Received: 1 November 2017 | Revised: 31 May 2018 | Accepted: 13 June 2018

DOI: 10.1111/geb.12803

RESEARCH PAPER

WILEY Global Ecology and Biogeography A Journal of Macroecology

Pan-tropical prediction of forest structure from the largest trees

Jean-François Bastin[1,2,3,4] | Ervan Rutishauser[4,5] | James R. Kellner[6,7] | Sassan Saatchi[8] | Raphael Pélissier[9] | Bruno Hérault[10,11] | Ferry Slik[12] | Jan Bogaert[13] | Charles De Cannière[3] | Andrew R. Marshall[14,15,16] | John Poulsen[17] | Patricia Alvarez-Loyayza[18] | Ana Andrade[19] | Albert Angbonga-Basia[20] | Alejandro Araujo-Murakami[21] | Luzmila Arroyo[22] | Narayanan Ayyappan[23,24] | Celso Paulo de Azevedo[25] | Olaf Banki[26] | Nicolas Barbier[9] | Jorcely G. Barroso[26] | Hans Beeckman[27] | Robert Bitariho[28] | Pascal Boeckx[29] | Katrin Boehning-Gaese[30,31] | Hilandia Brandão[32] | Francis Q. Brearley[33] | Mireille Breuer Ndoundou Hockemba[34] | Roel Brienen[35] | Jose Luis C. Camargo[19] | Ahimsa Campos-Arceiz[36] | Benoit Cassart[37,38] | Jérôme Chave[39] | Robin Chazdon[40] | Georges Chuyong[41] | David B. Clark[42] | Connie J. Clark[17] | Richard Condit[43] | Euridice N. Honorio Coronado[44] | Priya Davidar[22] | Thalès de Haulleville[13,27] | Laurent Descroix[45] | Jean-Louis Doucet[13] | Aurelie Dourdain[46] | Vincent Droissart[9] | Thomas Duncan[47] | Javier Silva Espejo[48] | Santiago Espinosa[49] | Nina Farwig[50] | Adeline Fayolle[13] | Ted R. Feldpausch[51] | Antonio Ferraz[8] | Christine Fletcher[36] | Krisna Gajapersad[52] | Jean-François Gillet[13] | Iêda Leão do Amaral[32] | Christelle Gonmadje[53] | James Grogan[54] | David Harris[55] | Sebastian K. Herzog[56] | Jürgen Homeier[57] | Wannes Hubau[27] | Stephen P. Hubbell[5,58] | Koen Hufkens[29] | Johanna Hurtado[59] | Narcisse G. Kamdem[60] | Elizabeth Kearsley[61] | David Kenfack[62] | Michael Kessler[63] | Nicolas Labrière[10,64] | Yves Laumonier[10,65] | Susan Laurance[66] | William F. Laurance[66] | Simon L. Lewis[35] | Moses B. Libalah[60] | Gauthier Ligot[13] | Jon Lloyd[67,68] | Thomas E. Lovejoy[68] | Yadvinder Malhi[69] | Beatriz S. Marimon[70] | Ben Hur Marimon Junior[70] | Emmanuel H. Martin[71,72] | Paulus Matius[73] | Victoria Meyer[8] | Casimero Mendoza Bautista[74] | Abel Monteagudo-Mendoza[75] | Arafat Mtui[76] | David Neill[77] | Germaine Alexander Parada Gutierrez[21] | Guido Pardo[78] | Marc Parren[79] | N. Parthasarathy[23] | Oliver L. Phillips[35] | Nigel C. A. Pitman[79] | Pierre Ploton[9] | Quentin Ponette[37] | B. R. Ramesh[23] | Jean-Claude Razafimahaimodison[80] | Maxime Réjou-Méchain[9] |

Samir Gonçalves Rolim[81] | Hugo Romero Saltos[82] | Luiz Marcelo Brum Rossi[81] | Wilson Roberto Spironello[32] | Francesco Rovero[76] | Philippe Saner[83] | Denise Sasaki[84] | Mark Schulze[85] | Marcos Silveira[86] | James Singh[87] | Plinio Sist[10,88] | Bonaventure Sonke[60] | J. Daniel Soto[21] | Cintia Rodrigues de Souza[24] | Juliana Stropp[89] | Martin J. P. Sullivan[35] | Ben Swanepoel[34] | Hans ter Steege[25,90] | John Terborgh[91,92] | Nicolas Texier[93] | Takeshi Toma[94] | Renato Valencia[95] | Luis Valenzuela[75] | Leandro Valle Ferreira[96] | Fernando Cornejo Valverde[97] | Tinde R. Van Andel[25] | Rodolfo Vasque[77] | Hans Verbeeck[61] | Pandi Vivek[22] | Jason Vleminckx[98] | Vincent A. Vos[78,99] | Fabien H. Wagner[100] | Papi Puspa, Warsudi[101] | Verginia Wortel[102] | Roderick J. Zagt[103] | Donatien Zebaze[60]

[1]Department of Environmental Systems Science, Institute of Integrative Biology, ETH Zürich, Zürich, Switzerland

[2] Carbon Cycle and Ecosystem Group VS, NASA, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

[3]Landscape Ecology and Plant Production System, Université libre de Bruxelles. CP264-2, Bruxelles, Belgium

[4]Carboforexpert (carboforexpert.ch), Hermance, Switzerland

[5]Smithsonian Tropical Research Institute, Balboa, Panama

[6]Department of Ecology and Evolutionary Biology, Brown University, Providence, Rhode Island

[7]Institute at Brown for Environment and Society, Brown University, Providence, Rhode Island

[8]NASA, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

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

[10]Cirad, UR Forest & Societies, Montpellier, France

[11]INPHB (Institut National Polytechnique Félix Houphouet Boigny), Yamoussoukro, Ivory Coast

[12]Faculty of Science, Universiti Brunei Darusallam, Gadong, Brunei Darussalam

[13]Gembloux Agro-Bio Tech, Université de Liège, Gembloux, Belgium

[14]CIRCLE, Environment Department, Wentworth Way, University of York, York, United Kingdom

[15]Tropical Forests and People Research Centre, University of the Sunshine Coast, Sunshine Coast, Queensland, Australia

[16]Flamingo Land Ltd, Kirby Misperton, United Kingdom

[17]Nicholas School of the Environment, Duke University, Durham, North Carolina, USA

[18]Field Museum of Natural History, Chicago, Illinois, USA

[19]Biological Dynamics of Forest Fragment Project (BDFFP – INPA/STRI), Manaus, Brazil

[20]Institut Facultaire des Sciences Agronomiques de Yangambi, Yangambi, Democratic Republic of the Congo

[21]Museo de Historia Natural Noel Kempff Mercado, Santa Cruz, Bolivia

[22]Department of Ecology and Environmental Sciences, Pondicherry University, Pondicherry, India

[23]French Institute of Pondicherry (IFP), Pondicherry, India

[24]Embrapa Amazônia Ocidental, Manaus, Brazil

[25]Naturalis Biodiversity Centre, Leiden, The Netherlands

[26]Universidade Federal do Acre, Campus Floresta, Acre, Brazil

[27]Service of Wood Biology, Royal Museum for Central Africa, Tervuren, Belgium

[28]Institute of Tropical Forest Conservation, Mbarara University of Science and Technology, Mbarara, Uganda

[29]Isotope Bioscience Laboratory – ISOFYS, Ghent University, Ghent, Belgium

[30]Senckenberg Biodiversity and Climate Research Centre (BiK-F), Frankfurt am Main, Germany

[31]Department of Biological Sciences, Goethe Universität, Frankfurt am Main, Germany

[32]National Institute for Amazonian Research (INPA), Manaus, Brazil

[33]School of Science and the Environment, Manchester Metropolitan University, Manchester, United Kingdom

[34]Wildlife Conservation Society, New York, New York, USA

[35]School of Geography, University of Leeds, Leeds, United Kingdom

[36]Malaysia Campus, Jalan Broga, Semenyih, Malaysia
[37]UCL-ELI, Earth and Life Institute, Université catholique de Louvain, Louvain-la-Neuve, Belgium
[38]Ecole Régionale Post-universitaire d'Aménagement et de Gestion Intégrés des Forêts et Territoires Tropicaux, Kinshasa, Democratic Republic of the Congo
[39]Laboratoire Evolution et Diversité biologique, CNRS & Université Paul Sabatier, Toulouse, France
[40]Department of Ecology and Evolutionary Biology, University of Connecticut, Storrs, Connecticut, USA
[41]Department of Botany and Plant Physiology, University of Buea, Buea, Cameroon
[42]Department of Biology, University of Missouri-St Louis, St Louis, Missouri, USA
[43]Field Museum of Natural History and Morton Arboretum, Chicago, Illinois, USA
[44]Coronado, Inst. de Investigaciones de la Amazonia Peruana, Iquitos, Peru
[45]ONF pôle R&D, Cayenne, France
[46]Cirad, UMR EcoFoG (AgroParisTech, CNRS, Inra, Universite des Antilles, Universite de la Guyane), Kourou, French Guiana
[47]Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon
[48]Departamento de Biología, Universidad de La Serena, La Serena, Chile
[49]Universidad Autónoma de San Luis Potosí, San Luis Potosí, México
[50]Department of Conservation Ecology, Philipps-Universität Marburg, Marburg, Germany
[51]Geography, College of Life and Environmental Sciences, University of Exeter, Exeter, United Kingdom
[52]Conservation International Suriname, Paramaribo, Suriname
[53]Faculty of Science, Department of Plant Biology, University of Yaounde I, Yaoundé, Cameroon
[54]Smith College Botanic Garden, Northampton, Massachusetts
[55]Royal Botanic Garden Edinburgh, Edinburgh, United Kingdom
[56]Museo de Historia Natural Alcide d'Orbigny, Cochabamba, Bolivia
[57]Plant Ecology, University of Goettingen, Goettingen, Germany
[58]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, California, USA
[59]Organization for Tropical Studies, La Selva, Costa Rica
[60]Plant Systematic and Ecology Laboratory, Higher Teacher's Training College, University of Yaoundé I, Yaoundé, Cameroon
[61]CAVElab – Computational and Applied Vegetation Ecology, Ghent University, Ghent, Belgium
[62]CTFS-ForestGEO, Smithsonian Tropical Research Institute, Washington, USA
[63]Department of Systematic and Evolutionary Botany, University of Zurich, Zurich, Switzerland
[64]AgroParisTech, Doctoral School ABIES, Paris, France
[65]Center for International Forestry Research, Bogor, Indonesia
[66]Centre for Tropical Environmental and Sustainability Science, College of Science and Engineering, James Cook University, Cairns, Queensland, Australia
[67]Department of Life Sciences, Imperial College London, Ascot, United Kingdom
[68]Department of Environmental Science and Policy, George Mason University, Fairfax, Virginia
[69]Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, United Kingdom
[70]Universidade do Estado de Mato Grosso, Campus de Nova Xavantina, Nova Xavantina, Brazil
[71]Udzungwa Ecological Monitoring Centre, Udzungwa Mountains National Park, Mang'ula, Tanzania
[72]Sokoine University of Agriculture, Morogoro, Tanzania
[73]Faculty of Forestry, Mulawarman University, Samarinda, Indonesia
[74]Escuela de Ciencias Forestales, Unidad Académica del Trópico, Universidad Mayor de San Simón, Sacta, Bolivia
[75]Jardín Botánico de Missouri, Oxapampa, Peru
[76]MUSE - Museo delle Scienze, Trento, Italy
[77]Universidad Estatal Amazónica, Puyo, Pastaza, Ecuador
[78]Universidad Autónoma del Beni, Riberalta, Bolivia
[79]Science and Education, The Field Museum, Chicago, Illinois
[80]Centre ValBio, Ranomafana, Madagascar
[81]Embrapa Florestas, Colombo, Brazil
[82]School of Biological Sciences and Engineering, Yachay Tech University, Urcuquí, Ecuador
[83]Department of Evolutionary Biology and Environmental Studies, University of Zurich, Zurich, Switzerland
[84]Fundação Ecológica Cristalino Alta Floresta, Mato Grosso, Brazil
[85]H. J. Andrews Experimental Forest, Blue River, Oregon
[86]Museu Universitário, Universidade Federal do Acre, Rio Branco, Brazil

[87]Guyana Forestry Commission, Georgetown, Guyana
[88]Forests and Societies, Univ. Montpellier, CIRAD, Montpellier, France
[89]Institute of Biological and Health Sciences, Federal University of Alagoas, Maceió, Brazil
[90]Systems Ecology, Free University, Amsterdam, The Netherlands
[91]Florida Museum of Natural History and Department of Biology, University of Florida – Gainesville, Gainesville, Florida
[92]Department of Biology, James Cook University, Cairns, Queensland, Australia
[93]Laboratoire d'Evolution Biologique et Ecologie, Faculté des Sciences, Université libre de Bruxelles, Bruxelles, Belgium
[94]Forestry and Forest Products Research Institute, Tsukuba, Japan
[95]Escuela de Ciencias Biológicas, Pontificia Universidad Católica del Ecuador, Quito, Ecuador
[96]Coordenação de Botânica, Museu Paraense Emilio Goeldi, Belém, Brazil
[97]Andes to Amazon Biodiversity Program, Madre de Dios, Peru
[98]Department of Integrative Biology, University of California, Berkeley, California
[99]Centro de Investigación y Promoción del Campesinado – Norte Amazónico, Riberalta, Bolivia
[100]Remote Sensing Division, National Institute for Space Research – INPE, São José dos Campos, Brazil
[101]The Center for Reforestation Studies in the Tropical Rain Forest (PUSREHUT), Mulawarman University, East Kalimantan, Indonesia
[102]Center for Agricultural Research in Suriname (CELOS), Paramaribo, Suriname
[103]Tropenbos International, Wageningen, The Netherlands

Correspondence
Jean-François Bastin, Institute of Integrative Biology, Department of Environmental Systems Science, ETH Zürich, 8092 Zürich, Switzerland.
Email: bastin.jf@gmail.com

Abstract

Aim: Large tropical trees form the interface between ground and airborne observations, offering a unique opportunity to capture forest properties remotely and to investigate their variations on broad scales. However, despite rapid development of metrics to characterize the forest canopy from remotely sensed data, a gap remains between aerial and field inventories. To close this gap, we propose a new pan-tropical model to predict plot-level forest structure properties and biomass from only the largest trees.

Location: Pan-tropical.

Time period: Early 21st century.

Major taxa studied: Woody plants.

Methods: Using a dataset of 867 plots distributed among 118 sites across the tropics, we tested the prediction of the quadratic mean diameter, basal area, Lorey's height, community wood density and aboveground biomass (AGB) from the *i*th largest trees.

Results: Measuring the largest trees in tropical forests enables unbiased predictions of plot- and site-level forest structure. The 20 largest trees per hectare predicted quadratic mean diameter, basal area, Lorey's height, community wood density and AGB with 12, 16, 4, 4 and 17.7% of relative error, respectively. Most of the remaining error in biomass prediction is driven by differences in the proportion of total biomass held in medium-sized trees (50–70 cm diameter at breast height), which shows some continental dependency, with American tropical forests presenting the highest proportion of total biomass in these intermediate-diameter classes relative to other continents.

Main conclusions: Our approach provides new information on tropical forest structure and can be used to generate accurate field estimates of tropical forest carbon stocks to support the calibration and validation of current and forthcoming space missions. It will reduce the cost of field inventories and contribute to scientific understanding of tropical forest ecosystems and response to climate change.

KEYWORDS
carbon, climate change, forest structure, large trees, pan-tropical, REDD+, tropical forest ecology

1 | INTRODUCTION

The fundamental ecological function of large trees is well established for tropical forests. They offer shelter to many organisms (Lindenmayer, Laurance, & Franklin, 2012; Remm & Lõhmus, 2011), regulate forest dynamics, regeneration (Harms, Wright, Calderón, Hernández, & Herre, 2000; Rutishauser, Wagner, Herault, Nicolini, & Blanc, 2010) and total biomass (Stegen et al., 2011), and are important contributors to the global carbon cycle (Meakem et al., 2018). Being major components of the canopy, the largest trees may also suffer more than sub-canopy and understorey trees from climate change, because they are directly exposed to variations in solar radiation, wind strength, temperature seasonality and relative air humidity (Bennett, McDowell, Allen, & Anderson-Teixeira, 2015; Laurance, Delamônica, Laurance, Vasconcelos, & Lovejoy, 2000; Lindenmayer et al., 2012; Meakem et al., 2018; Nepstad, Tohver, Ray, Moutinho, & Cardinot, 2007; Thomas, Kellner, Clark, & Peart, 2013). Given that they are visible from the sky, large trees are ideal for monitoring forest responses to climate change via remote sensing (RS; Asner et al., 2017; Bennett et al., 2015).

Large trees encompass a disproportionate fraction of total aboveground biomass (AGB) in tropical forests (Chave, Riera, Dubois, & Riéra, 2001; Lutz et al., 2018), with some variations in their relative contribution to the total AGB among the tropical regions (Feldpausch et al., 2012). In Central Africa, the largest 5% of trees in a forest sample plot (i.e., the 5% of trees with the largest diameter at 130 cm) store 50% of forest plot AGB on average (Bastin et al., 2015). Consequently, the density of large trees largely explains variation in forest AGB at local (Clark & Clark, 1996), regional (Malhi et al., 2006; Saatchi, Houghton, Dos Santos Alvalá, Soares, & Yu, 2007) and continental scales (Slik et al., 2013; Stegen et al., 2011). Detailing the contribution of each single tree to the diameter structure, we showed previously that plot-level AGB can be predicted from a few large trees (Bastin et al., 2015), with the measurement of the 20 largest trees per hectare being sufficient to estimate plot-level biomass with < 15% error in reference to ground estimates. These findings suggested that a substantial gain of cost-effectiveness might be achieved by focusing forest inventories on the largest trees rather than all size classes. Likewise, it suggested that RS approaches could focus on the measurement of the largest trees, instead of properties of the entire forest stand.

Several efforts are underway to close the gap between RS of forest biomass and field surveys (Coomes et al., 2017; Jucker et al., 2017). However, existing RS approaches typically require ground measurement of all trees ≥ 10 cm in diameter (*D*) for calibration (Asner & Mascaro, 2014; Asner et al., 2012). Collecting such data in the field is costly and time consuming, which therefore limits the spatial representativeness of available plot networks. Besides, extrapolation methods of ground-based biomass estimations on RS data still face important limits. For instance, using mean canopy height extracted from active sensors (Ho Tong Minh et al.., 2016; Mascaro, Detto, Asner, & Muller-Landau, 2011) or canopy grain derived from optical images (Bastin et al., 2014; Ploton et al., 2017; Ploton, Pélissier, & Proisy, 2012; Proisy, Couteron, & Fromard, 2007), the biomass is predicted with an error of only 10–20% compared with ground-based estimates. However, this good level of accuracy is limited to the extent of the RS scene used, which decreases considerably in the upscaling step necessary for national or global maps (Xu et al., 2017). A promising development to alleviate this spatial restriction lies in the 'universal approach', proposed by Asner et al. (2012) and further adapted by Asner and Mascaro (2014), in which plot-level biomass is predicted by a linear combination of ground-based and remotely sensed metrics. The 'universal approach' relies upon canopy height metrics derived from radar or LiDAR (top of canopy height, TCH), and basal area (BA; i.e., the cumulated cross-sectional area of the stems) and community wood density (i.e., weighted by basal area, WD_{BA}) derived from field inventories. Plot AGB is then predicted as follows (Asner et al., 2012):

$$AGB = aTCH^{b1}BA^{b2}WD_{BA}^{b3} \quad (1)$$

Although generally performing better than approaches based solely on RS of tree height (Coomes et al., 2017), this model relies on exhaustive ground measurements (i.e., wood density and the basal area of all trees > 10 cm in diameter at 130 cm, neither of which is measured using any existing remotely sensed data).

Recent advances in RS allow the identification of single trees in the canopy (Ferraz, Saatchi, Mallet, & Meyer, 2016), estimation of adult mortality rates for canopy tree species (Kellner & Hubbell, 2017), description of the forest diameter structure (Stark et al., 2015), depiction of crown and gap shapes (Coomes et al., 2017) and even the identification of some functional traits of canopy species (Asner et al., 2017). Given that routine retrieval of some canopy tree metrics is within reach, in the present study we test the capacity of the largest trees (i.e., trees that can potentially be derived using RS), to predict plot-level biomass. To this end, we adapted Equation (1) as follows:

$$AGB = a(Dg_{LTi}H_{LTi}WD_{LTi})^{b1} \quad (2)$$

where for the *i*th largest trees, Dg_{LT} is the quadratic mean diameter, H_{LT} the mean height, and WD_{LT} the mean wood density among the *i*th largest trees.

Using a large database of forest inventories gathered across the tropics (Figure 1), including secondary and old-growth forest plots, we test the ability of the largest trees to predict various metrics estimated at 1-ha plot level, namely the mean quadratic diameter, the basal area, the Lorey's height (i.e., plot-average height weighted by basal area), the community wood density (i.e., plot-average wood density weighted by basal area) and mean aboveground live biomass (Supporting Information Figure 1). By testing different numbers of largest trees as predictors, we aim to propose a threshold for the minimal number of largest trees required to predict forest plot metrics at a pan-tropical level with no bias and low uncertainty (i.e., error < 20%). Although previous work focused on estimating biomass in Central African forests (Bastin et al., 2015), the present study aims at generalizing the

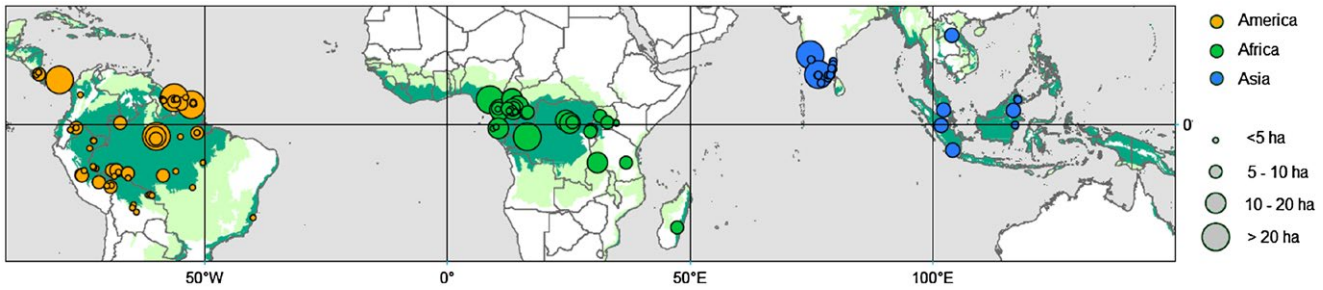

FIGURE 1 Geographical distribution of the plot database. We used 867 plots of 1 ha from 118 sites. Dots are coloured according to floristic affinities (Slik et al., 2015), with America, Africa and Asia in orange, green and blue, respectively. They are also sized according the total area surveyed in each site. In the background, moist forests are displayed in dark green and dry forest in light green

potential of large trees to predict these different plot metrics at continental and pan-tropical scales. Taking advantage of a unique dataset gathered across the tropics (867 1-ha plots), we also investigate major differences in forest structure across the three main tropical regions: the Americas, Africa and Asia. We further discuss how this approach can be used to guide innovative RS techniques and increase the frequency and representativeness of ground data to support global calibration and validation of current and planned space missions. These include the NASA Global Ecosystem Dynamics Investigation (GEDI), NASA-ISRO Synthetic Aperture Radar (NISAR) and ESA P-band radar (BIOMASS) (Dubayah et al., 2014; Le Toan et al., 2011). This study is a step forwards in bringing together RS and field sampling techniques for quantification of terrestrial carbon stocks in tropical forests.

2 | MATERIALS AND METHODS

2.1 | Database

For this study, we compiled standard forest inventories conducted in 867 1-ha plots from 118 sites across the three tropical regions (Figure 1), including mature and secondary forests. Each site comprises all the plots in a given geographical location (i.e., within a 10-km radius and collected by a principal investigator and their team). These consisted of 389 plots in America (69 sites), 302 plots in Africa (35 sites) and 176 plots in Asia (14 sites). Data were provided by principal investigators (see Supporting Information Table 1) and through datasets available on the following networks: TEAM, CTFS (www.forestgeo.si.edu/; Condit et al., 2012) and ForestPlots (https://www.forestplots.net/) for AfriTRON (the African Tropical Rainforest Observation Network; www.afritron.org) and RAINFOR (the Amazon forest inventory network; networks.

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

$$\mathrm{Ln}(H) = 0.893 - E + 0.760\ln(D) - 0.0340\ln(D)^2 \quad (3)$$

where *D* is the diameter at 130 cm and *E* is a measure of environmental stress (Chave et al., 2014). For sites with tree height measurements ($n = 20$), we developed local *D*–*H* models, using a Michaelis–Menten function (Molto et al., 2014). We used these local models to validate the predicted Lorey's height (i.e., plot-average height weighted by BA) from the largest trees, of which height has been estimated with a generic *H*–*D* model [Equation (3), Chave et al., 2014].

We estimated plot biomass as the sum of the biomass of live tree with diameter at 130 cm of ≥ 10 cm, using the following pan-tropical allometric model (Réjou-Méchain, Tanguy, Piponiot, Chave, & Hérault, 2017):

$$\mathrm{AGB} = \exp\left\{ -2.024 - 0.896E + 0.920\ln(\mathrm{WD}) + 2.795\ln(D) - 0.0461\left[\ln\left(\mathrm{D}^2\right)\right] \right\} \quad (4)$$

2.2 | Plot-level metric estimation from the largest trees

The relationship between each plot metric, namely basal area (BA), the quadratic mean diameter (*D*g), Lorey's height (H_{BA}; the mean height weighted by the basal area) and the community wood density (WD_{BA}; the mean wood density weighted by the basal area), and

those derived from largest trees was determined using an iterative procedure following Bastin et al. (2015). Trees were first ranked by decreasing diameter in each plot. An incremental procedure (i.e., including a new tree at each step) was used to sum or average information of the *i* largest trees for each plot metric. Each plot-level metric was predicted by the respective metric derived from the *i*th largest trees. For each increment, the ability (goodness of fit) of the *i* largest trees to predict a given plot metric was tested through a linear regression. To avoid overfitting, a leave-one-out procedure was used to develop independent site-specific models (n = 118). Specifically, the model to be tested at a site was developed with data from all other sites. Errors were then estimated as the relative root mean square error (rRMSE) computed between observed and predicted values (X):

$$\mathrm{rRMSE} = \sqrt{\sum \frac{(\mathrm{Xobs} - \mathrm{Xpred})^2}{n}} / -X \quad (5)$$

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

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

$$\mathrm{BA} = b_1 \Sigma \mathrm{BA}_i^{\gamma 1} \quad (6)$$

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

$$Dg = a_2 + b_2 Dg_i \quad (7)$$

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

$$\mathrm{WD}_{BA} = a_4 + b_4 \overline{\mathrm{WD}_i} \quad (9)$$

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

The quality of the predictions of plot-level metrics from the largest trees is quantified using the rRMSE between measured and predicted values and displayed along the cumulated number of largest trees. Model coefficients are estimated for each metric derived from the largest trees (N_{LT}) and averaged across the 118 models (see Supporting Information Table 2).

Mean rRMSE is plotted as a continuous variable, and its variation is presented as a continuous area between the 5th and the 95th percentiles of observed rRMSE.

2.3 | The optimal number of largest trees for plot-level biomass estimation

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

2.4 | Investigating residuals: What the largest trees do not explain

To understand the limits of predicting AGB through N_{LT}, we also investigated the relationship between AGB residuals and key structural and environmental variables using linear modelling. Forest structure was investigated through the total stem density (*N*), the quadratic mean diameter (*D*g), Lorey's height (H_{BA}) and community wood density (WB_{BA}). As environmental data, we used the mean annual rainfall and the mean temperature computed over the last 10 years at each site using the Climate Research Unit data (New et al., 1999; New, Lister, Hulme, & Makin, 2002), along with rough information on soil types (Carré, Hiederer, Blujdea, & Koeble, 2010). Major soil types were computed from the soil classification of the Harmonized World Soil Database into IPCC (Intergovernmental Panel on Climate Change) soil classes. In addition, considering observed differences in forest structure across tropical continents (Feldpausch et al., 2011, 2012) and recent results on pan-tropical floristic affinities (Slik et al., 2015), we tested for an effect of continent (America, Africa and Asia) on the AGB residuals. Differences in forest structure and AGB among continents were also illustrated through the analysis of their distribution.

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

3 | RESULTS

3.1 | Plot-level metrics

Plot metrics averaged at the site level (867 plots, 118 sites) present important variations within and between continents. In our

database, the quadratic mean diameter varies from 15 to 42 cm^2/ha, the basal area from 2 to 58 m^2/ha, Lorey's height from 11 to 33 m, and the wood density weighted by the basal area from 0.48 to 0.84 g/cm^3 (Supporting Information Figure 1). Such important differences between minimal and maximal values are observed because our database covers sites with various forest types, from young forest colonizing savannas to old-growth forest. However, most of our sites are found in mature forests, as shown by the relatively high average and median value of each plot metric (average AGB = 302 Mg/ha; Supporting Information Figure 1). In general, highest values of AGB are found in Africa, driven by the highest values of basal area and highest estimations of Lorey's height. The highest values of wood density weighted by basal area are found in America.

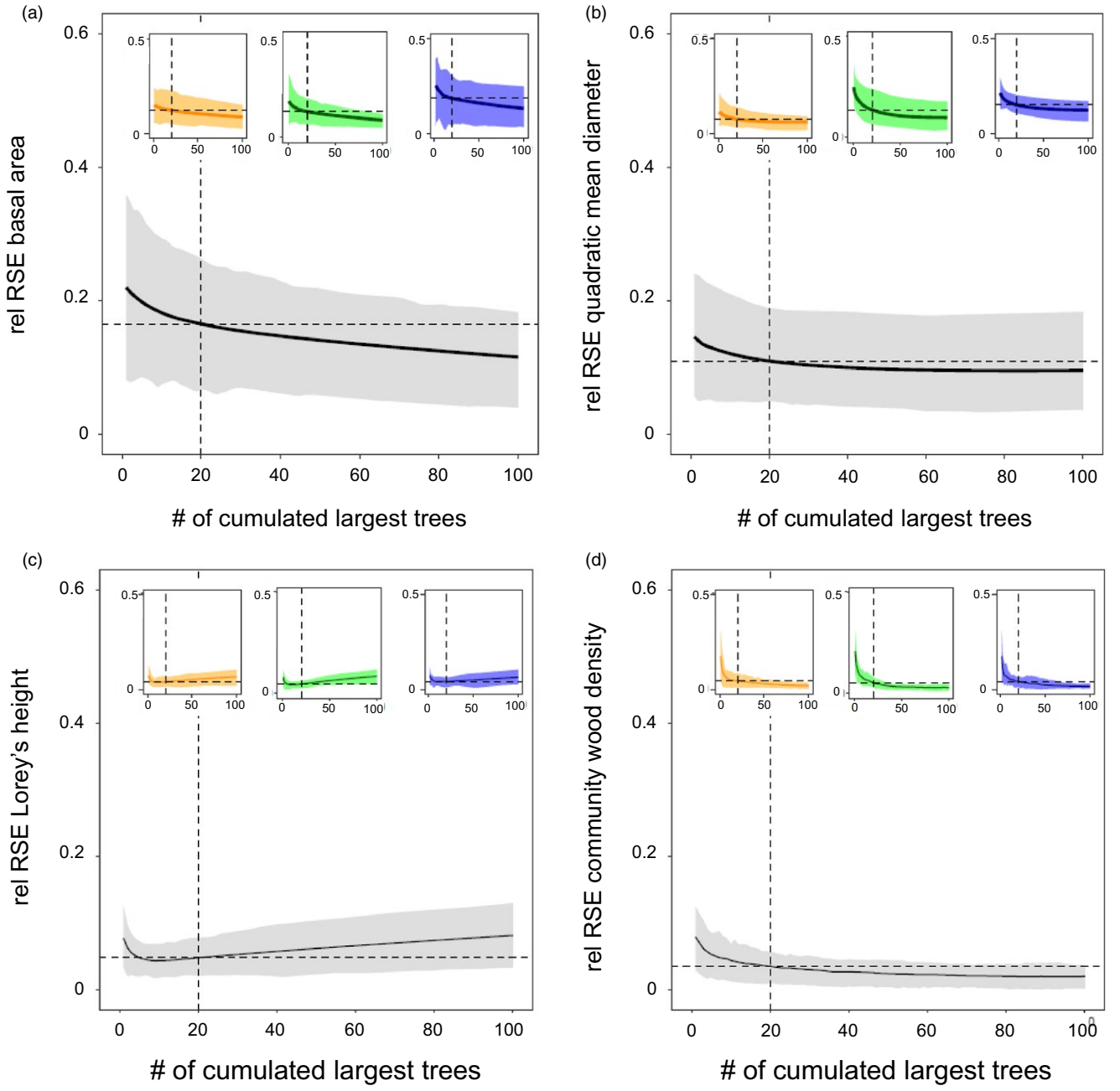

FIGURE 2 Quality of the prediction of plot metrics from largest trees. Variation of the relative root mean square error (rRMSE) of the prediction of plot metric from *i* largest trees versus the cumulative number of largest trees for: (a) basal area, (b) quadratic mean diameter, (c) Lorey's height, and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

3.2 | Plot-level estimation from the *i* largest trees

Overall, plot metrics at the 1-ha scale were well predicted by the largest trees, with qualitative agreement among global and continental models (Figure 2). When using the 20 largest trees to predict basal area (BA) and quadratic mean diameter (*D*g), the mean rRMSE was < 16 and 12%, respectively (Figure 3a,b). Lorey's height (H_{BA}) and wood density weighted by basal area (WD_{BA}) were even better predicted (Figure 3c,d), with mean rRMSE of 4% for the 20 largest trees. The prediction of Lorey's height from the largest trees using the local diameter–height model (Supporting Information Figure 2a) yielded results similar to those obtained using Equation (3) of Chave et al. (2014). More importantly, it also yielded similar results to the prediction of Lorey's height from the largest trees using plots where all the trees were measured on the ground (Supporting Information Figure 2b). This suggests that our conclusions are robust to the uncertainty introduced by height–diameter allometric models.

3.3 | Aboveground biomass prediction from the largest trees

We selected 20 as the number of largest trees to predict plot metrics. The resulting model predicting AGB (in megagrams per hectare) based on the 20 largest trees is:

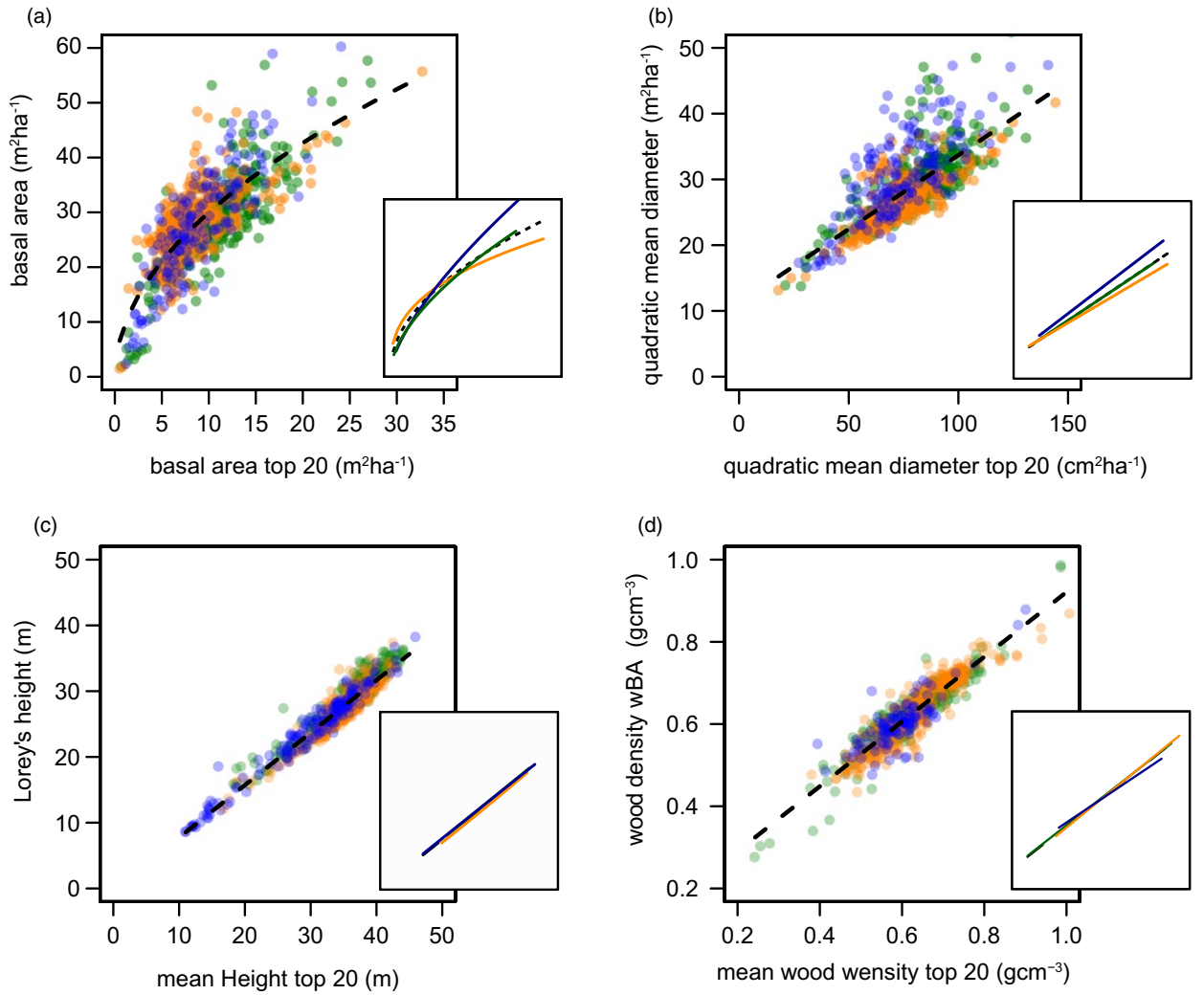

FIGURE 3 Prediction of plot metrics (*y* axis) from the 20 largest trees (*x* axis). Results are shown for: (a) basal area, (b) quadratic mean diameter, (c) Lorey's height, and (d) wood density weighted by the basal area. Each dot corresponds to a single plot, coloured in orange, green and blue for America,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

$$AGB = 0.0735 \times (Dg_{20} H_{20} WD_{20})^{1.1332} \ (rRMSE = 0.179;\ R^2 = 0.85; Akaike\ Criterion = -260.18) \quad (10)$$

Given that the exponent was close to one, we also developed an alternative and more operational model with the exponent constrained to one, given by:

$$AGB = 0.195 \times (Dg_{20} H_{20} WD_{20}) \ \left(rRMSE = 0.177;\ R^2 = 0.85;\ AIC = -195\right) \quad (11)$$

Ground measurements of plot AGB were predicted by our N_{LT} model with the exponent constrained to one, with a total error of 17.9% (Figure 4), a value which encompasses the error of the N_{LT} model and the error related to the allometric model chosen. The leave-one-out cross-validation procedure yielded similar results (rRMSE = 0.19; R^2 = 0.81), validating the use of the model on independent sites.

3.4 | Determining the cause of residual variations

The explanatory variables all together explain *ca.* 37% of the variance in AGB at both plot and site levels when omitting the diameter structure, and *ca.* 63% at site level when included (Figure 5). In general, forest structure and particularly the stem density explained most of the residuals (Table 1; weights: 79 and 54% at plot and site level, respectively). The stem density was followed by a continental effect (weights: 18, 28 and 1% for Africa, America and Asia, respectively) and by the effect of H_{BA} and WD_{BA} (respective weights: 1 and 0% at the plot level, 0 and 11% at the site level, and 23 and 0% when accounting for the diameter structure at the site level). Inclusion of the diameter structure provided the best explanation of residuals, with 63% of variance explained, and a weight of 69% for the first axis of the PCA (Supporting Information Figure 3). This first axis of the PCA was related to the general abundance of trees at a site, and in particular, medium-sized trees (40–60 cm). Among environmental variables, only rainfall was significantly related to the residuals at the site level when the diameter structure was considered (2%).

3.5 | Differences among continents

Although diameter structure explained a large fraction of the residual variance of our global model, there was a marked difference in forest structure across continents (Figure 6). Consequently, we investigated differences between continents in the distribution of residuals of the pan-tropical model (Figure 6a), in the relative contribution of the 20 largest trees to plot total biomass (Figure 6b) and in the contribution to the total AGB per diameter class (Figure 6c–f). To this end, we considered the following four classes of diameter at 130 cm: 10–30, 30–50, 50–70 and > 70 cm. The results show that the prediction of biomass from the 20 largest trees using the pan-tropical model tends to be slightly overestimated in Africa (+3%) and underestimated in America (−3%) and in Asia (−5%) (Figure 6a). The proportion of biomass is higher in the high-diameter class (> 70 cm) in Africa, in intermediate-diameter classes (between 30 and 70 cm) in America, and is equally distributed among the different diameter classes in Asia (Figure 6c,d).

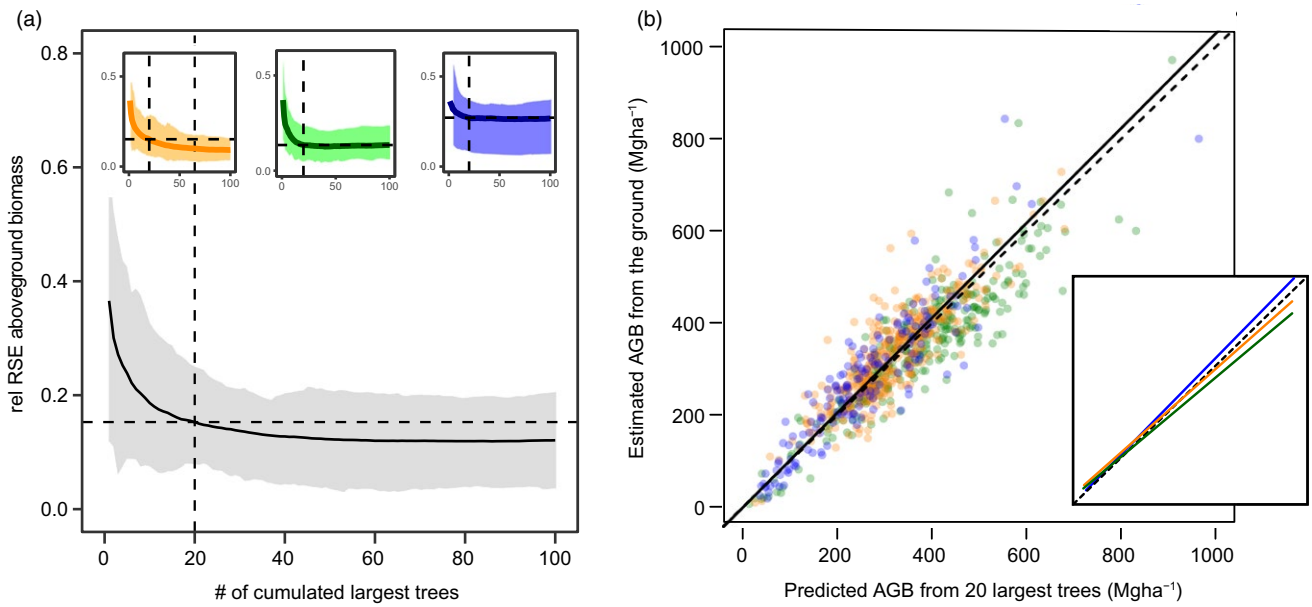

FIGURE 4 Quality of the prediction of aboveground biomass (AGB) from largest trees plot metrics. Variation of the relative root mean square error (rRMSE) of the prediction of AGB from *i* largest trees versus the cumulative number of largest trees (a) and detailed prediction of AGB from plot metrics of the 20 largest trees (b). Results are shown for the 867 plots in the three continents, coloured orange, green and blue for America, Africa and Asia, respectively. The regression line of the model is shown as a continuous black line, and the dashed black line shows a 1:1 relationship. The figure shows an unbiased prediction of AGB across the 867 plots, with slight but significant differences between the three continents

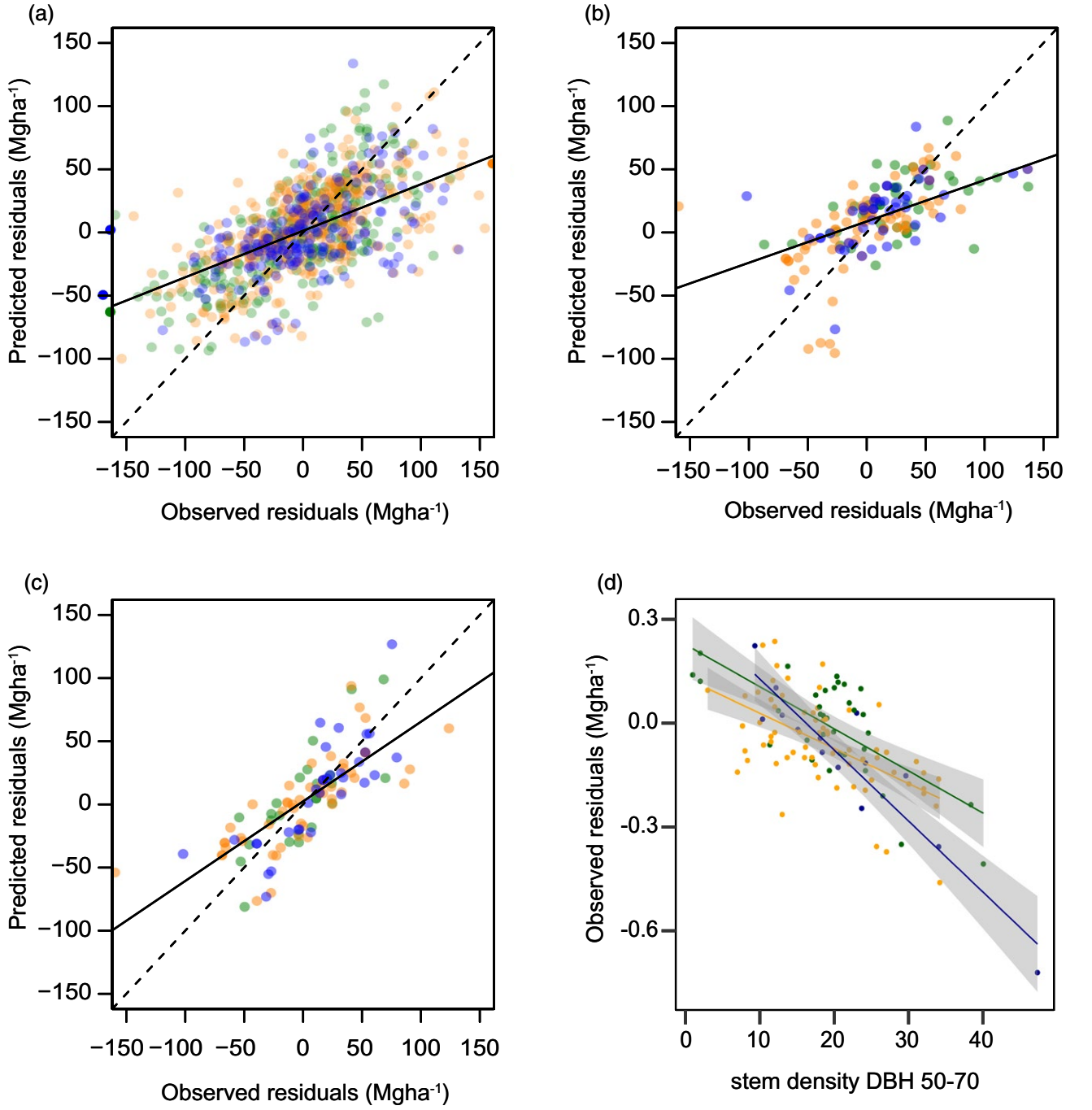

FIGURE 5 Predicted versus observed residuals of aboveground biomass predicted from the 20 largest trees. Residuals are explored at three different levels: (a) plot, (b) site (without considering the diameter structure as an explanatory variable), (c) site (considering the diameter structure), and (d) along the stem density of medium-sized trees. America, Africa and Asia are coloured in orange, green and blue, respectively. The panels show a good prediction of residuals in (a) and (b), driven by stem density, and a less biased prediction in (c), driven by the diameter structure. Variances of observed residuals are also well explained by the stem density of medium-sized trees (d), which mainly drive the first axis of the principal components analysis

4 | DISCUSSION

4.1 | The largest trees, convergences and divergences between continents

Sampling a few largest trees per hectare generally allows an unbiased prediction of four key descriptors of forest structures across the tropics. There is generally no improvement in predicting biomass, quadratic mean diameter, Lorey's height (H_{BA}) or community wood density beyond the first 10–20 largest trees (Figures 2,3 and 3a). But when a forest plot presents an abundant number of large trees (Figure 5d), increasing the number of trees sampled does improve the accuracy of the model. This is attributable to the fact that

TABLE 1 Weight of each variable retained for the explanation of aboveground biomass residuals

Level of residual	Parameter	Weight
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

Note. Weights are calculated as a type II sum of squares, which measures the decreased residual sum of squares attributable to an added variable once all the other variables have been introduced into the model. Results are shown for the exploration of residuals at the plot and at the site level, with and without consideration of the diameter structure. Weights are dominated by structural variables, and in particular, the stem density and the diameter structure. Height, wood density and continent have also a non-negligible influence on residuals. PCA = principal components analysis.

the higher the total AGB in a plot, the lower the proportion of total AGB encompassed by the largest trees. This is particularly true for BA, for which rRMSE continues to decrease up to 100 largest trees (Figure 2a). In contrast, Lorey's height predictions are altered when a large number of trees are included (Figure 2c; i.e., when smaller, often suppressed, trees draw the average down; Farrior, Bohlman, Hubbell, & Pacala, 2016). This might explain why the prediction of AGB does not mirror that of basal area (Figures 2,3b and 3a) and suggests that the number of largest trees should be set independently from each predictor considered. Interestingly, the evolution of relative error in AGB prediction as a function of the number of largest trees considered does not follow the same path between continents. For instance, the error of prediction saturates more quickly in Africa and Asia than in America. Investigation of residuals showed that the diameter structure (Figure 5c; Supporting Information Figure 3b), and in particular, the number of medium-sized trees (Figure 5d), drives variability in AGB predictions. It is therefore not surprising to see that in our dataset the site with higher levels of underestimations is the one with the highest number of medium-sized trees, which is found in Asia in the Western Ghats of India.

The good performance of models based on the 20 largest trees in predicting Lorey's height and community wood density at the site level was not surprising. Both metrics were weighted by basal area, driven *de facto* by the largest trees. Nonetheless, their consistency across sites and continents was not expected, which emphasize the generality of our approach.

The predictability of plot-level forest structure metrics from the largest trees implies that characteristics of smaller trees do not vary in a completely independent manner from those of the larger trees. For example, plots where the largest trees have a low basal area tend to have low plot-level basal area (Figure 3a), meaning that the total size of the smaller trees is sufficiently constrained that it does not compensate for the small size of the largest trees. Such constraints could arise through size–frequency distributions being set by allometric scaling rules (Enquist, West, & Brown, 2009) or could be

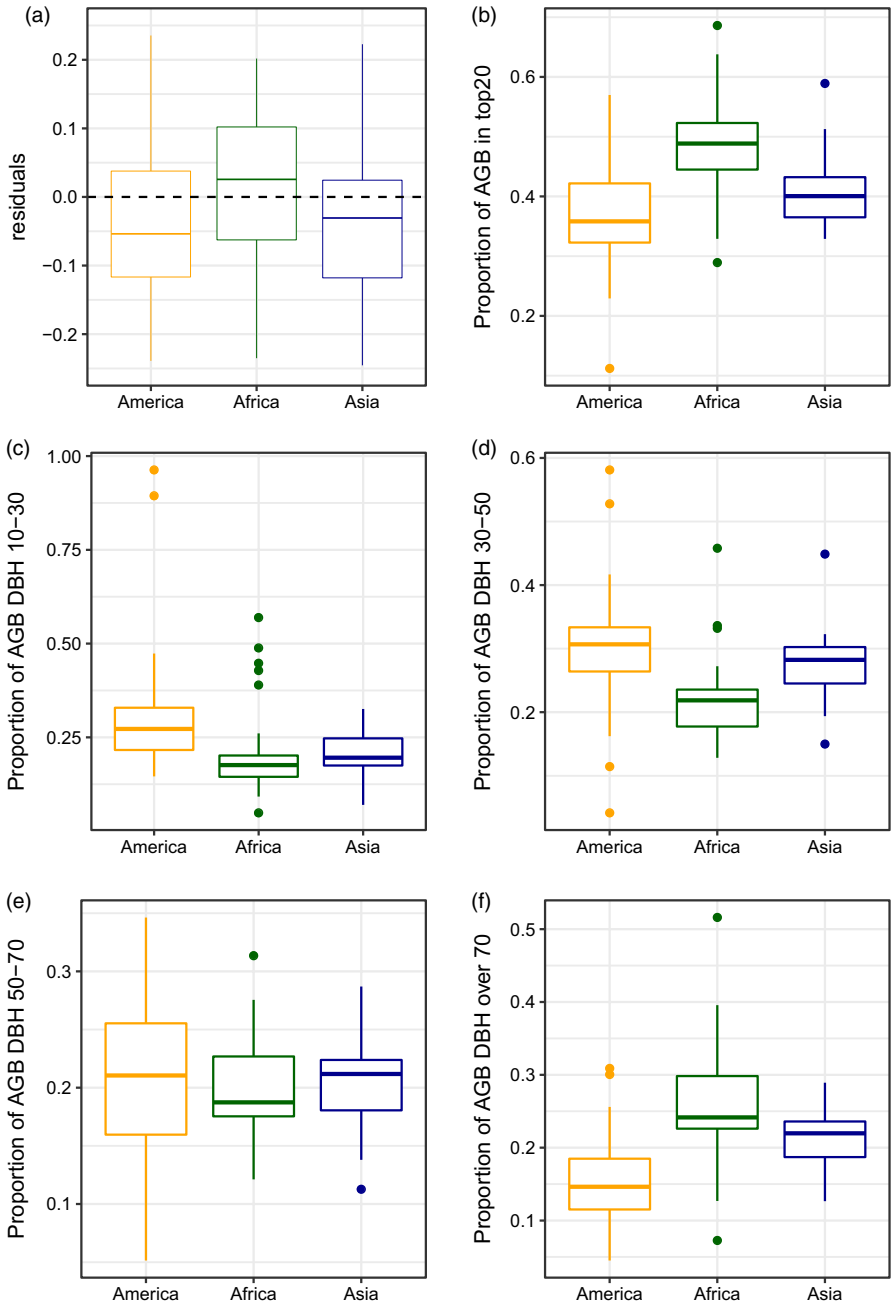

FIGURE 6 Comparison across continents of aboveground biomass (AGB) prediction per site and their contribution to different shares of the diameter structure. Africa, Asia and America are coloured in green, blue and orange, respectively. (a) The distribution of the residuals of pan-tropical AGB prediction from the 20 largest trees shows that predictions are slightly overestimated in Africa (+3%), and slightly underestimated in Asia (−3%) and America (−5%). (b) The proportion of AGB in the 20 largest trees is highest in Africa (48%), followed by Asia (40%) and America (35%). (c–f) The decomposition across four diameter classes [i.e., 10–30, 30–50, 50–70 and > 70 cm diameter at breast height (DBH)] of their relative share of the total biomass shows that most of the biomass is found in the large trees in Africa and in the small to medium trees in America. Asia presents a more balanced distribution of biomass across the diameter structure

attributable to the largest trees responding in the same way as the remaining smaller trees to environmental drivers.

Despite the general consistency of these relationships across continents, slight differences are evident when comparing the pan-tropical model residuals across continents (Figure 6; Supporting Information Figure 4). These differences indicate biogeographical variation in forest structure. In America, our pan-tropical model tends slightly to underestimate basal area (mean: −5%) and overestimate Lorey's height (mean: +3%) (Supporting Information Figure 4). This suggests that large trees make up a smaller proportion of basal

area in America and that for a given diameter we find higher trees (Supporting Information Figure 2), which confirms that the shape of height–diameter allometries varies between continents (Banin et al., 2012; Sullivan et al., 2018). In Africa, large trees (i.e., diameter at breast height > 70 cm) are more abundant and account for a large fraction of plot biomass (Figure 6f). This supports previous observations that African forests are characterized by fewer but larger stems (Feldpausch et al., 2012; Lewis et al., 2013), whereas forests in the Americas have more stems but generally have lower biomass (Sullivan et al., 2017). In Asia, the distribution of the biomass across diameter classes appears more balanced (Figure 6c–f). Such differences in forest structure, albeit limited, suggest that tropical forests differ between continents in terms of dynamics, carbon cycling, response and feedback to climate and resilience to external forcings (e.g., climate change, forest degradation and deforestation).

Interestingly, although a recent global phylogenetic classification of tropical forests groups American with African forests versus Asian forests (Slik et al., 2018), our study of forest structure properties tends more to single out American forests, and in particular, to highlight the contrast between African and American forests. Although this deserves further investigation, it might reveal the lack of a close relationship between forest structure properties and phylogenic similarity, which echoes recent results on the absence of a relationship between tropical forest diversity and biomass (Sullivan et al., 2017).

4.2 | Largest trees: A gateway to global monitoring of tropical forests

Revealing the predictive capacity held by the largest trees, our results constitute a major step forwards to monitor forest structures and biomass stocks. The largest trees in tropical forests can therefore be used to make accurate predictions of various ground-measured properties (i.e., the quadratic mean diameter, the basal area, Lorey's height and community wood density), whereas previous work has predicted only biomass 'estimates' (e.g., Bastin et al., 2015; Slik et al., 2013). The advantages of our approach are as follows: (a) it allows us to describe forest structure independently of any biomass allometric model; (b) it allows us to integrate environmentally based variations in the *D*–*H* relationship, known to vary locally (Feldpausch et al., 2011; Kearsley et al., 2013); and (c) it is also relatively insensitive to differences in floristic composition and community wood density (Poorter et al., 2015).

Furthermore, the 'largest trees' models were developed for each plot-level metric and for any number of largest trees. Thus, they do not rely on any arbitrary threshold of tree diameter. Note that the optimal number of largest trees to be measured (i.e., 20) was set for demonstration and can vary depending on the needs and capacities of each country or project (see Supporting Information Table 2). In the same way, local models could integrate locally developed biomass models, when available. Consequently our approach: (a) can be used in young or regenerating unmanaged forests with a low 'largest tree' diameter threshold; and (b) is compatible with recent RS approaches able to single out canopy trees and describe their crown and height metrics (Coomes et al., 2017; Ferraz et al., 2016).

4.3 | Aboveground biomass model from the largest trees: A multiple opportunity

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

Developing countries willing to implement Reduction of Emissions from Deforestation and Forest Degradation (REDD+) activities will also report on their carbon emissions and develop a national reference level (IPCC, 2006; Maniatis & Mollicone, 2010). However, most tropical countries lack the capacity to assume multiple, exhaustive and costly forest carbon inventories (Romijn, Herold, Kooistra, Murdiyarso, & Verchot, 2012). By measuring only a few large trees per hectare, our results show that it is possible to obtain unbiased estimates of aboveground carbon stocks in a time- and cost-efficient manner. Assuming that 400–600 trees with D > 10 cm are measured in a typical 1-ha sample plot, monitoring only 20 trees is a significant improvement. Although finding the 20 largest trees in a plot of several hundred individuals requires evaluating > 20 trees, in practice, a conservative diameter threshold could be defined to ensure that the 20 largest trees are sampled. An alternative could also be found in the development of a relascope-based approach adapted to detection of the largest trees in tropical forests. Using such approach would facilitate rapid field sampling in extensive areas to produce large-scale AGB estimates. Those could fulfil the needs for calibration and validation of current and forthcoming space missions focused on AGB.

Our findings also point towards the potential effectiveness of using RS techniques to characterize canopy trees for inferring the attributes of entire forest stands. Remote sensing data could be used for direct measurement (e.g., tree-level metrics, such as height, crown width and crown height) of the largest trees as a potential alternative to indirect development of complex metrics (e.g., mean canopy height, texture) used to extrapolate forest properties. Although the use of a single-tree approach has shown some limitations to extrapolate plot metrics (Coomes, Šafka, Shepherd, Dalponte, & Holdaway, 2018), we have yet to investigate its potential to identify largest trees. Some further refinements are needed, but most of the tools required to develop 'largest trees' models are readily available. In particular, Ferraz et al. (2016) developed an automated procedure to locate single trees based on airborne LiDAR data, to measure their height and crown area. Crown area could then be linked to basal area, because the logarithm of crown area is consistently correlated, with a slope of 1.2–1.3, to the logarithm of tree diameter across the tropics (Blanchard et al., 2016). Regarding wood density, hyperspectral signature and high-resolution topography offers a promising way to assess functional traits remotely (e.g., Asner et al., 2017; Jucker et al., 2018), which could potentially provide proxies of wood density. Alternative approaches could focus on the development of plot-level AGB prediction by replacing the basal area of the largest trees with their crown metrics. Although the measurement of crown areas has

yet to be generalized when inventorying plots, several biomass allometric models already partition trunk and crown mass (Coomes et al., 2017; Jucker et al., 2017; Ploton et al., 2016).

The main limitation of our approach lies in the limited inference that can be made on the understorey and sub-canopy trees. We show that most of the remaining variance is explained by variations in diameter structures, and in particular, in the total stem density. Interestingly, stem density was generally identified as a poor predictor of plot biomass in tropical forests (Lewis et al., 2013; Slik et al., 2010). However, our results show that stem density explains most of the remaining variance (Supporting Information Table S1). This suggests that, in addition to trying to understand large-scale variations in large trees and other plot metrics that can be quantified directly from RS, we should also put more effort into understanding variation in smaller trees, which mainly drives total stem density and the total floristic diversity. Smaller trees are also essential to characterize forest dynamics and understand changes in carbon stocks. Several options are nonetheless possible from RS, considering the variation in LiDAR point density below the canopy layer (D'Oliveira, Reutebuch, McGaughey, & Andersen, 2012), the distribution of leaf area density (Stark et al., 2015, 2012 ; Tang & Dubayah, 2017; Vincent et al., 2017) or the use of multitemporal LiDAR data to obtain information on forest gap generation dynamics and, consequently, on forest diameter structure (Farrior et al., 2016; Kellner, Clark, & Hubbell, 2009).

4.4 | Large trees in degraded forests

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

4.5 | Conclusion: Towards improved estimates of tropical forest biomass

The acquisition, accessibility and processing capabilities of RS data with a very high spatial, spectral and temporal resolution has increased exponentially in recent years (Bastin et al., 2017). However, to develop accurate global maps we will have to obtain a larger number of field plots and develop new ways to use RS data. Our results provide a step forwards for both by: (a) drastically decreasing the number of individual tree measurements required to obtain an accurate, yet less precise, estimate of plot biomass; and (b) opening the way to direct measurement of plot metrics measured from RS to estimate plot biomass.

As highlighted by Clark and Kellner (2012), new biomass allometric models relating plot-level biomass measured from destructive sampling and plot-level metrics measured from RS products should be developed, as an alternative to current tree-level allometric models. Such an effort will greatly reduce operational costs and uncertainties surrounding terrestrial carbon estimates and, consequently, will help developing countries in the development of national forest inventories and aid the scientific community in better understanding the effect of climate change on forest ecosystems.

ACKNOWLEDGMENTS

J.-F.B. was supported for data collection by the FRIA-FNRS (Fond National pour la Recherche Scientifique), ERAIFT (Ecole Régionale Post-Universitaire d'Aménagement et de Gestion Intégrés des Forêts Tropicales), World Wide Fund for Nature (WWF) (...) WWF and by the CoForTips project (ANR-12-EBID-0002); T.d.H. was supported by the COBIMFO project (Congo Basin integrated monitoring for forest carbon mitigation and biodiversity) funded by the Belgian Science Policy Office (Belspo); C.H.G. was supported by the 'Sud Expert Plantes' project of French Foreign Affairs, CIRAD and SCAC. Some of the data in this paper were provided by the RAINFOR Network, the AfriTRON network, TEAM Network, the partnership between Conservation International, The Missouri Botanical Garden, The Smithsonian Institution and The Wildlife Conservation Society and the Gordon and Betty Moore Foundation. We acknowledge data contributions from the TEAM network not listed as co-authors (upon a voluntary basis). We thank Jean-Phillipe Puyravaud, Estação Científica Ferreira Penna (MPEG) and the Andrew Mellon Foundation and National Science Foundation (DEB 0742830). The forest plots in Nova Xavantina and Southern Amazonia, Brazil were funded by grants from Project PELD-CNPq/FAPEMAT (403725/2012-7; 441244/2016-5; 164131/2013); CNPq-PPBio (457602/2012-0); productivity grants (CNPq/PQ-2) to B. H. Marimon-Junior and B. S. Marimon; Project USA-NAS/PEER (#PGA-2000005316) and Project ReFlor FAPEMAT 0589267/2016. Finally, we thank Helen Muller-Landau for her careful revision and comments on the manuscript.

CONFLICT OF INTEREST

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

AUTHOR CONTRIBUTIONS

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

DATA ACCESSIBILITY

Data for plots in the CTFS network are available through the online portal at https://www.forestgeo.si.edu; in the Forestplot network at https://www.forestplots.net/; and in the TEAM network at https://www.teamnetwork.org/.

ORCID

Koen Hufkens https://orcid.org/0000-0002-5070-8109
Jean-François Bastin https://orcid.org/0000-0003-2602-7247
Ervan Rutishauser https://orcid.org/0000-0003-1182-4032

REFERENCES

Asner, G. P., Martin, R. E., Knapp, D. E., Tupayachi, R., Anderson, C. B., Sinca, F., … Llactayo, W. (2017). Airborne laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science*, *355*, 385–389.

Asner, G. P., & Mascaro, J. (2014). Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. *Remote Sensing of Environment*, *140*, 614–624.

Asner, G. P., Mascaro, J., Muller-Landau, H. H. C., Vieilledent, G., Vaudry, R., Rasamoelina, M., … van Breugel, M. (2012). A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia*, *168*, 1147–1160.

Baccini, A., Walker, W., Carvalho, L., Farina, M., Sulla-Menashe, D., & Houghton, R. A. (2017). Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science*, *358*, 230–234.

Banin, L., Feldpausch, T. R., Phillips, O. L., Baker, T. R., Lloyd, J., Affum-Baffoe, K., … Lewis, S. L. (2012). What controls tropical forest architecture? Testing environmental, structural and floristic drivers. *Global Ecology and Biogeography*, *21*, 1179–1190.

Bastin, J.-F., Barbier, N., Couteron, P., Adams, B., Shapiro, A., Bogaert, J., … De Cannière, C. (2014). Aboveground biomass mapping of African forest mosaics using canopy texture analysis: Toward a regional approach. *Ecological Applications*, *24*, 1984–2001.

Bastin, J. F., Barbier, N., Réjou-Méchain, M., Fayolle, A., Gourlet-Fleury, S., Maniatis, D., … Bogaert, J. (2015). Seeing Central African forests through their largest trees. *Scientific Reports*, *5*, 13156.

Bastin, J.-F., Berrahmouni, N., Grainger, A., Maniatis, D., Mollicone, D., Moore, R., … Castro, R. (2017). The extent of forest in dryland biomes. *Science*, *356*, 635–638.

Bennett, A. C., McDowell, N. G., Allen, C. D., & Anderson-Teixeira, K. J. (2015). Larger trees suffer most during drought in forests worldwide. *Nature Plants*, *1*, 15139.

Blanchard, E., Birnbaum, P., Ibanez, T., Boutreux, T., Antin, C., Ploton, P., … Couteron, P. (2016). Contrasted allometries between stem diameter, crown area, and tree height in five tropical biogeographic areas. *Trees*, *30*, 1953–1968.

Carré, F., Hiederer, R., Blujdea, V., & Koeble, R. (2010). *Background guide for the calculation of land carbon stocks in the biofuels sustainability scheme drawing on the 2006 IPCC Guidelines for National Greenhouse Gas Inventories* (p. 128). Luxembourg: Publications Office of the European Union

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

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M. S., Delitti, W. B. C., … Vieilledent, G. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology*, *20*, 3177–3190.

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

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

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

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

Coomes, D. A., Dalponte, M., Jucker, T., Asner, G. P., Banin, L. F., Burslem, D. F. R. P., … Qie, L. (2017). Area-based vs tree-centric approaches to mapping forest carbon in Southeast Asian forests from airborne laser scanning data. *Remote Sensing of Environment*, *194*, 77–88.

Coomes, D. A., Šafka, D., Shepherd, J., Dalponte, M., & Holdaway, R. (2018). Airborne laser scanning of natural forests in New Zealand reveals the influences of wind on forest carbon. *Forest Ecosystems*, *5*, 10.

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

Dubayah, R., Goetz, S. J., Blair, J. B., Fatoyinbo, T. E., Hansen, M., Healey, S. P., … Swatantran, A. (2014). The global ecosystem dynamics investigation. *American Geophysical Union, Fall Meeting 2014, Abstract no. U14A-07.*

Enquist, B. J., West, G. B., & Brown, J. H. (2009). Extensions and evaluations of a general quantitative theory of forest structure and dynamics. *Proceedings of the National Academy of Sciences of the USA*, *106*, 7046–7051.

Farrior, C. E., Bohlman, S. A., Hubbell, S., & Pacala, S. W. (2016). Dominance of the suppressed: power-law size structure in tropical forests. *Science*, *351*, 155–157.

Feldpausch, T. R., Banin, L., Phillips, O. L., Baker, T. R., Lewis, S. L., Quesada, C. A., … Lloyd, J. (2011). Height-diameter allometry of tropical forest trees. *Biogeosciences*, *8*, 1081–1106.

Feldpausch, T. R., Lloyd, J., Lewis, S. L., Brienen, R. J. W., Gloor, M., Monteagudo Mendoza, A., … Phillips, O. L. (2012). Tree height integrated into pantropical forest biomass estimates. *Biogeosciences*, *9*, 3381–3403.

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

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

Ho Tong Minh, D., Le Toan, T., Rocca, F., Tebaldini, S., Villard, L., Réjou-Méchain, M., … Chave, J. (2016). SAR tomography for the retrieval of forest biomass and height: Cross-validation at two tropical forest sites in French Guiana. *Remote Sensing of Environment*, *175*, 138–147.

IPCC. (2006). *2006 IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme* (H. S. Eggleston, L. Buendia, K. Miwa, T. Ngara, & K. Tanabe, Eds.). Hayama, Japan: IGES.

Jucker, T., Bongalov, B., Burslem, D. F. R. P., Nilus, R., Dalponte, M., Lewis, S. L., … Coomes, D. A. (2018). Topography shapes the structure, composition and function of tropical forest landscapes. *Ecology Letters*, *21*, 989–1000.

Jucker, T., Caspersen, J., Chave, J., Antin, C., Barbier, N., Bongers, F., … Coomes, D. A. (2017). Allometric equations for integrating remote sensing imagery into forest monitoring programmes. *Global Change Biology*, *23*, 177–190.

Kearsley, E., de Haulleville, T., Hufkens, K., Kidimbu, A., Toirambe, B., Baert, G., … Verbeeck, H. (2013). Conventional tree height-diameter

relationships significantly overestimate aboveground carbon stocks in the Congo Basin. *Nature Communications*, *4*, 2269.

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

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

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

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

Le Toan, T., Quegan, S., Davidson, M. W. J., Balzter, H., Paillou, P., Papathanassiou, K., … Ulander, L. (2011). The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle. *Remote Sensing of Environment*, *115*, 2850–2860.

Lewis, S. L., Sonké, B., Sunderland, T., Begne, S. K., Lopez-Gonzalez, G., van der Heijden, G. M. F., … Zemagho, L. (2013). Above-ground biomass and structure of 260 African tropical forests. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *368*, 20120295.

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

Lutz, J. A., Furniss, T. J., Johnson, D. J., Davies, S. J., Allen, D., Alonso, A., … Zimmerman, J. K. (2018). Global importance of large-diameter trees. *Global Ecology and Biogeography*, *27*, 849–864.

Malhi, Y., Wood, D., Baker, T. R., Wright, J., Phillips, O. L., Cochrane, T., … Vinceti, B. (2006). The regional variation of aboveground live biomass in old-growth Amazonian forests. *Global Change Biology*, *12*, 1107–1138.

Maniatis, D., & Mollicone, D. (2010). Options for sampling and stratification for national forest inventories to implement REDD+ under the UNFCCC. *Carbon Balance and Management*, *5*, 9.

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

Meakem, V., Tepley, A. J., Gonzalez-Akre, E. B., Herrmann, V., Muller-Landau, H. C., Wright, S. J., … Anderson-Teixeira, K. J. (2018). Role of tree size in moist tropical forest carbon cycling and water deficit responses. *New Phytologist*, *219*, 947–958.

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

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

New, M., Hulme, M., Jones, P., New, M., Hulme, M., & Jones, P. (1999). Representing twentieth-century space-time climate variability. Part I: Development of a 1961–90 mean monthly terrestrial climatology. *Journal of Climate*, *12*, 829–856.

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

Ploton, P., Barbier, N., Couteron, P., Antin, C. M., Ayyappan, N., Balachandran, N., … Pélissier, R. (2017). Toward a general tropical forest biomass prediction model from very high resolution optical satellite images. *Remote Sensing of Environment*, *200*, 140–153.

Ploton, P., Barbier, N., Takoudjou Momo, S., Réjou-Méchain, M., Boyemba Bosela, F., Chuyong, G., … Pélissier, R. (2016). Closing a gap in tropical forest biomass estimation: Taking crown mass variation into account in pantropical allometries. *Biogeosciences*, *13*, 1571–1585.

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

Poorter, L., van der Sande, M. T., Thompson, J., Arets, E. J. M. M., Alarcón, A., Álvarez-Sánchez, J., … Peña-Claros, M. (2015). Diversity enhances carbon storage in tropical forests. *Global Ecology and Biogeography*, *24*, 1314–1328.

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

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

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

Romijn, E., Herold, M., Kooistra, L., Murdiyarso, D., & Verchot, L. (2012). Assessing capacities of non-Annex I countries for national forest monitoring in the context of REDD+. *Environmental Science & Policy*, *19–20*, 33–48.

Rutishauser, E., Hérault, B., Baraloto, C., Blanc, L., Descroix, L., Sotta, E. D., … Sist, P. (2015). Rapid tree carbon stock recovery in managed Amazonian forests. *Current Biology: CB*, *25*, R787–R788.

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

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

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

Slik, J. W. F., Aiba, S.-I., Brearley, F. Q., Cannon, C. H., Forshed, O., Kitayama, K., … Van Valkenburg, J. L. C. H. (2010). Environmental correlates of tree biomass, basal area, wood specific gravity and stem density gradients in Borneo's tropical forests. *Global Ecology and Biogeography*, *19*, 50–60.

Slik, J. W. F., Alvarez-loayza, P., Alves, L. F., Ashton, P., Balvanera, P., Bastian, M. L., … Bernacci, L. (2015). An estimate of the number of tropical tree species. *Proceedings of the National Academy of Sciences of the USA*, *112*, E4628–E4629.

Slik, J. W. F., Franklin, J., Arroyo-Rodríguez, V., Field, R., Aguilar, S., Aguirre, N., … Zang, R. (2018). Phylogenetic classification of the world's tropical forests. *Proceedings of the National Academy of Sciences of the USA*, *115*, 1837–1842.

Slik, J. W. F., Paoli, G., McGuire, K., Amaral, I., Barroso, J., Bastian, M., … Zweifel, N. (2013). Large trees drive forest aboveground biomass variation in moist lowland forests across the tropics. *Global Ecology and Biogeography*, *22*, 1261–1271.

Stark, S. C., Enquist, B. J., Saleska, S. R., Leitold, V., Schietti, J., Longo, M., … Oliveira, R. C. (2015). Linking canopy leaf area and light environments with tree size distributions to explain Amazon forest demography. *Ecology Letters*, *18*, 636–645.

Stark, S. C., Leitold, V., Wu, J. L., Hunter, M. O., de Castilho, C. V., Costa, F. R. C., … Saleska, S.R. (2012). Amazon forest carbon dynamics predicted by profiles of canopy leaf area and light environment. *Ecology Letters*, *15*, 1406–1414.

Stegen, J. C., Swenson, N. G., Enquist, B. J., White, E. P., Phillips, O. L., Jørgensen, P. M., … Núñez Vargas, P. (2011). Variation in above-ground forest biomass across broad climatic gradients. *Global Ecology and Biogeography*, *20*, 744–754.

Sullivan, M. J. P., Lewis, S. L., Hubau, W., Qie, L., Baker, T. R., Banin, L. F., … Phillips, O. L. (2018). Field methods for sampling tree height for tropical forest biomass estimation. *Methods in Ecology and Evolution*, *9*, 1179–1189.

Sullivan, M. J. P., Talbot, J., Lewis, S. L., Phillips, O. L., Qie, L., Begne, S. K., … Zemagho, L. (2017). Diversity and carbon storage across the tropical forest biome. *Scientific Reports*, *7*, 39102.

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

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

Vincent, G., Antin, C., Laurans, M., Heurtebize, J., Durrieu, S., Lavalley, C., & Dauzat, J. (2017). Mapping plant area index of tropical evergreen forest by airborne laser scanning. A cross-validation study using LAI2200 optical sensor. *Remote Sensing of Environment*, *198*, 254–266.

Xu, L., Saatchi, S. S., Shapiro, A., Meyer, V., Ferraz, A., Yang, Y., … Ebuta, D. (2017). Spatial distribution of carbon stored in forests of the Democratic Republic of Congo. *Scientific Reports*, *7*, 15030.

Zanne, A. E., Lopez-Gonzalez, G., Coomes, D. A., Ilic, J., Jansen, S., Lewis, … Chave, J. (2009). Global wood density database. *Dryad*, *235*, 33.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

BIOSKETCH

Jean-Francois Bastin is a postdoctoral Fellow of the Crowther Laboratory at the Institute of Integrative Biology, Department of Environmental Systems Science, ETH-Zurich. He is an ecologist and a geographer using remote sensing to study the effect of global change on terrestrial ecosystems.

Ervan Rutishauser is a postdoctoral Fellow at the Smithsonian Tropical Research Institute, broadly interested in understanding environmental resilience to natural or human-induced disturbances. He aims at providing rigorous and practical evidence to trigger a shift towards better ressource management and conservation.

How to cite this article: Bastin J-F, Rutishauser E, Kellner JR, et al. Pan-tropical prediction of forest structure from the largest trees. *Global Ecol Biogeogr*. 2018;00:1–18. https://doi.org/10.1111/geb.12803