SCIENTIFIC DATA

OPEN

DATA DESCRIPTOR

Harmonized global maps of above and belowground biomass carbon density in the year 2010

Seth A. Spawn[1,2]✉, Clare C. Sullivan[1,2], Tyler J. Lark[2] & Holly K. Gibbs[1,2]

Remotely sensed biomass carbon density maps are widely used for myriad scientific and policy applications, but all remain limited in scope. They often only represent a single vegetation type and rarely account for carbon stocks in belowground biomass. To date, no global product integrates these disparate estimates into an all-encompassing map at a scale appropriate for many modelling or decision-making applications. We developed an approach for harmonizing vegetation-specific maps of both above and belowground biomass into a single, comprehensive representation of each. We overlaid input maps and allocated their estimates in proportion to the relative spatial extent of each vegetation type using ancillary maps of percent tree cover and landcover, and a rule-based decision schema. The resulting maps consistently and seamlessly report biomass carbon density estimates across a wide range of vegetation types in 2010 with quantified uncertainty. They do so for the globe at an unprecedented 300-meter spatial resolution and can be used to more holistically account for diverse vegetation carbon stocks in global analyses and greenhouse gas inventories.

Background & Summary

Terrestrial ecosystems store vast quantities of carbon (C) in aboveground and belowground biomass[1]. At any point in time, these stocks represent a dynamic balance between the C gains of growth and C losses from death, decay and combustion. Maps of biomass are routinely used for benchmarking biophysical models[2–4], estimating C cycle effects of disturbance[5–7], and assessing biogeographical patterns and ecosystem services[8–11]. They are also critical for assessing climate change drivers, impacts, and solutions, and factor prominently in policies like Reducing Emissions from Deforestation and Forest Degradation (REDD+) and C offset schemes[12–14]. Numerous methods have been used to map biomass C stocks but their derivatives often remain limited in either scope or extent[12,15]. There thus remains a critical need for a globally harmonized, integrative map that comprehensively reports biomass C across a wide range of vegetation types.

Most existing maps of aboveground biomass (AGB) and the carbon it contains (AGBC) are produced from statistical or data-driven methods relating field-measured or field-estimated biomass densities and spaceborne optical and/or radar imagery[12,15,16]. They largely focus on the AGB of trees, particularly those in tropical landscapes where forests store the majority of the region's biotic C in aboveground plant matter. Land cover maps are often used to isolate forests from other landcover types where the predictive model may not be appropriate such that forest AGB maps intentionally omit AGB stocks in non-forest vegetation like shrublands, grasslands, and croplands, as well as the AGB of trees located within the mapped extent of these excluded landcovers[17]. Non-forest AGB has also been mapped to some extent using similar approaches but these maps are also routinely masked to the geographic extent of their focal landcover[18–21]. To date, there has been no rigorous attempt to harmonize and integrate these landcover-specific, remotely sensed products into a single comprehensive and temporally consistent map of C in all living biomass.

Maps of belowground biomass (BGB) and carbon density (BGBC) are far less common than those of AGB because BGB cannot be readily observed from space or airborne sensors. Consequently, BGB is often inferred from taxa-, region-, and/or climate-specific "root-to-shoot" ratios that relate the quantity of BGB to that of AGB[22–24]. These ratios can be used to map BGB by spatially applying them to AGB estimates using maps of their respective strata[5]. In recent years, more sophisticated regression-based methods have been developed to predict

[1]Department of Geography, University of Wisconsin-Madison, Madison, WI, USA. [2]Center for Sustainability and the Global Environment (SAGE), Nelson Institute for Environmental Studies, University of Wisconsin-Madison, Madison, WI, USA. ✉e-mail: spawn@wisc.edu

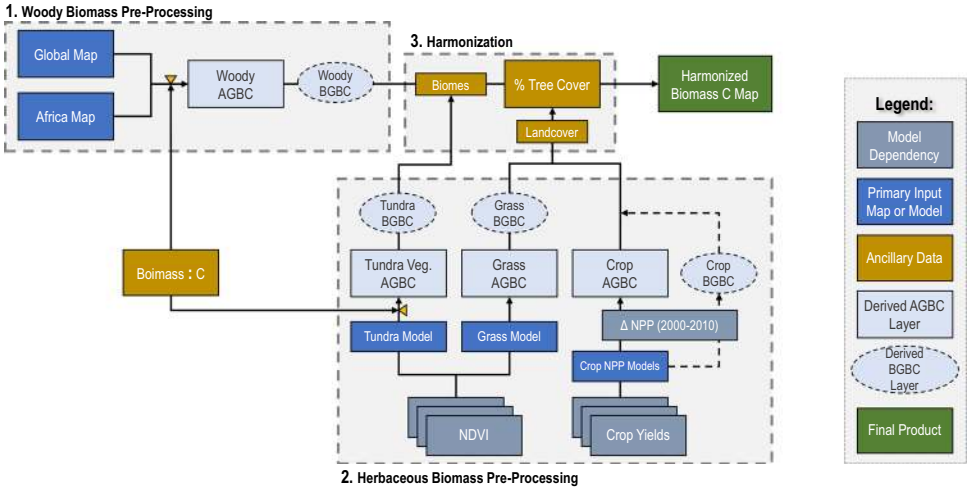

Fig. 1 Generalized, three-step workflow used to create harmonized global biomass maps. In step one, woody AGB maps are prepared, combined, converted to AGBC density and used to create separate but complementary maps of BGBC. In step two, a similar workflow is used to generate matching maps of AGBC and BGBC for tundra vegetation, grasses, and annual crops. In step three, all maps are combined using a rule-based decision tree detailed in Fig. 3 to generate comprehensive, harmonized global maps. All input data sources are listed and described in Table 1.

root-to-shoot ratios of some landcover types based on covariance with other biophysical and/or ecological factors[25,26]. When applied spatially, these methods can allow for more continuous estimates of local BGB[5,27]. Like AGBC, though, few attempts have been made to comprehensively map BGBC for the globe.

Despite the myriad of emerging mapping methods and products, to date, the Intergovernmental Panel on Climate Change (IPCC) Tier-1 maps by Ruesch and Gibbs[28] remains the primary source of global AGBC and BGBC estimates that transcend individual landcover types. These maps, which represents the year 2000, were produced prior to the relatively recent explosion of satellite-based AGB maps and they therefore rely on an alternative mapping technique called "stratify and multiply"[15], which assigns landcover-specific biomass estimates or "defaults" (often derived from field measurements or literature reviews) to the corresponding classified grid cells of a chosen landcover map[12]. While this approach yields a comprehensive wall-to-wall product, it can fail to capture finer-scale spatial patterns often evident in the field and in many satellite-based products[12,15]. The accuracy of these maps is also tightly coupled to the quality and availability of field measurements[29] and the thematic accuracy and discontinuity of the chosen landcover map.

Given the wealth of landcover-specific satellite based AGB maps, a new harmonization method akin to "stratify and multiply" is needed to merge the validated spatial detail of landcover-specific remotely sensed biomass maps into a single, globally harmonized product. We developed such an approach by which we (i) overlay distinct satellite-based biomass maps and (ii) proportionately allocate their estimates to each grid cell ("overlay and allocate"). Specifically, we overlay continental-to-global scale remotely sensed maps of landcover-specific biomass C density and then allocate fractional contributions of each to a given grid cell using additional maps of percent tree cover, thematic landcover and a rule-based decision tree. We implement the new approach here using temporally consistent maps of AGBC as well as matching derived maps of BGBC to generate separate harmonized maps of AGBC and BGBC densities. In addition, we generate associated uncertainty layers by propagating the prediction error of each input dataset. The resulting global maps consistently represent biomass C and associated uncertainty across a broad range of vegetation in the year 2010 at an unprecedented 300 meter (m) spatial resolution.

Methods

Our harmonization approach (Fig. 1) relies on independent, landcover-specific biomass maps and ancillary layers, which we compiled from the published literature (Table 1). When published maps did not represent our epoch of interest (i.e. grasslands and croplands) or did not completely cover the necessary spatial extent (i.e. tundra vegetation), we used the predictive model reported with the respective map to generate an updated version that met our spatial and temporal requirements. We then used landcover specific root-to-shoot relationships to generate matching BGBC maps for each of the input AGBC maps before implementing the harmonization procedure. Below we describe, in detail, the methodologies used for mapping AGBC and BGBC of each landcover type and the procedure used to integrate them.

Aboveground biomass carbon density maps. *Woody tree biomass.* Since the first remotely sensed woody AGB maps were published in the early 1990s, the number of available products has grown at an extraordinary pace[16] and it can thus be challenging to determine which product is best suited for a given application. For our purposes, we relied on the GlobBiomass AGB density map[30] as our primary source of woody AGB estimates due to its precision, timestamp, spatial resolution, and error quantification. It was produced using a combination

Data source	Description	Use
Santoro *et al.*[30] (GlobBiomass)	Global, remotely sensed map of woody AGB in living trees with DBH greater than 10 cm and masked to pixels containing Landsat-identified tree cover in 2010[34]. Native resolution of 100 m. Includes accompanying standard error of predictions layer.	Woody AGBC mapping
Bouvet *et al.*[35]	Continental, remotely sensed map of woody AGB in living trees of any size in Africa. Unmasked and includes shrublands. Native resolution of 25 m. RMSE of prediction of 17.0 Mg ha^{-1}.	Woody AGBC mapping
CCI Landcover 2010[37]	Landcover map produced at 300 m spatial resolution by the European Space Agency's Climate Change Initiative. Represents the year 2010.	Woody AGBC mapping; Woody BGBC mapping; Map Harmonization
Xia *et al.*[19]	Non-linear regression relating grassland AGBC density to AVHRR NDVI. Previously used for global mapping. RMSE = 0.3 Mg ha^{-1}.	Grassland AGBC mapping
MODIS NDVI (16 Day)[50,51]	16-Day global MODIS Aqua and Terra NDVI composite images. Native resolution of 250 m. Accessed in Google Earth Engine[81].	Grassland AGBC mapping
Fensholt and Proud[52]	Coefficients for calibrating MODIS to AVHRR NDVI values.	Grassland AGBC mapping
Berner *et al.*[18]	Non-linear regression model relating tundra AGBC density to Landsat ETM derived NDVI. Previously used to map Alaskan Tundra.	Tundra AGBC mapping
MODIS Surface Reflectance (Daily)[46,47]	Daily global NDVI composite images derived from MODIS Aqua and Terra surface reflectance images. Native resolution of 250 m. Accessed in Google Earth Engine[81].	Tundra AGBC mapping
Steven *et al.*[48]	Coefficients for calibrating MODIS to Landsat ETM NDVI values.	Tundra AGBC mapping
Monfreda *et al.*[20]	Globally gridded yield maps for 70 annually harvested herbaceous commodity crops (Online-only Table 2) representing c. 2000 at 8 km resolution. Crop-specific parameters used to calculate cropland ANPP.	Cropland AGBC mapping; Cropland BGBC mapping
Wolf *et al.*[21]	Crop-specific parameters and model used to calculate cropland ANPP.	Cropland AGBC mapping; Cropland BGBC mapping
Ramankutty *et al.*[55]	Map of global cropland area c. 2000 that complements the global crop yield maps of Monfreda *et al.*[20] at 8 km resolution.	Cropland AGBC mapping; Cropland BGBC mapping
MODIS ANPP[58]	Remotely sensed global maps of modelled MODIS Terra ANPP (2000–2015) at 1 km native resolution. Accessed in Google Earth Engine[81].	Cropland AGBC mapping; Cropland BGBC mapping
Reich *et al.*[25]	Multiple regression model predicting BGB of trees from AGB using environmental covariates.	Woody BGBC mapping
Potopov *et al.*[60]	Global map of "Intact forested landscapes" which are defined as large contiguous forest patches not influenced by human activity. User's accuracy = 92%.	Woody BGBC mapping
Harris *et al.*[61]	Spatial database of planted trees with incomplete global coverage	Woody BGBC mapping
FRA[62]	FAO Global Forest Resource Assessment – national statistics on the spatial extent of natural and planted forests and other woody vegetation.	Woody BGBC mapping
FAOSTAT	FAOSTAT database – national statistics on the planted area of tree crops. (http://www.fao.org/faostat)	Woody BGBC mapping
Fick and Hijmans[59] (WorldClim version 2)	Global map of mean annual temperature (MAT) between 1980–2000. 1 km native resolution. RMSE = 1.12 °C.	Woody BGBC mapping; Tundra BGBC mapping
Wang *et al.*[26]	Regression model predicting the root-to-shoot ratios of tundra plants from MAT.	Tundra BGBC mapping
Mokany *et al.*[22]	Mean and standard error of field-measured root-to-shoot ratios for natural landcover types, stratified by climatic zone.	Woody BGBC mapping Grassland BGBC mapping;
Kottek *et al.*[43]	Updated version of the Köppen-Geiger climate classification.	Woody AGBC mapping; Woody BGBC mapping; Grassland BGBC mapping
Martin *et al.*[42]	Mean and standard error of field measured biomass carbon fractions globally stratified by climatic zone and plant phylogeny.	Woody AGBC mapping; Woody BGBC mapping; Tundra AGBC mapping; Tundra BGBC mapping
MODIS Treecover[63]	Global map of percent tree cover in 2010 from MODIS Terra at a native 250 m resolution. Includes estimated standard deviation of each grid cell's prediction. Accessed in Google Earth Engine[81].	Map Harmonization
Resolve2017 Biomes[65]	Updated polygonal extents of the Olson biome classification[83]. Accessed in Google Earth Engine[81].	Map Harmonization

Table 1. Data sources used to generate harmonized global maps of above and belowground biomass carbon density.

of spaceborne optical and synthetic aperture radar (SAR) imagery and represents the year 2010 at a 100 m spatial resolution – making it the most contemporary global woody AGB currently available and the only such map available for that year. Moreover, GlobBiomass aims to minimize prediction uncertainty to less than 30% and a recent study suggests that it has high fidelity for fine-scale applications[31].

The GlobBiomass product was produced by first mapping the growing stock volume (GSV; i.e. stem volume) of living trees, defined following Food and Agriculture Organization (FAO) guidelines[32] as those having a diameter at breast height (DBH) greater than 10 centimeters (cm). AGB density was then determined from GSV by applying spatialized biomass expansion factors (BEFs) and wood density estimates. These factors were mapped using machine learning methods trained from a suite of plant morphological databases that compile thousands of field measurements from around the globe[33]. The resulting AGB estimates represent biomass in the living structures (stems, branches, bark, twigs) of trees with a DBH greater than 10 cm. This definition may thereby overlook AGB of smaller trees and/or shrubs common to many global regions. Unlike other maps, though, the

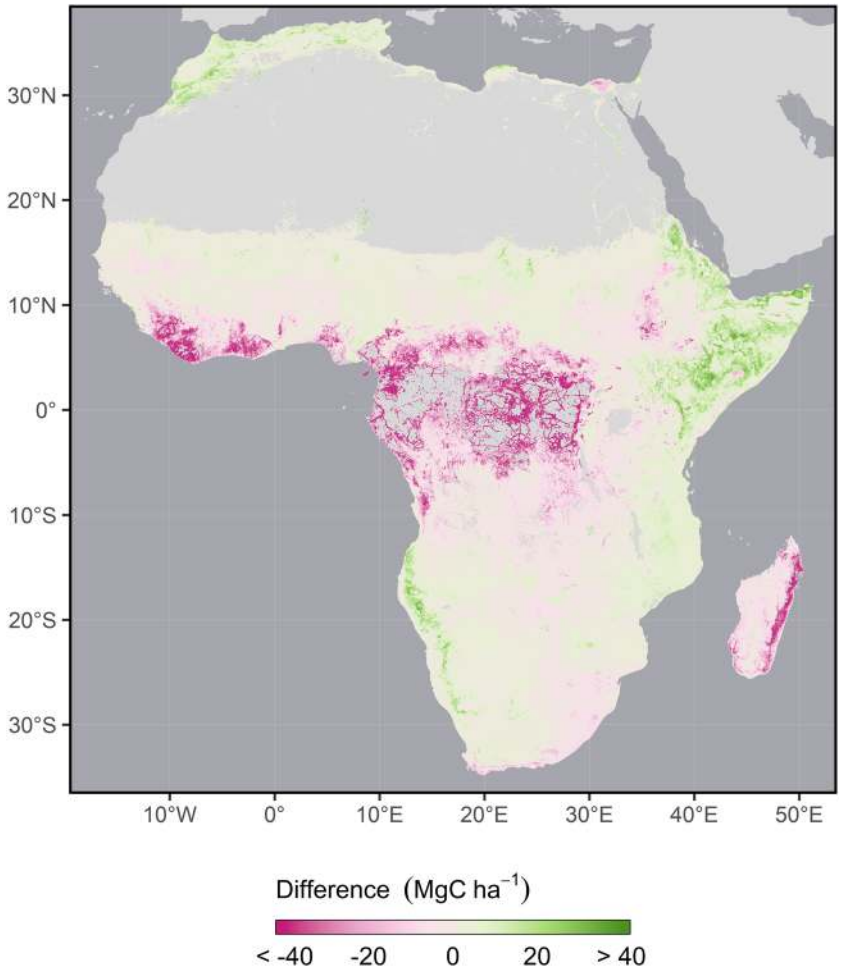

Fig. 2 Difference between underlying woody aboveground biomass maps in Africa. Maps considered are the GlobBiomass[30] global map and the Bouvet[35] map of Africa. Both maps were aggregated to a 300 m spatial resolution and converted to C density prior to comparison using the same schema. The difference map was subsequently aggregated to a 3 km spatial resolution and reprojected for visualization. Negative values denote lower estimates by Bouvet *et al.*[35], while positive values denote higher estimates.

GlobBiomass product employs a subpixel masking procedure that retains AGB estimates in 100 m grid cells in which any amount of tree cover was detected in finer resolution (30 m) imagery[34]. This unique procedure retains AGB estimates in tree-sparse regions like savannahs, grasslands, croplands, and agroforestry systems where AGB is often overlooked[17], as well as in forest plantations. The GlobBiomass product is the only global map that also includes a dedicated uncertainty layer reporting the standard error of prediction. We used this layer to propagate uncertainty when converting AGB to AGBC density, modelling BGBC, and integrating with C density estimates of other vegetation types.

Bouvet *et al.*[35] – some of whom were also participants of the GlobBiomass project – independently produced a separate AGB density map for African savannahs, shrublands and dry woodlands circa 2010 at 25 m spatial resolution[35] (hereafter "Bouvet map"), which we included in our harmonized product to begin to address the GlobBiomass map's potential omission of small trees and shrubs that do not meet the FAO definition of woody AGB. This continental map of Africa is based on a predictive model that directly relates spaceborne L-band SAR imagery – an indirect measure of vegetation structure that is sensitive to low biomass densities[36] – with region-specific, field-measured AGB. Field measurements (n = 144 sites) were compiled from 7 different sampling campaigns – each specifically seeking training data for biomass remote sensing – that encompassed 8 different countries[35]. The resulting map is not constrained by the FAO tree definition and is masked to exclude grid cells in which predicted AGB exceeds 85 megagrams dry mater per hectare (Mg ha^{-1}) – the threshold at which the SAR-biomass relationship saturates. To avoid extraneous prediction, it further excludes areas identified as "broadleaved evergreen closed-to-open forest", "flooded forests", "urban areas" and "water bodies" by the European Space Agency's Climate Change Initiative (CCI) Landcover 2010 map[37] and as "bare areas" in the Global Land Cover (GLC) 2000 map[38]. While the Bouvet map is not natively accompanied by an uncertainty layer, its authors provided us with an analytic expression of its uncertainty (*SE*; standard error of prediction) as a function of estimated AGB (Eq. 1) which we used to generate an uncertainty layer for subsequent error propagation.

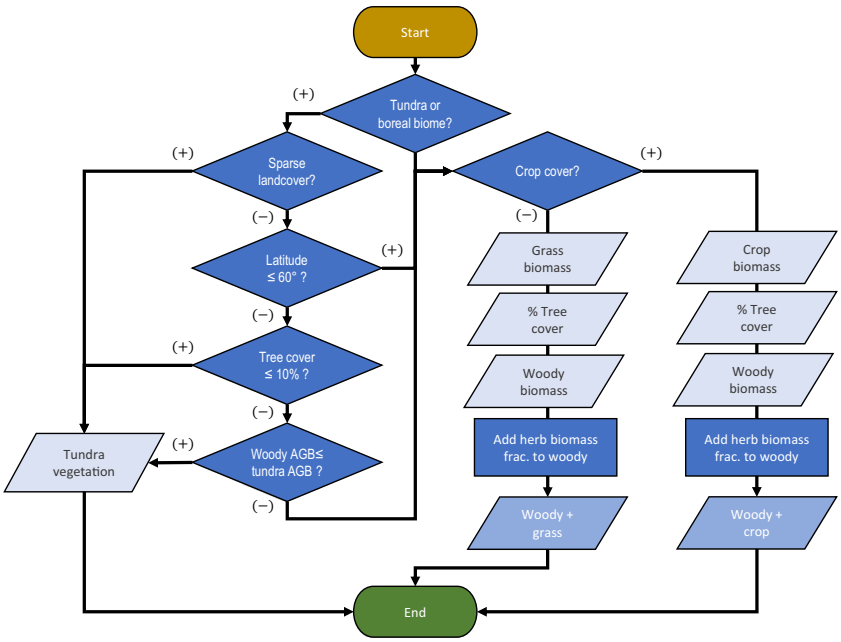

Fig. 3 Decision tree used to allocate landcover-specific biomass estimates to each grid cell of our harmonized global products.

$$SE = 1.0551 \cdot AGB - 0.007 \cdot AGB^2 - 0.0000273 \cdot AGB^3;\ 0 \le AGB \le 85\ \text{Mg ha}^{-1} \quad (1)$$

We combined the GlobBiomass and Bouvet products to generate a single woody biomass map by first upscaling each map separately to a matching 300 m spatial resolution using an area-weighted average to aggregate grid cells, and then assigning the Bouvet estimate to all overlapping grid cells, except those identified by the CCI Landcover 2010 map as closed or flooded forest types (Online-only Table 1) which were not within the dryland domain of the Bouvet map. While more complex harmonization procedures based on various averaging techniques have been used by others[39,40], their fidelity remains unclear since they fail to explicitly identify and reconcile the underlying source of the inputs' discrepancies[41]. We thus opted to use a more transparent ruled-based approach when combining these two woody biomass maps, which allows users to easily identify the source of a grid cell's woody biomass estimate. Given the local specificity of the training data used to produce the Bouvet map, we chose to prioritize its predictions over those of the GlobBiomass product when within its domain. In areas of overlap, the Bouvet map values tend to be lower in moist regions and higher in dryer regions (Fig. 2), though, where used, these differences rarely exceed ±25 megagrams C per hectare (MgC ha^{-1}).

We then converted all woody AGB estimates to AGBC by mapping climate and phylogeny-specific biomass C concentrations from Martin *et al.*[42]. Climate zones were delineated by aggregating classes of the Köppen-Gieger classification[43] (Table 2) to match those of Martin *et al.*[42]. Phylogenetic classes (angiosperm, gymnosperm and mixed/ambiguous) were subsequently delineated within each of these zones using aggregated classes of the CCI Landcover 2010 map (Online-only Table 1). Martin *et al.*[42] only report values for angiosperms and gymnosperms so grid cells with a mixed or ambiguous phylogeny were assigned the average of the angiosperm and gymnosperm values and the standard error of this value was calculated from their pooled variance. Due to residual classification error in the aggregated phylogenetic classes, we weighted the phylogeny-specific C concentration within each climate zone by the binary probability of correctly mapping that phylogeny (i.e. user's accuracy)[44] using Eq. 2

$$\mu_c = \mu_m p_m + \frac{(\mu_n + \mu_o)}{2}(1 - p_m) \quad (2)$$

where, within each climate zone, μ_c is the mean probability-weighted C concentration of the most probable phylogeny, μ_m is the mean C concentration of that phylogeny from Martin *et al.*[42], p_m is the user's accuracy of that phylogeny's classification (Table 3), and μ_n and μ_o are the mean C concentrations of the remain phylogenetic classes from Martin *et al.*[42]. Standard error estimates for these C concentrations were similarly weighted using summation in quadrature (Eq. 3)

$$\sigma_c = \sqrt{2\sigma_m^2 + \left(\frac{1 - p_m}{2}\sigma_n\right)^2 + \left(\frac{1 - p_m}{2}\sigma_o\right)^2 + (p_m\sigma_m)^2} \quad (3)$$

where σ_c is the probability-weighted standard error of the most probable phylogeny's C concentration and σ_m, σ_n and σ_o are the standard errors of the respective phylogeny-specific C concentrations from Martin *et al.*[42]. Probability-weighted C concentrations used are reported in Table 4.

Mapped, probability-weighted C estimates were then arithmetically applied to AGB estimates. Uncertainty associated with this correction was propagated using summation in quadrature of the general form (Eq. 4)

$$\sigma_f = \sqrt{\left(\frac{\partial f}{\partial i}\sigma_i\right)^2 + \left(\frac{\partial f}{\partial j}\sigma_j\right)^2 + \ldots + \left(\frac{\partial f}{\partial k}\sigma_k\right)^2} \quad (4)$$

where $\mu_f = f(i, j, \ldots, k)$, σ_f is the uncertainty of μ_f, and σ_i, σ_j, …, σ_k, are the respective uncertainty estimates of the dependent parameters (standard error unless otherwise noted). Here, μ_f, is the estimated AGBC of a given grid cell, and is the product of its woody AGB estimate, and its corresponding C concentration.

Tundra vegetation biomass. The tundra and portions of the boreal biome are characterized by sparse trees and dwarf woody shrubs as well as herbaceous cover that are not included in the GlobBiomass definition of biomass. AGB density of these classes has been collectively mapped by Berner *et al.*[18,45] for the North Slope of Alaska from annual Landsat imagery composites of the normalized difference vegetation index (NDVI) and a non-linear regression-based model trained from field measurements of peak AGB that were collected from the published literature (n = 28 sites). Berner *et al.*[18] note that while these field measurements did not constitute a random or systematic sample, they did encompass a broad range of tundra plant communities. In the absence of a global map and due the sparsity of high quality Landsat imagery at high latitudes, we extended this model to the pan-Arctic and circumboreal regions using NDVI composites created from daily 250 m MODIS Aqua and Terra surface reflectance images[46,47] that were cloud masked and numerically calibrated to Landsat ETM reflectance – upon which the tundra model is based – using globally derived conversion coefficients[48]. We generated six separate 80th percentile NDVI composites circa 2010 – one for each of the MODIS missions (Aqua and Terra) in 2009, 2010 and 2011 – following Berner *et al.*[18]. We chose to use three years of imagery (circa 2010) rather than just one (2010) to account for the potential influence that cloud masking may exert upon estimates of the 80th NDVI percentile in a single year. We then applied the tundra AGB model to each composite, converted AGB estimates to AGBC by assuming a biomass C fraction of 49.2% (SE = 0.8%)[42] and generated error layers for each composite from the reported errors of the AGB regression coefficients and the biomass C conversion factor using summation in quadrature as generally described above (Eq. 4). A single composite of tundra AGBC circa 2010 was then created as the pixelwise mean of all six composites. We also generated a complementary uncertainty layer representing the cumulative standard error of prediction, calculated as the pixelwise root mean of the squared error images in accordance with summation in quadrature. Both maps were upscaled from their native 250 m spatial resolution to a 300 m spatial resolution using an area weighted aggregation procedure, whereby pixels of the 300 m biomass layer was calculated as the area weighted average of contained 250 m grid cells, and the uncertainty layer was calculated – using summation in quadrature – as the root area-weighted average of the contained grid cells squared.

Grassland biomass. Grassland AGBC density was modelled directly from maximum annual NDVI composites using a non-linear regression-based model developed by Xia *et al.*[19] for mapping at the global scale. This model was trained by relating maximum annual NDVI as measured by the spaceborne Advanced Very High-Resolution Radiometer (AVHRR) sensor to globally distributed field measurements of grassland AGBC that were compiled from the published literature (81 sites for a total of 158 site-years). Like the tundra biomass training data, these samples did not constitute a random or systematic sample but do encompass a comprehensive range of global grassland communities. Given the inevitable co-occurrence of trees in the AVHRR sensor's 8 km resolution pixels upon which the model is trained, it's predictions of grassland AGBC are relatively insensitive to the effects of co-occurring tree cover. We thereby assume that its predictions for grid cells containing partial tree cover represent the expected herbaceous AGBC density in the absence of those trees. Maximum model predicted AGBC (NDVI = 1) is 2.3 MgC ha^{-1} which is comparable to the upper quartile of herbaceous AGBC estimates from global grasslands[49] and suggests that our assumption will not lead to an exaggerated estimation. For partially wooded grid cells, we used modelled grassland AGBC density to represent that associated with the herbaceous fraction of the grid cell in a manner similar to Zomer *et al.*[17] as described below (See "*Harmonizing Biomass Carbon Maps*").

We applied the grassland AGBC model to all grid cells of maximum annual NDVI composites produced from finer resolution 16-day (250 m) MODIS NDVI imagery composites circa 2010[50,51]. Here again, three years of imagery were used to account for potential idiosyncrasies in a single year's NDVI composites resulting from annual data availability and quality. As with AGB of tundra vegetation, annual composites (2009–2011) were constructed separately from cloud-masked imagery collected by both MODIS missions (Aqua and Terra; n = 6) and then numerically calibrated to AVHRR reflectance using globally derived conversion coefficients specific to areas of herbaceous cover[52]. We then applied the AGBC model to each of these composites and estimated error for each composite from both the AVHRR calibration (standard deviation approximated from the 95% confidence interval of the calibration scalar) and the AGBC model (relative RMSE) using summation in quadrature. A single map of grassland AGBC circa 2010 was then created as the pixelwise mean of all six composites and an associated error layer was created as the pixelwise root mean of the squared error images. Both maps were aggregated from their original 250 m resolution to 300 m to facilitate harmonization using the area-weighted procedure described previously for woody and tundra vegetation (see section 1.2).

Cropland biomass. Prior to harvest, cropland biomass can also represent a sizable terrestrial C stock. In annually harvested cropping systems, the maximum standing biomass of these crops can be inferred from annual net primary productivity (ANPP). While spaceborne ANPP products exist, they generally perform poorly in

KG Code	KG Class	KG Main Climate	KG Precipitation	KG Temperature	Mokany *et al.*[22] Grassland Class	Martin *et al.*[42] Carbon Domain
1	Af	Equatorial	Fully Humid	—	Tropical/Subtropical	Tropical
2	Am	Equatorial	Monsoonal	—	Tropical/Subtropical	Tropical
3	As	Equatorial	Summer Dry	—	Tropical/Subtropical	Tropical
4	Aw	Equatorial	Winter Dry	—	Tropical/Subtropical	Tropical
5	BSh	Arid	Steppe	Hot Arid	Temperate	Subtropical/Mediterranean
6	BSk	Arid	Steppe	Cold Arid	Temperate	Temperate
7	BWh	Arid	Desert	Hot Arid	Temperate	Subtropical/Mediterranean
8	BWk	Arid	Desert	Cold Arid	Temperate	Temperate
9	Cfa	W. Temperate	Fully Humid	Hot Summer	Temperate	Subtropical/Mediterranean
10	Cfb	W. Temperate	Fully Humid	Warm Summer	Temperate	Temperate
11	Cfc	W. Temperate	Fully Humid	Cool Summer	Temperate	Temperate
12	Csa	W. Temperate	Summer Dry	Hot Summer	Temperate	Subtropical/Mediterranean
13	Csb	W. Temperate	Summer Dry	Warm Summer	Temperate	Temperate
14	Csc	W. Temperate	Summer Dry	Cool Summer	Temperate	Temperate
15	Cwa	W. Temperate	Winter Dry	Hot Summer	Temperate	Subtropical/Mediterranean
16	Cwb	W. Temperate	Winter Dry	Warm Summer	Temperate	Subtropical/Mediterranean
17	Cwc	W. Temperate	Winter Dry	Cool Summer	Temperate	Temperate
18	Dfa	Snow	Fully Humid	Hot Summer	Cool Temperate	Temperate
19	Dfb	Snow	Fully Humid	Warm Summer	Cool Temperate	Temperate
20	Dfc	Snow	Fully Humid	Cool Summer	Tundra	Boreal
21	Dfd	Snow	Fully Humid	Ext. Continental	Tundra	Boreal
22	Dsa	Snow	Summer Dry	Hot Summer	Cool Temperate	Temperate
23	Dsb	Snow	Summer Dry	Warm Summer	Cool Temperate	Temperate
24	Dsc	Snow	Summer Dry	Cool Summer	Cool Temperate	Boreal
25	Dsd	Snow	Summer Dry	Ext. Continental	Cool Temperate	Boreal
26	Dwa	Snow	Winter Dry	Hot Summer	Cool Temperate	Temperate
27	Dwb	Snow	Winter Dry	Warm Summer	Cool Temperate	Temperate
28	Dwc	Snow	Winter Dry	Cool Summer	Tundra	Boreal
29	Dwd	Snow	Winter Dry	Ext. Continental	Tundra	Boreal
30	EF	Polar	—	Polar Frost	Tundra	Boreal
31	ET	Polar	—	Polar Tundra	Tundra	Boreal
32	Ocean	—	—	—	Tropical/Subtropical	Global

Table 2. Reclassification table of the Köppen-Gieger climate classification. The Köppen-Gieger (KG) climate classification[43] was used to stratify grassland root-to-shoot ratios from Mokany *et al.*[22] as described in Table 5 and biomass carbon concentrations by Martin *et al.*[42] as described in Table 4.

	Gymno.	Mixed	Angio.	User's Acc.
Gymno.	0.0180	0.0124	0.0024	55%
Mixed	0.0020	0.9241	0.0086	99%
Angio.	0.0006	0.0076	0.0243	75%
Prod. Acc.	87%	98%	69%	**Overall**: 97%

Table 3. Area-weighted confusion matrix for "likely forest phylogeny" classes. Classes were aggregated from the CCI landcover map[37] (Online-only Table 1) and the associated confusion matrix for the year 2010 (Tables 4–5 in version 2.5 of the D3.4-PUG CCI Landcover Product User Guide[84]). User's accuracies were used to propagate uncertainty of phylogenetic classification when converting biomass density to carbon density and in woody BGBC calculations.

croplands[53,54]. Instead, cropland ANPP is more commonly derived from crop yields[20,21,53]. We used globally gridded, crop-specific yields of 70 annually harvested herbaceous commodity crops circa 2000 by Monfreda *et al.*[20] – the only year in which these data were available. These maps were produced by spatially disaggregating crop-yield statistics for thousands of globally distributed administrative units throughout the full extent of a satellite-based cropland map[20]. These maps were combined with crop-specific parameters (Online-only Table 2) to globally map AGBC as aboveground ANPP for each crop following the method of Wolf *et al.*[21]. This method can be simplified as (Eq. 5)

$$AGBC = y\omega(0.451h^{-1} + 1.025c - 0.451) \quad (5)$$

Climatic Zone	Phylogeny	Mean	SE
Tropical	Angio.	0.454	0.003
	Mixed	0.452	0.004
	Gymno	0.450	0.008
Subtropical/Mediterranean	Angio.	0.465	0.006
	Mixed	0.478	0.008
	Gymno	0.484	0.009
Temperate	Angio.	0.472	0.005
	Mixed	0.483	0.006
	Gymno	0.489	0.006
Boreal	Angio.	0.488	0.013
	Mixed	0.480	0.011
	Gymno	0.476	0.009
Global	Angio.	0.471	0.011
	Mixed	0.476	0.016
	Gymno	0.479	0.012

Table 4. Climate and phylogeny specific biomass C fractions used to convert biomass density estimates to carbon density. C fractions were taken from Martin *et al.*[42] and weighted by the aggregated probability of correct phylogenetic classification (i.e. user's accuracy) from Table 3. Climate zones are spatially defined in Table 2.

Taxa	Climate	Strata/Taxa Map	Mean	SE
Savannah	All	CCI Landcover	0.642	0.111
Shrub	All	CCI Landcover	1.837	0.589
Grassland	Tropical/Subtropical	Köppen-Gieger	1.887	0.304
	Temperate	Köppen-Gieger	4.224	0.518
	Cool Temperate	Köppen-Gieger	4.504	1.337
	Tundra	Köppen-Gieger	4.804	1.188

Table 5. Root-to-shoot ratios used to map BGB of select landcover types. Root-to-shoot ratios and their standard errors were taken from Mokany *et al.*[22]. Grassland stratification classes correspond with those reported as "Mokany Grassland Class" in Table 2.

where y is the crop's yield (Mg ha^{-1}), ω is the dry matter fraction of its harvested biomass, h is its harvest index (fraction of total AGB collected at harvest) and c is the carbon content fraction of its harvested dry mass. This simplification assumes, following Wolf *et al.*[21], that 2.5% of all harvested biomass is lost between the field and farmgate and that unharvested residue and root mass is 44% C.

Total cropland AGBC density was then calculated as the harvested-area-weighted average of all crop-specific AGBC estimates within a given grid cell. Since multiple harvests in a single year can confound inference of maximum AGBC from ANPP, we further determined the harvest frequency (f) of each grid cell by dividing a cell's total harvested area (sum of the harvested area of each crop reported within a given grid cell) by its absolute cropland extent as reported in a complementary map by Ramankutty *et al.*[55]. If f was greater than one, multiple harvests were assumed to have occurred and AGBC was divided by f to ensure that AGBC estimates did not exceed the maximum standing biomass density.

Since the yields of many crops and, by association, their biomass have changed considerably since 2000[56,57], we calibrated our circa 2000 AGBC estimates to the year 2010 using local rates of annual ANPP change (MgC ha^{-1} yr^{-1}) derived as the Theil-Sen slope estimator – a non-parametric estimator that is relatively insensitive to outliers – of the full MODIS Terra ANPP timeseries (2000–2015)[58]. Total ANPP change between 2000 and 2010 for each grid cell was calculated as ten times this annual rate of change. Since MODIS ANPP represents C gains in both AGB and BGB, we proportionately allocated aboveground ANPP to AGBC using the total root-to-shoot ratio derived from the circa 2000 total crop AGBC and BGBC maps (described below). Since error estimates were not available for the yield maps or the crop-specific parameters used to generate the circa 2000 AGBC map, estimated error of the circa 2010 crop AGBC map was exclusively based on that of the 2000–2010 correction. The error of this correction was calculated as the pixel-wise standard deviation of bootstrapped simulations (n = 1000) in which a random subset of years was omitted from the slope estimator in each iteration. The 8 km resolution circa 2000 AGBC map and error layer were resampled to 1 km to match the resolution of MODIS ANPP using the bilinear method prior to ANPP correction and then further resampled to 300 m to facilitate harmonization.

Woody crops like fruit, nut, and palm oil plantations were not captured using the procedure just described and their biomass was instead assumed to be captured by the previously described woody biomass products which retained biomass estimates in all pixels where any amount of tree cover was detected at the sub-pixel level (see section 1.1).

Belowground biomass carbon maps. Matching maps of BGBC and associated uncertainty were subsequently produced for each of the landcover-specific AGBC maps using published empirical relationships.

With the exception of savannah and shrubland areas, woody BGBC was modelled from AGBC using a multiple regression model by Reich *et al.*[25] that considers the phylogeny, mean annual temperature (MAT), and regenerative origin of each wooded grid cell and that was applied spatially using maps of each covariate in a fashion similar to other studies[5,27]. Tree phylogeny (angiosperm or gymnosperm) was determined from aggregated classes of the CCI Landcover 2010 map[37] (Online-only Table 1) with phylogenetically mixed or ambiguous classes assumed to be composed of 50% of each. MAT was taken from version 2 of the WorldClim bioclimatic variables dataset (1970–2000) at 1 km resolution[59] and resampled to 300 m using the bilinear method. Since there is not a single global data product mapping forest management, we determined tree origin – whether naturally propagated or planted – by combining multiple data sources. These data included (i) a global map of "Intact Forest Landscapes" (IFL) in the year 2013[60] (a conservative proxy of primary, naturally regenerating forests defined as large contiguous areas with minimal human impact), (ii) a Spatial Database of Planted Trees (SDPT) with partial global coverage[61], (iii) national statistics reported by the FAO Global Forest Resources Assessment (FRA) on the extent of both naturally regenerating and planted forests and woodlands within each country in the year 2010[62], and (iv) national statistics reported by the FAOSTAT database (http://www.fao.org/faostat) on the planted area of plantation crops in 2010. Within each country, we assumed that the total area of natural and planted trees was equal to the corresponding FRA estimates. If the FAOSTAT-reported area of tree crops exceeded FRA-reported planted area, the difference was added to FRA planted total. All areas mapped as IFL were assumed to be of natural origin and BGB was modelled as such. Likewise, besides the exceptions noted below, all tree plantations mapped by the SDPT were assumed to be of planted origin. In countries where the extent of the IFL or SDPT maps fell short of the FRA/FAOSTAT reported areas of natural or planted forests, respectively, we estimated BGBC in the remaining, unknown-origin forest grid cells of that country ($BGBC_u$), as the probability-weighted average of the planted and natural origin estimates using Eq. 6

$$BGBC_u = BGBC_p\left(\frac{\Delta_p}{\Delta_p + \Delta_n}\right) + BGBC_n\left(\frac{\Delta_n}{\Delta_p + \Delta_n}\right) \tag{6}$$

where $BGBC_p$ and $BGBC_n$ are the respective BGBC estimates for a grid cell assuming entirely planted and natural origin, respectively, and Δ_p and Δ_n are the respective differences between (i) the FRA/FAOSTAT and (ii) mapped extent of planted and natural forest within the given grid cell's country. While the mapped extent of IFL forests within a given country never exceeded that country's FRA reported natural forest extent, there were infrequent cases (n = 22 of 257) in which the mapped extent of tree plantations exceeded the corresponding FRA/FAOSTAT estimate of planted forest area. In these cases, we down-weighted the BGB estimates of SDPT forests in a similar fashion such that the weight of their planted estimate (ω_p) was equal to the quotient of (i) the FRA/FAOSTAT planted area and (ii) the SDPT extent within the country, and the weight of the natural origin estimate applied to the SDPT extent (ω_n) was equal to $1 - \omega_p$.

A BGBC error layer was then produced using summation in quadrature from the standard error estimates of the model coefficients, the AGBC error layer, the relative RMSE of MAT (27%), and the derived global uncertainty of the phylogeny layer. Phylogeny error was calculated as the Bernoulli standard deviation (δ) of the binary probability (p) of correct classification (i.e. "area weighted user's accuracy"[44]; Table 3) using Eq. 7.

$$\delta = \sqrt{p(1-p)} \tag{7}$$

Since savannahs and shrublands are underrepresented in the regression-based model[25], their BGBC was instead estimated using static root-to-shoot ratios reported by Mokany *et al.*[22], which are somewhat conservative in comparison to the IPCC Tier-1 defaults[23,24] put favoured for consistency with methods used for grasslands (see below). Error was subsequently mapped from that of the AGBC estimates and the root-to-shoot ratios applied (Table 5).

BGBC of tundra vegetation was mapped from AGBC using a univariate regression model derived by Wang *et al.*[26] that predicts root-to-shoot ratio as a function of MAT. We applied the model using the WorldClim version 2 MAT map[59] and propagated error from the AGBC estimates, the relative RMSE of MAT and the standard error of regression coefficients. Where tundra AGB exceeded 25 Mg ha^{-1} – the maximum field-measured shrub biomass reported by Berner *et al.*[18] – vegetation was considered to include trees and the Reich *et al.*[25] method described earlier for woody vegetation was used instead.

In the absence of a continuous predictor of grassland root-to-shoot ratios, we applied climate specific root-to-shoot ratios from Mokany *et al.*[22] to the corresponding climate regions of the Köppen-Gieger classification[43] (Table 2). Here, again, these ratios vary slightly from the IPCC Tier-1 defaults[23,24] but were chosen for their greater sample size and specificity. Grassland BGBC error was mapped from the error of the AGBC estimates and the respective root-to-shoot ratios.

Cropland BGBC was again estimated from crop-specific yields and morphological parameters (Online-only Table 2) following Wolf *et al.*[21] and Eq. 8

$$BGBC = 0.451\ yrh^{-1} \tag{8}$$

where y is the crop's yield (Mg ha^{-1}), r is the root-to-shoot ratio of the crop, and h is its harvest index. Here again we assume that 2.5% of all harvested biomass is lost between the field and farmgate and that root biomass is 44% C, following Wolf *et al.*[21]. BGBC error was mapped from the error of the 2000-to-2010 ANPP correction for BGBC allocation as described above for cropland AGBC.

Harmonizing biomass carbon maps. The AGBC and BGBC maps were harmonized separately following the same general schema (Fig. 3). Given that our harmonized woody biomass map contains biomass estimates for grid cells in which any amount of tree cover was detected at the subpixel level (see section 1.1), we conserved its estimates regardless of the landcover reported by the 2010 CCI map in order to more fully account for woody biomass in non-forested areas[17]. We then used the MODIS continuous vegetation fields percent tree cover map for 2010[63] to allocate additional biomass density associated with the most probable herbaceous cover (grass or crop) to each grid cell in quantities complementary to that of the grid cell's fractional tree cover estimate (Eq. 9)

$$\mu_T = \mu_w + \mu_h(1 - q) \tag{9}$$

where μ_T is the total biomass estimate of a grid cell, μ_w is the woody biomass estimate for the grid cell, μ_h is its herbaceous biomass estimate, and q is the MODIS fractional tree cover of the grid cell. Since MODIS tree cover estimates saturate at around 80%[64], we linearly stretched values such that 80% was treated as complete tree cover (100%). Moreover, we acknowledge that percent cover can realistically exceed 100% when understory cover is considered but we were unable to reasonably determine the extent of underlying cover from satellite imagery. As such, our approach may underestimate the contribution of herbaceous C stocks in densely forested grid cells. The most likely herbaceous cover type was determined from the CCI Landcover 2010 map, which we aggregated into two "likely herbaceous cover" classes – grass or crop – based on the assumed likelihood of cropland in each CCI class (Online-only Table 1). However, due to inherent classification error in the native CCI Landcover map, when determining the herbaceous biomass contribution we weighted the relative allocation of crop and grass biomass to a given grid cell based on the probability of correct classification by the CCI map (i.e. "user's accuracy", Table 6) of the most probable herbaceous class (p_i) such that μ_h can be further expressed as (Eq. 10)

$$\mu_h = \mu_i p_i + \mu_j(1 - p_i) \tag{10}$$

where μ_i is the predicted biomass of the most probable herbaceous class, and μ_j is that of the less probable class.

The uncertainty of a grid cell's total AGBC or BGBC estimate (σ_T) was determined and mapped from that of its components (μ_w and μ_h) by summation in quadrature which can be simplified as (Eq. 11)

$$\sigma_T = \sqrt{\sigma_w^2 + \left(\mu_h q \sqrt{\left(\frac{\sigma_h}{\mu_h}\right)^2 + \left(\frac{\sigma_q}{q}\right)^2}\right)^2} \tag{11}$$

where σ_w is the error of the grid cell's estimated μ_w, σ_h is the error of its estimated μ_h, and σ_q is the error of its q. Here, σ_h can be further decomposed and expressed as Eq. 12 to account for the accuracy weighted allocation procedure expressed previously (Eq. 10)

$$\sigma_h = \sqrt{\left(\mu_i p_i \sqrt{\left(\frac{\sigma_i}{\mu_i}\right)^2 + \left(\frac{\delta_i}{p_i}\right)^2}\right)^2 + \left(\mu_j(1 - p_i) \sqrt{\left(\frac{\sigma_j}{\mu_j}\right)^2 + \left(\frac{\delta_i}{p_i}\right)^2}\right)^2} \tag{12}$$

where σ_i is the error of the estimated biomass density of the most probable herbaceous class, δ_i is the estimated standard deviation of that class's Bernoulli probability (p; Eq. 7), and σ_j is the error of the estimated biomass density of the less probable herbaceous subclass.

Exceptions to the above schema were made in the tundra and boreal biomes – as delineated by the RESOLVE Ecoregions 2017 biome polygons[65] – where thematic overlap was likely between the woody and tundra plant biomass maps. A separate set of decision rules (Fig. 3) was used to determine whether grid cells in these biomes were to be exclusively allocated the estimate of the tundra plant map or that of the fractional allocation procedure described above. In general, any land in these biomes identified as sparse landcover by the CCI landcover map (Online-only Table 1) was assigned the tundra vegetation estimate. In addition, lands north of 60° latitude with less than 10% tree cover or where the tundra AGBC estimate exceeded that of the woody AGBC estimate were also exclusively assigned the tundra vegetation estimate. Lands north of 60° latitude not meeting these criteria were assigned the woody value with the additional contribution of grass.

Subtle numerical artefacts emerged from the divergent methodologies employed north and south of 60°N latitude. These were eliminated by distance weighting grid cells within 1° of 60°N based on their linear proximity to 60°N and then averaging estimates such that values at or north of 61°N were exclusively based on the northern methodology, those at 60°N were the arithmetic average of the two methodologies and those at or south of 59°N were exclusively based on the southern methodology. This produced a seamless, globally harmonized product that integrates the best remotely sensed estimates of landcover-specific C density. Water bodies identified as class "210" of the CCI 2010 landcover map were then masked from our final products.

Data Records

Data layers (n = 4, Table 7) for the maps of AGBC and BGBC density (Fig. 4) as well as their associated uncertainty maps which represent the combined standard error of prediction (Fig. 5) are available as individual 16-bit integer rasters in GeoTiff format. All layers are natively in a WGS84 Mercator projection with a spatial resolution of approximately 300 m at the equator and match that of the ESA CCI Landcover Maps[37]. Raster values are in units megagrams C per hectare (MgC ha^{-1}) and have been scaled by a factor of ten to reduce file size.

	Crop	Non-Crop	User's Acc.
Crop	0.0314	0.0081	79%
Non-crop	0.0344	0.9261	96%
Prod. Acc.	48%	99%	**Overall**: 96%

Table 6. Area-weighted confusion matrix for "likely herbaceous" classes. Aggregated "likely herbaceous" classes were aggregated from the ESA CCI 2010 landcover map as described in Online-only Table 1. Class accuracies were taken from the ESA CCI matrix for the year 2010 as reported in Tables 4–5 in version 2.5 of the D3.4-PUG CCI Landcover Product User Guide[84] and area-weighted following Olofsson *et al.*[44]. Area weighted user's accuracies were used to propagate uncertainty associated with herbaceous biomass allocation.

These data are accessible through the Oak Ridge National Laboratory (ORNL) DAAC data repository (https://doi.org/10.3334/ORNLDAAC/1763)[66]. In addition, updated and/or derived vegetation-specific layers that were used to create our harmonized 2010 maps are available as supplemental data on *figshare*[67].

Technical Validation

Our harmonized products rely almost exclusively upon maps and models that have been rigorously validated by their original producers and were often accompanied by constrained uncertainty estimates. Throughout our harmonization procedure, we strived to conserve the validity of each of these products by minimizing the introduction of additional error and by tracking any introductions, as described above, such that the final error layers represent the cumulative uncertainty of the inputs used. Ground truth AGB and BGB data are almost always collected for individual landcover types. Consequently, we are unable to directly assess the validity of our integrated estimates beyond their relationships to individual landcover-specific estimates and the extents to which they were modified from their original, previously-validated form prior to and during our harmonization procedure.

Modifications to independent biomass layers. Temporal and spatial updates made to existing landcover-specific maps of non-tree AGB resulted in relatively small changes to their predictions. For example, we used numerically calibrated MODIS imagery to extend the Landsat-based tundra plant AGB model beyond its native extent (the North Slope of Alaska) to the pan-Arctic region since neither a comparable model nor a consistent Landsat time series were available for this extent. We assessed the effects of these assumptions by comparing our predictions for the North Slope with those of the original map[18] (Fig. 6a). Both positive and negative discrepancies exist between ours and the original, though these rarely exceed ± 2 MgC ha^{-1} and no discernibly systematic bias was evident.

Our updated map of grassland biomass carbon in the year 2010 was similarly made by applying the original AVHRR-based model to calibrated MODIS imagery. This too resulted in only subtle changes to the original biomass map (Fig. 6b) that were rarely in excess of 0.5 MgC ha^{-1}. In most areas, our estimates were higher than those of Xia *et al.*[19] who mapped the mean AGBC density between 1986 and 2006. Most of these elevated estimates corresponded with areas in which significant NDVI increases ("greening") have been reported while notably lower estimates in the Argentine Monte and Patagonian steppe biomes of southern South America, likewise, correspond with areas of reported "browning"[68,69]. Both greening and browning trends are well documented phenomena and have been linked to climatic changes[70]. Moreover, we further compared AGBC estimates from both the original Xia *et al.*[19] map and our 2010 update to AGBC field measurements coordinated by the Nutrient Network that were collected from 48 sites around the world between 2007 and 2009[49]. The RMSE (0.68 MgC ha^{-1}) of our updated map was 10% less that of the Xia *et al.* map for sites with less than 40% tree cover. Likewise, our 2010 estimates were virtually unbiased (bias $= -0.01$ MgC ha^{-1}) in comparison to the Xia map (bias $= 0.25$ MgC ha^{-1}). While still noisy, these results suggest that our temporal update improved the overall accuracy of estimated grassland AGBC.

Finally, cropland biomass carbon maps were also updated from their native epoch (2000) to 2010 using pixel-wise rates of MODIS ANPP change over a ten-year period. While MODIS ANPP may be a poor snapshot of crop biomass in a single year, we assumed that its relative change over time reflects real physiological shifts affecting the cropland C cycle. This correction also resulted in only small differences that rarely exceeded ± 2 MgC ha^{-1} and that, spatially, correspond well with observed declines in the yields of select crops that have been linked to climate change[71,72] (Fig. 6c). Nonetheless, updated global yield maps comparable to those available for 2000 would greatly improve our understanding of the interactions between climate change, crop yields, and C dynamics.

Belowground biomass estimates. Belowground biomass is notoriously difficult to measure, model, and also to validate. We accounted for the reported uncertainty of nearly every variable considered when estimating belowground biomass and pixel-level uncertainty, but we were unable to perform an independent validation of our harmonized estimates at the pixel level due to a paucity of globally consistent field data. To complete such a task, a globally orchestrated effort to collect more BGB samples data across all vegetation types is needed.

Given this lack of data, we instead compared the estimated uncertainty of our BGBC maps to that of our AGBC estimates to infer the sources of any divergence (Fig. 5). As expected, our cumulative BGBC uncertainty layer generally reveals greater overall uncertainty than our AGBC estimates, with BGBC uncertainty roughly twice that of AGBC throughout most of the globe. The highest absolute uncertainty was found in biomass rich forests. Arid woodlands, especially those of the Sahel and eastern Namibia, generally had the greatest relative BGBC uncertainty, though their absolute uncertainty was quite small (generally less than 3 MgC ha^{-1}). Here, biomass estimates of sparse woody vegetation were primarily responsible for heightened relative uncertainty.

Raster Layer	Description	Units	Scale Factor
agbc_2010.tif	Aboveground living biomass carbon stock density in 2010	MgC ha^{-1}	0.1
bgbc_2010.tif	Belowground living biomass carbon stock density in 2010	MgC ha^{-1}	0.1
agbc_2010_uncert.tif	Cumulative uncertainty (standard error) of aboveground living biomass carbon stock density in 2010 estimates	MgC ha^{-1}	0.1
Bgbc_2010_uncert.tif	Cumulative uncertainty (standard error) of belowground living biomass carbon stock density in 2010 estimates	MgC ha^{-1}	0.1

Table 7. Description of gridded data layers. Data layers should be multiplied by the scale factor to get raster values with units MgC ha^{-1}.

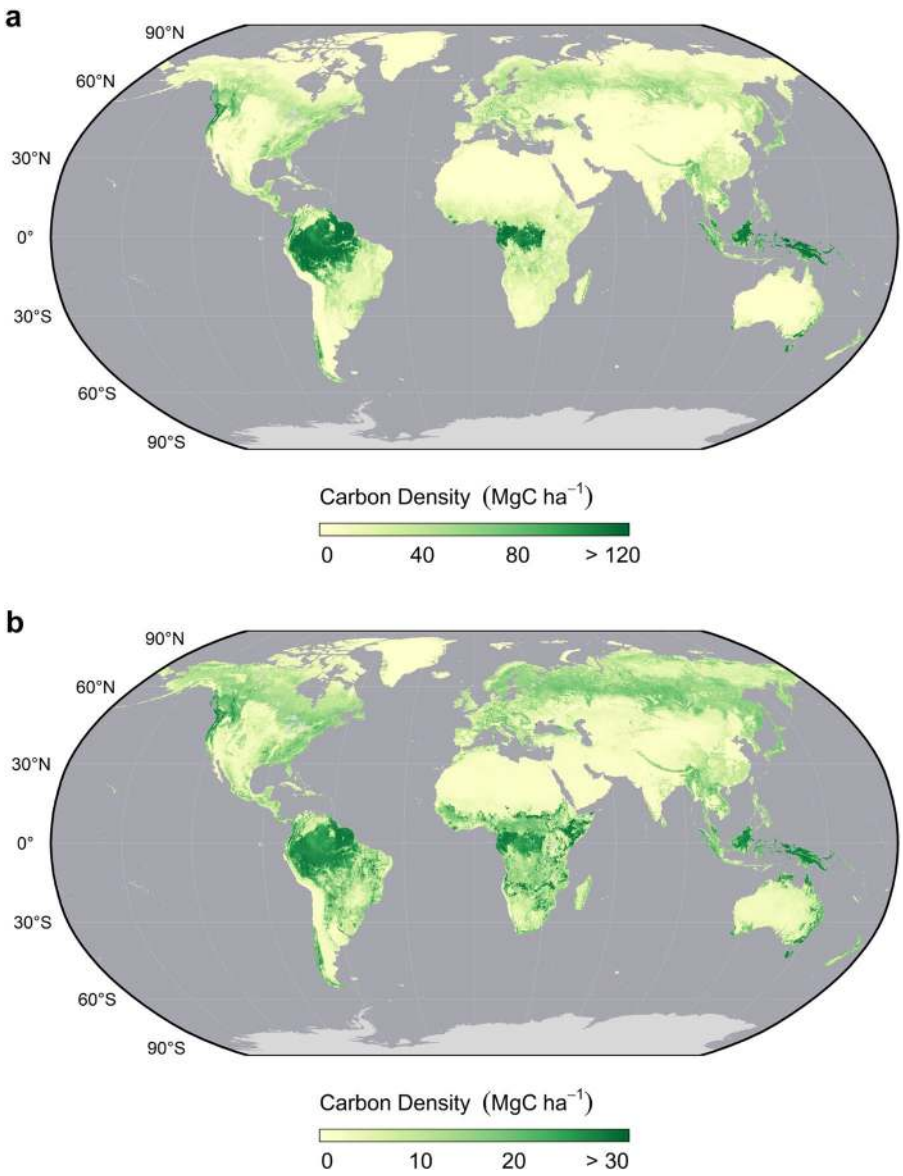

Fig. 4 Globally harmonized maps of above and belowground living biomass carbon densities. (**a**) Aboveground biomass carbon density (AGBC) and (**b**) belowground biomass carbon density (BGBC) are shown separately. Maps have been aggregated to a 5 km spatial resolution and reprojected here for visualization.

High relative and absolute BGBC uncertainty were also associated with predictions in select mountainous forests (e.g. east central Chile) as well as forested areas in and around cities. These patterns were largely driven by AGB uncertainty in the GlobBiomass product.

Biomass harmonization. The GlobBiomass global woody AGB map produced by Santoro *et al.*[30] comprises the backbone of our integrated products and, with few exceptions, remains largely unchanged in our final AGBC

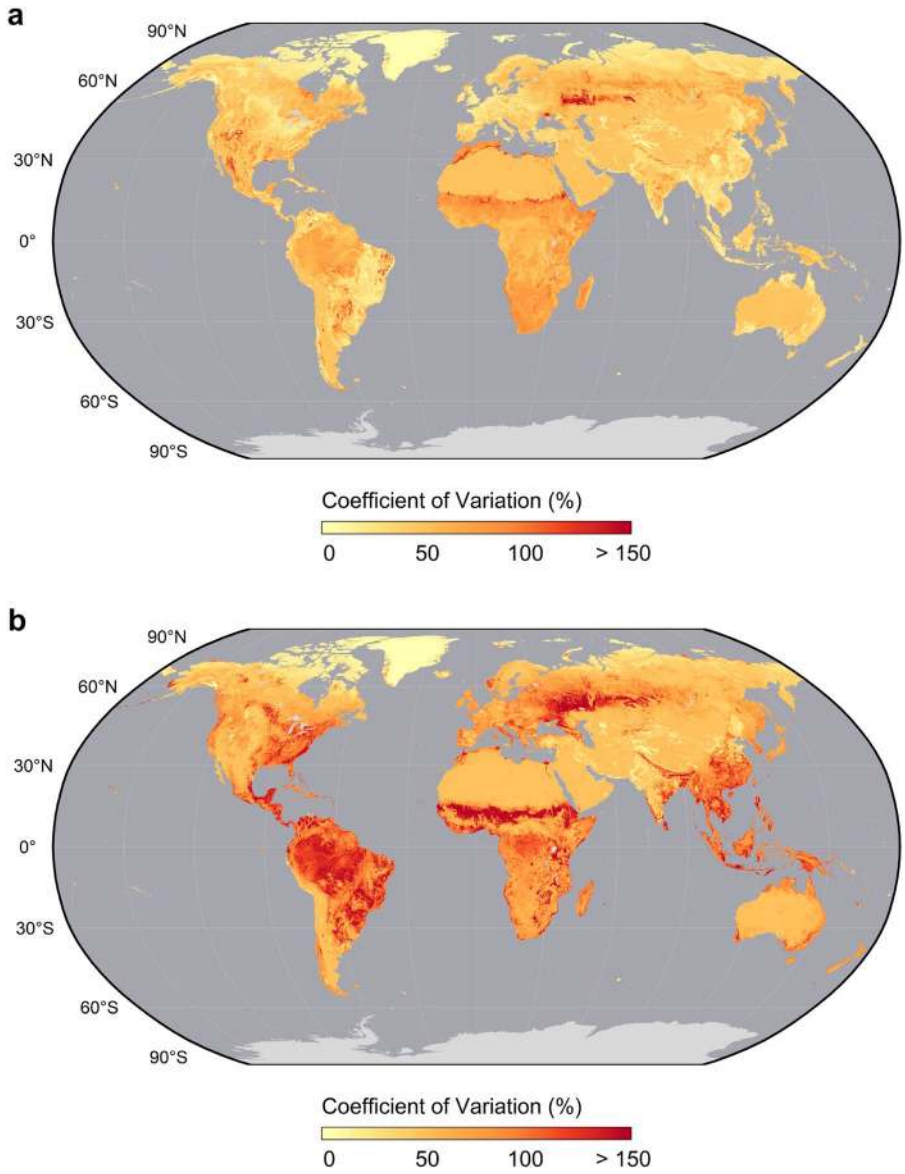

Fig. 5 Uncertainty of grid cell level above and belowground biomass carbon density estimates. Uncertainty is shown here as the coefficient of variation (%; standard error layer divided by mean estimate layer) of estimated AGBC (**a**) and BGBC (**b**) densities after harmonization. Maps have been aggregated to a 5 km spatial resolution and projected for visualization.

map. The native version of the GlobBiomass map is accompanied by an error layer describing the uncertainty of each pixel's biomass estimate and this too forms the core of our integrated uncertainty layers. In areas with tree cover, the global average error of GlobBiomass estimates is 39 Mg ha^{-1} or 50% with greater relative uncertainty in densely forested areas, along the margins of forested expanses like farm fields and cities, and in similar areas with sparse tree cover.

Adding additional grass or crop biomass in complementary proportion to a grid cell's tree cover often did not exceed the estimated error of the original GlobBiomass map (Fig. 7). Grid cells exceeding GlobBiomass's native uncertainty comprise less than 40% of its total extent. Exceptions were primarily found in grassland and cropland dominated regions where tree cover was generally sparse, and, consequently, the herbaceous biomass contribution was relatively high. Even so, the absolute magnitude of these additions remains somewhat small (less than 2.3 MgC ha^{-1} for grassland and 15 MgC ha^{-1} for cropland).

Larger deviations from GlobBiomass were also present in areas of both dryland Africa and the Arctic tundra biome, where we used independent layers to estimate woody biomass. In African drylands, GlobBiomass likely underestimates woody biomass by adopting the conservative FAO definition (DBH > 10 cm), which implicitly omits the relatively small trees and shrubs that are common to the region. The Bouvet map of Africa that we used to supplement these estimates is not bound by this constraint, was developed from region-specific data, and predicts substantially higher AGB density throughout much of its extent with comparatively high accuracy (RMSE = 17.1 Mg ha^{-1})[35].

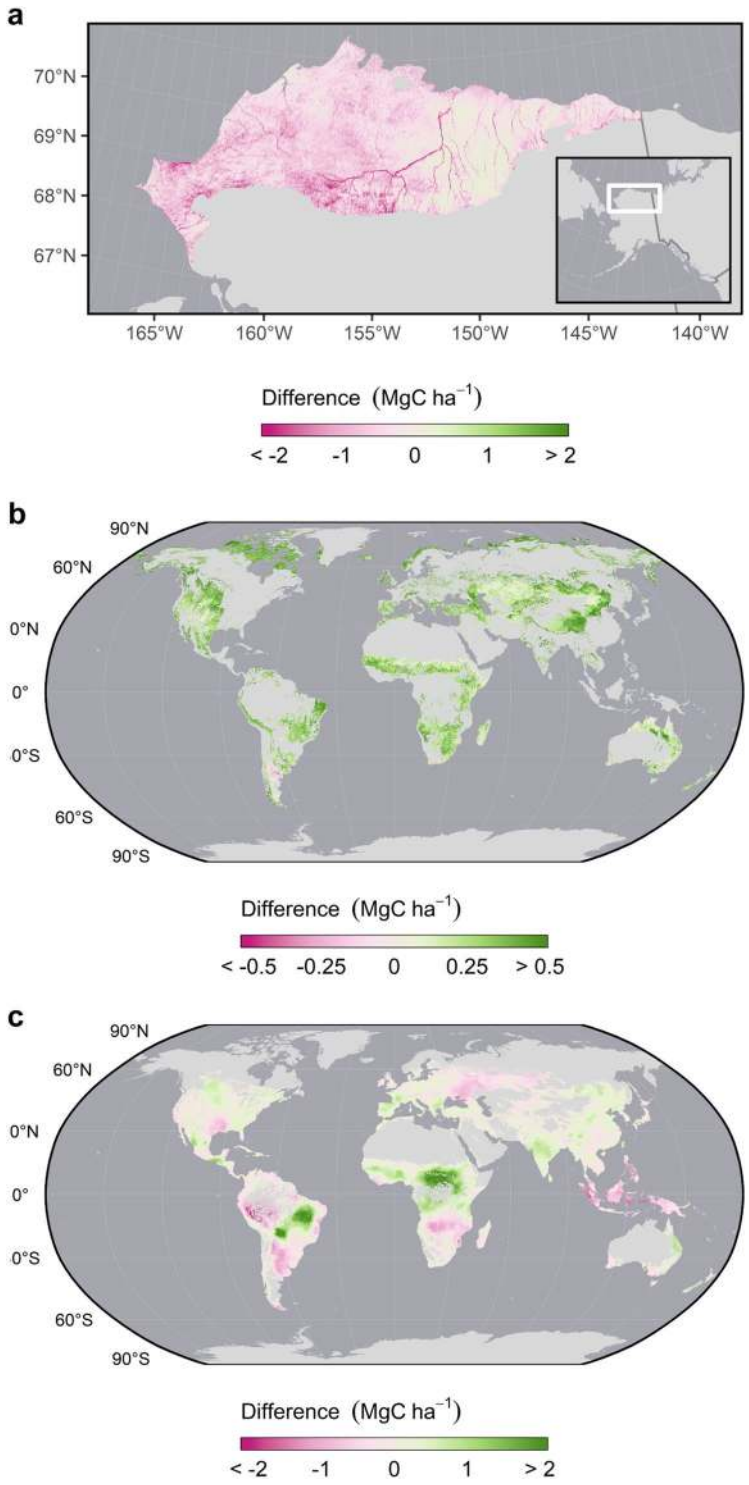

Fig. 6 Differences between landcover-specific AGBC estimates from the original published maps and the modified versions used as inputs to create the 2010 harmonized global maps. Tundra vegetation AGBC (**a**) is compared to the Landsat-based map of Berner *et al.*[45] for the north slope of Alaska after converting it to units MgC ha^{-1}. Here, the comparison map was subsequently aggregated to a 1 km resolution and reprojected for visualization. Grassland AGBC (**b**) is compared to the AVHRR-based map of Xia *et al.*[19] which represents the average estimate between 1982–2006. For visualization, the map was aggregated to a 5 km resolution and subsequently reprojected after being masked to MODIS IGBP grasslands in the year 2006[85] following Xia *et al.*[19]. As such, this map does not necessarily represent the spatial distribution of grid cells in which grassland estimates were used. Cropland AGBC (**c**) is compared to the original circa 2000 estimates to assess the effects of the 2000-to-2010 correction. The map is masked to the native extent of the combined yield maps and aggregated to a 5 km resolution for visualization. For all maps, negative values indicate that our circa 2010 estimates are lower than those of the earlier maps while positive values indicate higher estimates.

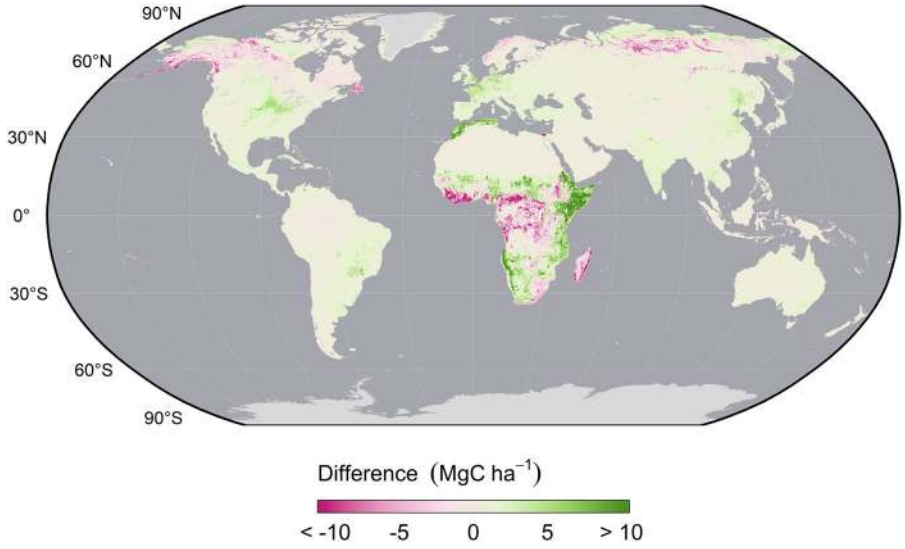

Fig. 7 Differences between the final harmonized AGBC map and GlobBiomass AGBC. GlobBiomass AGB was aggregated to a 300 m spatial resolution and converted to C density prior to comparison. Negative values indicate areas where the new map reports lower values than GlobBiomass while positive value denote higher estimates.

GlobBiomass also included sporadic biomass estimates throughout much of the Arctic tundra biome. Trees are generally scarce throughout this biome, which is instead dominated by dwarf shrubs and herbaceous forbs and graminoids, so given GlobBiomass's adherence to FAO guidelines, its predictions here may be spurious. We thus prioritized the estimates of the independent model developed specifically to collectively predict biomass of both woody and herbaceous tundra vegetation. These estimates were generally higher than GlobBiomass but agreed well with independent validation data from North America (RMSE = 2.9 Mg ha^{-1})[18].

Comparison with the IPCC Tier-1 global biomass carbon map. While far from a perfect comparison, the only other map to comprehensively report global biomass carbon density for all landcover types is the IPCC Tier-1 map for the year 2000 by Ruesch and Gibbs[28]. As previously described, this map was produced using an entirely different method ("stratify and multiply") and distinct data sources[23] and represents an earlier epoch. However, the map is widely used for myriad applications, and it may thus be informative to assess notable differences between it and our new products.

Ruesch and Gibbs[28] report total living C stocks of 345 petagrams (PgC) in AGBC and 133 PgC in BGBC for a total of 478 PgC, globally. Our estimates are lower at 287 PgC and 122 PgC in global AGBC and BGBC, respectively, for a total of 409 PgC in living global vegetation biomass. Herbaceous biomass in our maps comprised 9.1 and 28.3 PgC of total AGBC and BGBC, respectively. Half of all herbaceous AGBC (4.5 PgC) and roughly 6% of all herbaceous BGBC (1.7 PgC) was found in croplands. Moreover, we mapped 22.3 and 6.1 PgC, respectively, in the AGB and BGB of trees located within the cropland extent. These trees constituted roughly 7% of all global biomass C and are likely overlooked by both the Ruesch and Gibbs map[28] and by remotely sensed forest C maps that are masked to forested areas. Zomer *et al.*[17] first highlighted this potential discrepancy in the Ruesch and Gibbs map[28] when they produced a remarkably similar estimate of 34.2 Pg of overlooked C in cropland trees using Tier-1 defaults. However, their estimates were assumed to be in addition to the 474 PgC originally mapped by Ruesch and Gibbs[28]. Here, we suggest that the 28.4 PgC we mapped in cropland trees is already factored into our 409 PgC total.

Our AGBC product predicts substantially less biomass C than Ruesch and Gibbs[28] throughout most of the pantropical region and, to a lesser extent, southern temperate forests (Fig. 8a). This pattern has been noted by others comparing the Ruesch and Gibbs map[28] to other satellite-based biomass maps[73] and may suggest that the IPCC default values used to create it[23] are spatially biased. In addition, well-defined areas of high disagreement emerge in Africa that directly correspond with the FAO boundaries of the "tropical moist deciduous forest" ecofloristic zone and suggest that this area, in particular, may merit critical review. Moreover, the opposite pattern is observed in this same ecofloristic zone throughout South America. Our map also predicts greater AGBC throughout much of the boreal forest as well as in African shrublands and the steppes of South America.

We observed similar, though less pronounced discrepancies, when comparing BGBC maps (Fig. 8b). Notably, our map predicts substantially more BGBC throughout the tundra biome – a previously underappreciated C stock that has recently risen to prominance[74] – the boreal forest, African shrublands and most of South America and Australia. However, we predict less BGBC in nearly all rainforests (Temperate and Tropical). These differences and their distinct spatial patterns correspond with the vegetation strata used to make the IPCC Tier-1 map[28] and suggest that the accuracy of the "stratify and multiply" method depends heavily upon the quality of the referenced and spatial data considered. Inaccuracies in these data may, in turn, lead to false geographies. Integrating, continuous spatial estimates that better capture local and regional variation, as we have done, may thus greatly improve our understanding of global carbon geographies and their role in the earth system.

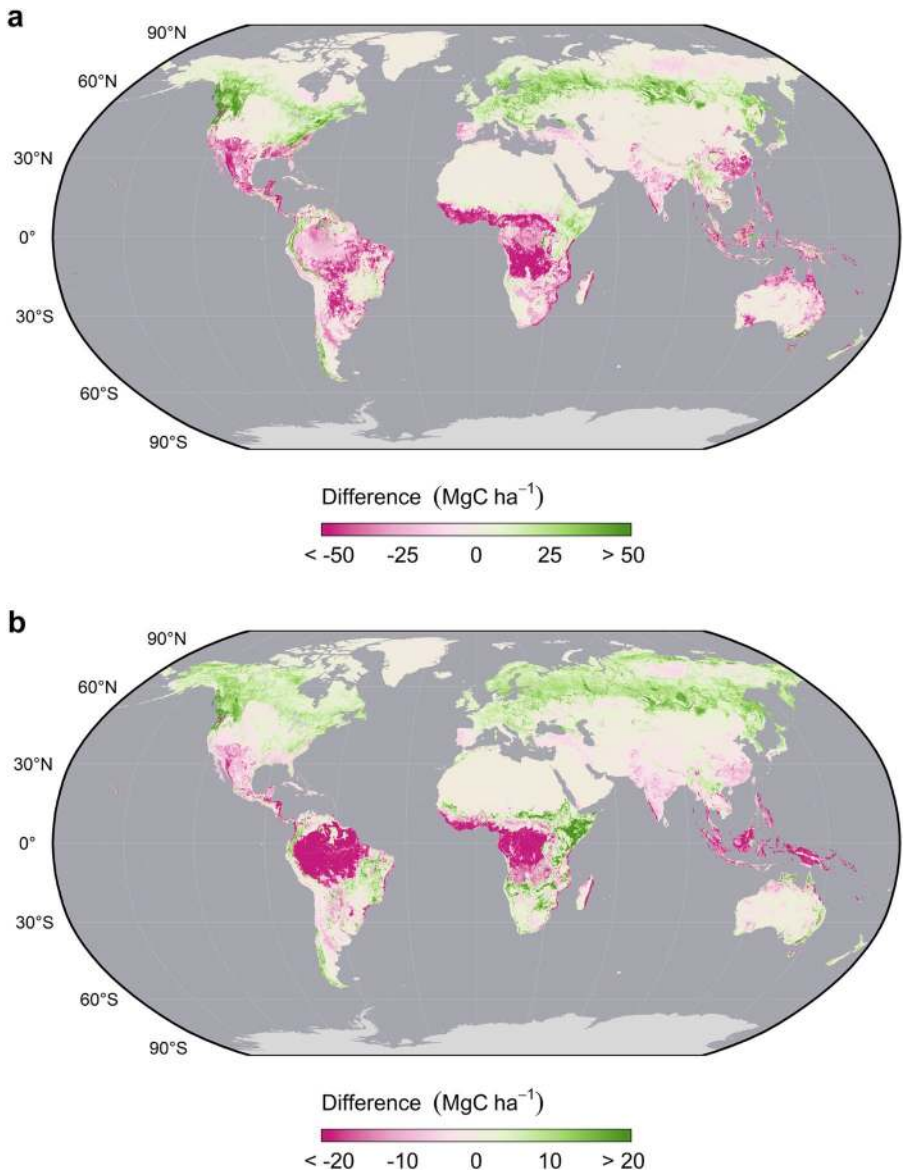

Fig. 8 Differences between the 2010 harmonized global maps of above and belowground biomass carbon density and those of the IPCC Tier-1 product by Ruesch and Gibbs for 2000. Comparisons of AGBC (**a**) and BGBC (**b**) maps are shown separately. Negative values indicate that the circa 2010 estimates are comparatively lower while positive values indicate higher estimates.

Congruence with IPCC Tier-2 and Tier-3 nationally reported woody carbon stocks. The error and variance between our woody biomass estimates – when aggregated to the country level – and comparable totals reported in the FRA were less for comparisons made against FRA estimates generated using higher tier IPCC methodologies than for those based on Tier-1 approaches (Fig. 9). Across the board for AGBC, BGBC, and total C comparisons, the relative RMSE ($RMSE_{CV}$) of our estimates, when compared to estimates generated using high tier methods, was roughly half of that obtained from comparisons with Tier-1 estimates (Table 8). Likewise, the coefficient of determination (R^2) was greatest for comparisons with Tier-3 estimates. For each pool-specific comparison (AGBC, BGBC, and total C), the slopes of the relationships between Tier-1, 2, and 3 estimates were neither significantly different from a 1:1 relationship nor from one another ($p > 0.05$; ANCOVA). Combined, these results suggest that our maps lead to C stock estimates congruent with those attained from independent, higher-tier reporting methodologies.

To explore this association at a finer regional scale, we also compared our woody C estimates to the United States Forest Service's Forest Inventory Analysis[75] (FIA) and found similarly strong congruence for AGBC and Total C stocks but subtle overestimates for BGBC (Fig. 9). The FIA is a Tier-3 inventory of woody forest biomass C stocks that is based on extensive and statistically rigorous field sampling and subsequent upscaling, We used data available at the state level for the year 2014 – again, the only year in which we could obtain data partitioned by AGBC and BGBC. Like our FRA comparison, we found a tight relationship between our woody AGBC totals and

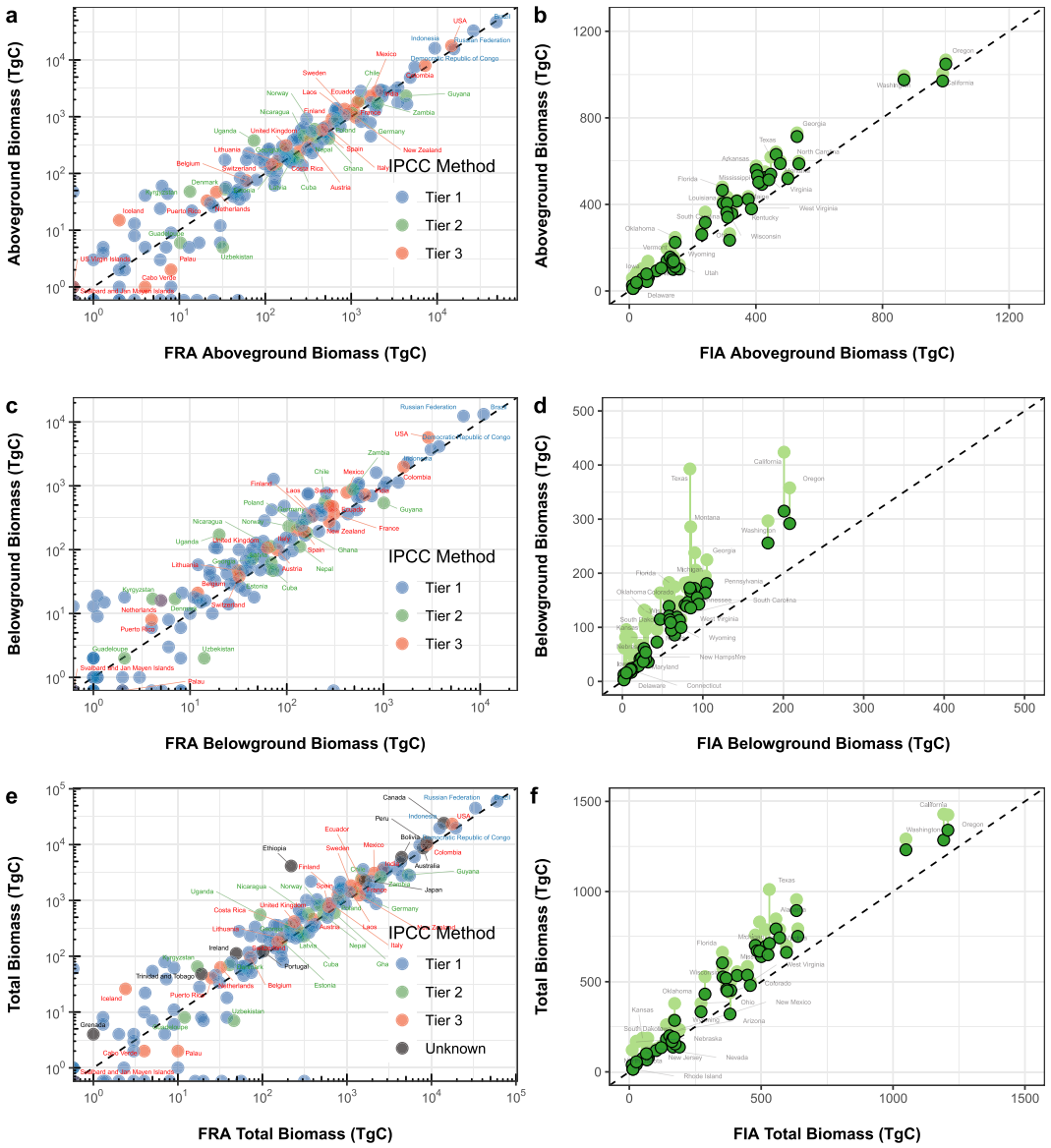

Fig. 9 Comparison of woody biomass density estimates to corresponding estimates of the FAO's FRA and the USFS's FIA. National woody AGBC totals derived from the woody components of our harmonized maps are compared to national totals reported in the 2015 FRA[62] (**a**) in relation to the IPCC inventory methodology used by each country. Likewise, we derived woody AGBC totals for US states and compared them to the corresponding totals reported by the 2014 FIA[75] (**b**), a Tier-3 inventory. We also show the additional effect of considering non-woody C – as is reported in our harmonized maps – in light green. Similar comparisons were made between our woody BGBC estimates and the corresponding estimates of both the FRA (**c**) and FIA (**d**). We further summed our woody AGBC and BGBC estimates and compared them to the total woody C stocks reported by both the FRA (**e**) and FIA (**f**).

those reported by the FIA (Fig. 9b; $RMSE_{CV} = 25.7\%$, $R^2 = 0.960$, slope = 1.10, n = 48). Our woody BGBC estimates, though, were systematically greater than those reported by the FIA (Fig. 9d; $RMSE_{CV} = 86.4\%$, $R^2 = 0.95$, slope = 1.51, n = 48). This trend has been noted by others[27] and suggests that the global model that we used to estimate woody BGBC may not be appropriate for some finer scale applications as is foretold by the elevated uncertainty reported in our corresponding uncertainty layer (Fig. 5b). Our total woody C (AGBC + BGBC) estimates (Fig. 9f), however, agreed well with the FIA ($RMSE_{CV} = 34.1\%$, $R^2 = 0.961$, slope = 1.17, n = 48) and thus reflect the outsized contribution of AGBC to the total woody C stock. When the contribution of herbaceous C stocks is further added to these comparisons, our stock estimates intuitively increase in rough proportion to a state's proportional extent of herbaceous cover. The effect of this addition is particularly pronounced for BGBC estimates due to the large root-to-shoot ratios of grassland vegetation.

The relative congruence of our results with higher-tier stock estimates suggests that our maps could be used to facilitate broader adoption of higher-tier methods among countries currently lacking the requisite data and those seeking to better account for C in non-woody biomass. This congruence spans a comprehensive range of

Reporting Method	N	Slope	R^2	$RMSE_{CV}$ (%)
Aboveground Woody Biomass Carbon				
Tier-1	136	0.983	0.884	23.6
Tier-2	18	0.949	0.819	13.0
Tier-3	25	0.999	0.963	10.5
"High Tier" (2 & 3)	43	0.987	0.931	11.7
Belowground Woody Biomass Carbon				
Tier-1	135	1.016	0.856	32.4
Tier-2	18	0.928	0.766	20.6
Tier-3	23	1.000	0.944	16.6
"High Tier" (2 & 3)	41	0.981	0.895	18.5
Total Woody Biomass Carbon				
Tier-1	136	0.983	0.853	26.4
Tier-2	18	0.946	0.816	13.2
Tier-3	25	0.997	0.960	11.2
"High Tier" (2 & 3)	43	0.984	0.927	12.1

Table 8. Statistical comparison of woody biomass carbon totals derived from the 2010 harmonized maps and those reported by the FRA in relation to the IPCC inventory methodology used. Statistics for AGBC, BGBC, and total C correspond to relationships depicted in Fig. 9a,c,f, respectively.

biophysical conditions and spatial scales ranging from small states to large nations. Moreover, a recent study suggests that the fidelity of the underlying GlobBiomass AGB map may extend to even finer scales[31]. While our BGBC estimates may differ from some fine-scale estimates (Fig. 9d), their tight agreement with high tier BGBC totals at the national level (Fig. 9c) suggests that they may still be well suited for many national-scale C inventories – especially for countries lacking requisite high tier data. Use of our maps is unlikely to introduce error in excess of that currently implicit in Tier-1 estimates. Credence, though, should be given to the associated uncertainty estimates. To facilitate wider adoption of higher-tier methodologies, our maps could be used to derive new, region-specific default values for use in Tier-2 frameworks[76] or to either represent or calibrate 2010 baseline conditions in Tier-3 frameworks. In so doing, inventories and studies alike could more accurately account for the nuanced global geographies of biomass C.

Usage Notes

These maps are intended for global applications in which continuous spatial estimates of live AGBC and/or BGBC density are needed that span a broad range of vegetation types and/or require estimates circa 2010. They are loosely based upon and share the spatial resolution of the ESA CCI Landcover 2010 map[37], which can be used to extract landcover specific C totals. However, our products notably do not account for C stored in non-living C pools like litter or coarse woody debris, nor soil organic matter, though these both represent large, additional ecosystem C stocks[77–79]. Our maps are explicitly intended for global scale applications seeking to consider C in the collective living biomass of *multiple* vegetation types. For global scale applications focused exclusively on the C stocks of a single vegetation type, we strongly encourage users to instead use the respective input map or model referenced in Table 1 to avoid potential errors that may have been introduced by our harmonization procedure. For AGB applications over smaller extents, users should further consider whether locally specific products are available. If such maps are not available and our maps are considered instead, credence should be given to their pixel-level uncertainty estimates. As mentioned above, the biomass of shrublands was only explicitly accounted for in Africa and the Arctic tundra, since neither broad-scale maps nor models generalizable to other areas were available in the existing literature. As such, we caution against the use of our maps outside of these areas when shrubland biomass is of particular interest or importance. Moreover, in contrast to the estimates for all other vegetation types considered, which we upscaled to a 300 m resolution, cropland C estimates were largely based on relatively coarse 8 km resolution data that were downscaled using bilinear resampling to achieve a 300 m spatial resolution. As such, these estimates may not adequately capture the underlying finer-scale spatial variation and should be interpreted with that in mind. Likewise, we reiterate that some BGBC estimates may differ from locally derived Tier-3 estimates, and attention should thus be given to our reported pixel-level uncertainty for all applications. Finally, our maps should not be used in comparison with the IPCC Tier-1 map of Ruesch and Gibbs (2008) to detect biomass change between the two study periods due to significant methodological differences between these products.

Code availability

Cropland biomass maps were created in the R statistical computing environment[80]. All other coding was done in Google Earth Engine[81] (GEE), wherein our workflow consisted of 18 interconnected scripts. All code can be found on GitHub (https://github.com/sethspawn/globalBiomassC) and permanently archived by Zenodo[82].

Received: 3 July 2019; Accepted: 14 February 2020;
Published online: 06 April 2020

References

1. Houghton, R. A., Hall, F. & Goetz, S. J. Importance of biomass in the global carbon cycle. *J. Geophys. Res. Biogeosciences* **114** (2009).
2. Huntzinger, D. N. *et al*. The North American Carbon Program Multi-Scale Synthesis and Terrestrial Model Intercomparison Project – Part 1: Overview and experimental design. *Geosci. Model Dev* **6**, 2121–2133 (2013).
3. Schwalm, C. R. *et al*. Toward "optimal" integration of terrestrial biosphere models. *Geophys. Res. Lett.* **42**, 4418–4428 (2015).
4. Li, W. *et al*. Land-use and land-cover change carbon emissions between 1901 and 2012 constrained by biomass observations. *Biogeosciences* **14**, 5053–5067 (2017).
5. Spawn, S. A., Lark, T. J. & Gibbs, H. K. Carbon emissions from cropland expansion in the United States. *Environ. Res. Lett.* **14**, 045009 (2019).
6. Harris, N. L. *et al*. Baseline Map of Carbon Emissions from Deforestation in Tropical Regions. *Science* **336**, 1573–1576 (2012).
7. Baccini, A. *et al*. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. *Science* **358**, 230–234 (2017).
8. Strassburg, B. B. N. *et al*. Global congruence of carbon storage and biodiversity in terrestrial ecosystems. *Conserv. Lett* **3**, 98–105 (2010).
9. West, P. C. *et al*. Trading carbon for food: Global comparison of carbon stocks vs. crop yields on agricultural land. *Proc. Natl. Acad. Sci.* **107**, 19645–19648 (2010).
10. Carvalhais, N. *et al*. Global covariation of carbon turnover times with climate in terrestrial ecosystems. *Nature* **514**, 213–217 (2014).
11. Brandão, A. *et al*. Estimating the Potential for Conservation and Farming in the Amazon and Cerrado under Four Policy Scenarios. *Sustainability* **12**, 1277 (2020).
12. Gibbs, H. K., Brown, S., Niles, J. O. & Foley, J. A. Monitoring and estimating tropical forest carbon stocks: making REDD a reality. *Environ. Res. Lett.* **2**, 045023 (2007).
13. Fargione, J. E. *et al*. Natural climate solutions for the United States. *Sci. Adv.* **4**, eaat1869 (2018).
14. Griscom, B. W. *et al*. Natural climate solutions. *Proc. Natl. Acad. Sci.* **114**, 11645–11650 (2017).
15. Goetz, S. J. *et al*. Mapping and monitoring carbon stocks with satellite observations: a comparison of methods. *Carbon Balance Manag* **4**, 2 (2009).
16. Xiao, J. *et al*. Remote sensing of the terrestrial carbon cycle: A review of advances over 50 years. *Remote Sens. Environ.* **233**, 111383 (2019).
17. Zomer, R. J. *et al*. Global Tree Cover and Biomass Carbon on Agricultural Land: The contribution of agroforestry to global and national carbon budgets. *Sci. Rep* **6**, 29987 (2016).
18. Berner, L. T., Jantz, P., Tape, K. D. & Goetz, S. J. Tundra plant above-ground biomass and shrub dominance mapped across the North Slope of Alaska. *Environ. Res. Lett.* **13**, 035002 (2018).
19. Xia, J. *et al*. Spatio-Temporal Patterns and Climate Variables Controlling of Biomass Carbon Stock of Global Grassland Ecosystems from 1982 to 2006. *Remote Sens* **6**, 1783–1802 (2014).
20. Monfreda, C., Ramankutty, N. & Foley, J. A. Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. *Glob. Biogeochem. Cycles* **22**, GB1022 (2008).
21. Wolf, J. *et al*. Biogenic carbon fluxes from global agricultural production and consumption. *Glob. Biogeochem. Cycles* **29**, 1617–1639 (2015).
22. Mokany, K., Raison, R. J. & Prokushkin, A. S. Critical analysis of root: shoot ratios in terrestrial biomes. *Glob. Change Biol.* **12**, 84–96 (2006).
23. IPCC 2006. *2006 IPCC Guidelines for National Greenhouse Gas Inventories*. vol. 4 (IPCC National Greenhouse Gas Inventories Programme, 2006).
24. IPCC 2019. *2019 Refinement to the 2006 IPCC Guidlines for National Greenhouse Gas Inventories*. vol. 4 (IPCC National Greenhouse Gas Inventories Programme, 2019).
25. Reich, P. B. *et al*. Temperature drives global patterns in forest biomass distribution in leaves, stems, and roots. *Proc. Natl. Acad. Sci.* **111**, 13721–13726 (2014).
26. Wang, P. *et al*. Belowground plant biomass allocation in tundra ecosystems and its relationship with temperature. *Environ. Res. Lett.* **11**, 055003 (2016).
27. Russell, M. B., Domke, G. M., Woodall, C. W. & D'Amato, A. W. Comparisons of allometric and climate-derived estimates of tree coarse root carbon stocks in forests of the United States. *Carbon Balance Manag* **10**, 20 (2015).
28. Ruesch, A. & Gibbs, H. New IPCC Tier-1 Global Biomass Carbon Map for the Year 2000. *Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory*, http://cdiac.ess-dive.lbl.gov (2008).
29. Schimel, D. *et al*. Observing terrestrial ecosystems and the carbon cycle from space. *Glob. Change Biol.* **21**, 1762–1776 (2015).
30. Santoro, M. *et al*. GlobBiomass - global datasets of forest biomass. *PANGAEA* https://doi.org/10.1594/PANGAEA.894711 (2018).
31. Huang, W. *et al*. High-Resolution Mapping of Aboveground Biomass for Forest Carbon Monitoring System in the 3 Tri-State Region of Maryland, Pennsylvania and Delaware, USA. *Environ. Res. Lett.* **14**, 095002 (2019).
32. Food and Agricultural Organization. *FRA 2015 Terms and Definitions*. (Food and Agricultural Organization of the United Nations, 2012).
33. Quegan, S. *et al*. DUE GlobBiomass: D6 - Global Biomass Map Algorithm Theoretical Basis Document. *GlobBiomass*, http://globbiomass.org/wp-content/uploads/DOC/Deliverables/D6_D7/GlobBiomass_D6_7_Global_ATBD_v2.pdf (2017).
34. Hansen, M. C. *et al*. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **342**, 850–853 (2013).
35. Bouvet, A. *et al*. An above-ground biomass map of African savannahs and woodlands at 25 m resolution derived from ALOS PALSAR. *Remote Sens. Environ.* **206**, 156–173 (2018).
36. Le Toan, T., Beaudoin, A., Riom, J. & Guyon, D. Relating forest biomass to SAR data. *IEEE Trans. Geosci. Remote Sens* **30**, 403–411 (1992).
37. European Space Agency. 300 m Annual global land cover time series from 1992 to 2015. *European Space Agency - Climate Change Initiative*, http://maps.elie.ucl.ac.be/CCI/viewer/download.php (2017).
38. Bartholomé, E. & Belward, A. S. GLC2000: a new approach to global land cover mapping from Earth observation data. *Int. J. Remote Sens.* **26**, 1959–1977 (2005).
39. Avitabile, V. *et al*. An integrated pan-tropical biomass map using multiple reference datasets. *Glob. Change Biol.* **22**, 1406–1420 (2016).
40. Englund, O. *et al*. A new high-resolution nationwide aboveground carbon map for Brazil. *Geo Geogr. Environ.* **4**, e00045 (2017).
41. Scholze, M., Buchwitz, M., Dorigo, W., Guanter, L. & Quegan, S. Reviews and syntheses: Systematic Earth observations for use in terrestrial carbon cycle data assimilation systems. *Biogeosciences* **14**, 3401–3429 (2017).
42. Martin, A. R., Doraisami, M. & Thomas, S. C. Global patterns in wood carbon concentration across the world's trees and forests. *Nat. Geosci.* **11**, 915 (2018).
43. Kottek, M., Grieser, J., Beck, C., Rudolf, B. & Rubel, F. World Map of the Köppen-Geiger climate classification updated. *Meteorol. Z.* **15**, 259–263 (2006).
44. Olofsson, P. *et al*. Good practices for estimating area and assessing accuracy of land change. *Remote Sens. Environ.* **148**, 42–57 (2014).

45. Berner, L. T., Jantz, P., Tape, K. D. & Goetz, S. J. ABoVE: Gridded 30-m Aboveground Biomass, Shrub Dominance, North Slope, AK, 2007–2016. *Oak Ridge National Laboratory Distributed Active Archive Center* https://doi.org/10.3334/ORNLDAAC/1565 (2018).
46. Vermote, E. F. & Wolfe, R. MYD09GQ MODIS/Aqua Surface Reflectance Daily L2G Global 250 m SIN Grid V006. *NASA EOSDIS Land Processes Distributed Active Archive Center*, https://doi.org/10.5067/MODIS/MYD09GQ.006 (2015).
47. Vermote, E. F. & Wolfe, R. MOD09GQ MODIS/Terra Surface Reflectance Daily L2G Global 250 m SIN Grid V006. *NASA EOSDIS Land Processes Distributed Active Archive Center*, https://doi.org/10.5067/MODIS/MOD09GQ.006 (2015).
48. Steven, M. D., Malthus, T. J., Baret, F., Xu, H. & Chopping, M. J. Intercalibration of vegetation indices from different sensor systems. *Remote Sens. Environ.* **88**, 412–422 (2003).
49. Adler, P. B. *et al*. Productivity Is a Poor Predictor of Plant Species Richness. *Science* **333**, 1750–1753 (2011).
50. Didan, K. MYD13Q1 MODIS/Aqua Vegetation Indices 16-Day L3 Global 250 m SIN Grid V006. *NASA EOSDIS Land Processes Distributed Active Archive Center*, https://doi.org/10.5067/MODIS/MYD13Q1.006 (2015).
51. Didan, K. MOD13Q1 MODIS/Terra Vegetation Indices 16-Day L3 Global 250 m SIN Grid V006. *NASA EOSDIS Land Processes Distributed Active Archive Center* https://doi.org/10.5067/MODIS/MOD13Q1.006 (2015).
52. Fensholt, R. & Proud, S. R. Evaluation of Earth Observation based global long term vegetation trends — Comparing GIMMS and MODIS global NDVI time series. *Remote Sens. Environ.* **119**, 131–147 (2012).
53. Li, Z. *et al*. Comparing cropland net primary production estimates from inventory, a satellite-based model, and a process-based model in the Midwest of the United States. *Ecol. Model.* **277**, 1–12 (2014).
54. Turner, D. P. *et al*. Evaluation of MODIS NPP and GPP products across multiple biomes. *Remote Sens. Environ.* **102**, 282–292 (2006).
55. Ramankutty, N., Evan, A. T., Monfreda, C. & Foley, J. A. Farming the planet: 1. Geographic distribution of global agricultural lands in the year 2000. *Glob. Biogeochem. Cycles* **22**, GB1003 (2008).
56. Grassini, P., Eskridge, K. M. & Cassman, K. G. Distinguishing between yield advances and yield plateaus in historical crop production trends. *Nat. Commun.* **4**, 2918 (2013).
57. Gray, J. M. *et al*. Direct human influence on atmospheric CO_2 seasonality from increased cropland productivity. *Nature* **515**, 398–401 (2014).
58. Running, S. W., Mu, Q. & Zhao, M. MOD17A3H MODIS/Terra Net Primary Production Yearly L4 Global 1 km SIN Grid V055. *NASA EOSDIS Land Processes Distributed Active Archive Center*, https://lpdaac.usgs.gov/products/mod17a3v055/ (2015).
59. Fick, S. & Hijmans, R. WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *Int. J. Climatol.* **37**, 4302–4315 (2017).
60. Potapov, P. *et al*. The last frontiers of wilderness: Tracking loss of intact forest landscapes from 2000 to 2013. *Sci. Adv.* **3**, e1600821 (2017).
61. Harris, N. L., Goldman, E. D. & Gibbes, S. Spatial Database of Planted Trees (SDPT Version 1.0). *World Resources Institute*, https://www.wri.org/publication/planted-trees (2019).
62. Food and Agricultural Organization. *Global Forest Resources Assessment 2015: Desk Reference*. (Food and Agricultural Organization of the United Nations, 2015).
63. Dimiceli, C. *et al*. MOD44B MODIS/Terra Vegetation Continuous Fields Yearly L3 Global 250 m SIN Grid V006. *NASA EOSDIS Land Processes Distributed Active Archive Center*, https://doi.org/10.5067/MODIS/MOD44B.006 (2015).
64. Sexton, J. O. *et al*. Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. *Int. J. Digit. Earth* **6**, 427–448 (2013).
65. Dinerstein, E. *et al*. An Ecoregion-Based Approach to Protecting Half the Terrestrial Realm. *BioScience* **67**, 534–545 (2017).
66. Spawn, S. A. & Gibbs, H. K. Global Aboveground and Belowground Biomass Carbon Density Maps for the Year 2010. *Oak Ridge National Laboratory Distributed Active Archive Center*, https://doi.org/10.3334/ORNLDAAC/1763 (2019).
67. Spawn, S. A., Sullivan, C. C., Lark, T. J. & Gibbs, H. K. Harmonized global maps of above and belowground biomass carbon density in the year 2010. *figshare* https://doi.org/10.6084/m9.figshare.c.4561940 (2020).
68. Gao, Q. *et al*. Climatic change controls productivity variation in global grasslands. *Sci. Rep* **6**, 26958 (2016).
69. de Jong, R., Verbesselt, J., Schaepman, M. E. & de Bruin, S. Trend changes in global greening and browning: contribution of short-term trends to longer-term change. *Glob. Change Biol* **18**, 642–655 (2012).
70. Gonsamo, A., Chen, J. M. & Lombardozzi, D. Global vegetation productivity response to climatic oscillations during the satellite era. *Glob. Change Biol.* **22**, 3414–3426 (2016).
71. Ray, D. K. *et al*. Climate change has likely already affected global food production. *Plos One* **14**, e0217148 (2019).
72. Lobell, D. B., Schlenker, W. & Costa-Roberts, J. Climate Trends and Global Crop Production Since 1980. *Science* **333**, 616–620 (2011).
73. Hu, T. *et al*. Mapping Global Forest Aboveground Biomass with Spaceborne LiDAR, Optical Imagery, and Forest Inventory Data. *Remote Sens* **8**, 565 (2016).
74. Iversen, C. M. *et al*. The unseen iceberg: plant roots in arctic tundra. *New Phytol.* **205**, 34–58 (2015).
75. USDA Forest Service. Forest Inventory and Analysis National Program: Standard Tables of Forest Caron Stock Estimates by State. *Forest Inventory and Analysis National Program*, https://www.fia.fs.fed.us/forestcarbon/index.php (2014).
76. 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? *Environ. Res. Lett.* **9**, 124008 (2014).
77. Jobbágy, E. G. & Jackson, R. B. The Vertical Distribution of Soil Organic Carbon and Its Relation to Climate and Vegetation. *Ecol. Appl.* **10**, 423–436 (2000).
78. Scharlemann, J. P., Tanner, E. V., Hiederer, R. & Kapos, V. Global soil carbon: understanding and managing the largest terrestrial carbon pool. *Carbon Manag* **5**, 81–91 (2014).
79. Domke, G. M., Woodall, C. W., Walters, B. F. & Smith, J. E. From Models to Measurements: Comparing Downed Dead Wood Carbon Stock Estimates in the U.S. Forest Inventory. *Plos One* **8**, e59949 (2013).
80. R Core Team. R: A Language and Environment for Statistical Computing, https://www.R-project.org/ (2017).
81. Gorelick, N. *et al*. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **202**, 18–27 (2017).
82. Spawn, S. A. sethspawn/globalBiomassC. *Zenodo* https://doi.org/10.5281/zenodo.3647567 (2020).
83. Olson, D. M. *et al*. Terrestrial Ecoregions of the World: A New Map of Life on EarthA new global map of terrestrial ecoregions provides an innovative tool for conserving biodiversity. *BioScience* **51**, 933–938 (2001).
84. European Space Agency. Land Cover CCI Product User Guide Version 2, D3.4-PUG, v2.5. *European Space Agency - Climate Change Initiative*, http://maps.elie.ucl.ac.be/CCI/viewer/download/ESACCI-LC-PUG-v2.5.pdf (2016).
85. Friedl, M. A. & Sulla-Menashe, D. MCD12Q1 MODIS/Terra + Aqua Land Cover Type Yearly L3 Global 500 m SIN Grid V006. *NASA EOSDIS Land Processes Distributed Active Archive Center*, https://doi.org/10.5067/MODIS/MCD12Q1.006 (2019).
86. Jing, Q., Bélanger, G., Baron, V. & Bonesmo, H. Modeling the Biomass and Harvest Index Dynamics of Timothy. *Agron. J.* **103**, 1397–1404 (2011).
87. West, T. O. *et al*. Cropland carbon fluxes in the United States: increasing geospatial resolution of inventory-based carbon accounting. *Ecol. Appl.* **20**, 1074–1086 (2010).
88. Unkovich, M., Baldock, J. & Forbes, M. Variability in harvest index of grain crops and potential significance for carbon accounting: examples from Australian agriculture. *Adv. Agron* **105**, 173–219 (2010).

89. Hay, R. K. M. Harvest index: a review of its use in plant breeding and crop physiology. *Ann. Appl. Biol.* **126**, 197–216 (1995).
90. Larcher, W. *Physiological Plant Ecology: Ecophysiology and Stress Physiology of Functional Groups*. (Springer-Verlag, 2003).
91. Hakala, K., Keskitalo, M. & Eriksson, C. Nutrient uptake and biomass accumulation for eleven different field crops. *Agric. Food Sci* **18**, 366–387 (2009).
92. Bolinder, M. A., Janzen, H. H., Gregorich, E. G., Angers, D. A. & VandenBygaart, A. J. An approach for estimating net primary productivity and annual carbon inputs to soil for common agricultural crops in Canada. *Agric. Ecosyst. Environ* **118**, 29–42 (2007).
93. Mackenzie, B. A. & Van Fossen, L. Managing Dry Grain In Storage. In *Agricultural Engineers' Digest* vol. 20 (Purdue University Cooperative Extension Service, 1995).
94. Goodwin, M. Crop Profile for Dry Bean in Canada. *Agriculture and Agri-Food Canada*, http://publications.gc.ca/collections/collection_2009/agr/A118-10-4-2005E.pdf (2005).
95. Schulte auf'm Erley, G., Kaul, H.-P., Kruse, M. & Aufhammer, W. Yield and nitrogen utilization efficiency of the pseudocereals amaranth, quinoa, and buckwheat under differing nitrogen fertilization. *Eur. J. Agron.* **22**, 95–100 (2005).
96. Bjorkman, T. Northeast Buckwheat Growers Newsletter No. 19. *Cornell University NYSAES*, http://www.hort.cornell.edu/bjorkman/lab/buck/NL/june05.php (2005).
97. Kyle, G. P. *et al. GCAM 3.0 Agriculture and Land Use: Data Sources and Methods*, https://doi.org/10.2172/1036082 (2011).
98. Bastin, S. & Henken, K. Water Content of Fruits and Vegetables. *University of Kentucky Cooperative Extension Service*, https://www.academia.edu/5729963/Water_Content_of_Fruits_and_Vegetables (1997).
99. Smil, V. Crop Residues: Agriculture's Largest HarvestCrop residues incorporate more than half of the world's agricultural phytomass. *BioScience* **49**, 299–308 (1999).
100. Squire, G. R. *The physiology of tropical crop production*. (C.A.B. International, 1990).
101. Williams, J. R. *et al*. EPIC users guide v. 0509. *Texas A & M University Blackland Research and Extension Center*, http://epicapex.tamu.edu/files/2013/02/epic0509usermanualupdated.pdf (2006).
102. Okeke, J. E. Cassava varietal improvement for processing and utilization in livestock feeds. In *Cassava as Livestock Feed in Africa* (International Institute of Tropical Agriculture, 1992).
103. Pongsawatmanit, R., Thanasukarn, P. & Ikeda, S. Effect of Sucrose on RVA Viscosity Parameters, Water Activity and Freezable Water Fraction of Cassava Starch Suspensions. *ScienceAsia* **28**, 129–134 (2002).
104. Gigou, J. *et al*. Fonio Millet (Digitaria Exilis) Response to N, P and K Fertilizers Under Varying Climatic Conditions in West. *AFRICA. Exp. Agric* **45**, 401–415 (2009).
105. Food and Agricultural Organization. FAOSTAT 2001: FAO statistical databasees. *FAOSTAT*, http://www.fao.org/faostat/en/#data (2006).
106. Bolinder, M. A., Angers, D. A., Bélanger, G., Michaud, R. & Laverdière, M. R. Root biomass and shoot to root ratios of perennial forage crops in eastern Canada. *Can. J. Plant Sci.* **82**, 731–737 (2002).
107. Deferne, J. & Pate, D. W. Hemp seed oil: A source of valuable essential fatty acids. *J. Int. Hemp Assoc* **3**, 4–7 (1996).
108. Islam, Md. R. *et al*. Study of Harvest Index and Genetic Variability in White Jute (Corchorus capsularis) Germplasm. *J. Biol. Sci.* **2**, 358–360 (2002).
109. Ahad, A. & Debnath, C. N. Shoot Root Ratio of Jute Varieties and the Nature of Association Between Root Characteristics and the Yield of Dry Matter and Fiber. *Bangladesh J. Agric. Res* **13**, 17–22 (1988).
110. Mondal, S. S., Ghosh, A. & Debabrata, A. Effect of seeding time of linseed (Linum usitatissimum) in rice (Oryza sativa)-based paira cropping system under rainfed lowland condition. *Indian J. Agric. Sci* **75**, 134–137 (2005).
111. Ayaz, S., Moot, D. J., Mckenzie, B. A., Hill, G. D. & Mcneil, D. L. The Use of a Principal Axis Model to Examine Individual Plant Harvest Index in Four Grain Legumes. *Ann. Bot.* **94**, 385–392 (2004).
112. Goudriaan, J. & Van Laar, H. H. Development and growth. In *Modelling Potential Crop Growth Processes: Textbook with Exercises* (eds. Goudriaan, J. & Van Laar, H. H.) 69–94 (Springer Netherlands, 1994).
113. National Research Council. *Nutrient Requirements of Nonhuman Primates: Second Revised Edition*. (The National Academies Press, 2003).
114. Roth, C. M., Shroyer, J. P. & Paulsen, G. M. Allelopathy of Sorghum on Wheat under Several Tillage Systems. *Agron. J.* **92**, 855–860 (2000).
115. Heidari Zooleh, H. *et al*. Effect of alternate irrigation on root-divided Foxtail Millet (Setaria italica). *Aust. J. Crop Sci* **5**, 205–2013 (2011).
116. Brück, H., Sattelmacher, B. & Payne, W. A. Varietal differences in shoot and rooting parameters of pearl millet on sandy soils in Niger. *Plant Soil* **251**, 175–185 (2003).
117. Oelke, E. A., Putnam, D. H., Teynor, T. M. & Oplinger, E. S. Quinoa. In *Alternative Field Crops Manual* (University of Wisconsin-Extension, Cooperative Extension, 1992).
118. Robertson, M. J., Silim, S., Chauhan, Y. S. & Ranganathan, R. Predicting growth and development of pigeonpea: biomass accumulation and partitioning. *Field Crops Res* **70**, 89–100 (2001).
119. Armstrong, E. Desiccation & harvest of field peas. In *Pulse management in Southern New South Wales* (State of New South Wales Agriculture, 1999).
120. Fischer, R. A. (Tony) & Edmeades, G. O. Breeding and Cereal Yield Progress. *Crop Sci.* **50**, S-85–S-98 (2010).
121. Atlin, G. N. *et al*. Developing rice cultivars for high-fertility upland systems in the Asian tropics. *Field Crops Res* **97**, 43–52 (2006).
122. Bueno, C. S. & Lafarge, T. Higher crop performance of rice hybrids than of elite inbreds in the tropics: 1. Hybrids accumulate more biomass during each phenological phase. *Field Crops Res* **112**, 229–237 (2009).
123. Yang, J. & Zhang, J. Crop management techniques to enhance harvest index in rice. *J. Exp. Bot* **61**, 3177–3189 (2010).
124. Ziska, L. H., Namuco, O., Moya, T. & Quilang, J. Growth and Yield Response of Field-Grown Tropical Rice to Increasing Carbon Dioxide and Air Temperature. *Agron. J.* **89**, 45–53 (1997).
125. Mwaja, V. N., Masiunas, J. B. & Weston, L. A. Effects of fertility on biomass, phytotoxicity, and allelochemical content of cereal rye. *J. Chem. Ecol.* **21**, 81–96 (1995).
126. Bruinsma, J. & Schuurman, J. J. The effect of spraying with DNOC (4,6-dinitro-o-cresol) on the growth of the roots and shoots of winter rye plants. *Plant Soil* **24**, 309–316 (1966).
127. Yau, S. K., Sidahmed, M. & Haidar, M. Conservation versus Conventional Tillage on Performance of Three Different Crops. *Agron. J.* **102**, 269–276 (2010).
128. Hojati, M., Modarres-Sanavy, S. A. M., Karimi, M. & Ghanati, F. Responses of growth and antioxidant systems in Carthamustinctorius L. under water deficit stress. *Acta Physiol. Plant.* **33**, 105–112 (2011).
129. Oelke, E. A. *et al*. Safflower. In *Alternative Field Crops Manual* (University of Wisconsin-Extension, Cooperative Extension, 1992).
130. Perez, R. Chapter 3: Sugar cane. In *Feeding pigs in the tropics* (Food and Agricultural Organization of the United Nations, 1997).
131. Van Dillewijn, C. *Botany of Sugarcane*. (Chronica Botanica Co, 1952).
132. Pate, F. M., Alvarez, J., Phillips, J. D. & Eiland, B. R. *Sugarcane as a Cattle Feed: Production and Utilization*. (University of Florida Extension Institute of Food and Agricultural Sciences, 2002).

Acknowledgements
We gratefully acknowledge all the data producers, without whom this work would not be possible. We especially thank Maurizio Santoro *et al*., Alexandre Bouvet *et al*., Jiangzhou Xia *et al*., Logan T. Berner *et al*., Chad Monfreda *et al*., and Julie Wolf *et al*. whose AGB estimates comprise the core of our harmonized products and, in many cases, whose feedback greatly improved the quality of their inclusion. We are also grateful to the thoughtful feedback of three anonymous reviewers whose suggestions greatly improved the quality of our products and the clarity of our manuscript. Funding for this project was generously provided by the David and Lucile Packard Foundation and the National Wildlife Federation.

Author contributions
S.A.S. designed the harmonization procedure, compiled and standardized individual biomass layers, conducted all mapping, and led manuscript development. C.C.S., T.J.L. and H.K.G. assisted with conceptualization, and manuscript development.

Competing interests
The authors declare no competing interests.

Additional information
Correspondence and requests for materials should be addressed to S.A.S.

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.