

Predicting the geographical distribution of plant communities in complex terrain – a case study in Fushian Experimental Forest, northeastern Taiwan

Chi-Ru Chang, Pei-Fen Lee, Mei-Ling Bai and Tzer-Tong Lin

Chang, C.-R., Lee, P.-F., Bai, M.-L. and Lin, T.-T. 2004. Predicting the geographical distribution of plant communities in complex terrain – a case study in Fushian Experimental Forest, northeastern Taiwan. – *Ecography* 27: 577–588.

Ecosystem management and biodiversity conservation are usually implemented using information of several targeted species or cover-types and usually do not include information about communities. This is not because community-level information is unimportant for management purposes, but because the detailed fieldwork required for gathering community-level information at the scale for ecosystem management is usually impractical. We propose two methods to estimate the geographical distribution of plant communities with the objectives of covering large areas with minimal field efforts. The first method estimates the geographical distribution of plant communities by combining clustering methods with vegetation modeling, and the second extrapolates the geographical distribution of gradients in plant communities by combining gradient analysis with vegetation modeling. Vegetation modeling with clustering methods can be used to allocate sites with potentially higher alpha diversity, with the benefit of having a list of species associated with the clustered type. Vegetation modeling with gradient analysis can be used to identify regions with potentially the highest beta diversity by means of selecting regions with the widest range or highest variability in major DCA axes scores, and thereby help to preserve the scope of environmental conditions that lead to diversity in species assemblages. This is especially important because biological entities such as species, communities, or even ecosystems may cease to exist in the long run, and the preservation of processes that lead to biodiversity will eventually become more meaningful. We conclude that new methods to study and manage the processes that contribute to biodiversity at all scales should be and can be developed.

C.-R. Chang, Dept of Landscape Architecture, Chinese Culture Univ., 55 Hwakang Road, Yangmingshan, Shihlin, Taipei 111, Taiwan. – P.-F. Lee (correspondence: leepf@ntu.edu.tw), Inst. of Ecology and Evolutionary Biology, National Taiwan Univ., 1, Sec. 4, Roosevelt Road, Taipei 106, Taiwan. – M.-L. Bai, Dept of Zoology, National Taiwan Univ., 1, Sec. 4, Roosevelt Road, Taipei 106, Taiwan. – T.-T. Lin, Fushan Experimental Forest, Taiwan Forest Research Inst., P.O. Box 132, Ilan, Taiwan (present address of M.-L.B., Amdtstr. 14116, 37075 Göttingen, Germany).

Describing the geographical distribution of vegetation is integral to the understanding and management of ecosystems and the conservation of biodiversity. It provides clues to the ecological processes involved in creating and maintaining the current vegetation composition and structure (Hubbell et al. 1999), and can serve

as a basis for the design of survey, research, and management strategies for the conservation of biodiversity (Margules 1989, Ahearn et al. 1990, Scott et al. 1993, Anon. 1994a, Gaston 1994, Alexander 1995). With the recent development and increasing availability of remotely sensed data and geographic information systems

Accepted 22 March 2004

Copyright © ECOGRAPHY 2004
ISSN 0906-7590

(GIS), ecological studies now have the tools to help broaden their horizons. For example, geographic information systems, when combined with species survey data or remotely sensed data of vegetation types, have been compared to broad-scale environmental variables such as terrain or soil types and anthropogenic factors to infer processes associated with the geographical distribution of vegetation patterns (e.g. Austin 1998, Tappeiner et al. 1998), to predict the occurrences of vegetation at a broader scale through vegetation modeling (e.g. Davis and Goetz 1990, Ostendorf and Reynolds 1998), and to provide a scientific basis for the management and conservation of biodiversity in plants (e.g. Austin 1998) and animals (e.g. Smith et al. 1997, Deangelis et al. 1998, Jarvis and Robertson 1999).

Remotely sensed data, including aerial photographs and satellite images, when compared to field surveys, have the advantage of being easier to obtain, and can be recorded at much higher frequencies and over larger areas. When applicable, remotely sensed data can greatly improve the availability of vegetation data (Roberts et al. 1997, Asner et al. 1998). Unfortunately, because of the resolution of remotely sensed data, most applications tend to classify vegetation into types or classes. Such a classification usually lends itself to the description and differentiation of more distinct vegetation, where the between-class differences are much higher than the within-class variation. This may not be a major problem in actively managed forests or forests dominated by a few species, where the boundaries between different patches are more evident. This may also be less of a problem when the topic of interest falls within a broader spatial scale where the gradient is significant enough to clearly separate vegetation into distinct types, and the finer scale differences are of less interest to the researcher or manager. However, there are certainly cases where the diversity and fine-grained heterogeneity of the forests can cause difficulties in using remotely sensed data to separate different vegetation types into distinct patches (e.g. Plumb 1991).

The study of plant communities, for example, often requires information on much finer-scale details, and the changes in spectral signals among communities are often both slight and gradual. This problem is further aggravated when tropical and subtropical rain forests are the topic of interest. These forests have high species diversity, a variety of life-forms, and a very complex forest structure. In these systems, the diversity and continuity of vegetation can make image-based delineation of distinct patches a difficult task, even when different types can be recognized on the ground. Both aerial photographs and satellite images often cannot differentiate these forests into communities and types (Heller and Ulliman 1983, Westman et al. 1989, Apan 1997). As a result, developments in remote sensing and geographic information systems have been mostly

applied to and benefited regional-scale vegetation studies but not community-scale efforts.

While it is true that community-level studies are generally targeted at understanding the ecological processes that maintain species composition and biodiversity, and therefore have to rely heavily on field surveys to document details about species composition and the environments in which they occur, the same level of intensity may not be required for management efforts or for applied problems, such as the description and geographical identification of potential fauna habitats. Similarly, the vast efforts needed to acquire community-level information when remotely sensed data cannot meet our needs should not be an excuse to set management objectives that exclude the conservation of plant communities or more importantly, the different processes that lead to different plant assemblages.

In this paper, we propose two methods to handle problems concerning the geographical distribution of plant communities when their heterogeneity exceeds the resolution of remote-sensing techniques, with the objectives of covering large areas with minimal field efforts. We do this by combining multivariate techniques commonly used in community ecology with geographic information systems. The first method estimates the geographical distribution of plant communities by combining clustering methods with vegetation modeling, and the second displays the geographical distribution of gradients in plant communities by combining gradient analysis with vegetation modeling. A search algorithm is also applied to the results from the second method in search of a fixed-sized area with the maximum beta-diversity. The methods are described using data from a subtropical forest in northeastern Taiwan, but can easily be applied to other areas.

Study area

Fushan Experimental Forest is an International Long-Term Ecological Research (ILTER) site located 40 km southeast of Taipei in northeastern Taiwan, east Asia (24°46'N, 121°43'E) (Fig. 1). The forest covers 1097.9 ha, spanned from north to south by a branch of the Snow Mountain Range, with its highest parts to the north. To the east and west of this ridge are two rivers: Tsu-keng creek to the southeast, and Har-pen creek to the west (Chang et al. 1986). The forest is predominantly in mountainous terrain, covering 386–1451 m in elevation (Bai et al. 2000). The climate is warm and moist, with an annual average temperature of 18.3°C, a mean annual precipitation of 4067 mm, and an average humidity of 96%. Typhoons frequent the area during summer, with maximum wind gusts over 35 m s⁻¹ (Mabery et al. 1998).

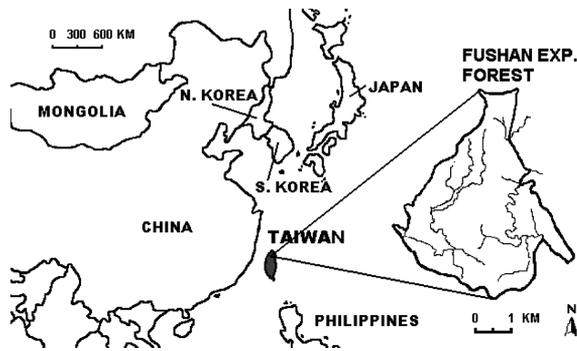


Fig. 1. Location of the Fushan Experimental Forest and the Taiwan island. The two rivers in the area are Tsu-keng creek to the southeast, and Har-pen creek to the west.

The soil is strongly acidic (pH 3.8–5.0), with a moderate quantity of organic matter in the sublayer (2–5%). Yellow soil is found on ridges and stable slopes, colluvial soil is found on midhills and foothills, and lithosols on very steep slopes or stream beds (Lin et al. 1996).

The vegetation can be classified as a subtropical rain-forest, grading into warm temperate rain-forests 1000 m a.s.l. Dominant tree species are mostly broadleaf, belonging to the families Fagaceae, Lauraceae, Theaceae, Ebenaceae, and Juglandaceae. Chinkapin *Castanopsis cuspidata* var. *carlesii* is the most dominant species. Other common tree species include *Machilus thunbergii*, *Engelhardia roxburghiana*, *Litsea acuminata*, and *Machilus zuihoensis*. The understory is composed of dense shrubs such as *Blastus cochinchinensis* and tree ferns such as *Cyathea podophylla*. The ground-layer is covered with ferns, with *Diplazium dilatatum* being the most dominant (Kuo 1994, Lin et al. 1996, Lin 1997, Chang 1998).

Methods

Vegetation data

Trees of the Fushan forest were surveyed between August 1992 and July 1994 (Lin 1997). Because the purpose of this survey was a preliminary study to document species composition across the landscape, a rough sampling scheme was used. The landscape was first stratified into relatively homogeneous areas in situ, resulting in 53 patches with undefined boundaries. In the proximity of the centre of these relatively homogeneous patches, 50 × 10 m sample plots were laid out and the centre of their locations mapped. For each plot, all trees taller than 1.3 m were counted, identified, and their diameter at breast-height (DBH) measured.

Environmental data

Although climate data were routinely collected for the area, they were not spatially explicit. Therefore, spatially explicit environmental data were derived from existing maps of terrain, streams, and soil (Table 1). The spatial location of each grid cell was recorded using the Transverse Mercator projection coordinate system (mE and mN).

A digital elevation model (DEM) was created using Arc/Info (Anon. 1997) based on a 1:5000 isoclines map of the area. Secondary information derived from the DEM includes topographic information and solar radiation.

Topographic information derived from the DEM includes slope (degrees), aspect (degrees), and topographic location. Because aspect is a circular variable, we transformed it into two derived variables, southness and westness, where

Table 1. List of environmental variables and their highly inter-correlated variables.

	Explanation	Units	Most correlated variables		
			1st	2nd	3rd
TMX	Location in the east-west direction	mE			
TMY	Location in the north-south direction	mN			
ELEV	Elevation	m	MDR (0.995)*	MTR (0.992)	VDR (0.662)
SLP	Slope	degree	WLS (-0.502)		
RIVD	Distance to river	m	TOPO (0.710)	WLS (0.680)	LS (-0.520)
TOPO	Topographic position	-1–1	WLS (0.801)	RIVD (0.710)	LS (-0.708)
SOUTH	Southness	0–180	VTR (-0.736)	VDR (-0.530)	
WEST	Westness	0–180			
YS	Yellow soil		LS (-0.542)		
CS	Colluvial soil		LS (-0.932)		
LS	Lithosol		CS (-0.932)	TOPO (-0.708)	YS (-0.542)
MDR	Mean annual direct radiation	cal cm ⁻² d ⁻¹	ELEV (0.995)	TOPO (0.646)	WLS (0.639)
MTR	Mean annual total radiation	cal cm ⁻² d ⁻¹	ELEV (0.992)	TOPO (0.646)	WLS (0.636)
VDR	Variance in annual direct radiation		VTR (0.925)	ELEV (0.662)	MDR (0.587)
VTR	Variance in annual total radiation		VDR (0.925)	SOUTH (-0.736)	
WLS	Whole light sky-space	%	TOPO (0.801)	RIVD (0.680)	ELEV (0.642)

* Correlated variables are listed in the order of the strength of their correlations, as shown in the brackets. Only the variables with correlations >0.5 are shown.

southness = $180 - |\text{aspect} - 180|$

so that south would be 180 and north 0; and

westness = $|180 - |\text{aspect} - 270||$

so that west would be 180 and east 0. Topographic location is expressed as an index:

$$\frac{\text{Distance from valley} - \text{Distance from ridge}}{\text{Distance from valley} + \text{Distance from ridge}}$$

This index has values between -1 and $+1$, with higher values indicating upper slope areas closer to ridges, and lower values closer to valleys.

Solar radiation information includes whole light sky-space and mean annual direct, diffuse, and total solar radiation. Whole light sky-space (WLS) is defined as the proportion of sky-space not blocked by surrounding terrain (Hsia and Wang 1985). With this method, a circle is used to represent the sky-space (a hemisphere), with the angles indicating aspect, and the radius divided into 90 units to represent the angle at which our view of the sky is blocked. In this study, 16 viewing aspects were used, and the points with the largest blocked angle along each viewing aspect were used to define a polygon representing the sky-space not blocked by surrounding terrain. WLS is calculated by dividing the area of this polygon by the area of the full circle, with a minimum value of 0 and a maximum of 1. Thus, a smaller value of WLS represents a smaller proportion of the sky-space being blocked by surrounding terrain, and thus receives more solar radiation, and vice versa. The remaining solar radiation variables (in $\text{cal cm}^{-2} \text{d}^{-1}$) were calculated with SolarImg (Harmon and Marks 1995), a program developed for the H. J. Andrews Experimental Forest in Oregon, USA. This program estimates solar radiation input to a site based on the site's latitude, elevation, aspect, slope, and cloud cover, following the methods of Bonan (1989) and Nikolov and Zeller (1992).

Stream data were digitized using a 1:5000 stream map of the area. Distance to streams (m) is defined as the minimum Euclidean distance of a site to streams, and was calculated using Arc/Info.

Soil data were derived from soil survey maps by Lin et al. (1996). The original soil-series data were used to create maps containing information on soil types as suggested by the authors (yellow soil, colluvial soils, and lithosols), parent material (shale or slate), and soil taxonomic order (ultisol, inceptisol, and entisol).

All digital maps were transformed into 12.5×12.5 m grid maps, and environmental variables were retrieved for the grid cells corresponding to the mapped centre of the vegetation sample plots. These variables were then checked for correlations, and their major gradients summarized with Principal Components Analysis (PCA) using the correlation matrix and the program PC-ORD (McCune and Mefford 1999).

Vegetation analysis and modeling

Detrended Correspondence Analysis (DCA, Hill and Gauch 1980) was used to examine the major vegetation gradients in the data, using PC-ORD. To minimize the influence of local species, species that were present in $<5\%$ of the sample plots were removed from the analysis. Graphic examination, correlation statistics, and regression analyses were then used to assess the importance of environmental factors in determining the major species DCA axes, using the statistical program S-Plus (Anon. 2001). All above-mentioned environmental variables were used in the analyses, among which the edaphic factors were classified as types, and treated as dummy variables. Stepwise multivariate regression was also used to build models that predict major axes DCA scores using various environmental variables. The resulting models were displayed using Arc/Info to describe the gradients of plant communities across the landscape.

To find fixed-sized areas with the maximum beta-diversity, the major DCA axes are used as indicators of the most important gradients. In our example, we used an algorithm that searched for the 49 ha quadrat (ca 50 ha, a size often used for large permanent plots) with the maximum range of DCA scores in the first three axes. To do this, we first calculated the ranges (maximum minus the minimum values) of the three DCA axes scores across the full landscape. We then systematically moved a 49 ha window (700×700 m) across the landscape, and calculated the range for each of the three DCA axes scores within the window. These three values were then each divided by their corresponding full-landscape ranges, resulting in three values indicating the percentage of the total range of DCA scores included in that window. These three values were then added and saved as an index of beta-diversity for the specific window. The window with the maximum index of beta-diversity was then saved and reported. We chose to use the first three DCA axes because the results of our studies (Bai et al. 2000 and Chang et al. unpubl.) suggest that the contributors to these three axes are all valid, but applications to other sites will have to adjust this criterion according to the researcher's knowledge of their sites. For comparison, we also searched for the 4 ha quadrat (200×200 m window) with maximum beta-diversity using the same methods.

Clustering of vegetation into classes of community types was performed with the TWINSpan (two-way indicator species analysis) algorithm, also using the program PC-ORD. Pseudo-species cut levels were set at 0, 2, 5, 10, and 20. Discriminant analysis was then used to build a model that classifies the sample sites into different vegetation types according to their environmental factors. A best-fit model was selected using a forward-selection algorithm using the statistical program SAS (Anon. 1994b). The resulting model was used in conjunction with environmental layers in the

geographical information systems to estimate the geographical distribution of plant communities, and the characteristics of the misclassified sites analyzed using classification trees using the program S-plus.

The potential contributors of alpha-biodiversity were analyzed by calculating the number of species occurring at a sample plot, and using graphic examination, correlation statistics, and generalized linear model to assess the difference in biodiversity among vegetation associations and the importance of environmental factors in determining site alpha-biodiversity. The generalized linear model was estimated using S-plus with a Poisson error distribution and a log-link function, and tested with an ANOVA with a χ^2 distribution. The final model was selected similar to a backward-selection algorithm, by initially adding all predictor variables, sequentially removing all insignificant predictor variables and rearranging the order of predictor variables in the order of significance, until all the predictor variables are significant and arranged in the order of significance.

Results

Species composition

A total of 92 tree species belonging to 37 families were encountered. The families with the most number of species were Lauraceae (13 species), Symplocaceae (7 species), and Theaceae (7 species). The forest is relatively evenly-dominated, with five of the most dominant species needed to account for 50% of the total basal area. *Castanopsis cupsidata* var. *carlesii*, *Machilus thunbergii*, and *Litsea acuminata* were the most dominant and widespread species.

Environmental gradients

Because many of the environmental variables are based on the different elements of topography, they are highly correlated (Table 1). For example, topographic position is highly correlated to elevation, distance to river, and the different representations of solar radiation, especially

whole light sky-space. The different representations of solar radiation are highly correlated to each other and to topographic factors, with whole light sky-space most correlated to topographic position, mean annual direct and mean annual total solar radiation and variance in annual direct solar radiation most correlated to elevation, and variance in annual total solar radiation most correlated to southness. Also, because the soil type lithosol is closely related to river bottoms, it is correlated to topographic position and distance to streams.

A PCA of the environmental variables also identifies this strong correlation between many of the environmental variables (Table 2), with the first PCA axis determined by variables related to the topographic location of the site, such as elevation, topographic position, distance to streams, the various representations of solar radiation, and the highly topography-related lithosol. The second, third, and sixth PCA axes are determined by soil types, with the second axis determined by soils with slate as parent material, the third by soils with shale as parent material, and the sixth by yellow soils. The fourth and fifth axes are determined by aspect, with the fourth axis determined by southness and the variance in annual total solar radiation, and the fifth by westness.

Gradient analysis

The first DCA axis separated the valley sites from the rest, and the second axis separated the mid-slope sites from the ridge sites (Fig. 2). The DCA eigenvalues were not high (Table 3), but the eigenvalue of axis 1 was considerably higher than those of axes 2 and 3, indicating that most of the extractable variance in species composition can be accounted by the first DCA axis.

When DCA scores were compared to environmental variables, the first DCA axis was strongly related to topographic position and soil types, the second axis was most related to elevation and factors related to total annual solar radiation input, such as whole light sky-space and mean annual direct and mean annual total solar radiation, and the third axis most related with southness, slope, and the variance in annual total solar

Table 2. Results from principal components analysis of environmental variables from the sample plots. The variable codes follow those of Table 1.

	PCA1	PCA2	PCA3	PCA4	PCA5	PCA6
Eigenvalue	8.539	4.905	3.912	3.597	2.512	2.164
% variance	23.720	13.624	10.868	9.990	6.977	6.010
Cumulative variance	23.720	37.344	48.212	58.202	65.179	71.189
Loadings (in order of importance):						
Variable 1	Lithosol-0.3019	Slate-0.4046	Shale-0.3421	SOUTH-0.4581	WEST-0.3747	Yellow-0.3286
Variable 2	TOPO-0.2772	MTR-0.216	TMX-0.3111	VTR-0.4295	SLP-0.2309	WEST-0.2174
Variable 3	ELEV-0.2563	MDR-0.2149	Yellow-0.242	VDR-0.337	Colluvial-0.2084	TMY-0.1687
Variable 4	MDR-0.2485	ELEV-0.2094	RIVD-0.2408	TMY-0.2422	WLS-0.2002	Colluvial-0.1446
Variable 5	MTR-0.2456	WEST-0.2089	WLS-0.2092	SLP-0.1311	MTR-0.1907	TMX-0.1102

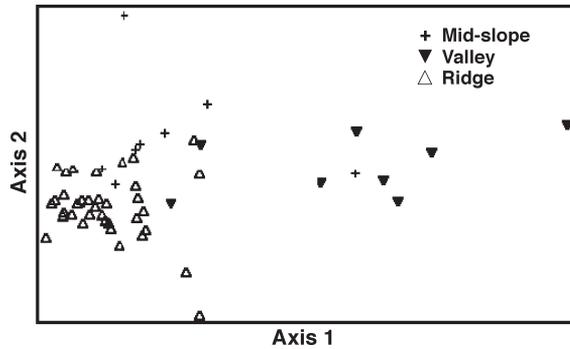


Fig. 2. Sites plotted on DCA axes 1 and 2 and displayed according to their topographic position index. Valley sites (−1 to −0.34) are shown as filled downward-pointing triangles; mid-slope sites (−0.33 to 0.33) are shown as crosses; and ridge sites (0.34 to 1) are shown as open triangles.

radiation. Stepwise multiple-regression of the DCA axes resulted in the following models:

$$\text{DCA1} = 2868.36 - 122.69 \text{ TOPO} - 2.77 \text{ SLP} - 46.16 \text{ VDR}$$

$$\text{DCA2} = 37364.92 - 186.5754 \text{ WLS} - 0.0135 \text{ TMY}$$

$$\text{DCA3} = 132.6485 + 0.2252 \text{ SOUTH} - 1.0539 \text{ SLP}$$

where TOPO is topographic position, SLP is slope, VDR is variance in annual direct solar radiation, WLS is whole light sky-space, TMY is location in the north-south direction, and SOUTH is southness. The variables in the models are listed in the order of the stepwise selection process. The coefficients were all statistically significant at $\alpha=0.05$, and the models explained 69%, 29%, and 16% of the variance in DCA axis 1, 2, and 3 scores, respectively.

Table 3. Results from gradient analysis.

	DCA		
	Axis 1	Axis 2	Axis 3
Eigenvalue	0.702	0.247	0.181
Pearson correlations with selected variables*			
Elevation	−0.591	−0.435	0.011
Slope	−0.230	0.269	−0.242
Distance to river	−0.483	−0.240	−0.047
Topographic position	−0.768	−0.402	0.040
Southness	−0.021	−0.240	0.273
Westness	0.158	0.007	0.186
Yellow soil	−0.233	−0.074	−0.163
Colluvial soil	−0.566	−0.233	−0.126
Lithosols	0.623	0.264	0.098
Total solar radiation	−0.580	−0.440	0.041
Variance in total solar radiation	−0.256	−0.089	−0.215

* Correlation values >0.5 are highlighted in bold-type.

Vegetation classification and community analysis

The first level division from the TWINSpan results (Fig. 3) separates the valley sites from the upland sites. In the second level division, the upland sites are divided into ridge sites and mid-slope sites (division 2), and in the valley group, site 31 is singled out based on the presence of *Callicarpa formosana* (division 3). The third level division further separates the level two groups into subgroups that cannot be clearly identified. We therefore chose to use three groups that corresponded to topography, with T00 corresponding to ridge-tops, T01 midslope, and T1 to valleys.

Of the three vegetation types (Table 4), the valley community (T1) is the most distinct, and is characterized by the dominance of mesic species (e.g. *Machilus japonica* var. *kusanoi* and *Phoebe formosana*), the relative abundance of gap species (*Lagerstroemia subcostata* and *Mallotus japonicus*), and the absence of xeric species (e.g. *Meliosma squamulata* and *Cyclobalanopsis longinux*). The ridge (T00) and mid-slope (T01) communities are quite similar in their species composition, both having *Castanopsis cuspidata* var. *carlesii* and *Machilus thunbergii* as their most dominant species. The ridge-type community, however, is characterized by the presence of xeric species *Myrsine sequinii* and *Ilex pubescens*, and the absence of mesic species *Oreocnide pedunculata* and *Machilus japonica* var. *kusanoi*, and gap species *Lagerstroemia subcostata*. The mid-slope

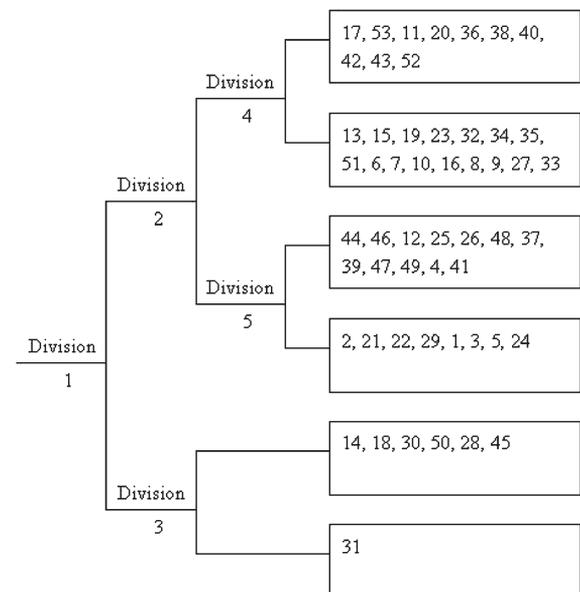


Fig. 3. Dendrogram from TWINSpan results. The numbers shown in the boxes are site numbers. Indicator species for the first division are *Machilus japonica* var. *kusanoi* 1(+) and *Oreocnide pedunculata* 2(+), for the second division, *Myrsine sequinii* 1(−) and *Elaeocarpus japonicus* 5(−), and for the third division, *Callicarpa formosana* 1(+).

Table 4. Analysis of vegetation in different community types. The number of unique species is the number of species found only in the specific community type, and the number of uniquely missing species is the number of species that are missing only in the specific community type. The most dominant species are listed in the order of dominance.

	All sites	Ridge-type (T00)	Mid-slope (T01)	Valley-type (T1)
Total number of species	92 (100%)	75 (81.5%)	64 (69.6%)	42 (45.7%)
Average number of species/plot	26 (28.3%)	31.5 (34.2%)	22 (23.9%)	18 (19.6%)
Number of species shared		28 (30.4%)		
with mid-slopes	–	53 (57.6%)	–	–
with valley-types	–	28 (30.4%)	36 (39.1%)	–
Number of unique species	–	22 (23.9%)	3* (3.3%)	6 (6.5%)
Number of uniquely-absent species	–	8 (8.7%)	0	25 (27.2%)
Most dominant species				
#1	<i>Castanopsis cuspidata</i> var. <i>carlesii</i>	<i>Castanopsis cuspidata</i> var. <i>carlesii</i>	<i>Castanopsis cuspidata</i> var. <i>carlesii</i>	<i>Machilus kusanoi</i>
#2	<i>Machilus thunbergii</i>	<i>Machilus thunbergii</i>	<i>Machilus thunbergii</i>	<i>Phoebe formosana</i>
#3	<i>Litsea acuminata</i>	<i>Meliosma squamulata</i>	<i>Litsea acuminata</i>	<i>Machilus zuihoensis</i>

* All species unique to the mid-slope community type occurred in only one site.

community, on the other hand, had few uniquely present species and no uniquely absent species, and can only be characterized by the relative abundance of *Pasania hancei* var. *ternaticupula*, a species rarely seen on xeric ridges. The three community types were not mutually exclusive, with the eight most widespread species present in all three types. The overall species distribution across the landscape can therefore be characterized as relatively even within communities and continuous across communities.

Vegetation pattern modeling

Using the DCA prediction models from the stepwise-regression and the layers of maps containing the predictor variables, we derived predicted species gradient maps containing the predicted values of DCA major axis scores across the landscape (Figs 4b–d). In the map of predicted DCA axis 1 scores (Fig. 4b), the vegetation gradient closely followed the topography of the terrain, with sites having high DCA axis 1 scores following the valleys of the two river systems, sites having low DCA axis 1 scores following the ridges, and the slopes bearing the intermediate gradients. In the map of predicted DCA axis 2 scores (Fig. 4c), the vegetation gradient reflected the openness of the surrounding terrain, with low DCA axis 2 scores concentrated in the open and high-elevation areas towards the north closer to Mount Ah-Yuh and along ridge-tops, and the high DCA axis 2 scores concentrated in the lower-elevation and steeper Tsukeng creek to the southeast. The predicted DCA axis 3 values (Fig. 4d) reflected the aspect, especially the southness of the aspect, with low DCA axis 3 scores concentrated on the north-facing slopes along the foothills of Mount Hungtzu, south of the forest.

The 49 ha quadrat with maximum potential beta-diversity is shown by the brown quadrangle lines in Figs 4a–d. Because it was designed to capture the 49 ha area with maximum ranges in the first three DCA axes, it

is located in an area with the maximum change in terrain at the scale of 700 m. A different search criterion, for example, using a 4 ha quadrat instead of 49 ha, results in a maximum potential beta-diversity area (black lines, Figs 4a–d) in a totally different area.

The prediction of plant communities using discriminant analysis also resulted in a set of functions using topographic position, slope, and variance in annual direct solar radiance as predictor variables (Table 5), with the order of variable-selection identical to the DCA axis 1 prediction model. The classification model had 81% accuracy, and the jackknife test resulted in a similar accuracy at 79%. Using the layers of maps containing the predictor variables, we derived a predicted plant community map over the landscape (Fig. 5). In this map, the vegetation also closely followed the topography of the terrain, with type T1 following the valleys of the two river systems, type T00 following the ridges, and type T01 as a transition between types T00 and T1. An analysis of the misclassified sites using classification trees suggested that the misclassified sites were mostly on south to west facing slopes in the northern half of the study area.

Analysis of alpha-diversity

A total of 92 tree species were found in the survey (Table 4). The average sample plot had 26 species, but the most species-rich plot had 46 species while the least species-rich plot had only 12 species. Ridge-type communities had the highest alpha-diversity, whereas valley-type communities had the lowest (Table 4). The ridge-community sites also had the largest number of unique species (23.9%). Most environmental variables showed a significant relationship (at significance level = 0.05) with the number of species, but the variables with the strongest R-squared were topographic position, whole light sky-space, and elevation. The prediction of the number of species per sample plot from environmental

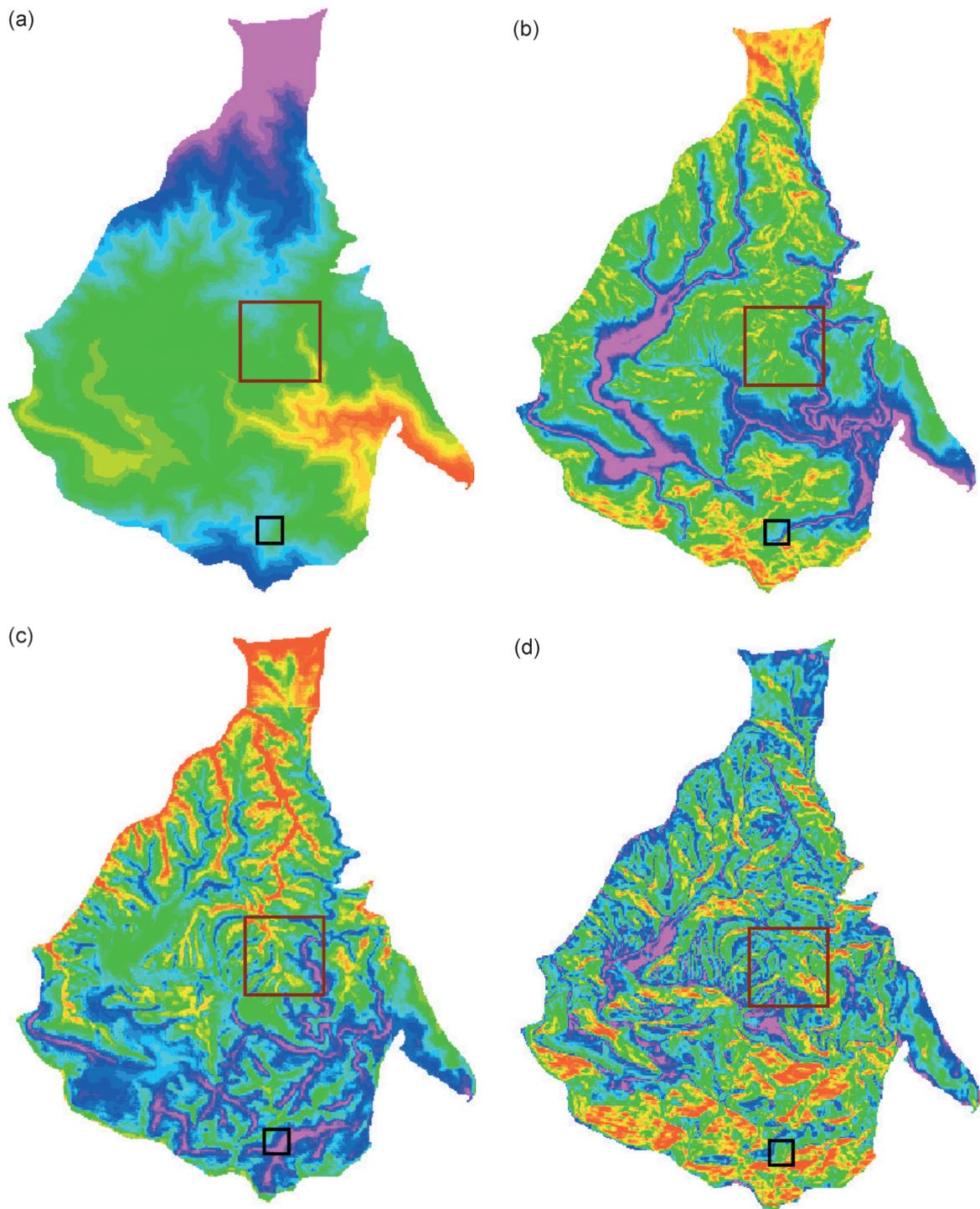


Fig. 4. Maps of terrain, predicted DCA scores, and location of quadrats with highest beta diversity: a) location of the 49 ha and 4 ha quadrats with highest beta diversity mapped onto a digital elevation model (DEM); b) predicted DCA1 scores; c) predicted DCA2 scores; d) predicted DCA3 scores. In all four maps, elevation (m) or DCA scores are coloured using a purple to red gradient, with purple indicating high values, and red indicating low values. The locations of the 49 ha quadrats with highest beta diversity are marked with brown lines, and the 4 ha quadrats with black lines.

Table 5. Results from discriminant analysis of three vegetation classes.

Misclassification rate	19%	
Jackknifed misclassification rate	21%	
Wilks' lambda	0.1790	
P value	<0.0001	
Canonical discriminant functions		
Eigenvalues	4.151	0.085
Factor scores (variables in order of selection):		
Topographic position	1.050	0.003
Slope	0.715	0.836
Variance in annual direct solar radiance	0.500	-0.366

variables using a generalized linear model with a Poisson error distribution and a log-link function resulted in the following model:

$$\text{NUMSP} = 225.82 + 0.2781244 \text{ TOPO} - 0.00009 \text{ TMY} + 0.0001 \text{ TMX}$$

where NUMSP is the number of species per plot, TOPO is topographic position, and TMX and TMY are x-y coordinates. This function suggests that addition to higher alpha-diversity towards the ridges, alpha-diversity also increases towards the southeast.

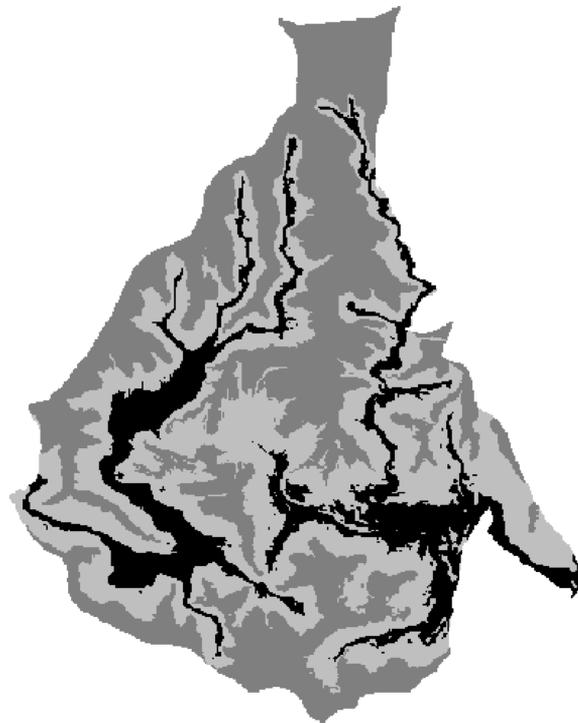


Fig. 5. Predicted vegetation community map using classification functions from discriminant analysis. Areas in medium-shade are ridge communities (T00), in black, valley communities (T1), and in light-shades, mid-slope communities (T01).

Discussion

Gradients in vegetation composition in the Fushan area

In mountainous terrain, vegetation is often closely associated with topography. In this study, both gradient analysis and classification methods found topographic position to be the most important factor in determining species composition for the Fushan area. An independent study with 140 sample plots across a 1.4 ha permanent plot that spanned an elevational range of 690–820 m (Wang et al. 2000) also found that the trees of Fushan forest can be separated into ridge-type, mid-slope, and valley-type communities despite the slightly lesser scope of the study. Topographic position may reflect the moisture gradient or wind exposure associated with topography, or in this study area to soil types, which happened to be strongly associated with topography. Other important gradients found include the total amount of solar radiation input to the site or other factors associated with elevation, such as the temperature gradient (DCA2), and the variance or annual fluctuation in the amount of solar radiation input to the site (DCA3). Even though principal components analysis (PCA) of the environmental variables showed that soil parent material was an important factor in the environmental gradients, it was not important in explaining the gradients in vegetation composition.

Limitations of this study

Most landscape-scale studies are limited by the scope and scale of their data, and the results of this study are no exception. For example, our results are limited by the resolution of the grid cells. Heterogeneity finer than the 12.5 m grid scale cannot be captured in the modeling process, and the effects of sharp changes in topography that span only a few grid cells may also be smoothed out. Because the initial vegetation sampling efforts were designed for an objective different from the current study, only the center of the plots were mapped, and therefore the location of all pixels falling within the plots could not be determined. For the same reasons, the sample plots were not randomly selected prior to the onset of the survey. As a result, the locations were biased towards the middle 37% of the full elevational zone and slightly towards the south-west of the full experimental forest area, whereas the steeper areas and the more open areas higher on the ridges were not sufficiently represented in the vegetation samples. Therefore, caution must be exercised in the extrapolation towards these less represented areas, or when the scale of interest diverges from that of the study.

Alpha versus beta diversity

When only the alpha-biodiversity of a local site is considered, the results of our analysis show that the highest alpha-diversity sites are on ridge positions. Not only did the ridge sites support 81.5% of the total tree species, they also had 22 species (24%) that were uniquely found in such community types. Results from the generalized linear model also indicated an increased alpha-diversity towards the southeast. This suggests that there may be other spatial processes involved in maintaining biodiversity, perhaps through the spatial dispersal of seeds or habitat-related patchiness across the landscape (Condit et al. 2000).

When the species composition among the three vegetation associations were analyzed, however, the results show that if only the most species-rich ridge-community sites were preserved, the remaining 18.5% of the species not found in such communities will not be included. Also, in topographically complex terrain such as the Fushan forest, a selection criterion focusing on specific community types will result in areas with very complex shapes and high edge-to-area ratios (Fig. 5). In other words, the identification of biodiversity hotspots often used in regional analysis (e.g. Iverson and Prasad 1998, Kress et al. 1998) may not be suitable for the conservation of biodiversity at the scale of communities. At this scale, a conservation strategy that targets high beta diversity can and should be considered. By mapping out predicted values of major DCA axes scores through the combination of vegetation modeling and gradient analysis, geographic information systems can easily identify regions with the widest range and highest variability in their major DCA axes scores. Through this process, regions with potentially the highest beta diversity can then be identified (Figs 4a–d). This method is also preferred when the objective is to preserve the factors and processes that lead to biodiversity. Note should be taken, however, that different search criteria (different number of DCA axes used, maximum range of DCA versus maximum variance in DCA, or the size of quadrat used) may result in maximum beta-diversity quadrats in totally different areas (e.g. black lines vs brown lines, Figs 4a–d), and therefore these criteria should be predetermined based on management needs.

Importance of gradients in conserving biodiversity

The notions of communities and gradients in ecology have in the past two decades declined in importance whereas broad scale issues such as land use and environmental change have increased. Part of this change can be contributed to findings from palaeoecological studies that have recognized the temporary nature of the concept “community”. As opposed to being a permanent entity, observations of past communities that

no longer exist have suggested that communities are actually temporary assemblages of several species whose habitat ranges concur with the contemporaneous environmental conditions of a specific location. Such a community may cease to exist as the environment changes to a different set of conditions, with each species moving in a different direction (Davis 1983). Therefore, under changing environmental conditions, an objective that aims at protecting specific communities will eventually become futile as these “communities” disintegrate.

A more recent revivification of the gradient concept (Gosz 1992) broadened the application of gradient analysis beyond communities to ecosystems, and turned its focus onto the gradients themselves, with its significance as the physical environment that constrains biological interactions. This paper also concludes that modeling efforts that include spatial gradients provide the foundation for management programs from a local to a regional scale. We concur with this view in arguing that the most important contribution of gradient analysis is its role as a bridge between the physical environment and the biota in identifying the most important gradients in the physical environment that contribute to the diversity of species assemblages, and also different forms of life. Results from this study show that important dimensions in the physical environment (major PCA axes) may not be sensed as equally important by plant species (not shown in major DCA axes.) Gradient analysis, therefore, can be used to help identify and thereby help preserve the important processes and factors that contribute to biological diversity.

Conclusions

Biomes are dynamic. As time passes and environments change, new species assemblages may form and old ones disappear. Over a longer time span, new species may evolve and old species become extinct. Conservation objectives that target specific entities such as species or ecosystems may be suitable for a relatively short term, but will eventually become futile as these entities themselves cease to exist. We therefore suggest that it is equally important, if not more so, to set objectives that preserve the processes and factors that contribute to biodiversity. Such processes and factors are often multitude in scale, and at each level, the contributors to biodiversity can be very different and may require different study methods. For example, a permanent plot design with high-intensity survey is definitely the ideal approach for a community-level study, but when a larger landscape is the scope of interest, a nested design with smaller clustered plots stratified in larger regular spacing (Urban et al. 2002) may be a more suitable alternative method for analyzing vegetation cross-scales. Similarly, because the various management tools are

each best suited for a specific scale, management strategies and tools should also be adaptive to the intended scale. For example, remote sensing techniques have a resolution best suited for the detection and monitoring of ecosystems at a regional scale, but may not be useful for the detection of plant communities. Therefore, for the management of communities, different techniques will have to be used. The results of our study suggest that, combined with a sampling scheme that captures the variation in vegetation across the landscape, vegetation modeling that incorporates information from multivariate analysis and geographic information systems can be a useful method for the description and management of biodiversity at the level of plant communities. Vegetation modeling with clustering methods can be used to allocate sites with higher potential alpha diversity, with the benefit of having a list of species associated with the clustered type. Vegetation modeling with gradient analysis, on the other hand, can be used to identify regions with the highest potential beta diversity by means of selecting areas with the widest range or highest variability in several major DCA axes scores, and thereby can help to preserve the largest range of environmental conditions that lead to diversity in species assemblages. We hope that fellow researchers will join our efforts in developing new methods to study and manage biodiversity at all scales.

Acknowledgements – The authors thank Yu-Joe Hsia, Chang-Fu Hsieh, and Hsiang-Hwa Wang for helpful comments on an earlier manuscript, and Fu-Wen Horng for providing the soils map used in this study. This study was funded in part by the National Science Council, Republic of China (NSC88-2621-B-002-022-A10, NSC89-2621-B-002-016, NSC89-2621-B-002-031-A10).

References

- Ahearn, S. C., Smith, J. L. D. and Wee, C. 1990. Framework for a geographically referenced conservation database: case study Nepal. – *Photogramm. Eng. Rem. Sens.* 56: 1477–1481.
- Alexander, M. 1995. Management planning in relation to protected areas. – *Parks* 5: 2–11.
- Anon. 1994a. Guidelines for country studies on biological diversity. – United Nations Environment Programme, Nairobi, Kenya.
- Anon. 1994b. SAS. – The SAS Inst., Cary, NC.
- Anon. 1997. ARC/INFO 7.2. – Environmental Systems Research Institute (ESRI), Redlands, CA.
- Anon. 2001. S-Plus 6 for Windows Guide to Statistics. Vol. 2. – Insightful Corporation, Seattle, WA.
- Apan, A. A. 1997. Land cover mapping for tropical forest rehabilitation planning using remotely-sensed data. – *Int. J. Rem. Sens.* 18: 1029–1049.
- Asner, G. P., Wessman, C. A. and Schimel, D. S. 1998. Heterogeneity of savanna canopy structure and function from imaging spectrometry and inverse modeling. – *Ecol. Appl.* 8: 1022–1036.
- Austin, M. P. 1998. An ecological perspective on biodiversity investigations: examples from Australian eucalypt forests. – *Ann. Mo. Bot. Gard.* 85: 2–17.
- Bai, M.-L., Lee, P.-F. and Chang, C.-R. 2000. Landscape classification of Fushan experimental forest in Taiwan. – *Quart. J. Chinese For.* 33: 291–312, in Chinese with English summary.
- Bonan, G. B. 1989. A computer model of the solar radiation, soil moisture, and thermal regimes in boreal forests. – *Ecol. Modell.* 45: 275–306.
- Chang, F.-H. et al. 1986. The survey on the natural resources of Ha-pen area. – *Ecol. Res. Rep. No. 13*, Council of Agriculture, Taiwan, in Chinese.
- Chang, Y.-H. 1998. The study of the relationship between the distribution of pteridophytes and environmental factors at Fushan area in north Taiwan, and estimating the utility of ferns as the microenvironmental indicators. – M.s. thesis, National Taiwan Univ., in Chinese.
- Condit, R. et al. 2000. Spatial patterns in the distribution of tropical tree species. – *Science* 288: 1414–1418.
- Davis, F. W. and Goetz, S. 1990. Modeling vegetation pattern using digital terrain data. – *Landscape Ecol.* 4: 69–80.
- Davis, M. B. 1983. Quaternary history of deciduous forests of eastern North America and Europe. – *Ann. Mo. Bot. Gard.* 70: 550–563.
- Deangelis, D. L. et al. 1998. Landscape modeling for Everglades ecosystem restoration. – *Ecosystems* 1: 64–75.
- Gaston, K. J. 1994. Spatial patterns of species description: how is our knowledge of the global insect fauna growing? – *Biol. Conserv.* 67: 37–40.
- Gosz, J. R. 1992. Gradient analysis of ecological change in time and space: implications for forest management. – *Ecol. Appl.* 2: 248–261.
- Harmon, M. E. and Marks, B. 1995. Programs to estimate the solar radiation for ecosystem models. – H. J. Andrews LTER electronic publication, <<http://www.fsl.orst.edu/lter/data/software/solarrad/solarrad.txt>>.
- Heller, R. C. and Ulliman, J. J. 1983. Forest resource assessments. – In: Estes, J. E. (ed.), *Manual of remote sensing*. Vol. 2: interpretation and applications, 2nd ed. American Society of Photogrammetry, pp. 2229–2324.
- Hill, M. O. and Gauch, H. G. 1980. Detrended correspondence analysis, an improved ordination technique. – *Vegetatio* 42: 47–58.
- Hsia, Y.-J. and Wang, W.-S. 1985. Calculation of potential solar irradiance on slopes. – Research Note No. 001, Taiwan Forestry Research Inst., in Chinese with English summary.
- Hubbell, S. P. et al. 1999. Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. – *Science* 283: 554–557.
- Iverson, L. R. and Prasad, A. 1998. Estimating regional plant biodiversity with GIS modelling. – *Div. Distribut.* 4: 49–61.
- Jarvis, A. M. and Robertson, A. 1999. Predicting population sizes and priority conservation areas for 10 endemic Namibian bird species. – *Biol. Conserv.* 88: 121–131.
- Kress, W. J. et al. 1998. Amazonian biodiversity: assessing conservation priorities with taxonomic data. – *Biodiv. Conserv.* 7: 1577–1587.
- Kuo, C.-M. 1994. Long-term ecological research in Fushan forest – understory vegetation. – In: Peng, C.-I. and Chou, C.-H. (eds), *Biodiversity and terrestrial ecosystems*. Inst. of Botany, Academia Sinica Monograph Ser. No. 14, pp. 383–399.
- Lin, K.-C. et al. 1996. Soil survey and classification of the Fushan Experimental Forest. – *Taiwan J. For. Sci.* 11: 159–174, in Chinese with English summary.
- Lin, T.-T. 1997. Vegetation of Fushan forest. – In: Chou, C.-H. (ed.), *Introduction of Taiwan long-term ecological research, a special issue of Taiwan LTER Newsletter*, No. 1. National Science Council, Republic of China, pp. 15–16.
- Mabery, C. M. et al. 1998. Typhoon disturbance and stand-level damage patterns at a subtropical forest in Taiwan. – *Biotropica* 30: 238–250.
- Margules, C. R. 1989. Introduction to some Australian developments in conservation evaluation. – *Biol. Conserv.* 50: 1–11.

- McCune, B. and Mefford, M. J. 1999. PC-ORD – Multivariate analysis of ecological data, ver. 4. – MJM software design, Gleneden Beach, OR.
- Nikolov, N. T. and Zeller, K. F. 1992. A solar radiation algorithm for ecosystem dynamic models. – *Ecol. Modell.* 61: 149–168.
- Ostendorf, B. and Reynolds, J. F. 1998. A model of arctic tundra vegetation derived from topographic gradients. – *Landscape Ecol.* 13: 187–201.
- Plumb, G. A. 1991. Assessing vegetation types of Big Bend National Park Texas USA for image-based mapping. – *Vegetatio* 94: 115–124.
- Roberts, D. A., Green, R. O. and Adams, J. B. 1997. Temporal and spatial patterns in vegetation and atmospheric properties from AVIRIS. – *Rem. Sens. Environ.* 62: 223–240.
- Scott, J. M. et al. 1993. GAP Analysis: a geographical approach to protection of biological diversity. – *Wildl. Monogr.* 123: 1–41.
- Smith, A. P., Horning, N. and Moore, D. 1997. Regional biodiversity planning and lemur conservation with GIS in western Madagascar. – *Conserv. Biol.* 11: 498–512.
- Tappeiner, U., Tasser, E. and Tappeiner, G. 1998. Modelling vegetation patterns using natural and anthropogenic influence factors: preliminary experience with a GIS based model applied to an Alpine area. – *Ecol. Modell.* 113: 225–237.
- Urban, D. et al. 2002. Extending community ecology to landscapes. – *Ecoscience* 9: 200–202.
- Wang, H.-H. et al. 2000. Vegetation classification and ordination of a permanent plot in the Fushan experimental forest, northern Taiwan. – *Taiwan J. For. Sci.* 15: 411–428, in Chinese with English summary.
- Westman, W. E., Strong, L. L. and Wilcox, B. A. 1989. Tropical deforestation and species endangerment: the role of remote sensing. – *Landscape Ecol.* 3: 97–110.