

Supplementary Materials for

Rethinking the role of intraspecific variability in species coexistence

Camille Girard-Tercieux^{*,1} Isabelle Maréchaux¹ Adam T. Clark²
James S. Clark^{3,4} Benoît Courbaud⁴ Claire Fortunel¹ Joannès Guillemot⁵
Georges Künstler⁴ Guerric le Maire⁵ Raphaël Pélissier^{1,6} Nadja Rüger^{7,8,9} and
Ghislain Vieilledent¹

* Correspondance to camillegirardtercieux@gmail.com

[1] AMAP, Univ. Montpellier, CIRAD, CNRS, INRAE, IRD, Montpellier, France

[2] Institute of Biology, Karl-Franzens University of Graz, Graz, Austria

[3] Nicholas School of the Environment, Duke University, Durham (NC), USA

[4] Univ. Grenoble Alpes, INRAE, LESSEM, F-38402 St-Martin-d'Hères, France

[5] Eco&Sols, Univ. Montpellier, CIRAD, INRAE, IRD, Institut Agro, Montpellier, France

[6] Department of Ecology, French Institute of Pondicherry, Puducherry, India

[7] Department of Economics, University of Leipzig, Leipzig, Germany

[8] German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig, Leipzig, Germany

[9] Smithsonian Tropical Research Institute, Balboa, Panama

Contents

1 Body of evidence	2
2 Analysis of tropical forest inventory data	3
2.1 Datasets	3
2.2 Data preparation	6
2.3 Estimation of IV	7
2.4 Moran's I analysis	14
2.5 Is variance within conspecifics inferior to variance within heterospecifics ?	15
2.6 Other tests and verifications	16
3 Analysis of an Eucalyptus clonal plantation dataset	17
3.1 The dataset	17
3.2 Competition index	20
3.3 Statistical growth model	20
3.4 Results of the model and variance partitioning	21
3.5 Conclusion	28
4 Virtual experiment with two species	29
4.1 Perfect knowledge model : generating response values.	29
4.2 Partial knowledge model : interpreting the response values with only one explanatory variable.	33
4.3 Spatial autocorrelation of individual response	43
4.4 Similarity between conspecific individuals compared to heterospecific individuals and consequences for species coexistence.	44
5 Code used to generate Figure 1 in the main text: <i>Reinterpreting observed intraspecific variability (IV): from niche widening to niche projection into a high-dimensional environment.</i>	45
6 Code implementation	52
References	52

1 Body of evidence

We used literature and data analyses of various nature to support our opinions. Figure 1.1 illustrates how these elements articulate together.

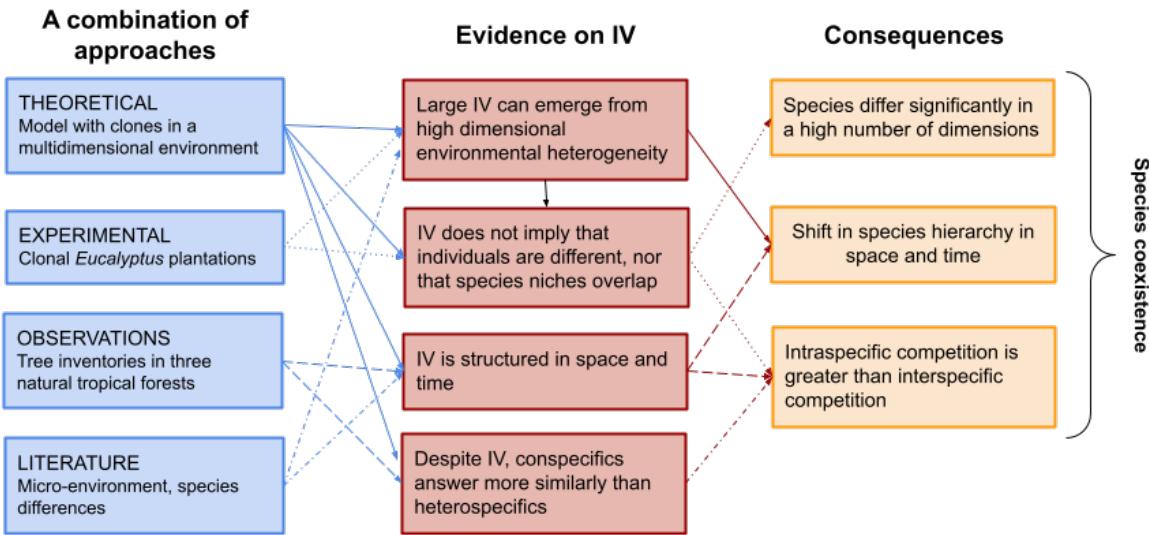


Figure 1.1: The body of evidence used to support the opinion of the paper

2 Analysis of tropical forest inventory data

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.

2.1 Datasets

2.1.1 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 Héault & Piponiot (2018). 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 2.1 shows the location of the trees in the Paracou site.

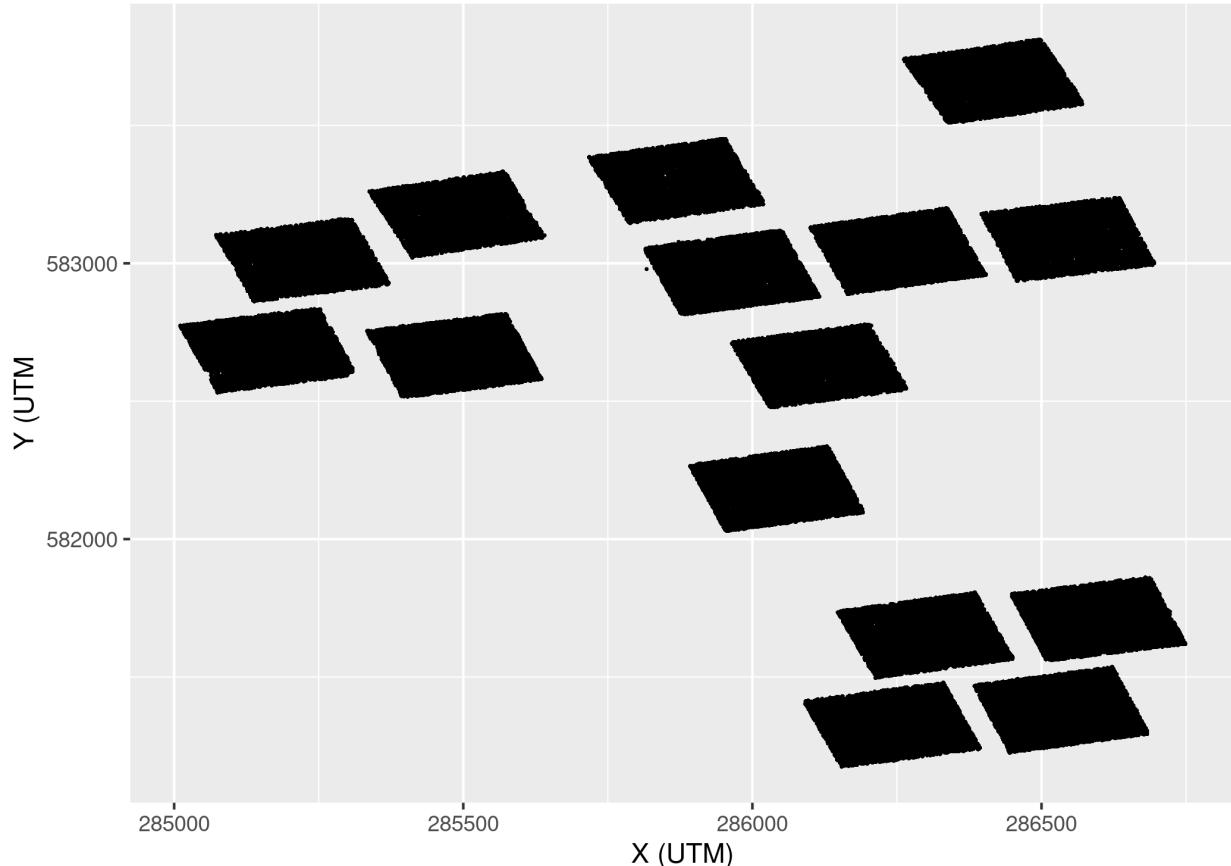


Figure 2.1: Plots at the Paracou site. Each point is a tree.

2.1.2 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 Kadamkal Reserve Forest, in the Pushpagiri Wildlife Sanctuary, in Karnataka state, India (Pélissier *et al.* 2011). It is a low altitude (500-600 m) wet evergreen monsoon Dipterocarp forest (Bec *et al.* 2015). 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 (Pascal & Pelissier 1996).

Figure 2.2 shows the location of the trees in the Uppangala site.

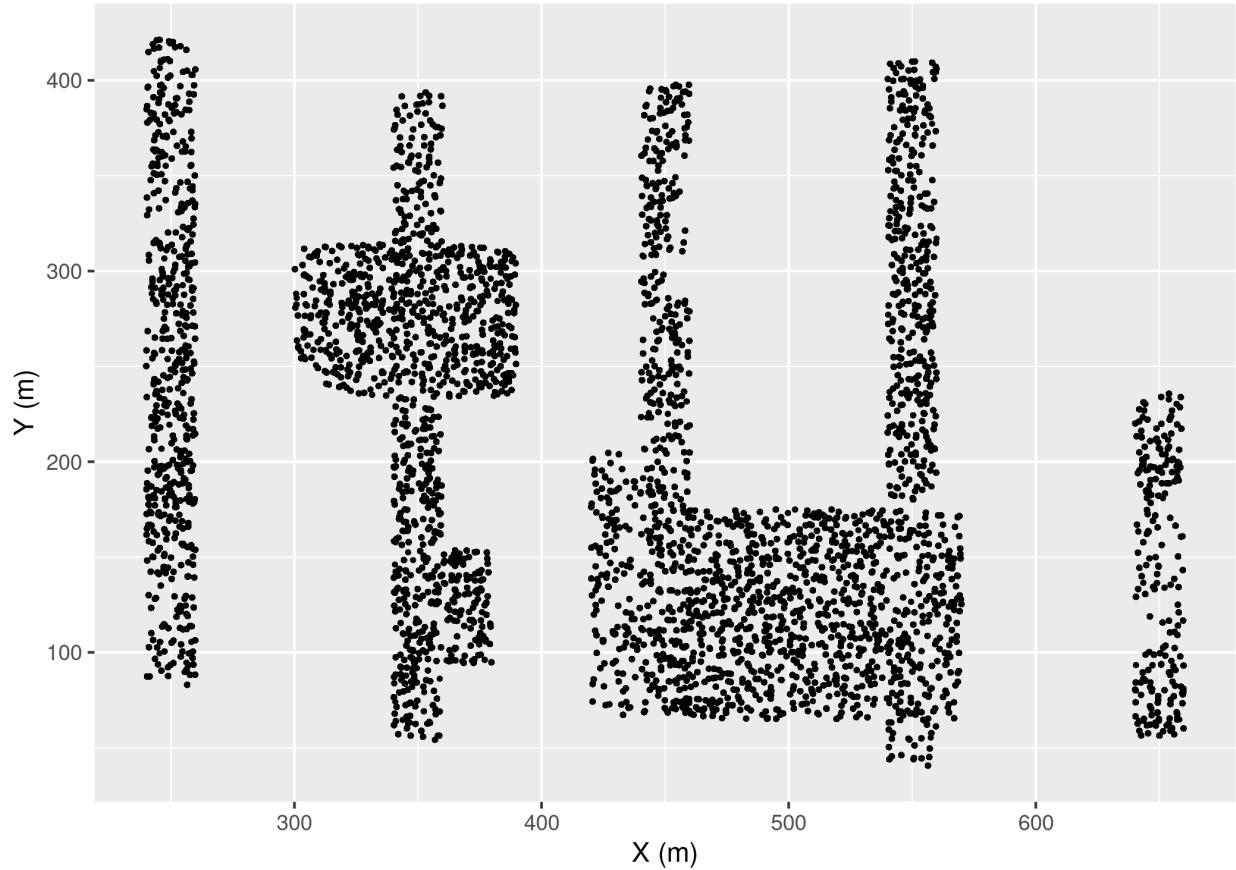


Figure 2.2: Plots at Uppangala site. Each point is a tree.

2.1.3 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 (Condit *et al.* 2019).

Figure 2.3 shows the location of the selected trees in the BCI site.

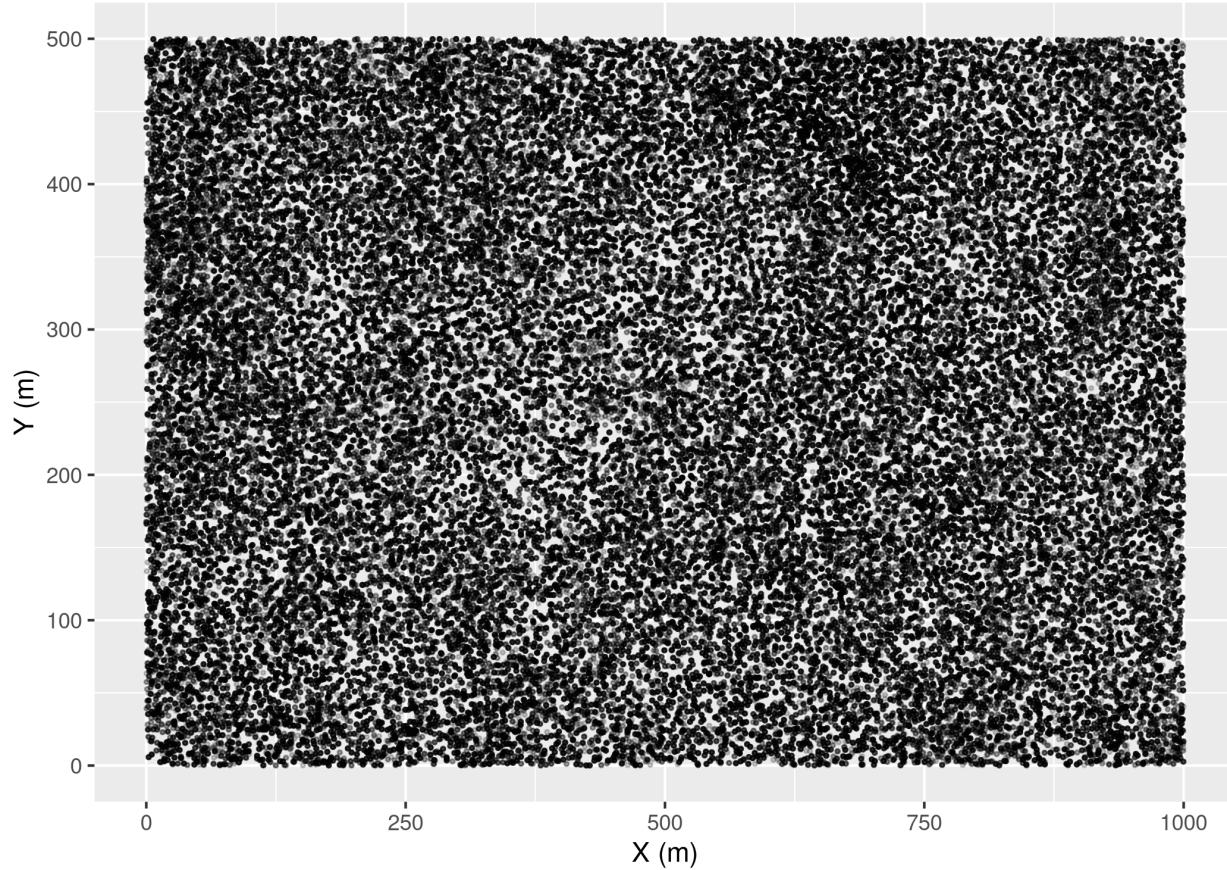


Figure 2.3: The 50 ha plot of the Barro Colorado Island site. Each point is a tree.

2.2 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 107670 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.

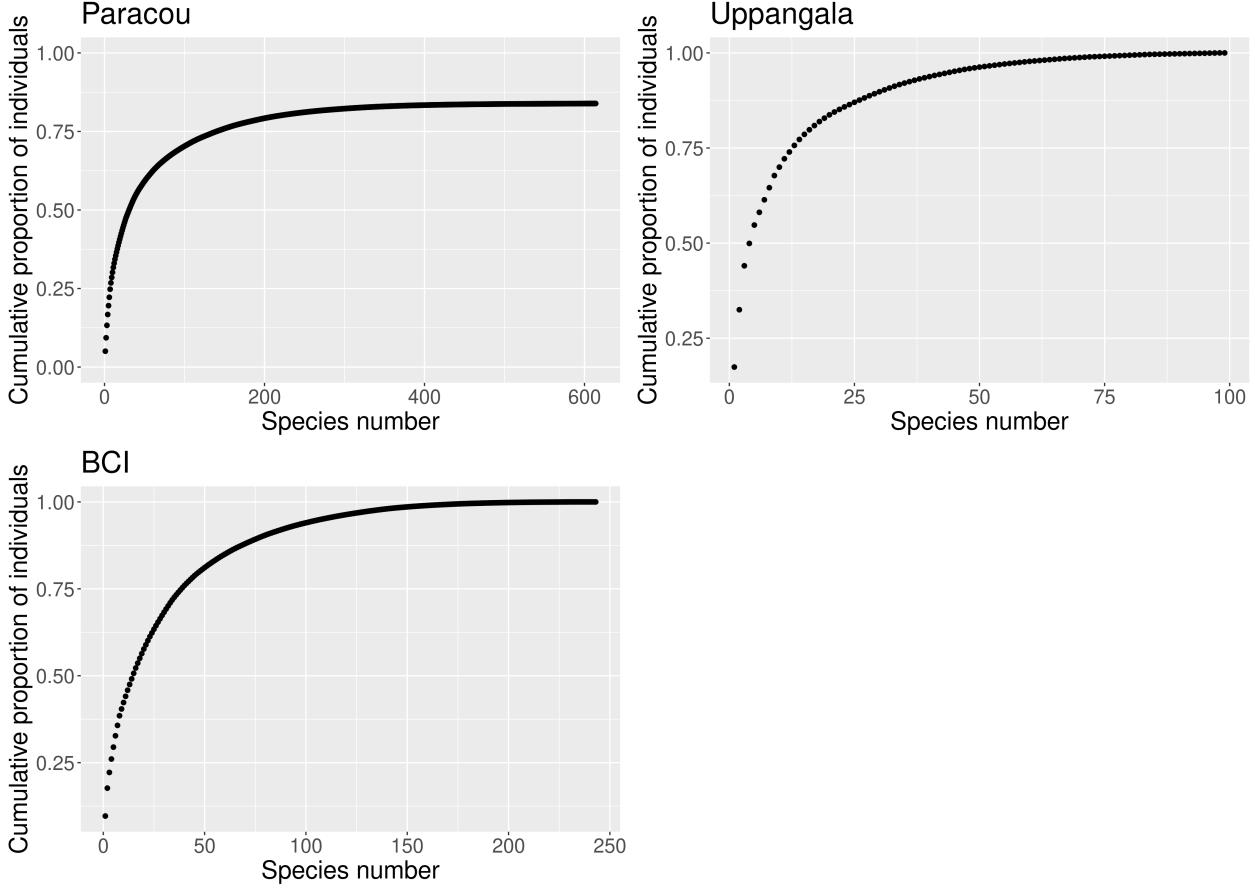


Figure 2.4: Abundance diagrams for the three study sites. For the Paracou site, only individuals that were identified to the species level are accounted for. The three communities include few dominant species and many rare species (long asymptote).

2.3 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 (Bürkner 2017, 2018) 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

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

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

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

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

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

Hyperpriors

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

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

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

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.

2.3.1 Paracou

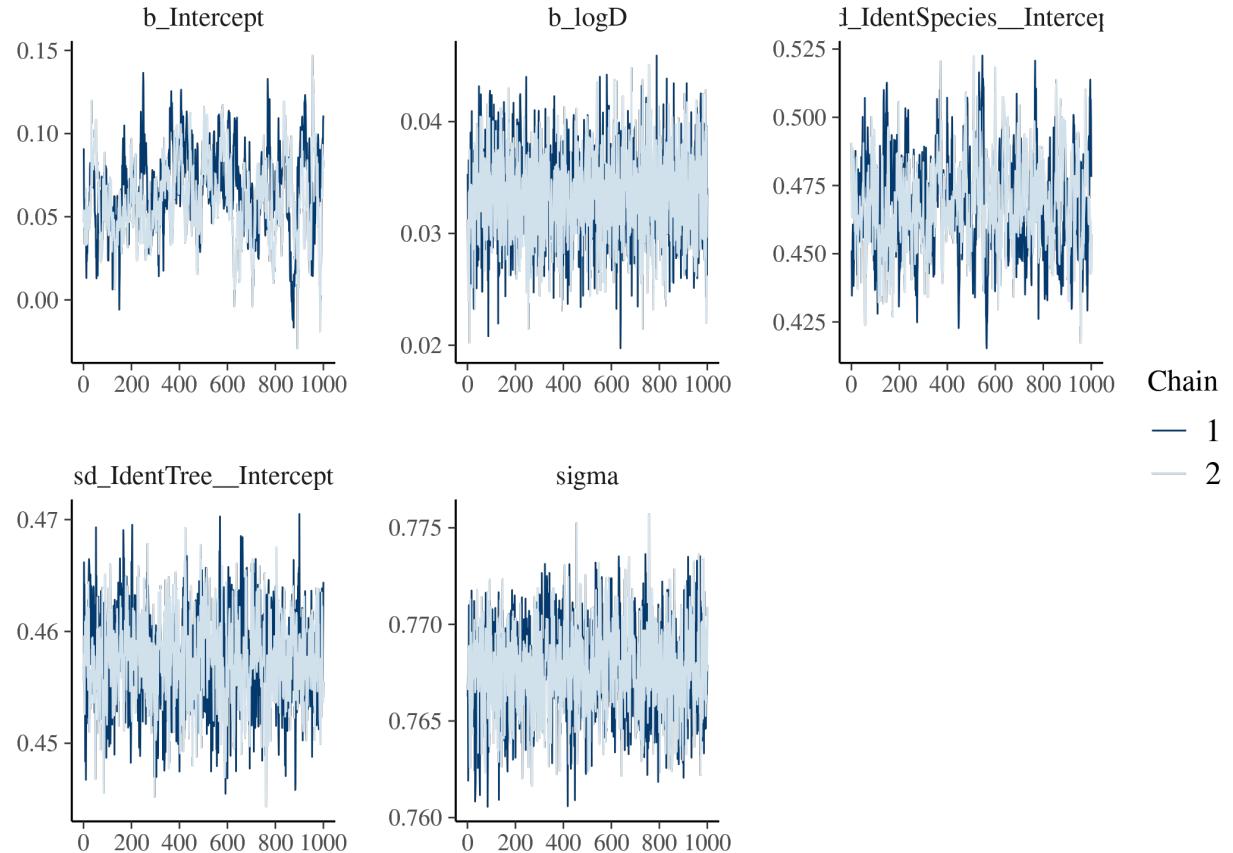


Figure 2.5: Trace of the posterior estimates for the Paracou site

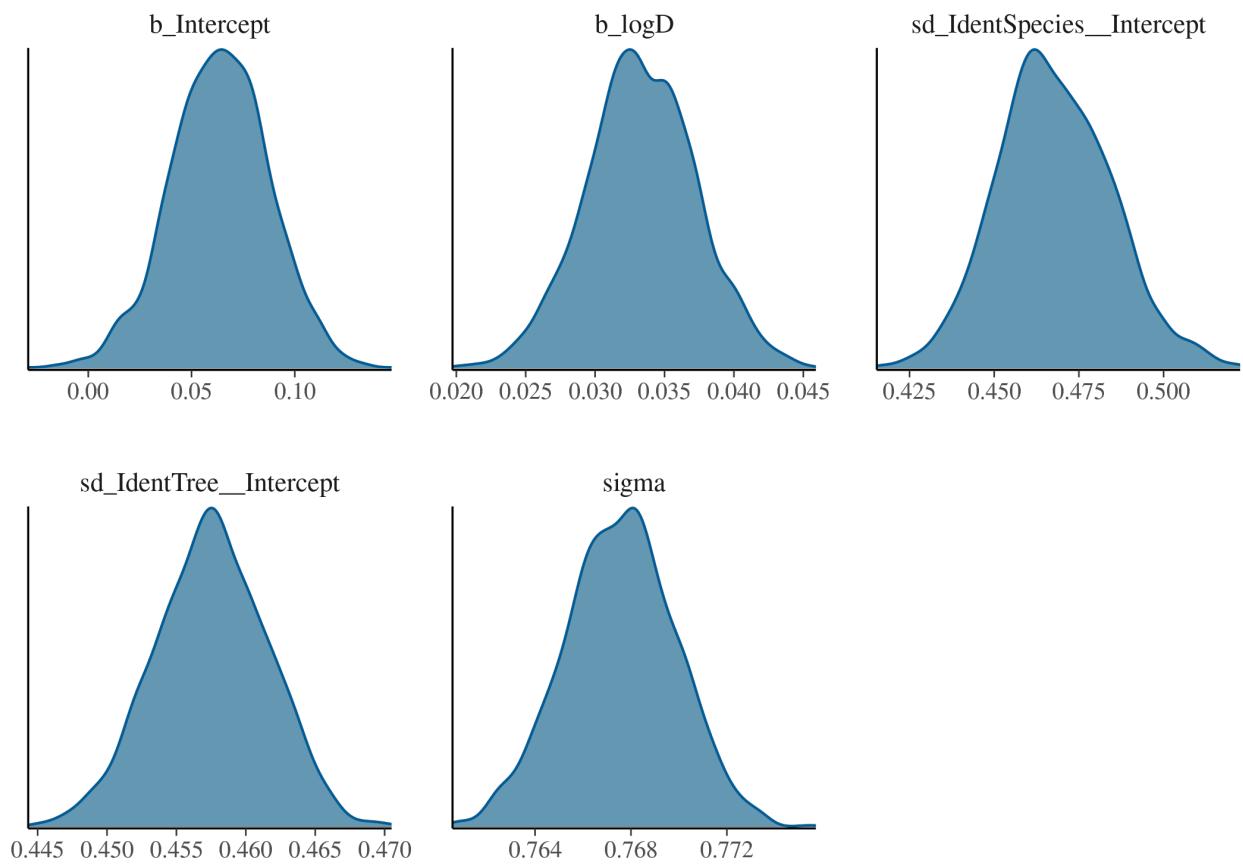


Figure 2.6: Density of the posterior estimates for the Paracou site

2.3.2 Uppangala

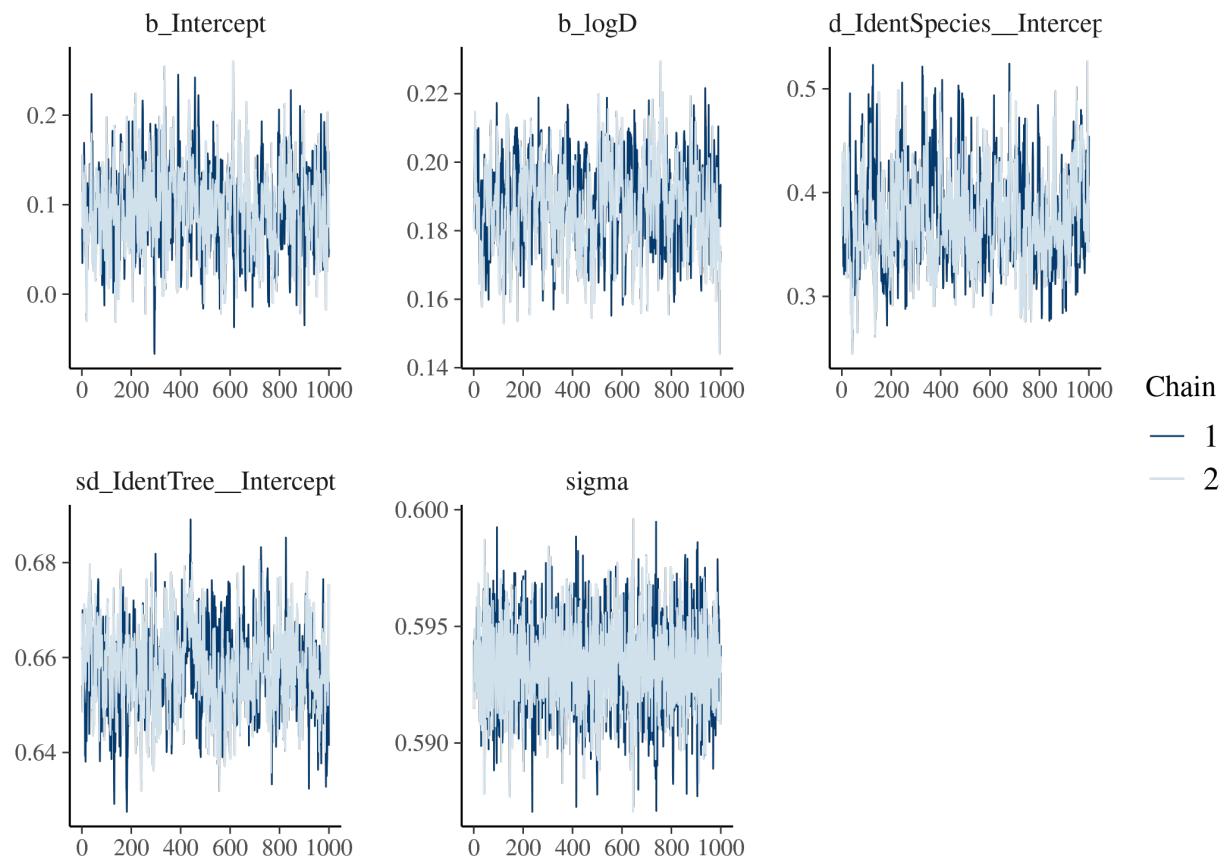


Figure 2.7: Trace of the posterior estimates for the Uppangala site

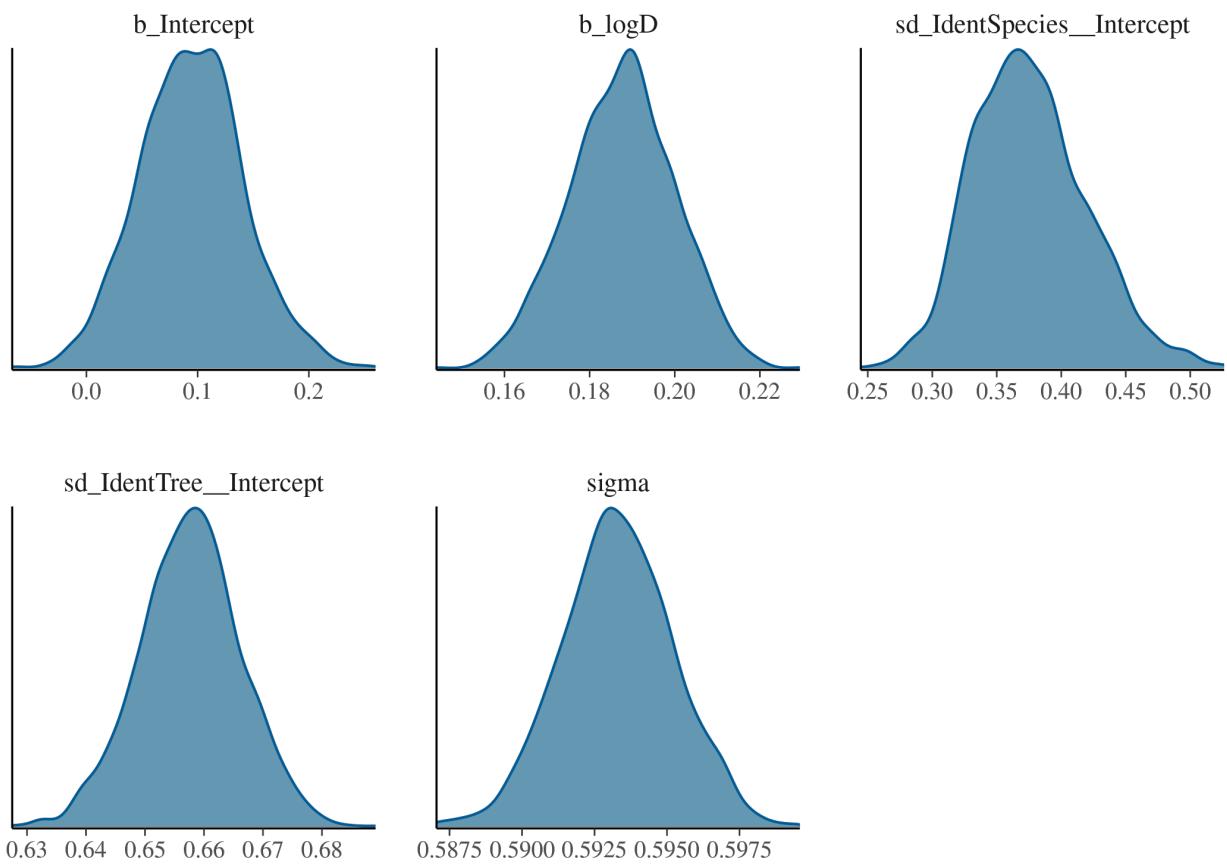


Figure 2.8: Density of the posterior estimates for the Uppangala site

2.3.3 BCI

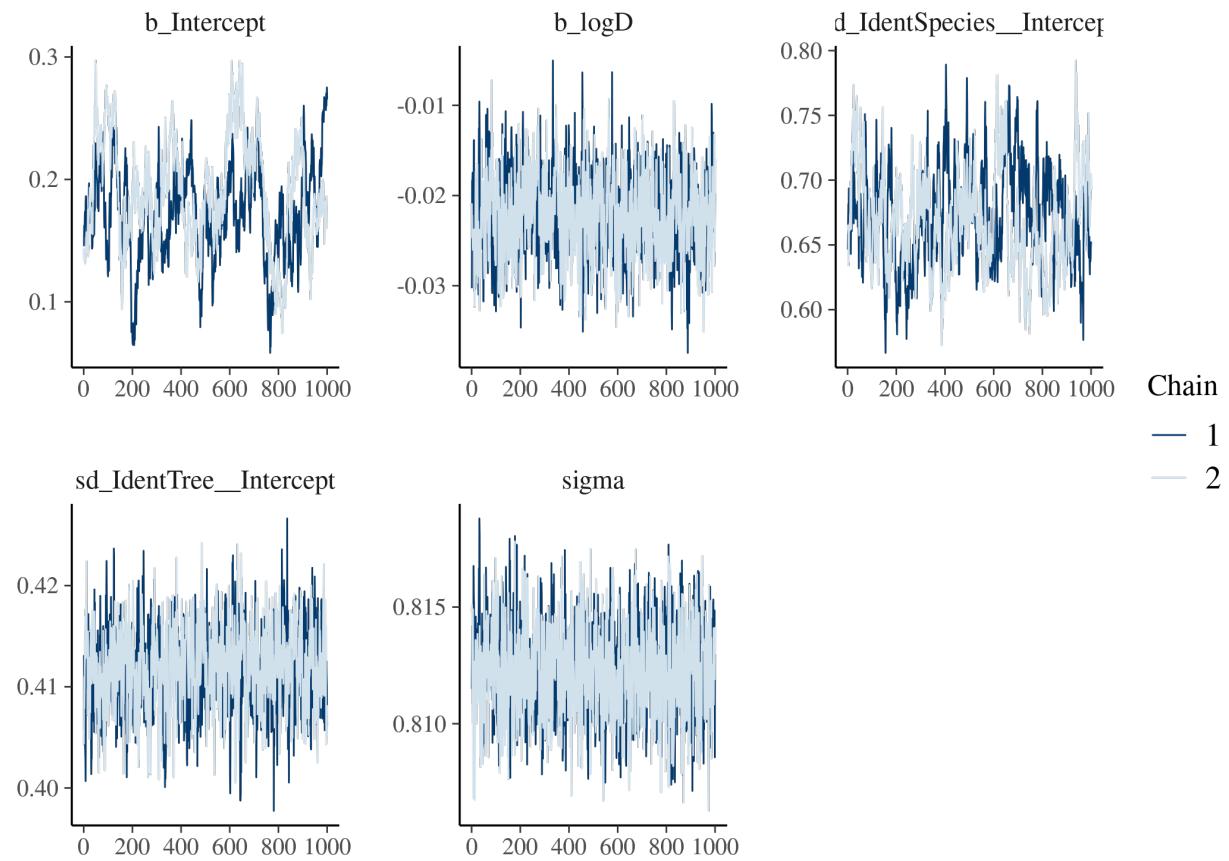


Figure 2.9: Trace of the posterior estimates for the BCI site

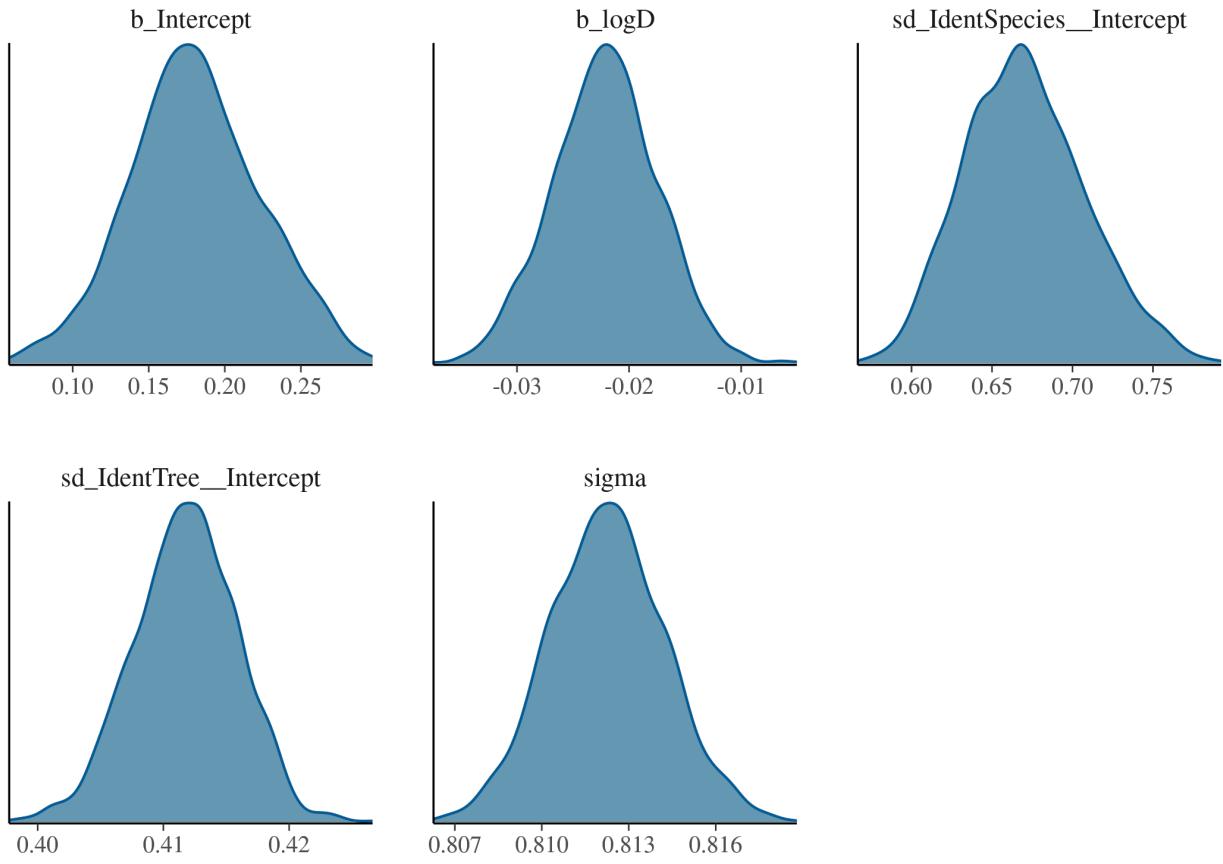


Figure 2.10: Density of the posterior estimates for the BCI site

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

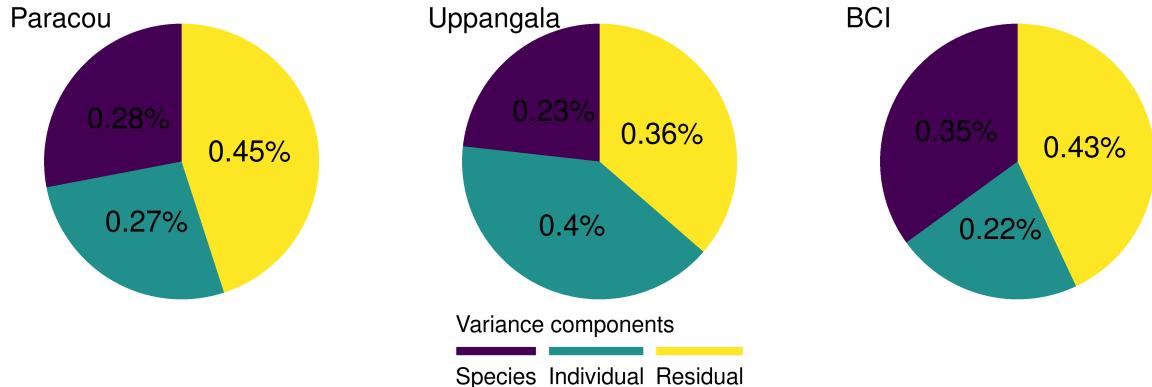


Figure 2.11: Variance partitioning of individual growth in three tropical forest sites (see Table above).

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 2.11).

2.4 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 (Paradis & Schliep 2019) 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 (Molino & Sabatier 2001). This resulted in a high heterogeneity and spatial structuration of the environment. When restricting our analysis to undisturbed control plots (see 2.6.1 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 (Hubbell 1999), while the mature forest in Uppangala has been reported as barely disturbed (Pélissier *et al.* 2011). 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

2.5 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. Semivariances 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

2.6 Other tests and verifications

2.6.1 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 %, for which individual growth is (i) significantly spatially autocorrelated, (ii) not significantly spatially autocorrelated, or (iii) that have too few individuals to allow a test to be made. 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.

2.6.2 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.

3 Analysis of an Eucalyptus clonal plantation dataset

In this analysis we tested whether intraspecific variability (IV) in observed individual tree growth can emerge from the environment only. To this aim, we used a clonal experimental setup. The EUCLUX common garden is a clonal trial where 14 Eucalyptus genotypes were grown in a replicated, statistically-sound design. One of its main goals is to determine and compare the productivity of each genotype. Our hypothesis was that IV in tree growth mainly results from responses to environmental factors, and not only from intrinsic genetic factors. Therefore, we used the EUCLUX dataset in order to quantify IV (within genotypes) for growth, i.e. in a dataset where genetically-driven IV is nil. Following our hypothesis, we expected to detect IV in growth within single genotypes.

3.1 The dataset

The EUCLUX experiment is located in Brazil, in the state of São Paulo. It includes 14 genotypes of 5 different species or hybrids, planted at the same date and grown in 10 replicated blocks of 100 trees each, which were monitored during 6 years. Each replicated block is distant by up to 1.5 km in a 200 ha stand showing some small changes in soil properties. The details of the setup can be found in (Maire *et al.* 2019).

We used the DBH measured during 6 complete censuses in order to compute mean annual growth in mm/year at five points in time. We computed the mean annual growth of each tree in mm/y as the difference between two consecutive censuses divided by the time between the two censuses. In case of mortality of the tree between two censuses, the data was discarded. We computed the neperian logarithm of diameter and growth (with a constant for growth in order to avoid undefined values).

The dataset included 64125 growth estimates corresponding to 13531 trees.

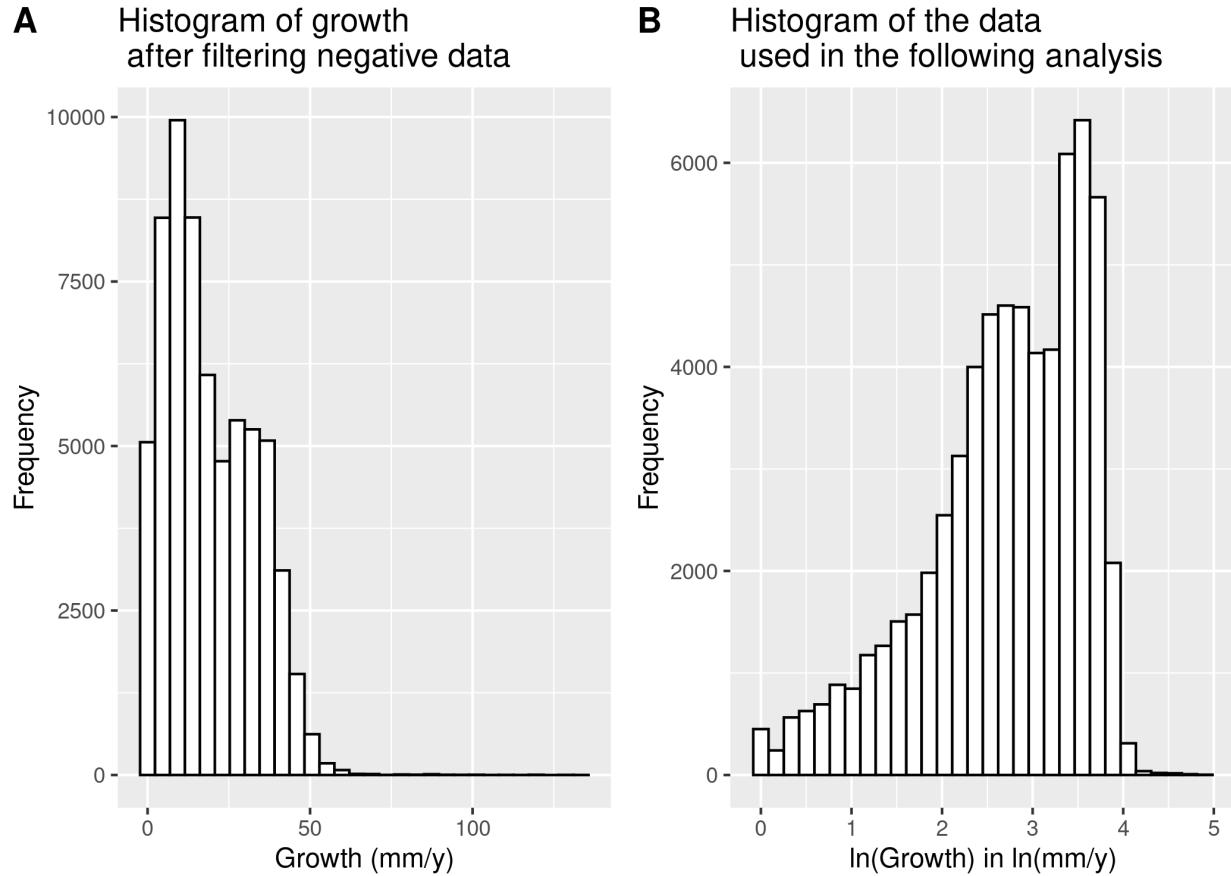


Figure 3.1: The original data without negative growth values (A) and the log-transformed growth (B).

Figure 3.1 shows the distribution of growth after removing negative values, and the latter with log-transformed values.

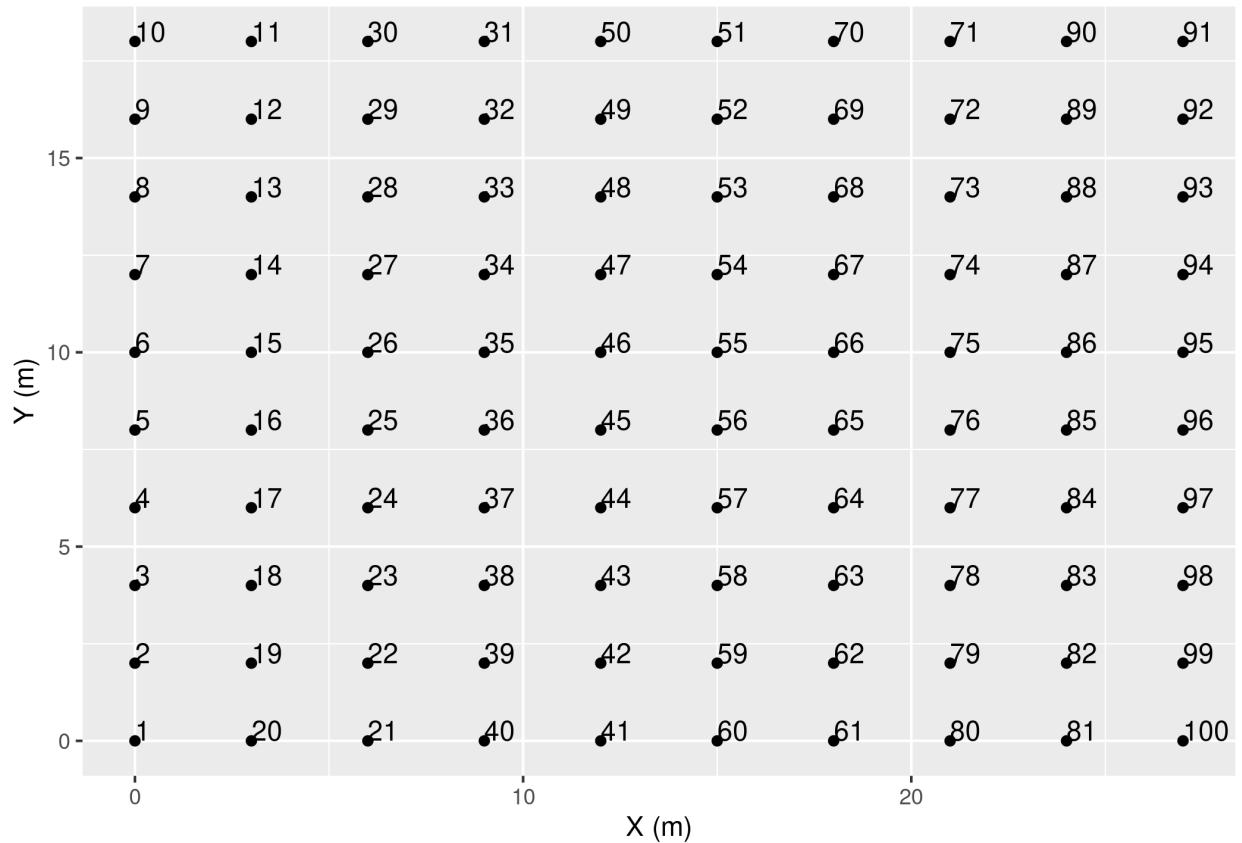


Figure 3.2: Design of a plot. Each point is a tree and the associated number is the tag of the tree.

Figure 3.2 shows the disposition of the trees in a single plot. There are 14 genotypes times 10 repetitions, so 140 plots with this same design.

3.2 Competition index

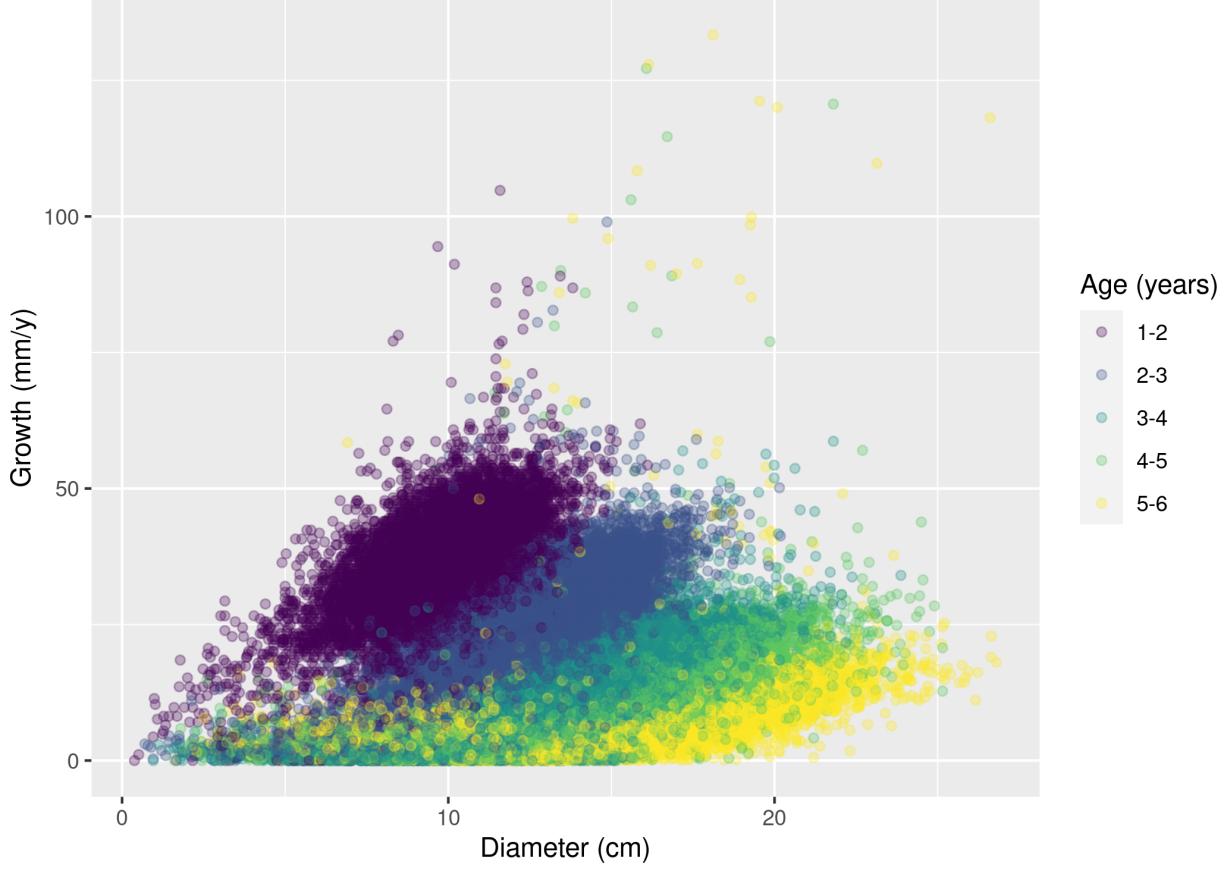


Figure 3.3: Plot of the growth versus the diameter. Each colour represents a tree age.

Figure 3.3 shows the age of the trees has a big influence on the values of growth but also on the relationship between growth and diameter : the slope is smaller with time, indicating that for the same diameter, growth is slower through time. This is likely an effect of competition for light and possibly underground resources, since as the trees grow their capacity to capture resources increases. Therefore, we computed a competition index $C_{i,t}$ to integrate this effect in the growth model. The competition index is computed for each tree which is not on the edge of a plot. It is the sum of the basal areas (BA) of the 8 direct neighbours divided by the area of the rectangle that comprises all the direct neighbours. It was then log-transformed. Dead neighbours are considered as having a null BA.

$$C_{i,t} = \sum BA_{neighbours(i,t)}$$

3.3 Statistical growth model

In order to partition the variance of individual growth data, we built a model incorporating a fixed effect on the intercept (β_0), on the slope of diameter D (β_1), and on the competition index C (β_2) and several random effects, namely temporal (date of census, b_t), individual (tree identifier, b_i), spatial (block, b_b), and genotype (b_g).

This model was fitted in a hierarchical Bayesian framework using the brms package (Bürkner 2017, 2018)

with 10000 iterations, a warming period of 5000 iterations and a thinning of 5. We obtained 1000 estimates per parameter.

Variables were scaled to help the convergence of the model.

$$\ln(G_{it} + 1) = (\beta_0 + b_i + b_b + b_g + b_t) + \beta_1 \times \ln(D_{it}) + \beta_2 \times \ln(C_{it}) + \epsilon_{it}$$

Priors

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

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

$$\beta_2 \sim \mathcal{N}(\text{mean} = 0, \text{var} = 1), iid$$

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

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

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

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

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

Hyperpriors

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

3.4 Results of the model and variance partitioning

After convergence of the model (checked visually), we examined the proportion of the residual variance (variation of the response variables that is not explained by the covariates , i.e. in the unexplained part of the statistical model) related to each random effect, and this enabled us to perform a residual variance partitioning.

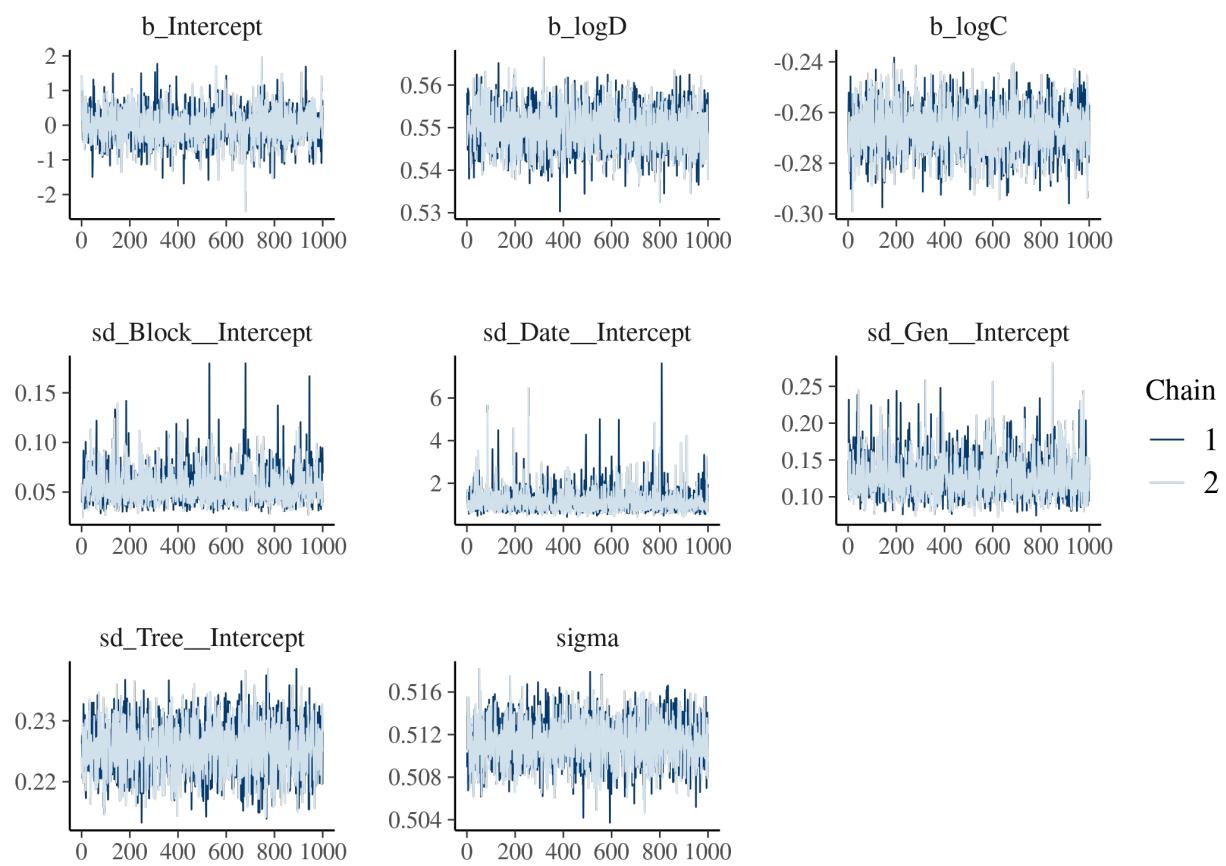


Figure 3.4: Trace of the posteriors of the inferred parameters

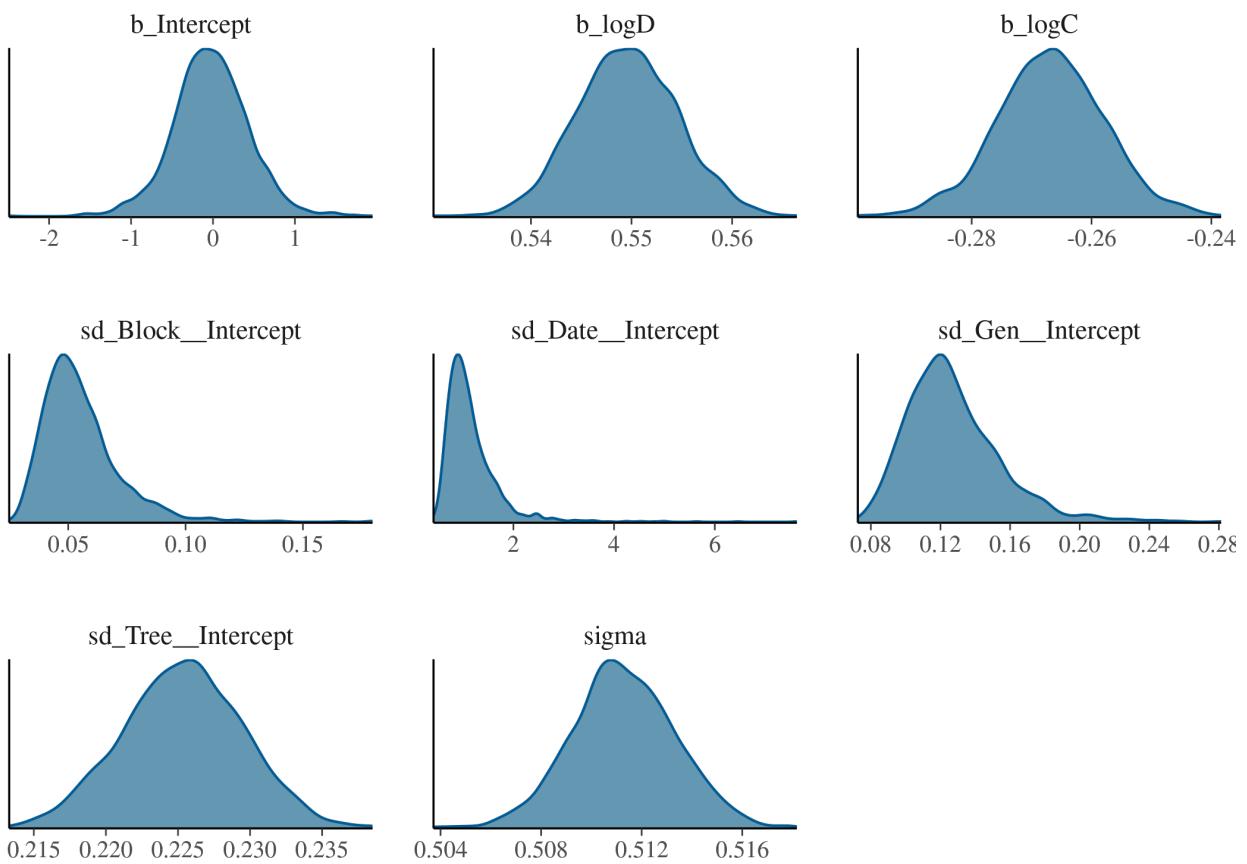


Figure 3.5: Density of the posteriors of the inferred parameters

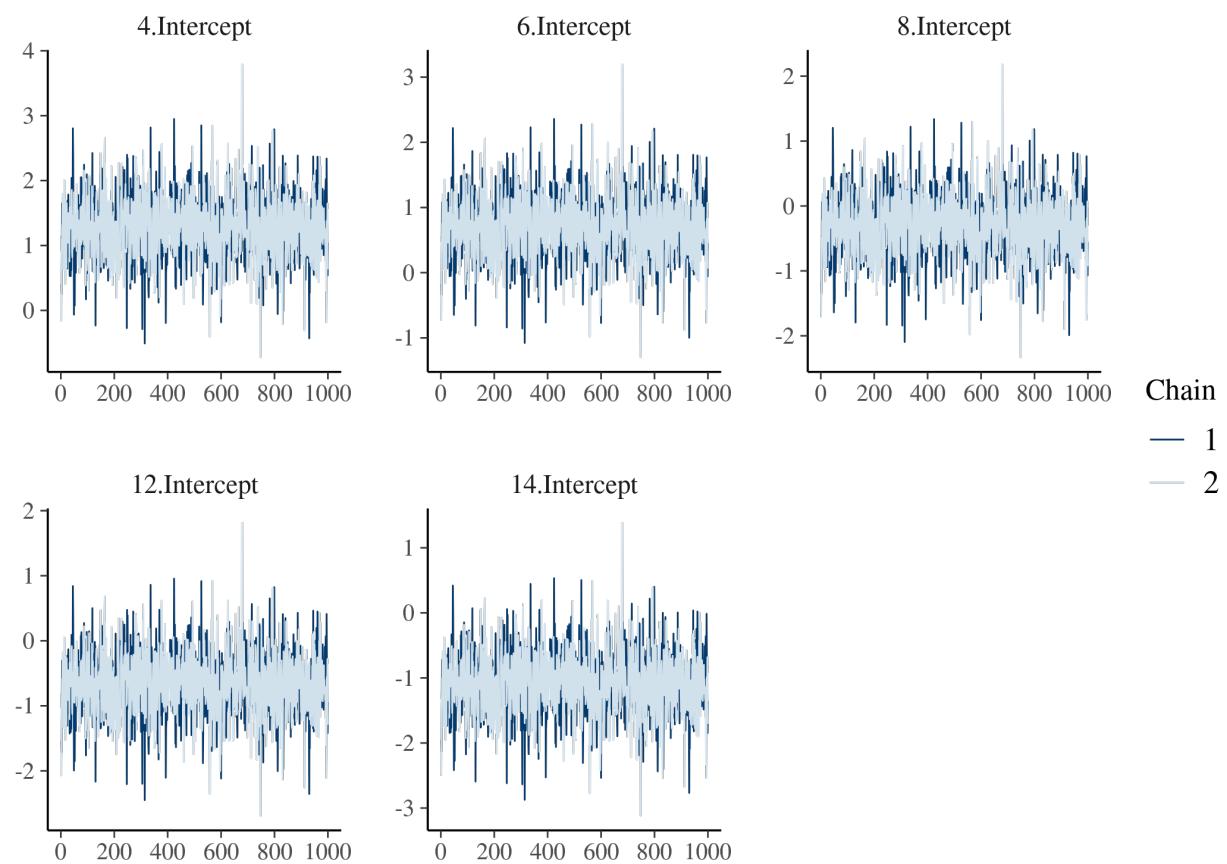


Figure 3.6: Trace of the temporal random effects

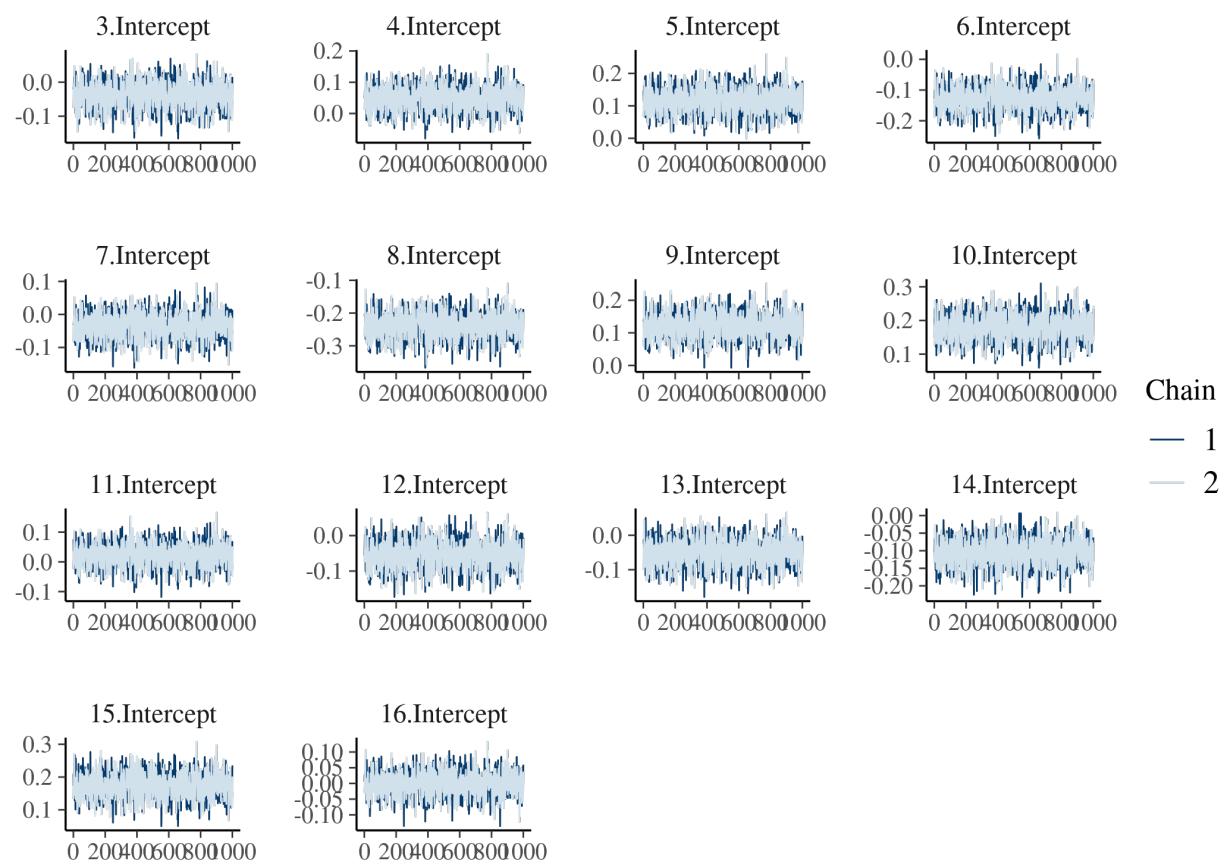


Figure 3.7: Trace of the genotype random effects

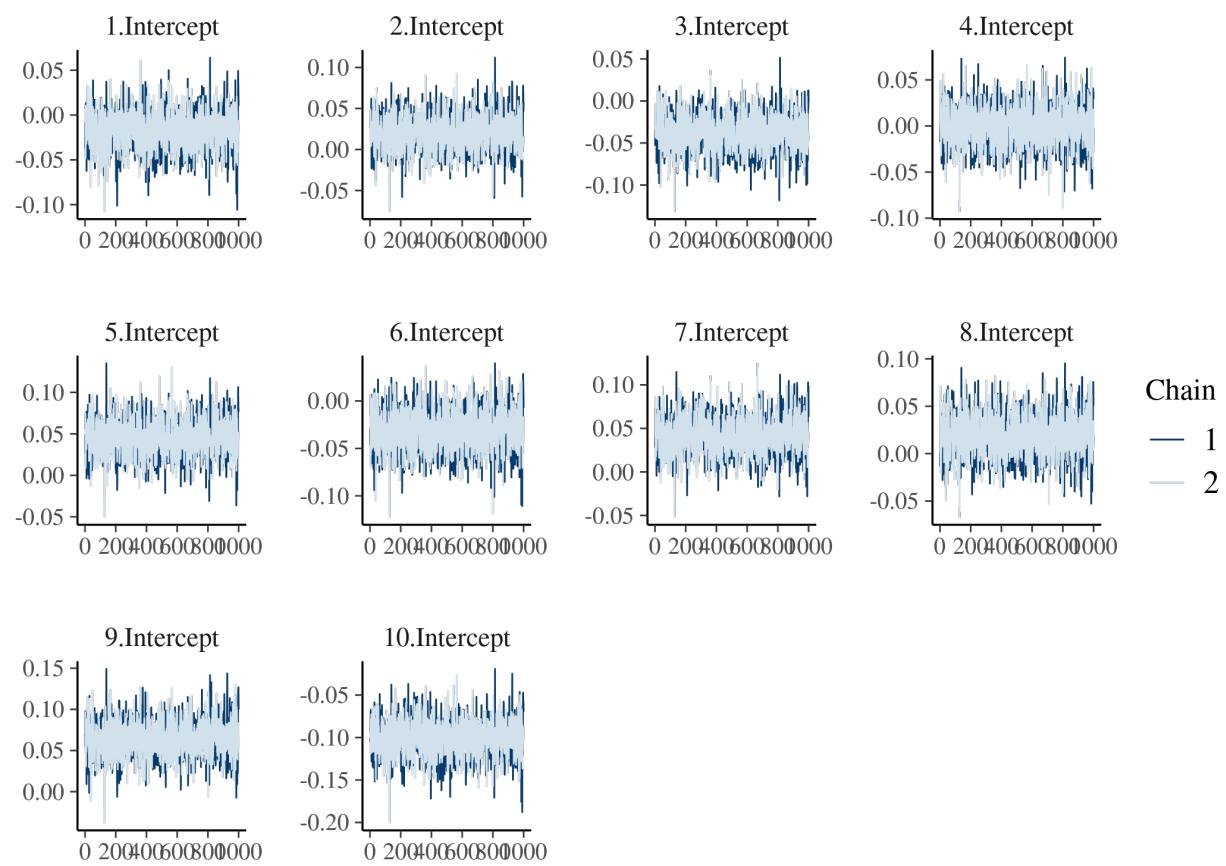
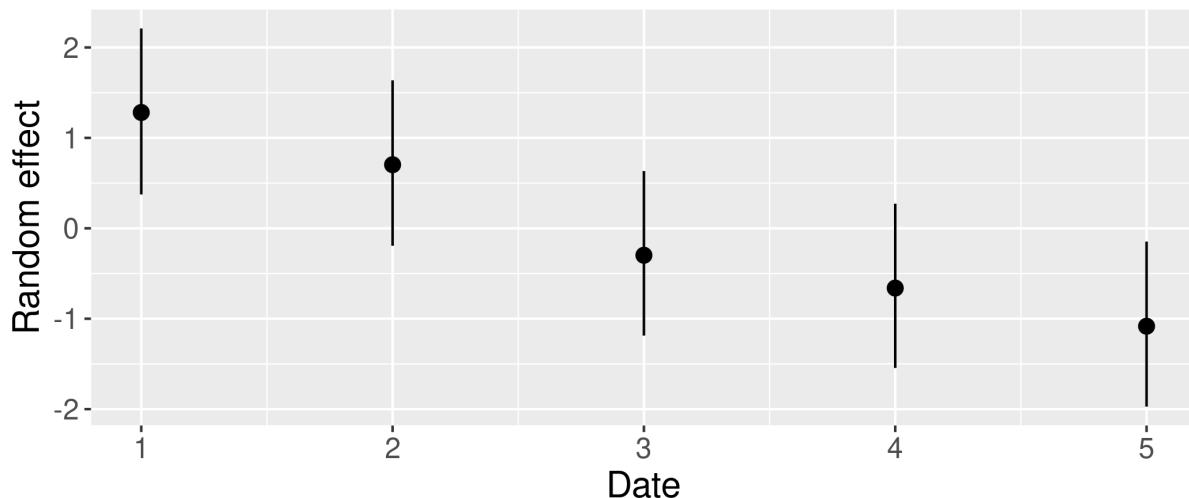
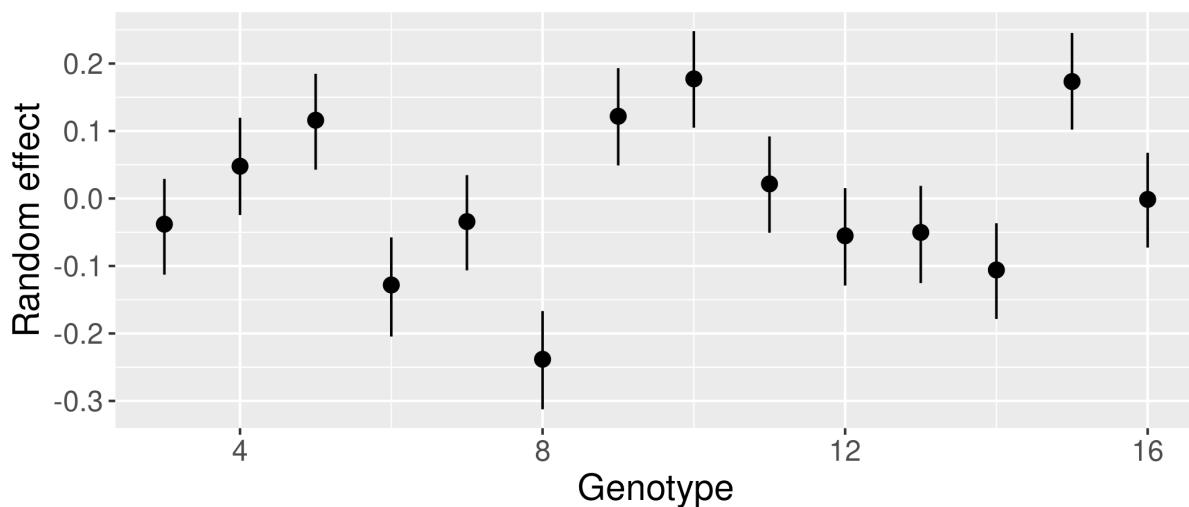


Figure 3.8: Trace of the spatial (block) random effects.

A Temporal random effects



B Genotype random effects



C Block random effects

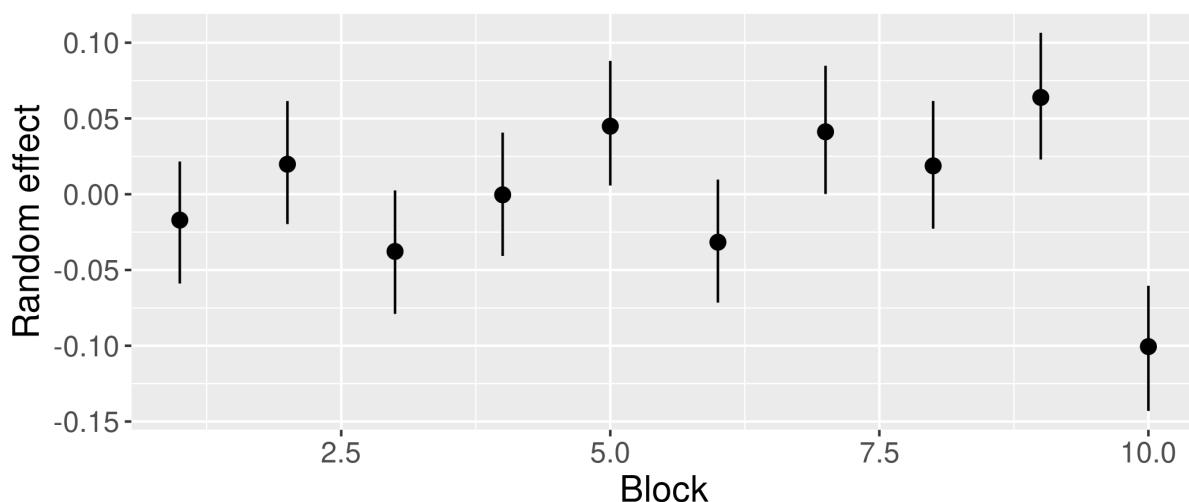


Figure 3.9: Mean values and 95% confidence interval of the temporal, genetic and spatial and random effects.

Table 6: Mean posteriors of the model and their estimation errors.

	Intercept (β_0)	Diameter (β_1)	Competition (β_2)	Individual variance (V_i)	Block variance (V_b)	Genetic variance (V_g)	Temporal variance (V_t)	Residuals variance (V)
Estimate	-2.5e-02	5.5e-01	-2.7e-01	2.3e-01	5.5e-02	1.3e-01	1.2e+00	5.1e-01
Estimation error	4.6e-01	5e-03	9e-03	4.1e-03	1.6e-02	2.7e-02	5.5e-01	2e-03

We found that the two most important contributors to variance were the date and individual identity. High estimation error for the intercept and the temporal random effect must be noted. The proportion of variance represented by each random effect and the residual variance were computed to visualise the variance partition.

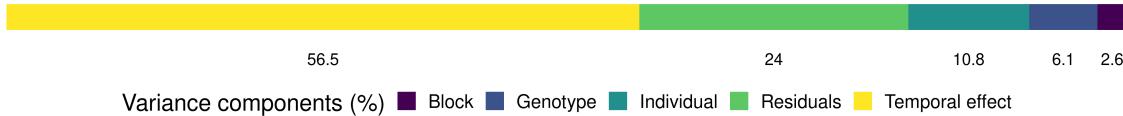


Figure 3.10: Proportion of each variance component of the unexplained variance.

The model showed that individual tree growth was a function of tree size and competition with neighbouring trees, and that variance around this model was mostly due to a temporal effect as well as an individual effect (Table 6, Figure 3.2). The effect of the genotype was quite small, and the effect of the block was even smaller. The temporal random effects declined with time (Figure 3.9, panel A), showing that the effect of the date on growth is negative (the older the trees become, the less they can grow). We attribute this tendency to competition. Therefore, the competition index C did not fully capture the effect of competition on growth. Another explanation is that the diameter slope was not able to fully capture the decrease of tree growth with size, maybe due to geometrical effects of distributing growth around increasing diameter or physiological constraints linked to height. The temporal effect explained the highest fraction of variance (Table 6, Figure 3.2). This could be due to the negative effect of competition for light, water, and/or nutrients on growth, which increases with the growth of the trees planted at high densities, and to physiological changes occurring with age.

Importantly, variability between individuals within genotypes (10.8%) is higher than between genotypes (6.1%). This shows that there is an observed individual variability within genotypes even if trees are clones. Estimated individual variability within genotypes can only be caused by exogenous factors in this particular case. Individual effects can be due to the micro-environment where each tree thrives in, but also to some individual history, such as seedling manipulation and plantation.

The block had the littlest impact with 2.5% of the variance explained. This is probably due to the fact that environmental conditions between blocks are quite homogeneous. As the experimental design aimed at minimizing environmental variations and selected productive genotypes able to accommodate several environmental conditions (Maire *et al.* 2019), this dataset is a strongly conservative test case for our hypothesis.

3.5 Conclusion

Overall, we found that there is IV within clonal tree plantations (Figure 3.11). This shows that IV can be due only to the environmental factors varying between individuals, proving that the source of IV is not necessarily intrinsic.

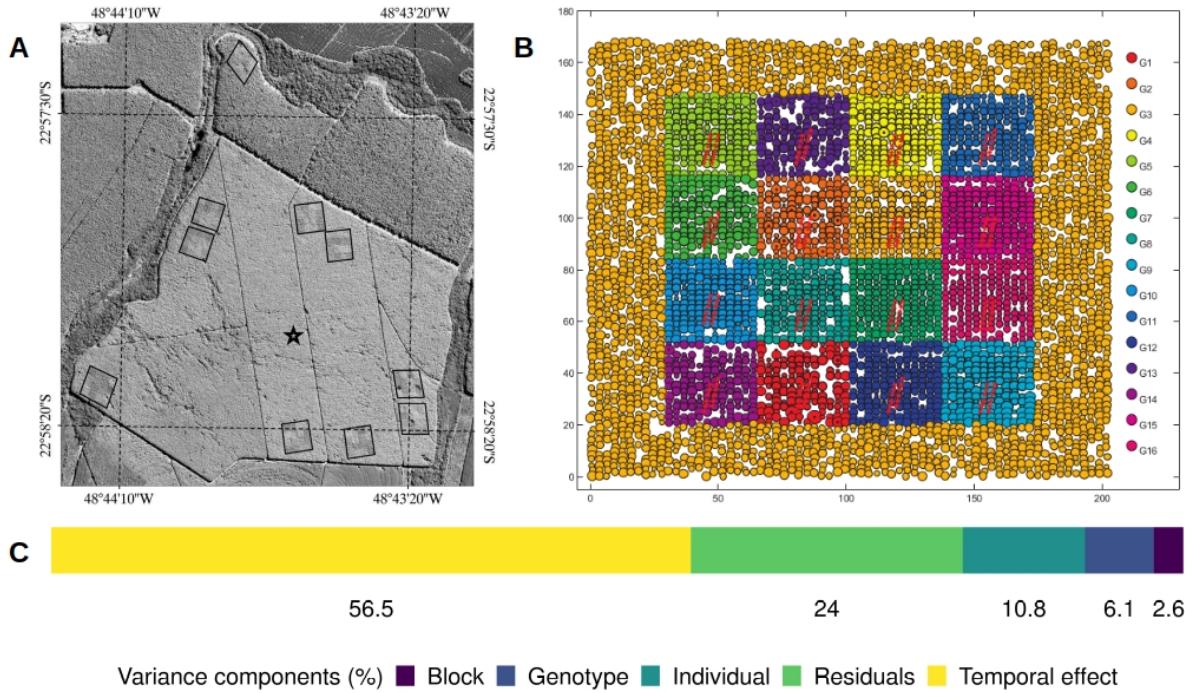


Figure 3.11: The EUCLUX setup and the variance partitioning. Site (A; each square represents a block) and organisation of a block (B; each coloured square represents a genotype) in the EUCLUX experiment. 16 clones are represented in B, but only 14 were used since the last two are from seed-origin and thus not genetically identical. (C) Variance partitioning of tree growth for a common garden experiment with various Eucalyptus clones

4 Virtual experiment with two species

Intraspecific variability (IV) is often seen as an intrinsic property of individuals, associated with genetics, which is unstructured in space and time. Here, we investigated whether it can result from the species responses to the spatial variations of the environment. Our hypothesis was that the individuals can be clones, i.e. have no genetic differences between them, and still have different measured attributes because the environment in which they strive varies, as shown in the previous analysis of a clonal dataset. The response of an individual results from the integrated effects of the environmental conditions which the individual has encountered during its life (local environmental variation, also called microsite effect) and its genetic features. Moreover, the differences between individuals due to environmental variation does not imply a broader overlap of the species niches. Therefore, this type of IV has no direct effect on species coexistence.

Here, we designed and conducted a virtual experiment that aims at illustrating that IV, or “individual effects,” can result from variation in unobserved environmental variables (Clark *et al.* 2007), therefore accounting for multidimensional species differences which cannot be observed on a few niche axes.

4.1 Perfect knowledge model : generating response values.

To do so, we first considered a “perfect-knowledge” model that depicts the ecological response (e.g. growth) of individual clones for two different species as a function of a set of environmental variables. This model represents the perfect knowledge of the process as it occurs in the field and thus includes no residuals :

$$Y_{ijt} = \beta_{0j} + \beta_{1j} * \ln(X_{1ijt}) + \beta_{2j} * X_{2ijt} + \dots + \beta_{Nj} * X_{Nijt} \quad \text{"Perfect knowledge model"}$$

where $1 \leq i \leq 500$ are the individuals ; $1 \leq j \leq 2$ are the species ; Y is a response variable, for instance growth ; $X1$ to XN are explanatory variables, for instance environmental variables, that can vary with time t ; and the $\beta_{k,i}$, $1 \leq k \leq N$ depict the species-specific responses to these environmental variables. As we consider clones, individuals of the same species respond the same way to environmental variables, and variation in Y among individuals within species is due to differences in the environment where each individual thrives. The first explanatory variable was natural log-transformed to represent a log-log relationship as would be the relationship between tree growth and light for example.

We fixed the species parameters of the “Perfect knowledge model” as follows, using 10 environmental variables ($N=10$) :

Parameters of species 1 : $\beta_0 = 0.25$, $\beta_1 = 0.15$, β_2 to β_6 were chosen randomly between -0.05 and 0.05 and β_7 to β_{10} were chosen randomly between 0 and 0.05.

Parameters of species 2 : $\beta_0 = 0.2$, $\beta_1 = 0.1$, β_2 to β_{10} were chosen the same way as for species 1.

The difference between the species is imposed by those parameters. Here, species 1 was more competitive on average thanks to its higher intercept (β_0) and response to the first environmental variable (β_1). The first environmental variable ($X1$) has a higher weight in the computation of the response variable, as would be the most limiting factor for the response variable.

To account for genetic variability in our generated data, we added an IV in species parameters by sampling each individual parameter in a normal distribution centered on the species mean parameter and with a standard deviation of a quarter of the species mean parameter.

To represent the spatialised environment, we build a 2D matrix of dimension 500×500 for each environmental variable at a time t_0 , by randomly generating them with spatial autocorrelation. We here considered that all environmental variables were independent. Each variable was simulated by using the gstat R package (Pebesma 2004 ; Gräler *et al.* 2016), enabling to create autocorrelated random fields. A spherical semivariogram model was used for each of the ten environmental variables, with a mean of 0 for each explanatory variable (beta = 0), a sill of 1 (psill = 1) for each and a range of 50 (range = 50).

We then considered that Y had been measured at two times, t_0 and t_1 , for each individual as it would be under periodic forest plot censuses for instance. A pair of coordinates within this spatialised environment was randomly assigned to each of the 500 individuals. We considered that some of the environmental variables (light, temperature, humidity, nutrient availability for instance) had changed between t_0 and t_1 and others had not (slope, altitude for instance). For the first environmental variable which had the stronger impact on Y values and another randomly drawn environmental variable, we computed values at t_1 as $t_0 + \epsilon$, $\epsilon \sim \mathcal{N}(0, 0.1)$ (they randomly increase or decrease). For two other environmental variables randomly drawn, $X_{t_1} = X_{t_0} + |\epsilon|$, $\epsilon \sim \mathcal{N}(0, 0.1)$ (they increase) and for two other environmental variables, $X_{t_1} = X_{t_0} - |\epsilon|$, $\epsilon \sim \mathcal{N}(0, 0.1)$ (they decrease).

This led to two repeated measurements of Y for each individual of each species, i.e. 2000 values of $Y_{i,j,t}$, with $i = [1; 2]$; $j = [1, 50]$; $t = [t_0; t_1]$.

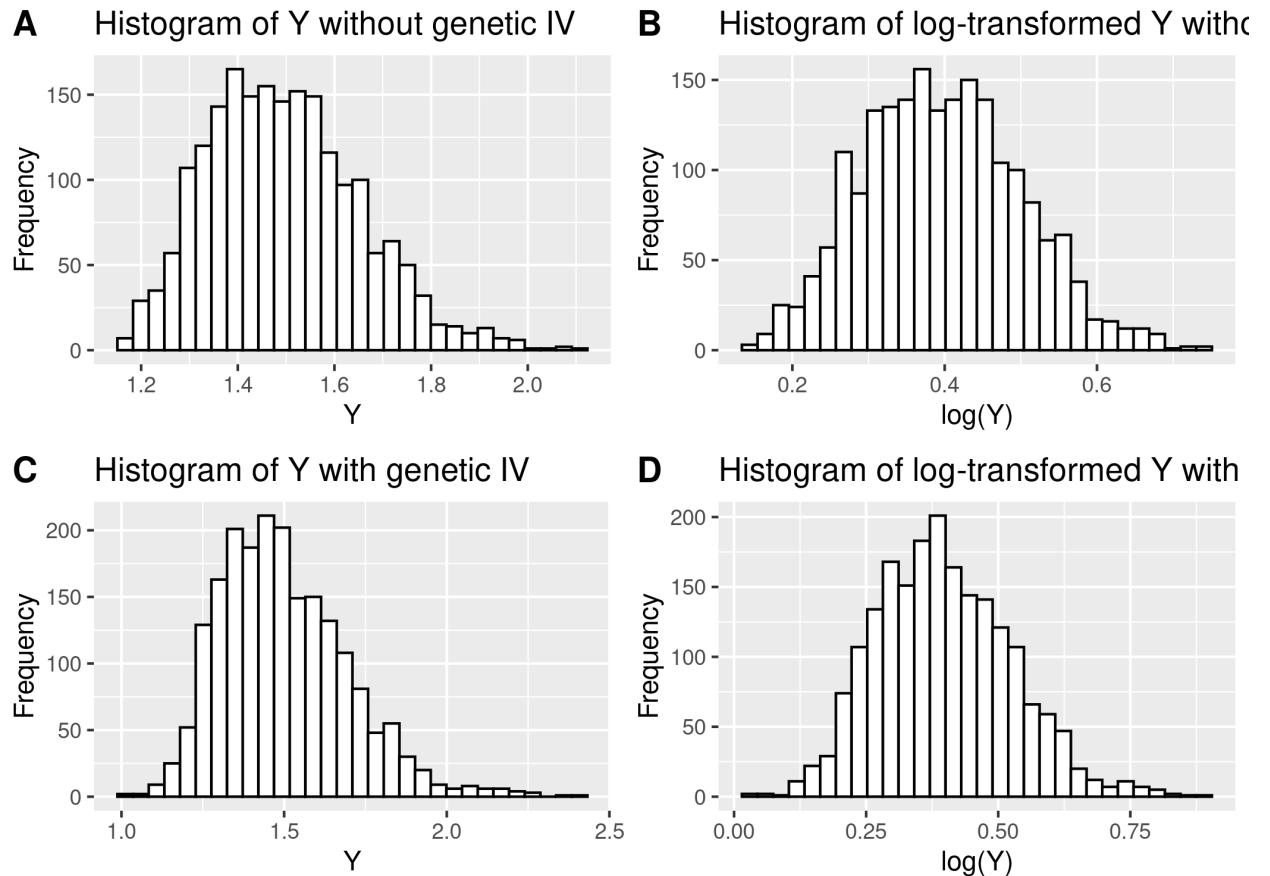


Figure 4.1: Histogram of raw and log-transformed Y with and without genetic IV

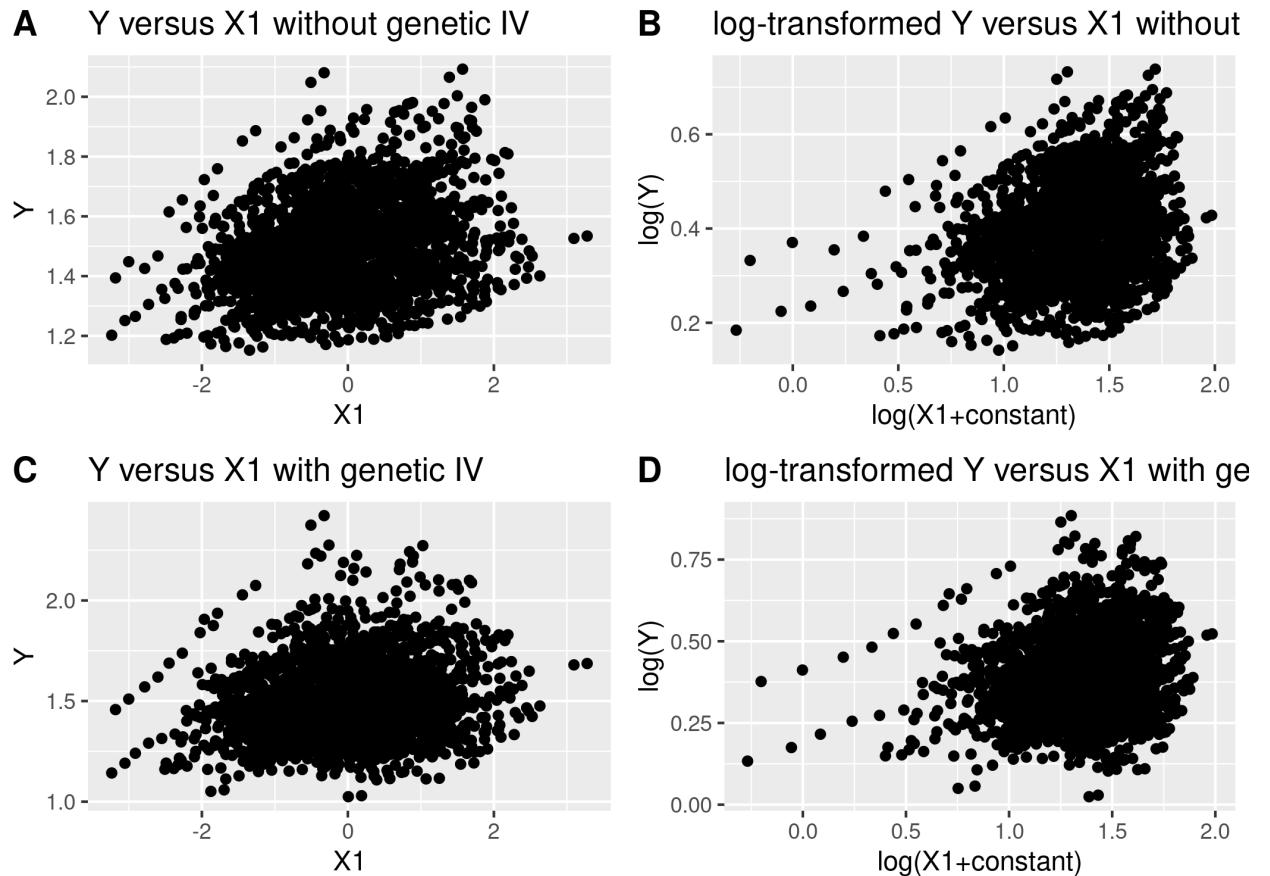


Figure 4.2: Raw and log-transformed Y versus X1 plots with and without genetic IV

Figure 4.1 shows the distribution of the data and Figure 4.2 shows the relationship between the response Y and the environmental variable X1.

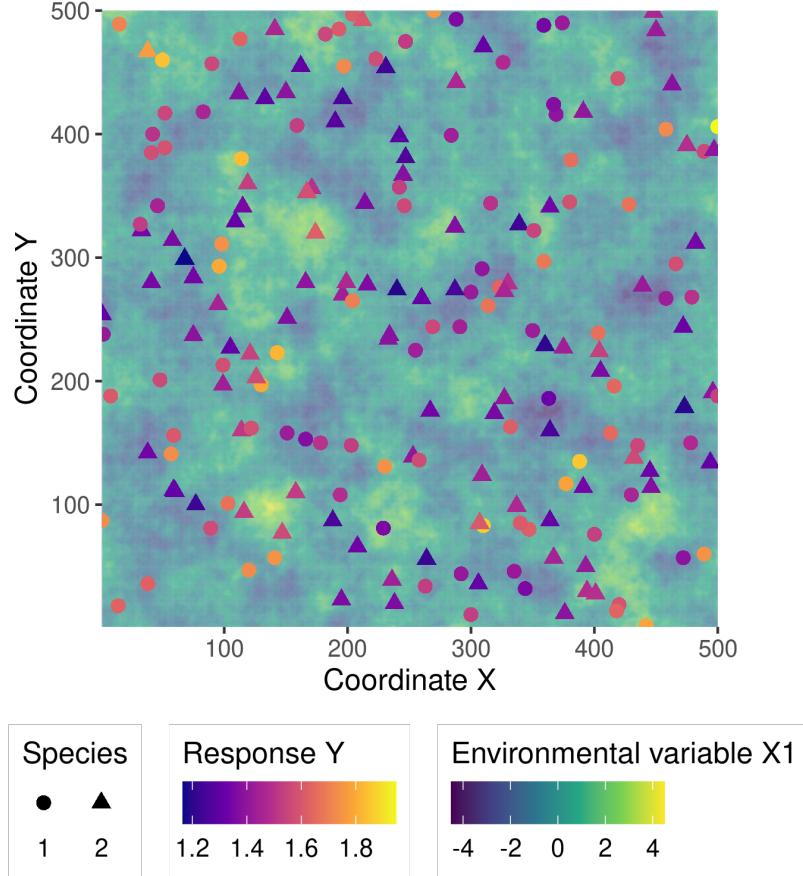


Figure 4.3: Virtual landscape representing X_1 and the individual with their respective Y value.

Figure 4.3 shows that microsite effects, which are the effects of the local multidimensional environment on the response variable, can result in local reversals of the competitive hierarchy between species. Indeed, the local outcome of competition can be opposite to the mean hierarchy : at one point of the space-time, an individual of Species 1 can overcome an individual of Species 2, whilst Species 1 is, on average, the fittest of the two species. Consequently, niche multidimensionality and variation of the environment in space allow the coexistence of Species 1 and Species 2.

4.2 Partial knowledge model : interpreting the response values with only one explanatory variable.

These two virtual datasets (with and without genetic variability) were then analysed with a “Partial knowledge model,” which represents the model derived by ecologists when using datasets derived from an incomplete characterisation of the environment : only a few variables (here only one, X_1) are actually measured and taken into account.

$$\ln(Y_{ijt}) = [\beta'_{0j} + b_{0i}] + \beta'_{1j} \ln(X_{1ijt}) + \varepsilon_{ijt} \quad \text{“Partial knowledge model”}$$

Priors :

$$\varepsilon_{ijt} \sim \mathcal{N}(0, V_j), j = [1, 2], \text{iid}$$

$$\beta'_{kj} \sim \mathcal{N}_2(0, 1), k = [0, 1], j = [1, 2], \text{iid}$$

$$b_{0i} \sim \mathcal{N}(0, V_{bj}), i = [1, 500], j = [1, 2], \text{iid}$$

Second level priors :

$$V_j \sim \mathcal{IG}(10^{-3}, 10^{-3}), j = [1, 2], \text{iid}$$

$$V_{bj} \sim \mathcal{IG}_2(10^{-3}, 10^{-3}), j = [1, 2], \text{iid}$$

This model includes a random individual effect on the intercept (b_{0i}) to account for the variability among individuals within species. We ran this model twice, with the datasets generated with and without genetic IV. These models were fitted in a hierarchical Bayesian framework using the brms package (Bürkner 2017, 2018) with 100000 iterations, a warming period of 50000 iterations and a thinning of 50. We obtained 1000 estimates per parameter.

We visualised the convergence and the results of the models thanks to trace and density plots.

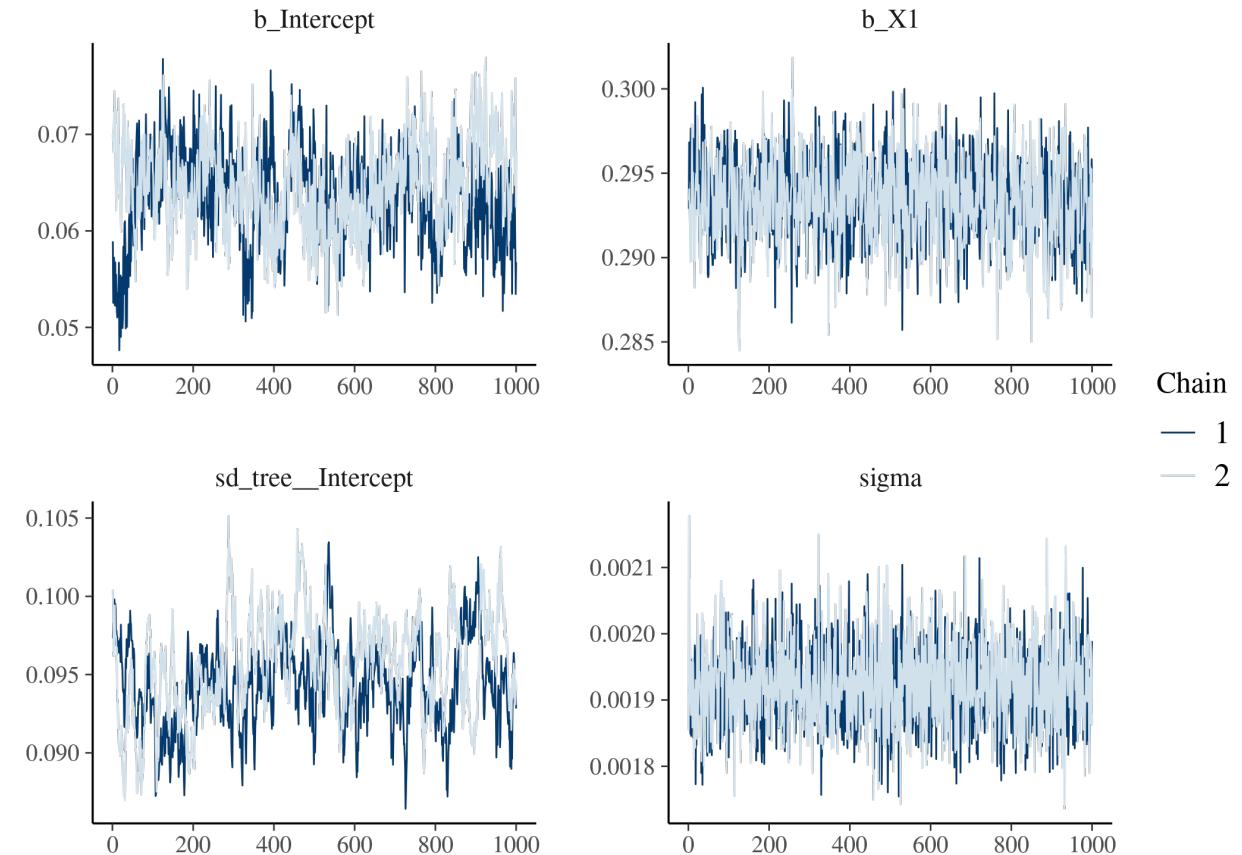


Figure 4.4: Trace of the posterior of the model for Species 1 without genetic IV

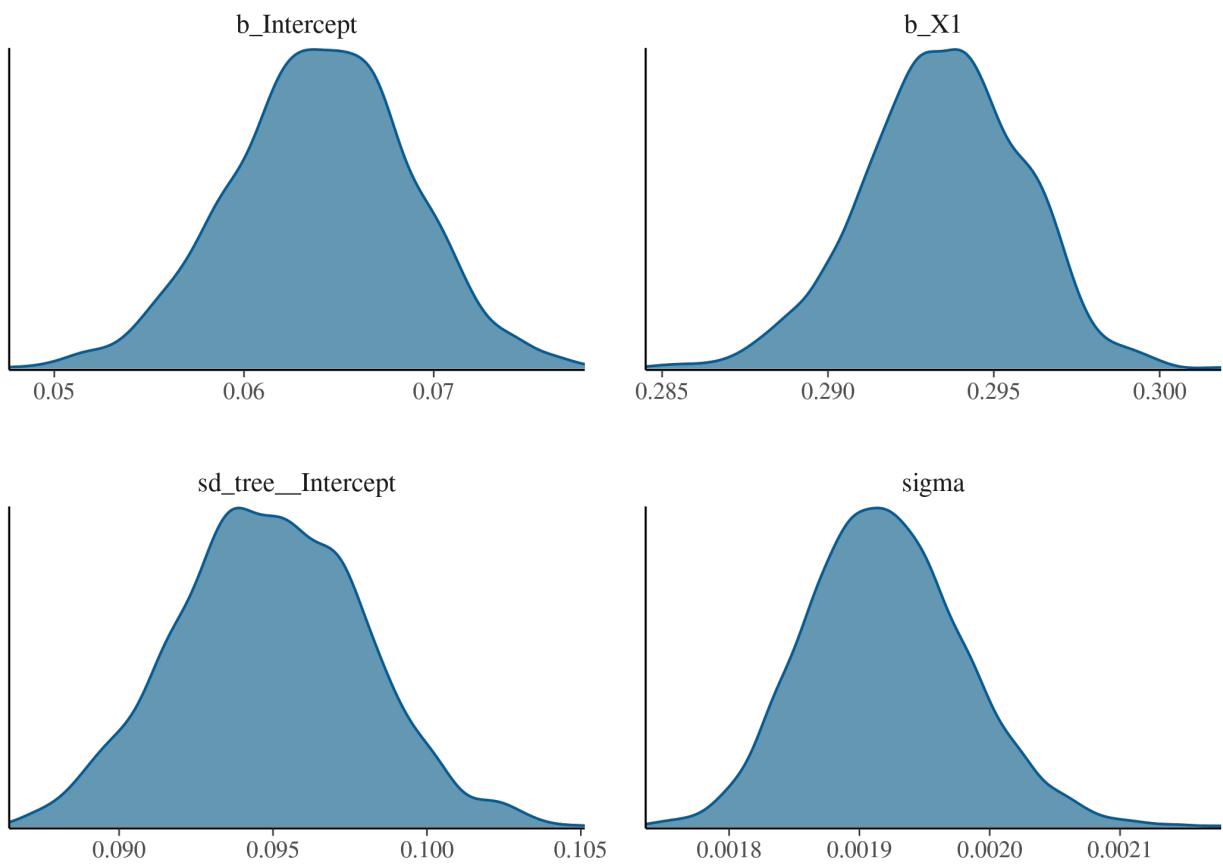


Figure 4.5: Density of the posterior of the model for Species 1 without genetic IV

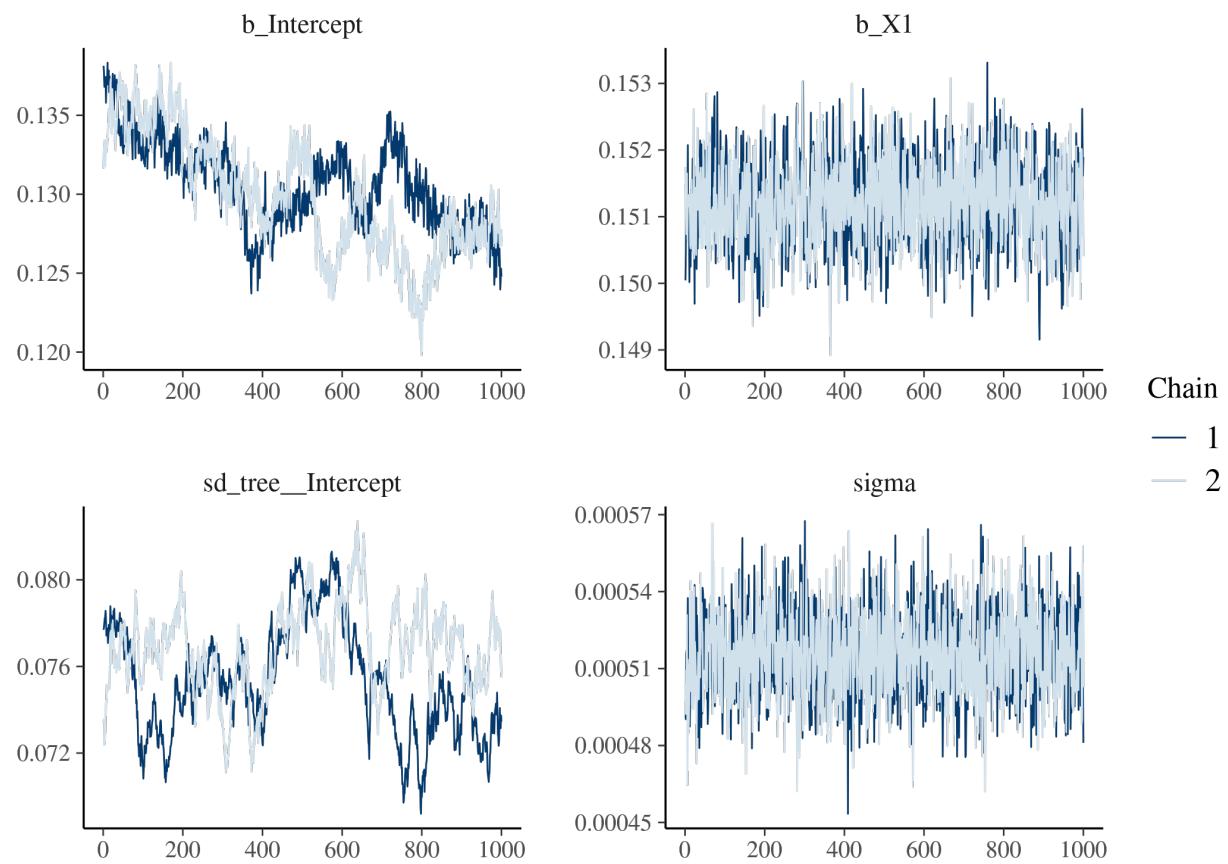


Figure 4.6: Trace of the posterior of the model for Species 2 without genetic IV

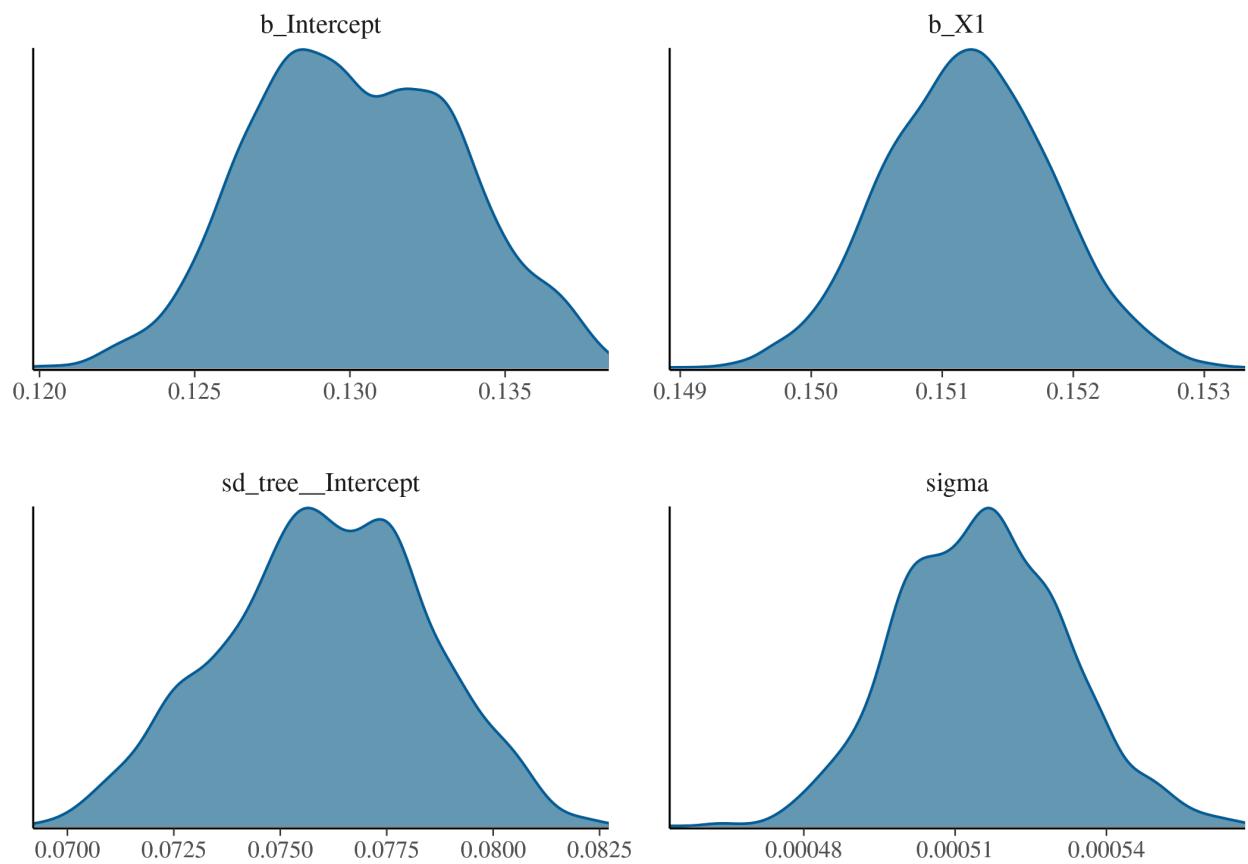


Figure 4.7: Density of the posterior of the model for Species 2 without genetic IV

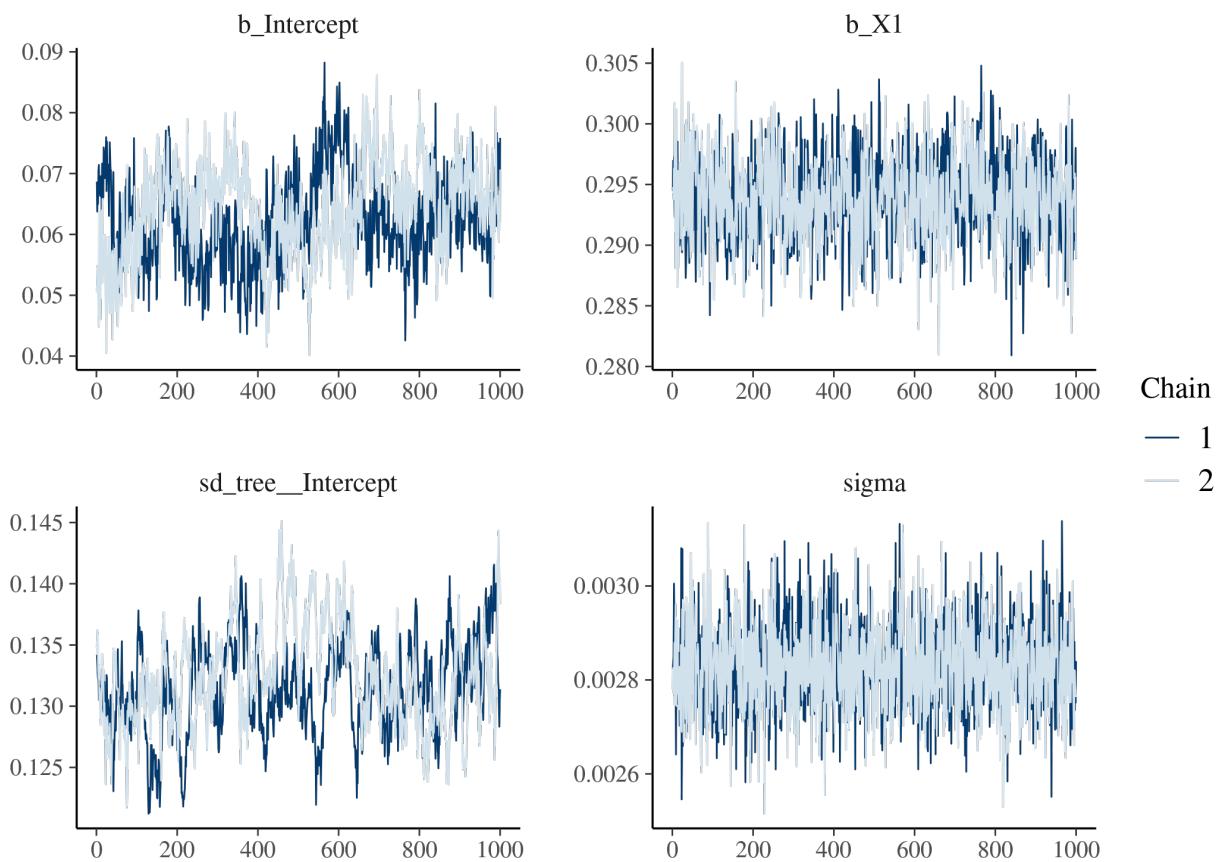


Figure 4.8: Trace of the posterior of the model for Species 1 with genetic IV

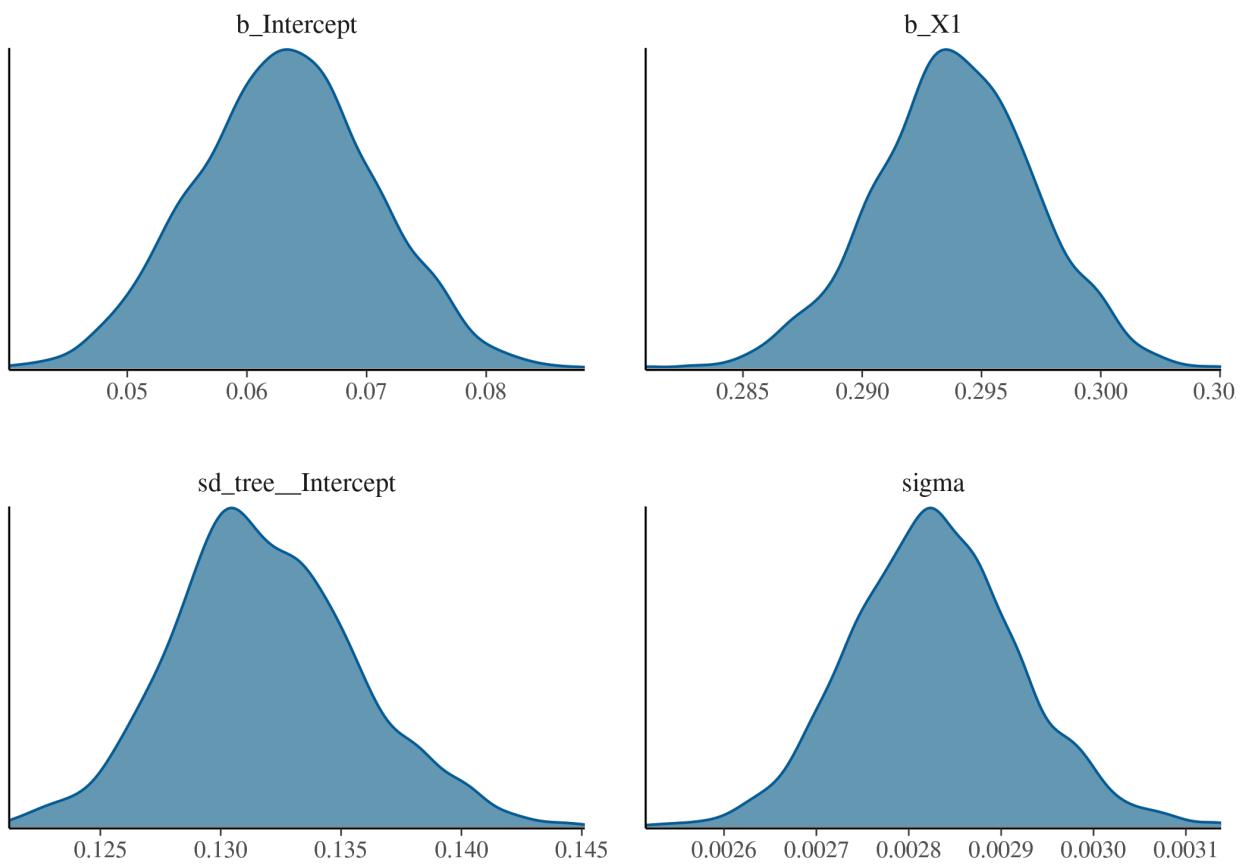


Figure 4.9: Density of the posterior of the model for Species 1 with genetic IV

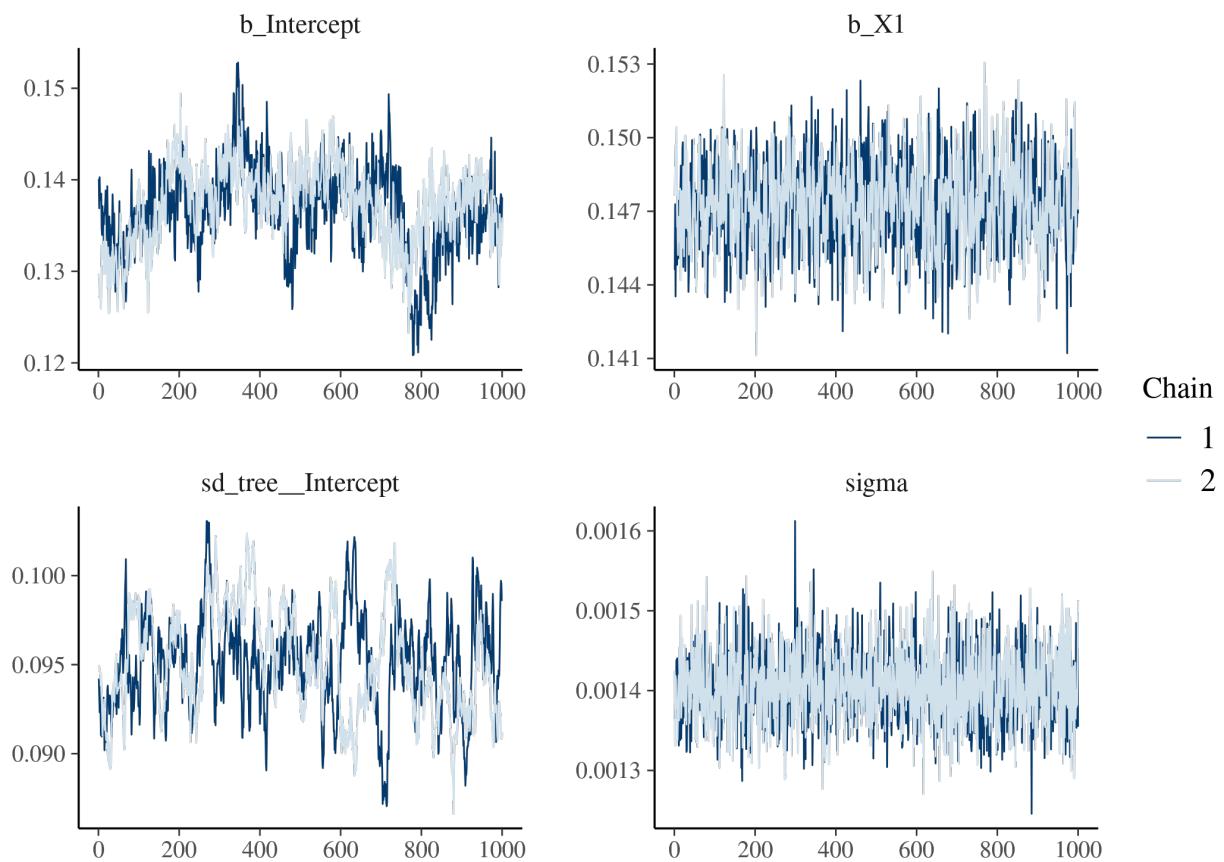


Figure 4.10: Trace of the posteriors of the model for Species 2 with genetic IV

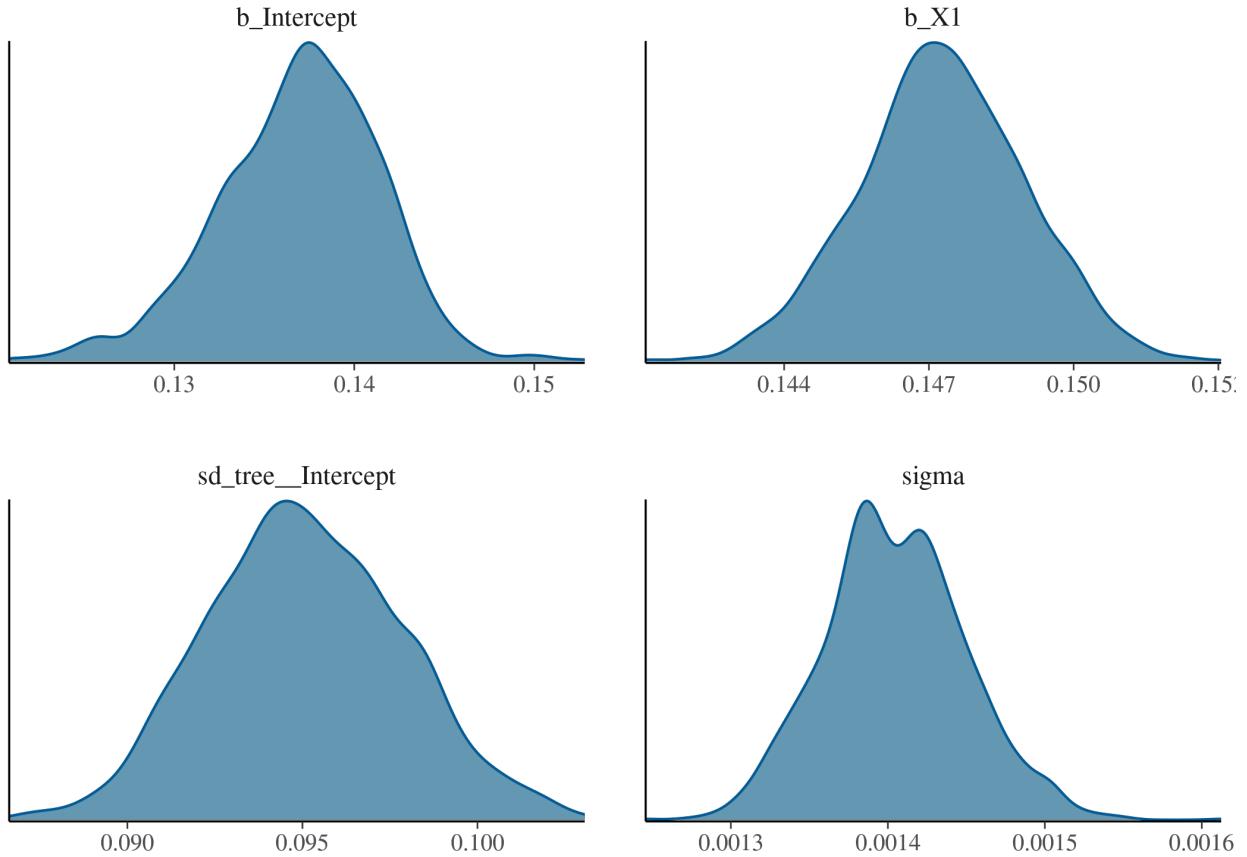


Figure 4.11: Density of the posteriors of the model for Species 2 with genetic IV

```
## Warning: Parts of the model have not converged (some Rhats are > 1.05). Be careful when analysing the
## results! We recommend running more iterations and/or setting stronger priors.
```

Table 7: Mean posteriors and their estimation error

	β_0	β_1	V_b	V
Species 1				
Estimate	6.4e-02	2.9e-01	9.5e-02	1.9e-03
Estimation error	4.7e-03	2.4e-03	3e-03	6.1e-05
Species 2				
Estimate	1.3e-01	1.5e-01	7.6e-02	5.2e-04
Estimation error	3.3e-03	6.3e-04	2.4e-03	1.7e-05

We inferred a high IV even in the absence of genetic IV (Table 7). Therefore, observed IV does not necessarily reveal intrinsic (mainly genetic) IV, but can also reveal hidden dimensions of the environment. We used the model parameters to plot the relationship between Y and $X1$ accounting for intraspecific variability.

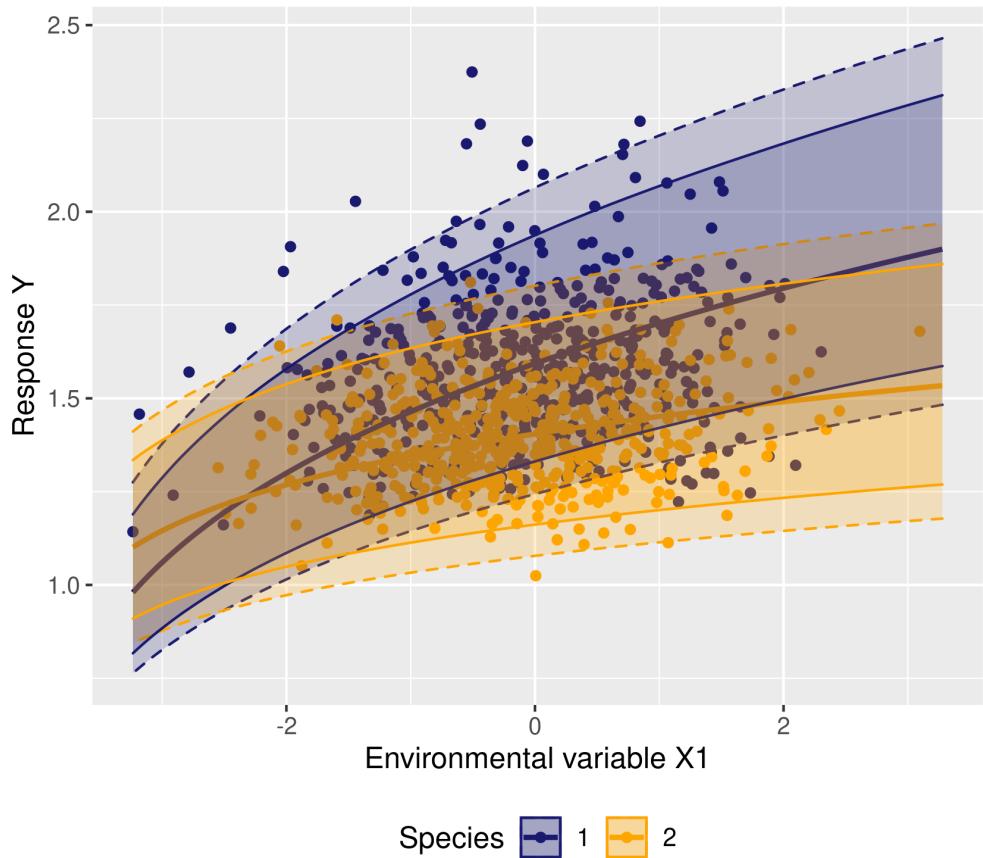


Figure 4.12: Plot of the real values - points - and estimated mean - bold line - and 95 % confidence interval - thin lines - of Y versus X1. The dotted lines correspond to the 95 % interval due to genetic IV.

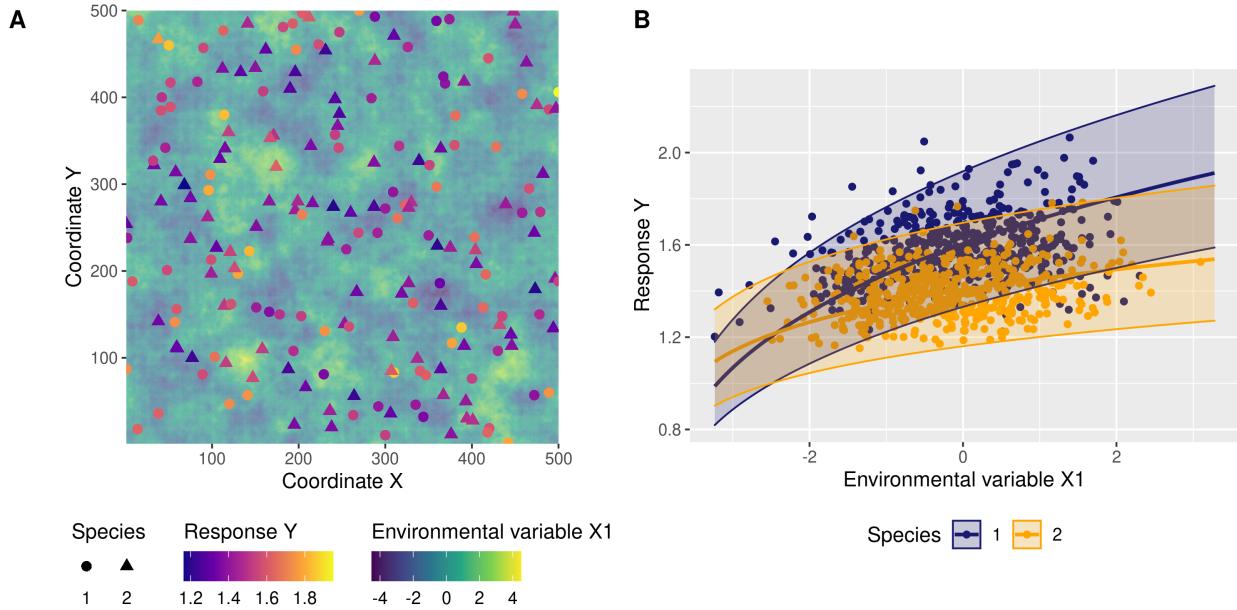


Figure 4.13: Response to an environmental variable and inferred intraspecific variability. (A) Positions of a sample of $I=600$ individuals from $J=2$ species in a landscape defined by a square grid of $C \times C$ cells ($C=500$). The background color indicates the value of the environmental variable X_1 (e.g., light) on each cell at date t . The response (e.g., growth) of each individual, which depends on the environment within each cell (Eq. I), is also indicated by a color scale. (B) Response as a function of the observed environmental variable X_1 for the two species. Points represent the data $\{Y_{ijt}, X_{1ijt}\}$. Thick lines represent the predictive posterior means for the two species. The envelopes delimited by two thin lines represent the 95% credible intervals of the predictive posterior marginalized over individuals (taking into account V_{bj}). The envelopes thus represent the intraspecific variability which is due to the $N-1$ unobserved environmental variables.

In Figure 4.12 and panel A of Figure 4.13, the solid and bold lines represent the mean rate of the response variable (e.g. growth) of Species 1 (blue) and Species 2 (orange) as computed with the parameters retrieved from the model with or without genetic variability respectively. The plain lines represent the 95% interval of the posteriors from the model without genetic variability and the dotted lines show the 95% confidence interval of the posteriors from the model with genetic variability. Figure 4.12 shows that genetic IV increases the overlap between the response of Species 1 and Species 2.

Figure 4.13 represents the virtual landscape of X_1 and the corresponding individual response and the plot of Y versus X_1 (without genetic variability) next to each other, helping to visualise this virtual experiment.

In the model without genetic IV, the unobserved variation in the environment resulted in an important inferred intraspecific variability. The performances of the two species can intersect, which means that the competitive hierarchy among the two species can be locally reversed depending on the micro-environmental variation, offering opportunities for the coexistence of the two species in a variable environment.

4.3 Spatial autocorrelation of individual response

We then analysed the spatial structure of the response variable at the individual scale. We hypothesised that as environmental variables are spatially autocorrelated, and the response variable is the result of these variables, then the response variable should also be spatially autocorrelated. To test this, we computed Moran's I test. It is computed with the Moran.I function of the ape R package (Paradis & Schliep 2019).

Table 8: Results of the Morans's I test for both species separately and together.

	Moran's I index	P-value
Species 1	6.7e-02	0e+00
Species 2	6.9e-02	0e+00
All individuals	4.3e-02	0e+00

Table 8 shows that the individual response variable is largely spatially autocorrelated. This is due to the spatial autocorrelation in the environmental variables. In this simple and controlled experiment, this seems natural. However, in a less controlled environment, detecting spatial autocorrelation in a response variable could be the sign of the spatial structure of the underlying environmental variables. Nonetheless, we acknowledge that the genetic structure of the population associated with limited dispersion for example could also lead to spatial autocorrelation in the response variable of the individuals, but this is likely to happen on a much larger geographical scale.

4.4 Similarity between conspecific individuals compared to heterospecific individuals and consequences for species coexistence.

Finally, we used semivariograms to visualise the spatial autocorrelation of the response variable and to test whether the individual growth was more similar within conspecifics than within heterospecifics. The semivariograms were computed and modelled with the variogram and the fit.variogram functions of the gstat R package (Gräler *et al.* (2016)) respectively. The variogram models were spherical.

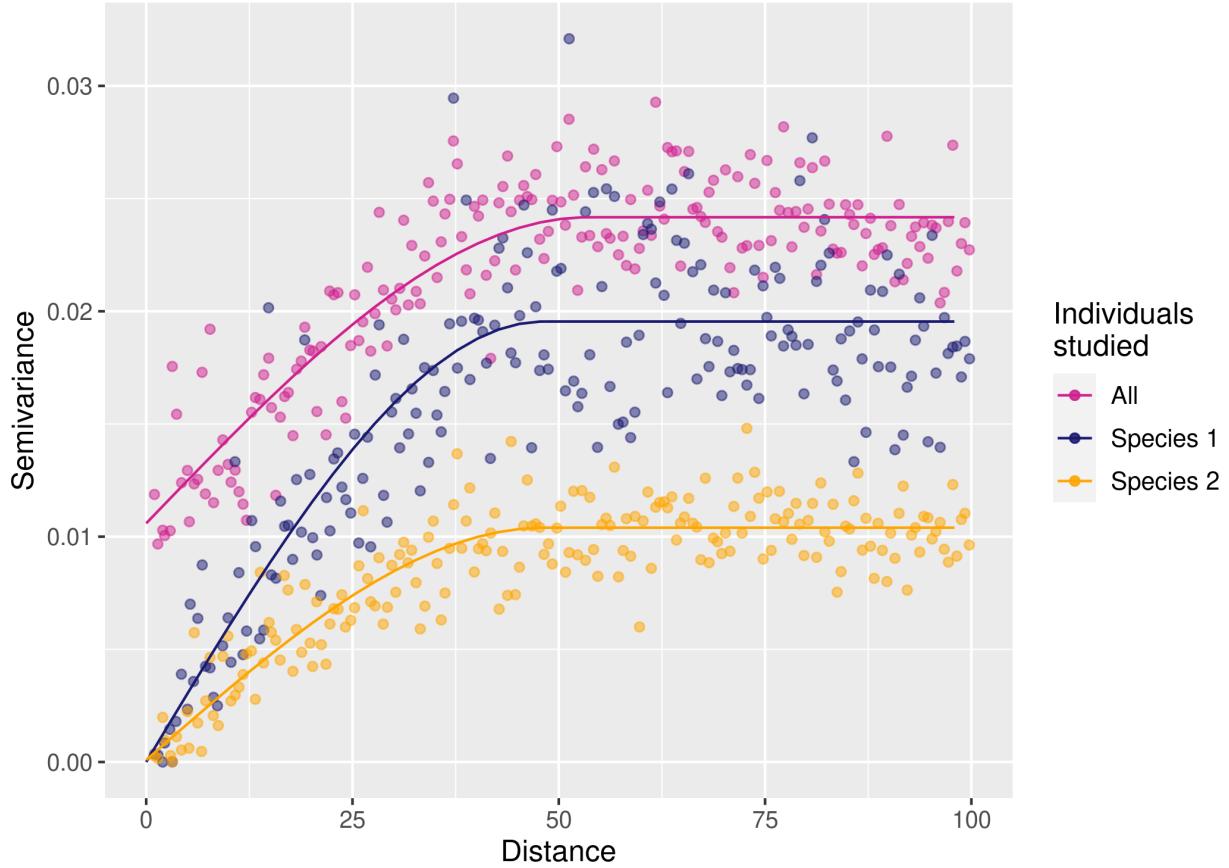


Figure 4.14: Spatial autocorrelation of response Y across individuals within and between species ($J=2$). This semivariogram represents the semivariance of the individual mean responses Y as a function of the distance between individuals. The increasing curves evidence spatial autocorrelation in Y (similar results using Moran's I test). The semivariance of all individuals taken together (purple curve) is higher than the semivariance of conspecifics for the two species (red and blue curves), which means that intraspecific variability is lower than interspecific variability.

As the semivariance between individuals of both species was higher than the semivariance between conspecifics (Figure 4.14), the individual response variable was more similar between conspecifics than heterospecifics. Considering this response variable as a proxy of the species performance, we argue that this would imply that intraspecific competition is greater than interspecific competition, which is the main condition for stable coexistence. Therefore, **stable** coexistence is possible even with high intraspecific variability, especially when this variability is not intrinsic but due to environmental heterogeneity that is structured in space.

5 Code used to generate Figure 1 in the main text: *Reinterpreting observed intraspecific variability (IV): from niche widening to niche projection into a high-dimensional environment.*

We obtain this figure by first generating a multivariate normal distribution. We then fix one of the two variables and obtain the mean and the standard deviation due to the second variable at this value.

```

library(ggplot2)
library(mvtnorm)
library(dplyr)
library(ggpubr)

#Environmental variables
minX = -5
minY = -10
maxX = 10
maxY = 40

data.grid <- expand.grid(x = seq(minX,maxX, length.out=200),
                         y = seq(minY, maxY, length.out=200))

#Species performance
mu1 = c(0, 0)
mu2 = c(3, 20)
Sigma1 = matrix (c(5, 0, 0, 5), 2, 2) #x and y axes are independent (0)
Sigma2 = matrix (c(10, 0, 0, 10), 2, 2)
p1 <- data.frame(X=data.grid$x,
                  Y=data.grid$y,
                  Perf=mvtnorm::dmvnorm(x=data.grid,
                                         mean = mu1,
                                         sigma = Sigma1))
p2 <- data.frame(X=data.grid$x,
                  Y=data.grid$y,
                  Perf=mvtnorm::dmvnorm(x=data.grid,
                                         mean = mu2,
                                         sigma = Sigma2))

x0 = -2

p1_E0 = data.frame(X=rep(x0, 200),
                     Y=seq(minY, maxY, length.out=200),
                     Perf=mvtnorm::dmvnorm(x=cbind(
                     rep(x0, 200),
                     seq(minY, maxY, length.out=200)),
                     mean = mu1,
                     sigma = Sigma1))

mean_p1_E0 = mean(p1_E0$Perf)

p2_E0 = data.frame(X=rep(x0, 200),
                     Y=seq(minY, maxY, length.out=200),
                     Perf=mvtnorm::dmvnorm(
                     x=cbind(
                     rep(x0, 200),
                     seq(minY, maxY, length.out=200)),
                     mean = mu2,
                     sigma = Sigma2))

mean_p2_E0 = mean(p2_E0$Perf)

```

```

x1 = 3

p1_E1 = data.frame(X=rep(x1, 200),
                     Y=seq(minY, maxY, length.out=200),
                     Perf=mvtnorm::dmvnorm(
                       x=cbind(
                         rep(x1, 200),
                         seq(minY, maxY, length.out=200)),
                       mean = mu1,
                       sigma = Sigma1))

mean_p1_E1 = mean(p1_E1$Perf)

p2_E1 = data.frame(X=rep(x1, 200),
                     Y=seq(minY, maxY, length.out=200),
                     Perf=mvtnorm::dmvnorm(
                       x=cbind(
                         rep(x1, 200),
                         seq(minY, maxY, length.out=200)),
                       mean = mu2,
                       sigma = Sigma2))

mean_p2_E1 = mean(p2_E1$Perf)

y0=0
y1=20

marginals_p1 <- p1%>%
  dplyr::group_by(X)%>%
  dplyr::summarise(mean=mean(Perf),
                    median=quantile(Perf, 0.5),
                    CI_2_5=quantile(Perf, 0.025),
                    CI_97_5=quantile(Perf, 0.975),
                    SD=sd(Perf))

marginals_p2 <- p2%>%
  dplyr::group_by(X)%>%
  dplyr::summarise(mean=mean(Perf),
                    median=quantile(Perf, 0.5),
                    CI_2_5=quantile(Perf, 0.025),
                    CI_97_5=quantile(Perf, 0.975),
                    SD=sd(Perf))

A <- ggplot2::ggplot(data=data.frame(Sp=as.factor(c("1","2"))),
                      Mean=c(mean_p1_E0, mean_p2_E0),
                      Sd=c(sd(p1_E0$Perf), sd(p2_E0$Perf))),
                      aes(x=Sp, y=Mean))+
  geom_point(aes(colour=Sp))+
  scale_color_manual(values=c("midnightblue", "orange"))+
  coord_cartesian(xlim=c(1, 2), ylim = c(-0.004, 0.009), clip="off")+
  theme(plot.background = ggplot2::element_blank(),
        panel.grid.major = ggplot2::element_blank(),
        panel.grid.minor = ggplot2::element_blank(),

```

```

panel.border = ggplot2::element_blank(),
text = ggplot2::element_text(size = 12),
legend.position="none",
axis.text=element_blank(),
axis.ticks=element_blank(),
plot.margin = unit(c(5.5, 5.5, 10, 10), "pt"),
axis.title.y = element_text(vjust=4),
axis.title.x = element_text(vjust=-3))+

ggplot2::labs(y="Performance at E0", x = "Species")+
ggplot2::geom_segment(ggplot2::aes(x = 0.4,
                                    y = mean_p1_E0,
                                    xend = 1,
                                    yend = mean_p1_E0),
                      linetype="dashed",
                      size=0.1)+

ggplot2::geom_segment(ggplot2::aes(x = 0.4,
                                    y = mean_p2_E0,
                                    xend = 2,
                                    yend = mean_p2_E0),
                      linetype="dashed",
                      size=0.1)+

ggplot2::geom_segment(ggplot2::aes(x = 1,
                                    y = mean_p1_E0,
                                    xend = 1,
                                    yend = -0.0045),
                      linetype="dashed",
                      size=0.1)+

ggplot2::geom_segment(ggplot2::aes(x = 2,
                                    y = mean_p2_E0,
                                    xend = 2,
                                    yend = -0.0045),
                      linetype="dashed",
                      size=0.1)+

annotate("text", label="p1", x = 0.34, y = mean_p1_E0, size=3)+
annotate("text", label="p2", x = 0.34, y = mean_p2_E0, size=3)+
annotate("text", label="Species 1", x = 1, y = -0.0053, size=3)+
annotate("text", label="Species 2", x = 2, y = -0.0053, size=3)

B <- ggplot2::ggplot(data=data.frame(Sp=as.factor(c("1","2")),
                                      Mean=c(mean_p1_E0, mean_p2_E0),
                                      Sd=c(sd(p1_E0$Perf), sd(p2_E0$Perf))),
                      ggplot2::aes(x=Sp, y=Mean))+

ggplot2::geom_point(aes(colour=Sp))+

ggplot2::geom_errorbar(ggplot2::aes(ymin=Mean-Sd,
                                       ymax=Mean+Sd,
                                       colour=Sp),
                       width=0.1,
                       size=0.5)+

ggplot2::scale_color_manual(values=c("midnightblue", "orange"))+
ggplot2::coord_cartesian(xlim=c(1, 2), ylim = c(-0.004, 0.009), clip="off")+
ggplot2::theme(plot.background = ggplot2::element_blank(),
               panel.grid.major = ggplot2::element_blank(),
               panel.grid.minor = ggplot2::element_blank(),

```

```

panel.border = ggplot2::element_blank(),
text = ggplot2::element_text(size = 12),
legend.position="none",
axis.text=element_blank(),
axis.ticks=element_blank(),
plot.margin = unit(c(5.5, 5.5, 10, 10), "pt"),
axis.title.x = element_text(vjust=-3),
axis.title.y = element_text(vjust=4))+

ggplot2::labs(y="Performance at E0", x = "Species")+
ggplot2::geom_segment(ggplot2::aes(x = 0.4,
                                    y = mean_p1_E0,
                                    xend = 1,
                                    yend = mean_p1_E0),
                      linetype="dashed",
                      size=0.1)+

ggplot2::geom_segment(ggplot2::aes(x = 0.4,
                                    y = mean_p2_E0,
                                    xend = 2,
                                    yend = mean_p2_E0),
                      linetype="dashed",
                      size=0.1)+

ggplot2::geom_segment(ggplot2::aes(x = 1,
                                    y = mean_p1_E0,
                                    xend = 1,
                                    yend = -0.0045),
                      linetype="dashed",
                      size=0.1)+

ggplot2::geom_segment(ggplot2::aes(x = 2,
                                    y = mean_p2_E0,
                                    xend = 2,
                                    yend = -0.0045),
                      linetype="dashed",
                      size=0.1)+

annotate("text", label="p1", x = 0.34, y = mean_p1_E0, size=3)+
annotate("text", label="p2", x = 0.34, y = mean_p2_E0, size=3)+
annotate("text", label="Species 1", x = 1, y = -0.0053, size=3)+
annotate("text", label="Species 2", x = 2, y = -0.0053, size=3)

C <- ggplot2::ggplot(data=marginals_p1, aes(x=X, y=mean))+
  ggplot2::geom_line(colour="midnightblue")+
  ggplot2::geom_line(data=marginals_p2, colour="orange")+
  ggplot2::coord_cartesian(xlim=c(minX, maxX), ylim = c(-0.004, 0.009), clip = "off")+
  ggplot2::geom_point(ggplot2::aes(x=x0, y=mean_p1_E0), colour="midnightblue") +
  ggplot2::geom_point(ggplot2::aes(x=x0, y=mean_p2_E0), colour="orange")+
  ggplot2::geom_errorbar(ggplot2::aes(x=x0,
                                       ymin=mean_p1_E0-sd(p1_E0$Perf),
                                       ymax=mean_p1_E0+sd(p1_E0$Perf)),
                         colour="midnightblue",
                         width=0.5,
                         size=0.2)+

  ggplot2::geom_errorbar(ggplot2::aes(x=x0,
                                       ymin=mean_p2_E0-sd(p2_E0$Perf),
                                       ymax=mean_p2_E0+sd(p2_E0$Perf)),
                         colour="orange",
                         width=0.5,
                         size=0.2)

```

```

            colour="orange",
            width=0.5,
            size=0.2)+

ggplot2::theme(plot.background = ggplot2::element_blank(),
               panel.grid.major = ggplot2::element_blank(),
               panel.grid.minor = ggplot2::element_blank(),
               panel.border = ggplot2::element_blank(),
               text = ggplot2::element_text(size = 12),
               legend.position="none",
               axis.text=element_blank(),
               axis.ticks=element_blank(),
               plot.margin = unit(c(5.5, 5.5, 10, 10), "pt"),
               axis.title.x = element_text(vjust=-3),
               axis.title.y = element_text(vjust=4))+

ggplot2::labs(x="x", y = "Performance")+
ggplot2::labs(x="x", y = "Performance")+
ggplot2::geom_segment(ggplot2::aes(x = x0,
                                    y = -0.0045,
                                    xend = x0,
                                    yend = max(mean_p1_E0,mean_p2_E0)),
                      linetype="dashed",
                      size=0.1)+

ggplot2::geom_segment(ggplot2::aes(x = x1,
                                    y = -0.0045,
                                    xend = x1,
                                    yend = max(mean_p1_E1,mean_p2_E1)),
                      linetype="dashed",
                      size=0.1)+

ggplot2::geom_segment(ggplot2::aes(x = x0-3.7,
                                    y = mean_p1_E0,
                                    xend = x0,
                                    yend = mean_p1_E0),
                      linetype="dashed",
                      size=0.1)+

ggplot2::geom_segment(ggplot2::aes(x = x0-3.7,
                                    y = mean_p2_E0,
                                    xend = x0,
                                    yend = mean_p2_E0),
                      linetype="dashed",
                      size=0.1)+

annotate("text", label="p1", x = x0-4.1, y = mean_p1_E0, size=3)+
annotate("text", label="p2", x = x0-4.1, y = mean_p2_E0, size=3)+
annotate("text", label="x0", x = x0, y = -0.0053, size=3)+
annotate("text", label="x1", x = x1, y = -0.0053, size=3)

```

```

D <- ggplot2::ggplot(data=p1, ggplot2::aes(x=X, y=Y, z=Perf)) +
  ggplot2::geom_contour(data=p1, colour="midnightblue")+
  ggplot2::geom_contour(data=p2, ggplot2::aes(x=X, y=Y, z=Perf), colour="orange")+
  ggplot2::coord_cartesian(xlim=c(minX, maxX), ylim=c(-5.1, 28), clip="off")+
  ggplot2::geom_segment(ggplot2::aes(x = -5.7,
                                    y = y0,
                                    xend = 10,

```

```

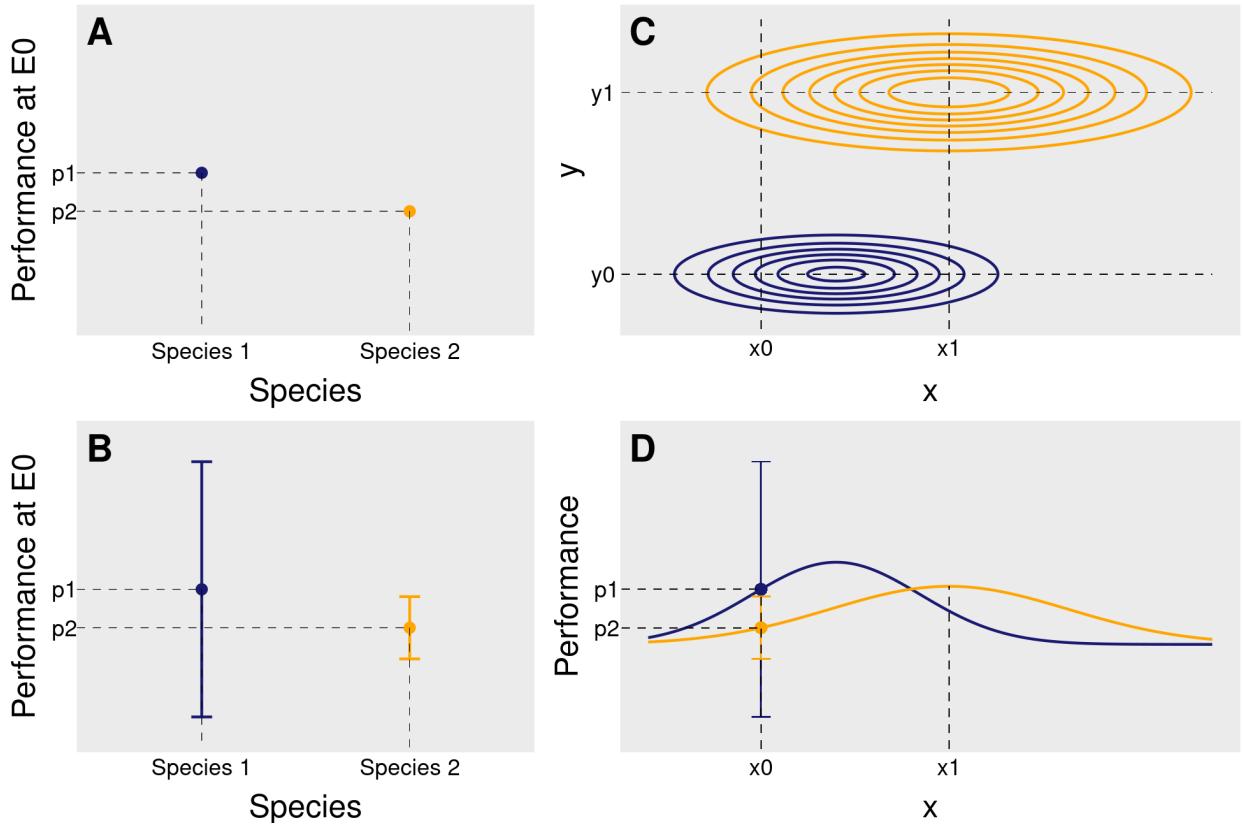
            yend = y0),
            linetype="dashed",
            size=0.1)+
ggplot2::geom_segment(ggplot2::aes(x = -5.7,
                                    y = y1,
                                    xend = 10,
                                    yend = y1),
                      linetype="dashed",
                      size=0.1)+
ggplot2::geom_segment(ggplot2::aes(x = x0,
                                    y = -6,
                                    xend = x0,
                                    yend = 28),
                      linetype="dashed",
                      size=0.1)+
ggplot2::geom_segment(ggplot2::aes(x = x1,
                                    y = -6,
                                    xend = x1,
                                    yend = 28),
                      linetype="dashed",
                      size=0.1)+
annotate("text", label="x0", x = x0, y = -8.1, size=3)+
annotate("text", label="x1", x = x1, y = -8.1, size=3)+
annotate("text", label="y0", x =-6.2, y = y0, size=3)+
annotate("text", label="y1", x =-6.2, y = y1, size=3)+
ggplot2::theme(plot.background = ggplot2::element_blank(),
               panel.grid.major = ggplot2::element_blank(),
               panel.grid.minor = ggplot2::element_blank(),
               panel.border = ggplot2::element_blank(),
               text = ggplot2::element_text(size = 12),
               legend.position="none",
               axis.text=element_blank(),
               axis.ticks=element_blank(),
               plot.margin = unit(c(5.5, 5.5, 10, 10), "pt"),
               axis.title.x = element_text(vjust=-3),
               axis.title.y = element_text(vjust=4))+
ggplot2::labs(x="x", y = "y")

```

```

Fig_1 <- ggpublisher::ggarrange(A, D, B, C,
                                labels=c("A", "C", "B", "D"),
                                widths = c(1,1.3),
                                hjust=c(-3, -3, -3, -3),
                                vjust=c(2, 2, 2, 2))
ggplot2::ggsave(file=here::here("outputs", "Fig_1.png"),
                Fig_1,
                width= 16.6,
                height=16.6/1.5,
                units='cm',
                dpi=300,
                bg="white")

```



6 Code implementation

The analyses were conducted using the R language (R Core Team 2021) in the Rstudio environment (RStudio Team 2021). The tables were made with the kableExtra package (Zhu 2021), the figures with the package ggplot2 (Wickham 2009), and the code uses other packages of the Tidyverse (Wickham *et al.* 2019) (tidyR (Wickham 2021), dplyr (Wickham *et al.* 2021), readr (Wickham & Hester 2020), lubridate (Golemund & Wickham 2011), magrittr (Bache & Wickham 2020), glue (Hester 2020)) and other R packages (here (Müller 2020), bayesplot (Gabry & Mahr 2021), ggpibr (Kassambara 2020), sp (Pebesma & Bivand 2005; Bivand *et al.* 2013), ggnewscale (Campitelli 2021), gstat Gräler *et al.* (2016)). Some functions were implemented in C++ thanks to the R packages Rcpp (Eddelbuettel & François 2011 ; Eddelbuettel 2013 ; Eddelbuettel & Balamuta 2018) and RcppArmadillo (Eddelbuettel & Sanderson 2014). The pdf and html documents were produced thanks to the R packages rmarkdown Xie *et al.* (2020), knitr Xie (2014) and bookdown (Xie 2017).

References

- Allaire, J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., *et al.* (2020). Rmarkdown: Dynamic documents for R.
- Bache, S.M. & Wickham, H. (2020). *Magrittr: A forward-pipe operator for r*.
- Bec, J.L., Courbaud, B., Moguédec, G.L. & Pélissier, R. (2015). Characterizing tropical tree species growth strategies: Learning from inter-individual variability and scale invariance. *PLoS ONE*, 10.
- Bivand, R., Pebesma, E.J. & Gómez-Rubio, V. (2013). *Applied spatial data analysis with R*. Use R! second edition. Springer, New York Heidelberg Dordrecht London.

- Bürkner, P.-C. (2017). Brms : An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80.
- Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, 10, 395.
- Campitelli, E. (2021). *Ggnewscale: Multiple fill and colour scales in 'ggplot2'*.
- Clark, J.S., Dietze, M., Chakraborty, S., Agarwal, P.K., Ibanez, I., LaDeau, S., et al. (2007). Resolving the biodiversity paradox. *Ecology Letters*, 10, 647–659.
- Condit, R., Pérez, R., Aguilar, S., Lao, S., Foster, R. & Hubbell, S. (2019). Complete data from the Barro Colorado 50-ha plot: 423617 trees, 35 years.
- Eddelbuettel, D. (2013). *Seamless R and C++ integration with Rcpp*. Use R! Springer, New York.
- Eddelbuettel, D. & Balamuta, J.J. (2018). Extending r with C++: A Brief Introduction to Rcpp. *The American Statistician*, 72, 28–36.
- Eddelbuettel, D. & François, R. (2011). Rcpp : Seamless R and C++ Integration. *Journal of Statistical Software*, 40.
- Eddelbuettel, D. & Sanderson, C. (2014). RcppArmadillo: Accelerating R with high-performance C++ linear algebra. *Computational Statistics & Data Analysis*, 71, 1054–1063.
- Gabry, J. & Mahr, T. (2021). Bayesplot: Plotting for bayesian models.
- Gourlet-Fleury, S., Guehl, J.-M., Laroussinie, O. & (Group), E. (Eds.). (2004). *Ecology and management of a neotropical rainforest: Lessons drawn from Paracou, a long-term experimental research site in French Guiana*. Elsevier, Paris.
- Gräler, B., Pebesma, E. & Heuvelink, G. (2016). Spatio-Temporal Interpolation using gstat. *The R Journal*, 8, 204.
- Grolemund, G. & Wickham, H. (2011). Dates and Times Made Easy with lubridate. *Journal of Statistical Software*, 40.
- Hester, J. (2020). Glue: Interpreted string literals.
- Héault, B. & 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.
- Hubbell, S.P. (1999). Light-Gap Disturbances, Recruitment Limitation, and Tree Diversity in a Neotropical Forest. *Science*, 283, 554–557.
- Kassambara, A. (2020). Ggpubr: 'ggplot2' based publication ready plots.
- Maire, G. le, Guillemot, J., Campoe, O.C., Stape, J.-L., Laclau, J.-P. & Nouvellon, Y. (2019). Light absorption, light use efficiency and productivity of 16 contrasted genotypes of several Eucalyptus species along a 6-year rotation in Brazil. *Forest Ecology and Management*, 449, 117443.
- Molino, J.-F. & Sabatier, D. (2001). Tree Diversity in Tropical Rain Forests: A Validation of the Intermediate Disturbance Hypothesis. *Science*, 294, 1702–1704.
- Müller, K. (2020). Here: A simpler way to find your files.
- Paradis, E. & Schliep, K. (2019). Ape 5.0: An environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics*, 35, 526–528.
- Pascal, J.-P. & Pelissier, R. (1996). Structure and Floristic Composition of a Tropical Evergreen Forest in South-West India. *Journal of Tropical Ecology*, 12, 191–214.
- Pebesma, E.J. (2004). Multivariable geostatistics in S: The gstat package. *Computers & Geosciences*, 30, 683–691.
- Pebesma, E.J. & Bivand, R.S. (2005). Classes and methods for spatial data in R. *R News*, 5, 9–13.

- Pélissier, R., Pascal, J.-P., Ayyappan, N., Ramesh, B.R., Aravajy, S. & Ramalingam, S.R. (2011). Tree demography in an undisturbed Dipterocarp permanent sample plot at Uppangala, Western Ghats of India: *Ecological Archives* E092-115. *Ecology*, 92, 1376–1376.
- R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- RStudio Team. (2021). *RStudio: Integrated development environment for r*. RStudio, PBC, Boston, MA.
- Wickham, H. (2009). *Ggplot2: Elegant graphics for data analysis*. Use R! Springer, New York.
- Wickham, H. (2021). *Tidyr: Tidy messy data*.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., et al. (2019). Welcome to the Tidyverse. *Journal of Open Source Software*, 4, 1686.
- Wickham, H., François, R., Henry, L. & Müller, K. (2021). *Dplyr: A grammar of data manipulation*.
- Wickham, H. & Hester, J. (2020). *Readr: Read rectangular text data*.
- Xie, Y. (2014). Knitr: A Comprehensive Tool for Reproducible Research in R. In: *Implementing reproducible research*, The R series (eds. Stodden, V., Leisch, F. & Peng, R.D.). CRC Press, Taylor & Francis Group, Boca Raton.
- Xie, Y. (2015). *Dynamic documents with R and Knitr*. Second edition. CRC Press/Taylor & Francis, Boca Raton.
- Xie, Y. (2017). *Bookdown: Authoring books and technical publications with R Markdown*. CRC Press, Boca Raton, FL.
- Xie, Y. (2021). *Knitr: A general-purpose package for dynamic report generation in r*.
- Xie, Y., Allaire, J.J. & Grolemund, G. (2019). *R Markdown: The definitive guide*. CRC Press, Taylor; Francis Group, Boca Raton.
- Xie, Y., Dervieux, C. & Riederer, E. (2020). *R markdown cookbook*. The R Series. 1st edn. Taylor; Francis, CRC Press, Boca Raton.
- Zhu, H. (2021). *kableExtra: Construct complex table with 'kable' and pipe syntax*.