

PUBLICATION INFORMATION

This is the author's version of a work that was accepted for publication in the Global Ecology and Biogeography journal. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in <https://doi.org/10.1111/geb.12803>.

Digital reproduction on this site is provided to CIFOR staff and other researchers who visit this site for research consultation and scholarly purposes. Further distribution and/or any further use of the works from this site is strictly forbidden without the permission of the Global Ecology and Biogeography Journal.

You may download, copy and distribute this manuscript for non-commercial purposes. Your license is limited by the following restrictions:

1. The integrity of the work and identification of the author, copyright owner and publisher must be preserved in any copy.
2. You must attribute this manuscript in the following format:

This is a pre-print version of an article by Bastin, J.F., Rutishauser, E., Kellner, J.R., Saatchi, S., Pélissier, R., Hérault, B., Slik, F., Bogaert, J., De Cannière, C., Marshall, A.R. et.al., 2018. **Pan-tropical prediction of forest structure from the largest trees**. *Global ecology and biogeography*, 27(11): 1366-1383. <https://doi.org/10.1111/geb.12803>



1 **Title**

2 Pan-tropical prediction of forest structure from the largest trees

3 **Authors**

4 Jean-François Bastin^{1,2,3,4}, Ervan Rutishauser^{4,5}, James R. Kellner^{6,7}, Sassan Saatchi⁸,
5 Raphael Pélissier⁹, Bruno Hérault^{10,11}, Ferry Slik¹², Jan Bogaert¹³, Charles De Cannière²,
6 Andrew R. Marshall^{14,15,16}, John Poulsen¹⁷, Patricia Alvarez-Loyayza¹⁸, Ana Andrade¹⁹,
7 Albert Angbonga-Basia²⁰, Alejandro Araujo-Murakami²¹, Luzmila Arroyo²², Narayanan
8 Ayyappan^{23,24}, Celso Paulo de Azevedo²⁵, Olaf Banki²⁶, Nicolas Barbier⁹, Jorcely G.
9 Barroso²⁶, Hans Beeckman²⁷, Robert Bitariho²⁸, Pascal Boeckx²⁹, Katrin Boehning-
10 Gaese^{30,31}, Hilandia Brandão³², Francis Q. Brearley³³, Mireille Breuer Ndoundou Hockemba³⁴,
11 Roel Brienen³⁵, Jose Luis C. Camargo¹⁹, Sto³⁶, Benoit Cassart^{37,38}, Jérôme Chave³⁹, Robin
12 Chazdon⁴⁰, Georges Chuyong⁴¹, David B. Clark⁴², Connie J. Clark¹⁷, Richard Condit⁴³,
13 Euridice N. Honorio Coronado⁴⁴, Priya Davidar²², Thalès de Haulleville^{13,27}, Laurent
14 Descroix⁴⁵, Jean-Louis Doucet¹³, Aurelie Dourdain⁴⁶, Vincent Droissart⁹, Thomas Duncan⁴⁷,
15 Javier Silva Espejo⁴⁸, Santiago Espinosa⁴⁹, Nina Farwig⁵⁰, Adeline Fayolle¹³, Ted R.
16 Feldpausch⁵¹, Antonio Ferraz⁸, Christine Fletcher³⁶, Krisna Gajapersad⁵², Jean-François
17 Gillet¹³, Iêda Leão do Amaral³², Christelle Gonmadje⁵³, James Grogan⁵⁴, David
18 Harris⁵⁵, Sebastian K. Herzog⁵⁶, Jürgen Homeier⁵⁷, Wannes Hubau²⁷, Stephen P. Hubbell^{58,59},
19 Koen Hufkens²⁹, Johanna Hurtado⁶⁰, Narcisse G. Kamdem⁶¹, Elizabeth Kearsley⁶², David
20 Kenfack⁶³, Michael Kessler⁶⁴, Nicolas Labrière^{10,65}, Yves Laumonier^{10,66}, Susan Laurance⁶⁷,
21 William F. Laurance⁶⁸, Simon L. Lewis³⁵, Moses B. Libalah⁶¹, Gauthier Ligot¹³, Jon Lloyd^{67,68},
22 Thomas E. Lovejoy⁶⁹, Yadvinder Malhi⁷⁰, Beatriz S. Marimon⁷¹, Ben Hur Marimon Junior⁷¹,
23 Emmanuel H. Martin⁷², Paulus Matius⁷³, Victoria Meyer⁸, Casimero Mendoza Bautista⁷⁴, Abel
24 Monteagudo-Mendoza⁷⁵, Arafat Mtui⁷⁶, David Neill⁷⁷, Germaine Alexander Parada
25 Gutierrez⁷⁸, Guido Pardo⁷⁹, Marc Parren⁸⁰, N. Parthasarathy²³, Oliver L. Phillips³⁵, Nigel C.A.
26 Pitman⁸⁰, Pierre Ploton⁹, Quentin Ponette³⁷, B.R. Ramesh²³, Jean-Claude
27 Razafimahaimodison⁸¹, Maxime Réjou-Méchain⁹, Samir Gonçalves Rolim⁸², Hugo Romero
28 Saltos⁸³, Luiz Marcelo Brum Rossi⁸², Wilson Roberto Spironello³², Francesco Rovero⁷⁶,

29 Philippe Saner⁸⁴, Denise Sasaki⁸⁵, Mark Schulze⁸⁶, Marcos Silveira⁸⁷, James Singh⁸⁸, Plinio
30 Sist^{10,89}, Bonaventure Sonke⁶¹, J.Daniel Soto⁹⁰, Cintia Rodrigues de Souza²⁴, Juliana
31 Stropp⁹¹, Martin J.P. Sullivan³⁵, Ben Swanepoel³⁴, Hans ter Steege^{25,92}, John
32 Terborgh^{93,94}, Nicolas Texier⁹⁵, T.Toma⁹⁶, Renato Valencia⁹⁷, Luis Valenzuela⁷⁵, Leandro
33 Valle Ferreira⁹⁸, Fernando Cornejo Valverde⁹⁹, Tinde R Van Andel²⁵, Rodolfo Vasque⁷⁷, Hans
34 Verbeeck⁶², Pandi Vivek²², Jason Vleminckx¹⁰⁰, Vincent A.Vos^{79,101}, Fabien H.Wagner¹⁰²,
35 Warsudi¹⁰³, Verginia Wortel¹⁰⁴, Roderick J. Zagt¹⁰⁵, Donatien Zebaze⁶¹

36 1. Institute of Integrative Biology, Department of Environmental Systems Science, ETH
37 Zürich, 8092 Zürich, Switzerland

38 2. Landscape Ecology and Plant Production System, Université libre de Bruxelles.
39 CP264-2, B-1050 Bruxelles, Belgium

40 3. Affiliated during analysis and writing at NASA, Jet Propulsion Laboratory, California
41 Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

42 4. Carboforexpert (carboforexpert.ch), 1248 Hermance, Switzerland

43 5. Smithsonian Tropical Research Institute, Box 0843-03092, Balboa, Ancon, Panama

44 6. Department of Ecology and Evolutionary Biology, Brown University, Providence, RI
45 02912, USA

46 7. Institute at Brown for Environment and Society, Brown University, Providence, RI
47 02912, USA

48 8. NASA, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove
49 Drive, Pasadena, CA 91109, USA

50 9. AMAP Lab, IRD, CIRAD, CNRS, INRA, Univ. Montpellier, Montpellier, France

51 10. Cirad, UR Forest & Societies, 34398 Montpellier Cedex 5, France

- 52 11. INPHB (Institut National Polytechnique Félix Houphouet Boigny), Yamoussoukro,
53 Ivory Coast
- 54 12. Faculty of Science, Universiti Brunei Darussalam, Gadong, Brunei Darussalam
- 55 13. Gembloux Agro-Bio Tech, Université de Liège, B-5030 Gembloux, Belgium
- 56 14. CIRCLE, Environment Department, Wentworth Way, University of York, Heslington,
57 York, YO10 5NG, UK
- 58 15. Tropical Forests and People Research Centre, University of the Sunshine Coast, QLD
59 4556, Australia
- 60 16. Flamingo Land Ltd., Kirby Misperton, YO17 6UX, UK
- 61 17. Nicholas School of the Environment, Duke University, PO Box 90328, Durham, NC
62 27708, USA
- 63 18. Field Museum of Natural History, Chicago, USA.
- 64 19. Biological Dynamics of Forest Fragment Project (BDFFP - INPA/STRI), Manaus -
65 Amazonas, Brazil
- 66 20. Institut Facultaire des Sciences Agronomiques de Yangambi. DRC
- 67 21. Museo de Historia Natural Noel Kempff Mercado, Santa Cruz, Bolivia
- 68 22. Department of Ecology and Environmental Sciences, Pondicherry University, Kalapet,
69 Pondicherry 605014, India
- 70 23. French Institute of Pondicherry (IFP), 11 Saint Louis Street, Pondicherry 605 001,
71 India
- 72 24. Embrapa Amazônia Ocidental, Brazil
- 73 25. Naturalis Biodiversity Centre, PO Box 9517, 2300 RA Leiden, The Netherlands
- 74 26. Universidade Federal do Acre, Campus Floresta, Cruzeiro do Sul, Acre, Brazil

- 75 27. Service of Wood Biology, Royal Museum for Central Africa, Tervuren, Belgium
- 76 28. Institute of Tropical Forest Conservation, Mbarara University of Science and
77 Technology, Uganda.
- 78 29. Isotope Bioscience Laboratory – ISOFYS, Ghent University, Belgium
- 79 30. Senckenberg Biodiversity and Climate Research Centre (BiK-F), Frankfurt am Main,
80 Germany
- 81 31. Dept of Biological Sciences, Goethe Universität, Frankfurt am Main, Germany
- 82 32. National Institute for Amazonian Research (INPA), Manaus, Amazonas, Brazil
- 83 33. School of Science and the Environment, Manchester Metropolitan University, Chester
84 Street, Manchester, M1 5GD, UK
- 85 34. Wildlife Conservation Society, New York, USA
- 86 35. School of Geography, University of Leeds, Leeds, UK
- 87 36. Malaysia Campus, Jalan Broga, Semenyih 43500, Selangor, Malaysia
- 88 37. UCL-ELI, Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve
89 BE-1348, Belgium
- 90 38. Ecole Régionale Post-universitaire d'Aménagement et de Gestion Intégrés des Forêts
91 et Territoires Tropicaux, Kinshasa, DRC
- 92 39. Laboratoire Evolution et Diversité biologique, CNRS & Université Paul Sabatier,
93 Toulouse 31062, France
- 94 40. Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs,
95 Connecticut 06268-3043, USA
- 96 41. Department of Botany and Plant Physiology, University of Buea, Cameroon
- 97 42. Department of Biology, University of Missouri-St Louis, Missouri, USA

- 98 43. Field Museum of Natural History and Morton Arboretum, Illinois, USA
- 99 44. Coronado, Inst. de Investigaciones de la Amazonia Peruana, Iquitos, Peru
- 100 45. ONF pôle R&D, Cayenne, France
- 101 46. Cirad, UMR EcoFoG (AgroParisTech, CNRS, Inra, Universite des Antilles, Universite
102 de la Guyane), Kourou, French Guiana
- 103 47. Department of Botany and Plant Pathology, Oregon State University, Corvallis, OR
104 97331, USA
- 105 48. Departamento de Biología, Universidad de La Serena, Casilla 554 La Serena, Chile
- 106 49. Universidad Autónoma de San Luis Potosí, San Luis Potosí, México
- 107 50. Department of Conservation Ecology, Philipps-Universität Marburg, Karl-von-Frisch-
108 Straße 8, 35032 Marburg, Germany
- 109 51. Geography, College of Life and Environmental Sciences, University of Exeter, Exeter,
110 EX4 4RJ, UK
- 111 52. Conservation International Suriname, Paramaribo, Suriname
- 112 53. Department of Plant Biology, Faculty of science, University of Yaounde I, BP 812
113 Yaoundé, Cameroon
- 114 54. Mount Holyoke College Botanic Garden, South Hadley, MA 01075, USA
- 115 55. Royal Botanic Garden Edinburgh, Edinburgh EH3 5LR, UK
- 116 56. Museo de Historia Natural Alcide d'Orbigny, Cochabamba, Bolivia
- 117 57. Plant Ecology, University of Goettingen, Untere Karspuele 2, 37073 Goettingen,
118 Germany
- 119 58. Department of Ecology and Evolutionary Biology, University of California, Los
120 Angeles, California 90095, USA

- 121 59. Smithsonian Tropical Research Institute, Apartado 0843-03092, Balboa, Republic of
122 Panama
- 123 60. Organization for Tropical Studies, Costa Rica
- 124 61. Plant Systematic and Ecology Laboratory, Higher Teacher's Training College,
125 University of Yaoundé I, P.O. Box 047, Yaoundé, Cameroon.
- 126 62. CAVElab – Computational and Applied Vegetation Ecology, Ghent University,
127 Belgium
- 128 63. CTFS-ForestGEO, Smithsonian Tropical Research Institute, MRC 166, NMNH, P.O.
129 Box 37012, Washington, DC 20013-7012, USA
- 130 64. Department of Systematic and Evolutionary Botany, University of Zurich,
131 Zollikerstrasse 107, Zurich 8008, Switzerland
- 132 65. AgroParisTech, Doctoral School ABIES, 19 Avenue du Maine, 75732 Paris Cedex 15,
133 France
- 134 66. Center for International Forestry Research, Jl. CIFOR, Situ Gede, Bogor Barat 16115,
135 Indonesia
- 136 67. Centre for Tropical Environmental and Sustainability Science, College of Science and
137 Engineering, James Cook University, Cairns, Queensland 4870, Australia.
- 138 68. Department of Life Sciences, Imperial College London, SL5 7PY, Ascot, UK
- 139 69. Department of Environmental Science and Policy, George Mason University, Fairfax,
140 VA, USA
- 141 70. Environmental Change Institute, School of Geography and the Environment,
142 University of Oxford, Oxford, UK
- 143 71. Universidade do Estado de Mato Grosso, Campus de Nova Xavantina, Nova
144 Xavantina, MT, Brazil

- 145 72. Udzungwa Ecological Monitoring Centre, Udzungwa Mountains National Park,
146 Tanzania, Sokoine University of Agriculture, Morogoro, Tanzania
- 147 73. Escuela de Ciencias Forestales, Unidad Académica del Trópico, Universidad Mayor
148 de San Simón, Sacta, Bolivia
- 149 74. Faculty of Forestry, Mulawarman University, Indonesia
- 150 75. Jardín Botánico de Missouri, Oxapampa, Pasco, Peru.
- 151 76. MUSE - Museo delle Scienze, Trento, Italy
- 152 77. Universidad Estatal Amazónica, Puyo, Pastaza, Ecuador
- 153 78. Museo de Historia Natural Noel Kempff Mercado, Santa Cruz, Bolivia
- 154 79. Universidad Autónoma del Beni, Riberalta, Bolivia
- 155 80. Science and Education, The Field Museum, 1400 South Lake Shore Drive, Chicago,
156 Illinois 60605–2496, USA
- 157 81. Centre ValBio, Ranomafana, Madagascar
- 158 82. Embrapa Florestas, Colombo/PR, Brazil
- 159 83. Yachay Tech University, School of Biological Sciences and Engineering. Urcuquí,
160 Ecuador
- 161 84. Department of Evolutionary Biology and Environmental Studies, University of Zurich,
162 CH-8057 Zurich, Switzerland
- 163 85. Fundação Ecológica Cristalino Alta Floresta, Brazil
- 164 86. HJ Andrews Experimental Forest, PO Box 300, Blue River, OR 97413, USA
- 165 87. Museu Universitário, Universidade Federal do Acre, Rio Branco 69910-900, Brazil
- 166 88. Guyana Forestry Commission, Georgetown, Guiana

- 167 89. Forests and Societies, Univ. Montpellier, CIRAD, Montpellier, France
- 168 90. Museo de Historia Natural Noel Kempff Mercado, Santa Cruz, Bolivia
- 169 91. Institute of Biological and Health Sciences, Federal University of Alagoas, Maceió,
170 Brazil
- 171 92. Systems Ecology, Free University, De Boelelaan 1087, Amsterdam, 1081 HV,
172 Netherlands.
- 173 93. Florida Museum of Natural History and Department of Biology, University of Florida -
174 Gainesville, Gainesville, FL 32611, USA
- 175 94. Department of Biology, James Cook University, Cairns, Australia
- 176 95. Laboratoire d'Evolution Biologique et Ecologie, Faculté des Sciences, Université libre
177 de Bruxelles, CP160/12, 1050 Bruxelles, Belgium
- 178 96. Forestry and Forest Products Research Institute, Matsunosato 1, Tsukuba 305-8687,
179 Japan
- 180 97. Escuela de Ciencias Biológicas, Pontificia Universidad Católica del Ecuador, Quito,
181 Ecuador
- 182 98. Coordenação de Botânica, Museu Paraense Emilio Goeldi, Belém, Brazil
- 183 99. Andes to Amazon Biodiversity Program, Madre de Dios, Peru
- 184 100. Department of Integrative Biology, University of California, Berkeley, 1005 Valley Life
185 Sciences Building 3140, Berkeley, CA 94720-3140, USA
- 186 101. Centro de Investigación y Promoción del Campesinado - Norte Amazónico, Riberalta,
187 Bolivia
- 188 102. Remote Sensing Division, National Institute for Space Research - INPE, São José
189 dos Campos 12227-010, SP, Brazil

- 190 103. The Center for Reforestation Studies in the Tropical Rain Forest (PUSREHUT),
191 Mulawarman University, Jln. Kihajar Dewantara Kampus Gunung Kelua, Samarinda 75123,
192 East Kalimantan, Indonesia
- 193 104. Center for Agricultural Research in Suriname (CELOS), Suriname
- 194 105. Tropenbos International, PO Box 232, Wageningen 6700 AE, The Netherlands

195 **Abstract**

196 **Aim.** Large tropical trees form the interface between ground and airborne observations,
197 offering a unique opportunity to capture forest properties remotely. However, despite rapid
198 development of metrics to characterize the forest canopy from remotely sensed data, a gap
199 remains between aerial and field inventories. To close this gap, we propose a new pan-tropical
200 model to predict plot-level forest structure properties and biomass from just the largest trees,
201 as a proxy for the whole plot inventory.

202 **Location.** Pan-tropical

203 **Method.** Using a dataset of 867 plots distributed among 118 sites across the tropics, we tested
204 the ability to predict quadratic mean diameter, basal area, Lorey's height and community wood
205 density from the i^{th} largest trees, i.e. testing the cumulative information gathered from these i
206 trees ranked by decreasing diameter. These tests served as a basis to select the optimal
207 number of the largest trees and further predict plot-level biomass from a single model.

208 **Result.** Focusing on readily available information captured by airborne remote sensing, we
209 show that measuring the largest trees in tropical forests enables unbiased predictions of plot
210 and site-level forest structure. The 20 largest trees per hectare predicted quadratic mean
211 diameter, basal area, Lorey's height and community wood density with 12%, 16%, 4% and 4%
212 of relative error. Building on this result, we developed a new model to predict plot-level AGB
213 from measurements of the 20 largest trees. This model allows an independent and unbiased
214 prediction of biomass with 17.7% of error compared to ground estimates. Most of the remaining
215 error is driven by differences in the proportion of total biomass held in medium size trees (50-
216 70 cm), which shows some continental dependency with American tropical forests presenting
217 the highest levels of total biomass share in these intermediate diameter classes.

218 **Conclusion.** Our approach provide new information on tropical forest structure and can be
219 employed to generate accurately field estimates of tropical forest carbon stocks to support the
220 calibration and validation of current and forthcoming space missions. It will reduce the cost of
221 programs to monitor, report, and verify forest resources, and will contribute to scientific
222 understanding of tropical forest ecosystems and response to climate change.

223 **Introduction**

224 The fundamental ecological function of large trees is well established for tropical forests. They
225 offer shelter to a multiple organisms (Remm & Löhmus, 2011; Lindenmayer *et al.*, 2012),
226 regulate forest dynamics, regeneration (Harms *et al.*, 2000; Rutishauser *et al.*, 2010) and total
227 biomass (Stegen *et al.*, 2011), and are important contributor to the global carbon cycle
228 (Meakem *et al.*, 2017). Being major components of the canopy, the largest trees also suffer
229 more than sub-canopy and understory trees from climate change, as they are directly exposed
230 to variations in solar radiation, wind strength, temperature seasonality and relative air humidity
231 (Laurance *et al.*, 2000; Nepstad *et al.*, 2007; Lindenmayer *et al.*, 2012; Thomas *et al.*, 2013;
232 Bennett *et al.*, 2015; Meakem *et al.*, 2017). Because they are visible from the sky, large trees
233 are ideal for monitoring forest responses to climate change via remote sensing (Bennett *et al.*,
234 2015; Asner *et al.*, 2017).

235 Large trees encompass a disproportionate fraction of total above-ground biomass (AGB) in
236 tropical forests (Chave *et al.*, 2001), with some variations in their relative contribution to the
237 total AGB among the tropical regions (Feldpausch *et al.*, 2012). In Central Africa, the largest
238 5% of trees, i.e. the 5% of trees with the largest diameter at 130 cm per area, store 50% of
239 forest aboveground biomass on average (Bastin *et al.*, 2015). Consequently, the density of
240 large trees largely explains variation in AGB at local (Clark & Clark, 1996), regional (Malhi *et*
241 *al.*, 2006; Saatchi *et al.*, 2007), and continental scales (Stegen *et al.*, 2011; Slik *et al.*, 2013).

242 Detailing the contribution of each single tree to the diameter structure, we showed previously
243 that plot-level AGB can be predicted from a few large trees (Bastin *et al.*, 2015), with the
244 measurement of the 20 largest trees per hectare being sufficient to estimate plot-level biomass
245 with less than 15% errors in reference to ground estimates. These findings opened the
246 possibility of measuring the largest trees to cost-effectively monitor forest biomass in Central
247 Africa, rather than conducting full inventories of all size classes. Similarly, they suggested that
248 remote sensing (RS) approaches should focus on the measurement of the largest trees,
249 instead of properties of the entire forest.

250 Several efforts are underway to close the gap between remote sensing and field surveys (e.g.
251 Jucker et al. 2016a, Coomes et al. 2017). However, field inventories still rely on exhaustive
252 data collection, while remote sensing surveys provide a limited alternative for the following
253 reasons. Existing RS approaches that provide predictions of biomass with less than 20% error
254 for 1 ha plot size are either specific to the relationship between forest type and image/scene
255 properties (Barbier *et al.*, 2011; Asner *et al.*, 2012; Barbier & Coutron, 2015), or require
256 ground measurement of all trees above or equal to 10 cm of D for calibration (Asner *et al.*,
257 2012; Asner & Mascaro, 2014). Using mean canopy height extracted from active sensors
258 (Mascaro *et al.*, 2011; Ho Tong Minh *et al.*, 2016), or canopy grain derived from optical images
259 (Proisy *et al.*, 2007; Ploton *et al.*, 2012, 2017; Bastin *et al.*, 2014), the biomass is predicted
260 from remote sensing with a typical error of only 10-20% compared to ground-based estimates,
261 but is limited to the extent of the scene used. An interesting development to alleviate this spatial
262 restriction lies in the ‘universal approach’, proposed by Asner et al. (2012) and further adapted
263 in Asner and Mascaro (2014), in which plot-level biomass is predicted by a linear combination
264 of ground-based and remotely-sensed metrics. The ‘universal approach’ relies upon canopy
265 height metrics derived from radar or LiDAR (top of canopy height, TCH), and basal area (BA,
266 i.e. the cross-sectional stem area) and community wood density (i.e. weighted by basal area,
267 WD_{BA}) derived from full field inventories. AGB is then predicted as follows (Asner *et al.*, 2012):
268 $AGB = aTCH^{b1}BA^{b2}WD_{BA}^{b3}(1)$
269 While generally performing better than approaches based solely on remote sensing of tree
270 height (Coomes *et al.*, 2017), this model largely relies on exhaustive ground measurements
271 (i.e. wood density and basal area of all trees above 10 cm of diameter at 130 cm, neither of
272 which is measured using any existing remotely sensed data).

273 Recent advances in remote sensing allow the identification of single trees in the canopy (Ferraz
274 *et al.*, 2016), estimation of adult mortality rates for canopy tree species (Kellner & Hubbell,
275 2017), description of the forest diameter structure (Stark *et al.*, 2015), depiction of crown and
276 gap shapes (Coomes *et al.*, 2017), and even identification of some functional traits of canopy
277 species (Asner *et al.*, 2017). Building upon this work, we test the capacity of metrics from the

278 largest trees that can be potentially derived using remote sensing to predict plot-level biomass
279 (i.e. the summed AGB of all live trees $D \geq 10$ cm in a plot). To this end, we tested the following
280 model:

$$281 \text{ AGB} = a(D_{\text{LT}}^2 H_{\text{LT}} W_{\text{LT}})^{b_1} \quad (2)$$

282 Where for the i^{th} largest trees, D_{LT} is the quadratic mean diameter, H_{LT} the mean height, and
283 W_{LT} the mean wood density averaged among the i^{th} largest trees.

284 Using a large database of forest inventories gathered across the Tropics (Figure 1), including
285 secondary and old growth forest plots, we test the ability of the largest trees to provide
286 information on various metrics estimated at 1-ha plot level, such as the mean quadratic
287 diameter, the basal area (BA), the Lorey's height (i.e. plot-average height weighted by BA), the
288 community wood density (i.e. plot-average wood density weighted by BA) and mean above-
289 ground live biomass (AGB) (supplementary figure 1). While previous work focused on
290 estimating biomass in Central African forests (Bastin *et al.*, 2015), the present study aims at
291 generalizing the potential of large trees in predicting these different plot metrics at continental
292 and pan-tropical scales. Taking advantage of a unique dataset gathered across the tropics (XX
293 ha, YYY plots), we also investigate major differences in forest structure across the three main
294 tropical regions, South America, Africa and South East Asia. We further discuss how this
295 approach can be used to guide innovative RS techniques and increase the frequency and
296 representativeness of ground data to support global calibration and validation of current and
297 planned space missions. These include the NASA Global Ecosystem Dynamics Investigation
298 (GEDI), NASA-ISRO Synthetic Aperture Radar (NISAR), and ESA P-band radar (BIOMASS).
299 This study is a step forward in bringing together remote sensing and field sampling techniques
300 for quantification of terrestrial C stocks in tropical forests.

301 **Material & Methods**

302 **Database**

303 For this study, we compiled standard forest inventories conducted in 867 1-hectare plots from
304 118 sites across the three tropical regions (Figure 1), including mature and secondary forests.
305 Each site comprises all the plots in a given geographical location, i.e. within a 10 km radius
306 and collected by a PI and its team. These consisted of 389 plots in America (69 sites), 302
307 plots in Africa (35 sites) and 176 plots in Asia (14 sites). Data were provided by Principal
308 Investigators (see supplementary Table 1), and through datasets available at ForestPlots
309 (<https://www.forestplots.net/>), TEAM (<http://www.teamnetwork.org/>) and CTFS
310 (<http://www.forestgeo.si.edu/>) networks.

311 We selected plots located between 23°N and 23°S, including tropical islands, with an area of
312 at least 1-ha to ensure stable intra-sample variance in basal area (Clark & Clark, 2000). Plots
313 in which at least 90% of the stems were identified to species, and in which all stems with the
314 diameter at 130 cm greater than or equal to 10 cm had been measured were included. Wood
315 density, here recorded as the wood dry mass divided by its green volume, was assigned to
316 each tree using the lowest available taxonomic level of botanical identifications (i.e. species or
317 genus) and the corresponding average wood density recorded in the Global Wood Density
318 Database (GWDD, Chave *et al.*, 2009; Zanne *et al.*, 2009). Botanical identification was
319 harmonized through the Taxonomic Names Resolution Service
320 (<http://tnrs.iplantcollaborative.org>), for both plot inventories and the GWDD. For trees not
321 identified to species or genus (~5%), we used plot-average wood density. We estimated
322 heights of all trees using Chave *et al.*'s (2014) pan-tropical diameter-height model which
323 accounts for heterogeneity in the D-H relationship using an environmental proxy:

$$324 \ln(H) = 0.893 - E + 0.760 \ln(D) - 0.0340 \ln(D)^2 \quad (3)$$

325 Where D is the diameter at 130cm and E is a measure of environmental stress (Chave *et al.*,
326 2014). For sites with tree height measurements (N=20), we developed local D-H models, using
327 a Michaelis-Menten function (Molto *et al.*, 2014). We used these local models to validate the

328 predicted Lorey's height (i.e. plot average height weighted by BA) from the largest trees, of
329 which height has been estimated with a generic H-D model (equation 3, Chave et al. 2014).

330 We estimated plot biomass as the sum of the biomass of live tree with diameter at 130 cm
331 superior or equal to 10 cm, using the following pan-tropical allometric model (Réjou-Méchain
332 *et al.*, 2017):

$$333 \text{AGB} = \exp(-2.024 - 0.896E + 0.920 \ln(\text{WD}) + 2.795 \ln(D) - 0.0461(\ln(D^2))) \quad (4)$$

334 **Plot-level metric estimation from the largest trees**

335 The relationship between each plot metric, namely basal area (BA), the quadratic mean
336 diameter (Dg), Lorey's height (H_{BA} ; the mean height weighted by the basal area) and the
337 community wood density (WD_{BA} ; the mean wood density weighted by the basal area), and
338 those derived from largest trees was determined using an iterative procedure following Bastin
339 et al. (2015). Trees were first ranked by decreasing diameter in each plot. An incremental
340 procedure (i.e. including a new tree at each step) was used to sum or average information of
341 the i largest trees for each plot metric. Specifically, each plot-level metric was predicted by the
342 respective metric derived from the i^{th} largest trees. For each increment, the ability (goodness
343 of fit) of the i largest trees to predict a given plot-metric was tested through a linear regression.
344 To avoid overfitting, a Leave-One-Out procedure was used to develop independent site-
345 specific models (N=118). Specifically, the model to be tested at a site was developed with data
346 from all other sites. Errors were then estimated as the relative root mean square error (rRMSE)
347 computed between observed and predicted values (X):

$$348 \text{rRMSE} = \bar{X} \sum \sqrt{\frac{(\text{X}_{\text{obs}} - \text{X}_{\text{pred}})^2}{n}} \quad (5)$$

349 The form of the regression model (i.e. linear, exponential) was selected to ensure a normal
350 distribution of the residuals.

351 To estimate plot basal area, we used a simple power-law constrained on the origin, as linear
352 model resulted in non-normal residuals. Plot-level basal area (BA) was related to the basal
353 area for the i largest trees (BA_i) using:

$$354 \text{BA} = b_1 \sum \text{BA}_i^{Y_1} \quad (6)$$

355 To estimate the quadratic mean diameter, Lorey's height and the wood density of the
356 community, we used simple linear models relating the plot-level metrics and the value of the
357 metrics for the i largest trees:

$$358 \quad D_g = a_2 + b_2 D_{gi} \quad (7)$$

$$359 \quad H_{BA} = a_3 + b_3 \overline{H_i} \quad (8)$$

$$360 \quad WD_{BA} = a_4 + b_4 \overline{WD_i} \quad (9)$$

361 Both Lorey's height (H_{BA}) and the average height ($\overline{H_i}$) of the i^{th} largest trees depend on the
362 same D-H allometry, which always contains uncertainty whether we use a local, a continental
363 or a pan-tropical model. To test the dependence of the prediction of H_{BA} from $\overline{H_i}$ on the
364 allometric model, we used measurement from Malebo in the Democratic Republic of the
365 Congo, where all heights were measured on the ground (see supplementary figure 2).

366 The quality of the predictions of plot-level metrics from the largest trees is quantified using the
367 relative root mean square error (rRMSE) between measured and predicted values, and
368 displayed along the cumulated number of largest trees (Figure 2). Model coefficients are
369 estimated for each metric derived from the largest trees (N_{LT}) and averaged across the 118
370 models (see supplementary table 2).

371 Mean rRMSE is plotted as a continuous variable, while its variation is presented as a
372 continuous area between 5th and the 95th percentiles of observed rRMSE (Figure 2).

373 **The optimal number of largest trees for plot-level biomass estimation**

374 The optimal number of largest trees N_{LT} was determined from the prediction of each plot-level
375 metric considered above, i.e. keeping a small number of trees while ensuring a low level of
376 error for each structural parameter. We then predicted plot-level biomass from the N_{LT} model
377 (equation 2). The final error was calculated by propagating the entire set of errors related to
378 equation 4 (Réjou-Méchain *et al.*, 2017) in the N_{LT} model (i.e. error associated to each allometric
379 model used). The model was then cross-validated across all plots ($N=867$).

380 **Investigating residuals: what the largest trees do not explain**

381 To understand the limits of predicting AGB through N_{LT} , we further investigated the relationship
382 between AGB residuals and key structural and environmental variables using linear modelling.
383 Forest structure was investigated through the total stem density (N), the quadratic mean
384 diameter (D_g), Lorey's height (H_{BA}) and community wood density (WB_{BA}). As environmental
385 data, we used the mean annual rainfall and the mean temperature computed over the last 10
386 years at each site using the Climate Research Unit data (New *et al.*, 1999, 2002), along with
387 rough information on soil types (Carré. *et al.*, 2010). Major soil types were computed from the
388 soil classification of the Harmonized World Soil Database into IPCC (intergovernmental panel
389 on climate change) soil classes. In addition, considering observed differences in forest
390 structure across tropical continents (Feldpausch *et al.*, 2011) and recent results on pan-tropical
391 floristic affinities (Slik *et al.*, 2015), we tested for an effect of continent (America, Africa and
392 Asia) on the AGB residuals.

393 The importance of each variable was evaluated by calculating the type II sum of squares that
394 measures the decrease in residual sum of squares due to an added variable once all the other
395 variables have been introduced into the model (Langsrud, 2003). Residuals were investigated
396 at both plot and site levels, the latter analyzed to test for any influence of the diameter structure,
397 which is usually unstable at the plot level due to the dominance of large trees on forest metrics
398 at small scales (Clark & Clark, 2000). Here we use a principal component analysis (PCA) to
399 summarize the information held in the diameter structure by ordinating the sites along the
400 abundance of trees in each diameter class (from 10 to +100 cm by 10 cm bins).

401

402 **Results**

403 **Plot-level metrics**

404 Plot metrics averaged at the site level (867 plots, 118 sites) present important variations within
405 and between continents. In our database, the quadratic mean diameter varies from 15 to 42
406 $\text{cm}^2\text{ha}^{-1}$, the basal area from 2 to 58 m^2ha^{-1} , Lorey's height from 11 to 33 m and the wood
407 density weighted by the basal area from 0.48 to 0.84 gcm^{-3} (Supplementary figure 1). Such
408 important differences between minimal and maximal values are observed because our
409 database cover sites with various forest types, from young forest colonizing savannas to old
410 growth forest. However, most of our sites are found in mature forests, as shown by relatively
411 high average and median value of each plot metric (average aboveground biomass = 302
412 Mgha^{-1} ; supplementary figure 1). In general, highest values of aboveground biomass are found
413 in Africa, driven by highest values of basal area and highest estimations of Lorey's height.
414 Highest values of wood density weighted by basal area are found in America.

415 **Plot-level estimation from the *i* largest trees**

416 Overall, plot metrics at 1 ha scale were well predicted by the largest trees, with qualitative
417 agreement among global and continental models (Figure 2).

418

419 When using the 20 largest trees to predict basal area (BA) and quadratic mean diameter (D_g),
420 the mean rRMSE was < 16% and 12%, respectively (Figs 3a and 3b). Lorey's height (H_{BA}) and
421 wood density weighted by basal area (WD_{BA}) were even better predicted (Figs 3c and 3d), with
422 mean rRMSE of 4% for the 20 largest trees. The prediction of Lorey's height from the largest
423 trees using local diameter-height model (supplementary Figure 2a) yielded results similar to
424 those obtained using equation 3 of Chave et al. (2014). More importantly, it also yielded similar
425 results to prediction of Lorey's height from the largest trees using plots where all the trees were
426 measured on the ground (supplementary figure 2b). This suggests that our conclusions are
427 robust to the uncertainty introduced by height-diameter allometric models.

428 **AGB prediction from the largest trees**

429 We selected “20” as the number of largest trees to predict plot metrics. The resulting model
430 predicting AGB (Mg ha⁻¹) based on the 20 largest trees is:

$$431 \text{ AGB} = 0.0735 \times (\text{Dg}_{20}\text{H}_{20}\text{WD}_{20})^{1.1332} \text{ (rRMSE=0.179; R}^2\text{=0.85; AIC= -260.18) (10)}$$

432 Because the exponent was close to 1, we also developed an alternative and more operationa
433 l model with the exponent constrained to 1, given by:

$$434 \text{ AGB} = 0.195 \times (\text{Dg}_{20}\text{H}_{20}\text{WD}_{20}) \text{ (rRMSE=0.177; R}^2\text{=0.85; AIC=-195) (11)}$$

435 Ground measurements of plot AGB were predicted by our N_{LT} model with the exponent
436 constrained to 1, with a total error of 17.9% (Figure 4), a value which encompass the error of
437 the N_{LT} model and the error related to the allometric model chosen. The Leave-One-Out cross-
438 validation procedure yielded similar results (rRMSE=0.19; R²=0.81), validating the use of the
439 model on independent sites.

440 **Determining the cause of residual variations**

441 The explanatory variables all together explain about 37% of the variance in AGB both at plot
442 and site levels when omitting the diameter structure, and about 63% at site level when included
443 (Fig. 5). In general, forest structure and particularly the stem density explained most of the
444 residuals (table 1; weights: 79% and 54% at plot- and site-level respectively). The stem density
445 was followed by a continental effect (weights: 18%, 28% and 1%, respectively for Africa, South
446 America and Asia) and by the effect of H_{BA} and WD_{BA} (respective weights: 1% and 0% at the
447 plot level, 0% and 11% at the site level, and 23% and 0% when accounting for the diameter
448 structure at the site level). Inclusion of the diameter structure provided the best explanation of
449 residuals, with 63% of variance explained, and a weight of 69% for the first axis of the PCA
450 (supplementary figure 3). This first axis of the PCA was related to the general abundance of
451 trees at a site, and in particular medium-sized trees (40-60cm). Among environmental
452 variables, only rainfall was significantly related to the residuals at the site level when the
453 diameter structure was considered (2%).

454 Discussion

455 The largest trees, convergences and divergences between continents

456 Sampling a few largest trees per hectare generally allows an unbiased prediction of four key
457 descriptors of forest structures across the Tropics. There is generally no improvement in
458 predicting basal area, quadratic mean diameter, Lorey's height (H_{BA}) or community wood
459 density beyond the first 10-to-20 largest trees (Figure 2, Figure 3a). In some cases, e.g. when
460 a forest plot presents an abundant number of large trees (Figure 5d), increasing the number
461 of trees sampled improves the model's accuracy. This is the case for BA for which rRMSE
462 continues to decrease up to 100 largest trees (Figure 2a). In contrast, Lorey's height
463 predictions are altered when a large number of trees are included (Figure 2c), i.e. when
464 smaller, often suppressed, trees draw the average down (Farrion *et al.*, 2016). This might
465 explain why the prediction of AGB does not mirror that of basal area (Figure 2b, Figure 3a),
466 and suggest that the number of largest trees shall be set independently to each predictor
467 considered. Interestingly, the evolution of relative error in AGB prediction as a function of the
468 number of largest trees considered does not follow the same path between continents. For
469 instance, the error of prediction saturates more quickly in Africa and Asia than Asia, where
470 high variations of residuals are observed. Investigation of residuals showed that the diameter
471 structure (Figure 5c, supplementary Figure 3b), and in particular the number of medium size
472 trees (Figure 5d), drives variability in AGB predictions. It is therefore not surprising to see that
473 in our dataset the site with higher levels of underestimations is the one with the highest number
474 of medium size trees, which is found in Asia in the Western Ghats of India.

475 The good performance of models based on the 20 largest trees in predicting Lorey's height
476 and community wood density at site level was not surprising. Both metrics were indeed
477 weighted by basal area, driven *de facto* by the largest trees. Their consistency across sites
478 and continents was not expected though. This suggests that the relationship between the 20
479 largest trees and descriptors of forest structures is stable across the tropics, and prove the
480 generality of our approach. Slight differences are however noticeable when comparing the
481 distribution of the pan-tropical model residuals across continents (Figure 6, supplementary

482 figure 4). In America, our pan-tropical model tend to slightly underestimate basal area (mean:
483 -5%) and overestimate Lorey's height (mean: +3%) (supplementary figure 4), suggesting
484 peculiar forest structures (i.e. higher tree height for a given diameter, and lower fractions of
485 large trees, supplementary figure 2). In Asia, and in particular in Africa, large (i.e. DBH > 50
486 cm) trees are more abundant and encompass a large fraction of plot biomass. The basal area
487 tends to be slightly overestimated in Africa, resulting in average to a 3% overestimation of AGB
488 (Figure 6a).

489 Interestingly, while a recent global phylogenetic classification of tropical forest groups
490 American with African forests vs. Asian forests (Slik *et al.*, 2018), our results tend more to
491 single out American forests. Although this deserves further investigations, it might reveal a lack
492 of close relationship between forest structure properties and phylogenetic similarity, which
493 echoes recent results on the absence of relationship between tropical forest diversity and
494 biomass (Sullivan *et al.*, 2017).

495 **Largest trees, a gateway to global monitoring of tropical forests**

496 Revealing the predictive capacity held by the largest trees, our results constitute a major step
497 forward to monitor forest structures and biomass stocks. The largest trees in tropical forests
498 can therefore be used to accurately predict and efficiently infer various ground-measured
499 properties (i.e. the quadratic mean diameter, the basal area, Lorey's height and community
500 wood density), while previous work has predicted only biomass "estimates" (e.g. Slik *et al.*,
501 2013; Bastin *et al.*, 2015). This approach allows us to (i) describe forest structure independently
502 of any biomass allometric model (ii) and cover local variations in D-H relationship, known to
503 vary locally (Feldpausch *et al.*, 2011; Kearsley *et al.*, 2013;). It is also (iii) relatively insensitive
504 to differences in floristic composition and community wood density (Poorter *et al.*, 2015).

505 Furthermore, the "largest trees" models were developed for each plot-level metric and for any
506 number of largest trees. Thus, they do not rely on any arbitrary threshold of tree diameter. Note
507 that the optimal number of largest trees to be measured (i.e. 20) was set for demonstration
508 and can vary depending on the needs and capacities of each country or project (see
509 supplementary table 2). In the same way, local models could integrate locally-developed

510 biomass models, when available. Consequently our approach (i) can be used in young or
511 regenerating un-managed forests with a low “largest tree” diameter threshold and (ii) is
512 compatible with recent remote sensing approaches able to single out canopy trees and
513 describe their crown and height metrics (Ferraz *et al.*, 2016; Coomes *et al.*, 2017).

514 **Aboveground biomass model from the largest trees, a multiple opportunity**

515 Globally, the N_{LT} model for the 20 largest trees allows plot biomass to be predicted with 17.9%
516 error. This result is a pan-tropical validation of results obtained in Central Africa (Bastin *et al.*,
517 2015). It opens new perspectives towards cost-effective methods to monitor forest structures
518 and carbon stocks through largest trees metrics, i.e. metrics of objects directly intercepted by
519 remote-sensing products.

520 Developing countries willing to implement a Reduction of Emissions from Deforestation and
521 Forest Degradation (REDD+), shall also report on their carbon emissions (CE) and develop a
522 national CE reference level (IPCC, 2006; Maniatis & Mollicone, 2010). However, most tropical
523 countries lack capacities to assume multiple, exhaustive and costly forest carbon assessment
524 (Romijn *et al.*, 2012). By measuring only a few large trees per hectare, our results show that
525 it is possible to obtain unbiased estimates of aboveground C stocks in a time and cost-efficient
526 manner. Assuming that 400 to 600 trees $D > 10$ cm are measured in a typical 1-ha sample
527 plot, monitoring only 20 trees is a significant improvement. Although finding the 20 largest trees
528 in a plot of several hundred individuals requires evaluating more than 20 trees, in practice, a
529 conservative diameter threshold could be defined to ensure that the 20 largest trees are
530 sampled. An alternative approach could also be found in the development of relascope-based
531 approach adapted to detection of the largest trees in tropical forests. Using such approach
532 would facilitate rapid field sampling in extensive areas to produce large scale AGB estimates.
533 Those could fulfil the needs in calibration and validation of current and forthcoming space
534 missions focused on aboveground biomass.

535 Our findings also points towards the potential effectiveness of using remote sensing
536 techniques to characterize canopy trees. Here, remote sensing data could be used for direct
537 measurement (e.g. tree level metrics such as height, crown width, crown height) of the largest

538 trees instead of indirect development of complex metrics (e.g. mean canopy height, texture)
539 used to extrapolate forest properties. While some further refinements are needed, most of the
540 tools required to develop “largest trees” models are readily available. In particular, Ferraz et
541 al. (2016) developed an automated procedure to locate canopy trees based on airborne LiDAR
542 data, to measure their height and crown area. Crown area could further be linked to basal area,
543 as the logarithm of crown area is consistently correlated with a slope of 1.2-1.3 to the logarithm
544 of tree diameter across the tropics (Blanchard *et al.*, 2016). Regarding wood density,
545 hyperspectral signature offers a promising way to assess functional traits remotely (e.g. Asner
546 *et al.*, 2017) which could potentially be tested to detect variability in wood properties.
547 Alternative approaches could focus on the development of plot-level AGB prediction by
548 replacing the basal area of the largest trees with their crown metrics. While the measurement
549 of crown areas have yet to be generalized when inventorying plots, several biomass allometric
550 models already partition trunk and crown mass (Jucker *et al.*, 2016; Ploton *et al.*, 2016;
551 Coomes *et al.*, 2017).

552 The main limitation of our approach lies in the understory and sub-canopy trees. We show that
553 most of the remaining variance is explained by variations in diameter structures, and in
554 particular among the total stem density. Interestingly, stem density was generally identified as
555 a poor predictor of plot biomass in tropical forests (Slik *et al.*, 2010; Lewis *et al.*, 2013).
556 However, our results show that stem density explains most of the remaining variance (Table
557 S1). This suggests that, in addition to trying to understand large-scale variations in large trees
558 and other plot metrics, which can be directly quantified from remote sensing, we should also
559 put more effort into understanding variation in smaller trees, which mainly drives total stem
560 density and the total floristic diversity. Smaller trees are also essential to characterize forest
561 dynamics and understand changes in carbon stocks. Several options are nonetheless possible
562 from remote sensing, considering the variation in lidar point density below the canopy layer
563 (D’Oliveira *et al.*, 2012), the distribution of leaf area density (Stark *et al.*, 2012, 2015; Tang &
564 Dubayah, 2017) or the use of multitemporal lidar data to get information on forest gap

565 generation dynamics and consequently on forest diameter structure (Kellner *et al.*, 2009;
566 Farrior *et al.*, 2016).

567 **Large trees in degraded forests**

568 If large trees are a key feature of unmanaged forests, they are conspicuously absent from
569 managed or degraded forests. Indeed, large trees are targeted by selective or illegal logging,
570 and are the first to disappear or to suffer from incidental damages when tropical forests are
571 exploited for timber (Sist *et al.*, 2014). The loss of largest trees drastically changes forest
572 structures and diameter distributions, and their loss is likely to counteract the consistency in
573 forest structures observed through this study. Understanding how, or whether, managed
574 forests deviate from our model predictions could help characterize forest degradation, which
575 accounts for a large fraction of carbon loss worldwide (Baccini *et al.*, 2017), acknowledging
576 that rapid post-disturbance biomass recovery (Rutishauser *et al.*, 2015) will remain hard to
577 capture.

578 **Conclusion – towards improved estimates of tropical forest biomass**

579 The acquisition, accessibility and processing capabilities of very high spatial, spectral and
580 temporal resolution remote sensing data has increases exponentially in recent years (Bastin
581 *et al.*, 2017). However, to develop accurate global maps, we will have to obtain a greater
582 number of field plots and develop new ways to use remote sensing data. Our results provide
583 a step forward for both by (i) decreasing drastically the number of individual tree measurements
584 required to get an accurate, yet less precise, estimate of plot biomass and (ii) opening the door
585 to direct measurement of plot metrics measured from remote sensing to estimate plot biomass.
586 As highlighted by Clark and Kellner (2012), new biomass allometric models relating plot-level
587 biomass measured from destructive sampling and plot-level metric measured from remote-
588 sensing products should be developed, as an alternative to current tree-level allometric
589 models. Such an effort will lead largely to lower operational costs and uncertainties surrounding
590 terrestrial C estimates, and consequently, will help developing countries in the development of
591 national forest inventories and aid the scientific community in better understanding the effect
592 of climate change on forest ecosystems.

593 **Acknowledgments**

594 J.-F.B. was supported for data collection by the FRIA (FNRS), ERAIFT (WBI), WWF and by
595 the CoForTips project (ANR-12-EBID-0002); T.d.H. was supported by the COBIMFO project
596 (Congo Basin integrated monitoring for forest carbon mitigation and biodiversity) funded by the
597 Belgian Science Policy Office (Belspo); C.H.G was supported by the “Sud Expert Plantes”
598 project of French Foreign Affairs, CIRAD and SCAC. Part of data in this paper was provided
599 by the TEAM Network, the partnership between Conservation International, The Missouri
600 Botanical Garden, The Smithsonian Institution and The Wildlife Conservation Society, and
601 these institutions and the Gordon and Betty Moore Foundation. This is [number to be
602 completed] publication of the technical series of the Biological Dynamics of Forest Fragment
603 Project (INPA/STRI). We acknowledge data contributions from the TEAM network not listed as
604 co-authors (upon voluntary basis). We thank Jean-Phillipe Puyravaud, Estação Científica
605 Ferreira Penna (MPEG) and the Andrew Mellon Foundation and National Science Foundation
606 (DEB 0742830). And finally, we thank Helen Muller-Landau for her careful revision and
607 comments of the manuscript.

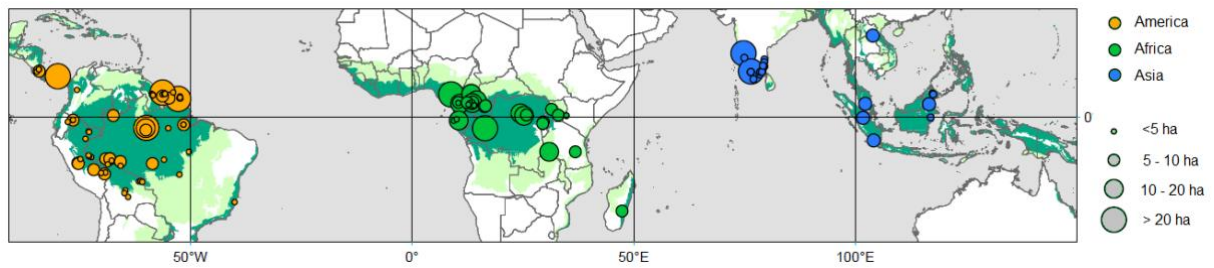
608 **Contributions**

609 J.F.Bastin and E.Rutishauser conceptualized the study, gathered the data, performed the
610 analysis and wrote the manuscript. All the co-authors contributed by sharing data and
611 reviewing the main text. A.R.Marshall, J.Poulsen and J.Kellner revised the English.

612 **Conflict of interest**

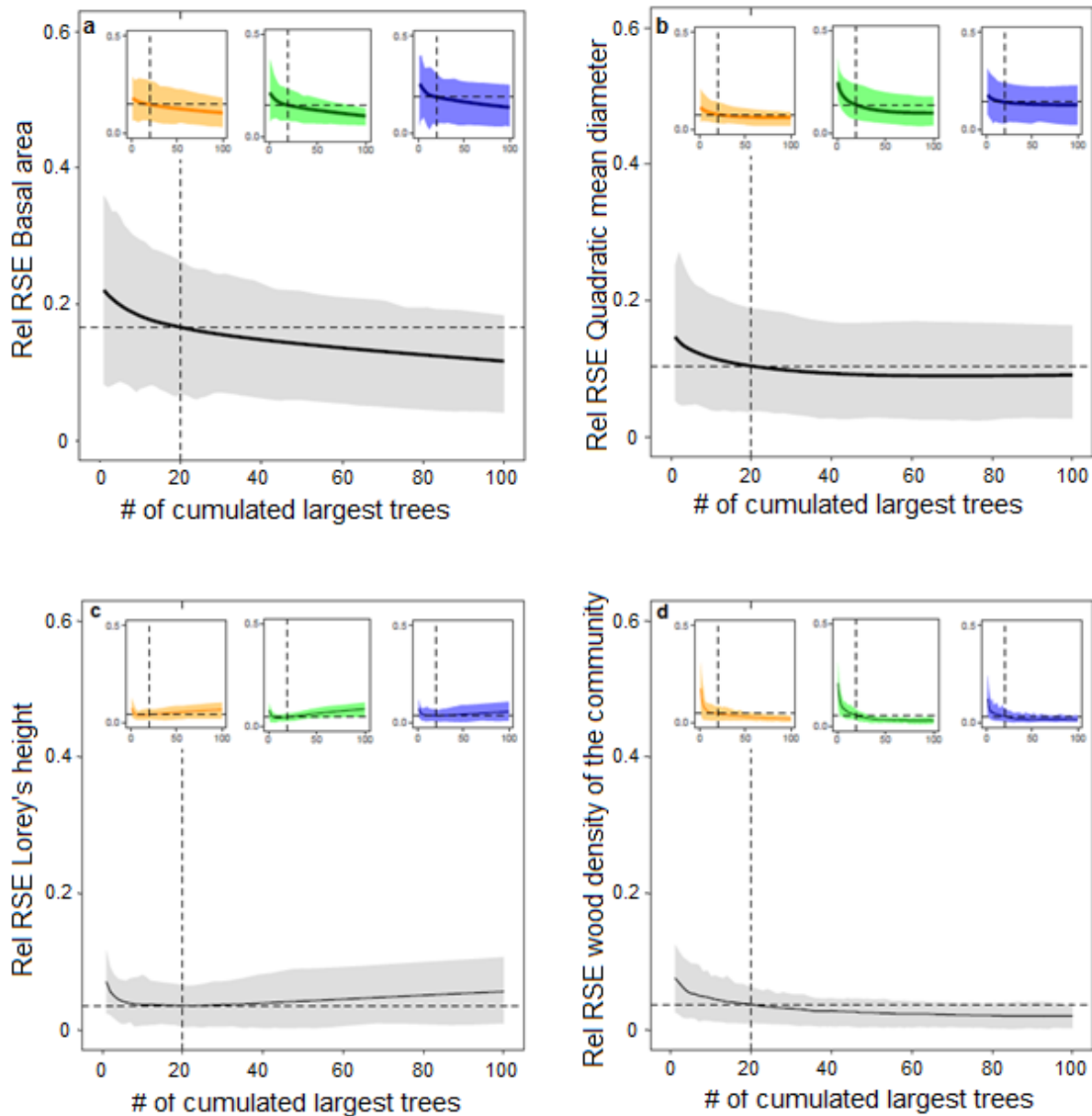
613 The authors declare there is no conflict of interest associated to this study.

614 **Figures**



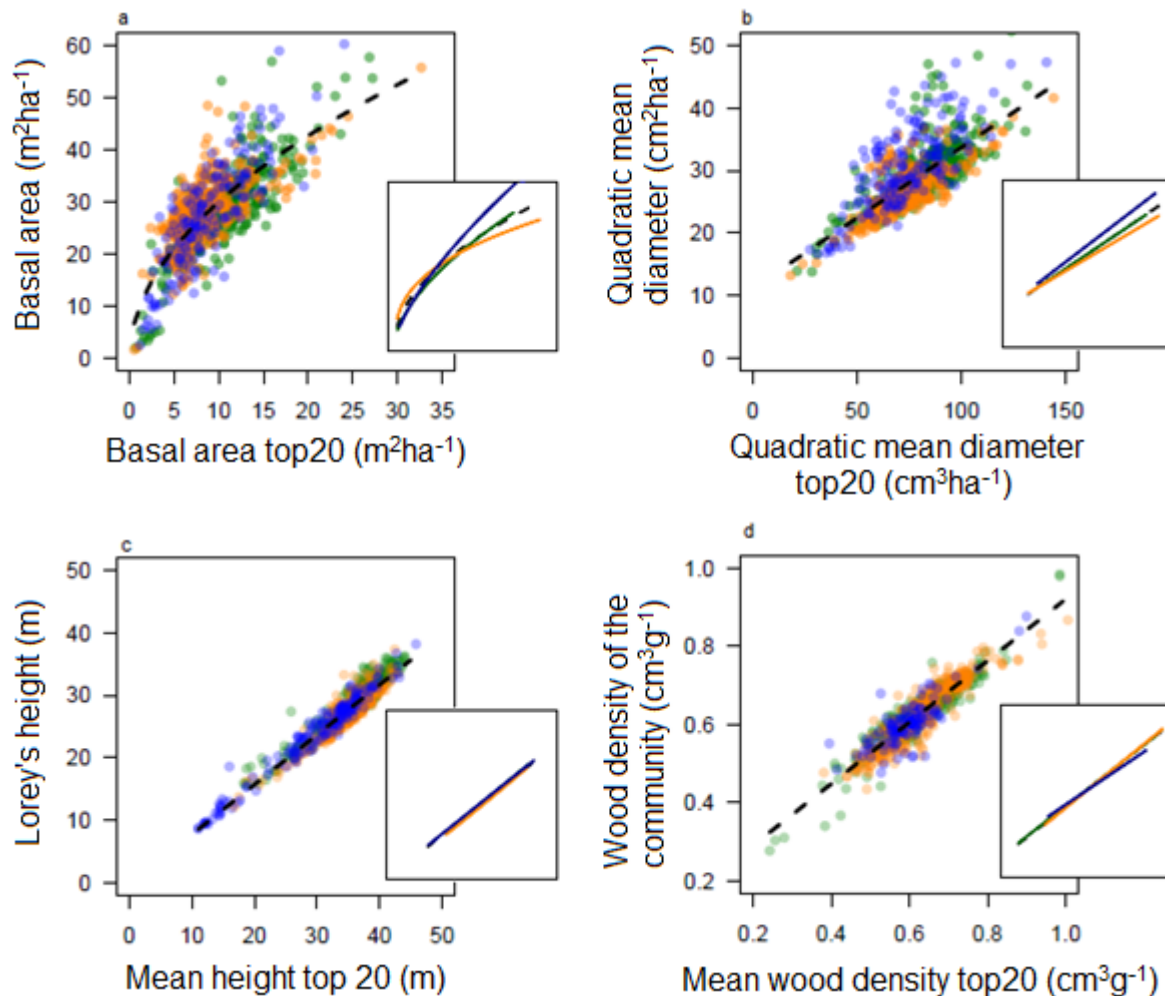
615

616 **Figure 1. Geographic distribution of the plot database.** We used 867 plots of 1 hectare
617 from 118 sites. Dots are colored according to floristic affinities (Slik et al. 2015), with America,
618 Africa and Asia respectively in orange, green and blue. They are also sized according the total
619 area surveyed in each site.



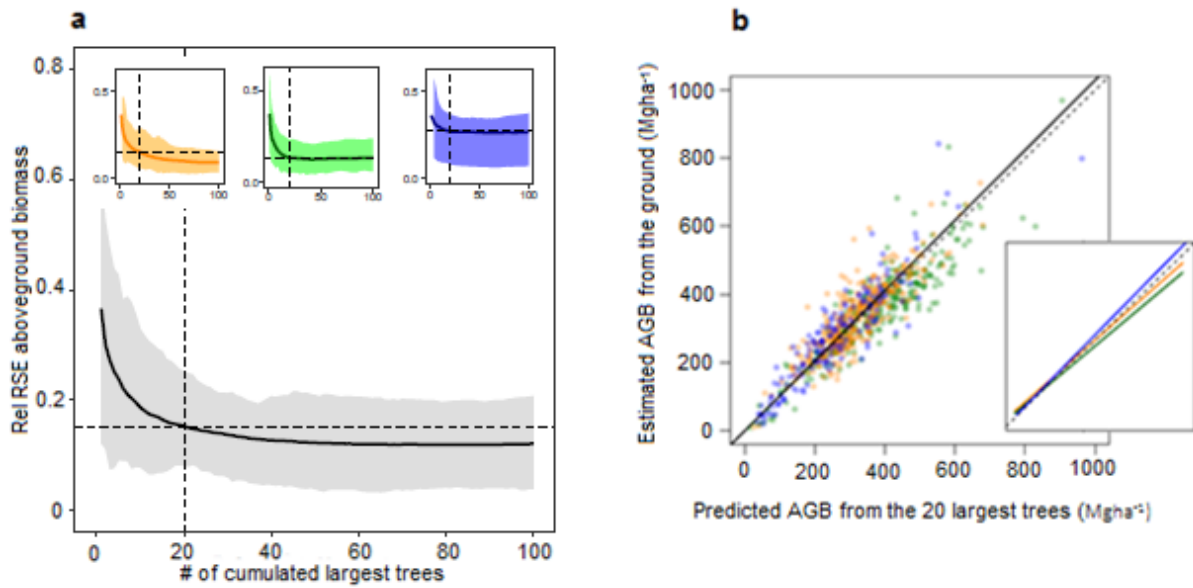
620

621 **Figure 2. Quality of the prediction of plot metrics from largest trees.** Variation of the
 622 relative Root Mean Square Error (rRMSE) of the prediction of plot metric from i largest trees
 623 versus the cumulative number of largest trees for (a) basal area, (b) quadratic mean diameter,
 624 (c) Lorey's height and (d) wood density weighted by the basal area. Results are displayed at
 625 the pan-tropical level (main plot in grey) and at the continental level (subplots; orange =
 626 America; green = Africa; blue = Asia). The solid line and shading shows the mean rRMSE and
 627 the 5th and the 95th percentiles. Dashed lines represent the mean rRMSE observed for each
 628 model, when considering the 20 largest trees.



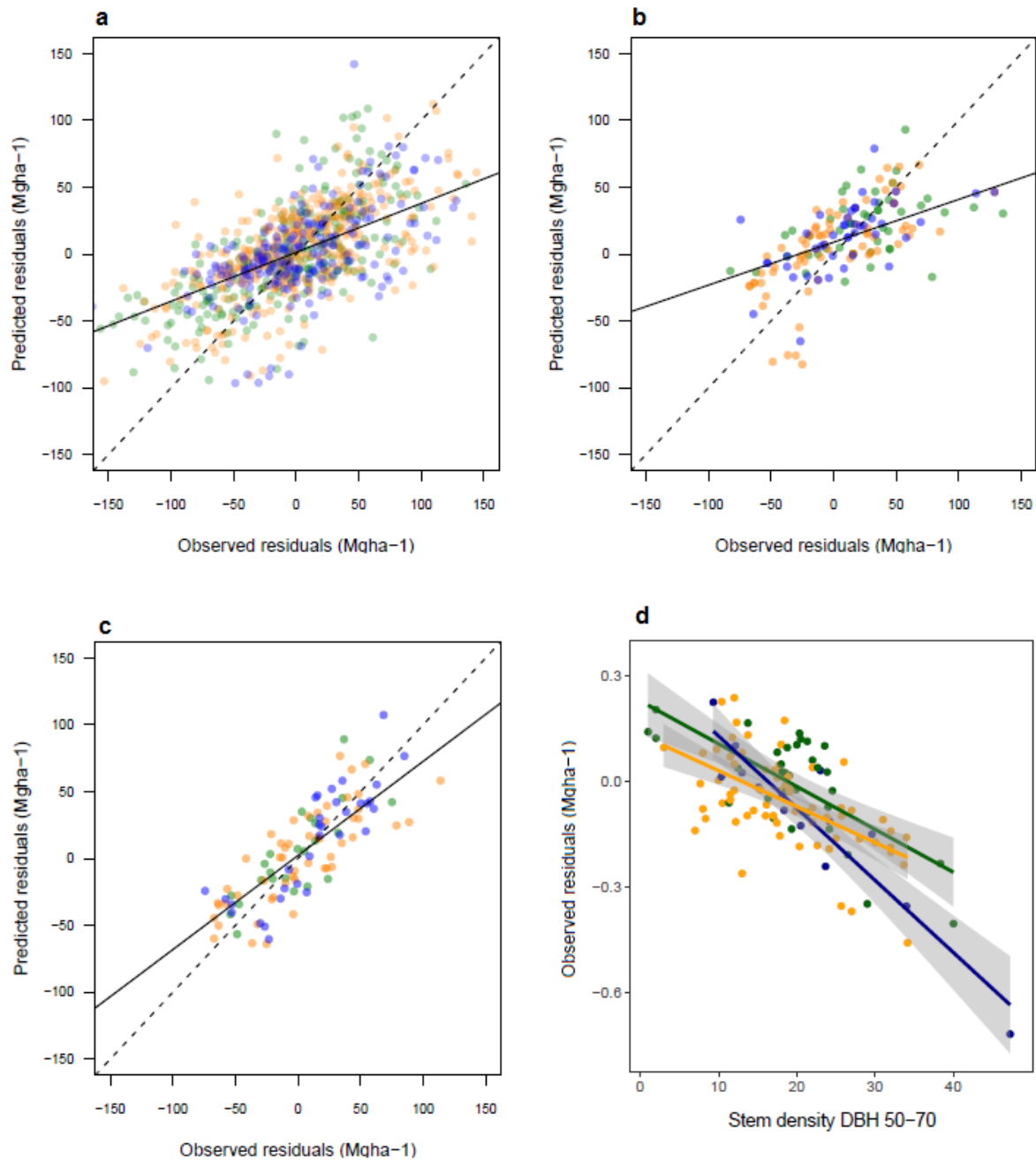
629

630 **Figure 3. Prediction of plot metrics (y-axis) from the 20 largest trees (x-axis).** Results are
 631 shown for (a) basal area, (b) quadratic mean diameter, (c) Lorey's Height and (d) wood density
 632 weighted by the basal area. Each dot corresponds to a single plot, colored in orange, green
 633 and blue for America, Africa and Asia respectively. Both pan-tropical (black dashed lines) and
 634 continental (coloured lines) regression models are displayed. These results show that
 635 substantial part of remaining variance, i.e. not explained by largest trees, is found when
 636 predicting the basal area and the quadratic mean diameter, with slight but significant
 637 differences between continents.



638

639 **Figure 4. Prediction of AGB from plot metrics of the 20 largest trees.** Results are shown
 640 for the 867 plots, among the three continents colored orange, green and blue for America,
 641 Africa and Asia respectively. The regression line of the model is shown as a continuous black
 642 line while the dashed black line shows a 1:1 relationship. The figure shows an unbiased
 643 prediction of AGB across the 867 plots, with slight but significant differences between the 3
 644 continents.



645

646 **Figure 5. Predicted vs. observed residuals of above ground biomass predicted from the**

647 **20 largest trees.** Residuals are explored at three different levels: (a) plot, (b) site [without

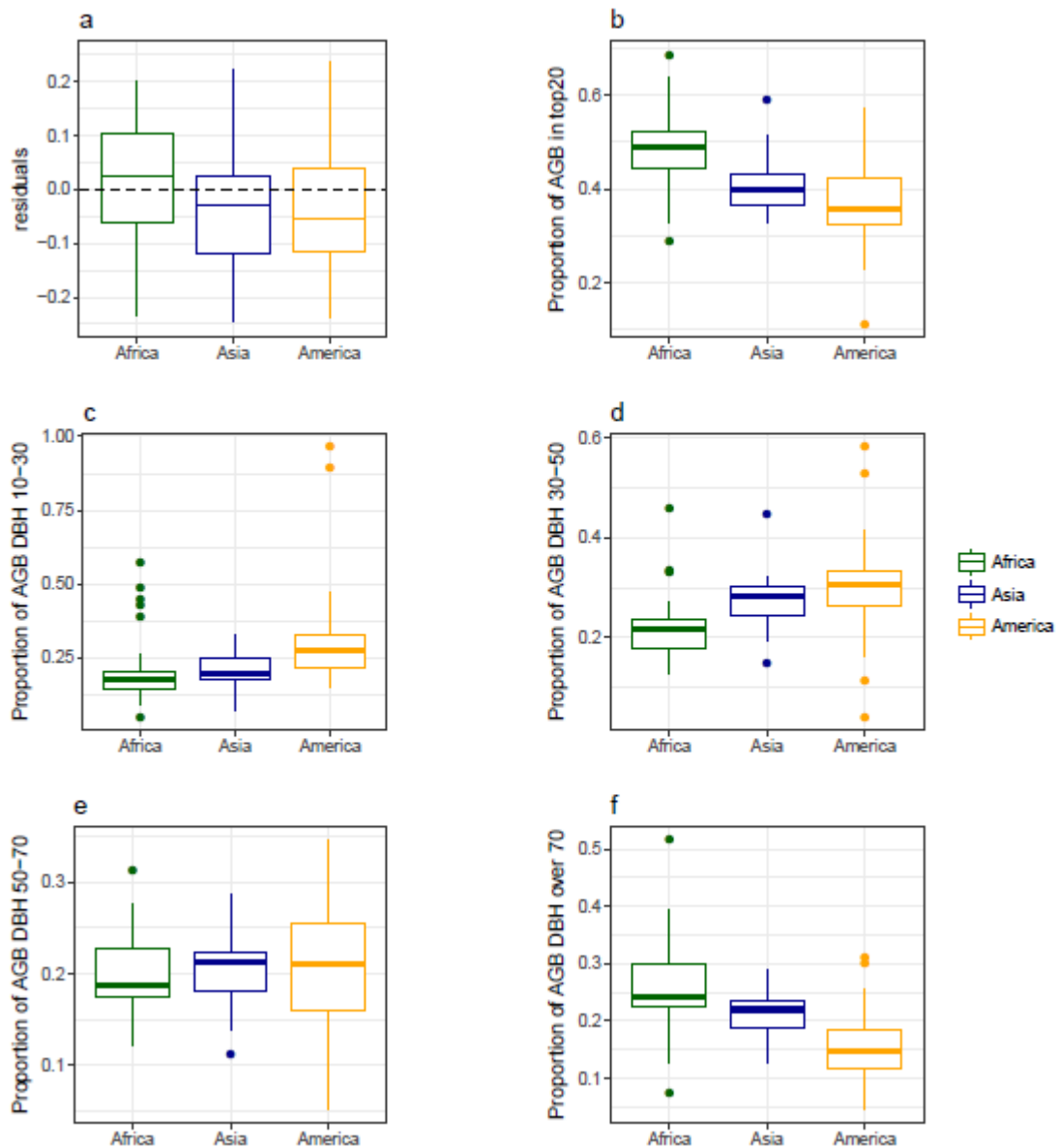
648 considering the diameter structure as an explanatory variable], (c) site [considering the

649 diameter structure] and (d) along the stem density of medium size trees. America, Africa and

650 Asia are colored in orange, green and blue respectively. The figures show a good prediction

651 of residuals in (a) and (b), driven by stem density, and a less biased prediction in (c), driven by

652 the diameter structure. Variance of observed residuals are also well explained by the stem
653 density of medium size trees (d), which mainly drive the first axis of the PCA.



654

655 Figure 6. Comparison across continents of aboveground biomass prediction per site and their
 656 contribution to different share of the diameter structure. Africa, Asia and America, are colored
 657 in green, blue and orange, respectively. The distribution of the residuals of pan-tropical
 658 aboveground biomass prediction from the 20 largest trees (a) shows predictions are slightly
 659 overestimated in Africa (+2%), and slightly underestimated in Asia (-2%) and America (-6%).
 660 The proportion of aboveground biomass in the 20 largest trees (b) is highest in Africa (48%),
 661 followed by Asia (40%) and America (35%). The decomposition across four diameter classes
 662 (c-f, i.e. from 10 to 30, 30 to 50, 50 to 70 and beyond 70 cm) of their relative share of the total
 663 biomass shows that most of the biomass is found in the large trees in Africa, and in the small

664 to medium trees in America. Asia presenting a more balanced distribution of biomass across
665 the diameter structure.

666 **Tables**

667 **Table 1. Weight of each variable retained for the explanation of AGB residuals.** Weights
 668 are calculated as a type II sum of squares, which measures the decreased residual sum of
 669 squares due to an added variable once all the other variables have been introduced into the
 670 model. Results are shown for the exploration of residuals at the plot and at the site level, with
 671 and without consideration of the diameter structure. Weights are dominated by structural
 672 variables, and in particular the stem density and the diameter structure. Height, wood density
 673 and continent have also a non-negligible influence on residuals.

Level of residual	Parameter	Weight	674
Plot	Stem density*	79	
	Continent*	18	
	Lorey's height*	1	
	Major soil types	1	
	Temperature	1	
	Wood density weighted by the basal area	0	
	Rainfall	0	
Site without diametric structure	Stem density*	54	
	Continent*	28	
	Wood density weighted by the basal area*	11	
	Rainfall	3	
	Major soil types	3	
	Temperature	2	
	Lorey's height	0	
Site with diametric structure	PCA axis 1*	69	
	Lorey's height*	23	
	Rainfall*	3	
	Major soil types	3	
	Continent	1	
	Temperature	1	
	Wood density weighted by the basal area	0	
	PCA axis 2	0	

675 **References**

- 676 Asner, G.G.P., Mascaro, J., Muller-Landau, H.H.C., Vieilledent, G., Vaudry, R., Rasamoelina,
677 M., Hall, J.S. & van Breugel, M. (2012) A universal airborne LiDAR approach for tropical
678 forest carbon mapping. *Oecologia*, **168**, 1147–1160.
- 679 Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C.B., Sinca, F., Vaughn,
680 N.R. & Llactayo, W. (2017) Airborne laser-guided imaging spectroscopy to map forest
681 trait diversity and guide conservation. *Science*, **355**.
- 682 Asner, G.P. & Mascaro, J. (2014) Mapping tropical forest carbon: Calibrating plot estimates
683 to a simple LiDAR metric. *Remote Sensing of Environment*, **140**, 614–624.
- 684 Baccini, A., Walker, W., Carvalho, L., Farina, M., Sulla-Menashe, D. & Houghton, R.A. (2017)
685 Tropical forests are a net carbon source based on aboveground measurements of gain
686 and loss. *Science (New York, N.Y.)*, **358**, 230–234.
- 687 Barbier, N. & Coutron, P. (2015) Attenuating the bidirectional texture variation of satellite
688 images of tropical forest canopies. *Remote Sensing of Environment*, **171**, 245–260.
- 689 Barbier, N., Proisy, C., Véga, C., Sabatier, D. & Coutron, P. (2011) Bidirectional texture
690 function of high resolution optical images of tropical forest: An approach using LiDAR
691 hillshade simulations. *Remote Sensing of Environment*, **115**, 167–179.
- 692 Bastin, J.-F., Barbier, N., Coutron, P., Adams, B., Shapiro, A., Bogaert, J., De Cannière, C.,
693 De Cannière, C. & De Cannière, C. (2014) Aboveground biomass mapping of African
694 forest mosaics using canopy texture analysis: toward a regional approach. *Ecological*
695 *Applications*, **24**, 1984–2001.
- 696 Bastin, J.-F., Barbier, N., Réjou-Méchain, M., Fayolle, A., Gourlet-Fleury, S., Maniatis, D., De
697 Haulleville, T., Baya, F., Beeckman, H., Beina, D., Coutron, P., Chuyong, G., Dauby,
698 G., Doucet, J.-L., Droissart, V., Dufrêne, M., Ewango, C., Gillet, J.F., Gonmadje, C.H.,
699 Hart, T., Kavali, T., Kenfack, D., Libalah, M., Malhi, Y., Makana, J.-R., Pélissier, R.,
700 Ploton, P., Serckx, A., Sonké, B., Stevart, T., Thomas, D.W., De Cannière, C. &
701 Bogaert, J. (2015) Seeing Central African forests through their largest trees. *Scientific*
702 *Reports*, **5**.

703 Bastin, J.-F., Berrahmouni, N., Grainger, A., Maniatis, D., Mollicone, D., Moore, R., Patriarca,
704 C., Picard, N., Sparrow, B., Abraham, E.M., Aloui, K., Atesoglu, A., Attore, F., Bassüllü,
705 Ç., Bey, A., Garzuglia, M., García-Montero, L.G., Groot, N., Guerin, G., Laestadius, L.,
706 Lowe, A.J., Mamane, B., Marchi, G., Patterson, P., Rezende, M., Ricci, S., Salcedo, I.,
707 Diaz, A.S.-P., Stolle, F., Surappaeva, V. & Castro, R. (2017) The extent of forest in
708 dryland biomes. *Science*, **356**, 635–638.

709 Bennett, A.C., McDowell, N.G., Allen, C.D. & Anderson-Teixeira, K.J. (2015) Larger trees
710 suffer most during drought in forests worldwide. *Nature Plants*.

711 Blanchard, E., Birnbaum, P., Ibanez, T., Boutreux, T., Antin, C., Ploton, P., Vincent, G.,
712 Pouteau, R., Vandrot, H., Hequet, V., Barbier, N., Droissart, V., Sonké, B., Texier, N.,
713 Kamdem, N.G., Zebaze, D., Libalah, M. & Couteron, P. (2016) Contrasted allometries
714 between stem diameter, crown area, and tree height in five tropical biogeographic
715 areas. *Trees*, **30**, 1953–1968.

716 Carré, F., Hiederer, R., Blujdea, V. & Koeble, R. (2010) *Background Guide for the*
717 *Calculation of Land Carbon Stocks in the Biofuels Sustainability Scheme Drawing on*
718 *the 2006 IPCC Guidelines for National Greenhouse Gas Inventories. EUR 24573 EN.*,
719 Luxembourg.

720 Chave, J., Coomes, D., Jansen, S., Lewis, S.L., Swenson, N.G. & Zanne, A.E. (2009)
721 Towards a worldwide wood economics spectrum. *Ecology letters*, **12**, 351–66.

722 Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M.S., Delitti, W.B.C.,
723 Duque, A., Eid, T., Fearnside, P.M., Goodman, R.C., Henry, M., Martínez-Yrizar, A.,
724 Mugasha, W.A., Muller-Landau, H.C., Mencuccini, M., Nelson, B.W., Ngomanda, A.,
725 Nogueira, E.M., Ortiz-Malavassi, E., Pélissier, R., Ploton, P., Ryan, C.M., Saldarriaga,
726 J.G. & Vieilledent, G. (2014) Improved allometric models to estimate the aboveground
727 biomass of tropical trees. *Global change biology*, **20**, 3177–3190.

728 Chave, J., Riera, B., Dubois, M.-A. & Riéra, B. (2001) Estimation of biomass in a neotropical
729 forest of French Guiana : spatial and temporal variability. *Journal of Tropical Ecology*,
730 **17**, 79–96.

731 Clark, D.B. & Clark, D.A. (1996) Abundance, growth and mortality of very large trees in
732 neotropical lowland rain forest. *Forest Ecology and Management*, **80**, 235–244.

733 Clark, D.B. & Clark, D.A. (2000) Landscape-scale variation in forest structure and biomass in
734 a tropical rain forest. *Forest Ecology and Management*, **137**, 185–198.

735 Clark, D.B. & Kellner, J.R. (2012) Tropical forest biomass estimation and the fallacy of
736 misplaced concreteness. *Journal of Vegetation Science*, **23**, 1191–1196.

737 Coomes, D.A., Dalponte, M., Jucker, T., Asner, G.P., Banin, L.F., Burslem, D.F.R.P., Lewis,
738 S.L., Nilus, R., Phillips, O.L., Phua, M.-H. & Qie, L. (2017) Area-based vs tree-centric
739 approaches to mapping forest carbon in Southeast Asian forests from airborne laser
740 scanning data. *Remote Sensing of Environment*, **194**, 77–88.

741 D’Oliveira, M.V.N., Reutebuch, S.E., McGaughey, R.J. & Andersen, H.-E. (2012) Estimating
742 forest biomass and identifying low-intensity logging areas using airborne scanning lidar
743 in Antimary State Forest, Acre State, Western Brazilian Amazon. *Remote Sensing of*
744 *Environment*, **124**, 479–491.

745 Farris, C.E., Bohlman, S.A., Hubbell, S. & Pacala, S.W. (2016) Dominance of the
746 suppressed: Power-law size structure in tropical forests. *Science*, **351**.

747 Fayolle, A., Loubota Panzou, G.J., Drouet, T., Swaine, M.D., Bauwens, S., Vleminckx, J.,
748 Biwole, A., Lejeune, P. & Doucet, J.-L. (2016) Taller trees, denser stands and greater
749 biomass in semi-deciduous than in evergreen lowland central African forests. *Forest*
750 *Ecology and Management*, **374**, 42–50.

751 Feldpausch, T.R., Banin, L., Phillips, O.L., Baker, T.R., Lewis, S.L., Quesada, C. a., Affum-
752 Baffoe, K., Arets, E.J.M.M., Berry, N.J., Bird, M., Brondizio, E.S., de Camargo, P.,
753 Chave, J., Djangbletey, G., Domingues, T.F., Drescher, M., Fearnside, P.M., França,
754 M.B., Fyllas, N.M., Lopez-Gonzalez, G., Hladik, a., Higuchi, N., Hunter, M.O., Iida, Y.,
755 Salim, K. a., Kassim, a. R., Keller, M., Kemp, J., King, D. a., Lovett, J.C., Marimon,
756 B.S., Marimon-Junior, B.H., Lenza, E., Marshall, a. R., Metcalfe, D.J., Mitchard, E.T. a.,
757 Moran, E.F., Nelson, B.W., Nilus, R., Nogueira, E.M., Palace, M., Patiño, S., Peh, K.S.-
758 H., Raventos, M.T., Reitsma, J.M., Saiz, G., Schrodt, F., Sonké, B., Taedoumg, H.E.,

759 Tan, S., White, L., Wöll, H. & Lloyd, J. (2011) Height-diameter allometry of tropical forest
760 trees. *Biogeosciences*, **8**, 1081–1106.

761 Feldpausch, T.R., Lloyd, J., Lewis, S.L., Brienen, R.J.W., Gloor, E., Monteagudo Mendoza,
762 a., Lopez-Gonzalez, G., Banin, L., Abu Salim, K., Affum-Baffoe, K., Alexiades, M.,
763 Almeida, S., Amaral, I., Andrade, a., Aragão, L.E.O.C., Araujo Murakami, a., Arets,
764 E.J.M.M., Arroyo, L., Baker, T.R., Bánki, O.S., Berry, N.J., Cardozo, N., Chave, J.,
765 Comiskey, J. a., Dávila, E. a., de Oliveira, a., DiFiore, a., Djagbletey, G., Domingues,
766 T.F., Erwin, T.L., Fearnside, P.M., França, M.B., Freitas, M. a., Higuchi, N., Iida, Y.,
767 Jiménez, E., Kassim, a. R., Killeen, T.J., Laurance, W.F., Lovett, J.C., Malhi, Y.,
768 Marimon, B.S., Marimon-Junior, B.H., Lenza, E., Marshall, a. R., Mendoza, C.,
769 Metcalfe, D.J., Mitchard, E.T. a., Nelson, B.W., Nilus, R., Nogueira, E.M., Parada, a.,
770 Peh, K.S.-H., Pena Cruz, a., Peñuela, M.C., Pitman, N.C. a., Prieto, a., Quesada, C.
771 a., Ramírez, F., Ramírez-Angulo, H., Reitsma, J.M., Rudas, a., Saiz, G., Salomão,
772 R.P., Schwarz, M., Silva, N., Silva-Espejo, J.E., Silveira, M., Sonké, B., Stropp, J.,
773 Taedoumg, H.E., Tan, S., ter Steege, H., Terborgh, J., Torello-Raventos, M., van der
774 Heijden, G.M.F., Vásquez, R., Vilanova, E., Vos, V., White, L., Wilcock, S., Woell, H. &
775 Phillips, O.L. (2012) Tree height integrated into pan-tropical forest biomass estimates.
776 *Biogeosciences Discussions*, **9**, 2567–2622.

777 Ferraz, A., Saatchi, S., Mallet, C. & Meyer, V. (2016) Lidar detection of individual tree size in
778 tropical forests. *Remote Sensing of Environment*, **183**, 318–333.

779 Gibbs, H.K., Brown, S., Niles, J.O. & Foley, J. a (2007) Monitoring and estimating tropical
780 forest carbon stocks: making REDD a reality. *Environmental Research Letters*, **2**, 1–13.

781 Harms, K.E., Wright, S.J., Calderón, O., Hernández, a & Herre, E. a (2000) Pervasive
782 density-dependent recruitment enhances seedling diversity in a tropical forest. *Nature*,
783 **404**, 493–5.

784 Ho Tong Minh, D., Le Toan, T., Rocca, F., Tebaldini, S., Villard, L., Réjou-Méchain, M.,
785 Phillips, O.L., Feldpausch, T.R., Dubois-Fernandez, P., Scipal, K. & Chave, J. (2016)
786 SAR tomography for the retrieval of forest biomass and height: Cross-validation at two

787 tropical forest sites in French Guiana. *Remote Sensing of Environment*, **175**, 138–147.

788 IPCC (2006) *2006 IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by*
789 *the National Greenhouse Gas Inventories Programme*, IGES. (ed. by H.S. Eggleston),
790 L. Buendia), K. Miwa), T. Ngara), and K. Tanabe) Japan, Japan.

791 Jucker, T., Caspersen, J., Chave, J., Antin, C., Barbier, N., Bongers, F., Dalponte, M., van
792 Ewijk, K.Y., Forrester, D.I., Haeni, M., Higgins, S.I., Holdaway, R.J., Iida, Y., Lorimer, C.,
793 Marshall, P.L., Momo, S., Moncrieff, G.R., Ploton, P., Poorter, L., Rahman, K.A.,
794 Schlund, M., Sonké, B., Sterck, F.J., Trugman, A.T., Usoltsev, V.A., Vanderwel, M.C.,
795 Waldner, P., Wedeux, B.M.M., Wirth, C., Wöll, H., Woods, M., Xiang, W., Zimmermann,
796 N.E. & Coomes, D.A. (2016) Allometric equations for integrating remote sensing
797 imagery into forest monitoring programmes. *Global Change Biology*, n/a-n/a.

798 Kearsley, E., de Haulleville, T., Hufkens, K., Kidimbu, A., Toirambe, B., Baert, G., Huygens,
799 D., Kebede, Y., Defourny, P., Bogaert, J., Beeckman, H., Steppe, K., Boeckx, P. &
800 Verbeeck, H. (2013) Conventional tree height-diameter relationships significantly
801 overestimate aboveground carbon stocks in the Congo Basin. *Nature communications*.

802 Kellner, J.R., Clark, D.B. & Hubbell, S.P. (2009) Pervasive canopy dynamics produce short-
803 term stability in a tropical rain forest landscape. *Ecology Letters*, **12**, 155–164.

804 Kellner, J.R. & Hubbell, S.P. (2017) Adult mortality in a low-density tree population using
805 high-resolution remote sensing. *Ecology*, **98**, 1700–1709.

806 Langsrud, Ø. (2003) ANOVA for unbalanced data: Use Type II instead of Type III sums of
807 squares. *Statistics and Computing*, **13**, 163–167.

808 Laurance, W.F., Delamônica, P., Laurance, S.G., Vasconcelos, H.L. & Lovejoy, T.E. (2000)
809 Conservation: Rainforest fragmentation kills big trees. *Nature*, **404**, 836–836.

810 Lewis, S.L.L., Sonké, B., Sunderland, T., Begne, S.K.S.K., Lopez-Gonzalez, G., Heijden,
811 G.M.F. Van Der, Phillips, O.L.O.L., Affum-Baffoe, K., Baker, T.R.T.R., Banin, L., Bastin,
812 J.-F.J.-F., Beeckman, H., Boeckx, P., Bogaert, J., De Cannière, C., Chezeaux, V., Clark,
813 C.J.C.J., Collins, M., Djagbletey, G., Droissart, V., Doucet, J.-L.J.-L., Ewango,
814 C.E.N.C.E.N., Fauset, S., Feldpausch, T.R.R., Foli, E.G., Gillet, J.-F.J.-F., Hamilton,

815 A.C.A.C., Harris, D.J.D.J., Hart, T.B.T.B., de Haulleville, T., Hladik, A., Hufkens, K.,
816 Huygens, D., Jeanmart, P., Jeffery, K.J., Kamdem, M.-N.D., Kearsley, E., Leal, M.E.E.,
817 Lloyd, J., Lovett, J.C.J.C., Makana, J.-R.J.-R., Malhi, Y., Marshall, A.R.R., Ojo, L., Peh,
818 K.S.-H.K.S.-H., Pickavance, G., Poulsen, J.R., Reitsma, M., Sheil, D., Simo, M., Steppe,
819 K., Taedoumg, H.E.E., Talbot, J., Taplin, J.J.R.D., Taylor, D., Thomas, S.C.S.C.,
820 Toirambe, B., Verbeeck, H., Vleminckx, J., White, L.J.T.J.T., Willcock, S., Woell, H.,
821 Zemagho, L., van der Heijden, G.M.F., Phillips, O.L.O.L., Affum-Baffoe, K., Baker,
822 T.R.T.R., Banin, L., Bastin, J.-F.J.-F., Beeckman, H., Boeckx, P., Bogaert, J., De
823 Cannière, C., Chezeaux, E., Clark, C.J.C.J., Collins, M., Djangbletey, G., Djuikouo,
824 M.N.K., Droissart, V., Doucet, J.-L.J.-L., Ewango, C.E.N.C.E.N., Fauset, S., Feldpausch,
825 T.R.R., Foli, E.G., Gillet, J.-F.J.-F., Hamilton, A.C.A.C., Harris, D.J.D.J., Hart, T.B.T.B.,
826 de Haulleville, T., Hladik, A., Hufkens, K., Huygens, D., Jeanmart, P., Jeffery, K.J.,
827 Kearsley, E., Leal, M.E.E., Lloyd, J., Lovett, J.C.J.C., Makana, J.-R.J.-R., Malhi, Y.,
828 Marshall, A.R.R., Ojo, L., Peh, K.S.-H.K.S.-H., Pickavance, G., Poulsen, J.R., Reitsma,
829 J.M., Sheil, D., Simo, M., Steppe, K., Taedoumg, H.E.E., Talbot, J., Taplin, J.J.R.D.,
830 Taylor, D., Thomas, S.C.S.C., Toirambe, B., Verbeeck, H., Vleminckx, J., White,
831 L.J.T.J.T., Willcock, S., Woell, H. & Zemagho, L. (2013) Above-ground biomass and
832 structure of 260 African tropical forests. *Philosophical transactions of the Royal Society*
833 *of London. Series B, Biological sciences*, **368**.

834 Lindenmayer, D.B., Laurance, W.F. & Franklin, J.F. (2012) Global decline in large old trees.
835 *Science*, **338**, 1305–1306.

836 Malhi, Y., Wood, D., Baker, T.R., Wright, J., Phillips, O.L., Cochrane, T., Meir, P., Chave, J.,
837 Almeida, S., Arroyo, L., Higuchi, N., Killeen, T.J., Laurance, S.G., Laurance, W.F.,
838 Lewis, S.L., Monteagudo, A., Neill, D. a., Vargas, P.N., Pitman, N.C. a., Quesada, C.A.,
839 Salomao, R., Silva, J.N.M., Lezama, A.T., Terborgh, J., Martinez, R.V. & Vinceti, B.
840 (2006) The regional variation of aboveground live biomass in old-growth Amazonian
841 forests. *Global Change Biology*, **12**, 1107–1138.

842 Maniatis, D. & Mollicone, D. (2010) Options for sampling and stratification for national forest

843 inventories to implement REDD+ under the UNFCCC. *Carbon balance and*
844 *management*, **5**, 9.

845 Mascaro, J., Detto, M., Asner, G.P. & Muller-Landau, H.C. (2011) Evaluating uncertainty in
846 mapping forest carbon with airborne LiDAR. *Remote Sensing of Environment*, **115**,
847 3770–3774.

848 Meakem, V., Tepley, A.J., Gonzalez-Akre, E.B., Herrmann, V., Muller-Landau, H.C., Wright,
849 S.J., Hubbell, S.P., Condit, R. & Anderson-Teixeira, K.J. (2017) Role of tree size in
850 moist tropical forest carbon cycling and water deficit responses. *New Phytologist*.

851 Molto, Q., Hérault, B., Boreux, J.-J., Daullet, M., Rousteau, A. & Rossi, V. (2014) Predicting
852 tree heights for biomass estimates in tropical forests – a test from French Guiana.
853 *Biogeosciences*, **11**, 3121–3130.

854 Nepstad, D.C., Tohver, I.M., Ray, D., Moutinho, P. & Cardinot, G. (2007) Mortality of large
855 trees and lianas following experimental drought in an Amazon forest. *Ecology*, **88**,
856 2259–69.

857 New, M., Hulme, M., Jones, P., New, M., Hulme, M. & Jones, P. (1999) Representing
858 Twentieth-Century Space–Time Climate Variability. Part I: Development of a 1961–90
859 Mean Monthly Terrestrial Climatology. *Journal of Climate*, **12**, 829–856.

860 New, M., Lister, D., Hulme, M. & Makin, I. (2002) A high-resolution data set of surface
861 climate over global land areas. *Climate Research*, **21**, 1–25.

862 Ploton, P., Barbier, N., Coutron, P., Antin, C.M., Ayyappan, N., Balachandran, N., Barathan,
863 N., Bastin, J.-F., Chuyong, G., Dauby, G., Droissart, V., Gastellu-Etchegorry, J.-P.,
864 Kamdem, N.G., Kenfack, D., Libalah, M., Mofack, G., Momo, S.T., Pargal, S., Petronelli,
865 P., Proisy, C., Réjou-Méchain, M., Sonké, B., Texier, N., Thomas, D., Verley, P.,
866 Zebaze Dongmo, D., Berger, U. & Pélissier, R. (2017) Toward a general tropical forest
867 biomass prediction model from very high resolution optical satellite images. *Remote*
868 *Sensing of Environment*, **200**.

869 Ploton, P., Barbier, N., Takoudjou Momo, S., Réjou-Méchain, M., Boyemba Bosela, F.,
870 Chuyong, G., Dauby, G., Droissart, V., Fayolle, A., Goodman, R.C., Henry, M.,

871 Kamdem, N.G., Mukirania, J.K., Kenfack, D., Libalah, M., Ngomanda, A., Rossi, V.,
872 Sonké, B., Texier, N., Thomas, D., Zebaze, D., Couteron, P., Berger, U. & Pélissier, R.
873 (2016) Closing a gap in tropical forest biomass estimation: taking crown mass variation
874 into account in pantropical allometries. *Biogeosciences*, **13**, 1571–1585.

875 Ploton, P., Pélissier, R. & Proisy, C. (2012) Assessing aboveground tropical forest biomass
876 using Google Earth canopy images. *Ecological Applications*, **22**, 993–1003.

877 Poorter, L., van der Sande, M.T., Thompson, J., Arets, E.J.M.M., Alarcón, A., Álvarez-
878 Sánchez, J., Ascarrunz, N., Balvanera, P., Barajas-Guzmán, G., Boit, A., Bongers, F.,
879 Carvalho, F.A., Casanoves, F., Cornejo-Tenorio, G., Costa, F.R.C., de Castilho, C. V.,
880 Duivenvoorden, J.F., Dutrieux, L.P., Enquist, B.J., Fernández-Méndez, F., Finegan, B.,
881 Gormley, L.H.L., Healey, J.R., Hoosbeek, M.R., Ibarra-Manríquez, G., Junqueira, A.B.,
882 Levis, C., Licona, J.C., Lisboa, L.S., Magnusson, W.E., Martínez-Ramos, M., Martínez-
883 Yrizar, A., Martorano, L.G., Maskell, L.C., Mazzei, L., Meave, J.A., Mora, F., Muñoz, R.,
884 Nych, C., Pansonato, M.P., Parr, T.W., Paz, H., Pérez-García, E.A., Rentería, L.Y.,
885 Rodríguez-Velazquez, J., Rozendaal, D.M.A., Ruschel, A.R., Sakschewski, B., Salgado-
886 Negret, B., Schiatti, J., Simões, M., Sinclair, F.L., Souza, P.F., Souza, F.C., Stropp, J.,
887 ter Steege, H., Swenson, N.G., Thonicke, K., Toledo, M., Uriarte, M., van der Hout, P.,
888 Walker, P., Zamora, N. & Peña-Claros, M. (2015) Diversity enhances carbon storage in
889 tropical forests. *Global Ecology and Biogeography*, **24**, 1314–1328.

890 Proisy, C., Couteron, P. & Fromard, F. (2007) Predicting and mapping mangrove biomass
891 from canopy grain analysis using Fourier-based textural ordination of IKONOS images.
892 *Remote Sensing of Environment*, **109**, 379–392.

893 Réjou-Méchain, M., Tanguy, A., Piponiot, C., Chave, J. & Hérault, B. (2017) biomass : an r
894 package for estimating above-ground biomass and its uncertainty in tropical forests.
895 *Methods in Ecology and Evolution*, **8**, 1163–1167.

896 Remm, J. & Löhmus, A. (2011) Tree cavities in forests – The broad distribution pattern of a
897 keystone structure for biodiversity. *Forest Ecology and Management*, **262**, 579–585.

898 Romijn, E., Herold, M., Kooistra, L., Murdiyarso, D. & Verchot, L. (2012) Assessing

899 capacities of non-Annex I countries for national forest monitoring in the context of
900 REDD+. *Environmental Science & Policy*, **19–20**, 33–48.

901 Rutishauser, E., Hérault, B., Baraloto, C., Blanc, L., Descroix, L., Sotta, E.D., Ferreira, J.,
902 Kanashiro, M., Mazzei, L., d'Oliveira, M.V.N., de Oliveira, L.C., Peña-Claros, M., Putz,
903 F.E., Ruschel, A.R., Rodney, K., Roopsind, A., Shenkin, A., da Silva, K.E., de Souza,
904 C.R., Toledo, M., Vidal, E., West, T.A.P., Wortel, V. & Sist, P. (2015) Rapid tree carbon
905 stock recovery in managed Amazonian forests. *Current biology : CB*, **25**, R787-8.

906 Rutishauser, E., Wagner, F., Herault, B., Nicolini, E.-A. & Blanc, L. (2010) Contrasting above-
907 ground biomass balance in a Neotropical rain forest. *Journal of Vegetation Science*,
908 672–682.

909 Saatchi, S.S., Houghton, R. a., Dos Santos Alvalá, R.C., Soares, J. V. & Yu, Y. (2007)
910 Distribution of aboveground live biomass in the Amazon basin. *Global Change Biology*,
911 **13**, 816–837.

912 Sist, P., Mazzei, L., Blanc, L. & Rutishauser, E. (2014) Large trees as key elements of
913 carbon storage and dynamics after selective logging in the Eastern Amazon. *Forest
914 Ecology and Management*, **318**, 103–109.

915 Slik, J.W.F., Aiba, S.-I., Brearley, F.Q., Cannon, C.H., Forshed, O., Kitayama, K., Nagamasu,
916 H., Nilus, R., Payne, J., Paoli, G., Poulsen, A.D., Raes, N., Sheil, D., Sidiyasa, K.,
917 Suzuki, E. & Van Valkenburg, J.L.C.H. (2010) Environmental correlates of tree biomass,
918 basal area, wood specific gravity and stem density gradients in Borneo's tropical forests.
919 *Global Ecology and Biogeography*, **19**, 50–60.

920 Slik, J.W.F., Alvarez-loayza, P., Alves, L.F., Ashton, P., Balvanera, P., Bastian, M.L.,
921 Bellingham, P.J., Berg, E. Van Den, Bernacci, L., Conceição, P., Blanc, L., Böhning-
922 gaese, K., Boeckx, P., Boyle, B., Bradford, M., Brearley, F.Q., Hockemba, B.,
923 Bunyavejchewin, S., Matos, C.L., Castillo-santiago, M., Eduardo, L.M., Chai, S., Chen,
924 Y., Colwell, R.K., Robin, C.L., Clark, C., Clark, D.B., Deborah, A., Culmsee, H., Damas,
925 K., Dattaraja, H.S., Dauby, G., Davidar, P., Dewalt, S.J., Doucet, J., Duque, A., Durigan,
926 G., Eichhorn, K.A.O., Pedro, V., Eler, E., Ewango, C., Farwig, N., Feeley, K.J., Ferreira,

927 L., Field, R., Ary, T., Filho, D.O., Fletcher, C., Forshed, O., Fredriksson, G., Gillespie, T.,
928 Amarnath, G., Griffith, D.M., Grogan, J., Gunatilleke, N., Harris, D., Harrison, R., Hector,
929 A., Homeier, J., Imai, N., Itoh, A., Jansen, P.A., Joly, C.A., Jong, B.H.J. De,
930 Kartawinata, K., Kearsley, E., Kelly, D.L., Kenfack, D., Kitayama, K., Kooyman, R.,
931 Larney, E., Laurance, S., Laurance, W.F., Michael, J., Leao, I., Letcher, S.G., Lindsell,
932 J., Lu, X., Mansor, A., Marjokorpi, A., Martin, E.H., Meilby, H., Melo, F.P.L., Metcalfe,
933 D.J., Vincent, P., Metzger, J.P., Millet, J., Mohandass, D., Juan, C., Nagamasu, H.,
934 Nilus, R., Ochoa-gaona, S., Paudel, E., Permana, A., Maria, T.F., Rovero, F., Rozak,
935 A.H., Santos, B.A., Santos, F., Sarker, S.K., Satdichanh, M., Schmitt, C.B., Schöngart,
936 J., Tabarelli, M., Tang, J., Targhetta, N., Theilade, I., Thomas, D.W., Tchouto, P.,
937 Hurtado, J., Valkenburg, J.L.C.H. Van, Do, T. Van, Verbeeck, H., Adekunle, V., Vieira,
938 S.A., Alvarez-loayza, P., Alves, L.F., Berg, E. Van Den & Bernacci, L. (2015) An
939 estimate of the number of tropical tree species. *Proceedings of the National Academy of*
940 *Sciences*, **112**, E4628–E4629.

941 Slik, J.W.F., Franklin, J., Arroyo-Rodríguez, V., Field, R., Aguilar, S., Aguirre, N., Ahumada,
942 J., Aiba, S.-I., Alves, L.F., Anitha, K., Avella, A., Mora, F., Aymard, G.A.C., Báez, S.,
943 Balvanera, P., Bastian, M.L., Bastin, J.-F., Bellingham, P.J., Van Den Berg, E., Da
944 Conceição Bispo, P., Boeckx, P., Boehning-Gaese, K., Bongers, F., Boyle, B.,
945 Brambach, F., Brearley, F.Q., Brown, S., Chai, S.-L., Chazdon, R.L., Chen, S., Chhang,
946 P., Chuyong, G., Ewango, C., Coronado, I.M., Cristóbal-Azkarate, J., Culmsee, H.,
947 Damas, K., Dattaraja, H.S., Davidar, P., DeWalt, S.J., Diñ, H., Drake, D.R., Duque, A.,
948 Durigan, G., Eichhorn, K., Eler, E.S., Enoki, T., Ensslin, A., Fandohan, A.B., Farwig, N.,
949 Feeley, K.J., Fischer, M., Forshed, O., Garcia, Q.S., Garkoti, S.C., Gillespie, T.W.,
950 Gillet, J.-F., Gonmadje, C., Granzow-De La Cerda, I., Griffith, D.M., Grogan, J.,
951 Hakeem, K.R., Harris, D.J., Harrison, R.D., Hector, A., Hemp, A., Homeier, J., Hussain,
952 M.S., Ibarra-Manríquez, G., Hanum, I.F., Imai, N., Jansen, P.A., Joly, C.A., Joseph, S.,
953 Kartawinata, K., Kearsley, E., Kelly, D.L., Kessler, M., Killeen, T.J., Kooyman, R.M.,
954 Laumonier, Y., Laurance, S.G., Laurance, W.F., Lawes, M.J., Letcher, S.G., Lindsell, J.,

955 Lovett, J., Lozada, J., Lu, X., Lykke, A.M., Bin Mahmud, K., Mahayani, N.P.D., Mansor,
956 A., Marshall, A.R., Martin, E.H., Matos, D.C.L., Meave, J.A., Melo, F.P.L., Mendoza,
957 Z.H.A., Metali, F., Medjibe, V.P., Metzger, J.P., Metzker, T., Mohandass, D., Munguía-
958 Rosas, M.A., Muñoz, R., Nurtjahy, E., De Oliveira, E.L., Onrizal, Parolin, P., Parren, M.,
959 Parthasarathy, N., Paudel, E., Perez, R., Pérez-García, E.A., Pommer, U., Poorter, L.,
960 Qi, L., Piedade, M.T.F., Pinto, J.R.R., Poulsen, A.D., Poulsen, J.R., Powers, J.S.,
961 Prasad, R.C., Puyravaud, J.-P., Rangel, O., Reitsma, J., Rocha, D.S.B., Rolim, S.,
962 Rovero, F., Rozak, A., Ruokolainen, K., Rutishauser, E., Rutten, G., Mohd Said, M.N.,
963 Saiter, F.Z., Saner, P., Santos, B., Dos Santos, J.R., Sarker, S.K., Schmitt, C.B.,
964 Schoengart, J., Schulze, M., Sheil, D., Sist, P., Souza, A.F., Spironello, W.R., Sposito,
965 T., Steinmetz, R., Stevart, T., Suganuma, M.S., Sukri, R., Sultana, A., Sukumar, R.,
966 Sunderland, T., Supriyadi, Suresh, H.S., Suzuki, E., Tabarelli, M., Tang, J., Tanner,
967 E.V.J., Targhetta, N., Theilade, I., Thomas, D., Timberlake, J., De Morisson Valeriano,
968 M., Van Valkenburg, J., Van Do, T., Van Sam, H., Vandermeer, J.H., Verbeeck, H.,
969 Vetaas, O.R., Adekunle, V., Vieira, S.A., Webb, C.O., Webb, E.L., Whitfeld, T., Wich, S.,
970 Williams, J., Wiser, S., Wittmann, F., Yang, X., Yao, C.Y.A., Yap, S.L., Zahawi, R.A.,
971 Zakaria, R. & Zang, R. (2018) Phylogenetic classification of the world's tropical forests.
972 *Proceedings of the National Academy of Sciences of the United States of America*, **115**.
973 Slik, J.W.F., Paoli, G., McGuire, K., Amaral, I., Barroso, J., Bastian, M., Blanc, L., Bongers,
974 F., Boundja, P., Clark, C., Collins, M., Dauby, G., Ding, Y., Doucet, J.-L., Eler, E.,
975 Ferreira, L., Forshed, O., Fredriksson, G., Gillet, J.-F., Harris, D., Leal, M., Laumonier,
976 Y., Malhi, Y., Mansor, A., Martin, E., Miyamoto, K., Araujo-Murakami, A., Nagamasu, H.,
977 Nilus, R., Nurtjahya, E., Oliveira, Á., Onrizal, O., Parada-Gutierrez, A., Permana, A.,
978 Poorter, L., Poulsen, J., Ramirez-Angulo, H., Reitsma, J., Rovero, F., Rozak, A., Sheil,
979 D., Silva-Espejo, J., Silveira, M., Spironelo, W., ter Steege, H., Stevart, T., Navarro-
980 Aguilar, G.E., Sunderland, T., Suzuki, E., Tang, J., Theilade, I., van der Heijden, G., van
981 Valkenburg, J., Van Do, T., Vilanova, E., Vos, V., Wich, S., Wöll, H., Yoneda, T., Zang,
982 R., Zhang, M.-G. & Zweifel, N. (2013) Large trees drive forest aboveground biomass

983 variation in moist lowland forests across the tropics. *Global Ecology and Biogeography*,
984 **22**, 1261–1271.

985 Stark, S.C., Enquist, B.J., Saleska, S.R., Leitold, V., Schiatti, J., Longo, M., Alves, L.F.,
986 Camargo, P.B. & Oliveira, R.C. (2015) Linking canopy leaf area and light environments
987 with tree size distributions to explain Amazon forest demography. *Ecology Letters*, **18**,
988 636–645.

989 Stark, S.C., Leitold, V., Wu, J.L., Hunter, M.O., de Castilho, C. V., Costa, F.R.C., McMahon,
990 S.M., Parker, G.G., Shimabukuro, M.T., Lefsky, M.A., Keller, M., Alves, L.F., Schiatti, J.,
991 Shimabukuro, Y.E., Brand??o, D.O., Woodcock, T.K., Higuchi, N., de Camargo, P.B.,
992 de Oliveira, R.C. & Saleska, S.R. (2012) Amazon forest carbon dynamics predicted by
993 profiles of canopy leaf area and light environment. *Ecology Letters*, **15**, 1406–1414.

994 Stegen, J.C., Swenson, N.G., Enquist, B.J., White, E.P., Phillips, O.L., Jørgensen, P.M.,
995 Weiser, M.D., Monteagudo Mendoza, A. & Núñez Vargas, P. (2011) Variation in above-
996 ground forest biomass across broad climatic gradients. *Global Ecology and*
997 *Biogeography*, **20**, 744–754.

998 Sullivan, M.J.P., Talbot, J., Lewis, S.L., Phillips, O.L., Qie, L., Begne, S.K., Chave, J., Cuni-
999 Sanchez, A., Hubau, W., Lopez-Gonzalez, G., Miles, L., Monteagudo-Mendoza, A.,
1000 Sonké, B., Sunderland, T., Ter Steege, H., White, L.J.T., Affum-Baffoe, K., Aiba, S.-I.,
1001 De Almeida, E.C., De Oliveira, E.A., Alvarez-Loayza, P., Dávila, E.Á., Andrade, A.,
1002 Aragão, L.E.O.C., Ashton, P., Aymard, G.A., Baker, T.R., Balinga, M., Banin, L.F.,
1003 Baraloto, C., Bastin, J.-F., Berry, N., Bogaert, J., Bonal, D., Bongers, F., Brienen, R.,
1004 Camargo, J.L.C., Cerón, C., Moscoso, V.C., Chezeaux, E., Clark, C.J., Pacheco, Á.C.,
1005 Comiskey, J.A., Valverde, F.C., Coronado, E.N.H., Dargie, G., Davies, S.J., De
1006 Caniere, C., Djuikouo, M.N., Doucet, J.-L., Erwin, T.L., Espejo, J.S., Ewango, C.E.N.,
1007 Fauset, S., Feldpausch, T.R., Herrera, R., Gilpin, M., Gloor, E., Hall, J.S., Harris, D.J.,
1008 Hart, T.B., Kartawinata, K., Kho, L.K., Kitayama, K., Laurance, S.G.W., Laurance, W.F.,
1009 Leal, M.E., Lovejoy, T., Lovett, J.C., Lukasu, F.M., Makana, J.-R., Malhi, Y.,
1010 Maracahipes, L., Marimon, B.S., Junior, B.H.M., Marshall, A.R., Morandi, P.S., Mukendi,

1011 J.T., Mukinzi, J., Nilus, R., Vargas, P.N., Camacho, N.C.P., Pardo, G., Peña-Claros, M.,
1012 Pétronelli, P., Pickavance, G.C., Poulsen, A.D., Poulsen, J.R., Primack, R.B., Priyadi,
1013 H., Quesada, C.A., Reitsma, J., Réjou-Méchain, M., Restrepo, Z., Rutishauser, E.,
1014 Salim, K.A., Salomão, R.P., Samsedin, I., Sheil, D., Sierra, R., Silveira, M., Slik,
1015 J.W.F., Steel, L., Taedoumg, H., Tan, S., Terborgh, J.W., Thomas, S.C., Toledo, M.,
1016 Umunay, P.M., Gamarra, L.V., Vieira, I.C.G., Vos, V.A., Wang, O., Willcock, S. &
1017 Zemagho, L. (2017) Diversity and carbon storage across the tropical forest biome.
1018 *Scientific Reports*, **7**.

1019 Tang, H. & Dubayah, R. (2017) Light-driven growth in Amazon evergreen forests explained
1020 by seasonal variations of vertical canopy structure. *Proceedings of the National*
1021 *Academy of Sciences of the United States of America*, **114**, 2640–2644.

1022 Thomas, R.Q., Kellner, J.R., Clark, D.B. & Peart, D.R. (2013) Low mortality in tall tropical
1023 trees. *Ecology*, **94**, 920–929.

1024 West, G.B., Enquist, B.J., Brown, J.H., West, G.B., Brown, J.H., Enquist, B.J. & Brown, J.H.
1025 (2009) Extensions and evaluations of a general quantitative theory of forest structure
1026 and dynamics. *Proceedings of the National Academy of Sciences of the United States*
1027 *of America*, **106**, 7040–7045.

1028 Zanne, A.E., Lopez-Gonzalez, G., Coomes, D.A., Ilic, J., Jansen, S., Lewis, S.L., Miller, R.B.,
1029 Swenson, N.G., Wiemann, M.C. & Chave, J. (2009) Global wood density database.
1030 *Dryad*, **235**, 33.

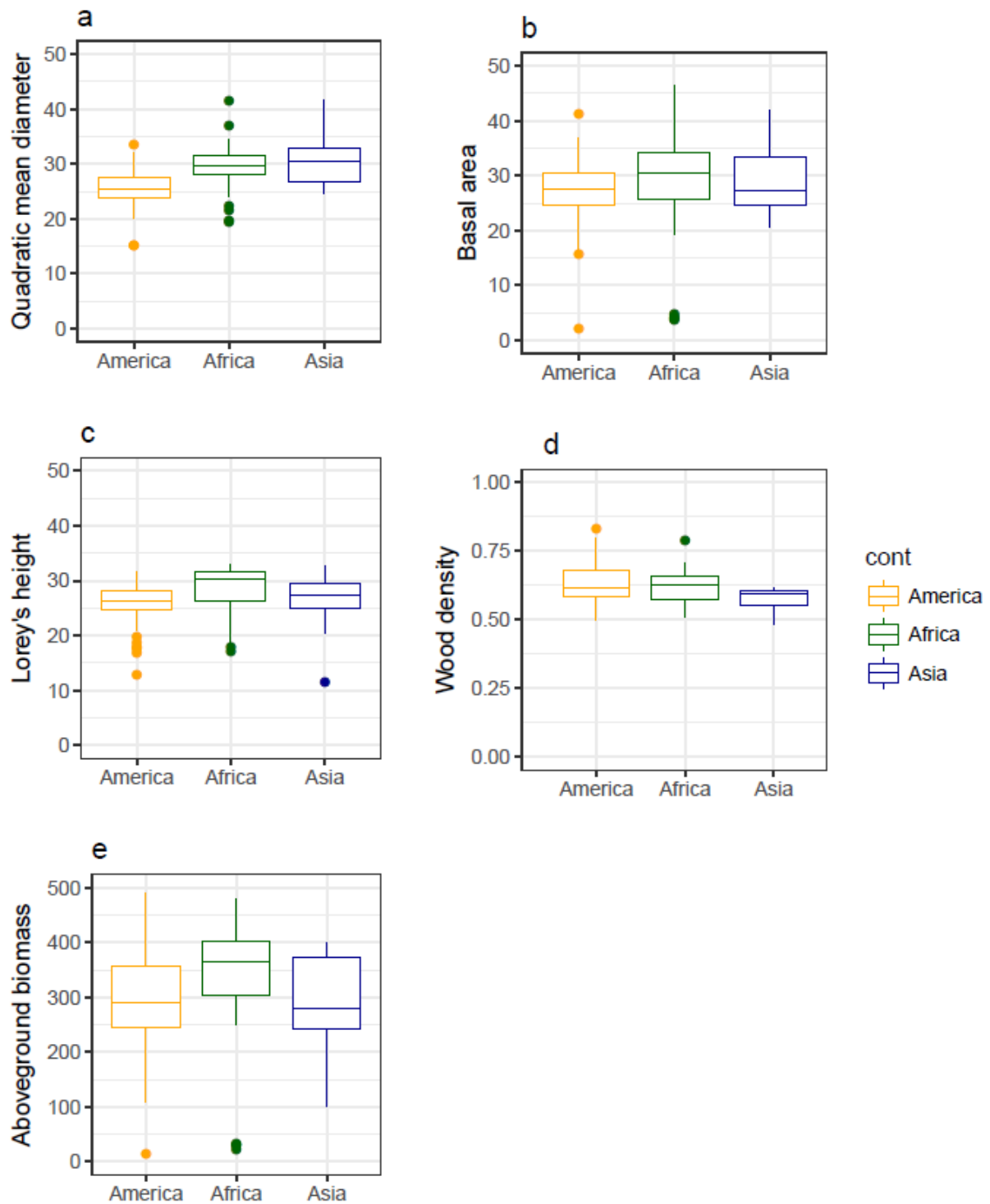
1031

1032 **Supplementary information.**

1033 **Supplementary table 1. Plot, Site and PIs**

1034 **Supplementary table 2. Coefficients of plot level structure prediction from the *ith***

1035 **largest trees.**



1036
1037

Supplementary figure 1. Cross-continent comparison of plot-metrics distribution

1038

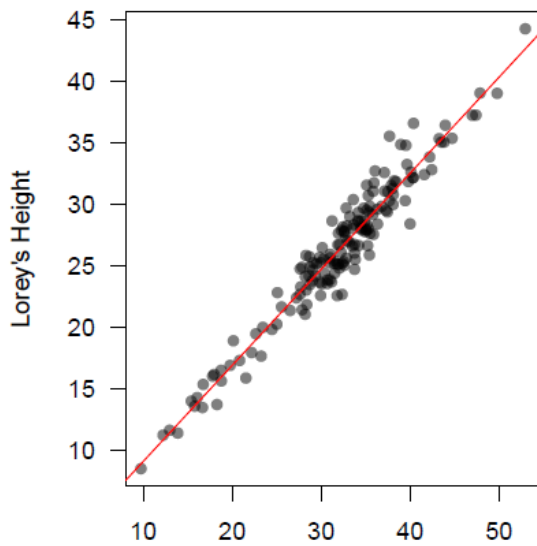
averaged at the site level. Figures illustrates respectively the distribution of the values for

1039

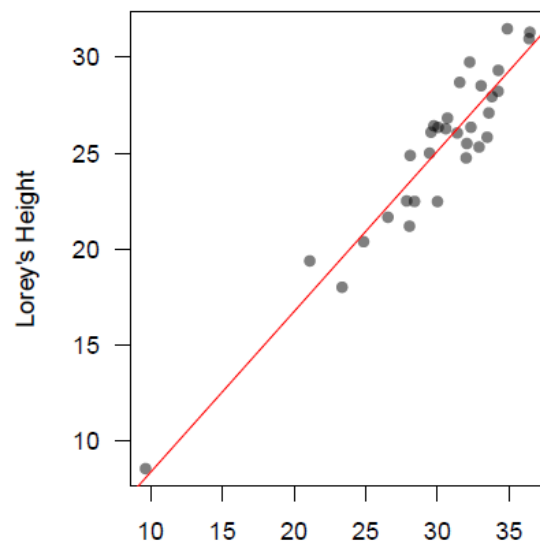
the quadratic mean diameter (a), basal area (b), Lorey's height (c), wood density (d) and

1040

aboveground biomass (e).



Mean Height - 20 largest trees (m) - local HD



Mean Height - 20 largest trees (m) - Observed H

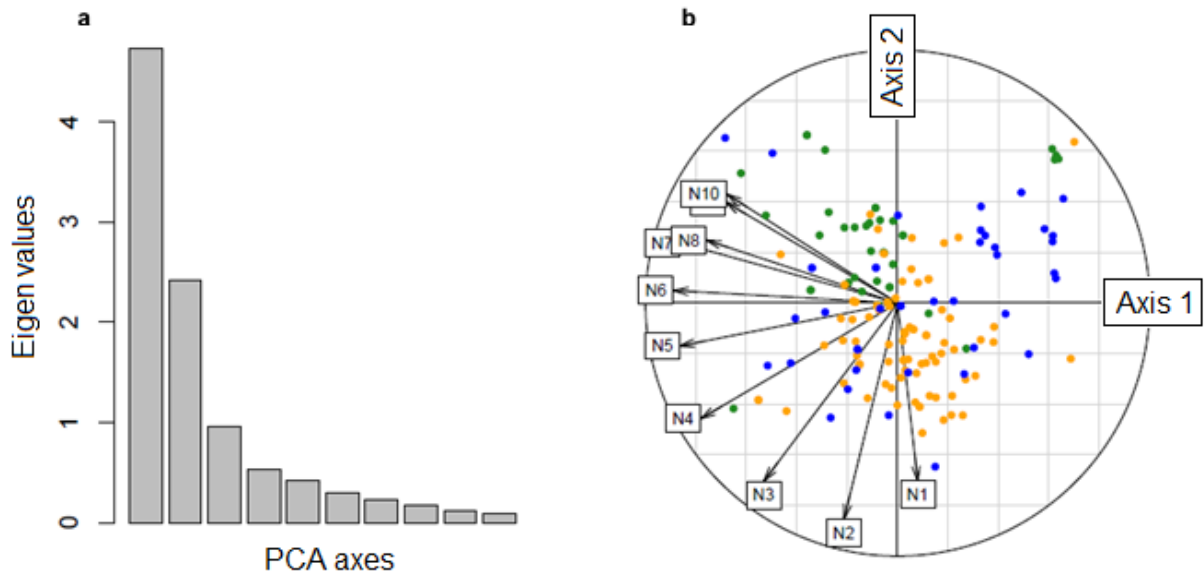
1041

1042 **Supplementary figure 2. Lorey's Height prediction from the 20 largest trees.** Figures

1043 show the results using (i) local D-H allometries for 20 sites (left subfigure) and (ii) using plots

1044 where height is measured on all trees in Malebo site in the Democratic Republic of the Congo

1045 (right subfigure).



1046

1047 **Supplementary figure 3. PCA on the diameter structure and corresponding mean**

1048 **distribution for high contributions of axis 1 and axis 2.** (A) Illustration of top and low

1049 percentile observed for each axis, with diameter distributions represented as the relative

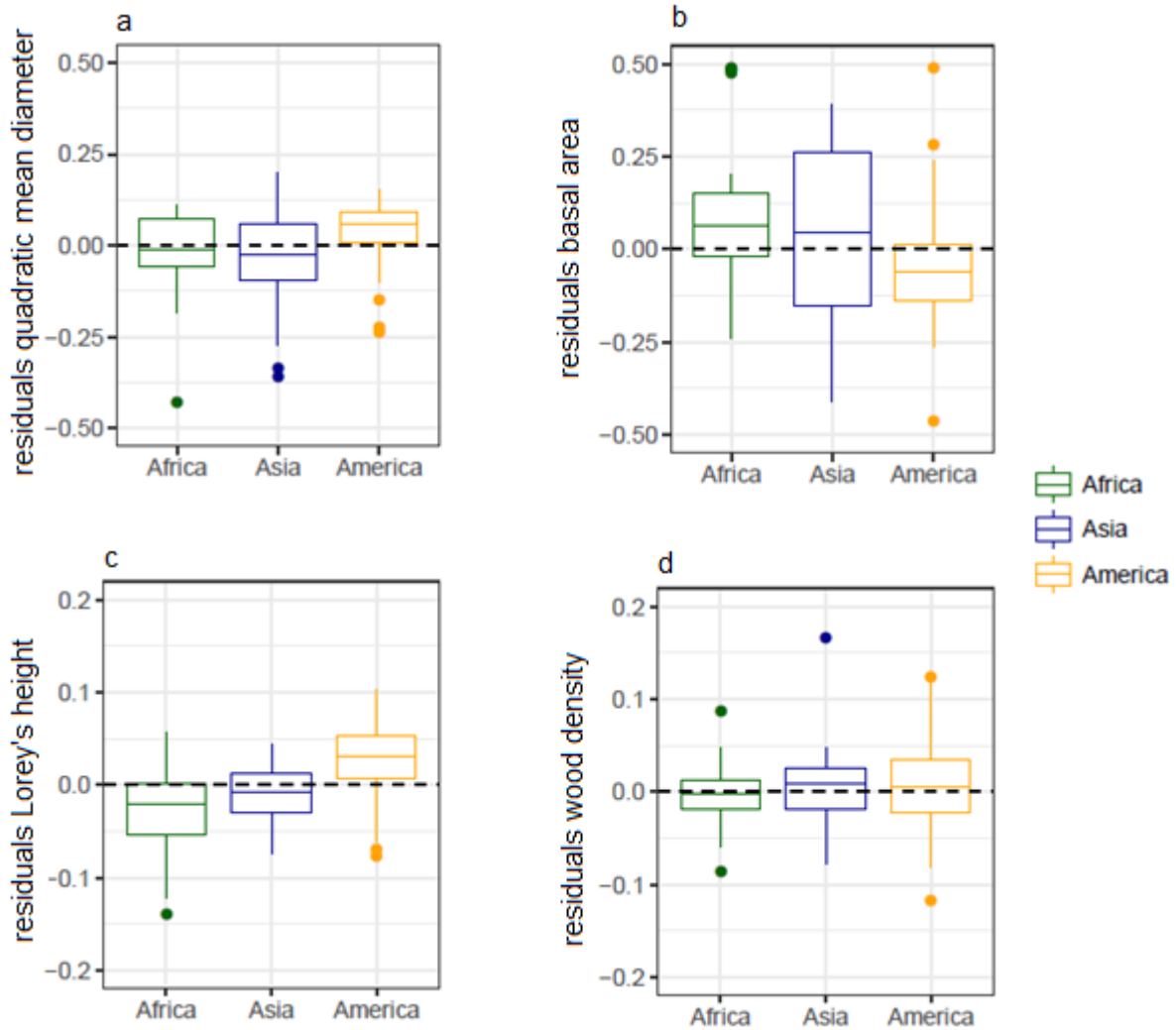
1050 difference with the average observed distribution.(B) Biplot with contribution to the PCA of all

1051 the diameter classes, with the respective position of each site in the space defined by axis1

1052 and 2. Axis 1 is driven by differences in global abundance of trees and axis 2 is driven by a

1053 difference of balance between abundance of small vs. large trees. Colors represent continent,

1054 with Africa, America and Asia respectively in green, orange and blue.



1055
1056

Supplementary figure 4. Cross-continent comparison of the relative residuals from the

1057

prediction of plot-metrics from the 20 largest trees. The relative residuals are generally

1058

low (<10%). Systematic small differences can however be found in America, where the

1059

quadratic mean diameter and Lorey's height tend to be slightly overestimated and the basal

1060

area slightly underestimated.