



**HAL**  
open science

## Ground Data are Essential for Biomass Remote Sensing Missions

Jérôme Chave, Stuart Davies, Oliver Phillips, Simon Lewis, Plinio Sist, Dmitry Schepaschenko, John Armston, Tim Baker, David Coomes, Mathias Disney, et al.

► **To cite this version:**

Jérôme Chave, Stuart Davies, Oliver Phillips, Simon Lewis, Plinio Sist, et al.. Ground Data are Essential for Biomass Remote Sensing Missions. *Surveys in Geophysics*, 2019, 40 (4), pp.863-880. 10.1007/s10712-019-09528-w . hal-02403077

**HAL Id: hal-02403077**

**<https://hal.science/hal-02403077>**

Submitted on 5 Dec 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# 1 **Ground data are essential for biomass remote sensing missions**

2

3 *Ms for submission to Surveys in Geophysics special issue and ISSI book (n°71) "Forest Biomass*  
4 *and Structure from Space"*

5

6 Jérôme Chave<sup>1</sup>, Stuart J. Davies<sup>2</sup>, Oliver L Phillips<sup>3</sup>, Simon L. Lewis<sup>3</sup>, Plinio Sist<sup>4</sup>, Dmitry  
7 Schepaschenko<sup>5</sup>, John Armston<sup>6</sup>, Tim R. Baker<sup>3</sup>, David Coomes<sup>7</sup>, Mathias Disney<sup>8</sup>, Laura  
8 Duncanson<sup>6</sup>, Bruno Héroult<sup>4,9</sup>, Nicolas Labrière<sup>1</sup>, Victoria Meyer<sup>10</sup>, Maxime Réjou-Méchain<sup>11</sup>,  
9 Klaus Scipal<sup>12</sup>, Sassan Saatchi<sup>10</sup>

10 Affiliations:

- 11 1. Université Toulouse 3 Paul Sabatier, CNRS, ENFA, UMR 5174 Evolution et Diversité  
12 Biologique (EDB), F-31062 Toulouse, France
- 13 2. Center for Tropical Forest Science-Forest Global Earth Observatory, Smithsonian  
14 Tropical Research Institute, Washington, DC, USA
- 15 3. School of Geography, University of Leeds, Leeds LS2 9JT, U.K.
- 16 4. CIRAD-ES, UPR Forests and Societies, University of Montpellier – CIRAD, 34398  
17 Montpellier Cedex 5, France
- 18 5. International Institute for Applied Systems Analysis, Schlossplatz 1 - A-2361 Laxenburg,  
19 Austria
- 20 6. Department of Geographical Sciences, University of Maryland, College Park, MD 20742,  
21 USA
- 22 7. Department of Plant Sciences, Forest Ecology and Conservation group, University of  
23 Cambridge, Cambridge, UK
- 24 8. Department of Geography, University College London, London WC1E 6BT, UKNERC  
25 National Centre for Earth Observation (NCEO), UK
- 26 9. Institut National Polytechnique Félix Houphouët-Boigny (INP-HB), Department of  
27 Forest, Water and Environment, Yamoussoukro, Ivory Coast
- 28 10. Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109 USA
- 29 11. AMAP, IRD, CNRS, CIRAD, INRA, Univ Montpellier, Montpellier, France.
- 30 12. ESA-ESTEC, Noordwijk, 2201 AZ, The Netherlands

31

32 Corresponding author: Jérôme Chave (jerome.chave@univ-tlse3.fr)

33 **Acknowledgements**

34

35 We thank the organizers of the ISSI Bern meeting in November 2017 for stimulating discussions,  
36 and for their invitation to submit this manuscript. We gratefully acknowledge funding by  
37 “Investissement d’Avenir” programs managed by Agence Nationale de la Recherche (CEBA,  
38 ref. ANR-10-LABX-25-01), from CNES, and from ESA (IFBN project  
39 4000114425/15/NL/FF/gp, as part of the BIOMASS mission program).

40

41 **Abstract**

42 Several remote-sensing missions will soon produce carbon maps over all terrestrial ecosystems.  
43 These missions are critically dependent on accurate and representative *in situ* datasets for the  
44 training of their algorithms and product validation. Long-term ground-based forest monitoring  
45 systems are limited, especially in the tropics. Ground-based observation systems are critical for  
46 the remote-sensing missions, and they need to be maintained at least over the lifetime of the  
47 planned missions. Here we propose a strategy for a coordinated and global network of *in situ*  
48 data that would benefit biomass remote sensing missions. To produce accurate ground-based  
49 biomass estimates, strict data quality must be guaranteed to users and ground sites need to be  
50 regularly re-visited. It is more rewarding to invest ground resources at sites where there currently  
51 is a guarantee of a long-term commitment locally, and where a core set of data is already  
52 available. We call these ‘supersites’. Long-term funding for such an inter-agency endeavour  
53 remains a critical challenge, and we here provide costing estimates to facilitate dialogue among  
54 stakeholders. One critical requirement is to ensure *in situ* data availability over the lifetime  
55 of remote-sensing missions. To this end, principal investigators of the sites should be involved  
56 early on, and long-term funding should be assured.

57

58 **Keywords:** biomass; calibration; forest; in situ data; validation

## 59 **1 Introduction**

60

61 The global carbon cycle is being altered by anthropogenic activities: carbon dioxide and other  
62 economy-related greenhouse gas emissions have steadily increased since the 1960s (Le Quéré et  
63 al. 2018). This has already had detectable consequences on the mean temperature of our planet  
64 (IPCC 2013). Land ecosystems hold a large potential for carbon storage. For example, it has  
65 been estimated that allowing Neotropical secondary forests to regenerate, without further human  
66 intervention, may enable Latin America and the Caribbean to be carbon-neutral for decades  
67 (Chazdon et al. 2016). Also, protecting intact forests is essential to ensuring carbon storage and  
68 many other ecosystem services (Pan et al. 2011). Thus, conserving existing intact forests, in  
69 combination with restoring and managing sustainably degraded forests is almost certain to be a  
70 key action to help meet the Paris Accord targets. The idea of financially incentivizing local and  
71 national initiatives to spare forest land and favour reforestation has thus received further  
72 attention, as evidenced by the United Nations' Reduced Emissions from Deforestation and forest  
73 Degradation (REDD+) program.

74 The REDD+ framework is predicated on the ability to measure the differential amount of  
75 carbon stored in land ecosystems as a result of a change in policy compared to a defined  
76 business-as-usual scenario. This presupposes that instruments and methods are in place for  
77 monitoring, reporting and verification of land carbon budgets, yet there remain great challenges  
78 in this area. In many temperate countries, which have largely built their political system around  
79 wood as a key commodity, elaborate systems of forest resource assessment and management  
80 were established early on, and they have been operated by national forest services. Thus,  
81 nationally determined carbon contributions are relatively reliable in the temperate zone, where  
82 forest biomass stocks are based on well-established sample-based forest inventories (Fridman et  
83 al. 2014). However, the political history of many tropical or subtropical countries has been such  
84 that national forest inventory systems are either young or absent, in spite of efforts by the FAO to  
85 set up such systems in several countries since the 1990s (Schimel et al. 2015). This situation is  
86 now changing, with national forest inventories being developed in Brazil and the Democratic  
87 Republic of Congo (Xu et al. 2017).

88 Remotely sensed approaches to estimate carbon stocks have emerged as a solution to this  
89 problem, and several missions are planned in the 2018-2022 period, including BIOMASS (P-  
90 band radar satellite by ESA), NISAR (L-band radar by NASA and ISRO), and GEDI (lidar  
91 onboard the ISS by NASA). These missions will not measure carbon stocks directly, but they  
92 will use proxies of forest structure, volume, and biomass components that correlate with the  
93 aboveground carbon stocks. Canopy height is measured by lidar and polarimetric interferometry,  
94 and tall forests tend to hold more carbon than shorter ones. The second physical quantity related  
95 to forest carbon store is the wood volume and water content which influence the backscattered  
96 electromagnetic energy measured at P-band (~70 cm) or L-band (~25 cm) wavelengths (LeToan  
97 et al. 2011, Saatchi et al. 2011, Shugart et al. 2010). Thus, these missions will collect data that  
98 can be empirically related to forest carbon content.

99 Because forest carbon stores are indirectly inferred from satellite sensors, with  
100 questionable assumptions about their dependence on forest structure and water content, it is  
101 essential that the planned missions make use of accurate ground estimates of carbon stocks to  
102 train their inversion algorithms and validate their products. However, estimating biomass on the  
103 ground is a challenge in itself and ecologists and foresters have struggled with this problem for a  
104 long time. Inevitably, providing inaccurate carbon stock estimates to the Earth Observation (EO)  
105 community will result in uncertain (and potentially biased) carbon maps, and this would have  
106 serious downstream effects on the usefulness of these maps in policy. For instance, even though  
107 pantropical biomass maps inferred from remote sensing have been available for some time now  
108 (Saatchi et al. 2011, Baccini et al. 2012), the IPCC has been reluctant to recommend their  
109 widespread use over national inventories because of possible calibration issues. Here, we offer a  
110 perspective from the ground up, and propose a strategy for gathering reliable ground-based  
111 measurements and biomass estimates that will be useful to the various Earth Observation  
112 missions aimed at quantifying forest structure and carbon stock at a global scale.

113 Overarching principles are summarized here, and echo meetings jointly held on ground  
114 data and upcoming land Earth Observation missions (NASA-ESA-Smithsonian Workshop, 2016;  
115 ISSI ESA meeting, Bern, 2017). First, the focus of all of these missions is primarily tropical.  
116 Many forested extra-tropical countries already have a forest inventory assessment in operation.  
117 In contrast, ground-based monitoring systems are sorely lacking in the tropics (Schimel et al.  
118 2015). Forest extent in the tropics is still very substantial, by far the most living biomass is

119 located in the tropics (63% of carbon in intact tropical forests, against 15% in boreal forests and  
120 13% in temperate forests, according to a comprehensive estimate, Pan et al. 2011). Second, in  
121 order to map change in forest ecosystems, ground sites need to be regularly re-visited (Frolking  
122 et al. 2009). It is impossible financially and logistically to maintain thousands of sites without  
123 long-lasting governmental or international support. These observation systems need to be  
124 maintained at least over the lifetime of the planned missions, but it is likely that they will find  
125 even greater value if made permanent through binding agreements – here we provide costing  
126 estimates to facilitate the discussion of this question, while acknowledging that informed  
127 recommendations for the calibration and validation of the missions are dependent on the nature  
128 of the algorithm, on the resolution of the data, and on the mission duration, and are therefore  
129 beyond the scope of the present study. Third, estimating biomass correctly in situ remains a  
130 delicate business, and strict data quality control must be guaranteed to users.

131

## 132 **2 Principles of ground-based biomass estimation**

133

134 Aboveground biomass (AGB) is the total amount of dry matter of live trees held aboveground in  
135 a plot. It is a crucial parameter for a range of applications, including greenhouse gas accounting,  
136 forest fire assessment, management of the timber industry, monitoring of land-use change, and  
137 ecosystem science. Currently, accurate AGB estimates can be obtained only by labour-intensive  
138 fieldwork (plot inventories) conducted by trained operators. The AGB of each tree is estimated  
139 from measured variables using an allometric model  $AGB=f(\rho,D,H)$ , where  $\rho$  is the stem wood  
140 density of the focal tree, D its trunk diameter, and H its total height. Trunk diameter is usually  
141 measured at breast height (130 cm aboveground) in forests, but in non-forest habitats, trees tend  
142 to branch low, and standard measurements are lower, e.g. at 10 cm aboveground in Australia  
143 (Paul et al. 2016). For the largest trees, diameter needs to be measured above buttresses (Sillett et  
144 al. 2019). Such allometric models are constructed from destructively harvested trees in which  
145 AGB and the other variables are all measured directly. Oven-dry biomass is approximately 47%  
146 carbon, the conversion from AGB to aboveground carbon stores is easy, and the two notions  
147 often used interchangeably. Importantly, a thorough and recent analysis of wood carbon content  
148 showed that carbon content varies predictably among plant functional types, and ranges from

149 41% to 51%, and this source of variation should be carefully considered (Martin et al. 2018).  
150 Since measuring wood density for each tree would be too labour-intensive, recent studies have  
151 used species-mean wood density values, taken from publicly available lookup tables. It follows  
152 that the reliable taxonomic identification of each tree is essential for an accurate estimate of its  
153 AGB.

154 The tree-level AGB estimates are then summed across trees and over the plot, to produce  
155 an aboveground biomass value at the plot level, usually expressed in Mg/ha of dry matter. AGB  
156 is not measured in forests, but estimated (Clark and Kellner 2012), and errors due to the  
157 estimation of height, wood density, and choice of the allometric model usually result in a plot-  
158 based AGB uncertainty that is non-negligible. AGB is estimated to within 50% absolute error for  
159 a single tree (Chave et al. 2014). This absolute error, when propagated at stand level, is around  
160 10% at the 1-ha scale (Réjou-Méchain et al. 2014).

161 In most current applications, live belowground biomass is usually inferred from AGB  
162 using standard root-shoot ratios (Mokany et al. 2006, Paul et al. 2019). Although this poses  
163 important and specific challenges, the issue of estimating biomass components other than AGB  
164 (belowground biomass, soil carbon, coarse woody debris) is not covered in this contribution.

165 In the past few decades, estimation of carbon stocks and AGB in the tropics has been  
166 based on permanent forest plots with measurements of tree trunk diameters, and often tree  
167 mapping and species identification of the trees. AGB can then be estimated per tree, and then  
168 summed over all trees in a stand (Brown and Lugo 1982, Brown 1997). Permanent sampling  
169 forest plots thus are the basic unit of biomass measurement. Their size ranges from 0.01 ha  
170 (10x10 m) to almost 100 ha (1000 x 1000 m). Within these plots, all trees above a given trunk  
171 diameter threshold (usually 10 cm) are censused. These trees are mapped (usually to the nearest  
172 m, or ideally better – but unfortunately quite often much less accurately), they are marked  
173 permanently with a tag, and their trunk diameter is measured at a standard position on the trunk  
174 (usually 130 cm above ground, or 50 cm above buttresses or irregularities, if present). The point  
175 of measurement is noted by a permanent paint mark. A sample of trees may also have their  
176 height dimension measured, and as far as possible, all trees are identified taxonomically. The  
177 numbers of trees  $\geq 10$  cm in one hectare of mature moist forest varies from 300 to 1000. The



178 number of tree species per hectare is up to 30 in temperate forests, and up to 300 in the tropics  
179 (Gentry 1988). Thus, species identification can be a considerable challenge.

180           Temperate and boreal forests present different challenges than tropical forests, and  
181 several temperate countries can rely on National Forest Inventories constructed using statistical  
182 sampling theory. These are designed to provide inferences of biomass and other commodity  
183 values at regional or jurisdictional levels using a large number of small plots (Smith 2002).  
184 These plots can be used to validate the Earth Observation biomass products, but their use for  
185 algorithm development and training is more limited due to plot size and potentially large plot-  
186 level uncertainty of ground biomass. The challenge of mobilizing temperate-country ground  
187 data is an important one, but because most of the world's high biomass forests are in the tropics,  
188 we here emphasize the tropical zone. We also note that vast regions like extratropical Asia are  
189 missing NFIs, and it would be important to better account for in situ forest information in these  
190 regions.

191           Aerial lidar scanning (ALS) has been intensively used for estimating tropical forest  
192 biomass (Drake et al. 2002, Asner et al. 2010), and the literature suggests that if ALS data can be  
193 calibrated locally with permanent sampling plots, the resulting biomass maps are unbiased and  
194 reliable (they have a relative uncertainty of less than 20% at the 1-ha scale). Establishing high-  
195 resolution biomass maps at 1000-ha (10 km<sup>2</sup>) scale would result in a 100-fold increase over plot  
196 data, and the 1000-ha scale is typically the area surveyed around permanent field stations in the  
197 tropics: this means that the sites are within walking distance and the ALS-derived biomass map  
198 can be thoroughly ground-truthed. However, since forests are constantly changing, algorithm  
199 training and validation of ALS data is impossible without near-contemporaneous fieldwork.

200

### 201 **3 High-quality carbon estimates require long-term study sites**

202

#### 203 3.1 Forest dynamics and tree inventories

204 Collections of small plots offer a representative sample of the landscape-scale variability of  
205 biomass, but lack the temporal dimension that is also critically important for understanding the  
206 system. Forest changes include (i) secular changes in mature forests driven externally by climate

207 (e.g. increasing/declining growth rate), (ii) sudden, stochastic changes (e.g., drought-, flood-,  
208 wind-, pest-, fire-induced mortality), (iii) successional development (e.g., savanna thickening  
209 into forest, floodplain forests). Anthropogenic impacts on forests are equally important, complex,  
210 and ubiquitous. As a result, it is not enough to use forest biomass estimates at a given point in  
211 time, but forests should be measured repeatedly in situ. Most existing permanent census plots are  
212 recensused by trained teams of foresters and botanists every 4-5 years, funds permitting, a  
213 monitoring revisit frequency that is sufficient if tree turnover rates are around 1-3%/yr.

214 Another reason for measuring forest stands repeatedly is that some measurement errors  
215 may affect biomass estimates far more than others. A small number of large trees hold a large  
216 fraction of the biomass in a stand, and these are the most difficult to measure in the field.  
217 Assume a 100 × 100 m stand of tropical forest contains around 500 trees of trunk diameter  
218 greater than 10 cm, and the oven-dry aboveground biomass is 300 tons or more, a typical  
219 situation in moist tropical forests. Thus, on average, a tree weighs 0.6 ton. However, the  
220 distribution of tree weights is hugely skewed, since according to one study conducted in intact  
221 tropical forests, 41% of the aboveground biomass was held in trees above 60 cm in trunk  
222 diameter (Lutz et al. 2018). In tropical forests, historical permanent plots were often established  
223 by botanists to explore plant diversity. Initially, little attention may have been paid to carefully  
224 measuring the largest trees, and plots were often located based on convenience more than based  
225 on a sampling protocol. Clark and Clark (2000) provided the first comprehensive study  
226 comparing different carbon sampling strategies in the tropical forest of La Selva, Costa Rica.  
227 They showed that measuring trunk diameter above buttresses was key to a proper estimate of  
228 AGB (see also Condit 1998).

229 It is essential to realize that for many sites, the history of the plots is complex and data  
230 quality may have changed over time. Therefore, the issue is not only to process pre-existing data,  
231 but also to critically appraise the field collection protocols to ensure that legacy data are made  
232 available and associated with an uncertainty assessment that accounts for the history of data  
233 acquisition, that varies greatly from site to site and among groups of data collectors.

234 In the tropics, a major contribution is required from developing country scientists and  
235 technicians. The ground data they produce are hard-won, and need to be repeated at regular  
236 intervals. The participants in our plot networks span hundreds of tropical forested localities

237 across up to 50 different nations. Such work involves obvious logistical complexities of  
238 organizing ground data collection, institutional collaborations, intellectual property, permits and  
239 health and safety protocols to allow remote fieldwork and plant collection across so many  
240 countries while complying with protected area regulations. In addition, a key challenge is to  
241 harmonize datasets and differing existing ground biomass protocols. Consequently, because of  
242 this major effort and the heavy dependence on often specialized human labour, local researchers  
243 in charge of ground-based measurements must be involved as scientific collaborators, and field  
244 teams adequately trained, equipped, insured, and paid.

245

### 246 3.2 Calibration and validation strategies for Earth Observation missions

247 Several biomass EO missions are currently in the process of developing their algorithms or  
248 preparing their validation plans. This includes the GEDI mission, launched in December 2018  
249 (NASA), the NISAR mission (NASA-ISRO, launch in 2021) and the BIOMASS mission (ESA,  
250 launch in 2022). The ground data already collected as part of these efforts are remarkably  
251 similar, even if the requirements differ slightly.

252 The major requirement is that ground biomass values be available, based on intensive tree  
253 inventories, and reliable biomass estimation methods. EO missions have included requirements  
254 about quality assessment of these plot-based biomass estimates, because improperly estimated  
255 ground biomass values are not rare, and failure to account for unreliable data will result in  
256 serious problems in the calibration and validation plan.

257 The three teams in charge of ground data management for GEDI, NISAR and BIOMASS  
258 have recently shared their metadata. The GEDI science team, the most advanced, has assembled  
259 a dataset of 105 sites. These data were contributed by a variety of projects, and are thus in-kind  
260 contributions. They span the major biomes, and represent almost 1400 ha of surveyed plots, of  
261 which 40% are in the Neotropics (557 ha), 12% in Africa (173 ha), and 7% in tropical Asia (108  
262 ha). The NISAR mission cal/val team has assembled data for 77 sites, with quite some overlap  
263 with that of the GEDI science team. BIOMASS is the least advanced, including 6 sites, two in  
264 French Guiana (Neotropics) and four in Gabon, Africa (Labrière et al. 2018), and a total sampled  
265 area of 227 ha. In addition to permanent plot data, all three missions include airborne lidar  
266 scanning (ALS) in their ground dataset. ALS has been shown to be a critically component to EO

267 missions, because it provides invaluable information on forest structure, in a carefully  
268 georeferenced format, and this can be used to upscale plot-based biomass estimates to landscape-  
269 scale biomass maps.

270 In comparison with these datasets, forest plot networks include far more information  
271 (Table 1). For instance, the Smithsonian Institute’s ForestGEO coordinates 245 ha of forest  
272 across 4 sites in Africa, and over 250 ha in tropical Asia (Anderson-Teixeira et al. 2015). The  
273 ForestPlots network (including Rainfor and AfriTRON, plus the Asian project T-Forces),  
274 managed by University of Leeds, coordinates no less than 400 ha of plots in Amazonia alone  
275 (Mitchard et al. 2014), and 315 ha in tropical Africa (Lewis et al. 2013). These two networks  
276 have almost no overlap, and they do not include independent large projects such as forest  
277 management experiments now coordinated by the Tropical managed Forest Observatory  
278 (TmFO), with almost 1200 ha of forests permanently monitored (Sist et al. 2015). Also, a  
279 network of secondary forest plots has been established in the Neotropics and coordinates effort  
280 on forest regeneration (Chazdon et al. 2016). Our estimate is that the area of tropical forests that  
281 are currently monitored globally is in excess of 2500 ha by these four networks, and not  
282 accounting for many more projects. This however remains a minuscule fraction of the total area  
283 covered by forest worldwide, and the biomass estimation challenge is therefore one of upscaling.

284

### 285 3.3 Super-sites

286 Based on our knowledge of available data within the partners, it would be more cost-effective to  
287 prioritize a limited number of ecologically representative sites around the world. We call this the  
288 “supersites” concept. Such sites combine intensive and long-term fieldwork data, airborne  
289 vegetation monitoring, and ancillary information, such that reliable landscape-scale biomass  
290 estimation is possible (Figure 2). This idea of establishing long-term sampling sites with EO  
291 applications in mind is fundamentally the same as that of the US long-term ecological research  
292 sites (LTER) in place since the 1970s, and the International Biological Program (Golley 1993).  
293 The Committee on Earth Observing Satellites (CEOS) Working Group on Land Product  
294 Validation has officially endorsed a supersite concept not only for biomass, but to identify and  
295 promote the collection of validation data for the wide range of Essential Climate Variables

296 products that are currently available or expected in the coming years (see also Duncanson et al.  
297 *this volume*).

298 Meeting high data quality in the tropics is possible at or near research stations with  
299 existing infrastructure and resources, and with the potential to upgrade datasets, funds  
300 permitting. Based on previous experience with the TropiSAR campaign (French Guiana; Dubois-  
301 Fernandez et al. 2012), and AfriSAR mission (Gabon), we propose that these supersites be  
302 selected based on the following specific requirements: (1) Availability of at least 10 already  
303 established 1-ha permanent sampling plots, ideally well-distributed across the landscape,  
304 capturing local gradients of biomass. The plots should be established according the best tropical  
305 forestry standards (see RAINFOR or CTFS protocols; e.g. Condit 1998); (2) Availability of tree  
306 height measurements at each of these plots (for all trees or at least a representative sample of  
307 trees); (3) Availability or potential future collection of ALS coverage over at least 1000 ha,  
308 flown over the permanent plots, with minimal quality requirements (ie such that 1-m canopy  
309 elevation models can be constructed); (4) Availability of a weather station and, optionally,  
310 automated soil moisture monitoring (ideally encompassing the landscape-scale variation of soil  
311 moisture).

312 We also propose to implement terrestrial lidar scanning (TLS) surveys of the permanent  
313 plots. TLS surveys are no substitute for forest tree inventories, but they have the potential to  
314 complement them usefully: they provide an accurate measure of tree volume at tree scale, a  
315 reliable measure of total tree height, and an accurate correction of stem geolocation (relative, at  
316 stand scale). They also give access to the details of forest structure, that may be important in  
317 modelling canopy reflectance at these sites (Calders et al. 2018). This considerably increases the  
318 quality of the key plot data on which all of the other estimates rely. The resulting tree volume  
319 data can also be used to augment existing allometric relationships used to generate tree biomass  
320 estimates, particularly across a much larger range of tree size and including many more large  
321 trees (Disney et al. 2018). Tree volume and tree weight may differ significantly in the case of  
322 hollow trunks, and large trees tend to be more often hollow than small ones (Nogueira et al.  
323 2006, Réjou-Méchain et al. *this volume*). Recently, drone-based alternatives for terrestrial lidar  
324 scanning have been proposed and they present the additional advantages of scanning the canopy  
325 tops, and of producing already stitched point clouds, over large areas typically several hectares

326 (Brede et al. 2017). While this technology still requires development, it would be important to  
327 explore its applicability to the establishment and monitoring of forest supersites.

328

#### 329 3.4 Drivers of biomass stocks and geographical coverage

330 Forest structure varies at all spatial scales. Determining the optimal sampling strategy for Earth's  
331 forests requires research into the drivers of biomass stocks, which in turn depends on spatially  
332 explicit maps of forest structure that will not be available until the new mission products come  
333 online. It is difficult to segment tropical forests worldwide into forest types that would both make  
334 ecological sense and would be optimal for the training of biomass retrieval algorithms. For  
335 instance, two forests may have a similar structure, yet display species with different wood  
336 densities, resulting in very different biomass estimates (Phillips et al., *this volume*). Also, the  
337 forest lower canopy may play a significant role in the radar backscattering properties, and, like  
338 wood density, this is not readily assessed remotely.

339 We therefore provide ecologically-informed guiding principles for the selection of sites.  
340 Tropical forests vary in their structure and floristic composition, and this in turn impacts their  
341 biomass storage capacity (Malhi et al. 2004, Stegen et al. 2011). The four main driving factors of  
342 this variation are soil fertility, moisture supply, elevation, and disturbance regime. Thus, forests  
343 often hold less biomass on very infertile or very fertile soils (Castilho et al. 2006). Also, dry  
344 tropical forests have less biomass, but there is also potentially a hump-shaped distribution of  
345 biomass with respect to annual precipitation whereby ever-wet forests tend to have lower  
346 biomass stocks than moist forests (Brown and Lugo 1982). Elevation is another important factor,  
347 and biomass usually declines with increasing altitude, although some exceptions exist, for  
348 example when trees of the oak family are present (Phillips et al. 2016). Finally, disturbed forests  
349 have a lower biomass than undisturbed ones. The foremost cause of disturbance in the tropics is  
350 anthropic, but other causes exist including wildfires, wind storms, insect predation or diseases,  
351 and the frequency and intensity of natural disturbance exerts a critical control on intact forest  
352 wood density and biomass (Keeling and Phillips 2007, Johnson et al. 2016).

353 In addition, tropical forests have almost zero floristic overlap between the Neotropics  
354 (South America), Africa, Asia, and Oceania (including Papua New Guinea and Australia), with  
355 each biogeographic region having thousands of tree species whose architecture and unique

356 identity helps to determine forest structure and biomass in that region. Assuming that three  
357 conditions (low, medium, high) are selected for each of the four major gradients in each of the  
358 four continents, the number of possibilities is  $4 \times 3^4 = 324$ .

359         Practically, when selecting sites for algorithm training or product validation, it is essential  
360 to include the full range of variability in biomass, i.e. high-biomass forests, typically moist  
361 tropical forests with biomass stocks more than 300-400 tons/ha, and up to 600 tons/ha, but also  
362 low-biomass forests, typically less than 100-200 tons/ha. For instance, a relatively young  
363 secondary forest of ca. 20 years regrowing from clear-cutting holds about 100 tons/ha in tropical  
364 areas (assuming an accumulation rate of 5 tons/ha/yr). Also, woodlands store 30-150 tons/ha. It  
365 would be important to include both secondary vegetation in the study landscapes, and to select  
366 dry vegetation types. These vegetation types are particularly important for the NISAR mission,  
367 which aims at estimating biomass up to 100 Mg/ha, above which L-band backscatter signals  
368 saturate with respect to biomass.

369

## 370 **4 Building on long-term forest plots**

371

### 372 4.1 The Forest Observation System

373 Permanent plots provide the most accurate method for forest biomass estimation, which not only  
374 depends on biometric variables, but also on wood density (species-dependent). Many sites across  
375 the forested tropics have on the order of ten 1-ha plots, scattered around a landscape, because this  
376 sampling intensity is manageable. Much larger sampling intensities do exist but they are rare.  
377 Further, plots are often not established randomly in space.

378         The European Space Agency has funded the Forest Observation System (FOS) as an  
379 effort to coordinate in situ activities in relation with the BIOMASS mission. The FOS includes  
380 several large international consortia who are addressing the issues of ground data sharing and  
381 standardization: ForestPlots.net (including RAINFOR, AfriTRON, and T-FORCES; led from the  
382 University of Leeds), ForestGEO (including CTFS; Smithsonian Institution). These consortia  
383 both have a solid record in tackling key scientific questions, in engaging a community of  
384 collaborators and in standardizing forestry data. For up to 40 years now, they have been devoted

385 to coordinating long-term research with permanent sampling forest plots. They have (i)  
386 established permanent sampling plots in tropical and temperate forests, (ii) encouraged and  
387 carried out extensive plant collection and identification, (iii) proposed robust protocols for  
388 accurate tree mapping, and measurement, (iv) monitored existing plots repeatedly, and (v)  
389 established databases with a special emphasis on data quality control at the tree level, and have  
390 successfully incorporated historical databases.

391 Two additional networks of permanent forest plots have now been invited to join the  
392 Forest Observation System: TmFO (Tropical managed Forest Observatory; Sist et al. 2015) and  
393 AusCover (CSIRO). We are aware that many more groups of scientists and networks of plots  
394 have been established, but when examining inclusion of new sites into the FOS, it is essential to  
395 consider upstream quality assessment. It is preferable to build upon projects that have already  
396 established a data sharing policy, quality assessment procedures, and instruments for  
397 communication with principal investigators at each of the sites.

398 NASA and ESA are also in the process of establishing a Multi-mission Analysis and  
399 Algorithm platform (MAAP), which will house field plot, airborne lidar, and spaceborne  
400 datasets, including data from NISAR, GEDI and BIOMASS (Albinet et al. *this issue*). This will  
401 be a virtual open and collaborative environment, bringing together data, cloud-based computing  
402 resources, and collaborative tools. It will establish a collaboration framework between ESA and  
403 NASA to share data, science algorithms and computing resources in order to foster and  
404 accelerate scientific research conducted by NASA and ESA scientists. We intend for the Forest  
405 Observation System to become an integral part of this multi-mission analysis platform,  
406 facilitating provision of field plot data such as from existing plots and new supersite data  
407 acquisitions.

408 Gathering calibration and validation data relevant to biomass for the Earth Observation  
409 community faces a number of challenges, and the FOS aims to address the most important ones.  
410 We here list the priorities: (1) ensuring the respect of intellectual property rights, (2) providing  
411 site principal investigators with a knowledge of the scientific challenges undertaken with their  
412 data, and (3) ensuring that datasets included in FOS are of the highest possible quality and are  
413 representative of all forest ecosystems.



414 A key aspect of the collaboration is that the intellectual property of the primary data  
415 remains with the site's principal investigator. This principle is upheld in the FOS data sharing  
416 policy (under preparation). Official data sharing policies are found, for instance for the  
417 Smithsonian Institute (white paper 'Sharing Smithsonian Digital Scientific Research Data from  
418 Biology', March 2011), for the RAINFOR project (white paper 'Ethical Code, Data Sharing &  
419 Publication Policy for RAINFOR Participants', June 2009) and in TmFO's Memorandum of  
420 Understanding. Within the FOS, plot consortia are acting on behalf of the site principal  
421 investigators. Importantly, data providers are not asked to provide their primary (tree-by-tree)  
422 data. The data shared in the FOS are stand-level descriptors, including aboveground biomass  
423 estimates, that are obtained from a standardized procedure.

424 One of the most frequent complaints voiced by site principal investigators is that the data  
425 they are providing serve projects downstream that they are not made aware of. This is to a large  
426 extent a communication problem, and one that can be solved through constant interaction with  
427 site principal investigators through a mailing list.

428

#### 429 4.2 Plot data requirements

430 Minimal data requirements are here discussed. These data should be produced by the partners  
431 and provided to the Forest Observation System database.

432 A minimum set of site descriptors are included in the metadata. These include: (a) the name and  
433 contact (email) of the plot principal investigator(s); they should agree to be mentioned in the  
434 database (for privacy protection, this information is made available online in the password-  
435 protected part of the database); (b) the name of the partner institution(s) and individual in charge  
436 of data management; (c) the names of the funding bodies; and (d) some characteristic  
437 photographs of the forest.

438 The following plot information is important: (a) plot coordinates, which should be  
439 checked for the geodetic system and be provided in WGS84; GPS coordinates should be of high  
440 accuracy, typically to within 10 m (but ideally with surveying GPS to within cm), so as to  
441 facilitate co-registration with other data sets (ALS, TLS and EO); plot coordinates should ideally  
442 refer to the centre and the four corners of the plot; (b) collection date and periodicity; number

443 and date of censuses carried out should also be known; the census number for which AGB data  
444 are provided should be given; (c) the total sampled area, i.e. the horizontal projection of the on-  
445 ground sampled area (i.e. topography effects are ignored), and plot geometry; most plots are  
446 squares or rectangular; (d) the dataset should also document the relief (slope, exposition); in  
447 situations where aerial LiDAR is available, this usually provides accurate measurements of  
448 ground relief; (e) forest type (i.e. wet, moist, dry forests) and successional status should be  
449 documented. Note that networks have already faced the issue of post-field data  
450 standardization/filtering. However, it is not established that they all have settled to a common  
451 practice.

452 We also report on metadata for the tree inventory itself: (a) the number of trees  $\geq 10$  cm in  
453 trunk diameter; note that trees  $< 10$  cm and other life forms are usually excluded in AGB  
454 estimates in case their contribution to AGB is less than 5%; (b) a quality assessment index  
455 should be devised, reporting on whether points of measurement have been properly recorded for  
456 each tree; (c) an index reporting on the quality of taxonomic identification will also be needed; as  
457 a rough measure, the proportion of trees identified to species level, genus level, and family level,  
458 is reported. In tropical forests, identification of less than 50% of the trees to species level is far  
459 from unusual. Careful botanical identification by botany experts results in identification rates of  
460  $> 90\%$  of the trees, but may entail climbing trees to collect and significant down-stream  
461 identification effort with botanists and herbaria; (d) plot-averaged wood density is the basal-area  
462 weighted wood density of the trees in a plot. For plots with reliable taxonomic identifications,  
463 this may be deduced from census data and species-average wood density values; (e) mean  
464 canopy height of the plot, as inferred from direct tree height measurements or from airborne  
465 LiDAR measurements; if necessary, several canopy height metrics should be provided; quality-  
466 control metric: height of the largest measured tree, trunk diameter of the largest measured tree.

467 Finally, above ground biomass and confidence intervals are computed and provided at the  
468 plot scale, following an agreed single methodology across partners; the methodology will be  
469 made accessible for each database release, and partners should be prepared to adapt to changes in  
470 the methodology. An efficient strategy is to jointly develop a statistical routine such that several  
471 database formats can be accommodated, and that perform the tasks of calculating biomass and  
472 canopy height at each site. The R statistical software is recommended because it is free, already  
473 widely used in the ecological research community, and networks such as ForestGEO or

474 ForestPlots already have developed R routines for parsing the datasets and performing quality  
475 checks. We have established a package called BIOMASS that calculates biomass values and  
476 propagates uncertainty from tree measurement to stand-level estimates (Réjou-Méchain et al.  
477 2017). This package is flexible and makes it possible to use user-supplied conditions or  
478 allometric equations.

#### 479 4.3 Candidate supersites, and their coverage of environmental gradients

480 There are around 50-100 supersites already potentially available worldwide, and here we discuss  
481 a list of 78 sites included as priority sites by the ESA-NASA cross-mission working group. All  
482 sites share a number of basic features including a long-term presence of scientists, existing forest  
483 monitoring programs, and willingness to collaborate in international scientific projects on the  
484 part of the principal investigators.

485 Taken together, these sites encompass much of the variability in forest types, and within  
486 each 1000-ha region of interest, these sites display a large spectrum in biomass ranges and  
487 disturbance histories. Figure 3 illustrates the location of sites that could be prioritized as  
488 supersites. A majority of the supersites are located in the tropics, reflecting the more pressing  
489 need for data in tropical forest environments. However, several sites were also selected outside  
490 of the tropical belt.

491 We also illustrate the coverage of these sites in terms of biomes and bioclimatic conditions  
492 (Figure 4). The 78 sites currently being considered for the network of supersites span broad  
493 bioclimatic conditions, and although they are mostly located below 1000 m in elevation, a few  
494 sites (n=6) are above this limit. As seen in Figure 4, the current list of supersites does not include  
495 many dry forests, semi-deciduous tropical forests, or boreal forests. Also, warm temperate forests  
496 are currently under-sampled in this dataset. Finally, a large proportion of the sites are currently  
497 located in areas with less than 1% disturbance from 2000 to 2017 (28 out of 78) but some are in  
498 highly disturbed landscapes. One example is the STREK site in Indonesia (TmFO), in which  
499 over 60% of the surrounding landscape has been deforested since 2000, another example being  
500 the Pasoh plot (ForestGEO) with over 40% of deforestation since 2000.

501

## 502 **5 Conclusion. Building a ground-based Earth Observation mission**

503

504 There has been a shift over the past few decades toward freely available Earth Observation data,  
505 and NASA and ESA have adopted open data policies with the aim of accelerating science and  
506 applications (Turner et al., 2015). Earth Observation data are typically delivered free at the point  
507 of use, using technology that has cost space agencies and their funding governments hundreds of  
508 millions of dollars or euros to develop and launch. Once the sensors are installed in orbit they  
509 continue to supply data at, relatively, limited recurrent cost. Because nations have provided the  
510 core investment, they can rightly insist that Earth Observation data are provided for free to the  
511 entire scientific community (although conditions of use may vary among space agencies).

512 In situ ground-based datasets stand in stark contrast with this situation, because the more  
513 reliable data are obtained by human specialists, who are paid for gathering them, verifying them,  
514 and maintaining databases over decades. In addition to data collection costs, data curation and  
515 coordination is also costly, and these costs do not come down with time. Most of the ground  
516 forest stand data available as of 2018, and summarized above, were collected and processed  
517 through long-term collaborations and with funds mobilized by the scientific community of many  
518 countries, and for a multitude of purposes. Few were collected with the express purpose of  
519 calibrating or validating remote imagery. It should not be assumed that the level of funding  
520 provided to these science projects will persist with the same intensity from 2019 to 2029. The  
521 majority of the principal investigators reside in countries with limited support from national  
522 science funding, hence relying on international collaboration to sustain their activities. It is  
523 reasonable to suppose that if the substantial future ground effort proposed in this chapter is to be  
524 effectively used to support remote-sensing missions, then it needs to be funded to do so.

525 It was estimated by FOS partners that the full cost of recensusing a single 1-ha plot in  
526 high-diversity tropical forests is on the order of 15 k€ (2016 economic conditions). This reflects  
527 the entire cost from concept to delivery of the highest possible quality data with accurate tree  
528 dimensions and identification. For instance, to fully recensus 600 1-ha plots across all four  
529 tropical continents included in the ForestPlots database, the full cost would be 9 M€ per  
530 remeasurement cycle. A similar figure is to be expected for the ForestGEO network. These costs  
531 are indicative, but result from decades of experience in establishing and maintaining tropical

532 forest plots across the world. If satellite estimates of biomass are to be of high quality and serve  
533 the widest use, these costs should be factored in the calibration and validation strategies of EO  
534 missions.

535 A coordinated and global ground-based monitoring of forests would benefit several  
536 sectors of science and the society, and would be of direct use to biomass-related spaceborne  
537 missions. It would allow to collect and maintain ground-based databases for the lifetime of the  
538 currently planned missions, and potentially for longer periods. In addition, this ‘ground mission’  
539 would help consolidate the remote-sensing/ecology nexus, helping bridge the gap between these  
540 two scientific communities and accelerate the valorisation of both ground and remotely-sensed  
541 data. Finally, the study sites could be valorised beyond the currently planned biomass missions.  
542 For instance, several missions are committed to measuring photosynthetic activity through solar-  
543 induced fluorescence, or aim to monitor biodiversity using hyperspectral imagery. In these  
544 situations, it is also essential to validate the concept of these missions at a set of reference sites  
545 that can be appropriately accessed, equipped and maintained.

546

## 547 **References**

548 Ashton PS (1964) *Ecological Studies in the Mixed Dipterocarp Forests of Brunei State*,  
549 Clarendon Press Oxford.

550 Asner GP, Powell GV, Mascaro J, Knapp DE, Clark JK, Jacobson J et al. (2010) High-resolution  
551 forest carbon stocks and emissions in the Amazon. *Proceedings of the National Academy of*  
552 *Sciences*, 107:16738-16742.

553 Anderson-Teixeira KJ, Davies SJ, Bennett AC, Gonzalez-Akre EB, Muller-Landau HC, Wright  
554 SJ et al. (2015) CTFIS-Forest GEO: a worldwide network monitoring forests in an era of global  
555 change. *Global Change Biology*, 21:528-549.

556 Avitabile V, Herold M, Heuvelink GB, Lewis SL, Phillips OL, Asner GP et al. (2016) An  
557 integrated pan-tropical biomass map using multiple reference datasets. *Global Change Biology*,  
558 22:1406-1420.

559 Baccini A, Goetz SJ, Walker WS, Laporte NT, Sun M, Sulla-Menashe D, Hackler J et al. (2012)  
560 Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density  
561 maps. *Nature Climate Change* 2:182–185.

562 Brede B, Lau A, Bartholomeus HM, Kooistra L (2017) Comparing RIEGL RiCOPTER UAV  
563 LiDAR derived canopy height and DBH with terrestrial LiDAR. *Sensors*, 17:2371.

564 Brown S, Lugo AE (1982) The storage and production of organic matter in tropical forests and  
565 their role in the global carbon cycle. *Biotropica*, 161-187.

566 Calders K, Origo N, Burt A, Disney MI, Nightingale J, Raunonen P, Åkerblom M, Malhi Y,  
567 Lewis P (2018) Realistic Forest Stand Reconstruction from Terrestrial LiDAR for Radiative  
568 Transfer Modelling, *Remote Sensing*, 10:933.

569 Castilho CV, Magnusson WE, de Araújo RN, Luizao RC, Luizao FJ, Lima AP, Higuchi N.  
570 (2006) Variation in aboveground tree live biomass in a central Amazonian Forest: Effects of soil  
571 and topography. *Forest Ecology and Management*. 234:85-96

572 Chave J, Réjou-Méchain M, Búrquez A, Chidumayo E, Colgan MS, Delitti WB et al. (2014)  
573 Improved allometric models to estimate the aboveground biomass of tropical trees. *Global  
574 Change Biology*, 20:3177-3190.

575 Chazdon RL, Broadbent EN, Rozendaal DM, Bongers F, Zambrano AMA, Aide TM et al.  
576 (2016) Carbon sequestration potential of second-growth forest regeneration in the Latin  
577 American tropics. *Science Advances*, 2(5), e1501639.

578 Clark DB, Clark DA (2000) Landscape-scale variation in forest structure and biomass in a  
579 tropical rain forest. *Forest Ecology and Management*, 137:185-198.

580 Clark DB, Kellner JR (2012) Tropical forest biomass estimation and the fallacy of misplaced  
581 concreteness. *Journal of Vegetation Science*, 23:1191-1196.

582 Condit, R. (1998) *Tropical forest census plots: methods and results from Barro Colorado Island,  
583 Panama and a comparison with other plots*. Springer Science & Business Media.

584 Disney MI, Boni Vicari M, Calders K, Burt A, Lewis S, Raunonen P, Wilkes P (2018)  
585 Weighing trees with lasers: advances, challenges and opportunities, *Royal Society Interface  
586 Focus*, 8.

587 Drake JB, Dubayah RO, Clark DB, Knox RG, Blair JB, Hofton MA et al. (2002) Estimation of  
588 tropical forest structural characteristics using large-footprint lidar. *Remote Sensing of*  
589 *Environment*, 79:305-319.

590 Dubois-Fernandez PC, Le Toan T, Daniel S, Oriot H, Chave J, Blanc L et al. (2012) The  
591 TropiSAR airborne campaign in French Guiana: Objectives, description, and observed temporal  
592 behavior of the backscatter signal. *IEEE Transactions on Geoscience and Remote Sensing*, 50:  
593 3228-3241.

594 Fridman J, Holm S, Nilsson M, Nilsson P, Ringvall, A. H., & Ståhl, G. (2014) Adapting National  
595 Forest Inventories to changing requirements—the case of the Swedish National Forest Inventory  
596 at the turn of the 20th century. *Silva Fennica*, 48(3), 29.

597 Chambers JQ, Shugart HH, Hurtt GC (2009) Forest disturbance and recovery: A general review  
598 in the context of spaceborne remote sensing of impacts on aboveground biomass and canopy  
599 structure. *Journal of Geophysical Research: Biogeosciences*, 114(G2).

600 Gentry, A. H. (1988) Changes in plant community diversity and floristic composition on  
601 environmental and geographical gradients. *Annals of the Missouri Botanical Garden*, 1-34.

602 Golley FB (1993) *A history of the ecosystem concept in ecology: more than the sum of the parts*.  
603 Yale University Press.

604 Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SAA, Tyukavina A et al. (2013)  
605 High-resolution global maps of 21st-century forest cover change. *science*, 342:850-853.

606 IPCC (2018) Special Report on Global Warming of 1.5°C. Available at  
607 <https://www.ipcc.ch/sr15/>

608 Johnson MO, Galbraith D, Gloor M, De Deurwaerder H, Guimberteau M, Rammig A et al.  
609 (2016) Variation in stem mortality rates determines patterns of above-ground biomass in  
610 Amazonian forests: implications for dynamic global vegetation models. *Global Change Biology*,  
611 22:3996-4013.

612 Keeling, HC, & Phillips, OL. (2007) The global relationship between forest productivity and  
613 biomass. *Global Ecology and Biogeography*, 16:618-631.

614 Labrière N, Tao S, Chave J, Scipal K, Le Toan T, Abernethy K et al. (2018) In Situ Reference  
615 Datasets From the TropiSAR and AfriSAR Campaigns in Support of Upcoming Spaceborne  
616 Biomass Missions. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote*  
617 *Sensing*, (99), 1-11.

618 Le Quéré C, Andrew RM, Friedlingstein P, Sitch S, Pongratz J, Manning AC et al. (2017) Global  
619 carbon budget 2017. *Earth System Science Data*, 10:405-448, 2018.

620 Le Toan T, Quegan S, Davidson MWJ, Balzter H, Paillou P, Papathanassiou K et al. (2011) The  
621 BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon  
622 cycle. *Remote Sensing of Environment*, 115:2850-2860.

623 Lewis SL, Sonké B, Sunderland T, Begne SK, Lopez-Gonzalez G, Van Der Heijden GM et al.  
624 (2013) Above-ground biomass and structure of 260 African tropical forests. *Phil. Trans. R. Soc.*  
625 *B*, 368:20120295.

626 Martin, A. R., Doraisami, M., & Thomas, S. C. (2018). Global patterns in wood carbon  
627 concentration across the world's trees and forests. *Nature Geoscience*, 11(12), 915.

628 Malhi Y, Wood D, Baker TR, Wright J, Phillips OL, Cochrane T, Meir P et al. (2004) The  
629 regional variation of aboveground live biomass in old-growth Amazonian forests. *Global*  
630 *Change Biology* 12:1107-1138

631 Mitchard ET, Feldpausch TR, Brienen RJ, Lopez-Gonzalez G, Monteagudo A, Baker TR, et al.  
632 (2014) Markedly divergent estimates of Amazon forest carbon density from ground plots and  
633 satellites. *Global Ecology and Biogeography*, 23:935-946.

634 Mokany K, Raison RJ, Prokushkin AS (2006) Critical analysis of root:shoot ratios in terrestrial  
635 biomes. *Global Change Biology* 12:84–96.

636 NASA-ESA-Smithsonian Workshop on Calibration and Validation of Upcoming Satellite  
637 Missions on Forest Structure and Biomass, Washington DC, 2016  
638 ([https://nisar.jpl.nasa.gov/files/nisar/NISAR\\_Vegetation\\_Biomass\\_Workshop\\_Report.pdf](https://nisar.jpl.nasa.gov/files/nisar/NISAR_Vegetation_Biomass_Workshop_Report.pdf)).

639 Nogueira EM, Nelson BW, Fearnside PM (2006). Volume and biomass of trees in central  
640 Amazonia: influence of irregularly shaped and hollow trunks. *Forest Ecology and Management*,  
641 227:14-21.



642 Pan Y, Birdsey RA, Fang J, Houghton R, Kauppi PE, Kurz WA et al. (2011) A large and  
643 persistent carbon sink in the world's forests. *Science*, 1201609.

644 Paul KI, Roxburgh SH, Chave J, England JR, Zerihun A, Specht A et al. (2016) Testing the  
645 generality of above-ground biomass allometry across plant functional types at the continent  
646 scale. *Global Change Biology*, 22, 2106-2124.

647 Paul KI, Larmour J, Specht A, Zerihun A, Ritson P, Roxburgh SH et al. (2019) Testing the  
648 generality of below-ground biomass allometry across plant functional types. *Forest Ecology and  
649 Management*, 432:102-114.

650 Phillips J, Duque A, Scott C, Wayson C, Galindo G, Cabrera E et al. (2016) Live aboveground  
651 carbon stocks in natural forests of Colombia. *Forest Ecology and Management*, 374:119-128.

652 Réjou-Méchain M, Muller-Landau HC, Detto M, Thomas SC, Toan TL, Saatchi SS et al. (2014)  
653 Local spatial structure of forest biomass and its consequences for remote sensing of carbon  
654 stocks. *Biogeosciences*, 11, 5711.

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

658 Saatchi SS, Harris NL, Brown S, Lefsky M, Mitchard ET, Salas W, Zutta BR et al. (2011)  
659 Benchmark map of forest carbon stocks in tropical regions across three continents. Proceedings  
660 of the National Academy of Sciences USA 108:9899–9904.

661 Saatchi S, Marlier M, Chazdon RL, Clark DB, Russell AE (2011) Impact of spatial variability of  
662 tropical forest structure on radar estimation of aboveground biomass. *Remote Sensing of  
663 Environment*, 115:2836-2849.

664 Schimel D, Pavlick R, Fisher JB, Asner GP, Saatchi S, Townsend P et al. (2015) Observing  
665 terrestrial ecosystems and the carbon cycle from space. *Global Change Biology*, 21:1762-1776.

666 Shugart HH, Saatchi S, Hall FG (2010) Importance of structure and its measurement in  
667 quantifying function of forest ecosystems. *Journal of Geophysical Research: Biogeosciences*,  
668 115(G2).

669 Sillett SC, Van Pelt R, Carroll AL, Campbell-Spickler J, Coonen EJ, Iberle B (2019) Allometric  
670 equations for Sequoia sempervirens in forests of different ages. *Forest Ecology and*  
671 *Management*, 433, 349-363.

672 Sist P, Rutishauser E, Peña-Claros M, Shenkin A, Hérault B, Blanc L et al. (2015) The Tropical  
673 managed Forests Observatory: a research network addressing the future of tropical logged  
674 forests. *Applied Vegetation Science*, 18:171-174.

675 Smith, W. B. (2002) Forest inventory and analysis: a national inventory and monitoring program.  
676 *Environmental Pollution*, 116:S233-S242.

677 Stegen JC, Swenson NG, Enquist BJ, White EP, Phillips OL, Jørgensen PM, Weiser MD,  
678 Mendoza AM, Vargas PN (2011) Variation in above-ground forest biomass across broad  
679 climatic gradients. *Global Ecology and Biogeography*. 20:744-54.

680 Turner W, Rondinini C, Pettorelli N, Mora B, Leidner AK, Szantoi Z et al. (2015) Free and  
681 open-access satellite data are key to biodiversity conservation. *Biological Conservation*,  
682 182:173-176.

683 Xu L, Saatchi SS, Shapiro A, Meyer V, Ferraz A, Yang Y et al. (2017) Spatial distribution of  
684 carbon stored in forests of the Democratic Republic of Congo. *Scientific Reports*, 7:15030.

685

686

687

688 **Tables**

689

690

691 Table 1. List of characteristics for some of the international tropical forest monitoring networks  
 692 currently in operation.

<b>Network</b>	<b>Countries</b>	<b>Plots</b>	<b>Plot Sample Area (ha)</b>	<b>Trees</b>	<b>Species</b>	<b>Measurements</b>	<b>Forest Types</b>	<b>Regional Focus</b>
<b>ForestGEO</b>	26	65	16 – 120	6.5 million	12,000	20 million	Primary	Global
<b>RAINFOR</b>	9	400	0.2 – 9, mostly 1	280,000	5,500	2 million	Primary	South America
<b>AfriTRON</b>	11	320	0.2 – 10, mostly 1	170,000	1,800	600,000	Primary	Africa
<b>TmFO</b>	10	517	0.25 - 27, mostly 1	300,000-400,000	-	~6 million	Logged	Pantropical

693

694

695

696 **Figure captions**

697

698 **Fig 1** Permanent plot in intact lowland tropical forest at the Nouragues station, French Guiana.  
699 Left: Large, buttressed, trees encompass the majority of the aboveground biomass stock. Right:  
700 A pioneer tree (*Cecropia sciadophylla* Mart.) has grown from the top of a palm, causing issues  
701 of trunk diameter measurement. Situations like this one are resolved only with proper field  
702 protocols

703

704 **Fig 2** The supersite concept. Relatively few sites with long-term investment by plot principal  
705 investigators, and the potential to upgrade the sites. The background is taken from Ashton (1964,  
706 Kuala Belalong, Brunei)

707

708 **Fig 3** Potential location of 78 candidate supersites. Proposed sites were selected to maximize  
709 geographical coverage, environmental and forest structure conditions, and logistical constraints  
710 of maintaining long-term sites. The background is Avitabile et al. (2016) carbon stock map

711

712 **Fig 4** (a) Environmental coverage of the 78 candidate supersites in bioclimatic space (Whittaker  
713 diagram). (b) Distribution of the candidate supersites across the range of elevation (in m above  
714 sea level); drought stress, as measured by the climate water deficit: larger values represent more  
715 stressed environments (sites above 500 are usually ascribed to xeric habitats). In both panels,  
716 disturbance intensity is displayed with large red dots representing highly disturbed sites, while  
717 small yellow dots represent undisturbed habitats. Disturbance intensity was measured by the  
718 proportion of forest pixels lost between 2000 and 2017 in a buffer of 5-km radius around the  
719 supersites, using the global 30-m resolution Landsat (Hansen et al. 2013)

720

721

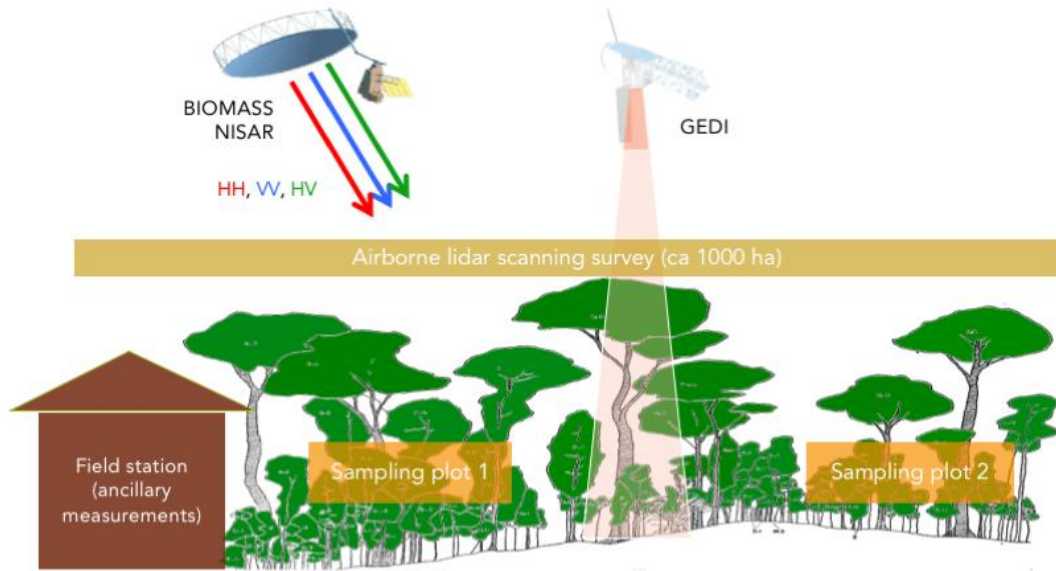


722

723

724

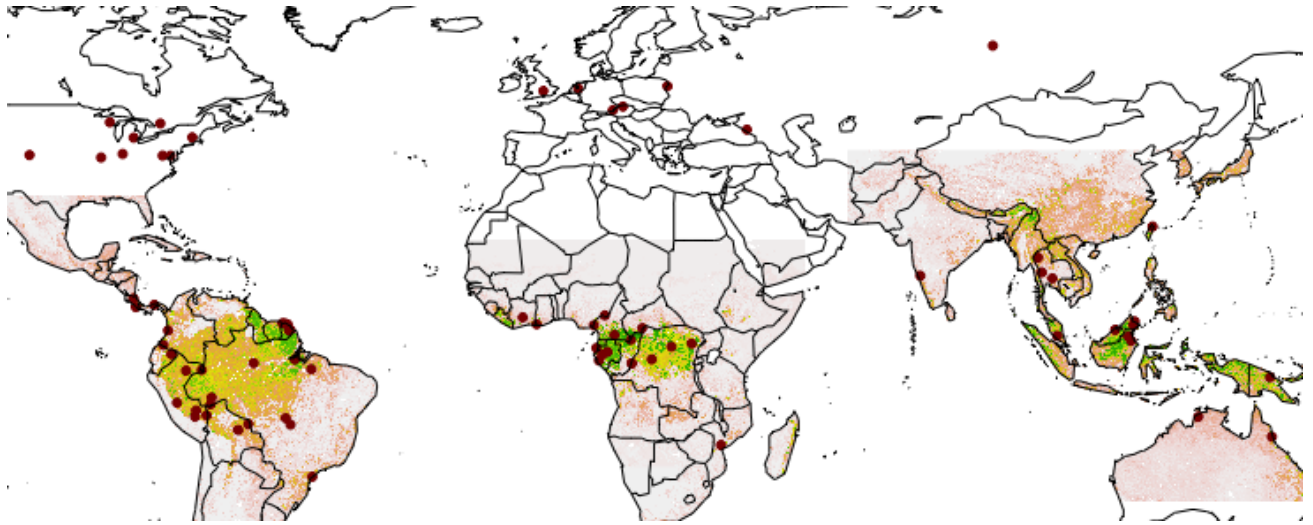
725



726

727

728



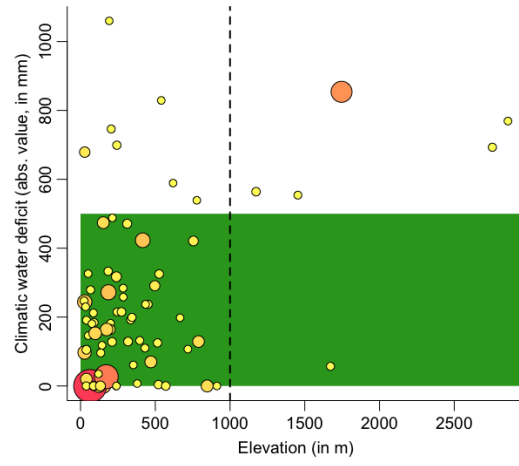
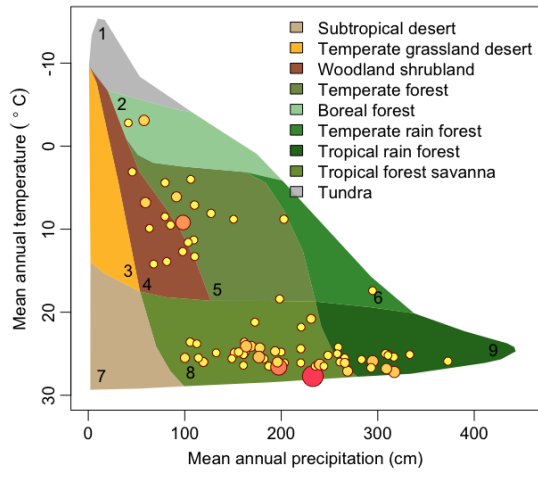
729

730

731

732

733



734

735