Ardeola 53(2), 2006, 213-227

BIAS IN AVIAN SAMPLING EFFORT DUE TO HUMAN PREFERENCES: AN ANALYSIS WITH CATALONIAN BIRDS (1900 - 2002)

Xavier FERRER* [1], Luis M. CARRASCAL**, Oscar GORDO*** and Joan PINO****

SUMMARY.—*Bias in avian sampling effort due to human preferences: an analysis with Catalonian birds (1900 - 2002)*.

Aims: To investigate the sources of spatial bias in the bird records of Catalonia from 1900 to 2002, with the aim of providing generalized recommendations for using other databases, and setting up broad inventory projects. The paper examines the influence of environmental variables, human distribution and ornithological preferences of birdwatchers on past avian sampling effort in Catalonia, a contrasting region in the north east of Spain.

Location: Catalonia (Spain).

Methods: The relationship between time (field days) devoted to sampling birds in 10 x 10 Km UTM squares (from the records of VertebraCat database, across 5 study periods from 1900 to 2002) and a set of environmental, human distribution and bird species richness variables was analysed. These relationships were analyzed by means of Partial Least Squares Regression (PLSR) and multiple regression analyses.

Results: A partial least squares regression analysis accounted for 39.4 % of the spatial variation in time devoted to avian sampling per UTM 10 x 10 km square in the most recent (1983 - 2002) period. The major pattern shows that large visit frequencies were mainly associated with coastal areas with a large cover of wetlands, marshes and sand dunes, high human population density, dense transportation, and also high covers of irrigated croplands, urban and industrial environments. Another group of highly visited UTM squares was also covered with a large surface of protected areas, mostly located in mountainous Catalonian sectors, as opposed to those UTM squares in lowland areas mainly covered by non-irrigated extensive croplands. Catalonian ornithologists and birdwatchers also showed an uneven preference for different groups of bird species, as they mainly relied on migratory and endangered species.

Conclusions: The illustrated biased pattern of ornithological field work in Catalonia casts doubts about the usefulness of biodiversity indices obtainable from databases of observation records without a random-stratified sampling approach. However, these problems might be overcome by including a variable of sampling bias such as the number of records or days of field work.

Key words: Birds, birdwatcher preferences, Catalonia, environmental characteristics, sampling biases.

* Departamento de Biologia Animal (Vertebrats), Facultat de Biologia, Universitat de Barcelona. Diagonal 645, Barcelona, E-08028 Spain.

** Departamento de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Naturales (CSIC), José Gutiérrez Abascal, 2, Madrid, E-28006 Spain.

*** Departamento de Ecología Evolutiva, Museo Nacional de Ciencias Naturales (CSIC), José Gutiérrez Abascal, 2, Madrid, E-28006 Spain.

**** Center for Ecological Research and Forestry Applications (CREAF), Universitat Autònoma de Barcelona, Bellaterra, Barcelona, E-08193 Spain.

1 Corresponding author: xferrer@ub.edu

RESUMEN.—*Sesgo en el esfuerzo de muestreo ornitológico debido a las preferencias humanas: un análisis con la avifauna de Cataluña (1900 - 2002).*

Objetivos: Investigar la influencia de variables ambientales, de distribución humana y de preferencias ornitológicas sobre el muestreo histórico de la avifauna de Cataluña desde 1900 hasta 2002, y proponer recomendaciones generales para el uso de otras bases de datos y llevar a cabo inventarios a escala regional.

Localidad: Cataluña (España).

Métodos: Se determinó el tiempo (días de campo) dedicado al muestreo de aves en cuadrículas UTM de 10 x 10 km a partir de los registros de VertebraCat, distinguiéndose cinco períodos de estudio entre 1900 y 2002. Para cada UTM se obtuvieron también un conjunto de variables ambientales, geográficas y de distribución humana, así como la riqueza de especies de aves de diversos grupos fenológicos y de conservación. La asociación entre la intensidad de prospección y las variables ambientales se analizó mediante regresión por mínimos cuadrados parciales para el período comprendido entre 1983 y 2002. Los residuos de esta regresión se correlacionaron con las variables de riqueza de distintos grupos de aves.

Resultados: Un 39,4 % del total de la varianza en el tiempo destinado al muestreo ornitológico por UTM de 10 x 10 km para el período 1983 - 2002 se pudo explicar por las variables ambientales y geográficas. Las frecuencias elevadas de prospección ornitológica se asociaron con áreas cercanas a la costa con una elevada cobertura de marismas, zonas húmedas, dunas costeras y cultivos de regadío, una alta densidad de población humana y cobertura de zonas industriales, y una densa red de comunicaciones viarias. También fueron muy visitadas algunas UTM con una elevada proporción del área incluida en espacios protegidos y localizadas principalmente en sectores de montaña, en contra de lo observado en UTM de zonas bajas y cubiertas por cultivos extensivos de secano. Los ornitólogos y observadores de aves de Cataluña mostraron, además, una marcada preferencia por la observación de aves migradoras y de especies amenazadas.

Conclusiones: Este sesgo en el muestreo ornitológico de Cataluña introduce dudas sobre la utilidad de los índices de biodiversidad extraídos de bases de datos construidas a partir de registros sin un muestreo aleatorio-estratificado. Sin embargo, estos problemas pueden solucionarse introduciendo en los análisis variables que recojan el posible sesgo muestral, como serían el número de registros o de días de campo.

Palabras clave: Aves, preferencias ornitológicas, Cataluña, características ambientales, sesgos de muestreo.

INTRODUCTION

Area conservation priorities are based upon an exhaustive knowledge of biodiversity patterns in large geographic areas, as this information is needed to identify hot-spots compared to other less outstanding areas. This approach requires a thorough exploration of species present throughout a region, implying a regular, stratified or random survey (i.e., uniform sampling) of all areas. Insufficient time invested in sampling underestimates the number of species present in previously non-established important biodiversity areas or hot-spots, particularly if hard-to-detect or low density organisms are involved (Williams *et al.*, 2002). This could pose an important problem if recording effort does not reach the asymptote of the relationship between number of species and sampling time. This phenomenon is not a main concern today in scientific research in topics like landscape ecology or biogeography where data are obtained according to sampling protocols (but see Kodric-Brown and Brown, 1993). Nevertheless, it may have been a problem in the past or could constitute one in actual biodiversity databases which benefit from scattered information gathered through the participation of many observers that do not follow clearly defined methodological standards over long time periods (see Nelson *et al.*, 1990, for Brazilian botanical collections; Reddy and Dávalos, 2003, for African birds; Parnell *et al.*, 2003, for Thailand plants). A number of bird surveys involving birdwatchers and field ornithologists have analysed the role of sampling effort and spatial distribution on the number of species observed (Sharrock, 1966; Fraser, 1997; Field *et al.*, 2002; Estrada *et al.*, 2005). The majority of them failed to demonstrate that there is a lack of association between biodiversity indices and inventory effort.

Several papers have shown significant correlations between human population density and several measurements of species richness or abundance of different taxa in different parts of the world (Fjeldså and Rahbek, 1998; Balmford *et al.*, 2001a, 2001b; Dobson *et al.*, 2001; Araújo, 2003, not significant for birds; Chown *et al.*, 2003; Gaston and Evans, 2004; Khun *et al.*, 2004). Moreover, important areas for biodiversity, selected with presence data or population sizes of endangered species, have consistently more people and transportation infrastructures than expected by chance (Araújo *et al.*, 2002; Carrascal and Lobo, 2003). These results, *a priori* surprising, are explained considering ecological, historical and sampling-limitation arguments. Despite some historical factors, some city areas were preferentially located in pre-existing biodiversity hotspots of high productivity or glacier refugia both for humans and plants or animals (Fjeldså and Rahbek, 1998; Araújo, 2003; Gaston and Evans, 2004). Nevertheless, these associations could also emerge as a consequence of sampling bias in biodiversity inventories. Some of the most important candidates to bias biodiversity estimations over large regions are the distribution of human population centres and the network of roads and railways (Freitag *et al.*, 1998; Reddy and Dávalos, 2003). Due to logistic difficulties or time constraints (e.g., available time for amateurs), inventories are expected to be carried out near dense population areas and sampling frequency should be determined with the availability of transportation networks penetrating the territory, except in cases of well known biodiversity hotspots (Şekercioğlu, 2002).

This paper examines the influence of environmental variables, human distribution and ornithological preferences of birdwatchers on avian sampling effort carried out in the recent past (twentieth century) in a very contrasting region of the north east of the Spain (Catalonia). This approach does not cast doubts about the actual knowledge about bird distribution in Catalonia, because there is a thorough and exhaustive Atlas of the breeding birds (Estrada *et al.*, 2005) that has been aware of the problem and ensured a sampling effort enough to reach the asymptotic species richness of all areas in this region.

MATERIALS AND METHODS

Study area

Catalonia is a region of 31,900 km^2 located in the Mediterranean north-eastern coast of Spain (Fig. 1), in the boundary between the Eurosiberian and the Mediterranean regions. This region includes Eurosiberian, alpine and Mediterranean biogeographical sectors, a large variety of habitats, and spans from coastal Mediterranean areas to high mountains in the Pyrenees. Its climate ranges between dry-hot to very cold and wet, human population density varies enormously from the densely populated Barcelona to several nearly uninhabited areas. The Pyrenees determine a sharp topographic-climatic gradient, from Mediterranean or Eurosiberian-type biomes to Subalpine and Alpine types. The region exhibits highly variable environmental conditions considering the relief (sea level to 3143 m in the Pyrenees) and geographical situation, having Mediterranean, Atlantic and even Saharan climatological influences (Ninyerola *et al.*, 2000). Rainfall decreases, and average temperature increases, southwards. There is also a climate

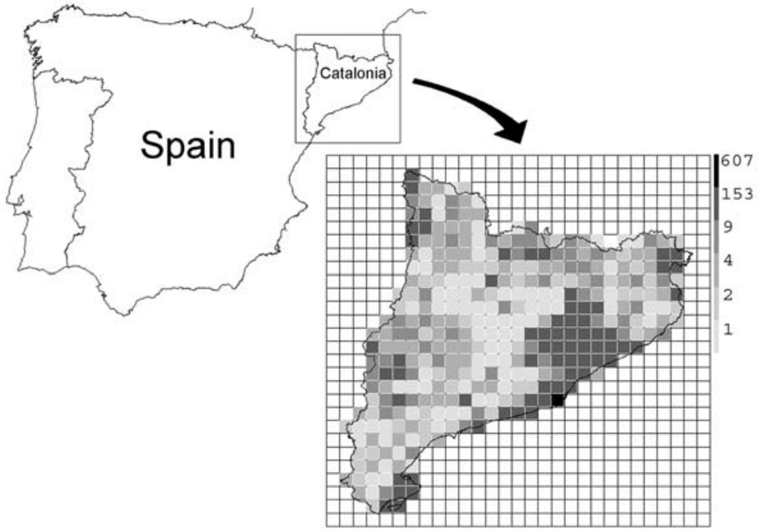

FIG. 1.—Sampling effort (number of field work days) per UTM 10 x 10 km square for the period 1983 - 2002 in Catalonia according to VertebraCat database. UTM with less than 25 % of Catalonian land cover were not considered.
[Esfuerzo de muestreo por cuadrícula UTM de 10 x 10 km para el período 1983 - 2002 en Cataluña según la base de datos VertebraCat. No se han considerado las cuadrículas UTM con menos de un 25 % de su cobertura en Cataluña.]

gradient from the moist temperature climate of the eastern coast, to the inland dry, continental, climate. This large environmental heterogeneity determines a high landscape and botanical diversity (Pausas *et al*., 2003). The central coast of Catalonia corresponds to the city of Barcelona and its surrounding area, which is thought to be one of the most industrialized and populated areas on the northern Mediterranean coast (MMAMB, 1995).

The ornithological knowledge of this region is very good, due to a large number of birdwatchers (2000 - 2500 in 2002) that have produced two well-performed breeding bird atlases (one in 1980-82 -Muntaner *et al*., 1984- and another in 1999 - 2002 -Estrada *et al*., 2005). The total number of bird species observed in Catalonia is 390 homologated species (J. Clavell, *pers. comm.*). Breeding birds account for 231 species (78.3 % of total Iberian breeding bird species). Migrant bird species are important because the Catalonian coast is on the NE-SW leading line axis for Central Europe migratory birds (Moreau, 1956). The number of alien species (i.e., those introduced by humans, either deliberately or not) is also large, accounting for 14.7 % of observed species in the region.

Environmental variables

A number of variables concerning geographical position, climate, topography, geology, land-use, and human settlement of Catalonia (i.e., independent variables) were obtained for the study area. These variables were selected considering their important role in avian biogeographical patterns in the Iberian Peninsula (González *et al.*, 1990; Carrascal *et al.*, 1993; Suárez-Seoane et al., 2002; Carrascal and Lobo, 2003; Carrascal and Díaz, 2003; Brotons *et al.*, 2005). All the GIS procedures involving the set up of these variables were performed using

MiraMon (Pons, 2002). Because of the limitations of spatial resolution imposed by the response variable (i.e., number of days of ornithological research), all these variables were calculated for the UTM 10 km grid.

Climate variables were calculated from the climatic models of Catalonia, set up by Ninyerola *et al*. (2000) at a spatial resolution of 180 m, using the existing network of meteorological stations and Digital Elevation Models (DEM). We calculated the mean values per UTM square of mean annual temperature, annual rainfall and mean annual solar radiation. Topographic variables were obtained from the official DEM (30 m pixel size) and the 1:50.000 topographic map of Catalonia, both generated by the Cartographic Institute of Catalonia (ICC). The mean and range (maximum and minimum) elevations per UTM square were calculated from the DEM. The topographic map was also used to calculate the proportion of Catalonian land and the length of coastline included in each UTM square, using GIS methods of coverage combination. UTM squares with less than 25 % of land area within Catalonian region were rejected, which provided a set of 337 UTM in final analyses.

Landscape variables were set up using the Land Use Map of Catalonia, generated by the ICC in 1997 by processing multitemporal data captured by the Thematic Mapper sensor of the Landsat satellite. The resulting 30 m pixel grid, which accounts for 21 land cover classes, has been generated every 5 years from 1987 using a standard procedure (Viñas and Baulies, 1995). This map was used to calculate the percentage of each UTM square corresponding to each land cover class, and the land cover diversity (using the Shannon index) of each UTM based on these classes.

Frequency of fires was also included as a variable of landscape disturbance. The mean number of fires per UTM cell in the period 1973 - 1997 was calculated from the periodicity fire map of Díaz-Delgado and Pons (2001). This map considers all fires larger than 30 Ha that have affected Catalonia at a pixel size of 30 m.

Several human settlement variables were also included in the analyses. Mean population density was considered to evaluate the extent of metropolitan occupation. The official urban settlement map of Catalonia (http://www.gencat.net/mediamb/sig/sig-a.htm) that includes the most recent (1999) population census was used to calculate the mean population density per 10 x 10 km UTM square. Miramon distance algorithms were used to calculate the mean distance to the main roads and motorways for each UTM square, applied on a selection of the motorways and major roads from the 1:50.000 topographic map of Catalonia.

Finally, the effect of natural protected areas for each UTM was included in the study considering the proportion of each UTM included into areas of special protection (national and natural parks and reserves; http://mediambient.gencat.net/cat/el_medi/parcs_de_catalunya/).

Sampling measurements

Bird sampling effort was measured as the number of days reported in VertebraCat database for each 10 x 10 km UTM square. VertebraCat is a biodiversity database on vertebrate status and distribution within Catalonia, based on published information up to 2002 (http://biodiver.bio.ub.es/biocat/homepage.html). The aim of this dataset is to compile biogeographic information about biodiversity of fishes, amphibians, reptiles, birds and mammals in Catalonia, to show, not only species distributions, but also their condition of conservation, degree of threat, rarity or endemic level. The VertebraCat is based on materials from various sources (more than 1500 bibliographical references). The main part of the database has been computerized from Spanish bibliography (articles, local monographs, doctoral thesis, and books). This information has been checked by X. F., excluding faunistic quotations belonging to the archives of the *Institut Català d'Ornitologia* since 1996, and the records used in the recent Catalonian breeding bird Atlas (Estrada *et al*., 2005) All records/identifications available (both from professional and amateur observers) have been verified and entered in the database. In order to analyse changes in temporal patterns of sampling five time intervals were defined, from 1900 to 2002: 1900 - 1962, 1963 - 1972, 1973 - 1982, 1983 - 1992 and 1993 - 2002.

To analyze the environmental and human distribution determinants of bird sampling in Catalonia the number of field days in the last two decades (1983 - 2002) was utilised because this is the most intensely prospected period and there is a nearly total overlap with environmental information derived from GIS databases. The total number of field days in this period was 3,620.

Variables of bird species richness

For each UTM 10 x 10 km square the checklist of bird species reported in VertebraCat in the last 20 years (1983 - 2002) was obtained the bird species in Catalonia were classified according to their origin, phenology and conservation status according to the available literature (Tucker and Heath, 1994; Copete, 1997; Clavell, 2002). First, species were separated into two groups considering whether they were recently introduced by man (namely alien or exotic species) or are native ones. Native species were classified into five phenological categories: accidental, migratory, wintering, aestival and sedentary. Species were included in only one phenological category. Accidental species are non-human introduced species that occur occasionally from the Catalan Rarity Checklist (Copete, 1997). Migratory birds include those species that only cross Catalonia in their northward and/or southward migration routes (a few species in this category have very scarce wintering or breeding populations in the study area, lower than five pairs per year). Wintering species are those that are exclusively present in the study region in winter time (November - February). Estival birds include species that are regular breeders in Catalonia (April - July), but leave the region in winter time. Finally, sedentary species included those present throughout the year in the study region. The conservation status of the native species was assessed considering only the most endangered species in Europe (SPEC categories 1 and 2 in Tucker and Heath, 1994).

In summary, nine variables related to richness of birds in each UTM for these 20 years were generated as: total number, alien, accidental, migratory, wintering, breeding, sedentary, highly endangered species, and breeding species in the most recent sampling date (2002).

Statistical analyses

Associations between environmental and geographical variables and time devoted to bird inventories were analyzed by means of Partial Least Squares Regression (PLSR). This is an extension of the multiple linear regression analysis where the effects of several predictors of any type on one dependent (response) variable can be analyzed. In partial least squares regression, associations are established with factors extracted from predictor variables that maximize the explained variance in the dependent variable. These factors are defined as a lineal combination of original variables, so the original multidimensionality is reduced to a lower number of factors to detect structure in the relationships between predictor variables, and between these factors and the response variable. The extracted factors are orthogonal (i.e., independent of each other) and they account for the successive lower proportions of original variance. For more details on this statistical exploratory technique, see StatSoft (2001) and Tobias (1995). The meaning of PLSR components was derived from the correlations with original variables significant at $P < 0.001$.

Geographical position variables (mean latitude and longitude of each 10 x 10 km UTM square) were considered in the analyses to define geographical gradients and to control for spatial non-independence (i.e., autocorrelation). The influence of the spatial location and proximity of the UTM 10 x 10 km sample units were included in the analyses by means of a two-order polynomial of latitude and longitude, thus performing a trend surface analysis (Legendre, 1993). The residuals of the PLSR model did not show a clear spatial autocorrelation pattern (1.7 % of variation in the residuals was explained by a two-order polynomial of latitude and longitude, $P > 0.05$).

Residuals of this PLSR model were also correlated with several measurements of avian species richness, to test if there were ornithological preferences for different groups according to phenology and conservation status.

Other statistical tests (correlations and multiple regression analysis) are presented when used.

RESULTS

Time devoted to avian sampling was associated across study periods (average of 10 correlation coefficients among the 5 periods: 0.519, n = 337 UTM squares; range: 0.107 - 0.800). The highest correlation was obtained between 1983 - 1992 and 1993 - 2002 periods ($r = 0.800$, $P < 0.001$).

Table 1 shows the results of partial least squares regression analysis with the spatial variation in time devoted to avian sampling per UTM 10 x 10 km square in the most recent 1983 - 2002 period, that makes up the large amount of inventories. Three highly significant components account for 39.4 % of spatial variation in time invested to gather avian information (see Table 1 for intensity of association and significance of these components with sampling investment). The first one (29.6 % of the original variance) associates large visit frequencies to those coastal UTM squares with high human population density, densely covered with any kind of motor roads, urban and industrial environments, and also with a large cover of marshes, coastal sand dunes and irrigation croplands (some of them included within protected areas; Fig. 2). The second component (6.2 % of variance) related highly visited UTM squares to those covered with a large surface of protected areas, and located in mountainous Catalonian sectors (positively associated with altitudinal range, average altitude and cover of subalpine grasslands, and negatively with mean annual temperature). They are opposed to those UTM squares in lowland areas mainly covered by non-irrigated extensive croplands and vineyards (with the exception of localities with large cover of irrigated croplands and marshes that also reached high levels of time devoted to field work). Finally, the third component (3.7 % of original variance) associates a low avian sampling to those UTM squares mainly located at the south-western part of the study area (negatively correlated with longitude-latitude terms), covered with a large surface of vineyards (as opposed, again, to the highly inventoried areas with high landscape diversity, covered by large surface of protected areas, marshlands and irrigated croplands, near large population centers). In other words, UTM squares that were more intensively sampled than expected on the basis of their population density, accessibility, and interest for ecotourism, had higher numbers of endangered, breeding and/or sedentary species.

The residuals of the PLSR model were significantly and positively related to avian species richness, considering different groups according to phenology and conservation status (Table 2). The highest correlations were obtained for endangered species (European SPEC's 1 and 2), breeding (aestival) and sedentary species, while the lowest correlations were found in accidental and alien (introduced)

TABLE 1

Results of the partial least squares regression analyses analyzing the spatial variation in the number of days devoted to avian sampling in the period 1983 - 2002 in 337 UTM 10 x 10 km squares in Catalonia. The figures shown are Pearson correlation coefficients between the predictor variables and three components (COMP's 1 to 3). Variables highly associated ($P < 0.001$) with the components are in bold type. *r*: Pearson correlation coefficients between avian sampling (response variable) and each component. R^2 (%): percentage of variance in the spatial variation in avian sampling accounted for each component.
*[Resultados del análisis de la regresión por mínimos cuadrados parciales que explora la variación espacial en el número de días destinados a muestreo ornitológico en el período 1983 - 2002 para 337 cuadrados UTM de 10 x 10 km en Cataluña. Los números corresponden a correlaciones de Pearson entre las variables predictoras y tres componentes (COMP 1 a 3). Las variables muy asociadas (*P *< 0.001) con las componentes aparecen en negrita.* r*: coeficiente de correlación de Pearson entre el muestreo ornitológico (variable respuesta) y cada componente.* R^2 *(%): porcentaje de la varianza en la variación espacial del muestreo ornitológico explicado por cada componente.]*

Variable	**COMP 1**	**COMP 2**	**COMP 3**
Longitude (UTM X) *[Longitud (UTM X)]*	0.16	-0.10	**-0.21**
Latitude (UTM Y) *[Latitud (UTM Y)]*	-0.04	0.08	-0.15
X^2	0.16	-0.10	**-0.21**
Y^2	-0.04	0.08	-0.15
X * Y	0.15	-0.09	**-0.22**
Altitude range (max-min) *[Rango altitudinal (máx-mín)]*	-0.03	**0.28**	0.04
Mean altitude *[Altitud media]*	-0.07	0.26	0.02
Percentage of land *[% de tierra]*	-0.16	0.07	**0.24**
Coastal length *[Longitud de la costa]*	**0.30**	0.05	-0.02
Mean solar radiation *[Radiación solar media]*	-0.13	-0.15	0.05
Mean annual temperature (ºC) *[Temperatura media annual (ºC)]*	0.07	**-0.25**	0.00
Annual rainfall (mm) *[Precipitación annual (mm)]*	0.04	0.16	-0.16
Fire frequency *[Frecuencia de fuegos]*	-0.05	-0.08	0.00
% of area included into special protected areas *[% del área incluida en areas de especial protección]*	**0.18**	**0.43**	**0.38**
Population density (inhab./ha) *[Densidad poblacional (hab./Ha)]*	**0.26**	0.07	0.02
Mean distance to roads (m) *[Distancia media a carreteras (m)]*	-0.05	0.05	-0.12
Land cover diversity (Shannon index) *[Diversidad de cobertura vegetal (Índice de Shannon)]*	0.13	0.13	**0.24**
Continental water (% of cover) *[Cobertura de aguas continentales (%)]*	0.17	0.18	0.10
Glaciers and perpetual snow (% of cover) *[Cobertura de glaciares y nieves perpetuas (%)]*	0.03	0.09	-0.09
Roads and rail networks (% of cover) *[Cobertura de carreteras y vias de tren (%)]*	**0.29**	0.00	0.00
Discontinuous urban fabric *[Fábricas urbanas discontinuas]*	**0.31**	0.06	0.09
Continuous urban fabric (% of cover) *[Cobertura de fábricas urbanas continuas (%)]*	**0.37**	0.14	0.20
Industrial or commercial units (% of cover) *[Cobertura de unidades comerciales o industriales (%)]*	**0.26**	-0.07	-0.16

TABLE 1 (CONT.)

Variable	COMP 1	COMP 2	COMP 3
Non-irrigated arable crops (% of cover) *[Cobertura de campos de cultivo de secano (%)]*	-0.13	**-0.23**	-0.11
Irrigated arable crops (% of cover) *[Cobertura de campos de cultivo de regadio (%)]*	**0.20**	**0.30**	**0.47**
Non-irrigated fruit trees (% of cover) *[Cobertura de frutales de secano (%)]*	-0.13	-0.06	0.15
Irrigated orchards (% of cover) *[Cobertura de frutales de regadio]*	-0.07	-0.16	-0.11
Vineyards (% of cover) *[Cobertura de viñedos (%)]*	-0.14	**-0.28**	**-0.24**
Supra-forest grasslands (% of cover) *[Cobertura de herbazales(%)]*	0.06	**0.25**	0.04
Scrublands (% of cover) *[Cobertura arbustiva (%)]*	-0.15	0.01	-0.05
Sclerophyllous forests (% of cover) *[Cobertura de bosques esclerófilos (%)]*	-0.02	-0.09	-0.17
Deciduous forests (% of cover) *[Cobertura de bosques caducifolios (%)]*	0.01	0.10	0.02
Coniferous forests (% of cover) *[Cobertura de bosques de coníferas (%)]*	-0.01	0.07	-0.04
Wetlands (% of cover) *[Cobertura de humedales (%)]*	**0.28**	**0.27**	**0.23**
Dunes and beaches (% of cover) *[Cobertura de dunas y playas (%)]*	**0.20**	0.06	-0.09
r	0.54	0.30	0.24
R^2 (%)	29.6	6.2	3.7
P	< 0.001	< 0.001	< 0.001

species. Migratory and wintering species attained correlation coefficients in between these two extremes.

By means of a multiple regression analysis, the groups significantly associated with residual variability in the PLSR analysis are migratory and endangered species (R^2 = 39.6 %, $F_{7,329}$ = 30.76, P < 0.001; Fig. 3). Therefore, the explained variability in time devoted to avian inventories across UTM squares in Catalonia was 63.4 %; the addition of 39.4 % (derived from the PLSR analysis) plus 24.0 % (39.6 % out of the unexplained 60.6 % residual variance).

DISCUSSION

Biodiversity patterns (i.e., number of species in its simplest formulation) should be checked for potential biases that may result from variations in sampling effort (Williams *et al*., 2002), such as those related to spatial dispersion of research institutions or main centres of residency of observers, and the availability of transportation networks favouring an homogeneous inventory of the territory. The results of this paper point out that even in a relatively small region with high density of bird observers (approximately 70 per 1000 km^2), and without striking transportation constraints, avian sampling effort was not determined at random within the Catalonian territory up to the year 2000. Spatial features of the territory explained approximately 40 % of the observed variability in time invested to avian field work throughout Catalonian region, violating the principles of random or stratified sampling. Reddy and Dávalos (2003) found that the location and intensity of collecting passerine birds have historically been heavily influenced by accessibility: sampling localities showed dense,

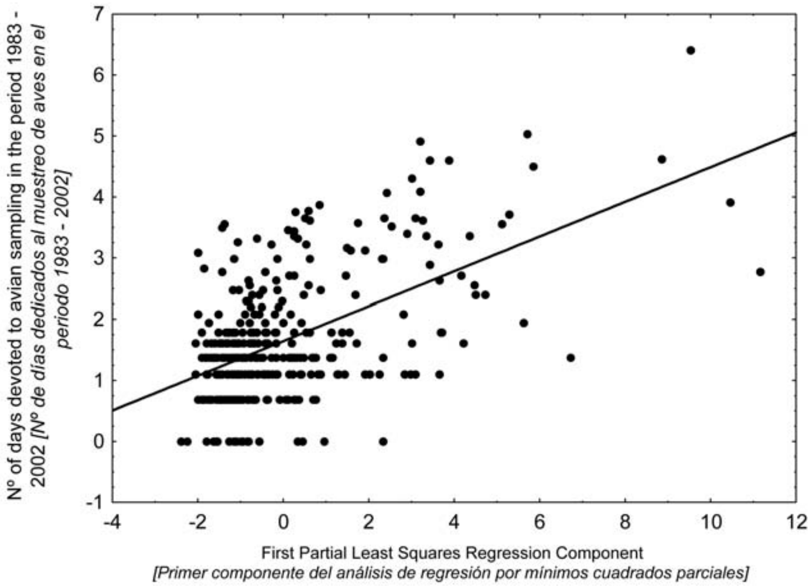

FIG. 2.—Relationship between the number of days devoted to avian sampling in each UTM 10 x 10 km square (in ln) in the period 1983 - 2002 and the first component of the partial least squares regression analysis (see Table 1). Sample size = 337 UTMs; $r = 0.544$, $P < 0.001$.
[Relación entre el número de días destinados a muestreo ornitológico por cuadrícula UTM de 10 x 10 km (en ln) en el período 1983 - 2002 y la primera componente del análisis de la regresión por mínimos cuadrados parciales (ver Tabla 1). Tamaño de la muestra = 337 UTMs; r *= 0,544,* P *< 0,001.]*

significant aggregation around city limits, and along rivers and roads (see also Nelson *et al.*, 1990, for botanical collections). Moreover, there were also striking preferences for some groups of species considering their population levels, phenological status or main habitat preferences in the region, as migrant and endangered bird species were emphasized (explaining nearly 25 % of variation in time devoted to avian inventories).

The positive effect of urban area on avian inventories is easily explained considering the strong association with the distribution of birdwatchers' and ornithologists' residence (correlation between the number of bird observers living per UTM square with urban area: $r = 0.904$, $n = 337$, $P < 0.001$). Although several authors have found positive relationships between human population density and vertebrate or plant diversity, and have interpreted them on ecological and historical backgrounds (Fjeldså and Rahbek, 1998; Balmford *et al.*, 2001a, 2001b; Dobson *et al.*, 2001; Chown *et al.*, 2003; Gaston and Evans, 2004; Khun *et al.*, 2004; but see Araujo, 2003), our results indicate that this association can also be explained considering that observers mainly stay within a short distance of inhabited areas or research facilities, generating the observed pattern (see Nelson *et al.*, 1990; Reddy and Dávalos, 2003).

Wetland environments have attracted considerable attention by ornithologists to document bird distribution patterns, and by birdwatchers as areas that maximize the number of different species observed. The positive selection for marshland habitats in Catalonia also has an historical basis because the wetlands were the most endangered habitats in Spain during the first half of the twentieth century (Fernández, 2004). The positive effect of coastal

TABLE 2

Pearson correlation coefficients (*r*) between residuals of the partial least squares regression analysis (PLSR, Table 1) and the number of species per UTM 10 x 10 km square in six ornithological groups according to phenology and conservation status (all correlation coefficients are highly significant, $P < 0.001$). The results of the multiple regression analysis relating these groups to PLSR residuals are also shown (beta: standardized regression coefficient).
*[Coeficientes de correlación de Pearson (*r*) entre los residuos del análisis de la regresión por mínimos cuadrados parciales (PLSR, Tabla 1) y el número de especies por UTM de 10 x 10 km para 6 grupos ornitológicos según la fenología y el estado de conservación (todas las correlaciones son muy significativas,* P *< 0,001). Se proporcionan también los resultados de análisis de la regresión múltiple que relacionan estos grupos con los residuos del PLSR (beta: coeficiente de regresión parcial estandarizado).]*

	r	**Beta**	***P***
Accidental species *[Especies accidentals]*	0.277	-0.21	0.073
Alien (introduced) species *[Especies introducidas]*	0.258	-0.10	0.141
Migratory species *[Especies migratorias]*	0.398	0.40	0.002
Wintering species *[Especies invernantes]*	0.457	-0.08	0.435
Breeding species *[Especies reproductoras]*	0.596	0.05	0.823
Sedentary species *[Especies sedentarias]*	0.575	0.19	0.239
Endangered species (SPEC's 1+2) *[Especies en peligro (SPEC 1+2)]*	0.602	0.37	0.015

environments on avian inventories can be understood in the Mediterranean sector of the Iberian Peninsula considering their location on the migratory routes of western Palaearctic birds (Moreau, 1956; Bernis, 1966; Pérez-Tris and Santos, 2004) and the addition of some bird species within the Catalonian avifauna especially attractive for birders (e.g., *Larus audouinii, Larus genei*, five species of the genus *Sterna*). The important effect of irrigated croplands (mainly rice fields) in determining inventory patterns is possibly the consequence of a spatial coincidence of these environments with other wetland habitats located in estuaries and deltas, and because they are an important habitat for waders and waterfowl as wintering or stopover areas in the Mediterranean (Ferrer and Martínez-Vilalta, 1987; Fasola and Ruíz, 1997).

The avoidance pattern for vineyards and extensive, non-irrigated, cereal croplands could be explained considering their general lack of habitat suitability for birds due to agricultural practices, the low species richness in these formations, and because they are inhabited by generalists that are also found in other open habitats (Martínez-Vilalta, 1986; Suárez, 2004). These facts made them non-attractive environments for birdwatchers and ornithologists in the past, although this tendency has been reverted in the last years because steppe-land birds have undergone population declines throughout Europe (Tucker and Heath, 1996), thus being the focus of more attention.

Mountainous and Subalpine Catalonian sectors, mainly located at the Pyrenees, were frequently visited by bird observers. This fact could be explained considering the attractiveness of alpine landscapes for people in dry Mediterranean areas (Múgica, 1994). These sectors are covered by a large extension of protected areas and have a large number of distinctive, exclusive, alpine avifauna. The attraction of birdwatchers to protected areas has also been found by Reddy and Dávalos (2003) in African passerines. Apart from these conser-

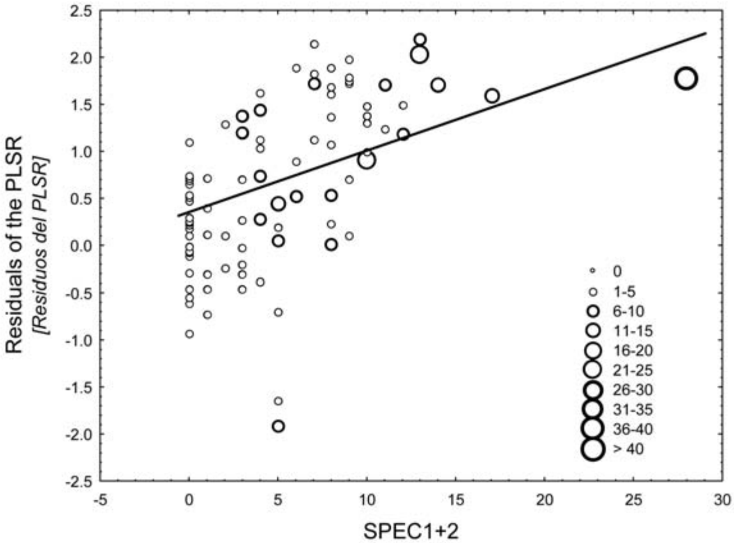

FIG. 3.—Relationship between the residuals of the partial least squares regression model in Table 1 (residuals of PLSR) analyzing the spatial variation in the time devoted to avian sampling, and the number of endangered species (SPEC1+2; species categorized as SPEC 1 or SPEC 2 by Tucker and Heath, 1994) and migratory species (represented by the size of the points) per UTM 10 x 10 km square. Sample size is 337 UTMs.
[Relación entre los residuos del análisis de la regresión por mínimos cuadrados parciales de la Tabla 1 (residuos del PLSR) que analizan la variación espacial en el tiempo destinado al muestreo ornitológico, y el número de especies amenazadas (SPEC1+2; especies clasificadas como SPEC 1 o SPEC 2 por Tucker and Heath, 1994) y especies migradoras (representadas por el tamaño de los puntos) por UTM de 10 x 10 km. Tamaño de muestra = 337 UTMs.]

vation and biodiversity facts, there must be other subtle preferences for these areas.

Apart from these geographical, environmental and human population determinants, some ornithological preferences were related to avian inventories. Catalonian ornithologists preferred areas where the observation of endangered or migrant species was more likely. On the other hand, areas with higher concentrations of migratory bird species are of great interest for amateur ornithologists because migrants are a classical target in Spanish ornithology (Rogers, 1988; De Juana, 2003; Pérez-Tris and Santos, 2004).

The illustrated biased pattern of ornithological field work in this area casts doubts about the usefulness of biodiversity indices obtainable from databases of observational records without a random-stratified sampling approach (Graham *et al.*, 2004). This could pose an important problem if recording effort does not reach the asymptote of the relationship between number of species and sampling time, and could be magnified in other less developed areas of rough terrain with less logistic facilities or with a lower density of zoologists (Peterson *et al.*, 1998). Nevertheless, these problems may be overcome by testing for the presence of a sampling bias (see also Reddy and Dávalos, 2003). The amount of variability explained by environmental factors on time invested in bird inventories is a direct measure of that bias. The date of zoological records, always present in all museum collections or databases, can be used as a simple index of sampling intensity defining the sampling pattern. Knowing the factors relating the number of records or days of field work to environmental traits, new sampling campaigns can be defined in low visited areas to more properly attain a random or stratified sampling protocol of regional biodiversity. This approach could be useful in inventories based on faunistic records (registering the date and locality of observation) for defining biodiversity patterns in many taxa lacking a well-established base of professional researchers or amateur observers (e.g., bats and other small mammals, butterflies, beetles, or freshwater and terrestrial molluscs).

This paper do not intends to model how biases in sampling effort affect avian inventory patterns. Instead it documents the fact that there are predictable between-sites differences in sampling effort and that such differences should be taken into account before using raw inventory data to draw conclusions about spatial patterns of avian diversity.

ACKNOWLEDGEMENTS.—We wish to acknowledge the Departament de Medi Ambient of Generalitat de Catalunya for funding the VertebraCat database and the related web of BioCat (Biodiversity data bank of Catalonia) and also Salvador Grau, Francesc Diego and Xavier Font. Also thanks to: Miquel Vila, Eva Ribes, Cristina Sánchez, Pilar Olalla, Teresa Pardo, Lynx Eds., Sociedad Española de Ornitologia, Institut Català d'Ornitologia, Vito Pedrocchi, Sergi Herrando, Joan Estrada, Museu de Zoologia de Barcelona, Lluís Brotons, David Vieites and Ricard Gutiérrez. Claire Jasinski kindly reviewed the language of the manuscript.

BIBLIOGRAPHY

ARAÚJO, M. B. 2003. The coincidence of people and biodiversity in Europe. *Global Ecology and Biogeography*, 12: 5-12.

ARAÚJO, M. B., WILLIAMS, P. H. and TURNER, A. 2002. A sequential approach to minimise threats within selected conservation areas. *Biodiversity and Conservation*, 11: 1011-1024.

BALMFORD, A., MOORE, J. L., BROOKS, T., BURGESS, N., HANSEN, L. A., LOVETT, J. C., TOKUMINE, S., WILLIAMS, P., WOODWARD, F. I. and RAHBEK, C. 2001a. People and biodiversity in Europe - response. *Science*, 293: 1591-1592.

BALMFORD, A., MOORE, J. L., BROOKS, T., BURGESS, N., HANSEN, L. A., WILLIAMS, P. and RAHBEK, C. 2001b. Conservation conflicts across Africa. *Science*, 291: 2616-2219.

BERNIS, F. 1966. *Aves migradoras ibéricas. Vol. I, Fasc. 1º a 4º.* Sociedad Española de Ornitología. Madrid.

BROTONS, L., PONS, P. and HERRANDO, S. 2005. Colonization of dynamic Mediterranean landscapes: where do birds come from after fire? *Journal of Biogeography*, 32: 789-798.

BURROUGH, P. A. and MCDONELL, R. A. 1998. *Principles of Geographical Information Systems*. Oxford University Press. New York.

CARRASCAL, L. M AND; DÍAZ, L. 2003. Asociación entre distribución continental y regional. Análisis con la avifauna forestal y de medios arbolados de la península Ibérica. *Graellsia*, 59: 179-207

CARRASCAL, L. M. and LOBO, J. M. 2003. Respuestas a viejas preguntas con nuevos datos: estudio de los patrones de distribución de la avifauna española y consecuencias para su conservación. In, R. Martí and J. C. Del Moral (Eds.): *Atlas de las Aves Reproductoras de España*, pp. 645-662 and 718-721. Dirección General de la Conservación de la Naturaleza-Sociedad Española de Ornitología. Madrid.

CARRASCAL, L. M., BAUTISTA, L. M. and LÁZARO, E. 1993. Geographical variation in the density of the white stork *Ciconia ciconia* in Spain: influence of habitat structure and climate. *Biological Conservation*, 65: 83-87.

CHOWN, S. L., VAN RENSBURG, B. J., GASTON, K. J., RODRIGUES, A. S. L. and VAN JAARSVELD, A. S. 2003. Energy, species richness, and human population size: Conservation implications at a national scale. *Ecological Applications*, 13: 1233-1241.

CLAVELL, J. 2002. *Catàleg dels ocells del Països Catalans*. Lynx Edicions, Bellaterra.

COPETE, J. L. 2000. *Anuari d'ornitologia de Catalunya. 1997*. Grup Català d'Anellament. Barcelona.

DE JUANA, E. 2003. Report on rare birds in Spain in 2001. *Ardeola*, 50: 123-149.

DÍAZ-DELGADO, R. and PONS, X. 2001. Spatial patterns of forest fires in Catalonia (northeast of Spain) along the period 1975-1995: Analysis of vegetation recovery after fire. *Forest Ecology and Management*, 147: 67-74.

DOBSON, A. P., RODRIGUEZ, J. P. and ROBERTS, W. M. 2001. Synoptic tinkering: integrating strategies for large-scale conservation. *Ecological Applications*, 11: 1019-1026.

ESTRADA, J.; PEDROCCHI, V., BROTONS, L. and HERRANDO, S. 2005. *Atlas dels ocells nidificants de Catalunya 1999-2002*. Institut Català d'Ornitologia. Barcelona.

FASOLA, M. and RUÍZ, X. 1997. Rice farming and waterbirds: integrated management in an artificial landscape. In, D. Pain and M. W. Pienkowski (Eds): *Farming and birds in Europe*, pp. 211-235. Academic Press. London.

FERNÁNDEZ, J. 2004. *50 años en defensa de las aves (1954-2004)*. Sociedad Española de Ornitología. Madrid.

FERRER, X. and MARTÍNEZ-VILALTA, A. 1987. Le delta de l'Ebre: un milieu aquatique réglé par la culture du riz. *L'Oiseau et R.F.O*, 57: 13-22.

FIELD, S. A., TYRE, A. J. and POSSINGHAM, H. P. 2002. Estimating bird species: How should repeat surveys be organized in time? *Austral Ecology*, 27: 624-629.

FJELDSÅ, J. and RAHBEK, C. 1998. Continent-wide conservation priorities and diversification processes. In, G. M. Mace, A. Balmford and J. Ginsberg (Eds): *Conservation in a changing world*, pp. 139-160. Cambridge University Press. Cambridge.

FRASER, P. 1997. How many rarities are we missing? Weekend bias and length of stay revisited. *British Birds*, 90: 94-101.

FREITAG, S., HOBSON, C., BIGGS, H. C. and VAN JAARSVELD, A. S. 1998. Testing for potential survey bias: the effect of roads, urban areas and nature reserves on a southern African mammal data set. *Animal Conservation*, 1: 119-127.

GASTON, K. J. and EVANS, K. L. 2004. Birds and people in Europe. *Proceedings of the Royal Society of London, Series B*, 271: 1649-1655.

GONZÁLEZ, L. M., BUSTAMANTE, J. and HIRALDO, F. (1990). Factors influencing the present distribution of the Spanish imperial eagle *Aquila adalberti*. *Biological Conservation*, 51: 311-319.

GRAHAM, C. H., FERRIER, S., HUETTMAN, F., MORITZ, C. and PETERSON, A. T. 2004. New developments in museum-based informatics and applications in biodiversity analysis. *Trends in Ecology and Evolution*, 19: 497-503.

KODRIC-BROWN, A. and BROWN, J. H. 1993. Incomplete data sets in community ecology and biogeography: a cautionary tale. *Ecological Applications*, 3: 736-742.

KUHN, I., BRANDL, R. and KLOTZ, S. 2004. The flora of German cities is naturally species rich. *Evolutionary Ecology Research*, 6: 749-764.

LEGENDRE, P. 1993. Spatial autocorrelation: trouble or new paradigm? *Ecology*, 74: 1659-1673.

MARTÍ, R. and DEL MORAL, J. C. 2003. *Atlas de las Aves Reproductoras de España*. Dirección General de Conservación de la Naturaleza- Sociedad Española de Ornitología. Madrid.

MARTÍNEZ-VILALTA, A. 1986. Característiques generals del poblament ornític dels Països Catalans. In, X. Ferrer (Ed.): *Història Natural dels Països Catalans Vol.12 (Ocells)*, pp. 361-366. Fundació Enciclopèdia Catalana. Barcelona.

MMAMB (Mancomunitat de Municipis de l'Àrea Metropolitana de Barcelona). 1995. *Metropolitan dynamics in the Barcelona area and region*. AMB, Barcelona.

MOREAU, R. E. 1956. The Iberian Peninsula and migration. *Bird Study*, 3:1-25.

MÚGICA, M. 1994. *Modelos de demanda paisajística y uso recreativo de los espacios naturales*. Serie Documentos nº 16: 165. Centro de Investigación de Espacios Naturales Protegidos Fernando González Bernáldez. Madrid.

MUNTANER, J., FERRER, X. and MARTÍNEZ-VILALTA, A. 1984. *Atlas dels ocells nidificants de Catalunya i Andorra*. Ketres editora. Barcelona.

NELSON, B. W., FERREIRA, C. A. C., DA SILVA, M. F. and KAWASAKI, M. L. 1990 Endemism centres, refugia and botanical collection density in Brazilian Amazonia. *Nature*, 345: 714-716.

NINYEROLA, M., PONS, X. and ROURE, J. M. 2000. A methodological approach of climatological modelling of air temperature and precipitation through GIS techniques. *International Journal of Climatology*, 20: 1823-1841.

PARNELL, J. A. N., SIMPSON, D. A., MOAT, J., KIRKUP, D. W., CHANTARANOTHAI, P., BOYCE, P. C., BYGRAVE, P., DRANSFIELD, S., JEBB, M. H. P., MACKLIN, J., MEADE, C., MIDDLETON, D. J., MUASYA, A. M., PRAJAKSOOD, A., PENDRY, C. A., POOMA, R., SUDDEE, S. and WILKIN, P. 2003. Plant collecting spread and densities: their potential impact on biogeographical studies in Thailand. *Journal of Biogeography*, 30: 193-209.

PAUSAS, J. G., CARRERAS, J., FERRÉ A. and FONT, X. 2003. Coarse-scale plant species richness in relation to environmental heterogeneity. *Journal of Vegetation Science*, 14: 661-668.

PÉREZ-TRIS, J. and SANTOS, T. 2004. El estudio de la migración de aves en España: trayectoria histórica y perspectivas de futuro. *Ardeola*, 51:71-89.

PETERSON, A. T., NAVARRO-SIGÜENZA, A. G. and BENÍTEZ-DÍAZ, H. 1998. The need for continued scientific collecting; a geographic analysis of Mexican bird specimens. *Ibis*, 140: 288-294.

PONS, P. and BAS, J. M. 2005. Open-habitat birds in recently burned areas: The role of the fire extent and species' habitat breadth. *Ardeola*, 52: 119-131.

PONS, X. 2002. *MiraMon. Sistema d'Informació Geogràfica i software de Teledetecció*. Center for Ecological Research and Forestry Applications (CREAF). Bellaterra.

REDDY, S. and DÁVALOS, L. M. 2003. Geographical sampling bias and its implications for conservation priorities in Africa. *Journal of Biogeography*, 30: 1719-1727.

ROGERS, M. J. 1988. Report on rare birds in Great-Britain in 1987. *British Birds*, 81: 535-596.

Şekercioğlu, Ç. H. 2002. Impacts of birdwatching on human and avian communities. *Environmental Conservation*, 29: 282-289.

SHARROCK, J. T. R. 1966. Weekend bias, length of stay of rarities and proportion of rarities undetected. *British Birds*, 59: 556-558.

STATSOFT, INC. 2001. STATISTICA (data analysis software system). Tulsa, Oklahoma, USA.

SUÁREZ, F. 2004. Aves y agricultura en España peninsular: Una revisión sobre el estado actual de conocimiento y una previsión sobre su futuro. In: J. L. Tellería (Ed): *Ornitología hoy. Volumen homenaje a F. Bernis*, pp. 223-265. Universidad Complutense. Madrid.

SUÁREZ-SEOANE, S., OSBORNE, P. E. and ALONSO, J. C. 2002. Large-scale habitat selection by agricultural steppe birds in Spain: Identifying species-habitat responses using generalized additive models. *Journal of Applied Ecology,* 39: 755-771.

TOBIAS, R. D. 1995. An introduction to partial least squares regression. *SUGI Proceedings*. Available online at: http://support.sas.com/rnd/app/papers/pls.pdf.

TUCKER, G. and HEALTH, M. F. 1994. *Birds in Europe: their conservation status*. Birdlife International (Birdlife Conservation Series nº 3). Cambridge.

VIÑAS, O. and BAULIES, X. 1995. 1:250000 Land-use map of Catalonia (32.000 km^2) using multitemporal Landsat-TM data. *Journal of International Remote Sensing*, 16: 129-146.

WILLIAMS, P. H., MARGULES, C. R. and HILBERT, D. W. 2002 Data requirements and data sources for biodiversity priority area selection. *Journal of Biosciences*, 27: 327-338.

[Recibido: 15-01-06]
[Aceptado: 03-06-06]

Xavier Ferrer is Professor at the Department of Animal Biology, University of Barcelona. His research interests are biogeography, mainly of birds, and applied ornithology. He approaches the History of Ornithology in Spain. **Luis M. Carrascal** is a Tenured Scientist at the Department of Biodiversity and Evolutionary Biology of the Museo Nacional de Ciencias Naturales. His research is focused on macroecology, biogeographical ecology of the South-Western Palaearctic avifauna and modelling patterns of bird species abundance/occurrence. **Oscar Gordo**'s interests are focused on the effects of climate change in the phenology of organisms. He is currently working on temporal and spatial phenological patterns of trans-Saharan birds from the Iberian Peninsula. **Joan Pino**, Ph.D., is an expert in plant biology, landscape ecology and GIS. His major research topics are landscape classification, analyses of the association between biodiversity and landscape structure and dynamics, and application of landscape ecology to nature conservation in metropolitan areas.