ELSEVIER

Forest Ecology and Management 172 (2003) 89–108

Forest Ecology and Management

www.elsevier.com/locate/foreco

Drawing ecological insights from a management-oriented forest inventory in French Guiana

P. Couteron[a,*], R. Pélissier[b], D. Mapaga[c], J.-F. Molino[b], L. Teillier[d]

[a]ENGREF/UMR botAnique et BioinforMatique de l'Architecture des Plantes (AMAP), Boulevard de la Lironde, TA40/PS2, 34398 Montpellier Cédex 05, France
[b]IRD/UMR botAnique et BioinforMatique de l'Architecture des Plantes (AMAP), Boulevard de la Lironde, TA40/PS2, 34398 Montpellier Cédex 05, France
[c]IRAF, BP 2246 Libreville, Gabon
[d]Sylvafrica SA, filiale d'ONF-International, BP 1888 Libreville, Gabon

Received 19 September 2001; received in revised form 23 April 2002; accepted 19 May 2002

Abstract

Reliable ecological information at the landscape scale is generally lacking for tropical rain forests, although extensive areas have been sampled by forest inventory to estimate timber resources. We used the data provided by a 12,240 ha management-oriented forest inventory in the lowland rain forest of French Guiana to document species/environment relationships and to characterise the spatial variation of the floristic composition. The forest inventory encompassed 22,023 trees larger than 7.5 cm diameter measured in 411 0.3 ha sampling plots spread over a systematic grid with 500 m × 400 m spacing between plot centres. In each plot, all the trees above 37.5 cm diameter have been recorded, while the trees between 7.5 and 37.5 cm diameter have been recorded in smaller sub-plots. Each sampling plot was characterised using semi-quantitative ecological descriptors relating to topography, remnants of lateritic cuirasses, presence of hydromorphic soils, etc. Preliminary analyses revealed that most of these variables could be accounted for by topographical categories subdivided in relation to presence/absence of hydromorphic soils. However, stand structure, expressed by the distribution of trees in diameter classes, proved fairly independent on such categories.

Floristic information was based on a refined vernacular nomenclature (291 taxa) with collection of herbarium specimens that allowed us to equate 59 taxa with known botanical species. We produced a reduced floristic table expressing the distribution of the 59 botanical species within the 411 plots, and a complete floristic table featuring the distribution of all the vernacular taxa. The main floristic gradients were extracted from these tables via correspondence analysis (CA) and non-symmetric correspondence analysis (NSCA), a complementary approach that emphasises frequent species. Analogous constrained ordinations (CAIV, i.e. canonical correspondence analysis, and NSCAIV), based on the approximation of the floristic tables by ecological variables (i.e. topography and stand structure) were also used.

All analyses yielded consistent results pointing towards a major floristic gradient closely linked to topography, and to secondary gradients related to stand structure. The environmental variables had significant and non-redundant explanatory powers for floristic composition. Two main geographical partitions of the forest were revealed, one reasonably accounted for by environmental variables, the other remaining insufficiently explained.

In tropical rain forests, inventories could be a valuable source of ecological information, at the price of reasonable effort oriented towards enhanced vernacular nomenclatures (collection of herbarium specimens, training of tree-spotters), in situ

---

[*] Corresponding author. Present address: ENGREF, 648 rue J.F. Breton, B.P. 44494, 34093 Montpellier Cédex 05, France.
Tel.: +33-467-047-126; fax: +33-467-047-101.
E-mail address: couteron@engref.fr (P. Couteron).

PII: S0378-1127(02)00310-9

recording of simple environmental variables (e.g. slope, topography), and data analyses based on complementary ordination methods.

Keywords: Forest inventory; French Guiana; Neotropics; Ordination; Tropical rain forest

1. Introduction

The lack of reliable ecological information on tropical rain forests at a critical mesoscale (ca. 1–100 km^2; Clark et al., 1998, 1999), is currently recognised as a major hindrance for connecting local results, established at a plot or transect level, with floristic syntheses carried out at regional scales of ca. 10^4 to 10^7 km^2 (e.g. Ter Steege et al., 2000a,b). Such a lack of information is largely due to the cost of landscape scale sampling designs in tropical rain forests (Greig-Smith, 1983), that exceeds the means devoted to ecological investigations. On the other hand, vast areas in the humid tropics have been sampled by forest services or logging companies, aimed at appraising timber resources. Large data sets about the structure and composition of forest stands, that are gathered through inventories, could serve as a valuable source of mesoscale information, enabling a broader understanding of species/environment relationships, and opening the way to the definition and mapping of forest types (Condit, 1996) at a scale compatible with management/conservation issues.

Consequently, drawing ecological insights from inventory data should be considered as an important challenge for tropical rain forest ecology. However, forest inventories have been scarcely used for ecological research (Sheil, 1995) mainly because the corresponding floristic data may often have appeared as incomplete or imprecise, with regard to academic standards. In fact, the floristic information usually collected by local tree-spotters, aimed at timber resource evaluation, is often based on commercial and/or vernacular names. This does not mean that such data are necessarily useless, since many of the vernacular names refer to a particular botanical species (Oldeman, 1968; Richards, 1996, pp. 495–496). In addition, most forest inventory protocols are based on individual tree records over a network of sampling units (Shiver and Bonders, 1996) whose stand structure can be simply described through the frequency distribution of trees in different size (diameter) classes. Moreover, in most cases, simple characteristics of the sampling units, such as slope angle or the presence of particular herbaceous species, are also available and can be used as simple ecological descriptors. Compilation of these data can thus allow the realisation of forest typologies through a comparison of floristic and ecological variables.

In 1994, within the framework of a pilot management project of the Office National des Forêts (ONF), an inventory covering 12,240 ha has been launched in Counami forest (French Guiana) and carried out by CIRAD-Forêt. Although, the main objective was a classical timber evaluation of the principal commercial species, particular attention was given to the sampling protocol (Teillier, unpublished) which featured both a careful floristic identification, using a refined vernacular nomenclature (with collection of herbarium specimens), and a detailed characterisation of the sampling units, through a set of ecological variables. As a consequence, the resulting data base was more accurate and consistent than would generally be expected from tropical forest inventories in French Guiana.

The present paper analyses the relationship between floristic composition and simple ecological variables on the basis of the forest inventory data of Counami. Our overall approach stems from the following assumptions: (i) the most common vernacular names could be translated into their botanical equivalents using herbarium samples and information obtained in comparable forests (Richards, 1996, pp. 495–496); (ii) the substratum (sensu lato) and the local topography influence the floristic composition (Ter Steege et al., 1993; Collinet, 1997; Sabatier et al., 1997; Pélissier et al., 2002a); (iii) the structure in diameter classes allows to characterise the gap-phase regeneration stages and their corresponding floristic composition (Ashton and Hall, 1992; Poorter et al., 1994; Riéra et al., 1998).

The hypothesis that several ecological variables may be relevant predictors for the floristic composition of the forest stands has a simple translation in terms of multivariate analysis of data tables. Classically, main floristic gradients, that summarise the distribution of species among the sampling units, are obtained through an analysis of the plots by species table (Hill, 1973). Methods of tables coupling (Ter Braak, 1987; Sabatier et al., 1989) then allow an ordination of species along explicit gradients determined by ecological variables, whose explanatory power can, furthermore, be evaluated. Hence, the main purpose of this paper is to identify, for the Counami forest, patterns of floristic variations that would be consistent and robust in regard to the method of ordination used and/or to the fraction of the species pool taken into consideration. We also intend to assess the extent to which such floristic structures may be explained by ecological variables characterising both the physical environment (topography, water saturation, etc.) and the structure of the stand (distribution of diameter classes). In doing so, our final aim is to provide an understanding and a zonation of the forest that could both serve management purposes and guide further scientific investigations.

2. Materials

2.1. Study site and sampling design

Counami forest is located at about 143 km north-west of the main town of Cayenne in French Guiana. It is an unlogged lowland rain forest, growing under a humid tropical climate, with annual rainfall ranging between 2750 and 3000 mm, and scattered over 9 months (Blancaneaux, 2001). The soils are of ferralitic type, developed on granitic or shistose parent-rocks (Milési et al., 1995). The current erosion-induced transformation along topographical catenas of the initial ferralitic cover has been established by several authors in French Guiana (see for an introduction Sabatier et al., 1997 or Pélissier et al., 2002a). This process leads to important discrepancies in the soil drainage between the hilltops and the foot-slopes, which often experience more or less prolonged periods of water saturation (hydromorphic soils).

The forest inventory was based on a systematic sampling design featuring 411 rectangular plots (75 m × 40 m, i.e. 0.3 ha) that were oriented according to an azimuth of 19°, with distances between centres of 500 m in this direction and 400 m on a perpendicular axis (i.e. one plot per 20 ha). Only the trees with a diameter (dbh; at breast height or 30 cm above the buttresses if any) larger or equal to 37.5 cm have been identified and measured in the whole plot area. Trees between 22.5 and 37.5 cm dbh (upper bound exclusive) have only been recorded in one half of the plot, whilst smaller trees between 7.5 and 22.5 cm have been recorded in a central circular sub-plot of radius 11.28 m (i.e. 0.04 ha). In the following, all the trees $\geq$7.5 cm dbh are used to characterise the structure and composition of the sampling units. They represented a total of 22,023 individuals.

2.2. Floristic data

The trees were identified by an experienced team of local tree-spotters supervised by a forest officer (one of us, L. Teillier). Identifications followed a vernacular nomenclature that has been refined and completed for the requirements of this inventory. It comprised 291 taxa, that were either commercial, vernacular, or species/genera names. Apart from the difficulties of species identifications in rain forests, and particularly in Amazonia, the use of a vernacular nomenclature poses different kinds of problems presented by Oldeman (1968) and Richards (1996); among them are: (i) the accuracy of the vernacular names that sometimes encompass related taxa instead of a single botanical species; (ii) the ability of a given tree-spotter to properly discriminate among vernacular categories in the field. The first point means that different botanical species can be referred to the same vernacular name, and the second point means that a particular botanical species can be found under different vernacular names.

In order to identify which of the vernacular categories lend themselves to translation into known botanical species, we used the herbarium specimens collected at Counami (Teillier, unpublished data) and the data from two field surveys undertaken by two experienced botanists at Counami and at the neighbouring forest research facility of Paracou (Molino and Sabatier, unpublished data), on the basis of the vernacular names given by the same team of tree-spotters. Using these data, we had at our disposal a total of 1764 individuals whose vernacular names and species-level identifications were known (species nomenclature follows Boggan et al., 1997). They represented 351 botanical species and 179 of the 291 vernacular names used in the forest inventory. By comparison, in a 10 ha plot at Piste de St. Elie in French Guiana, Sabatier et al. (1997) identified a total of 459 botanical species among 6134 individuals with a dbh above 10 cm.

We then arranged the data in the form of a table expressing all the correspondences between the botanical species (351 rows) and the vernacular names (179 columns). We computed for each vernacular category, v, represented by at least two individuals, an index of homogeneity corresponding to Simpson's (1949) formula:

$$\hat{\lambda}_v = \sum_{i=1}^{N_s} \frac{n_{iv}(n_{iv}-1)}{n_{\cdot v}(n_{\cdot v}-1)}$$

where the summation is over all N_s botanical species ($N_s = 351$), and where n_{iv} denotes the number of individuals identified as belonging to species i and to category v, in which $n_{\cdot v}$ individuals have been recorded. The term λ_v is interpreted as the probability that two individuals bearing the same vernacular name v belong to the same botanical species. The greater the homogeneity index, the more reliable is the translation of a vernacular name into a botanical species. Similarly, an index λ_s was computed from the rows of the table to express the probability of having two individuals of a given botanical species, s, being recorded under the same vernacular name.

2.3. Ecological descriptors

Each sampling unit was characterised with a set of ecological descriptors: (i) the topographical position, coded in seven classes along a sequence leading from ridges (1) to bottomlands (7); (ii) the average slope angle; (iii) traces on the ground of an ancient lateritic cuirass, coded in two classes (absence = 1 or presence = 2) of either blocks ≥20 cm, stones between 20 and 2 cm, or gravel between 2 cm and 2 mm; (iv) traces of prolonged periods of flooding (dark-brown soil, existence of respiratory roots) as one binary variable (absence = 1 or presence = 2); (v) the frequency (absent = 1, rare = 2 or abundant = 3) of two perennial herbs of the Rapateaceae family: *Rapatea* sp., used as an indicator of a water-table at a depth less than 50 cm during the dry season (Gonzalès and Ferry, 1998) and *Spathanthus* sp. used as an indicator of thinned soils (Ferry, personal communication), often linked to an impeded vertical drainage (Sabatier et al., 1997). All ecological variables were recorded with regard to modal plot condition, i.e. ignoring potential intra-plot spatial variation (see Section 2.1).

To assess the relevance of the ecological descriptors and to avoid using redundant variables in analyses, we first performed a normalised principal component analysis, considering all the descriptors as semi-quantitative variables, including the slope angle coded as, 1: slope $< 2^\circ$ ($N = 163$); 2: slope between 2 and 15° ($N = 121$); 3: slope $\geq 15^\circ$ ($N = 127$). The two principal axes of this analysis accounted for 61.7% of the total variance of the table. All the variables were well represented by these two axes (Fig. 1a), which shows positive correlations, on one hand between the presence of ground traces of an ancient lateritic cuirass (blocks, stones and gravel) and high slope angles (Pearson's correlation coefficients: $0.25 < r < 0.44$) and, on the other hand, between traces of surface flooding, the abundance of *Rapatea* sp. and the bottomlands ($r > 0.52$). These two sets of variables were negatively correlated along axis 1, while the abundance of *Spathanthus* sp. was positively correlated with axis 2 ($r = 0.73$). *Rapatea* sp. and *Spathanthus* sp. were frequently abundant in the same plots at the bottom of the topographical sequence. Nevertheless, the latter species was also present in some sloping plots and even on the hilltops (Fig. 1b and c). These species were thus more sensitive indicators of an impeded drainage than the traces of surface flooding, mainly observed in thalwegs. The codification of the topographical variable was thus refined by the introduction of the presence of *Spathanthus* sp. as an indicator of slightly hydromorphic soil conditions on upper-slopes and hilltops, and the presence of *Rapatea* sp. as an indicator of hydromorphic soil conditions on foot-slopes and bottomlands (Table 1). In the following, we will use the general term of topography for this new synthetic variable coded in 12 classes. The map of the distribution of the topographical classes (not shown) exhibited a concentration of plots belonging to classes 70 and 71 (large thalwegs) in the alluvial valley of the Counamama river, near the south-eastern boundary of the forest.

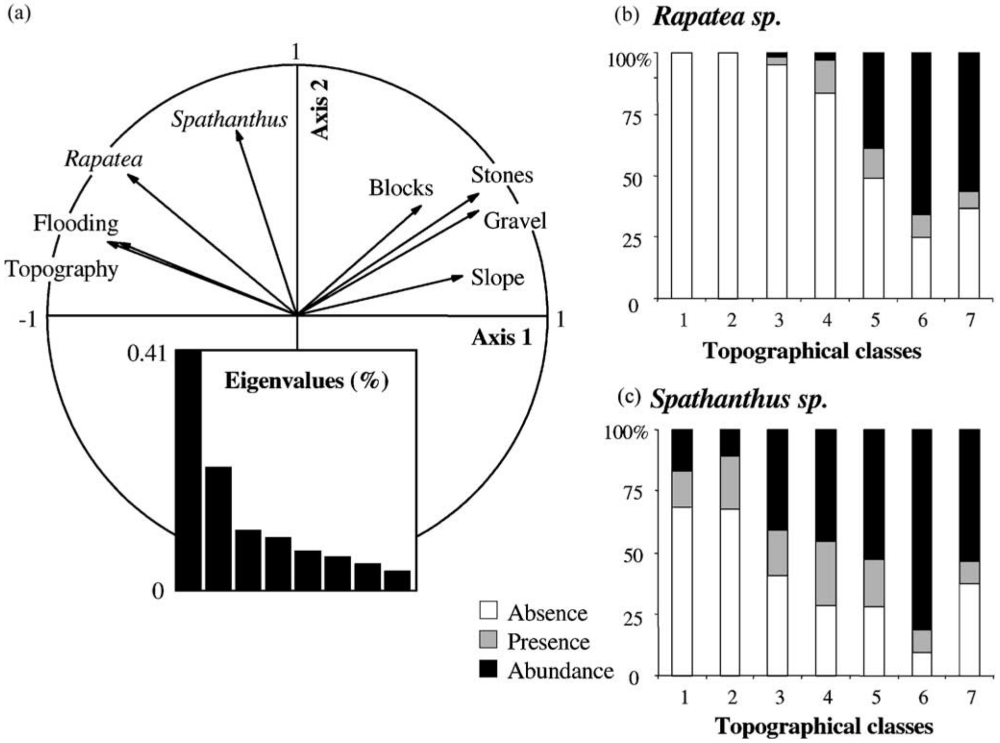

Fig. 1. Analysis of the ecological variables: (a) PCA correlation circle; (b) and (c) frequencies among the topographical classes (coded as in Table 1) *Rapatea* sp. and of *Spathanthus* sp. abundance classes.

Table 1
Codification of the topographical variable, taking into account initial topographical classes and the presence of *Spathanthus* sp. and *Rapatea* sp., as indicators of more or less hydromorphic soil conditions

Initial code	*N*	*Spathanthus*	*Rapatea*	Final code	*N*
1-Flat hilltop	48	0		10-Non-hydromorphic flat hilltop	33
		1		21-Slightly-hydromorphic hilltop	19
2-Ridge	28	1			
		0		20-Non-hydromophic ridge	25
3-Upper-slope	61	1		31-Slightly-hydromorphic upper-slope	25
		0		30-Non-hydromorphic upper-slope	36
4-Middle-slope	73	1		41-Slightly-hydromorphic middle slope	21
		0		40-Non-hydromorphic middle-slope	52
5-Foot-slope	57		1	51-Hydromorphic foot-slope	26
			0	50-Slightly-hydromorphic foot-slope	31
6-Narrow thalweg	32		1	61-Hydromorphic narrow thalweg	24
			0	70-Non-hydromorphic thalweg	49
7-Large thalweg	112		0		
			1	71-Hydromorphic large thalweg	71

The number of plots per class is indicated by *N*.

[a] 0: Absence; 1: presence.

2.4. Stand structure

We used the frequency distribution of diameter classes within the plots as a descriptor of local stand structure. In order to avoid classes of minimal frequency, we defined 14 classes, of 5 cm in range, from 7.5 to 37.5 cm dbh, and of 10 cm in range from 37.5 to 107.5 cm dbh, and set the last class as having all the trees ≥107.5 cm dbh.

We then performed a correspondence analysis (CA; see Section 3.1) on the contingency table, partitioning the trees from the 411 sampling plots into the 14 dbh classes. Note that having some dbh classes recorded on larger areas than others has no effect on CA results. The structure of the table was mainly captured by the first axis (Fig. 2a), which accounted for about 14% of the total variance of the table, and contrasted the plots relatively rich in large trees (classes 7–14; dbh $\geq$ 37.5 cm) to those rich in small trees (classes 1–6; dbh $<$ 37.5 cm). Mapping plot coordinates on axis 1 did not reveal any particular geographical partition of the forest, but a mosaic of neighbouring plots with similar diameter distributions (Fig. 2b). Mapping the basal area of the plots also showed no clear large-scale pattern within the forest (Fig. 2c).

In order to test whether the diameter distribution was linked to the refined topographical variable, we performed several one-way ANOVAs, taking plot coordinates on axis 1 and plot basal areas, both having nearly normal frequency distributions, as dependent variables. Basal area appeared more linked to the topographical variable ($F = 2.944$, d.f. $= 11$, $P = 0.001$) than did the diameter distribution ($F = 1.916$, d.f. $= 11$, $P = 0.036$). For the diameter distribution, pairwise comparisons using Tukey–Kramer HSD (Sokal and Rohlf, 1995) revealed only slightly significant differences between the mean plot

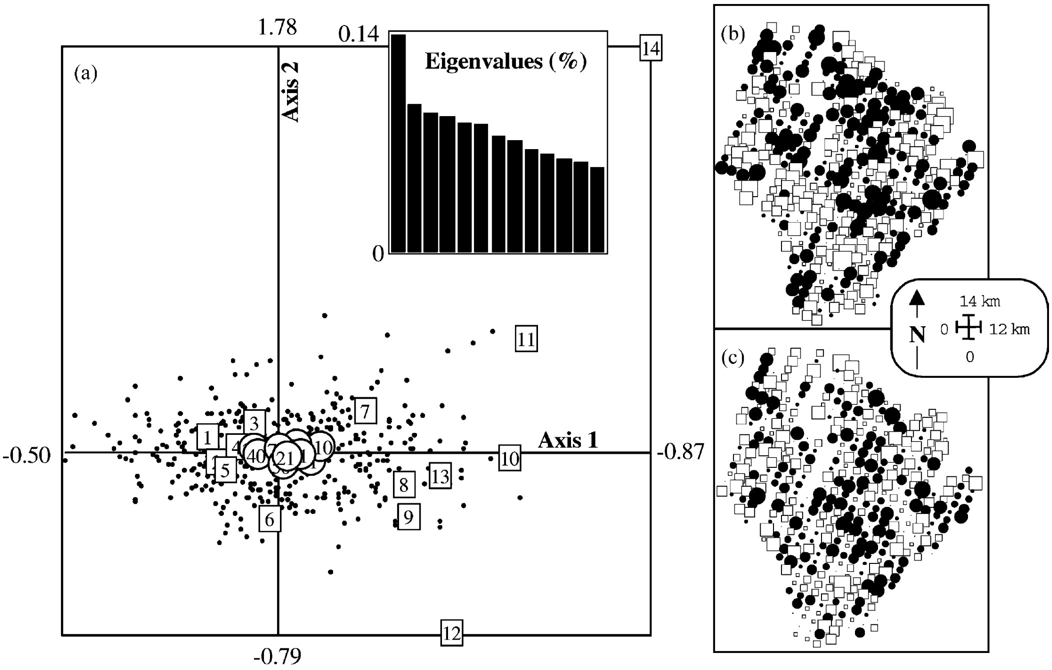

Fig. 2. Analysis of the stand structure (14 dbh classes): (a) distribution of plots in the main CA plane. Squared and encircled numbers indicate the mean position of the dbh and topographical classes, respectively. The dbh classes are of 5 cm width from 1 (7.5–12.5 cm) to 6 (32.5–37.5 cm), of 10 cm width from 7 (37.5–47.5 cm) to 13 (97.5–107.5 cm). Class 14 groups together all trees over 107.5 cm dbh. The topographical classes are coded as in Table 1; (b) forest map with indication of plot coordinates on axis 1; (c) forest map with mention of the plot basal areas (centred). In (b) and (c) the squares and circles indicate negative and positive values, respectively, while the size of the symbol is proportional to the departure from zero.

coordinates of some situations devoid of hydromorphic soil conditions, i.e. the flat hilltops (coded 10) and the upper-slopes (coded 30; $P = 0.044$) on one hand, and the middle-slopes (coded 40; $P = 0.034$) on the other hand. There was thus limited correlation between the distribution in dbh classes and the topographical context. In subsequent analyses, these variables will be, hence, considered as complementary ecological representations to be compared with the floristic composition of the stand.

3. Methods of multivariate analysis

3.1. Analysis of the plot by species (floristic) table

Correspondence analysis (CA Hill, 1973) is a widely used method to study contingency tables crossing taxonomic units (e.g. species) and sampling units (plots). CA produces a simultaneous ordination of the sites and species, enabling an indirect gradient analysis (Hill, 1973). To do so, CA summarises the contingency table in terms of departures from the null hypothesis of independence between rows and columns, i.e. random distribution of species in plots, and the sum of the eigenvalues that are yielded by CA is proportional to the total χ^2 of the contingency table. In the following, this quantity will be referred to as variance (sensu lato), since it expresses the total departure from an expected result (null hypothesis). In a given cell of the table, corresponding to the ith row and the jth column, the departure from independence is measured by $z_{ij} = (p_{ij} - p_{i\cdot}p_{\cdot j})/p_{i\cdot}p_{\cdot j}$, where p_{ij} is the relative frequency of the cell, and $p_{i\cdot}$ and $p_{\cdot j}$ the marginal relative frequencies of row i and column j, respectively. Synthetically, CA can be viewed as the analysis of the statistical triplet $(\boldsymbol{Z}, \boldsymbol{D_J}, \boldsymbol{D_I})$, where $\boldsymbol{Z}$ is the table of general term z_{ij}, and $\boldsymbol{D_I}$ and $\boldsymbol{D_J}$ are the diagonal matrices containing the marginal weights $p_{i\cdot}$ and $p_{\cdot j}$, respectively (see Dolédec et al., 1996 or Dolédec et al., 2000 for an introduction to triplet notation in ecological analyses). CA positions the sampling plots along an ordination axis in a way that maximises the separation between species distributions (optimal between-species discrimination). Symmetrically, the species positions maximises the variance of the sampling plots (optimal between-sites discrimination). Particular correspondences between poor plots and scarce species often strongly influence the results of CA (Ter Braak, 1987; Gimaret-Carpentier, 1999).

To avoid this problem, non-symmetric correspondence analysis (NSCA; Lauro and D'Ambra, 1984; Gimaret-Carpentier et al., 1998) offers an alternative standpoint to study a contingency table crossing species and plots. While plots and species play an equivalent role in CA, NSCA analyses either the plots or species profiles. Therefore, there are two different possible NSCAs for a given contingency table. In the present paper, we shall consider only NSCA on plots profiles, since we are mainly looking for a typology of plots based on their floristic composition. The corresponding analysis considers the statistical triplet $(\boldsymbol{L}, \mathbf{Id}_{\boldsymbol{J}}, \boldsymbol{D_I})$, where $\boldsymbol{L}$ is the table of general term $l_{ij} = p_{ij}/p_{i\cdot} - p_{\cdot j} = (p_{ij} - p_{i\cdot}p_{\cdot j})/p_{i\cdot}$, and $\mathbf{Id}_{\boldsymbol{J}}$ is the identity matrix of size J, the number of columns (species) of the contingency table. The l_{ij} measures, for each species, j, the departure between its mean relative frequency and its conditional frequency in a given plot i. As in CA, each plot is positioned at the mean position of the species it contains, but the mean plot position is not necessary at the origin. Species are not positioned at the mean plot position, because their position is weighted by their relative frequency, i.e. $p_{\cdot j}$.

CA, which is a compromise between NSCA on species profiles and NSCA on plots profiles, summarises the contingency table by gradients that are common for species and plots, to the detriment of particular structures related either to species or plots. A fundamental criterion to appreciate the relative relevance of the two methods is the reliability of the information conveyed by the absence of a given species in a particular plot (Chessel and Gimaret-Carpentier, 1998; Pélissier et al., 2002b). To consider the absence of a species as relevant as its presence, we must be sure that this species, if present, has inevitably been recorded (or had a high probability to have been recorded taking into account possible misidentifications or omissions). In such a case, NSCA has to be preferred because the closer to 0.5 is a species' relative frequency, the stronger is its weight in the analysis. Otherwise, presence is the sole relevant information and CA should be preferred (Pélissier et al., 2002b).

During the inventory of Counami forest, all the trees have been thoroughly recorded within the plots. The risk of omission was thus minimal and, therefore, the absence of a well-known species is relevant information that militates in favour of using NSCA. But the absence of a badly identified species is difficult information to interpret, arguing in favour of CA. Analysing a data set characterised by a rigorous counting and a certain taxonomic imprecision thus justifies the use of the two methods, whose results could be profitably compared.

3.2. Coupling the floristic analysis with environmental variables

The contingency table containing the floristic information can be coupled with an external table containing environmental variables in the general framework of principal component analysis on instrumental variables (PCAIV), also known as redundancy analysis (Rao, 1964; Sabatier et al., 1989). Such an approach provides ordination axes that directly visualise patterns of floristic variation along explicit environmental variables (direct gradient analysis). For a plot by species table analysed through CA, the method is known as canonical correspondence analysis (CCA; Ter Braak, 1987) though correspondence analysis on instrumental variables may have been preferable (CAIV; Lebreton et al., 1988).

CAIV is defined from the CA statistical triplet as ($\hat{\boldsymbol{Z}}$, $\boldsymbol{D_J}, \boldsymbol{D_I}$), where $\hat{\boldsymbol{Z}}$ is the approximation of $\boldsymbol{Z}$ by the table $\boldsymbol{X}$ of the explanatory variables. This is equivalent to the projection of $\boldsymbol{Z}$ onto the sub-space engendered by the linear combinations of the variables in $\boldsymbol{X}$, which can be viewed as a sub-space of synthetic environmental variables. CAIV has all the properties of CA, as presented in Section 3.1, and can be viewed as a form of discriminant analysis (Lebreton et al., 1988).

The general framework of instrumental variables also applies to the approach of the contingency table through NSCA, and the present paper will give an illustration of this possibility which, as far as we know, seems to have been unemployed before. We shall call non-symmetric correspondence analysis on instrumental variables (NSCAIV), the analysis of the statistical triplet ($\hat{\boldsymbol{L}}$, $\mathbf{Id}_{\boldsymbol{J}}$, $\boldsymbol{D_I}$), where $\hat{\boldsymbol{L}}$ is approximated by the linear combinations of the variables in $\boldsymbol{X}$.

Using either CAIV and NSCAIV, we can also analyse the residual table $\boldsymbol{W} = \boldsymbol{Z} - \hat{\boldsymbol{Z}}$, $\boldsymbol{W} = \boldsymbol{L} - \hat{\boldsymbol{L}}$ respectively) via an orthogonal analysis from the statistical triplet ($\boldsymbol{W}, \boldsymbol{D_J}, \boldsymbol{D_I}$) or ($\boldsymbol{W}, \mathbf{Id}_{\boldsymbol{J}}, \boldsymbol{D_I}$) respectively. This enables to know which are the floristic patterns that are not accounted for by the environmental variables in $\boldsymbol{X}$. Furthermore, each of the residual tables can be compared with another explanatory table, say $\boldsymbol{X_2}$, of additional environmental variables. This enables exploration of the effect of $\boldsymbol{X_2}$, after the effect of $\boldsymbol{X}$ has been eliminated (partial analyses; Ter Braak, 1988; Sabatier et al., 1989).

For any approach based on instrumental variables, statistical significance of the portion of initial variance that is captured by the approximated table can be tested, using the Monte Carlo method (Manly, 1991; Fraile et al., 1993). The null hypothesis—no relation between the floristic table and the environmental table $\boldsymbol{X}$—is tested by computing the total variance of the approximated table, after a random permutation of the rows of $\boldsymbol{X}$. The process is iterated m times and the proportion of permutations yielding an approximated variance above the observed one gives the probability of the null hypothesis. We took $m = 1000$ permutations to conduct the tests in this paper. All analyses have been performed using the ADE-4 software (Thioulouse et al., 1997).

4. Results

4.1. Identification of reliable botanical species

The computation of the homogeneity index λ_v identified 66 vernacular categories with $\lambda_v = 1$; 71 with $\lambda_v > 0.75$; 79 with $\lambda_v > 0.66$; and 92 with $\lambda_v > 0.5$. We then established the correspondences with the most abundant species of each vernacular category. Eight species appeared to be dominant in two different vernacular categories, which have been merged after verification that the merging increased the homogeneity index, λ_s, for the most abundant species in the new vernacular category. After pooling, we obtained 46 correspondences with both λ_v and λ_s equalling 1; 53 with λ_v and λ_s above 0.75; 59 with λ_v and λ_s above 0.66; and 70 with λ_v and λ_s above 0.5. Using *Eperua falcata*, one of the most abundant and easily identified species in this forest as a reference, we considered the 59 correspondences displaying values above 0.66 for both homogeneity indices λ_v and λ_s as reliable. This list has subsequently been checked with the help of experienced botanists (D. Sabatier and one of us, J.-F. Molino).

One correspondence was discarded because we suspected that its vernacular name referred to several congeneric species that are difficult to distinguish in the field. One new correspondence was added to the list for a non-tested palm species (*Euterpe oleracea*), which can be unambiguously identified by the field workers. In the following, we shall call the 59 categories designated by their botanical equivalents, botanical species, and the remaining 225, categories whose botanical equivalence is uncertain vernacular species. The word species will be used to indicate either a botanical or a vernacular species.

4.2. Ordinations from the floristic tables

We performed both CA and NSCA on the reduced floristic table of 411 plots by 59 botanical species, and on the complete floristic table of 411 plots by 59 botanical + 225 vernacular species. In both CAs, the first eigenvalues showed a regular decreasing pattern and accounted for a low proportion of total variance (χ^2) of the corresponding floristic table: 7.7% for axis 1 and 26.7% for the first five axes from the reduced floristic table; 3.6% for axis 1 and 12.7% for the five main axes from the complete floristic table. In both these analyses, all the species determining the main axes had low frequencies (Fig. 3a), a fact that is a frequent outcome of CA. In spite of this, axes 2 and 3 of the reduced floristic table expressed geographical partitions of the forest (Fig. 3b and c) that were homologous to the patterns displayed by axes 1 and 4 of the complete floristic table (not presented). The main pattern was created by plots situated near the south-eastern boundary of the forest, in the bottomlands fringing the alluvial valley of the Counamama river (Fig. 3b). These plots were marked by their richness in species adapted to soil water saturation (e.g. *E. oleracea* $N = 142$; *Symphonia globulifera* $N = 342$; *Jessenia bataua* $N = 70$). Some plots in the Counamama valley were also distinguished by axis 3, although this axis expressed mainly an opposition between plots situated on both sides of a diagonal going from north-west to south-east (Fig. 3c), and which had, at this stage, no clear ecological interpretation.

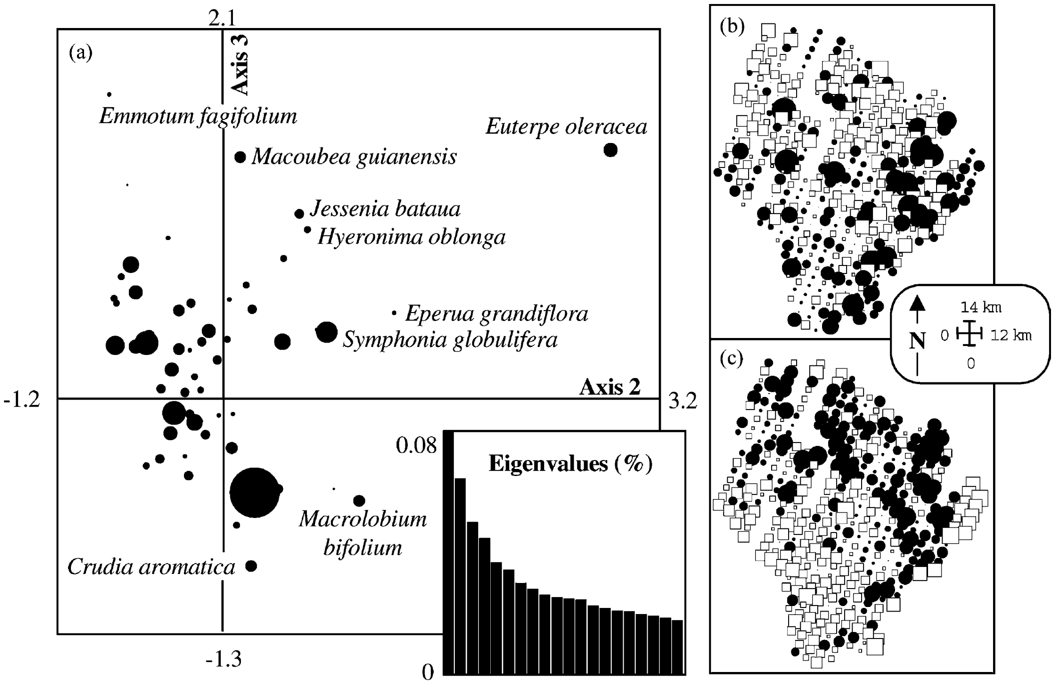

Fig. 3. CA of the reduced floristic table (59 botanical species): (a) plane of axes 2 and 3 for species (circles are proportional to species frequencies); (b) and (c) forest maps with indication of plot coordinates on axes 2 and 3, respectively (square: negative coordinate; circle: positive coordinate; the size of the symbol expresses the distance to the origin).

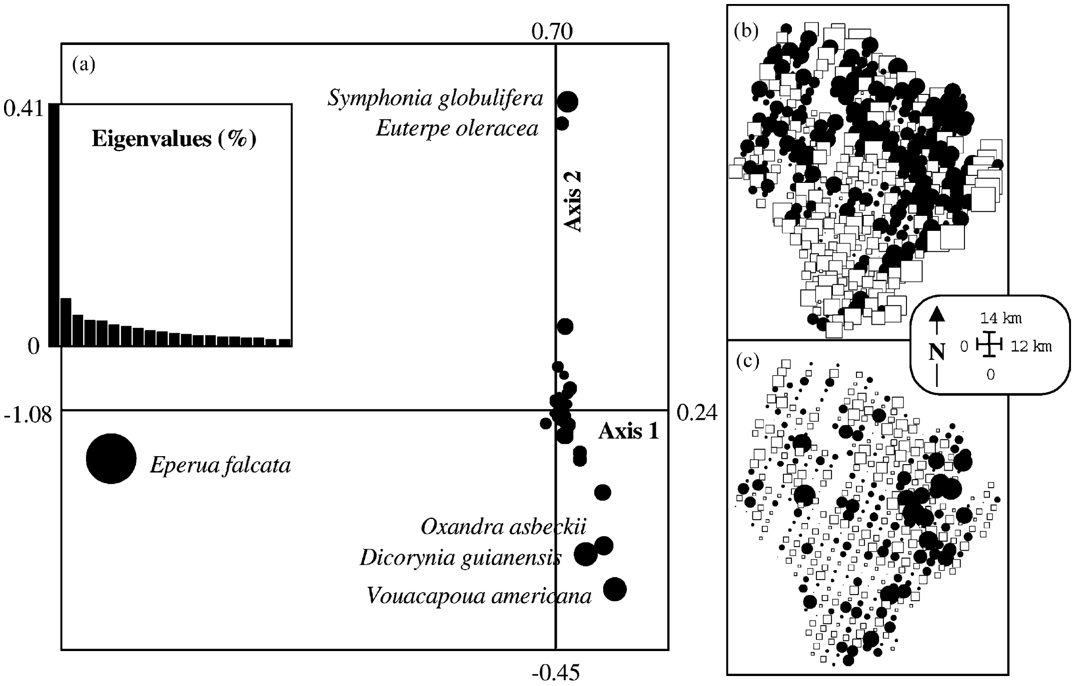

Fig. 4. NSCA of the reduced floristic table (59 botanical species); (a) plane of axes 1 and 2 for species (circles are proportional to species frequencies); (b) and (c) forest maps with mention of plot coordinates on axes 1 and 2, respectively (square: negative coordinate; circle: positive coordinate; the size of the symbol expresses the distance to the origin).

Both NSCAs exhibited a prominent first axis, representing 40.7% (48.6% for the first two axes) and 17.9% (41.6% for the first four axes) of total variance of the reduced and complete floristic tables, respectively. In NSCAs, the main axes are mainly determined by abundant species: axis 1 was largely dominated by *E. falcata* ($N = 2187$) in both analyses (Fig. 4a). In NSCA of the reduced floristic table, axis 2 showed a gradient from species confined to the upper-slopes and hilltops (*Vouacapoua americana* $N = 414$; *Dicorynia guianensis* $N = 421$; *Oxandra asbeckii* $N = 268$) to species restricted to the bottomlands (*S. globulifera* $N = 342$; *E. oleracea* $N = 142$) (Fig. 4a). In NSCA of the complete floristic table, axes 2 and 3 were mainly determined by two frequent vernacular species: Maho noir ($N = 2392$) and Maho rouge ($N = 1539$), and a topography-related gradient of species appeared only on axis 4. However, for both NSCAs, mapping plot coordinates on axis 1 yielded a forest partition (Fig. 4b) which, though similar to CAs results, relied on the abundance of the most frequent species (*E. falcata*) instead of the presence of scarcer species. Similarly, axis 2 of the reduced floristic table and axis 4 of the complete floristic table exhibited a geographical partition homologous to the one displayed by CAs (Fig. 3c), and isolating the south-eastern part of the forest (Fig. 4c).

4.3. Explanatory power of the ecological variables

The explanatory power of the topography and the stand structure were evaluated by comparing total variance of the floristic table with variance after approximation by the tables containing the explanatory variables (Table 2). In doing so, we wanted to know to which extent the floristic gradients revealed by CAs and NSCAs could be accounted for by topography and/or stand structure. In all analyses, topography had a better explanatory power (ranging from 5.60 to 11.07% of variance of the floristic table), than the stand structure (ranging from 3.87 to 6.57% of variance of the floristic table). Taking into account both sets of variables accounted for between 9.33 and 17.54% of total variance of the floristic table.

Table 2
Percentages of total variance of the floristic tables explained by the ecological variables

	I_{topo}	I_{dbh}	$I_{\text{topo/dbh}}$	$I_{\text{dbh/topo}}$	$I_{\text{topo+dbh}}$	Residual
CA (complete table)	5.60***	3.87***	5.45***	3.72**	9.33***	90.67
CA (reduced table)	6.52***	4.03*	6.29***	3.80	10.32***	89.68
NSCA (complete table)	8.54***	5.66***	8.39***	5.51***	14.05***	85.95
NSCA (reduced table)	11.07***	6.57***	10.96***	6.46***	17.54***	82.46

I_{topo} and I_{dbh} correspond to the variance after approximation by the topographical and stand structure variables, respectively. $I_{\text{topo/dbh}}$ and $I_{\text{dbh/topo}}$ correspond to the variance after approximation by the topographical (stand structure) variable, once the effect of the other variable has been eliminated. $I_{\text{topo+dbh}}$ corresponds to the approximated variance on the complete table of both ecological variables. Residuals are computed as $100 - I_{\text{topo+dbh}}$. Notice that in all cases: $I_{\text{topo+dbh}} = I_{\text{topo}} + I_{\text{dbh/topo}} = I_{\text{dbh}} + I_{\text{topo/dbh}}$, while $I_{\text{topo+dbh}} - (I_{\text{topo}} + I_{\text{dbh}})$ quantifies to which extent the two variables may be redundant. Levels of statistical significance of the permutation tests: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

This proportion increased when the vernacular species were excluded (reduced floristic table), or when the low frequency species played a less important role in variance definition (NSCAs instead of CAs).

Another interesting point was that the proportion of variance of the floristic table explained by the stand structure after the effect of topography has been eliminated ($I_{\text{dbh/topo}}$)—or conversely explained by the topography once the effect of stand structure has been removed ($I_{\text{topo/dbh}}$)—remained important, representing about 40%–60%, respectively—of total variance jointly explained by the two sets of variables. This means that topography and stand structure were non-redundant for predicting floristic composition.

Notice that the proportion of explained variance for a given species is exactly the same in CA and NSCA: only the corresponding proportion for the whole set of species, and thus the total variance, differs between the two analyses (see Section 3.1). Fig. 5a shows that some species were well explained by the topographical variable (*E. oleracea*, *O. asbeckii*, *S. globulifera*, *D. guianensis*) or the stand structure variable (*Hevea guianensis*, *Sextonia rubra*, *Qualea rosea*), while species well explained by both variables proved rare (*E. falcata*, *V. americana*). This illustrates the fact that topography and stand structure were largely non-redundant. In Fig. 5b, the frequent vernacular species Maho rouge and Maho noir, which mainly determined NSCA axes 2 and 3, appeared badly explained by both explanatory variables, probably because these categories refer to several botanical species that behave differently with regard to ecological factors. For instance, in the nearby forest of Piste de St. Elie, two species (*Eschweilera coriacea* and *E. micrantha*) that belong to the broad categories of Mahos displayed complementary patterns of soil affinities (Sabatier et al., 1997). On the other hand, some other vernacular species, although potentially composite, were well explained by the topographical variable: Mutusi marécage, Tosso passa marécage, Maho cochon, Koko, Yayamadou marécage. Among them, only the last has not been sufficiently collected to be tested for species homogeneity, though being probably homogeneous (Sabatier, personal communication).

4.4. *Ordinations from the approximated tables*

We analysed the approximation of the floristic table by both topographic and stand structure tables in order to depict floristic gradients directly linked to these two ecological variables. CAIV conducted on the reduced floristic table showed a prominent first axis (31.1% of total variance of the approximated table) expressing a gradient in accordance with the topographical sequence (Fig. 6a): classes related to thalwegs (61, 70 and 71) were found in the positive part of axis 1, while the hilltops and upper-slopes (10–41) were grouped on the negative side of axis 1. Axes 2 and 3 (12.41 and 10.73%, respectively, of total variance of the approximated table) were mainly determined by the dbh classes (Fig. 6b): axis 2 showed an opposition between intermediate classes (5–9: 27.5–67.5 cm dbh) and the other dbh classes, and axis 3 between small (1–5: 7.5–32.5 cm dbh) and large (9–14: over 57.5 cm dbh) trees. As in the initial CAs, the species displayed at the ends of axes were the less abundant ones. Furthermore, mapping ordination scores of plots did not reveal any clear geographical partition of the

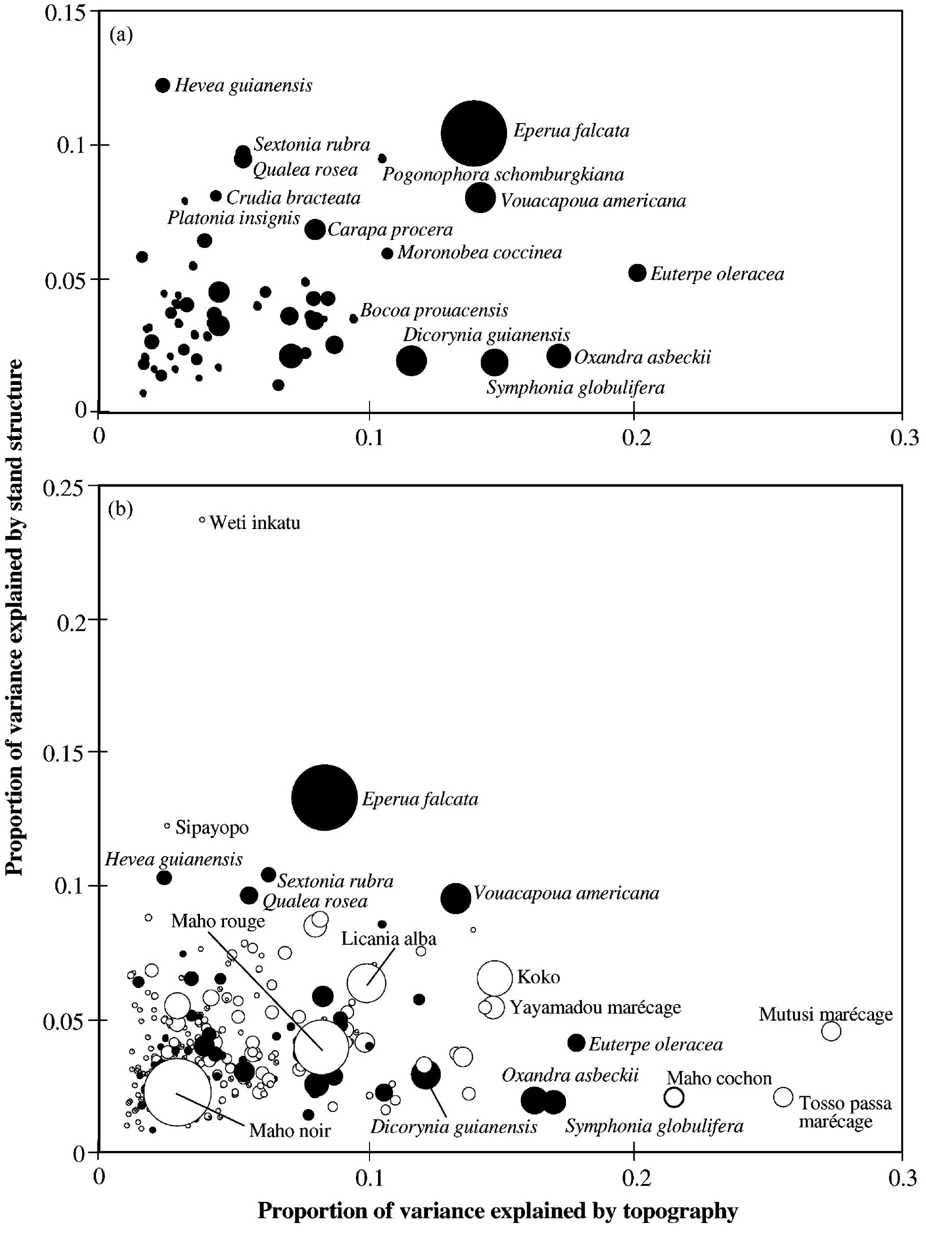

Fig. 5. Proportion of species variance explained by the topographical and stand structure variables. (a) Reduced floristic table: 59 botanical species. (b) Complete floristic table: 59 botanical species (black) + 225 vernacular species (white). Circles are proportional to species frequencies.

forest, but mainly clusters of similar neighbouring plots (not shown). CAIV of the complete floristic table exhibited the same topographical gradient along axis 1 (24.91% of total variance of the approximated table). However, axis 2 of this analysis (9.89% of total variance) was largely determined by the topographical class 70 (thalweg with non-hydromorphic soils) corresponding mostly to situations in the alluvial valley of the Counamama that are sufficiently far away from the river itself. These plots stood as having very scarce

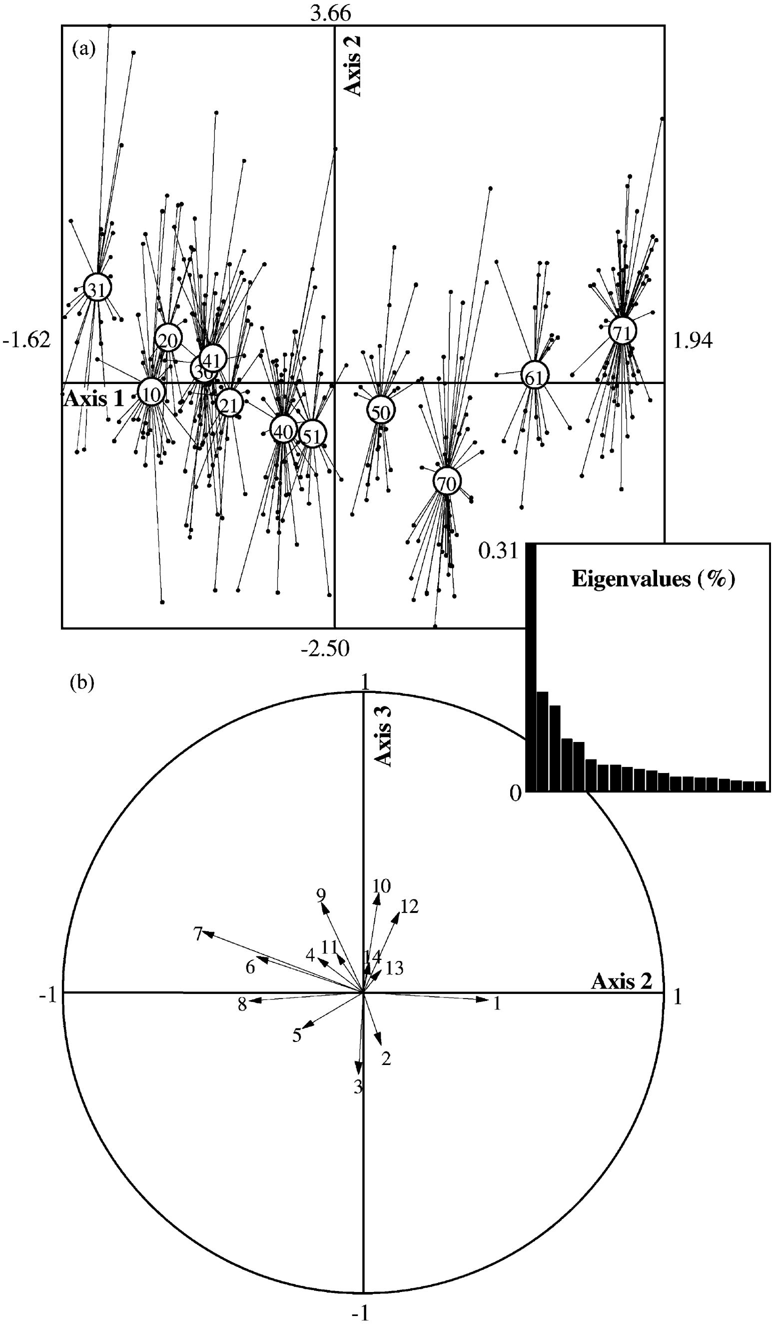

Fig. 6. CAIV of the reduced floristic table (59 botanical species) coupled with 12 topographical variables and 14 dbh classes. (a) Plane of axes 1 and 2: encircled numbers represent the mean position of the topographical classes (coded as in Table 1). (b) Correlation of the dbh classes with the axes 2 and 3 (dbh classes are defined as in Fig. 2).

vernacular species (Sipayopo $N = 1$; Chawari rivière $N = 7$; Anaoula $N = 18$). The dbh classes in this analysis were mainly correlated with axes 3 and 4 (7.94 and 6.64% of total variance of the approximated table, respectively).

In NSCAIV conducted on the reduced floristic table, axes 1 and 2 accounted for 61.36 and 17.46% of total variance of the approximated table. The topographical gradient appeared in diagonal in the plane determined by the two axes (Fig. 7a), while a direction perpendicular to this gradient opposed medium-sized trees (classes 5–9) vs. small and large trees (Fig. 7b). The very abundant species *E. falcata* ($N = 2179$) displayed the largest correlation with both axes. This first plane illustrated, once again, the fact that topography and stand structure behaved as independent explanatory variables. In fact, in most topographical situations, a clear gradient in abundance of *E. falcata* was observed in relation to the relative abundance of dbh classes 6–8 (Fig. 7c). Because the most abundant vernacular species (Maho noir $N = 2392$ and Maho rouge $N = 1539$) were composite and badly explained by both ecological variables (see Fig. 5b), they had little influence on the NSCAIV conducted on the complete floristic table. Consequently, the results were similar to the ones obtained from the reduced table (Fig. 7).

4.5. Ordinations from the residual tables

Floristic patterns that could have remained unexplained by the topography and stand structure variables have been investigated via multivariate analyses of the residual tables (orthogonal analyses). Residual tables are obtained from the initial floristic tables after subtraction of their approximations. Orthogonal CAIV conducted on the reduced floristic table showed a prominent first axis, accounting for 8.06% of total variance of the residual table (Fig. 8a), and which expressed a pattern due to the presence of scarce species in particular plots that was already apparent via axis 1 of the initial CA (see Section 4.1). This means that this pattern, which did not determine any clear geographical partition of the forest, was badly explained by the ecological variables. On the contrary, the most significant geographical pattern that isolated the Counamama valley along axis 2 in initial CA (see Fig. 3b), was no longer apparent in the orthogonal CAIV. Hence, although portions of explained variance in Table 2 may have appeared low, the ecological variables were able to account for a large part of this important floristic structure.

Maps of axes 2 and 3 in orthogonal CAIV (Fig. 8b and c), pointed towards a north-west/south-east opposition as the main residual structure. It corresponded to the geographical pattern that appeared along axis 3 in initial CA (see Fig. 3c), and which remained insufficiently explained by both the topographical and stand structure variables. Similar conclusions were reached from the orthogonal CAIV of the complete floristic table (results not presented).

In orthogonal NSCAIV of the reduced floristic table, axis 1 (37.09% of total variance of the residual table) also stood out from subsequent axes (Fig. 8d). As in initial NSCA and NSCAIV, this axis was mainly dominated by *E. falcata*. However, mapping coordinates of plots along axis 1 of the orthogonal NSCAIV showed that the north-west/south-east pattern displayed along axis 1 in the initial NSCA (see Fig. 4b) has been largely attenuated (Fig. 8e). This indicated that the forest partition related to the distribution of *E. falcata*, was fairly explained by the joint effect of topography and stand structure. Hence, these variables appeared more efficient to explain spatial variations of the most abundant species, than the homologous geographical pattern based on an association of less frequent species. Axis 2 of the same analysis (6.30% of variance of the residual table) was dominated by *S. globulifera* and *E. oleracea*, as in initial NSCA and NSCAIV. However, the strong residual pattern appearing on this axis (Fig. 8f) reveals that the local singularities expressed along axis 2 in the initial NSCA (see Fig. 4c), have been insufficiently explained by both ecological variables.

In orthogonal NSCAIV conducted on the complete floristic table, the first three axes accounted for 16.28, 9.71 and 8.25%, respectively, of total variance of the residual table. As in initial NSCAIV, *E. falcata* was highly correlated with the residual axis 1, while the vernacular species Maho noir and Maho rouge, were highly correlated with axes 2 and 3. Geographical maps of plot coordinates on these axes (not shown), indicated that both vernacular species, which were badly explained by the ecological variables (see Fig. 5b), had important residual structures. Maho noir and Maho rouge formed large unexplained clusters in

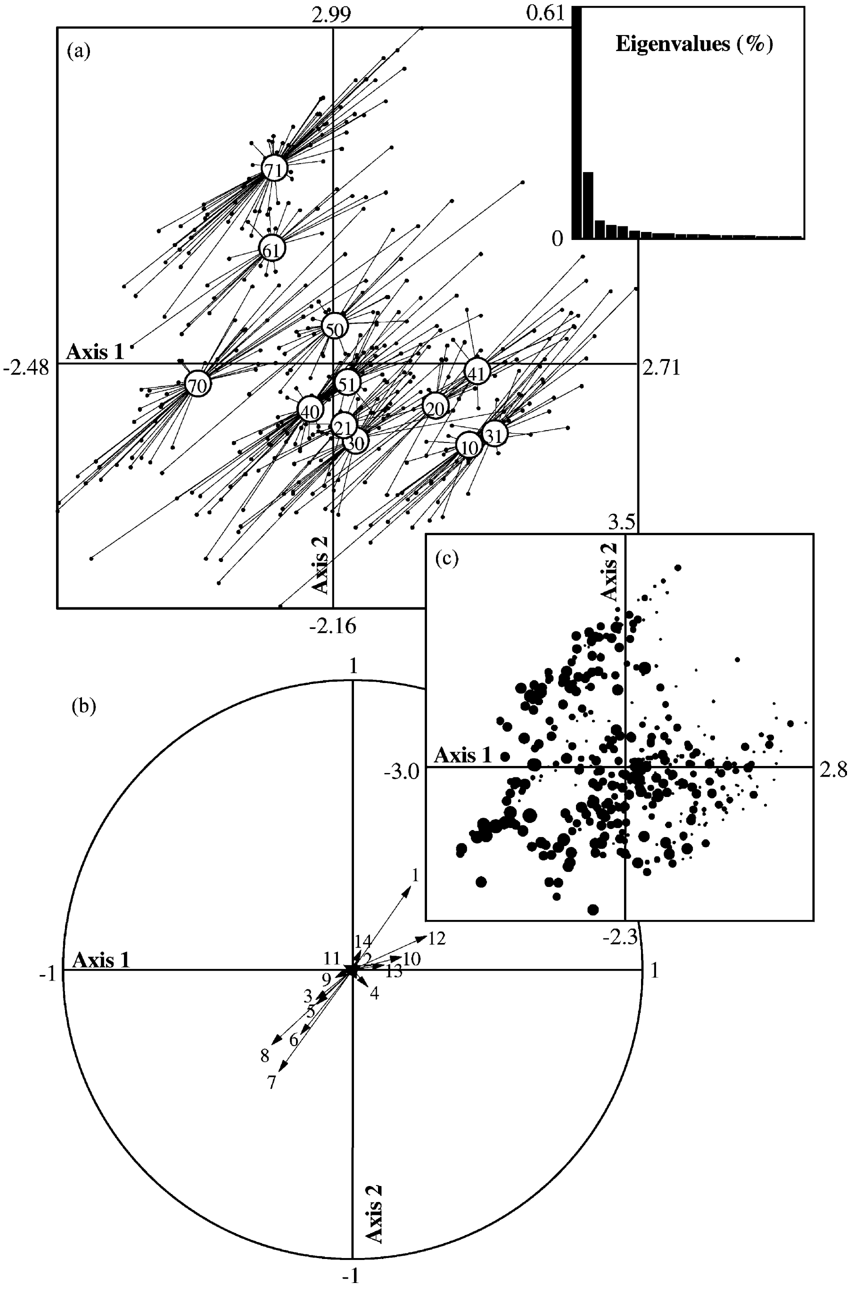

Fig. 7. NSCAIV of the reduced floristic table (59 botanical species) coupled with 12 topographical variables and 14 dbh classes. (a) Plane of axes 1 and 2: encircled numbers represent the mean position of the topographical classes (coded as in Table 1). (b) Correlation of the dbh classes with the axes 1 and 2 (dbh classes are defined as in Fig. 2). (c) Plane of axes 1 and 2 with mention of the frequency of *Eperua falcata*.

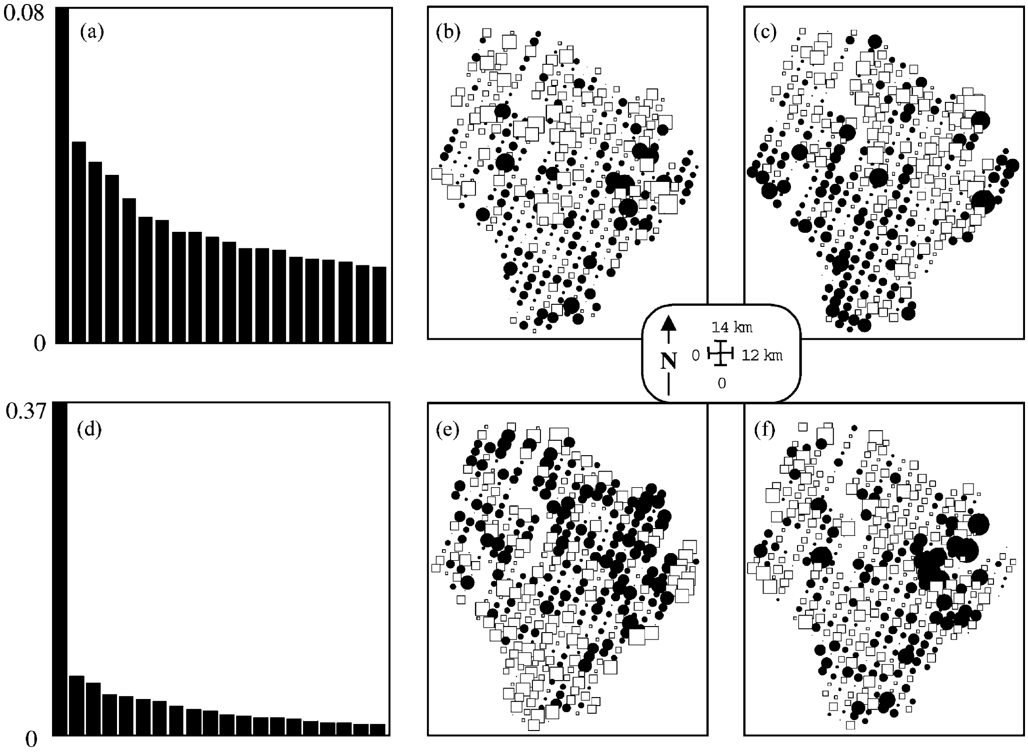

Fig. 8. Orthogonal analyses of the reduced floristic table (59 botanical species): (a–c) orthogonal CAIV; (d–f) orthogonal NSCAIV. (a) and (d) eigenvalues of orthogonal analyses (%); (b) and (c) forest maps with mention of plot coordinates on axes 2 and 3, respectively, of orthogonal CAIV; (e) and (f) forest maps with mention of plot coordinates on axes 1 and 2, respectively, of orthogonal NSCAIV. In (b), (c), (e) and (f): squares: negative coordinate; circles: positive coordinate; the size of the symbol expresses the distance to the origin.

the northern third (residual axis 2), and the southern two-third of the forest (residual axis 3), respectively.

5. Discussion

5.1. *Evidence of robust ecological patterns*

Analysis of the floristic data of Counami forest allowed the detection of three main mesoscale floristic units, that can be considered as robust with respect to the method of ordination and to the set of species under consideration. Indeed, they were always revealed among the four main axes of the initial analyses dealing with the floristic tables. (i) The first unit corresponded to the south-western forest boundary, situated on the alluvial terrain of the Counamama river, which has been identified on the basis of its richness in the most common species adapted to periodic soil water saturation, mainly *E. oleracea* and *S. globulifera* (NSCA of the reduced floristic table; see Fig. 4), in mixture with less common species, like *J. bataua*, *Eperua grandiflora* or *Macrolobium bifolium* (CA of the reduced floristic table; see Fig. 3). Some relatively important vernacular species, though composite, were also found associated to this floristic cortège: mainly Mutusi marécage, Tosso passa marécage and Yayamadou marécage (NSCA of the complete floristic table). This floristic entity is also characterised by the scarcity of otherwise abundant species, such as *V. americana*, *D. guianensis* and *O. asbeckii* (NSCA of the reduced floristic table; see Fig. 4), which probably cannot survive prolonged periods of soil water saturation, and some common vernacular species, like Koko and Licania (NSCA of the complete floristic table). (ii) The second floristic unit corresponds to the south-western part of the forest, dominated by the very common *E. falcata*

(CA and NSCA of the reduced floristic table; see Figs. 3 and 4), which was locally found associated with the abundant vernacular species Maho rouge (NSCA of the complete floristic table). (iii) In opposition, the north-eastern part of the forest is mainly characterised by the rarity of *E. falcata* (NSCA of the reduced floristic table; see Fig. 4), locally compensated by the presence of large clusters of the common vernacular species Maho noir (NSCA of the complete floristic table). The contrast between these two floristic units was virtually not perceptible in the field, and its detection via both scarce and frequent species demonstrates that using complementary multivariate approaches—an option which is sometimes seen as excessive sophistication—can pay off.

On the contrary, the floristic uniqueness of the valley of the Counamama river was an expected result which is, furthermore, consistent with the results obtained for recent alluvial soils at a landscape scale in La Selva, Costa Rica (Clark et al., 1999). At Counami, however, it has been only partly explained by the refined topographical variable: the corresponding geographical pattern disappeared completely from the analysis of the residual tables by orthogonal CAIVs, but remained strongly apparent from orthogonal NSCAIVs (see Fig. 8). This residual pattern is mainly due to the presence of *S. globulifera*. This species, though appearing as relatively homogeneous from our preliminary testing ($\lambda_v = 0.8$ and $\lambda_s = 1$), is probably mixed up with a not-described vicariant species that grows on less hydromorphic soil conditions. Such a result illustrates the fact that quantitative analyses dealing with species distributions may be a valuable contribution to taxonomic issues.

5.2. Topography and stand structure as ecological descriptors

Insights provided by the topographical variable went beyond a trivial contrast of areas liable to occasional flooding versus those from non-flooded situations. Indeed, three of the topographical classes related to thalwegs (61, 70 and 71) were always clearly discriminated through their floristic composition. This conclusion parallels results from Peruvian Amazonia, which showed that a substantial proportion of tree species displayed marked preferences for particular landscape units, most of them defined within an alluvial valley from historical river dynamics (Pitman et al., 1999). On the other hand, there was no clear discrimination between topographical classes relating to slopes and hilltops (CAIVs and NSCAIVs; see Figs. 6 and 7), and the use of *Spathanthus* sp. as an indicator of thinned and hydromorphic soils proved to be inefficient, as an explanation for floristic variations. Nevertheless, it must be kept in mind that the sample plots were large, and often edaphically heterogeneous. In comparable forests in French Guiana, more detailed analyses linking individual trees to soil conditions demonstrated the existence of significant floristic gradients, according to soil types, even in non-flooded locations (Sabatier et al., 1997; Pélissier et al., 2002a). Similarly, at a landscape scale in La Selva, Costa Rica, Clark et al. (1998, 1999) found significant edaphic influences on species distributions among never-flooded soil units. Our topographical variable seems, nevertheless, to be excessively detailed in slopes and hilltops, while foot-slopes and thalwegs still require an effort at understanding and codification that could benefit from the work undertaken at the neighbouring forest station of Paracou (Gonzalès and Ferry, 1998). At Counami, it would be useful to elaborate the edaphic features behind the class “thalweg with non-hydromorphic soils” (coded 70) that was significantly influential in all analyses constrained by the topographical variable. This class, though probably internally heterogeneous, represents 10% of the sample plots, and the majority of the plots located in the valley of the Counamama river.

Both topography and stand structure explained a significant part of the variance of the initial floristic tables, and can be considered as fairly good predictors for the distribution of some species (Fig. 5). However, the explained portion of variance remained very low for most species (<5%), especially for the scarcer ones, and rather low for the whole tables (9–17%; Table 2). Clark et al. (1999) found, nevertheless, figures that were lower (1.5–2.2%), yet significant, when using soil types and topography as predictors for floristic composition of small plots at La Selva, Costa Rica. In mega-diverse tropical forests, one should expect most ecological descriptors to account for a limited part of floristic variation, due to the complexity of biotic interactions (Condit, 1996), to the presence of a large number of scarce species, and also to the fact that most non-random distributions correspond rather to differences in degree of occupancy than to complete restriction to particular ecological conditions (Clark et al., 1999; Pitman et al., 1999). Hence, diverse and complementary quantitative analyses, among which multivariate techniques, are needed to reveal and explain so a subtle pattern of floristic variation. Overall, beyond percentages of explained vs. unexplained variance, what really matters is the nature of floristic variation that is not accounted for by the ecological descriptors under consideration. This crucial question can be explicitly addressed via the analysis of the residual tables, which may contains either meaningful patterns of geographic variation (as on Fig. 8c and e; see also Pélissier et al., 2002a) or nothing more than a high amplitude noise, displaying no obvious pattern in both geographical and ordination spaces.

We have explicitly used the stand structure, described through the distribution of diameters, to explain floristic variation. This led to a result which has never been reported before, namely that stand structure and topography can be non-redundant predictors for floristic composition. The stand structure appeared pertinent to explaining, at least partially, the distribution of some abundant species (see Fig. 5), among which the very common *E. falcata*, which partitioned the forest into north-eastern and south-western portions (initial CAs and NSCAs; see Figs. 3 and 4). This result is all the more interesting, since Counami forest showed neither evidence of recent large canopy-gaps (Couteron, unpublished results from aerial photographs), nor any subdivision between old-growth and secondary succession stages. Although, a significant relationship between soil types and stand structures has been established at plot or transect level in French Guiana (Lescure and Boulet, 1985; Sabatier et al., 1997; Paget, 1999), and at landscape scale in Costa Rica (Clark and Clark, 2000), no clear link between topography and stand structure emerged from our analyses. Factors that could explain the mesoscale variation of the stand structure in Counami forest thus remain to be investigated. At this point, a plausible hypothesis is that stand structure reflects the stage of forest reconstitution after tree-falls that create gaps of limited size. If so, the frequency and the size distribution of such gaps would, nevertheless, not depend much on topography, in contrast to what has been reported by Hartshorn (1978) and Poorter et al. (1994), but in agreement with the results of Hubbell et al. (1999).

5.3. *Drawing ecological insights from tropical forest inventories*

Our results confirm that, in spite of a certain scepticism among tropical ecologists, a meaningful ecological analysis can be carried out from forest inventory data. At Counami, some botanical species, well identified by field workers, like *E. falcata* or *E. oleracea*, exhibited significant ecological preferences that could be compared with the results obtained from more accurate studies carried out on smaller sampling plots by experienced botanists (e.g. Sabatier et al., 1997; Paget, 1999; Pélissier et al., 2002a). Even some abundant vernacular species, like Mutusi marécage or Koko, though composite in terms of botanical species, showed consistent mesoscale ecological behaviours. Such a result could be useful for management perspectives, such as rapid stand appraisal, as well as for suggesting future research directions. Hence, although a precise botanical identification would be always desirable, it is neither always possible, particularly over vast forest areas, nor always necessary to the detection of relevant ecological patterns. A compromise must be sought between what would be desirable to record and what is possible with the available human and material resources (Rondeux, 1999). From this point of view, collections of herbarium specimens, identified under their local vernacular names and subsequently checked for botanical equivalence by experienced botanists, should be encouraged (Richards, 1996, pp. 495–496). Similarly, detailed soil studies will never cover vast forest areas. Therefore, identifying simple qualitative descriptors of the main soil features is highly advisable. Moreover, some ecological variables could be filled in a posteriori, using topographical maps, geological maps or aerial photographs, while being ground-truthed through selective returns to the field. Finally, the use of complementary multivariate methods, like CA(IV) and NSCA(IV), should be more systematic in order to extract, from somewhat imprecise inventory data, trends that would be sufficiently robust to allow ecological patterns and processes to be safely inferred.

Putting together all these elements can permit extraction of valuable ecological information, such as species/environment relationships or forest zonation, from forest inventory data. Hence, such an objective should be considered as an essential element of scaling up insights gained from local case studies to mesoscale and regional perspectives on tropical rain forest organisation.

Acknowledgements

The inventory of Counami forest was carried out by CIRAD-Forêt as a contribution to a project of improved forest management, which was conducted by the technical section of the Office National des Forêts (ONF) and by Silvolab-Guyane. The work related to the DEA thesis of D. Mapaga was supported by the ENGREF centre of Kourou under the supervision of M. Fournier. D. Sabatier (IRD) greatly helped put vernacular categories into botanical species. We are particularly grateful to the field workers of CIRAD, M. Baisie, F. Kago, O. Ngwete, P. Naisso, M. Passou and R. Santé who, under the supervision of one of us (L. Teillier) and of J.-M. Bouvet (both at CIRAD), have tirelessly collected valuable and highly consistent data.

References

Ashton, P.S., Hall, P., 1992. Comparisons of structure among mixed Dipterocarp forests of north-western Borneo. J. Ecol. 80, 459–481.

Boggan, J., Funk, V., Kelloff, C., Hoff, M., Cremers, G., Feuillet, C., 1997. cheklist of the Plants of the Guianas (Guyana, Surinam, French Guiana), 2nd Edition. Wiley, New York, p. 238.

Blancaneaux, P., 2001. Le Climat Guyanais. In: Barret, J. (Ed.), Atlas Illustré de la Guyane. CNES, IESG, IRD, Région Guyane, pp. 46–49.

Chessel, D., Gimaret-Carpentier, C., 1998. ADE-4, Fiche thématique 4.9: Analyse canonique et liste d'occurrences d'espèces. http://biomserv.univ-lyon1.fr/ADE-4.html, p. 37.

Clark, D.B., Clark, D.A., 2000. Lanscape-scale variation in forest structure and biomass in a tropical rain forest. For. Ecol. Manage. 137, 185–198.

Clark, D.B., Clark, D.A., Read, J.M., 1998. Edaphic variation and the mesoscale distribution of tree species in a neotropical rain forest. J. Ecol. 86, 101–112.

Clark, D.B., Palmer, M.W., Clark, D.A., 1999. Edaphic factors and the landscape-scale distributions of tropical rain forest trees. Ecology 80, 2662–2675.

Collinet, F., 1997. Essai de regroupement des principales espèces structurantes d'une forêt dense humide d'après l'analyse de leur répartition spatiale (Forêt de Paracou, Guyane). Thèse de Doctorat, Université Claude Bernard, Lyon, France, p. 313.

Condit, R., 1996. Defining and mapping vegetation types in megadiverse tropical forests. Trends Ecol. E 11, 4–5.

Dolédec, S., Chessel, D., Ter Braak, C.J.F., Champely, S., 1996. Matching species traits to environmental variables: a new three-table ordination method. Environ. Ecol. Stat. 3, 143–166.

Dolédec, S., Chessel, D., Gimaret-Carpentier, C., 2000. Niche separation in community analysis: a new method. Ecology 81, 2914–2927.

Fraile, L., Escoufier, Y., Raibaut, A., 1993. Analyse des correspondances de données planifiées: étude de la chémotaxie de la larve infestante d'un parasite. Biometrics 49, 1142–1153.

Gimaret-Carpentier, C., 1999. Analyse de la biodiversité à partir d'une liste d'occurrences d'espèces: nouvelles méthodes d'ordination appliquées à l'étude de l'endémisme dans les Ghats occidentaux. Thèse de Doctorat, Université Claude Bernard, Lyon, France, p. 242.

Gimaret-Carpentier, C., Chessel, D., Pascal, J.-P., 1998. Non-symmetric correspondence analysis: an alternative for species occurrences data. Plant Ecol. 138, 97–112.

Gonzalès, S., Ferry, B., 1998. Etude phyto-écologique des bas-fonds de Paracou. Multigrade, ENGREF, Kourou, p. 40.

Greig-Smith, P., 1983. Quantitative Plant Ecology. Blackwell Scientific Publications, Oxford, p. 256.

Hartshorn, G.S., 1978. Tree falls and tropical forest dynamics. In: Tomlinson, P.B., Zimmermann, M.H. (Eds.), Tropical Trees as Living Systems. Cambridge University Press, Cambridge, pp. 617–638.

Hill, M.O., 1973. Reciprocal averaging: an eigenvector method of ordination. J. Ecol. 6, 237–249.

Hubbell, S.P., Foster, R.B., O'Brien, S.T., Harms, K.E., Condit, R., Wechsler, B., Wright, S.J., Loo de Lao, S., 1999. Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. Science 283, 554–557.

Lauro, N., D'Ambra, L., 1984. L'analyse non-symétrique des correspondances. In: Diday, E., et al. (Ed.), Data Analysis and Informatics. III. Elsevier, Amsterdam, pp. 433–446.

Lebreton, J.-D., Chessel, D., Prodon, R., Yoccoz, N., 1988. L'analyse des relations espèces-milieu par l'analyse canonique des correspondances. I. Variables de milieu quantitatives. Acta Oecol. 9, 53–67.

Lescure, J.-P., Boulet, R., 1985. Relationships between soil and vegetation in a tropical rain forest in French Guiana. Biotropica 17, 155–164.

Manly, B.F.J., 1991. Randomization and Monte Carlo Methods in Biology. Chapman & Hall, London, p. 281.

Milési, J.P., Egal, E., Ledru, P., Vernhet, Y., Thiéblemont, D., Cocherie, A., Tegyet, M., Martel-Jantin, B., Lagny, P., 1995. Les minéralisations du nord de la Guyane Française dans leur cadre géologique. Chronique de la Recherche Minière 518, 5–58.

Oldeman, R.A.A., 1968. Sur la valeur des noms vernaculaires des plantes en Guyane française. Bois For. Trop. 117, 17–23.

Paget, D., 1999. Etude de la diversité spatiale des écosystèmes forestiers guyanais: réflexion méthodologique et application. Thèse de Doctorat, ENGREF, Nancy, France, p. 155.

Pélissier, R., Dray, S., Sabatier, D., 2002a. Within-plot relationships between tree species occurrences and hydrological soil constraints: an example in French Guiana investigated through canonical correlation analysis. Plant Ecol., in press.

Pélissier, R., Couteron, P., Dray, S., Sabatier, D., 2002b. Consistency between ordination techniques and diversity measurements: two alternative strategies for species occurrence data. Ecology, in press.

Pitman, N.C.A., Terborgh, J., Silman, M.R., Nuñez, P., 1999. Tree species distributions in an upper Amazonian forest. Ecology 80, 2651–2661.

Poorter, L., Jans, L., Bongers, E., Rompaey, R.S.A.R.V., 1994. Spatial distribution of gaps along 3 catenas in the moist forest of Taï National Park, Ivory-Coast. J. Trop. Ecol. 10, 385–398.

Rao, C.R., 1964. The use and interpretation of principal component analysis in applied research. Sankhya Ser. A 26, 329–357.

Richards, P.W., 1996. The Tropical Rain Forest. Cambridge University Press, Cambridge, p. 575.

Riéra, B., Pélissier, R., Houllier, F., 1998. Caractérisation d'une mosaïque forestière et de sa dynamique en forêt tropicale humide sempervirente. Biotropica 30, 251–260.

Rondeux, J., 1999. Forest inventories and biodiversity. Unasylva 196, 35–41.

Sabatier, R., Lebreton, J.-D., Chessel, D., 1989. Principal component analysis with instrumental variables as a tool for modeling composition data. In: Coppi, R., Bolasco, S. (Eds.), Multiway Data Analysis. Elsevier, Amsterdam, pp. 341–352.

Sabatier, D., Grimaldi, M., Prévost, M.-F., Guillaume, J., Godron, M., Dosso, M., Curmi, P., 1997. The influence of soil cover organization on the floristic and structural heterogeneity of a Guianan rain forest. Plant Ecol. 131, 81–108.

Sheil, D., 1995. A critique of permanent plot methods and analysis with examples from Budongo forest, Uganda. For. Ecol. Manage. 77, 11–34.

Shiver, B.D., Bonders, B.E., 1996. Sampling Techniques for Forest Resource Inventory. Wiley, New York, p. 356.

Sokal, R.R., Rohlf, F.J., 1995. Biometry. Freeman, New York, p. 887.

Simpson, E.H., 1949. Measurement of diversity. Nature 163, 688.

Ter Braak, C.J.F., 1987. The analysis of species–environment relationships by canonical correspondence analysis. Vegetatio 69, 69–77.

Ter Braak, C.J.F., 1988. Partial canonical correspondence analysis. In: Bock, H.H. (Ed.), Classification and Related Methods of Data Analysis. Elsevier, Amsterdam, pp. 551–558.

Ter Steege, H., Jetter, V.G., Polak, A.M., Werger, M.J.A., 1993. Tropical rain forest and soil factors in a watershed area in Guyana. J. Vegetat. Sci. 4, 705–716.

Ter Steege, H., Sabatier, D., Castellanos, H., Van Andel, T., Duivenvoorden, J., Adalardo de Oliveira, A., Ek, R., Lilwah, R., Maas, P., Mori, S., 2000a. A regional perspective: analysis of amazonian floristic composition and diversity that includes the Guiana shield. In: Ter steege, H. (Ed.), Plant Diversity in Guyana. The Tropenbos Foundation, Wageningen, The Netherlands, pp. 19–34.

Ter Steege, H., Sabatier, D., Castellanos, H., Van Andel, T., Duivenvoorden, J., Adalardo de Oliveira, A., Ek, R., Lilwah, R., Maas, P., Mori, S., 2000b. An analysis of the floristic composition and diversity of Amazonian forests including those of the Guiana shield. J. Trop. Ecol. 16, 801–828.

Thioulouse, J., Chessel, D., Dolédec, S., Olivier, J.-M., 1997. ADE-4: a multivariate analysis and graphical display software. Stat. Comput. 7, 75–83.