

BIODIVERSITY IN FOREST ECOSYSTEMS

The role of regional and local scale predictors for plant species richness in Mediterranean forests

G. BACARO^{1,2}, D. ROCCHINI^{1,2}, I. BONINI¹, M. MARIGNANI^{2,3}, S. MACCHERINI^{1,2}, & A. CHIARUCCI^{1,2}

¹Dipartimento di Scienze Ambientali “G. Sarfatti”, Università di Siena, Siena, Italy, ²Terradata s.r.l. environmetrics, Siena, Italy, and ³Dipartimento di Biologia Vegetale, Università di Roma “La Sapienza”, Roma, Italy

Abstract

Both local and regional predictors play a role in determining plant community structure and composition. Climate, soil features as well as different local history and management affect forest understorey and tree species composition, but to date their specific role is relatively unknown. Few studies have addressed the importance of these predictors, especially in the Mediterranean area, where environmental conditions and human impacts have generated heterogeneous forest communities. In this study, the relationships between environmental variables and species richness of different groups of vascular plants (vascular species, woody species and open habitat species) and bryophytes were investigated in Tuscan forests. A total of 37 environmental variables were used by generalised linear model fitting in order to find parsimonious subsets of environmental factors (predictors) that are able to explain species diversity patterns at the local scale. Moreover, the role of regional and local variable groups on species richness of the considered plant groups was estimated by using the variance partitioning approach. We found that local variables, such as forest management and structure, explained more variance than regional variables for total species richness, open habitat species richness and bryophyte species richness. On the other hand, regional variables (such as elevation) played a central role for woody species richness.

Keywords: *Biodiversity, bryophytes, plant communities, species diversity, vegetation, woody species*

Introduction

Plant community structure and diversity are influenced by many interacting biotic and abiotic factors. At the local scale, ecological variables and disturbance are key factors interacting with plant strategies to determine vegetation composition and structure (Ribichich & Protomastro 1998). Ecological factors such as light, temperature, water regime or the degree of habitat patchiness at the landscape, or regional, level can be used as predictors for explaining species richness and its variability (Austin 2002). Local plant species richness can vary considerably across a landscape, suggesting that local factors may be more important than those acting at larger scale (Huston 1999; Pausas & Austin 2001; Ohlemüller et al. 2006). However, Ohlemüller et al. (2004) pointed out that only a few studies focused on comparing the effects

of local and regional factors on plant species richness at the local scale.

From an applicative point of view, understanding the processes controlling species diversity at the local scale can help to promote suitable environmental policies and forest management practices. In addition, the knowledge of correlates of species diversity can help to set up proxies that can help large-scale monitoring of plant species diversity (Austin 2002).

In Tuscany (central Italy) forests exceed 10⁶ ha, corresponding to about 50% of the total surface, and more than 220,000 ha of forests belong to the Regional Administration of Tuscany or other public Institutions (Chiarucci & Bonini 2005). The Tuscan forests, forming a heterogeneous landscape with many different forest types, have always been considered a fundamental resource, for timber production and soil protection. However, consideration of forest

ecosystems for environment and landscape issues has increased in recent years (Bartolozzi et al. 2002).

This work aims to address the relationships between plant species richness and predictor variables in these forest ecosystems. More specifically, this study aims to model the contribution of different groups of environmental variables in controlling the patterns of species richness at the local scale for different groups of vascular plants as well as bryophytes.

Methods

Study sites and sampling design

Six forest estates (hereafter “forests”) owned by the regional administration of Tuscany were chosen as study sites (Figure 1).

Each forest is formed by contiguous or non-contiguous woodlands managed as a single unit. These forests cover a total of 37,240 ha and range in size from 2098 to 10,311 ha. They are located from the lowlands of the coastline to the Apennine mountains and their vegetation is formed by different plant communities, varying from the evergreen Mediterranean forests dominated by *Quercus ilex*, along the coastlines, to the *Fagus sylvatica* and *Abies alba* forests of mountain sites; conifer plantations are

present everywhere (see Chiarucci & Bonini 2005 for details).

Forest structure and composition were surveyed by a design-based probabilistic sample of sites. The sites used in this survey were selected as a random sub-sample of the sites adopted for the second-stage sampling of the Forest Inventory of Tuscany. The number of sites corresponded to that of an existing network for the monitoring of tree crown conditions (109 sites), with the number of sites proportional to the true forest surface of each area, i.e., excluding open patches therein. In each site, once located with a high precision GPS, a 20 × 20 m plot was delimited (Chiarucci & Bonini 2005).

Response variables

The following data were collected within each 20 × 20 m plot: (i) total list of vascular plant species; (ii) DBH of each individual plant with DBH ≥ 3 cm; (iii) a list of bryophyte species. A total of 12,829 stems with DBH ≥ 3 cm belonging to 61 species were recorded. Vascular plant species richness (woody plus understorey plants) was 420. Within each plot, bryophytes were sampled by using 24 “bryoplots”: these were located in four clusters, selected by a restricted random procedure, positioned in two different aspects (north and south) of three different micro-habitats (trunk, tree basis, soils) with respect to a reference tree (see Chiarucci et al. 2007).

From these data, four different response variables were extracted per plot: (1) *vascular plant species richness*, including tree, shrub and grass species (range 3–69); (2) *woody species richness*, calculated as the number of species found with DBH > 3 cm (range 1–14); (3) *open habitat species richness*, calculated as the number of species with an ecological optimum in open habitats (range 0–25). For this purpose, all species were classified as forest or non-forest species (according to Pignatti 1982; Honnay et al. 1998; Hermy et al. 1999; Guirado et al. 2006). The presence of these species in forest habitats can be used as an indicator of disturbance and of the number of competitors that might hamper the survival of true forest species (or ancient forest plant species, sensu Hermy 1994; Hermy et al. 1999); (4) *bryophyte species richness*, calculated as the pooled number of bryophytes recorded in the 24 bryoplots sampled within each plot (range 0–22).

Predictor variables

A continuum of environmental variables exists, from indirect variables, with no direct impact on plant



Figure 1. Distribution of the six investigated forest estates in Tuscany.

growth or survival processes, to direct causal variables. Predictions become increasingly robust and less location-specific as the predictors become more relevant to biological processes. However, the need to use predictors for which values are available for all the sites may limit the choice of variables (Austin 2002). Here, 37 predictor variables were obtained, by direct measurement or GIS derivation for each plot, and these were classified into three groups (Table I), according to a spatial scale criterion (regional vs. local predictors) and on their allegedly similar effect on plants:

Regional predictors. Seventeen regional-scale predictors were used:

1. *Climatic (13 variables).* This group included direct measures of climatic features, such as rainfall and temperature, or derived indices (e.g. “summer humidity”, i.e. the ratio between the sum of summer rainfall and the mean summer temperature). Since a high degree of co-variation existed among these variables, a Principal Component Analysis (PCA) was performed on this set of variables to avoid multicollinearity and aliasing effects in the generalised linear model (GLM) modelling process (Taylor et al. 2002). The first three PCA factors explained more than 99% of the variance (Table II); thus the 13 climatic variables were reduced to three new variables (the first three PCA factors).
2. *Elevation.* As a topographic parameter, elevation is an important variable representing a broad-scale constraint for life and dispersion of plants.
3. *Geographic location (two variables).* Latitude and longitude were used to test for geographic patterns in species richness (Palmer et al. 2002; Ohlemüller et al. 2004).
4. *Forest estate.* This categorical variable (six classes) was included to check for spatial dependence in species richness distribution.

Local scale – Abiotic predictors. Seven predictors were assigned to this group:

1. *Geo-morphological features (three variables).* This category included the classic slope, but also the “folded aspect” and the “potential direct incident radiation” (Radiation in Table I, calculated according to McCune & Keon 2002).
2. *Soil features (four categorical variables).* Derived from thematic maps (1:250,000), these variables included pH, water availability, organic matter and limestone content.

Local scale – Forest structure and management predictors. Thirteen predictors were included in this group:

1. *Forest stand structure (seven variables).* These variables indicated the forest structure in the plots, from the number of stems with $DBH \geq 3$ cm to the total basal area, but also included the percentage cover of the tree, shrub and herb layers. The cumulative cover of co-dominant tree species (Kikvidze & Ohsawa 2002) was also calculated.
2. *Categorical forest variables (two variables).* These variables included the forest management and forest physiognomy.
3. *Land use (four variables).* Using a Corine Land Cover map (EEA 2005) the surface of forest and agricultural areas (I and II Corine Levels) was assessed in a circular buffer of 2500 m centred on the plot. As a measure of habitat diversity, the Shannon index H' was computed using the area occupied by each 2nd level Corine Land Cover type in the same area. These variables were used as proxies of habitat diversity and sources of competitive plants (Hermy et al. 1999; Ohlemüller et al. 2004).

Model development and evaluation

Ecologists use different techniques to investigate relations between response variables and sets of predictors (Crawley 1993; Guisan et al. 2002; Rushton et al. 2004). Regression methods are widely used, but the assumptions required by these techniques are hardly respected by real data (Legendre & Legendre 1998; Guisan et al. 2002). An example is the departure from normality of biological count data such as the number of species that can follow a Poisson distribution (Oksanen & Minchin 2002; Ohlemüller et al. 2004). More flexible methods are the GLM (McCullagh & Nelder 1989). In fact, when the response variable is assumed to follow a given distribution, the choice of a link function and its associated variance function can easily be done, since GLMs are built to accommodate a large variety of data type and response variable distributions (Dalthorp 2004), including Gaussian, Poisson and Negative Binomial (Engler et al. 2004)

A weighted analysis of deviance was here used to find the relationships between the dependent variables and the predictor variables. This analysis is analogous to an analysis of variance, except for assuming non-normal error distribution (McCullagh & Nelder 1989; Pausas et al. 2003). Since the response variables were count data, Poisson error

Table I. Predictor variables and categories in which they were grouped.

Predictor categories and variables	Measurement unit	Min	Max	Mean
Regional predictors				
<i>i) Climatic variables</i>				
Total annual rainfall	mm	715.9	2278.8	1290.8
Maximum rainfall in a month	mm	96.7	309.4	171.0
Minimum rainfall in a month	mm	16.8	75.9	45.5
Σ summer rainfall	mm	88.0	302.9	188.6
Σ winter rainfall	mm	198.6	698.2	370.6
Mean annual temperature	°C	3.2	14.1	9.6
Maximum temperature in the hottest month	°C	12.2	23.1	18.6
Minimum temperature in the coldest month	°C	-4.4	5.9	1.4
Σ summer temperatures	°C	88.0	302.9	52.6
Σ winter temperatures	°C	-12.0	20.2	6.3
Σ temperatures months with mean $T > 5^\circ$	°C	47.5	169.1	109.0
Ratio (annual precipitation/annual temperature)	mm/°C	51.0	681.5	173.2
Ratio (summer precipitation/summer temperature)	mm/°C	1.35	8.74	3.92
<i>ii) Elevation</i>	m	20	1659	763.3
<i>iii) Geographical location</i>				
Latitude (datum WGS84)	°	42°56'09"	44°07'27"	43°33'59"
Longitude (datum WGS84)	°	10°21'25"	11°54'31"	11°08'17"
<i>iv) Forest estate: six levels</i>	See Figure 1			
Local scale predictors – Abiotic variables				
<i>i) Geo-Morphological plot features</i>				
Slope	Radiants	0.02	0.14	0.048
Folded aspect	Radiants	0	3.14	1.53
Radiation	MJ*cm ⁻² *year ⁻¹	1.1	2.78	2.32
<i>ii) Soil features</i>				
pH	Three levels (5, 6; 6,7; 7,8)			
Organic matter	Three levels (1, 2; 3, 4; 5, 6)			
Water availability	Three levels (51–80; 81–110; 111–140)			
% Calcareous component	Three levels (0–4.9; 5–9.9; 10–14. 9)			
Local scale predictors – Forest management, structure and land use variables				
<i>i) Forest stand structure</i>				
Total basal area for trees with DBH > 3	cm ²	937.3	36,134.4	14,354
Total tree density	n. stems	6	502	117.6
% Cover of the tree layer	%	0	100	72.2
% Cover of the shrub layer	%	0	80	15
% Cover of the herbaceous layer	%	0	95	17.1
% Cover of co-dominant trees species	%	63.2	100	92.1
Total % cover of vegetation	%	50	100	86.9
<i>ii) Categorical forest variables</i>				
Forest management: two levels	Two classes (Natural; artificial)			
Forest physiognomy: three levels	Three classes (Evergreen broadleaved; deciduous broadleaved; coniferous)			
<i>iii) Land use</i>				
First level Corine land cover (3): Forest and natural areas percentage	%	36.6	100	96.1
First level Corine land cover (2): Agricultural areas percentage	%	0	57.6	3.5
Second level Corine land cover (3.1): Forests percentage	%	0	100	93.5
Shannon Index (H) on relative cover percentage for First Level Corine		0	1.09	0.39

distribution was selected as a fitting parameter in GLM specifications. However, when the overdispersion parameter was higher than one (variance much

higher than the mean, Pausas et al. 2003), the response variable was normalised by a Box-Cox Transformation (Box & Cox 1964) and the model

Table II. Results of principal components analysis.

	Axis 1	Axis 2	Axis 3
	89.47%	97.56%	99.36%
Cumulative explained variance	Coefficients of correlation variables-factors		
Total annual rainfall	-0.921	-0.384	-0.001
Max <i>P</i> for month	-0.908	-0.411	0.031
Min <i>P</i> for month	-0.950	-0.168	-0.241
Σ summer <i>P</i>	-0.956	-0.225	-0.173
Σ winter <i>P</i>	-0.901	-0.415	0.044
Mean annual <i>T</i>	0.959	-0.281	-0.016
Max <i>T</i> hottest month	0.959	-0.268	-0.032
Min <i>T</i> coldest month	0.960	-0.267	0.039
Σ summer <i>T</i>	0.956	-0.279	-0.033
Σ winter <i>T</i>	0.958	-0.279	0.018
Σ <i>T</i> months with <i>T</i> > 5°	0.942	-0.305	0.079
Annual <i>P</i> /annual <i>T</i>	-0.923	-0.081	0.360
Summer <i>P</i> /summer <i>T</i>	0.994	-0.053	0.047
Correlation with elevation	-0.906	-0.112	-0.156

In bold: Climatic variables showing the highest correlation coefficients with each axis.

re-built with new distribution parameters. Before fitting the model, the response variables were fitted to the 2nd order polynomial and linear values of predictors: when polynomial fitting significantly decreased the residual sum of squares, the predictor was included as polynomial.

A null model considering only the response variable without any predictive variable was first developed to know the accurate value of the intercept for the regression model. Second, a full model including all the predictor variables (except those directly related to the response variables – e.g. the cover of the herbaceous layer to predict total plant species richness) was built; by means of the full model, the total variance of response variables explained by all the predictor variables was described and the proportion of unexplained variance (stochastic variation or lack of proper predictors) estimated. Third, through an iterative stepwise model selection by the Akaike Information Criterion (AIC, Hastie & Pregibon 1992), a minimal adequate model was estimated: this model explains the highest proportion of variation by using the minimum number of variables. Significance of each variable retained in the minimal adequate model was tested using the *F*-statistic for “Normal” models and using the χ^2 statistic for “Poisson” models (Crawley 1993; Ohlemüller et al. 2004). An *F*-test was performed in order to check statistical differences between the Full and Minimal models.

As a measure of “goodness of fit” for each minimal GLM, the adjusted D^2 (D_{adj}^2) was calculated

(see Guisan & Zimmermann 2000). The D_{adj}^2 ranges from 0 (no fit) to 1 (perfect fit). To avoid the occurrence of multicollinearity in the predictors retained in minimal models, the Generalised Variance Inflation Factor (Fox & Monette 1992) was computed. A leave-one-out cross-validation estimate of prediction error (Davison & Hinkley 1997) was also calculated. The coefficient of correlation (*r*) between the values predicted by the minimal model and the corresponding observed values of the response variables was computed. Lastly, the variance explained by each group of predictors (local and regional) was quantified by partial regression analysis (excluding the co-variation between the three groups) with a variance partitioning procedure (Borcard et al. 1992; Anderson & Gribble 1998; Legendre & Legendre 1998). The partitioning of variation for species richness includes the following components: pure regional variables; pure local-abiotic; pure local-forest structure and management; combined regional/local-abiotic factor, combined regional/local-forest structure and management; combined local-abiotic factor/local-forest structure and management; combined regional/local components, and unexplained.

Analyses were performed by using the R software (R Development Core Team 2007).

Results

A high number of significant correlations between environmental variables and the species richness of the considered plant groups were observed (Tables III–VI), even though a fair rate of redundancy was present because of autocorrelation among variables.

Vascular plant species richness

The GLM for vascular plant species richness using Poisson error distribution showed a high degree of overdispersion (dispersion parameter > 3). Consequently, data were normalised by a Box-Cox transformation ($\gamma = 0.4774$ where γ is the Box-Cox power). Moreover, a graphical examination of data distribution by mean of a QQ-plot revealed a normal-like distribution (not shown). The normalised response variable was used in a GLM with Gaussian error and identity link function. A total of 24 predictors were placed in the Full Model. Total variance explained by the minimal adequate model, after stepwise routine, was approximately 58% ($D_{adj}^2 = 0.581$, Table III). Eight explanatory variables were retained in the minimal model, but only five provided a significant reduction of deviance ($p < 0.001$) in explaining total species richness (Table III). In particular, the highest deviance reduction was achieved by elevation (fitted as

polynomial, Table III) with species richness related in a humped-back manner with elevation (coefficient values were both negative; peak of species richness with respect to elevation was expected between 500 m and 700 m a.s.l.). Forest physiognomy was

the second most important predictor (categorical), associated with pH classes. Predictors related to the stand structure were also found to be significant, with species richness modelled as a decreasing function of density and cover of trees. All values of

Table III. Summary of the GLM full and minimal models for the response variables vascular plant species richness.

Response Variable	Mean	Min	Max	No. of variables involved in the full GLM				Model features
Vascular plant species richness	27	3	69	24				Gaussian distribution, "identity" link function
Variables retained in the minimal GLM								
ID Var.	Deviance reduction	Coefficient value	VIF	$p(\mathbf{F})$	Mean error (Leave-one-out cross validation)	D_{adj}^2	r	Difference between full and minimal GLM $p(\mathbf{F})$
1. Elevation	141.62	(-, -)	1.89	***	3.20	0.581	0.792	0.993
2. Forest physiognomy	99.30		1.29	***				
3. Tree density	98.37	-	1.60	***				
4. Soil pH	55.69		1.16	***				
5. Cover % tree layer	28.48	-	1.28	**				

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

Table IV. Summary of the GLM full and minimal models for the response variable woody species richness.

Response variable	Mean	Min	Max	No. of variables involved in the full GLM				Model features
Woody species richness	4.7	1	14	21				Poisson distribution, "log" link function
Variables retained in minimal GLM								
ID Var.	Deviance reduction	Coefficient value	VIF	$p(\chi^2)$	Mean error (Leave-one-out cross validation)	D_{adj}^2	r	Difference between full and minimal GLM $p(\chi^2)$
1. Elevation	147.74	-	1.24	***	1.77	0.832	0.913	0.999
2. Slope	31.56	+	1.24	***				
3. Radiation	2.42	-	1.00	0.12				

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

Table V. Summary of the GLM full and minimal models for the response variable open habitat species richness.

Response variable	Mean	Min	Max	No. of variables involved in the full GLM				Model features
Open habitat species richness	2.3	0	25	25				Poisson distribution, "log" link function
Variables retained in minimal GLM								
ID Var.	Deviance reduction	Coefficient value	VIF	$p(\chi^2)$	Mean Error (Leave-one-out cross validation)	D_{adj}^2	r	Difference between full and minimal GLM $p(\chi^2)$
1. Tree density	141.00	(-, +)	1.63	***	24.35	0.686	0.879	0.640
2. Cover % tree layer	110.54	(-, -)	2.76	***				
3. Forest estate	67.42		2.60	***				
4. PCA axis 1	49.84	+	1.60	***				
5. Latitude	26.30	+	1.53	***				
6. Folded aspect	11.01	+	1.39	***				
7. Forest physiognomy	10.19		1.35	***				

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

Downloaded By: [Bacaro, G.] At: 09:44 19 November 2008

VIF were below 10, indicating the lack of multicollinearity in predictors. The minimal model for species richness was statistically not different from the full model (Table III). The correlation between observed and predicted values was high ($r = 0.792$, $p < 0.001$; Figure 2a). Analysed as group of predictors, the stand structural variables at the local scale represented the highest explanatory power with respect to all others groups (Table VII).

Woody species richness

The overdispersion parameter for this model was lower than the threshold value of 1. Remarkably, only two predictors were found to be statistically significant. Elevation, in particular, caused a significant reduction in deviance (Table IV) and negatively shaped trees species richness. Slope, ranked as the second most important predictor in

Table VI. Summary of the GLM full and minimal models for the response variable bryophyte species richness.

Response variable	Mean	Min	Max	No. of variables involved in the full GLM			Model features	
Bryophyte species richness	8.2	0	22	26			Gaussian distribution, "identity" link function	
Variables retained in minimal GLM								
ID Var.	Deviance reduction	Coefficient value	VIF	$p(\mathbf{F})$	Mean error (Leave-one-out cross validation)	D_{adj}^2	r	Difference between Full and Minimal GLM $p(\mathbf{F})$
1. Basal area	158.78	(+, -)	1.20	***	2.71	0.640	0.812	0.959
2. Forest physiognomy	140.89		1.21	***				
3. Radiation	118.09	-	1.06	***				
4. Slope	61.66	+	1.05	***				

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

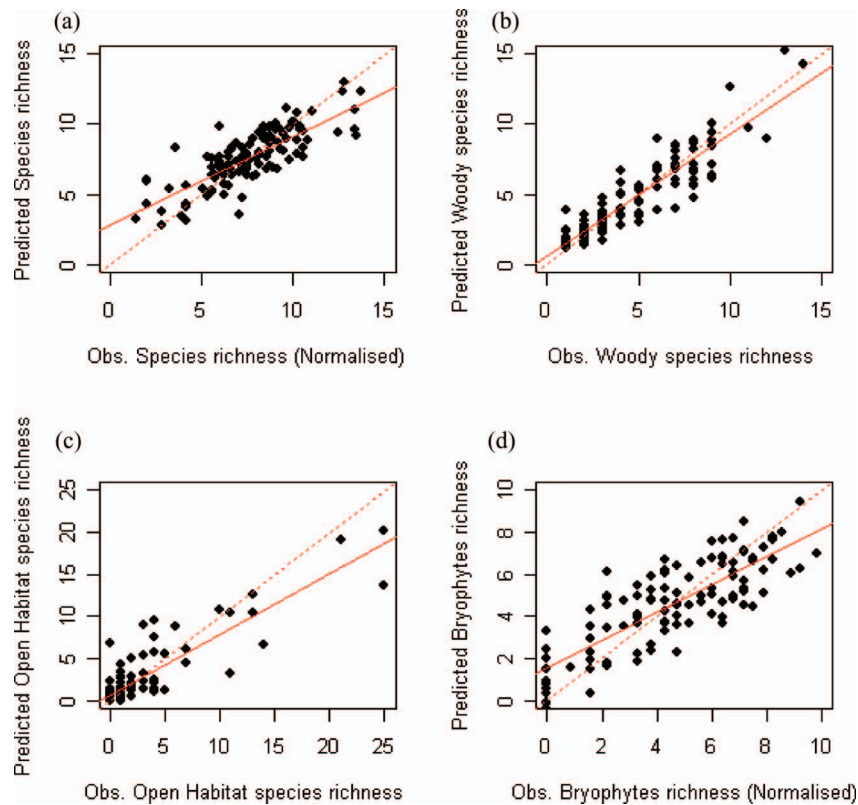


Figure 2. Relationship between the number of species predicted from the generalised linear models and that observed in the field for (a) total species richness (normalised by Box-Cox transformation), (b) woody species richness, (c) open habitat species richness and (d) bryophyte species richness (normalised by Box-Cox transformation). In each graph, continuous lines represent the calculated regression between predicted and observed response variables, while dotted lines denote the expected regression lines with intercept 0 and slope 1 (perfect prediction).

Table VII. Comparison between the amount of variance explained by each group of predictors in the minimal models for the different response variables.

% Explained variance	Minimal model (%)	Regional predictors* (%)	Local predictors – Abiotic variables* (%)	Local predictors – Forest structure* (%)
Vascular plant species richness	58	7	9	37
Woody species richness	83	20	19	1
Open habitat species richness	67	3	1	24
Bryophyte species richness	64	1	2	40

*Represents the non-shared effect of a single group of variables (i.e. species richness variation explained by regional or local predictors independent of other measured variables). In bold: the group of predictors explaining the highest proportion of variance for each response variable.

the minimal model, was positively correlated with woody species richness. The percentage of the explained variance by the minimal model was about 83%, a value statistically not different from that obtained for the full model ($p=0.999$, Table IV). The cross-validation of these predicted data resulted in a low mean error (± 1.755), in line with the calculated coefficient of correlation between observed and predicted data ($r=0.913$, $p < 0.001$; Figure 2b). The calculated VIF values were substantially below 10. Regional and Local (abiotic) predictors explained more or less the same amount of total variation (20%) and showed a high degree of covariation considering that the minimal model reached 83% of the total variation (Table VII).

Open habitat species richness

This response variable was 0 in a high number of plots. A Poisson distribution was therefore assumed to fit a general model and checked for overdispersion. About 69% of the total deviance was explained by the minimal model (Table V) but, different from the first two models, the stepwise procedure retained seven predictors in the minimal model. Excluding tree density and cover of the tree layer (highest deviance reduction, Table V), the other five predictors showed very similar values of explained variance. No statistical differences were found between the minimal and full models and, as for the previously described models, predictors were not correlated ($VIF < 10$). Mean (cross-validated) error was very high (24.350), even though the coefficient of correlation was quite high ($r=0.879$, $p < 0.001$; Figure 2c). Predictors in the forest structure group accounted for the highest amount of explained variance (Table VII).

Bryophytes species richness

Totally, 26 predictors were used in the Full model. The Poisson fitting showed high overdispersion (> 2) and a Box-Cox transformation was performed ($\gamma=0.6303$) to normalise the data. Four variables

were retained in the minimal model (Table VI). Basal area had the highest explanatory power (fitted as a polynomial term), while the categorical “forest physiognomy” was the second best (Table VI). All the variables introduced in the minimal model were not correlated with each other ($VIF < 10$). Bryophytes species richness was negatively related to radiation and positively related to slope. The total explained variance was 64% and the mean error calculated by the leave-one-out routine was 2.706 (Table VI). Predicted values of bryophytes resulted in a well-shaped linear relationship with the observed values ($r=0.829$, $p < 0.001$, Figure 2d). Finally, variables related to forest structure and management explained the largest amount of variation in the number of bryophyte species when fitted as a single group (Table VII).

Discussion

Vascular plant species richness

About 58% of the total variation in vascular plant species richness was explained by the variables used in the GLM. This value is not low if compared with the proportion of variance explained by this type of models at the regional scale (e.g. Pausas 1994; Ohmann & Spies 1998; Pausas et al. 2003) and with little loss of information compared with the model based on all available variables. For example, for plant species richness in New Zealand forest fragments, the total amount of variance explained by Ohlemüller et al. (2004) by using at least six predictors was only 36%.

In our model, the best predictor was a regional one, the only geo-morphological variable retained in the minimal model, elevation. A unimodal pattern of species richness was observed in relation to elevation. Many studies showed a negative relationship between species richness and elevation in various ecosystems (Stevens 1992; Pausas 1994; Rey-Benayas 1995). This has often been referred to a negative correlation between elevation and energy input (Pausas et al. 2003). However, in other studies a mid-domain peak of species richness distribution for intermediate

elevation values was also found (Currie 1991; Lomolino 2001; Colwell et al. 2004; Maccherini 2006). In a first analysis of these data, Chiarucci and Bonini (2005) found a unimodal relationship between species richness and elevation; albeit with a relatively low proportion of explained variance (16.3%).

Forest physiognomy was the second most important predictor of total plant species richness. By using species-area relationships, Chiarucci et al. (2001) found significant differences in the number of species among different forest types in Tuscan forests. They pointed out that mountain forests dominated by *Fagus sylvatica* have low species richness due to the temperature stress in winter. On the contrary, the *Quercus cerris* forests of hilly sites, with markedly lower climatic limitations, showed the highest species richness at all the spatial scales (from 1 to 2500 m²). These results, coupled with the soil features predictor (variable pH classes), confirm and describe, in quantitative terms, the fact that oak woods of hilly sites on soil with a pH close to neutrality or slightly acidic are the most plant species rich forest types in Italy (Blasi et al. 2004).

Two forest stand structure variables (local scale predictors), i.e. tree density and the cover of the tree layer, were included into the model and they were both negatively related to species richness. Intuitively, increasing total tree canopy cover corresponds to a decrease in resource availability, such as light, for many understorey plant species (Weiher 2003; Decocq et al. 2004). Weiher and Howe (2003) found that at large spatial scales species richness was higher with about 40% of tree canopy cover, while above this threshold plant species richness decreased. Possibly, the decline of species richness with increasing tree density was due to factors other than light availability, such as the previous coppicing activity – a standard management practice in many forest types of this area (Bernetti 1987). Moreover, tree density can be an indirect measure of root biomass and nutrient over-cropping by tree individuals. The resulting competition for resources among individuals in dense stands can lead to the elimination of less competitive species and a decline of species diversity (Guo 2003).

Noteworthy, while in this model the best predictor was a regional one, local scale predictors linked to forest structure had a high explanatory power. This is likely due to the “multipart” nature of the total plant richness variable, i.e. the sum of species richness for different functional plant groups, with potentially different responses to the same predictive variables.

Woody species richness

The models for woody species richness were much simpler than those obtained for the total species

richness but also more effective, explaining 83% of the total variance. Woody species richness showed an almost linear decrease in relation to elevation. This fact was already noted in a previous analysis of the same forests by Chiarucci and Bonini (2005), and was also reported for the tree flora of other countries, such as New Zealand (Ohlemüller & Wilson 2000, but see also review by Pausas & Austin 2001). In general, many authors assumed elevation as a surrogate for temperature (Pausas 1994; Rey-Benayas 1995). At higher altitudes, no or few broadleaved species can occur due to low temperatures, while the low-altitudinal forests are often rich in broadleaved shrubs and small trees. In fact, the shorter growing season at higher elevations prevents deciduous species from completing their annual cycle. However, general conclusions should be avoided given the hard link existing between temperature and water availability (Pausas & Austin 2001). With temperature probably being a major causal factor, detailed analyses of the elevation-temperature relationship could help to clarify this mechanism (Ohlemüller & Wilson 2000). In fact, the different environmental factors may not be singularly optimal but their combination is such to promote the maximum coexistence among species (Lomolino 2001).

Slope represented another explanatory morphological factor used to describe the topographical complexity of an area. The relationship between slope and species richness is not as well-described as for other environmental factors; one explanation takes into account the distribution of nutrients along the terrain profile: given the greater homogeneity of chemical, physical and biological plant growth factors, a smooth terrain profile should provide fewer niches than a steeper (and presumably uneven) profile, therefore having a lower species richness as well (Everson & Boucher 1998).

Although radiation was not defined in our model as statistically significant, it was retained in the minimal model for its contribution to deviance reduction, confirming its importance as previously documented. For example, Pausas (1994) found a negative relationship between radiation and plant species richness in Pyrenean *Pinus sylvestris* forests. If assumed as a surrogate of water availability (Pausas 1994), high radiation implies high values of evapotranspiration and, in general, drier environmental conditions (Smith & Huston 1989). In fact, the expected pattern results in low values of woody species richness where radiation is high.

In summary, for this model, a predominant role of regional factors and abiotic local factors in explaining woody species richness was observed. As pointed out by Zobel (1997) and Ohlemüller et al. (2004), tree species richness at a site can be limited by local

(e.g. number of niches) and by regional (the amplitude of the species pool) processes. A high geo-morphological differentiation at a larger scale can act as a factor that creates conditions leading to a large regional species pool. The results of the current model are likely to support the idea that the local features of the tree community at the stand scale considered (e.g. density, cover, basal area) had only a small effect on tree species richness.

Open habitat species richness

This model was the most complex, including a relatively high number (seven) of explanatory variables. The main effect was attributable to forest stand structure variables (local factors): tree density, followed by canopy cover of the tree layer, both with a negative effect. This is consistent with a number of previous studies showing how the cover and biomass of forest understorey vegetation often dramatically increase with canopy opening, and reporting, for example, the immediate dominance of the species favoured by disturbance (e.g. *Rubus* sp.pl.) after thinning (Decocq et al. 2004). This pattern was also described, with an increase in open habitat indicator species after coppicing, in Mediterranean forests (Pignatti & Pignatti 1968). Disturbance may provide opportunities for many species of open habitats that are often aggressive ruderals, to quickly spread into the forest (Beese & Bryant 1999). The results of this study support the hypothesis of the overstorey control on understorey diversity through controlling light availability (Estevan et al. 2007), and the positive effect of solar radiation in the model provides the evidence for the affinity of open habitat species with light availability.

The other variables, mainly regional predictors, gave a similar contribution to the model. The number of open habitat species was related to the forest estate and physiognomy, suggesting a spatial covariation of this variable. From an ecological point of view, these differences may depend on the different management practices, not included in the model. Nevertheless, the observed influence of the latitude variable represented an ecological gradient linked to forest estates and physiognomy.

The model of open habitat species was the only one which significantly included climatic components. Given the negative correlation of the first PCA axis with the rainfall components of climate (Table II), our model depicted an increasing number of open habitat indicator species in drier climatic sites. Other studies (e.g. Pysek et al. 2002; Ohlemüller et al. 2004) reported a positive relation between dry and warm conditions and the degree of invasibility by alien species in indigenous forest vegetation, as well as a decrease of alien species with

elevation. The observed pattern in our model may be explained, as previously noted, with the high degree of affinity of open habitat species for high radiation values.

Both local and regional factors played an important role in shaping the number of non-forest species. However, local factors contributed more to an overall deviance reduction (Table VII), in agreement with other observations (Grashof-Bokdam 1997). To improve the model, it would be necessary to develop a better classification of species responses, as available for Central and Northern Europe (Honnay et al. 1998; Hermy et al. 1999), since no similar classification is presently available for Mediterranean ecosystems.

Bryophytes richness

The model explaining bryophyte species richness was rather simple and efficient. One local predictor describing the forest structure, i.e. total basal area, and the categorical variable forest physiognomy were the most important factors. These two factors are related to forest management that is known to significantly affect bryophyte species diversity. For example, in a study on 36 Estonian forests, Vellak and Ingerpuu (2005) reported that even moderate management eliminates a significant number of species from the bryophyte communities. On the other hand, Humphrey et al. (2002) reported that bryophyte species richness was similar in plantations and semi-natural stands, while the pine plantations had a less species rich bryophyte flora than oak woodlands. In this study, since the majority of the stands classified as “artificial” were conifer plantations, the results well agree with those by Humphrey et al. (2002). Moreover Gustafsson et al. (1992) have reported that bryophyte species preferred broad-leaved woodlands with a long continuity, occurrence of large old trees, and with a high number of shrub and tree species, while bryophyte species richness decreased in managed forest plantations. However, the most important effect was the basal area of the stand, for which a unimodal response in terms of bryophyte species richness was observed, with a maximum at about 25–40 m² ha⁻¹ of basal area. After this threshold, bryophyte species richness decreased rather constantly. Although the analysis of species composition was not the goal of this study, it is likely that a more specialised shade-tolerant bryophyte flora characterises the plots with higher basal area.

Solar radiation was found to be the fourth most important factor in explaining bryophyte species richness. Although there are no experimental studies on the effects of solar radiation on bryophyte species richness in Mediterranean forests, it is intuitive that

solar radiation could be a very restrictive factor for the richness of this group in this area. Summer drought is an important limiting factor for many Tuscan forests (Bernetti 1987), and bryophytes are certainly more prone to suffer from this effect. In fact most bryophytes are “poikilohydric”, lacking the sophisticated water-control mechanisms of vascular plants, so high temperatures combined with infrequent precipitation in summer, result in high stress for bryophytes.

Conclusions

The power of modelling processes depends on the selection of appropriate predictors. Luoto et al. (2002) raised concerns about stepwise modelling procedures since they do not necessarily select the ecologically most important variables, thus forcing a careful pre-selection of the predictors to be used. Moreover the estimation of species richness is not independent from the spatial scale components, such as grain and extent. These play a crucial role and their effects on the statistical results should be considered, since they affect the conclusions about ecological patterns and processes (Dalthorp 2004).

The models obtained in the present study showed a large amount of unexplained variance; to what extent this is due to environmental predictors not included or to stochastic factors is unknown. Errors in the data due to sampling problems (see e.g. Baffetta et al. 2007) or to GIS-derived variables may also contribute to unexplained errors (Pausas et al. 2003). Several authors (Minchin 1989; Pausas 1994; Austin 2002) reported markedly different types of species richness patterns and, subsequently, models for different life-forms. In this debate, our models for different functional or taxonomic plant groups provide another example (after Richerson & Lum 1980; Minchin 1989; Moore & Keddy 1989; Cox & Larson 1993; Pausas 1994) of Peet's (1978) hypothesis that patterns of species richness cannot be congruent when dealing with such different groups.

In any case, the proportion of explained variance was in line with that of similar models developed for a range of geographic areas (Pausas 1994; Ohmann & Spies 1998; Pausas et al. 2003), suggesting that these models can be used as baselines for the interpretation of species diversity patterns in Tuscan forests or for future testing of specific hypotheses. This is particularly important in applied research, such as forest monitoring programmes, which are often based on non-probabilistic samples (Ferretti & Chiarucci 2003) and on the need for specific references for model-based inference and generalisation.

Acknowledgements

This work is part of the research project MONITO, funded by the Regional Administration of Tuscany and the European Community under the Regulations EEC 2157/92 and 3528/86. Many people contributed to this work and we wish to express our gratitude to all of them: A. Gabellini, F. Casini, A. Riva, L. Frati, M. Cetoloni, F. Ghini, D. Viciani, M. Boddi and G. Nicoletti who participated in the field sampling. Special thanks are due to Ramona Magno who allowed us to analyse the climatic data-set of the project “DESERNET” INTERREG IIIB – MEDOCC – Regione Toscana – IBIMET – CNR. Comments by M. Diekmann and an anonymous referee on previous drafts of this paper are greatly appreciated.

Reference

- Anderson MJ, Gribble NA. 1998. Partitioning the variation among spatial, temporal and environmental components in a multivariate data set. *Aust J Ecol* 23:158–167.
- Austin MP. 2002. Spatial prediction of species distribution: An interface between ecological theory and statistical modelling. *Ecol Model* 157:101–118.
- Baffetta F, Bacaro G, Fattorini L, Rocchini D, Chiarucci A. 2007. Multi-stage cluster sampling for estimating average species richness at different spatial grains. *Comm Ecol* 8:119–127.
- Bartolozzi L, Bonini I, Boretti R, Bussotti F, Cenni E, Chiarucci A, Cozzi A, De Dominicis V, Ferretti M, Grossoni P, Landi G, Leonzio C, Vignozzi G. 2002. Forest ecosystem monitoring in Tuscany (Italy): Past activities, present status and future perspectives. *J Limnol* 61:129–136.
- Beese WJ, Bryant AA. 1999. Effect of alternative silvicultural systems on vegetation and bird communities in coastal montane forests of British Columbia, Canada. *For Ecol Manage* 115:231–242.
- Bernetti G. 1987. I boschi della Toscana. Giunta Regionale Toscana. Bologna: Edagricole.
- Blasi C, Di Pietro R, Filesi L. 2004. Syntaxonomical revision of *Quercetalia pubescenti-petraeae* in the Italian Peninsula. *Fitosociologia* 41:87–164.
- Borcard D, Legendre P, Drapeau P. 1992. Partialling out the spatial component of ecological variation. *Ecology* 73:1045–1055.
- Box GEP, Cox DR. 1964. An analysis of transformations. *JRSS B* 26:211–246.
- Chiarucci A, Bonini I. 2005. Quantitative floristics as a tool for the assessment of plant diversity in Tuscan forests. *For Ecol Manage* 212:160–170.
- Chiarucci A, D' Auria F, Bonini I. 2007. Is vascular plant species diversity a predictor of bryophyte species diversity in Mediterranean forests? *Biodivers Conserv* 16:525–545.
- Chiarucci A, De Dominicis V, Wilson JB. 2001. Structure and floristic diversity in permanent monitoring plots in forest ecosystems of Tuscany. *For Ecol Manage* 141:201–210.
- Colwell RK, Rahbek C, Gotelli NJ. 2004. The mid-domain effect and species richness patterns: What have we learned so far? *Am Nat* 163:E1–E23.
- Cox JE, Larson DW. 1993. Environmental relations of the bryophyte and vascular components of a talus slope plant community. *J Veg Sci* 4:553–560.

- Crawley MJ. 1993. GLIM for ecologist. Oxford: Blackwell.
- Currie DJ. 1991. Energy and large-scale patterns of animal- and plant-species richness. *Am Nat* 137:27–49.
- Dalthorp D. 2004. The generalized linear model for spatial data: Assessing the effects of environmental covariates on population density in the field. *Entomol Exp Appl* 111:117–131.
- Davison AC, Hinkley DV. 1997. Bootstrap methods and their application. Cambridge: Cambridge University Press.
- Decocq G, Aubert M, Dupont F, Alard D, Saguez S, Wattez-Franger A, De Foucault B, Delelis-Dusollier A, Bardat J. 2004. Plant diversity in a managed temperate deciduous forest: Understorey response to two silvicultural systems. *J Appl Ecol* 41:1065–1079.
- EEA. 2005. IMAGE2000 and CLC2000. Products and methods CORINE land cover updating for the year 2000. Italy: Ispra.
- Engler R, Guisan A, Rechsteiner L. 2004. An improved approach for predicting the distribution of rare and endangered species from occurrence and pseudo-absence data. *J Appl Ecol* 41:263–274.
- Estevan H, Lloret F, Vayreda J, Terradas J. 2007. Determinants of woody species richness in Scot pine and beech forests: Climate, forest patch size and forest structure. *Acta Oecol* 31:325–331.
- Everson DA, Boucher H. 1998. Tree species-richness and topographic complexity along the riparian edge of the Potomac River. *For Ecol Manage* 109:305–314.
- Ferretti M, Chiarucci A. 2003. Design concepts adopted in long-term forest monitoring programs in Europe—problems for the future? *Sci Total Environ* 310:171–178.
- Fox J, Monette G. 1992. Generalized collinearity diagnostics. *J Am Stat Assoc* 87:178–183.
- Grashof-Bokdam C. 1997. Forest species in an agricultural landscape in the Netherlands: Effects of habitat fragmentation. *J Veg Sci* 8:21–28.
- Guirado M, Pino J, Roda F. 2006. Understorey plant species richness and composition in metropolitan forest archipelagos: Effects of forest size, adjacent land use and distance to the edge. *Global Ecol Biogeogr* 15:50–62.
- Guisan A, Edwards J, Hastie T. 2002. Generalized linear and generalized additive models in studies of species distributions: Setting the scene. *Ecol Model* 157:89–100.
- Guisan A, Zimmermann NE. 2000. Predictive habitat distribution models in ecology. *Ecol Model* 135:147–186.
- Guo Q. 2003. Plant abundance: The measurement and relationship with seed size. *Oikos* 101:639–642.
- Gustafsson L, Fiskesjö A, Hallingbäck T, Ingelög T, Pettersson B. 1992. Semi-natural deciduous broadleaved woods in southern Sweden – habitat factors of importance to some bryophyte species. *Biol Conserv* 59:175–181.
- Hastie T, Pregibon M. 1992. Generalized linear models. In: Chambers TJ, Hastie T, editors. *Statistical models*. S. Cole: Wadsworth and Brooks.
- Hermý M. 1994. Effects of former land use on plant species diversity and pattern in European deciduous woodlands. In: Boyle TJB, Boyle CEB, editors. *Biodiversity, temperate ecosystems and global change*. Berlin: Springer Verlag. pp. 123–144.
- Hermý M, Honnay O, Firbank L, Grashof-Bokdam C, Lawesson JE. 1999. An ecological comparison between ancient and other forest plant species of Europe, and the implications for forest conservation. *Biol Conserv* 91:9–22.
- Honnay O, Hermý M, Coppin P. 1998. Effects of area, age and diversity of forest patches in Belgium on plant species richness, and implications for conservation and reforestation. *Biol Conserv* 87:73–84.
- Humphrey JW, Davey S, Peace AJ, Ferris R, Harding K. 2002. Lichens and bryophyte communities of planted and semi-natural forests in Britain: The influence of site type, stand structure and deadwood. *Biol Conserv* 107:165–180.
- Huston MA. 1999. Local processes and regional patterns: Appropriate scales for understanding variation in the diversity of plants and animals. *Oikos* 86:393–401.
- Kikvidze Z, Ohsawa M. 2002. Measuring the number of co-dominants in ecological communities. *Ecol Res* 17:519–525.
- Legendre P, Legendre L. 1998. *Numerical ecology*. 2nd English ed. Elsevier: Amsterdam, The Netherlands.
- Lomolino MV. 2001. Elevation gradients of species-density: Historical and prospective views. *Global Ecol Biogeogr* 10:3–13.
- Luoto M, Toivonen T, Heikkinen RK. 2002. Prediction of total and rare plant species richness in agricultural landscapes from satellite images and topographic data. *Landscape Ecol* 17:195–217.
- Maccherini S. 2006. Factors associated with species richness in a remnant calcareous grassland. *Grass Sci* 52:181–184.
- McCullagh P, Nelder JA. 1989. *Generalized linear models*. London: Chapman and Hall.
- McCune B, Keon D. 2002. Equations for potential annual direct incident radiation and heat load. *J Veg Sci* 13:603–606.
- Minchin PR. 1989. Montane vegetation of the Mt. Field massif, Tasmania: A test of some hypotheses about properties of community patterns. *Plant Ecol* 83:97–110.
- Moore DRJ, Keddy PA. 1989. The relationship between species richness and standing crop in wetlands: The importance of scale. *Vegetatio* 79:99–106.
- Ohlemüller R, Bannister P, Dickinson CJ, Walker BH, Anderson BJ, Wilson JB. 2004. Correlates of vascular plant species richness in fragmented indigenous forests: Assessing the role of local and regional factors. *Comm Ecol* 5:45–54.
- Ohlemüller R, Walker S, Wilson JB. 2006. Local vs regional factors as determinants of the invasibility of indigenous forest fragments by alien plant species. *Oikos* 112:493–501.
- Ohlemüller R, Wilson JB. 2000. Vascular plant species richness along latitudinal and altitudinal gradients: A contribution from New Zealand temperate rainforests. *Ecol Lett* 3:262–266.
- Ohmann JL, Spies TA. 1998. Regional gradient analysis and spatial pattern of woody plant communities of Oregon forests. *Ecol Monogr* 68:151–182.
- Oksanen J, Minchin PR. 2002. Continuum theory revisited: What shape are species responses along ecological gradients? *Ecol Model* 157:119–129.
- Palmer MW, Earls PG, Hoagland BW, White PS, Wohlgemuth T. 2002. Quantitative tools for perfecting species lists. *Environmetrics* 13:121–137.
- Pausas JG. 1994. Species richness patterns in the understorey of Pyrenean *Pinus sylvestris* forest. *J Veg Sci* 5:517–524.
- Pausas JG, Austin MP. 2001. Patterns of species richness in relation to different environments: An appraisal. *J Veg Sci* 12:153–166.
- Pausas JG, Carreras J, Ferré A, Font X. 2003. Coarse-scale plant species richness in relation to environmental heterogeneity. *J Veg Sci* 14:661–668.
- Peet RK. 1978. Forest vegetation of the Colorado front range: Patterns of species diversity. *Plant Ecol* 37:65–78.
- Pignatti E, Pignatti S. 1968. Die Auswirkungen von Kahlschlag und Brand auf das Quercetum ilicis von Süd-Toskana, Italien. *Folia Geobot Phytotx* 3:17–46.
- Pignatti S. 1982. *Flora d'Italia*. Bologna: Edagricole.
- Pysek P, Jarosik V, Kucera T. 2002. Patterns of invasion in temperate nature reserves. *Biol Conserv* 104:1324.
- R Development Core Team. 2007. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Rey-Benayas JM. 1995. Patterns of diversity in the strata of boreal montane forest in British Columbia. *J Veg Sci* 6:95–98.

- Ribichich MA, Protomastro J. 1998. Woody vegetation structure of xeric forest stands under different edaphic site conditions and disturbance histories in the Biosphere Reserve 'Parque Costero del Sur', Argentina. *Plant Ecol* 139:189–201.
- Richerson PJ, Lum K. 1980. Patterns of plant species diversity in California: Relation to weather and topography. *Am Nat* 116:504–536.
- Rushton SP, Ormerod SJ, Kerby G. 2004. New paradigms for modelling species distributions? *J Appl Ecol* 41:193–200.
- Smith T, Huston M. 1989. A theory of spatial and temporal dynamics of plant communities. *Vegetatio* 83:49–69.
- Stevens GC. 1992. The elevational gradient in altitudinal range: An extension of Rapoport's latitudinal rule to altitude. *Am Nat* 140:893–911.
- Taylor AH, Allen JI, Clark PA. 2002. Extraction of a weak climatic signal by an ecosystem. *Nature* 416:629–632.
- Vellak K, Ingerpuu N. 2005. Management effects on bryophytes in Estonian forests. *Biodivers Conserv* 14:3255–3263.
- Weiher E. 2003. Species richness along multiple gradients: Testing a general multivariate model in oak savannas. *Oikos* 101:311–316.
- Weiher E, Howe A. 2003. Scale-dependence of environmental effects on species richness in oak savannas. *J Veg Sci* 14:917–920.
- Zobel M. 1997. The relative of species pools in determining plant species richness. An alternative explanation of species co-existence? *Trends Ecol Evol* 12:266–269.