Partitioning beta diversity in a subtropical broad-leaved forest of China

Pierre Legendre,[1,6] Xiangcheng Mi,[2] Haibao Ren,[2] Keping Ma,[2] Mingjian Yu,[3] I-Fang Sun,[4] and Fangliang He[5]

[1]*Département de Sciences Biologiques, Université de Montréal, C.P. 6128, Succursale Centre-ville, Montréal, Québec H3C 3J7 Canada*
[2]*Institute of Botany, Chinese Academy of Sciences, Beijing*
[3]*College of Life Sciences, Zhejiang University, Hangzhou*
[4]*Center for Tropical Ecology and Biodiversity, Tunghai University, Taichung*
[5]*Department of Renewable Resources, 751 General Services Building, University of Alberta, Edmonton, Alberta T6G 2H1 Canada*

Abstract. The classical environmental control model assumes that species distribution is determined by the spatial variation of underlying habitat conditions. This niche-based model has recently been challenged by the neutral theory of biodiversity which assumes that ecological drift is a key process regulating species coexistence. Understanding the mechanisms that maintain biodiversity in communities critically depends on our ability to decompose the variation of diversity into the contributions of different processes affecting it. Here we investigated the effects of pure habitat, pure spatial, and spatially structured habitat processes on the distributions of species richness and species composition in a recently established 24-ha stem-mapping plot in the subtropical evergreen broad-leaved forest of Gutianshan National Nature Reserve in East China. We used the new spatial analysis method of principal coordinates of neighbor matrices (PCNM) to disentangle the contributions of these processes. The results showed that (1) habitat and space jointly explained ~53% of the variation in richness and ~65% of the variation in species composition, depending on the scale (sampling unit size); (2) tree diversity (richness and composition) in the Gutianshan forest was dominantly controlled by spatially structured habitat (24%) and habitat-independent spatial component (29%); the spatially independent habitat contributed a negligible effect (6%); (3) distributions of richness and species composition were strongly affected by altitude and terrain convexity, while the effects of slope and aspect were weak; (4) the spatial distribution of diversity in the forest was dominated by broad-scaled spatial variation; (5) environmental control on the one hand and unexplained spatial variation on the other (unmeasured environmental variables and neutral processes) corresponded to spatial structures with different scales in the Gutianshan forest plot; and (6) five habitat types were recognized; a few species were statistically significant indicators of three of these habitats, whereas two habitats had no significant indicator species. The results suggest that the diversity of the forest is equally governed by environmental control (30%) and neutral processes (29%). In the fine-scale analysis (10 × 10 m cells), neutral processes dominated (43%) over environmental control (20%).

Key words: beta diversity; Chinese Forest Biodiversity Monitoring Network; Gutianshan National Nature Reserve; PCNM analysis; spatial scale; stem-mapped forest plot; variation partitioning.

INTRODUCTION

Large permanent forest plots with precise stem maps have proven invaluable for understanding the coexistence of species, studying diversity patterns, testing ecological theories, monitoring the dynamics of stand structure and function, and conserving and managing biodiversity (Condit 1995, Condit et al. 1998, 2006, Hubbell 2001, He and Legendre 2002, Ibáñez et al. 2003, Losos and Leigh 2004). The best-known mapped plots are the forest dynamism and diversity plots of the Center for Tropical Forest Science (CTFS), which are distributed from tropical Africa to Asia to Central America. Spatially referenced tree demographic data from these plots have profoundly advanced our understanding of the structure, composition, diversity, and dynamics of tropical forests (Losos and Leigh 2004). However, the geographical bias towards tropical forests has raised an important question: are the findings from the tropics also applicable to forests in other regions? No answer can be given to this question at the moment, as similar large-scale forest plots have not yet been established in non-tropical areas (or, when established, data collection is still in progress), while results from other sources are most often not comparable.

The 24-ha permanent plot in Gutianshan Forest Reserve of subtropical China was established, in part, to answer the above question. The Gutianshan plot (*shan* meaning *mountain*) is part of the Chinese Forest Biodiversity Monitoring Network, which consists of five large (20 to 25 ha in size) stem-mapped plots in China along a latitudinal gradient from temperate, subtropical to tropical forests. The primary scientific goals of this network are to (1) collect long-term diversity and dynamics data for understanding and synthesizing spatial and temporal macroecological patterns of tree species along the latitudinal gradient; (2) test biodiversity theories about mechanisms (e.g., density dependence, competition, niche differentiation, fluctuating recruitment or storage effects, dispersal limitation, etc.) that are considered important for the promotion of biodiversity; (3) monitor the change of forest communities in the face of climate change and human disturbances through repeated censuses on the permanent plots; (4) provide key information on composition, growth, and dynamics of forest stands, which provide a basis for sustainable forest management; (5) describe the biodiversity in natural old-growth Chinese forests to serve as a benchmark for the assessment of exploited secondary forests; and (6) provide data collected in standardized form for the comparison of permanent forest plots throughout the world. The latitude-based Chinese network is complementary to the longitudinal CTFS network, forming a truly global broad-scale forest biodiversity monitoring network. In this network, the Gutianshan plot represents the evergreen broad-leaved forests which are typical of the middle subtropics of China (Wu 1980).

This study will focus on the spatial structures found in the trees of the 24-ha Gutianshan plot, or Gutian plot for short. Spatial structures are of paramount importance in community studies because their presence indicates that some process has been at work to create them. Basically, two families of mechanisms can generate spatial structures (Legendre and Legendre 1998, Fortin and Dale 2005). First, variation in environmental conditions may be responsible for the spatial structures found in species assemblages through species-habitat associations. This is an application of the classical environmental control model (Whittaker 1956, Bray and Curtis 1957, Hutchinson 1957). If the environmental variables are spatially structured, their structure will be reflected in the species distributions through induced spatial dependence. Second, spatial structures may also be generated by the species assemblages themselves, and in particular by dispersal limitation which can produce aggregated patterns through the so-called neutral mechanisms, which assume individuals of every species to have the same set of demographic rates (Hubbell 2001, Borda de Agua et al. 2007), leading to spatial autocorrelation in the species data. Partitioning the variation in community structure among sampling units between environmental and

spatial components according to this framework provides a useful ground for testing and separating niche from neutral mechanisms in biodiversity studies (Harms et al. 2001, Tilman 2004, Karst et al. 2005, Gravel et al. 2006, Laliberté et al. 2008).

We will limit the study of environmental control to the effect of topographic variables on the spatial distribution of species richness and tree species. Species distributions are often seen to correlate with different topographical features (Whittaker 1956, Harms et al. 2001). Unlike soil properties, topography is not a direct environmental variable but an indirect, or proxy variable that comprehensively characterizes the overall quality of a habitat. Part of the spatial structure that is related to environmental causes may be reflected by topographic variables, but not necessarily all of it. For example, some environmental variables such as soil pH may be acting at the scale of meters, not at the scale of topographic variation. So, when partitioning species variation between spatial and topographic explanatory variables, some uncertainty will remain as to the interpretation of the spatially structured variation that is not explained by the topographic variables. Nevertheless, a large spatial effect that is not explained by topographic variables may be indicative of the operation of other factors such as neutral mechanisms.

The objectives of this paper are to (1) describe the biodiversity found in the natural evergreen subtropical forest of the 24-ha Gutian plot and (2) test hypotheses about the processes (environmental control and neutral) that may be responsible for the beta diversity observed in the plot, by partitioning the effects of topography and space on the distribution of species at different spatial scales (size of the sampling units, or cells). This study will contribute to understanding the spatial organization of the tree biodiversity at multiple scales in the Gutian plot and the roles that habitat heterogeneity and unmeasured spatial processes, including dispersal limitation, play in shaping that tree community. Ultimately, we will show that niche and neutral processes do not have to diametrically oppose each other; they actually worked together side by side to regulate beta diversity in our study area.

Materials and Methods

The nature reserve.—The Gutianshan National Nature Reserve, approximately 81 km^2 in area, is located in Kaihua County, at the extreme west of Zhejiang Province, East China (29°10′19″–29°17′41″ N, 118°03′50″–118°11′12″ E). The reserve was set up in 1975, in the Yangtse River basin, to preserve a portion of the old-growth evergreen broad-leaved forest in the region. About 57% of the reserve is natural forest. 1426 species of seed plants belonging to 648 genera and 149 families have been inventoried in the reserve. Eighteen species are found in the Chinese list of rare and endangered species. Annual mean temperature in the region is 15.3°C; annual mean precipitation, calculated

from data from 1958 to 1986, is 1964 mm (Yu et al. 2001). Most of the precipitation occurs between March and September. The vegetation is representative of the typical subtropical evergreen broad-leaved forest (Yu et al. 2001). *Castanopsis eyrei* (Champ. ex Benth.) Tutch. (Fagaceae) and *Schima superba* Gardn. et Champ. (Theaceae), which are broadly distributed in subtropical China, are the dominant species in evergreen broad-leaved forests and in this plot.

Location and description of the study plot.—The 24-ha forest plot under study (29°15′6″–29°15′21″ N, 118°07′1″–118°07′24″ E) forms a rectangle of 600 × 400 m. The smaller (400-m) side of the plot is oriented at about 5° west of true north. (The orientation of the plot was calculated from the latitude and longitude coordinates of the four corners of the plot, obtained from GPS.) Based on dendrochronological evidence (Wang Xiaochun and Zhang Qibin, *unpublished data*) and consultation with inhabitants of the region, the Gutian plot contains secondary forest about 160–180 years old that was heavily disturbed by agriculture and charcoal production about 80 years ago. At the present time, most of the forest is in the middle and late successional stages.

All trees with diameter at breast height (dbh) $\geq$ 1 cm were tagged, identified, measured, and georeferenced during the summer of 2005. The plot was very rugged: altitude varied from 446.3 to 714.9 m above sea level whereas the 20-m cell slopes varied from 13° to 62°. It took nine months for a field team comprising 20 scientists, graduate students, technicians, and workmen to map and collect the data about 140 676 individual trees belonging to the 49 families and 159 species identified in the plot. The species are listed in Appendix A. The nomenclature follows Zheng (2005). For the present study, the trees were grouped into cells 10 × 10, 20 × 20, 40 × 40, and 50 × 50 m in size, which allow a division of the whole plot into cells of equal sizes. There were only 24 (100 × 100 m) cells, too few for statistical analysis. The cell sizes were used to study the change of beta diversity with scale (grain size) in the 24-ha Gutian plot. This study mainly focuses on the results of the 20 × 20 m cell analysis.

Statistical analyses.—The study of the spatial distribution of species richness and community composition includes all 159 tree species found in the plot. Our interest is to model the variation of richness and community composition in terms of topography and the spatial structure represented by principal coordinates of neighbor matrices (PCNM) eigenfunctions.

Species richness within 20 × 20 m cells was mapped, and then its spatial variation was analyzed by variation partitioning (Borcard et al. 1992, Borcard and Legendre 1994, Legendre and Legendre 1998, Peres-Neto et al. 2006, Legendre 2007) with respect to topographic and spatial variables. Four topographic attributes were measured in the field: altitude (the difference in altitude between the highest and lowest cells was 253 m), terrain

convexity (with values from −16.6 to 18.6 m), slope (with values from 12.8° to 62.0°), and aspect (with values from 93.9° to 269.2°) for the 20 × 20 m cells (Fig. 1). Following Harms et al. (2001) and Valencia et al. (2004), elevation of a cell was defined as the mean of the elevation values at its four corners. Convexity was the elevation of the cell of interest minus the mean elevation of the eight surrounding cells. For the edge cells, convexity was the elevation of the centre point minus the mean of the four corners. Slope was the mean angular deviation from horizontal of each of the four triangular planes formed by connecting three of its corners. Aspect refers to the direction to which a slope faces. Altitude, convexity, and slope were used to construct third-degree polynomial equations, for a total of nine monomials; the monomials with exponents allow the modeling of nonlinear relationships between the topographic predictors and the response variables (richness or species composition). Aspect is a circular variable; sin(aspect) and cos(aspect) were computed in order to use aspect in linear models. This resulted in 11 variables in total in the expanded topographic data table.

For the 20 × 20 m cells, PCNM eigenfunctions (Borcard and Legendre 2002, Borcard et al. 2004, Dray et al. 2006) were computed across the 600 points of the spatial grid. PCNM eigenfunctions represent a spectral decomposition of the spatial relationships among the grid cells; they describe all spatial scales that can be accommodated in the sampling design. They are obtained by principal coordinate analysis (PCoA) of a truncated geographic distance matrix among the sampling sites, as explained in the above-mentioned papers. In the present study, all distances larger than the distance between the centers of diagonally adjacent cells were replaced by four times that value before PCoA; 339 PCNM eigenfunctions with positive eigenvalues were generated. The PCNMs were then used as explanatory variables to analyze the spatial variation of the tree community composition data. The Supplement presents R language scripts to compute the PCNM eigenfunctions for all resolutions (i.e., cell sizes) used in this paper and display them on maps. Forward selection (with permutation tests, at the 5% significance level, of the increase in R^2 at each step) was applied to the PCNM table in order to determine if the spatial structure was mostly broad-, middle-, or fine-scaled.

Beta diversity can be defined in many different ways (Koleff et al. 2003). For example, Whittaker (1960, 1972) described the well-known index $\beta = S/\bar{\alpha}$, where S is the number of species in the whole area of interest while $\bar{\alpha}$ is the mean number of species observed per cell. He also proposed to apply the same formula to Shannon diversity H' instead of species richness. The index proposed by Legendre et al. (2005) is used in this study. This index is the sum, over all species and over all sites, of the squared abundance deviations from the species means. This index is a direct measure of the variation in
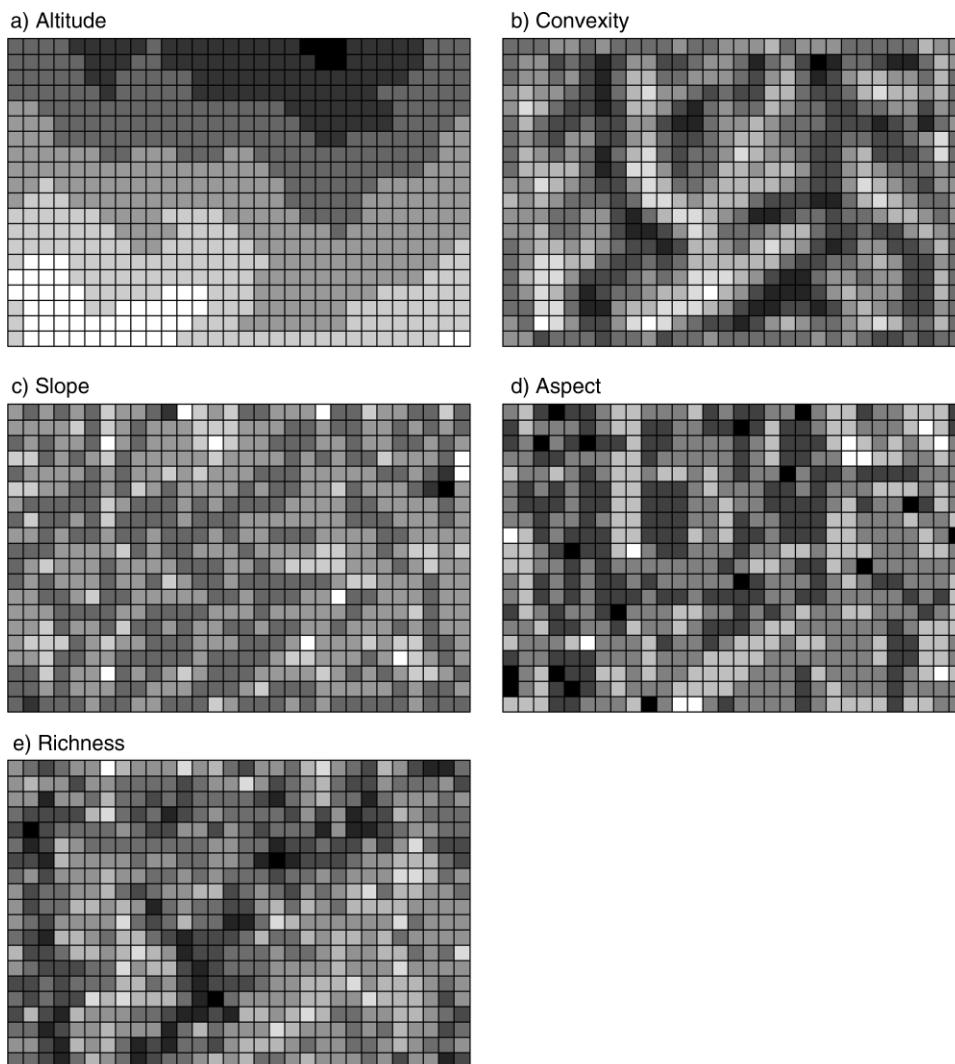
Fig. 1. Maps of the four topographic variables and species richness at the scale of 20 × 20 m. (a) Altitude from 11.8 m (white) to 264.6 m (black) above the lowest corner of the plot, or 458.1 m (white) to 710.9 m (black) above sea level. (b) Convexity from −16.6 m (white) to 18.6 m (black). There is an edge effect in the convexity map. (c) Slope from 12.8° (white) to 62.0° (black). (d) Aspect from 93.8° (white) to 269.2° (black). (e) Richness from 19 (white) to 54 (black) species.

species composition among the sites in the area of interest, which corresponds to the concept of beta diversity.

The variation of the community composition data was partitioned between the topographic and PCNM variables using canonical redundancy analysis (RDA; Rao 1964). A multivariate regression tree (MRT; De'ath 2002) was computed to delineate habitat types that were similar in topographic conditions and in species composition. Indicator species analysis (Dufrêne and Legendre 1997) was conducted to identify the species that were statistically significant indicators of these habitat types.

To examine the effect of the cell size on the partitioning of diversity variation, the above analysis at the 20 × 20 m scale was repeated for three other scales: 10 × 10, 40 × 40, and 50 × 50 m. Here we will only report the variation partitioning results across scales.

The canonical analyses, variation partitioning, and tests of significance of the fractions were computed using the "vegan" library (Oksanen et al. 2007) of the R statistical language (R Development Core Team 2007). The multivariate regression tree was computed using the "mvpart" library (De'ath 2006), whereas indicator species analysis was computed using the "labdsv" library (Roberts 2006). PCNM eigenfunctions were created using the package "spacemakeR" of Stéphane Dray (Laboratoire de Biométrie et Biologie Évolutive, UMR CNRS 5558, Université Lyon I, France) in R; forward selection was computed using the "packfor" package of the same author.

TABLE 1. Variation partitioning results for different cell sizes.

Cell size (m)	n	[a]	[b]	[c]	[d]	[a + b]	[b + c]
Species richness							
10 × 10	2400	0.0047	0.0651	0.3598	0.5705	0.0698	0.4249
20 × 20	600	0.0198	0.2380	0.3383	0.4039	0.2578	0.5763
40 × 40	150	0.1073	0.2548	0.2507	0.3872	0.3621	0.5055
50 × 50	96	0.0507	0.3233	0.1163	0.5098	0.3740	0.4396
Community composition							
10 × 10	2400	0.0034	0.1953	0.4296	0.3718	0.1987	0.6249
20 × 20	600	0.0295	0.2779	0.3483	0.3444	0.3074	0.6262
40 × 40	150	0.1185	0.2936	0.2487	0.3292	0.4121	0.5423
50 × 50	96	0.1308	0.2719	0.2602	0.3371	0.4027	0.5321

Notes: Fractions [a]–[d] (adjusted R^2 statistics, R_a^2): [a] = variation explained by the environmental variables and not spatially structured, [b] = variation explained by the environmental variables and spatially structured, [c] = spatially structured variation not explained by the environmental variables, [d] = residual variation. Fraction [b] is the intersection (not the interaction) of the amounts of variation explained by linear models of the two explanatory tables. Topographic variables used to compute fraction [a + b]: altitude, convexity, and slope were represented by third-degree polynomials; aspect was represented by sin(aspect) and cos(aspect). Principal coordinates of neighbor matrices (PCNM) eigenfunctions were the explanatory variables used to compute fraction [b + c].

RESULTS

Scale variation of beta diversity.—Table 1 presents the results of variation partitioning across scales (different cell sizes) for richness and species composition. The total unexplained variation ([d]) remains fairly constant across scales. This is equivalent to saying that the total proportion of explained variation ([a + b + c]) is invariant across scales. While this is true, the components [a], [b], and [c] vary with scales: the increase in one corresponds to a decrease in other components. As the size of the sampling units increases, the spatially structured variation [b + c] decreases while the topography-controlled variation [a + b] increases.

Spatial variation of species richness, 20 × 20 m cells.— Species richness (Fig. 1e) ranged from 19 to 53 in the 20 × 20 m cells. The forward selection procedure retained 66 PCNM eigenfunctions for modeling species richness (adjust R^2, $R_a^2 = 0.575$, which is nearly the same as the value for all 339 PCNMs without any selection, $R_a^2 = 0.576$). As shown in Fig. 2a, most of these are among the first 100 of the 339 PCNM functions. The first PCNMs represent broad-scale variation; see The Supplement.

The variation partitioning of richness is shown in Fig. 3a. The partitioning is based on adjusted R^2 statistic, R_a^2, as recommended by Peres-Neto et al. (2006); 57.6% of the variation of richness is spatially structured and explained by the PCNM eigenfunctions; 41.3% of that amount is also explained by the four topographic variables. The effect of topography is highly spatialized (92.3% of the topographic variation). Fig. 4 presents maps of the fitted values of richness, obtained by multiple regression, corresponding to the total explained variation [a + b + c], the variation explained by the three topographic variables [a + b], and the spatially structured fraction [c] unexplained by the presently available topographic variables.

Spatial variation of the community composition, 20 × 20 m cells.—The forward selection procedure retained 179 PCNM base functions for modeling community composition ($R_a^2 = 0.625$, which is nearly the same as the value for all 339 PCNMs without any selection, $R_a^2 = 0.626$). As shown in Fig. 2b, most of these are among the first 180 of the 339 PCNM eigenfunctions. These PCNMs represent broad- to middle-scale variation; see the Supplement.

The variation partitioning results are described in Fig. 3b. 62.6% of the variation (R_a^2) of the community composition data is spatially structured and explained by the PCNM eigenfunctions. Nearly half of that (44.4%) is also explained by the four topographic variables. Similar to the variation of richness, the effect of topography on species composition is highly spatialized (90.4% of the topographic variation). Fig. 5a–c presents maps of the fitted values of the three most significant canonical axes corresponding to the total explained variation [a + b + c]. Fig. 5d–f shows maps of the significant canonical axes corresponding to the variation explained by the three topographic variables, [a + b], as well as the spatially structured fraction [c] unexplained by the presently available topographic variables.

An interesting property of PCNM eigenfunctions is that their variances correspond to their spatial scales. Since the principal coordinate analysis orders them by decreasing variances, they are also ordered by decreasing spatial scales (Borcard and Legendre 2002, Borcard et al. 2004). We analyzed the fitted values corresponding to fractions [a + b] (caused by environmental control processes corresponding to topography) and [c] (caused by unmeasured environmental variables and neutral processes) of variation partitioning by successive blocks of 25 PCNMs (Fig. 6). Other divisions of the PCNMs into blocks produced similar results. We found that fraction [a + b] of the species composition data, which is the portion fitted to the topographic

a) The 66 selected PCNM eigenfunctions for the species richness variable

b) The 179 selected PCNMs for the community composition data table
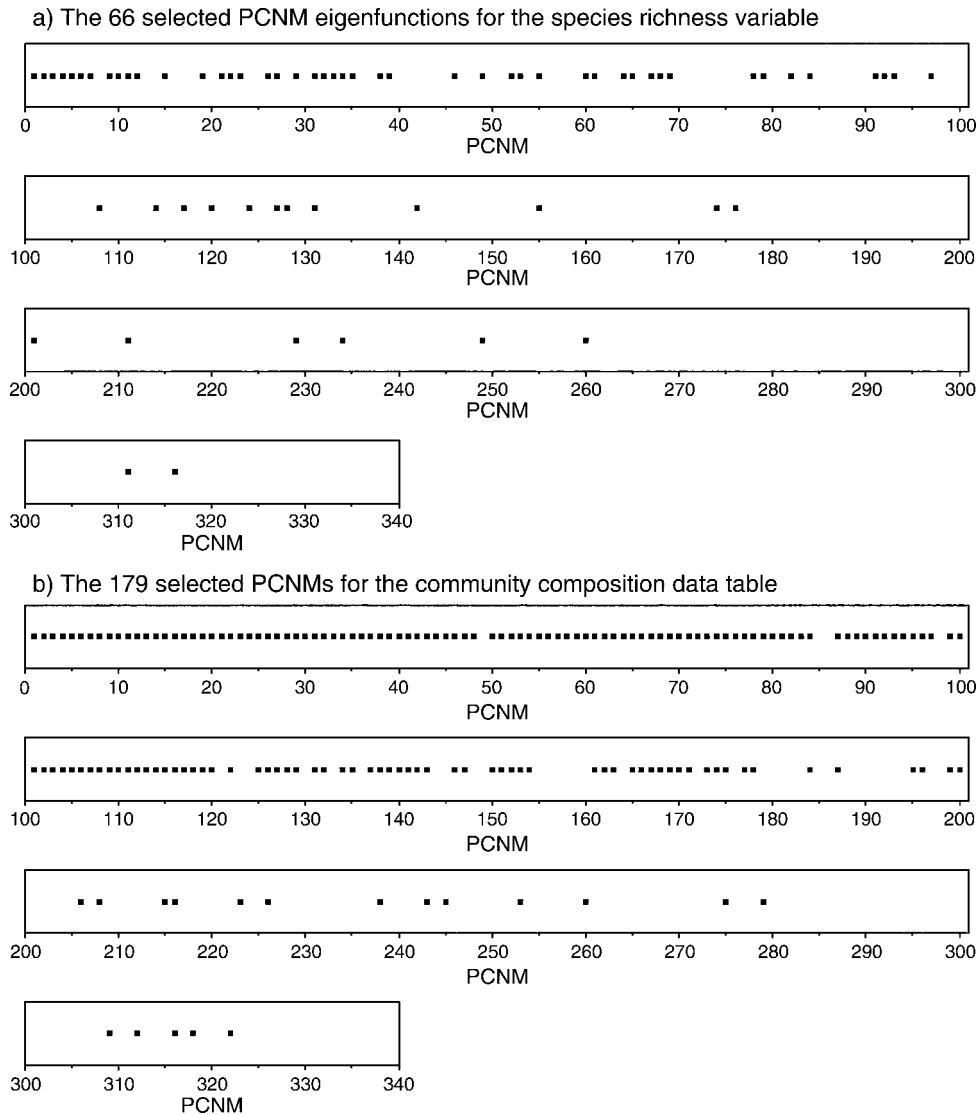
FIG. 2. Among the 339 principal coordinates of neighbor matrices (PCNM) eigenfunctions, those that were selected by forward selection are represented by square dots. (a) For species richness, most of the 66 selected PCNMs are among the first 100 (first row). (b) For community composition, most of the 179 selected PCNMs are among the first 180 (first two rows). Fifty-six PCNMs are common to the two lists.

variables, corresponded mostly to broad-scaled spatial structures (PCNMs #1–25: $R_a^2 = 0.565$), whereas fraction [c] corresponded to a mixture of broad- and finer-scaled structures (PCNMs 1–175, in blocks 1–7, had $R_a^2 > 0$).

Interpretation of the canonical axes, 20 × 20 m cells.— Partitioning richness variation among the groups of topographic variables shows the dominant effect of altitude and convexity: the altitude set of three monomials accounts for 8.4% of the richness variation (semipartial R^2) and the convexity set for 11.6%, whereas the other two sets account for very little (slope set 2.5%, aspect set 0.2%). Species richness is negatively correlated with altitude ($r = -0.086$) and convexity ($r = -0.388$) and positively correlated with slope ($r = 0.204$), meaning that

richness is higher at lower altitude and on sloping ground that is not strongly convex. The map of the fitted values of fraction [a + b] of richness (Fig. 4b) reflects the pattern of valleys which can be seen in Fig. 1b, c, and the highest values of richness are at low altitude (Fig. 1a).

For the analysis of species composition in 20 × 20 m cells, Fig. 5 shows the maps of the fitted values of the six most significant canonical axes corresponding to the total explained variation [a + b + c], the topographic variation [a + b], and the spatially structured fraction [c] unexplained by the topographic variables. In the case of [a + b + c] and for the 20 × 20 m cells, axis 1 is strongly correlated with the three monomials of altitude (multiple $R^2 = 0.533$) and convexity (multiple $R^2 = 0.217$). Axis 2 is more weakly correlated with altitude (multiple $R^2 =$
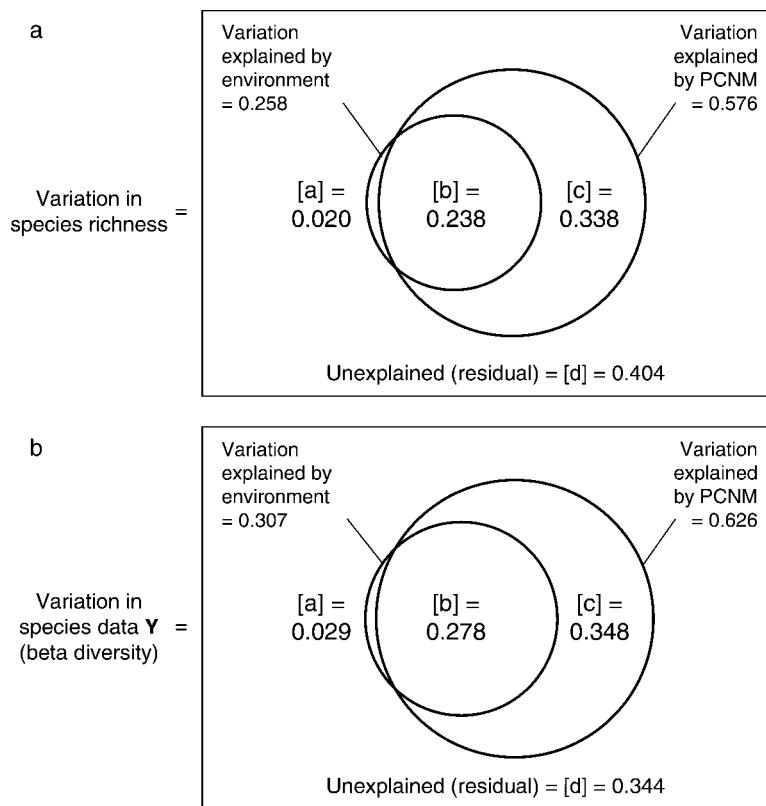
FIG. 3. Variation partitioning results. The two figure panels show Venn diagrams representing the partition of the variation of (a) species richness and (b) community composition, between two sets of explanatory variables: topographic variables (left circle) and PCNM eigenfunctions (right circle). Each box represents 100% of the variation in the corresponding response variable (a, species richness) or data table (b, species data $\mathbf{Y}$). The reported fractions (names as in Table 1) are adjusted R^2 statistics (R_a^2).

0.218) and convexity (multiple $R^2 = 0.147$). Axis 3 is correlated with convexity (multiple $R^2 = 0.105$). The slope and aspect sets of monomials are weakly correlated with these three axes. The main north–south crest of the plot, which has special vegetation types, stands out in the maps of these axes (Fig. 5a–c). Several species are associated with altitude: 11 species have their variance fitted at 30% or more by the canonical axes that are linear combinations of the three monomials of altitude; of these species, *Camellia chekiang-oleosa* (abbreviation in Appendix A: CamChe), *Corylopsis glandulifera* var. *hypoglauca* (CorGla), *Lyonia ovalifolia* var. *hebecarpa* (LyoOva), *Quercus serrata* var. *brevipetiolata* (QueSer), *Rhododendron mariesii* (RhoMar), *Rhododendron simsii* (RhoSim), *Schima superba* (SchSup) are positively correlated and *Adinandra millettii* (AdiMil), *Elaeocarpus japonicus* (ElaJap), *Machilus grijsii* (MacGri), *Tarenna mollissima* (TarMol) are negatively correlated with altitude. Other species are associated with convexity (three species have their variance fitted at 20% or more by the canonical axes that are linear combinations of the three monomials of convexity: *Pinus massoniana* (PinMas) and *Quercus serrata* var. *brevipetiolata* (QueSer) are positively corre-

lated and *Camellia fraterna* (CamFra) is negatively correlated with convexity.

Because fraction [a + b] is fitted to the topographic variables, one can look for relationships with these variables in Fig. 5d, e. Axis 1 is strongly correlated with altitude (multiple $R^2 = 0.515$) and convexity (multiple $R^2 = 0.210$). Axis 2 is correlated with altitude (multiple $R^2 = 0.270$) and convexity (multiple $R^2 = 0.124$). Axis 3 is more lightly correlated with altitude (multiple $R^2 = 0.111$) and convexity (multiple $R^2 = 0.193$). The slope and aspect sets of monomials are weakly correlated with these three axes. Fig. 5f shows canonical axis 1 and fraction [c]. Fraction [c] represents spatially structured variation that is not explained by the present set of topographic variables; it is thus unrelated to these variables. The multiple correlation of that axis with the 11 topographic monomials is zero.

Habitat types, 20 × 20 m cells.—Multivariate regression tree analysis (MRT) produced five groups (habitat types); they are presented on a map of the plot in Fig. 7. Groups 1 and 2 are separated from groups 3–5 by the altitude breakpoint of 196.6 m above the lowest corner of the plot (or 642.9 m above sea level); groups 1 and 2 are found in the lower portions of the plot. Group 1 ($n_1 = 237$), in the valleys, is separated from group 2 ($n_2 = $

a) Fitted species richness, fraction [a + b + c].
Gray scale from 19.9 (white) to 53.7 (black) species.
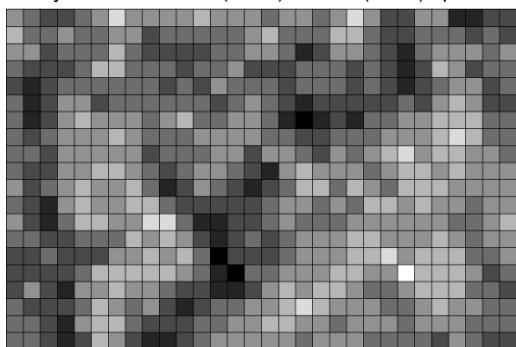
b) Fitted species richness, fraction [a + b].
Gray scale from 26.4 (white) to 50.1 (black) species.
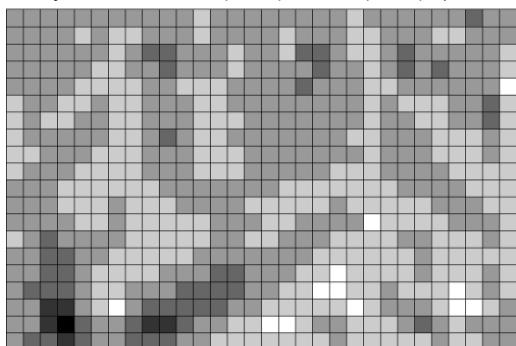
c) Fitted species richness, fraction [c].
Gray scale from 18.9 (white) to 48.9 (black) species.
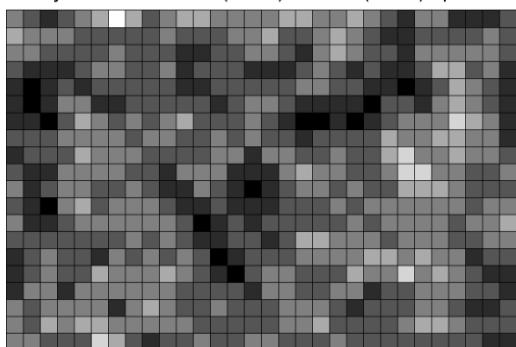
Fig. 4. Maps of the study plot showing fractions of variation of richness represented by the fitted values of three multiple regressions. There are five levels of gray scale in each panel: (a) [a + b + c], fitted values of the regression on the topographic variables and PCNM base functions; (b) fraction [a + b], fitted values of the regression on the topographic variables alone; (c) fraction [c], fitted values of the regression on PCNM base functions while controlling for the topographic variables.

269), on the mid-altitude ridges, by a breakpoint in the convexity variable at 0.585. Group 5 ($n_5 = 44$) contains the most convex cells (convexity ≥ 3.3); groups 3 and 4 occupy the less convex high-altitude cells. Group 3 ($n_3 = 42$) contains cells located in lower altitude than group 4 ($n_4 = 8$) (breakpoint at altitude 236.5 m above the lowest corner of the plot, or 682.8 m above sea level).

The nine species that are statistically significant indicators of habitat types 1, 4, and 5 are presented in Appendix B. No statistically significant indicator species were found for habitat types 2 and 3.

DISCUSSION

The answer to the question about how beta diversity is maintained in a community depends on our ability to decompose the variation of diversity into the contributions of different processes affecting it. Such processes are potentially numerous (e.g., life history traits, reproductive and dispersal behavior, soil properties, climatic variation, etc.), depending on the mechanisms of interest in a study and data availability. The purpose of our study was to show that environmental control was not diametrically opposed to the mechanisms, including neutral, that generate fine-scaled spatial variation unexplained by the environmental variables acting at broad scales. Our results showed that the two types of processes actually worked together, side by side, to regulate beta diversity in the Gutian forest plot.

In the present study, we focused on identifying the effects of topography as well as the spatially structured processes that do not covary with topography. The Gutian plot has the roughest terrain among all the plots that have so far been mapped in tropical forests, and also in non-tropical areas of China. Topography should contribute significantly to our understanding of how diversity is maintained on such a terrain.

Spatial scale (cell size)

PCNM analysis provided a powerful tool for analyzing the spatial variation in species composition (beta diversity) in the Gutian plot. The PCNM results have shown that the magnitudes of spatial variation of richness on the one hand and community composition on the other are very similar (Fig. 3). In both cases, the dominant structure is broad-scaled (Fig. 2) even though the terrain is highly variable. Because community composition is a less synthetic description of forest cells than richness, it is more powerful for detecting significant relationships with PCNM variables (179 significant PCNMs for community composition compared to 66 for richness), but the 66 richness-related PCNMs are largely a subset of the 179 community-related PCNMs; the two lists have 56 PCNM variables in common. This indicates that the unmeasured spatially structured variables explaining richness also contribute to the explanation of the variation in species composition.

A striking finding of our scaling analysis (Table 1) is that the total proportion of explained variation (or complementarily the total unexplained variation, fraction [d]) in species richness and community composition is nearly invariant to changes in sampling scale (cell size). This is due to the scale-dependent tradeoff between environmental-control effects (fraction [a + b]) and the effect of unobserved variables, including neutral pro-
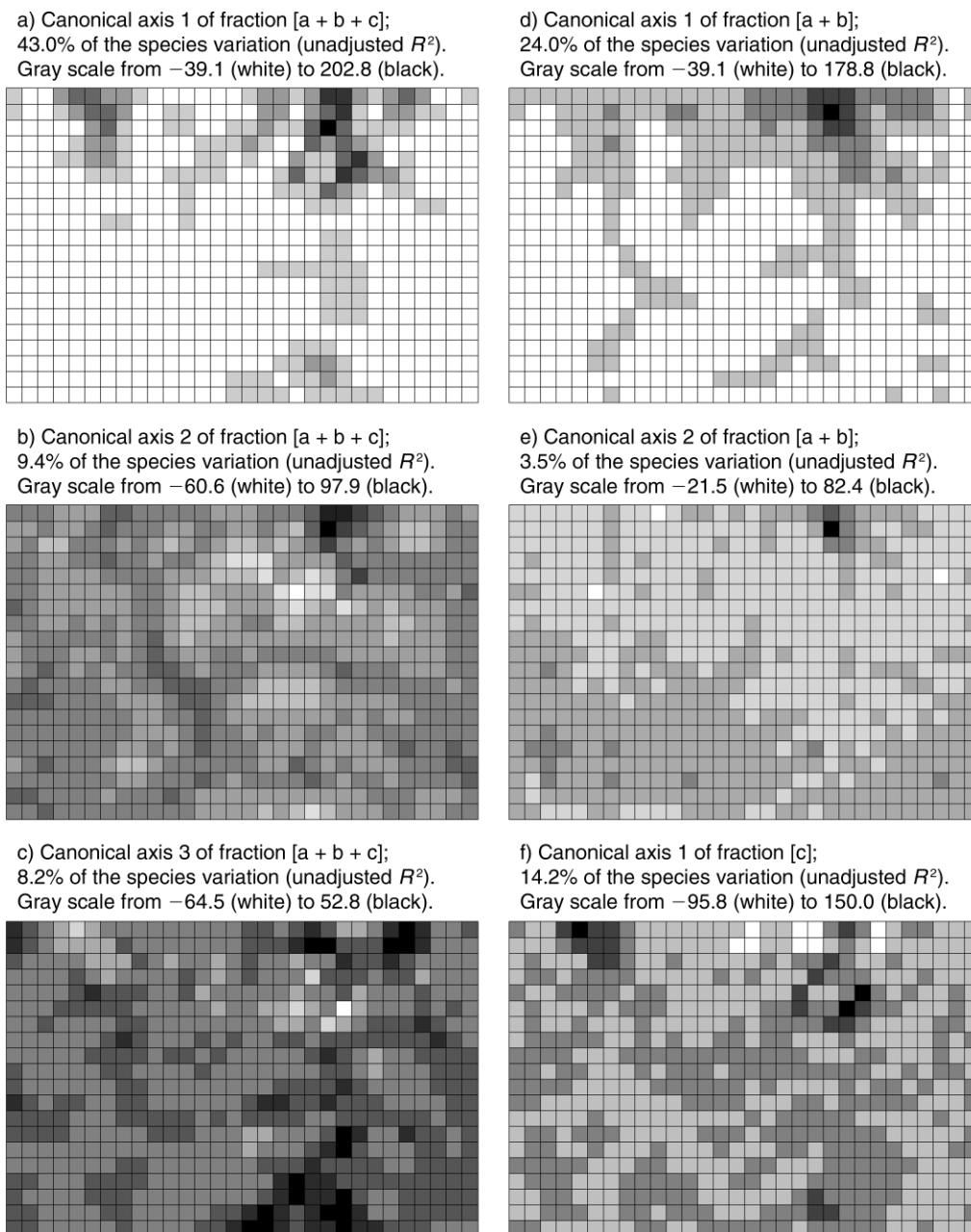
FIG. 5. Maps of the fitted values of the six most significant canonical axes (a–f) corresponding to the total explained variation in richness [a + b + c], the topographic variation [a + b], and the spatially structured fraction [c] unexplained by the presently available topographic variables.

cesses, whose effects come out in the pure spatial structure [c]. The increase in cell size homogenizes ("averages") the effect of topographic variables in individual cells, thus increasing fraction [a]. In contrast, the spatially structured variation ([b + c]) decreases with scale (cell size). That decrease is due to the fact that as cell size increases above 20 × 20 m, the PCNMs are reduced in numbers from 339 for 20 × 20 m cells to 84 for 40 × 40 m cells and 55 for 50 × 50 m cells. The broad-scaled PCNMs remain the same, but the PCNMs

computed for large cell sizes capture less and less of the intermediate-scale spatial variation, which is important to explain the overall variation of the vegetation data (Fig. 2).

Environmental control and neutral processes

A large portion (~60%) of the variation of species richness and community composition in the Gutian plot is determined by topography and the PCNMs (Fig. 3). This portion can be further divided into purely
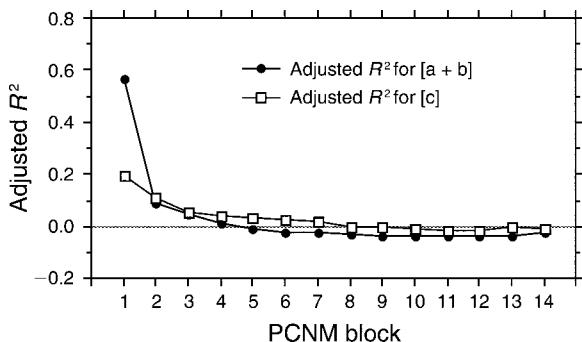
FIG. 6. Adjusted R^2 of the explanation provided by successive blocks of PCNM eigenfunctions for the variation captured by fractions [a + b] and [c] of the species composition for 20 × 20 m cells. Block 1, PCNMs numbered 1–25, etc.; Block 14, the last 14 PCNMs, numbered 326–339.

topographic variation ([a]), the spatially structured variation ([c]), and a combination of the two ([b]). It is not surprising to see that much of the topographic variation is spatially structured because of the strong south–north altitudinal gradient of the plot and the obvious spatial patterns in convexity, slope, and aspect (Fig. 1). It is a salient feature that the spatially structured component ([b + c]) explains such a large proportion of variation in community composition (58% and 63% in Fig. 3a, b). Component [c] represents the contributions of unobserved variables that are not correlated with topography but are spatially structured plus the spatially structuring effect of community dynamics. They include soil properties as well as other

habitat and ecological mechanisms such as directional dispersal of propagules.

It is important to note that about 40% of the variation is undetermined (fraction [d] in Fig. 3). Several reasons may be invoked to explain this high proportion of unaccounted variation, e.g., other nonspatially structured biological or environmental factors that were not measured in the field. Another plausible explanation is that it may be due to stochastic processes. The latter explanation has theoretical connection to the neutral theory of macroecology which assumes that the dynamics of populations are primarily driven by ecological drift and dispersal, with or without limitation, and are not habitat dependent. Dispersal has a spatial signature and produces variation in fractions [c] and [d] whereas the effect of drift comes out in fraction [d]. It is likely that the variation decomposition shown in Fig. 3 will be altered if other environmental (e.g., soil chemistry) or biological (e.g., species traits) variables are recorded and included in the analysis (John et al. 2007). However, given the complexity of the topography of the plot, we suspect this would not change the finding that stochastic variation is a significant component of the Gutian plot. This prediction remains to be tested; we are currently sampling soil data for that purpose.

In summary, our results suggest that both deterministic (topography and other spatially structured environmental variables) and stochastic processes are substantial determinants of the distribution of the tree diversity in the Gutian plot. This general result seems robust to the effect of the spatial scale of observation
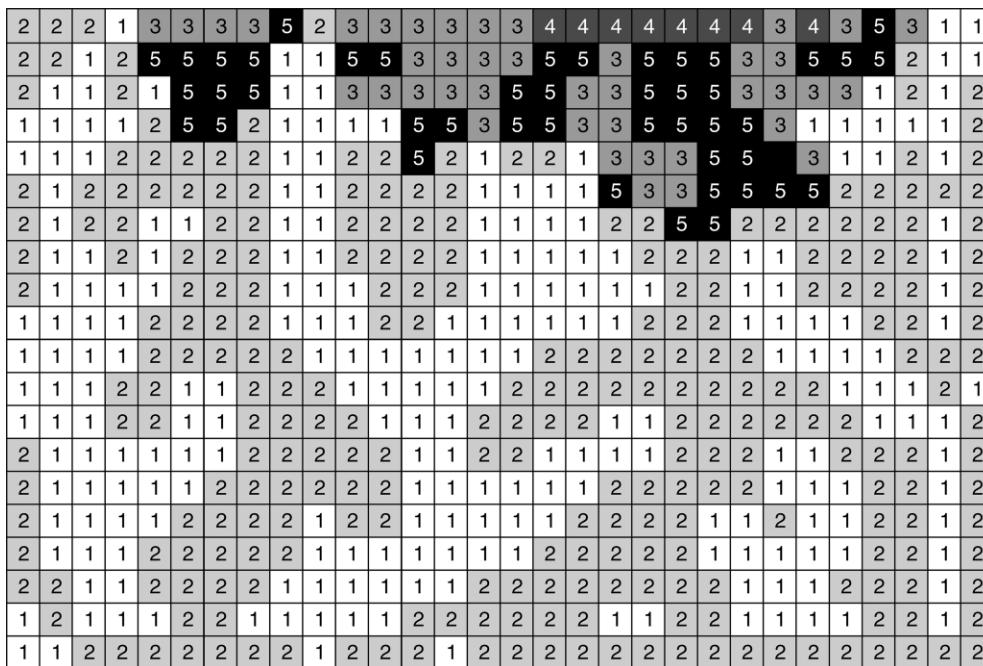
FIG. 7. Map of the 20 × 20 m cells classified into five habitat types (groups 1–5 in different shades of gray) by multivariate regression tree analysis.

(Table 1). Overall, the variation accounted for by topography ([a + b]) increases with mapping scales, whereas the purely spatially structured variation ([c]), which is a signature of neutral processes and the undetermined variation, decreases, as expected by the homogenizing effect of large cells.

The topographic factors that control richness and species composition are similar, with altitude and convexity playing dominant roles, and slope and aspect playing weak roles. The results reported in Fig. 6 show, however, that environmental control (fraction [a + b]) was correlated to PCNM eigenfunctions describing broader scales than fraction [c] which corresponds to unmeasured environmental variables and neutral processes. These two sources of variation cannot be distinguished at the present time.

Habitat types and indicator species

The Gutian plot can be divided into five habitat types in terms of topographic variation, with the valleys and ridges characterizing the plot (corresponding to habitats 1 and 2 in Fig. 6). It is, however, interesting to observe that these types may or may not represent distributions of unique species composition. For example, two species are found to be statistically significant indicators of the valleys (Fig. 6, Appendix B), but no species are good indicators of the ridges. This result is different from that of Barro Colorado Island BCI which was classified by Harms et al. (2001) into six habitat types, with many species found to have strong habitat associations. An explanation for the lack of species–habitat association in some habitats is that the forest is largely secondary and was disturbed about 50 years ago; the northeast part of the plot was once burned.

MRT habitat type 1 in the low valleys is significantly correlated with hygrophilous and shade-tolerant species, such as *Camellia fraterna* and *Neolitsea aurata* var. *chekiangensis*, whereas low ridges (habitat type 2) are always characterized by species such as *Schima superba*, *Castanopsis eyrei*, *Pinus massoniana*, and similar species (these were not, however, statistically significant indicators, Appendix B). The cells around high slopes (habitat types 4 and 5, which include part of the high ridges) were severely disturbed by fire in the 1960s. As a result, the average diameter at breast height of individuals in these disturbed habitats is much smaller; the individuals are also denser than in other habitats. Therefore the forest in the high slopes (habitat type 4) is significantly correlated with pioneer species, such as *Quercus serrata* var. *brevipetiolata, Lyonia ovalifolia* var. *hebecarpa*, and *Rhododendron mariesii*. Sometimes the forest on ridges (habitat types 4 and 5) and especially on high ridges (habitat type 5), was more easily disturbed by fire caused by lightning strokes at the local scale; parts of the high ridges were also disturbed by fire around high slopes. Therefore, the vegetation on high ridges (habitat type 5) is significantly correlated with disturbance-related species such as *Albizia kalkor, Lindera reflexa*, and *Platycarya strobilacea*, and is also indicated by *Sorbus folgneri* which grows along high-elevation ridge. The mid-slope (habitat type 3) is the transitional area between high ridges and low valleys and ridges; no species was found to be indicative of this habitat, but most species can grow in that habitat.

Future work

The method of analysis described in this paper will be used to compare permanent broad-scale stem-mapped forest plots throughout the world, from different regions and latitudes. These comparisons will allow us to study spatial macroecological patterns of tree species along the latitudinal gradient and test biodiversity theories about the mechanisms that are important for the maintenance of biodiversity. The comparisons will focus on important questions such as the degree of neutrality of each forest and the factors controlling beta diversity: are they the same everywhere, or do they differ with region and latitude? We will also be able to determine if soil chemistry data, where available, are correlated and thus redundant with the topographic variables or if they bring additional information to explain beta diversity.

LITERATURE CITED

Borcard, D., and P. Legendre. 1994. Environmental control and spatial structure in ecological communities: an example using oribatid mites (Acari, Oribatei). Environmental and Ecological Statistics 1:37–53.

Borcard, D., and P. Legendre. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. Ecological Modelling 153:51–68.

Borcard, D., P. Legendre, C. Avois-Jacquet, and H. Tuomisto. 2004. Dissecting the spatial structure of ecological data at multiple scales. Ecology 85:1826–1832.

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

Borda de Agua, L., S. P. Hubbell, and F. He. 2007. Scaling biodiversity under neutrality. Pages 347–375 *in* D. Storch, J. Brown, and P. Marquet, editors. Scaling biodiversity. Cambridge University Press, Cambridge, UK.

Bray, R. J., and J. T. Curtis. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecological Monographs 27:325–349.

Condit, R. 1995. Research in large, long-term tropical forest plots. Trends in Ecology and Evolution 10:18–22.

Condit, R., et al. 2006. The importance of demographic niches to tree diversity. Science 313:98–101.

Condit, R., R. Sukumar, S. P. Hubbell, and R. B. Foster. 1998. Predicting population trends from size distributions: a direct

test in a tropical tree community. American Naturalist 152: 495–509.

De'ath, G. 2002. Multivariate regression trees: a new technique for modeling species–environment relationships. Ecology 83: 1105–1117.

De'ath, G. 2006. mvpart: multivariate partitioning. R package version 1.2-4. ⟨http://cran.r-project.org/⟩

Dray, S., P. Legendre, and P. R. Peres-Neto. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). Ecological Modelling 196:483–493.

Dufrêne, M., and P. Legendre. 1997. Species assemblages and indicator species: the need for a flexible asymmetrical approach. Ecological Monographs 67:345–366.

Fortin, M.-J., and M. R. T. Dale. 2005. Spatial analysis: a guide for ecologists. Cambridge University Press, Cambridge, UK.

Gravel, D., C. D. Canham, M. Beaudet, and C. Messier. 2006. Reconciling niche and neutrality: the continuum hypothesis. Ecology Letters 9:399–409.

Harms, K. E., R. Condit, S. P. Hubbell, and R. B. Foster. 2001. Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. Journal of Ecology 89:947–959.

He, F., and P. Legendre. 2002. Species diversity patterns derived from species–area models. Ecology 83:1185–1198.

Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography. Princeton University Press, Princeton, New Jersey, USA.

Hutchinson, G. E. 1957. Concluding remarks. Cold Spring Harbor Symposia on Quantitative Biology 22:415–427.

Ibáñez, R., R. Condit, G. Angehr, S. Aguilar, T. Garcia, R. Martínez, A. Sanjur, R. Stallard, S. J. Wright, A. S. Rand, and S. Heckadon. 2003. An ecosystem report on the Panama Canal: monitoring the status of the forest communities and the watershed. Environmental Monitoring and Assessment 80:65–95.

John, R., J. W. Dalling, K. E. Harms, J. B. Yavitt, R. F. Stallard, M. Mirabello, S. P. Hubbell, R. Valencia, H. Navarrete, M. Vallejo, and R. B. Foster. 2007. Soil nutrients influence spatial distributions of tropical tree species. Proceedings of the National Academy of Sciences (USA) 104:864–869.

Karst, J., B. Gilbert, and M. J. Lechowicz. 2005. Assembling fern communities: the roles of chance and environmental determinism at local and intermediate spatial scales. Ecology 86:247–2486.

Koleff, P., K. J. Gaston, and J. J. Lennon. 2003. Measuring beta diversity for presence–absence data. Journal of Animal Ecology 72:367–382.

Laliberté, E., A. Paquette, P. Legendre, and A. Bouchard. 2008. Assessing the scale-specific importance of niches and other spatial processes on beta diversity: a case study from a temperate forest. Oecologia, in press. [doi: 10.1007/s00442-008-1214-8]

Legendre, P. 2007. Studying beta diversity: ecological variation partitioning by multiple regression and canonical analysis. Journal of Plant Ecology [formerly Acta Phytoecologica Sinica] 31:976–981. [In Chinese].

Legendre, P., D. Borcard, and P. R. Peres-Neto. 2005. Analyzing beta diversity: partitioning the spatial variation of community composition data. Ecological Monographs 75: 435–450.

Legendre, P., and L. Legendre. 1998. Numerical ecology. Second English edition. Elsevier Science BV, Amsterdam, The Netherlands.

Losos, E. C., and E. G. Leigh, editors. 2004. Tropical forest diversity and dynamism: findings from a large-scale plot network. Chicago University Press, Chicago, Illinois, USA.

Oksanen, J., R. Kindt, P. Legendre, and R. B. O'Hara. 2007. vegan: community ecology package. R package version 1.9-25. ⟨http://cran.r-project.org/⟩

Peres-Neto, P. R., P. Legendre, S. Dray, and D. Borcard. 2006. Variation partitioning of species data matrices: estimation and comparison of fractions. Ecology 87:2614–2625.

R Development Core Team. 2007. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ⟨http://www.R-project.org⟩

Rao, C. R. 1964. The use and interpretation of principal component analysis in applied research. Sankhyaá, Series A 26:329–358.

Roberts, D. W. 2006. labdsv: Laboratory for Dynamic Synthetic Vegephenomenology. R package version 1.2-2. ⟨http://cran.r-project.org/⟩

Tilman, D. 2004. Niche tradeoffs, neutrality, and community structure: a stochastic theory of resource competition, invasion, and community assembly. Proceedings of the National Academy of Sciences (USA) 101:10854–10861.

Valencia, R., R. B. Foster, G. Villa, R. Condit, J.-C. Svenning, C. Hernández, K. Romoleroux, E. Losos, E. Magård, and H. Balslev. 2004. Tree species distributions and local habitat variation in the Amazon: large forest plot in eastern Ecuador. Journal of Ecology 92:214–229.

Whittaker, R. H. 1956. Vegetation of the Great Smoky Mountains. Ecological Monographs 26:1–80.

Whittaker, R. H. 1960. Vegetation of the Siskiyou mountains, Oregon and California. Ecological Monographs 30:279–338.

Whittaker, R. H. 1972. Evolution and measurement of species diversity. Taxon 21:213–251.

Wu, Z. 1980. Vegetation of China. Science Press, Beijing, China. [In Chinese.]

Yu, M.-J., Z.-H. Hu, J.-P. Yu, B.-Y. Ding, and T. Fang. 2001. Forest vegetation types in Gutianshan Natural Reserve in Zhejiang. Journal of Zhejiang University (Agriculture and Life Science) 27:375–380. [In Chinese.]

Zheng, C. 2005. Keys of seed plants in Zhejiang. Zhejiang Science and Technology Press, Hangzhou, China. [In Chinese.]

APPENDIX A

Latin names of the 49 families and 159 species found in the Gutianshan plot (*Ecological Archives* E090-046-A1).

APPENDIX B

Indicator values for the nine significant species found in three multivariate regression tree (MRT) groups, and distribution maps of individuals of those species (*Ecological Archives* E090-046-A2).

SUPPLEMENT

R-language code to compute the coordinates of the Gutianshan plot cells for the various cell sizes used in the paper, plot maps indicating cell positions, compute PCNM eigenfunctions, and plot maps of all PCNM eigenfunctions (*Ecological Archives* E090-046-S1).