scientific **data**

Check for updates

OPEN

DATA DESCRIPTOR

Intergovernmental Panel on Climate Change (IPCC) Tier 1 forest biomass estimates from Earth Observation

Neha Hunka[1✉], Laura Duncanson[1], John Armston[1], Ralph Dubayah[1], Sean P. Healey[2], Maurizio Santoro[3], Paul May[4], Arnan Araza[5,6], Clement Bourgoin[7], Paul M. Montesano[8,9], Christopher S. R. Neigh[8], Hedley Grantham[10,11], Peter Potapov[1], Svetlana Turubanova[1], Alexandra Tyukavina[1], Jessica Richter[12], Nancy Harris[12], Mikhail Urbazaev[1], Adrián Pascual[1], Daniela Requena Suarez[13], Martin Herold[5,13], Benjamin Poulter[8], Sylvia N. Wilson[14], Giacomo Grassi[7], Sandro Federici[15], Maria J. Sanz[16,17] & Joana Melo[7]

Aboveground biomass density (AGBD) estimates from Earth Observation (EO) can be presented with the consistency standards mandated by United Nations Framework Convention on Climate Change (UNFCCC). This article delivers AGBD estimates, in the format of Intergovernmental Panel on Climate Change (IPCC) Tier 1 values for natural forests, sourced from National Aeronautics and Space Administration's (NASA's) Global Ecosystem Dynamics Investigation (GEDI) and Ice, Cloud and land Elevation Satellite (ICESat-2), and European Space Agency's (ESA's) Climate Change Initiative (CCI). It also provides the underlying classification used by the IPCC as geospatial layers, delineating global forests by ecozones, continents and status (primary, young ($\leq$20 years) and old secondary ($>$20 years)). The approaches leverage complementary strengths of various EO-derived datasets that are compiled in an open-science framework through the Multi-mission Algorithm and Analysis Platform (MAAP). This transparency and flexibility enables the adoption of any new incoming datasets in the framework in the future. The EO-based AGBD estimates are expected to be an independent contribution to the IPCC Emission Factors Database in support of UNFCCC processes, and the forest classification expected to support the generation of other policy-relevant datasets while reflecting ongoing shifts in global forests with climate change.

[1]Department of Geographical Sciences, University of Maryland, 4600 River Road, Riverdale, 20737, Maryland, USA. [2]US Forest Service, Rocky Mountain Research Station, Riverdale, UT, 84405, USA. [3]GAMMA Remote Sensing Research and Consulting AG, Worbstrasse 225, Gümligen, Switzerland. [4]South Dakota Mines, Rapid City, South Dakota, 57701, USA. [5]Laboratory of Geo-Information and Remote Sensing, Wageningen University and Research, Droevendaalsesteeg, 46708, PB Wageningen, The Netherlands. [6]Earth Systems and Global Change Gr, Wageningen University & Research, Droevendaalsesteeg, 46708, PB Wageningen, The Netherlands. [7]European Commission, Joint Research Centre, Via E. Fermi 2749-TP 261, I-21027, Ispra (VA), Italy. [8]NASA Goddard Space Flight Center, Greenbelt, 20771, Maryland, USA. [9]ADNET Systems, Inc., Bethesda, 20817, Maryland, USA. [10]Bush Heritage Australia, Melbourne, 3000, Australia. [11]Centre for Ecosystem Science, University of New South Wales, Sydney, 2052, Australia. [12]World Resources Institute, 20002, Washington DC, USA. [13]Helmholtz Center Potsdam GFZ German Research Centre for Geosciences, Remote Sensing and Geoinformatics, Telegrafenberg, Potsdam, 14473, Germany. [14]United States Geological Survey, 12201 Sunrise Valley Drive, Reston, Virginia, USA. [15]Institute for Global Environmental Strategies, IGES, Hayama, 240-0112, Japan. [16]Basque Centre for Climate Change (BC3), Sede Building, 1, 1st floor, Scientific Campus of the University of the Basque Country, 48940, Leioa, Spain. [17]Ikerbasque Foundation, Euskadi Pl., 5,Abando, 48009, Bilbao, Spain. ✉e-mail: nhunka@umd.edu

Background & Summary

Aboveground dry woody Biomass Density (AGBD) maps produced with Earth Observation (EO) data have a large potential to periodically provide a transparent, consistent and replicable picture of the state of the world's forests. Such map estimates, calibrated and validated with *in situ* forest data, may provide crucial support to nations and stakeholders reporting to the United Nations Framework Convention on Climate Change (UNFCCC). They may, for example, serve as regular contributions of complete and comparable aggregate forest biomass estimates in the Global Stocktake (GST) under the Paris Agreement[1,2]. Yet, a recent synthesis of national-level reporting has concluded that only a handful of countries currently (as of 2024) use EO-based AGBD maps in their assessments for UNFCCC reporting[3], highlighting a gap that is partially attributed to the lack of provision of map estimates in line with policy guidance[4].

To overcome this challenge, as part of the Committee on Earth Observing Satellites (CEOS) Agriculture, Forestry and Other Land Use (AFOLU) Roadmap[5,6], a Biomass Harmonization activity[7] was initiated. The activity is an effort to release EO-based forest AGBD estimates in a manner comparable and consistent with those required for policy. The provision of mono-temporal estimates of default Tier 1 AGBD in natural forests, as in the Intergovernmental Panel on Climate Change (IPCC) Guidelines for National Greenhouse Gas Inventories (NGHGIs)[8,9,see Table 4.7], forms the first of the objectives under this activity. The current IPCC Tier 1 defaults were developed for countries lacking domestic data for national/sub-national assessments of carbon stock and fluxes, and as independent reference for use in technical assessment of NGHGIs. These default values are the most comprehensive assessment of AGBD for various ecoregions across most continents, compiled primarily from *in-situ* forest plots available until 2019. However, the paucity and uneven distribution of field plots implies that some ecoregions remain under-sampled, particularly in the tropics and subtropics, and are hence provided with crude average AGBD estimates (e.g. tropical dry forests in Africa), or are sourced from only a handful of references (e.g. primary, tropical moist deciduous forests in Africa)[10]. Ecoregions for which estimates are sourced from EO data (e.g. subtropical humid and dry forests in Africa) are currently based on a single version of an AGBD map[11] representing forest conditions in 2010.

The timely provision of EO-based AGBD estimates from recent, ongoing and upcoming space missions, which are modeled with the calibration of their data using ground plots, can synchronously enable the release of worldwide up-to-date values for the IPCC Tier 1 defaults[12]. Estimates can now be routinely derived from various spaceborne instruments that are specifically optimized to characterize carbon-relevant differences in vegetation height and structure. For example, the National Aeronautics and Space Administration's (NASA) Global Ecosystem Dynamics Investigation (GEDI)[13] has been used to estimate AGBD in ecoregions ranging from tropical rainforests to low-latitude boreal forests with data collected up to 2023[14]. Where coverage is absent, i.e. in high-latitude boreal forests, AGBD estimates based on the Ice, Cloud and land Elevation Satellite (ICESat-2) lidar[15] combined with Harmonized Landsat Sentinel-2 optical data[16] and topographic data from Europe's Copernicus Programme[17], are available[18]. Further, data from these lidar missions, as a part of ongoing research, can be exploited for AGBD gain/loss assessments. Similarly, the European Space Agency's (ESA) Climate Change Initiative (CCI) releases gridded estimates of global annual AGBD[19]. These maps are generated by first calibrating ICESat-2 data with (sub)national AGBD averages from field plots or published by national forest inventories, and then relating AGBD to data from current earth-orbiting synthetic aperture radar (SAR) sensors[20]. In contrast to the static snapshot of forest status/conditions in the current IPCC guidelines[8,9,see Table 4.7], such periodically updated EO-based datasets calibrated with *in situ* data can reflect dynamic forest conditions, including the effects of land use change and climate change on forest AGBD with time.

This article develops and demonstrates a framework to estimate IPCC's Tier 1 AGBD from data from spaceborne instruments as part of an international coordination effort. Specifically, it aims to provide (1) a complete dataset of EO-derived AGBD estimates (years circa 2020-2022) in the same practical and concise format of the IPCC Tier 1 table for natural forests[9,Table 4.7], derived from GEDI and ICESat-2, and ESA CCI Biomass, and (2) the underlying global stratification used in the IPCC table, classifying forests by Geographic Ecological Zones (GEZ)[21], continents and forest status/condition, i.e. primary (intact old growth or with no active anthropogenic intervention), young secondary (≤20 years) and old secondary (>20 years) forests. Dataset (1) is designed to contribute to the IPCC's Emission Factors Database[22], as a further step towards supporting UNFCCC-related estimation and reporting processes. The dataset is relevant to both compilers of NGHGIs and technical expert reviewers of submissions to the UNFCCC, as it is presented in a manner directly comparable to that used in IPCC generic methodologies. Dataset (2) is derived through a synthesis of existing EO-derived forest tree cover, height, age and land-use classification maps, such that their synergistic and combinatorial information is reduced to a single layer of forest status/condition[9,as in Table 4.7]. This layer is expected to support the derivation of mono-, bi- and multi-temporal EO-based AGBD loss/gain estimates, which will be enabled by forthcoming data from several ongoing (e.g. GEDI, ICESat-2, Sentinel and Advanced Land Observing Satellite (ALOS)) and planned (e.g. ESA BIOMASS[23] and NASA-ISRO Synthetic Aperture Radar (NISAR)[24]) missions. Overall, as and when improvements in EO-based AGBD estimates are available (e.g. with the use of national or global validation sites[25,26] and/or open-access *in situ* plots[27]), the framework presented in this study allows for timely updates to IPCC's Tier 1 AGBD values for natural forests.

Methods

Forest definitions. The IPCC, with the aim of ensuring consistency to NGHGIs, defines forests as "all land with woody vegetation consistent with thresholds used to define Forest Land" in the respective NGHGI, as well as lands with potential to "reach the threshold values"[28]. The UNFCCC provides a more specific explication, defining forests as areas of "minimum 0.05-1 ha, with minimum crown closure at maturity of 10–30%, and minimum tree height at maturity of 2–5 m"[29]. Young growing stands and areas expected to revert to these minima thresholds are also included under the definition of forests. Under this umbrella of definitions, the default IPCC Tier 1 AGBD

values include forests with at least 10% tree cover, classified in a combination of GEZ[21], continental boundaries and status/condition age classes, i.e. primary forests, young secondary forests (≤20 years) and old secondary forests (>20 years)[9,Table 4.7]. A single mean AGBD estimate and associated uncertainty is provided for most classes, derived from a meta-analysis of existing plot-level ground reference data where available, and the GlobBiomass product[11,30] where unavailable[10].

To replicate the IPCC Tier 1 table with AGBD estimates derived from data collected from spaceborne instruments, it was first required to stratify global forests into the same classes as the IPCC. Three considerations were required here - (1) the underlying classifications be produced independent of any existing EO-based AGBD map products, (2) information from the data used to create the classifications be harnessed to most closely match IPCC definition of forested lands, and (3) emphasis be placed on the reduction of potential commission errors rather than omission errors in the classification, since a single, and as precise as feasible, AGBD value is required for each class. Importantly, we do not formulate any new definitions of forested land, but leverage definitions from datasets to most closely align with the IPCC definition. Finally, EO-based estimates of mean AGBD from the GEDI and ICESat-2, and ESA CCI Biomass map, are derived for each delineated class. The following methods describe the generation of IPCC global forest classification, and the derivation of AGBD estimates for each class from the GEDI and ICESat-2 space mission data, and the ESA CCI Biomass map.

IPCC global forest classification. To establish a global map of forest status/condition classes, we conduct a Boolean analysis of a suite of existing datasets, including satellite-derived forest tree cover, height, age and land use classification layers (described in Table 1). In this approach, first, layers that identify a potential forest status/condition class (e.g. primary forests) are merged, and second, layers that identify sources of disagreement (e.g. presence of plantations or deforestation detected in the delineated primary forests) are used to remove areas of potential commission errors. Most of the maps capture conditions in the years between circa 2000 and 2020, allowing the distinction of young (≤20 years) and old secondary forests (>20 years). Due to the diverse sources of these independently-created layers, cases of duplicity in their native data is inevitable; for example, Landsat time-series are used frequently in products that detect disturbances to tree cover, albeit with different underlying definitions and assumptions[31]. This, however, has no impact on our analysis. Since our objective stresses on the reduction of commission errors, we simply take a stringent conservative approach that eliminates any area that is unlikely to be a forest in any given dataset. In this way, the approach exploits the combinatorial strength of multiple layers affirming that a patch of land must be either a primary, young secondary or old secondary forest, or else be eliminated from the classification. The global classification is run on the collaborative open-science could-computing system, the ESA-NASA Multi-mission analysis and algorithm platform (MAAP). MAAP has capabilities to host relevant data, processing algorithms, and computing capabilities in a common cloud environment, linked to a public GitHub repository (see Code Availability section).

For this study, layers were sourced from various host servers onto the MAAP (see references in Table 1), spatially aligned and overlaid by resampling to 10°×10° tile extents, with pixel sizes of 1 arc-second per pixel (~30 m per pixel at the equator) in EPSG:4326 projection (geographic latitude/longitude). Datasets provided as vectors were rasterized to this resolution. Table 1 provides the resampling techniques, dataset descriptions and abbreviations used to denote the inputs, and the index or pixel value of relevant classes within, merged as elaborated in the following sections:

Primary forests. The primary forest class (PF) was established using datasets identifying intact/primary forests, with a high forest integrity index, the presence of tree cover and forests heights of 5 m or more, and no known forest loss events, planted forests or plantations (Eq. (1)).

$$\begin{aligned}
PF = EWC \wedge (&(FLII > 9.6) \vee PHTF \vee IF \vee (JRCFT = 10) \wedge (FH_{2000} \geq 5 \text{ m}) \wedge (FH_{2020} \geq 5 \text{ m}) \\
&\wedge (TCLY = 0) \wedge (JRCFC = 0) \wedge \neg (SDPT > 0) \wedge \neg (JRCFT \geq 81 \wedge JRCFT \leq 86) \\
&\wedge \neg (GFM = 32 \vee GFM = 40 \vee GFM = 53))
\end{aligned}$$
(1)

where $\wedge$ is "and", $\vee$ is "or" and $\neg$ is "not" in Boolean notation. Abbreviations of the layers in Eq. (1) are provided in Table 1 and logical function illustrated in Fig. 1a.

Young secondary forests. The young secondary forest class (YS) captured pixels that were estimated to have had changes in forest height or cover between 2000 and 2020, excluding planted forests and plantations. It was established using datasets indicating heights of 5 m or more in 2020 and with either heights shorter than 5 m in 2000, or heights of 5 m or more in 2000 but having experienced tree cover loss after year 2000. These conditions ensure that various forest change dynamics that can occur in a 20-year period are captured, albeit pixels with tree cover loss occurring after year 2018 were ignored to minimize the risk of commission errors. Areas previously classified as primary forests (Eq. (1)) are also excluded. For the boreal ecoregions, where a separate Landsat-based dataset of stand ages is available (BFA in Table 1), forests younger than or equal to 20 years in 2020 were added to this class (Eq. (2)), with abbreviations in Fig. 1b and Table 1).

$$\begin{aligned}
YS = EWC \wedge (&((FH_{2020} \geq 5 \text{ m}) \wedge (FH_{2000} < 5 \text{ m}) \vee (JRCFT = 32 \vee JRCFT = 33)) \\
&\vee ((FH_{2020} \geq 5 \text{ m}) \wedge (FH_{2000} \geq 5 \text{ m}) \wedge ((TCLY > 2000) \wedge (TCLY < 2018))) \\
&\vee ((FH_{2020} \geq 5 \text{ m}) \wedge (FH_{2000} \geq 5 \text{ m}) \wedge ((JRCFC > 2000) \wedge (JRCFC < 2018))) \\
&\vee (BFA \leq 20 \text{ years}) \wedge \neg (SDPT > 0) \wedge \neg (JRCFT \geq 81 \wedge JRCFT \leq 86) \\
&\wedge \neg (GFM = 32 \vee GFM = 40 \vee GFM = 53))
\end{aligned}$$
(2)

Dataset	Abbreviation (in eqs. (1)–(3))	Description	Native pixel/unit size (resampling method)	Year(s) represented
ESA World Cover 2021[45]	EWC	Global classification of land use, derived by the integration of Sentinel 1 and Sentinel 2 imagery. For this study, the class with values of 10.0 is used to mask areas with tree cover.	~10 m (resampled to ~30 m using "mode" function)	2021
Global Forest Landscape Integrity Index[69]	FLII	The degree of anthropogenic modification in forests, provided on a continuous scale of 0 to 10. Values of high forest integrity (scores $\geq$ 9.6) define "interiors and natural edges of more or less unmodified naturally regenerated (non-planted) forest ecosystems, comprised entirely or almost entirely of native species."	~300 m (resampled to ~30 m using "average" function)	2019
Intact Forest Landscapes[70]	IFL	The extent of forest landscapes that exhibit no signs of human activity or habitat fragmentation, large enough to maintain native biological diversity, derived from Landsat imagery.	Vector, delineating contiguous blocks of a minimum size of 500 km², at least 10 km broad and 2 km wide in corridors/appendages (rasterized to ~30 m)	2020
Primary Humid Tropical Forests[71]	PHTF	The extent of primary forests, including primary forest fragments and forests close to and within areas of human activity, derived from Landsat imagery.	~30 m (no resampling necessary)	2001
Global Forest Height[44]	FH	Global forest canopy height and height change estimates (in meters), developed through the integration of the GEDI lidar forest structure metrics and Landsat analysis-ready data time-series.	~30 m (no resampling necessary)	2000 and 2020
Joint Research Centre Tropical Moist Forest transition map[47]	JRCFT	Status and transition stages in tropical moist forests derived using the Landsat time series. The map provides a classification of closed evergreen or semi-evergreen forest without observed disturbance (value 10), old forest regrowth (deforested before 2002 followed by permanent vegetative regrowth, value 31) and young forest regrowth (deforested between 2002 and 2018 followed by 3 years minimum of vegetative regrowth, values 32 and 33). Plantations (mainly rubber and oil palm, values 81–86) are identified by visual interpretation of high-resolution imagery[47] and various external data sources[72–77].	~30 m (no resampling necessary, aligned to grid using "nearest neighbour")	1982 to 2022
Boreal Forest Age[78]	BFA	For Boreal regions (47° to 73° north), estimates of forest stand age (up to 36 years since year 2000) from Landsat-derived locally-calibrated estimates of tree canopy cover.	~30 m (no resampling necessary, aligned to grid using "average" function)	Annual, 1984 to 2020
Global Tree Cover Loss Year[79]	TCLY	Spatially explicit areas of forest loss, with pixel values denoting the year of occurrence of loss, derived using Landsat time-series imagery.	~30 m (no resampling necessary)	Annual, 2000 to 2020
Joint Research Centre Tropical Moist Forest change map[47]	JRCFC	Spatially-explicit areas of deforestation and forest degradation in tropical moist forests, with pixel values denoting the year of occurrence, derived using Landsat time-series imagery. Forest degradation from selective logging, fire or natural events is defined as temporary disturbances occurring for up to 2.5 years, while deforestation is a longer-term disturbance where disruptions to tree cover (absence of tree foliage cover within a 30-m pixel) are visible for more than 2.5 years.	~30 m (no resampling necessary, aligned to grid using "nearest neighbour")	Annual, 1982 to 2022
Spatial Database of Planted Trees (v2.0)[46]	SDPT	Global differentiation of planted forests and tree crops from natural and seminatural forests, compiled from national governments, nongovernmental organizations, independent researchers, or a combination of sources.	~30 m (no resampling necessary)	2020
Global Forest Management[52]	GFM	Spatially explicit information on forest management, including short rotation plantations for timber (value 32), oil palm plantations (value 40) and agroforestry (value 53), generated with PROBA-V satellite imagery, expert and crowdsourcing campaigns.	~100 m (resampled to ~30 m using "nearest neighbour")	2015

Table 1. Datasets compiled to determine global forest status/condition classes, i.e. primary forests, young secondary forests ($\leq$20 years) and old secondary forests (>20 years). Datasets in geographic latitude/longitude coordinates are reported with approximate pixel sizes at equator.

where $\wedge$ is "and", $\vee$ is "or" and $\neg$ is "not" in Boolean notation. Abbreviations are provided in Table 1 and the logical function illustrated in Fig. 1b.

Old secondary forests. The old secondary forest class (OS) captured the remainder of pixels with forests upon excluding the primary and young secondary forest classes (Eq. (3)). These pixels must have been observed to have forest heights of 5 m or more in both 2000 and 2020, with no tree cover loss nor forest disturbances detected after 2000, nor any planted forests or plantations. Areas previously classified as primary forests or young secondary forests (Eqs. (1) and (2)) were also excluded. For the boreal ecoregions, forests between 20–36 years of age were added to this class.

$$OS = EWC \wedge (FH_{2020} \geq 5 \text{ m}) \wedge (FH_{2000} \geq 5 \text{ m}) \vee (JRCFT = 31)$$
$$\wedge (TCLY = 0) \wedge (JRCFC \leq 2000) \vee (BFA > 20 \text{ years} \wedge BFA \leq 36 \text{ years})$$
$$\wedge \neg (SDPT > 0) \wedge \neg (JRCFT \geq 81 \wedge JRCFT \leq 86)$$
$$\wedge \neg (GFM = 32 \vee GFM = 40 \vee GFM = 53), \tag{3}$$

where $\wedge$ is "and", $\vee$ is "or" and $\neg$ is "not" in Boolean notation. Abbreviations are provided in Table 1 and the logical function illustrated in Fig. 1c.

Biomass estimates based on spaceborne lidar.   *NASA GEDI estimates of forest aboveground biomass.*   The GEDI mission is a spaceborne lidar instrument mounted on the International Space Station (ISS) to acquire high-precision estimates of forest height and structure globally. Its data is provided as discrete ~25-m footprint-level samples of waveform lidar over the globe up to ±~51.6° latitude. Relative height metrics extracted from these data are calibrated with field and airborne lidar to model AGBD and associated uncertainty. These footprint-level AGBD estimates are provided in the publicly-available Level 4A (L4A) product, which is cross-validated with plot-level data in various geographic regions of the world[32,33]. For estimation of mean AGBD for any grid cell (e.g. 1 × 1 km), a closed-form statistical estimation approach called hybrid inference is used, where GEDI L4A predictions are combined within the frame of a designed sample to infer both mean AGBD and the variance around that estimated mean[34]. The statistical variance estimators used in this approach are composed of terms representing both sampling uncertainty (on the basis of sampling theory applied to the spatial distribution of GEDI footprints) and model uncertainty (derived from the covariance matrix produced around the terms in the parametric models used to predict footprint-level AGBD)[35]. GEDI's hybrid inference estimators do not account for non-uniform sampling probability caused by the variations in the ISS orbit or cloud cover. Hence, for estimation of mean AGBD over a large arbitrary shaped region (e.g., a country or forest class in this study), estimates are aggregated from smaller area grid cells (6 × 6 km) where the probability of sampling can be considered constant. In the aggregation of these grid cells, additional uncertainty terms (expressed as the sum of covariances) are derived to account for spatially correlated model and sampling errors introduced by the same L4A model used in multiple grid cells and the same GEDI tracks crossing multiple grid cells, respectively[36].

The GEDI AGBD estimates (in the latest product at the time of writing) are based on observations from mission week 19 to 223, i.e. April 2019 to March 2023[14,37]. For these AGBD estimates, as well as those generated in this study, only high-quality waveforms to which the L4A models are applicable (e.g. leaf-on conditions for deciduous broadleaf forests) are selected for inclusion in the sample for a region[34,36]. The total uncertainty of GEDI AGBD estimates are expressed as the standard error (SE) of the mean.

High northern latitude estimates of forest aboveground biomass.   High northern latitude estimates of forest aboveground biomass density are compiled for the circumpolar boreal forest north of ~51. 6° using a dataset built specifically for this domain. These estimates were mapped using a suite of spaceborne data products derived from lidar (ICESat-2), multi-spectral and SAR sensors[18]. The multi-spectral and SAR-derived data products were assembled from open source archives of Harmonized Landsat Sentinel-2 (HLS)[16] and Copernicus GLO30[17] datasets. The final map resolves mean and uncertainty of modelled AGBD as 30-m gridded estimates continuously across North America and Eurasia.

The procedure to derive the gridded map involved multiple model building and application steps. These first included (1) a calibration of NASA's ICESat-2 30-m lidar samples collected from 2019 to 2021 using field plot measurements of AGBD and simulations of ICESat-2's ATL08 canopy heights from coincident airborne lidar, following[32]. This yielded models of AGBD from ICESat-2 forest heights, which were (2) applied to all quality-filtered ICESat-2 lidar canopy height estimates, which yielded samples of AGBD across the circumboreal domain. Data are filtered to include only growing season nighttime data, which can significantly reduce the sample for high northern latitudes where much of the growing season is in near 24 hours of daylight. In areas with low availability of quality (low sun angle) data, data were sampled from across the boreal domain. Then, (3) gridded models are then developed predicting 30 m AGBD as a function of a suite of optical and SAR-based covariates, with ICESat-2 estimates of AGBD serving as training data in 90 km local random forests models. ICESat-2 training data are filtered to include only growing season nighttime data, which can significantly reduce the sample for high northern latitudes where much of the growing season is in near 24 hours of daylight. In areas with low availability of quality (<5° sun angle) data, ICESat-2 data were sampled from across the full boreal domain. Finally, (4) these models were applied to each 90 km tile across the circumboreal and north of ~51. 6° to predict spatially-continuous 30-m estimates of AGBD. Uncertainties were estimated per pixel by propagating uncertainties from models in steps 1 and 3 through large-compute bootstrapping. The per pixel uncertainty is the standard error of a suite of predictions based on model refitting from random samples in steps (1) and (3).

Combining GEDI and high northern latitudes estimates.   For each forest class that overlaps both the domain of GEDI (up to 51.6° latitude) and the high northern latitudes, only a heuristic of the mean and associated error is estimated for the purposes of this study.

$$\hat{y}_T = \frac{\hat{y}_G \cdot A_G + \hat{y}_H \cdot A_H}{A_G + A_H}$$

(4)

where $\hat{y}_T$ is the estimated AGBD for the entire area of the forest class, $\hat{y}_G$ is the GEDI estimated mean AGBD over area A_G, and $\hat{y}_H$ is the high northern latitude estimated mean AGBD over area A_H. And,

$$\hat{\sigma}_T = \sqrt{\frac{A_G^2 \cdot \hat{\sigma}_G^2 + A_H^2 \cdot \hat{\sigma}_H^2}{(A_G + A_H)^2}}$$

(5)

where $\hat{\sigma}_T$ is the estimated standard error of mean AGBD over the entire area of the forest class, $\hat{\sigma}_G$ is the GEDI estimated standard error over area A_G, and $\hat{\sigma}_H$ is the high northern latitude estimated standard error over area A_H.
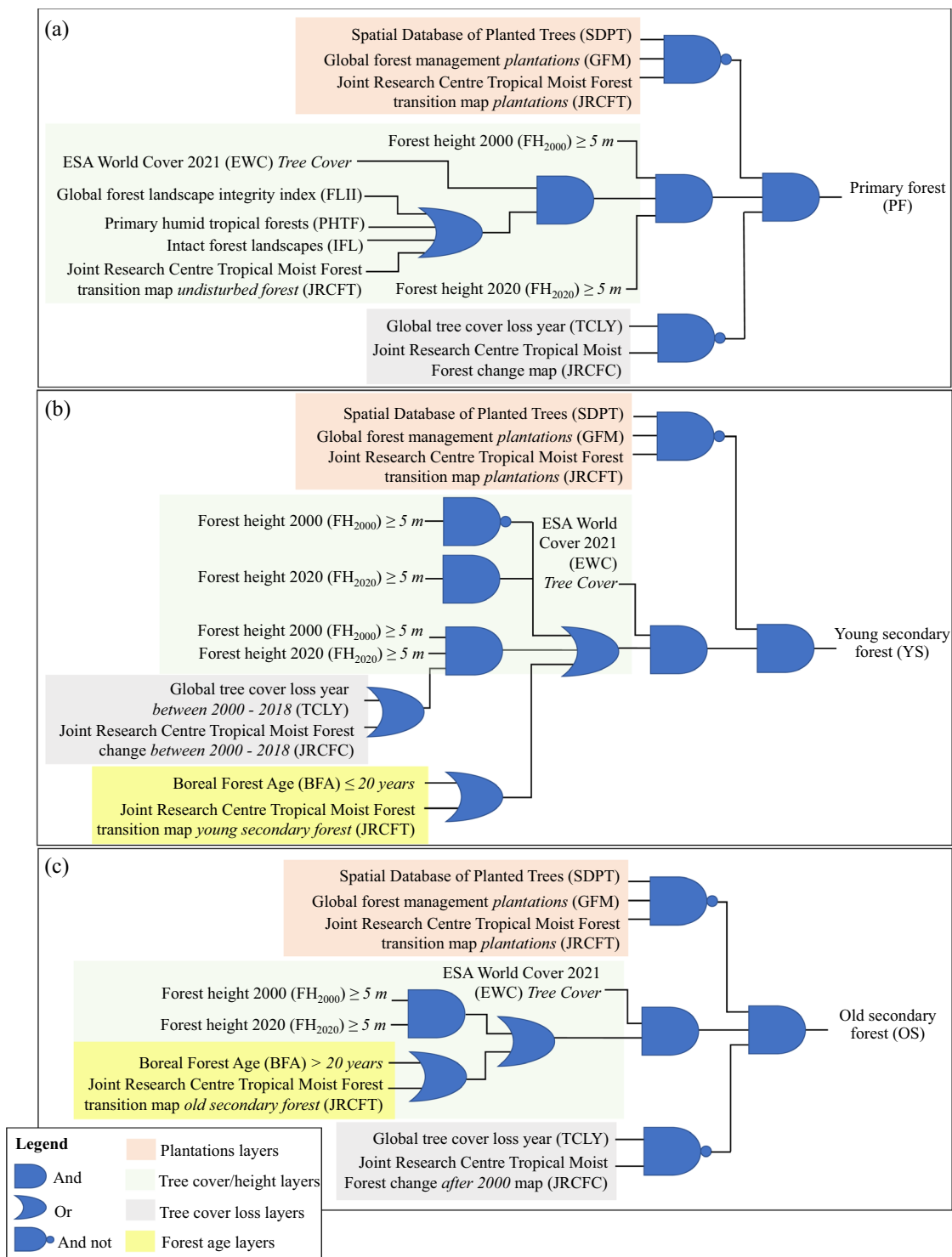
Fig. 1 Boolean circuit board diagram illustrating the logical combinations of various datasets and layers used to delineate (**a**) primary, (**b**) young secondary, and (**c**) old secondary forests. The approach is further outlined in Eqs. (1)–(3), with the datasets described in Table 1.

Biomass estimates based on spaceborne radar.　*ESA CCI estimates of forest aboveground biomass.*　The CCI Biomass dataset consists of annual estimates of AGBD of woody vegetation with a pixel size of 1 ha. This study uses the map product of the year 2020 available in version 4 of the Climate Research Data Package[19]. AGBD was estimated from synthetic aperture radar (SAR) observations of the backscattered intensity, recorded by the Advanced Land Observing Satellite 2 (ALOS-2) Phased Array L-band SAR 2 (PALSAR-2) and the Sentinel-1A and -1B satellite units[20]. A physics-aware model of SAR backscatter was inverted for each observation to obtain estimates of AGBD (ibid.). These estimates were then synthesized to a single value with a weighted average that exploited the sensitivity of the radar backscatter to AGBD, which depended on seasonal conditions and

wavelength of the SAR system. The retrieval was aided by auxiliary data layers acting as a constraint to the inversion (structural models and maximum AGBD based on spaceborne lidar observations from ICESat-2, land cover, tree canopy density etc.). The precision of an AGBD estimate was quantified by an standard error, obtained by propagating individual errors of the SAR backscatter measurements and the retrieval model parameters throughout the retrieval approach. Few studies have assessed the CCI map accuracy to AGBD estimates from a global collection of National Forest Inventories and research plots, highlighting that networks of regularly monitored forest biomass reference sites may better reflect regional performance[38,39].

For each forest class delineated in this study, the uncertainty of average AGBD accounted for both the spatial variance and the spatial covariance of the retrieval error (equation (6)).

$$\delta(AGB_{\mathrm{ave}})^2 = \frac{1}{N}\sum_{i=1}^{N}\delta(AGB_i)^2 + \frac{2}{N}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N} \mathrm{Cov}\,(\delta AGB_i, \delta AGB_j) \tag{6}$$

where

$$\mathrm{Cov}\,(\delta AGB_i, \delta AGB_j) = r_{\mathrm{spat},ij} \cdot \delta AGB_i \cdot \delta AGB_j \tag{7}$$

In equation (6), $\delta(AGB_{\mathrm{ave}})^2$ and $\delta(AGB_i)^2$ represent the variance of the average AGBD and the AGBD estimate at a given pixel within the averaging window, respectively. Variable $r_{spat,ij}$ represents the correlation of errors between the estimates of AGBD from pixel i and pixel j within the averaging window. The error correlation was modelled with an exponential function based on experimental assessment that used several local AGBD maps derived from airborne lidar as reference (Eq. (8))

$$r = 1 - e^{-k \cdot D} \tag{8}$$

Because of the non-uniform distribution of these reference AGBD maps across landscapes and biomes, a single global value corresponding to an average exponential was eventually used ($k = 0.0445$).

Comparison of IPCC and EO estimates of AGBD. The two sets of AGBD estimates (IPCC Tier 1 defaults, and EO-based estimates from GEDI/ICESat-2 and CCI) are a priori independent, derived using different sources of inputs and methods. None can be considered to be the true value of global forest AGBD. The standard deviation of the IPCC values captures the variability in the many compiled datasets to generate Tier 1 defaults[9,Table 4.7], and the uncertainty of the EO-based estimates captures variance resulting from the finite set of training data in their underlying models and AGBD samples for each forest class. Hence, the comparisons of the two sets of estimates, while accounting for their uncertainties, is limited to the following:

- Identifying whether there are systematic differences between the AGBD estimates when *grouped* by either continent, GEZ or forest status/condition. For each such group, we use a regression model (Eq. (9)), with the response variable equal to the difference in AGBD estimates (for example, EO-based mean AGBD estimate minus the IPCC value) in ith forest class, such that

$$\hat{y}_i = \alpha + \delta_i + \varepsilon_i; \quad \varepsilon_i \sim N(0, \sigma^2); \quad i \in \{1, \ldots, n\}. \tag{9}$$

Variable $\hat{y}$ is the difference in estimated AGBD. Parameter α is the unknown constant systematic difference between the two estimates, and is the primary parameter of interest. Variable δ_i is the sampling error, i.e. the variation in the two estimates as a result of their finite respective samples. The variance of δ_i is fixed at the sum of the respective estimate variances. Finally, ε_i is the *inherent* error, depicting the random (but mean zero) discrepancies that would exist between the two estimates even with limitless respective samples. Variable ε_i has unknown variance σ^2. This mixed effects model is a simple form of a Fay-Herriot regression[40,41], containing both the sampling and inherent error. By accounting for the sampling error, inference is based on underlying systematic deviations between the two sets of AGBD estimates and not on random deviations occurring simply due to finite sampling. We are interested in the estimate of α, its associated error and its significance (p-value), which would indicate whether there is sufficient evidence to infer that systematic differences exist between the two sets of AGBD estimates.

- Identifying *individual* forest classes with potentially significant differences between the sets of AGBD estimates. For each ith class, we compute a test statistic (t) as the ratio of difference in mean estimates (for example, EO-based mean AGBD estimate minus the IPCC value) to the total estimated variance (Eq. (10)),

$$t = \frac{\hat{y}_{EO(i)} - \hat{y}_{IPCC(i)}}{\sqrt{\widehat{\mathrm{Var}}[\hat{y}_{EO(i)}] + \widehat{\mathrm{Var}}[\hat{y}_{IPCC(i)}]}} \tag{10}$$

where $\hat{y}_{EO(i)}$ and $\hat{y}_{IPCC(i)}$ are the EO-based estimates and IPCC values, with variance of $\widehat{\mathrm{Var}}[\hat{y}_{EO(i)}]$ and $\widehat{\mathrm{Var}}[\hat{y}_{EO(i)}]$ respectively. Instead of setting a single cut-off threshold confidence level (e.g. 95% confidence interval), which would conclusively demarcate which classes are and are not systematically different, we simply report the t value. Large absolute t values indicate the differences in mean are large relative to the sampling uncertainties of the estimates.

Various forms of estimates of uncertainty in AGBD are provided by the IPCC, i.e. SD, confidence intervals and default values[9,see Table 4.7]. For the above comparisons, if no other information is provided, the IPCC default of ±90% confidence interval is used, but converted to one SD by dividing by 1.96[42].

Data Records

The AGBD in natural forests as reported by the IPCC[9,Table 4.7], and estimated by EO in this study, are provided as a Microsoft Excel table in Supplementary Table 1 and presented more concisely in Supplementary Table 2. The global forest status/condition produced published at NASA's Oak Ridge National Laboratory (ORNL) Distributed Active Archive Center (DAAC) with open access to the public[43]. The dataset is provided as Cloud Optimized GeoTIFFs (COGs) in 10°×10° tiles, with unsigned 16-bit values and a spatial resolution of 1 arc-second per pixel (~30 m per pixel at the equator) in EPSG:4326 projection (geographic latitude/longitude coordinates). Data may be viewed and analyzed in any common Geographic Information System (GIS) software that can read and display COGs. The data is also accessible through a SpatioTemporal Asset Catalogs (STAC) endpoint on the NASA-ESA Multi-mission analysis and algorithm platform (MAAP).

Technical Validation

Overview of key results. *Global forest classification.* Our approach to delineating primary, young secondary (≤20 years) and old secondary (>20 years) forests is stringent. It is a conservative combination of many underlying layers (Table 1) with an emphasis on the reduction of potential false positives, thereby reducing the inclusion of ambiguous areas in any class (see Methods, Fig. 1 and Eqs. (1)–(3)). Across the world's natural forests (excluding planted forests), approximately 1678 Mha of primary forests, 1265 Mha of old secondary forests and 316 Mha of young secondary forests were identified (examples in Fig. 2). Of these, primary forests were most abundant in South America (~565 Mha) and Asia (~557 Mha), while young secondary forests were most widespread over Asia (~108 Mha) and North America (~71 Mha) (Fig. 3, and Supplementary Table 1 provides area estimated per class). By ecozones, in line with expectations, tropical rainforests and the boreal mountains/woodlands, included some of the largest fractions of primary forests, while various temperate and subtropical ecozones included the largest fractions of old secondary forests. The largest fractions of young secondary forests were found in tropical shrublands and rainforests, as well as subtropical humid forests. The implications of our choice of method for delineating the world's forests, and the adverse aspects of not validating these delineations due to the unavailability of any suitable validation datasets, are addressed in the Limitations section.

Global estimates of forest AGBD. The global coverage of the two EO-based datasets (i.e. the GEDI/ICESat-2 and ESA CCI Biomass map) allows the estimation of AGBD for every combination of continent, GEZ and forest status/condition, filling gaps in the current IPCC Tier 1 default values (Supplementary Table 1 and 2). The comparisons of these estimates to the IPCC values is, logically, limited to only those forest classes where the latter dataset has values. Generally, the trends of AGBD across all classes are captured well by the EO-datasets; like the IPCC values, the GEDI/ICESat-2 dataset estimates that primary Asian tropical rainforests and mountain systems harbour some of the highest AGBD globally, while the CCI dataset estimates that primary African rainforests harbour the highest AGBD. Both datasets estimate that young secondary tropical/subtropical dry forests, steppe and shrublands harbour the least AGBD on average. Kernel density estimates of footprint/pixel-level mean AGBD across the geographical extents of tropical rainforests and subtropical dry forests are displayed in Fig. 4, and the standard deviation (and range quartiles) for all forest classes are provided in Fig. 5 and Supplementary Table 1. Most classes display a wide range of estimates, which is expected from the vast continental-scale coverage of these ecoregions. A notable observation upon the examination of these ranges, however, is that often the estimated AGBD in young secondary forests approximately equals or exceeds that of old secondary forests. Possible reasons for this are addressed in the Limitations section.

Separately, to examine the statistical differences between the two sets of AGBD, i.e. IPCC values and EO-based estimates, a mixed-effect model was used (allowing inference while accounting for the variance in the compared datasets as well as the model's parameters (Eq. (9))). Here, in the case of the EO-based AGBD estimates, the "variance" is a descriptor of uncertainty in the mean; it refers to the error in a single mean AGBD for a forest class resulting from estimation by modeling with EO data. It is compared to the list of mean AGBD of classes provided by the IPCC and their associated uncertainty[9,see Table 4.7], with the objective of identifying systematic differences (if any) with the data available at hand. Results of the mixed-effect model are discussed for each of the two EO-based estimates (GEDI/ICESat-2 and ESA CCI Biomass) separately.

- **GEDI/ICESat-2 Tier 1 estimates:** Broadly, our model results show that there is not sufficient evidence to indicate that GEDI/ICESat-2 estimates exhibit significant systematic differences to the IPCC values over all forest classes globally ($p > 0.05$). Similarly, for most groups of forest classes (i.e. the classes grouped by either continent, GEZ or forest status/conditions), there is not sufficient evidence to suggest that there are significant systematic differences to IPCC values (Table 2). The exception to this result is the group of classes including young secondary forests, where the GEDI/ICESat-2 estimates are on average 12.08 ± 4.95 Mg/ha higher than those of the IPCC and found to be significantly different ($p = 0.018$, see $\hat{\alpha}$ and associated error in Table 2, estimated from Eq. (9)).
- **ESA CCI Biomass Tier 1 estimates:** Broadly, model results indicate that the ESA CCI Biomass estimates exhibit systematic differences to the IPCC values at the confidence level of 95% ($p = 0.03$), although with an estimated mean difference of only 10.29 ± 4.77 Mg/ha over all forest classes globally. Significant differences are observed when forest classes are grouped in (1) the continents of Africa, Asia and South America, (2) old secondary forests, or in (3) subtropical dry forests and temperate continental forests, although estimates differ

by less than 37 Mg/ha, often with a large associated uncertainties in these groups (see $\hat{\alpha}$ and associated error in these groups, Table 2 and Eq. (9)).

The general global trends mentioned above must be interpreted carefully. Since the EO-based estimates are derived from modeling EO data, they are relatively *precise* owing to the dense samples of lidar footprints in the case of GEDI, and wall-to-wall pixels' coverage of each forest class in the case of the circumboreal ICESat-2 and CCI Biomass map. Such high precision could cause our comparisons to return statistically significant deviations between EO-based and IPCC based values, despite the differences in means being as low as 15 Mg/ha. However, compare also the high precision of the EO-based estimates to the reported IPCC values. Mean AGBD values in the latter are reported only with a standard deviation or confidence interval, which is a best-effort representation of the wide distribution of estimates within classes, compiled from various literature and datasets that may or may not be fully representative of the underlying population[12]. Hence, conversely, for groups with wide intervals of reported uncertainty, there is often insufficient evidence to infer that there are significant differences between IPCC Tier 1 values and EO-based estimates. This interpretation is important; for groups where p-values are insignificant (see Table 2), we cannot draw any definitive conclusions on how the EO-based estimated AGBDs compare to the IPCC Tier 1 defaults.

The individual, class-wise, departures of the EO-based mean AGBD estimates from the IPCC values are displayed in Fig. 6a-b. The wide intervals with which IPCC values are reported are also important when interpreting these differences. For example, the mean GEDI estimates of AGBD in primary African tropical rainforests are approximately 200 Mg/ha lower than those reported by the IPCC (Fig. 6a), but the wide interval of the IPCC value (404.2 ± 120.4 Mg/ha) implies that we cannot conclusively infer that this difference is statically significant ($t = 1.73$, Eq. (10)), visualized in Fig. 6c). In fact, only a handful of forest classes individually exhibit a test statistic (t) that is likely to indicate a significant difference of the GEDI/ICESat-2 estimates to IPCC values. For example, if a significance level of $\alpha = 0.05$ is chosen, only a few areas of fragmented forest classes (such as North American Boreal mountain system of young secondary forests) will exhibit statistically significant deviances. Similarly, between the CCI estimates and IPCC values, mean class-wise differences are low (Fig. 6b) and significant differences are most likely to occur in small areas of fragmented forests in Asia and Oceania (Fig. 6d).

Limitations. Currently, satellite data are well positioned to provide estimates of forest AGBD that are congruent with the IPCC guidelines and the reporting requirements of the UNFCCC. This research, under the CEOS Biomass Harmonization effort, has developed and demonstrated an EO-based framework for providing estimates of IPCC Tier 1 default AGBD for primary, young secondary ($\leq$20 years) and old secondary ($>$20 years) natural forests globally. By using the latest available satellite datasets (Table 1), it establishes the groundwork to provide regular, up-to-date, policy-relevant information on the state of the world's natural forests, adaptable to new datasets as they become available. Below, we first discuss the shortcomings of the underlying forest classes, and then those related to the generation of EO-based Tier 1 AGBD estimates.

Reflection on the global forest classification. The approximately 3.26 billion ha of delineated forest is likely to be a conservative estimate of the true global forested land, resulting from the conservative combination of underlying datasets and layers to acquire a single and, to the extent possible, precise mean AGBD in each IPCC-defined forest class. We present a few examples to illustrate this point. First, although the Global Forest Height dataset[44] (see Table 1) delineates ~4.02 billion ha of forest with height $\geq$5 m, only 91.5% of this area overlaps with the ESA World Cover 2021 (EWC)[45] forest class with >10% tree cover. Our use of a 5 m height threshold together with the ESA tree cover mask (to minimize potential commission errors) hence willfully eliminates ~338 Mha from our classification. This also implies that the inclusion of areas with tree cover <20% (varying in different ecozones) cannot be assured in our classification. Second, our approach also purposefully excludes planted forests, which are treated separately from natural forests in the IPCC guidelines[9,see Table 4.8]. Simply the use of the Spatial Database of Planted Trees (SDPT)[46], not considering the other layers identifying plantations, eliminates 329 Mha of planted trees across the globe from our classification. Third, young growing stands or areas expected to revert to forests, which would fall under the UNFCCC and IPCC definition of "forests" (see Methods, Forest definitions), are also likely to be eliminated because of the lack of reliable underlying datasets/layers that would identify such lands. Similarly, although anthropologically degraded forests defined and identified by the Tropical Moist Forest change map[47] are excluded, naturally disturbed areas (e.g. by fires, windbreaks, drought etc.) may have also been eliminated by this layer (see Table 1). Natural fires in tropical regions are relatively rare in comparison to anthropogenic disturbances[48], and distinguishing the two remains an ongoing EO-related research topic[49–51]. Natural fires in the boreal ecozones may as well be problematic in this regard. Finally, conversely, some areas that may not precisely fall under natural forests may be included in our study. For example, we cannot assure the exclusion of planted trees (e.g. agroforestry) from secondary forests, which are present in the other layers but not identified by the SDPT[46], by the Tropical Moist Forest Transition map[47] or by the Global Forest Management map[52].

Within the lands demarcated as forests in this study, old secondary forests posed the biggest challenge due to the lack of any ancillary datasets that specifically identify this status/condition. Hence, instead, the old secondary forest classes comprise mostly of areas that default to this category because they are neither primary nor young secondary forests (see Methods, Eqs. (1)–(3)). The expected consequence, therefore, is that this group of forest classes may contain a large variation of forest statuses/conditions and may potentially be the least accurate and reliable group. But, this accuracy itself remains untested; despite our analysis presenting a comprehensive compilation of the latest published datasets on forest conditions, the nonexistence of any independent sample of global data (let alone that from a probabilistic distribution), that would enable the validation of these delineations, is a major constraint. Our approach relies purely on the accuracy of the sourced underlying layers,
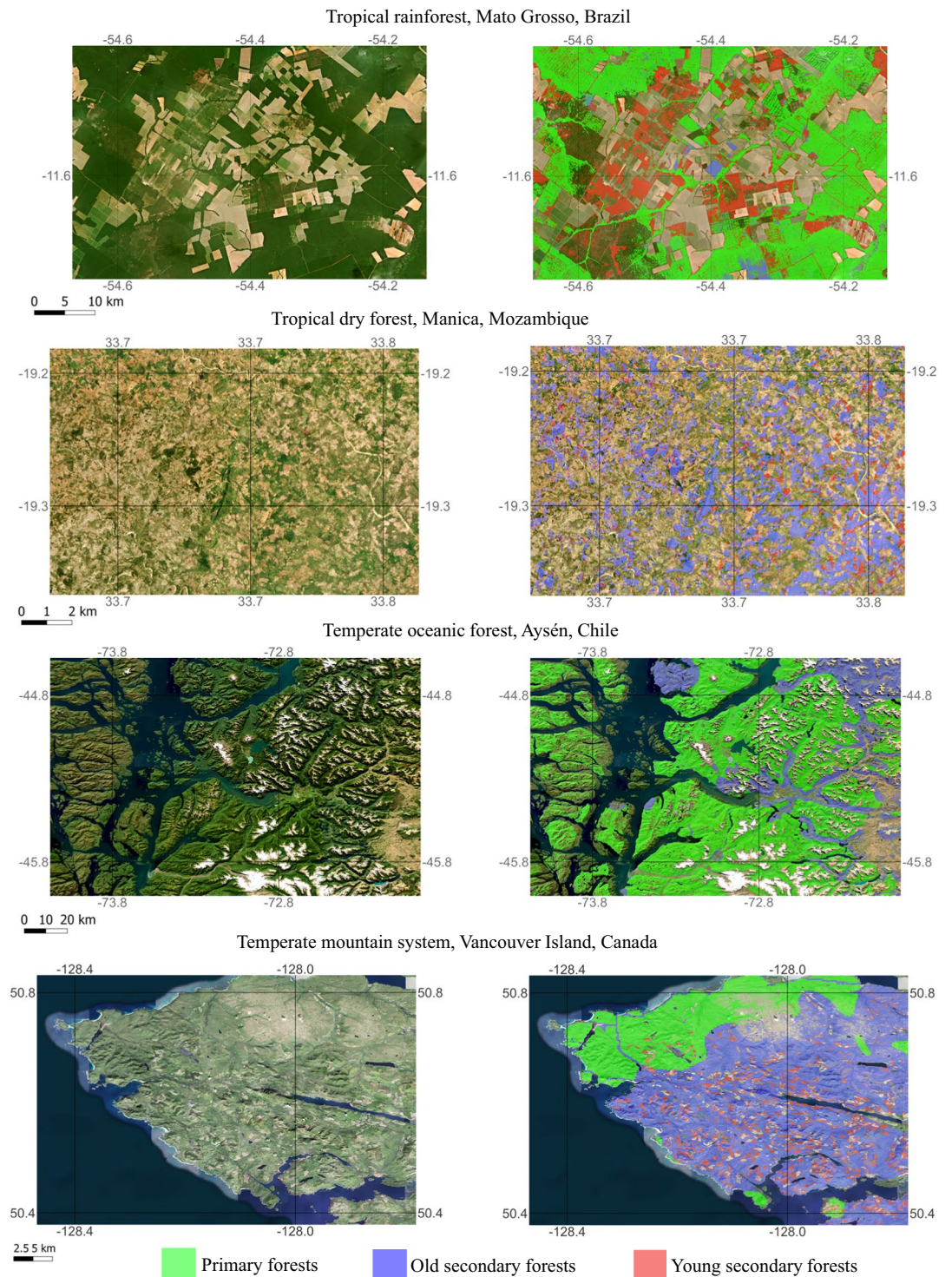
Fig. 2 Visualizations of forest status/condition classes in selected Geographic Ecological Zones (GEZ). The left column displays high resolution optical images (Planet[80] or ESRI World Imagery[81]), overlaid by the forest classes delineated in this study in the right column. Mato Grosso (Brazil) is dominated by cattle ranches and large-scale agriculture, Manica province (Mozambique) is dominated by shifting agriculture in dry *miombo* woodlands, Aysén (Chile) is dominated by undisturbed temperate evergreen and deciduous forests, and Vancouver Island (Canada) is dominated by areas of active logging and old secondary forests, as well as vast expanses of temperate old-growth forests.

strengthened by the condition that multiple layers must affirmatively indicate that a forest is either primary, young secondary or default to the old secondary group. Worth noting is that even the broader definitions of

(a)

(b)
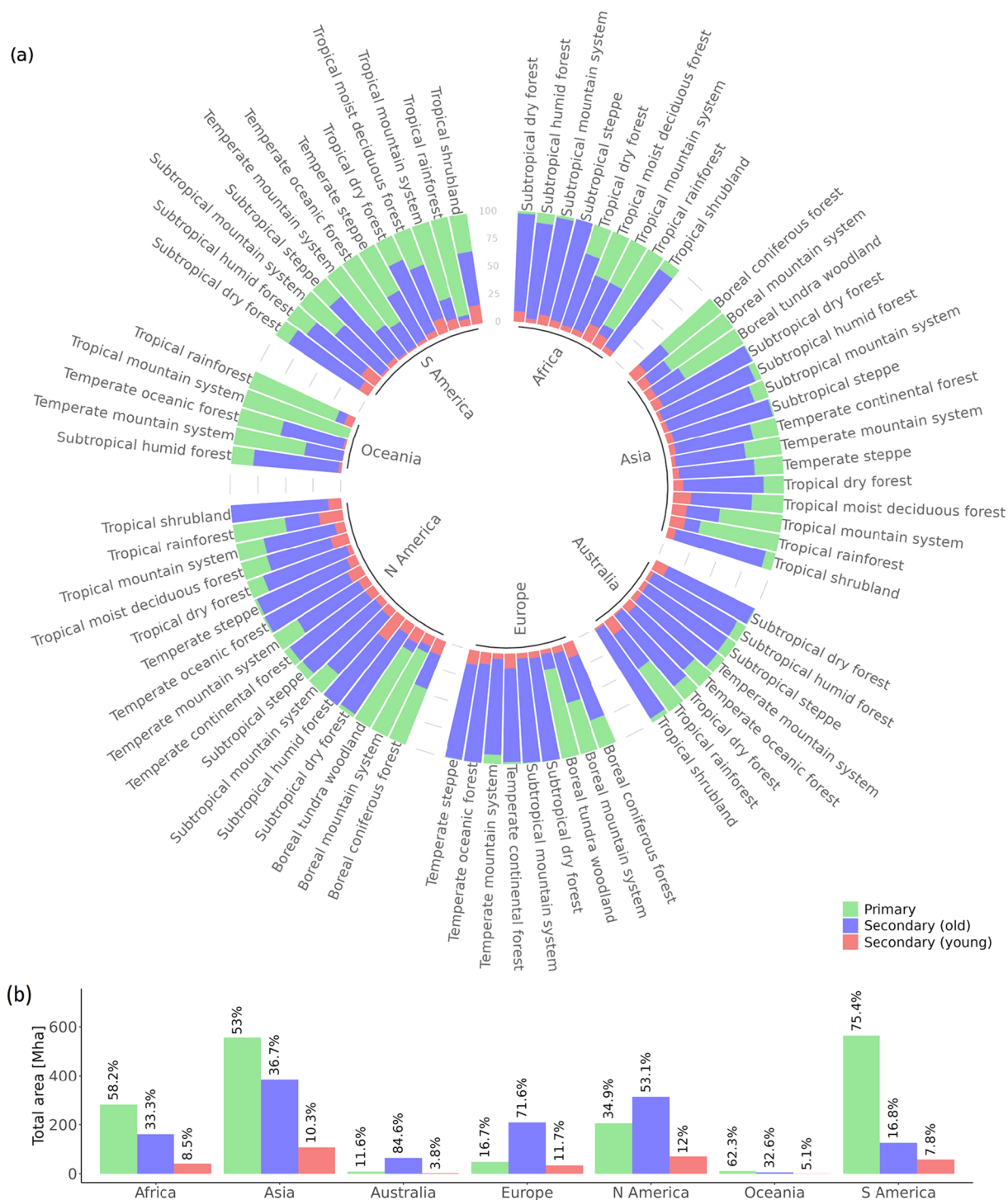
Fig. 3 (**a**) Relative area (as % of total area) of forest status/condition classes in each Geographic Ecological Zone (GEZ) and continent. (**b**) Total and relative area of each forest status/condition class by continent.

these classes remain debated in literature; for example, whether forests are "primary" may relate to their biodiversity and ecological function, structural complexity, endemism, age, and/or the lack of habitat fragmentation, which in in turn influences which and how EO-based biophysical predictors may be used to identify these forests[53]. Currently, the lack of training or validation data precludes the rigorous testing of conformation to any such definitions, the testing of the accuracy of the forest classification (i.e. how representative the classes are of true global forest conditions), or a sensitivity analysis of the subsequent EO-based estimated AGBD to changing forest conditions/statuses.
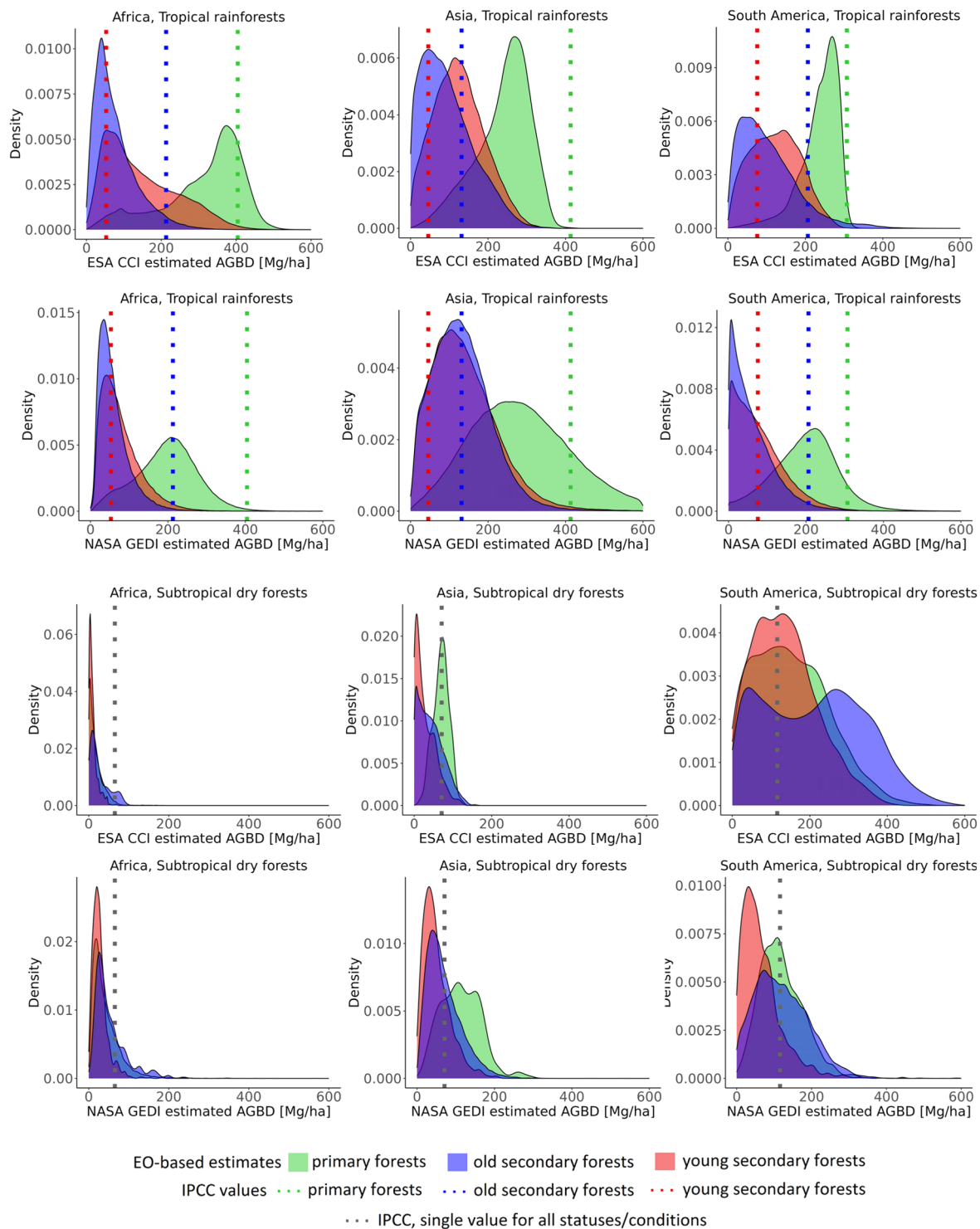
Fig. 4 Kernel density estimates of mean aboveground biomass density (AGBD) at footprint/pixel-level from EO across the geographical extents of tropical rainforests and subtropical dry forests of Africa, Asia and South America (interquartile ranges are provided in Supplementary Table 1). Mean IPCC Tier 1 AGBD values are shown for reference.

Reflection on IPCC and EO-based Tier 1 AGBD estimates. Several challenges were faced simply in the presentation of this first iteration of EO-based Tier 1 AGBD estimates. Some of these challenges stem from limitations in the current formulation and values of the IPCC Tier 1 table[9,Table 4.7], which complicate their comparison to any newer estimates. First, as reported in the overview of key results, the wide confidence intervals with which the IPCC Tier 1 values are currently reported can render tests to identify systematically significant departures to the EO-based datasets insufficient. The large variability in the population, and the unassessed representativeness
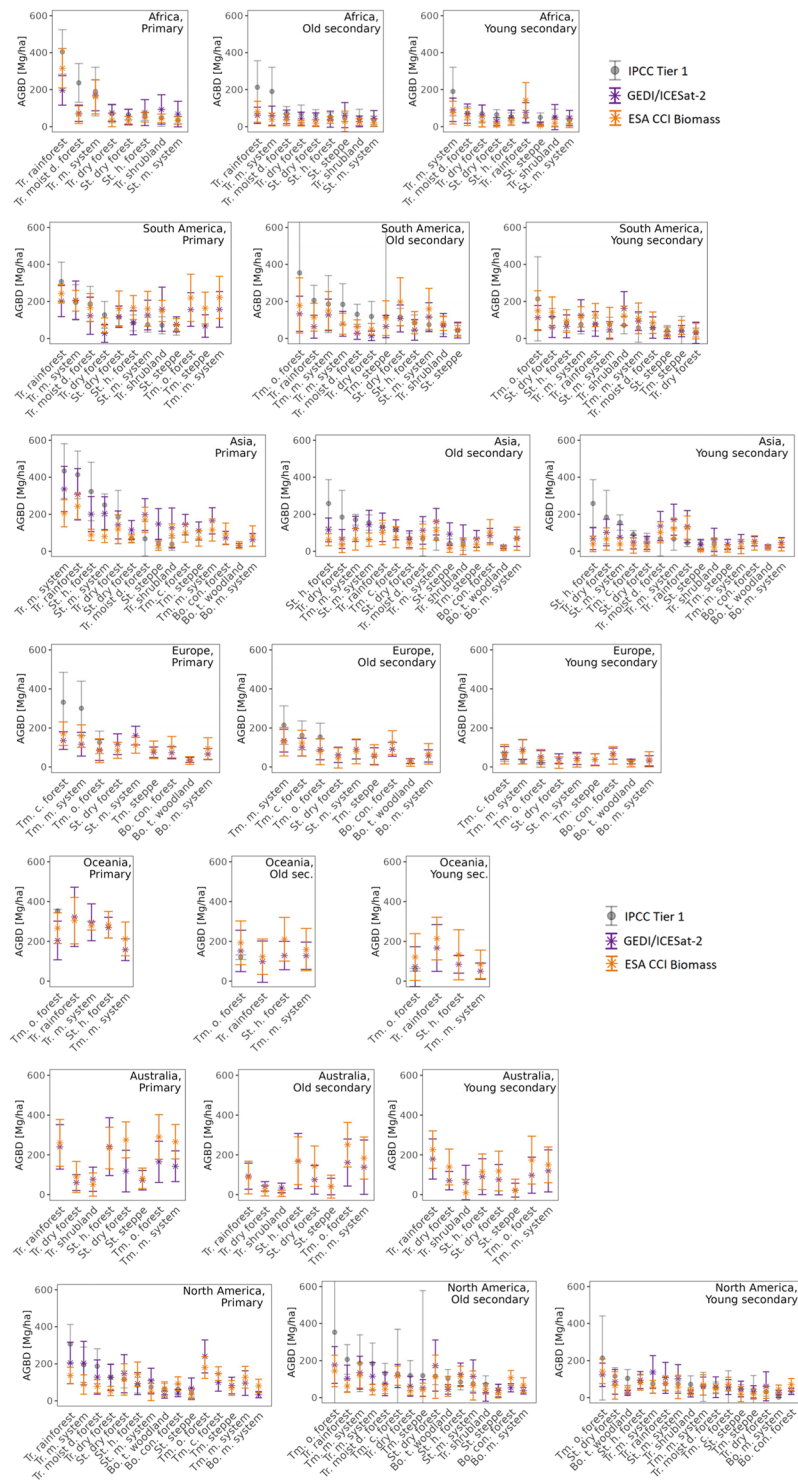
Fig. 5 IPCC Tier 1 default and Earth Observation (EO) estimates of mean aboveground biomass density (AGBD). Bars denote standard deviation reported for the IPCC values, and standard deviation of footprint/pixel-level mean AGBD estimated from EO across the geographical extents of each ecozone. Abbreviations are as follows: Primary="P", Young Secondary ="YS", Old Secondary = "OS", tropical="Tr.", subtropical="St.", temperate="Tm.", boreal="Bo.", mountain="m.", deciduous="d.", continental="c.", coniferous="con.", oceanic="o.", tundra="t." and humid="h.".Abbreviations are as follows: Primary="P", Young Secondary ="YS", Old Secondary = "OS", tropical="Tr.", subtropical="St.", temperate="Tm.", boreal="Bo.", mountain="m.", deciduous="d.", continental="c.", coniferous="con.", oceanic="o.", tundra="t." and humid="h.".

of limited samples used to derive the Tier 1 defaults, implies that our statements on systematic deviations of EO-based estimates must be interpreted with caution; they are a best effort of a global-scale comparison based

Group	N	GEDI and ICESat-2				ESA CCI			
All classes		$\hat{\alpha}$	SE[$\hat{\alpha}$]	t	p	$\hat{\alpha}$	SE[$\hat{\alpha}$]	t	p
—	140	−1.66	4.46	−0.37	0.71	−10.29	4.77	−2.16	0.03*
Continent		$\hat{\alpha}$	SE[$\hat{\alpha}$]	t	p	$\hat{\alpha}$	SE[$\hat{\alpha}$]	t	p
Africa	26	−4.41	7.77	−0.57	0.575	−22.04	6.19	−3.56	0.002*
Asia	30	10.09	10.12	1.00	0.327	−31.94	11.23	−2.84	0.008*
Europe	9	15.88	9.07	1.75	0.118	22.66	10.22	2.22	0.057
North America	39	5.29	6.93	0.76	0.450	−9.35	8.03	−1.16	0.252
Oceania	3	−55.02	60.18	−0.91	0.457	16.34	42.05	0.39	0.735
South America	33	−1.19	8.82	−0.13	0.894	22.70	7.85	2.89	0.007*
Forest status/condition		$\hat{\alpha}$	SE[$\hat{\alpha}$]	t	p	$\hat{\alpha}$	SE[$\hat{\alpha}$]	t	p
Primary	40	−3.46	11.58	−0.30	0.767	−11.29	9.83	−1.15	0.258
Old secondary	50	−5.97	5.62	−1.06	0.294	−22.83	8.15	−2.80	0.007*
Young secondary	50	12.08	4.95	2.44	0.018*	−0.04	7.06	−0.01	0.996
Geographic Ecological Zones		$\hat{\alpha}$	SE[$\hat{\alpha}$]	t	p	$\hat{\alpha}$	SE[$\hat{\alpha}$]	t	p
Tropical dry forest	12	−9.81	14.86	−0.66	0.523	−19.61	13.41	−1.46	0.172
Tropical moist deciduous forest	12	−13.61	14.22	−0.96	0.359	−15.25	13.76	−1.11	0.292
Tropical mountain system	12	25.00	25.92	0.96	0.355	−11.99	22.92	−0.52	0.611
Tropical rainforest	12	−0.36	15.13	−0.02	0.982	−10.46	25.04	−0.42	0.684
ropical shrubland	11	25.39	13.53	1.88	0.090	4.12	12.47	0.33	0.748
Subtropical dry forest	12	−1.90	11.26	−0.17	0.869	−22.20	9.51	−2.34	0.039*
Subtropical humid forest	12	−0.85	11.11	−0.08	0.940	−5.00	9.79	−0.51	0.619
Subtropical mountain system	12	15.30	12.33	1.24	0.241	−18.33	16.83	−1.09	0.299
Subtropical steppe	11	13.91	10.24	1.36	0.204	−16.25	7.65	−2.12	0.060
Temperate continental forest	7	−2.24	9.84	-0.23	0.828	−36.86	7.94	−4.64	0.004*
Temperate mountain system	8	−32.91	30.20	−1.09	0.312	−22.75	33.13	−0.69	0.514
Temperate oceanic forest	10	−49.60	33.14	−1.50	0.169	0.08	26.74	0.00	0.998
Temperate steppe	4	−12.23	56.70	−0.22	0.843	−4.63	53.50	−0.09	0.937
Boreal tundra woodland	3	−47.63	22.87	−2.08	0.173	−22.68	22.49	−1.01	0.419

Table 2. Results of mixed-effects models that compare the IPCC Tier 1 default to the EO-based mean aboveground biomass density (AGBD), while filtering the variance associated with each estimate in both datasets. Results are provided for all forest classes, and separately per continent, per forest status/condition and per Geographic Ecological Zone (GEZ). N is the number of samples, α is the estimate of the intercept, and SE is the standard error of α (Eq. (9)). Values of $p < 0.05^*$ indicate a significant systematic difference between IPCC values and EO-based estimates.

on the data at hand, which in some classes have untraceable sources and/or estimation approaches[10,12]. Second, and more importantly, while the comparison of IPCC Tier 1 defaults to EO-based estimates highlights ecozones with the largest discrepancies, it cannot be interpreted as a validation of any dataset; the *accuracy*, i.e. the closeness to the true value of global forest AGBD, even of the IPCC values, remains unknown and untested without an independent sample of reference data from ground-measured plots. These shortcomings are further compounded if all lands with at least 10% tree cover canopy[9,see footnote of Table 4.7] have not been included for all forest classes, nor the true population probabilistically sampled, for the generation of IPCC Tier 1 AGBD values. In fact, regardless of how diligently any new estimates (whether EO-based or not) attempt to generate Tier 1 estimates, any comparisons to current IPCC values will inevitably be hindered by these inherent limitations.

Similarly, there are limitations in the EO-based AGBD estimates; the two EO-based datasets used in this study are known to themselves have disagreements, both in their estimates of AGBD and associated variance, and in their uncertainty assessment frameworks. An in-depth comparison of the two is outside the scope of this work, and interested readers are referred to Hunka *et al.*[4]. Here, we point out limitations relevant to the manner in which global forests are classified in this study. Take, for example, the observation of nearly equal or higher estimates in young secondary forests compared to old secondary forests in some ecoregions (seen in Fig.4). This could be the result of the lack of accuracy of the demarcated classes (particularly old secondary forests), or that of the fact that a 20-year difference is too short a duration in the life-cycle of forest recovery after disturbance[54]. But, it could also be a result of the systematic overestimation of EO-based AGBD known to occur in forests with low AGBD ranges[38,39]. For biomass maps based on satellite radar, for example, the influence of soil moisture and surface roughness can cause models to overestimate AGBD[55]. This, in turn, might have resulted in systematically higher AGBD in young secondary tropical rainforests in the ESA CCI Biomass map. Such drawbacks must be viewed under the context that newly available EO-based AGBD estimates posit not only relatively new scientific advancements, but also expose exploratory findings and previously unknown challenges in biomass estimation from space. We take another example of the GEDI AGBD estimates to elaborate this point. The GEDI mission was designed around a statistical framework to enable both precise and accurate estimates of AGBD
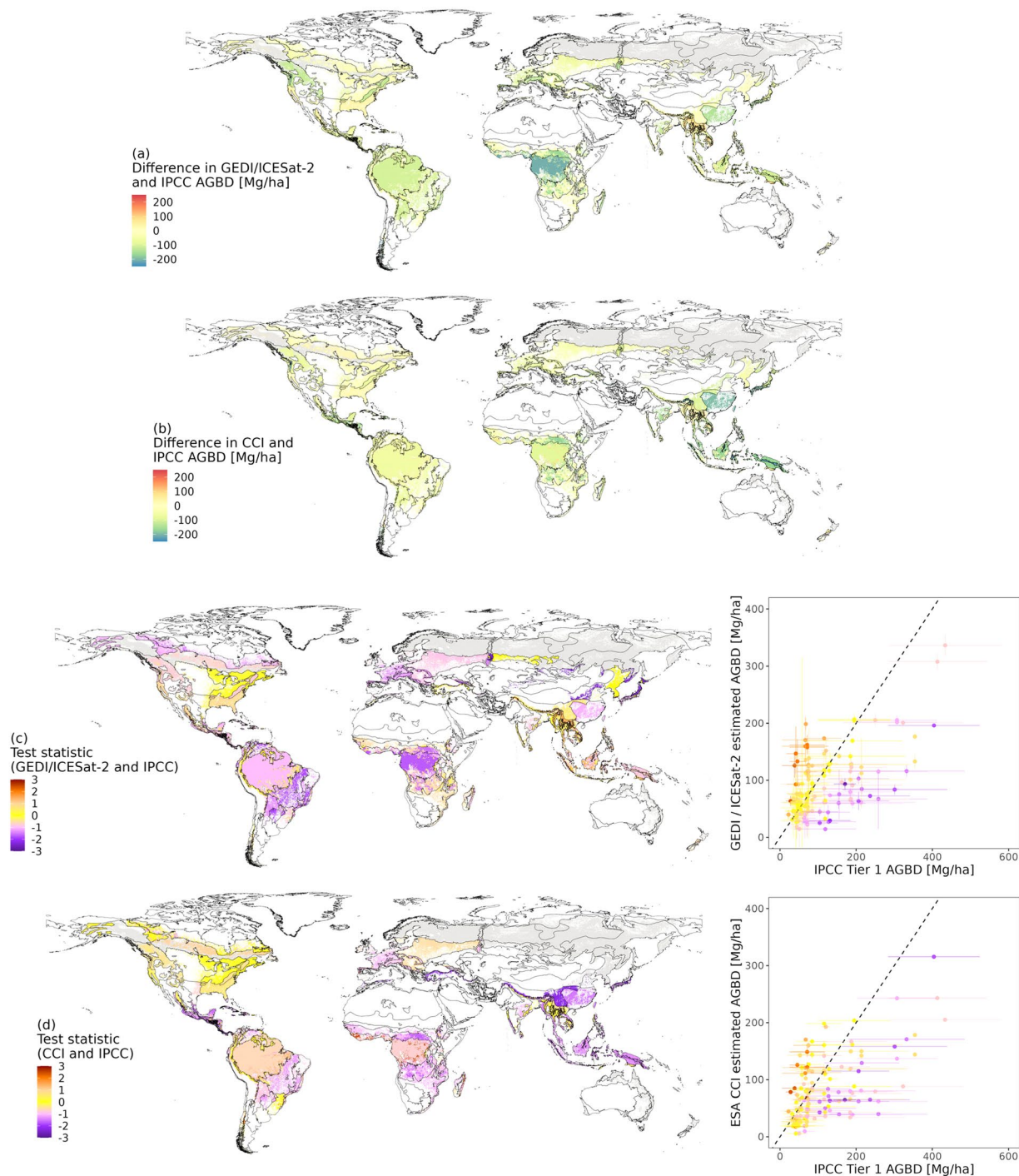
Fig. 6 (**a**) GEDI/ICESat-2 and (**b**) ESA CCI estimates minus the IPCC Tier 1 default values of aboveground biomass density (AGBD) for each combination of continent, forest status/condition and Geographic Ecological Zone (GEZ[21], black boundaries). (**c-d**) Spatial distribution of the estimated test statistic (*t*, Eq. (10)), i.e. ratio of difference in the mean estimates of AGBD to the total variance of the two datasets. Grey values on the maps show areas classified as forests in this study, for which IPCC Tier 1 defaults are missing[9,Table 4.7]. Scatter plots show mean AGBD estimates of the datasets, with error bars denoting standard deviation (in the case of IPCC), standard error on mean (in the case of EO estimates) and colours corresponding to *t*.

(for areas >1 km) through its use of hybrid estimation[36,56]. In incorporating multiple sources of uncertainty in its variance estimation approach, it is congruent with the various components of uncertainty covered by the IPCC guidelines[57]. However, Dubayah *et al.* (2022) caution that the accuracy of their mean estimates rests on the assumption that model-training data apply to the area where AGBD is being predicted, and note that as new training data allows model refinement, AGBD confidence intervals may or may not overlap with those of

previous estimates. Recently, Ståhl *et al.* (2024) have argued that EO-based estimates of forest AGBD may be, paradoxically, *model*-unbiased but *design*-biased, i.e. they may be unbiased compared to their *expected* values but biased compared to *true*, and fixed but unknown, values of forest AGBD[58].

Discussions such as these notably highlight the continuing research effort towards the improvement of AGBD estimation from remotely sensed data. The differences in the EO-based AGBD estimates do not render them inapt for policy by default; instead, they have spurred dedicated research on optimizing their accuracies[13,38,59], integrating them with national forest inventories[60–62], and encouraged transparency and harmonizing approaches in their dissemination for policy-relevant uses[4,5]. As aforementioned, it is not currently possible to assess the relative improvement in EO-based AGBD over time without a globally representative biomass reference network, and thus this limits the scope of current efforts. However, several efforts are working toward a fiducial reference system for biomass (e.g. GEO-TREES[25,26]), which would enable a step change toward true product validation and thus accelerate progress in AGBD mapping and associated uptake for policy applications. Recommendations in this regard are provided in the Future Work and Dataset Iterations section.

Future work and dataset iterations. The analyses presented in the article focus on replicating a rather prescriptive delineation of global natural forests that matches the classes in the IPCC Tier 1 defaults[9]. This need not be a restriction for subsequent research. In future work, the EO-based Tier 1 AGBD estimates must be routinely updated with new and incoming data, and can be further refined to country-level estimates with nationally-delineated ecozones, initiating practicality and consistency for use in various National Greenhouse Gas Inventories (NGHGIs). The current limitations of the underlying global forest classification stem primarily from area mismatches between the various input datasets, or simply the lack of datasets that demarcate naturally disturbed areas, sparse vegetation or old secondary forests. What may be done to overcome these limitations if policy reporting continues to demand the use of similarly prescriptive forest classes?

Besides the improvements in the accuracy and algorithms to produce the various input datasets, the critical constraint of an absence of independent global validation datasets must be overcome. The launches of new satellite missions (ESA BIOMASS[23] and NISAR[24]), and the expectation of improved EO-based AGBD estimates with their data *alongside* a globally-distributed forest biomass reference system through GEO-TREES[25,26], hold some promise that validation datasets may be available in the near future. To support this, the open-science framework presented in this study may offer an additional opportunity - areas of disagreement between the various input datasets can be identified and targeted for validation, or areas with sparse vegetation or uncertain classes of old secondary forests can be prioritized in ground reference data collection. Open access to the *in situ* plot reference datasets is crucial[27], it should contain information on observed forest ages, disturbance history and estimated AGBD across a range of ecological regions, and be used with consistent validation protocols[63,64]. Such reference information will be a valuable dataset for validating the EO-based IPCC Tier 1 estimates, and subsequently contribute to the generation of other policy-relevant datasets. As and when input datasets or AGBD estimates improve and align, future work may separately present EO-based estimation of biomass in the ~800 Mha of the remainder of the world's forests, hence providing a holistic assessment that covers the entirety of the ~4 billion ha of forests identified globally[65]. Such an effort would also initiate alignment with the 2020 Global Forest Resources Assessments of the UN Food and Agriculture Organization (FAO)[65], hence releasing a more consistent and clear overview of space-based AGBD estimates to the public.

Under the CEOS Biomass Harmonization objectives, we expect that regularly updated and validated EO-based datasets will effectively aid nations and policy, and can potentially contribute to UNFCCC processes[6,66]. This is especially relevant as newer datasets are likely to capture ongoing biome shifts associated with climate change. With this foresight, the methodology and data processing for the forest classification produced in this study are designed in a simple, open-science, Boolean scheme that may be revised as new versions of any underlying maps or new geospatial datasets are publicly released (see Code Availability and associated scripts for full details). Transparency and the replicability of this methodology are given utmost importance; not only does it enable trust in the global forest classification but also allows the receipt of feedback on potential areas of improvement for better alignment with scientific and policy needs. Furthermore, as new EO-based AGBD map estimates are generated, or if the IPCC were to adopt new global ecological zoning[67], open access to the framework and codes for generation of forest age classifications eases computation, and thereby, the presentation of the new EO-based AGBD estimates in line with IPCC guidelines. The presented framework and results of this article are a pioneering international effort from CEOS, presenting AGBD maps in a format practical for policy and potentially adoptable, upon review, in the IPCC Emissions Factors Database[22].

Code availability

Data was processed on the NASA-ESA Multi-mission Algorithm and Analysis Platform (MAAP), which has public code repositories. The associated code for this article can be accessed in the Biomass Harmonization folder and is published on Zenodo[68]. A "README.txt" file is provided as an introduction to the steps followed for the classification of global forests. For a single example tile extent of 10 × 10 degrees (tile 00N_000E), the web-links of all input data sources and their pre-processing (resampling, gridding and alignment) is provided. A script for their classification into primary, young secondary and old secondary forests is provided separately. Batch processing is supported by Amazon Web Services (AWS) on the NASA MAAP. On local machines (without cloud-computing capabilities), single tile extents of 10 × 10 degrees can be run in sequence following the single-tile example provided in the repository.

References

1. UNFCCC. Global Stocktake. Tech. Rep. https://unfccc.int/topics/global-stocktake (2023).
2. UNFCCC. Technical dialogue of the first global stocktake. Synthesis report by the co-facilitators on the technical dialogue. Tech. Rep. https://unfccc.int/documents/631600 (2023).
3. Melo, J., Baker, T., Nemitz, D., Quegan, S. & Ziv, G. Satellite-based global maps are rarely used in forest reference levels submitted to the UNFCCC. *Environmental Research Letters* **18**, 034021, https://doi.org/10.1088/1748-9326/acba31 (2023).
4. Hunka, N. *et al*. On the NASA GEDI and ESA CCI biomass maps: aligning for uptake in the UNFCCC global stocktake. *Environmental Research Letters* **18**, 124042, https://doi.org/10.1088/1748-9326/ad0b60 (2023).
5. Ochiai, O. *et al*. Towards a roadmap for space-based observations of the land sector for the UNFCCC global stocktake. *iScience* **26**, 106489, https://doi.org/10.1016/j.isci.2023.106489 (2023).
6. Poulter, B. *et al*. CEOS Roadmap for Space-Based Support of Agriculture, Forestry and Other Land Use (AFOLU) Emissions and Removals of Greenhouse Gases (V1.0). CEOS (2023).
7. NASA. Biomass Earthdata Dashboard BETA: The Biomass Harmonization Activity (2021).
8. IPCC. 2006 IPCC Guidelines for National Greenhouse Gas Inventories, Volume 4: Agriculture, Forestry and Other Land Use (2006).
9. IPCC. 2019 Refinement to the 2006 IPCC Guidelines For National Greenhouse Gas Inventories. Chapter 4: Forest Land https://www.ipcc-nggip.iges.or.jp/public/2019rf/pdf/4_Volume4/19R_V4_Ch04_Forest%20Land.pdf. (2019).
10. Rozendaal, D. M. A. *et al*. Aboveground forest biomass varies across continents, ecological zones and successional stages: refined ipcc default values for tropical and subtropical forests. *Environmental Research Letters* **17**, 014047, https://doi.org/10.1088/1748-9326/ac45b3 (2022).
11. Santoro, M. GlobBiomass - global datasets of forest biomass, https://doi.org/10.1594/PANGAEA.894711 (2018).
12. Langner, A., Achard, F. & Grassi, G. Can recent pan-tropical biomass maps be used to derive alternative Tier 1 values for reporting REDD+ activities under UNFCCC? *Environmental Research Letters* **9**, 124008, https://doi.org/10.1088/1748-9326/9/12/124008 (2014).
13. Dubayah, R. *et al*. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth's forests and topography. *Science of Remote Sensing* **1**, 100002, https://doi.org/10.1016/j.srs.2020.100002 (2020).
14. Dubayah, R. *et al*. GEDI L4B Gridded Aboveground Biomass Density, Version 2.1, https://doi.org/10.3334/ORNLDAAC/2299 (2023).
15. Markus, T. *et al*. The Ice, Cloud, and land Elevation Satellite-2 (ICESat-2): Science requirements, concept, and implementation. *Remote Sensing of Environment* **190**, 260–273, https://doi.org/10.1016/j.rse.2016.12.029 (2017).
16. Claverie, M. *et al*. The Harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sensing of Environment* **219**, 145–161, https://doi.org/10.1016/j.rse.2018.09.002 (2018).
17. ESA. Copernicus DEM - Global and European Digital Elevation Model (COP-DEM). https://doi.org/10.5270/ESA-c5d3d65 (2022).
18. Duncanson, L. *et al*. Aboveground biomass density for high latitude forests from icesat-2, 2020, https://doi.org/10.3334/ORNLDAAC/2186 (2023).
19. Santoro, M. & Cartus, O. ESA Biomass Climate Change Initiative (Biomass_cci): Global datasets of forest above-ground biomass for the years 2010, 2017, 2018, 2019 and 2020, v4. https://catalogue.ceda.ac.uk/uuid/af60720c1e404a9e9d2c145d2b2ead4e (2023).
20. Santoro, M., Cartus, O., Kay, H., Lucas, R. & Quegan, S. Algorithm Theoretical Basis Document (ATBD), Year 4, Version 4.0 (European Space Agency, 2023).
21. FAO. GEONETWORK, Global ecological zones (2nd ed.). https://data.apps.fao.org/catalog/iso/2fb209d0-fd34-4e5e-a3d8-a13c241eb61b (2012).
22. IPCC. EFDB: Emission Factor Database. https://www.ipcc-nggip.iges.or.jp/EFDB/main.php (2022).
23. Quegan, S. *et al*. The European Space Agency BIOMASS mission: Measuring forest above-ground biomass from space. *Remote Sensing of Environment* **227**, 44–60, https://doi.org/10.1016/j.rse.2019.03.032 (2019).
24. NASA-JPL. NASA-ISRO SAR (NISAR) Mission Science Users' Handbook. https://nisar.jpl.nasa.gov/documents/26/NISAR_FINAL_9-6-19.pdf (2019).
25. Labrière, N. *et al*. Toward a forest biomass reference measurement system for remote sensing applications. *Global Change Biology* **29**, 827–840, https://doi.org/10.1111/gcb.16497 (2023).
26. GEO-TREES. Geo-trees initiative. https://geo-trees.org/project/ (2023).
27. Liang, J. & Gamarra, J. G. P. The importance of sharing global forest data in a world of crises. *Scientific Data* **7**, https://doi.org/10.1038/s41597-020-00766-x (2020).
28. IPCC. 2019 Refinement to the 2006 IPCC Guidelines For National Greenhouse Gas Inventories. Chapter 3: Consistent Representation of Lands https://www.ipcc-nggip.iges.or.jp/public/2019rf/pdf/4_Volume4/19R_V4_Ch03_Land%20Representation.pdf (2019).
29. IPCC. Good Practice Guidance for Land Use, Land-Use Change and Forestry (GPG-LULUCF), Chapter 4: Supplementary Methods and Good Practice Guidance Arising from the Kyoto Protocol. https://www.ipcc-nggip.iges.or.jp/public/gpglulucf/gpglulucf_files/Chp4/Chp4_1_to_24.pdf (2003).
30. Santoro, M. *et al*. The global forest above-ground biomass pool for 2010 estimated from high-resolution satellite observations. *Earth System Science Data* **13**, 3927–3950, https://doi.org/10.5194/essd-13-3927-2021 (2021).
31. Sims, M. *et al*. Global Forest Watch: Differences Between Global Forest Watch's Tree Cover Loss Data and JRC's Tropical Moist Forest Data Explained. https://www.globalforestwatch.org/blog/data-and-research/tree-cover-loss-and-tropical-moist-forest-data-compared/ (2024).
32. Duncanson, L. *et al*. Aboveground biomass density models for NASA's Global Ecosystem Dynamics Investigation (GEDI) lidar mission. *Remote Sensing of Environment* **270**, 112845, https://doi.org/10.1016/j.rse.2021.112845 (2022).
33. Kellner, J. R., Armston, J. & Duncanson, L. Algorithm Theoretical Basis Document for GEDI Footprint Aboveground Biomass Density. *Earth and Space Science* **10**, e2022EA002516, https://doi.org/10.1029/2022EA002516 (2023).
34. Healey, S. P., Patterson, P. L. & Armston, J. Algorithm Theoretical Basis Document (ATBD) for GEDI Level-4B (L4B) Gridded Aboveground Biomass Density. Version 2.1. https://daac.ornl.gov/daacdata/gedi/GEDI_L4B_Gridded_Biomass_V2_1/comp/GEDI_L4B_ATBD_V2.0.pdf (ORNL Distributed Active Archive Center, 2023).
35. Patterson, P. L. *et al*. Statistical properties of hybrid estimators proposed for GEDI -NASA's global ecosystem dynamics investigation. *Environmental Research Letters* **14**, 065007, https://doi.org/10.1088/1748-9326/ab18df (2019).
36. Dubayah, R. *et al*. GEDI launches a new era of biomass inference from space. vol. 17, 095001, https://doi.org/10.1088/1748-9326/ac8694 (IOP Publishing, 2022b).
37. Armston, J. *et al*. GEDI L4B Country-Level Summaries of Aboveground Biomass, Version 1, https://doi.org/10.3334/ORNLDAAC/2321 (2023).
38. Araza, A. *et al*. A comprehensive framework for assessing the accuracy and uncertainty of global above-ground biomass maps. *Remote Sensing of Environment* **272**, 112917, https://doi.org/10.1016/j.rse.2022.112917 (2022).
39. Santoro, M. *et al*. Design and performance of the climate change initiative biomass global retrieval algorithm. *Remote Sensing of Environment* **10**, 100169, https://doi.org/10.1016/j.srs.2024.100169 (2024).
40. Fay III, R. E. & Herriot, R. A. Estimates of income for small places: an application of James-Stein procedures to census data. *Journal of the American Statistical Association* **74**, 269–277 (1979).
41. Rao, J. N. & Molina, I. Small area estimation. https://doi.org/10.1002/9781118735855 (John Wiley & Sons, 2015).

42. Harris, N. *et al*. Global maps of twenty-first century forest carbon fluxes. *Nature Climate Change* **11**, 234–240, https://doi.org/10.1038/s41558-020-00976-6 (2021).

43. Hunka, N. *et al*. Classification of global forests for IPCC aboveground biomass Tier 1 estimates, 2020, *ORNL DAAC, Oak Ridge, Tennessee, USA*, https://doi.org/10.3334/ORNLDAAC/2345 (2024).

44. Potapov, P. *et al*. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sensing of Environment* **253**, 112165, https://doi.org/10.1016/j.rse.2020.112165 (2021).

45. Zanaga, D. *et al*. ESA WorldCover 10 m 2021 v200, https://doi.org/10.5281/zenodo.7254221 (2022).

46. Richter, J. *et al*. Spatial database of Planted Trees (SDPT version 2.0). Technical Note. https://doi.org/10.46830/writn.23.00073 (2024).

47. Vancutsem, C. *et al*. Long-term (1990–2019) monitoring of forest cover changes in the humid tropics. *Science Advances* **7**, eabe1603, https://doi.org/10.1126/sciadv.abe1603 (2021).

48. Tyukavina, A. *et al*. Global trends of forest loss due to fire from 2001 to 2019. Frontiers in Remote Sensing **3**, https://doi.org/10.3389/frsen.2022.825190 (2022).

49. Long, T. *et al*. 30 m resolution global annual burned area mapping based on landsat images and google earth engine. Remote Sensing **11**, https://doi.org/10.3390/rs11050489 (2019).

50. Sebald, J., Senf, C. & Seidl, R. Human or natural? landscape context improves the attribution of forest disturbances mapped from landsat in central europe. *Remote Sensing of Environment* **262**, 112502, https://doi.org/10.1016/j.rse.2021.112502 (2021).

51. Slagter, B. *et al*. Monitoring direct drivers of small-scale tropical forest disturbance in near real-time with sentinel-1 and -2 data. *Remote Sensing of Environment* **295**, 113655, https://doi.org/10.1016/j.rse.2023.113655 (2023).

52. Lesiv, M. *et al*. Global forest management data for 2015 at a 100 m resolution. *Scientific Data* **9**, 199, https://doi.org/10.1038/s41597-022-01332-3 (2022).

53. Bruening, J. M., Dubayah, R. O., Pederson, N., Poulter, B. & Calle, L. Definition criteria determine the success of old-growth mapping. *Ecological Indicators* **159**, 111709, https://doi.org/10.1016/j.ecolind.2024.111709 (2024).

54. Cole, L. E. S., Bhagwat, S. A. & Willis, K. J. Recovery and resilience of tropical forests after disturbance. Nature Communications **5**, https://doi.org/10.1038/ncomms4906 (2014).

55. Santoro, M. *et al*. Retrieval of growing stock volume in boreal forest using hyper-temporal series of Envisat ASAR ScanSAR backscatter measurements. *Remote Sensing of Environment* **115**, 490–507, https://doi.org/10.1016/j.rse.2010.09.018 (2011).

56. Saarela, S. *et al*. Comparing frameworks for biomass prediction for the global ecosystem dynamics investigation. *Remote Sensing of Environment* **278**, 113074, https://doi.org/10.1016/j.rse.2022.113074 (2022).

57. IPCC. 2019 Refinement to the 2006 IPCC Guidelines For National Greenhouse Gas Inventories. Chapter 3: Uncertainties. https://www.ipcc-nggip.iges.or.jp/public/2019rf/pdf/1_Volume1/19R_V1_Ch03_Uncertainties.pdf (2019).

58. Ståhl, G. *et al*. Why ecosystem characteristics predicted from remotely sensed data are unbiased and biased at the same time - and how this affects applications. *Forest Ecosystems* **11**, 100164, https://doi.org/10.1016/j.fecs.2023.100164 (2024).

59. Emick, E. *et al*. An approach to estimating forest biomass while quantifying estimate uncertainty and correcting bias in machine learning maps. *Remote Sensing of Environment* **295**, 113678, https://doi.org/10.1016/j.rse.2023.113678 (2023).

60. Málaga, N. *et al*. Precision of subnational forest AGB estimates within the Peruvian Amazonia using a global biomass map. *International Journal of Applied Earth Observation and Geoinformation* **115**, 103102, https://doi.org/10.1016/j.jag.2022.103102 (2022).

61. May, P., McConville, K. S., Moisen, G. G., Bruening, J. & Dubayah, R. A spatially varying model for small area estimates of biomass density across the contiguous united states. *Remote Sensing of Environment* **286**, 113420, https://doi.org/10.1016/j.rse.2022.113420 (2023).

62. Hunka, N. *et al*. National forest biomass assessments enhanced with earth observation to aid climate policy needs. https://doi.org/10.2139/ssrn.4910141 (2024).

63. Duncanson, L. *et al*. The importance of consistent global forest aboveground biomass product validation. *Surv. Geophys.* **40**, 979–999 (2019).

64. Duncanson, L. *et al*. Aboveground woody biomass product validation : good practices protocol (Land Product Validation Subgroup (WGCV/CEOS), 2021).

65. FAO. 2020 Global forest resources assessment Main Report (Rome: FAO). www.fao.org/documents/card/en/c/ca9825en (2020).

66. CEOS. Committee on Earth Observation Satellites (CEOS) Greenhouse Gas Roadmap. https://ceos.org/news/ghg-roadmap/ (2023).

67. Audebert, P. *et al*. Ecological zoning for climate policy and global change studies. Nature Sustainability https://doi.org/10.1038/s41893-024-01416-5 (2024).

68. Hunka, N., Shinde, R., Mandel, A. & Duncanson, L. Global Forest Classification for Intergovernmental Panel on Climate Change (IPCC) Tier 1 natural forest biomass estimates, https://doi.org/10.5281/zenodo.13346188 (2024).

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

70. Potapov, P. *et al*. The last frontiers of wilderness: Tracking loss of intact forest landscapes from 2000 to 2013. *Science Advances* **3**, e1600821, https://doi.org/10.1126/sciadv.1600821 (2017).

71. Turubanova, S., Potapov, P. V., Tyukavina, A. & Hansen, M. C. Ongoing primary forest loss in Brazil, Democratic Republic of the Congo, and Indonesia. *Environmental Research Letters* **13**, 074028, https://doi.org/10.1088/1748-9326/aacd1c (2018).

72. Harris, N., Goldman, E. D. & Gibbes, S. Spatial database of Planted Trees (SDPT version 1.0) (2019).

73. Petersen, R. *et al*. Mapping tree plantations with multispectral imagery: preliminary results for seven tropical countries. https://www.wri.org/research/mapping-tree-plantations-multispectral-imagery-preliminary-results-seven-tropical (2016).

74. Vijay, V., Pimm, S. L., Jenkins, C. N. & Smith, S. J. The impacts of oil palm on recent deforestation and biodiversity loss. *PLOS ONE* **11**, e0159668, https://doi.org/10.1371/journal.pone.0159668 (2016).

75. Descals, A. *et al*. High-resolution global map of smallholder and industrial closed-canopy oil palm plantations. *Earth System Science Data* **13**, 1211–1231, https://doi.org/10.5194/essd-13-1211-2021 (2021).

76. Descals, A. *et al*. High-resolution global map of closed-canopy coconut palm. *Earth System Science Data* **15**, 3991–4010, https://doi.org/10.5194/essd-15-3991-2023 (2023).

77. Gaveau, D., Salim, M. A., Husnayaen & Manurung, T. Industrial and smallholder Oil Palm plantation expansion in Indonesia from 2001 to 2019, https://doi.org/10.5281/ZENODO.6069211 (2022).

78. Feng, M. *et al*. ABoVE: Tree Canopy Cover and Stand Age from Landsat, Boreal Forest Biome, 1984-2020, https://doi.org/10.3334/ORNLDAAC/2012 (2022).

79. Hansen, M. C. *et al*. High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853, https://doi.org/10.1126/science.1244693 (2013).

80. Planet. Level 0 - Open and Public Good: Public view-only access of the Visual Mosaics (Bi-annual Archive and Monthly Monitoring) through Purpose Ally sites.). https://www.planet.com/nicfi/ (2021).

81. ESRI. World Imagery. https://www.arcgis.com/home/item.html?id=10df2279f9684e4a9f6a7f08febac2a9 (2009).

Acknowledgements

Author contributions

N. Hunka, L.D. and J.M. conceptualized the methodological approach. P.P., S.T. and A.T. refined the approach and design of the forest status/condition classification. C.B., P. Montesano, C.N., H.G., P.P., S.T., A.T., J.R. and N. Harris provided input datasets required as inputs to the forest classification. L.D., J.A., R.D., S.H., M.S., A.A., M.U. and A.P. contributed to the estimation of space-based Tier 1 aboveground biomass density. L.D., D.R.S., M.H., B.P., S.N.W., G.G., S.F., M.J.S. and J.M. formulated the policy-related impact and implications of the research. L.D. and N. Hunka acquired funding and coordinate the CEOS Biomass Harmonization activity. N. Hunka and P. May analysed the results, and N. Hunka wrote the manuscript. All authors reviewed and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41597-024-03930-9.

Correspondence and requests for materials should be addressed to N.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.