Species richness coincidence: conservation strategies based on predictive modelling

DIRK MAES[1,*], DIRK BAUWENS[1], LUC DE BRUYN[1,2], ANNY ANSELIN[1],
GLENN VERMEERSCH[1], WOUTER VAN LANDUYT[1], GEERT DE KNIJF[1]
and MARIUS GILBERT[3,4]

[1]*Institute of Nature Conservation, Kliniekstraat 25, B-1070 Brussels, Belgium;* [2]*Department of Biology, University of Antwerp, Groenenborgerlaan 171, B-2020 Antwerpen, Belgium;* [3]*Laboratoire de Biologie Animale et Cellulaire, CP 160/12, Université Libre de Bruxelles, 50 av. F. D. Roosevelt, B-1050 Brussels, Belgium;* [4]*Fonds National de la Recherche Scientifique, Brussels, Belgium;* *Author for correspondence (e-mail: dirk.maes@instnat.be; fax: +32-2-558-18-05)*

Abstract. The present-day geographic distribution of individual species of five taxonomic groups (plants, dragonflies, butterflies, herpetofauna and breeding birds) is relatively well-known on a small scale (5×5 km squares) in Flanders (north Belgium). These data allow identification of areas with a high diversity within each of the species groups. However, differences in mapping intensity and coverage hamper straightforward comparisons of species-rich areas among the taxonomic groups. To overcome this problem, we modelled the species richness of each taxonomic group separately using various environmental characteristics as predictor variables (area of different land use types, biotope diversity, topographic and climatic features). We applied forward stepwise multiple regression to build the models, using a subset of well-surveyed squares. A separate set of equally well-surveyed squares was used to test the predictions of the models. The coincidence of geographic areas with high predicted species richness was remarkably high among the four faunal groups, but much lower between plants and each of the four faunal groups. Thus, the four investigated faunal groups can be used as relatively good indicator taxa for one another in Flanders, at least for their within-group species diversity. A mean predicted species diversity per mapping square was also estimated by averaging the standardised predicted species richness over the five taxonomic groups, to locate the regions that were predicted as being the most species-rich for all five investigated taxonomic groups together. Finally, the applicability of predictive modelling in nature conservation policy both in Flanders and in other regions is discussed.

Introduction

One of the major challenges for conservation biology is to stop the ongoing and accelerating decline of biodiversity (Pimm et al. 1995). However, limited funding and the constantly growing number of threatened species call for prioritisation. One of the ways to increase efficiency in nature conservation is to direct efforts towards species-rich sites ('biodiversity hotspots' – Myers et al. 2000). This strategy would prevent the extinction of a larger number of species per unit protected area (Reid 1998). Several authors have delineated the most diverse or most threatened areas world-wide or on a continental scale (e.g., Pearson and Cassola 1992; Dobson et al. 1997). However, most conservation policies are restricted to country or region

boundaries and applying the concept of delineating species-rich sites on smaller scales would considerably improve the efficacy of national or local nature conservation policies (Prendergast et al. 1993). A problem of this approach is that species-rich sites of different taxonomic groups do not necessarily coincide, a finding that calls into question the utility of the concept of 'indicator taxa' for conservation policy purposes (Prendergast et al. 1993; van Jaarsveld et al. 1998).

Few countries or regions have sufficiently fine-scaled species distribution data of different taxonomic groups to allow tests for the coincidence of local species richness at a scale where nature conservation is generally applied in the field. Moreover, differences among taxonomic groups in geographic scope of the collected data and in survey efforts can seriously bias a straightforward delineation of species-rich sites (Prendergast et al. 1993, 1999). Predictive modelling, applying multiple regression techniques on distribution data and a set of environmental variables, has been proposed as a useful tool to 'correct' for differences in mapping intensity and unequal area coverage (Maddock and Du Plessis 1999). This approach uses data of a limited number of well-surveyed sites to model species diversity for a given taxonomic group as a function of environmental data. After appropriate validation, the model is used to obtain predictions of local species richness, which are less biased owing to differences in mapping intensity and incomplete survey coverage. This method was found to be successful in predicting species richness at different scales for a variety of taxonomic groups: terrestrial vertebrates in American national parks (Edwards et al. 1996); mammals in the North American continent (Badgley and Fox 2000); butterflies in countries like France (Dennis et al. 2002; Dennis and Shreeve 2003), Belgium (Maes et al. 2003) or in the Great Basin (Mac Nally et al. 2003). However, these studies were mostly focused on single taxonomic groups and analysing the degree of species richness coincidence among taxonomic groups has so far only been carried out with uncorrected and biased data (Maddock and Du Plessis 1999).

Here, we used fine-scale distribution data (5×5 km grid cell size) and the method outlined above to build separate predictive models of five taxonomic groups (higher plants, dragonflies, butterflies, herpetofauna and breeding birds) in Flanders, accounting for incomplete geographic coverage and variation in survey intensity. Spatial coincidence in the predicted local species richness patterns are explored and discussed in relation to the relevance of biodiversity indicator species and to conservation strategies and policy (e.g., prioritisation of areas for conservation – Myers et al. 2000).

Materials and methods

Study area

Flanders (total area 13,512 km^2) is one of the federal regions of Belgium, covering the northern part of the country (Figure 1). It exhibits the typical features of a western industrialised region (OECD 1998): extensive industry, infrastructure, house
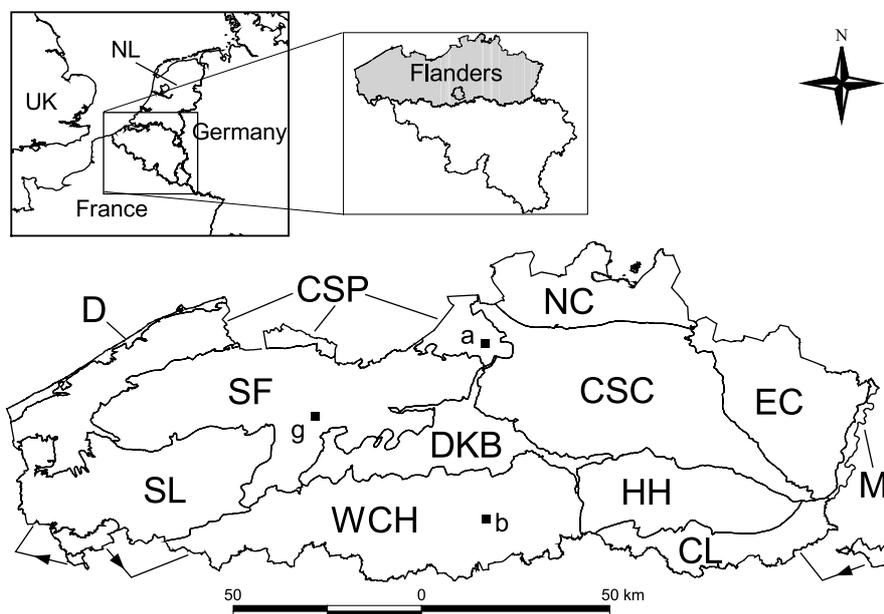
Figure 1. Delimitation of the ecological regions in Flanders and geographic location of Flanders within Western Europe and Belgium (insets). The following ecological regions were considered (De Blust 2001): Dunes (D); Sandy Loam (SL); Sandy Flanders (SF); Coastal and Scheldt polders (CSP); Western and Central hills (WCH); Dender-Klein Brabant (DKB); North Campine (NC); Central and South Campine (CSC); East Campine (EC); Hageland-Haspengouw (HH); Calcareous-Loam (CL); Meuse valley (M). Also shown is the location of the cities of Antwerp (a), Brussels (b) and Ghent (g).

building and agriculture, and a very high human population density (431 citizens/km^2 – Van Hecke and Dickens 1994).

Nature conservation is one of the political competencies that were transferred from the Federal to the Flemish Government. The total area of officially recognised nature reserves in Flanders is limited (i.e., 1.6% of the total territory – Decleer and Vanroose 2001); 1019 km^2 and 978 km^2 are designated as Habitat Directive (EU Directive 92/43/EEG) and Bird Directive (EU Directive 79/409/EEG) areas respectively, of which 366 km^2 overlap (Dries 2002). Based on general features of the landscape and geomorphology 12 ecological regions were distinguished in Flanders (Figure 1 – De Blust 2001).

Distribution of species diversity

Data on the distribution of individual species were obtained from different databases for five taxonomic groups: higher plants (Van Landuyt et al. 2000), dragonflies (De Knijf and Anselin 1996), butterflies (Maes and Van Dyck 2001), herpetofauna (i.e., amphibians and reptiles; Bauwens and Claus 1996) and breeding birds

Table 1. (a) Number of records, number of surveyed squares (5 × 5 km; total $N = 585$), number of indigenous species and the survey period for the different taxonomic groups; (b) number of squares in the training (T) and evaluation (E) sets within each of the different ecological regions used to build and evaluate the multiple regression models of five taxonomic groups.

(a)

	Plants	Dragonflies	Butterflies	Herpetofauna	Breeding birds
No. of records	2,270,000	40,000	210,000	15,000	323,000
No. of investigated squares	585	483	575	563	563
No. of indigenous species	1125	58	64	19	163
Survey period	1972–2002	1980–2002	1991–2002	1974–2002	1999–2002

(b)

	T	E	T	E	T	E	T	E	T	E
Calcareous-Loam	7	3	5	2	7	3	7	2	7	3
Central and South Campine	21	8	23	8	23	8	23	9	23	8
Coastal and Schelde polders	11	3	12	6	12	5	11	3	11	5
Dender-Klein Brabant	8	3	8	3	8	3	8	3	8	3
Dunes	1	1	1	1	1	1	1	1	1	1
East Campine	13	3	13	4	12	4	13	4	13	4
Hageland-Haspengouw	10	6	10	4	10	3	10	3	10	3
Meuse	1	1	1	1	1	1	1	1	1	1
North Campine	10	4	10	4	10	3	10	3	9	3
Sandy Flanders	23	9	22	7	23	8	23	9	23	8
Sandy Loam	13	4	14	4	15	5	13	4	15	4
Western and Central hills	26	9	26	10	26	9	26	9	25	9
Total	144	54	145	54	148	53	146	51	146	51

(Vermeersch and Anselin, unpublished data). Distribution data were collected by a large number of volunteers attributing observations to grid cells of 5 × 5 km of the Universal Transverse Mercator (UTM) projection for the dragonfly, butterfly and breeding bird mapping schemes and to grid cells of 4 × 4 km of the "Institut Floristique de la Belgique et Luxembourg" (IFBL) projection for the plants and herpetofauna mapping schemes. Prior to analyses, we converted IFBL grid cells to UTM grid cells, hereafter called squares, by overlaying both projections in the geographical information system Arcview GIS 3.2 (Esri, Redlands, CA, USA). Only squares having >25% of their area within Flanders were used for the analyses ($N = 585$).

Table 1 shows some basic information for the different survey schemes. For each taxonomic group, we also obtained information on the number of visits made to each square, allowing assessment of sampling intensity. Distribution records for four taxonomic groups cover at least 90% of the squares subjected to the analyses (Table 1a). The data for dragonflies have a lower coverage (78%), but there is nevertheless a sufficient number of well investigated squares to develop a predictive model for species richness in Flanders (Lobo and Martín-Piera 2002; Luoto et al. 2002).

Table 2. Symbols and data source for the environmental variables used in the multiple regression models for species diversity.

Variable	Symbol	Data source
Biotope data		
Urban area	Ur	CORINE codes beginning with 1
Agricultural land	Ar	CORINE codes beginning with 2
Deciduous woodland	Dw	CORINE code 311
Mixed woodland	Mw	CORINE code 313
Coniferous woodland	Cw	CORINE code 312
Natural grassland	Gr	CORINE code 321
Heathland and bog	Hb	CORINE code 322 + 412
Shrub	Sh	CORINE code 324
Salt marshes	Sm	CORINE code 421 + 423
Dunes	Du	CORINE code 331
Water courses	Wc	CORINE code 511 + 522
Water bodies	Wb	CORINE code 512
Marsh	Ma	CORINE code 411
Biotope Diversity Index	BDI	Shannon diversity index of biotopes
Climate data		
Number of frost days	F	Interpolated point data of the RMIB[a]
Yearly rainfall	R	Interpolated point data of the RMIB[a]
Sum of sunhours	S	Interpolated point data of the RMIB[a]
Maximum temperature	T	Interpolated point data of the RMIB[a]
Topographic data		
Mean elevation	El	Digital elevation model Belgium
Range elevation	RE	Digital elevation model Belgium

[a] Royal Meteorological Institute of Belgium.

Species diversity was estimated in each square and for each taxonomic group as the number of indigenous species recorded during the survey period.

Environmental variables

We extracted data on the distribution of different land uses from the CORINE land cover map for Europe (CEC 1994). The 44 land use categories distinguished on this map were lumped into 13 land use types that are present in Flanders (Table 2). For each square we estimated the area occupied by the different land use types using the GIS. In addition, we estimated biotope diversity per square using the Shannon diversity index (Magurran 1988).

Climate data were obtained from the Royal Meteorological Institute of Belgium for the period 1996–2001. Point climate data were interpolated in the squares by universal kriging (Isaaks and Srivastava 1989) when a clear spatial structure could be modelled or, alternatively, using the 'inverse weighted distance' interpolation method. Universal kriging with a linear drift was used to interpolate yearly cumulated rainfall (mm; 186 locations), yearly average maximum temperature (°C;

114 locations) and yearly cumulated number of frost days (114 locations). The 'inverse weighted distance' method was used to interpolate yearly-cumulated sunshine exposure (hours, 22 locations). Spatial interpolations were carried out using the software Surfer 8.0 (Golden Software Inc., Golden, CO, USA).

Topographic variables (mean elevation; elevation range [i.e., the difference between the highest and lowest elevation]) were derived from a digital elevation model for Belgium (1996, National Geographical Institute, resolution 20 m) and estimated for each square using the GIS.

Modelling species diversity

Prediction of species diversity over the entire territory of Flanders was based upon multiple regression models developed on a subset of the surveyed squares and relating species diversity to environmental variables. However, multiple regressions require that the predictor variables are mutually independent. Therefore, we first examined the correlations among the 20 environmental variables, using data for all squares ($N = 585$). Unbiased correlation levels of significance were obtained using the method proposed by Clifford et al. (1989) and modified by Dutilleul (1993) that quantifies the reduction in degrees of freedom due to spatial autocorrelation in the two variables. We also corrected the levels of significance using the multiple testing adjustment procedure of Legendre and Legendre (1998). In addition, we performed a principal components analysis (PCA) on all environmental variables to examine whether it was appropriate to substitute the original variables by a reduced set of component variables.

We built and evaluated a predictive model per taxonomic group on a subset of the squares ($N = \pm 200$). Specifically, we selected the 30% best-surveyed (i.e., most visited) squares within each of the 12 ecological regions (Figure 1 and Table 1b). This procedure accounts for differences in area, mapping intensity and species richness in the different ecological regions. Three-fourths of these squares were used to build the model (hereafter called *training set*), while the remaining fourth of the squares were used to evaluate the model (hereafter called *evaluation set*). Attributing the well-surveyed squares to either the training or to the evaluation set was based on a random selection within each ecological region.

Species diversity for each of the five taxonomic groups was modelled using forward stepwise multiple regression. The putative presence of curvilinear relationships between the predictor variables and species richness was taken into account by incorporating the quadratic terms of the predictor variables (Nicholls 1989). The function of the independent variable that accounted for the largest reduction in deviance (F-ratio test, $p < 0.05$) was first incorporated into the model (Crawley 1993). Next, all the remaining predictor variables were tested in the same way until inclusion was no longer significant. At each step, all previously entered variables were tested for their significance and removed from the model if they were no longer significant. For each linear model, the co-variogram of the residuals was estimated to check for spatial autocorrelation (Overmars et al. 2003). No

spatial autocorrelation was present in the residuals of the predictive models for dragonflies and breeding birds. For plants, butterflies and herpetofauna the covariance among residuals due to spatial autocorrelation was modelled using the SAS MIXED procedure by the function

$$\text{Cov}(e_i, e_j) = \sigma^2 f(d_{ij})$$

where e_i is the error corresponding to the i-th observation, d_{ij} is the distance between the spatial location of the i-th and j-th residual and f is the spatial covariance function (Littell et al. 1996). The spatial covariance function was adjusted by modelling the experimental co-variogram of the multiple regression residuals using the spherical model (Isaaks and Srivastava 1989). The spatial covariance model parameters were identified using a combination of 'fit-by-eye' and least squares approaches, selecting the model providing the best fit.

The final regression model for each species group was used to predict species diversity within the squares in the evaluation set. We stress that these squares were not used to build the model. We assessed the goodness of fit of the predictive model by the Spearman rank correlation between the predicted and observed species diversity in the squares of the evaluation set. After its validation, the regression model for a species group was used to predict species richness in all squares.

The mean predicted species diversity in each square was estimated as the mean of the standardised predicted species richness (SSR) over the five taxonomic groups; SSR is calculated as

$$\text{SSR} = \frac{(x_i - x_{\min})}{(x_{\max} - x_{\min})}$$

where x_i is the predicted species richness in the i-th square and $x_{\min}$ and $x_{\max}$ are the minimum and maximum predicted species richness, respectively (Gower 1971).

Results

Correlations of observed species diversity among taxonomic groups

A rather restricted number of squares ($N = 244$; 38% of total) was surveyed at least three times for all five taxonomic groups together. We used this subset of squares to calculate correlations of the observed species diversity among the five groups (Table 3). The lowest correlation was between plant and butterfly species diversity, while species diversity of dragonflies is highly correlated with that of butterflies and breeding birds.

Geographic patterns of observed species diversity for dragonflies, butterflies, herpetofauna and breeding birds showed a large concentration of species-rich squares in the Campine regions (NE Flanders – Figures 2b–e). Plant species diversity did not show such a pronounced pattern and species-rich squares were more scattered over Flanders (Figure 2a).

Table 3. Pairwise correlation coefficients (Spearman rank correlation) for the observed species diversity among the different taxonomic groups. The correlations are based on data for the 244 squares that were surveyed at least three times for each of the species groups.

	Plants	Dragonflies	Butterflies	Herpetofauna
Dragonflies	0.270***	–		
Butterflies	0.174**	0.508***	–	
Herpetofauna	0.286***	0.336***	0.229***	–
Birds	0.306***	0.527***	0.306***	0.286***

$**p < 0.01$; $***p < 0.001$.

Correlations among environmental variables

The analysis of the collinearity among the environmental variables revealed that only 17 out of 179 possible correlations (i.e., 9%) were judged statistically significant after correcting for spatial autocorrelation and multiple testing (Table 4). Biotope diversity was the variable that was most frequently correlated with other environmental characteristics. The area of agricultural land was negatively correlated with the area occupied by most other land use types. It should be noted that very few significant correlations were found between the land use variables and the climatic and topographic variables (Table 4).

A principal components analysis on the correlation matrix of all environmental variables ($N = 20$) yielded six component axes that had eigenvalues >1; together they represented 65% of the total variance. The extraction of 14 axes was required to retain 90% of the original variance. Hence, the principal components analysis did not achieve a meaningful reduction of the dimensionality of our data set. Therefore we opted to use the original environmental variables as independent variables in the ensuing multiple regression analyses.

Modelling species diversity

The predictive models (i.e., multiple regression analyses) for all five species groups were highly significant (R^2 varies between 37–66%; Table 5). More important, the Spearman rank correlations between predicted and observed species richness in the squares of the evaluation set were also highly significant (Table 5), indicating that the regression models provided reasonably accurate predictions of species diversity in the evaluation set.

The number of variables used to build the predictive models ranged from two (butterflies) to nine (plants). Biotope diversity was the only variable that entered in the models for all five species groups, either as a linear or as a quadratic term (Table 5). Residuals of the models for plants, butterflies and herpetofauna showed some degree of spatial autocorrelation and the models therefore included a model of
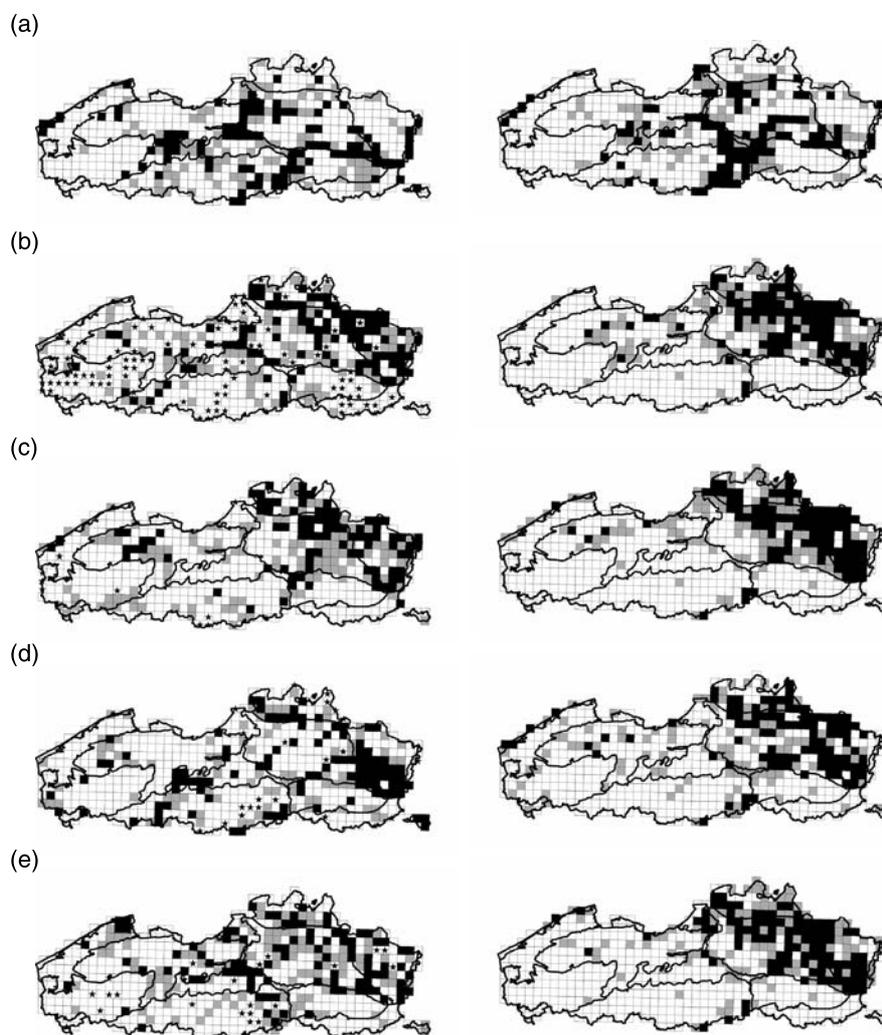
(a)

(b)

(c)

(d)

(e)

Figure 2. Geographic patterns of observed (left) and predicted (right) species richness for the different taxonomic groups: plants (a), dragonflies (b), butterflies (c), herpetofauna (d) and birds (e). In black the top 100 most species rich squares, in grey the next 100 most species rich squares; stars indicate squares that were not surveyed.

covariance among residuals to obtain unbiased estimates and levels of significance (Keitt et al. 2002).

We applied the models to obtain estimates of species diversity for each species group over the total area of Flanders (Figure 2). Species richness of plants was predicted to be high in the dune areas, near the cities of Ghent and Antwerp, in the transition zones between several ecological regions and in the valleys of the rivers

Table 4. Pearson correlation coefficients among the twenty environmental variables for all squares in Flanders ($N = 585$). In bold: correlations that were significant ($p < 0.10$) after applying a modified t-test that accounts for spatial autocorrelation and a correction for multiple testing. Freq = number of squares in which the biotope type occurs; . = insufficient data. Abbreviations of the environmental variables are given in Table 2.

	Freq	Ur	Ar	Dw	Mw	Cw	Gr	Hb	Sh	Sm	Du	Wc	Wb	Ma	BDI	F	R	S	T	El
Ur	584	—																		
Ar	584	**−0.472**	—																	
Dw	340	0.113	−0.197	—																
Mw	203	**−0.238**	−0.258	−0.066	—															
Cw	262	−0.036	**−0.332**	−0.008	0.196	—														
Gr	8	0.388	−0.563	.	.	.	—													
Hb	139	−0.049	−0.231	−0.070	0.149	0.350	0.145	—												
Sh	101	−0.016	**−0.427**	0.012	0.079	0.321	.	**0.517**	—											
Sm	8	−0.470	0.063	.	.	.	−0.997	0.279	0.003	—										
Du	22	0.191	**−0.588**	0.013	−0.328	0.225	0.761	0.178	0.132	0.666	—									
Wc	83	0.046	−0.182	−0.072	0.115	0.230	0.335	0.333	0.155	0.043	−0.832	—								
Wb	175	0.109	−0.186	−0.005	−0.184	−0.020	−0.719	0.155	0.043	0.291	−0.655	−0.050	—							
Ma	60	0.069	−0.197	0.100	0.124	−0.323	.	0.150	0.403	.	−0.834	0.223	0.046	—						
BDI	585	**0.473**	**−0.476**	**0.356**	**0.536**	**0.546**	0.812	**0.447**	**0.625**	0.515	0.378	0.348	0.182	**0.522**	—					
F	585	0.233	−0.259	0.150	0.330	0.315	0.283	0.175	0.282	0.327	−0.297	−0.132	0.003	0.208	0.578	—				
R	585	−0.045	−0.123	−0.004	0.188	0.140	0.267	0.122	0.244	0.126	−0.370	0.198	−0.013	0.158	0.225	0.179	—			
S	585	−0.283	−0.036	−0.046	0.025	0.125	−0.120	0.007	0.194	0.057	0.201	0.033	0.195	0.101	−0.095	−0.194	0.020	—		
T	585	0.188	−0.036	−0.037	0.289	−0.006	0.192	0.157	0.204	0.166	−0.287	0.276	−0.011	−0.013	0.321	0.397	0.044	−0.211	—	
El	585	0.130	−0.012	0.094	0.299	0.201	−0.251	0.025	0.039		−0.423	−0.389	−0.228	−0.170	0.038	0.133	0.439	−0.338	0.162	—
RE	585	0.133	0.024	0.125	0.098	0.083	0.047	−0.062	−0.046	−0.328	−0.182	−0.152	−0.196	−0.031	−0.049	0.096	−0.420	−0.360	−0.003	**0.837**

Table 5. Parameter estimates for the environmental variables (linear and/or quadratic terms) that entered in the multiple regression analyses of species diversity within each species group. The number between brackets denotes the order in which the variable was entered into the model. Codes for the environmental variables are given in Table 2.

	Plants	Dragonflies	Butterflies	Herpetofauna	Birds
Ur	0.061***(4)	–	–	–	–
Ur^2	–	–	–	−0.010*(2)	–
Ar	–	−0.074*(8)	–	–	–
Dw	–	0.022*(9)	–	0.070**(3)	–
Dw^2	0.005***(3)	–	–	−0.018*(4)	–
Mw	–	0.043***(2)	–	–	0.012**(3)
Mw^2	−0.001$^{n.s.}$(2)	–	0.007***(1)	–	–
Gr^2	0.015***(5)	–	–	–	−0.013**(2)
Wc	–	−0.028*(7)	–	−0.019*(5)	–
Wb	0.011***(6)	–	–	–	–
Wb^2	–	0.015***(3)	–	–	0.005**(4)
BDI	0.240***(1)	0.659**(1)	0.320***(2)	–	0.297***(1)
BDI^2	–	–	–	1.269***(1)	–
S	−1.126***(7)	–	–	–	–
S^2	–	−0.573***(4)	–	–	–
T	−2.104$^{n.s.}$(8)	–	–	–	–
RE	–	0.414**(6)	–	–	–
RE^2	–	−0.296**(5)	–	–	–
Model R^2	0.656***	0.596***	0.442***	0.368***	0.475***
Spearman r evaluation set	0.639***	0.779***	0.648***	0.492***	0.454***

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

Table 6. Pairwise correlation coefficients (Spearman rank correlation) for the predicted species diversity among the different taxonomic groups. The correlations are based on predicted species richness in all squares ($N = 585$). Between brackets: the number of squares in common in the top 100 most species-rich squares.

	Plants	Dragonflies	Butterflies	Herpetofauna
Dragonflies	0.563*** (35)	–		
Butterflies	0.589*** (24)	0.872*** (77)	–	
Herpetofauna	0.489*** (33)	0.656*** (75)	0.707*** (77)	–
Birds	0.606*** (37)	0.898*** (79)	0.937*** (76)	0.666*** (69)

*** $p < 0.001$.

Dijle and Scheldt (Figure 2a). Species diversity of dragonflies, butterflies, herpetofauna and breeding birds was predicted to be high mainly in the Campine regions and in some scattered squares in the Sandy Flanders region (Figures 2b–e).

Correlations of the geographic pattern of predicted species diversity among the five species groups is given in Table 6. On average, correlations among the four faunal groups were clearly higher than correlations between each of the faunal
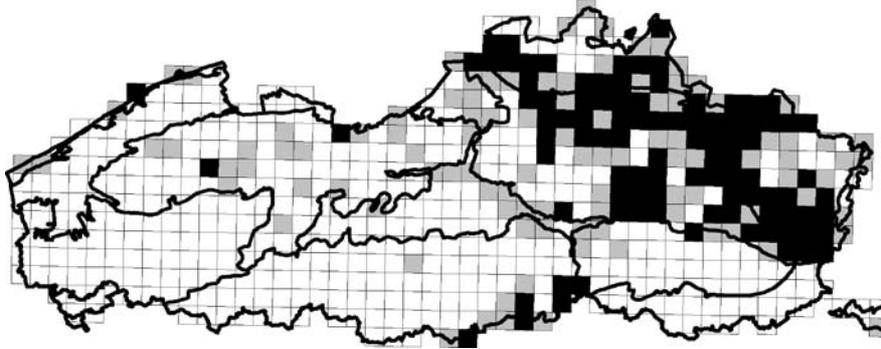
Figure 3. Geographic pattern of the average predicted species richness. In black the top 100 most species rich squares, in grey the next 100 most species rich squares.

groups and predicted plant species richness (average Spearman r among the four faunal groups = 0.789; average Spearman r of plants with the four faunal groups = 0.562).

The distribution pattern of the mean predicted species diversity for all squares in Flanders showed a prominent concentration of species-rich squares in the Campine regions (Figure 3). Other areas with a less pronounced aggregation of squares with a predicted high mean species diversity were found in the Dunes region, the south-eastern part of the Western and Central Hills region, the northern part of the Dender–Klein Brabant region and scattered over the Sandy Flanders region (Figure 3).

Discussion

Despite considerable efforts that were put into the separate survey schemes, only a relatively low number of squares was adequately surveyed for all of the species groups in common. This hampers exploring the coincidence among species groups in the geographic patterns of observed species diversity. Therefore, we adopted an alternative approach using predictive modelling of species diversity. We here discuss some methodological aspects, the main results of the predictive models and the relevance of our findings for nature conservation.

Reliability of species distribution data

Our approach ideally requires the simultaneous collection of data on species diversity for each of the taxonomic groups and for the environmental variables. Although collecting periods for the different data sets inevitably differed, they

overlap largely (Table 1), such that it is reasonable to assume that any discrepancies did not affect the outcome of our analyses.

The largest difference in survey periods was between breeding birds, which were studied recently and in a short time-span (i.e., 1999–2002), and the other species groups. For the latter taxa, it was necessary to lump information collected over a longer period to obtain adequate geographic coverage of the data. A drawback of accumulating survey data over long periods is that they may include data on species that went extinct after the initial years of the mapping period. Hence, for the taxonomic groups that were surveyed over relatively long time periods (i.e., plants, dragonflies and herpetofauna), our data may have overestimated the present-day species diversity in some of the squares. However, the majority of the recent local species extinctions in Flanders occurred in the period 1950–1970 (Bauwens and Claus 1996; De Knijf and Anselin 1996; Maes and Van Dyck 2001), that is, before the start of the mapping schemes of the taxonomic groups studied here. Hence, the limited number of local species extinctions that took place during the survey periods should not have had a substantial impact on our estimates of local species richness.

Geographic variation in sampling intensity is inevitable in survey schemes carried out by volunteers and may induce biases in the analyses. To minimise such biases, we built and evaluated the predictive models using data from the most frequently visited squares within the different ecological regions of Flanders. Moreover, the numbers of squares included in the analyses were proportional to the area of each of the ecological regions, such that the selected squares were distributed homogeneously over Flanders. This procedure reduces any biases induced by geographic variation in local species richness.

Modelling species diversity

Nature conservancy policy makers throughout the world have to base conservation strategies on incomplete and/or biased data (Samways 1993; Lobo et al. 1997), even in relatively well-surveyed countries or regions such as NW Europe (Dennis and Hardy 1999). Bias in the available data is caused by the unequal distribution of recording intensity (Dennis et al. 1999; Dennis and Thomas 2000). This may lead to non-optimal use of limited resources in nature conservation by wrongly prioritising the designation or acquisition of areas for conservation (Pearson and Carroll 1998; Myers et al. 2000). Through modelling techniques, we can upgrade biased and incomplete distribution databases by assessing the potential conservation value of unsurveyed or clearly undersurveyed sites (Lobo et al. 1997; Maddock and Du Plessis 1999).

The multiple regression analyses indicated that geographic variation of species diversity within each taxonomic group could be explained, albeit to a variable degree, by geographic variation in environmental variables in Flanders. Local species richness of all five taxonomic groups was more often correlated with land use variables than with climatic or topographic variables, a result that contrasts

other predictive models of species richness (e.g., Sparks et al. 1995; Fleishman et al. 2001). Compared to large-scale studies (e.g., continents) where variation in climate and topography is far more pronounced (e.g., Kerr et al. 1998; Badgley and Fox 2000), Flanders has little geographic variation in climate (e.g., mean maximum temperature ranges from 13.7–15.1 °C) and topography (elevation ranges from 1 to 237 m). Our analyses revealed that species richness of all five taxonomic groups in Flanders was positively correlated with biotope diversity. This finding emphasises the importance of the presence of different biotopes for species richness (Kerr and Packer 1997; Weibull et al. 2000).

To be reliable for nature conservation purposes, predictive modelling should always include a testing phase, preferably using an evaluation data set that is independent from the data used to build the models (Mac Nally 2000). Thus, we compared predicted to observed local species richness for a different set of squares, which were selected using the same criteria as for the selection of the squares in the training set. Complete independence between the training and evaluation set was probably not achieved here as both sets of squares were extracted from the same larger data set (Guisan and Zimmermann 2000). However, the restricted area of small regions like Flanders impedes achievement of truly geographic independence between training and evaluation data sets.

Correlations between observed and predicted species diversity were highly significant in all five species groups. The correlations were particularly high for the models of plants, dragonflies and butterflies and the geographic patterns of observed and predicted local species richness patterns largely coincided. Although significant, correlations between observed and predicted species diversity were lower for herpetofauna and birds. Species richness of the herpetofauna in Flanders may be more difficult to model because of the low number of species involved ($N = 19$) and the relatively large among-species differences in habitat preferences. The relatively large scale on which birds interact with their environment – vagrancy is higher in birds than in the other taxonomic groups – may make it more difficult to build a predictive model with the environmental variables used.

The significant rank correlations showed that the predictive models produced acceptable estimates of the rank order of species diversity within each species group. However, the distribution of the models residuals indicated that the models systematically underestimated the number of species in species-rich squares and overestimated species diversity in species-poor squares (cf. Lobo and Martín-Piera 2002). This indicates that the predictive models are not fully able to fit the interactions between local species richness and the environment on the scale used here (grid cells of 5×5 km). This can be due to the high degree of fragmentation of the Flemish landscape (EEA 2002) which renders predictive modelling more difficult, or to variables not accounted for in the present model (Moller and Jennions 2002). Possible missing variables are interaction terms between variables, biotope quality (in the present analyses only biotope quantity is entered) or higher order terms of the environmental variables (Legendre and Legendre 1998). Inclusion of higher order terms or interaction terms increases the models complexity and makes model interpretations difficult and/or spurious (Bio et al. 2002). However, this does not

invalidate our approach for conservation-oriented applications, which are based on relative differences in species diversity among areas, rather than on absolute numbers.

Coincidence of species diversity among taxonomic groups

Globally, reliable distribution data are available for at most a limited number of taxonomic groups. The lack of data for 'unpopular' species groups usually results in nature conservation strategies that are based upon data for a limited number of taxa (Prendergast et al. 1993). To overcome the problem of time-consuming – and hence expensive – surveys for a wide range of taxonomic groups, conservationists and policy makers apply the concept of indicator taxa, where one taxon is used as a surrogate for many others (Simberloff 1998; Caro and O'Doherty 1999). In many countries and regions, birds and plants have been used as indicator taxa (e.g., Niemela and Baur 1998; Blair 1999; Bibby 1999; Pharo et al. 1999). However, different studies have shown that the coincidence of species richness across taxa can be very low (e.g., Prendergast et al. 1993; van Jaarsveld et al. 1998; Andelman and Fagan 2000). These results call into question the reliability of the concept of indicator taxa for conservation purposes.

Our results indicate that correlations of geographic patterns of both observed and predicted species-richness among species groups were relatively high in Flanders compared to those reported in other studies carried out at larger (e.g., Prendergast et al. 1993; van Jaarsveld et al. 1998) or finer scales (e.g., Vessby et al. 2002). This does not appear to be related to the study scale, but more likely to the high pressure on land use which constraints species distribution to a restricted number of semi-natural sites in Flanders. The coincidence of the predicted local species richness was especially high among the four faunal groups, which had on average 76 squares in common among the top 100 most species-rich squares. This indicates that, in Flanders, the four investigated faunal groups can be used as fairly good indicator taxa for one another. On the other hand, the geographic coincidence in predicted species diversity is much lower between plants and each of the four faunal groups, with on average only 32 squares in common among the top 100 most species-rich squares. Hence, plant species diversity cannot be considered as a useful indicator of faunal species richness within Flanders.

Prioritising areas for conservation

The prioritisation and subsequent designation of areas for conservation often lacks quantitative scientific underpinning and is frequently based on 'best professional judgements' or on personal experience of local conservationists (Pullin and Knight 2001). For instance, until present, no attempts were made to integrate overall analyses of species distribution data into the designation of important conservation areas in Flanders. Rather, designation of most conservation areas in Flanders was

based on the presence of certain (threatened) biotopes, with strong constraints imposed by political and socio-economic considerations. It should therefore be questioned to what extent these conservation policies are expected to contribute to the objective of preserving species diversity at its present-day level. To answer this question, we explore the extent of geographic overlap between recently designated (complexes of) conservation areas and the patterns of predicted local species richness.

A first conservation policy that was recently (2001) implemented in Flanders was the designation of ca. 1630 km^2 as 'Natura 2000' areas, in compliance to the Habitat and Bird Directives of the European Union. The 'Natura 2000' network aims at preserving species diversity on a European scale and prioritises the conservation of internationally threatened species and biotopes. Because very few internationally threatened species occur in Flanders, the designation of 'Natura 2000' areas was primarily based on the presence of certain biotope types. Overlays of the map of designated 'Natura 2000' sites with the map of the mean predicted local species richness revealed that only 43 squares of the top 100 predicted most species-rich squares overlapped with the squares that contained at least 500 ha of 'Natura 2000' sites.

The regional government of Flanders recently (2003) also approved the designation of ca. 900 km^2 of conservation areas to create the 'Flemish Ecological Network' (FEN). One of the explicit objectives of this policy is the maintenance of the present-day species richness in Flanders. Nevertheless, selection of the FEN areas was mainly based on the presence of certain biotopes and no systematic consideration was given to species diversity. Overlays of the map of designated FEN sites with the map of the mean predicted local species richness revealed that only 46 of the top 100 predicted most species-diverse squares overlapped with the squares that contained at least 500 ha of FEN sites.

Although both conservation programmes (Natura 2000 and FEN) differ greatly in the geographic scale of their objectives, their designated areas overlap to almost identical extent with sites with the mean predicted species richness in Flanders. This is presumably a consequence of the fact that both programmes used similar criteria to designate areas, even though they have diverging objectives. Second, although the designated areas overlap to some extent with the sites with a high mean predicted local species richness, less than one half of the predicted most species-rich squares was adequately incorporated in the schemes. Consequently, we question whether the 'FEN' will achieve its objective of maintaining local species diversity at its present-day levels.

This example illustrates how decisions on the designation of conservation areas greatly benefit from predictive modelling performed at a local scale. We strongly believe that policy makers in Flanders, but also in other parts of the world, should make more use of modelling techniques to produce predicted maps of species richness among taxa as a proactive conservation tool because it allows to better target sites with a collective high species richness for different taxonomic groups. Furthermore, the simultaneous use of taxonomic groups representing organisms at different trophic levels (from nectar feeders to predators) and from both aquatic and

terrestrial biotopes at different levels of geographic scale (from very small [plants, invertebrates] to very large [birds of prey]), assures a representative sample for a wide variety of other uninvestigated taxa (cf. Vanderklift et al. 1998). Different avenues for application in future conservation programmes are under investigation at present: the prediction of local species richness on an even smaller scale then the one presented here (e.g., 1×1 km squares or in the optimal scenario, parcels – Fleishman et al. 2003) and the incorporation of taxonomic groups for which distribution data are less complete than for the ones studied here. Detailed land cover classification will become more readily available through remote sensing, such that it will become possible to perform similar analyses on relatively large regions (Kerr and Ostrovsky 2003; Turner et al. 2003). Future analyses should further explore the minimum number of squares and survey visits needed to adequately model species richness in a given region.

Acknowledgements

References

Andelman S.J. and Fagan W.F. 2000. Umbrellas and flagships: efficient conservation surrogates or expensive mistakes? Proceedings of the National Academy of Science USA 97: 5954–5959.

Badgley C. and Fox D.L. 2000. Ecological biogeography of North American mammals: species diversity and ecological structure in relation to environmental gradients. Journal of Biogeography 27: 1437–1467.

Bauwens D. and Claus K. 1996. Verspreiding van amfibieën en reptielen in Vlaanderen. De Wielewaal, Turnhout, Belgium.

Bibby C.J. 1999. Making the most of birds as environmental indicators. Ostrich 70: 81–88.

Bio A.M.F., De Becker P., De Bie E., Huybrechts W. and Wassen M. 2002. Prediction of plant species distribution in lowland river valleys in Belgium: modelling species response to site conditions. Biodiversity and Conservation 11: 2189–2216.

Blair R.B. 1999. Birds and butterflies along an urban gradient: surrogate taxa for assessing biodiversity? Ecological Applications 9: 164–170.

Caro T.M. and O'Doherty G. 1999. On the use of surrogate species in conservation biology. Conservation Biology 13: 805–814.

Clifford P., Richardson S. and Hémon D. 1989. Assessing the significance of the correlation between two spatial processes. Biometrics 45: 123–134.

Crawley M.J. 1993. GLIM for Ecologists. Blackwell Scientific Publications, Oxford, UK.

De Blust G. 2001. De ecoregio's. In: Kuijken E., Boeye D., De Bruyn L., De Roo K., Dumortier M., Peymen J., Schneiders A., van Straaten D. and Weyembergh G. (eds) Natuurrapport 2001. Toestand van de natuur in Vlaanderen: cijfers voor het beleid. Instituut voor Natuurbehoud, Brussel, Belgium, pp. 13–17.

Decleer K. and Vanroose S. 2001. Verwerving van natuurgebieden. In: Kuijken E., Boeye D., De Bruyn L., De Roo K., Dumortier M., Peymen J., Schneiders A., van Straaten D. and Weyembergh G. (eds)

Natuurrapport 2001. Toestand van de natuur in Vlaanderen: cijfers voor het beleid. Instituut voor Natuurbehoud, Brussel, Belgium, pp. 183–194.

De Knijf G. and Anselin A. 1996. Een gedocumenteerde Rode lijst van de libellen van Vlaanderen. Instituut voor Natuurbehoud, Brussel, Belgium.

Dennis R.L.H. and Hardy P.B. 1999. Targeting squares for survey: predicting species richness and incidence of species for a butterfly atlas. Global Ecology and Biogeography Letters 8: 443–454.

Dennis R.L.H. and Shreeve T.G. 2003. Gains and losses of French butterflies: tests of predictions, under-recording and regional extinction from data in a new atlas. Biological Conservation 110: 131–139.

Dennis R.L.H. and Thomas C.D. 2000. Bias in butterfly distributions maps: the influence of hot spots and recorder's home range. Journal of Insect Conservation 4: 73–77.

Dennis R.L.H., Sparks T.H. and Hardy P.B. 1999. Bias in butterfly distribution maps: the effects of sampling effort. Journal of Insect Conservation 3: 33–42.

Dennis R.L.H., Shreeve T.G., Sparks T.H. and Lhonore J.E. 2002. A comparison of geographical and neighbourhood models for improving atlas databases. Biological Conservation 108: 143–159.

Dobson A.P., Rodriguez J.P., Roberts W.M. and Wilcove D.S. 1997. Geographic distribution of endangered species in the United States. Science 275: 550–553.

Dries L. 2002. Natura 2000 in Vlaanderen: een schakel in een Europees netwerk. Ministerie van de Vlaamse Gemeenschap – Afdeling Natuur i.s.m. WWF & Natuurpunt, Brussel, Belgium.

Dutilleul P. 1993. Modifying the t-test for assessing the correlation between two spatial processes. Biometrics 49: 305–314.

Edwards T.C., Deshler E.T., Foster D. and Moisen G.G. 1996. Adequacy of wildlife habitat relation models for estimating spatial distributions of terrestrial vertebrates. Conservation Biology 10: 263–270.

EEA 2002. Environmental signals 2002. Benchmarking the millennium. European Environment Agency, Copenhagen, Denmark.

Fleishman E., Austin G.T. and Murphy D.D. 2001. Biogeography of Great Basin butterflies: revisiting patterns, paradigms, and climate change scenarios. Biological Journal of the Linnean Society 74: 501–515.

Fleishman E., Mac Nally R. and Fay J.P. 2003. Validation tests of predictive models of butterfly occurrence based on environmental variables. Conservation Biology 17: 806–817.

Gower J.C. 1971. A general coefficiënt of similarity and some of its properties. Biometrics 27: 857–871.

Guisan A. and Zimmermann N.E. 2000. Predictive habitat distribution models in ecology. Ecological Modelling 13: 147–186.

Isaaks E.H. and Srivastava R.M. 1989. An introduction to applied geostatistics. Oxford University Press, Oxford, UK.

Keitt T.H., Bjornstad O.N., Dixon P.M. and Citron-Pousty S. 2002. Accounting for spatial pattern when modeling organism-environment interactions. Ecography 25: 616–625.

Kerr J.T. and Ostrovsky M. 2003. From space to species: ecological applications for remote sensing. Trends in Ecology and Evolution 18: 299–305.

Kerr J.T. and Packer L. 1997. Habitat heterogeneity as a determinant of mammal species richness in high-energy regions. Nature 385: 252–254.

Kerr J.T., Vincent R. and Currie D.J. 1998. Lepidopteran richness patterns in North America. Ecoscience 5: 448–453.

Legendre P. and Legendre L. 1998. Numerical Ecology. Elsevier, Amsterdam, The Netherlands.

Littell R.C., Milliken G.A., Stroup W.W. and Wolfinger R.D. 1996. SAS System for Mixed Models. SAS Institute Inc., Cary, NC.

Lobo J.M. and Martín-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. and JayRobert P. 1997. Taxonomic databases as tools in spatial biodiversity research. Annales de La Societe Entomologique de France 33: 129–138.

Luoto M., Toivonen T. and Heikinnen R.K. 2002. Prediction of total and rare plant species richness in agricultural landscapes from satellite images and topographic data. Landscape Ecology 17: 195–217.

Mac Nally R. 2000. Regression and model-building in conservation biology, biogeography and ecology: the distinction between – and reconciliation of – 'predictive' and 'explanatory' models. Biodiversity and Conservation 9: 655–671.

Mac Nally R., Fleishman E., Fay J.P. and Murphy D.D. 2003. Modelling butterfly species richness using mesoscale environmental variables: model construction and validation for mountain ranges in the Great Basin of western North America. Biological Conservation 110: 21–31.

Maddock A. and Du Plessis M.A. 1999. Can species data only be appropriately used to conserve biodiversity? Biodiversity and Conservation 8: 603–615.

Maes D. and Van Dyck H. 2001. Butterfly diversity loss in Flanders (north Belgium): Europe's worst case scenario? Biological Conservation 99: 263–276.

Maes D., Gilbert M., Titeux N., Goffart P. and Dennis R. 2003. Prediction of butterfly diversity hotspots in Belgium: a comparison of statistically-focused and land use-focused models. Journal of Biogeography 30: 1907–1920.

Magurran A.E. 1988. Ecological diversity and its measurement. Princeton University Press, Princeton, New Jersey.

Moller A.P. and Jennions M.D. 2002. How much variance can be explained by ecologists and evolutionary biologists? Oecologia 132: 492–500.

Myers N., Mittermeier R.A., Mittermeier C.G., daFonseca G.A.B. and Kent J. 2000. Biodiversity hotspots for conservation priorities. Nature 403: 853–858.

Nicholls A.O. 1989. How to make biological surveys go further with generalised linear models. Biological Conservation 50: 51–75.

Niemela J. and Baur B. 1998. Threatened species in a vanishing habitat: plants and invertebrates in calcareous grasslands in the Swiss Jura mountains. Biodiversity and Conservation 7: 1407–1416.

OECD 1998. Environmental performance reviews Belgium. OECD Editions, Paris, France.

Overmars K.P., de Koning G.H.J. and Veldkamp A. 2003. Spatial autocorrelation in multi-scale land use models. Ecological Modelling 164: 257–270.

Pearson D.L. and Carroll S.S. 1998. Global patterns of species richness: spatial models for conservation planning using bioindicator and precipitation data. Conservation Biology 12: 809–821.

Pearson D.L. and Cassola F. 1992. World-wide species richness patterns of tiger beetles (Coleoptera: Cicindelidae): indicator taxon for biodiversity and conservation studies. Conservation Biology 6: 376–391.

Pharo E.J., Beattie A.J. and Binns D. 1999. Vascular plant diversity as a surrogate for bryophyte and lichen diversity. Conservation Biology 13: 282–292.

Pimm S.L., Russell G.J., Gittleman J.L. and Brooks T.M. 1995. The future of biodiversity. Science 269: 347–350.

Prendergast J.R., Quinn R.M. and Lawton J.H. 1999. The gaps between theory and practice in selecting nature reserves. Conservation Biology 13: 484–492.

Prendergast J.R., Quinn R.M., Lawton J.H., Eversham B.C. and Gibbons D.W. 1993. Rare species, the coincidence of diversity hotspots and conservation strategies. Nature 365: 335–337.

Pullin A.S. and Knight T.M. 2001. Effectiveness in conservation practice: pointers from medicine and public health. Conservation Biology 15: 50–54.

Reid W.V. 1998. Biodiversity hotspots. Trends in Ecology and Evolution 13: 275–280.

Samways M.J. 1993. Insects in biodiversity conservation: some perspectives and directives. Biodiversity and Conservation 2: 258–282.

Simberloff D. 1998. Flagships, umbrellas, and keystones: is single-species management passé in the landscape era? Biological Conservation 83: 247–257.

Sparks T.H., Dover J.W., Warren M.S. and Cox R. 1995. How well can we model the distribution of butterflies at the landscape scale? In: Griffith G.H. (ed) Landscape Ecology: Theory and Applications. IALE (UK), Aberdeen, pp. 24–31.

Turner W., Spector S., Gardiner N., Fladeland M., Sterling E. and Steininger 2003. Remote sensing for biodiversity science and conservation. Trends in Ecology and Evolution 18: 306–314.

Van Hecke E. and Dickens C. 1994. Bevolking. In: Verbruggen A. (ed) Leren om te keren, Milieu- en natuurrapport Vlaanderen. Vlaamse Milieumaatschappij en Garant Uitgevers N.V., Leuven/Apeldoorn, pp. 45–61.

1364

van Jaarsveld A.S., Freitag S., Chown S.L., Muller C., Koch S., Hull H. et al., 1998. Biodiversity assessment and conservation strategies. Science 279: 2106–2108.

Vanderklift M.A., Ward T.J. and Phillips J.C. 1998. Use of assemblages derived from different taxonomic levels to select areas for conserving marine biodiversity. Biological Conservation 86: 307–315.

Van Landuyt W., Heylen O., Vanhecke L., Van den Bremt P. and Baeté H. 2000. Verspreiding en evolutie van de botanische kwaliteit van ecotopen: gebaseerd op combinaties van indicatorsoorten uit Florabank. Flo.Wer vzw, Instituut voor Natuurbehoud, Nationale Plantentuin, Universiteit Gent, Brussel/Meise/Gent, Belgium.

Vessby K., Söderström B., Glimskär A. and Svensson B. 2002. Species-richness correlations of six different taxa in Swedish seminatural grasslands. Conservation Biology 16: 430–439.

Weibull A.C., Bengtsson J. and Nohlgren E. 2000. Diversity of butterflies in the agricultural landscape: the role of farming system and landscape heterogeneity. Ecography 23: 743–750.