SUPPLEMENTARY MATERIAL

Supplementary Material S1: Variable selection algorithm

We consider the following model for the height H_i of a tree i in the forest plot p, for i=1,…,n. The plots are described by d variables $x_1, \ldots, x_d$. The indicators i_{α_j} and i_{β_j} indicates the presence (1) or absence (0) of the j-th variable in the model. The parameters to be inferred are θ_{α_j}, i_{α_j}, θ_{β_j}, i_{β_j}, and σ for j=1,…,d.

$$H_i = \frac{1}{1/\alpha_p + \beta_p/DBH_i} \times \varepsilon_i, \qquad \varepsilon_i \sim LN(0, \sigma^2)$$

$$\alpha_p = \exp\left(\theta_{\alpha_0} + \sum_j \theta_{\alpha_j} i_{\alpha_j} x_{j,p}\right), \qquad \beta_p = \exp\left(\theta_{\beta_0} + \sum_j \theta_{\beta_j} i_{\beta_j} x_{j,p}\right)$$

where LN is the log-normal distribution.

Estimations of α_p and β_p are available from the site-specific model. A rough maximum likelihood estimates (MLE) of θ_{α_j} and θ_{β_j} can be obtained with a linear model linking the site specific α_p and β_p to the variables x_j.

We did weighted regression because we wanted our model being more accurate for large trees. For the tree i, we used the weight w_i as a proxy of the tree biomass:

$$w_i = DBH_i^{\,2} \times H_i$$

$Hobs_i$ and $Hpred_i$ are respectively the observed height and the predicted height for the tree i:

$$Hpred_i = \frac{1}{1/\alpha_p + \beta_p/DBH_i}$$

The likelihood function is given by:

$$\mathcal{L}\left(\theta_\alpha, i_\alpha, \theta_\beta, i_\beta, \sigma^2 \middle| DBH, H\right) = \prod_{i=1}^{n} dlnorm(Hobs_i, Hpred_i, \sigma^2) w_i$$

Where dlnorm is the density function of the lognormal distribution. The model inference was done though the Bayesian paradigm; parameters were attributed standard low-informative priors.

For our model a Kuo-Mallick algorithm is defined by:

Repeat:

For each variable x_j in a random order:

Compute the MLE of $\theta_{\alpha_{MVE}}$ in the current model including the variable j.

Generate $\theta_{\alpha_j}{}^*$ from a normal proposition distribution centered on $\theta_{\alpha_{MVE}}$ with variance 0.1

Reject or accept the proposition with the Metropolis ratio

Compute the likelihood ratio:

$$r = \frac{\mathcal{L}\left(\theta_\alpha, i_\alpha(i_{\alpha_j} = 1), \theta_\beta, i_\beta,\ \sigma^2 \middle| DBH, H\right)}{\mathcal{L}\left(\theta_\alpha, i_\alpha(i_{\alpha_j} = 0), \theta_\beta, i_\beta,\ \sigma^2 \middle| DBH, H\right)}$$

Generate i_{α_j} from a Bernoulli distribution $\mathcal{B}\left(p = \frac{1}{1+r}\right)$

(Note: the intercepts θ_{α_0} and θ_{β_0} are always included in the model).

With the same process, update θ_β and i_β.

Compute

$$V = \sum_i w_i \left(\log\left(H_{obs_i}\right) - \log\left(H_{pred_i}\right)\right)^2$$

Generate σ^2 from an inverse-gamma distribution Inv$G(prior + \frac{nind}{2}, prior + \frac{V}{2})$.

We discard the beginning of the chains (burn-in) and use a thinning to reduce autocorrelation.

Supplementary Material 2: Forest plots description

Plot ID	log_area_ drain	slope (rad)	TRI_20	alt_hydro (m)	rainfall (mm)	dry season index (month)	BA (m^2/ha)	prop_stem _1 (%)	prop_stem _2 (%)	prop_stem _3 (%)	prop_stem _4 (%)	α	β	Fresh AGB (t/ha)
38	9,00	0,24	205,48	245,53	3040,05	2,53	10,05	88,71	10,75	0,54	0,00	19,21	2,43	71,29
11	9,00	0,18	104,44	115,12	3167,78	2,73	29,77	91,50	7,52	0,65	0,33	18,27	3,00	202,86
48	9,00	0,05	236,65	285,70	2661,64	2,45	19,20	70,05	25,60	3,38	0,97	38,78	2,55	225,50
2	9,00	0,01	6,89	11,95	2370,99	3,32	28,30	77,97	20,05	1,49	0,50	33,70	3,55	301,47
36	11,40	0,16	106,10	0,00	3039,15	2,53	28,85	71,43	23,21	4,29	1,07	44,45	1,50	328,77
46	10,10	0,14	21,05	0,00	2665,66	2,44	37,21	77,22	16,71	5,57	0,51	33,77	1,77	386,27
3	10,61	0,01	7,59	7,75	2383,91	3,30	38,52	78,81	15,25	4,13	1,81	29,11	3,24	413,68
35	9,00	0,07	62,30	47,97	3038,09	2,53	37,18	86,54	8,76	3,21	1,50	35,91	2,31	426,39
18	9,00	0,50	88,42	112,86	3167,48	2,73	48,09	79,88	18,34	1,48	0,30	27,61	3,52	459,64
A11	9,00	0,14	15,72	22,78	2302,08	3,36	34,47	55,68	29,55	10,91	3,86	44,14	1,98	464,88
B4	9,00	0,06	18,82	22,87	2378,35	2,99	34,14	56,58	29,61	12,06	1,75	50,02	1,79	465,22
37	15,34	0,01	87,80	0,00	3031,37	2,52	39,59	84,71	10,35	3,76	1,18	37,56	2,16	481,16
16	13,59	0,01	44,92	0,13	3168,00	2,72	44,74	72,79	21,48	3,58	2,15	42,36	1,57	485,12
21	9,00	0,03	21,27	28,50	2400,63	3,25	39,19	59,47	29,07	9,25	2,20	43,34	1,90	496,64
4	9,00	0,01	16,87	21,73	2376,31	3,33	40,49	68,91	22,12	7,69	1,28	37,14	2,29	498,31
17	9,69	0,11	23,53	6,64	3167,87	2,72	49,56	58,99	30,90	7,30	2,81	41,43	1,10	503,94
15	9,00	0,06	28,37	37,65	3167,19	2,70	44,64	82,86	13,37	2,45	1,32	46,54	1,85	521,55
P018	10,39	0,12	8,47	3,22	2545,89	3,70	38,16	58,10	28,29	9,07	4,54	53,44	1,43	533,19
9	9,00	0,11	8,95	11,27	2403,37	3,25	47,90	65,43	27,16	5,68	1,73	31,94	2,32	537,01
M1711	9,00	0,10	72,27	86,47	3137,00	2,62	39,97	60,61	26,26	8,28	4,85	49,53	1,47	540,11
5	9,00	0,02	12,64	18,44	2370,64	3,33	40,17	60,61	27,65	7,58	4,17	44,10	2,18	562,40
41	9,00	0,09	31,20	38,10	2665,35	2,44	43,32	64,02	27,13	5,49	3,35	49,76	1,85	568,14
10	10,10	0,06	41,11	4,02	3167,97	2,73	40,79	63,60	26,15	7,07	3,18	45,55	2,05	568,49
P006	11,08	0,05	10,04	3,26	2558,00	3,66	46,23	57,23	31,80	8,74	2,23	40,53	1,89	588,20
31	9,00	0,06	95,62	19,67	3032,44	2,52	45,28	63,57	26,12	5,84	4,47	38,63	2,43	593,74
NL11	10,39	0,29	64,65	27,60	3026,66	2,51	39,05	61,87	26,77	5,88	5,48	61,26	1,50	594,04
33	10,10	0,15	87,10	22,37	3037,09	2,53	40,34	62,24	25,73	5,39	6,64	60,39	1,41	601,11
12	9,00	0,11	63,77	76,47	3167,96	2,72	62,92	75,07	21,98	2,41	0,54	31,67	2,39	610,27
7	11,08	0,04	10,42	0,50	2390,11	3,29	52,05	65,95	24,70	6,95	2,40	37,90	2,34	626,77
LV1	9,69	0,10	20,68	18,42	3131,50	2,77	43,93	56,70	30,93	9,97	2,41	56,67	1,88	640,62
NH20	9,00	0,02	92,19	24,25	3032,96	2,52	45,10	54,62	30,52	9,64	5,22	56,82	1,45	673,33
T1	9,00	0,14	24,13	28,80	3093,72	2,86	43,66	50,49	32,28	10,44	6,80	52,33	1,94	706,67
34	9,00	0,20	68,92	79,82	3036,05	2,52	47,05	61,11	24,81	7,41	6,67	50,92	1,77	707,59
32	9,69	0,16	84,49	34,60	3033,48	2,52	48,85	55,56	28,97	9,52	5,95	48,90	1,74	720,90

Plot ID	log_area_ drain	slope (rad)	TRI_20	alt_hydro (m)	rainfall (mm)	dry season index (month)	BA (m^2/ha)	prop_stem _1 (%)	prop_stem _2 (%)	prop_stem _3 (%)	prop_stem _4 (%)	α	β	Fresh AGB (t/ha)
6	9,00	0,11	19,24	25,65	2393,45	3,28	49,76	56,02	30,29	7,47	6,22	41,83	1,85	732,48
42	9,00	0,05	83,95	19,54	2662,55	2,45	52,48	60,85	30,16	6,35	2,65	49,82	2,10	744,26
14	9,00	0,08	36,23	40,92	3167,50	2,72	52,09	59,32	27,46	7,46	5,76	43,33	2,18	756,10
8	10,39	0,04	12,37	0,00	2391,19	3,27	62,19	60,64	28,71	8,42	2,23	35,01	2,23	768,01
13	9,00	0,17	75,73	91,94	3167,65	2,73	65,86	63,85	27,06	5,84	3,25	39,73	2,01	808,75
45	15,96	0,03	54,92	0,00	2663,14	2,44	55,21	50,94	32,45	12,08	4,53	53,88	1,35	817,43
44	9,00	0,02	15,21	3,89	2664,67	2,44	58,30	63,25	24,10	7,23	5,42	46,05	1,58	833,63
43	9,00	0,18	72,41	42,34	2666,06	2,45	59,77	64,37	25,75	7,36	2,53	51,28	1,79	841,05

Plot ID: Identification of the forest plots. ID starting with letters are the 1-ha plots; ID with numbers only are the 0.5-ha Gentry plots

log_area_drain: Logarithm of the drained area

TRI_20: Terrain Ruggedness Index

alt_hydro: altitude above the closest stream of the hydraulic basin

BA: Basal Area

prop_stem_1: proportion of stems between 10cm and 20 cm DBH

prop_stem_2: proportion of stems between 20cm and 40 cm DBH

prop_stem_3: proportion of stems between 40cm and 60 cm DBH

prop_stem_4: proportion of stems above 60cm DBH

α: mean value of the alpha parameter

β: mean value of the beta parameter