

Analysis of tropical forest inventory data

Contents

1 Datasets	2
1.1 Paracou	2
1.2 Uppangala	3
1.3 BCI	4
2 Data preparation	5
3 Estimation of IV	6
3.1 Paracou	7
3.2 Uppangala	8
3.3 BCI	10
4 Moran's I analysis	13
5 Is variance within conspecifics inferior to variance within heterospecifics ?	14
6 Other tests and verifications	16
6.1 Undisturbed plots in the Paracou site	16
6.2 Smaller stems in the BCI site	16
7 Code implementation	16
References	17

```
knitr::opts_chunk$set(echo=FALSE, fig.pos='H')
```

In this section, we aimed at exploring the intraspecific variability of tree growth in three hyperdiverse mature forests, encompassing contrasting climates and disturbance regimes across the Tropics. First, for each dataset, we estimated the intraspecific variance of annualized growth using a hierarchical bayesian model. Then, we tested if spatial structure in individual mean growth could be detected. Finally, we tested if intraspecific variability was higher than interspecific variability in individual mean growth.

Note that we here used data on tree growth, which is one component of individual fitness and an imperfect proxy of tree competitive ability, and our analyses could be further tested with other traits in the future.

1 Datasets

1.1 Paracou

The Paracou forest is located in French Guiana. It is one of the best-studied lowland tropical forests in the Guiana Shield region. It belongs to the Caesalpiniaceae facies and has amongst the highest alpha-diversity in the Guiana shield with 150-200 tree species per hectare in inventories of trees with diameter at breast height (DBH) ≥ 10 cm. The studied 93.75-ha area harbours 768 species, 254 genera and 65 families of plants. The Guiana Shield is characterized by Pre-Cambrian granitic and metamorphic geological formations, highly eroded. It is associated with gently undulating landscapes and a very dense hydrographic system. Paracou forest lies in a hilly area, on a formation called “série Armina” characterized by schists and sandstones and locally crossed by veins of pegmatite, aplite and quartz. The topography of the site consists of small hills separated by narrow (< 5 m wide) sandy waterbeds. The altitude varies from 5 to about 45 m above sea level [2]. The mean annual temperature is 26 °C and winds are generally weak. There is a well marked dry season (from mid-August to mid-November) and a long rain season with a short drier period between March and April (mean annual rainfall of 3,041 mm). Different study programs have been led at the Paracou site (<https://paracou.cirad.fr/>). Here, we used data from a disturbance experiment, where 15 plots of 6.25 ha were exposed to four different logging intensities between 1986 and 1988. Since then, cartesian coordinates, DBH, species identity and survival of each tree with a DBH ≥ 10 cm have been collected every one or two years, during the dry season (mid-August to mid-November). The final dataset contains ca. 46,000 tree measurements.

Figure 1.1 shows the location of the trees in the Paracou site.

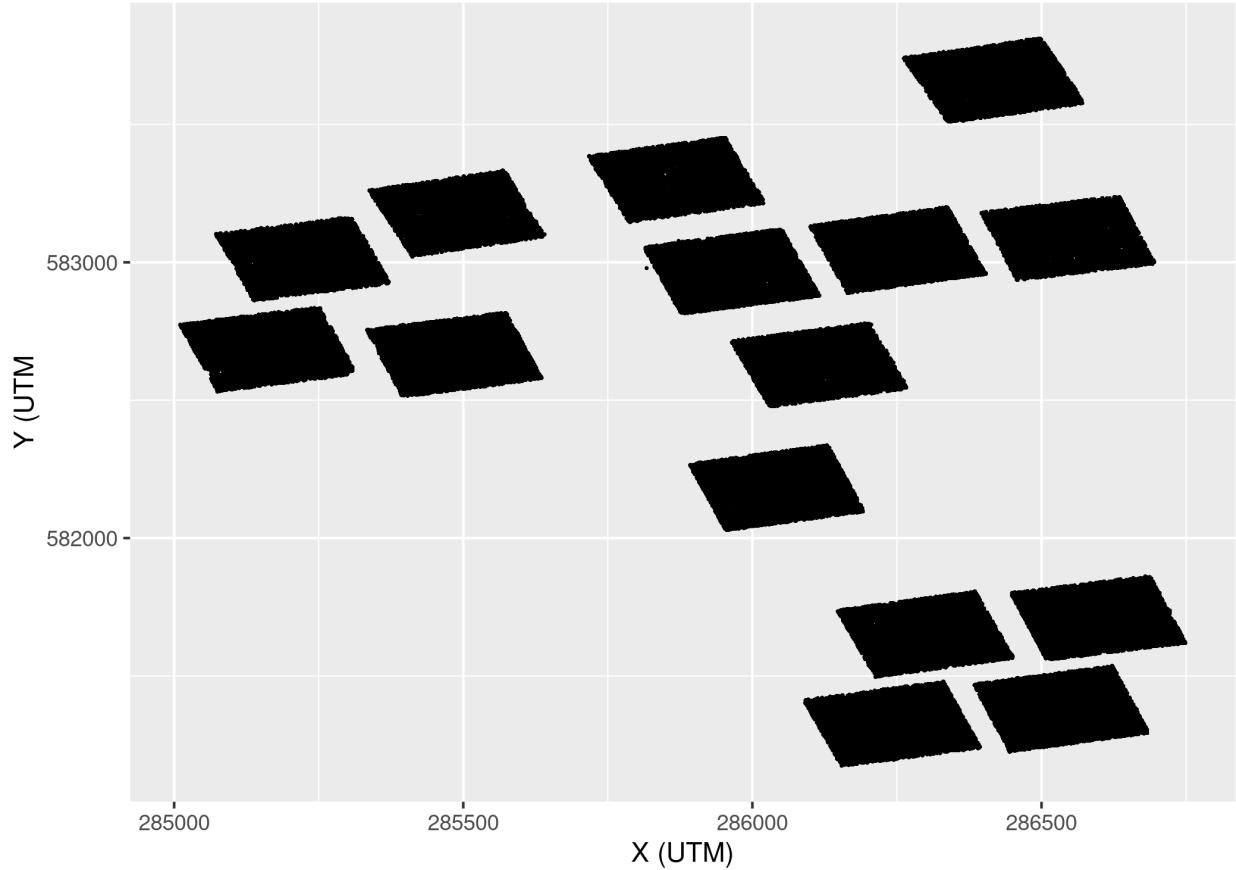


Figure 1.1: Plots at the Paracou site. Each point is a tree.

1.2 Uppangala

The Uppangala Permanent Sample Plot (UPSP) is located in South-East Asia, in the Western Ghats of India, and was established in 1989 by the French Institute of Pondicherry in the Kadamakal Reserve Forest, in the Pushpagiri Wildlife Sanctuary, in Karnataka state, India [3]. It is a low altitude (500-600 m) wet evergreen monsoon Dipterocarp forest [4]. The studied area, of 5.07 ha, is quite steep, with a mean slope angle of about 30–35°. The plots consist in five North–South oriented transects that are 20-m wide, 180-to 370-m long, and 100-m apart center to center, in addition to three rectangular plots which overlap the transects. The transects gather data from 1990 to 2011 and the rectangular plots from 1993 to 2011. The trees with GBH (Girth at Breast Height) ≥ 30 cm (equivalent to ca. 9.5 cm DBH) were measured every 3 to 5 years. The final dataset contains measurements of 3,870 trees of 102 species (including 2 morphospecies). This forest is considered as one of the rare undisturbed tropical forests in the world [5].

Figure 1.2 shows the location of the trees in the Uppangala site.

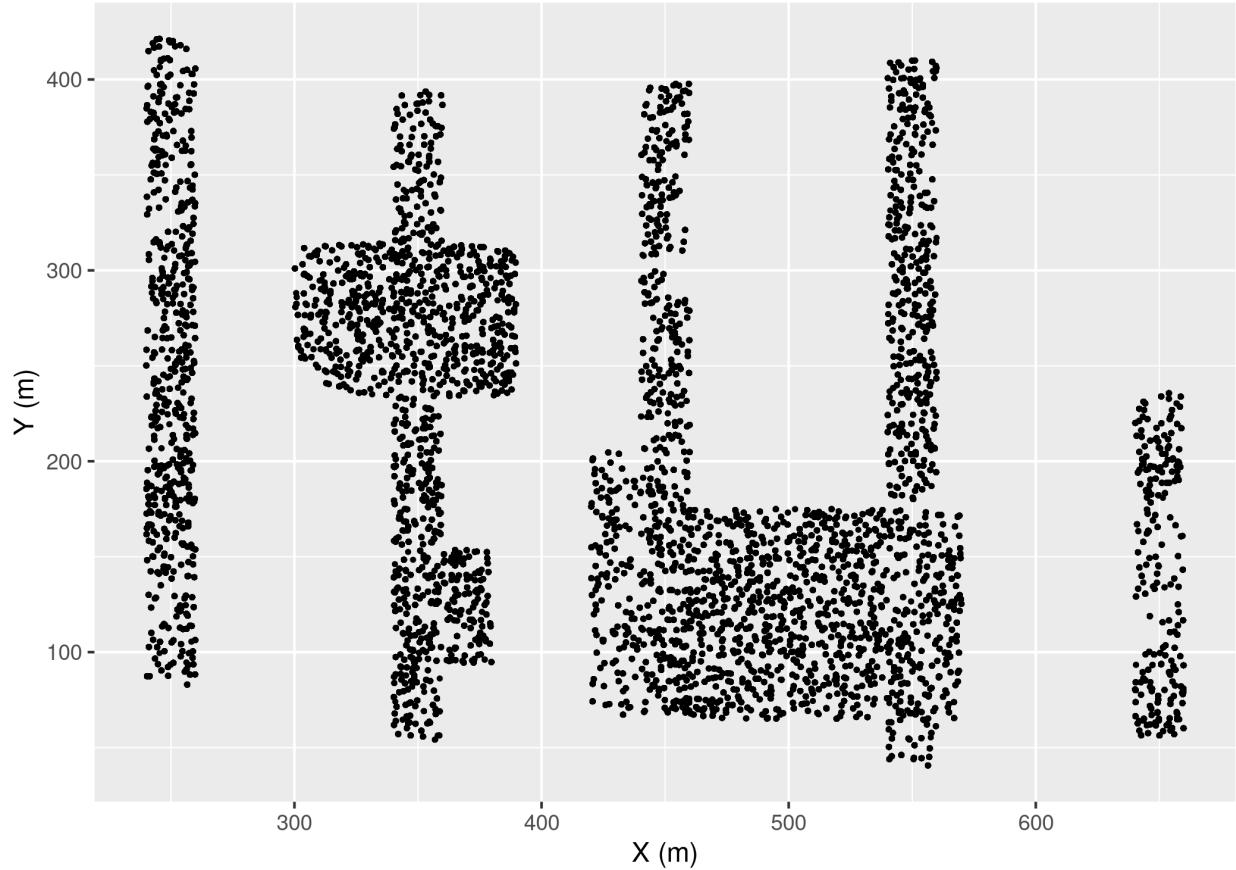


Figure 1.2: Plots at Uppangala site. Each point is a tree.

1.3 BCI

The Barro Colorado Island site is located in central America, in Panama, covered by a lowland tropical moist forest. The zone became an island after a valley was flooded in order to build the Panama Canal, in 1913. It is nowadays considered as the most intensively studied tropical forest in the world. The studied site is a 50 ha plot (500×1000 m). It has an elevation of 120 m and is quite flat (most slopes are gentler than 10°). Complete censuses of all trees with DBH ≥ 1 cm have been performed every 5 years since 1980. The dataset contains measurements of 328 tree species and 423617 trees, and when only taking into account trees with DBH ≥ 10 cm, it contains measurements of 255 species and 37224 trees. The dataset is available at [6].

Figure 1.3 shows the location of the selected trees in the BCI site.

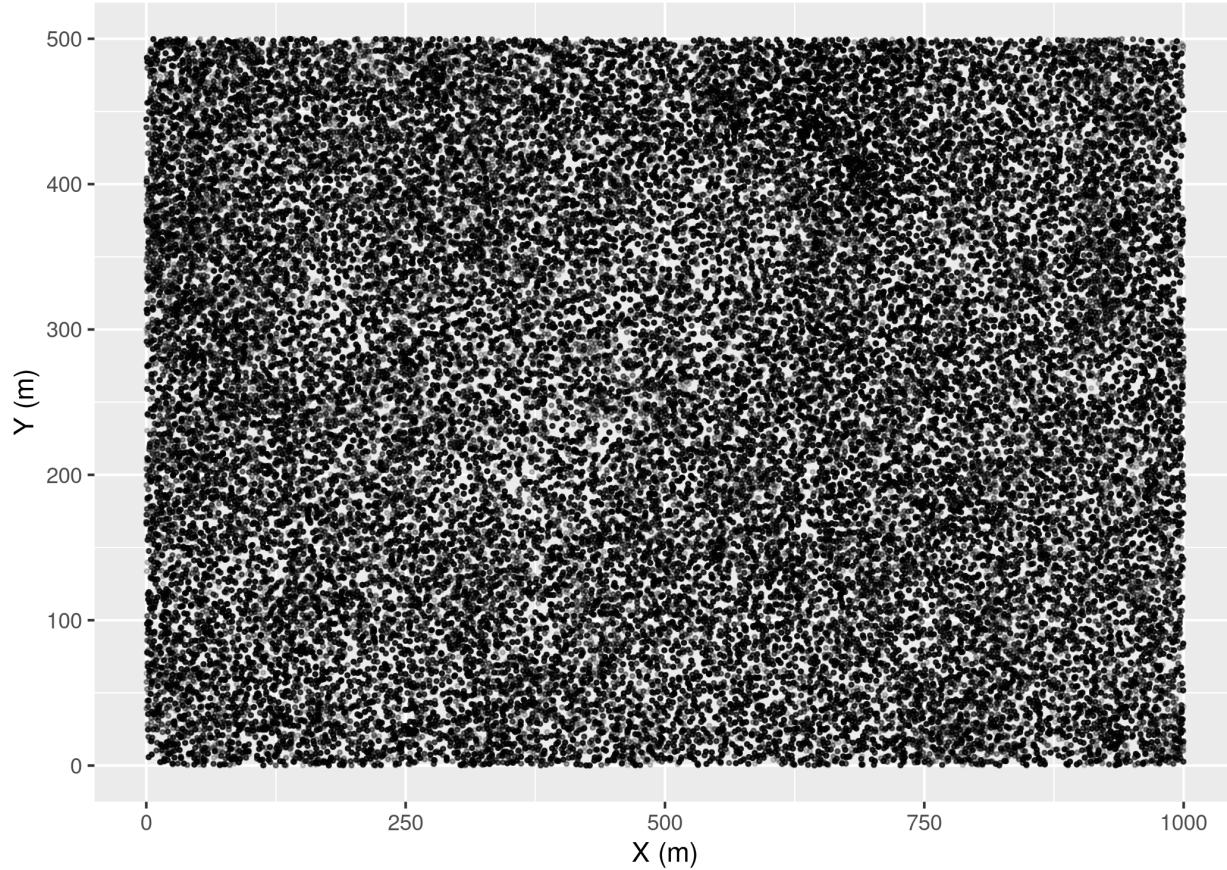


Figure 1.3: The 50 ha plot of the Barro Colorado Island site. Each point is a tree.

2 Data preparation

For the Paracou site, we used the plots 1 to 15 of the Paracou site. We removed all trees of undetermined species and all measures of the years before the perturbations were performed and the biodiversity plots were added (1985-1991).

For the Uppangala site, we removed the census dates which were not common for all plots.

For the BCI site, we removed the trees with missing DBH measures and trees with $\text{DBH} < 10 \text{ cm}$ for consistency with the Paracou and Uppangala sites.

For all three datasets, we computed annualized growth between two censuses as the difference of DBH (Diameter at Breast Height) between two consecutive censuses, divided by the time period between those two censuses. We then removed estimates with a growth $< -2 \text{ mm/year}$ or $> 100 \text{ mm/year}$.

for the Paracou site, after removing `nrow(Paracou_no_mean[which(Paracou_no_mean$Sp=="Indet_sp"),])` growth estimates for 12629 trees of indeterminate species, this leads to a dataset with 945571 growth estimates for 65805 trees of 614 species.

for the Uppangala site, this leads to a dataset with 57921 growth estimates for 3725 trees of 102 species.

for the BCI site, this leads to a dataset with 167618 growth estimates for 30386 trees of 244 species.

For all three datasets, we finally computed the mean annual growth for each individual tree as the difference of DBH between the first and the last time a tree was measured, divided by the period between those two measurements.

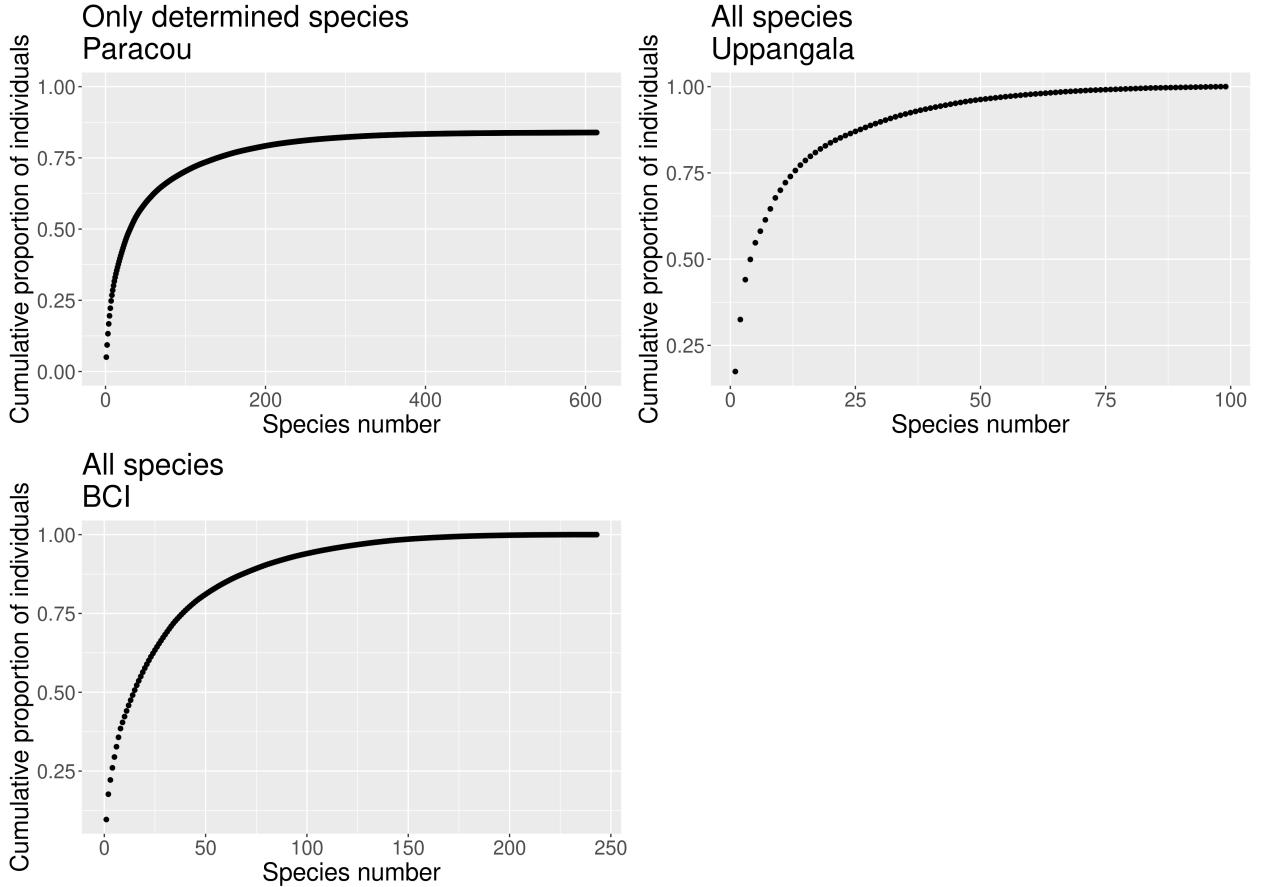


Figure 2.1: Abundance diagrams for the three study sites. For the Paracou site, only individuals that were identified to the species level are accounted for. The three communities include few dominant species and many rare species (long asymptote).

3 Estimation of IV

To quantify the relative importance of intraspecific and interspecific variability for each site, we used Stan to run a Bayesian hierarchical model, with a species random effect and an individual random effect on the intercept. In order to compare to classic growth models, and estimate an intraspecific variability within diameter classes, we also added an intercept and a diameter fixed effect. We used the brms package [7,8] to implement the model. We scaled the data to improve convergence time.

$$\ln(G_{ijt} + 2) = [\beta_0 + b_{0j} + d_{0i}] + \beta_1 \times \ln(D_{ijt}) + \epsilon_{ijt},$$

Priors

$$\beta_0 \sim \mathcal{N}(\text{mean} = 0, \text{var} = 1), \text{iid}$$

$$\beta_1 \sim \mathcal{N}(\text{mean} = 0, \text{var} = 1), \text{iid}$$

$$b_{0j} \sim \mathcal{N}(\text{mean} = 0, \text{var} = V_b), \text{iid}$$

$$d_{0i} \sim \mathcal{N}(\text{mean} = 0, \text{var} = V_d), \text{iid}$$

$$\epsilon_{ijt} \sim \mathcal{N}(\text{mean} = 0, \text{var} = V), iid$$

Hyperpriors

$$V_b \sim \text{IG}(\text{shape} = 10^{-3}, \text{rate} = 10^{-3}), iid$$

$$V_d \sim \text{IG}(\text{shape} = 10^{-3}, \text{rate} = 10^{-3}), iid$$

$$V \sim \text{IG}(\text{shape} = 10^{-3}, \text{rate} = 10^{-3}), iid$$

Where i is the individual and j the species. The model included two fixed effects (β_0 and β_1), a random species effect b_{0j} on the intercept and a random individual effect d_{0i} on the intercept. V_d is the intraspecific variance and V_b is the interspecific variance.

3.1 Paracou

```
## No divergences to plot.
```

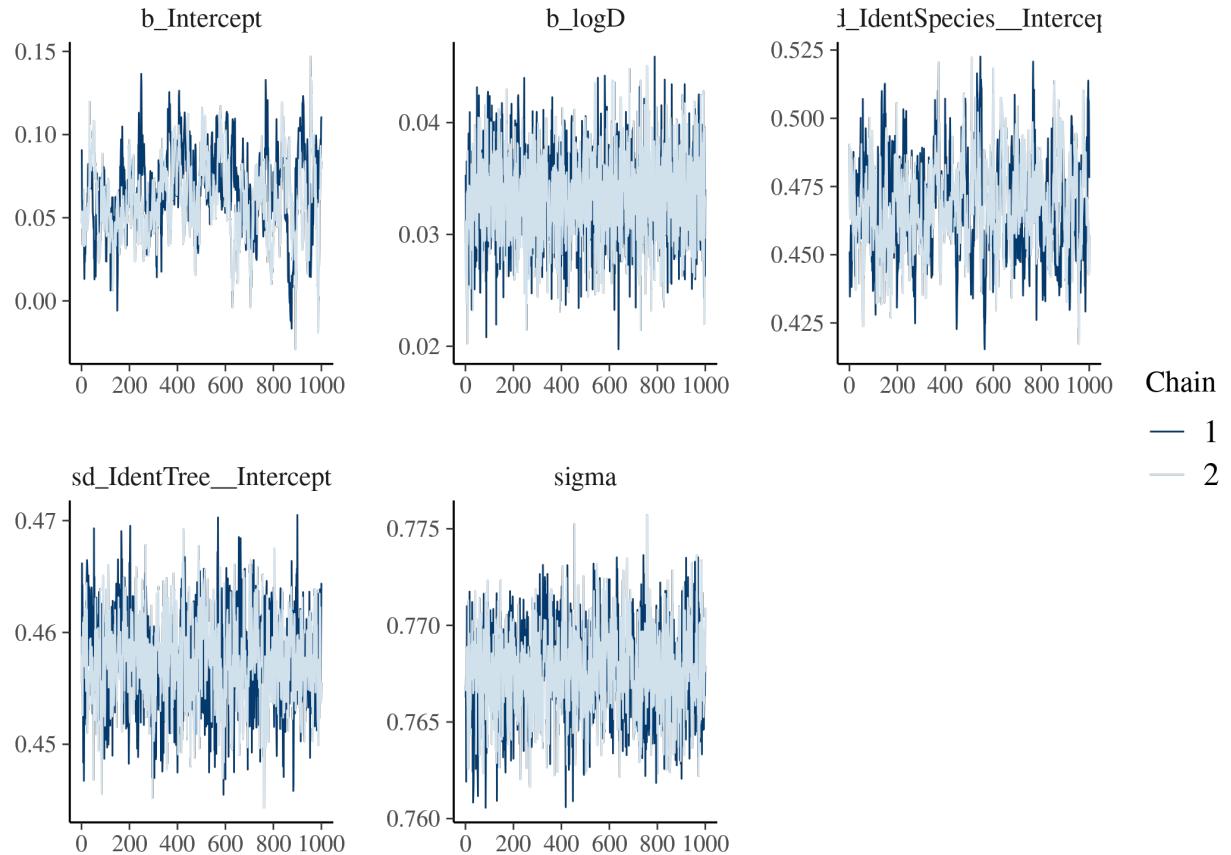


Figure 3.1: Trace of the posterior estimates for the Paracou site

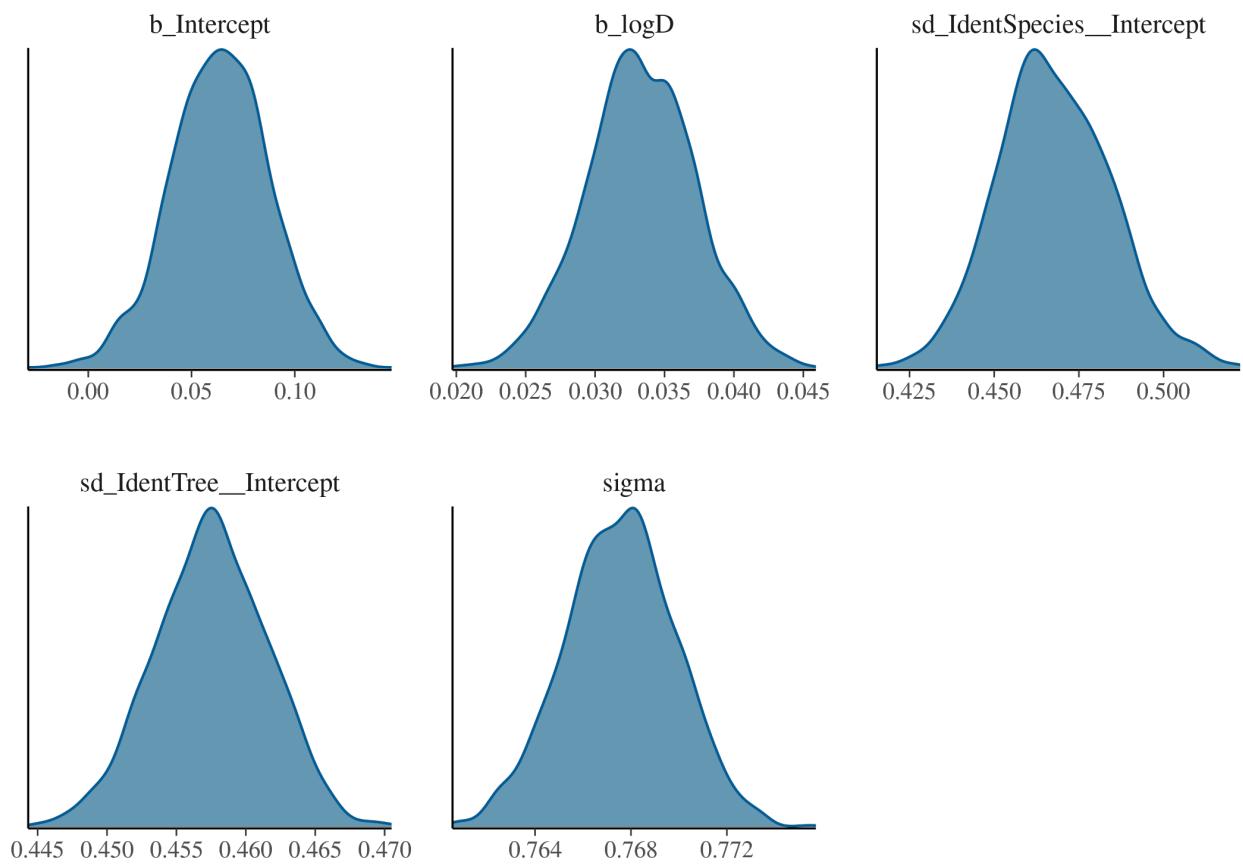


Figure 3.2: Density of the posterior estimates for the Paracou site

3.2 Uppangala

```
## No divergences to plot.
```

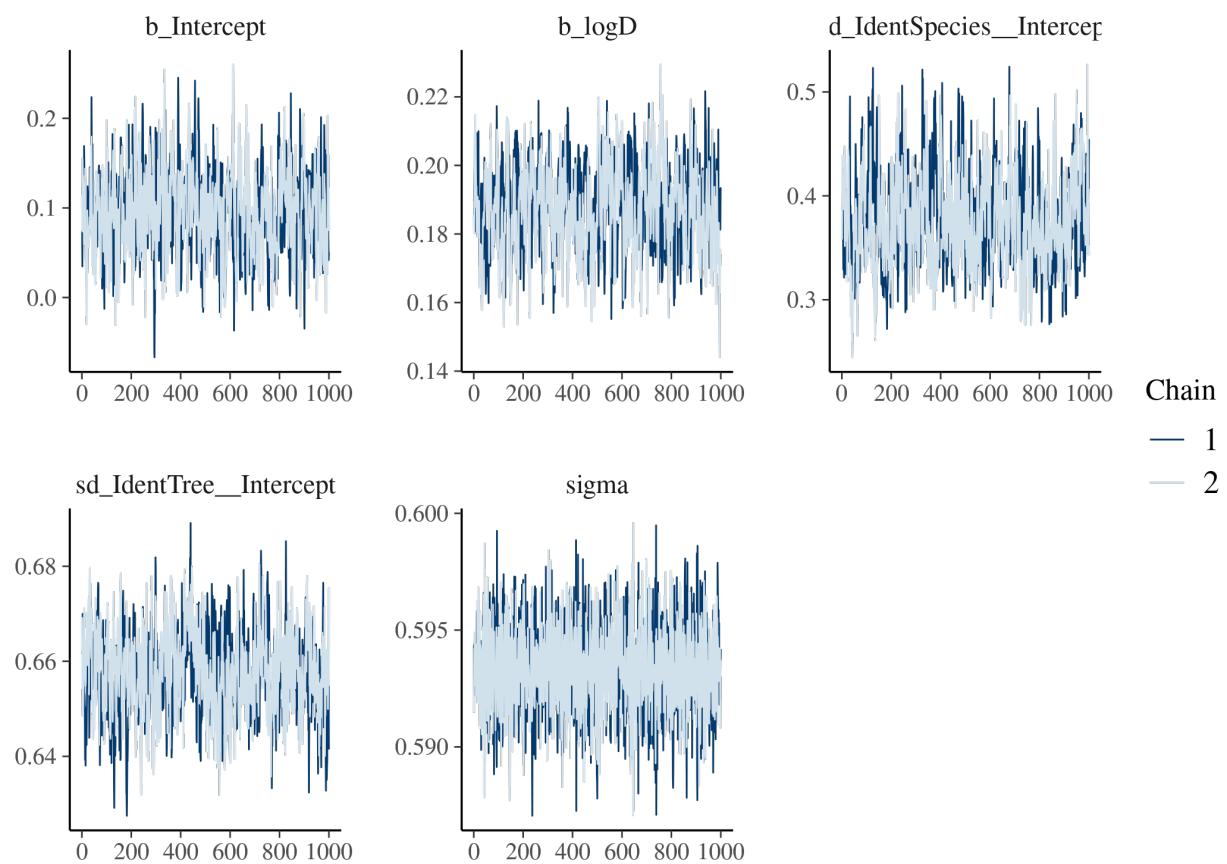


Figure 3.3: Trace of the posterior estimates for the Paracou site

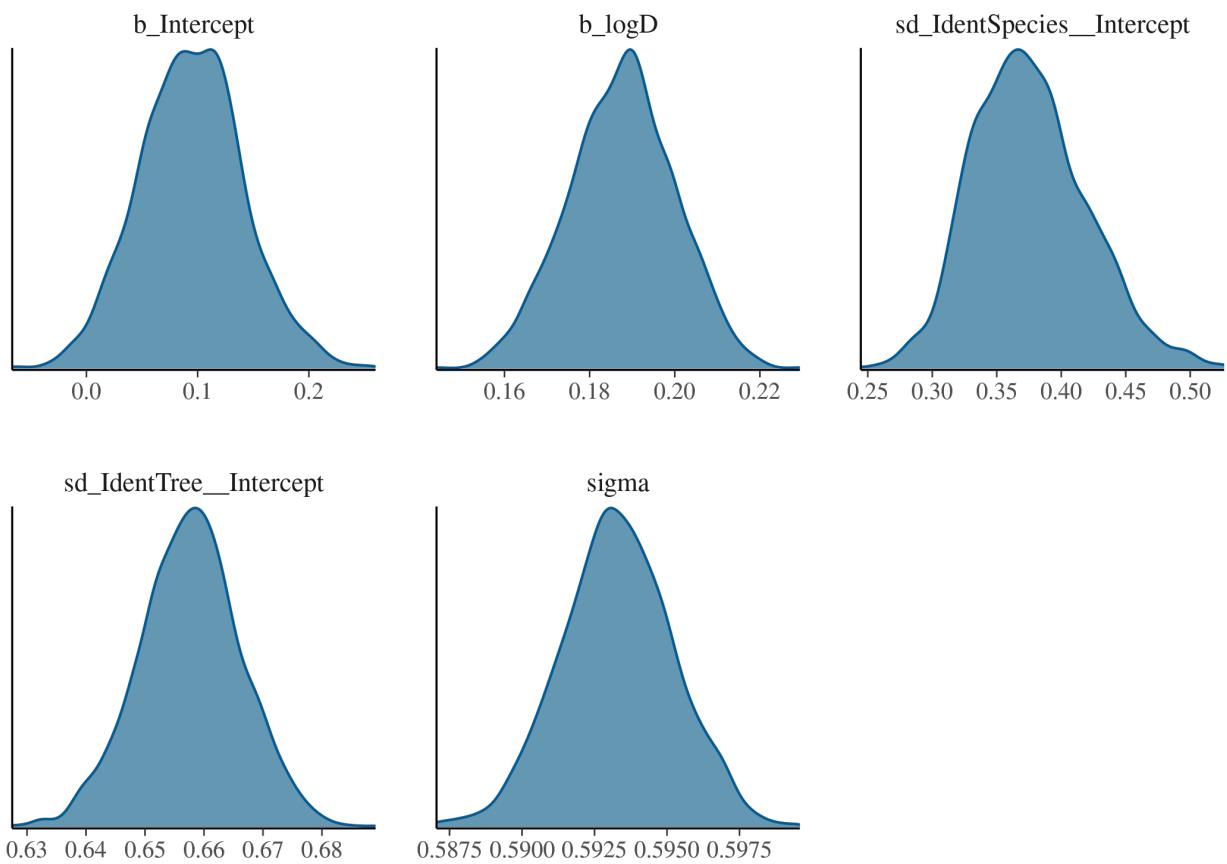


Figure 3.4: Density of the posterior estimates for the Uppangala site

3.3 BCI

```
## No divergences to plot.
```

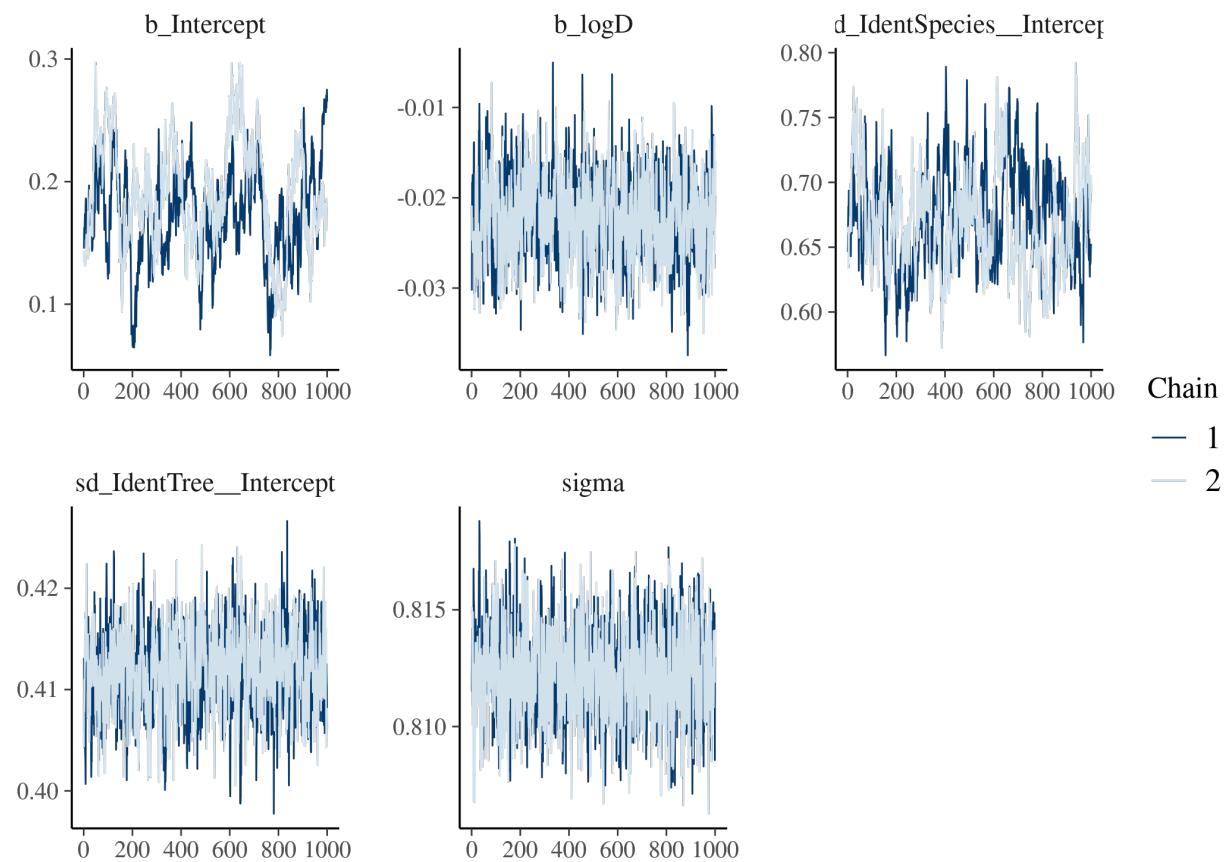


Figure 3.5: Trace of the posterior estimates for the BCI site

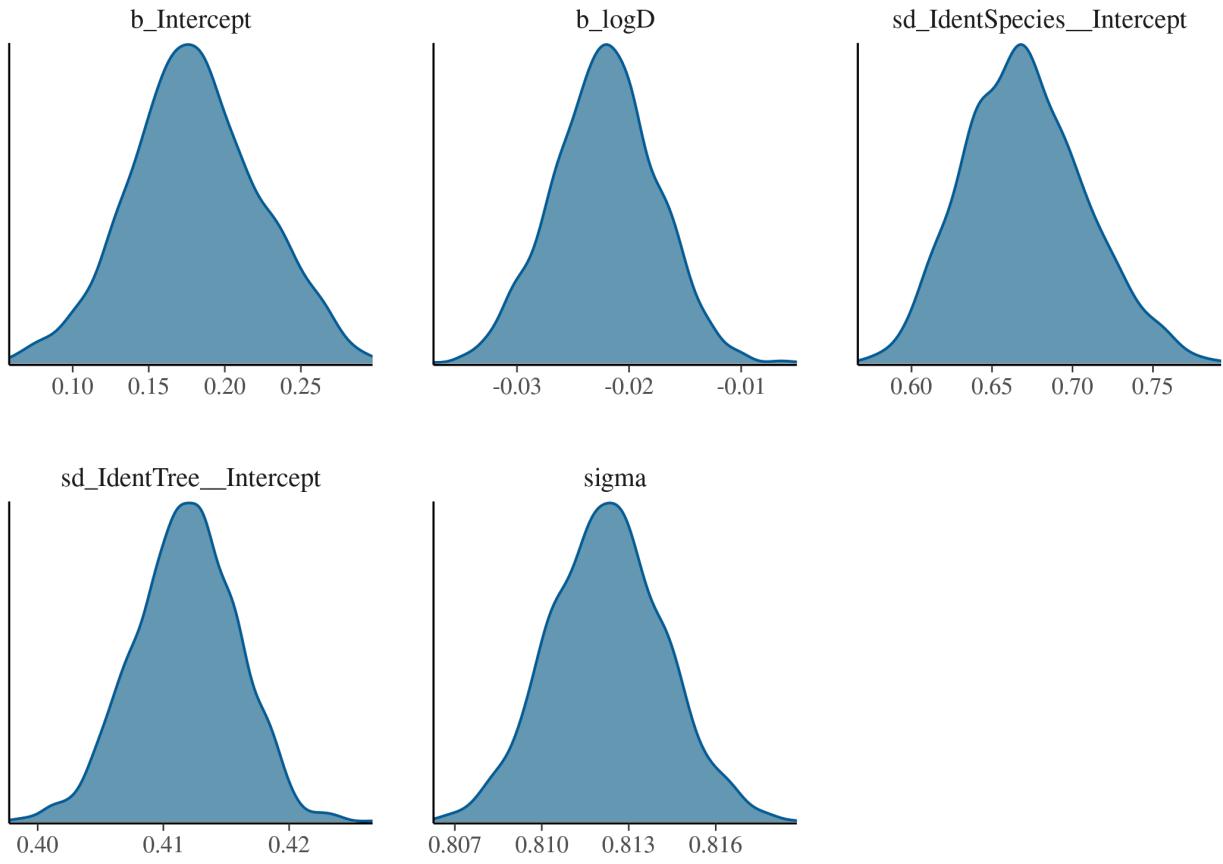


Figure 3.6: Density of the posterior estimates for the BCI site

Table 1: Variance partitioning of individual growth in three tropical forest sites. Shown are the mean posterior estimates and their estimated error from the growth model fitted for each site

	Intercept (β_0)	Diameter (β_1)	Species variance (V_b)	Individual variance (V_d)	Residual variance (V)
Paracou					
Estimate	6.4e-02	3.3e-02	4.7e-01	4.6e-01	7.7e-01
Estimation error	2.4e-02	4e-03	1.7e-02	4.1e-03	2.3e-03
Uppangala					
Estimate	9.5e-02	1.9e-01	3.8e-01	6.6e-01	5.9e-01
Estimation error	4.7e-02	1.2e-02	4.4e-02	8.6e-03	1.9e-03
BCI					
Estimate	1.8e-01	-2.2e-02	6.7e-01	4.1e-01	8.1e-01
Estimation error	4.3e-02	4.6e-03	3.7e-02	4.1e-03	2e-03

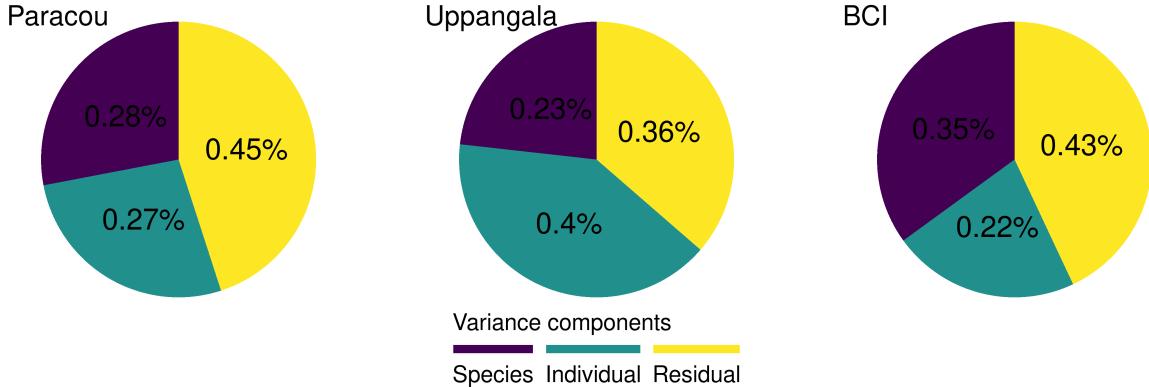


Figure 3.7: Variance partitioning of individual growth in three tropical forest sites (see Table above).

Overall, a large part of variability in tree growth was imputable to individual effects in the three sites, even after accounting for the effect of diameter on tree growth, showing a high intraspecific variability in growth in this tropical forests (Figure 3.7).

4 Moran's I analysis

In order to test if individual tree growth was spatially autocorrelated within each site, we performed a Moran's I test on individual mean growth for each species with more than 5 individuals. For very abundant species (with more than 3000 individuals), we sampled 3000 individuals with a uniform probability. Moran's I test function was re-written from the ape package's Moran.I function with a one-tailed test [9] in order to be able to select the pairs of neighbours that are less than 100-m apart and thus assess local spatial autocorrelation.

We then computed the proportion of species with a significantly and positively spatial autocorrelated growth (with a 0.05 alpha risk) and the proportion of species with non significant spatial autocorrelation, as well as the corresponding proportion of individuals. The proportions are computed relative to the species on which the test was performed.

29%, 18%, and 17% of the species showed a positive and significant spatial autocorrelation in tree growth, in Paracou, Uppangala, and BCI respectively. (Table 2).

Although the three forest sites are located in very different areas across the tropics, with contrasting biogeographical history, climate, regime of perturbations and topography, they all showed a substantial spatial autocorrelation of tree growth intraspecific variability, which supports the generality of our results. The magnitude of spatial autocorrelation varied across the three sites however, with Paracou presenting the biggest proportion and Uppangala the lowest, which can be related to each site specificities. The plots of the Paracou site were subjected to disturbance treatments that created numerous large gaps [10]. This resulted in a high heterogeneity and spatial structuration of the environment. When restricting our analysis to undisturbed control plots (see 6.1 below), the proportion of detected spatial autocorrelation in conspecific growth drops from 78% to 51% individuals, supporting the hypothesis that the more important spatial autocorrelation in Paracou is due to the stronger spatial structuration of the micro-environment associated with disturbance treatments. Conversely, the disturbance regime in BCI is known to produce frequent small gaps [11], while the mature forest in Uppangala has been reported as barely disturbed [3]. We acknowledge that the hilly topography of Uppangala could make our estimates of distances among individuals based on pairs of tree coordinates provided by field inventory less accurate than in the other two flatter sites.

Table 2: Spatial autocorrelation of the growth of conspecific individuals in three tropical forest sites. Shown are the proportion of species, and of corresponding individuals, in percent, for which individual growth is significantly positively spatially autocorrelated. The spatial autocorrelation of individual growth was tested using Moran's I index.

	Significant	Not significant
Paracou		
Proportion of species	28.7	71.3
Proportion of individuals	77.7	22.3
Uppangala		
Proportion of species	18.5	81.5
Proportion of individuals	45.3	54.7
BCI		
Proportion of species	16.9	83.1
Proportion of individuals	51.4	48.6

5 Is variance within conspecifics inferior to variance within heterospecifics ?

In order to test if the performance of conspecifics was locally more similar than that of heterospecifics in the three tropical forests, for each species with more than five individuals and more than five heterospecific neighbours within 100-m radii around these individuals , we computed the semivariances of growth of individuals that are less than 100-m apart, considering conspecifics only on the one hand, and heterospecifics on the other hand. In the first case, semivariance was estimated as the mean of the squared difference in individual mean growth for all pairs of conspecific individuals. In the second case, semivariance was estimated as the mean of the squared difference in individual mean growth for all pairs of individuals with an individual of the focal species and one of another species. For each species, we compared the semivariances obtained with conspecifics only to the ones obtained with heterospecifics using a Mann-Whitney test.

For each site, the proportion of species with a significantly higher or lower semivariance than the one estimated with heterospecifics or for which the difference was not significant, and the corresponding proportion of individuals were then computed.

In all three tropical forest sites however, growth semivariance was locally significantly higher among heterospecifics than among conspecifics in a large proportion of the tested cases (62%, 42% and 47% of the species, and 89%, 58% and 75% of the individuals in Paracou, Uppangala and BCI respectively; Table 3).

Note that, to control for a potential effect of species abundance on the values of semivariances with heterospecifics, we replicate the analysis by computing the semivariances with heterospecifics by sampling a maximum of ten individuals per species. The results were qualitatively unchanged (Table 4).

Table 3: Comparison of local intra- and interspecific variability in individual growth for three tropical forest sites. The variability was estimated with the semivariance and the comparison was performed with a Mann-Whitney's test. The semivariances were computed for all species with > 5 individuals and > 5 heterospecific neighbours within 100 m in the same plot, and considering pairs of individuals that were less than 100 m apart and in the same plot. Shown are the proportion of species, and of corresponding individuals, for which (i) intraspecific variability was significantly lower than interspecific variability, (ii) intraspecific variability was significantly higher than interspecific variability, or (iii) the difference between inter- and intraspecific variabilities was not significant.

	Intraspecific variability < interspecific variability (i)	Intraspecific variability ~ interspecific variability (ii)	Intraspecific variability > interspecific variability (iii)
Paracou			
Proportion of species	61.7	39.7	0.333
Proportion of individuals	89.1	10.9	0.0699
Uppangala			
Proportion of species	42.2	62.2	4.44
Proportion of individuals	57.7	23.6	18.8
BCI			
Proportion of species	46.6	44.3	9.09
Proportion of individuals	75.2	17.2	7.6

Table 4: Comparison of local intra- and interspecific variability in individual growth for three tropical forest sites, when controlling for species abundance. See Table 3. Semi-variances with heterospecifics were here computed by sampling a maximum of 10 individuals per species.

	Intraspecific variability < interspecific variability (i)	Intraspecific variability ~ interspecific variability (ii)	Intraspecific variability > interspecific variability (iii)
Paracou			
Proportion of species	61.7	37	1.33
Proportion of individuals	86.3	12.1	1.56
Uppangala			
Proportion of species	42.2	55.6	2.22
Proportion of individuals	61.2	20.4	18.4
BCI			
Proportion of species	50	47.7	2.27
Proportion of individuals	71.2	27.8	1.02

6 Other tests and verifications

6.1 Undisturbed plots in the Paracou site

One of our main hypotheses was that since many environmental variables are spatially structured and that tree growth is largely influenced by environmental variables, tree growth should be spatially structured. Our analysis using Moran's I test showed that tree growth was indeed spatially structured at our study sites. The Paracou dataset offers the opportunity to test this hypothesis further. Indeed, some plots were disturbed in the early eighties, creating artificial gaps, whereas others were not disturbed. As the creation of gaps results in a strong spatial structure of the light available under the canopy, we hypothesized that growth should be less structured in plots that were not disturbed.

Table 5: Spatial autocorrelation of the growth of conspecific individuals in three tropical forest sites. Shown are the proportion of species, and of corresponding individuals, in percent, for which individual growth is significantly positively spatially autocorrelated. The spatial autocorrelation of individual growth was tested using Moran's I index.

	Significant	Not significant
Proportion of species	18	82
Proportion of individuals	54.7	45.3

Considering the undisturbed plots only, the proportion of individuals of species with a significant Moran's I test was 55%, much lower than when including the disturbed plots (89%, see Table 2; Table 5). This corroborates our hypothesis that the openness of the canopy triggered important spatial structure in disturbed plots due to light gaps, and as in these gaps, trees tend to grow faster, the spatial structure of growth is stronger.

6.2 Smaller stems in the BCI site

In order to be able to compare all three datasets together, we removed all stems that had a DBH inferior to 10 cm in the BCI inventories, which include all stems with $DBH \geq 1$ cm. We replicated our analysis on the spatial autocorrelation of tree growth in the complete dataset and found that the spatial structure in individual growth was even more significant.

7 Code implementation

The whole analysis was conducted using the R language [12] in the Rstudio environment [13]. The tables were made with the kableExtra package [14], the figures with the package ggplot2 [15], and the code uses other packages of the Tidyverse [16] (tidyR [17], dplyr [18], readr [19], lubridate [20], glue [21], magrittr [22]) and other R packages (here [23], sp [25]). Some functions were implemented in C++ thanks to the R packages Rcpp [26,27,28] and RcppArmadillo [29]. The pdf and html documents were produced thanks to the R packages rmarkdown [32], knitr [35] and bookdown [36].

References

- 1 Gourlet-Fleury, S. *et al.*, eds. (2004) *Ecology and management of a neotropical rainforest: Lessons drawn from Paracou, a long-term experimental research site in French Guiana*, Elsevier.
- 2 Héault, B. and Piponiot, C. (2018) Key drivers of ecosystem recovery after disturbance in a neotropical forest: Long-term lessons from the Paracou experiment, French Guiana. *Forest Ecosystems* 5, 2
- 3 Pélissier, R. *et al.* (2011) Tree demography in an undisturbed Dipterocarp permanent sample plot at Uppangala, Western Ghats of India: *Ecological Archives* E092-115. *Ecology* 92, 1376–1376
- 4 Bec, J.L. *et al.* (2015) Characterizing tropical tree species growth strategies: Learning from inter-individual variability and scale invariance. *PLoS ONE* 10,
- 5 Pascal, J.-P. and Pelissier, R. (1996) Structure and Floristic Composition of a Tropical Evergreen Forest in South-West India. *Journal of Tropical Ecology* 12, 191–214
- 6 Condit, R. *et al.* Complete data from the Barro Colorado 50-ha plot: 423617 trees, 35 years. (2019), Dryad
- 7 Bürkner, P.-C. (2017) Brms : An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software* 80,
- 8 Bürkner, P.-C. (2018) Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal* 10, 395
- 9 Paradis, E. and Schliep, K. (2019) Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics* 35, 526–528
- 10 Molino, J.-F. and Sabatier, D. (2001) Tree Diversity in Tropical Rain Forests: A Validation of the Intermediate Disturbance Hypothesis. *Science* 294, 1702–1704
- 11 Hubbell, S.P. (1999) Light-Gap Disturbances, Recruitment Limitation, and Tree Diversity in a Neotropical Forest. *Science* 283, 554–557
- 12 R Core Team (2021) *R: A language and environment for statistical computing*, R Foundation for Statistical Computing.
- 13 RStudio Team (2021) *RStudio: Integrated development environment for r*, RStudio, PBC.
- 14 Zhu, H. (2021) *kableExtra: Construct complex table with 'kable' and pipe syntax*,
- 15 Wickham, H. (2009) *Ggplot2: Elegant graphics for data analysis*, Springer.
- 16 Wickham, H. *et al.* (2019) Welcome to the Tidyverse. *Journal of Open Source Software* 4, 1686
- 17 Wickham, H. (2021) *Tidyr: Tidy messy data*,
- 18 Wickham, H. *et al.* (2021) *Dplyr: A grammar of data manipulation*,
- 19 Wickham, H. and Hester, J. (2020) *Readr: Read rectangular text data*,

- 20 Golemund, G. and Wickham, H. (2011) Dates and Times Made Easy with **lubridate**. *Journal of Statistical Software* 40,
- 21 Hester, J. (2020) *Glue: Interpreted string literals*,
- 22 Bache, S.M. and Wickham, H. (2020) *Magrittr: A forward-pipe operator for r*,
- 23 Müller, K. (2020) *Here: A simpler way to find your files*,
- 24 Pebesma, E.J. and Bivand, R.S. (2005) Classes and methods for spatial data in R. *R News* 5, 9–13
- 25 Bivand, R. *et al.* (2013) *Applied spatial data analysis with R*, Second edition. Springer.
- 26 Eddelbuettel, D. and François, R. (2011) Rcpp : Seamless R and C++ Integration. *Journal of Statistical Software* 40,
- 27 Eddelbuettel, D. (2013) *Seamless R and C++ integration with Rcpp*, Springer.
- 28 Eddelbuettel, D. and Balamuta, J.J. (2018) Extending *r* with C++: A Brief Introduction to Rcpp. *The American Statistician* 72, 28–36
- 29 Eddelbuettel, D. and Sanderson, C. (2014) RcppArmadillo: Accelerating R with high-performance C++ linear algebra. *Computational Statistics & Data Analysis* 71, 1054–1063
- 30 Allaire, J. *et al.* Rmarkdown: Dynamic documents for R. (2020)
- 31 Xie, Y. *et al.* (2019) *R Markdown: The definitive guide*, CRC Press, Taylor; Francis Group.
- 32 Xie, Y. *et al.* (2020) *R markdown cookbook*, (1st edn) Taylor; Francis, CRC Press.
- 33 Xie, Y. (2021) *Knitr: A general-purpose package for dynamic report generation in r*,
- 34 Xie, Y. (2015) *Dynamic documents with R and Knitr*, Second edition. CRC Press/Taylor & Francis.
- 35 Xie, Y. (2014) Knitr: A Comprehensive Tool for Reproducible Research in R. In *Implementing reproducible research* (Stodden, V. et al., eds), CRC Press, Taylor & Francis Group
- 36 Xie, Y. (2017) *Bookdown: Authoring books and technical publications with R Markdown*, CRC Press.