Supporting Information

Danum gaps

Table of contents

1 Climber cutting **1**

2 Species taxonomy **3**

3 Survival analysis of logged forest seedlings in their first census **4**

4 Estimating missing values of basal diameter **5**

5 Prior choice **6**

6 Posterior predictive checks **9**

1 Climber cutting

Three of the Sabah Biodiversity Experiment plots included in this study received a complete climber cutting treatment as part of a different experiment (plots numbered 5, 11 and 14), where all lianas $\geq$ 10 cm in height were cut at the base in 2011 and 2014 [see @obrienPositiveEffectsLiana2019].

Including the climber cutting treatment as a fixed effect in our growth and survival models did not change our results.

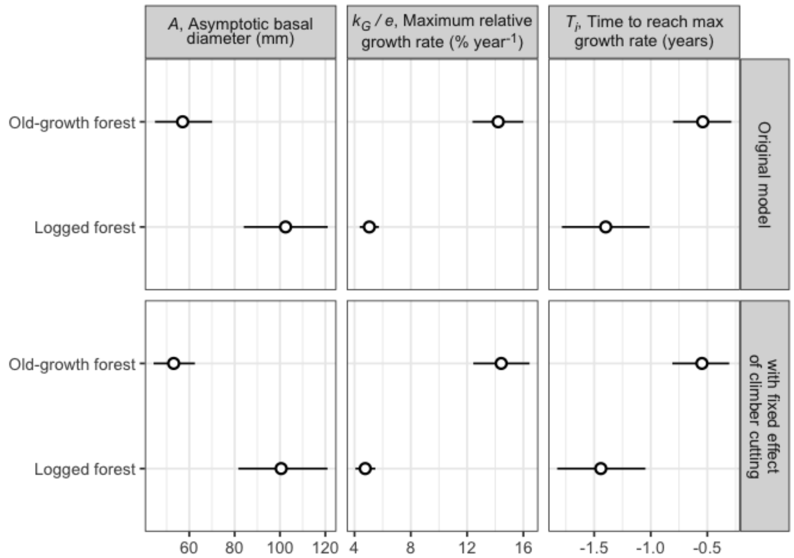

Figure 1: Comparison of parameter estimates for the growth model with and without a fixed effect of climber cutting. Facet columns indicate the parameters of the growth function with estimates for the old-growth and logged forest types on the y axis. Facet rows show results from the original model presented in the main text, and a model with the addition of a fixed effect for climber cutting treatment. Point shows the mean of the posterior distribution with lines dentoting the 95% credible interval.

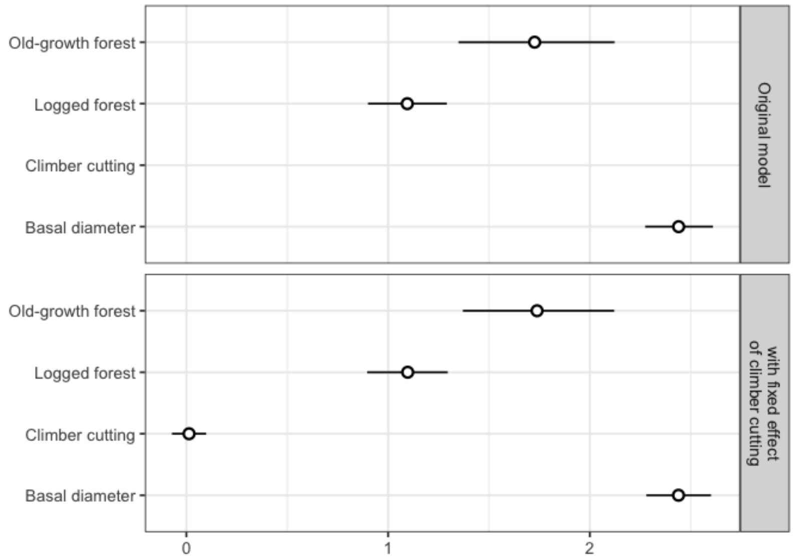

Figure 2: Comparison of effect size from survival models with and without a fixed effect of climber cutting. Facet rows show results from the original model presented in the main text, and a model with the addition of a fixed effect for climber cutting treatment. Point shows the mean of the posterior distribution with lines dentoting the 95% credible interval.

2 Species taxonomy

In the main text we use species names which are consistent with prior publications involving the Sabah Biodiversity Experiment. Here, we list the current accepted names for reference. Species names were matched to a static copy of The World Flora Online (WFO) v.2024.12 zenodo.14538251 using the function `WFO.match` from the R package {WorldFlora}. The function `WFO.one` was then used to find one unique matching name for each submitted name. using the argument `priority = "Accepted"`, it first limits candidates to accepted names, with a possible second step of eliminating accepted names that are synonyms.

Table 1: Taxonomy of species included in the study

Original name	New accepted name	WFO ID	Authorship
Dipterocarpus conformis	Dipterocarpus conformis	wfo-0000651311	Slooten
Dryobalanops lanceolata	Dryobalanops lanceolata	wfo-0000657521	Burck

Original name	New accepted name	WFO ID	Authorship
Hopea sangal	Hopea sangal	wfo-0000724645	Korth.
Parashorea malaanonan	Parashorea malaanonan	wfo-0000396696	(Blanco) Merr.
Parashorea tomentella	Parashorea tomentella	wfo-0000396691	(Symington) Meijer
Shorea argentifolia	Rubroshorea argentifolia	wfo-1000050969	(Symington) P.S.Ashton & J.Heck.
Shorea beccariana	Rubroshorea beccariana	wfo-1000050992	(Burck) P.S.Ashton & J.Heck.
Shorea faguetiana	Richetia faguetiana	wfo-1000051017	(F.Heim) P.S.Ashton & J.Heck.
Shorea gibbosa	Richetia gibbosa	wfo-1000051019	(Brandis) P.S.Ashton & J.Heck.
Shorea johorensis	Rubroshorea johorensis	wfo-1000050945	(Foxw.) P.S.Ashton & J.Heck.
Shorea leprosula	Rubroshorea leprosula	wfo-1000050975	(Miq.) P.S.Ashton & J.Heck.
Shorea macrophylla	Rubroshorea macrophylla	wfo-1000050993	(de Vriese) P.S.Ashton & J.Heck.
Shorea macroptera	Rubroshorea macroptera	wfo-1000050964	(Dyer) P.S.Ashton & J.Heck.
Shorea ovalis	Rubroshorea ovalis	wfo-1000051055	(Korth.) P.S.Ashton & J.Heck.
Shorea parvifolia	Rubroshorea ovata	wfo-1000050977	(Dyer ex Brandis) P.S.Ashton & J.Heck.

3 Survival analysis of logged forest seedlings in their first census

57% of all individuals were never recorded as alive (5,868 out of 10,272) i.e., they died in the period between planting and their first census. This occurs exclusively in the logged forest since unlike in the old-growth forest, diameter measurements were not taken at the time of planting.

We excluded these individuals from the main analysis since they lack data both on size and exact age, hence the seedlings cannot be directly compared. For the first cohort of seedlings planted into the logged forest, between 18 and 20 months elapsed between planting and censusing, in that time 70% of seedlings died (5,034 out of 7,222). For the second cohort (replacing dead seedlings from cohort one), between 7 and 28 months elapsed between planting and censusing, and 30% died (834 out of 2,746).

Nevertheless, this data might provide some insight on which species are more likely to survive transplanting from the nursery to the logged forest over these time periods, and we present a simple analysis here.

Survival of seedlings from cohort one and two were assessed in separate models with identical formulae. We fit the models in the `R` package `brms` with a Bernoulli response distribution, where zero indicated the seedling was dead in it's first census and one indicated it was alive. The predictor was a fixed effect of species. We use a regularising prior of `normal(0, 2)`, a common choice for Bernoulli models. Results are presented in Figure 3.

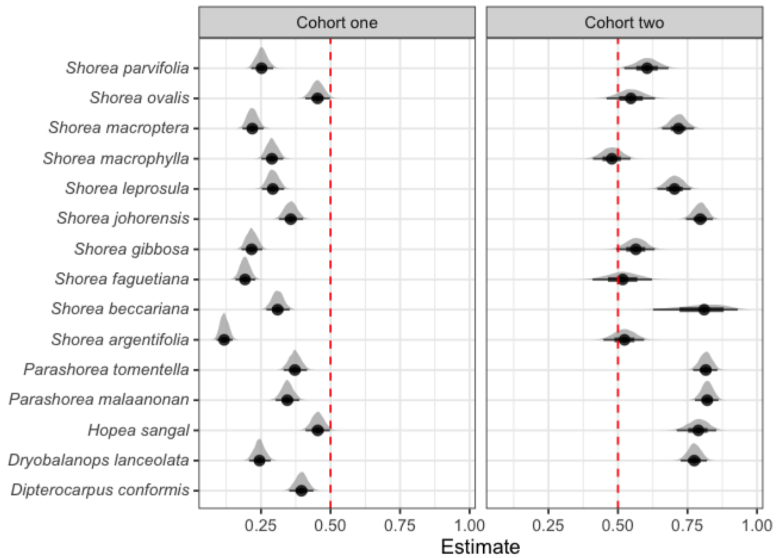

Figure 3: Posterior estimates of survival probability in the time period between planting and the first census of cohort one and two seedlings in the logged forest. The posterior median is represented by a point with the line showing the 66% and 95% credible interval. The dashed red line shows where the likelihood of survival is 50%. No *Dipterocarpus conformis* seedlings were planted in the second cohort due to insufficient avalibility.

4 Estimating missing values of basal diameter

There are 190 missing values of basal diameter in our data (0.73% of all records for living trees). For these missing values, we estimated basal diameter from diameter at breast height using a known allometric equation [@cushmanImprovingEstimatesBiomass2014];

$$d_{h2} = \frac{D_{h1}}{exp(b_1 h1 - h2)}$$

where d_{h2} is the estimated diameter (mm) at height h_2 (m), D_{h1} is the known diameter (mm) at height h_1 (m), and b_1 is a taper parameter.

We chose the value for our taper parameter $b_1 = 3.91$ by optimising on trees where we had diameter measurements both at the base and at breast height, to find the value which gave the lowest Root Mean Squared Error.

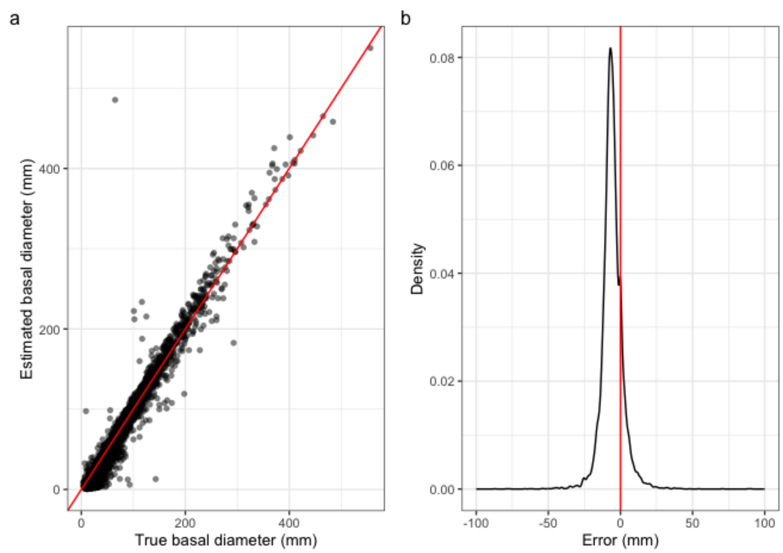

Figure 4: In (a), estimated values of basal diameter are plotted against their true values. The red line indicates where the estimate is equal to the true value. (b) Shows the distribution of the error around zero (red line).

5 Prior choice

We set weakly informative priors on each of the three parameters describing growth (equation 1) for the population-level effect of forest type. We left the predictor variables on their original scale, since the parameters have meaningful values. We used a log-normal prior for A, maximum size, with = 6, = 1. The world's tallest tropical tree, found at our primary forest study site in Danum Valley, has a DBH of 212 cm (Shorea faguetiana) (Shenkin et al., 2019). We expect a right-skewed distribution of adult size, with few very large individuals and a lower limit of 0 mm. The Student's-t distribution is generally recommended for weakly informative priors as it has heavier tails than a normal distribution and hence is more robust to outliers (Ghosh et al., 2017). We used Student's-t priors with degrees of freedom set to five for the remaining population-level priors. The prior for k, the growth rate coefficient, was set with = 0, = 0.5 and a lower bound of 0, since the basal diameter of dipterocarp seedlings in other locations grow at a rate of X-X (cite). Finally, the prior for delay, the time at inflection of the curve in years, was set with = 0, = 5. Since seedlings were germinated in the nursery no more than X months before transplanting to the forest, delay cannot be less than -X years and is very unlikely to be greater than 10 years. For the group-level effects of species and seedling ID we used the default priors provided by the R package {brms}: a Student-t prior with three degrees of freedom, = 0 and = 2.5.

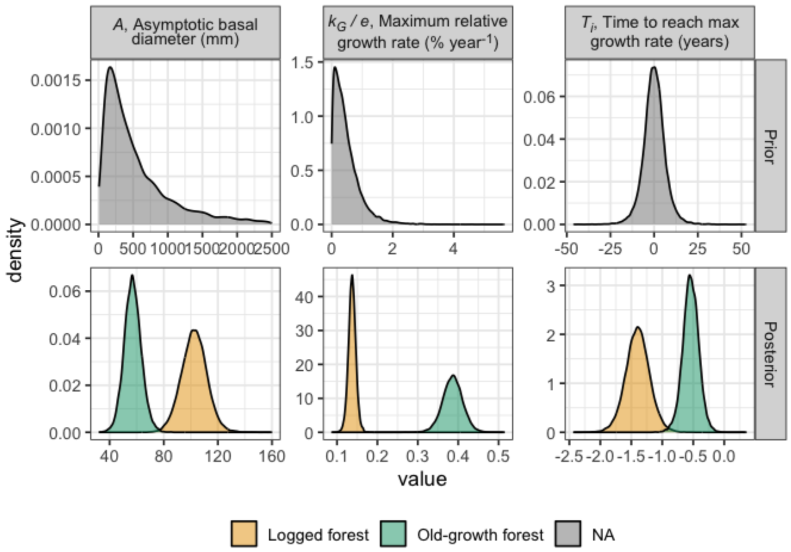

Figure 5: Prior and posterior distributions for the growth model. Facet columns indicate parameters estimated by the model. Facet rows indicate the prior and posterior distributions. Note that the same priors were set for both forest types, and that axes are not preserved across panels.

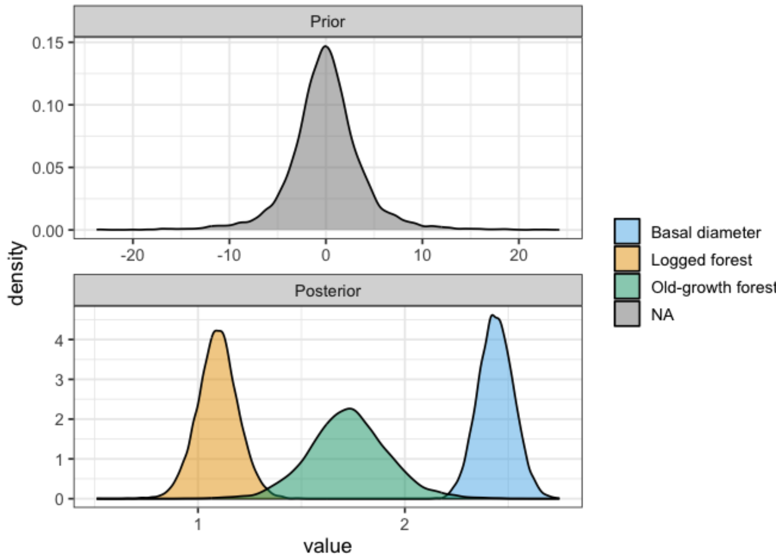

Figure 6: Prior and posterior distributions for the survival model. Prior and posterior distributions are displayed in separate panels and colours represent regression coefficients for the population-level, or "fixed" effects. Note that the same prior was set across all the population-level parameters, and that axes are not preserved across panels.

6 Posterior predictive checks

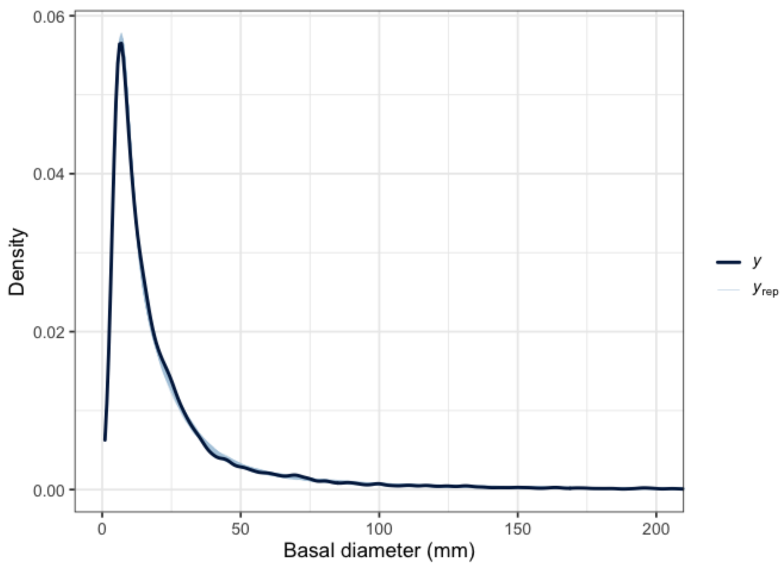

Figure 7: Posterior predictive check for the growth model, where the observed data is in dark blue (y) and 50 draws from the posterior distribution are in light blue ($yrep$).

Table 2: Posterior estimates of fixed effects from the growth model

Parameter	Forest type	Posterior mean	l-95% CI	u-95% CI	Rhat	Bulk effective sample size	Tail effective sample size
A	Logged	102.462	84.025	121.043	1.003	5261.080	8376.864
A	Old-growth	57.068	44.889	70.088	1.008	2961.325	10236.973
k_G/e	Logged	5.054	4.371	5.729	1.002	7524.692	9803.324
k_G/e	Old-growth	14.209	12.385	16.002	1.008	6262.388	13648.914
T_i	Logged	-1.398	-1.785	-1.011	1.001	9006.186	11873.259
T_i	Old-growth	-0.541	-0.803	-0.288	1.002	4856.719	9402.731

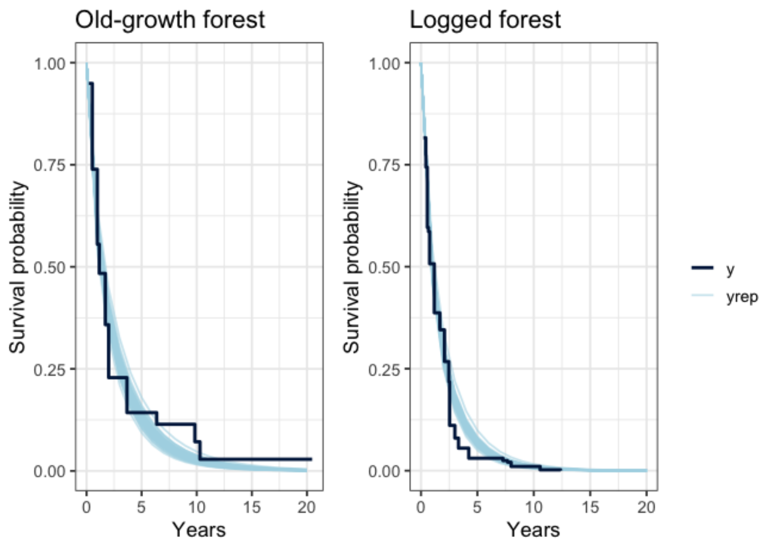

Figure 8: Posterior predictive check for the survival model, where the observed data is in dark blue (y) and 50 draws from the posterior distribution are in light blue ($yrep$). Conditional Kaplan-Meier curves, represented by y, were estimated for seedlings at representive values of basal diameter (the 0.25, 0.5 and 0.75 quartiles +/- the standard distribution multiplied by 0.01). Posterior preditions ($yrep$), were made for seedlings of these exact sizes (the 0.25, 0.5 and 0.75 quartiles).

Table 3: Posterior estimates of fixed effects from the survival model

Parameter	Posterior mean	l-95% CI	u-95% CI	Rhat	Bulk effective sample size	Tail effective sample size
Logged forest	1.054	0.842	1.255	1.001	2090.704	3589.735
Old-growth forest	1.709	1.307	2.106	1.000	2583.691	4432.720
Basal diameter	2.600	2.432	2.775	1.000	10644.645	7872.663