Modelling the species richness distribution for French Aphodiidae (Coleoptera, Scarabaeoidea)

Jorge M. Lobo, Pierre Jay-Robert and Jean-Pierre Lumaret

The species richness distribution of the French Aphodiidae was predicted using Generalized Linear Models to relate the number of species to spatial, topographic and climate variables. The entire French territory was studied, divided into 301 0.72×0.36 degree grid squares; the model was developed using 66 grid squares previously identified as well sampled. After eliminating nine outliers, the final model accounted for 74.8% of total deviance with a mean Jackknife predictive error of 10.5%. Three richest areas could be distinguished: the western head (Brittany), southwestern France, and, to a lesser extent, the northeastern region. Sampling effort should now be focused on the western head, where no square was correctly sampled, and on southwestern France, which was recognised as a diversity hotspot, both for Aphodiidae and for Scarabaeidae. The largest fraction of variability (37%) in the number of species was accounted for by the combined effect of the three groups of explanatory variables. After controlling for the effect of significant climate and topographic variables, spatial variables still explain 27% of variation in species richness, suggesting the existence of a spatial pattern in the distribution of species richness (greater diversity in western France) that can not be explained by the environmental variables considered here. We hypothesize that this longitudinal spatial pattern is due to the relevance of a western colonization pathway along the glacial-interglacial cycles, as well as by the barrier effect played by the Alps.

J. M. Lobo (mcnj117@mncn.csic.es), Dept de Biodiversidad y Biología Evolutiva, Museo Nacional de Ciencias Natuxales (CSIC), Cl José Gutiérrez Abascal 2, E-28006 Madrid, Spain. – P. Jay-Robert and J.-P. Lumaret, Dept de Biologie, UFR III, Univ. Paul Valéry, UMR 5175, Route de Mende, F-34199 Montpellier Cedex 5, France.

Recently there has been growing interest in the potential of different statistical modelling methods to forecast the distribution of species in the absence of exhaustive data (Manel et al. 1999a, b, Pearce and Ferrier 2000, Vayssières et al. 2000, Hirzel et al. 2001, Guisan et al. 2002, Stockwell and Peterson 2002). Application of such techniques could, perhaps, provide reliable knowledge about the distribution of biological diversity in a territory for conservation or scientific purposes (Ferrier 2002). An extensive literature employs predictive models using presence-only data or presence/absence data to estimate the probability of occurrence where no species information exist (for example, Mitchell 1991, Pereira and Itami 1991, Buckland and Elston 1993, Iverson and Prasad 1998a, Peterson et al. 1999, Parker 1999). Thus, to infer the distribution of a biodiversity attribute like the species richness of a taxonomic group it is necessary to overlay the predicted individual distributions. Unfortunately, it is almost impossible to elaborate reliable individual models if species geographical information is scarce and biased (for example, in the case of rare or endemic species; see Dennis and Hardy 1999, Soberón et al. 2000 or Ferrier 2002). If rare species are responsible for an important percentage of total species richness

(Gaston 1996), ignoring them homogenizes the real differences among localities in the diversity of species, diminishing the conservation utility of the derived predictive maps. In this case the most reasonable alternative may be to directly model the collective properties of biodiversity as species richness (Margules et al. 1987, Bojorquez-Tapia et al. 1996, Leathwick et al. 1998, Iverson and Prasad 1998a,b, Zimmermann and Kienast 1999, Hortal et al. 2001, Lobo and Martín-Piera 2002, Lobo et al. 2002), rarity or faunistic composition and use these estimates as surrogates in conservation planning (Austin 1998, Ferrier 2002, Ferrier et al. 2002) or study environmental or historical factors related with the distribution of these attributes (Lobo et al. 2002).

Taking into account information about species richness from previously defined well-surveyed squares, a predictive model for the Aphodiidae species richness in France is designed using environmental data as predictive variables and a GLM modelling technique. Thanks to a long-standing naturalist tradition, the species richness distribution of some insect groups is relatively well known in France. Aphodiidae, which have been studied for a long time (Paulian and Baraud 1982, Baraud 1992), is one such group (Lumaret 1990, Lumaret et al. 1996). Using the French Scarabaeoidea Laparosticti database (37 300 records), the asymptotic relationship between the number of species and the number of records in 0.72×0.36 degree grid squares (about 2230 km^2 each) was analysed for each one of the five biogeographic subprovinces. These asymptotic curves allow one to obtain an estimate of the number of species present in the squares of each subregion. Subsequently, the number of database records required to obtain at least 70% of the estimated number of species was calculated for each subregion; squares with such a number of database records (66 squares, 22% of 301) were considered well-inventoried (Lobo et al. 1997). As no large inventorying is planned at present, it would now be particularly useful to obtain a predictive map of the distribution of Aphodiidae species richness. The species richness distribution obtained will be mapped and discussed.

In addition, the most important determinants of French Aphodiidae species richness were estimated by decomposing its variation into spatial, topographic, and climatic components in order to provide correlations to possible causal factors that have led to the current species richness distribution. When the aim is to forecast, the frequent collinearity between explanatory environmental variables is not critical, because the objective is to maximize the explained variation of the data (Legendre and Legendre 1998). However, if we want to explore the influence of these same explanatory variables on species number variation, the collinearity may bias model parameter estimation, hindering the always difficult task of identifying the likely causal relationships under-

lying multivariate environmental and spatial correlations (Levin 1992, Legendre and Legendre 1998, Shipley 1999). As manipulative experiments and causal modelling are almost impossible approximations in the case of larger spatial scale studies, factors correlated with the number of species are frequently the starting point for the identification of factors that might possibly explain geographical species richness patterns (see Richerson and Lum 1980, Austin et al. 1984, Owen 1989, Tonteri 1994, Birks 1996, Heikkinen and Neuvonen 1997, Fraser 1998, Qian 1998, Wohlgemuth 1998, Kerr and Packer 1999, Lobo et al. 2001, Lobo et al. 2002). However, if different explanatory variable groups are analysed jointly, it is convenient to estimate, by partial regression analysis, the pure influence of each variable, not correlated with that of the other variables, and the joint influence of the various pairs of explanatory variables (Borcard et al. 1992, Legendre 1993, Pinel-Alloul et al. 1995, Qinghong and Bråkenhielm 1995, Heikkinen and Birks 1996, Anderson and Gribble 1998, Legendre and Legendre 1998, Lobo et al. 2001, Lobo et al. 2002). Although such a type of variation decomposition would not necessarily identify causal factors, it may facilitate the formulation of hypotheses about the processes that may have generated the observed patterns.

Materials and methods

Selected environmental variables

The French territory (France mainland and Corsica island) was divided into 301 0.72×0.36 degree grid squares (Fig. 1). For each square, six climatic, two topographic and two spatial variables were used. All variables are continuous and were standardised to eliminate the effect of differences in measurement scale. The sources and the method of obtaining this environmental and spatial information were described in Lobo et al. (2002). The variables employed are: mean annual temperature (TM), annual temperature variation (TV), maximum monthly mean temperature (T), minimum monthly mean temperature (t), mean annual precipitation (P), annual precipitation variation (PV), mean elevation (E), altitude range (AR), central latitude (LAT) and central longitude (LON).

Model building

Generalised Linear Models (GLM; McCullagh and Nelder 1989, Dobson 1990) have been used to summarize the relationship between number of Aphodiidae species and the above-mentioned explanatory variables (e.g., Nicholls 1989, Austin et al. 1990, 1996, Heikkinen and Neuvonen 1997, Guisan et al. 2002). GLMs are an extension of linear models that allow for non-linearity in
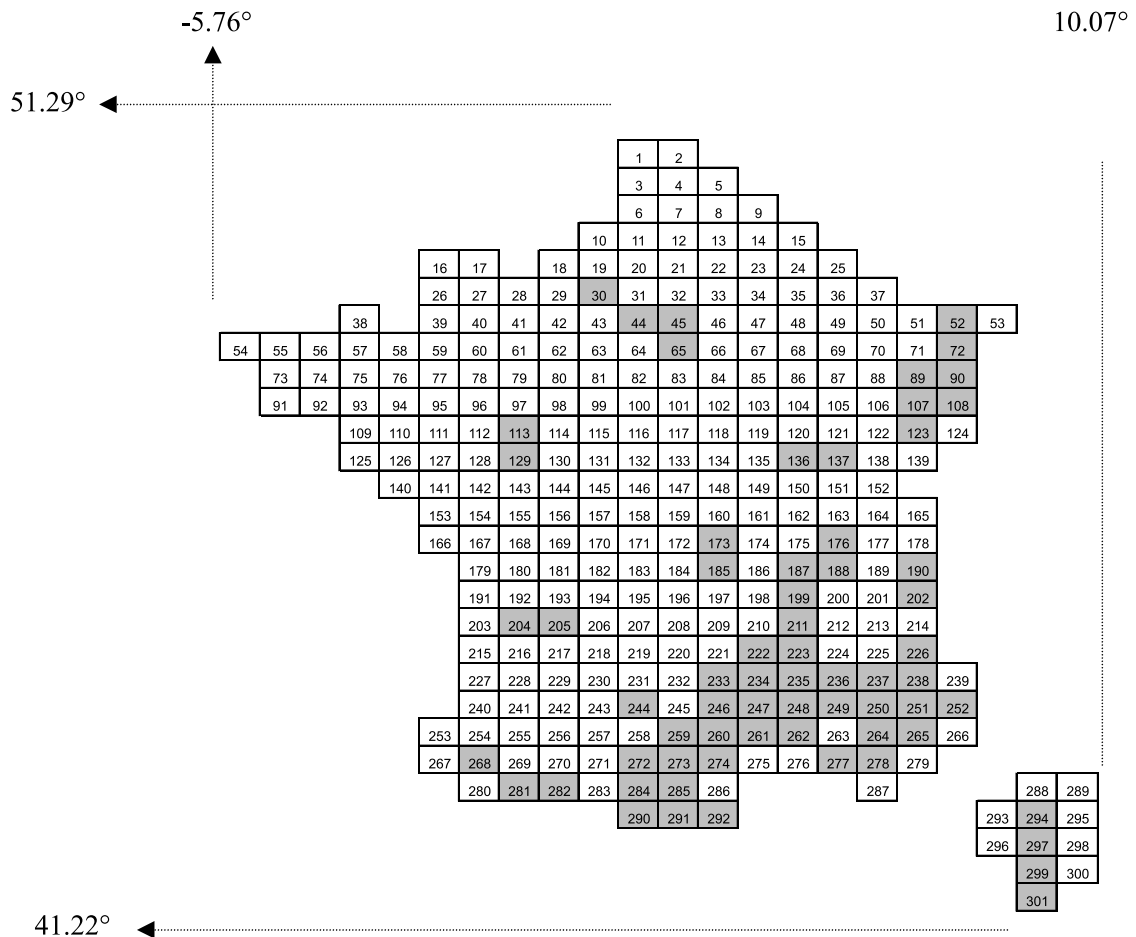
146

Fig. 1. Continental France and Corsica Island, divided into 301 squares of 0.72 × 0.36 degrees. Broken lines represent maximum and minimum latitude and longitude. Shaded squares are the 66 well-sampled grid squares (Lobo et al. 1997).

the data and also a wide range of model specification distributions other than the normal distribution of the random component. In our case the assumed relationship between the number of species and the explanatory variables is logarithmic (the link function), while a Poisson error distribution for the number of Aphodiidae species was assumed (see Crawley 1993).

The total number of Aphodiidae in the 66 well-surveyed 0.72 × 0.36 degree grid squares (Lobo et al. 1997) was first related separately to each environmental variable, selecting either a linear, a quadratic or a cubic function of each environmental variable in order to consider possible curvilinear relationships (Austin 1980). A forward-stepwise procedure was used to enter the variables into the model (see Nicholls 1989, Austin et al. 1996). First, the linear, quadratic or cubic function of the variable that accounts for the most important change in deviance was selected for inclusion. Next, all the remaining variables were tested for significance, adding them sequentially, according to estimated importance. The procedure was iteratively repeated until no more

statistically significant explanatory variables remained (p ≤ 0.05). After each inclusion, the significance of the terms previously selected was also tested, submitting the new model to a backward selection procedure in order to exclude the terms that had become non-significant. The significance of all pairwise interaction terms between the explanatory variables (44 interaction terms) was also tested (see Margules et al. 1987), adding them one by one to the model in order of contribution to change in deviance. This procedure was iteratively repeated until no more significant interaction terms remained. Finally, third-degree polynomial equation of the central latitude (Lat) and longitude (Lon) of each square were included in the model (Trend Surface Analysis, see Legendre 1993) because spatial structures can incorporate the effects caused by other historical, biotic or environmental variables not otherwise taken into consideration (Legendre and Legendre 1998). Latitude and longitude were centred on their respective means. We performed a backward stepwise regression with the nine terms of the equation as predictor variables and the number of

Aphodiidae species as dependent variable in order to remove the non-significant spatial terms. The significant spatial terms (p ≤ 0.05) were retained and included in the previous model. The Statistica package (Anon. 1999) was used for all statistical computations.

Autocorrelation, residuals and model validation

On a geographical scale most environmental variables are autocorrelated, so that broad-scale spatial structure becomes an important source of information that must be taken into account to improve the predictive power of models (Smith 1994, Legendre and Legendre 1998). When autocorrelated variables are used, it is critical to examine the spatial independence of errors (Cliff and Ord 1981, Odland 1988), for if the residuals are spatially autocorrelated, the model should probably be enlarged to include some spatially structured variables not otherwise taken into account by the model (Cliff and Ord 1981, Odland 1988). Thus, the autocorrelation of the residuals of the final model was checked by calculating Moran's I spatial autocorrelation statistic with a Bonferroni-corrected significance level (Cliff and Ord 1981, Sawada 1999).

The adequacy of the final model was improved by detecting the outliers that disproportionately influence the fit of the model, examining the standard errors of coefficients, plotting the residuals vs the fitted values, plotting the residuals in a normal probability plot, and by examining the leverage of the observations (a measure of the distance of each observation from the centroid of the multi-dimensional space defined by the variables included in the model; Nicholls 1989, Dobson 1990).

A Jackknife procedure was applied in order to validate the final model, so from the data set of the 66 well-sampled grid squares, the model was recalculated 66 times leaving out one square in turn. Each one of the regression models based on the n − 1 grid squares was then applied to that excluded square in order to predict species richness scores for each one of the squares. The predictive power of the model was checked both by examining the correlation between the recorded and the Jackknife-predicted species richness, and by estimating the percentage error for each observation. The mean of the 66 estimates of error provides a measure of the prediction error associated with the model and data used.

A more extensive discussion of the statistical considerations contemplated for the construction of this model can be found in Lobo et al. (2002).

Variation partitioning

The significant terms of the explanatory variables selected in the first step of the regression analysis were further analysed to determine the comparative influence of spatial (S), topographic (T), and climatic (C) variables in variation of Aphodiidae species richness. The total variation within the number of species was decomposed among these three types of explanatory variable and the percentage of explained deviance calculated for eight different components (Fig. 2A): pure effect of spatial variation alone (a), pure effect of topography alone (b), pure effect of climate alone (c), combined variation due to the joint effect of spatial and topographic components (d), combined variation due to the joint effect of spatial and climate components (e), combined variation due to the joint effect of topographic and climate components (f), combined variation due to the joint effect of the three components (g), and variation not explained by the independent variables included in the analysis (U). The decomposition of the variation in the species richness into the three sets of explanatory variables was carried out by means of a partial regression analysis (Legendre and Legendre 1998). A complete explanation of this process of variation decomposition can be obtained in Borcard et al. (1992), Qinghong and Bråkenhielm (1995), Anderson and Gribble (1998) or Lobo et al. (2001).

Results

The predictive model

In the model using all the well-surveyed squares (n = 66), three environmental variables, minimum mean temperature, mean annual precipitation, and precipitation variation (Table 1) were statistically significant when tested
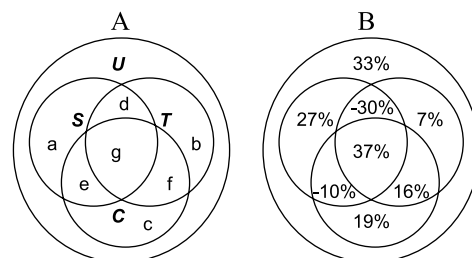

Fig. 2. A. Diagram showing the variation in the dependent variable (number of species) due to three groups of explanatory variables: S (space), T (topography), and C (climate); U is the unexplained variation; a, b and c are the pure effects of spatial variation, topography, and climate, respectively; d, e and f are the combined variation due to the joint effect of space and topography, the joint effect of space and climate, and the joint effect of topography and climate, respectively; g is the combined variation due to the joint effect of the three variables. B. Variation partitioning in the number of French Aphodiidae species between the three set of explanatory variables. The decomposition of the variation has been carried out by partial regression analysis (Legendre and Legendre 1998) using Generalized Linear Models (Crawley 1993) and the percentage of total explained deviance as the measure of explained variation.

Table 1. Deviance and change in the deviance from a full model for Aphodiidae species number associated with explanatory variables in the 0.72 × 0.36 degree grid squares of France. The linear, quadratic or cubic functions of each variable that account for a significant change in the deviance (p < 0.05) have been selected. The sign column indicates the sign for each selected term of each function. Spatial variables were chosen after removing the non-significant terms of a third degree polynomial equation of latitude and longitude by a backward stepwise selection. *** p < 0.001; ** p < 0.01; * p < 0.05.

Variables	Code	Selected terms	Deviance	DF	Change in deviance	F	Sign
Full model			182.73	65			
Longitude	Lon }	$Lat + Lat^3 + Lon^2 + Lon^3$					
Latitude	Lat ʃ		122.13	61	60.59	30.26***	$- + - +$
Mean elevation	E	E	174.69	64	8.04	2.95	
Altitude range	AR	AR	174.64	64	8.09	2.97	
Mean annual temperature	TM	TM	182.73	64	0.00	0.00	
Annual temperature variation	TV	TV	182.62	64	0.10	0.04	
Maximum monthly mean temperature	T	T	180.43	64	2.30	0.81	
Minimum monthly mean temperature	t	$t + t^2$	164.63	63	18.10	6.93*	$+ -$
Mean annual precipitation	P	P	160.43	64	22.30	8.90**	$-$
Annual precipitation variation	PV	$PV + PV^2$	162.49	63	20.24	7.85**	$- +$

separately as linear, quadratic or cubic functions. With regard to the spatial variables, the backward stepwise regression of the third-degree polynomial terms of latitude and longitude selected the linear and cubic terms of latitude, and the quadratic and cubic terms of longitude (Table 1).

As the linear term of mean annual precipitation accounted for the most important change in deviance, this variable was the first included in the model (Table 2A). In the second step, the quadratic function of minimum monthly mean temperature was selected but the backward procedure eliminated the linear term of this function. Next, the quadratic function of annual precipitation variation was selected but the linear term of this last function was excluded by the backward procedure. None of the remaining environmental variables was significant when added to the model. After fitting the significant environmental variables, the importance of interaction terms added one-by-one was tested. The interaction between minimum monthly mean temperature and annual precipitation variation (t × PV) accounted for the most significant change in deviance. The product of mean elevation and maximum monthly mean temperature (E × T) was the second significant interaction term that entered in the model. The inclusion of the interaction between latitude and mean annual precipitation (Lat × P) removed the quadratic terms of annual precipitation variation and minimum monthly mean temperature (Table 2A). Finally, the significant spatial terms (Table 1) were added to the model, and only the quadratic term of longitude was selected. Thus, model 1 was:

$$S = \exp (c + P + t \times PV + E \times T + Lat \times P + Lon^2)$$

where S is the total number of predicted Aphodiidae species and c is the intercept. This model explains 46.2% of total deviance (see coefficients and standard errors in Table 2A).

A Jackknife test procedure on this model showed that the correlation between observed and Jackknife-predicted values is significant and positive (Pearson correlation coefficient: r = 0.655, p < 0.001; Spearman rank correlation: rs = 0.615, p < 0.0001). However, the mean Jackknife predictive error of the observations was relatively high (21.2%; 95% confidence interval from 11.4% to 31.1%). Moreover, three squares showed positive standard residuals higher than +1.5 (fitted values lower than observed) and seven squares showed negative standard residuals lower than −1.5 (fitted values higher than observed, Fig. 3). The nine squares (out of ten) with high residuals (standard Pearson residual > 1.5) and small leverage scores (< 0.15; environmentally redundant cells) were deleted (squares 108, 129, 187, 211, 236, 260, 261, 274 and 301 in Fig. 1) and the modelling procedure done again. These squares contain surprisingly more or fewer species than predicted by the explanatory variables, probably because the number of database records for each square may not always reflect similar sampling effort.

The successive steps of the second model calculation (n = 57) were generally different, but four out of the five selected terms were shared by the two models (Table 2B). The interaction between annual precipitation variation and annual temperature variation was selected, whereas the addition of spatial terms to the model did not significantly increase the percentage of explained deviance. Thus, model 2 was:

$$S = \exp (3.521 - 0.234 P - 0.143 t \times PV + 0.019 PV$$
$$\times TV + 0.082 E \times T - 0.003 Lat \times P)$$

This model explains a higher percentage of the total deviance than model 1 (74.8% and 46.2%, respectively) (see full coefficients and standard errors in Table 2B). The Jackknife test on model 2 showed a higher correlation between observed and Jackknife-predicted values than model 1 (Pearson correlation coefficient: r = 0.843, p < 0.0001; Spearman rank correlation: rs = 0.815, p < 0.0001). The mean Jackknife predictive error of observations was reduced to 10.5% (95% confidence interval ranking between 8.0% and 13.0%), and the examination

Table 2. Summary of the stepwise forward selection of variables to build the models for the Aphodiidae species number in the 0.72 × 0.36 degree grid squares, (A) with the 66 squares considered as well surveyed (Lobo et al. 1997), and (B) eliminating the nine grid squares as outliers in the former model (n = 57). The change in deviance after including a term in the model has been tested by a F-ratio test with a probability < 0.05. Variable codes as in Table 1.

|	DF	Deviance	Change in deviance	F	% Explained deviance	Coefficients	Standard error
A							
Full model							
Intercept	65	182.73					
						3.534	0.032
Step 1							
P	64	160.43	22.30	8.90**	12.20	− 0.136	0.021
Step 2							
+t	63	154.64	5.79	2.36	15.37		
+t^2	62	141.35	13.29	5.83*	22.65		
− t	63	143.31	1.96	0.86	21.57		
Step 3							
+PV	62	140.86	0.48	0.21	21.84		
+PV^2	61	127.34	13.53	6.48*	29.24		
− PV	62	130.76	3.42	1.62	27.37		
Step 4							
+t × PV	61	115.17	15.59	8.26**	35.90	− 0.158	0.022
Step 5							
+E × T	60	105.19	9.98	5.69*	41.36	0.090	0.022
Step 6							
+Lat × P	59	97.88	7.32	4.41*	45.37	− 0.052	0.011
− PV^2	60	99.20	1.32	0.80	44.65		
− t^2	61	101.31	2.10	1.27	43.50		
Step 7							
+Lon^2	60	96.39	4.92	3.06*	46.19	− 0.008	0.003
B							
Full model							
Intercept	56	108.35					
						3.521	0.029
Step 1							
P	55	75.86	32.49	23.56***	30.00	− 0.234	0.030
+P^2	54	69.61	6.25	4.85*	35.75		
Step2							
+PV	53	64.44	5.17	4.25*	40.53		
+PV^2	52	59.11	5.33	4.69*	45.45		
− P^2	53	59.11	0.00	0.00	45.45		
− PV	54	61.47	2.36	2.07	43.27		
Step 3							
+t	53	58.75	2.72	2.46	45.78		
+t^2	52	55.11	3.64	3.43	49.14		
− t	53	56.47	1.35	1.27	47.89		
Step 4							
+P × PV	52	55.20	1.26	1.19	49.05		
− PV^2	53	55.62	0.42	0.40	48.67		
Step 5							
+t × PV	52	45.06	10.56	12.19***	58.41	− 0.143	0.023
− t^2	53	48.35	3.29	3.61	55.38		
Step 6							
+PV × TV	52	45.36	2.99	3.42	58.14	0.019	0.010
− P × PV	53	45.38	0.02	0.03	58.12		
Step 7							
+TV × E	52	39.17	6.22	8.25**	63.86		
Step 8							
+E × T	51	35.95	3.22	4.56*	66.83	0.082	0.023
− TV × E	52	36.10	0.15	0.21	66.69		
Step 9							
+Lat × P	51	27.37	8.73	16.27***	74.75	− 0.003	0.001

of residuals did not show especially aberrant features nor spatial autocorrelation. As none of the Moran's I values in the different lag-classes were significant at an 0.05 significance level with the Bonferroni correction, we can assume that no spatially structured variable remains to be included in the model. These results suggest that the model predicts reasonably well the Aphodiidae species richness in the cells more poorly sampled (see Fig. 4).

Variation partitioning

Without considering the interactions between explanatory variables, the amount of deviance explained by all

Fig. 3. Leverages and standardized Pearson residual scores of the 66 well-sampled grid squares (see Fig. 1) used to build the forecasting model described in Table 2A. The dotted lines delimit 10 squares with positive residuals higher than 1.5 or with negative residuals less than -1.5. The cells marked with squares were excluded from the final model calculation.
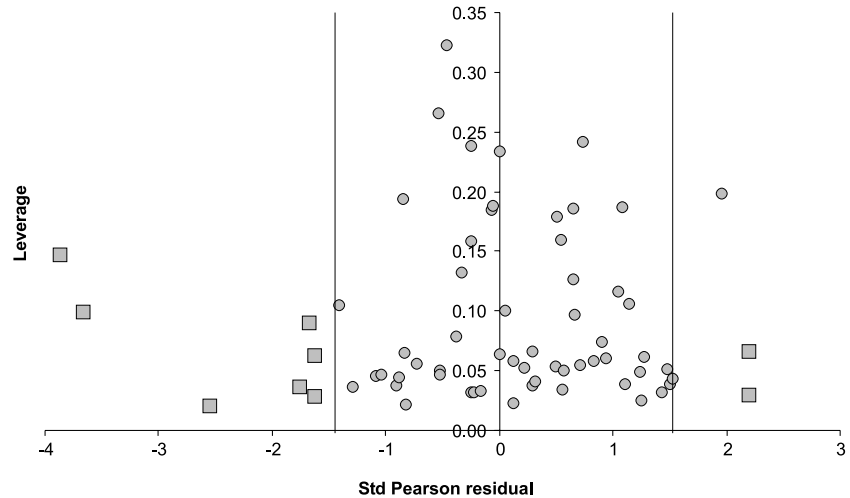
statistically significant climatic, topographic, and spatial variables was 67% (Table 1 and Fig. 2B). The decomposition of the variation between these three groups of variables showed that the largest fraction of the variability in the number of species (37%) was accounted for by the combined effect of the three groups of explanatory variables (spatial, climatic and topographic); it was not possible, therefore, to attribute such a role to any single one of them alone. The pure spatial component explained 27% of total variation in species richness. The effect of the climate component alone and the joint influence of climate and topography were comparable (19% and 16%, respectively), whereas the pure effect of topography was moderate (7%). The negative values of the combination of spatial and climate variables on the one hand, and of spatial and topographic variables on the other ($-10%$ and $-30%$, respectively) probably indicate that spatial and topographic or climatic variables have opposed effects on species richness (Legendre and Legendre 1998). Due to the great relevance of spatial variables, independently of topography and climate, the residuals of significant topographic and climate variables were mapped, showing a gradient in species richness increasing from the south-east to the north-west (Fig. 5).

Discussion

Little accurate geographic information on distribution of species richness is available, for the majority of insect groups. The distribution of biodiversity-related attributes (species richness, rarity, etc.) currently reflects the distribution of the areas investigated by entomologists (Dennis and Hardy 1999). France is a country with a very old tradition of entomological studies. The taxon-

omy and distribution of the French Aphodiidae insect group has been thoroughly investigated; its inventory was completed in 1990 (Lumaret 1990). However, we do not have an accurate map of the geographical variation in French species richness. This present paper demonstrates that with a reasonable image of the sampling effort distribution, it is possible to delimit the localities with reliable faunistic inventories (Lobo et al. 1997) and forecast a predictive map of species richness. New studies and regional surveys can be carried out (Gangloff 1991) but no large-scale survey of Aphodiidae is at present planned. The present model of species richness distribution thus provides a good substitute for information that could not be provided in the coming years.

Sampling all poorly explored or unexplored territories would be expensive and impractical. Supplementing the map recently published for the sister family Scarabaeidae (Lobo et al. 2002), our results demonstrate that delimiting the adequately sampled areas and elaborating forecasting models using simple environmental variables is quicker and cheaper. By such means we can rapidly produce an estimate of the spatial distribution of species richness, thereby obtaining testable geographic patterns. Predictive model techniques should be used to design effective sampling protocols, so new field work data from less inventoried regions will allow us to validate and continually improve the predicted figures. As the areas of high predicted species richness for Aphodiidae were not well inventoried, the sampling effort to validate and improve this model must actually be focused on two main regions: i) the western head which showed the highest estimated species richness per square (maximum of 59 species, i.e. >50% of the total number of French species), and where no square was correctly sampled (Fig. 1); ii) south western France, which was recognised as a hotspot of diversity both for all dung beetle species

Fig. 4. Predicted Aphodiidae species richness distribution in France and Corsica in the upper map. The species richness scores were forecasted by applying the final multiple regression model (see Table 2B) where the data of 9 outliers were omitted. The lower was interpolated at a 10 km resolution to represent the variation in the species richness. The three main French Aphodiidae hotspots were represented: the western head (A), southwestern France (B), and northeastern France (C).

(Lobo et al. 2002) and where one can expect to observe from 56 to 73 species per square, considering Aphodiidae and Scarabaeidae together.

The Aphodiidae species richness per square ranged between 15 and 59 species and the map of predicted species richness does not show an obvious geographic distribution pattern. However, three main Aphodiidae hotspots were observed: the western head (Brittany), southwestern France (between the Atlantic and the Central Massif) and, to a lesser extent, northeastern

Fig. 5. Geographic contour plot of the number of Aphodiidae species predicted by the pure spatial variables. Both latitude and longitude were submitted to a backward stepwise regression with significant topographic and climate variables as independent variables, retaining residuals as a measure of the purely spatial contribution. The number of species was then regressed with these residuals calculating the species richness predicted by these pure spatial variables. Bold solid lines delimit the well-sampled squares taken into account for this analysis (n = 57).
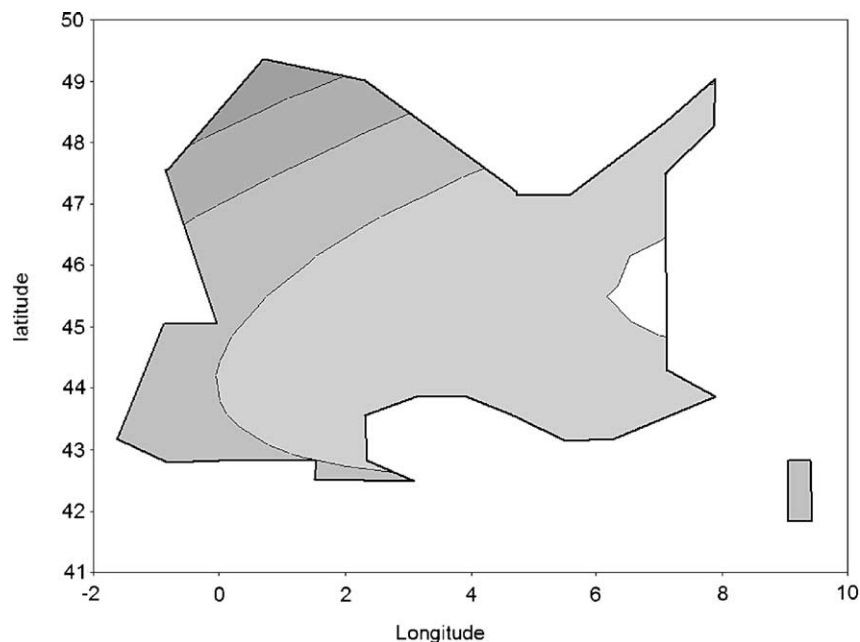
France (from the Belgian to the Swiss borders). These three areas, with a maximum of 59, 55, and 50 species per square, respectively, were traversed by a NW–SE trending poor-species-richness region that includes Corsica (approximately 20–40 species per square). Unlike the case of Scarabaeidae, the Mediterranean region was not of any particular interest and the species richness was not affected by the presence of mountain massifs. Thus, the geographic distribution of species richness found was more complex for Aphodiidae than for Scarabaeidae (Lobo et al. 2002).

The model obtained for Aphodiidae was satisfactory but with a lesser predictive power than the model obtained for Scarabaeidae (74.8% vs 86.2% of total deviance, respectively) (Lobo et al. 2002). The three types of variables considered (spatial, climatic and topographic) influence in a different way the observed distributional pattern of Aphodiidae species richness. As the largest fraction of variability in species richness was accounted for by the combined effect of spatial, climatic and topographic variables, the majority of the variation in French Aphodiidae species richness can be indifferently attributed to these three groups of variables. The pure climatic contribution explains a moderate fraction of Aphodiidae species richness (19%), the importance of the pure topographic fraction is less marked (7%), while the variability indifferently attributed to either climatic or topographic variables is 16%. Purely environmental contribution is relevant because of the influence of non-spatially-structured variables, which play their role on a regional scale. Thus, among the environmental variables

considered those of climate seem to be the most important in explaining Aphodiidae species richness, although their influence is modest. Taking into account the results of model 2, mean annual precipitation (P) alone is the most influential (36% of species richness variability), followed by the annual precipitation variation (PV; 17%). Both P and PV are curvilinearly related with the species richness; inversely with the number of Aphodiidae species until a value of around 1200 mm of annual precipitation, and until around 700 mm of annual precipitation variation. For greater values of P and PV the richness remains constant or even increases (Fig. 6).

Which underlying biological principles might be able to explain the relevance of precipitation and the limited importance of temperature and topographic variables? Whereas the latitudinal distributional pattern of the basically warm-adapted Scarabaeidae species richness is very simple and primarily related to temperature (Lobo et al. 2002), the explanatory factors in the case of Aphodiidae are probably more diverse. Aphodiidae is essentially a Holarctic group well-adapted to cold temperate conditions and dominating the assemblages of north temperate regions (Cambefort 1991, Hanski 1991). However, within the Aphodiidae there are also many monophyletic lineages with an Afrotropical origin (Dellacasa et al. 2001, Cabrero-Sañudo, pers. comm.). The heterogeneity in the environmental adaptations of Aphodiidae subgroups may be the main cause of the lack of influence of temperature and altitudinal variables, so the combination of warm low-altitude and cold
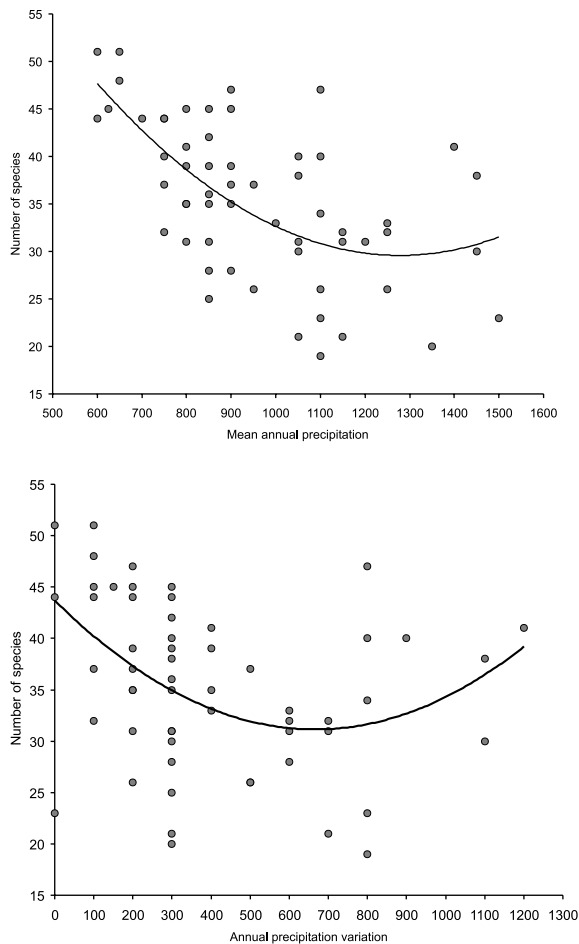
Fig. 6. Curvilinear relationships between the number of species in each square and the mean annual precipitation (above) or the annual precipitation variation (below). Solid lines are the fitted values derived from the quadratic functions of each variable.

high-altitude species may reduce the explanatory role of temperature and topographic variables. The importance of precipitation variables could be explained by the free-ranging life style of immature larval stages. If the rainfall increases and the rising increment of soil humidity hinders the survival of some of these species, total species richness would diminish until these soil-moisture-sensitive species stopped occurring.

Spatial variables seem to be the most important influence on the species richness distribution, and these probably reflect the role played by other ignored environmental variables, biotic interactions or past contingent events (Legendre and Legendre 1998). The purely spatial contribution to species occurs on the map of the residuals of the spatial polynomial terms, after controlling for the effect of significant climatic and topographic variables. Which variables can explain this spatial pattern? Examining the obtained residuals map it is difficult to imagine a non-considered environmental

variable, unrelated with temperature or precipitation, able to account for this non explained variation. Here we suspect that past events may be responsible for this spatially-structured pattern.

The Maghreb-Iberian Peninsula and the Balkan-Caucasus-Anatolia region were the two main centres of endemism for Aphodiidae in the Western Palaearctic region (Lumaret and Lobo 1996). There is abundant evidence that these regions supplied colonizers for northern areas, due to the refuge they provided during the warm inter-glacial intervals (Bennett et al. 1991, Hewitt 1996, Taberlet et al. 1998, Brewer et al. 2002). In the case of Aphodiidae, there is fossil support for both wide southwards and northwards dispersal during the glacial/interglacial phases (Coope and Angus 1975, Coope 1979, 1990, Foddari 1994). The current pattern of Aphodiidae diversity observed in France could be partially explained by the geographical location of the colonization routes followed by this family of Coleoptera during the quaternary climate changes. The European range expansion from Iberian glacial refuges seems to have been possible only in the Mediterranean and Atlantic end of the Pyrenees mountains (Olalde et al. 2002). This recolonization was much faster on the Atlantic coast (Brewer et al. 2002). Thus the higher diversity of western France could be due to the relevance of a western colonization pathway and also to the role played by the Alps in slowing the spread from eastern Europe. Different recent phylogeographic studies carried out mainly with *Quercus* species indicate that post-glacial colonization routes follow the Atlantic and North Sea coast (Brewer et al. 2002, Cottrell et al. 2002, Petit et al. 2002a, b), and the Alps acted very probably like a barrier for the colonization of western Europe from eastern refuges (Brewer et al. 2002, Hewitt 2001, Taberlet et al. 1998). Thus, we suggest that postglacial colonization is probably responsible of the present western surplus in Aphodiinae species richness.

Although the Alps act as a barrier for the southeastern species of Europe, they have also been a refuge and a centre of endemism for some cold-adapted dung beetle species (Jay-Robert et al. 1997). The purely spatial model predicted a very low number of species near the Alps, but the final model did not present this pattern. The signs of the significant explanatory variables show that species richness increases with environmental heterogeneity (PV, TV) and elevation (E). Thus, local environmental characteristics of the Alps (high altitude and high heterogeneity) would have contributed to increase the Aphodiidae species richness, allowing numerous species to coexist. While the Alps have some environmental characteristics that promote Aphodiidae diversity, they probably have contributed to decrease the species richness through their historical role as a barrier.

References

Anderson, M. J. and Gribble, N. A. 1998. Partitioning the variation among spatial, temporal and environmental components in a multivariate data set. – Aust. J. Ecol. 23: 158–167.

Anon. 1999. STATISTICA for Windows. Computer program manual. – Tulsa, OK.

Austin, M. P. 1980. Searching for a model for use in vegetation analysis. – Vegetatio 42: 11–21.

Austin, M. P. 1998. An ecological perspective on biodiversity investigations: examples from Australian eucalypt forests. – Ann. Missouri Bot. Gard. 85: 2–17.

Austin, M. P., Cunningham, R. B. and Fleming, P. M. 1984. New approaches to direct gradient analysis using environmental scalars and statistical curve-fitting procedures. – Vegetatio 55: 11–27.

Austin, M. P., Nicholls, A. O. and Margules, C. R. 1990. Measurement of the realized qualitative niche: environmental niches of five *Eucalyptus* species. – Ecol. Monogr. 60: 161–177.

Austin, M. P., Pausas, J. G. and Nicholls, A. O. 1996. Patterns of tree species richness in relation to environment in southeastern New South Wales, Australia. – Aust. J. Ecol. 21: 154–164.

Baraud, J. 1992. Coléoptères Scarabaeoidea d'Europe, Faune de France 78. – Fédération Française des Sociétés de Sciences Naturelles, Société Linnéenne de Lyon.

Bennett, K. D., Tzedakis, P. C. and Willis, K. J. 1991. Quaternary refugia of north European trees. – J. Biogeogr. 18: 103–115.

Birks, H. J. B. 1996. Statistical approaches to interpreting diversity patterns in the Norwegian mountain flora. – Ecography 19: 332–340.

Bojorquez-Tapia, L. A. et al. 1996. Identifying conservation priorities in México through geographic information systems and modelling. – Ecol. Appl. 5: 215–231.

Borcard, D., Legendre, P. and Drapeau, P. 1992. Partialling out the spatial component of ecological variation. – Ecology 73: 1045–1055.

Brewer, S. et al. 2002. The spread of deciduous *Quercus* throughout Europe since the last glacial period. – For. Ecol. Manage. 156: 27–48.

Buckland, S. T. and Elston, D. A. 1993. Empirical models for the spatial distribution wildlife. – J. Appl. Ecol. 30: 478–495.

Cambefort, Y. 1991. Biogeography and evolution. – In: Hanski, I. and Cambefort, Y. (eds), Dung beetle ecology. Princeton Univ. Press, pp. 51–67.

Cliff, A. D, and Ord, J. K. 1981. Spatial processes. Models and applications. – Pion.

Coope, G. R. 1979. Late Cenozoic Coleoptera: evolution, biogeography and ecology. – Annu. Rev. Ecol. Syst. 10: 247–267.

Coope, G. R. 1990. The invasion of northern Europe during the Pleistocene by Mediterranean species of Coleoptera. – In: Di Castri, F., Hansen, A. J. and Debussche, M. (eds), Biological invasions in Europe and the Mediterranean basin. Kluwer, pp. 203–215.

Coope, G. R. and Angus, R. B. 1975. An ecological study of a temperate interlude in the middle of the last glaciation, based on fossil Coleoptera from Isleworth, Middlesex. – J. Anim. Ecol. 44: 365–391.

Cottrell, J. E. et al. 2002. Distribution of chloroplast DNA variation in British oaks (*Quercus robur* and *Q. petraea*): the influence of postglacial colonization and human management. – For. Ecol. Manage. 156: 181–195.

Crawley, M. J. 1993. GLIM for ecologists. – Blackwell.

Dennis, R. H. and Hardy, P. B. 1999. Targeting squares for survey: predicting species richness and incidence of species for a butterfly atlas. – Global Ecol. Biogeogr. 8: 443–454.

Dellacasa, G., Bordat, P. and Dellacasa, M. 2001. A revisional essay of world genus-group taxa of Aphodiinae. – Mem. Soc. Entomol. Ital. 79: 1–482.

Dobson, A. J. 1990. An introduction to Generalized Linear Models. – Chapman and Hall.

Ferrier, S. 2002. Mapping spatial pattern in biodiversity for regional conservation planning: where to from here? – Syst. Biol. 51: 331–363.

Ferrier, S. et al. 2002. Extended statistical approaches to modelling spatial pattern in biodiversity in northeast New South Wales. II. Community-level modelling. – Biodiv. Conserv. 11: 2309–2338.

Foddari, D. 1994. Fossil arthropods from a full glacial site in north eastern Italy. – Quat. Res. 41: 336–342.

Fraser, R. H. 1998. Vertebrate species richness at the mesoscale: relative roles of energy and heterogeneity. – Global Ecol. Biogeogr. 7: 215–220.

Gangloff, L. 1991. Catalogue et atlas des Coléoptères d'Alsace. Tome 4: Lamellicornia, Scarabaeidae, Lucanidae. – Société alsacienne d'entomologie, Musée zoologique de l'université et de la ville de Strasbourg.

Gaston, K. J. 1996. Species richness: measure and measurement. – In: Gaston, K. J. (ed.), Biodiversity. A biology of numbers and difference. Blackwell, pp. 77–113.

Guisan, A., Edwards, T. C. and Hastie, J. T. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. – Ecol. Modell. 157: 89–100.

Hanski, I. 1991. North temperate dung beetles. – In: Hanski, I. and Cambefort, Y. (eds), Dung beetle ecology. Princeton Univ. Press, pp. 75–96.

Heikkinen, R. K. and Birks, H. J. B. 1996. Spatial and environmental components of variation in the distribution patterns of subarctic plant species at Kevo, N Finland – a case study at the meso-scale level. – Ecography 19: 341–351.

Heikkinen, R. K. and Neuvonen, S. 1997. Species richness of vascular plants in the subarctic landscape of northern Finland: modelling relationships to the environment. – Biodiv. Conserv. 6: 1181–1201.

Hewitt, G. M. 1996. Some genetic consequences of ice ages, and their role in divergence and speciation. – Biol. J. Linn. Soc. 58: 247–276.

Hewitt, G. M. 2001. Speciation, hybrid zones and phylogeography – or seeing genes in space and time. – Mol. Ecol. 10: 537–549.

Hirzel, A. H., Helfer, V. and Metral, F. 2001. Assesing habitat-suitability models with a virtual species. – Ecol. Modell. 145: 111–121.

Hortal, J., Lobo, J. M. and Martín-Piera, F. 2001. Forecasting insect species richness scores in poorly surveyed territories: the case of the Portuguese dung beetles (Col. Scarabaeinae). – Biodiv. Conserv. 10: 1343–1367.

Iverson, L. R. and Prasad, A. M. 1998a. Predicting abundance of 80 tree species following climate change in the eastern United States. – Ecol. Monogr. 68: 465–485.

Iverson, L. R. and Prasad, A. 1998b. Estimating regional plant biodiversity with a GIS modelling. – Div. Distr. 4: 49–61.

Jay-Robert, P., Lobo, J. M. and Lumaret, J. P. 1997. Elevational turnover and species richness variation in European mountain dung beetle assemblages. – Arc. Ant. Alp. Res. 29: 196–205.

Kerr, J. T. and Packer, L. 1999. The environmental basis of North American species richness patterns among *Epicauta* (Coleoptera: Meloidae). – Biodiv. Conserv. 8: 617–628.

Leathwick, J. R., Burns, B. R. and Clarkson, B. D. 1998. Environmental correlates of tree alpha-diversity in New Zealand primary forest. – Ecography 21: 235–246.

Legendre, P. 1993. Spatial autocorrelation: trouble or new paradigm? – Ecology 74: 1659–1673.

Legendre, P. and Legendre, L. 1998. Numerical ecology. – Elsevier.

Levin, S. A. 1992. The problem of pattern and scale in ecology. – Ecology 73: 1943–1967.

Lobo, J. M., Lumaret, J. P. and Jay-Robert, P. 1997. Les atlas faunistiques comme outils d'analyse spatiale de la biodiversité. – Ann. Soc. Ent. France (N.S.) 33: 129–138.

Lobo, J. M., Castro, I. and Moreno, J. C. 2001. Spatial and environmental determinants of vascular plant species richness distribution in the Iberian Peninsula and Balearic islands. – Biol. J. Linn. Soc. 72: 233–253.

Lobo, J. M., Lumaret, J. P. and Jay-Robert, P. 2002. Modelling the species richness distribution of French dung beetles (Coleoptera, Scarabaeidae) and delimiting the predictive capacity of different groups of explanatory variables. – Global Ecol. Biogeogr. 11: 265–277.

Lobo, J. M. and Martín-Piera, F. 2002. Searching for a predictive model for Iberian dung beetle species richness (Col., Scarabaeinae) using spatial and environmental variables. – Conserv. Biol. 16: 158–173.

Lumaret, J. P. 1990. Atlas des Scarabéides Laparosticti de France. – Secrétariat Faune-Flore/MNHN.

Lumaret, J. P. and Lobo, J. M. 1996. Geographic distribution of endemic dung beetles (Coleoptera, Scarabaeidae) in the Western Palaearctic region. – Biodiv. Lett. 3: 192–199.

Lumaret, J. P., Lobo, J. M. and Jay-Robert, P. 1996. Catalogue et Atlas des Scarabéides Laparosticti endémiques de France. – Service du Patrimoine Naturel/IEGB/MNHN – OPIE.

Manel, S. et al. 1999a. Alternative methods for predicting species distribution: an illustration with Himalaya river birds. – J. Appl. Ecol. 36: 734–747.

Manel, S., Dias, J. H. and Ormerod, S. J. 1999b. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. – Ecol. Modell. 120: 337–347.

Margules, C. R., Nicholls, A. O. and Austin, M. P. 1987. Diversity of *Eucalyptus* species predicted by a multi-variable environment gradient. – Oecologia 71: 229–232.

McCullagh, P. and Nelder, J. A. 1989. Generalized Linear Models. – Chapman and Hall.

Mitchell, N. D. 1991. The derivation of climate surfaces for New Zealand, and their application to the bioclimatic analysis of the distribution of kauri (*Agathis australis*). – J. Roy. Soc. New Zealand 21: 13–24.

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

Odland, J. 1988. Spatial autocorrelation. – Sage Publ.

Olalde, M. et al. 2002. White oaks phylogeography in the Iberian Peninsula. – For. Ecol. Manage. 156: 89–102.

Owen, J. G. 1989. Patterns of herpetofaunal species richness: relation to temperature, precipitation, and variance in elevation. – J. Biogeogr. 16: 141–150.

Parker, V. 1999. The use of logistic regression in modelling the distributions of bird species in Swaziland. – S. Afr. J. Zool. 34: 39–47.

Paulian, R. and Baraud, J. 1982. Faune des Coléoptères de France. II. Lucanoidea et Scarabaeoidea. – Lechevalier.

Pearce, J. and Ferrier, S. 2000. An evaluation of alternative algorithm for fitting species distribution models using logistic regression. – Ecol. Modell. 128: 127–147.

Pereira, J. M. C. and Itami, R. M. 1991. GIS-based habitat modelling using logistic multiple regression: a study of the Mt. Graham red squirrel. – Photogram. Eng. Remote Sen. 57: 1475–1486.

Peterson, A. T., Soberón, J. and Sánchez-Cordero, V. 1999. Conservatism of ecological niches in evolutionary time. – Science 285: 1265–1267.

Petit, R. J. et al. 2002a. Chloroplast DNA variation of oaks in France and the influence of forest fragmentation on genetic diversity. – For. Ecol. Manage. 156: 115–129.

Petit, R. J. et al. 2002b. Identification of refugia and post-glacial colonisation routes of European white oaks based on chloroplast DNA and fossil pollen evidence. – For. Ecol. Manage. 156: 49–74.

Pinel-Alloul, B., Niyonsenga, T. and Legendre, P. 1995. Spatial and environmental components of freshwater zooplankton structure. – Ecoscience 2: 1–19.

Qian, H. 1998. Large-scale biogeographic patterns of vascular plant richness in North America: an analysis at the generic level. – J. Biogeogr. 25: 829–836.

Qinghong, L. and Bråkenhielm, S. 1995. A statistical approach to decompose ecological variation. – Water Air Soil Poll. 85: 1587–1592.

Richerson, P. J. and Lum, K. L. 1980. Patterns of plant species diversity in California: Relation to weather and topography. – Am. Nat. 116: 504–536.

Sawada, M. 1999. ROOKCASE: An Excel 97/2000 Visual Basic (VB) add-in for exploring global and local spatial autocorrelation. – Bull. Ecol. Soc. Am. 80: 231–234.

Shipley, B. 1999. Testing causal explanations in organismal biology: causation, correlation and structural equation modelling. – Oikos 86: 374–382.

Smith, P. A. 1994. Autocorrelation in logistic regression modelling of species' distributions. – Global Ecol. Biogeogr. 4: 47–61.

Soberón, J. M., Llorente, J. B. and Oñate, L. 2000. The use of specimen-label databases for conservation purposes: an example using Mexican papilionid and pierid butterflies. – Biodiv. Conserv. 9: 1441–1466.

Stockwell, D. R. B. and Peterson, A. T. 2002. Effects of sample size on accuracy of species distribution models. – Ecol. Modell. 148: 1–13.

Taberlet, P. et al. 1998. Comparative phylogeography and postglacial colonization routes in Europe. – Mol. Ecol. 7: 453–464.

Tonteri, T. 1994. Species richness of boreal understorey forest vegetation in relation to site type and successional factors. – Ann. Zool. Fenn. 31: 53–60.

Vayssières, M. P., Plant, R. E. and Allen-Diaz, B. H. 2000. Classification trees: an alternative non-parametric approach for predicting species distributions. – J. Veg. Sci. 11: 679–694.

Wohlgemuth, T. 1998. Modelling floristic species richness on a regional scale: a case study in Switzerland. – Biodiv. Conserv. 7: 159–177.

Zimmermann, N. E. and Kienast, F. 1999. Predictive mapping of alpine grasslands in Switzerland: species versus community approach. – J. Veg. Sci. 10: 469–482.