



**HAL**  
open science

# The relative importance of dispersal limitation and habitat preference in shaping spatial distribution of saplings in a tropical moist forest: a case study along a combination of hydromorphic and canopy disturbance gradients

Grégoire Vincent, Jean-François Molino, Lucile Marescot, Karim Barkaoui, Daniel Sabatier, Vincent Freycon, Jean-Baptiste Roelens

## ► To cite this version:

Grégoire Vincent, Jean-François Molino, Lucile Marescot, Karim Barkaoui, Daniel Sabatier, et al.. The relative importance of dispersal limitation and habitat preference in shaping spatial distribution of saplings in a tropical moist forest: a case study along a combination of hydromorphic and canopy disturbance gradients. *Annals of Forest Science*, 2011, 68 (2), pp.357-370. 10.1007/s13595-011-0024-z . hal-00930759

**HAL Id: hal-00930759**

**<https://hal.science/hal-00930759>**

Submitted on 11 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# The relative importance of dispersal limitation and habitat preference in shaping spatial distribution of saplings in a tropical moist forest: a case study along a combination of hydromorphic and canopy disturbance gradients

Grégoire Vincent · Jean-François Molino ·  
Lucile Marescot · Karim Barkaoui · Daniel Sabatier ·  
Vincent Freycon · Jean Baptiste Roelens

Received: 17 April 2010 / Accepted: 26 August 2010 / Published online: 23 February 2011  
© INRA and Springer Science+Business Media B.V. 2011

## Abstract

• **Introduction** Various processes contribute to shaping the local assemblage of species in hyperdiverse tropical forest. The relative contribution of environmental factors and dispersal limitation in determining the spatial distribution of saplings at local scale is unclear.

• **Methods** We examined two types of environmental factors: (a) soil type reflecting drainage regime and (b) past logging damages reflecting light regime in a neotropical moist forest site. We used a logistic model to predict presence or absence of a given species in a network of elementary small plots.

• **Results** The effect of mapped environmental factors and a spatial correlation term were jointly estimated providing a direct measure of the relative role of habitat specialisation and dispersal limitation. At community level, dispersal limitation was the most important determinant of species

absence at local scale. The two environmental factors examined played a balanced role. Different species however showed different degrees of dispersal limitation and habitat specialisation. The distribution of a large proportion of species—the majority of the most abundant species—was significantly affected by at least one environmental factor. We provide a ranking of 49 species sensitive to canopy disturbance (from shade specialist to pioneer) and 41 species affected by seasonal flooding (either positively or negatively).

**Keywords** Species assemblage · Dispersal limitation · Tropical moist forest · Niche

## 1 Introduction

Excluding the most extreme environments such as swamps, white-sand podzols or flood plain forest, local tree species richness in neotropical moist forest is remarkably high. A single hectare of tropical rainforest may harbour 100 to 300 species of trees with stem diameter at breast height (dbh) above 10 cm (Gentry 1988; Valencia et al. 1994; ter Steege et al. 2000).

What are the key processes responsible for the maintenance of such a high local diversity is yet unclear. The respective importance of chance, dispersal limitation and species ecological niche in shaping local assemblage of species is still a matter of debate (Hubbell and Foster 1986; Condit et al. 2000; Pulliam 2000; Harms et al. 2001; Dalling et al. 2002; Phillips et al. 2003; Bell 2005).

Tropical tree species differ in their light requirement for growth and survival at seedling stage (Poorter and Arets 2003; Wright et al. 2003) and hence light regime is

---

Handling Editor: Erwin Dreyer

---

G. Vincent (✉) · L. Marescot · K. Barkaoui  
IRD-Unité Mixte de Recherche AMAP,  
Campus Agronomique-(CIRAD), BP 701, 97387 Kourou Cedex,  
French Guiana, France  
e-mail: gregoire.vincent@ird.fr

J.-F. Molino · D. Sabatier  
IRD, UMR AMAP,  
Montpellier 34000, France

J. B. Roelens  
CIRAD, UMR ECOFOG,  
97310 Kourou, France

V. Freycon  
CIRAD, Unité Propre de Recherche BSEF,  
34398 Montpellier, France

expected to affect species recruitment locally. Soil characteristics and notably water regime (drainage, water holding capacity) are also known to act as potentially important ecological filters for many tropical tree species (Sollins 1998). However, a complication arises when estimating the effect of environmental factors on species abundance as statistical models may be influenced by spatial autocorrelation in abundance, environmental variables or both. Neglecting spatial autocorrelation is likely to lead to incorrect conclusions regarding the importance of environmental variables as determinants of abundance (Keitt et al. 2002). In addition, spatial autocorrelation in abundance potentially conveys ecologically meaningful information on species dispersal strategy (Traissac 2003).

Spatial autocorrelation in abundance can be treated as a mere complication when exploring species habitat relationships. Alternatively, as in the present study, the focus can be on assessing the importance of spatial correlation in stem distribution versus environmental factors in determining presence or absence of a species in a particular location. To do so, we posit that the probability of finding a species in a particular microsite depends on overall population density at the site, microsite habitat suitability and species dispersal-colonisation efficiency. Formally, dispersal limitation is equated with the increase in the odds of finding a species in a particular microsite which is associated to the decrease of one unit distance to the nearest conspecific stem. In other words, high clumping indices are indicative of low dispersal efficiency and vice versa.

The first objective was therefore to examine the relative contribution of selected environmental factors and dispersal limitation in constraining the spatial distribution of species at local scale ( $<3 \text{ km}^2$ ). The second objective was simply to provide dearly lacking baseline data on species habitat preference at sapling stage.

## 2 Material and methods

### 2.1 Study site

This study was conducted in the lowland tropical rain forest of the Paracou experimental site, near Sinnamary, French Guiana. Elevation is between 5 and 50 m, and mean annual temperature is  $26^\circ\text{C}$ , with an annual range of  $1^\circ\text{C}$  to  $1.5^\circ\text{C}$ . Rainfall averaged  $2,875 \pm 540 \text{ mm} \cdot \text{a}^{-1}$  for the period 1986–2005, with a 3-month dry season (less than  $100 \text{ mm} \cdot \text{month}^{-1}$ ) from mid-August to mid-November. The landscape features a patchwork of hills (100–300 m in diameter and 20–35 m high) separated by narrow streams. Slopes range from 25% to 45%. The parent material is a metamorphic formation of the Precambrian shield, characterised by schists and

sandstones, locally crossed by veins of pegmatite, apatite and quartz. Soils are mostly acrisols (Epron et al. 2006), limited in depth by loamy aloterite, which has a low permeability, leading to lateral underground circulation of water during heavy rains. Upper permeable horizons are usually thickest on hilltops and shallowest in upper or mid position along slopes (Boulet 1990; Barthès 1991). Bleaching of surface horizons resulting from seasonal waterlogging is total along streams (Gleysols), but occurs also on the slopes, especially in their lowest parts.

Spatial variation of the soil cover was mapped following the method described by (Boulet 1978). The soil mapping units correspond to successive stages of evolution of a currently unbalanced ferrallitic cover. Five soil units occur at the experimental site. Those soil units are described in detail in Sabatier et al. (1997): Alt (first stage of the thinning out of the original microaggregated ferrallitic cover, vertical drainage is restricted by less porous red aloterite at a depth of less than 1.2 m), superficial lateral drainage (SLD) associated to further weathering of the initial ferrallitic soil, uphill system (UHS): yet a more evolved stage where soil saturation lasts longer in the rainy season when the slope is more gentle on flat top hills near the top; temporary perched water-table, down-hill system (DhS): lateral saturated throughflow of superficial horizons feed the seasonally fluctuating water table, surface hydromorphy (SH), associated to thalwegs, where water table seasonally or permanently (swamps) reaches the soil surface.

Between 1986 and 1988, four contrasted silvicultural treatments were applied on the 12 plots of the Paracou experimental site (Fig. 1 in the Electronic supplementary material in Molino and Sabatier 2001). The treatments were (Gourlet-Fleury et al. 2004): (1) no disturbance (i.e. control plots) “ $T_0$ ”; (2) traditional selective felling of commercial species, namely of ca. ten large trees,  $\geq 50\text{--}60 \text{ cm dbh}$ , per ha “ $T_1$ ”; (3) selective felling plus thinning by poison-girdling of all non-commercial species (NCS)  $\geq 40 \text{ cm dbh}$  “ $T_2$ ”; and (4) selective felling plus logging of NCS between 40 and 50 cm dbh (for fuelwood), followed by poison-girdling of all NCS  $\geq 50 \text{ cm dbh}$  “ $T_3$ ”. All logging damage (logging trails and canopy openings) were carefully mapped and digitized into a GIS system.

### 2.2 Tree inventory data

We used two datasets collected on the same site but with different characteristics. Dataset 1 is based on a complete botanical inventory of all stems tallied in a set of scattered plots jointly covering 5 ha. Dataset 2 is spatially intensive and comprises all stems of 15 selected species mapped over a compact area of 25 ha. Both datasets were inventories of self-supporting ligneous

plants with a stem diameter at 1.30 m above ground (dbh) between 2 and 10 cm (dataset 1) or 1 and 10 cm (dataset 2). The experimental plots from which the tree datasets were collected fell within a single convex area of about 3 km<sup>2</sup> (Fig. 1 in the Electronic supplementary material in Molino and Sabatier 2001).

The rationale for focusing on a subset of the tree community formed by small-sized individuals was that the floristic composition of small stems (predominantly composed of young trees) was more likely to capture the impact of recent disturbance than the floristic composition of the entire tree community. On the other hand the species-soil type association would probably have been stronger if adult trees had been considered. Indeed environmental filtering would have exerted its influence for a longer time and differential mortality would have played a stronger role (Webb and Peart 2000; Baraloto et al. 2007). By including all stems between 1 (or 2) and 10 cm we attempted to capture both type of species-habitat associations albeit each one possibly somewhat weakened.

### 2.3 Dataset 1

Ten 0.5-ha (20×250 m) transects were laid out in the Paracou experimental site (Fig. 1 in the Electronic supplementary material in Molino and Sabatier 2001). Those transects included all logging intensities experimented on the site. Between November 1994 and October 1998 (i.e. 8 to 12 years post logging) all 14,032 self-supporting stems with  $2\text{ cm} \leq \text{dbh} \leq 10\text{ cm}$  were identified to species level and mapped to the nearest m. Among the 512 species inventoried, 84 were tallied at least 40 times (i.e. were species with average density  $>8$  stems/ha). Those 84 species totalled 9,617 stems and were included in the analysis. Most of them (75) are “true” tree species, i.e. at least some individuals  $>10$  cm dbh have been recorded in Paracou or nearby. The other nine species might be considered as treelets in our area, although for six of them individuals with dbh  $>10$  cm have been recorded elsewhere.

### 2.4 Dataset 2

The second dataset covers 25 ha in four 6.25 ha square plots (P9 to P12 in Fig. 1 in the Electronic supplementary material in Molino and Sabatier (2001)). Three of the four plots were logged over in 1986–1988 according to one of the three silvicultural treatments described above. The fourth was left unlogged (control).

All stems with  $1\text{ cm} \leq \text{dbh} \leq 10\text{ cm}$  from 15 selected species were tagged, recorded and mapped to the nearest m. Species were chosen in order to encompass a wide range of

adult size, light requirement and dispersal mode (Flores 2005). The 28,619 stems were tallied between March and July 2002 i.e. ca. 15 years after logging.

Three transects of dataset 1 are included in the plots covered by dataset 2. Less than 2% of stems are common to both datasets.

### 2.5 Data analysis

We predicted presence or absence of a given species in a network of  $5 \times 5$  m quadrats on the basis of mapped environmental factors and a spatial correlation term.

All models were species specific, i.e. they aimed at predicting a particular species spatial distribution. No attention was paid to inter-specific spatial correlation.

When assessing the proportion of species sensitive to one or the other environmental factor at community level, we applied the same test to all the species in the dataset. Since many species were tested independently false rejection of  $H_0$  hypothesis may have occurred. Therefore, we applied the Dunn–Sidak correction for multiple independent tests to the individual experiment alpha values to ensure a family wise error rate of 0.05.

### 2.6 Factor coding

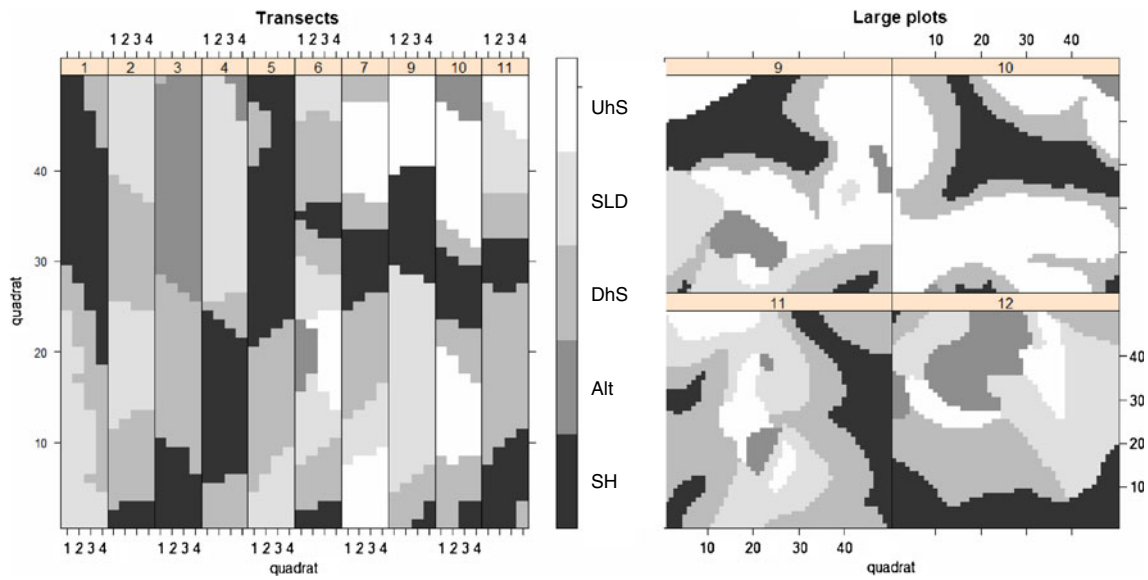
Soil typology describing the drainage regime had five classes (Fig. 1).

Logging damages which were mapped at the time of logging were used as a proxy for canopy disturbance. It should be noted that soil disturbance may be locally important following logging and may also affect local species composition. The  $\log_{10}$ -transformed distance to the nearest logging damage (whatever type of damage skidding trail or canopy gap) was used as a disturbance descriptor (Fig. 2).

In addition to the above predictors all models included a clumping index factor (or spatial autocorrelation factor) which was defined for any combination of a  $5 \times 5$  m quadrat and a species as the distance from that quadrat to the nearest quadrat harbouring a conspecific stem.

### 2.7 Randomization tests

Coincidental correlation between spatial structure in species distribution and environmental factors may lead to spurious species-habitat associations (Harms et al. 2001). This problem is reduced but not completely avoided by including a spatial correlation term. To further guard against such spurious species-habitat associations, torus translations between plots were used to generate random patterns of associations between plants and habitat—defined as a combination of soil type and distance to



**Fig. 1** Soil map (five classes) for dataset 1 (ten 250×20 m scattered plots) and dataset 2 (four 250×250 m noncontiguous plots); relative position of plots is arbitrary. Grey levels represent, from dark to white:

*SH* surface hydromorphy, *Alt* fair drainage, *DhS* downhill system, *SLD* superficial lateral drainage, *UhS* uphill system

canopy disturbance—while preserving the spatial structure of both the habitats and the plant populations. Typically, 999 random configurations were drawn. For each random association, the model was rerun. When random associations of maps lead to lower residual deviance of the full model than the original data in more than 5% of the cases the species-habitat was considered dubious.

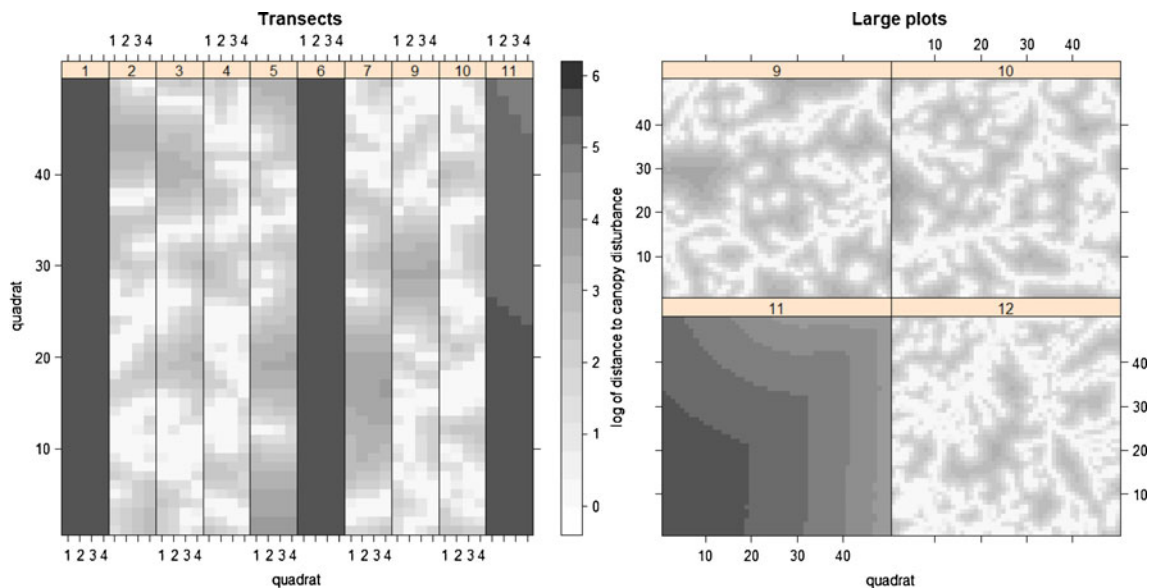
In dataset 1, torus translations were conducted along the main axis only to minimize edge effects such as unrealistic

adjacent environmental classes which may occur when wrapping up one edge onto the opposite.

For each simulated dataset, torus translations were drawn independently for each plot/transect.

2.8 Odds ratio

The logistic model of species’ presence/absence data was fitted via lrm procedure in the Design library (Harrell 2001) in



**Fig. 2** Canopy disturbance for the same sets of plots as Fig. 2. Grey levels represent log of distance in m to closest logging damage. Relative position of plots is arbitrary

*R* (<http://cran.rproject.org>). When a binary outcome variable is modelled using logistic regression, it is assumed that the logit transformation of the outcome variable has a linear relationship with the predictor variables. The odds of success are defined as the ratio of the probability of success of an event (such as finding a species in a habitat) over the probability of failure of the same event. The transformation from probability to odds is a monotonic transformation, meaning the odds increase as the probability increases or vice versa. Odds range from 0 to positive infinity. An odds ratio associated to some type of environment with a value inferior to 1 reflects the avoidance of a species for such an environment. On the contrary, an odds ratio superior to one indicates a positive species-environment association. Alternatively positive association with disturbance will translate into an odds ratio of the (log)distance to disturbance less than 1 and positive spatial correlation will be associated to odds ratio less than one (the larger the distance to a particular species the smaller the probability to find a conspecific).

### 3 Results

#### 3.1 Dataset 1

The 84 most abundant species made up for 69% of the sapling community. Summing-up the contribution of each factor to the explained deviance across those 84 species (tab 1,  $\chi^2$  columns), the overall relative contribution of each factor was respectively 13% for drainage regime, 17% for canopy disturbance, and 70% for clumping index.

Seventy-five species out of 84 (89%) were found to be significantly aggregated (column “PCLump” $<0.05$  and “Odds ratio Clump” $<1$ , Table 1). Applying the Dunn–Sidak correction to the level of significance (i.e.  $P<0.0006$ ) reduced the proportion of significantly aggregated species to 71%.

Forty-two (50%) species (Table 1) were found to be affected by environmental factors and more frequently so by proximity to past logging disturbance (33/42) than drainage regime (18/42). Following the Dunn–Sidak correction the proportion of species significantly affected by environmental factors dropped to 23%.

Torus translation test lead to the elimination of another two species. This left 17 species (20%) significantly affected by environmental factors. Of those 17 species 16 were affected by past disturbance and three by soil type.

Four of the seven most abundant species (>250 stems in 5 ha) are included (Table 1) in those 17 significant species  $\times$  habitat associations.

#### 3.2 Dataset 2

Summing-up the reduction of deviance associated with each factor across the 15 species of dataset 2 (Table 2,  $\chi^2$  columns), the relative contribution of each model term was respectively 16% for drainage regime, 8% for disturbance and 76% for clumping index.

Fourteen out of 15 species appeared to be significantly aggregated at the family wise error rate of 0.05 (“PCLump” $<0.0034$  and “OddClump” $<0$ , Table 2).

Thirteen species were found to be significantly affected by at least one environmental factor ( $P_{Env}<0.0034$ ). Ten species were retained as being significantly associated with a particular habitat after the torus translation test. Nine species were affected by soil type and seven by canopy disturbance regime. Hence, the spatial distribution of ten of the 15 species tested (i.e. 67%) can be considered as being affected by at least one environmental factor.

Rankings along ecological gradients are presented in Figs. 3 and 4. We used minus log of the average of the odds ratio over all non-SH soil types as an index of flood tolerance. This is a fairly crude summary of the sensitivity to restricted drainage and should be taken as such. Some discrepancies between the two datasets are noticeable, in particular for the hydromorphy ranking of three species (Fig. 3) all of which have less than 130 stems in dataset 1.

Similarly, when the disturbance effect was retained in the specific logistic model, we used the minus log of oddDist as an index of heliophily. More species were found to be favoured by canopy disturbance than negatively affected (Fig. 4).

### 4 Discussion

Considering all species in each dataset together, the clumping factor accounted for 70% and 76%, respectively, of the reduction of deviance in datasets 1 and 2. This indicated that dispersal limitation was the major determining factor of species distribution at the scale of the study presented here.

There were however differences between species and datasets in terms of the relative contribution of the factors tested. The lower contribution of the canopy disturbance factor in dataset 2 was probably a consequence of not having included any pioneer species in that dataset. Notably, a number of pioneer species in dataset 1 (Table 1) appear more constrained by the degree of canopy disturbance than dispersal limitation (*Cecropia obtusa* Trécul, *Goupia glabra* Aubl., *Inga cayennensis* Sagot ex Benth., *Inga umbellifera* (Vahl) Steud., *Loreya mespiloides* Miq.,...).

Consistent with previous findings (Condit et al. 2000), most species showed a significant degree of spatial correla-

**Table 1** Summary statistics for the logistic model of presence/absence of 84 relatively abundant species in Paracou

Taxon	Abund.	R <sup>2</sup>	Intercept	OddDHS	OddAlt	OddSLD	OddUHS	OddDist	OddClump	PSoil	PDist	PEnv	PClump	χ <sup>2</sup> Soil	χ <sup>2</sup> Dist	χ <sup>2</sup> Clump	PTorus
<i>Anaxagorea dolichocharpa</i> Sprague & Sandwith	470	0.24	9.56	NA	NA	NA	NA	1.10	0.13	NA	0.002	0.002	0.000	NA	9.83	128.24	0.000
<i>Apeiba glabra</i> Aubl.	73	0.12	0.39	NA	NA	NA	NA	0.73	0.51	NA	0.001	0.001	0.000	NA	10.57	16.97	0.001
<i>Bocopa prouaencensis</i> Aubl.	60	0.02	0.13	NA	NA	NA	NA	NA	0.60	NA	NA	NA	0.001	NA	NA	10.11	NA
<i>Capparis frondosa</i> Jacq.	56	0.13	0.64	0.49	0.24	0.41	0.29	NA	0.41	0.048	NA	0.048	0.000	9.61	NA	28.47	0.037
<i>Carapa procera</i> DC.	44	0.01	0.03	NA	NA	NA	NA	0.85	NA	NA	0.050	0.050	NA	NA	3.84	NA	0.122
<i>Catostemma fragrans</i> Benth.	108	0.02	0.03	NA	NA	NA	NA	1.20	NA	NA	0.000	0.000	NA	NA	13.91	NA	0.001
<i>Cecropia obtusa</i> Trécul	73	0.20	0.47	NA	NA	NA	NA	0.47	0.54	NA	0.000	0.000	0.001	NA	23.96	11.24	0.000
<i>Conceveiba guianensis</i> Aubl.	80	0.06	0.11	0.81	1.04	0.44	0.29	0.75	NA	0.032	0.000	0.000	NA	10.53	18.80	NA	0.027
<i>Cordia nodosa</i> Lam.	51	0.00	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>Cordia sagotii</i> I.M.Johnst.	53	0.15	0.44	NA	NA	NA	NA	0.72	0.49	NA	0.003	0.003	0.000	NA	9.06	21.53	0.034
<i>Couratari oblongifolia</i> Ducke & R.Knuth	63	0.11	0.07	3.19	5.09	5.04	4.64	NA	0.49	0.100	NA	0.100	0.000	7.79	NA	22.18	0.126
<i>Cupania rubiginosa</i> (Poir.) Radlk.	63	0.11	0.45	NA	NA	NA	NA	NA	0.40	NA	NA	NA	0.000	NA	NA	34.93	NA
<i>Dicorynia guianensis</i> Amshoff	55	0.19	0.27	1.00	0.00	2.65	2.94	NA	0.39	0.051	NA	0.051	0.000	9.45	NA	32.10	0.032
<i>Drypetes variabilis</i> Uittien	47	0.07	0.27	NA	NA	NA	NA	NA	0.45	NA	NA	NA	0.000	NA	NA	23.56	NA
<i>Duguetia calycina</i> Benoist	114	0.07	0.44	NA	NA	NA	NA	NA	0.45	NA	NA	NA	0.000	NA	NA	32.76	NA
<i>Duguetia yeshidan</i> Sandwith	122	0.61	5.50	NA	NA	NA	NA	1.71	0.11	NA	0.001	0.001	0.000	NA	10.38	37.42	0.002
<i>Eperua falcata</i> Aubl.	238	0.30	8.01	NA	NA	NA	NA	1.08	0.13	NA	0.063	0.063	0.000	NA	3.47	101.17	0.014
<i>Eperua grandiflora</i> (Aubl.) Benth.	46	0.22	0.63	NA	NA	NA	NA	NA	0.37	NA	NA	NA	0.000	NA	NA	36.96	NA
<i>Eschweilera congestiflora</i> (Benoist) Eyma	44	0.15	0.05	3.16	6.55	4.83	8.28	1.13	0.43	0.065	0.141	0.065	0.000	8.84	2.17	24.81	0.003
<i>Eschweilera corticea</i> (DC.) S.A.Mori	58	0.13	0.47	NA	NA	NA	NA	NA	0.40	NA	NA	NA	0.000	NA	NA	39.80	NA
<i>Eschweilera decolorans</i> Sandwith	43	0.28	1.39	NA	NA	NA	NA	NA	0.27	NA	NA	NA	0.000	NA	NA	44.04	NA
<i>Eschweilera sagotiana</i> Miers	206	0.04	0.60	NA	NA	NA	NA	NA	0.46	NA	NA	NA	0.000	NA	NA	31.04	NA
<i>Goupta glabra</i> Aubl.	99	0.22	0.40	1.05	2.27	0.59	0.72	0.46	0.65	0.054	0.000	0.000	0.028	9.30	35.50	4.81	0.000
<i>Guarea costata</i> A.Juss.	54	0.24	0.15	3.05	7.14	2.08	1.33	NA	0.42	0.003	NA	0.003	0.000	15.73	NA	27.54	0.015
<i>Guatteria punctata</i> (Aubl.) R. A.Howard	43	0.14	0.45	NA	NA	NA	NA	0.83	0.44	NA	0.069	0.069	0.000	NA	3.32	24.60	0.096
<i>Gustavia hexapetala</i> (Aubl.) Sm.	64	0.01	0.09	NA	NA	NA	NA	NA	0.70	NA	NA	NA	0.021	NA	NA	5.29	NA
<i>Hebepetalum humirifolium</i> (Planch.) Benth.	72	0.25	2.60	NA	NA	NA	NA	NA	0.20	NA	NA	NA	0.000	NA	NA	60.18	NA
<i>Henriettella flavescens</i> (Aubl.) Triana	198	0.20	3.24	0.74	0.18	0.60	0.22	0.88	0.29	0.007	0.009	0.007	0.000	14.07	6.89	70.50	0.242
<i>Inga cayennensis</i> Sagot ex Benth.	41	0.09	0.02	1.16	0.20	0.35	0.00	0.73	1.42	0.105	0.001	0.001	0.062	7.65	11.53	3.49	0.008
<i>Inga loubryana</i> Poncey	87	0.11	0.26	1.52	3.87	1.63	2.33	NA	0.42	0.013	NA	0.013	0.000	12.66	NA	31.72	0.008
<i>Inga pezizifera</i> Benth.	106	0.33	1.39	1.50	1.75	0.22	0.98	0.78	0.34	0.177	0.044	0.044	0.000	6.32	4.04	36.93	0.409

<i>Inga stipularis</i> DC.	136	0.11	0.62	NA	NA	NA	NA	NA	0.77	0.49	NA	0.000	0.000	0.000	NA	14.14	21.08	0.001
<i>Inga umbellifera</i> (Vahl) Steud.	45	0.07	0.13	NA	NA	NA	NA	NA	0.70	0.71	NA	0.001	0.001	0.028	NA	11.23	4.81	0.024
<i>Iryanthera hostmannii</i> (Benth.) Warb.	54	0.04	0.16	NA	NA	NA	NA	NA	NA	0.55	NA	NA	NA	0.000	NA	NA	14.08	NA
<i>Iryanthera sagotiana</i> (Benth.) Warb.	183	0.18	0.95	1.42	4.25	2.14	2.10	2.10	NA	0.27	0.000	NA	0.000	0.000	22.63	NA	63.63	0.030
<i>Jacaranda copaia</i> (Aubl.) D. Don	116	0.13	0.45	1.08	2.72	0.73	0.72	0.72	0.80	0.51	0.002	0.002	0.002	0.000	16.60	9.78	19.27	0.019
<i>Lactia procera</i> (Poepp.) Eichler	52	0.51	1.08	4.48	6.25	1.06	0.08	0.08	0.27	0.30	0.109	0.000	0.000	0.000	7.55	16.09	19.78	0.001
<i>Lecythis persistens</i> Sagot	390	0.08	2.27	0.78	0.26	0.69	0.57	0.57	0.94	0.37	0.003	0.045	0.003	0.000	16.41	4.02	45.04	0.232
<i>Licania alba</i> (Bernoulli) Cuatrec.	211	0.09	0.99	NA	NA	NA	NA	NA	1.06	0.34	NA	0.107	0.107	0.000	NA	2.60	56.51	0.275
<i>Licania canescens</i> Benoist	93	0.13	0.92	NA	NA	NA	NA	NA	NA	0.32	NA	NA	NA	0.000	NA	NA	52.54	NA
<i>Licania heteromorpha</i> Benth. var. <i>heteromorpha</i>	135	0.06	0.10	3.08	4.51	3.10	3.44	3.44	NA	0.56	0.008	NA	0.008	0.000	13.93	NA	17.54	0.001
<i>Licania membranacea</i> Sagot ex Laness.	206	0.03	0.18	0.99	1.59	1.68	1.51	1.51	1.09	0.62	0.052	0.023	0.023	0.001	9.37	5.15	11.07	0.032
<i>Licania sprucei</i> (Hook.f.) Fritsch	52	0.00	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>Loreya arborexscens</i> (Aubl.) DC.	62	0.38	6.54	NA	NA	NA	NA	NA	0.65	0.16	NA	0.018	0.018	0.000	NA	5.60	37.56	0.000
<i>Loreya mespiloides</i> Miq.	259	0.24	1.84	NA	NA	NA	NA	NA	0.55	0.40	NA	0.000	0.000	0.000	NA	48.18	22.97	0.000
<i>Mabea piriri</i> Aubl.	57	0.35	3.89	0.57	0.28	0.82	0.24	0.24	0.84	0.26	0.144	0.068	0.068	0.000	6.85	3.32	46.23	0.397
<i>Miconia acuminata</i> (Steud.) Naudin	324	0.41	17.31	NA	NA	NA	NA	NA	0.60	0.15	NA	0.000	0.000	0.000	NA	35.53	85.07	0.000
<i>Miconia minutiflora</i> (Bonpl.) DC.	46	0.36	1.40	NA	NA	NA	NA	NA	0.41	0.37	NA	0.000	0.000	0.000	NA	13.56	22.73	0.002
<i>Miconia tschudyoides</i> Cogn.	523	0.39	17.65	1.07	0.83	1.98	1.04	1.04	0.72	0.13	0.004	0.000	0.000	0.000	15.51	30.32	109.28	0.014
<i>Oxandra asbeckii</i> (Pulle) R.E.Fr.	459	0.17	5.40	NA	NA	NA	NA	NA	1.05	0.19	NA	0.124	0.124	0.000	NA	2.37	122.74	0.002
<i>Paypayrola guianensis</i> Aubl.	66	0.33	2.76	NA	NA	NA	NA	NA	1.19	0.15	NA	0.042	0.042	0.000	NA	4.15	59.53	0.024
<i>Pogonophora schomburgkiana</i> Miers ex Benth.	176	0.14	0.97	NA	NA	NA	NA	NA	1.08	0.32	NA	0.063	0.063	0.000	NA	3.46	60.37	0.047
<i>Poraqueiba guianensis</i> Aubl.	67	0.02	0.02	2.22	1.47	2.64	3.38	3.38	NA	NA	0.083	NA	0.083	NA	8.25	NA	NA	0.324
<i>Pouteria bangii</i> (Rusby) T.D. Penn.	41	0.00	0.09	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
<i>Pradosia cochlearia</i> (Lecomte) T.D. Penn.	92	0.09	0.10	3.05	3.21	2.89	3.16	3.16	1.13	0.45	0.158	0.031	0.031	0.000	6.62	4.66	25.99	0.039
<i>Protium giganteum</i> Engl. var. <i>crassifolium</i> (Engl.) Daly	47	0.07	0.26	NA	NA	NA	NA	NA	NA	0.46	NA	NA	NA	0.000	NA	NA	22.15	NA
<i>Protium opacum</i> Swart subsp. <i>rabelianum</i> Daly	174	0.11	0.86	0.59	1.04	0.32	0.66	0.66	0.83	0.52	0.007	0.000	0.000	0.000	14.26	13.34	18.17	0.027
<i>Qualea rosea</i> Aubl.	128	0.37	0.66	1.18	9.71	1.19	1.58	1.58	NA	0.32	0.000	NA	0.000	0.000	59.21	NA	42.80	0.053
<i>Recordoxylon spectiosum</i> (Benoist) Gazel ex Barneby	46	0.06	0.16	NA	NA	NA	NA	NA	NA	0.56	NA	NA	NA	0.000	NA	NA	17.94	NA
<i>Rinorea amapensis</i> Hekking	63	0.44	10.93	NA	NA	NA	NA	NA	NA	0.12	NA	NA	NA	0.000	NA	NA	51.67	NA
<i>Sandwithia guyanensis</i> Lanj.	209	0.59	5.45	2.27	7.72	2.01	2.27	2.27	1.32	0.09	0.023	0.000	0.000	0.000	11.38	12.98	50.13	0.029
<i>Schefflera decaphylla</i> (Sagot ex Seem.)	59	0.11	0.21	1.03	1.87	1.25	2.81	2.81	0.87	0.50	0.062	0.084	0.062	0.000	8.96	2.98	15.85	0.034



Table 1 (continued)

Taxon	Abund.	R <sup>2</sup>	Intercept	OddDhS	OddAlt	OddSLD	OddUhS	OddDist	OddClump	PSoil	PEnv	PClump	χ <sup>2</sup> Soil	χ <sup>2</sup> Dist	χ <sup>2</sup> Clump	PTorus
Harms																
<i>Simaba cedron</i> Planch.	83	0.02	0.16	NA	NA	NA	NA	NA	0.60	NA	NA	0.001	NA	NA	11.41	NA
<i>Sterculia pruriens</i> (Aubl.) K.Schum.	134	0.11	0.35	1.19	2.60	0.75	0.82	0.77	0.62	0.003	0.000	0.000	15.98	17.18	9.73	0.094
<i>Swartzia leblondii</i> R.S.Cowan	65	0.01	0.08	NA	NA	NA	NA	NA	0.72	NA	NA	0.038	NA	NA	4.29	NA
<i>Symphonia</i> sp. 1	79	0.10	0.03	6.56	4.99	10.80	7.46	NA	0.56	0.015	NA	0.015	12.35	NA	12.29	0.005
<i>Tachigali melinonii</i> (Harms) Zarucchi & Herend.	56	0.20	0.62	1.75	1.03	0.44	0.36	0.82	0.40	0.038	0.139	0.038	10.16	2.19	29.86	0.291
<i>Talisia guianensis</i> Aubl.	124	0.04	0.25	NA	NA	NA	NA	1.08	0.52	NA	0.100	0.100	NA	2.71	21.37	0.143
<i>Talisia hexaphylla</i> Vahl	66	0.03	0.13	NA	NA	NA	NA	NA	0.62	NA	NA	0.001	NA	NA	11.45	NA
<i>Talisia megaphylla</i> Sagot ex Radlk.	48	0.01	0.02	NA	NA	NA	NA	1.13	NA	NA	0.083	0.083	NA	3.01	NA	0.147
<i>Tapirira guianensis</i> Aubl.	119	0.20	1.59	NA	NA	NA	NA	0.76	0.32	NA	0.003	0.003	NA	8.60	42.91	0.001
<i>Tapirira obtusa</i> (Benth.) J.D.Mitch.	75	0.25	1.19	0.76	1.53	0.27	0.35	0.78	0.38	0.058	0.020	0.020	9.13	5.41	28.96	0.452
<i>Theobroma subincanum</i> Mart.	45	0.06	0.23	NA	NA	NA	NA	NA	0.47	NA	NA	0.000	NA	NA	21.71	NA
<i>Tovomita brasiliensis</i> (Mart.) Walp.	54	0.10	0.08	2.24	1.54	3.72	1.79	NA	0.55	0.081	NA	0.081	8.31	NA	18.77	0.191
<i>Tovomita brevistaminea</i> Engl.	108	0.01	0.14	NA	NA	NA	NA	NA	0.69	NA	NA	0.007	NA	NA	7.34	NA
<i>Tovomita</i> sp. 2	62	0.02	0.13	NA	NA	NA	NA	NA	0.60	NA	NA	0.001	NA	NA	10.35	NA
<i>Trichilia schomburgkii</i> C.DC.	87	0.19	0.53	NA	NA	NA	NA	1.19	0.33	NA	0.002	0.002	NA	9.56	56.46	0.004
<i>Unonopsis rufescens</i> (Baill.) R.E.Fr.	53	0.02	0.11	NA	NA	NA	NA	NA	0.63	NA	NA	0.006	NA	NA	7.67	NA
<i>Unonopsis stipitata</i> Diels	93	0.04	0.10	1.50	2.69	1.74	2.64	NA	0.62	0.065	NA	0.065	8.83	NA	11.49	0.249
<i>Virola michelii</i> Heckel	93	0.17	0.13	1.93	8.55	1.40	2.24	0.87	0.57	0.000	0.068	0.000	37.33	3.33	11.70	0.001
<i>Vismia guianensis</i> (Aubl.) Choisy	55	0.40	1.39	2.69	0.01	3.14	0.86	0.32	0.32	0.121	0.000	0.000	7.29	16.64	27.56	0.025
<i>Vismia sessilifolia</i> (Aubl.) Choisy	162	0.38	7.40	NA	NA	NA	NA	0.40	0.19	NA	0.000	0.000	NA	33.80	46.17	0.000
<i>Votomita guianensis</i> Aubl.	199	0.12	0.90	NA	NA	NA	NA	NA	0.40	NA	NA	0.000	NA	NA	50.49	NA
<i>Xylopia nitida</i> Dunal	255	0.35	5.73	1.34	2.83	0.95	1.63	0.61	0.20	0.005	0.000	0.000	14.70	29.59	64.87	0.000

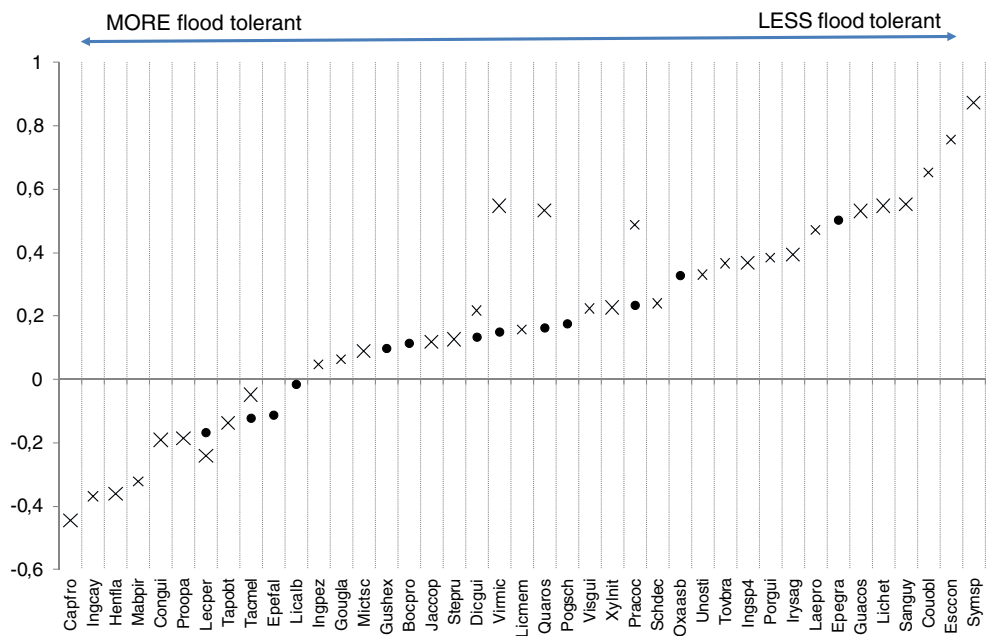
R<sup>2</sup> = Nagelkerke index. Odds ratio, P values and χ<sup>2</sup> statistics are given for the following factors or factor levels a/soil type: down-hill system (DhS), alloterite <1.2 m (Alt), superficial lateral drainage (SLD), up-hill system (UhS), superficial hydromorphy (SH) is taken to be the reference b/distance to logging damage (logscale) c/distance to conspecific (Clump, logscale). Env is the minimum of soil (drainage) and disturbance (Dist). Last column (P Torus), reports the frequency of lower residual deviances found when model is applied after random torus translation of environmental factors map with regard to stem map

**Table 2** Summary statistics for the logistic model of presence/absence of 15 selected species in one 25 ha block in the Paracou experimental site

Taxon	Abund.	R <sup>2</sup>	Interc.	OddDHS	OddAlt	OddSLD	OddUHS	OddDist	OddClump	PSoil	PDist	PEnv	PClump	χ <sup>2</sup> Soil	χ <sup>2</sup> Dist	χ <sup>2</sup> Clump	PTorus
<i>Bocoa prouaensis</i> Aubl.	1,023	0.02	0.22	1.14	1.27	1.35	1.43	1.07	0.55	0.019	0.001	0.001	0.000	11.79	10.46	60.19	0.008
<i>Carapa procera</i> DC.	270	0.02	0.15	NA	NA	NA	NA	0.89	0.53	NA	0.010	0.010	0.000	NA	6.63	31.70	0.377
<i>Dicorynia guianensis</i> Amshoff	940	0.18	2.21	1.23	1.13	1.79	1.28	NA	0.19	0.000	NA	0.000	0.000	20.90	NA	383.57	0.115
<i>Eperua falcata</i> Aubl.	1,455	0.35	22.52	1.07	0.59	0.80	0.63	0.93	0.09	0.000	0.003	0.000	0.000	21.22	8.87	469.02	0.349
<i>Eperua grandiflora</i> (Aubl.) Benth.	1,534	0.28	1.70	1.72	4.85	3.20	2.98	1.16	0.15	0.000	0.000	0.000	0.000	95.09	46.47	424.27	0.000
<i>Gustavia hexapetala</i> (Aubl.) Sm.	497	0.03	0.14	1.51	0.64	1.07	1.80	0.92	0.56	0.000	0.015	0.000	0.000	24.21	5.88	38.55	0.041
<i>Lecythis persistens</i> Sagot subsp. <i>persistens</i>	4,484	0.07	6.78	0.80	0.58	0.62	0.73	1.05	0.21	0.000	0.000	0.000	0.000	54.42	12.98	282.93	0.006
<i>Licania alba</i> (Bernoulli) Cuatrec.	2,713	0.03	1.31	1.21	0.67	0.97	1.00	NA	0.39	0.000	NA	0.000	0.000	26.45	NA	139.34	0.029
<i>Oxandra asbeckii</i> (Pulle) R.E. Fr.	5,761	0.23	9.10	1.57	1.87	2.57	2.50	1.20	0.10	0.000	0.000	0.000	0.000	152.28	167.27	551.48	0.000
<i>Pogonophora schomburgkiana</i> Miens ex Benth.	3,540	0.12	4.29	1.53	1.39	1.57	1.51	1.04	0.17	0.000	0.013	0.000	0.000	28.16	6.16	448.32	0.006
<i>Pradosia cochlearia</i> (Lecomte) T.D Penn.	1,387	0.04	0.24	1.62	1.77	1.64	1.83	1.12	0.48	0.000	0.000	0.000	0.000	26.16	30.04	85.75	0.000
<i>Qualtea rosea</i> Aubl.	1,701	0.38	26.24	1.09	1.86	1.17	1.69	0.93	0.08	0.000	0.001	0.000	0.000	24.41	10.41	547.69	0.491
<i>Sextonia rubra</i> (Mez) van der Werff	182	0.01	0.04	0.71	1.08	0.41	0.69	NA	0.77	0.034	NA	0.034	0.049	10.42	NA	3.88	0.027
<i>Tachigali melinonii</i> (Harms) Zarucchi & Herend.	894	0.20	4.21	1.01	0.87	0.43	0.72	0.83	0.19	0.001	0.000	0.000	0.000	19.35	20.30	318.31	0.013
<i>Virola michelii</i> Heckel	2,238	0.10	1.78	1.07	1.51	1.53	1.53	0.86	0.30	0.000	0.000	0.000	0.000	38.37	51.59	220.04	0.025

R<sup>2</sup> = Nagelkerke index. Odds ratio, P values, and χ<sup>2</sup> statistics are given for the following factors or factor levels a/ soil type; down-hill system (DhS), alloterite <1.2 m (Alt), superficial lateral drainage (SLD), up-hill system (UHS); superficial hydromorphy (SH) is taken to be the reference b/distance to logging damage (logscale) c/distance to conspecific (Clump, logscale). PEnv is the minimum of PSoil (drainage) and disturbance (PDist). Last column (PTorus) reports the frequency of lower residual deviances found when model is applied after random torus translation of environmental factors map with regard to stem map

**Fig. 3** Ranking of species tolerance to soil hydromorphy, expressed as minus log of the average of the odds ratio over all non-SH soil types (smaller values indicate larger tolerance). *Crosses* dataset 1 (transects), smaller size if PSoil >0.05; *dots* dataset 2 (large plots); see text for details. Species code uses the first three letter of genus name concatenated with the first three letters of species name (full names are given in tables)



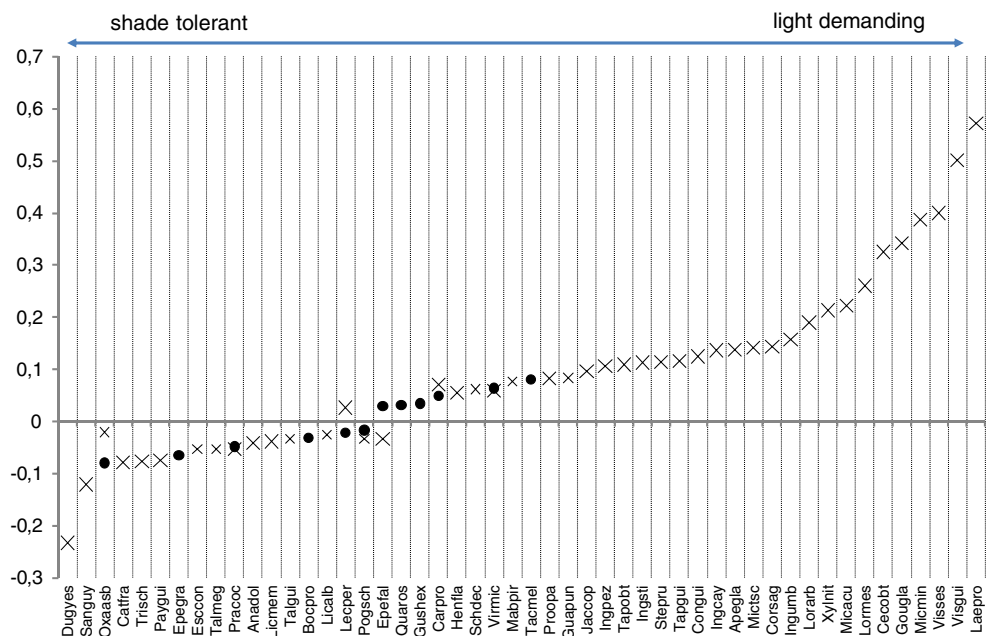
tion. OddClump which captures this spatial dependency appears to be consistent between the two datasets despite the fairly different spatial sampling pattern (Pearson’s coefficient correlation  $r=0.81$ ;  $P<0.001$ ; Bartlett Chi square test,  $n=13$ ). There was only one species for which spatial correlation was not picked-up in dataset 1 while it was in dataset 2 (*Carapa procera* DC.).

This large effect of spatial correlation may be related to the size of stems targeted in this study (and will also depend on the elementary quadrat size taken to be  $5 \times 5$  m in the present study). The spatial correlation is expected to be

stronger for saplings than for trees as population thinning will gradually weaken the pattern originating from individual trees seed dispersal kernel. Hence, the relative importance of spatial correlation and habitat specialisation is likely to differ between saplings and adults. It will also be affected by the overall degree of disturbance.

Introducing a clumping factor in the logistic model reduced the number of “significant” species-habitat association from 74% to 42% thereby strongly limiting the number of spurious associations. In one case of very restricted distribution (*Duguetia yeshidan* Sandwith, a

**Fig. 4** Ranking of species degree of heliophily, expressed as minus log of odd Dist (smaller value indicate lower tolerance to canopy disturbance). *Crosses* dataset 1 (transects), smaller size if Pdist>0.05; *dots* dataset 2 (large plots); see text for details. Species code uses the first three letter of genus name concatenated with the first three letters of species name (full names are given in Tables 1 and 2)



treelet found in two plots only) introducing this clumping factor allowed to reveal its sensitivity to canopy openness (Table 1) which was otherwise missed out (data not shown).

Spatial aggregation of conspecific stems may occur at different spatial scales indeed and may directly reflect dispersal limitation such as clumping of saplings around a mother tree, or larger aggregates known to occur for many tropical species (Traissac 2003). Aggregation may however also be induced by latent environmental factors. Statistically significant clumping can also be a consequence of imprecise or ill-coded predictors of environmental factors included in the model. For instance, *C. obtusa* is considered to be an extremely efficient disperser: its seeds are dispersed by bats and may survive many years in the soil until a canopy gap occurs and they will germinate. They are extremely abundant and widespread (de Foresta and Prévost 1986). We found that the distribution of *C. obtusa* was largely determined by canopy disturbance (Table 1) as expected but that the species still showed a significant degree of clumping in our model. Since the species tends to be clumped in canopy gaps this spatial correlation may be a consequence of unmapped natural gaps which were not included in the model. The presence of a group of stems not associated to a recorded disturbance zone will increase the deviance captured by the clumping factor. Mapping and coding of disturbance may also be prone to errors even in the logged-over areas. Consequently, spatial correlation might also have been inflated as a consequence of inaccurate canopy disturbance mapping.

The final rate of positive species  $\times$  habitat association was 20% in dataset 1 and 67% in dataset 2. The latter was characterised by a higher sampling pressure and a stronger spatial coherence (four large square plots covering 25 ha in total included in a 40 ha compact area, Fig. 1 in the Electronic supplementary material in Molino and Sabatier 2001). The true percentage of species-habitat association in dataset 1 is most probably much larger than the percentage reported here.

For a given area sampled, the statistical power of the test, i.e. the probability of detecting an existing species  $\times$  habitat association, will increase with species abundance. This is reflected by the highly significant negative rank correlation between species abundance and PEnv (Spearman's  $\rho = -0.401$ ;  $P < 0.001$ ) in dataset 1. It may be true that more common species are more frequently associated to particular habitat though there is no definite ecological argument why this should be the case. By including species with as few as 40 stems tallied in 5 ha in our analysis, we accepted a high risk of not rejecting  $H_0$  even though it was false (type II error).

Everything else being equal, the power of the statistical test will also increase with the area sampled. Higher sampling pressure also affected the higher discovery rate

in dataset 2. Positive associations detected in dataset 1 were consistently detected in dataset 2 for the 14 species common to the two datasets (column PSoil and PDist in Tables 1 and 2). However, lower sampling intensity in dataset 1 did limit the power of the analysis and half of the associations detected in dataset 2 were missed in dataset 1 (Tables 1 and 2): only five of the ten species found to be preferentially associated to a particular habitat were also detected in dataset 1. For some species, only one significant environmental factor was detected in dataset 1 while both were found to be significant when using dataset 2 (e.g. *Tachigali melinonii* (Harms) Zarucchi & Herend.).

Finally, as we applied a multiple test correction to maintain a family wise type I error rate of 5%, we further increased the type II error and more drastically so for dataset 1 which had more species tested.

As mentioned above the accuracy of spatial predictors as well as their precision might have been limiting. The large range of dbh and the rather long time since logging might have blurred the floristic changes which have occurred following logging.

Nonetheless, despite the aforementioned limitations, this study clearly indicates that ecological determinism applies to sapling distribution of a large proportion of species in the Paracou forest.

In addition to different optimum environmental settings, species clearly exhibited a variety of niche breadth as they were more or less strongly affected by the environmental factors examined here, some abundant species being apparently unaffected either by canopy openness or drainage regime.

Similar conclusions have been reached in a series of studies in the tropics. A study conducted in the Barro Colorado Island 50 ha plot (Harms et al. 2001) found that 64% of the 171 most abundant species showed significant association (either positive or negative) with a particular habitat. In the latter study, all stems above 1 cm dbh were censused. The rate of species-habitat association detected was remarkably high given that habitats were rather crudely defined as a combination of slope and elevation class (in addition to a small patch of secondary forest, stream-sides and a swamp which were treated as separate habitats). This high detection rate however does not include correction for multiple testing. No evidence was found that rarer and commoner species may vary in the degree to which they are associated with habitats.

John et al. (2007), using data collected in three neotropical forest plots in Colombia (La Planada, 25 ha), Ecuador (Yasuni, 25 ha), and again Panama (Barro Colorado Island, 50 ha) also showed that the spatial distributions of 29% to 40% of tree species were strongly associated with the distributions of soil nutrients at all three sites.

Webb and Peart (2000) sampled trees (in 28 plots totalling 4.5 ha and seedlings ( $< 1$  cm diameter) in 28 sub-plots

totalling 0.1 ha across a study area of 150 ha of rain forest in Borneo. Combining light (assessed by hemispherical photographs) and physiographic habitat (defined as topographic position with associated moisture and soil regime), 20 out of the 45 abundant species were associated with at least one habitat factor as either adults or seedlings. They also report stronger physiographic habitat specificity at tree stage than at seedling stage and, less expectedly, a fairly low degree of consistency between adult and seedling—physiographic habitat associations.

The species-habitat associations detected in the present study (as in the studies mentioned above) may be biased if unmapped latent environmental factors significantly affect species distributions or as a consequence of errors in mapping of the spatial predictors. A first test of robustness of our study is provided by comparing the results of both datasets for the shared species (Figs. 3 and 4). A few discrepancies between the two datasets indicate that the precise strength of some associations may be poorly estimated in some cases. Notably the abundance threshold of eight stems per ha that was used with dataset 1 may have been too permissive. However the overall consistency between the two datasets is good in as much as one is interested in positive or negative associations with environmental factors and clumping intensity. A second test of robustness of the species-habitat associations was conducted by comparing the species ranking along ecological gradients arrived at in the present study with previously published independent studies.

To assess the validity of the plant-soil associations detected, we compared our results with the findings of Pélissier et al. (2002) at a nearby site (Piste de Saint Elie Research station). This study led to the ranking of 118 tree species site along a main gradient of tolerance to prolonged water saturation. The study conducted at the site used a detailed soil classification in which nine units were distinguished, including well-drained soils (deep vertical drainage unit—DVD). Although simpler and reflecting a shorter gradient (DVD are rare at Paracou and absent from our dataset), our 5-unit classification is directly superimposable to this 9-unit classification. For each of our datasets, and for the species they respectively share with Pélissier et al.'s study (2002), we compared the ordering along Pélissier et al.'s gradient of tolerance with that obtained with oddAlt values in our study. OddAlt values reflect species' attractiveness (or repulsion) for the soils with the best drainage in Paracou (Alt). Among the 38 species in dataset 1 that appear in the Piste de St Elie dataset, ten are sensitive to soil type ( $P_{\text{Soil}} < 0.05$  and  $P_{\text{Torus}} < 0.05$ ). Their ranking according to oddAlt appeared significantly correlated to their ordering in Pélissier et al.'s analysis (Spearman's  $\rho = 0.71$ ,  $P = 0.02$ ).

As regards plant—canopy openness relationships, we compared our results to the list of pioneer and heliophilic tree species found in Paracou (Molino and Sabatier 2001). Among the 28 species of dataset 1 that appeared to be affected by the proximity to past logging disturbance ( $P_{\text{Dist}} < 0.05$ ), only those (and all of them) that were positively linked ( $\text{OddDist} < 1$ ) appeared in Molino and Sabatier's list (2001). Meanwhile, in dataset 2, the only two species that appear very significantly positively affected by the proximity to past logging disturbance ( $P_{\text{Dist}} < 10^{-5}$ ,  $\text{OddDist} = 0.83$  and  $0.86$ ) were also previously classified as pioneer or heliophilic (Molino and Sabatier 2001).

Overall our results regarding species sensitivity to canopy disturbance and drainage restriction are largely consistent with previously published observations. Minor discrepancies exist which can result from one or more of the following sources. Firstly, difference in strength of species-habitat association is likely to occur with change in development stage (Webb and Peart 2000). This may explain some of the observed differences in preferred drainage regime found with Pélissier et al.'s study (2002) which dealt exclusively with stem  $> 10$  dbh. Secondly, most other studies do not analyse jointly the two environmental gradients considered here, hence potentially leading to a biased estimate of the single gradient considered as those gradients are usually not independent. For instance poorly drained seasonally flooded bottomland have been shown to have a higher turnover rate (Madelaine et al. 2007) and a more open canopy (Vincent et al. 2010) than the surrounding forest. In our dataset, the reverse may be true since canopy disturbance was artificial (bottomlands with superficial hydromorphy may have been less accessible and hence have a lower level of canopy disturbance than more accessible parts). In any case taking into account both factors as we did should have improved the estimate of the effect of each factor. Thirdly, differences in the environmental gradient covered in the different studies may affect the results of the studies. Fourthly, as suggested earlier some species may be ill sorted due to low abundance. As a rule of thumb, the more abundant the species (Table 1) the more reliable the ecological profile may be.

## 5 Conclusions

Our results are consistent with a growing number of independent studies indicating that despite their high alpha-diversity local community assemblages of tropical saplings are strongly spatially organised. Dispersal limitation is the most constraining factor for a majority of species at the scale examined here (a patch of forest less than  $3 \text{ km}^2$ ) and for the community as a whole.

Canopy disturbance regime and drainage regime also both play a significant role for many species as respectively 20% and 67% of the species examined in datasets 1 and 2 were found to be sensitive to either drainage regime or canopy disturbance regime or both. The ecological factors examined most probably affected the distribution of a majority of the species occurring in this forest (including those rare species that were not tested). This is suggested by higher detection of association in abundant species and higher detection of association when sampling pressure is increased (dataset 2 vs. dataset 1).

Finally, the species preferences for soil type or degree of canopy disturbance found in our analyses are largely congruent with former studies suggesting that both the methodology used and the results obtained are robust. The ranking proposed should however be taken as a first attempt and is likely to be modified and refined in the future. In particular, it is recommendable to consider results on species-soil-type associations obtained from samples with less than 150 stems per species as preliminary. We hope, however, that this work will contribute to sorting out the extraordinary functional diversity of neotropical moist forests.

**Acknowledgements** The authors thank François Munoz for comments on an earlier draft and Vivien Rossi for helpful discussions on statistical issues. Many thanks to Lilian Blanc for making available sapling inventory data (dataset 2) and the canopy disturbance map of the CIRAD Paracou experimental site. This study benefited from partial funding by the European Regional Development Fund (convention n° 2828).

## References

- Baraloto C, Morneau F, Bonal D, Blanc L, Ferry B (2007) Seasonal water stress tolerance and habitat associations within four neotropical tree genera. *Ecology* 88(2):478–489
- Barthès B (1991) Influence des caractères pédologiques sur la répartition spatiale de deux espèces du genre *Eperua* (Caesalpinaceae) en forêt guyanaise. *Revue d'écologie* 46(4):303–320
- Bell G (2005) The co-distribution of species in relation to the neutral theory of community ecology. *Ecology* 86:14
- Boulet R (1978) Existence de systèmes à forte différenciation latérale en milieu ferrallitique guyanais: un nouvel exemple de couverture pédologique en déséquilibre. *Sci Sol* 2:75–82
- Boulet R (1990) Organisation des couvertures pédologiques des bassins versants ECEREX. Hypothèses sur leur dynamique. In: Sarrailh JM (ed) *Mise en valeur de l'écosystème forestier guyanais. Opération ECEREX*, INRA, CTFT, Paris, Nogent-sur-Mame. pp. 15–45
- Condit R, Ashton PS, Baker P, Bunyavejchewin S, Gunatilleke S, Gunatilleke N, Hubbell SP, Foster RB, Itoh A, LaFrankie JV, Lee HS, Losos E, Manokaran N, Sukumar R, Yamakura T (2000) Spatial patterns in the distribution of tropical tree species. *Science* 288:1414–1418
- Dalling JW, Muller-Landau HC, Wright SJ, Hubbell SP (2002) Role of dispersal in the recruitment limitation of neotropical pioneer species. *J Ecol* 90(4):714–727
- de Foresta H, Prévost MF (1986) Végétation pionnière et graines du sol en forêt guyanaise. *Biotropica* 18(4):279–286
- Epron D, Bosc A, Bonal D, Freycon V (2006) Spatial variation of soil respiration across a topographic gradient in a tropical rain forest in French Guiana. *J Trop Ecol* 22(05):565–574
- Flores O (2005) Déterminisme de la régénération chez quinze espèces d'arbres tropicaux en forêt guyanaise: les effets de l'environnement et de la limitation par la dispersion, PhD thesis, Biologie de l'Evolution et Ecologie. Université des Sciences et Techniques, Montpellier, p 166
- Gentry AH (1988) Tree species richness of upper Amazonian forests. *Proc Natl Acad Sci USA* 85(1):156–159
- Gourlet-Fleury S, Ferry B, Molino JF, Petronelli P, Schmitt L (eds) (2004) *Ecology and Management of a Neotropical Rainforest. Lessons drawn from Paracou, a long-term experimental research site in French Guiana*. Elsevier, Paris, pp 3–60
- Harms KE, Condit R, Hubbell SP, Foster RB (2001) Habitat associations of trees and shrubs in a 50-ha neotropical forest plot. *J Ecol* 89(6):947–959
- Harrell FE (2001) *Regression modeling strategies (with applications to linear models, logistic regression, and survival analysis)*. Springer, New York, p 568
- Hubbell SP, Foster RB (1986) Biology, chance, and the history and structure of tropical rain forest tree communities. In: Diamond J, Case TJ (eds) *Community ecology*. Harper and Row, New York, pp 314–329
- John R, Dalling JW, Harms KE, Yavitt JB, Stallard RF, Mirabello M, Hubbell SP, Valencia R, Navarrete H, Vallejo M, Foster RB (2007) Soil nutrients influence spatial distributions of tropical tree species. *Proc Natl Acad Sci USA* 104(3):864–869
- Keitt TH, Bjornstad ON, Dixon PM, Citron-Pousty S (2002) Accounting for spatial pattern when modeling organism-environment interactions. *Ecography* 25(5):616–625
- Madeline C, Pélissier R, Vincent G, Molino JF, Sabatier D, Prévost MF, de Namur C (2007) Mortality and recruitment in a lowland tropical rain forest of French Guiana: effects of soil type and species guild. *J Trop Ecol* 23(03):277–287
- Molino JF, Sabatier D (2001) Tree diversity in tropical rain forests: a validation of the intermediate disturbance hypothesis. *Science* 294(5547):1702–1704
- Pélissier R, Dray S, Sabatier D (2002) Within-plot relationships between tree species occurrences and hydrological soil constraints: an example in French Guiana investigated through canonical correlation analysis. *Plant Ecol* 162(2):143–156
- Phillips OL, Vargas PN, Monteagudo AL, Cruz AP, Chuspezans M-E, Washington Galiano S, Yli-Halla M, Rose S (2003) Habitat association among Amazonian Tree Species: a Landscape-Scale Approach. *J Ecol* 91(5):757–775
- Poorter L, Arets EJMM (2003) Light environment and tree strategies in a Bolivian tropical moist forest: an evaluation of the light partitioning hypothesis. *Plant Ecol* 166(2):295–306
- Pulliam HR (2000) On the relationship between niche and distribution. *Ecol Lett* 3(4):349–361
- Sabatier D, Grimaldi M, Prévost M-F, Guillaume J, Gordon 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
- Sollins P (1998) Factors influencing species composition in tropical lowland rain forest: does soil matter? *Ecology* 79:23–30
- ter Steege H, Sabatier D, Castellanos H, Andel T v, Duivenvoorden JF, Oliveira AAd, Ek RC, Lilwah R, Maas PJM, Mori SA (2000) An analysis of the floristic composition and diversity of Amazonian forests including those of the Guiana shield. *J Trop Ecol* 16:801–828

- Traissac S (2003) Dynamique spatiale de *Vouacapoua americana* (Aublet), arbre de forêt tropicale humide à répartition agrégée. Ph.D. thesis, Laboratoire de Biométrie et de Biologie Evolutive UMR 5558, Université Claude Bernard Lyon I, Lyon. p 230
- Valencia R, Balslev H, Miño G Paz y (1994) High tree alpha-diversity in Amazonian Ecuador. *Biodivers Conserv* 3(1):21–29
- Vincent G, Weissenbacher E, Sabatier D, Blanc L, Proisy C, Couteron P (2010) Détection des variations de structure de peuplements en forêt dense tropicale humide par Lidar aéroporté (Small foot-print airborne LiDAR proves highly sensitive to changes in structure of moist tropical forest). *Revue Française de Photogrammétrie et Télédétection* 191:42–51
- Webb CO, Peart DR (2000) Habitat associations of trees and seedlings in a Bornean rain forest. *J Ecol* 88(3):464–478
- Wright SJ, Muller-Landau HC, Condit R, Hubbell SP (2003) Gap-dependent recruitment, realized vital rates and size distributions of tropical trees. *Ecology* 84(12):3174–3185