

Benchmark map of forest carbon stocks in tropical regions across three continents

Sassan S. Saatchi^{a,b,1}, Nancy L. Harris^c, Sandra Brown^c, Michael Lefsky^d, Edward T. A. Mitchard^e, William Salas^f, Brian R. Zutta^{a,b}, Wolfgang Buermann^b, Simon L. Lewis^g, Stephen Hagen^f, Silvia Petrova^c, Lee White^h, Miles Silmanⁱ, and Alexandra Morel^j

^aJet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109; ^bInstitute of Environment, University of California, Los Angeles, CA 90095; ^cWinrock International, Ecosystem Services Unit, Arlington, VA 22202; ^dCollege of Natural Resources, Colorado State University, Fort Collins, CO 80523; ^eInstitute of Geography, School of Geosciences, University of Edinburgh, Edinburgh EH8 9XP, United Kingdom; ^fApplied GeoSolutions, Durham, NH 03824; ^gEarth and Biosphere Institute, University of Leeds, Leeds LS2 9JT, United Kingdom; ^hAgence Nationale des Parcs Nationaux, République Gabonaise, BP 20379 Gabon; ⁱDepartment of Biology, Wake Forest University, Winston-Salem, NC 27106; and ^jEnvironmental Change Institute, University of Oxford, Oxford OX1 3QY, United Kingdom

Edited* by Susan E. Trumbore, University of California, Irvine, CA, and approved May 5, 2011 (received for review December 28, 2010)

Developing countries are required to produce robust estimates of forest carbon stocks for successful implementation of climate change mitigation policies related to reducing emissions from deforestation and degradation (REDD). Here we present a “benchmark” map of biomass carbon stocks over 2.5 billion ha of forests on three continents, encompassing all tropical forests, for the early 2000s, which will be invaluable for REDD assessments at both project and national scales. We mapped the total carbon stock in live biomass (above- and belowground), using a combination of data from 4,079 in situ inventory plots and satellite light detection and ranging (Lidar) samples of forest structure to estimate carbon storage, plus optical and microwave imagery (1-km resolution) to extrapolate over the landscape. The total biomass carbon stock of forests in the study region is estimated to be 247 Gt C, with 193 Gt C stored aboveground and 54 Gt C stored belowground in roots. Forests in Latin America, sub-Saharan Africa, and Southeast Asia accounted for 49%, 25%, and 26% of the total stock, respectively. By analyzing the errors propagated through the estimation process, uncertainty at the pixel level (100 ha) ranged from $\pm 6\%$ to $\pm 53\%$, but was constrained at the typical project (10,000 ha) and national ($>1,000,000$ ha) scales at ca. $\pm 5\%$ and ca. $\pm 1\%$, respectively. The benchmark map illustrates regional patterns and provides methodologically comparable estimates of carbon stocks for 75 developing countries where previous assessments were either poor or incomplete.

forest biomass | forest height | microwave and optical imaging | error propagation | carbon cycling

Deforestation and forest degradation, located primarily in tropical regions, accounted for 12–20% of global anthropogenic greenhouse gas (GHG) emissions in the 1990s and early 2000s (1–4) and these processes also impact the future potential of forests to remove additional carbon from the atmosphere (5–7). Estimates of GHG emissions from deforestation require information on both the area of forest loss and the corresponding carbon stock of the land that is cleared (8, 9). Both are considered challenging to quantify accurately (10). Much of the emphasis to date has focused on improving spatially represented estimates of forest area loss (11, 12). To improve confidence in estimated emissions, equal emphasis is needed on improving spatially explicit estimates of carbon stored in forests, which remain uncertain in tropical regions (13). The largest proportion of this uncertainty is in estimates of aboveground biomass (14, 15), which accounts for 70–90% of forest biomass carbon (16), and its spatial variability that depends on factors such as climate, human and natural disturbance and recovery, soil type, and topographical variations (14, 17). Reducing the uncertainty in emissions estimates requires temporally constrained estimates of forest carbon content at a spatial scale that is fine enough to capture the variability over the landscape and is quantified at the scale of disturbance affecting the forest. Such information would improve project- and national-level carbon stock estimates as well as assist in the development of baseline information re-

quired for reducing emissions from deforestation and degradation (REDD) activities designed to curb greenhouse gas emissions from the land use sector (15, 18).

Efforts to estimate the distribution of biomass rely on remote sensing techniques due to the wide geographical extent of forests, difficult accessibility, and the limited utility of field inventories owing to the natural spatial variability of forest biomass (8, 9, 14). New remote sensing approaches using light detection and ranging (Lidar) and radio detection and ranging (radar) from airborne sensors have been successful in providing high-resolution estimates of forest carbon density for small areas (19–21). However, the spaceborne sensors needed to use these approaches for large-scale mapping and monitoring efforts will not be available before the end of this decade (22). Until then, cost-effective mapping of carbon stocks for project- and national-scale assessments will rely on a combination of satellite imagery and ground-based inventory samples of forest carbon density (14, 21).

Here, we report on our use of global forest height data measured by the Geoscience Laser Altimeter System (GLAS), onboard the Ice, Cloud, and land Elevation Satellite (ICESat) (23), in combination with other remote sensing data bases and ground data, to model the spatial distribution of aboveground standing biomass density (AGB) (in megagrams of mass per unit area) in forests across three continental regions for the early 2000s. Our approach includes >3 million Lidar shots collected along the ICESat orbital tracks. For calibration of GLAS Lidar height to AGB and for validation of AGB distribution, we use AGB data from 4,079 available inventory and research plots distributed over the study region. We estimate belowground biomass carbon in roots from AGB using tree allometry (24). Our approach results in a benchmark map of forest carbon density at 1-km resolution. The accuracy of carbon estimates for every pixel is evaluated by propagating individual components of uncertainty through the spatial analysis.

Results

Relating Forest Height to Biomass. AGB estimates were compiled for 4,079 geo-referenced in situ forest plots (>0.1 ha) distributed over three continents and restricted to inventory dates between 1995 and 2005 (Fig. S1 and Table S1). Of these data, 493 calibration plots (298 in Latin America, 75 in Africa, and 120 in Southeast Asia) were located under the GLAS Lidar shots or within the same forest stands (Table S2). Data for these plots

Author contributions: S.S.S., N.L.H., S.B., and W.S. designed research; S.S.S., N.L.H., S.B., M.L., and E.T.A.M. performed research; S.S.S., N.L.H., S.H., L.W., M.S., and A.M. contributed new reagents/analytic tools; S.S.S., N.L.H., S.B., M.L., E.T.A.M., W.S., B.R.Z., W.B., S.L.L., S.P., L.W., M.S., and A.M. analyzed data; and S.S.S., N.L.H., S.B., and S.L.L. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

¹To whom correspondence should be addressed. E-mail: saatchi@jpl.nasa.gov.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1019576108/-DCSupplemental.

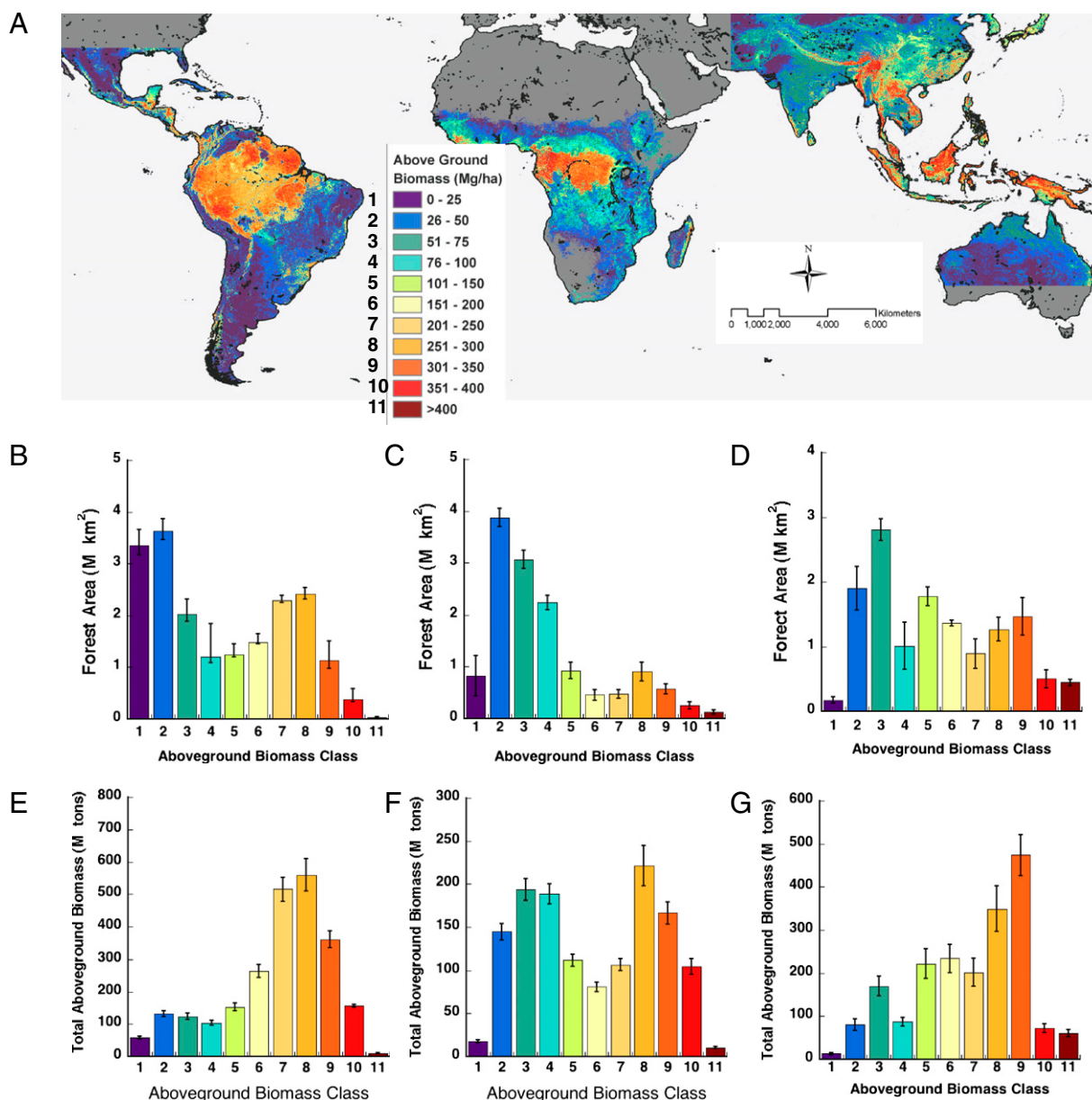


Fig. 2. Distribution of forest aboveground biomass (circa 2000). (A) Forest aboveground biomass is mapped at 1-km spatial resolution. The study region was bounded at 30° north latitude and 40° south latitude to cover forests of Latin America and sub-Saharan Africa and from 60° to 155° east and west longitude. The map was colored on the basis of 25–50 Mg ha⁻¹ AGB classes to better show the overall spatial patterns of forest biomass in tropical regions. Histogram distributions of forest area (at 10% tree cover) for each biomass class were calculated by summing the pixels over Latin America in B, Africa in C, and Asia in D. Similarly, total AGB for each class was computed by summing the values in each region with distributions provided for Latin America in E, Africa in F, and Asia in G. All error bars were computed by using the prediction errors ([SI Materials and Methods](#)) from spatial modeling.

map predicted much lower biomass compared with the new benchmark map, but with higher uncertainty. These forests were mapped with less uncertainty in the new benchmark map due to the extensive GLAS Lidar sampling of forest structure.

National Assessment of Carbon Stocks. To estimate total biomass carbon stocks and produce a forest carbon “benchmark” map against which future changes can be assessed, we calculated belowground biomass (BGB) as a function of AGB (BGB = $0.489 \text{ AGB}^{0.89}$) (Fig. S2) and estimated total carbon as 50% of total biomass (AGB + BGB) (27) (Fig. 3A). We defined forest extent using three fractional cover thresholds (10%, 25%, and 30% tree cover) based on the range of thresholds used by individual Parties to the United Nations Framework Convention on Climate Change (UNFCCC) proceedings (28) and estimated

total carbon stored in forests of each country in the study area (Table S3). In Table 1, we present the five countries per continental region that contain the highest forest carbon stocks, along with continental totals. From this analysis, we estimate the total forest biomass carbon stocks at 10% tree cover as 247 Gt C, with 193 Gt C in AGB and 54 Gt C in BGB. Forests in Latin America are the most extensive and contain ~49% of the total biomass carbon, followed by 26% in Asia and 25% in Africa. Applying a higher tree cover threshold (30%) eliminates large areas of savanna woodlands in Africa from the forest domain and reduces the total carbon stock to 208 Gt C (16% reduction with 163 Gt C for AGB and 45 Gt C for BGB). Among the countries analyzed, Brazil, the Democratic Republic of Congo, and Indonesia have the largest area of forest as well as the highest carbon stocks (62, 24, and 24 Gt C, respectively, at 10% tree cover). These esti-

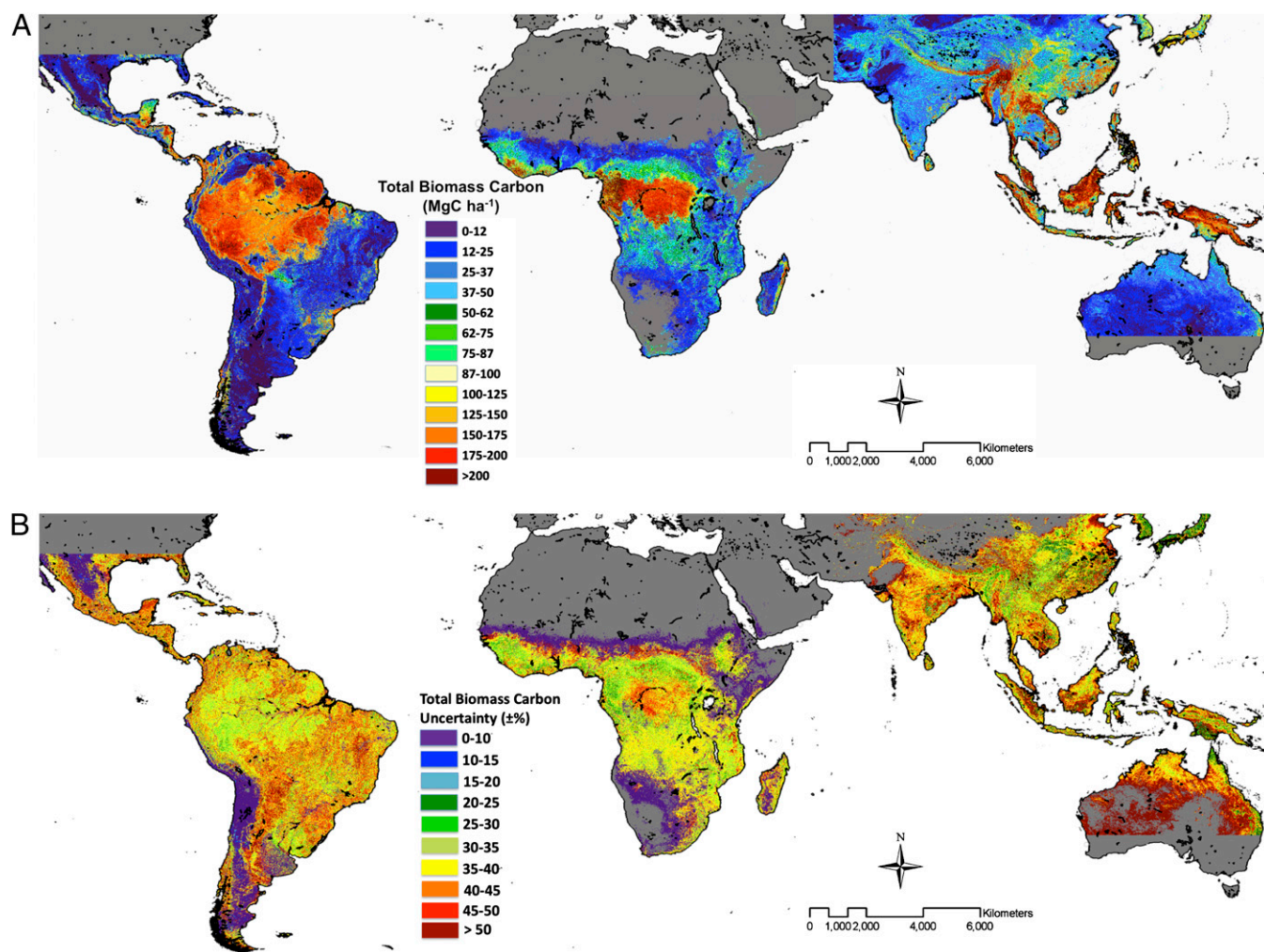


Fig. 3. Benchmark map of carbon stock and uncertainty. (A) Forest carbon stock defined as 50% of AGB + BGB is mapped at 1-km pixel resolution and colored on the basis of a 12–25 Mg C ha⁻¹ range to show the spatial patterns. (B) The uncertainty of the benchmark map is estimated using error propagation through a spatial modeling approach. The uncertainty is given in terms of plus or minus percent and it includes all errors associated with prediction from spatial modeling, estimation of Lorey's height from GLAS, estimation of AGB from Lorey's height, errors from pixel level variations, and errors associated with BGB estimation (*SI Materials and Methods*).

mates improve on forest carbon stock estimates reported previously (8, 13–15, 17, 18, 25) by providing a traceable and systematic approach to geographically locate the stock estimates for further monitoring and verification. The forest definitions chosen here using tree cover thresholds can readily change the estimates of total carbon and area-weighted carbon densities at national and regional scales.

Uncertainty Analysis. We assess the accuracy of the biomass carbon estimates by calculating the error as the difference between the true mean biomass value (bootstrapped samples of ground and Lidar-estimated AGB) and the predicted biomass value (mapped at 1-km grid cell resolution) and propagating these errors through the spatial modeling process (*SI Materials and Methods*). Errors in the distribution of forest aboveground biomass can be random or systematic in nature and can include the following: (i) *observation errors* associated with the uncertainty in estimates of Lorey's height from GLAS Lidar, errors associated with estimating AGB derived from GLAS Lidar height, and errors in estimating BGB from AGB (27); (ii) *sampling errors* associated with the spatial variability of AGB within a 1-km pixel and the representativeness and size of inventory plots and GLAS pixels over the landscape (29); and (iii) *prediction errors* associated with spatial analysis and mapping of AGB from significant

contributions from satellite imagery (Fig. S3) (14, 30). We combined these three types of errors (*SI Materials and Methods*) to quantify the uncertainty of total biomass carbon stock as the 95% bootstrapped confidence interval at the 1-km pixel level (Fig. 3B). The overall uncertainty in mapping AGB at the pixel scale averaged over all continental regions is estimated at $\pm 30\%$, but it is not uniform across regions or AGB ranges ($\pm 6\%$ to $\pm 53\%$) and depends on regional variations of forests, quality of remote sensing imagery, and sampling size and distribution of available ground and GLAS data. However, when averaged over all AGB ranges, regional uncertainties were comparable: $\pm 27\%$ over Latin America, $\pm 32\%$ over Africa, and $\pm 33\%$ over Asia (Fig. S4). The uncertainty in total carbon stock at the pixel scale averaged $\pm 38\%$ over all three continents after errors associated with BGB estimation were included in the analysis.

We computed the uncertainty around carbon estimates at national and regional scales by propagating errors associated with observation, including the errors associated with BGB estimates, sampling, and prediction. The uncertainty of carbon stock estimates at the national level was calculated as the square root of the sum of per-pixel errors for all pixels within the national boundary. This process reduced the relative errors as sample area increased. The national estimates were found to be constrained to within $\pm 1\%$ of the total carbon stock obtained

separated these from the other uncertainties considered in this study. By assuming that the systematic bias in estimating the carbon stock is approximately confined by the range of uncertainty of carbon predicted at the pixel level, we used the uncertainty bounds for each pixel to arrive at a range of estimates for total carbon (Fig. 4 and Table S3). However, these nonrandom errors associated with ground-based estimation of forest biomass will remain uncertain until consistent allometric equations within forest types or regions are developed.

All countries participating in a future policy mechanism to reduce emissions from deforestation and forest degradation will need to develop national- to regional-scale estimates of historic emissions (~2000–2010), which will be the starting point for generating reference emission levels. Most developing countries currently have limited data on carbon stocks of forests with which to estimate historic carbon emissions from past deforestation and degradation. Instead, countries often rely on estimates based on old or incomplete national forest inventories as reported by the Food and Agriculture Organization (31) or on Tier 1 biome-level estimates reported by the Intergovernmental Panel on Climate Change (32), neither of which are spatial in nature and thus do not allow for matching the carbon stock data with the areas undergoing change. The benchmark map presented here can assist country efforts by providing relatively fine-scale, spatially explicit and consistent forest carbon estimates that can be used with readily available remote sensing imagery to produce more robust estimates of historic emissions.

The benchmark map can also be used to assist countries in assessing the carbon emissions that are likely to be avoided by implementing different policies and programs aimed at reducing deforestation and forest degradation at regional and project scales. The map will assist developing country governments, land managers, policy makers, and civil society to become more informed about the likely result of their policies and programs in reducing national greenhouse gas emissions from the land-use sector.

Materials and Methods

Our methodological approach to mapping forest carbon stocks consists of four steps: (i) processing of ground and GLAS Lidar observations to sample forest structure and biomass over tropical regions, (ii) developing relations between

Lidar-derived Lorey's height and AGB and between AGB and BGB, (iii) mapping AGB at 1-km spatial resolution using satellite imagery to stratify tropical forest types and structure and the Maximum Entropy (MaxEnt) approach to spatially model AGB, and (iv) assessing the uncertainty in modeling the spatial distribution of AGB by validating the results and propagating the errors through the methodology to estimate the total carbon stock and its uncertainty at the national scale.

Ground data used to train the biomass prediction model were derived from various sources including published literature and national forest inventories collected by the authors and their colleagues. The plots covered a variety of forest types on each continent, including old growth moist and wet tropical forest, woodland savanna, dry forest, peat swamp forest, and forests recovering from past disturbance or clearing. To compensate for the lack of systematic spatial sampling of aboveground biomass from ground measurements, we included >3 million AGB values calculated from Lidar measurements of forest vertical structure. We used 493 calibration plots distributed over forests across three continents to convert Lorey's height inferred from Lidar measurements to AGB and used tree allometry to estimate BGB from AGB. We estimated the total biomass per plot as the sum of AGB and BGB estimates and converted the results to carbon content by using a conversion factor of 0.5. To scale our plot- and Lidar-based AGB results over the landscape, we used nonparametric spatial modeling using the MaxEnt model, which included three steps: (i) compilation of the spatially gridded remote sensing data, (ii) implementation of MaxEnt and the production of the AGB map, and (iii) estimation of prediction uncertainty. We estimated the overall uncertainty in the final benchmark map by combining the errors associated with four independent terms: measurement errors, allometry errors, sampling errors, and prediction errors. Detailed information on each step is provided in *SI Materials and Methods*.

ACKNOWLEDGMENTS. We thank J. Reitsma, H. Woell, J. Woods, A. Daniels, S. Chinekan, and the Forest Developments Authority (Liberia) for assistance in processing inventory plots. This work was carried out under Contract 7150484 between Winrock International and The World Bank Group. The original remote sensing data were distributed online and were compiled and processed at the Jet Propulsion Laboratory, California Institute of Technology from previous National Aeronautics and Space Administration grants. The integrated remote sensing and distribution model was developed at the University of California, Los Angeles, Institute of Environment from previous National Aeronautics and Space Administration and National Science Foundation grants. The Blue Moon Fund supported forest inventory plots in the Peruvian Amazon for Lidar calibration.

- Denman KL, et al. (2007) *IPCC Climate Change 2007: The Physical Science Basis*, eds Solomon S, et al. (Cambridge Univ Press), pp 499–587.
- Schrope M (2009) When money grows on trees. *Nat Rep Climate Change* 3:101–103.
- Van der Werf GR, et al. (2009) CO₂ emissions from forest loss. *Nat Geosci* 2:737–738.
- Dixon RK, et al. (1994) Carbon pools and flux of global forest ecosystems. *Science* 263: 185–190.
- Chave J, et al. (2008) Assessing evidence for a pervasive alteration in tropical tree communities. *PLoS Biol* 6:e45.
- Lewis SL, et al. (2009) Increasing carbon storage in intact African tropical forests. *Nature* 457:1003–1006.
- Field CB, Behrenfeld MJ, Randerson JT, Falkowski P (1998) Primary production of the biosphere: Integrating terrestrial and oceanic components. *Science* 281:237–240.
- Gibbs HK, Brown S, Niles JO, Foley JA (2007) Monitoring and estimating tropical forest carbon stocks: Making REDD a reality. *Environ Res Lett* 2:045023.
- Houghton RA (2005) Aboveground forest biomass and the global carbon balance. *Glob Change Biol* 11:945–958.
- Ramankutty N, et al. (2007) Challenges to estimating carbon emissions from tropical deforestation. *Glob Change Biol* 13:51–66.
- Hansen MC, et al. (2008) Humid tropical forest clearing from 2000 to 2005 quantified by using multitemporal and multiresolution remotely sensed data. *Proc Natl Acad Sci USA* 105:9439–9444.
- Achard F, et al. (2002) Determination of deforestation rates of the world's humid tropical forests. *Science* 297:999–1002.
- Houghton RA, Lawrence KT, Hackler JL, Brown S (2001) The spatial distribution of forest biomass in the Brazilian Amazon: A comparison of estimates. *Glob Change Biol* 7:731–746.
- Saatchi SS, et al. (2007) Distribution of aboveground live biomass in the Amazon basin. *Glob Change Biol* 13:816–837.
- Brown S, Gaston G (1996) *Tropical Africa: Land Use, Biomass, and Carbon Estimates for 1980* (CDIAC NDP-055, ORNL/CDIAC-92), CDIAC Communications, No. 23, Spring 1997 (Oak Ridge National Laboratory, Oak Ridge, TN), pp 12.
- Cairns MA, Brown S, Helmer EH, Baumgardner GA (1997) Root biomass allocation in the world's upland forests. *Oecologia* 111:1–11.
- Gibbs HK, Brown S (2000) Geographical distribution of woody biomass carbon stocks in tropical Africa: An updated database for 2000. Available at <http://cdiac.ornl.gov/epubs/ndp/ndp055/ndp05b.html> from the Carbon Dioxide Information Center, Oak Ridge National Laboratory, Oak Ridge, TN. Accessed 2009.
- Loarie SR, Asner GP, Field CB (2009) Boosted carbon emissions from Amazon deforestation. *Geophys Res Lett* 36:L14810.
- Saatchi S, et al. (2011) Impact of spatial variability of tropical forest structure on Radar estimation of aboveground biomass. *Remote Sens Environ*, 10.1016/j.rse.2010.02.022.
- Drake JB, et al. (2002) Estimation of tropical forest structural characteristics using large-footprint Lidar. *Remote Sens Environ* 79:305–319.
- Asner GP, et al. (2010) High resolution forest carbon stock and emissions in the Amazon. *Proc Natl Acad Sci USA* 107:16738–16742.
- Shugart HH, Saatchi S, Hall F (2010) A primer on the structure of forest and its measurements from space. *J Geophys Res Biogeosci* 115:G00E13.
- Lefsky MA (2010) A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. *Geophys Res Lett* 37:L15401.
- Laurance WF (2007) Forest destruction in tropical Asia. *Curr Sci* 93:1544–1550.
- Malhi Y, et al. (2006) The regional variation of aboveground live biomass in old-growth Amazonian forests. *Glob Change Biol* 12:1107–1138.
- Slik JWF, et al. (2010) Environmental correlates of tree biomass, basal area, wood specific gravity, and stem density gradients in Borneo tropical forests. *Glob Ecol Biogeogr* 19:50–60.
- Mokany K, Raison RJ, Prokushkin AS (2006) Critical analysis of root:shoot ratios in terrestrial biomes. *Glob Change Biol* 12:84–96.
- Annex to UNFCCC decision 16/CMP.1. Available at <http://unfccc.int/resource/docs/2005/cmp1/eng/08a03.pdf>. Accessed 2010.
- Chave J, et al. (2004) Error propagation and scaling for tropical forest biomass estimates. *Philos Trans R Soc Lond B Biol Sci* 359:409–420.
- Köhl M, Baldauf T, Plugge D, Plugge J (2009) Reduced emissions from deforestation and forest degradation (REDD): A climate change mitigation strategy on a critical track. *Carbon Balance Manag* 4:1–10.
- Food and Agriculture Organization (2010) *Global Forest Resources Assessment 2010*, FAO Forestry Paper 163 (Food and Agriculture Organization, Rome).
- Paustian K, Ravindranath NH, Van Amstel A (2006) (coordinating lead authors); Intergovernmental Panel on Climate Change (2006) *IPCC Guidelines for National Greenhouse Gas Inventories*, Agriculture, Forestry and Other Land Use, Vol 4, Prepared by the National Greenhouse Gas Inventories Programme, [Institute for Global Environmental Strategies (IGES), Tokyo, Japan]. Available at <http://www.ipcc-nggip.iges.or.jp/public/2006gl/vol1.htm>. Accessed 2010.

Supporting Information

Saatchi et al. 10.1073/pnas.1019576108

SI Materials and Methods

Overview. Our approach consists of four steps: (i) processing of ground and Geoscience Laser Altimeter System (GLAS) Lidar observations to sample forest structure and biomass over tropical regions, (ii) developing relations between Lidar-derived Lorey's height and AGB and between AGB and BGB, (iii) mapping forest biomass carbon (AGB + BGB) at 1-km spatial resolution using satellite imagery to stratify tropical forest types and structure and modeling the spatial distribution using the maximum entropy (MaxEnt) approach, and (iv) assessing the uncertainty in modeling the spatial distribution by validating the results and propagating the errors through the methodology to estimate the total carbon stock and its uncertainty at project and national scales.

(i) Processing of in situ and GLAS data. In situ forest inventory data. We assembled 4,079 plots that spanned a variety of forest types on each continent, including old growth moist and wet tropical forest, woodland savanna, dry forest, peat swamp forest, and forests recovering from past disturbance or clearing. These ground data used to train the biomass prediction model were derived from various sources including published literature and national forest inventories collected by the authors and their colleagues (Table S1). Although these measurements do not follow a uniform systematic inventory protocol (they vary in plot size, sampling scheme, allometric equations, and number of structural components), they have produced the largest known dataset on forest aboveground biomass density throughout the tropics. The plots included in this study met the following criteria: (i) All biomass measurements were made after 1995 and before 2005 so that the resulting biomass map would be representative of forest biomass circa the year 2000. (ii) Plots had a minimum size of 0.1 ha and the biomass densities represented all trees >10 cm in diameter. For pixels that included multiple plots, we used the average biomass value of all plots weighted by the plot size. We used a minimum area of 1 ha from one plot or a combination of smaller plots to sample the biomass density of 1-km pixels and eliminated all pixels with less plot area from the analysis. A total of 1,877 1-km pixels from inventory plots were used in the analysis (Fig. S1). The sampling errors were included in the uncertainty analysis (described below) of the forest biomass map.

ICESat GLAS Lidar data. To compensate for the lack of systematic spatial sampling of aboveground biomass from ground measurements, we included a method for estimating biomass from satellite Lidar measurement of forest vertical structure. Data from the GLAS, onboard the Ice, Cloud, and land Elevation Satellite (ICESat), acquired in 2003 and 2004 over tropical forests were used in this study. GLAS is a waveform sampling Lidar sensor; it emits short duration (5 ns) laser pulses toward the land surface and records the echo of those pulses as they reflect off the ground surface (1). When the surface is vegetated, the return echoes, or waveforms, are a function of the vertical distribution of vegetation and ground surfaces within the area illuminated by the laser (the footprint). For forests on level ground, stand height can be calculated as the difference between the elevation of the first returned energy minus the mean elevation of the ground return (waveform extent) (2). The vertical extent of each waveform increases as a function of terrain slope and footprint size (the area on the ground that is illuminated by the laser), as modified by the spatial pattern of ground surfaces visible to the laser. Over sloped terrain, information on the vertical extent of the waveform is insufficient to make estimates of tree height. Therefore, algorithms capable of retrieving information about

terrain slope, stand uniformity, and the vertical distribution of visible ground surfaces from the waveform itself were used for terrain slope corrections (3, 4).

To estimate forest height from Lidar waveform indexes, a model developed from several study areas in broadleaf stands in temperate and tropical forests was used (5). The model was calibrated with three study areas with ground estimates of height in the Amazon basin, located in the municipalities of Santarem, Para State; Manaus, Amazonas State; and Canarana, Mato Grosso State, all in Brazil. These sites represented areas with a range of dry season duration (months with <100 mm of rain). Dry season duration was selected as an environmental variable that was likely to summarize structural trends across the Amazon (6).

The index of height derived from field measurements and used in the processing of the global GLAS dataset is Lorey's height, the basal area weighted height of all trees >10 cm in diameter,

$$H_{\text{Lorey}} = \frac{\sum_{i=1}^N BA_i h_i}{\sum_{i=1}^N BA_i}, \quad [\text{S1}]$$

where BA_i and h_i are the basal and canopy heights of individual trees in a plot. Basal area weighting of tree heights increases the importance of the largest trees in a stand and represents the height of stands with tallest trees. Indexes of total waveform extent and the height of the 10th and 90th percentiles of waveform height were used with least-squares regression to estimate Lorey's height (5). An equation to estimate Lorey's height for broadleaf stands explained 83% of variance with an RMSE of 3.3 m or 17.3% of relative error (5).

We processed ~3 million Lidar Lorey's heights derived from GLAS shots distributed over tropical forests (523,985 in Africa, 1,583,557 in Latin America, and 974,392 in Asia). All GLAS heights derived for terrains with slopes >20% were deleted from the dataset because of potential errors in Lorey's height estimates (4). The terrain slope was calculated from shuttle radar topography mission (SRTM) data at 90 m spatial resolution coincident with Lidar footprints.

Estimation of AGB from Lidar data. Lorey's heights derived from GLAS data were converted to AGB using 493 calibration plots. The calibration plots were located under the GLAS footprints or within the same forest stand and represented the forest structure sampled by GLAS Lidar. Each plot had an area of at least 0.25 ha that was equal to or larger than the effective footprint area of GLAS (<0.5 ha) Lidar samples (5) and included tree height, basal area, and genera-specific wood density for all trees >10 cm in diameter at breast height (dbh). The plots covered a wide range of forest types. In Latin America, 298 plots were distributed in terra firme forests of the Central Amazon in Manaus (Amazonia), Santarem (Para), and Camarana (Mato Grosso); seasonal transitional forests of Rondonia (Ji Parana); and forests of the western Amazonia in Peru including terra firme, bamboo, and seasonally flooded forests (Manu, Tappopata). In Africa, 75 plots covered the forest-savanna transitions of Mbam-Djerrem (Cameroon); old growth forests in Budongo (Uganda), Cavalla and Grebo (Liberia), Ekobakoba, Oveng, and Dousalla (Gabon); and old growth and successional forests in Lope National Park (Gabon). In Asia, 120 plots covered old growth, secondary, and logged forests in Sabah (Malaysia) (Table S2).

We estimated the aboveground biomass for all individual trees in calibration plots using pan-tropical generalized allometric equations developed by Chave et al. (7). For woodland savanna

plots, the dry forest equation, $AGB = 0.112(\rho D^2 h)^{0.916}$, and for all other plots, the moist forest equation, $AGB = 0.0509(\rho D^2 h)$, were used to estimate AGB (ρ is the wood density, D is dbh, and h is the tree height). In these equations, tree heights were accounted for by direct measurements with clinometers and laser range finders for almost all trees ($DBH > 10$ cm), and diameter-based estimation was used for a small fraction where no direct height measurements were available (8).

The tree height measurements from the calibration plots were used to develop a relationship between Lorey's height and the estimated AGB at the plot level. The new allometric equations were derived separately for each continent and for the combined dataset (Fig. 1). After evaluating the equations by cross-validating with data from each continent, we concluded that continent-based equations provided the best model to convert Lorey's height to forest biomass. The equations were then used to convert GLAS Lidar height data to AGB for all available GLAS footprints over tropical forests and woodlands.

Differences in equations can be attributed to several factors: (i) The number of plots used in each continent and the sampling strategy to capture the forest types and structure are different. In Asia, the plots were equally distributed in old growth, secondary forests, logged forests, and plantations. In South America plots were mainly from old growth forests, with a small portion in secondary and swamp forests. Plots in Africa were almost equally divided between dense old growth forests and woodland savanna. (ii) Wood density is used in estimating the biomass of trees for each plot. Differences in wood density of large trees in each continent can bias the equation converting Lorey's height to biomass. Potentially, by including wood density, differences in continental equations could be reduced.

Next, we used the average of at least five AGB values of Lidar footprints in 1-km pixels to create 160,918 pixels (80,579 pixels in Africa, 37,931 pixels in America, and 42,408 pixels in Asia). The 1-km pixel data were subsequently divided into training data (93,188 pixels) and test data (67,730 pixels) for AGB modeling, validation, and uncertainty analysis. The geographical distribution of 1-km pixels with average Lorey's height along with available ground data is shown in Fig. S1.

BGB estimation from AGB. We encountered virtually no consistent measurements of belowground biomass in our data compilation efforts. This result was not surprising, as measurement methods for collecting belowground biomass data are laborious, time-consuming, and technically challenging to perform correctly. Instead, belowground biomass is usually estimated from aboveground biomass using regression equations developed from field data collected across multiple biomes (9). A synthesis of data from available literature, along with elimination of data collected using unclear or incorrect methods, provided a universal equation for estimating forest belowground biomass. We used this equation to estimate belowground biomass from aboveground biomass estimated for each 1-km forested pixel in our analysis,

$$BGB = 0.489AGB^{0.89}, \quad [S2]$$

where BGB is the belowground and AGB is the aboveground biomass in units of megagrams per hectare of dry weight. To develop an uncertainty in the above relationship, we acquired the data from Mokany et al. (9) and examined the variations in the ratio of below:aboveground biomass or root:shoot biomass ratios with respect to vegetation types used in the study. By excluding sites in forest plantations and grasslands and tundra, the RMSE in predicting the belowground biomass was 9.46 Mg ha^{-1} with relative error of $\sim 23.2\%$ (Fig. S2). Once belowground biomass was calculated from aboveground biomass, we calculated the total as the sum of above- and belowground estimates and converted the results to live tree carbon content by using a conversion factor of 0.5. In our analysis, we have ignored all

other components of the Intergovernmental Panel on Climate Change carbon pools such as dead wood and litter due to lack of data and models to estimate from aboveground biomass.

(ii) Spatial modeling of AGB. There are multiple ways of extrapolating the samples of forest biomass data to a gridded map. These include parametric approaches such as the use of regression models with spatial environmental data and nonparametric approaches such as interpolation, cokriging, classification or coloring by numbers, decision rule techniques as in random forest, MaxEnt, and several machine learning approaches (10–12). Parametric models are not suitable for extrapolating the biomass data because there is no reasonable relationship that exists between current satellite observations (passive optical and microwave) and tropical forest biomass. In addition, parametric models often violate conditions of independence and multivariate normality when complex ecological systems and environmental variables derived from different remote sensing measurements are involved. Nonparametric models are found to be more suitable in geospatial and geostatistic analysis because they are not affected by these violations, can integrate variables with different statistical distributions, and provide more stable and relevant information. Furthermore, forest structure and biomass often exhibit complex, nonlinear variations, autocorrelation, and variable interaction across temporal and spatial scales. In these cases, nonparametric approaches often greatly outperform the parametric methods (11).

Among nonparametric models, we selected MaxEnt after comparing its performance against two other methods (random forest and maximum-likelihood classification), using a set of training and test datasets. In all cases, MaxEnt outperformed the other methods in modeling the spatial distribution of biomass and in providing significantly better accuracy compared with an independent dataset.

MaxEnt is a general-purpose algorithm that generates predictions or inferences from an incomplete set of information. The MaxEnt approach is based on a probabilistic framework. It relies on the assumption that an incomplete empirical probability distribution, which is based on individual occurrences of a variable's point locality in geographical space (here, biomass sample points from ground and GLAS Lidar), can be approximated with a probability distribution that has maximum entropy (the MaxEnt distribution) subject to certain environmental constraints and that this distribution approximates potential geographic distribution (12, 13). For our purposes, we assume the unknown probability distribution P is defined over a finite set X (interpreted as the set of pixels within the study area), with the probability value of $P(x)$ for each point x . These probabilities sum to 1 over the space defined by X . The MaxEnt algorithm uses a Bayesian approach to approximate P with a probability distribution \hat{P} by maximizing the entropy of \hat{P} as

$$\text{Entropy} = \sum_{x \in X} \hat{P}(x) \ln [\hat{P}(x)], \quad [S3]$$

where \ln is the natural logarithm. Entropy is a nonnegative number with the maximum value equal to the natural log of the number of pixels in X and is a measure of constraints or choices of a probability distribution. The distribution \hat{P} with highest entropy (i.e., the least additional information is introduced through model assumptions), while still subject to the constraints of incomplete information, is considered the best distribution for inference. By using features that are continuous real-valued functions of X , such as remote sensing variables, and a set of sample points (training data) of X provided by the field data and GLAS Lidar points, the MaxEnt algorithm employs likelihood estimation procedures to find a probability distribution for all of the points in X that have similar statistics to the sample points. MaxEnt assumes a priori a uniform distribution and performs

a number of iterations in which the weights (of different features) are adjusted to maximize the average probability of the point localities (also known as the average sample likelihood), expressed as the training gain (13). These weights are then used to compute the MaxEnt distribution over the entire geographic space. In the context of the present study, MaxEnt can be applied to geographic locations of forest plots and remote sensing data to produce distributions expressing the suitability (probability) of each pixel as a function of the environmental variables of that pixel. A high value of the probability function for a particular pixel indicates that the pixel is predicted to be suitable for having similar characteristics to the training pixels (12). MaxEnt has a number of properties that make it useful for modeling forest biomass over landscapes (12, 13). These properties include a deterministic framework (and hence stability), high performance with relatively few sample points, good computing efficiency (which enables the use of large-scale, high-resolution data layers), continuous output (i.e., from least to most suitable conditions), and the ability to model complex responses to environmental or remote sensing variables. The MaxEnt model has a built-in jackknife option, which allows estimation of the significance of individual data layers in the computation of the final distributions.

The spatial modeling with MaxEnt included three steps: (i) compiling the spatially gridded remote sensing data, (ii) implementation of MaxEnt and the production of the AGB map, and (iii) estimation of prediction uncertainty.

Compiling remote sensing data. We compiled a set of remote sensing data and products from different earth observing sensors to derive metrics sensitive to vegetation cover, density, seasonality, moisture, roughness, and surface topography. The dataset included both optical and microwave satellite sensors. All optical data used in this study are from the moderate resolution imaging spectroradiometer (MODIS) aboard the Terra satellite. We included normalized difference vegetation index (NDVI) and leaf area index (LAI) products. MODIS products were downloaded from the Earth Resources Observation and Science Data Center (<http://e4ftl01.cr.usgs.gov/MOTA/>) from the latest iteration of MODIS product development (collection 5, MOD13A35) and processed for cloud cover and pixel quality.

We used LAI to develop landscape-scale data layers related to vegetation canopy structure and seasonality as an important feature to stratify the forest and the NDVI data as a measure of vegetation greenness that generally correlates well with ground measurements of gross photosynthesis (14).

Both NDVI and LAI data were processed to develop cloud-free monthly images for the years 2000–2001. For pixels that did not have any values because of MODIS quality flags and cloud contamination, we used the best data from the same months over the next few years (2001–2003) to fill gaps. The monthly data were then used to generate three LAI metrics, annual maximum, mean, and range (difference of maximum and minimum); and five NDVI metrics, maximum NDVI, mean NDVI, green NDVI (average of greenest or highest NDVI quarter), brown NDVI (average of brownest or lowest NDVI quarter), and NDVI green–brown (difference between highest and lowest NDVI quarters) (15).

As part of the microwave remote sensing measurements, we included global quick scatterometer (QSCAT) data available in 3-d composites at 2.25-km resolution (16). The 3-d time series data in the year 2000 were used to create average monthly composites at a 1-km resolution and were then further processed to produce four metrics that included annual mean and SD of radar backscatter at both HH and VV polarizations (H, horizontal send/receive; V, vertical send/receive). QSCAT radar measurements are at KU band (12 GHz) and are sensitive to surface or canopy roughness, moisture, and other seasonal attributes, such as phenological changes, although the relationship between QSCAT and specific forest variables is yet to be explored. For areas with

low vegetation biomass, such as woodlands and savanna, measurements at different polarizations correlate positively with the aboveground biomass. For areas with dense forest, backscatter measurements are sensitive to canopy roughness and moisture and contribute to measuring differences in forest types and canopy structure. In this study, we used the long-term mean and SD of QSCAT HH backscatter data and excluded the VV backscatter data because of its high correlation with the HH backscatter over tropical forests (16). In addition two seasonality metrics of average HH backscatter from dry and wet quarters were added to the list of input variables.

We added the SRTM digital elevation data, aggregated from ~90-m resolution to 1 km, in the pool of spatial data layers. In addition to the mean elevation, the SD of surface elevation was calculated during pixel aggregation from 90-m to 1-km resolution and was also included as a metric to represent landforms or geological features with different ruggedness or topographical variability (17, 18). We used the SRTM layers along with the rest of the remote sensing data to model the distribution of forest ecological variables. We also included the MODIS-derived vegetation continuous field (VCF) product as a measure of the percentage of tree canopy cover within each 1-km pixel resolution (19). The VCF product was not used in spatial modeling of AGB, but used later to estimate total carbon stock for forests above 10%, 25%, and 30% tree cover.

MaxEnt implementation. We ran the MaxEnt model using the locations of inventory plots and GLAS Lidar points over forests of all three tropical regions, using 14 remote sensing image layers (5 NDVI, 3 LAI, 4 QSCAT, and 2 SRTM metrics). The total number of biomass pixels from inventory plots was 1,877 and that from GLAS Lidar points was 160,918. Although, the number of inventory plots was much smaller than that of GLAS Lidar points, it provided biomass samples with less uncertainty and in regions where there were no GLAS Lidar tracks. From the combined dataset, we randomly selected 93,188 points to train the model and used the rest as an independent dataset to examine the accuracy of the resulting map. Aboveground biomass density from the training data were divided into 11 biomass classes in 25 Mg ha⁻¹ intervals for biomass ranging from 0 to 100 Mg ha⁻¹ and 50 Mg ha⁻¹ intervals for biomass >100 Mg ha⁻¹. Biomass classes were used partially to capture the errors associated with the GLAS-derived biomass values and partially because MaxEnt does not run with continuous data (13). However, biomass ranges for each class can be easily modified as long as there are enough samples (>100 locations) within each class bin. We ran MaxEnt for different biomass ranges to examine the impact on the final AGB map and we found the selected biomass ranges suitable for creating the optimum map. For each interval (biomass class), we performed a MaxEnt model run and generated 11 continuous probability distribution maps ranging from 0 to 100, with 0 as the least suitable pixel for the biomass class and 100 as the most suitable. The continuous probability maps were then combined and converted into a single biomass map by choosing the biomass value associated with the maximum probability weighted mean for each pixel, using the relationship

$$\hat{B} = \frac{\sum_{i=1}^N P_i^n B_i}{\sum_{i=1}^N P_i^n}, \quad [S4]$$

where P_i is the MaxEnt probability estimated for each biomass range class B_i (median value of the range), and \hat{B} is the predicted value of AGB for each pixel. The power of the probability n is chosen to weight the predicted value toward the maximum probability that is close to the true value when other class probabilities are small. After several iterations and cross-validation with the test data, we found that for $n = 3$ the distribution of AGB and the cross-validation converged to its optimum value.

This process also preserved the skewness in distributions for each pixel, which by definition is the third standard moment ($n = 3$) of the random variable (biomass) defined by the probability distributions at each pixel. Values $>n = 3$ did not change the estimated biomass at the pixel scale ($<0.001\%$ of pixels changed values in each continent). The estimator in Eq. S4 performed superior to other options such as the use of maximum probability class or the average ($n = 1$). In both cases, errors associated with the estimated biomass were large. In the case of average value ($n = 1$), areas of low and high biomass values were forced to the mean biomass of the distribution over the region. On the other hand, the choice of maximum probability class ignored the probability of other classes for the same pixel. We could not use any other simple and efficient estimator such as the maximum likelihood because at any pixel the estimated probabilities of the biomass did not have a normal distribution. For example, for a pixel with a biomass class range of 200–250 Mg ha^{-1} , MaxEnt predicts high probability for the correct biomass range, suggesting a skewed and nonnormal distribution of probabilities. In general, AGB of a cluster of pixels at the landscape scale may have normal distribution. However, the predicting probability of biomass at each pixel from the MaxEnt model will not be normal.

To include upper and lower bounds to the biomass estimation, we also created biomass maps using the maximum and minimum values of each biomass class range instead of the median value. These maps are used to estimate the upper and lower bounds for national- and regional-level carbon stocks assessments.

Estimation of prediction uncertainty. The use of nonparametric models to develop AGB distributions is often subject to prediction errors. Unlike parametric models such as regression techniques where uncertainty can be readily quantified from the sensitivity of the measurement variables to predicting parameters, in nonparametric models, the prediction errors are potentially large and have complex spatial distributions.

We examined the MaxEnt model performance using two indicators: (i) the fraction of predicted area and extrinsic omission rate at a selected threshold and (ii) the area under the receiver operator curve (AUC) as a measure of model performance across all thresholds (12, 13). For all model runs, AUC values ranged between 0.86 and 0.98, suggesting that the predictions were significantly better than random ($\text{AUC} = 0.5$), with high statistical significance (one-tailed $P < 0.001$). This result was obtained by bootstrapping 25% of the training data (93,188 points) $\sim 1,000$ times, suggesting a robust performance of the MaxEnt algorithm to capture the variations in environmental variables to predict the probability distributions of biomass range classes.

To examine the presence of other potential errors such as the effect of the number of training pixels on MaxEnt performance, we developed several simulation runs by varying the number of training pixels. MaxEnt performance was optimal when a minimum of 100 training pixels were used for each biomass range run. This requirement was readily met when hundreds to several thousands of pixels were used in each biomass range.

Significance of remote sensing variables. The contribution of 14 remote sensing variables in modeling the distribution of AGB was evaluated by the jackknife test on all input variables and they were grouped and reported separately by geographic region in four AGB ranges (Fig. S3). We selected these biomass ranges on the basis of the sensitivity of the remote sensing data to detecting the spatial variations of AGB (13). The jackknife test was performed for all model runs and the significance of layers was reported in terms of percentage of explaining the total variability of each AGB range. We grouped 14 remote sensing variables in four types of observations or features that related to vegetation greenness and seasonality (5 NDVI metrics), leaf area index (3 LAI metrics), canopy roughness and water content and its seasonality (4 QSCAT metrics), and surface elevation and slope variations (2 SRTM metrics).

SRTM data contributed less than other variables in modeling AGB, with the highest contribution in Asia due to extensive topographic features in southeast islands of Indonesia and Papua New Guinea and in highlands of Nepal, Bhutan, and eastern China. Among other variables, the NDVI metrics had the largest contribution in low biomass density forests ($<100 \text{ Mg ha}^{-1}$), the LAI metrics were important in midrange biomass density (100–200 and 200–300 Mg ha^{-1}), and the QSCAT metrics were important in all biomass ranges $>100 \text{ Mg ha}^{-1}$ with the largest contribution in modeling the high-density biomass forests. These contributions cannot be explained in terms of physical relationships between the remote sensing variables and the aboveground biomass density, yet they represent the relative significance of the remote sensing variables and their spatial and temporal characteristics to stratify the tropical landscapes in features that allow extrapolating the forest biomass density.

(iii) Uncertainty analysis. We calculated the overall uncertainty in the final benchmark map by quantifying the errors associated with the distribution of AGB and the estimation of BGB from AGB at the pixel and national or regional scales. We assumed no errors associated with converting the total biomass to carbon using a 50% conversion factor or with the original allometric equations that converted the ground forest structure to biomass. The total uncertainty of AGB at the pixel level is divided into four independent terms and all terms are reported in terms of percentage of relative uncertainty:

Measurement error ($\epsilon_{\text{measurement}}$) is associated with the estimation of Lorey's height from GLAS Lidar data. For broadleaf forests, the RMSE has been estimated to be 3.3 m with the relative error of $\sim 13.7\%$ over the entire height range (20).

Allometric error ($\epsilon_{\text{allometric}}$) refers to errors in estimating AGB from Lorey's height and was estimated from the relations developed from calibration plots. Regional allometric equations (Fig. 1) provided uncertainty in AGB estimation of 15.8% in Latin America, 21.7% in Africa, and 25.1% in Asia. The allometric errors are estimated for small areas representing the effective footprint of GLAS Lidar samples (0.25–0.5 ha).

Sampling error ($\epsilon_{\text{sampling}}$) has two components: (i) the error associated with the representativeness of ground sampling plots and GLAS Lidar shots of the true distribution of the AGB in each region and (ii) the error associated with the AGB of a 1-km pixel due to the spatial variability of forest structure and biomass in a 1-km grid cell. The first component has been accounted for while quantifying the prediction error associated with the MaxEnt modeling of AGB. To quantify the second error, we require forest structure and biomass data over 1-km^2 (100 ha) plot sizes. In the absence of reliable data to quantify the sampling errors, we approximated the errors using published data for 50 ha of forest on Barro Colorado Island, Panama as a representative example of undisturbed tropical forest (21). Using a sampling unit of 0.25 ha (i.e., the area of one GLAS Lidar footprint), the AGB is approximately normally distributed around its mean (6). The variation of AGB within the 50-ha plot is high and has the mean SD of 128 Mg ha^{-1} and a 95% confidence interval of 20.1 Mg ha^{-1} . Using the sampling size equation for 95% confidence interval, Chave et al. (6) found the number of 0.25-ha or 0.5-ha plots required to estimate the biomass of a 50-ha plot with $\pm 10\%$ uncertainty varies between 16 (at 0.5 ha) and 26 (at 0.25 ha). By assuming the biomass of a 1-km pixel (100-ha) plot has the same variations as that of the 50-ha plot, and by using a minimum of five samples (GLAS Lidar shots), the uncertainty of AGB estimation with 95% confidence will increase and be bounded between 17.8% ($10 \times \sqrt{16/5}$) and 22.8% ($10 \times \sqrt{26/5}$). By examining the spatial variability of AGB in smaller plots (5–10 ha) available in Rondonia Brazil, Peru, Lope National Park in Gabon, and Sabah in Indonesia, we

were able to verify that variability did not exceed $\pm 20\%$. A similar study by Chave et al. (22) in French Guiana using data from an 80-ha plot concluded that estimating the biomass of the entire plot with only a 1-ha plot will yield a typical error of $\pm 20\%$ due to spatial variability and sampling errors for capturing large trees.

Prediction error ($\epsilon_{\text{prediction}}$) was calculated from prediction probabilities of the MaxEnt model. We used two approaches to quantify the prediction uncertainty. To estimate the average error or uncertainty, we used the 67,730 pixels with AGB values representing the mean AGB of a 1-km pixel. These estimates provided an average relative error in percentage of the mean AGB for pixels not used in training the spatial model (Fig. S4). To estimate the spatial uncertainty that resulted from the MaxEnt model, we used the predicted probability for each biomass range to calculate the root mean squared error,

$$\sigma_{\hat{B}} = \sqrt{\frac{\sum_{i=1}^N (B_i - \hat{B})^2 P_i}{\sum_{i=1}^N P_i}}, \quad [\text{S5}]$$

with the relative uncertainty for each pixel given by: $\epsilon_{\text{prediction}} = \sigma_{\hat{B}} / \hat{B} \times 100$. This relation is used for the entire AGB map, using the values of MaxEnt probabilities P_i for each biomass range class B_i and the predicted biomass value \hat{B} . The prediction uncertainty includes both the sampling error associated with the representativeness of the training data of the actual spatial distribution of AGB and the model predictions.

Finally, we propagated the errors through the entire process by assuming all errors were independent and random and quantified the uncertainty in estimating AGB (ϵ_{AGB}) using

$$\epsilon_{\text{AGB}} = \left(\epsilon_{\text{measurement}}^2 + \epsilon_{\text{allometry}}^2 + \epsilon_{\text{sampling}}^2 + \epsilon_{\text{prediction}}^2 \right)^{1/2}, \quad [\text{S6}]$$

where all errors terms are the relative errors associated with the AGB at a pixel (23). The error propagation equation assumes that the errors are uncorrelated and on the average the covariant terms are neglected. To demonstrate the propagation of errors in modeling AGB of a 1-km pixel in Latin America, we use the measurement error of Lorey's height at 13.7% and calculate the relative error in AGB by multiplying the height relative error by 1.9701 using the equation given in Fig. 1B to arrive at 26.9%, allometric error of 15.7% to convert the Lorey's height to AGB, a maximum sampling error of 22.8%, and assuming a maximum prediction error of 21% at the pixel level, to arrive at the maximum relative error of AGB of 43.9%.

The uncertainty in total biomass carbon is calculated by combining the errors in AGB and BGB. The error in BGB is estimated from both the error in the prediction of BGB from AGB through the allometric Eq. S2 (23.2%) (9) and the error in AGB,

$$\epsilon_{\text{BGB}} = \left[(23.2)^2 + (0.89\epsilon_{\text{AGB}})^2 \right]^{1/2}, \quad [\text{S7}]$$

where ϵ_{BGB} is the relative uncertainty in BGB. The coefficient, 0.89 is the power of AGB in the allometric relationship used in Eq. S2 (9, 23). To estimate the uncertainty in total carbon at the pixel level, we use a similar equation to Eq. S6 to propagate the errors in estimating ϵ_{BC} , the relative uncertainty associated with the total biomass carbon (BC) defined by $\text{BC}(\text{Mg C ha}^{-1}) = (\text{AGB} + \text{BGB})/2$.

Uncertainty at the national scale. Uncertainty at the national or regional scale is estimated by increasing the sample area and propagating the errors from pixel scale to the national or regional scale. We calculate the national- or regional-level SE (Mg C ha^{-1}) by summing the errors from all pixels using

$$\sigma_{\text{national}} = \sqrt{\sum_{i=1}^N (\text{BC}_i \epsilon_{\text{BC}_i})^2}, \quad [\text{S8}]$$

where N is the number of pixels within the national boundary, and BC_i and ϵ_{BC_i} are the total carbon and its relative uncertainty associated with pixel i , respectively. Although the SE is large at the national scale, the relative error declines rapidly as N increases according to

$$\epsilon_{\text{national}} = \frac{\sigma_{\text{national}}}{\sum_{i=1}^N \text{BC}_i}. \quad [\text{S9}]$$

The national-scale uncertainty analysis was also performed for variable size ($N \times N$) windows over the carbon map to examine the reduction of relative uncertainty as sampling size increases. We found the relative uncertainty stayed bounded below $\pm 5\%$ for $n = 10$ (10,000 ha) and below $\pm 1\%$ for $n = 100$ (1,000,000 ha) over the entire carbon map.

Uncertainty from forest spatial and temporal dynamics. The benchmark map represents the state of the carbon stock in forests for the year circa 2000. The coarse spatial resolution and the temporal changes in forest cover at the subpixel level due to deforestation dynamics may introduce uncertainty in carbon estimates. As the spatial unit of the map is 1 km, the spatial model underestimates the carbon density of the remaining forest for a pixel with subpixel forest cover.

In addition, satellite image data used to develop the spatial distribution of forest biomass are primarily from the period of 2000–2001 with $>83\%$ of pixels from the year 2000, $\sim 16\%$ from the year 2001, and a small fraction of pixels ($<1\%$) extracted from subsequent years (2002–2003), which are largely over areas where forest cover has not changed for the period (25). However, the use of multitemporal imagery for modeling carbon distribution will introduce an uncertainty in carbon stocks of pixels that changed between 2000 and 2001. Both uncertainties from coarse spatial resolution and temporal imagery have almost no impact on evaluating the carbon stock at national or regional scales. However, they may introduce errors when estimating emissions from small-scale (<1 km) deforestation or degradation. These errors can be corrected by developing an area-correction factor for carbon density of pixels with subpixel forest cover change.

Comparison with published results. The comparison of the benchmark map with an earlier result over the Amazon Basin (ref. 14) was performed at three steps: (i) The earlier map was produced in biomass class range. The map was converted to continuous biomass numbers by choosing the median value of each biomass class at the pixel level and assigning zero values for bare and grass savanna classes. (ii) The difference between the two maps over the Amazon basin was calculated and values outside the uncertainty range of the benchmark map were colored for negative and positive differences. (iii) The maps were visually compared with the vegetation map of the Rio Negro basin of central Amazonia (25) where significant differences were observed over areas of swamp and white sand vegetation (Fig. S5). Lorey's height inferred from GLAS data over these forests is significantly lower than that over other terra firme forests in Amazonia because of their low stature and biomass density (26).

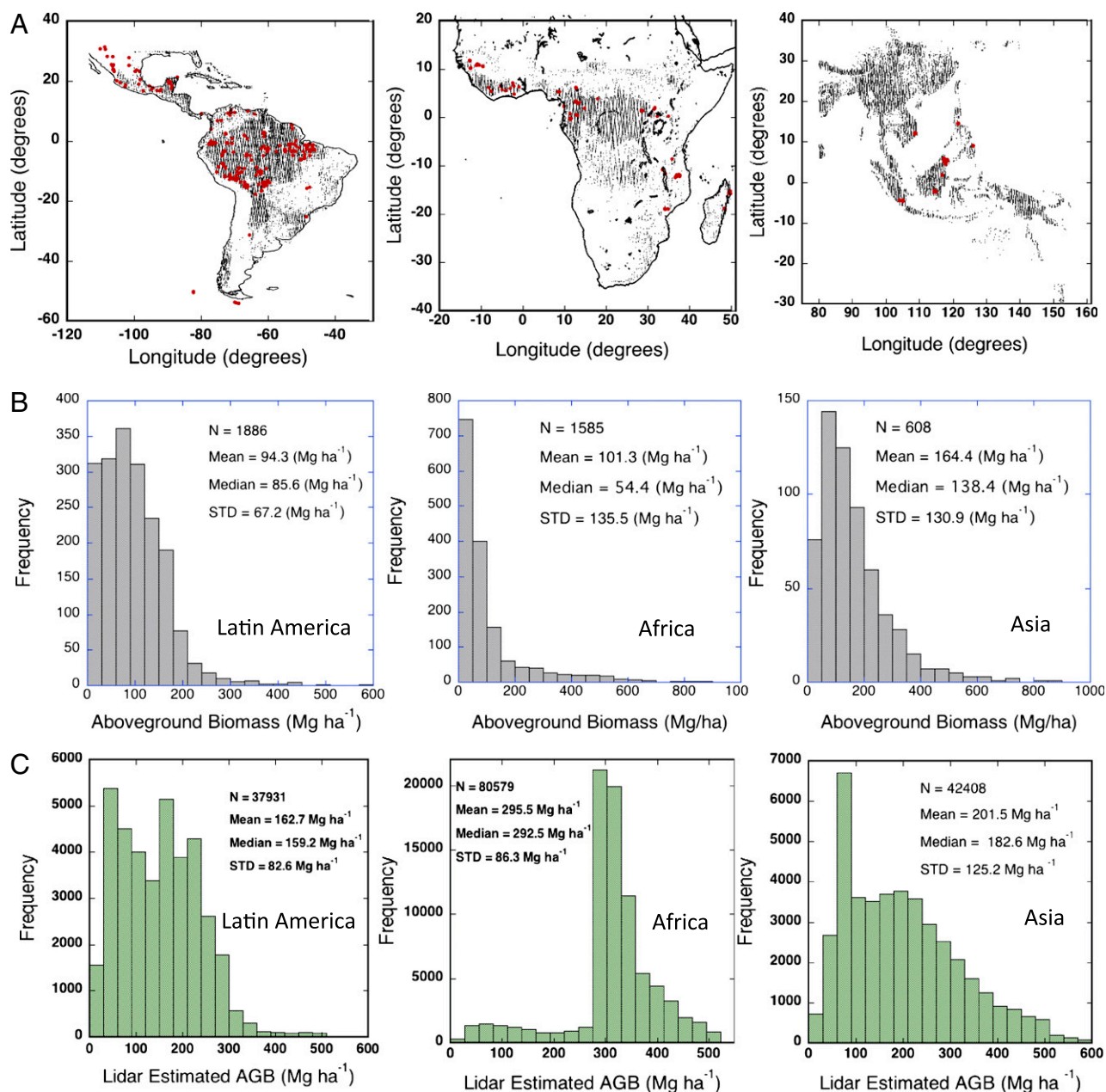


Fig. S1. Distribution of the location and magnitude of aboveground biomass values collected from in situ measurements and derived from GLAS Lidar data over tropical forests across three continents. (A) Geographical distribution of training and test data used in AGB spatial modeling with red dots representing the 1-km pixels associated with 4,079 in situ field plots and gray dots showing 160,918 pixels with AGB values derived from Lidar data across orbital tracks. (B) Distribution of AGB values from in situ data representing an uneven and unsystematic sampling of forest biomass across three continents. (C) Distribution of AGB values from GLAS Lidar data showing a widespread and representative sampling of AGB of forest types across three continents.

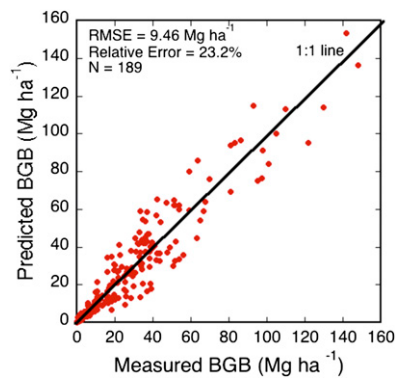


Fig. S2. Predicted versus measured BGB using an allometric equation and data from 189 field plots from Mokany et al. (9). The solid line shows the 1:1 relationship and the RMSE and the relative errors are calculated at 95% confidence interval from bootstrapping cross-validation. The data were obtained from Mokany et al. (9).

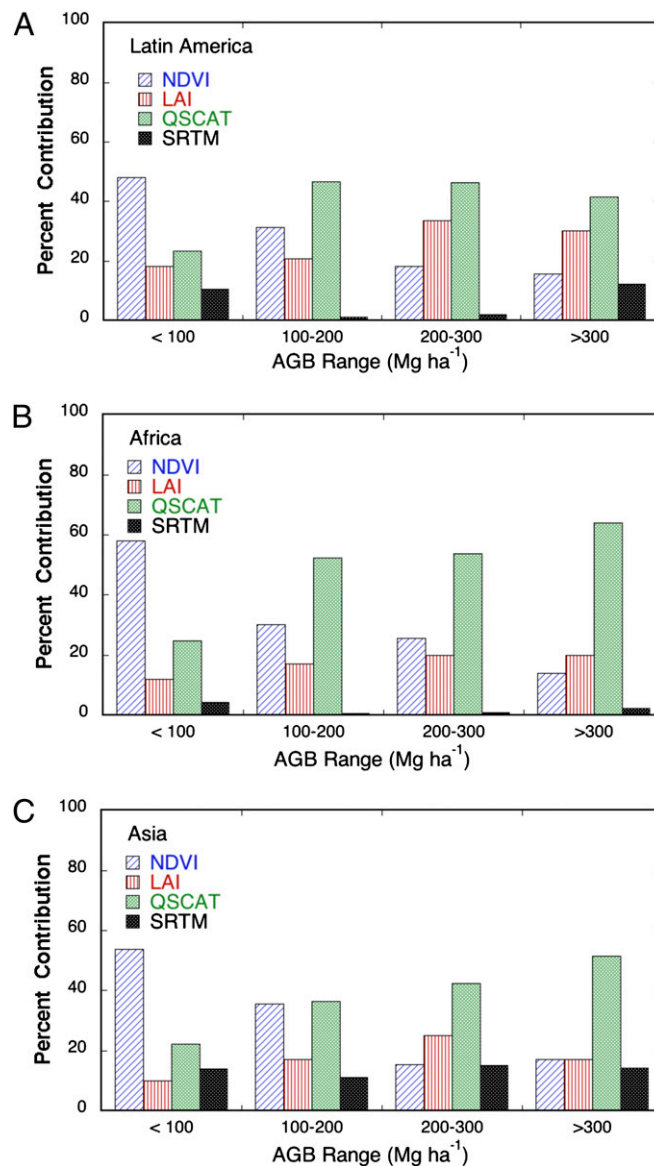


Fig. S3. Relative contribution of remote sensing variables in modeling the spatial variability of AGB. The contributions are calculated using jackknife analysis and shown for four categories of remote sensing observations and derived variables for (A) Latin America, (B) Africa, and (C) Asia. The remote sensing variables included in the analysis are five MODIS NDVI, three MODIS LAI, four QSCAT, and two SRTM derived metrics.

