Unstratified forests dominate the tropics especially in regions with lower fertility or higher temperatures

Christopher E. Doughty*[1], Camille Gaillard[1], Patrick Burns[1], Jenna Keany[1], Andrew Abraham[1], Yadvinder Malhi[2], Jesus Aguirre-Gutierrez[2], George Koch[3], Patrick Jantz[1], Alexander Shenkin[1], Hao Tang[4]

[1]School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ, USA

[2]Environmental Change Institute, School of Geography and the Environment, University of Oxford, Oxford, UK

[3]Department of Biology, Northern Arizona University, Flagstaff, AZ, USA

[4] Department of Geography, National University of Singapore, Singapore

*to whom correspondence should be addressed – chris.doughty@nau.edu

Keywords –GEDI, tropical forests, stratification, biomass

Abstract – The stratified nature of tropical forest structure had been noted by early explorers, but until recent use of satellite-based LiDAR (GEDI, or Global Ecosystems Dynamics Investigation LiDAR), there has been no way to quantify stratification across all tropical forests. Understanding stratification is important because by some estimates, a majority of the world's species inhabit tropical forest canopies. Stratification can modify vertical microenvironment, and thus can affect a species' susceptibility to global warming. A better understanding of structure could also improve predictions of biomass across the tropics. Here we find that, based on analyzing each GEDI 25m diameter footprint in tropical forests (after screening for human impact), most footprints (60-90%) do not have multiple layers of vegetation. We find stratification depends on the spatial resolution of the pixel (e.g. going from a 25m footprint to a 1 ha footprint will impact the results). However, with a 25m footprint, the most common forest structure has a minimum plant area index (PAI) at ~40m followed by an increase in PAI until ~15m followed by a decline in PAI to the ground layer (described hereafter as a one peak footprint). However, there are large geographic patterns to forest structure within the Amazon basin (ranging between 60-90% one peak) and between the Amazon (79±9sd) and SE Asia or Africa (72±14 v 73±11). The number of canopy layers is significantly correlated with tree height (r^2=0.12), forest biomass (r^2=0.14), maximum temperature (T_{max}) (r^2=0.05), vapor pressure deficit (VPD) (r^2=0.03) and soil fertility proxies (e.g. total cation exchange capacity -r^2=0.01). Certain boundaries, like the Pebas Formation and Ecoregions, clearly delineate continental scale structural changes. More broadly, deviation from more ideal conditions (e.g. lower fertility or higher temperatures) leads to shorter, less stratified forests with lower biomass.

Introduction

Early Western visitors describe tropical forests as *horror vacui* (nature abhorring a vacuum) since vegetation was “anxious to fill every available space with stems and leaves”, which was a change from more open temperate forests (Richards, 1952). However, a closer examination of tropical forests revealed structure or stratification with “a discernible, though complicated, arrangement in space” (Richards, 1952). Halle et al 1980 built on this with their influential work identifying twenty-three unique tree architecture types and delving into the drivers of forest architecture (Halle, Oldeman and Tomlinson, 1980). They recognized that because tropical forests had fewer hydraulic or cold temperature constraints, the tropics was a good place to study the potential for trees to fill vertical space. They developed theories using detailed 20 by 30m vertical profiles of old growth canopies where “trees of the present” occupy space in the upper canopy as well as in a second layer of increased light at 15-20m where sunflecks converge. This old growth forest architecture would result in a stratified or layered forest (artistically rendered in Figure 1) unlike younger pioneer forests with a single upper canopy strata. We define a stratified or multilayer forest as having two or more peaks in horizontal vegetation (e.g. overstory and midstory in Fig 1) with a lower amount of vegetation between them. Others have quantified stratification in different ways and found both temperate and tropical forests commonly have 2-3 tree layers (Baker and Wilson, 2000). However, tropical forest stratification has not been addressed previously at high spatial resolutions (e.g. 25m diameter) at the global scale.

More recently, the Global Ecosystems Dynamic Investigation (GEDI) on the International Space Station (ISS)-based LiDAR instrument (Dubayah *et al.*, 2020), allows us for the first time to peer into the structure of tropical forests in unprecedented resolution at a global scale. Prior to GEDI, there were other satellite lidar instruments (e.g. GLAS on ICESAT-1) used for measuring vegetation structure at large scale (Tang *et al.*, 2016; Tang and Dubayah, 2017), but these were lower resolution, much more sparse, and focused on polar regions. At a more regional scale, aircraft and terrestrial lidar have shown detailed individual tropical forest tree architectures. For instance, aircraft lidar in tropical Peru found that tree architecture or shape (height of peak canopy volume (P) divided by canopy height) was highly correlated with canopy height (Asner *et al.*, 2014) and in Panama others successfully predicted the tree size distributions with airborne lidar (Taubert *et al.*, 2021). At a global scale, Ehbrecht et al 2021 scaled up terrestrial laser scanning to show that forest structural complexity is a function of annual precipitation and precipitation seasonality (Ehbrecht *et al.*, 2021). Both simulation and sensitivity analysis suggest that high-quality GEDI data is able to provide measurements of similar accuracy for variables like plant area index (PAI) or species richness in the tropics when compared to aircraft and terrestrial lidar (Marselis *et al.*, 2018, 2020). These different lidar tools (that inform on structure from the individual tree to global scale) can help us to better understand forest stratification across the tropics globally.

Forest stratification may be due to genetic constraints that evolved over time (floristics) or trees not achieving their genetic heights (potential height under optimal environmental conditions). The debate about what sets the upper limits of tree height largely involves either hydraulic limitation (Koch et al 2004), mechanical limitation, or environmental factors such as wind speed (Jackson *et al.*, 2021). Certain factors drive heights such as the need to overtop competitors or disperse seeds while other factors reduce it such as hydraulic failure and vulnerability to wind. Environment alone could also directly impact tree height and structure, with hydraulic limitations, carbon deficiencies, or wind regimes causing trees to not being able to achieve their genetic height. There is a literature describing how the environment (soils or climate) impacts the species composition in tropical forests. For instance, Amazonian species composition may follow a south-west/north-east soil fertility gradient and a north-west/south-east precipitation gradient (ter Steege *et al.*, 2006). Soil cation concentrations are the primary driver of floristic variation for Amazonian trees (Tuomisto *et al.*, 2019) with climate being of secondary importance. However, in central African forests, climate is considered to be the driving factor of floristic patterns (Réjou-Méchain *et al.*, 2021).

Structure matters because it can give us new insights into forest biomass, which is the primary goal of GEDI. Currently the L4A product for tropical forests uses relative height (RH) RH98 and RH50 to predict a median AGBD of 300 Mg Ha^{-1} for tropical forests (0.66 r^2 and RMSE of 10.4) (Duncanson *et al.*, 2022). Ecological theory suggests that a stratified forest with more large emergent trees is indicative of an older forest (Halle, Oldeman and Tomlinson, 1980), which generally has higher biomass and carbon content. Therefore, incorporating canopy layers may improve prediction of tropical forest biomass. Trait theory suggests that canopy scale leaf traits may also be correlated with tree architecture (Violle *et al.*, 2007). For instance, plant leaf traits have been related to plot level architecture in the tropics and predicted with leaf spectral data (Doughty *et al.*, 2017). Remotely sensed canopy trait maps using Sentinel-2 for phosphorus, wood density and specific leaf area (SLA) among other traits for broad swaths of tropical forests (Aguirre-Gutiérrez *et al.*, 2021) and such optically derived leaf traits may be correlated with structure at the landscape scale.

Tropical forest structure matters because it is indicative of use: for example, tall canopies were a strong predictor of habitat use by Baldfaced saki monkeys (*Pithecia irrorata*) in the Peruvian Amazon (Palminteri and Peres, 2012) and structure data are increasingly being used in species distribution models (Burns *et al.*, 2020). However, structure is understudied because detailed pan-tropical structural data did not exist prior to GEDI, and yet it is where the bulk of the world's species exist (Stork, 2018) including over 75 % of all vertebrates and 60 % of neotropical mammal species (Kays and Allison, 2001). Stratification has been hypothesized to increase rates of pollination and dispersal, optimize light use, increase inter-canopy CO_2 concentrations, reduce leaf, fruit and flower predation, and increase forest structural integrity (Smith, 1973). Overall, structure also creates the habitat for all other forest dwelling species (Terborgh, 1992). For instance, figure one shows animals both impacting and being impacted by forest structure.

The structure of forests is also a principal factor in determining not just the mean environment experienced by forest-dwelling organisms, but also the diversity, extent, and variability of microenvironments. The extent and diversity of microenvironments directly affects the niches available to organisms, and hence the diversity of forest-dwelling organisms. For instance, Oliveira and Scheffers (2019) proposed an ‘arboreality hypothesis’ where species have increased ranges because they can take advantage of changing microclimates in different canopy layers as temperatures shift due to elevation and latitude. They further suggested that future warming may push arboreal species towards the cooler ground layer (Oliveira and Scheffers, 2019). Another study suggested that climate change may cause arboreal species in hot sparse canopies towards greater ground use (Eppley *et al.*, 2022). Detailed models now exist to predict canopy microclimate with forest structure as a possible input (Maclean and Klinges, 2021). Therefore, forest structure can help determine microhabitats which becomes even more critical as climate change progresses.

Here we use GEDI to understand tropical forest structure and address the following questions:

Q1 – Is the classic paradigm of “old growth” tropical forest architecture with multiple canopy layers correct (visually represented in Fig 1)?

Q2 – What drives the spatial distribution of canopy structure (e.g. total cation exchange capacity, environment e.g. maximum temperature and/or leaf traits)?

Methods

GEDI data – We used the vertical forest structure (L2A and L2B, Version 2) and biomass (L4A – see below) products from the GEDI instrument (Dubayah *et al.*, 2020) based on the ISS between 2019.04.18 and 2021.02.17 for tropical forest regions (Amazonia, Central Africa, and SE Asia). The L2A product has already been ground validated in tropical forests and that is not a goal of this paper (Marselis *et al.*, 2018, 2020; Liu, Cheng and Chen, 2021; Cobb *et al.*, 2023). We principally used the Plant Area Volume Density (PAVD) profile, defined as the Plant Area Index (PAI – which incorporates both leaf and wood) separated into 5-meter vertical bins. We applied a number of data filters to ensure quality such as: degrade flag = 0 (e.g. not in degraded altitude), L2A and L2B quality flags = 1 (simplified metric to only use highest quality data based on energy, sensitivity, amplitude, and real-time surface tracking quality), sensitivity >= 0.95, power beams during night and day and coverage beams during night only (nights are generally better to remove the negative impact of background solar illumination). To ensure that the footprints were in tropical forest regions, we applied three further data quality filters and two further data analysis filters.

Data quality filters:

1. We used the well-established SAR (synthetic aperture radar) dataset TanDEM-X (Krieger *et al.*, 2007) as a comparison to GEDI and removed GEDI data where elevation difference from GEDI is between +/- 100 m because if there is such a large difference, the GEDI data might be wrong.
2. We used the well-established Landsat dataset to ensure forest cover by only including data with tree cover >90% in the year 2010, defined as canopy closure for all vegetation taller than 5m (Hansen *et al.*, 2013).
3. We used the well-established MODIS dataset to further ensure forest cover. The GEDI footprint was classified as Plant Functional Type (PFT) Broadleaf Evergreen Tropical based on MODIS MCD12Q1v006 Product from 2021 (Friedl et al 2019) at 500m spatial resolution following the Land Cover Type 5 Classification scheme. We identified the 25m GEDI footprint within the 500m MODIS pixel for comparison.

Data analysis filters:

1. We screened out areas with tree heights <10m using the relative height metric 98% which was calculated as the height relative to ground elevation under which 98% percentage of waveform energy has been returned. To further ensure quality we vary this number in a sensitivity study 15, 20, and 25m (Fig S1).
2. We compared an index of forest integrity as determined by degree of anthropogenic modification https://www.forestintegrity.com/ (Grantham *et al.*, 2020) to our results (see below – Fig S2).

If the L2a GEDI footprint passed these filters, we then estimated the number of canopy layers (peaks – P). If there were two layers, we estimated the height (H) and depth (D) differences between the two peaks (Fig 1). Ecologically the number of peaks, as well as the height and depths between peaks will influence microclimate, vertical light environment, animal niche space, and biomass. In Figure 1, we show an example GEDI footprint and then classify it using a flow diagram on the right.

1. We first classified each footprint by the number of local maxima (change in first derivative - hereafter: peaks =P) (1-3+) using the Matlab (Mathworks) function "islocalmax" on each PAVD profile. If it had one peak it was classified as 1 peak (blue line Figs 2-3). If it had two peaks, we further classified it (see below). If it had three or more peaks it was classified as 3 peak (orange line Figs 2-3). We did not further classify waveforms with three or more peaks because they were rare (<1%).
2. If the waveform had two distinct peaks, we then classified whether P1 (the peak farther from the ground) had more PAVD than P2 (the peak closer to the ground). By distinct peaks we mean the peaks were more than 10m vertically apart. If the peaks were not distinct (e.g. H≤10m) then the peak was classified as 2p_even (black line Figs 2-3).
3. We then used the following equation to determine if there were a large (>50%) or small (<50%) difference in the depth (D) of the peaks (where ABS is the absolute value):

Equation 1 - $$D = \left(ABS \left(\frac{PAVD\ of\ P1 - PAVD\ of\ P2}{PAVD\ of\ P1} \right) \right) * 100$$

If P2>P1 with D less than 50% difference between the peaks, we classify it as 2p_eq_high (red line Figs 2-3), if D is more than 50% difference it is classified as 2p_eq_low (magenta line Figs 2-3). If P2<P1 with D less than 50% difference, it is classified as 2p_uneq_high (green line Figs 2-3), if D is more than 50% difference 2p_uneq_low (yellow line Figs 2-3).

Overall, there are seven distinct profiles, but we do not show results from three+ peak forests as they were rare (<1%). To calculate the percentage of one peak PAVD profiles (blue line Figs 2 and 3), we sum the number of one peak profiles, divided by all profiles within a 0.1 by 0.1 degrees size gridcell (resolution was chosen for visual clarity). We recognize that our thresholds for H and D are somewhat arbitrary, and therefore, in a sensitivity study we tested these thresholds by changing H from 5 to 15m and D from 40 and 60% but only found a change of ~1% in structural parameters on average (Fig S3). The biggest change resulting in ~2% change in structural parameters occurred by increasing H to 15m.

We downloaded the GEDI L4B above ground biomass density (AGBD) product from DAAC (https://daac.ornl.gov/cgi-bin/dsviewer.pl?ds_id=2017) and averaged it for each 0.1 by 0.1-degree pixel. With this data, we created a histogram of tree heights for each 0.1 by 0.1-degree subregion for all tree heights (RH98) that pass our filters. The peak of the histogram is classified as median rh98 tree height. For each 0.1 by 0.1-degree subregion, we estimate the total plant area index (PAI) as a proxy for commonly used metrics like leaf area index (LAI).

Measuring scale dependence with individual tree data –We recognize that vertical canopy layers may be a function of spatial resolution. To test the dependence of vertical layers on spatial scale, we use a database (Araujo-Murakami *et al.*, 2014; Doughty *et al.*, 2015) where, for a series of plots in six diverse regions of the Amazon basin, we estimate stratification by calculating crown area using measured tree diameter at breast height (DBH) and tree height for individual trees in Caxiuana 4 ha – 2250 trees >10cm DBH, Tambopata 2 ha – 1367 trees>10cm DBH, Iquitos 2 ha 1165 trees>10cm DBH, Tapajos – 18 ha- 1036 trees>25cm DBH, Bolivia 2 ha 974 trees>10 cm DBH, Tanguro 1 ha – 366 trees>10 cm DBH. Plot locations are shown as black dots in Figure 2. For each plot, we used tree height in each 5-meter tree height bin (5-35m) to estimate crown diameter following Asner et al 2002, shown below as Eq 2 where DBH is the diameter at breast height (cm) and crown diameter is in meters.

Equation 2 - Crown diameter (m) $= 9.3*\ln(\text{DBH (cm)}) - 22.2$;

We estimate crown area to ground area ratio for all trees in the plots (e.g. Iquitos 2 ha = 1165 trees>10 cm DBH) and on a subset of groups of 50 trees to better approximate the 25 m size of a GEDI footprint, as this is an approximate average number of trees >10cm DBH per 25m diameter circle in the tropics. For instance, a typical one hectare tropical forest plot would contain between 500-1000 trees with DBH>10cm (Malhi *et al.*, 2021) (~20 GEDI footprints if evenly spaced – which would not happen in practice) and each footprint, therefore, might contain 25-50 trees (with DBH>10cm). We then use the same "peak" procedure (Eq 1 - described above) to estimate % one peak as a percentage for each region. We estimate crown area to ground area for each 5-meter bin and vertically area summed. We also show median and maximum tree height for the plots. To test the values in equation 2 influence our results, we varied the slope in Eq 2 (9.3) by ±5% and show how this impacts the results in Figure 4. To test the dependence of structure on spatial resolution, we estimate % one peak for spatial resolutions of 10m (Fig S4), 25m and 1 ha (Fig 4).

Comparison data layers – We compared % one peak to several other climate, soils, leaf traits, and ecoregion maps listed below for the Amazon basin. We currently focus on the drivers of structure and validating GEDI for the Amazon region in this paper, but follow on papers may do a similar analysis for Africa and SE Asia. Each dataset had its own resolution, which we standardized to 0.1 by 0.1 degrees.

Soils – We used data from soilgrids https://www.soilgrids.org/ (Batjes, Ribeiro and van Oostrum, 2020). We focused on total cation exchange capacity at pH 7 from 0-5cm in units of mmol(c)/kg as previous studies had suggested this to be an important variable to explain floristic composition (Figueiredo *et al.*, 2018).

Climate – We averaged TerraClimate (Abatzoglou *et al.*, 2018) https://www.climatologylab.org/terraclimate.html data between 2000 and 2018 for Climatic water deficit (CWD) (the difference between monthly reference evapotranspiration calculated using the Penman Monteith approach and actual evapotranspiration), Vapor Pressure Deficit (VPD in kPa), Mean Monthly Precipitation (mm/month), potential evapotranspiration (PET) and maximum and minimum temperature (°C). These data were originally based on CRU Ts4.0 data and modified by Abatzoglou et al 2018.

Leaf traits - For plots in the GEM network (listed in Table 1) (Malhi *et al.*, 2021), we found the PAVD profile for the footprint closest to the plot as well as all footprints within a 0.03° grid around the plot coordinates. Most of these plots had in situ leaf traits measured to account for 70-80% of the basal area (of trees >10cm DBH) of 1 ha plots. Based on the above described field campaigns, (Aguirre-Gutiérrez *et al.*, 2021) used Sentinel-2 to create remotely sensed canopy trait maps for P=phosphorus %, WD = wood density $g.cm^{-3}$, and SLA=specific leaf area $m^2\ g^{-1}$.

We then compared the GEDI profile (% one peak) to the trait value predicted by those maps to that footprint.

Ecoregions - Ecoregions reflect the distributions of a broad range of fauna and flora across the entire planet and we use them as a proxy for plant biogeography https://www.sciencebase.gov/catalog/item/508fece8e4b0a1b43c29ca22 - (Olson *et al.*, 2001).

Statistical analysis –We used the matlab function "fitlm" to fit linear models and "fitnlm" for the non-linear models to compare variables such as soils data, environmental data, or leaf trait data (at 0.1 degree resolution) to GEDI structure data of what percent of all footprints in a 0.1 degree area have one peak. The P values listed are for the *t*-statistic of the two-sided hypothesis test.

Results

Most individual GEDI footprints in tropical forests do not have multiple layers (as in Fig 1) and instead have a single peak in vegetation density at ~15m, but this ranged geographically (regionally and between continents) between 60 to 90% (Figs 2-3). Within the Amazon basin (Figure 1), the broad geographic patterns were a large central region with low stratification, surrounded by another broad region with greater stratification bordered to the west by the Pebas formation (Higgins et al 2011), to the east by the Tapajos River, and the South at ~12°S. Another region of lower stratification occurred towards the southeast in the "arc of deforestation" and savanna transition zones. River floodplains also tended towards increased stratification. The Congo basin showed a broadly similar spatial orientation with a central area with lower stratification surrounded by regions with greater stratification. The floodplains again were areas with greater stratification. Southeast Asia, composed of mainly islands, showed greater stratification towards the island center. The island of New Guinea had increasing stratification moving northward.

A low PAI peak (e.g. ~15m) may also indicate forest disturbance due to selective logging or other human impact. For instance, there was selective logging in parts of Borneo (Riutta *et al.*, 2018) and this impacted structure by increasing the dominance of shorter pioneer one-peak forests (i.e. Bornean logged plots are 78 % one peak versus 44% for old growth forests). However, the filters we used (tree height, MODIS PFT, logging product) should remove most human impact (although there may be older legacy effects we cannot account for). We tested this by increasing the minimum tree height (between 15, 20 and 25m) and did not see a big impact on the broader results, although there were minor changes at the 25m threshold (Fig S1). We also show comparisons of percentage of one peak to a forest disturbance product (Grantham *et al.*, 2020), which showed large regions dominated by one peak forests in areas of minimal human disturbance (Fig S2).

On a subset of the Amazon (5 by 5° black box regions chosen to represent the broader region in Fig 2 and 3), we averaged the vertical profile for each footprint in each of six structural categories (see methods) and found "one peak" forests peaked in PAVD at 15m with a fairly linear decline going upwards until ~40m (Figure 2 blue line = mean -sd). The next most common profile type was 2p_eq_high (Figure 2 red line = mean -sd). It was a "2 peak" forest (at ~5% of the results), with an initial peak in PAI at 15m and a second lesser peak in PAI at 30m and a local minimum at 20m. Average forest height of this forest type exceeded the one peak forests with a maximum height at ~45m versus 40m. This forest type ranged between 1 to 10% of forest pixels and was more abundant in the Southeast and Northwest of the Amazon (Fig 6 – similar figure for Central Africa is Fig S5 and SE Asia Fig S6). The third most common forest structure (represented by the black line (2p_even) at 3.2%) had two close peaks at 15 and 25m, with a small nadir at 20m. This forest type had a PAI peak at 25m and was followed by a steep drop at ~40m. This forest type ranged between 1 and 5% across the Amazon and was widely dispersed throughout the Basin. The next most common 2-peak structure (magenta (2p_eq_low) in Fig 2) at ~3.2% of forest types with a peak at 15 m followed by a much weaker peak with less than 50% of the PAVD at 30m. This had a similar distribution to the "red (2p_eq_high)" line, but with an additional hotspot in the Southeast that was not present in the "red (2p_eq_high)" (Fig 6). The remaining forest types had greater PAVD in the upper canopy with peaks at ~30m.

To test the dependence of vertical layers on spatial scale, for six locations (shown as black dots in Fig 2), we used DBH, tree height, and a canopy diameter model (Asner et al 2002) to estimate that total vertically summed crown area/ground area averaged 1.8 $m^2.m^{-2}$ (0.96-2.3 min max). Averaging at the 1 ha scale for all trees >10cm DBH (>25cm for the Tapajos) in the plots (size ranging between 2 ha to 18 ha) showed a single peak that averaged 20m (between 17.5-22.5m) in crown area/ ground area (thick lines in Figure 4). This 20m height may be taller than the GEDI mean of 15m due to the absence of smaller 0-10cm DBH trees measured at the plots. We then subsampled 50 trees from each plot (a better approximation for the GEDI footprint size) and more stratification resulted. For these subsets, we calculated one peak/all data and found a low in Tambopata of 56% one peak to a high of 95% one peak in the Tapajos with the other sites ranging between 73 to 77% one peak, which is a good approximation of percentage one-peak across the Amazon basin (~79%) (Figure 2). If we average across the 6 sites at a spatial resolution of 25m (like GEDI) we find 75% one peak, but if we reduce the resolution to 10m, then % one peak drops to 65% (Fig S4), so the spatial resolution of the footprint clearly matters for our question. The Tapajos results must be viewed with caution because only large trees (>25cm DBH) were recorded, which led to a very high percentage one peak. According to Fig 2, Tambopata and the Tapajos are near regions divided between areas of high and low structure and most other plots are in areas of less structure (Figure 2). Therefore, spatial scale matters since averaging over a wider spatial area will hide individual forest structure.

How representative is the structure in plot networks compared to the broader Amazon? To answer this, we compare GEDI footprints (closest footprint and all footprints averaged within 0.03° radius of the plots) to a well-studied plot network (GEM - (Malhi *et al.*, 2021) in tables 1 and 2) and found the GEDI footprint nearest to the plots showed a gradient from the Western Amazon (90% one peak), Eastern Amazon (85%), Gabon (80%), to Borneo (50%). Averaging all nearby footprints showed similar (except for Gabon), but generally lower trends: Western Amazon (84%), Eastern Amazon (79%), Gabon (54%), and Borneo (61%). In Table one, we show data for each individual plot along with remotely sensed trait data (Aguirre-Gutiérrez *et al.*, 2021) calibrated from in situ measurements at the plot network, and we found a significant relationship between structure and SLA (r^2=0.12, P<0.05, %one peak=-68*SLA+1.4) but not with wood density and %P. However, this is a global analysis, and the signal is dependent on low SLA values along an elevation gradient where GEDI is less accurate because of difficulty in discerning the ground layer. In Borneo, the GEM plot network (Riutta *et al.*, 2018) is along a logging gradient with a clear change in structure (78 % one peak for logged plots versus 44% one peak for old growth forests). We found a significant increase in SLA (P<0.05) with disturbance and a close to significant increase in %P with disturbance (P=0.06).

We compared the average PAVD profiles from the entire Amazon to the average PAVD profiles for the entire SE Asia and Africa (average continental scale 0.1 by 0.1 degree pixels and not just the black boxes in figs 2 and 3). On average, the Amazon had greater percent of one peak forests (79±9sd) than either SE Asia or Africa (72±14 v 73±11). Median tree height (rh98) was lower in the Amazon at 25.6m than in Africa at 28.5m or SE Asia at 28.7m. In the black box regions shown in Fig 3 for Africa and SE Asia, one peak forests were most abundant (~70%) with a similar peak at 15m (Figure 5). In both the Africa and SE Asia subplots, both red (2p_eq_high) and magenta (2p_eq_low) structure types were much more common forest structures than in the Amazon, accounting for >20% of forest types vs <10% in the Amazon. The average curves changed shape with Amazon having more PAVD in the mid-canopy ~20m and Africa and SE Asia having more PAVD in the upper canopy ~30m. The less represented green and yellow structures increased by an absolute 3-4% over the Amazon and had much more PAVD (~0.05 increased PAVD) in the upper canopy (at ~30m height). River basins throughout the tropics had similar structural properties.

To explain the spatial patterns in the distributions of % one peak forests, we compared maps of percent one peak to a variety of datasets such as tree height (rh98), ecoregions, GEDI L4A AGBD, plant area index, number of footprints, climate (CWD, VPD, MMP, T_{min}), and total soils cation exchange capacity (Figure 7 – similar figure for Africa is Fig S7 and SE Asia Fig S8). The strongest correlations were with tree height and AGBD, with biomass a slightly better predictor for one peak forests (0.12 vs 0.14 r^2 respectively) (Figure 8). The AGBD L4A product is driven by tree height, so the similar strength of the correlations is not surprising, but there is a question of whether structure or tree height is a better predictor of biomass, which we discuss later. We compared meteorological data for VPD, PPT, CWD, PET, T_{max} and T_{min} to percent one peak and all were highly significant ($P<0.001$) but explained relatively little variance in the data. T_{max} explained the most at 5% of the variance, followed by VPD at 2.5% and the others explaining ~0.01 of the variance. Likewise, total cation exchange capacity was highly significant but again explained only about 1% of the variation (Figure 8). Other variables such as number of footprints was not related ($r^2<0.01$), but PAI explained ~4% of variance, which is again, likely related to tree height. We then combined all climate and soil variables which explained ~9% of variance and the key parameters were T_{max}, VPD followed by total cation exchange capacity.

Ecoregions, which may be a good proxy for floristics, delineated structure well for particular ecoregions. For instance, ecoregion 68 (Figure 7) had boundaries similar to boundaries of our structure dataset with a lower average value of percent one peak (75% vs 80%) than surrounding ecoregions. Another ecoregion with the boundary of the Pebas formation also delineated the structure data quite well. There were some regions that were partially delineated well but not entirely. For instance, even ecoregion 68 (Figure 7) had a sharp boundary in structure in the south not accounted for in the ecoregion.

Discussion

There are large (>10%) differences in forest structure within the Amazon basin (60-90% one peak) and between the Amazon (79±9sd) than SE Asia and Africa (72±14 v 73±11 respectively). We are confident that the spatial patterns of structural changes are not mainly due to modern human influence, because we carefully screened for human influence using several independent remote sensed products (MODIS PFT (Friedl et al 2019), a Landsat based deforestation product (Hansen *et al.*, 2013) and GEDI tree height itself from GEDI (Dubayah *et al.*, 2020). Plot data from undisturbed regions (Doughty *et al.*, 2015) (DBH and tree height) showed similar structural trends in old growth plots (Figure 4). Human influence, as measured through forest integrity (Grantham *et al.*, 2020), also did not explain our geographic patterns of structure (Fig S2). The finding that the majority of GEDI footprints had a single PAI peak at ~15m was initially surprising. However, several tropical aircraft lidar campaigns showed similar shape for the lowland tropics (a single peak when averaged over ~1 ha) but a slightly higher peak in PAI at ~20m (Asner and Mascaro, 2014; Asner *et al.*, 2014). We hypothesize that the difference in the height of peak PAI may be due the difference in "energy return" profiles or how to correct for the reduced energy reaching the understory and the difficulty of laser pulses in the lower canopy returning due to an abundance of plant material. Full waveform information from GEDI can help correct for this energy return. In addition, prior work comparing TLS, LVIS and simulated GEDI data has found high-quality GEDI profiles on average to be accurate (Marselis *et al.*, 2018, 2020). Finally, we are confident that the bulk of structural differences across the tropics are of natural origin because on top of the filters applied, some regions of the Amazon very far from human influence still had the dominance of one peak forests, such as the broad region north of Manaus in the Amazon (although there may be ancient legacy effects that we do not account for) (Fig S2).

The classic paradigm of “old growth” tropical forest architecture (visually represented in Fig 1 and figures in Halle et al 1980) is a generally closed upper canopy with large emergent trees at ~30-35m where PAI peaks followed by a second peak at 15m with slightly lower PAI. These PAI peaks at ~15 and 30m are occupied by “trees of the present” taking advantage of increased light cells (top of canopy and a second area of increased light at ~15m where lightflecks converge) (Halle, Oldeman and Tomlinson, 1980). This “classic paradigm” implies a stratified canopy that might be best represented by the green (2p_uneq_low) or yellow (2p_uneq_high) lines in Figs 2-5, but we find that this forest structure is relatively uncommon across the tropics making up just 3-6% of tropical forest area. In contrast, by far the most common PAVD profile across the tropics has a single peak in PAI density at 15m and this forest type likely reflects the absence of a closed upper canopy. In our color scheme (Figs 2-3), we can think of a gradually increasing proportion of vegetation percent in the upper canopy going from the highest PAI at the top with yellow (2p_uneq_high) (0.5% of total footprints), green (2p_uneq_low) (2%), red (2p_eq_high) (5%), magenta (2p_eq_low) (3%), and the lowest at blue (1 peak)(86%). Overall, these results show that a “stratified” forest with higher upper canopy closure is relatively rare across tropical forests.

Our structure maps broadly matched results from plot-based methods (Fig 4). We also found strong correlations between our structure maps and detailed maps of structure, floristics, climate and soils for a broad region of Central Africa from Fayolle et al 2014 where old growth *celtis* forest is associated with regions with more vertical layers (~60% 1 peak) while more degraded or young *celtis* forests with more pioneer species is associated with less structure (70% one peak) (Fayolle *et al.*, 2014). A floristic map for all of central Africa also showed correlations with our structure map (Réjou-Méchain *et al.*, 2021) with, for instance, north (more structure) to south (less structure) gradients in Central Africa (Figure 3) that match a transition in their figures from PCA 1, where floristics was controlled by a transition between cool, light-deficient forests and forests with high evapotranspiration rates, to PCA 2, where floristics were controlled more by seasonality and maximum temperature. In S.E. Asia, we compared our structure results to a logging gradient (Riutta *et al.*, 2018) with known structural changes and found GEDI footprints near Danum valley, where the tallest trees were found, also had some of the highest stratification (44% one peak) versus logged (78 % one peak) which gives further confidence in the results. Broadly, old growth forests in SE Asia have the highest levels of stratification and this may be partially due to the presence of Dipterocarps which are the tallest tropical trees (Shenkin *et al.*, 2019; Jackson *et al.*, 2021).

Most of our independent datasets of soils or climate (as well as our combined model) did not strongly capture the spatial patterns of forest structure in the Amazon basin (Figure 7). Tree height and AGBD did match these patterns (Figure 8), but those variables cannot be considered independent of structure. However, patterns shown in Figure 4c in Figueiredo et al 2018 are similar to the one we highlight in this study (Fig 1) (Figueiredo *et al.*, 2018). Figueiredo et al (2018) created species distribution models for 40 species across the Amazon basin using 19 bioclimatic variables, 19 soil variables, and four remote sensing variables (including GLAS derived canopy height (Simard *et al.*, 2011)). Overall, for most species, a combination of soils and climate variables explain most variance (similar to (Tuomisto *et al.*, 2019)) but single-variable models did poorly with an average of less than 8% of the variance explained. This broadly reflects our attempts to model structure with single variables. There was a tight correlation between regions with less structure (e.g. higher percentage of one peak) and areas where soils are the limiting factor to species occurrence, and regions with greater structure (i.e. lower percentage of 1peak) to areas where climate is the limiting factor to species occurrence. Perhaps deeper, more fertile soils allow for taller (either species or trees reaching their genetic height) and higher canopy closure forest types. Canopy height from the GLAS was the second most important variable for explaining species distributions, so it is possible that the Figueiredo et al (2018) map shows similar patterns to Fig 2 due to the inclusion of the height metric (a strong predictor of structure). A global study of forest structure based on upscaling terrestrial lidar with WorldClim2 datasets showed some correlations with our structure maps but also missed many of the regional changes (Ehbrecht *et al.*, 2021).

Ecoregions delineated boundaries in structural composition in a few key areas of the Amazon basin like the Pebas formation (Higgins *et al.*, 2011) and the Tapajos region in Para, Brazil (Figure 7). Higgins et al (2011) found a strong east-west gradient with an almost complete floristic turnover and an order of magnitude change in soil cation exchange capacity associated with the presence of the Pebas formation (Higgins *et al.*, 2011). This line marking the boundary of the Pebas formation also seems to strongly delineate forest structure with one peak forest more abundant east of this line with lower cation exchange capacity and two peak forests more abundant to the west with higher cation exchange capacity. There is a further boundary delineated by the very wide (12-16 km) Tapajos River with forests to the west having a higher percentage one peak vs the eastern forests. Interestingly, some ecoregions (like 68) matched well with boundaries of vegetation structure, except for a few key areas (like in the south of region 68 – fig 7). This may indicate that forest structure could be used in the future to improve upon current ecoregion boundaries.

What causes the dominance of one peak forests in the tropics and the spatial changes in these patterns? A forest with a fully closed emergent canopy layer would have canopy layers, but most forests likely lack a fully closed upper layer, leading to the dominance of the one peak forests. Rephrasing the initial question, we can instead ask: Is the rarity of a closed upper layer canopy (or relative rareness of large emergent trees) due to the environment (soils or climate) or floristics (species composition)? In practice it is difficult to disentangle the floristic and environmental and there is a large literature describing how the environment (soils or climate) impacts the species composition. For instance, Amazonian species composition may follow a south-west/north-east soil fertility gradient and a north-west/south-east precipitation gradient (ter Steege *et al.*, 2006). Soil cation concentrations is the primary driver of floristic variation for trees (Tuomisto *et al.*, 2019) with climate being of secondary importance at regional scales. Environment alone could also directly impact tree height and structure, with hydraulic limitations or nutrient deficiencies causing trees to not being able to achieve their genetic height. Soil depth can impact structure as shallow soils can cause stunted root growth leading to a thinner upper canopy structure (Halle, Oldeman and Tomlinson, 1980).

What may explain the continental scale differences in structure between the Amazon and other tropical regions? Previous authors have noted large continental scale differences in AGBD and tree height (Borneo>Central Africa>Amazon) that broadly match the trends we show in structure (Feldpausch *et al.*, 2011; Lewis *et al.*, 2013). For instance, the Congo basin had average AGB values of 429 Mg ha^{-1}, similar to Bornean forests (445 Mg ha^{-1}), and much higher than the Amazon (289 Mg ha^{-1}) (Lewis *et al.*, 2013). We show similar broad trends with the Amazon at 79±9sd % one peak and 25.6m height, SE Asia 72±14 and 28.7m height and Central Africa 73±11 and 28.5m. Lewis et al 2013 had hypothesized that AGBD differences between Amazon and Africa were due to different biomass residence times, the differences between Africa –Borneo differences were possibly due to NPP differences. However, tree height and biomass are structural attributes and do not explain the difference in continental structure.

To fully understand structural gradients across the Amazon, higher resolution aircraft lidar can be used. Asner et al 2014 flew aircraft lidar along an elevation and nutrient gradient in Peru and found that canopy height and shape (height of peak canopy volume divided by canopy height) had a high, negative correlation with gap density (Asner *et al.*, 2014). Perturbation, either up an elevation gradient or from high soil fertility to low, led to shorter forests with more gaps and a peak canopy volume at a lower height in the canopy. These changes are broadly correlated with our maps of percentage of one peak, with perturbation (up elevation gradients or fertility gradients) increasing percentage of one peak forests. We found canopy stratification decreased as T_{max} increased and soil fertility decreased (Fig 8). Therefore, our results support this paradigm that a movement away from ideal conditions may result in less structural complexity. Climate change will increase T_{max}, but it is unclear whether this would further reduce structural complexity of tropical forests in the future.

In addition to tree height, remotely sensed leaf traits were also related to structure near some of our plots. Increased stratification (lower percentage of one peak) was significantly correlated ($P<0.05$) with increases in SLA, but this was almost entirely driven by low SLA values in high elevation plots and removing these plots removed the significant correlation (Malhi *et al.*, 2021). Along a logging gradient in Borneo (Riutta *et al.*, 2018), less stratification as logging increased was significantly correlated with an increase in SLA and foliar concentrations of phosphorus, similar to other studies (Baraloto *et al.*, 2012) (Carreño-Rocabado *et al.*, 2016). However, Both et al 2019, a nearby field study, found a contrary result when comparing SLA along the forest gradient (Both *et al.*, 2019). Furthermore, Swinfield et al 2019 used high resolution aircraft hyperspectral data to predict SLA across the Bornean landscape (Swinfield *et al.*, 2019), but unlike most early studies (Doughty *et al.*, 2017) did not predict SLA accurately. Overall, we have reasons for caution for how well SLA can predict structure in tropical forests, but our abilities may improve in the future with hyperspectral satellites which could more accurately predict leaf traits at a global scale.

The primary goal of GEDI is to improve global predictions of biomass and incorporating structure could aid this goal. GEDI L4B was correlated ($r^2 = 0.12$ and 0.14) with both tree height (rh 98) and structure (% one peak). The GEDI algorithm uses tree height (rh 98) as a metric to predict biomass, and since tree height is correlated with structure, the similar strength of the correlations is not surprising (Duncanson *et al.*, 2022). However, there is a question of whether structure in addition to tree height can be used to improve biomass predictions. The dominance of one peak forests likely indicates more open upper canopy forests and Asner and Mascaro (2014) have shown these forest types make biomass prediction more challenging (Asner and Mascaro, 2014). The plot data used to calibrate GEDI for tropical regions were not widely distributed throughout Amazonia, especially in the regions where height and structure diverge (Fig 2). Understanding why height and structure diverge in these regions may be key towards understanding whether structure can improve biomass predictions in the future.

Overall, in the majority of tropical forest area, the upper canopy may be more open and tropical forest stratification is simpler than previously expected and this has important implications for predicting biomass. Furthermore, our results indicate that tropical forest canopies may be more open than previously thought which may expose animals to greater climate change related heat stress and require modifications to their behavior (Oliveira and Scheffers, 2019; Eppley *et al.*, 2022).

Table 1 – Structure and trait data for regions surrounding plots from the GEM network (Malhi *et al.*, 2021). The columns are global region, RAINFOR plot code, plot structure classification for the footprint closest to the plot coordinates and the height of this footprint (highest vertical bin). Next is the average % one peak for footprints within 0.03° of the coordinates surrounding the plot and the average height of area. The last three columns are regionally averaged remotely sensed trait data (P=phosphorus=%, WD = wood density g cm^{-3}, and SLA=specific leaf area - m^2 g^{-1})(Aguirre-Gutiérrez *et al.*, 2021).

Region	Rainfor code	Plot classification	height	% 1 peak near plot	Ave height	P	WD	SLA
SE Asia	DAN-04	magenta	80	21	61	10	0.61	0.01
SE Asia	DAN-05	blue	35	22	60	10	0.61	0.01
SE Asia	LAM-01	magenta	50	56	45	9	0.6	0.0105
SE Asia	LAM-02	magenta	50	44	51	10	0.59	0.0104
SE Asia	MLA-01	magenta	55	78	40	NaN	NaN	NaN
SE Asia	SAF-01	blue	45	88	43	10	0.58	0.0103
SE Asia	SAF-02	blue	40	71	44	10	0.59	0.0101
SE Asia	SAF-03	blue	40	80	44	10	0.58	0.0105
SE Asia	SAF-04	3-peak	95	53	62	10	0.6	0.0106
SE Asia	SAF-05	Blue	35	100	38	10	0.58	0.0102
W. Amazon	ALP11	yellow	45	82	41	10	0.61	0.01
W. Amazon	ALP30	blue	40	80	41	10	0.6	0.01
W. Amazon	SPD02	blue	45	78	47	10	0.6	0.009
W. Amazon	SPD01	blue	60	80	46	10	0.6	0.0091
W. Amazon	TRU08	blue	40	81	47	10	0.6	0.0089
W. Amazon	TRU07	blue	50	79	49	10	0.6	0.0089
W. Amazon	ESP01	blue	40	88	38	12	0.62	0.0075
W. Amazon	WAY01	blue	45	87	43	12	0.62	0.0074
W. Amazon	TRU03	blue	50	98	38	11	0.62	0.0076
W. Amazon	ACJ01	blue	30	89	39	12	0.62	0.0078
E. Amazon	CAX-03	blue	40	82	38	9	0.61	0.0102
E. Amazon	CAX-06	black	35	0	35	NaN	NaN	NaN
E. Amazon	STB-08	blue	45	69	45	9	0.61	0.0104
E. Amazon	STD-05	blue	40	81	35	8	0.65	0.0108
E. Amazon	STD-10	blue	40	94	38	9	0.62	0.0101
E. Amazon	STD-11	blue	30	85	39	8	0.61	0.0102
E. Amazon	STN-02	yellow	40	43	42	9	0.64	0.0104
E. Amazon	STN-04	blue	25	90	34	9	0.64	0.0103
E. Amazon	STN-06	blue	35	80	36	9	0.64	0.0102
E. Amazon	STN-09	blue	40	95	33	9	0.63	0.01
E. Amazon	STO-03	blue	45	70	44	8	0.66	0.0106
E. Amazon	STO-06	blue	35	89	44	8	0.65	0.0106
E. Amazon	STO-07	blue	40	73	44	8	0.66	0.0108

Gabon	IVI-01	blue	40	60	44	9	0.64	0.011
Gabon	IVI-02	blue	35	57	46	9	0.65	0.0109
Gabon	LPG-01	black	45	57	44	NaN	NaN	NaN
Gabon	LPG-02	blue	50	33	56	NaN	NaN	NaN
Gabon	MNG-04	blue	25	63	42	NaN	NaN	NaN

Table 2 – Percent one peak forest of all GEDI footprints closest to the GEM plots and within a 0.03° radius around the plot coordinates. Same as results from Table 1, but averaged (± sd) by continental region.

	W. Amazon	**E. Amazon**	**Gabon**	**SE Asia**
Nearest to plot	90%	85%	80%	50%
within a 0.03° radius	84 ± 28%	73 ± 26%	54 ± 12%	61 ± 27%

Figures

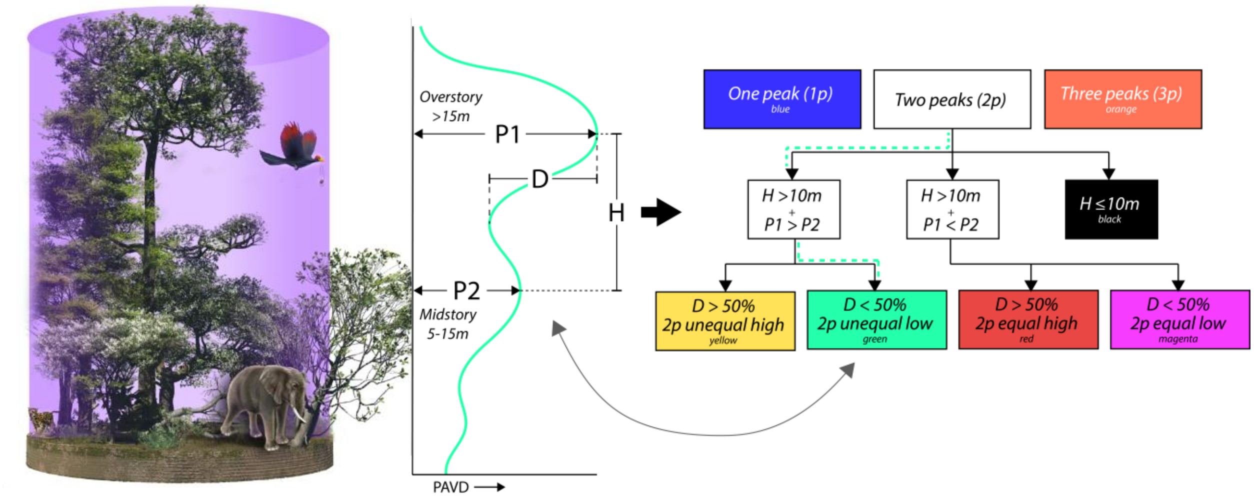

Figure 1 – Artistic rendition of a "typical" stratified tropical forest with the forest (left) within a 25m diameter GEDI pulse and the expected layered return of the profile (center). Animals in figure show how animals both impact and are impacted by canopy structure. (right) Flow chart diagram showing our procedure for delineating the profiles. Green dashed line shows how the example profile would be classified.

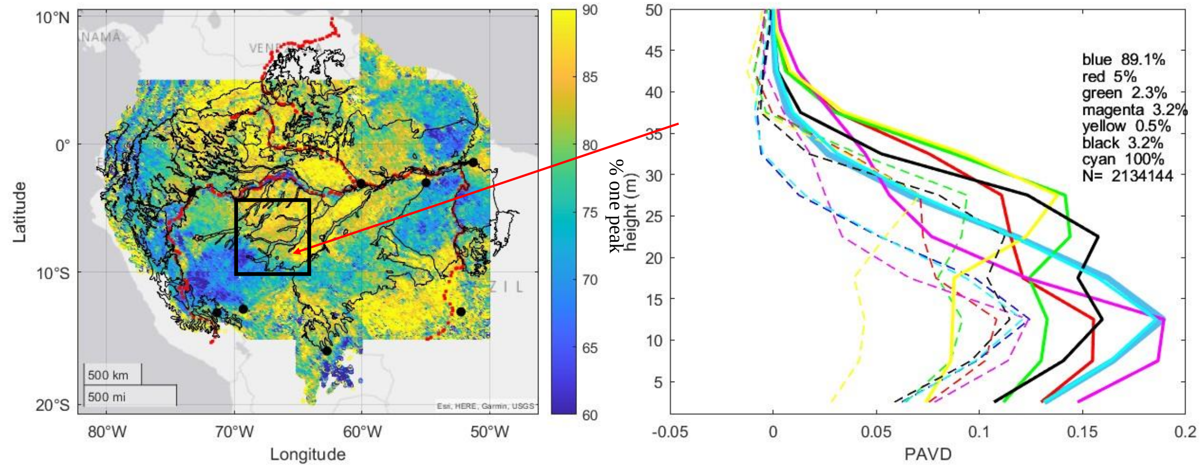

Figure 2 – (left) Each pixel represents the number of one peak footprints (as represented by the blue line on the left) divided by total number of GEDI footprints in a 0.1 by 0.1 degree region for Amazonia. Black lines are ecoregions for the Amazon region. Red lines are rivers and black dots are field plots used in Figure 4. (right) Average (solid) - sd (dashed) waveforms for the region in the black box. We give the total number of individual footprints analyzed and the percentage for each type. PAVD is plant area volume density. Cyan is the average waveform for all data (100%) in the black box.

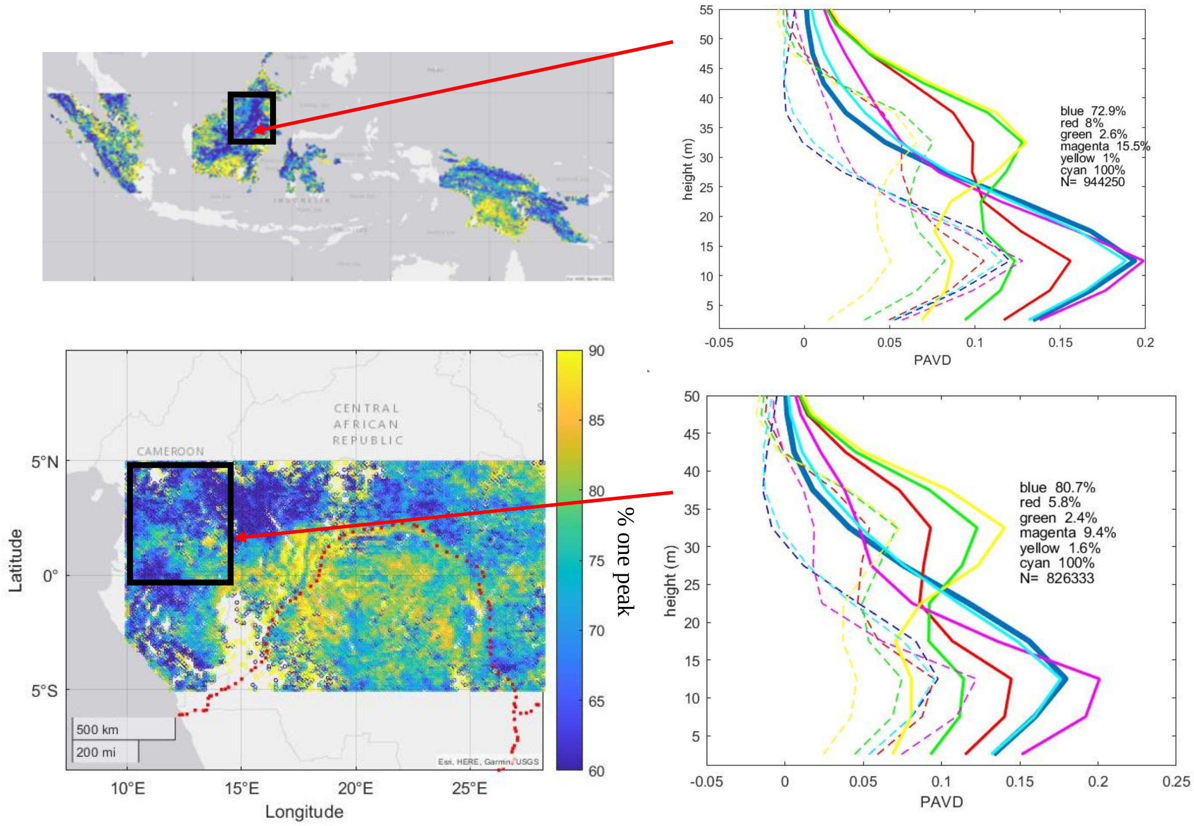

Figure 3 – (left) Each pixel represents the number of one peak footprints divided by total GEDI footprints in a 0.1 by 0.1 degree region for SE Asia (top) and Central Africa (bottom). Red lines are major rivers. (right) Average (solid) - sd (dashed) vertical footprints for the region in the black box. For each type we give the percentage and the total number of individual footprints analyzed. Averages representing <1% were removed. PAVD is plant area volume density. Cyan is the average waveform for all data (100%) in the black box.

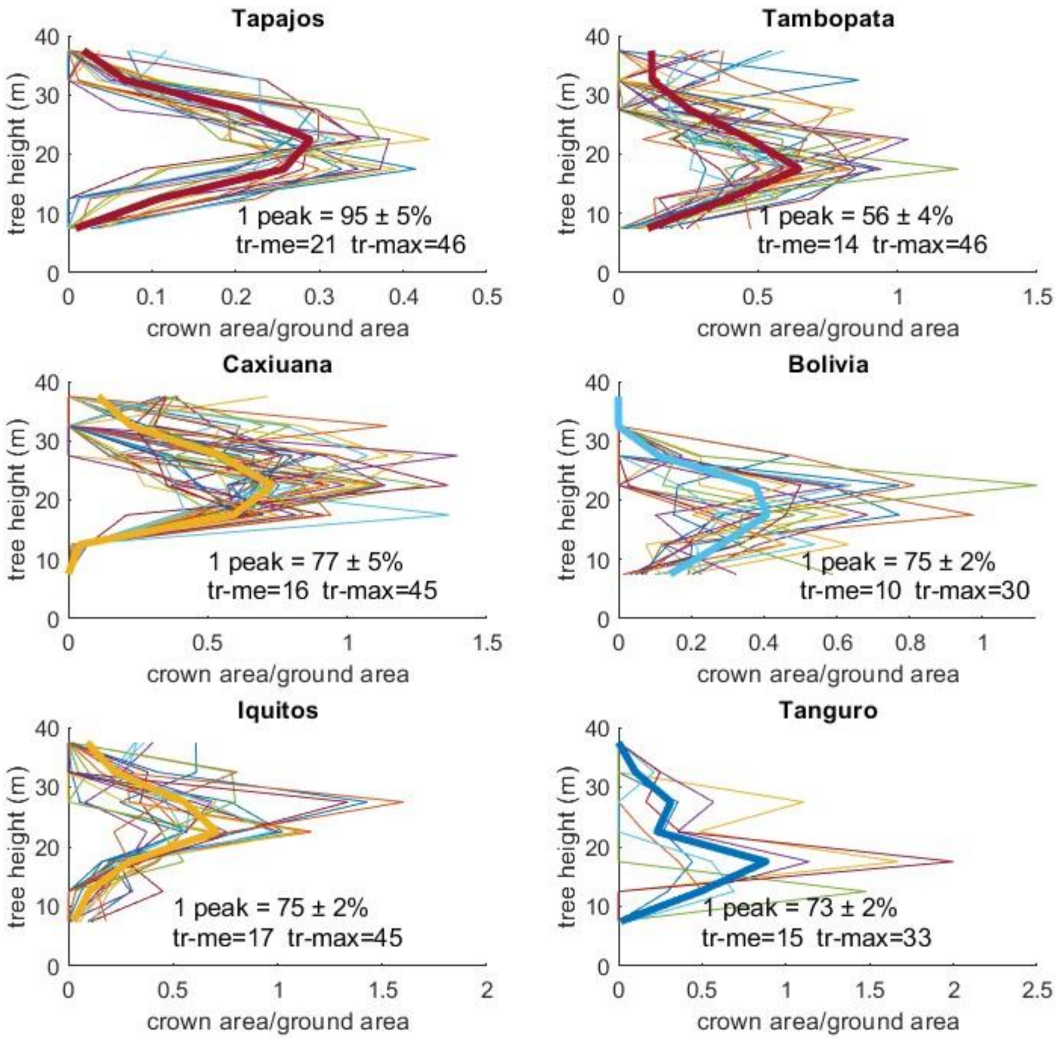

Figure 4 – Tree height versus crown area/ground area as estimated with plot level tree DBH and tree height for six regions as shown in Figure one (Tapajos – 18ha, Caxiuana 4 ha, Tambopata 2 ha, Iquitos 2 ha, Bolivia 2 ha, Tanguro 1 ha). Thin lines are groups of 50 trees and the bold line is the plot average. For each 5-meter tree height bin we estimate crown diameter following Asner et al 2002. We then use the same "peak" procedure as with GEDI data to estimate one vs two peak forests and show this as a percentage. The confidence intervals show results modifying the slope of the equation from Asner et al 2002 by 5%. We also show median (tr-me) and maximum tree height (tr-max) for the plots. Results from the Tapajos are for trees >25cm DBH only.

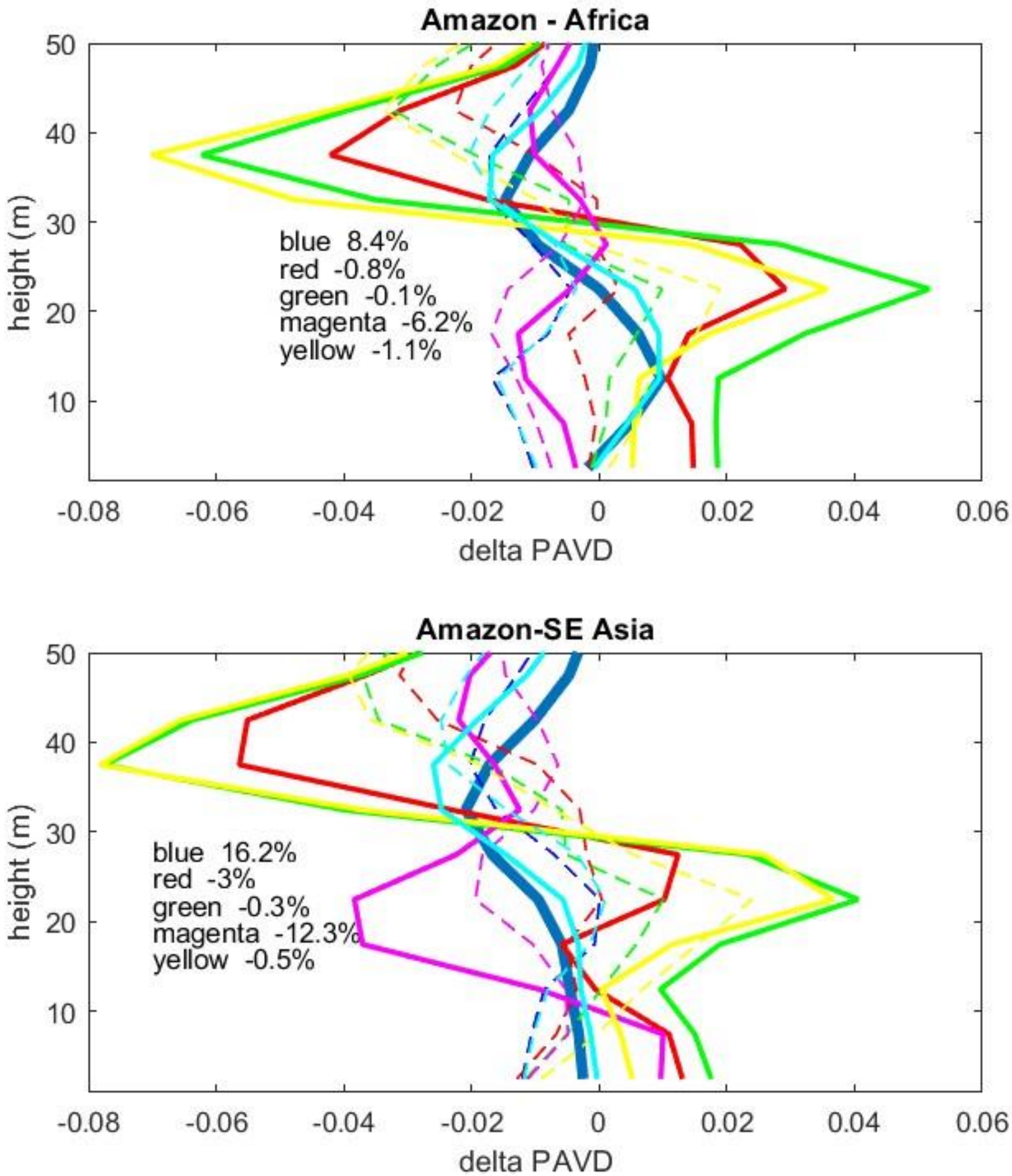

Figure 5 – The change in the average (solid) and sd (dashed) forest structure between the Amazon and Africa (top) and the Amazon and SE Asia (bottom) for the regions highlighted in black in Fig 2-3. The numbers are the listed differences in the percentage abundance. Cyan is not listed as it represents 100%.

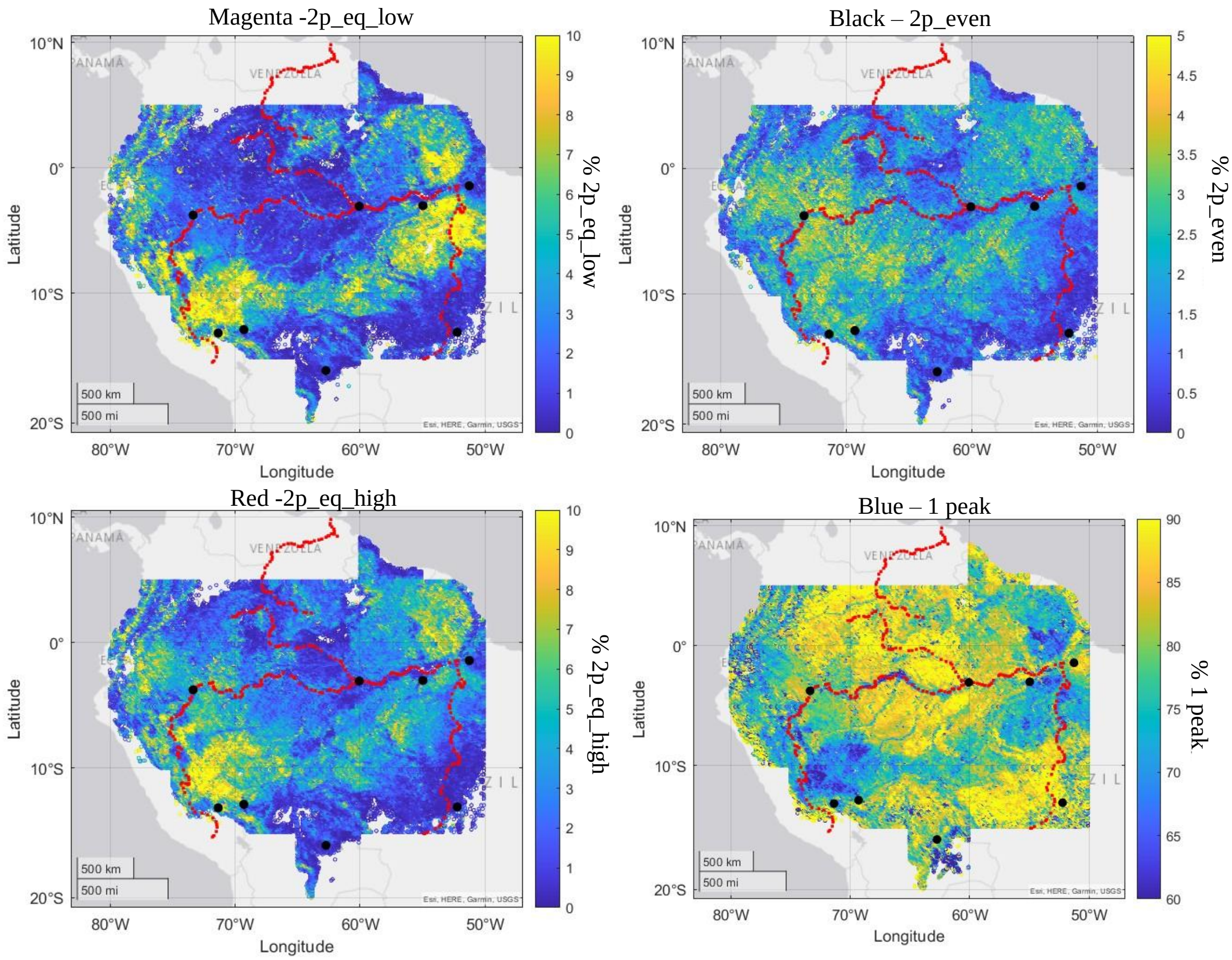

Figure 6 – Spatial distributions for the Amazon basin for different types of the "2 peak" forests. The color labels are associated with the colors of the lines in Figs 2-3. The colorbar scales are different between panels.

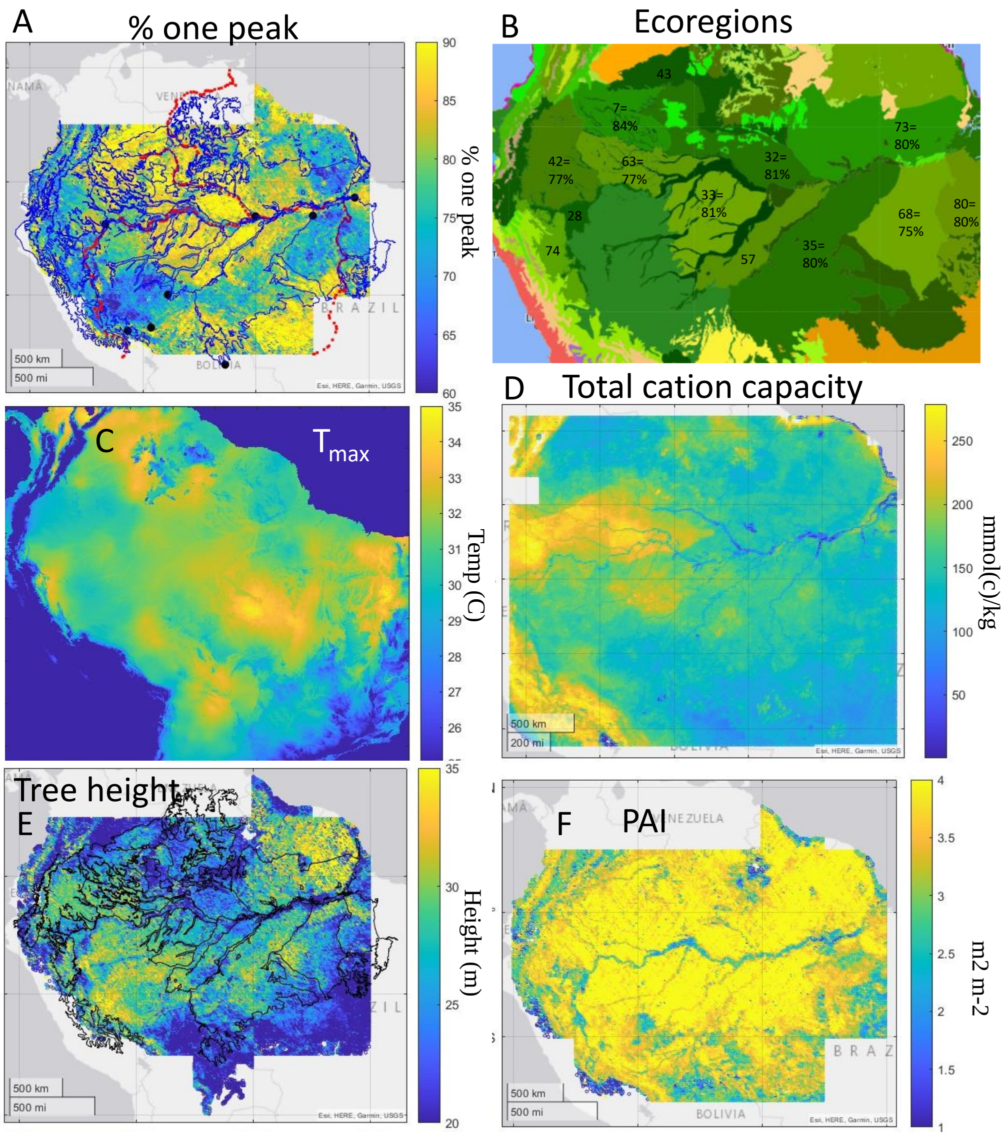

Figure 7 – Different data layers that were used for comparison with the percent one peak dataset. (A) Spatial distribution of the percentage of one peak forests (same as figure 1) with the ecoregions of the Amazon basin overlaid (Olsen et al 2001). (B) A map of the ecoregions alone shown above for clarity with percent one peak for each ecoregion. (C) Max temperature - T_{max} (°C), (D) total cation exchange capacity (mmol(c)/kg, (E) median tree height from rh98 GEDI with ecoregions, and (F) plant area index from GEDI.

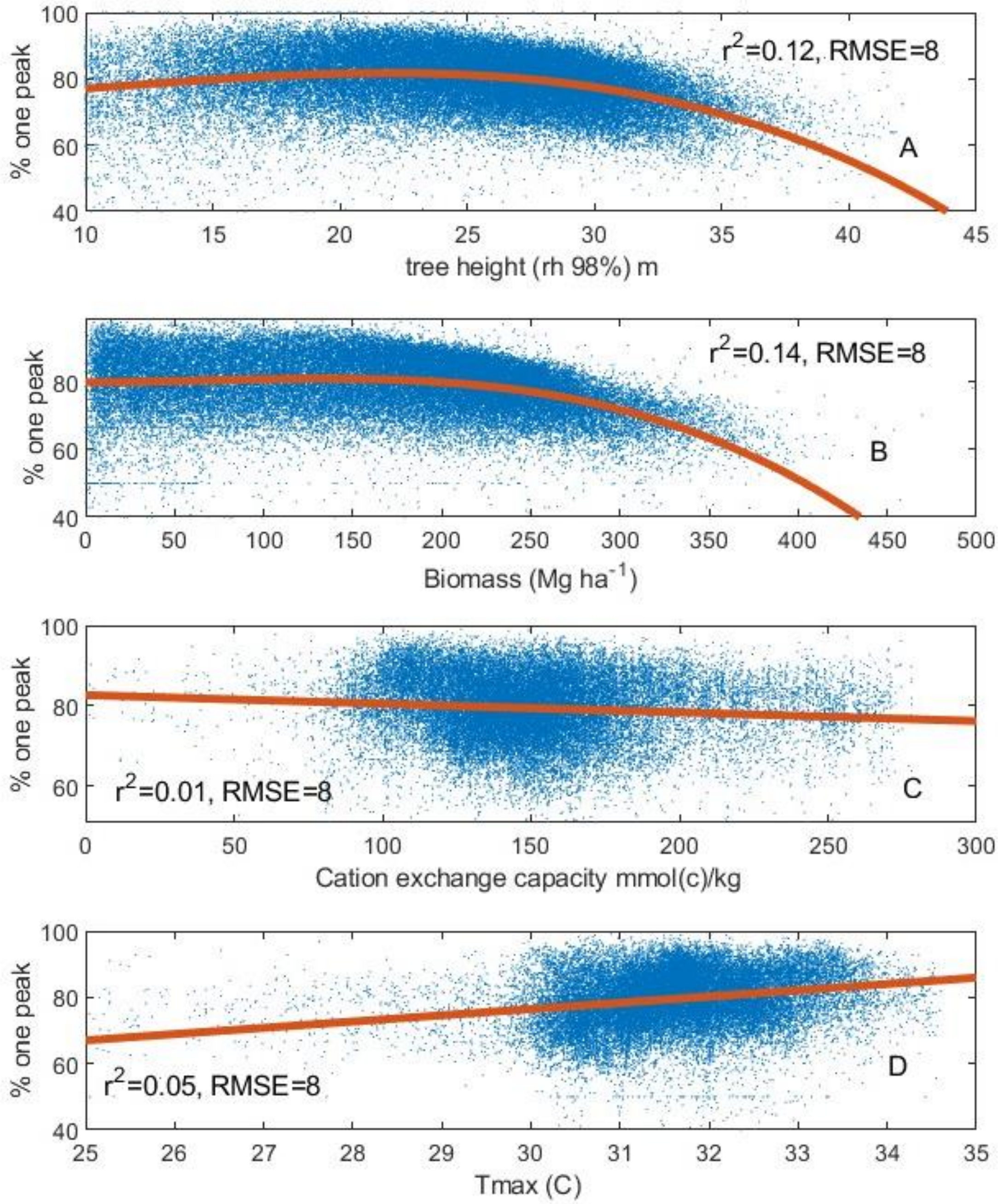

Figure 8 –(A) Tree height (rh98%), (B) AGBD from GEDI L4B, (C) cation exchange capacity (mmol(c)/kg and (D) T_{max} (°C) vs percent one peak forests for the Amazon basin. For each we show r^2 and RMSE.

References

Abatzoglou, J. T. *et al.* (2018) ‘TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015’, *Scientific Data*, 5(1), p. 170191. doi: 10.1038/sdata.2017.191.

Aguirre-Gutiérrez, J. *et al.* (2021) ‘Pantropical modelling of canopy functional traits using Sentinel-2 remote sensing data’, *Remote Sensing of Environment*, 252, p. 112122. doi: https://doi.org/10.1016/j.rse.2020.112122.

Araujo-Murakami, A. *et al.* (2014) ‘The productivity, allocation and cycling of carbon in forests at the dry margin of the Amazon forest in Bolivia’, *Plant Ecology & Diversity*. Taylor & Francis, 7(1–2), pp. 55–69. doi: 10.1080/17550874.2013.798364.

Asner, G. P. *et al.* (2014) ‘Landscape-scale changes in forest structure and functional traits along an Andes-to-Amazon elevation gradient’, *Biogeosciences*, 11(3), pp. 843–856. doi: 10.5194/bg-11-843-2014.

Asner, G. P. and Mascaro, J. (2014) ‘Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric’, *Remote Sensing of Environment*, 140, pp. 614–624. doi: https://doi.org/10.1016/j.rse.2013.09.023.

Baker, P. J. and Wilson, J. S. (2000) ‘A quantitative technique for the identification of canopy stratification in tropical and temperate forests’, *Forest Ecology and Management*, 127(1), pp. 77–86. doi: https://doi.org/10.1016/S0378-1127(99)00118-8.

Baraloto, C. *et al.* (2012) ‘Contrasting taxonomic and functional responses of a tropical tree community to selective logging’, *Journal of Applied Ecology*. John Wiley & Sons, Ltd, 49(4), pp. 861–870. doi: https://doi.org/10.1111/j.1365-2664.2012.02164.x.

Batjes, N. H., Ribeiro, E. and van Oostrum, A. (2020) ‘Standardised soil profile data to support global mapping and modelling (WoSIS snapshot 2019)’, *Earth System Science Data*, 12(1), pp. 299–320. doi: 10.5194/essd-12-299-2020.

Both, S. *et al.* (2019) ‘Logging and soil nutrients independently explain plant trait expression in tropical forests’, *New Phytologist*. doi: 10.1111/nph.15444.

Burns, P. *et al.* (2020) ‘Incorporating canopy structure from simulated GEDI lidar into bird species distribution models’, *Environmental Research Letters*. IOP Publishing, 15(9), p. 95002. doi: 10.1088/1748-9326/ab80ee.

Carreño-Rocabado, G. *et al.* (2016) ‘Land-use intensification effects on functional properties in tropical plant communities’, *Ecological Applications*. John Wiley & Sons, Ltd, 26(1), pp. 174–189. doi: https://doi.org/10.1890/14-0340.

Cobb, A. R. *et al.* (2023) ‘Improved terrain estimation from spaceborne lidar in tropical peatlands using spatial filtering’, *Science of Remote Sensing*, 7, p. 100074. doi: https://doi.org/10.1016/j.srs.2022.100074.

Doughty, C. E. *et al.* (2015) ‘Drought impact on forest carbon dynamics and fluxes in

Amazonia', *Nature*. doi: 10.1038/nature14213.

Doughty, C. E. *et al.* (2017) 'Can Leaf Spectroscopy Predict Leaf and Forest Traits Along a Peruvian Tropical Forest Elevation Gradient?', *Journal of Geophysical Research: Biogeosciences*. doi: 10.1002/2017JG003883.

Dubayah, R. *et al.* (2020) 'The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography', *Science of Remote Sensing*. doi: 10.1016/j.srs.2020.100002.

Duncanson, L. *et al.* (2022) 'Aboveground biomass density models for NASA's Global Ecosystem Dynamics Investigation (GEDI) lidar mission', *Remote Sensing of Environment*, 270, p. 112845. doi: https://doi.org/10.1016/j.rse.2021.112845.

Ehbrecht, M. *et al.* (2021) 'Global patterns and climatic controls of forest structural complexity', *Nature Communications*, 12(1), p. 519. doi: 10.1038/s41467-020-20767-z.

Eppley, T. M. *et al.* (2022) 'Factors influencing terrestriality in primates of the Americas and Madagascar', *Proceedings of the National Academy of Sciences*. Proceedings of the National Academy of Sciences, 119(42), p. e2121105119. doi: 10.1073/pnas.2121105119.

Fayolle, A. *et al.* (2014) 'A new insight in the structure, composition and functioning of central African moist forests', *Forest Ecology and Management*, 329, pp. 195–205. doi: https://doi.org/10.1016/j.foreco.2014.06.014.

Feldpausch, T. R. *et al.* (2011) 'Height-diameter allometry of tropical forest trees', *Biogeosciences*, 8(5), pp. 1081–1106. doi: 10.5194/bg-8-1081-2011.

Figueiredo, F. O. G. *et al.* (2018) 'Beyond climate control on species range: The importance of soil data to predict distribution of Amazonian plant species', *Journal of Biogeography*. John Wiley & Sons, Ltd, 45(1), pp. 190–200. doi: https://doi.org/10.1111/jbi.13104.

Grantham, H. S. *et al.* (2020) 'Anthropogenic modification of forests means only 40% of remaining forests have high ecosystem integrity', *Nature Communications*, 11(1), p. 5978. doi: 10.1038/s41467-020-19493-3.

Halle, F., Oldeman, R. and Tomlinson, P. (1980) *Tropical Trees and Forests: An Architectural Analysis*. New York: Springer.

Hansen, M. C. *et al.* (2013) 'High-Resolution Global Maps of 21st-Century Forest Cover Change', *Science*. American Association for the Advancement of Science, 342(6160), pp. 850–853. doi: 10.1126/science.1244693.

Higgins, M. A. *et al.* (2011) 'Geological control of floristic composition in Amazonian forests', *Journal of Biogeography*. John Wiley & Sons, Ltd, 38(11), pp. 2136–2149. doi: https://doi.org/10.1111/j.1365-2699.2011.02585.x.

Jackson, T. D. *et al.* (2021) 'The mechanical stability of the world's tallest broadleaf trees', *Biotropica*. John Wiley & Sons, Ltd, 53(1), pp. 110–120. doi: https://doi.org/10.1111/btp.12850.

Kays, R. and Allison, A. (2001) 'Arboreal tropical forest vertebrates: current knowledge and research trends BT - Tropical Forest Canopies: Ecology and Management: Proceedings of ESF

Conference, Oxford University, 12–16 December 1998’, in Linsenmair, K. E. et al. (eds). Dordrecht: Springer Netherlands, pp. 109–120. doi: 10.1007/978-94-017-3606-0_9.

Krieger, G. *et al.* (2007) ‘TanDEM-X: A Satellite Formation for High-Resolution SAR Interferometry’, *IEEE Transactions on Geoscience and Remote Sensing*, 45(11), pp. 3317–3341. doi: 10.1109/TGRS.2007.900693.

Lewis, S. L. *et al.* (2013) ‘Above-ground biomass and structure of 260 African tropical forests’, *Philosophical Transactions of the Royal Society B: Biological Sciences*. doi: 10.1098/rstb.2012.0295.

Liu, A., Cheng, X. and Chen, Z. (2021) ‘Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals’, *Remote Sensing of Environment*, 264, p. 112571. doi: https://doi.org/10.1016/j.rse.2021.112571.

Maclean, I. M. D. and Klinges, D. H. (2021) ‘Microclimc: A mechanistic model of above, below and within-canopy microclimate’, *Ecological Modelling*, 451, p. 109567. doi: https://doi.org/10.1016/j.ecolmodel.2021.109567.

Malhi, Y. *et al.* (2021) ‘The Global Ecosystems Monitoring network: Monitoring ecosystem productivity and carbon cycling across the tropics’, *Biological Conservation*, 253, p. 108889. doi: https://doi.org/10.1016/j.biocon.2020.108889.

Marselis, S. M. *et al.* (2018) ‘Distinguishing vegetation types with airborne waveform lidar data in a tropical forest-savanna mosaic: A case study in Lopé National Park, Gabon’, *Remote Sensing of Environment*, 216, pp. 626–634. doi: https://doi.org/10.1016/j.rse.2018.07.023.

Marselis, S. M. *et al.* (2020) ‘Evaluating the potential of full-waveform lidar for mapping pan-tropical tree species richness’, *Global Ecology and Biogeography*. John Wiley & Sons, Ltd, 29(10), pp. 1799–1816. doi: https://doi.org/10.1111/geb.13158.

Oliveira, B. F. and Scheffers, B. R. (2019) ‘Vertical stratification influences global patterns of biodiversity’, *Ecography*. John Wiley & Sons, Ltd, 42(2), p. 249. doi: https://doi.org/10.1111/ecog.03636.

Olson, D. M. *et al.* (2001) ‘Terrestrial Ecoregions of the World: A New Map of Life on Earth: A new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity’, *BioScience*, 51(11), pp. 933–938. doi: 10.1641/0006-3568(2001)051[0933:TEOTWA]2.0.CO;2.

Palminteri, S. and Peres, C. A. (2012) ‘Habitat Selection and Use of Space by Bald-Faced Sakis (Pithecia irrorata) in Southwestern Amazonia: Lessons from a Multiyear, Multigroup Study’, *International Journal of Primatology*, 33(2), pp. 401–417. doi: 10.1007/s10764-011-9573-0.

Réjou-Méchain, M. *et al.* (2021) ‘Unveiling African rainforest composition and vulnerability to global change’, *Nature*, 593(7857), pp. 90–94. doi: 10.1038/s41586-021-03483-6.

Richards, P. W. (1952) *The Tropical Rain Forest*. Cambridge University Press.

Riutta, T. *et al.* (2018) ‘Logging disturbance shifts net primary productivity and its allocation in Bornean tropical forests’, *Global Change Biology*. John Wiley & Sons, Ltd (10.1111), 24(7), pp. 2913–2928. doi: 10.1111/gcb.14068.

Shenkin, A. *et al.* (2019) 'The World's Tallest Tropical Tree in Three Dimensions', *Frontiers in Forests and Global Change*, 2. doi: 10.3389/ffgc.2019.00032.

Simard, M. *et al.* (2011) 'Mapping forest canopy height globally with spaceborne lidar', *Journal of Geophysical Research: Biogeosciences*. doi: 10.1029/2011JG001708.

Smith, A. P. (1973) 'Stratification of Temperature and Tropical Forests', *The American Naturalist*. The University of Chicago Press, 107(957), pp. 671–683. doi: 10.1086/282866.

ter Steege, H. *et al.* (2006) 'Continental-scale patterns of canopy tree composition and function across Amazonia', *Nature*, 443(7110), pp. 444–447. doi: 10.1038/nature05134.

Stork, N. E. (2018) 'How Many Species of Insects and Other Terrestrial Arthropods Are There on Earth?', *Annual Review of Entomology*. Annual Reviews, 63(1), pp. 31–45. doi: 10.1146/annurev-ento-020117-043348.

Swinfield, T. *et al.* (2019) 'Imaging spectroscopy reveals the effects of topography and logging on the leaf chemistry of tropical forest canopy trees', *Global Change Biology*. John Wiley & Sons, Ltd (10.1111), n/a(n/a). doi: 10.1111/gcb.14903.

Tang, H. *et al.* (2016) 'Characterizing leaf area index (LAI) and vertical foliage profile (VFP) over the United States', *Biogeosciences*, 13(1), pp. 239–252. doi: 10.5194/bg-13-239-2016.

Tang, H. and Dubayah, R. (2017) 'Light-driven growth in Amazon evergreen forests explained by seasonal variations of vertical canopy structure', *Proceedings of the National Academy of Sciences*. Proceedings of the National Academy of Sciences, 114(10), pp. 2640–2644. doi: 10.1073/pnas.1616943114.

Taubert, F. *et al.* (2021) 'Deriving Tree Size Distributions of Tropical Forests from Lidar', *Remote Sensing*. doi: 10.3390/rs13010131.

Terborgh, J. (1992) 'Maintenance of Diversity in Tropical Forests', *Biotropica*. [Association for Tropical Biology and Conservation, Wiley], 24(2), pp. 283–292. doi: 10.2307/2388523.

Tuomisto, H. *et al.* (2019) 'Discovering floristic and geoecological gradients across Amazonia', *Journal of Biogeography*. John Wiley & Sons, Ltd, 46(8), pp. 1734–1748. doi: https://doi.org/10.1111/jbi.13627.

Violle, C. *et al.* (2007) 'Let the concept of trait be functional!', *Oikos*. John Wiley & Sons, Ltd, 116(5), pp. 882–892. doi: https://doi.org/10.1111/j.0030-1299.2007.15559.x.

Supplementary Figures

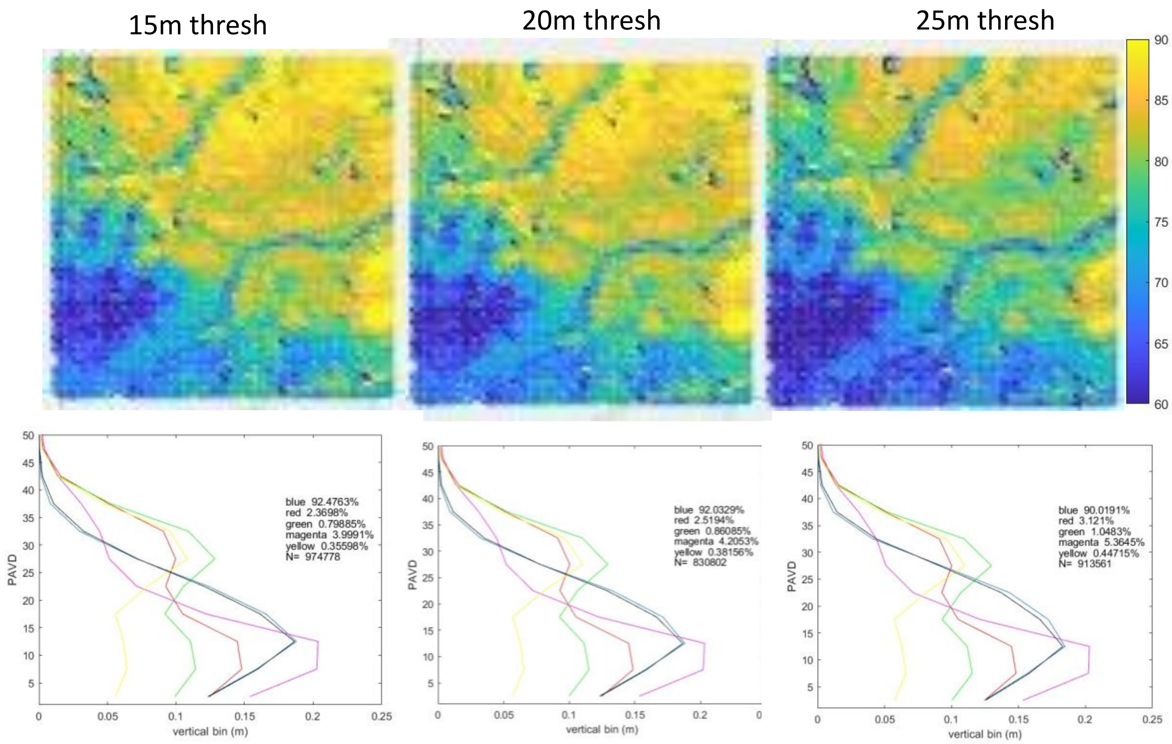

Figure S1 – (top) A map of % one peak forests in a 5 by 5 degree region of the Amazon where we modified our relative height metric 98% with a lower threshold of 15, 20, and 25m. (bottom) The different PAVD profiles for each threshold similar to fig 2.

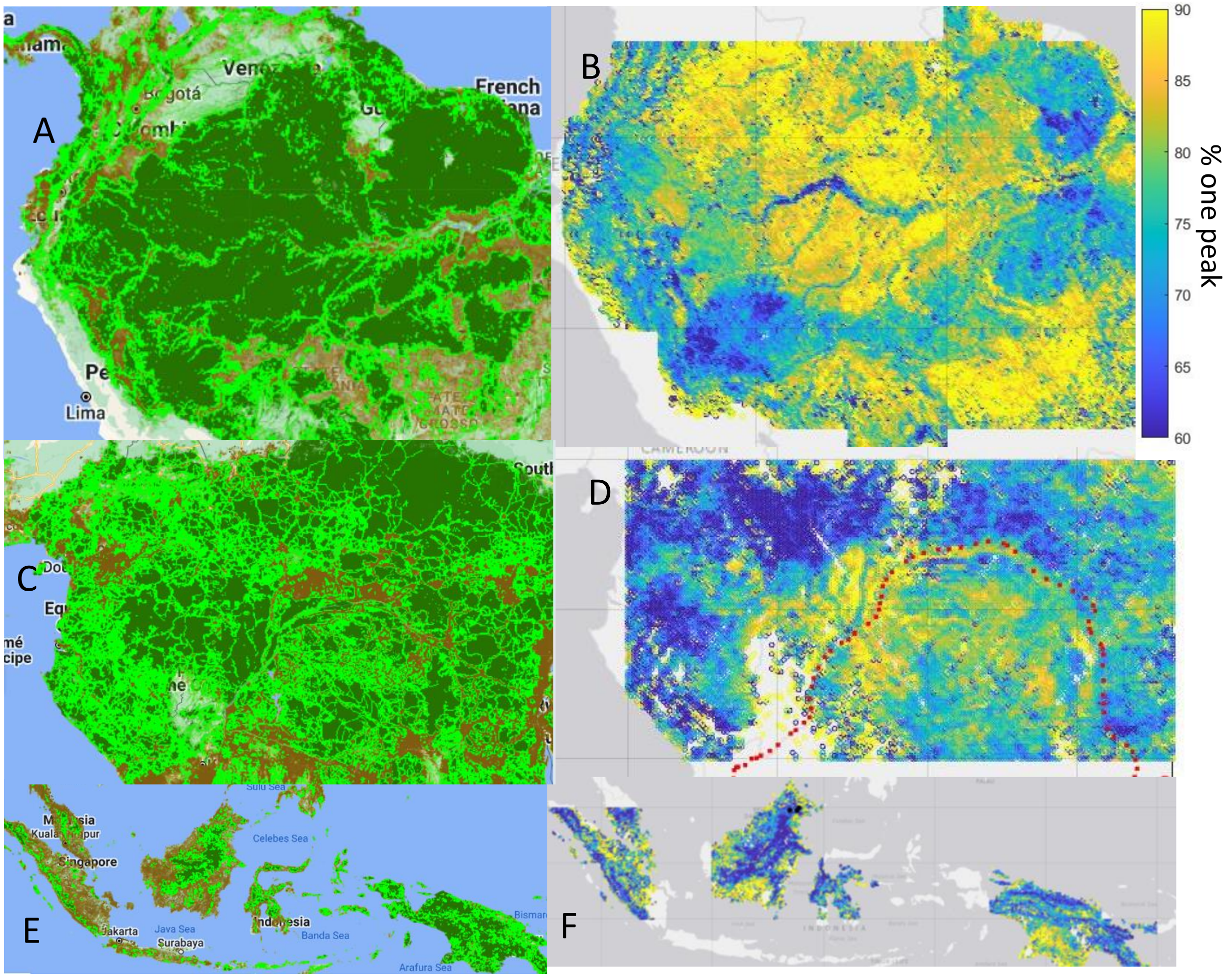

Figure S2 – A comparison of one peak forest types for (B) Amazonia, (D) Central Africa, and (F) SE Asia to an index of forest integrity as determined by degree of anthropogenic modification from [https://www.forestintegrity.com/ (Grantham *et al.*, 2020) for (A) Amazonia, (C) Central Africa, and (E) SE Asia where the darkest greens are areas with the least human disturbance.](https://www.forestintegrity.com/)

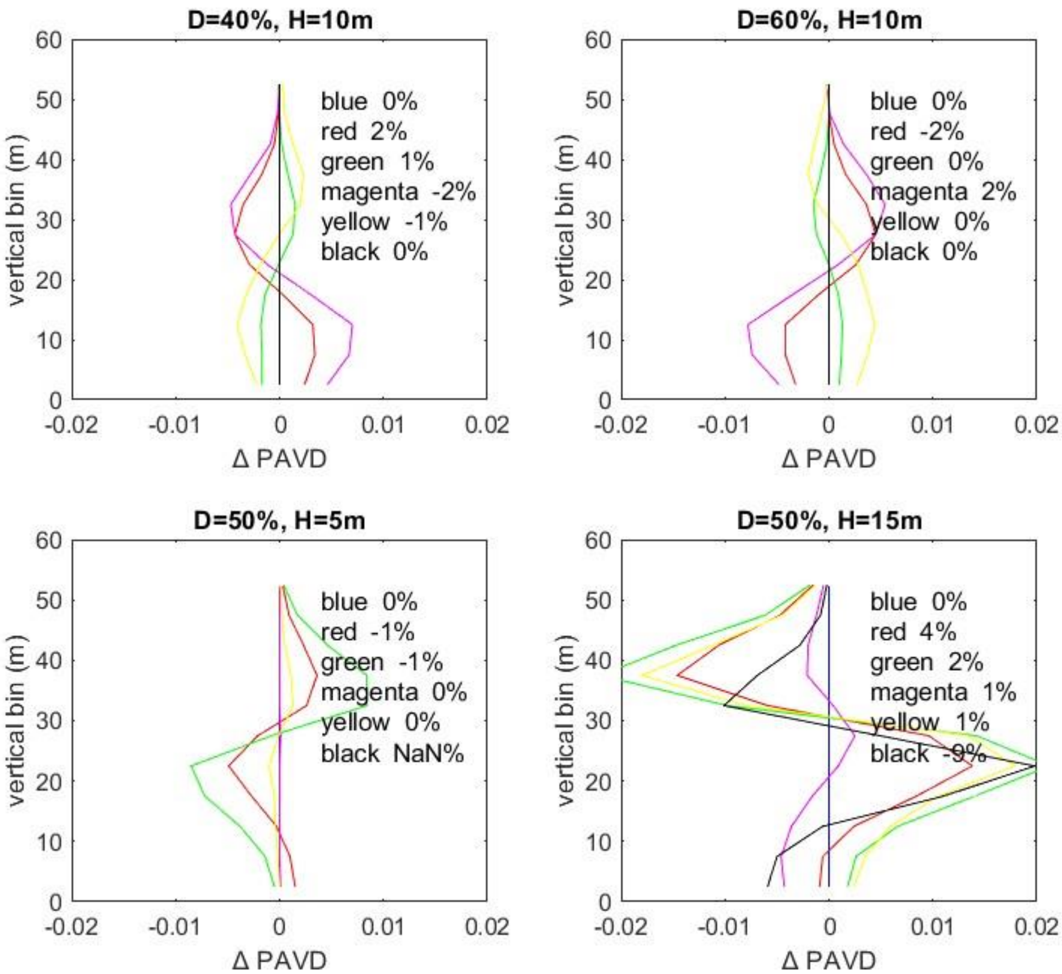

Fig S3 – We rerun the GEDI data from the black box in Central Africa shown in Fig 3 but changing the threshold parameters for D and H. We subtract the average PAVD waveform where the thresholds are D=50%, H=10m from the same dataset where D=40%, H=10m (top left), D=60%, H=10m (top right), D=50%, H=5m (bottom left), D=50%, H=15m (top left). The text is the percent change in number of those classifications that fit the category.

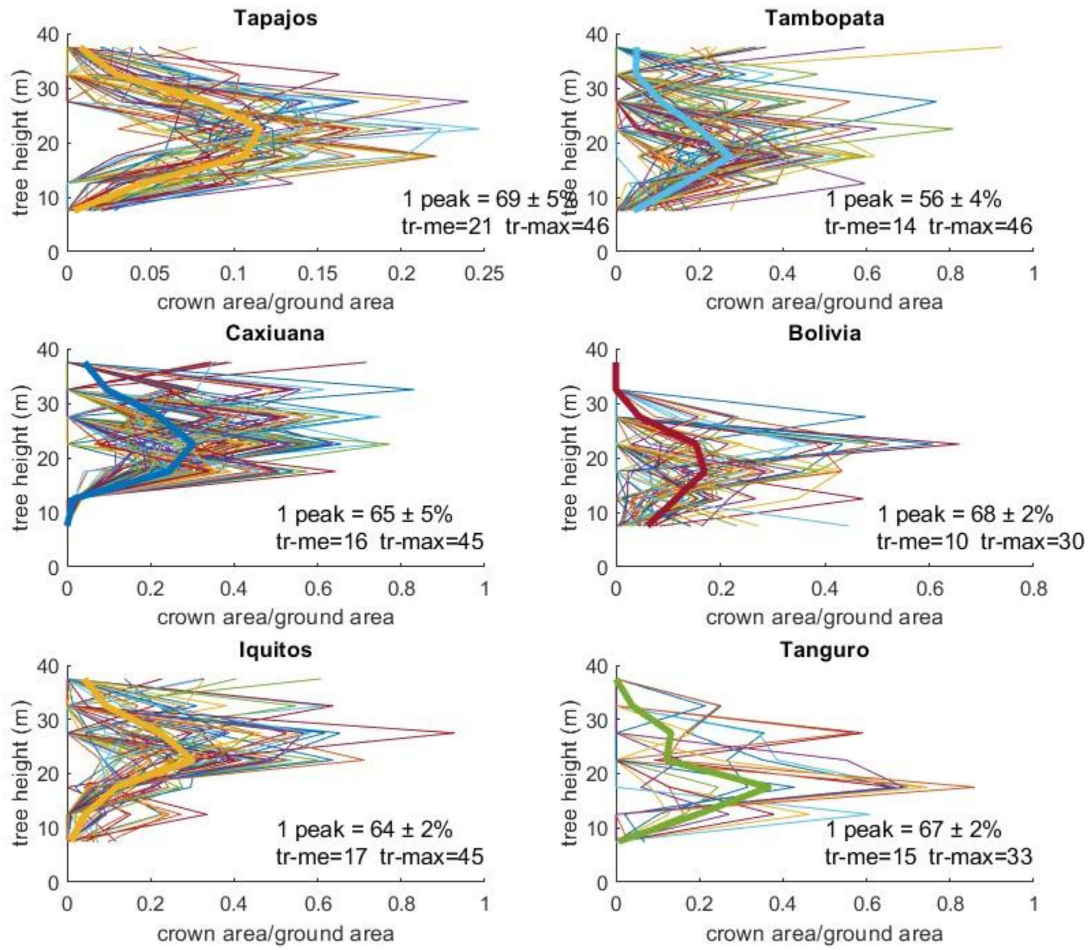

Figure S4 – Same as Fig 4 but instead of thin lines sampled at 25m spatial resolution we show at a 10m spatial resolution.

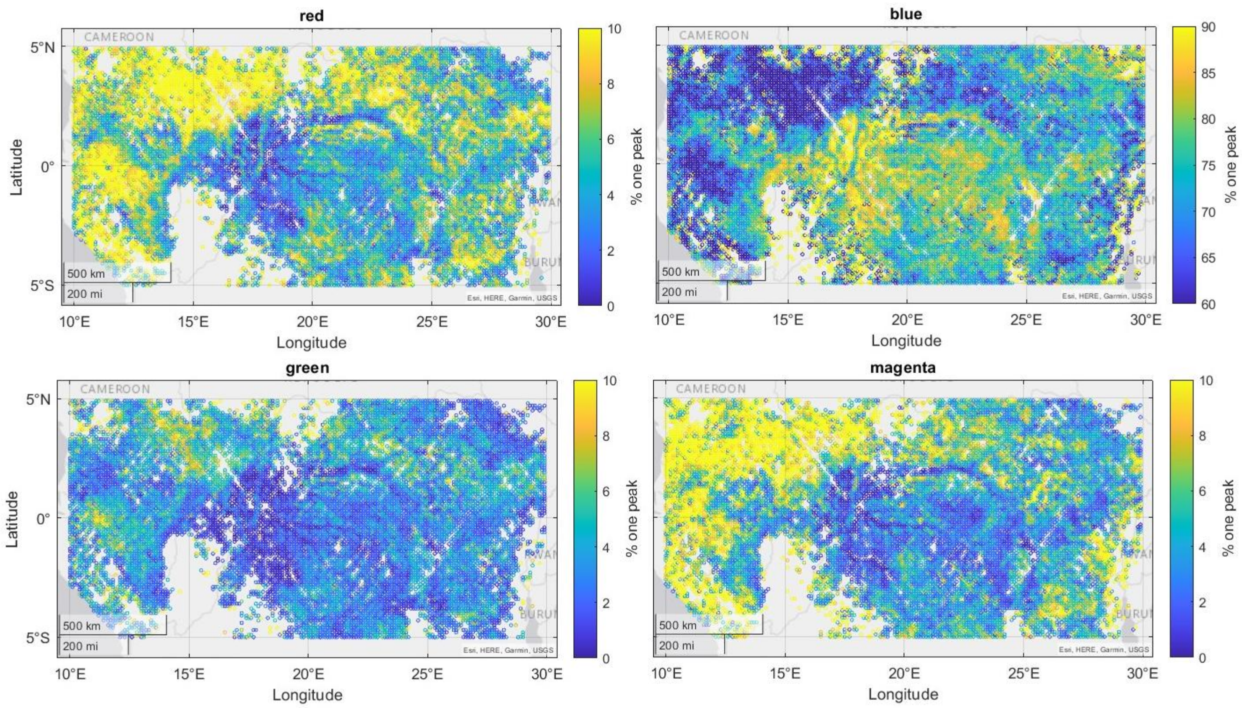

Figure S5 - Spatial distributions for different types of Central African "two peak" forests. The color labels are associated with the colors of the lines in Figs 2-3.

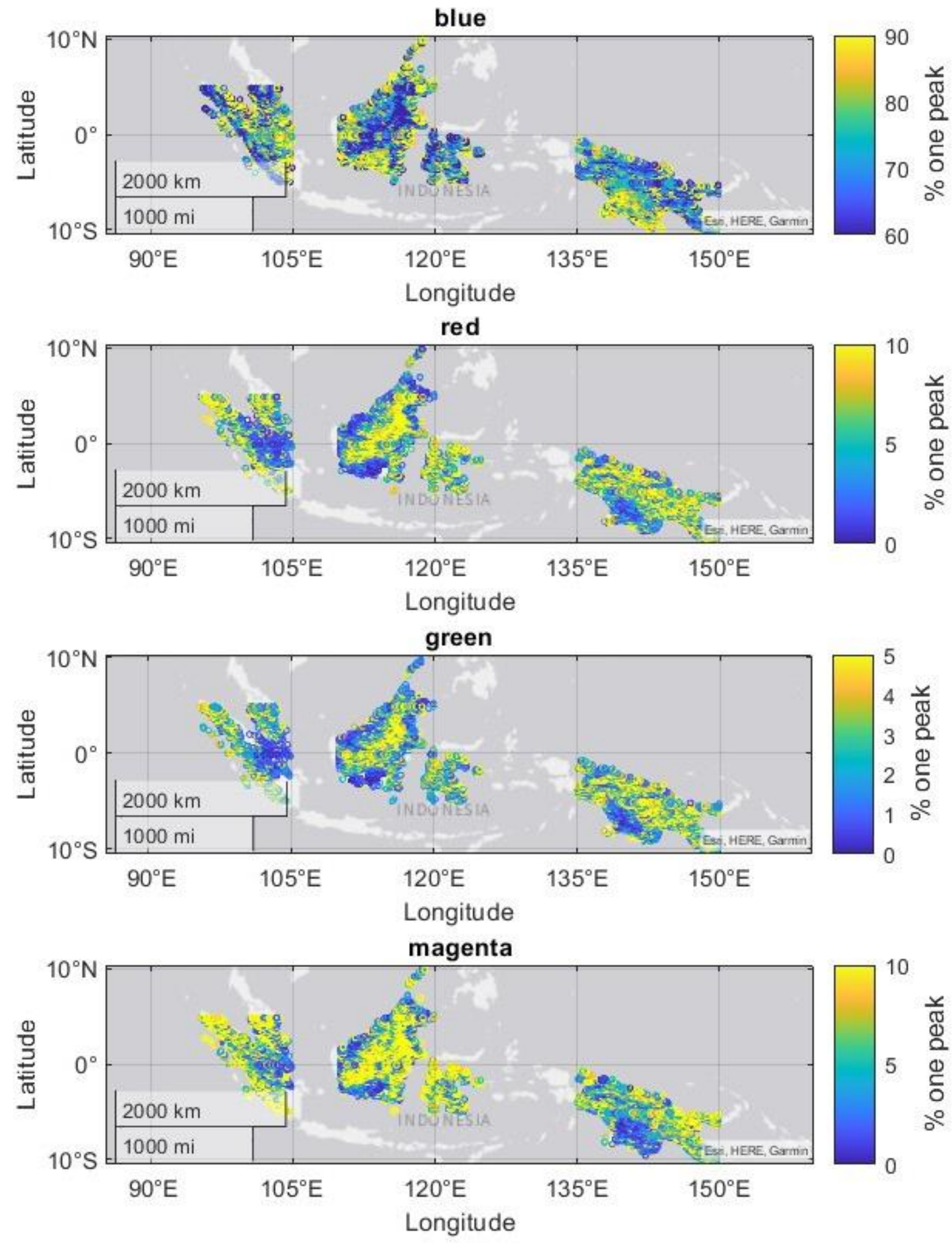

Figure S6 - Spatial distributions for different types of SE Asian "two peak" forests. The color labels are associated with the colors of the lines in Figs 2-3.

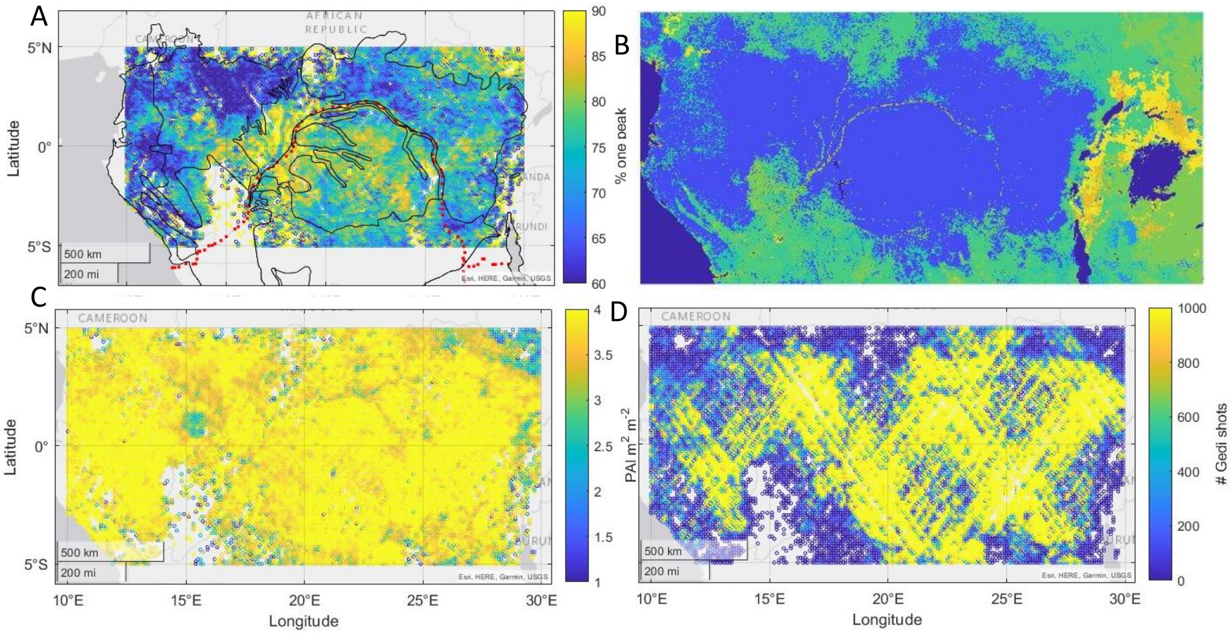

Figure S7 - Different data layers for Central Africa. (A) Spatial distribution of the percentage of 1 peak forests (same as figure 3) with ecoregions overlaid. (B) MODIS PFT classification with the light blue representing broadleaf tropical evergreen PFT. (C) Plant area index from GEDI and (D) # of GEDI shots per 0.1 by 0.1 pixel.

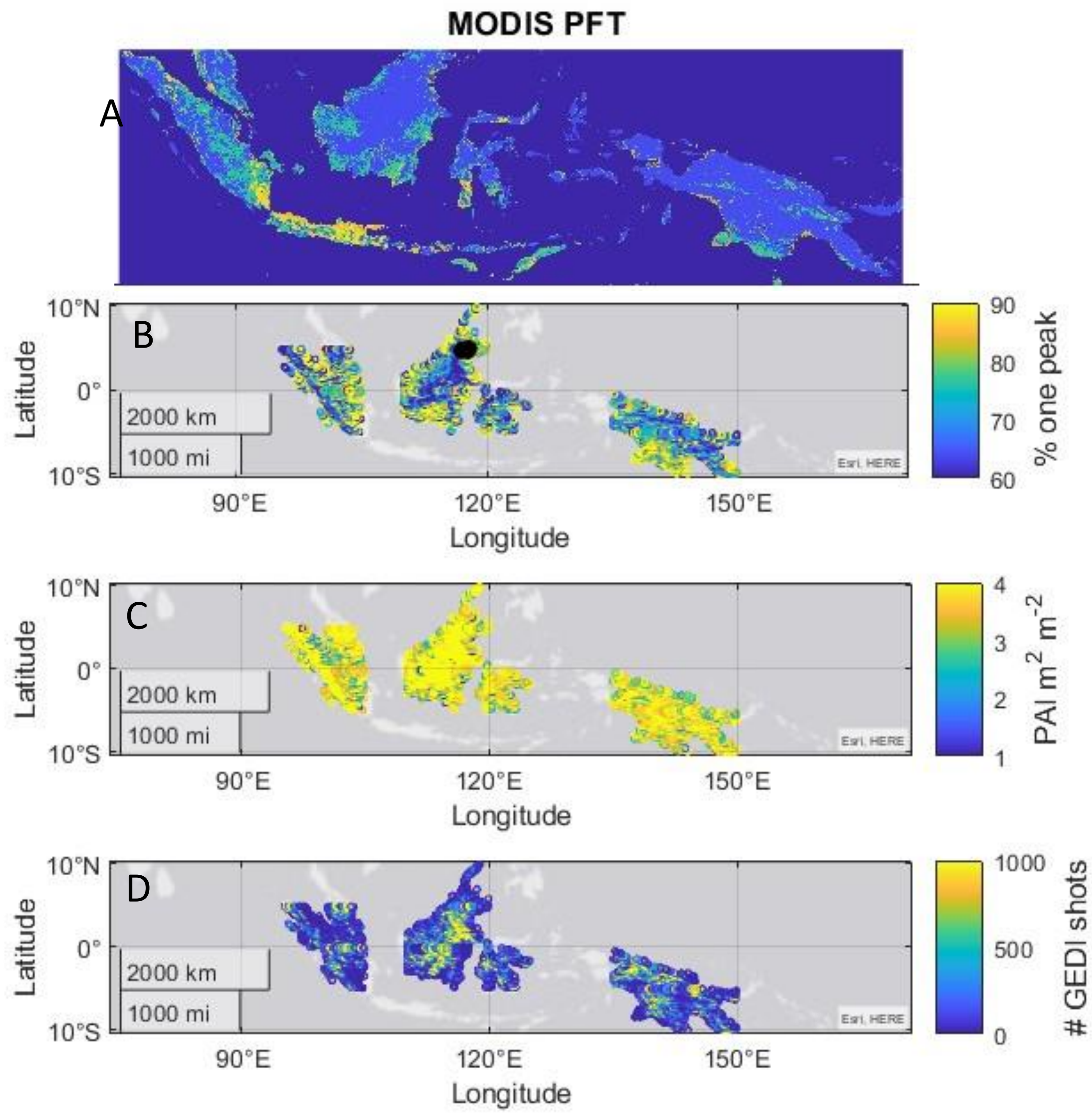

Figure S8 - Different data layers for SE Asia. (A) MODIS PFT classification with the light blue representing broadleaf tropical evergreen PFT. (B) Spatial distribution of the percentage of 1 peak forests (same as figure 3). (C) Plant area index from GEDI and (D) # of GEDI shots per 0.1 by 0.1 pixel.