Geostatistical modelling of regional bird species richness : exploring environmental proxies for conservation purpose

Giovanni Bacaro, Alessandro Chiarucci

BIOCONNET, BIOdiversity and CONservation NETwork, Dipartimento di Scienze, Ambientali"G. Sarfatti", Università di Siena, Via P. A. Mattioli 4, 53100, Siena, Italy.

Elisa Santi

IRPI-CNR, Via Madonna Alta 126, 06128 Perugia, Italy

Duccio Rocchini

Department of Biodiversity and Molecular Ecology, GIS and Remote Sensing Unit, Fondazione Edmund Mach, Research and Innovation Centre, Via E. Mach 1, 38010 S. Michele all'Adige, TN, Italy

Francesco Pezzo, Luca Puglisi

Centro Ornitologico Toscano, C.P. 470, 57100 Livorno, Italy

Abstract: Identifying spatial patterns in species diversity represents an essential task to be accounted for when establishing conservation strategies or monitoring programs. Predicting patterns of species richness by a model-based approach has recently been recognised as a significant component of conservation planning. Here, a spatially-explicit data-set on birds presence and distribution across the whole Tuscany region was analysed using geostatistical models. Species richness was calculated within 1x1 km grid cells and 10 environmental predictors were included in the analysis. A statistical model integrating spatial components of variation with predictive ecological factors of bird species richness was developed and used to obtain predictive regional maps of bird diversity hotspots.

Keywords: Bird richness, Conservation, Distribution maps, Natura 2000 Network, Predictive model, Semivariance, Spatial autocorrelation, Tuscany.

1. Introduction

The identification of spatial patterns in species diversity represents an essential task for biodiversity conservation strategies or monitoring programs. Recently, species distribution modeling emerged as a new approach to generate species distribution maps, on the basis of the relationship between species presence (or abundance) records and environmental variables. Typically, modeling methods attempt to predict the probability of occurrence of species as a function of a set of environmental variables. In particular, geostatistical modeling techniques, which have been developed mainly in the field of geography, are designed to model spatially dependent observations (Goovaerts 1997), but in recent years, such methodologies have been applied even in the ecological literature (Bacaro an Ricotta 2007). Birds are among the best-studied organisms, especially in Europe. They are often considered as excellent indicators of environmental changes and as good ecological proxies to assess the biodiversity values of an area. In this work, a geostatistical modelling approach was applied on the data provided by the "Monitoring Program of Breeding Birds of Tuscany", one of the most extensive regional bird monitoring programs in Italy. The aim of this paper is *i*) to describe the

spatial patterns of bird species richness and *ii*) to identify those environmental factors underlying these patterns. This latter point represents an important task in the ecological context since the environmental proxies driving bird richness could be used to decide conservation strategies.

2. Materials and Methods

Bird data: The bird species occurrence data were obtained from the Monitoring Program of Breeding Birds of Tuscany carried out by the Centro Ornitologico Toscano (www.centrornitologicotoscano.org) and based on Point Counts method (Bibby *et al.* 2000). Points were distributed according to a two stages sampling design: in randomly selected 10*10 km UTM cells, a number of 12-15 point counts were selected according to a second random sampling procedure. The original data set of geo-referenced observations was assembled to produce a regional map of bird species richness for cells of 1*1 km. Such a grid covering the whole Tuscany region resulted in 22060 cells, 3584 of which enclosed data on bird occurrences.

Putative determinants of bird species richness: for each 1*1 km cell, three sets of predictor/explanatory variables were derived and grouped according to a similarity criterion. I) Geographical features (4 predictors): the coordinates for each grid cell (Latitude and Longitude), elevation and distance to the sea were included in this group. II) Landscape feature and complexity (4 variables): Data on land cover were derived from the third level of the CORINE Land Cover Map. For each grid cell, the number of patches and the area (mean and standard deviation) covered by each land cover class was calculated. Landscape shape complexity was calculated by using the AWMSI (Area Weighted Mean Shape Index). The third level data of the CORINE Land Cover were used for calculating the Shannon index. III) Primary Productivity (2 variables): NDVI (Normalized Difference Vegetation Index) and its standar deviation were used on to discriminate between the amount of biomass characterising different vegetation types.

Geostatistical modelling: a combined multi-predictor model was developed in this study, and it was further used in conjunction with geostatistical techniques to predict birds diversity in 1 x 1 km grid cells across the whole Tuscany region. Statistical modelling process was organised into the following three parts: 1) Data transformation (normalization); 2) Building the generalized linear spatial model: once the response variable (number of bird species) at each grid cell within the Tuscany region was denoted as:

$$x_i, y_i \; : i = 1, \ldots, n \qquad (1)$$

where x_i identifies the spatial location (in two-dimensional space - longitude and latitude expressed in kilometres) and y_i is the bird richness value associated with the location x_i, a geostatistical (isotropic) model can be defined as:

$$Y_i = S \; x_i \; + Z_i : i = 1, \ldots, n \qquad (2)$$

where

$$\{ S \; x : x \qquad ^2 \} \qquad (3)$$

is a Gaussian process with a spatially varying mean *μ(x)* defined by a classical linear regression model. The described Gaussian process is also characterized by a variance given by:

$$\sigma^2 = Var \{ S \; x \} \qquad (4)$$

and by a positive-defined correlation function:

$$\rho \; u \; = Corr \{ S \; x \; , S \; x' \} \quad (5)$$

defining the way correlation function decays to zero for increasing distances occurring between observations at locations x and x'. Explanatory variables for modelling the large-scale variation in bird diversity were chosen via a model selection technique (AIC). Secondly, the residuals from the model were examined for spatial correlation and a suitable family of correlations was chosen. The estimates of the parameters in the trend surface (model spatial component) were updated using the quasi-Newton optimisation function (Byrd et al. 1995) followed by maximum likelihood estimation of the parameters of the covariance function using the residuals. 3) Universal kriging was used to predict expected bird richness (and its variation) in each 1x1 km grid cell across the whole Tuscany Region.

3. Results

The number of bird species per cell grid was normalized using a Box-Cox power of 0.184. Only 4 predictors were included in the predictive model (Table 1). The intercept of the estimated spatial varying mean resulted highly significant and was, consequently, included in the model.

Table 1: Description of explanatory variables (and their associated coefficients) included after stepwise selection in the spatial varying mean component (*** p<0.001).

Trend parameters (spatial varying mean)	Estimated Value
Intercept	3.066[***]
NDVI St.Dev.	0.811[***]
H' index	0.104[***]
Mean elevation	-0.001[***]
Distance sea	>0.001[***]
Spatial Parameters	
Nugget (τ^2)	0.147
Partial sill (σ^2)	0.270
Range (φ)	0.054
Practical Range	0.162
Normalisation parameter (Box-Cox power)	
lambda (λ)	0.184
Covariance Function Parameters (Matèrn)	
Order (k)	0.5 (exponential model)

The modeled spatial parameters highlighted that autocorrelation in bird richness value existed and strongly influenced the number of observed species. In particular, the practical range was reached after 16 km, indicating the absence of further correlative structure in data after this threshold (see Figure 1). Relatively to the covariance function used to model the empirical variogram, the k=0.5 parameter was selected (corresponding to fit an exponential theoretical variogram with respect to the observed data). Predicted values were significantly related with observed bird richness (R^2 = 0.448, p < 0.001). For comparison, a simple multiple regression model without the inclusion of the spatial component in the analysis, showed a lower R^2 value (R^2=0.15, p<0.001). Predicted bird richness (and its associated variance) across all the Tuscan region is shown in Figure 2.
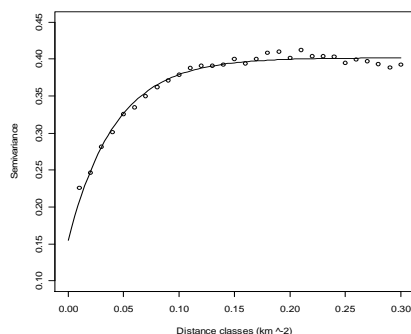
Figure 1: Plot of the empirical (circles) and fitted (solid line) semivariograms versus distance (km) obtained using the residuals after the spatial varying mean was subtracted by raw (normalized) data.
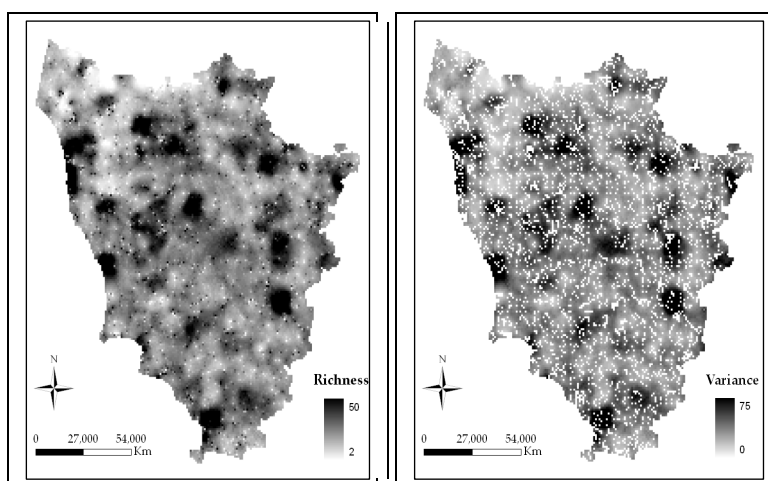

Figure 2: Regional pattern of bird species richness as expected under the described geostatistical model. a) Expected birds species richness and b) its expected variance.

4. Conclusions

By applying geostatistical models, a well-performing predictive model was obtained for the distribution of bird species richness in Tuscany by considering relatively few variables. Ancillary variables based on remotely sensed information (e.g., NDVI or Shannon *H'* derived from a classified image) can be used as powerful tools to model the spatial variation of bird species richness and locate biodiversity hotspots. Moreover, geostatistical models own the advantage to incorporate information of environmental co-variation and neighborhood effects, improving the quality of predictions.

References

Bacaro G, Ricotta C (2007) A spatially explicit measure of beta diversity. Community Ecol 8: 41-46.

Byrd RH, Lu P, Nocedal J, Zhu C (1995) A limited memory algorithm for bound constrained optimization.SIAM J. Scientific Computing16: 1190–1208.

Bibby CJ, Burgess ND, Hill DA, Mustoe SH (2000) Bird census techniques, 2nd edn. Academic Press, London.

Goovaerts P (1997) Geostatistics for natural resources evaluation. Oxford University Press, New York.