

1 Pantropical modelling of canopy functional traits using Sentinel-2 remote sensing data

2
3 Jesús Aguirre-Gutiérrez^{a,b,*}, Sami Rifai^a, Alexander Shenkin^a, Imma Oliveras^a, Lisa Patrick Bentley^c, Martin
4 Svátek^d, Cécile A. J. Girardin^a, Sabine Both^e, Terhi Riutta^a, Erika Berenguer^a, W. Daniel Kissling^f, David
5 Bauman^{a,g}, Nicolas Raab^a, Sam Moore^a, William Farfan-Rios^{h,i,j}, Axa Figueiredo^k, Simone Matias Reis^{a,l},
6 Josué Edzang Ndong^m, Fidèle Evouna Ondo^m, Natacha N'ssi Bengoneⁿ, Vianet Mihinhouⁿ, Marina Maria
7 Moraes de Seixas^o, Stephen Adu-Bredu^p, Kate Abernethy^{q,r}, Gregory P. Asner^s, Jos Barlow^{t,u}, David F.R.P.
8 Burslem^v, David A. Coomes^w, Lucas A. Cernusak^x, Greta C. Dargie^y, Brian J. Enquist^z, Robert M. Ewers^{aa},
9 Joice Ferreira^u, Kathryn J. Jeffery^{aa}, Carlos A. Joly^{ab}, Simon L. Lewis^{y,ac}, Ben Hur Marimon-Junior^l, Roberta
10 E. Martin^s, Paulo S. Morandi^l, Oliver L. Phillips^y, Carlos A. Quesada^{ad}, Norma Salinas^{a,ae}, Beatriz Schwantes
11 Marimon^l, Miles Silman^{af}, Yit Arn Teh^{ag}, Lee J. T. White^{n,q,r}, Yadvinder Malhi^a

12
13 ^aEnvironmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford,
14 UK

15 ^bBiodiversity Dynamics, Naturalis Biodiversity Center, Leiden, The Netherlands

16 ^cDepartment of Biology, Sonoma State University, 1801 East Cotati Avenue, Rohnert Park, CA 94928, USA

17 ^dDepartment of Forest Botany, Dendrology and Geobiocoenology, Faculty of Forestry and Wood
18 Technology, Mendel University in Brno, Brno, Czech Republic

19 ^eEnvironmental and Rural Science, University of New England, Armidale, 2351 NSW, Australia

20 ^fInstitute for Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam, Amsterdam, The
21 Netherlands

22 ^gLaboratoire d'Écologie Végétale et Biogéochimie, CP 244, Université Libre de Bruxelles, Brussels, Belgium

23 ^hLiving Earth Collaborative, Washington University in Saint Louis, St. Louis, MO, USA

24 ⁱCenter for Conservation and Sustainable Development, Missouri Botanical Garden, St. Louis, MO, USA

25 ^jHerbario Vargas (CUZ), Escuela Profesional de Biología, Universidad Nacional de San Antonio Abad del
26 Cusco, Cusco, Peru

27 ^kNational Institute of Amazonian Research – INPA. C.P. 2223, 69080-971, Manaus, AM, Brazil

28 ^lLaboratório de Ecologia Vegetal (LABEV), Universidade do Estado de Mato Grosso, Nova Xavantina, Brazil

29 ^mAgence Nationale des Parcs Nationaux, BP20379, Libreville, Gabon

30 ⁿMinistère des Eaux, des Forêts, de la Mer et de L'Environnement, Libreville, Gabon

31 ^oEmbrapa Amazônia Oriental, Trav. Dr. Enéas Pinheiro, s/n, CP 48, 66095-100 Belém, PA, Brazil

32 ^pCSIR-Forestry Research Institute of Ghana, University P.O. Box 63, Kumasi, Ghana

33 ^qInstitut de Recherche en Écologie Tropicale, Libreville, Gabon.

34 ^rBiological and Environmental Sciences, University of Stirling, Stirling, UK.

35 ^sCenter for Global Discovery and Conservation Science, Arizona State University, Tempe, AZ, United States
36 ^tLancaster Environment Centre, Lancaster University, Lancaster LA1 4YQ, UK
37 ^uMCT/Museu Paraense Emílio Goeldi, Av. Magalhães Barata 376, São Braz, 66040-170 Belém, PA, Brazil
38 ^vSchool of Biological Sciences, University of Aberdeen, Aberdeen, UK
39 ^wConservation Research Institute, Department of Plant Sciences, University of Cambridge, Cambridge CB2
40 3QZ, UK
41 ^xCollege of Science and Engineering, James Cook University, Cairns, Qld, 4878 Australia
42 ^yEcology and Global Change, School of Geography, University of Leeds, Leeds, West Yorkshire, UK
43 ^zDepartment of Ecology and Evolutionary Biology, University of Arizona, Tucson, Arizona, USA
44 ^{aa}Department of Life Sciences, Imperial College London, Ascot, UK
45 ^{ab}Universidade Estadual de Campinas, Instituto de Biologia, Departamento de Biologia Vegetal, Campinas,
46 São Paulo, Brazil
47 ^{ac}Department of Geography, University College London, London, UK
48 ^{ad}Coordenação de Dinâmica Ambiental, Instituto Nacional de Pesquisas da Amazônia, Manaus, Brazil
49 ^{ae}Sección Química, Pontificia Universidad Católica del Perú, Avenida Universitaria 1801, San Miguel, Lima
50 32, Peru
51 ^{af}Department of Biology, Wake Forest University, Winston-Salem, NC 27109, USA
52 ^{ag}School of Natural and Environmental Sciences, Newcastle University, Newcastle Upon Tyne, UK
53 *Corresponding author

54

55 **Abstract**

56 Tropical forest ecosystems are undergoing rapid transformation as a result of changing environmental
57 conditions and direct human impacts. However, we cannot adequately understand, monitor or simulate
58 tropical ecosystem responses to environmental changes without capturing the high diversity of plant
59 functional characteristics in the species-rich tropics. Failure to do so can oversimplify our understanding
60 of ecosystems responses to environmental disturbances. Innovative methods and data products are
61 needed to track changes in functional trait composition in tropical forest ecosystems through time and
62 space. This study aimed to track key functional traits by coupling Sentinel-2 derived variables with a
63 unique data set of precisely located *in-situ* measurements of canopy functional traits collected from 2434
64 individual trees across the tropics using a standardised methodology. The functional traits and vegetation
65 censuses were collected from 47 field plots in the countries of Australia, Brazil, Peru, Gabon, Ghana, and

66 Malaysia, which span the four tropical continents. The spatial positions of individual trees above 10 cm
67 diameter at breast height (DBH) were mapped and their canopy size and shape recorded. Using geo-
68 located tree canopy size and shape data, community-level trait values were estimated at the same spatial
69 resolution as Sentinel-2 imagery (i.e. 10 m pixels). We then used the Geographic Random Forest (GRF) to
70 model and predict functional traits across our plots. We demonstrate that key plant functional traits can
71 be measured at a pantropical scale using the high spatial and spectral resolution of Sentinel-2 imagery in
72 conjunction with climatic and soil information. Image textural parameters were found to be key
73 components of remote sensing information for predicting functional traits across tropical forests and
74 woody savannas. Leaf thickness ($R^2=0.52$) obtained the highest prediction accuracy among the
75 morphological and structural traits and leaf carbon content ($R^2 = 0.70$) and maximum rates of
76 photosynthesis ($R^2 = 0.67$) obtained the highest prediction accuracy for leaf chemistry and photosynthesis
77 related traits, respectively. Overall, the highest prediction accuracy was obtained for leaf chemistry and
78 photosynthetic traits in comparison to morphological and structural traits. Our approach offers new
79 opportunities for mapping, monitoring and understanding biodiversity and ecosystem change in the most
80 species-rich ecosystems on Earth.

81 Keywords: Plant traits, Sentinel-2, Tropical forests, Random Forest, Pixel-level predictions, Image texture

82 **1. Introduction**

83 Some of the most urgent questions in ecology and ecosystem science today focus on how communities
84 of organisms respond to global environmental changes (Naeem et al., 2009), how biodiversity and
85 ecosystem changes across the world can be consistently mapped and monitored (Navarro et al., 2017),
86 and how spatial, temporal and taxonomic variability in biodiversity influences ecosystem resilience to
87 climate change (Oliver et al., 2015). In terms of Earth system science, we need to understand and model
88 how the terrestrial biosphere will respond (and already is responding) to global environmental change,
89 and whether there are critical thresholds or “tipping points” beyond which major biomes may not be able
90 to recover. Nowhere is the challenge more urgent than in the species-rich tropical forest and woody

91 savanna biomes, which together are home to more than half of global biodiversity and over 60% of
92 terrestrial productivity (Beer et al., 2010). There is evidence that atmospheric change may have effects on
93 tropical forest productivity and tree functional composition (Esquivel-Muelbert et al., 2019, Hubau et al.,
94 2020). These effects may include a stimulation of productivity (perhaps due to rising CO₂) and/or a
95 degradation or dieback, possibly caused by increased seasonality and incurred intensity of extreme
96 drought events (Malhi et al., 2008, Malhi et al., 2018). Such events are partly responsible for the increased
97 tree mortality and decreased carbon residence time in tropical forests worldwide (McDowell et al., 2018).
98 However, to adequately understand such responses we need to capture and map the high diversity of
99 plant ecosystem function in the species-rich tropics and savannas.

100 Species functional traits are defined as the morphological, physiological or phenological attributes
101 which determine the fitness of organisms, their response to changes in the environment and their
102 influence on ecosystem functions (Kissling et al., 2018, Díaz & Cabido, 2001). Functional traits provide
103 tangible and mechanistic means of assessing the ability of communities to adapt to climate change
104 (Pacifi et al., 2015) and play a major role in determining ecosystem productivity, functioning and notably
105 nature's contribution to people (e.g. water and wood availability) (Díaz et al., 2019, Carmona et al., 2016).
106 Any tools or methods that facilitate quantification of functional traits across large spatial scales and at
107 high spatial resolution would be invaluable for quantifying ecosystem functioning and ecological
108 responses to disturbance at scales relevant for policy and management (Kissling et al., 2018). However, it
109 is still challenging to map functional trait diversity in tropical regions given the lack of plant trait data
110 available for most of those locations (Jetz et al., 2016). Additional challenges come from different and
111 often incompatible trait collection protocols and the lack of systematic high spatial, spectral and temporal
112 resolution remote sensing imagery that coincides with data for functional traits at the canopy level and
113 the lack of geo-located tree stems at the plot level. Thus, there is a need for spatially-explicit methods to
114 map and quantify plant functional traits at high spatial resolution in tropical forest and woody savanna
115 ecosystems.

116 Tracking functional traits can shed light on differences in ecosystem functioning across broad
117 spatial extents and therefore aid policy and decision making, e.g. for creating adequate biodiversity
118 conservation policies or for providing early warning of directional shifts in ecosystems. The key challenges
119 of any functional trait approach are scalability and monitoring: how can functional shifts in highly diverse
120 tropical forests and woody savannas be monitored and tracked over large spatial extents? Intensive field
121 sampling of plant functional traits at a pantropical scale is time-consuming and economically unviable.
122 There are large gaps in the availability of plant trait data globally, and the largest gaps are in the tropics
123 (Jetz et al., 2016). Large plant trait datasets aim to overcome this issue and have advanced our ability to
124 carry out plant functional trait analysis in an unprecedented way (Kattge et al., 2020, Gallagher et al.,
125 2020). However, as with any database, the plant trait values from such databases will represent the local
126 trait-environment relationships for the site where they were collected, which may not be the area of
127 interest. A key assumption in trait-based ecology is that the environment is filtering for an optimal set of
128 trait characteristics so that the resulting communities are adapted to the environment where they are
129 distributed (Fell & Ogle, 2018, Lebrija-Trejos et al., 2010, Lortie et al., 2004). Hence, we might expect an
130 optimal set of trait characteristics for a given location, which when analysed over time could quantify the
131 dynamics of community trait distributions or shifts in functional composition relating to environmental
132 changes (Enquist et al. 2015).

133 Recently, there has been an increasing investment into mapping plant functional trait
134 distributions given economic and data availability constraints such efforts have mostly focused on
135 hyperspectral imagery at local (Schneider et al., 2017) to regional scales (Asner et al., 2015, Asner et al.,
136 2016). However, high resolution hyperspectral imagery is not widely available (Clark, 2017, Szabó et al.,
137 2019). Landsat-8 imagery at coarser spatial (30 m pixel), spectral and temporal resolution than Sentinel-2
138 imagery has been used to map four traits over small (20 × 20 m) vegetation plots covering small spatial
139 extents (Wallis et al. 2019). The spatial mismatch between site-level trait sampling and the spatial
140 resolution of pixels may partly affect overall model predictions (Wallis et al. 2019). Other studies restricted
141 to European forests (Ma et al., 2019) show how Sentinel-2 imagery could be used to map functional trait

142 diversity in the comparatively low tree diversity forests of Europe (Ma et al., 2019) and to retrieve specific
143 leaf area from Landsat-8 imagery (Ali et al., 2017). However, the tropics present a different set of
144 challenges, such as the high species richness, low accessibility and comparatively low availability of trait
145 data, plus the low coverage of remote sensing data because of persistent high cloud cover. These
146 challenges have hampered developments in mapping plant functional trait distributions across most
147 tropical areas.

148 Satellite imagery with high spectral, spatial and temporal resolution is particularly needed in the
149 wet tropics (Asner et al., 2017), where clear days can be infrequent and several images may be required
150 to construct a cloud-free composite. The Copernicus mission from the European Space Agency's (ESA;
151 www.esa.int) aids in the improvement in this area. The Sentinel-2 multispectral imager satellites are part
152 of the Copernicus programme, which has the potential to provide new opportunities to evaluate canopy
153 traits remotely. Sentinel-2 has 13 spectral channels covering the visible, near-infrared, and short-wave
154 infrared, a spatial resolution of 10 m for visible and near-infrared, 20 m for short-wave infrared, revisit
155 period of 5 days and it provides open data availability. The improved spectral sampling (13 bands, 10
156 excluding the 60 m atmospheric bands) and fine spatial resolution of the Sentinel-2 images have the
157 potential to elucidate leaf chemistry, morphology, photosynthesis and water content at the pixel-level,
158 although this remains largely untested. Multispectral sensors do not provide the rich information available
159 from hyperspectral sensors, which have been used in numerous studies to map functional traits at small
160 spatial extents (Townsend et al., 2003, Laurin et al., 2016, Asner et al., 2015, Martin, M. E. et al., 2008).
161 However, high resolution open-access hyperspectral imagery is not currently available from space.
162 Although Landsat images have been used to predict a few functional traits at a local scale (Wallis et al.,
163 2019), the extended spectral, spatial and temporal capabilities of the state-of-the-art sensors onboard the
164 Sentinel-2 satellites provide greater potential for mapping functional trait diversity in tropical forest
165 ecosystems at large extents.

166 Here, we employ a unique and large dataset of *in-situ* plant canopy traits and vegetation census
167 data collected with a standardised protocol at multiple sites across the tropics to calibrate and validate

168 Sentinel-2 imagery for predicting community leaf trait composition. The data provide 14 standardized
169 measurements of *in-situ* collected plant functional traits, precisely geo-located and delineated individual
170 tree crowns and vegetation censuses from Australia, South East Asia, Africa and South America to model
171 and predict functional trait composition at the pixel-level. We investigate how functional traits of tropical
172 forests vary within and between these different tropical regions and whether Sentinel-2 spectral data in
173 conjunction with climatic and soil information provide sufficient information to predict such pixel-level
174 trait composition in long-term vegetation plots across the tropics. We hypothesised that there would be
175 differences in trait variation among sites and regions given the range of climatic and soil conditions across
176 the tropics. Given the high spectral and spatial resolution of Sentinel-2 imagery we further hypothesised
177 that raw spectral bands and textural information will prove to be key predictors of functional trait
178 distributions across the tropics. The very high spatial resolution and local origin of the input plant traits
179 and census dataset, which represent traits adapted to local environments, plus the use of the Sentinel-2
180 data will allow us to accurately predict plant functional trait distributions that are potentially generalisable
181 across the tropical forest biome.

182 **2. Methods**

183 *2.1 Vegetation plots*

184 We collected vegetation census data from 47 permanent vegetation plots that are part of the Global
185 Ecosystems Monitoring network (GEM; www.gem.tropicalforests.ox.ac.uk). These plots encompass wet
186 tropical forests, seasonally dry tropical forests, and tropical forest-savanna transitional vegetation. The
187 sampled vegetation plots have an area ranging from 0.1 to 1 ha, with most (61%) being 1 ha. The plots
188 used are located across four tropical continents and specifically in the countries of Australia, Brazil, Gabon,
189 Ghana, Malaysian Borneo (from here onwards referred to as Malaysia) and Peru (Table 1). In each plot all
190 woody plant individuals with a diameter ≥ 10 cm at breast height (DBH) or above buttress roots were
191 measured and their exact geographic location was recorded (see the 'Individual tree crowns' section

192 below for more details). In two plots (NXV-01 and NXV-10) in Nova Xavantina, here onwards referred to
193 as Brazil-NX, the DBH was measured near ground level as is standard in savanna monitoring protocols.

194 2.2 *Functional traits*

195 We collected plant functional trait measurements from all woody plants located in each of the 47
196 vegetation plots mentioned above (Table 2). All traits were gathered from the GEM network and were
197 collected following a standardized methodology across plots. Forest inventory data were used to stratify
198 tree species by basal area dominance, a proxy for canopy area dominance. The tree species that
199 contributed most to basal area abundance were sampled with 3-5 replicate individuals per species, with
200 a goal of sampling 60-80% of basal area across the sampling region. Eighty percent of basal area was often
201 achieved in low diversity sites (e.g. montane or dry forests) but only around 60% was achieved in some
202 high diversity sites (lowland humid rainforests). For each selected tree a sun and a shade branch were
203 sampled and in each branch 3-5 leaves were used for trait measurements. We only included the sun
204 exposed branches in our analysis because we were interested in the branches that could potentially be
205 receiving direct sun radiation and thus show direct spectral reflectance. This represented a total sample
206 of 2434 individual trees across the tropics (Table 1). The plant functional traits collected were those
207 related to photosynthetic capacity at both saturating CO₂ concentration (2000 ppm CO₂; A_{max}) and
208 ambient CO₂ concentration (400 ppm CO₂; A_{sat}); leaf chemistry (nitrogen, phosphorus, carbon, calcium,
209 potassium and magnesium content); and leaf morphological and structural traits (area, specific leaf area,
210 thickness, dry mass, fresh mass and water content). An overview of the methods for individual leaf
211 functional trait measurements is provided in the Supplementary Information (see full traits collection
212 protocol section). Further details of measurements for the Peruvian Andes campaign are given in Martin
213 et al. (2020) and Enquist et al. (2017), for the Malaysian campaign in Both et al. (2019), and for the Ghana
214 and Brazil campaigns in Oliveras et al. (2020) and Gvozdevaite et al. (2018).

215 Some individuals in the plots lacked functional trait values. To assign representative trait values
216 to unsampled individuals we did the following: 1) individuals from which traits were measured kept their

217 original trait information, 2) for individuals with no trait information we randomly sampled trait values
218 from other individuals from the same species present in the same plot, 3) if the species was not sampled
219 in the given plot then we randomly sampled an individual from the same species that had trait information
220 in other plots from the same region (Table 1). This protocol for trait value allocation allowed us to work
221 with the existing range of trait values at the species level and avoided to create average values per species
222 (Cadotte et al., 2011, Schneider et al., 2017). We did not assign trait values to the remaining individuals
223 belonging to species from which no trait collection was obtained at the regional level.

224 *2.3 Individual tree crowns*

225 Tree crown locations and structural attributes were recorded for each tree, where crown area and shape
226 were measured by direct crown field measurements in the case of plots in Malaysia and Peru (see protocol
227 below), or by means of regional level allometric equations developed by Shenkin et al. (2019) (all other
228 plots). In the latter case, the crown's shape was assumed to be circular. The direct field crown
229 measurements were as follows: all trees ≥ 10 cm DBH (i.e., 1.3 m from the ground) were mapped using a
230 ground-based Field-Map laser technology (IFER, Ltd., Jílové u Prahy, Czech Republic) (Hédli et al., 2009).
231 The Field-Map technology was based on a combination of Impulse 200 Standard laser rangefinder (with
232 in-built tilt sensor for measuring vertical angles), MapStar module II electronic compass (both Laser
233 Technology Inc., Colorado, USA), and the specialized mapping software Field-Map v. 11 (IFER, Czech
234 Republic). The technology was used to record spatial positions of tree stems in three-dimensional space
235 (x, y, z-coordinates) as well as to map individual horizontal projections of tree crowns in the plots. The
236 horizontal crown projection of every tree was obtained by measuring spatial positions (x and y-
237 coordinates) of series of points (ranging from 5 to 30 points depending on the size of the crown) at the
238 boundary of a crown projected to the horizontal plane. The shape of crown projection was subsequently
239 smoothed using the "smooth contour line" feature of Field-Map software v. 11. Heights of all trees with
240 DBH ≥ 10 cm were measured by the Impulse and TruPulse 360 R laser rangefinders (both Laser Technology
241 Inc., Colorado, USA). Thus, each individual crown was accurately geolocated rendering information about
242 its shape and vertical and horizontal position.

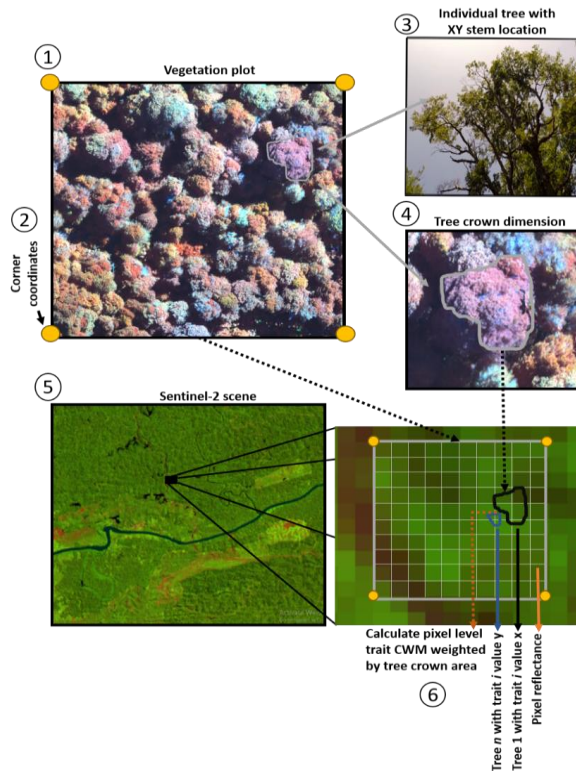
243 2.4 Calculating pixel-level trait composition

244 We calculated the community weighted mean of each trait for each 10 × 10 m subplot (matching the
245 highest pixel spatial resolution of the Sentinel-2 imagery) based on the mass ratio hypothesis, which
246 states that the most dominant species drive the ecosystem processes by means of their functional traits
247 (Grime, 1998). We first geolocated the vegetation plot, with its already mapped tree crowns (see
248 protocol above), to the Sentinel-2 imagery based on the corner coordinates of the plots. This is an
249 important step as geolocation errors between the vegetation plot and the correct location in the
250 satellite image could represent a large proportion of a given plot depending on the plots' area. Then for
251 each of the traits, t , and pixels, p , we calculated their community level weighted mean (CWM) using the
252 individual tree crown horizontal area as the weighting factor (Fig. 1) as follows:

253
$$CWM_{tp} = \frac{\sum_{i=1}^N CA_{ip} \times t_{ip}}{CA_p}$$

254 (Eq. 1)

255 Where CA_{ip} is the crown area of individual i in pixel p , t_{ip} is the trait value of individual i in pixel p , N is the
256 total number of individuals per pixel and CA_p is the crown area of pixel p . The crown contribution to the
257 CWM was weighted by its proportional cover of the corresponding pixel. The total number of pixels used
258 in our calculations are 403 for Australia, 449 for Brazil -NX (Nova Xavantina), 302 for Brazil -ST (Santarem),
259 464 for Gabon, 620 for Ghana, 976 for Malaysia and 1280 for Peru.



260

261 **Figure 1. Diagram summarising the steps followed to assign trait values per Sentinel-2 pixel.** 1) First the
 262 vegetation plots are defined based on the GEM (Global Environmental Monitoring) dataset and 2) from
 263 each vegetation plot the corner coordinates are extracted. 3) From each vegetation plot the XY position
 264 of each stem ≥ 10 cm DBH is extracted and 4) the crown horizontal area is calculated based on the protocol
 265 described in the methods section. 5) Then the Sentinel-2 imagery for the study area is processed to level
 266 2A using the ESA SNAP toolbox and 6) the vegetation plot is overlaid in the Sentinel-2 image based on its
 267 corner coordinates. In this last step (6) each pixel defines a ‘subplot’ which is the unit used to calculate
 268 the trait community weighted mean based on the crown area of the trees that are contained by that pixel.
 269 In 6) n refers to a given tree in a given pixel, trait i represents a given trait and x and y are values for that
 270 trait. The image used as an example in step (1) was taken by Jesus Aguirre-Gutierrez over a vegetation
 271 plot using a multispectral ALTUM camera mounted on an Inspire 1 drone.”

272

273 *2.5 Sentinel-2 data, vegetation indices and canopy texture parameters*

274 We gathered Sentinel-2 imagery that was closest in time and season to the sampling dates of functional
 275 traits and vegetation census across the tropics for each of the study locations (Table S1). The Sentinel-2
 276 imagery was first selected using the European Space Agency (ESA) ScienceHub (scihub.copernicus.eu)
 277 choosing images with high pixel quality and low cloud cover (<10%). Atmospheric, radiometric and
 278 topographic corrections were applied to the selected imagery (Level 1C) using the Sen2Cor algorithm in
 279 the Sentinel SNAP toolbox (step.esa.int). Our overlapping imagery with the vegetation plots appeared free
 280 of clouds and cirrus effects. The above-mentioned steps allowed us to obtain level 2A imagery with surface

281 reflectance values. We then resampled the 20 m bands to 10 m spatial resolution using bilinear
282 interpolation. The Sentinel-2 60 m resolution bands (B01, B09, B10) were not used as these are designed
283 for cirrus, water vapour and cloud detection (Table 3). Band 8A was not used as it covers an overlapping
284 spectral window with band 8 and has a lower spatial resolution. Since vegetation indices may increase
285 prediction accuracy when modelling community weighted traits (Wallis et al., 2019), we calculated three
286 of them (Table 3) which we hypothesised to inform trait distributions given their association with
287 chlorophyll and nutrient levels in the leaves and their use of the visible-to-red edge spectral bands.

288 Canopy structure may play an important role in separating different vegetation types and
289 differences in canopy spectral composition. To characterize canopy structure, we calculated the Grey
290 Level Co-Occurrence Matrix (GLCM) based texture features (Haralick et al., 1973). The desired texture
291 metrics are computed from a grey tone matrix that is spatially dependent. The co-occurrence matrix
292 depends on the angular relationship and distance between two neighbouring pixels and depicts the
293 number of occurrences of the relationship between a pixel and its neighbour. After trials with smaller
294 windows size (5 x 5) we opted to use a 9 x 9 pixel kernel window which was sufficient to render enough
295 canopy contrast information during the modelling step (see section 2.7 below) without taking large
296 periods of time for its calculation. The texture results obtained with the used kernel window was highly
297 correlated to the smaller kernel window ($Cor = 0.94$, $P = < 0.0001$). Based on the GLCM we calculated two
298 variables that are least correlated with each other, the Entropy and Correlation, for each of the vegetation
299 indices. While Entropy measures the homogeneity level for a given area, the Correlation measures
300 probability of occurrence of the specified pixel pairs across the image (Haralick et al., 1973, Wallis et al.,
301 2019). All remote sensing analyses related to the generation of vegetation indices and texture metrics
302 were carried out using the Sentinel SNAP toolbox (step.esa.int) and the R statistical environment (R
303 Development Core Team, 2014) with the 'Sen2R' package.

304 2.6 Environmental and soil data

305 Climatic, topographic and soil characteristics may vary across regions and could at least partly determine
306 the region's vegetation and intrinsic trait composition. We obtained information on these three
307 components for each sampling location. The three components were grouped as belonging to
308 environmental (climate) or soil-terrain (texture, pH, cation exchange capacity and topography) drivers
309 (Table 3).

310 For climate and for each sampling location we gathered gridded data on the mean annual climatic
311 water deficit (MCWD), which is a metric of drought intensity and severity, mean annual maximum
312 temperature (MATmax), solar radiation (SRAD) and soil moisture (SM) (Table 3). All climatic data with a
313 spatial resolution of ~4km were obtained from the TerraClimate gridded climate product (Abatzoglou et
314 al., 2018). To characterise the climatic conditions for each location we used a climatology of 30 years
315 (1986-2015) as suggested by the World Meteorological Organization (WMO;
316 www.wmo.int/pages/prog/wcp/ccl/faqs.php). We used the terrain slope to characterise the plot's
317 topography, as it has been shown that topography may shape the composition and structure of tropical
318 forests (Jucker et al., 2018) and may affect the vegetation spectral reflectance by modifying soil water and
319 nutrient availability. Terrain slope was calculated using a high-resolution digital elevation model, ~30m
320 pixel size at the equator, from the Shuttle Topography Mission (Farr et al., 2007). At most sites soil data
321 were sampled locally, and analysed to a standardised protocol in labs in either INPA, Manaus, Brazil or the
322 University of Leeds, UK, following the RAINFOR soil protocol (Quesada et al., 2012). From these data we
323 summarised plot level soil data averaged over the first 30 cm for texture (Sand% and Clay%), cation
324 exchange capacity (eCEC) and pH-H₂O (pH). Plot level texture data were not available for plots in Australia
325 and the NXV-10 plots and were thus derived from the SoilGrids dataset at 250m pixel spatial resolution
326 for those plots only (Hengl et al., 2017).

327 *2.7 Comparing community level trait distributions across regions*

328 We tested if and to what extent the community-level trait distributions differed among regions. We
329 square-root transformed the trait value to improve normality and applied an analysis of variance

330 (ANOVA). We then applied a Tukey's Honest Significant Difference (Tukey HSD) test to investigate the
331 significance of the differences between the means of the community weighted mean (CWM) trait values
332 among locations. The ANOVA and Tukey test were carried out using the 'stats' package for R (R
333 Development Core Team, 2014).

334 *2.7 Relating pixel-level trait composition to spectral reflectance, environment and soil conditions*

335 We modelled the community weighted mean (CWM) of each trait at the pixel-level (10×10 m) as a function
336 of the Sentinel-2 remote sensing, environmental and soil covariates (Table 3) using a 'spatial' version of
337 the machine learning Random Forest (RF) algorithm (Breiman, 2001) named Geographic Random Forests
338 (GRF) (Georganos et al., 2019). RF is a nonparametric algorithm that has been shown to be robust to
339 overfitting and variable inputs thanks to the bagging process and its random feature selection (Hastie et
340 al., 2009). Moreover, it has been extensively used to model and predict ecological and remote sensing
341 data within and across ecosystems (e.g. Asner et al., 2016, Van der Plas et al., 2018). In contrast to RF,
342 GRF disaggregates the underlying data in geographic space, in this case based on the spatial coordinates
343 of the Sentinel-2 pixels, building global and local sub-models (plot level), making the modelling framework
344 thus spatially explicit. The explicit inclusion of the spatial component (XY pixel location) in the models,
345 which are sequentially fitted with different sets of the training data (the bagging process) may contribute
346 to the observed reduced spatial autocorrelation of GRF in comparison to the common RF (Georganos et
347 al. 2019). In the GRF a global model is built as in other RF applications. However, GRF also generates a
348 local RF for each location, which includes a specified number of nearby observations, here defined by all
349 pixels in the vegetation plot (mostly 1 ha; Table 1), called 'neighbourhood', obtaining in this way metrics
350 of local and global model predictive power and variable importance. For model predictions, a fusion
351 between the global model (that uses more data) and local models (with low bias) can be applied,
352 weighting the contributions of the global and local models based on the parameters that increase the
353 predictive accuracy and decrease the model's Root Mean Square Error (RMSE). We used the spatial GRF
354 to fit a global model for each functional trait and also fit a specific model for each region (Australia, Brazil
355 -ST, Brazil -NX, Gabon, Ghana, Malaysia and Peru) using the SpatialML package in R.

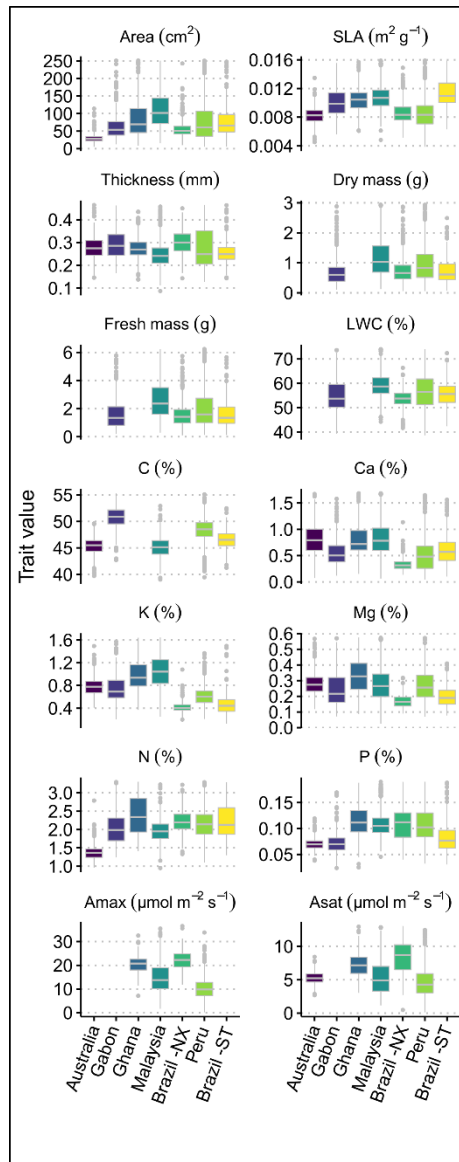
356 We performed an extensive set of model optimization and regularization procedures to reduce
357 over-fitting. For the CWM models we selected the number of trees to fit by 10-fold cross-validation
358 analysis with number of trees ranging between 500 and 1500 and the number of variables randomly
359 sampled as candidates at each split (mtry) ranging between 1 to 10, using in the final model the
360 combination of terms that generated the lowest RMSE. All covariates included in the models had pairwise
361 Pearson correlation coefficients $r \leq 0.82$ (Table 3). For the final global and local models, we used 80% of
362 the data for model fitting and the remaining 20% for model evaluation. Variable importance for each
363 model was computed as the decrease in node impurities from splitting on the variable, averaged over all
364 trees and derived from the Out of Bag (OOB) error. Then the resulting importance was standardised to a
365 0–1 scale for comparison purposes.

366 We carried out all analyses stated above with the full set of tree individuals present in each
367 vegetation plot with functional traits assuming that the contribution of small individuals to the trait CWM
368 value, and thus to the community reflectance at the pixel-level, would be minimal given the weighting
369 factor used (i.e. the individual's crown area). However, to underpin this we carried out all analyses on two
370 smaller datasets, one where the 25th and other where the 50th percentile of the smallest trees per region
371 were removed. All analyses were carried out in the R statistical environment with the 'caret', 'tidyverse'
372 and 'SpatialML' packages.

373 **3. Results**

374 *3.1 Variation in trait composition across tropical forests*

375 Most leaf functional traits exhibited significant differences across the tropics (Fig. 2) including wide trait
376 range variation within the same region (Fig. S8), with leaf fresh mass and leaf thickness being on average
377 less variable among locations (Table S2).



378

379 **Figure 2. Comparison of trait distributions across tropical regions.** The boxplots are based on the pixel-
 380 level (10×10 m) community trait values for each trait and region ($n=403$ for Australia, 449 for Brazil-NX,
 381 302 for Brazil-ST, 464 for Gabon, 620 for Ghana, 976 for Malaysia and 1280 for Peru). Horizontal lines in
 382 each boxplot show the median value and vertical lines are the whiskers that extend to the largest value
 383 or not further than 1.5 times the inter-quartile range. For some locations information for all traits was not
 384 available. For full details in significant differences in mean trait values among locations see Table S4. Brazil
 385 -NX: Nova Xavantina; Brazil -ST: Santarem.

386 Leaf chemistry and photosynthetic capacity (A_{\max} and A_{sat}) often showed significant differences
 387 among locations (Table S2). Drier locations as in Nova Xavantina (Brazil -NX) displayed trait adaptations
 388 to seasonal rainfall and temperature with on average thicker and smaller (30 ± 0.05 mm and 56.2 ± 24.7
 389 cm^2 respectively) leaves at the community level, with some of the highest community-level leaf nitrogen
 390 concentration (2.2 ± 0.3 %) and highest photosynthetic capacity (mean $A_{\max} = 21.9 \pm 4.3 \mu\text{mol m}^{-2} \text{s}^{-1}$, and
 391 $A_{\text{sat}} = 8.3 \pm 2.5 \mu\text{mol m}^{-2} \text{s}^{-1}$). In contrast, wetter regions such as Malaysia displayed on average some of

392 the biggest ($113.5 \pm 55 \text{ cm}^2$) and thinnest ($0.25 \pm 0.05 \text{ mm}$) leaves with high leaf water content (59.1 ± 5
393 %). The Peruvian altitudinal transect showed large variation in community-level traits values, which often
394 overlapped with trait values from all other sampled locations across the tropics (Fig. 2). For most nutrients,
395 leaf nutrient concentration was often highest in forests found in Ghana (e.g. $\text{K}\% = 0.97 \pm 0.27$ and $\text{Mg}\% =$
396 0.33 ± 0.1) and Malaysia ($\text{K}\% = 1.05 \pm 0.27$ and $\text{Mg}\% = 0.27 \pm 0.1$). Australian forests showed on average
397 some of the lowest community-level N ($1.3 \pm 0.21 \%$) and P ($0.07 \pm 0.01\%$) leaf concentrations.

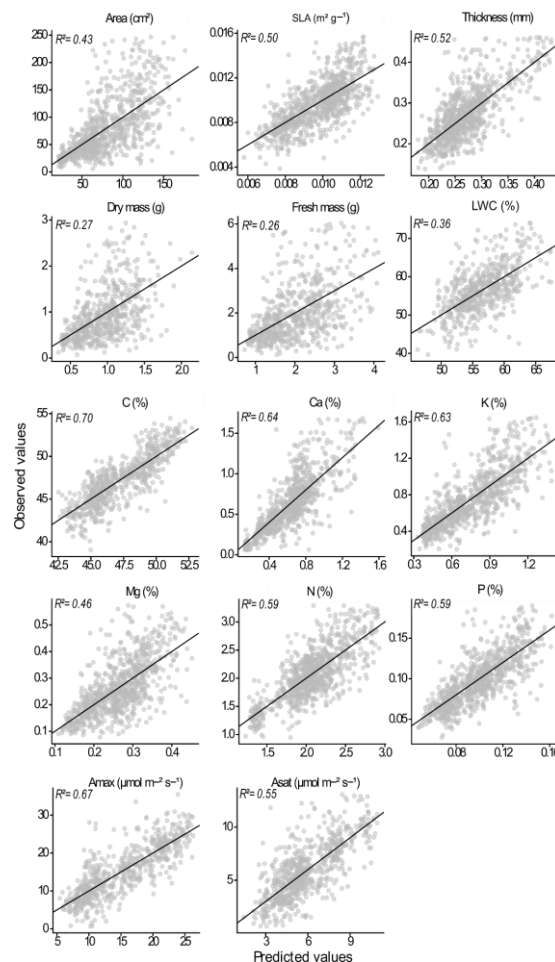
398 *3.2 Pantropical and local community level functional trait models*

399 The analyses carried out with the full dataset and the dataset where the 25th and 50th percentile of the
400 smallest trees per region were removed gave similar results for the global ($R^2 = 0.95$ and $R^2 = 0.97$
401 respectively; Table S3) and local ($R^2 = 0.81$ and $R^2 = 0.80$ respectively; Table S4) models of plant trait
402 distributions. Therefore, in the following we only present the results for the models carried out with the
403 full vegetation dataset.

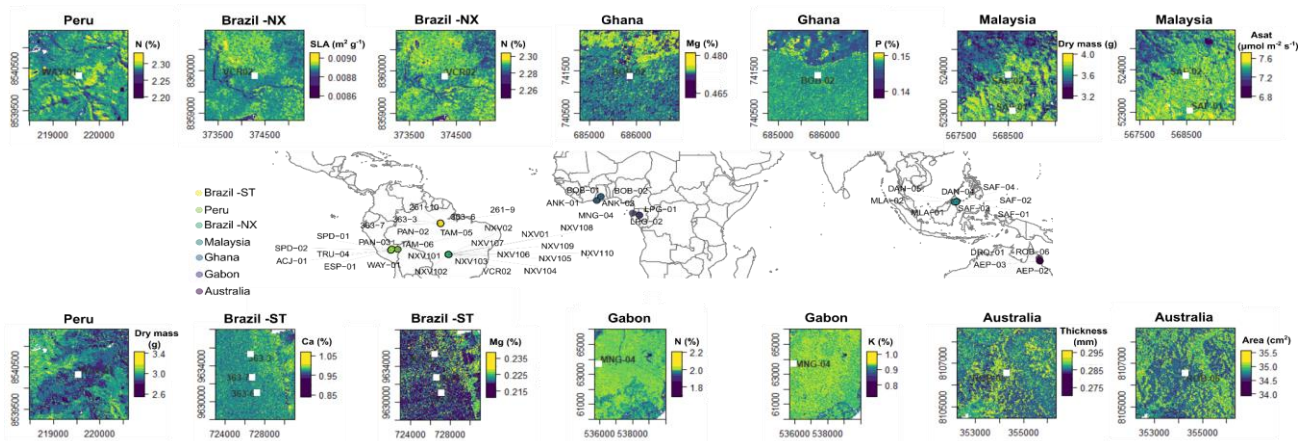
404 The accuracy of the pantropical prediction of functional traits ranged between a minimum of $R^2 =$
405 0.26 , for leaf fresh mass, and a maximum of $R^2 = 0.70$ for leaf carbon content (C%) based on the out-of-
406 sampled (testing) data across the tropics (Table 4). The predictive accuracies of leaf chemistry and
407 photosynthetic traits were often higher than for morphological and structural traits such as leaf dry mass
408 ($R^2 = 0.27$) and leaf area ($R^2 = 0.43$) (Fig. 3). At the pantropical level, the highest prediction accuracy was
409 obtained for leaf thickness ($R^2 = 0.52$) for morphological and structural traits, for leaf Ca (Ca%; $R^2 = 0.64$)
410 and leaf K (K%; $R^2 = 0.63$) for the chemical traits other than carbon. Leaf N and P concentrations were also
411 predicted with high accuracy ($R^2 = 0.59$). Leaf photosynthetic capacity traits, A_{max} and A_{sat} , showed some
412 of the highest prediction accuracies ranging from $R^2 = 0.55$ to 0.67 , respectively. Model spatial predictions
413 for several traits and locations are shown in Fig. 4 and others can be seen in Fig. S1-Fig. S7.

414 Models built for each tropical region and trait uncovered marked differences in prediction
415 accuracy among them (Fig. 5; Table 5 and Table S5). Leaf area prediction accuracy ranged from $R^2 = 0.04$
416 (Brazil -ST) to 0.35 (Australia), and that of specific leaf area (SLA) ranged from $R^2 = 0.06$ for Malaysia to

417 0.54 for Brazil -NX (Table S5). The local models showed a higher accuracy for predicting local level leaf
 418 chemical nutrients (up to $R^2 = 0.68$), especially for P, Ca, and N concentrations in comparison to
 419 morphological (e.g. leaf area and SLA) traits (Table 5; Fig. 5). Traits related to photosynthetic capacity
 420 showed an overall better prediction accuracy than leaf area and SLA with prediction values ranging
 421 between 0.36 (Peru) to 0.49 (Ghana) for A_{max} and up to 0.52 for A_{sat} (Brazil -NX; Fig. 5). On average the
 422 highest prediction accuracy across regions for a given trait were reached for leaf P concentration ($R^2 =$
 423 0.47) and A_{max} ($R^2 = 0.44$) and the locations with the highest average prediction accuracy across traits were
 424 the Nova Xavantina savanna (Brazil -NX, $R^2=0.40$) and the Peru elevation gradient ($R^2 = 0.38$; Table 5), both
 425 sites encompassing strong gradients in vegetation morphology and structure.



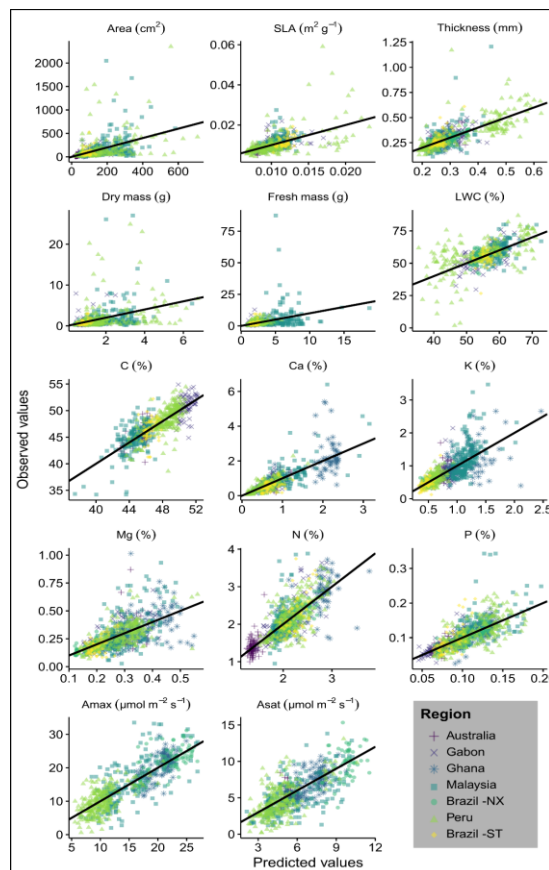
426
 427 **Figure 3. Model predictions to the 20% test data from the general model which was fitted with 80% of**
 428 **the trait data from across the tropics.** Grey dots are the observed against predicted trait values of the
 429 pixel-level (10 × 10 m) community weighted mean traits from the test dataset. The black line shows the
 430 1:1 relationship between observed and predicted values. Model prediction accuracy is shown in the top
 431 left. Full model results are shown in Table 4.



432

433 **Figure 4. Spatial predictions of trait distributions for a selected subset of plant traits and locations.**

434 The map (middle) shows the locations of vegetation plots that were used during the modelling
 435 framework. The spatial predictions (top and bottom rows) were obtained using the general models
 436 (Table 4) for each of the traits and locations at a 10×10 m pixel resolution. The approximated location
 437 of each vegetation plot used is shown as a white square within each spatial prediction map (for
 438 visualisation purposes white squares are not scaled to the plot's real size). Spatial predictions for other
 439 traits can be found in Fig. S1-S7.



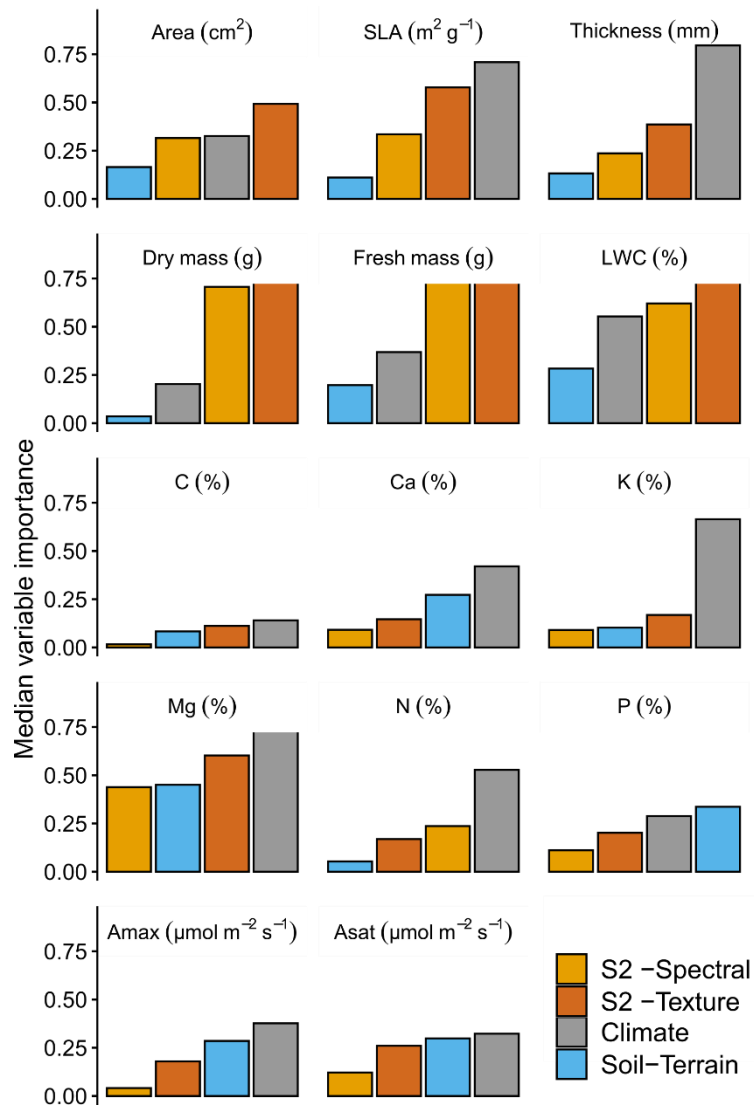
440

441 **Figure 5. Models predictions to the 20% test data from the regional models fitted with 80% of the trait**
 442 **data from each region across the tropics.** Each colour represents an individual regional model and the
 443 coloured symbols are the observed against predicted trait values of the pixel-level (10×10 m) community
 444 weighted mean traits from the test dataset. The black line shows the 1:1 relationship between observed
 445 and predicted values. Model prediction accuracy is shown in Table 5. Full model results are shown in Table
 446 S5. Brazil -NX: Nova Xavantina; Brazil -ST: Santarem.

447 3.3 Importance of spectral remote sensing, climatic and soil data for mapping trait distributions

448 We included Sentinel-2 band derived reflectance values, vegetation indices, their canopy texture
449 parameters, climatic and soil variables in the general trait models to predict community level traits at the
450 pixel-level (Table 3). The importance of these variables for predicting traits depended on the specific trait
451 being addressed (Fig. 6). In the global model, the remote sensing texture parameters were the first or
452 second major contributor for predicting nine of the functional traits across the tropics (Fig. 6 and Fig. S9).
453 Raw spectral variables were the second most important group for predicting four of such functional traits
454 but often lower in importance than the textural parameters. In the global model, soil and terrain factors
455 were on average some of the most important for predicting photosynthetic traits and foliar P
456 concentration. On average, climatic variables were important for predicting 11 out of 14 functional traits
457 but their contribution was lower for predicting leaf dry and fresh mass and leaf water content (Fig. 6).
458 However, it is evident that a combination of textural, spectral, climatic and soil information is required to
459 obtain the best general model predictions across functional traits and no single variable appears as the
460 most important across all traits (Fig. S9).

461 The local models provided a site-specific view of the most important remote sensing derived
462 variables, environmental and soil conditions for deriving community level traits composition (Fig. S10).
463 Sentinel-2 remote sensing related variables were more important for detecting leaf morphology and
464 nutrient values than environmental and soil related variables in 88% of the trait models (in 75 out of 85
465 possible traits by region combinations). Eighty-one percent of the time (69 location by trait combinations)
466 the canopy texture parameters were more important than the raw spectral reflectance factors. In 5.9%
467 and 4.7% of the possible trait and region combinations, climatic or soil-topography related variables
468 respectively were the most important for detecting community traits (Fig. S10).



469
 470 **Figure 6. Group median variable importance of spectral remote sensing, environmental and soil**
 471 **related variables for determining functional trait predictions in the global model.** Variable importance
 472 (Y axis) ranges from 0 (no importance) to 1 (highest importance) and represents the decrease in node
 473 impurities from splitting on the variable, averaged over all trees and derived from the Out of Bag error,
 474 the resulting value has been standardised to a 0-1 scale for comparison purposes. The spectral group (S2
 475 -Spectral) contains the select raw bands from the Sentinel-2 and the vegetation indices; Texture
 476 parameters (S2 -Texture) contain the Correlation and Entropy metrics from the grey level co-occurrence
 477 matrix obtained from the vegetation indices; Climate contains all climatic variables; Soil-Terrain contains
 478 all soil characteristics and slope. All variables are described in Table

479 **4. Discussion**

480 To the best of our knowledge, this is the first study evaluating the ability of Sentinel-2 satellite data to
 481 map plant functional traits across tropical ecosystems. Tropical forest trait mapping is fundamental for
 482 understanding of plant responses to global change, and notably the plant functional traits we predict in
 483 this study are relevant to plant species responses to a changing environment (Both et al., 2019, Nunes et

484 al., 2019, Soudzilovskaia et al., 2013, Aguirre-Gutiérrez et al., 2019). We have demonstrated that accurate
485 pixel-level (10×10 m) predictions of tropical forest functional trait distributions across the tropics can be
486 generated by making use of extensive *in-situ* collected plant functional traits, geo-located canopy
487 structure, vegetation censuses and high spectral and spatial resolution remote sensing data from the
488 Sentinel-2 satellites.

489 4.1 Tropical forest trait distributions

490 Plant functional traits are characteristics that aid species to thrive in their environment or adapt to new
491 conditions. Given such adaptations to specific environments it might be expected that trait variation
492 would be higher in regions that encompass more varied environmental conditions (Enquist et al., 2015).
493 Environmental adaptation is exemplified by the strong variation in values for most traits in Peru and
494 Malaysia. In Peru, the data represent a climatic and altitudinal gradient ranging from the lowland Amazon
495 in the Tambopata National Park at an elevation of 200-225 masl to plots in Acjanaco at above 3000 masl.
496 In Malaysia, the vegetation plots are distributed across a land-use gradient ranging from undisturbed to
497 heavily logged forests (Both et al., 2019). Environmental adaptation may be also shown by the observed
498 differences in trait distributions between different regions across the tropics (e.g. Australia and Gabon;
499 see also Fig. 2). The pixel-based community trait values in the Peruvian transect often extend across much
500 of the range in trait values observed in other locations (Fig. 2). We detected an overall significant
501 difference among locations in terms of morphological, chemical and photosynthetic traits (Table S2). This
502 wide variation in traits suggests the presence of local biotic and abiotic controls of trait distributions and
503 plant species adaptations that may differ among tropical regions. Such differences in trait composition
504 highlight the importance and the challenge of sampling as fully as possible the functional trait diversity
505 across different tropical forests. This is of pivotal importance when comparing forest responses to
506 changing environments across multiple regions. We thus suggest that further field trait survey campaigns
507 across the tropics are needed to improve pantropical trait predictions. As for the local biotic and abiotic
508 controls of trait distributions, for instance, it is widely known that African tropical forests are in general
509 less species diverse than their Asian and South American counterparts but that they have some of the

510 highest biomass carbon storage capacity per unit area (Sullivan et al., 2017). Tropical forests in West Africa
511 are in general drier in comparison to Amazonian tropical forests (Parmentier et al., 2007) and some African
512 regions such as Gabon have experienced increases in temperature and decreases in precipitation over the
513 last 30 years (Bush et al., 2020). Thus, such changes in climatic conditions as those observed in West
514 African tropical forest may also underlie variations in species composition and the locally observed
515 functional trait pool as shown in this study. It is also worth noting that two caveats of the community-
516 weighted mean trait approach may account for part of the unexplained trait variation. First, it makes the
517 assumption of a unique functional optimum in a given environment, while multiple optimal strategies –
518 potentially corresponding to contrasting trait values – could coexist (Laughlin et al., 2018). Secondly, it
519 does not account for the dynamic nature of communities, so that a community weighted mean at a given
520 time point might not encompass the optimum at equilibrium (Laughlin et al., 2018).

521 Morphological and structural traits such as leaf area, fresh and dry mass, leaf thickness, SLA and
522 LWC, represent trade-offs between energy acquisition, consumption and survival and form a main part of
523 the global spectrum of plant functioning (Díaz et al., 2016). Besides investigating the predictability of such
524 plant structural traits, we further analysed the potential for predicting leaf chemistry (C, K, Mg, Ca, N, P)
525 and photosynthesis related traits (A_{\max} and A_{sat}). Mapping chemical and photosynthetic traits at a
526 pantropical scale has the potential for increasing our understanding of how photosynthetic capacity shifts
527 across tropical regions and on possible impacts of a changing environment on tropical forests productivity
528 (Guan et al., 2015, Mueller et al., 2014).

529 *4.2 Sentinel-2 remote sensing for mapping community level trait distributions across the tropics*

530 In their pioneering work with hyperspectral imagery and simulated multispectral Sentinel-2 data over
531 Ghana, Laurin et al. (2016) demonstrated that Sentinel-2 imagery could be used to discriminate tropical
532 forest types and map plant functional types. The authors argued that the full band set and vegetation
533 indices derived from the Sentinel-2 would be advantageous for accurately mapping plant functional guilds
534 in the tropics. By using functional trait data collected *in situ* across tropical forests and modelling at high

535 spatial resolution (pixel-level) we show that most of our global trait distribution models present a high
536 predictive power for most traits analysed, with prediction accuracy on the testing datasets being highest
537 for predicting leaf chemical and photosynthetic capacity traits. However, we also show that the local level
538 trait models produced less accurate predictions than the global models, probably as a result of the
539 narrower range of in plant trait values found at the local region in comparison to across the regions,
540 something also shown by Wallis et al. (2019). The prediction accuracy obtained from our models using
541 Sentinel-2 multispectral data is similar and in some cases higher than that shown by recent studies that
542 make use of hyperspectral imagery and other multispectral sensors to map functional traits (Martin et al.,
543 2018, Asner et al., 2017, Asner et al., 2015). For instance our predictions on test data for leaf nitrogen,
544 phosphorus and carbon are comparable or higher than those obtained by other innovative studies in
545 Malaysia ($R^2 = 0.46, 0.44$ and 0.48 respectively; Martin et al., 2018), Peru ($R^2 = 0.48, 0.39$ and 0.44 ; Asner
546 et al., 2015) and temperate forests ($R^2 = 0.55, 0.22, 0.46$; Nunes et al., 2017), and closely related to those
547 obtained by Wallis et al. (2019) with other multispectral sensor for nitrogen and phosphorus ($R^2 = 0.65$
548 and 0.65). Specially the work of Asner et al. (2017) has shown how such plant trait predictions (with its
549 inherent accuracies) can be used for other applications such as to guide biodiversity conservation actions.
550 In our approach we resample the 20m spatial resolution bands from the Sentinel-2 to 10m pixels as to
551 work with the highest spatial resolution available for most spectral bands. Such resampling could in
552 principle have an effect on the match between the tree canopies' reflectance signal and the spectral signal
553 from the Sentinel-2 pixel and could thus influence the textural parameters, by for instance, detecting
554 lower heterogeneity.

555 Some of the leaf chemistry we modelled can be directly related to the reflectance obtained from
556 the Sentinel-2 remote sensor in the visible, infrared and red-edge regions which capture the leaf
557 biogeochemistry (Ustin & Gamon, 2010). For instance, it has been shown how carbon and carbon
558 containing metabolites peak in reflectance at around 550 nm (band 3 in the Sentinel-2) and at the lower
559 part of the 702–715 nm (Ely et al., 2019), which would be depicted best by the red-edge band 5 in the
560 Sentinel-2. such spectral behaviour captured by the Sentinel-2 bands contributed to the high prediction

561 accuracy of leaf carbon in our study. Our models show how Sentinel-2 imagery, and especially the canopy
562 texture parameters derived from it, can be especially useful for mapping traits related to leaf chemistry
563 (Fig. 2 and Fig. S9). Moreover, our high predictive accuracy for photosynthetic capacity (A_{max} , A_{sat}) is
564 consistent with studies carried out in other vegetation types (e.g. agroecosystems; Serbin et al., 2015)
565 where a strong association was shown between photosynthesis related traits and the red-edge spectral
566 region. Sentinel-2 has 3 bands over the red-edge spectral region (bands 5, 6, 7) and two over the near
567 infrared (bands 8 and 8a) with different bandwidths, which as shown by Shiklomanov et al. (2016) can be
568 advantageous for detecting foliar nutrients such as leaf N (Schlemmer et al., 2013) , as small differences
569 in wavelength position in different bands may impact their capacity to retrieve canopy trait
570 characteristics. Moreover, the strong relationship between photosynthetic capacity and spectral
571 reflectance can be partly captured from the leaf N signal, as leaf N concentrations are strongly associated
572 with photosynthetic capacity (Reich, 2012, Vincent, 2001). The N reflectance signal is often best obtained
573 in wavebands centred between 440 and 570 nm (Ferwerda et al., 2005).

574 In this study we leverage evidence on covariation among traits to estimate and predict values of
575 traits that have no clear physical effects on spectral reflectance. There is ample evidence of the existence
576 of covariation among plant traits, as for instance between leaf N concentration, specific leaf area and leaf
577 longevity (Walker et al., 2017). Such covariation among traits may in principle also represent covariation
578 in the spectral reflectance patterns across vegetation types (Ma et al., 2019), especially if such individuals
579 vary in leaf structural tissue that drive energy scattering and reflectance (Ollinger, 2011). Such covariation
580 between traits can be helpful for mapping functional trait diversity across large spatial extents that include
581 diverse vegetation types (Townsend et al., 2003, Both et al., 2019). We show that the spectral reflectance,
582 image textural parameters (Entropy and Correlation), climate and soil, are highly relevant for modelling
583 plant trait distributions across the tropics with high prediction accuracy. However, the canopy texture
584 parameters (Entropy and Correlation) are some of the most important for attaining high trait prediction
585 accuracies across plant functional traits (Sarker & Nichol, 2011, Wallis et al., 2019) and differences in

586 spectral, climatic and soil conditions between different regions are key components for improving model
587 predictions across broad spatial extents.

588 Image texture parameters were derived from the vegetation indices that we calculated, which in
589 turn were derived from the raw spectral bands of the Sentinel-2. Thus, the texture metrics besides taking
590 advantage of the high spectral resolution of the sensor also take advantage of its high spatial resolution.
591 Although the raw spectral bands of the Sentinel-2 were not as important for predicting some functional
592 traits as image texture, it is relevant to consider that texture values tend to differ based on the spatial
593 resolution of the underlying data on which they are based. A larger pixel (e.g. Landsat's 30 × 30 m pixels)
594 may thus mask differences in the landscape that could in principle be captured by the Sentinel 10 × 10 m
595 resolution texture generated metrics. This therefore highlights the relevance of Sentinel-2 imagery for
596 functional plant functional trait predictions in comparison to others with lower spectral and spatial
597 resolution. Image texture parameters can help characterise the upper surface of the vegetation, which in
598 our study is composed of varied sets of functional trait characteristics that confer them different spectral
599 responses. When such spectral differences are analysed with grey level co-occurrence matrices, the
600 generated image texture parameters (e.g. entropy and correlation) can help differentiate contrasting
601 vegetation in the landscape. The role of texture parameters for modelling biomass and functional traits
602 has also been recognised by other studies focusing not only on mapping functional traits along elevation
603 gradients but also for estimating standing biomass (Wallis et al., 2019). Moreover, such relevance of
604 texture parameters does not seem to be limited to the spatial resolution of the Sentinel-2 imagery as
605 shown when using high spatial resolution SPOT imagery for modelling forest aboveground biomass
606 (Hlatshwayo et al., 2019) and WorldView-3 for tree species identification (Ferreira et al., 2019), or lower
607 spatial resolution data as that from the Landsat (Wallis et al., 2019). Other added value of the Sentinel-2
608 in contrast to finer spatial resolution satellites (e.g. SPOT and WorldView-3) is its high revisit period, to
609 obtain cloud free imagery, and it's free availability. Moreover, soil properties can be informative when
610 modelling trait distributions across regions in the tropics as they partly drive the plant functional and
611 species compositional turnover (Prada et al., 2017, Asner et al., 2016). In our study different vegetation

612 plots appeared to be on soils with different parent materials resulting in varying cation exchange capacity,
613 pH and soil texture, and thus including differences between sites contributes to increasing the prediction
614 accuracy of trait distributions.

615 Although in the past it was thought not to be possible to map individual plant species or functional
616 traits (Price, 1994, Ustin & Gamon, 2010), the advent of remotely sensed data with high spectral, spatial
617 and temporal resolution has made it possible to extract information on the chemical and structural
618 composition of forest canopies even in highly biodiverse tropical forests. This has been demonstrated with
619 the use of hyperspectral sensors (Asner et al., 2017, Asner et al., 2015, Jetz et al., 2016), which often
620 collect hundreds of spectral bands at very high spatial and spectral resolutions but at relatively small
621 spatial extents and often without temporal replication. More research is needed to disentangle to what
622 extent hyperspectral data offers more information to that offered by the Sentinel-2 sensors for an
623 increased mapping accuracy of functional traits of tropical forests. As shown by Laurin et al. (Laurin et al.,
624 2016), results obtained with simulated Sentinel-2 data are highly comparable to those obtained from
625 hyperspectral imagery for mapping forest types, dominant tree species and functional guilds. Being able
626 to monitor functional traits at high spatial and temporal resolution with multispectral data ranging from
627 the visible to the shortwave infrared across the tropics and with freely available data opens new
628 opportunities for understanding the effects of environmental changes on biodiversity at a local scale. This
629 is because functional traits play a major role in determining ecosystem productivity and functioning, e.g.
630 carbon capture (Díaz et al., 2019, Carmona et al., 2016). Moreover, spatially explicit models of functional
631 traits shift across the tropics can help decipher how ecosystem functioning varies even among tropical
632 areas, providing a cost-effective pathway to identifying regions of high conservation value and hence aid
633 in the creation of locally adequate biodiversity conservation policies. Overall, our findings are of relevance
634 for informing biodiversity monitoring policies under ecosystem change as we show that accurate
635 predictions of relevant plant functional traits can be obtained in high biodiversity areas such as the tropics.
636 Our approach thus facilitates tracking possible shifts in trait distributions and composition across large
637 spatial extents as a response to environmental changes using the Sentinel-2 satellites.

638 **5. Conclusions**

639 Tropical forest ecosystems are witnessing a rapid transformation as a result of changing environmental
640 conditions and direct human impacts (Lewis et al., 2015, Taubert et al., 2018, Aguirre-Gutiérrez et al.,
641 2019). However, we cannot adequately understand or simulate tropical ecosystem responses to
642 environmental changes based solely on current ecosystem model approaches as these are unable to
643 capture the high diversity of plant ecosystem functions in the species-rich tropics. Neglect of functional
644 biodiversity can oversimplify the simulated response of an ecosystem to an environmental disturbance.
645 Here we show the high variation in functional traits that exists among tropical regions, which hints at the
646 different capabilities of such forests to respond to a changing environment. We demonstrate
647 opportunities for measuring the distribution of key functional traits across tropical forest ecosystems at
648 the pixel-level using the Sentinel-2 satellites, which if done across time could reveal areas where
649 functional shifts have occurred and likely where biodiversity conservation/amelioration measures are
650 needed. Although the Sentinel-2 satellites show high promise for this endeavour, our approach is limited
651 by the short time interval since they were launched (i.e. 2015) and the lower spectral resolution of
652 Sentinel-2 imagery in comparison to that derived from hyperspectral sensors. Methods and data products
653 are needed to track changes in functional composition in forest ecosystems across time and space. We
654 demonstrate a new approach to develop a rapid monitoring tool for capturing the effects of a changing
655 environment across the tropics. This new tool has the potential to contribute to a more robust and
656 evidence-based policy-making for conservation of tropical forest ecosystems.

657 **Authorship contribution statement**

658 J.A.G. conceived the study, designed and carried out the analysis and wrote the first draft of the paper.
659 Y.M. conceived and implemented the GEM Network, obtained funding for most of the GEM traits field
660 campaigns and commented on earlier versions of the manuscript. S.R. advised on statistical and remote
661 sensing analysis and commented on earlier versions of the manuscript. All co-authors participated in or
662 coordinated vegetation, trait data and/or soil data collection or processed field data. The authors named

663 between S.A.B. and L.J.T.W. are listed alphabetically. All co-authors commented on and approved the
664 manuscript.

665 **Declaration of competing interest**

666 The authors declare that they have no known competing financial interests or personal relationships that
667 could have appeared to influence the work reported in this paper.

668 **Acknowledgements**

669 This work is a product of the Global Ecosystems Monitoring (GEM) network (gem.tropicalforests.ox.ac.uk).
670 J.A.G. was funded by the Natural Environment Research Council (NERC; NE/T011084/1 and
671 NE/S011811/1) and the Netherlands Organisation for Scientific Research (NWO) under the Rubicon
672 programme with project number 019.162LW.010. The traits field campaign was funded by a grant to Y.M.
673 from the European Research Council (Advanced Grant GEM-TRAIT: 321131) under the European Union's
674 Seventh Framework Programme (FP7/2007-2013), with additional support from NERC Grant
675 NE/D014174/1 and NE/J022616/1 for traits work in Peru, NERC Grant ECOFOR (NE/K016385/1) for traits
676 work in Santarem, NERC Grant BALI (NE/K016369/1) for plot and traits work in Malaysia and ERC Advanced
677 Grant T-FORCES (291585) to Phillips for traits work in Australia. Plot setup in Ghana and Gabon were
678 funded by a NERC Grant NE/I014705/1 and by the Royal Society-Leverhulme Africa Capacity Building
679 Programme. The Malaysia campaign was also funded by NERC Grant NE/K016253/1. Plot inventories in
680 Peru were supported by funding from the US National Science Foundation Long-Term Research in
681 Environmental Biology program (LTREB; DEB 1754647) and the Gordon and Betty Moore Foundation
682 Andes-Amazon Program. Plots inventories in Nova Xavantina (Brazil) were supported by the National
683 Council for Scientific and Technological Development (CNPq), Long Term Ecological Research Program
684 (PELD), Proc. 441244/2016-5, and the Foundation of Research Support of Mato Grosso (FAPEMAT),
685 Project ReFlor, Proc. 589267/2016. During data collection, I.O. was supported by a Marie Curie Fellowship
686 (FP7-PEOPLE-2012-IEF-327990). GEM trait data in Gabon was collected under authorisation to Y.M. and
687 supported by the Gabon National Parks Agency. D.B. was funded by the Fondation Wiener-Anspach.

688 W.D.K. acknowledges support from the Faculty Research Cluster 'Global Ecology' of the University of
689 Amsterdam. M.S. was funded by a grant from the Ministry of Education, Youth and Sports of the Czech
690 Republic (INTER-TRANSFER LTT19018). Y.M. is supported by the Jackson Foundation. We thank the two
691 anonymous reviewers and Associate Editor G. Henebry for their insightful comments that helped
692 improved this manuscript.

- 693 References
- 694 Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., & Hegewisch, K.C. (2018). TerraClimate, a high-resolution
695 global dataset of monthly climate and climatic water balance from 1958–2015. *Scientific data*, *5*, 170191
- 696 Aguirre-Gutiérrez, J., Oliveras, I., Rifai, S., Fauset, S., Adu-Bredu, S., Affum-Baffoe, K., et al. (2019). Drier
697 tropical forests are susceptible to functional changes in response to a long-term drought. *Ecology*
698 *Letters*, *22*, 855-865
- 699 Ali, A.M., Darvishzadeh, R., & Skidmore, A.K. (2017). Retrieval of specific leaf area from landsat-8 surface
700 reflectance data using statistical and physical models. *IEEE Journal of selected topics in applied earth*
701 *observations and remote sensing*, *10*, 3529-3536
- 702 Asner, G.P., Martin, R.E., Anderson, C.B., & Knapp, D.E. (2015). Quantifying forest canopy traits: Imaging
703 spectroscopy versus field survey. *Remote Sensing of Environment*, *158*, 15-27
- 704 Asner, G.P., Knapp, D.E., Anderson, C.B., Martin, R.E., & Vaughn, N. (2016). Large-scale climatic and
705 geophysical controls on the leaf economics spectrum. *Proceedings of the National Academy of Sciences*,
706 *113*, E4043-E4051
- 707 Asner, G.P., Martin, R.E., Knapp, D.E., Tupayachi, R., Anderson, C.B., Sinca, F., et al. (2017). Airborne
708 laser-guided imaging spectroscopy to map forest trait diversity and guide conservation. *Science*, *355*,
709 385-389
- 710 Barnes, E.M., Clarke, T.R., Richards, S.E., Colaizzi, P.D., Haberland, J., Kostrzewski, M., et al. (2000).
711 Coincident detection of crop water stress, nitrogen status and canopy density using ground based
712 multispectral data. , *1619*
- 713 Beer, C., Reichstein, M., Tomelleri, E., Ciais, P., Jung, M., Carvalhais, N., et al. (2010). Terrestrial gross
714 carbon dioxide uptake: global distribution and covariation with climate. *Science*, *329*, 834-838
- 715 Both, S., Riutta, T., Paine, C.T., Elias, D.M., Cruz, R.S., Jain, A., et al. (2019). Logging and soil nutrients
716 independently explain plant trait expression in tropical forests. *New Phytologist*, *221*, 1853-1865
- 717 Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5-32
- 718 Bush, E.R., Jeffery, K., Bunnefeld, N., Tutin, C., Musgrave, R., Moussavou, G., et al. (2020). Rare ground
719 data confirm significant warming and drying in western equatorial Africa. *PeerJ*, *8*, e8732
- 720 Cadotte, M.W., Carscadden, K., & Mirotchnick, N. (2011). Beyond species: functional diversity and the
721 maintenance of ecological processes and services. *Journal of Applied Ecology*, *48*, 1079-1087
- 722 Carmona, C.P., de Bello, F., Mason, N.W., & Lepš, J. (2016). Traits without borders: integrating functional
723 diversity across scales. *Trends in ecology & evolution*, *31*, 382-394

- 724 Chave, J., Muller-Landau, H.C., Baker, T.R., Easdale, T.A., Steege, H.t., & Webb, C.O. (2006). Regional and
725 phylogenetic variation of wood density across 2456 neotropical tree species. *Ecological Applications*, *16*,
726 2356-2367
- 727 Clark, M.L. (2017). Comparison of simulated hyperspectral HypsIRI and multispectral Landsat 8 and
728 Sentinel-2 imagery for multi-seasonal, regional land-cover mapping. *Remote Sensing of Environment*,
729 *200*, 311-325
- 730 Daughtry, C., Walthall, C.L., Kim, M.S., De Colstoun, E.B., & McMurtrey Iii, J.E. (2000). Estimating corn
731 leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment*, *74*,
732 229-239
- 733 Díaz, S., Kattge, J., Cornelissen, J.H., Wright, I.J., Lavorel, S., Dray, S., et al. (2016). The global spectrum of
734 plant form and function. *Nature*, *529*, 167-171
- 735 Díaz, S., Settele, J., Brondízio, E., Ngo, H.T., Guèze, M., Agard, J., et al. (2019). Summary for policymakers
736 of the global assessment report on biodiversity and ecosystem services of the Intergovernmental
737 Science-Policy Platform on Biodiversity and Ecosystem Services.
738 https://www.ipbes.net/sites/default/files/downloads/spm_unedited_advance_for_posting_htn.pdf, -
739 ADVANCE UNEDITED VERSION –
- 740 Díaz, S. & Cabido, M. (2001). Vive la différence: plant functional diversity matters to ecosystem
741 processes. *Trends in ecology & evolution*, *16*, 646-655
- 742 Ely, K.S., Burnett, A.C., Lieberman-Cribbin, W., Serbin, S.P., & Rogers, A. (2019). Spectroscopy can predict
743 key leaf traits associated with source–sink balance and carbon–nitrogen status. *Journal of experimental*
744 *botany*, *70*, 1789-1799
- 745 Enquist, B.J., Norberg, J., Bonser, S.P., Violle, C., Webb, C.T., Henderson, A., et al. (2015). Scaling from
746 traits to ecosystems: developing a general trait driver theory via integrating trait-based and metabolic
747 scaling theories. In Anonymous *Advances in ecological research* (pp. 249-318). : Elsevier
- 748 Enquist, B.J., Bentley, L.P., Shenkin, A., Maitner, B., Savage, V., Michaletz, S., et al. (2017). Assessing
749 trait-based scaling theory in tropical forests spanning a broad temperature gradient. *Global Ecology and*
750 *Biogeography*, *26*, 1357-1373
- 751 Esquivel-Muelbert, A., Baker, T.R., Dexter, K.G., Lewis, S.L., Brienen, R.J., Feldpausch, T.R., et al. (2019).
752 Compositional response of Amazon forests to climate change. *Global Change Biology*, *25*, 39-56
- 753 Farr, T.G., Rosen, P.A., Caro, E., Crippen, R., Duren, R., Hensley, S., et al. (2007). The shuttle radar
754 topography mission. *Reviews of Geophysics*, *45*
- 755 Fell, M., & Ogle, K. (2018). Refinement of a theoretical trait space for North American trees via
756 environmental filtering. *Ecological Monographs*, *88*, 372-384
- 757 Ferreira, M.P., Wagner, F.H., Aragão, L.E., Shimabukuro, Y.E., & de Souza Filho, Carlos Roberto. (2019).
758 Tree species classification in tropical forests using visible to shortwave infrared WorldView-3 images and
759 texture analysis. *ISPRS journal of photogrammetry and remote sensing*, *149*, 119-131
- 760 Ferwerda, J.G., Skidmore, A.K., & Mutanga, O. (2005). Nitrogen detection with hyperspectral normalized
761 ratio indices across multiple plant species. *International Journal of Remote Sensing*, *26*, 4083-4095
- 762 Gallagher, R.V., Falster, D.S., Maitner, B.S., Salguero-Gómez, R., Vandvik, V., Pearse, W.D., et al. (2020).
763 Open Science principles for accelerating trait-based science across the Tree of Life. *Nature ecology &*
764 *evolution*, *4*, 294-303

765 Georganos, S., Grippa, T., Gadiaga, A.N., Linard, C., Lennert, M., Vanhuyse, S., et al. (2019).
766 Geographical Random Forests: A Spatial Extension of the Random Forest Algorithm to Address Spatial
767 Heterogeneity in Remote Sensing and Population Modelling. *Geocarto International*, 1-12

768 Grime, J.P. (1998). Benefits of plant diversity to ecosystems: immediate, filter and founder effects.
769 *Journal of Ecology*, 86, 902-910

770 Guan, K., Pan, M., Li, H., Wolf, A., Wu, J., Medvigy, D., et al. (2015). Photosynthetic seasonality of global
771 tropical forests constrained by hydroclimate. *Nature Geoscience*, 8, 284

772 Gvozdevaite, A., Oliveras, I., Domingues, T.F., Peprah, T., Boakye, M., Afriyie, L., et al. (2018). Leaf-level
773 photosynthetic capacity dynamics in relation to soil and foliar nutrients along forest–savanna
774 boundaries in Ghana and Brazil. *Tree physiology*, 38, 1912-1925

775 Haralick, R.M., Shanmugam, K., & Dinstein, I.H. (1973). Textural features for image classification. *IEEE*
776 *transactions on systems, man, and cybernetics*, 610-621

777 Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning, 2nd edn* New York.

778 Hawthorne, W.D. (1995). Ecological profiles of Ghanaian forest trees. *Tropical forestry papers*

779 Hédli, R., Svátek, M., Dančák, M., Rodzay, A.W., Salleh, A.B., & Kamariah, A.S. (2009). A new technique
780 for inventory of permanent plots in tropical forests: a case study from lowland dipterocarp forest in
781 Kuala Belalong, Brunei Darussalam. *Blumea-Biodiversity, Evolution and Biogeography of Plants*, 54, 124-
782 130

783 Hengl, T., de Jesus, J.M., Heuvelink, G.B., Gonzalez, M.R., Kilibarda, M., Blagotić, A., et al. (2017).
784 SoilGrids250m: Global gridded soil information based on machine learning. *PLoS one*, 12, e0169748

785 Hlatshwayo, S.T., Mutanga, O., Lottering, R.T., Kiala, Z., & Ismail, R. (2019). Mapping forest aboveground
786 biomass in the reforested Buffelsdraai landfill site using texture combinations computed from SPOT-6
787 pan-sharpened imagery. *International Journal of Applied Earth Observation and Geoinformation*, 74, 65-
788 77

789 Huang, W., Ratkowsky, D.A., Hui, C., Wang, P., Su, J., & Shi, P. (2019). Leaf fresh weight versus dry
790 weight: which is better for describing the scaling relationship between leaf biomass and leaf area for
791 broad-leaved plants? *Forests*, 10, 256

792 Hubau, W., Lewis, S.L., Phillips, O.L., Affum-Baffoe, K., Beeckman, H., Cuní-Sanchez, A., et al. (2020).
793 Asynchronous carbon sink saturation in African and Amazonian tropical forests. *Nature*, 579, 80-87

794 Jetz, W., Cavender-Bares, J., Pavlick, R., Schimel, D., Davis, F.W., Asner, G.P., et al. (2016). Monitoring
795 plant functional diversity from space. *Nature Plants*, 2, 16024

796 Jucker, T., Bongalov, B., Burslem, D.F., Nilus, R., Dalponte, M., Lewis, S.L., et al. (2018). Topography
797 shapes the structure, composition and function of tropical forest landscapes. *Ecology Letters*, 21, 989-
798 1000

799 Juneau, K.J., & Tarasoff, C.S. (2012). Leaf area and water content changes after permanent and
800 temporary storage. *PLoS One*, 7, e42604

801 Kattge, J., Bönsch, G., Díaz, S., Lavorel, S., Prentice, I.C., Leadley, P., et al. (2020). TRY plant trait
802 database–enhanced coverage and open access. *Global Change Biology*

803 Kissling, W.D., Walls, R., Bowser, A., Jones, M.O., Kattge, J., Agosti, D., et al. (2018). Towards global data
804 products of Essential Biodiversity Variables on species traits. *Nature ecology & evolution*, 2, 1531-1540

805 Laughlin, D.C., Strahan, R.T., Adler, P.B., & Moore, M.M. (2018). Survival rates indicate that correlations
806 between community-weighted mean traits and environments can be unreliable estimates of the
807 adaptive value of traits. *Ecology Letters*, *21*, 411-421

808 Laurin, G.V., Puletti, N., Hawthorne, W., Liesenberg, V., Corona, P., Papale, D., et al. (2016).
809 Discrimination of tropical forest types, dominant species, and mapping of functional guilds by
810 hyperspectral and simulated multispectral Sentinel-2 data. *Remote Sensing of Environment*, *176*, 163-
811 176

812 Lebrija-Trejos, E., Pérez-García, E.A., Meave, J.A., Bongers, F., & Poorter, L. (2010). Functional traits and
813 environmental filtering drive community assembly in a species-rich tropical system. *Ecology*, *91*, 386-398

814 Lewis, S.L., Edwards, D.P., & Galbraith, D. (2015). Increasing human dominance of tropical forests.
815 *Science*, *349*, 827-832

816 Lortie, C.J., Brooker, R.W., Choler, P., Kikvidze, Z., Michalet, R., Pugnaire, F.I., et al. (2004). Rethinking
817 plant community theory. *Oikos*, *107*, 433-438

818 Ma, X., Mahecha, M.D., Migliavacca, M., van der Plas, F., Benavides, R., Ratcliffe, S., et al. (2019).
819 Inferring plant functional diversity from space: the potential of Sentinel-2. *Remote Sensing of*
820 *Environment*, *233*, 111368

821 Malhi, Y., Rowland, L., Aragao, L. E. O. C., & Fisher, R.A. (2018). New insights into the variability of the
822 tropical land carbon cycle from the El Niño of 2015/2016. *Philosophical transactions of the Royal Society*
823 *of London. Series B, Biological sciences*, *373*, 10.1098/rstb.2017.0298

824 Malhi, Y., Roberts, J.T., Betts, R.A., Killeen, T.J., Li, W., & Nobre, C.A. (2008). Climate change,
825 deforestation, and the fate of the Amazon. *Science*, *319*, 169-172

826 Martin, M.E., Plourde, L.C., Ollinger, S.V., Smith, M., & McNeil, B.E. (2008). A generalizable method for
827 remote sensing of canopy nitrogen across a wide range of forest ecosystems. *Remote Sensing of*
828 *Environment*, *112*, 3511-3519

829 Martin, R.E., Chadwick, K.D., Brodrick, P.G., Carranza-Jimenez, L., Vaughn, N.R., & Asner, G.P. (2018). An
830 approach for foliar trait retrieval from airborne imaging spectroscopy of tropical forests. *Remote*
831 *Sensing*, *10*, 199

832 Martin, R.E., Asner, G.P., Bentley, L.P., Shenkin, A., Salinas, N., Huaypar, K.Q., et al. (2020). Covariance of
833 Sun and Shade Leaf Traits Along a Tropical Forest Elevation Gradient. *Frontiers in plant science*, *10*, 1810

834 McDowell, N., Allen, C.D., Anderson-Teixeira, K., Brando, P., Brienen, R., Chambers, J., et al. (2018).
835 Drivers and mechanisms of tree mortality in moist tropical forests. *New Phytologist*, *219*, 851-869

836 Mueller, T., Dressler, G., Tucker, C., Pinzon, J., Leimgruber, P., Dubayah, R., et al. (2014). Human land-use
837 practices lead to global long-term increases in photosynthetic capacity. *Remote Sensing*, *6*, 5717-5731

838 Naeem, S., Bunker, D.E., Hector, A., Loreau, M., & Perrings, C. (2009). *Biodiversity, Ecosystem*
839 *Functioning, and Human Wellbeing: An Ecological and Economic Perspective*. (pp. 388). : Oxford
840 University Press

841 Navarro, L.M., Fernández, N., Guerra, C., Guralnick, R., Kissling, W.D., Londoño, M.C., et al. (2017).
842 Monitoring biodiversity change through effective global coordination. *Current opinion in environmental*
843 *sustainability*, *29*, 158-169

844 Nunes, M., Davey, M., & Coomes, D. (2017). On the challenges of using field spectroscopy to measure
845 the impact of soil type on leaf traits.

846 Nunes, M.H., Both, S., Bongalov, B., Brelsford, C., Khoury, S., Burslem, D.F., et al. (2019). Changes in leaf
847 functional traits of rainforest canopy trees associated with an El Niño event in Borneo. *Environmental*
848 *Research Letters*, 14, 085005

849 Oliver, T.H., Heard, M.S., Isaac, N.J., Roy, D.B., Procter, D., Eigenbrod, F., et al. (2015). Biodiversity and
850 resilience of ecosystem functions. *Trends in Ecology & Evolution*, 30, 673-684

851 Oliveras, I., Bentley, L., Fyllas, N.M., Gvozdevaite, A., Shenkin, A.F., Prepah, T., et al. (2020). The
852 influence of taxonomy and environment on leaf trait variation along tropical abiotic gradients. *Frontiers*
853 *in Forests and Global Change*, 3, 18

854 Ollinger, S.V. (2011). Sources of variability in canopy reflectance and the convergent properties of
855 plants. *New Phytologist*, 189, 375-394

856 Pacifici, M., Foden, W.B., Visconti, P., Watson, J.E., Butchart, S.H., Kovacs, K.M., et al. (2015). Assessing
857 species vulnerability to climate change. *Nature Climate Change*, 5, 215-224

858 Parmentier, I., Malhi, Y., Senterre, B., Whittaker, R.J., ATDN, Alonso, A., et al. (2007). The odd man out?
859 Might climate explain the lower tree α -diversity of African rain forests relative to Amazonian rain
860 forests? *Journal of Ecology*, 95, 1058-1071

861 Prada, C.M., Morris, A., Andersen, K.M., Turner, B.L., Caballero, P., & Dalling, J.W. (2017). Soils and
862 rainfall drive landscape-scale changes in the diversity and functional composition of tree communities in
863 premontane tropical forest. *Journal of vegetation science*, 28, 859-870

864 Price, J.C. (1994). How unique are spectral signatures? *Remote Sensing of Environment*, 49, 181-186

865 Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H., & Sorooshian, S. (1994). A modified soil adjusted vegetation
866 index. *Remote Sensing of Environment*, 48, 119-126

867 Quesada, C.A., Phillips, O.L., Schwarz, M., Czimczik, C.I., Baker, T.R., Patiño, S., et al. (2012). Basin-wide
868 variations in Amazon forest structure and function are mediated by both soils and climate.
869 *Biogeosciences*, 9

870 R Development Core Team. (2014). R: A language and environment for statistical computing. R
871 foundation for statistical computing, Vienna, Austria. ISBN 3-900051-07-0, URL [http://www.R-](http://www.R-project.org)
872 [project.org](http://www.R-project.org).

873 Reich, P.B. (2012). Key canopy traits drive forest productivity. *Proceedings of the Royal Society B:*
874 *Biological Sciences*, 279, 2128-2134

875 Sarker, L.R., & Nichol, J.E. (2011). Improved forest biomass estimates using ALOS AVNIR-2 texture
876 indices. *Remote Sensing of Environment*, 115, 968-977

877 Schlemmer, M., Gitelson, A., Schepers, J., Ferguson, R., Peng, Y., Shanahan, J., et al. (2013). Remote
878 estimation of nitrogen and chlorophyll contents in maize at leaf and canopy levels. *International Journal*
879 *of Applied Earth Observation and Geoinformation*, 25, 47-54

880 Schneider, F.D., Morsdorf, F., Schmid, B., Petchey, O.L., Hueni, A., Schimel, D.S., et al. (2017). Mapping
881 functional diversity from remotely sensed morphological and physiological forest traits. *Nature*
882 *communications*, 8, 1441

883 Serbin, S.P., Singh, A., Desai, A.R., Dubois, S.G., Jablonski, A.D., Kingdon, C.C., et al. (2015). Remotely
884 estimating photosynthetic capacity, and its response to temperature, in vegetation canopies using
885 imaging spectroscopy. *Remote Sensing of Environment*, 167, 78-87

886 Shenkin, A., Bentley, L.P., Oliveras, I., Salinas, N., Adu-Bredu, S., Marimon, B.H., et al. (2019). The
887 Influence of Ecosystem and Phylogeny on Tropical Tree Crown Size and Shape. *BioRxiv*, 789255

888 Shiklomanov, A.N., Dietze, M.C., Viskari, T., Townsend, P.A., & Serbin, S.P. (2016). Quantifying the
889 influences of spectral resolution on uncertainty in leaf trait estimates through a Bayesian approach to
890 RTM inversion. *Remote Sensing of Environment*, *183*, 226-238

891 Soudzilovskaia, N.A., Elumeeva, T.G., Onipchenko, V.G., Shidakov, I.I., Salpagarova, F.S., Khubiev, A.B., et
892 al. (2013). Functional traits predict relationship between plant abundance dynamic and long-term
893 climate warming. *Proceedings of the National Academy of Sciences*, *110*, 18180-18184

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

896 Szabó, L., Burai, P., Deák, B., Dyke, G.J., & Szabó, S. (2019). Assessing the efficiency of multispectral
897 satellite and airborne hyperspectral images for land cover mapping in an aquatic environment with
898 emphasis on the water caltrop (*Trapa natans*). *International Journal of Remote Sensing*, *40*, 5192-5215

899 Taubert, F., Fischer, R., Groeneveld, J., Lehmann, S., Müller, M.S., Rödig, E., et al. (2018). Global patterns
900 of tropical forest fragmentation. *Nature*, *554*, 519

901 Townsend, P.A., Foster, J.R., Chastain, R.A., & Currie, W.S. (2003). Application of imaging spectroscopy to
902 mapping canopy nitrogen in the forests of the central Appalachian Mountains using Hyperion and
903 AVIRIS. *IEEE Transactions on Geoscience and Remote Sensing*, *41*, 1347-1354

904 Ustin, S.L., & Gamon, J.A. (2010). Remote sensing of plant functional types. *New Phytologist*, *186*, 795-
905 816

906 Van der Plas, F., Ratcliffe, S., Ruiz-Benito, P., Scherer-Lorenzen, M., Verheyen, K., Wirth, C., et al. (2018).
907 Continental mapping of forest ecosystem functions reveals a high but unrealised potential for forest
908 multifunctionality. *Ecology Letters*, *21*, 31-42

909 Vincent, G. (2001). Leaf photosynthetic capacity and nitrogen content adjustment to canopy openness in
910 tropical forest tree seedlings. *Journal of Tropical Ecology*, *17*, 495-509

911 Walker, A.P., McCormack, M.L., Messier, J., Myers-Smith, I.H., & Wullschleger, S.D. (2017). Trait
912 covariance: the functional warp of plant diversity? *New Phytologist*, *216*, 976-980

913 Walker, A.P., Beckerman, A.P., Gu, L., Kattge, J., Cernusak, L.A., Domingues, T.F., et al. (2014). The
914 relationship of leaf photosynthetic traits— V_{cmax} and J_{max} —to leaf nitrogen, leaf phosphorus, and
915 specific leaf area: a meta-analysis and modeling study. *Ecology and Evolution*, *4*, 3218-3235

916 Wallis, C.I., Homeier, J., Peña, J., Brandl, R., Farwig, N., & Bendix, J. (2019). Modeling tropical montane
917 forest biomass, productivity and canopy traits with multispectral remote sensing data. *Remote Sensing
918 of Environment*, *225*, 77-92

919 Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., et al. (2004). The worldwide
920 leaf economics spectrum. *Nature*, *428*, 821-827

921

923 **Figure 1. Diagram summarising the steps followed to assign trait values per Sentinel-2 pixel.** 1) First the
 924 vegetation plots are defined based on the GEM (Global Environmental Monitoring) dataset and 2) from
 925 each vegetation plot the corner coordinates are extracted. 3) From each vegetation plot the XY position
 926 of each stem ≥ 10 cm DBH is extracted and 4) the crown horizontal area is calculated based on the protocol
 927 described in the methods section. 5) Then the Sentinel-2 imagery for the study area is processed to level
 928 2A using the ESA SNAP toolbox and 6) the vegetation plot is overlaid in the Sentinel-2 image based on its
 929 corner coordinates. In this last step (6) each pixel defines a ‘subplot’ which is the unit used to calculate
 930 the trait community weighted mean based on the crown area of the trees that are contained by that pixel.
 931 In 6) n refers to a given tree in a given pixel, trait i represents a given trait and x and y are values for that
 932 trait. The image used as an example in step (1) was taken by Jesus Aguirre-Gutierrez over a vegetation
 933 plot using a multispectral ALTUM camera mounted on an Inspire 1 drone.”

934 **Figure 2. Comparison of trait distributions across tropical regions.** The boxplots are based on the pixel-
 935 level (10×10 m) community trait values for each trait and region ($n=403$ for Australia, 449 for Brazil-NX,
 936 302 for Brazil-ST, 464 for Gabon, 620 for Ghana, 976 for Malaysia and 1280 for Peru). Horizontal lines in
 937 each boxplot show the median value and vertical lines are the whiskers that extend to the largest value
 938 or not further than 1.5 times the inter-quartile range. For some locations information for all traits was not
 939 available. For full details in significant differences in mean trait values among locations see Table S4. Brazil
 940 -NX: Nova Xavantina; Brazil -ST: Santarem.

941 **Figure 3. Model predictions to the 20% test data from the general model which was fitted with 80% of**
 942 **the trait data from across the tropics.** Grey dots are the observed against predicted trait values of the
 943 pixel-level (10×10 m) community weighted mean traits from the test dataset. The black line shows the
 944 1:1 relationship between observed and predicted values. Model prediction accuracy is shown in the top
 945 left. Full model results are shown in Table 4.

946 **Figure 4. Spatial predictions of trait distributions for a selected subset of plant traits and locations.**
 947 The map (middle) shows the locations of vegetation plots that were used during the modelling
 948 framework. The spatial predictions (top and bottom rows) were obtained using the general models
 949 (Table 4) for each of the traits and locations at a 10×10 m pixel resolution. The approximated location
 950 of each vegetation plot used is shown as a white square within each spatial prediction map (for
 951 visualisation purposes white squares are not scaled to the plot’s real size). Spatial predictions for other
 952 traits can be found in Fig. S1-S7.

953 **Figure 5. Models predictions to the 20% test data from the regional models fitted with 80% of the trait**
 954 **data from each region across the tropics.** Each colour represents an individual regional model and the
 955 coloured symbols are the observed against predicted trait values of the pixel-level (10×10 m) community
 956 weighted mean traits from the test dataset. The black line shows the 1:1 relationship between observed
 957 and predicted values. Model prediction accuracy is shown in Table 5. Full model results are shown in Table
 958 S5. Brazil -NX: Nova Xavantina; Brazil -ST: Santarem.

959 **Figure 6. Group median variable importance of spectral remote sensing, environmental and soil**
 960 **related variables for determining functional trait predictions in the global model.** Variable importance
 961 (Y axis) ranges from 0 (no importance) to 1 (highest importance) and represents the decrease in node
 962 impurities from splitting on the variable, averaged over all trees and derived from the Out of Bag error,
 963 the resulting value has been standardised to a 0-1 scale for comparison purposes. The spectral group (S2
 964 -Spectral) contains the select raw bands from the Sentinel-2 and the vegetation indices; Texture
 965 parameters (S2 -Texture) contain the Correlation and Entropy metrics from the grey level co-occurrence
 966 matrix obtained from the vegetation indices; Climate contains all climatic variables; Soil-Terrain contains
 967 all soil characteristics and slope. All variables are described in Table 3.

Table 1. Collection details for vegetation plots and plant functional traits. A total of 2434 individual trees were sampled for functional traits.

Location	Species sampled for traits	Plot code	Size (ha)	Centroid coordinates		Date of collection	
				X	Y	Vegetation census	Traits
Australia	81	AEP-02	0.5	145.586	-17.146	2011	June-September 2015
		AEP-03	0.5	145.592	-17.088		
		DRO-01	0.9	145.430	-16.103		
		ROB-06	1	145.630	-17.121		
Ghana	63	ANK-01	1	-2.696	5.268	2013	October-March 2015/2016gramm
		ANK-03	1	-2.692	5.271		
		BOB-01	1	-1.339	6.691	2015	
		BOB-02	1	-1.319	6.704		
Gabon	41	LPG-01	1	11.574	-0.174	2014	February-March 2017
		LPG-02	1	11.615	-0.216		
		MNG-04	1	9.324	0.577	2016	
Brazil -NX	64	NXV-01	1	-52.352	-14.708	2014	March-May 2014
		NXV-02	1	-52.351	-14.701		
		VCR-02	1	-52.168	-14.832		
		NXV-10-1	0.1	-52.353	-14.713		
		NXV-10-2	0.1	-52.352	-14.713		
		NXV-10-3	0.1	-52.351	-14.713		
		NXV-10-4	0.1	-52.349	-14.713		
		NXV-10-5	0.1	-52.346	-14.713		
		NXV-10-6	0.1	-52.349	-14.712		
		NXV-10-7	0.1	-52.348	-14.711		
NXV-10-8	0.1	-52.347	-14.711				
NXV-10-9	0.1	-52.347	-14.711				
NXV-10-10	0.1	-52.346	-14.712				
Brazil -ST	136	261-10	0.25	-55.005	-3.019	2014	

		261-9	0.25	-55.015	-3.040		
		363-6	0.25	-54.956	-3.337		August-September
		363-3	0.25	-54.963	-3.297		2015
		363-7	0.25	-54.961	-3.321		
		ESP-01	1	-71.595	-13.176		
		PAN-02	1	-71.263	-12.650		
		SPD-01	1	-71.542	-13.047	2013	
		SPD-02	1	-71.537	-13.049		
		TRU-04	1	-71.589	-13.106		April-November
		WAY-01	1	-71.587	-13.191		2013
		ACJ-01	1	-71.632	-13.147		
		PAN-03	1	-71.274	-12.638	2014	
		TAM-05	1	-69.271	-12.830		
		TAM-06	1	-69.296	-12.839		
		SAF-01	1	4.732	117.619		
		SAF-02	1	4.739	117.617		
		SAF-03	1	4.691	117.588		
		SAF-04	1	4.765	117.700	2016	July-December
		DAN-04	1	4.951	117.796		2015
		DAN-05	1	4.953	117.793		
		MLA-01	1	4.747	116.970		
		MLA-02	1	4.754	116.950		

Brazil -NX: Nova Xavantina; Brazil -ST: Santarem; Malaysia: Malaysian Borneo.

Table 2. Description of plant functional traits collected across the tropics and their relevance under a changing environment.

Trait	Abbreviation	Units	Description	Relevance	References*
Leaf area	Area	cm ²	One-sided area of the leaf	Higher leaf area could result in higher levels of light capture and photosynthetic activity.	
Specific leaf area	SLA	m ² g ⁻¹	One-sided area of a leaf divided by dry mass	Relevant for photosynthetic capacity, light capture, water loss, net assimilation rate, leaf life span.	
Leaf thickness	Thickness	mm	Thickness of a fresh leaf	Trade-off between decreasing water transpiration at the expense of higher construction investment and probably lower photosynthetic efficiency in thicker leaves.	
Leaf nitrogen content	N	%	Content per unit dry leaf mass	Nutrient relevant for metabolic reactions, including light capture, related to photosynthetic capacity and growth. Restricted availability of some nutrients may limit plant carbon acquisition and growth.	(Walker et al., 2014, Wright et al., 2004, Juneau & Tarasoff, 2012, Díaz et al., 2016, Hawthorne, 1995, Chave et al., 2006, Huang et al., 2019)
Leaf phosphorus content	P	%			
Leaf carbon content	C	%			
Leaf calcium content	Ca	%			
Leaf potassium content	K	%			
Leaf magnesium content	Mg	%			
Leaf water content	LWC	%	Amount of water in the leaf relative to its dry and fresh mass	Leaf mass is a proxy of leaf biomass investment which may vary depending on environmental conditions and phenology of species. It has been considered relevant for photosynthetic potential.	
Leaf dry mass	Dry mass	g	Mass of a dry leaf		
Leaf fresh mass	Fresh mass	g	Mass of a fresh leaf		
A _{max}	A _{max}	μmol m ⁻² s ⁻¹	Light-saturated maximum rates of net photosynthesis at saturated CO ₂ (2000 ppm CO ₂)	Indicate the maximum CO ₂ assimilation and are thus indices of leaf photosynthetic capacity	
A _{sat}	A _{sat}	μmol m ⁻² s ⁻¹	Light-saturated rates of net photosynthesis at ambient CO ₂ concentration (2000 ppm CO ₂)		

*References are not exhaustive.

Table 3. Spectral remote sensing, environmental and soil related variables used during the modelling protocol. All climatic variables but slope were calculated using a climatology of 30 years (1986-2015). All soil variables were calculated for the top 30cm soil layer. Sentinel-2 band wavelengths (nm) are given in parenthesis after the band name.

Type	Variable	Description	References	
RS	B2 (490), B3 (560), B4 (665), B8 (842)	Sentinel-2 bands with spatial resolution of 10m	www.esa.int	
	B5 (705), B6 (740), B7 (783), B11 (1610), B12 (2190)	Sentinel-2 bands with spatial resolution of 20m		
	MCARI	Modified Chlorophyll Absorption in Reflectance Index	(Daughtry et al., 2000)	
	MSAVI2	Modified Soil Adjusted Vegetation Index 2	(Qi et al., 1994)	
	NDRE	Normalized Difference Red edge Index	(Barnes et al., 2000)	
	Texture		Entropy, calculated for vegetation indices	(Haralick et al. 1973)
			Correlation, calculated for vegetation indices	
	Climate	MCWD	Mean annual climatic water deficit	(Abatzoglou et al. 2018)
MATmax		Mean maximum annual temperature		
SM		Soil moisture as a water balance indicator		
SRAD		Downward Solar Radiation		
Soil-Terrain	eCEC	Cation Exchange Capacity ($\text{mmol}^+ / \text{kg}^{-1}$)	Plot level soil data from the Global Environmental Monitoring (GEM) database	
	pH	Soil pH (H ₂ O solution)		
	Clay (%)	Amount of clay (weight %)		
	Sand (%)	Amount of sand (weight %)		
	Slope	Terrain slope (30 m resolution)		(Farr et al. 2007)

Table 4. Statistical results on the test data (20% of full dataset) for the global trait distribution models. The prediction accuracy is shown by the R^2 score.

Type	Trait	MAE	RMSE	R^2
Morphological and structural	Area (cm ²)	28.32	39.854	0.43
	Dry mass (g)	0.349	0.48	0.27
	Fresh mass (g)	0.799	1.075	0.26
	SLA (m ² g ⁻¹)	0.001	0.001	0.50
	Thickness (mm)	0.034	0.046	0.52
Chemistry	LWC (%)	3.718	4.886	0.36
	C (%)	1.237	1.615	0.70
	Ca (%)	0.14	0.204	0.64
	K (%)	0.133	0.186	0.63
	Mg (%)	0.055	0.075	0.46
	N (%)	0.23	0.3	0.59
	P (%)	0.015	0.02	0.59
Photosynthetic	Amax ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	2.89	3.937	0.67
	Asat ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	1.297	1.734	0.55

MAE: Mean Absolute Error; RMSE: Root mean square error.

975

Table 5. Prediction accuracy (R^2) on the testing data among regions (shaded Region mean R^2 column) and functional traits (shaded Trait mean R^2 row). Not shaded values in the table show the prediction accuracy (R^2) on the test data per region and trait.

Location	P (%)	Amax ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	Ca (%)	N (%)	Thickness (mm)	Asat ($\mu\text{mol m}^{-2} \text{s}^{-1}$)	Mg (%)	C (%)	SLA ($\text{m}^2 \text{g}^{-1}$)	LWC (%)	K (%)	Dry mass (g)	Area (cm^2)	Fresh mass (g)	Region mean R^2
Australia	0.21	-	0.33	0.17	0.21	0.03	0.12	0.34	0.25	-	0.06	-	0.35	-	0.21
Brazil -NX	0.68	0.42	0.49	0.52	0.66	0.52	0.46	-	0.54	0.07	0.07	0.38	0.08	0.31	0.40
Brazil -ST	0.47	-	0.15	0.30	0.42	-	0.28	0.07	0.29	0.05	0.29	0.25	0.04	0.18	0.23
Gabon	0.60	-	0.39	0.50	0.23	-	0.52	0.22	0.15	0.38	0.24	0.22	0.11	0.11	0.31
Ghana	0.47	0.49	0.53	0.52	0.22	0.36	0.15	-	0.14	-	0.23	-	0.29	-	0.34
Malaysia	0.34	0.48	0.50	0.27	0.31	0.36	0.31	0.38	0.06	0.24	0.28	0.07	0.11	0.03	0.27
Peru	0.49	0.36	0.69	0.44	0.64	0.38	0.46	0.47	0.32	0.34	0.30	0.09	0.18	0.20	0.38
Trait mean R^2	0.47	0.44	0.44	0.39	0.38	0.33	0.33	0.30	0.25	0.22	0.21	0.20	0.17	0.17	

976 - :no data available; Brazil -NX: Nova Xavantina; Brazil -ST: Santarem.

977