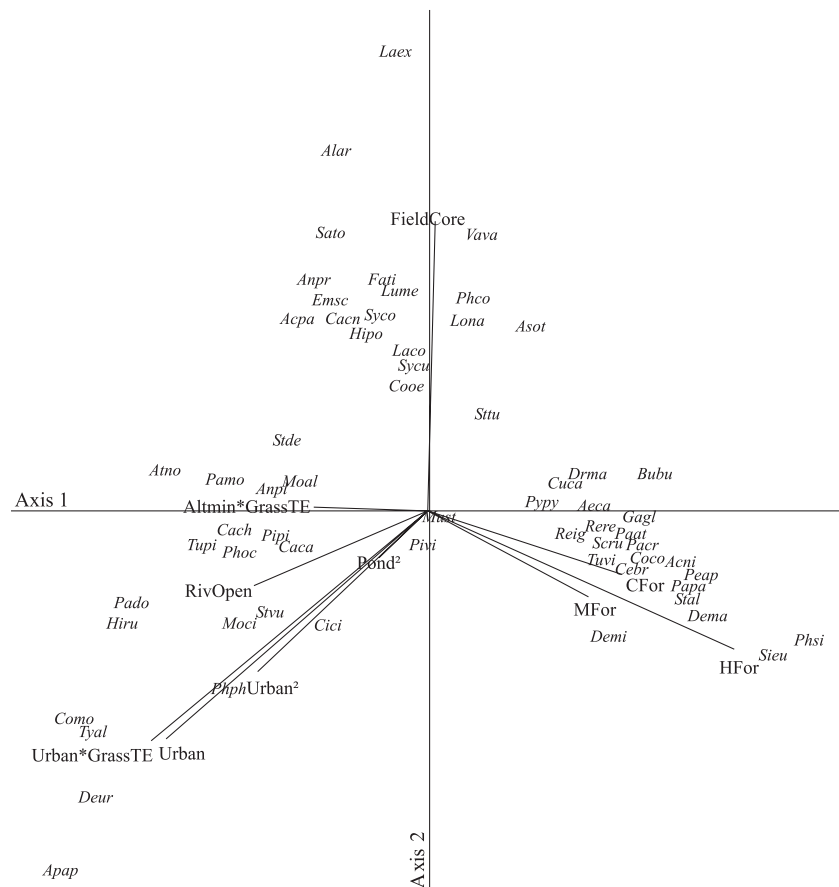


Figure 3 Ordination biplot depicting the first two axes of the 'environmental' partial Canonical Correspondence Analysis of the species assemblages. Landscape metrics are represented by lines and their acronyms (see Table 2), and species locations by their code, as indicated in Table 3.

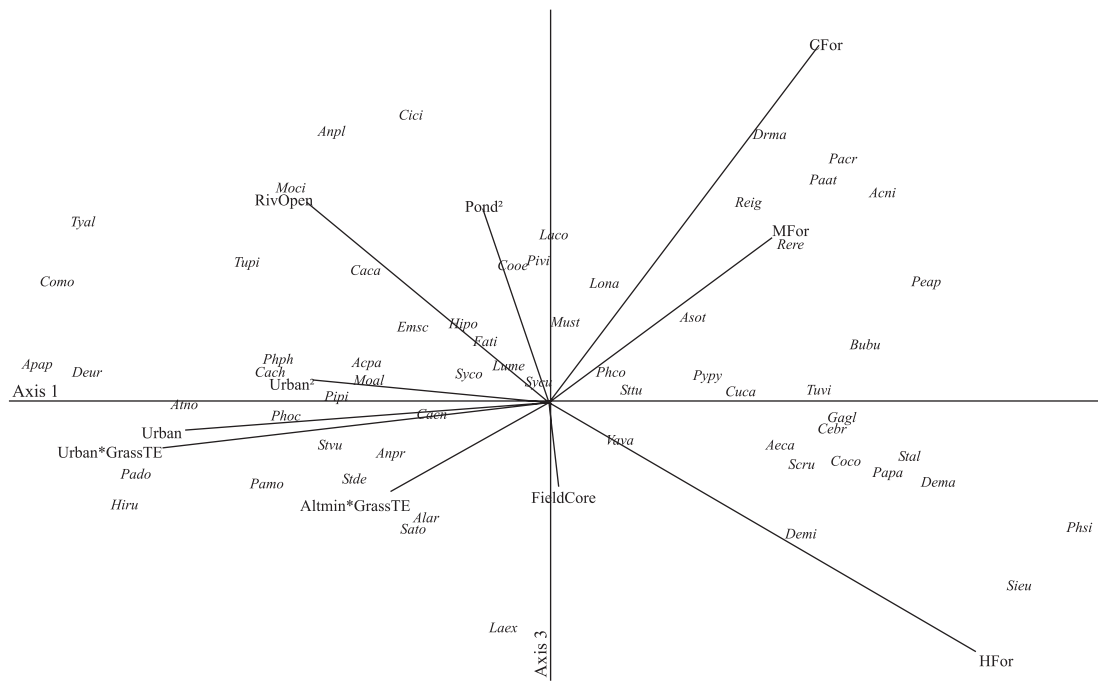


Figure 4 Ordination biplot depicting the first and third axes of the ‘environmental’ partial Canonical Correspondence Analysis of the species assemblages. Landscape metrics are represented by lines and their acronyms (see Table 2), and species locations by their code, as indicated in Table 3.

found in very open lowland or field habitats, such as great grey shrike (*Lanius excubitor*) or skylark (*Alauda arvensis*), including species of grassland and hedges, such as the European greenfinch and goldfinch (*Carduelis chloris* and *C. carduelis*). The third axis notably separates two forest habitats and species: hardwood forests on one side, with wood warbler (*Phylloscopus sibilatrix*) or eurasian nuthatch (*Sitta europaea*) for instance, and coniferous and mixed forests on the other side, with Eurasian sparrowhawk (*Accipiter nisus*), goldcrest and firecrest (*Regulus regulus* and *R. ignicapillus*) or crested and coal tits (*Parus cristatus* and *P. ater*). The third axis is also responsible for the separation of aquatic species: as seen on Figs 3 and 4, white-throated dipper (*Cinclus cinclus*) and grey wagtail (*Motacilla cinerea*) are clustered in the three-dimensional ordination space formed by the first three axes, and closely associated to rivers in open habitats and ponds.

DISCUSSION

Environmental and spatial effects

The spatial pattern of species assemblages was assessed by means of a cubic trend surface function of the geographical coordinates of atlas squares which ensured the extraction of not only linear gradients in the species assemblages, but also more local landscape characteristics such as patches or gaps. About 2.5% of the species assemblage variation (i.e. 7.4% of explained variation) is attributable to these spatial variables

without being related to the 10 selected environmental variables. This fraction varies from species to species, ranging from 0.1% for wood warbler (*Phylloscopus sibilatrix*) to 10% for coal tit (*Parus ater*) (Table 3). In the ‘spatial’ CCA (step 2), it can be observed that the percentage of variance explained by species is substantially higher than these latter figures for all species (mean of 5.48%), especially for forest species [for instance, 14.17% for great spotted woodpecker (*Dendrocopos major*) and 17.12% for eurasian nuthatch (*Sitta europaea*)]. This is because of the spatial arrangement of forests in large patches within the study area. This broad-scale pattern is well captured by the selected terms of geographical coordinates while more local patches are probably not detected (Meot et al., 1998) and require many additional terms. After factoring out the effect of environment (‘spatial’ partial CCA), however, space seems to be a minor descriptor of patterns of species assemblages undetected by the 10 environmental variables. The majority (about two-thirds) of the spatial distributions of species is because of the spatial arrangement of their habitats (see below for the interpretation of this fraction). The remaining ‘pure’ spatial variation reflects some contagious biological processes without relationship to the environmental variables included in the analysis, such as growth, predation, reproduction or social aggregation for instance (Borcard et al., 1992), creating spatial autocorrelation in the multivariate dependent data set (bird species matrix). At a broader scale, this fraction may reflect the role of spatial dynamics and/or historical patterns of dispersal in producing contemporary distributional patterns. This effect could be potentially strong

and well captured by a polynomial trend surface when larger areas are investigated, i.e. at national or continental scales, including several biogeographical regions and where some species reach their distributional limits (Storch *et al.*, 2003). In such cases, geographical coordinates could affect the shape and structure of distributional ranges. The low influence of 'pure' space in our ordinations seems to indicate that bird assemblages are weakly spatially-structured. The spatial structure of the birds assemblages could however have been missed by the polynomial trend surface. These results suggest that the use of a polynomial function of geographical coordinates of the sampling sites is appropriate for modelling broad-scale variation but may not be adequate for quantifying spatial patterns at finer resolution. Consequently, to detect spatial patterns over a wider range of scales, an improvement in the method is required.

Environment explains about 31.5% of the variation in the species assemblages, for which about 4% is shared by the polynomial function of geographical coordinates of the atlas squares. In similar studies (e.g. Borcard *et al.*, 1992; Borcard & Legendre, 1994; Hobson *et al.*, 2000), spatially-structured environmental variation can reach highly variable proportions, sometimes surpassing by far the 'pure' spatial and 'pure' environmental variations. The fairly low fraction observed here seems to suggest that species assemblages and environmental variables do not have a common spatial pattern and do not respond to similar underlying causes (Borcard *et al.*, 1992). Again, a large part of the spatially-structured environmental conditions could have been missed by the polynomial trend surface. The remaining 27.5% of the environmental variation reflects local effects of the environmental variables after partialling out the spatial component. This partition is rarely higher than 40–50% in other studies and is sometimes close to the value obtained here or smaller (e.g. Borcard *et al.*, 1992; Borcard & Legendre, 1994; Hobson *et al.*, 2000). As shown by the biplots, by the order of variable selection during manual forward procedure, by canonical coefficients and intraset correlations, and by the relative marginal and conditional effects of these variables (Table 4), avian assemblages are primarily driven by the degree of urbanization and the amount of forested and open areas. To a lesser extent, the composition of open habitats (grassland vs. fields) and of forests (coniferous/mixed forests vs. hardwood forests) are also influential, as well as aquatic habitats. Other environmental variables seem to be less important in structuring bird assemblages. For instance, hedges do not appear during the manual forward selection, maybe because they are not described with sufficient accuracy in terms of floristic composition or spatial arrangement. Likewise, landscape level metrics seem to be less important, probably because they are highly correlated with landscape composition and configuration metrics and thus, were not included during the forward selection procedure. Biologically, they also correspond to the habitat requirements of a minority of species. Nevertheless, the number of species per square is positively and signifi-

cantly explained by landscape level metrics such as *TE*, *Shdi* and *PR* (N. Titeux unpubl. data), indicating a strong relationship between landscape heterogeneity and species richness.

Unexplained variation interpretation

The unexplained variation of species assemblages is quite high (about 66%) and can be ascribed to the stochastic space–time fluctuations of the communities, to the fact that some species do not occupy all patches of suitable habitat ('unsaturation'), to unmeasured environmental variables or to spatially-structuring processes that have been missed by the selected geographical terms (see above for a discussion of the latter point). Incorporation of the following best explanatory variables in the analysis does not substantially increase the sum of all canonical eigenvalues in steps 1 or 3. Moreover, a CCA run including ubiquitous species, i.e. detected in more than 90% of atlas squares, and with the same explanatory variables shows that these species are clustered near the origin of the ordination space, indicating a weak species–environment relationships, and strengthening their omission from the species matrix. Moreover, inclusion of rare species (< 10% of atlas squares) does not increase the total explained variation.

This fairly low percentage of explained variation is not uncommon in ecological studies because species abundance or occurrence data are often very noisy (ter Braak, 1986; ter Braak & Smilauer, 1998; Guisan *et al.*, 1999). Other studies using partial CCA have obtained similar levels of unexplained variation, ranging from about 40% to 85% (see e.g. Borcard *et al.*, 1992; Borcard & Legendre, 1994; Pinelalloul *et al.*, 1995; Hobson *et al.*, 2000).

The unexplained variation could also be due to the species abundance data being recorded at a spatial scale that was probably not appropriate for all species: some species can react more strongly to microhabitat characteristics at a finer spatial resolution than 1 km² (MacFaden & Capen, 2002), e.g. tits or warblers, whilst others, e.g. the black woodpecker (*Dryocopus martius*) or the great grey shrike (*Lanius excubitor*) react to environmental factors at a broader spatial scale. Therefore, the abundance or occurrence of all species within 1 km squares may not be appropriate even if it represents an acceptable compromise (Hustings *et al.*, 1985; Jacob & Paquay, 1992). Considering avian habitat selection as a multi-scale phenomenon would certainly increase the relevance of the results but is technically challenging.

Furthermore, at this spatial scale of investigation wide ecological gradients or spatial structures can not be displayed within the study area because it is quite homogeneous in terms of climate, altitudinal extent and land cover, unlike the areas chosen for other studies (Pasinelli *et al.*, 2001; Storch *et al.*, 2003).

Finally, the fairly low percentage of explained inertia is certainly because of the poorly explained distribution of some species requiring very specific and local habitat

conditions, not summarized by the 59 landscape metrics, e.g. the red-backed shrike (*Lanius collurio*), the European turtle-dove (*Streptopelia turtur*) or the Eurasian cuckoo (*Cuculus canorus*). Species–environment relationships could not have been identified for several species, which tend to obscure the ordination. This suggests that an effort should be made to compute more relevant or additional landscape metrics which are difficult to map. For instance, the 59 metrics characterize habitat quantity but none of them describe habitat quality such as forest succession stages or degree of intensification of cultivation or pastures. An increase in the amount of explained variation of species assemblages would, therefore, be expected by computing additional and more relevant landscape metrics.

Nevertheless, it is impossible to quantify the fraction of currently unexplained variation that could potentially be explained by additional environmental variables (Borcard *et al.*, 1992), because a fair amount of unexplained variation is probably because of spatiotemporal stochasticity. Keeping this in mind, an ordination diagram that explains only a low percentage of species–environment relationships may still be quite informative (ter Braak & Smilauer, 1998; Guisan *et al.*, 1999). Therefore the canonical ordination technique provides an interesting overview of avian community (Guisan *et al.*, 1999), but requires improvement in the case of fine-scale analysis.

CONCLUSION

This study presents a holistic approach to assess the relative roles of environment and space in driving bird species distribution, to display substantial spatial and ecological gradients and to identify relevant spatial and environmental variables for bird assemblage composition using a fine-scale bird atlas data set. Although a traditional canonical ordination without partialling out the spatial effects would have provided similar ecological results, the factoring out procedure used here was essential because ecological phenomena arise from processes that are not exclusive and which can overlap considerably (Legendre, 1993). This is clearly the case in the context of atlas data where individual samples (squares) are not independent, as sites that are close to each other share both environmental conditions (spatially-structured explanatory variables) and species (because of spatial autocorrelation). Such an approach should be adopted in order to provide a general overview of the multivariate data set, prior to specific analyses. Nevertheless, the polynomial function of geographical coordinates of the sampling sites does not seem to be appropriate for detecting fine spatial patterns of communities. Consequently, the method should be improved to detect hidden spatial variations that are unexplained by present ecological gradients so as to extract this spatial component of the variation. In this way it is possible to ensure that the observed species–environment relationships are not because of underlying spatially-structured causes.

ACKNOWLEDGMENTS

We thank all the members of the COA (Centrale Ornithologique Aves, Belgium) who have participated in the mapping schemes. We are grateful to Dirk Maes and Mark Rounsevell for useful discussions and for commenting on a draft of the manuscript. We thank the Institut Géographique National for providing us with 1 : 50 000 maps.

REFERENCES

- Berry, J.K. & McGarigal, K. (1998) *FRAGSTATS*ARC: integrating ARC/INFO with the FRAGSTATS landscape analysis program*. ESRI 1998 User Conference, San Diego, CA, USA.
- Bibby, C.J., Burgess, N.D. & Hill, D.A. (1992) *Bird census techniques*. Academic Press, Londres.
- Böhning-Gaese, K. & Bauer, H.G. (1996) Changes in species abundance, distribution, and diversity in a central European bird community. *Conservation Biology*, **10**, 175–187.
- Borcard, D. & Legendre, P. (1994) Environmental control and spatial component in ecological communities: an example using Oribatid mites (*Acari, Oribatei*). *Environmental and Ecological Statistics*, **1**, 37–53.
- Borcard, D., Legendre, P. & Drapeau, P. (1992) Partialling out the spatial component of ecological variation. *Ecology*, **73**, 1045–1055.
- Boren, J.C., Engle, D.M., Palmer, M.W., Masters, R.E. & Criner, T. (1999) Land use change effects on breeding bird community composition. *Journal of Range Management*, **52**, 420–430.
- ter Braak, C.J.F. (1986) Canonical correspondence analysis: a new eigen vector technique for multivariate direct gradient analysis. *Ecology*, **67**, 1167–1179.
- ter Braak, C.J.F. & Smilauer, P. (1998) *CANOCO reference manual and user's guide to Canoco for windows: software for canonical community ordination (version 4)*. Microcomputer Power, Ithaca, NY, USA.
- Burrough, P.A. & McDonnell, R.A. (1998) *Principles of Geographical Information Systems*. Oxford University Press, Oxford.
- Chamberlain, D.E. & Fuller, R.J. (2000) Local extinctions and changes in species richness of lowland farmland birds in England and Wales in relation to recent changes in agricultural land-use. *Agriculture, Ecosystems & Environment*, **78**, 1–17.
- Chamberlain, D.E., Fuller, R.J., Garthwaite, D.G. & Impey, A.J. (2001) A comparison of farmland bird density and species richness in lowland England between two periods of contrasting agricultural practice. *Bird Study*, **48**, 245–251.
- Coppedge, B.R., Engle, D.M., Masters, R.E. & Gregory, M.S. (2001) Avian response to landscape change in fragmented southern Great Plains grasslands. *Ecological Applications*, **11**, 47–59.
- Dennis, R.L.H. & Hardy, P.B. (1999) Targeting squares for survey: predicting species richness and incidence of species

- for a butterfly atlas. *Global Ecology and Biogeography*, **8**, 443–454.
- Donald, P.F. & Fuller, R.J. (1998) Ornithological atlas data: a review of uses and limitations. *Bird Study*, **45**, 129–145.
- Gates, S., Gibbons, D.W., Lack, P.C. & Fuller, R.J. (1993) Declining farmland bird species: modelling geographical patterns of abundance in Britain. *Large-scale ecology and conservation biology* (ed. by P.J. Edwards, R.M. May and N.R. Webb), pp. 153–177. Blackwell Scientific Publications, Oxford.
- Guisan, A. & Zimmermann, N.E. (2000) Predictive habitat distribution models in ecology. *Ecological Modelling*, **135**, 147–186.
- Guisan, A., Weiss, S.B. & Weiss, A.D. (1999) GLM versus CCA spatial modeling of plant species distribution. *Plant Ecology*, **143**, 107–122.
- Gustafson, E.J. (1998) Quantifying landscape spatial pattern: what is the state of the art? *Ecosystems*, **1**, 143–156.
- Hagemeijer, W.J.M. & Blair, M.J. (1997) *The EBCC atlas of European breeding birds: their distribution and abundance*. Poyser, London/San Diego, 960 pp.
- Hobson, K.A., Kirk, D.A. & Smith, A.R. (2000) A multivariate analysis of breeding bird species of western and central Canadian boreal forests: stand and spatial effects. *Ecoscience*, **7**, 267–280.
- Hustings, M.F.H., Kwak, R.G.M., Opdam, P.F.M. & Reijnen, M.J.S.M. (1985) *Natuurbeheer in Nederland. Deel 3: vogelinventarisatie. Achtergronden, richtlijnen en verlaglegging*. Pudoc, Wageningen & Nederlandse Vereniging tot Bescherming van Vogels, Zeist.
- Jacob, J.P. & Paquay, M. (1992) *Oiseaux nicheurs de Famenne. Atlas de Lesse et Lomme*. Aves, Liège, 360 pp.
- Jongman, R.H.G., ter Braak, C.J.F. & van Tongeren, E.F.R. (1995) *Data analysis in community and landscape ecology*. Cambridge University Press, Cambridge.
- Legendre, P. (1993) Spatial autocorrelation – trouble or new paradigm? *Ecology*, **74**, 1659–1673.
- Legendre, P. & Legendre, L. (1998) *Numerical ecology*. Elsevier Health Sciences B.V., Amsterdam.
- Lobo, J.M. & Martin-Piera, F. (2002) Searching for a predictive model for species richness of Iberian dung beetle based on spatial and environmental variables. *Conservation Biology*, **16**, 158–173.
- Lobo, J.M., Lumaret, J.P. & Jay-Robert, P. (2002) Modelling the species richness distribution of French dung beetles (Coleoptera, Scarabaeidae) and delimiting the predictive capacity of different groups of explanatory variables. *Global Ecology and Biogeography*, **11**, 265–277.
- Luoto, M., Kuussaari, M. & Toivonen, T. (2002) Modelling butterfly distribution based on remote sensing data. *Journal of Biogeography*, **29**, 1027–1037.
- MacFaden, S.W. & Capen, D.E. (2002) Avian habitat relationships at multiple scales in a New England forest. *Forest Science*, **48**, 243–253.
- Maes, D., Gilbert, M., Titeux, N., Goffart, P. & Dennis, R.L.H. (2003) Prediction of butterfly diversity hotspots in Belgium: a comparison of statistically-focused and land use-focused models. *Journal of Biogeography*, **30**, 1907–1920.
- Meot, A., Legendre, P. & Borcard, D. (1998) Partialling out the spatial component of ecological variation: questions and propositions in the linear modelling framework. *Environmental and Ecological Statistics*, **5**, 1–27.
- Monti, D., Legendre, L., Therriault, J.C. & Demers, S. (1996) Horizontal distribution of sea-ice microalgae: environmental control and spatial processes (southeastern Hudson Bay, Canada). *Marine Ecology-Progress Series*, **133**, 229–240.
- Natuhara, Y. & Imai, C. (1996) Spatial structure of avifauna along urban–rural gradients. *Ecological Research*, **11**, 1–9.
- Osborne, P.E. & Tigar, B.J. (1992) Interpreting bird atlas data using logistic-models – a example from Lesotho, Southern Africa. *Journal of Applied Ecology*, **29**, 55–62.
- Palmer, M.W. (1993) Putting things in even better order – the advantages of canonical correspondence-analysis. *Ecology*, **74**, 2215–2230.
- Pasinelli, G., Naef-Daenzer, B., Schmid, H., Keller, V., Holzang, O., Graf, R. & Zbinden, N. (2001) An avifaunal zonation of Switzerland and its relation to environmental conditions. *Global Ecology and Biogeography*, **10**, 261–274.
- Pearson, R.G. & Dawson, T.P. (2003) Predicting the impacts of climate change on the distribution of species: are bioclimate envelope models useful? *Global Ecology and Biogeography*, **12**, 361–371.
- Pinelalloul, B., Niyonsenga, T. & Legendre, P. (1995) Spatial and environmental components of fresh-water zooplankton structure. *Ecoscience*, **2**, 1–19.
- Schmid, H., Luder, R., Naef-Daenzer, B., Graf, R. & Zbinden, N. (1998) *Atlas des oiseaux nicheurs de Suisse 1993–1996*. Station ornithologique suisse, Sempach, Suisse, 574 pp.
- Siriwardena, G.M., Crick, H.Q.P., Baillie, S.R. & Wilson, J.D. (2000) Agricultural land-use and the spatial distribution of granivorous lowland farmland birds. *Ecography*, **23**, 702–719.
- Storch, D., Konvicka, M., Benes, J., Martinkova, J. & Gaston, K.J. (2003) Distribution patterns in butterflies and birds of the Czech Republic: separating effects of habitat and geographical position. *Journal of Biogeography*, **30**, 1195–1205.
- Telfer, M.G., Preston, C.D. & Rothery, P. (2002) A general method for measuring relative change in range size from biological atlas data. *Biological Conservation*, **107**, 99–109.
- Timothy, J. & Sharrock, R. (1974) Minutes of the second meeting of the European Ornithological Committee. *Acta Ornithologica*, **14**, 404–411.
- Tobalske, C. & Tobalske, B.W. (1999) Using atlas data to model the distribution of woodpecker species in the Jura, France. *Condor*, **101**, 472–483.
- Trzcinski, M.K., Fahrig, L. & Merriam, G. (1999) Independent effects of forest cover and fragmentation on the distribution of forest breeding birds. *Ecological Applications*, **9**, 586–593.

Vanhinsbergh, D.P. & Chamberlain, D.E. (2001) Habitat associations of breeding Meadow Pipits *Anthus pratensis* in the British uplands. *Bird Study*, **48**, 159–172.

Woodward, A.A., Fink, A.D. & Thompson, F.R. (2001) Edge effects and ecological traps: effects on shrubland birds in Missouri. *Journal of Wildlife Management*, **65**, 668–675.

BIOSKETCHES

Nicolas Titeux is a PhD student at the Université catholique de Louvain and works on landscape ecology, especially on the relationships between landscape structure and bird communities.

Marc Dufrière is a research doctor and is responsible for coordinating the inventory and monitoring of species populations and nature conservation programmes in southern Belgium.

Jean-Paul Jacob is an industrial engineer and coordinates breeding bird atlases and monitoring schemes (birds, reptiles and amphibians) in southern Belgium and the Brussels area.

Marc Paquay is a forester and field naturalist observing regularly in the area of the *Atlas de Lesse et Lomme*.

Pierre Defourny is Professor in geomatics and leads the research team in remote sensing and GIS applied to environmental sciences and land-use planning at the Université catholique de Louvain.

Copyright of Journal of Biogeography is the property of Blackwell Publishing Limited and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.