Remote Sensing of Environment xxx (xxxx) xxx

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Mapping global forest canopy height through integration of GEDI and Landsat data

Peter Potapov [a,*], Xinyuan Li [a], Andres Hernandez-Serna [a], Alexandra Tyukavina [a], Matthew C. Hansen [a], Anil Kommareddy [a], Amy Pickens [a], Svetlana Turubanova [a], Hao Tang [a], Carlos Edibaldo Silva [a], John Armston [a], Ralph Dubayah [a], J. Bryan Blair [b], Michelle Hofton [a]

[a] Department of Geographical Sciences, University of Maryland, United States
[b] NASA Goddard Space Flight Center, United States

ARTICLE INFO

Keywords:
GEDI
Landsat
Lidar
Forest height
Forest monitoring
Time-series

ABSTRACT

Consistent, large-scale operational monitoring of forest height is essential for estimating forest-related carbon emissions, analyzing forest degradation, and quantifying the effectiveness of forest restoration initiatives. The Global Ecosystem Dynamics Investigation (GEDI) lidar instrument onboard the International Space Station has been collecting unique data on vegetation structure since April 2019. Here, we employed global Landsat analysis-ready data to extrapolate GEDI footprint-level forest canopy height measurements, creating a 30 m spatial resolution global forest canopy height map for the year 2019. The global forest height map was compared to the GEDI validation data (RMSE = 6.6 m; MAE = 4.45 m, $R^2 = 0.62$) and available airborne lidar data (RMSE = 9.07 m; MAE = 6.36 m, $R^2 = 0.61$). The demonstrated integration of GEDI data with time-series optical imagery is expected to enable multidecadal historic analysis and operational forward monitoring of forest height and its dynamics. Such capability is important to support global climate and sustainable development initiatives.

1. Introduction

Forests (defined here as wildland, managed, and planted tree cover) perform essential ecosystem functions that sustain the Earth climate and biosphere and provide a wide array of ecosystem services. Forest ecosystems are functioning as the largest terrestrial carbon storage (Dixon et al., 1994) and sink (Pan et al., 2011). Forest regrowth (natural restoration and tree plantations) is responsible for more than 60% of the forest carbon sink (Pugh et al., 2019), while timber extraction and forest clearing for other land uses are among the largest sources of anthropogenic carbon emissions (Houghton, 1999). Forest restoration, avoiding forest conversion, and sustainable forest management are considered as cost-effective natural climate solutions in stabilizing climate and preventing species extinction (Balmford et al., 2002; Gibson et al., 2011; Griscom et al., 2017).

Spatially exhaustive, multitemporal data on global forest ecosystem structure are required for carbon accounting and parameterizing global-scale biogeochemistry, hydrology, biodiversity, and climate models. Such data are crucial for practical applications including carbon emissions and sequestration reporting, forest degradation assessment, and quantification of the effectiveness of forest restoration initiatives. Specifically, forest canopy height is used to estimate forest aboveground biomass and timber volume, to monitor the effects of forest degradation, to measure the success of forest restoration, and to model other key ecosystem variables such as primary production and biodiversity (Asner et al., 2012; Goetz and Dubayah, 2011). Traditionally, tree height has been measured in the field, and more recently, using Airborne Laser Scanning (ALS) systems. Both methods are expensive and impractical over large geographic extents. Satellite-based forest canopy height earth observation data have previously been limited to NASA GLAS (ICESat) and more recently to ATLAS (ICESat-2) lidar instruments (Lefsky, 2010; Simard et al., 2011; Popescu et al., 2018). However, these instruments are not optimized for vegetation mapping (Montesano et al., 2015; Hansen et al., 2016).

The NASA Global Ecosystem Dynamics Investigation (GEDI) is a new spaceborne lidar instrument operating onboard the International Space Station (ISS) and has been collecting data since April 2019 (Dubayah et al., 2020a). GEDI is a full waveform lidar specifically designed for

* Corresponding author at: Department of Geographical Sciences, University of Maryland, LeFrak Hall, 7251 Preinkert Drive, College Park, MD 20742, United States.
E-mail address: potapov@umd.edu (P. Potapov).

https://doi.org/10.1016/j.rse.2020.112165
Received 21 August 2020; Received in revised form 16 October 2020; Accepted 23 October 2020
0034-4257/© 2020 Elsevier Inc. All rights reserved.

measuring vegetation structure. GEDI data consist of footprint-based measurements collected between ~51.6°N and ~51.6°S globally (data coverage is limited by the ISS orbit inclination). The instrument provides an unprecedented sampling density of forest structural properties and is anticipated to sample 4% of the Earth's land surface over a nominal two-year mission. The GEDI data are used to create a suite of vegetation structure and aboveground biomass products. The footprint-level vegetation canopy height and cover metrics for the first half year of GEDI observations (April–October 2019) are currently processed and available for analysis. Planned gridded products based solely on GEDI sample data will have a spatial resolution of 1 km at the end of two years of observation. Standard GEDI products support core scientific objectives of the GEDI mission: estimation of aboveground biomass of tropical and temperate forests, modeling land surface carbon budgets and understanding the effects of vegetation structure on biodiversity (Dubayah et al., 2020a).

Standard GEDI data products have several limitations for practical applications. The footprint-level products represent a point sample of a limited portion of the land area, leaving most of the land surface without observations. GEDI's transect sampling may not be suitable for targeting rare forest change events (i.e., selective logging), for practical implementation at local scale (e.g., within a logging concession), or tracking events through time. Integration of GEDI sample data with wall-to-wall global optical data time-series, e.g. from Landsat and Sentinel-2 satellites, enables the spatiotemporal extrapolation of GEDI information for consistent multidecadal monitoring of vegetation structure.

The methodology for extrapolating lidar-based sampled forest structure and biomass data using wall-to-wall optical remotely sensed imagery has been prototyped in different forest types and geographic extents. The GLAS-based canopy structure metrics and empirically derived footprint-level biomass measurements were extrapolated nationally (Chi et al., 2015), continentally (Baccini et al., 2008), and globally (Simard et al., 2011) using MODIS optical data. Landsat imagery was used to extrapolate GLAS-based forest canopy height (Hudak et al., 2002; Hansen et al., 2016) and biomass at local (Duncanson et al., 2010) and biome scales (Tyukavina et al., 2015). Integration of ALS and Landsat data enabled the multidecadal analysis of forest extent and structure change at national (Matasci et al., 2018) and regional (Potapov et al., 2019) scales. Multitemporal Landsat data were found to be beneficial compared to other optical data sources for modeling forest structure properties (Lefsky et al., 2001). Our earlier work (Tyukavina et al., 2015; Hansen et al., 2016; Potapov et al., 2019) demonstrated the importance of using Landsat-based land surface phenology metrics for forest structure modeling.

Here, we introduce a medium resolution (30 m per pixel) global map of forest canopy height for the year 2019 based on the integration of GEDI-derived canopy height metrics with Landsat multitemporal surface reflectance data. GEDI data available to date (April–October 2019) were used for model calibration. The optical multitemporal metrics were derived from Landsat analysis-ready data (Potapov et al., 2020). To quantify model uncertainty, we used set-aside GEDI validation data and ALS data collected over the USA, Mexico, the Democratic Republic of the Congo (DRC), and Australia. We also demonstrated the application of annual forest canopy height characterizations for multidecadal forest change assessment. The 2019 global forest canopy height map is available at https://glad.umd.edu/dataset/gedi/.

2. Data

2.1. GEDI canopy height metrics

The GEDI Level 2 data include two products: GEDI L2A Elevation and Height Metrics and GEDI L2B Canopy Cover and Vertical Profile Metrics, available from the NASA/USGS Land Processes Distributed Active Archive Center (Dubayah et al., 2020b, 2020c). The data were collected between April 18 and October 2, 2019. The data represent waveform return metrics for each 25 m diameter GEDI footprint. The footprint data are geolocated and have expected positional errors of 15–20 m (standard deviation). For each footprint, we extracted a set of relative height (RH) metrics RH75, RH90, RH95, and RH100 corresponding to 75th, 90th, 95th, and 100th percentile of energy return height relative to the ground. These data are based on the L2A processing algorithm that uses different settings to control the waveform interpretation (Beck et al., 2020). As a result, each GEDI observation has six different versions of each RH metric corresponding to varying algorithm settings. We excluded minimum and maximum values and calculated the mean of the remaining four values to obtain the RH metric value for each footprint. The metric values were gridded to the Landsat 30 m pixels. The Landsat pixel was selected by the location of the footprint center.

2.2. ALS data

Vegetation structure data collected using airborne lidar instruments (ALS) may be used as a proxy for field measurements (Lovell et al., 2003; White et al., 2016). We employed the ALS data as the reference to select the appropriate GEDI waveform metric for model calibration and for validation of modeled forest height. The ALS data used in our research were collected by the G-LiHT instrument over the United States and Mexico in 2013–2014 (Cook et al., 2013), by ALTM 3100 over the DRC in 2014–2015 (Xu et al., 2017), and by Riegl LMS-Q560 over Australia in 2012–2017 (TERN AusCover, 2020). The ALS data consist of transects and samples distributed to maximize the representation of different forest types. The source ALS observations (point clouds) were transformed into a canopy height model (CHM) which is calculated by taking the maximum above-ground height of the first return within each grid cell. The CHM spatial resolution was 1 m (the USA, Mexico, and Australia) and 2 m (the DRC). The CHM products were manually filtered to exclude buildings, power lines, and sensor artifacts. When the fine-resolution CHM data are aggregated to the 30 m Landsat pixel grid, several statistics can be extracted from the distribution of CHM values per pixel (mean, median, percentiles, etc.) (Matasci et al., 2018). The mean and median statistics do not allow discrimination between heterogeneous landscapes (e.g. savanna) and short closed-canopy forests, and the maximum value is prone to errors in lidar data. The 95th or 90th percentile values are preferable as they represent the top of the tree canopies while avoiding noise in lidar observations. Here, we defined the "forest canopy height" as the 90th percentile of ALS-based canopy height within the Landsat pixel, as in the approach of Potapov et al. (2019).

2.3. Landsat analysis-ready data time-series

The optical observation time-series are derived from Landsat analysis-ready data (GLAD ARD; Potapov et al., 2020). To generate GLAD ARD, we processed all Landsat Collection 1 (Tier 1) data from 1997 to 2019. For each Landsat image, we applied a set of decision tree models to obtain an observation quality layer that represents the per-pixel likelihood of observation contamination by clouds, haze, and cloud shadows. Using MODIS surface reflectance data as a normalization target, we processed Landsat spectral data into normalized surface reflectance. Reflectance normalization reduced the effects of atmospheric scattering and surface anisotropy and ensured spatiotemporal consistency of spectral data. While the normalized surface reflectance is not equal to surface reflectance derived using atmospheric transfer models, it was shown to be closely related to the MODIS Nadir BRDF-Adjusted Reflectance product (Potapov et al., 2020).

Individual Landsat images were subsequently aggregated into 16-day composites using the observation quality layer to prioritize clear-sky images. Each 16-day composite contains normalized surface reflectance values for blue, green, red, NIR, and two SWIR bands, brightness temperature, and the observation quality layer. The 16-day GLAD ARD are publicly available globally for the 1997–2019 interval (https://glad.

geog.umd.edu/ard/home). The data are stored in geographic coordinates with a spatial resolution of 0.00025° (approximately 30 m at the Equator) and organized as a set of 1 × 1° tiles. The tile format facilitates parallel data processing.

3. Methods

The global forest height mapping methodology was based on a per-pixel machine-learning algorithm (regression tree) that predicts forest height value from spatiotemporal multispectral Landsat data. The dependent variable for model calibration is the 95th percentile relative height metric from GEDI selected using ALS data as a reference. The independent variables are Landsat multitemporal metrics that were derived from the GLAD ARD time-series. The regression tree model was calibrated and applied to each individual Landsat GLAD ARD tile (1 × 1°) in a "moving window" mode. The global product was validated using set-aside GEDI observations and ALS-based forest height data.

3.1. Selecting GEDI calibration data

To select the highest quality training and validation data, we filtered the available GEDI data. We selected only the observations collected in power beam mode (that is, using full strength lasers), during the night (to limit the background noise effects of reflected solar radiation), with beam sensitivity ≥0.9 and where the range of predicted ground elevations among the six algorithms was ≤2 m. Beam sensitivity is a measure of signal-to-noise related to the canopy cover through which a waveform has sufficient energy to penetrate to the ground (Hofton and Blair, 2019). Data collected during the leaf-off season in temperate and boreal forests were excluded. This filtering resulted in 372 million GEDI samples suitable for model calibration and validation. We randomly separated the GEDI samples into training (90%) and validation (10%) data sets.

The distribution of filtered GEDI data was highly unequal (Fig. 1), ranging from just a few samples to nearly 200,000 samples per 1 × 1° Landsat GLAD ARD tile. The average number of GEDI samples per tile was 31,000. The highest number of samples per tile was observed in the northern part of the GEDI data range due to the convergence of ISS ground tracks. Cloud cover significantly affected the number of high-quality GEDI samples. GEDI data availability was relatively low (14,000 samples per tile on average) for tiles located within humid tropical forests.

We used a comparison of the reference forest canopy height, defined as the 90th percentile of ALS-based canopy height within the Landsat pixel with GEDI waveform return metrics to select the best GEDI metric as training data for the Landsat-based forest canopy height model. For this intercomparison we used Landsat pixels with both ALS data and high quality GEDI observations ($n = 23{,}491$). We analyzed GEDI relative height (RH) metrics RH75, RH90, RH95, and RH100 and found that the RH95 metric had the highest correlation with the 90th percentile of ALS-based canopy height per Landsat data pixel (Fig. 2; Table 1). Metric RH90 underestimated canopy height (mean difference 2.3 m) and RH100 overestimated (mean difference − 2.7 m) compared to RH95 (mean difference 0.7 m). The RH95 metric was therefore selected as the training variable to calibrate our global forest height model.

3.2. Generating Landsat multitemporal metrics

Land surface phenology information in the form of multitemporal spectral metrics has an established use in operational forest mapping and monitoring at continental and global scales (Hansen et al., 2013) and has been shown to facilitate forest height modeling (Tyukavina et al., 2015; Hansen et al., 2016; Potapov et al., 2019). Here we created a set of multitemporal spectral metrics to serve as consistent land surface phenology inputs for global annual forest height mapping and to overcome the inter-annual and seasonal inconsistency of clear-sky data availability, following the approach of Potapov et al. (2019, 2020). The metrics were derived for each year from corresponding annual time-series of normalized surface reflectance (23 16-day composites per year). The per-pixel 16-day time-series data is composed only of clear-sky observations during the snow-free season. To facilitate spatiotemporal consistency of the reflectance time-series, we implemented gap-filling of the missing 16-day data. First, we searched for data collected during the four preceding years. If no clear-sky data exist, we applied a linear regression to estimate missing reflectance values. The resulting annual source data consist of the spectral reflectance of all bands, selected normalized difference indices, spectral variability vegetation index (SVVI, Coulter et al., 2016), and tasseled cap greenness (Crist, 1985). A small portion of land area within humid tropical forests was not processed due to persistent cloud cover.

The annual multitemporal metrics represent a set of Landsat spectral reflectance distribution and phenology statistics extracted from time-series data. To generate the annual metrics, we ranked reflectance and index values and extracted minimum, maximum, and median values, averages between quartiles, and standard deviation per year. We also ranked observations by the value of a vegetation index and by surface temperature to extract spectral information corresponding to specific phenological stages, as described in Potapov et al. (2020). Phenology metrics that reflect salient points of the phenology cycle (start, peak, end of the season; growing season average and total) were based on the normalized difference vegetation index (NDVI) time-series. The full list of multitemporal metrics used for forest height modeling is presented in Fig. 3. To incorporate spatial features, in addition to each spectral metric we calculated the focal average of the metric value within the 3 × 3-pixel kernel.

Spectral metrics were supplemented by land surface elevation derived from the Shuttle Radar Topography Mission (SRTM) 90 m Digital Elevation Data downloaded from http://srtm.csi.cgiar.org

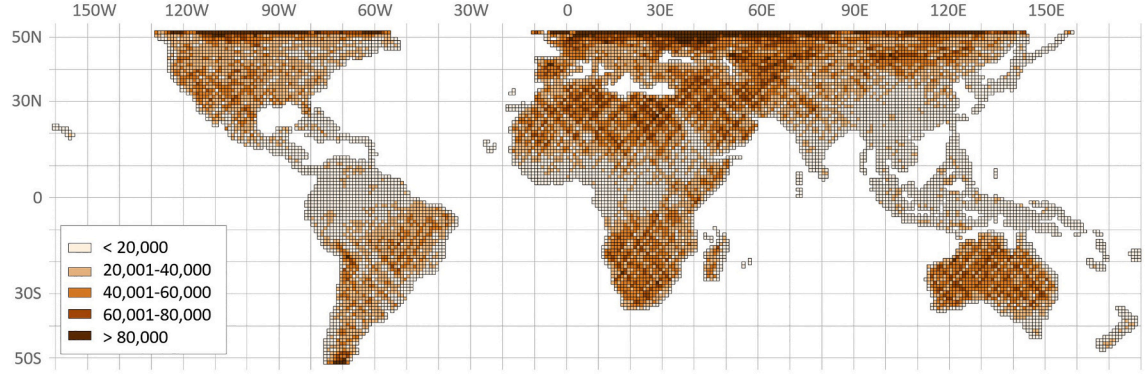


Fig. 1. Number of GEDI samples per 1 × 1° Landsat GLAD ARD tile. Only high-quality samples are included (372 million samples total).

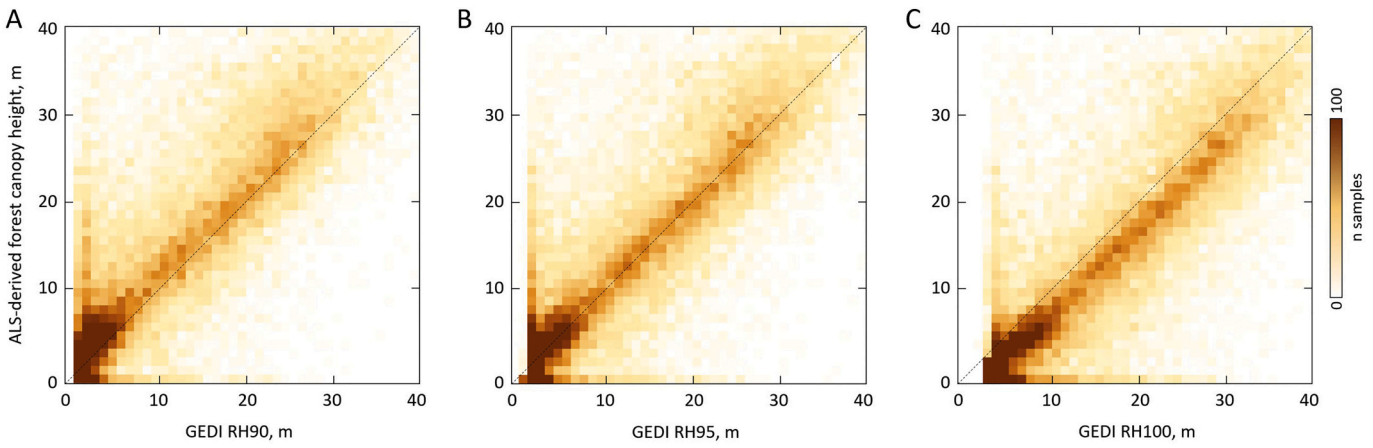


Fig. 2. Comparison of the 90th percentile ALS-based forest canopy height with GEDI RH90 (A), RH95 (B), and RH100 (C) metrics. The ALS data are collected in the USA, Mexico, the DRC, and Australia; n = 23,491.

Table 1
Comparison statistics between the 90th percentile ALS-based forest canopy height and GEDI relative height (RH) metrics. The ALS data are collected in the USA, Mexico, the DRC, and Australia.

	RH75	RH90	RH95	RH100
Root-mean-square error (RMSE)	8.8	6.9	6.5	7.2
Mean absolute error (MAE)	6.1	4.5	4.2	5.3
Mean difference	5.1	2.3	0.7	−2.7
Coefficient of determination (R^2)	0.65	0.7	0.71	0.7

(Jarvis et al., 2008).

3.3. Forest height model calibration

We implemented the bagged regression trees ensemble method (Breiman et al., 1984; Breiman, 1996) to model forest height. A regression tree is a machine learning algorithm that allows modeling of continuous variables. It has been successfully used for forest structure and biomass mapping at the regional to global scales (Peterson and Nelson, 2014; Ahmed et al., 2015; Tyukavina et al., 2015; Hansen et al., 2016; Matasci et al., 2018; Potapov et al., 2019). To reduce model overfitting, ensemble learning methods are usually employed, either in the form of a random forest (Matasci et al., 2018) or using bagging (bootstrap aggregation) techniques (Potapov et al., 2019). Here we implemented the latter as an ensemble of 25 regression trees each calibrated using an independent sample of training data; the map value represents the ensemble median.

We used GEDI RH95 as a dependent variable for the regression tree model. A standard definition of a "tree" used for carbon accounting and reporting is woody vegetation with the potential to reach a minimum height of 2–5 m (UNFCCC, 2006). Here, we define a tree as woody vegetation with the height of 3 m or above. Consistent with this definition, we assigned the forest height for the GEDI samples with RH95 < 3 m to zero.

Several issues of the original GEDI data affected model calibration. First, GEDI height data do not discriminate trees from buildings, confounding vegetation height mapping in urban areas. Second, we found an overestimation of vegetation height in areas of complex topography, especially on steep slopes (Fig. 4). To limit the adverse effects of these calibration data errors we used existing tree canopy cover products for the years 2000 and 2010 (Hansen et al., 2013) as a reference. For a training pixel that had no tree cover in both 2000 and 2010, we set the forest height value to zero if one of the following criteria was also true: (a) the pixel was classified as water according to the GLAD ARD quality layer; (b) the pixel was located within an urban area (defined using the global urban dataset of Florczyk et al., 2019) and the Landsat-based annual maximum NDVI value was below 0.5; (c) the SRTM elevation was above 1500 m and the slope was above 6°; or (d) the Landsat-based annual maximum NDVI value was below 0.3. These empirical thresholds were defined by visual analysis of Landsat image composites, NDVI metrics and GEDI observations within tropical and temperate forests. The slope threshold was informed by Simard et al. (2011).

The Landsat multitemporal metrics for the year 2019 and SRTM elevation served as independent variables for the regression tree model. In total, 546 metrics were used as independent variables for model calibration.

A random set of training samples was extracted from each tile within the GEDI data range to calibrate the regression tree model. To address the disproportionate GEDI training data distribution (Fig. 1), we implemented equal sampling, extracting 500 random samples from each tile. The sampling rate for non-forest training was 1.5 times higher than for forest training to reduce effect of RH95 overestimation errors. The sampling procedure was repeated to create 25 independent training sets for each tile.

For each GLAD ARD tile (1 × 1°), a separate regression tree ensemble model was calibrated using training data collected within the target and neighboring tiles. This "moving window" model calibration approach allowed us to create locally calibrated models while maintaining global map consistency. The neighboring tiles were defined using a 12° radius. The radius was selected via visual analysis of the products calibrated using different sets of neighboring tiles for test areas in temperate and tropical forests. In total, between 100,000 and 200,000 samples were used to calibrate each regression tree model. An ensemble of 25 trees was calibrated and applied to each tile to create the 2019 forest height map. Overall, we calibrated 11,860 individual ensemble models within the GEDI data range. The same approach was used to extrapolate mapping south of 52°S.

The prototype for the boreal forest height map, where GEDI data were absent (north of 52°N), was calibrated using training data extracted from the modeled forest height between 40°N and 52°N. We used continental models for the boreal forest height mapping. Three models were calibrated separately for (a) North America, (b) Europe and West Siberia, and (c) East Siberia and the Far East. The calibration was performed in two steps. First, models calibrated using GEDI forest height data were applied within each boreal region. Initial model results overestimated forest height within wetland areas and tundra. To address this error, we collected additional non-forest training in areas where our forest height map disagreed with the year 2010 tree canopy cover (Hansen et al., 2013). Then, the final boreal models were calibrated using the 40°N - 52°N training data augmented with non-forest training collected north of 52°N.

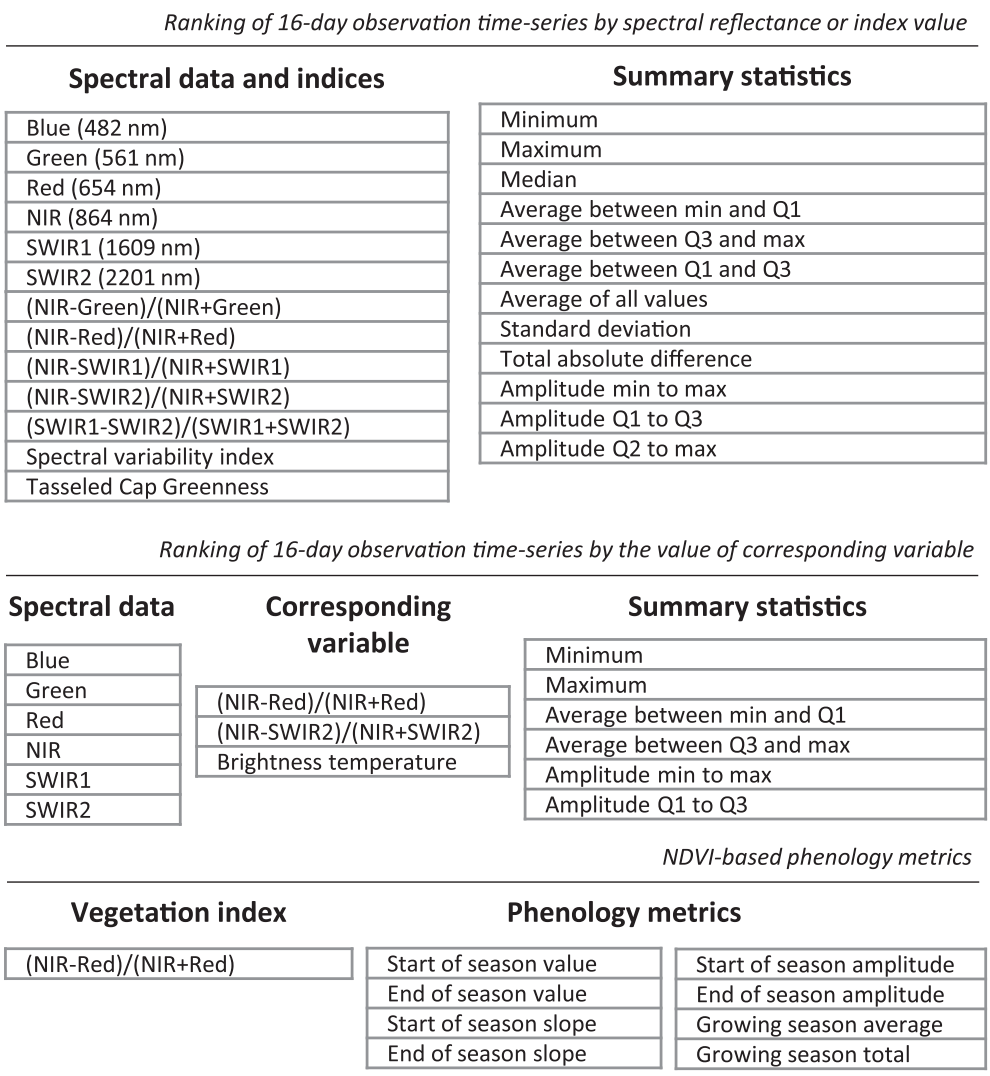


Fig. 3. Landsat multitemporal metrics. The first set of metrics represents summary statistics calculated from 16-day observation time-series ranked by the spectral reflectance or index value. The ranking is performed independently for each spectral band or index. The second set of metrics represents summary statistics calculated from 16-day observations ranked by the value of corresponding variable. The third set represents phenology metrics extracted from NDVI time-series. Q1, Q2, and Q3 stand for 1st, 2nd, and 3rd quartiles.

To demonstrate the forest height monitoring potential of the presented methodology, we applied the forest height model annually (2000–2019) for selected 1 × 1° tiles. The Landsat multitemporal metrics were calculated for each year following approach described in section 3.2. For each tile, we use the respective model calibrated using the 2019 data.

3.4. Product validation

A set-aside 10% of the GEDI observations served as the primary validation data set. The GEDI validation data are assumed to represent a probability sample within 52°N - 52°S extent. To address per-tile variation of global GEDI data distribution (Fig. 1), we extracted 1000 random samples from each tile (all samples were used for the tiles with less than 1000 samples). In total, we extracted 13.4 million validation samples. The RH95 metric gridded to Landsat pixel served as the reference value.

The ALS data available for selected areas in the USA, Mexico, the DRC, and Australia served as an additional validation data source. While these data do not represent a probability sample, they span different forest biomes and represent diverse forest types and ages. The disadvantage of the ALS data is that they were primarily collected over forests, while treeless ecosystems and heterogeneous landscapes (savannas, agroforestry) were not covered by airborne observations (specifically, in Africa). To reduce the effect of time difference between the ALS data collection (2012–2017) and Landsat modeling (2019), we excluded all Landsat pixels where forest loss was detected after 2009 using data from Hansen et al. (2013). We addressed the disproportionate allocation of ALS data between continents by selecting an equal number of random ALS samples from North America, Africa, and Australia (1.3 million samples total). The 90th percentile of CHM canopy height per Landsat pixel was used as the reference data for Landsat-modeled forest height map validation.

We quantified forest height model uncertainty by comparing GEDI RH95 value and ALS-derived canopy height with Landsat-derived products at the output pixel/grid scale within the Landsat-mapped forest class (i.e. pixels with modeled forest height ≥ 3 m). Inclusion of non-forest areas from the Landsat-based map would increase the overall agreement and improve the linear regression fit between the map and validation datasets (due to a large number of agreement samples within non-forest land), but would obscure the relationship between the product and validation data within forests. To quantify forest detection accuracy, we transformed both validation and Landsat-based data into binary forest and non-forest classes using a set of thresholds to define different forest classes (≥3 m, ≥5 m, ≥10 m). The ≥3 m threshold is consistent with tree definition used for the current analysis; the 5 m minimum height threshold is commonly used for forest area reporting (Sasaki and Putz, 2009); and the ≥10 m threshold was used to map tall forests only. Using the resulting confusion matrices, we calculated map accuracy statistics for each validation data source and each forest class

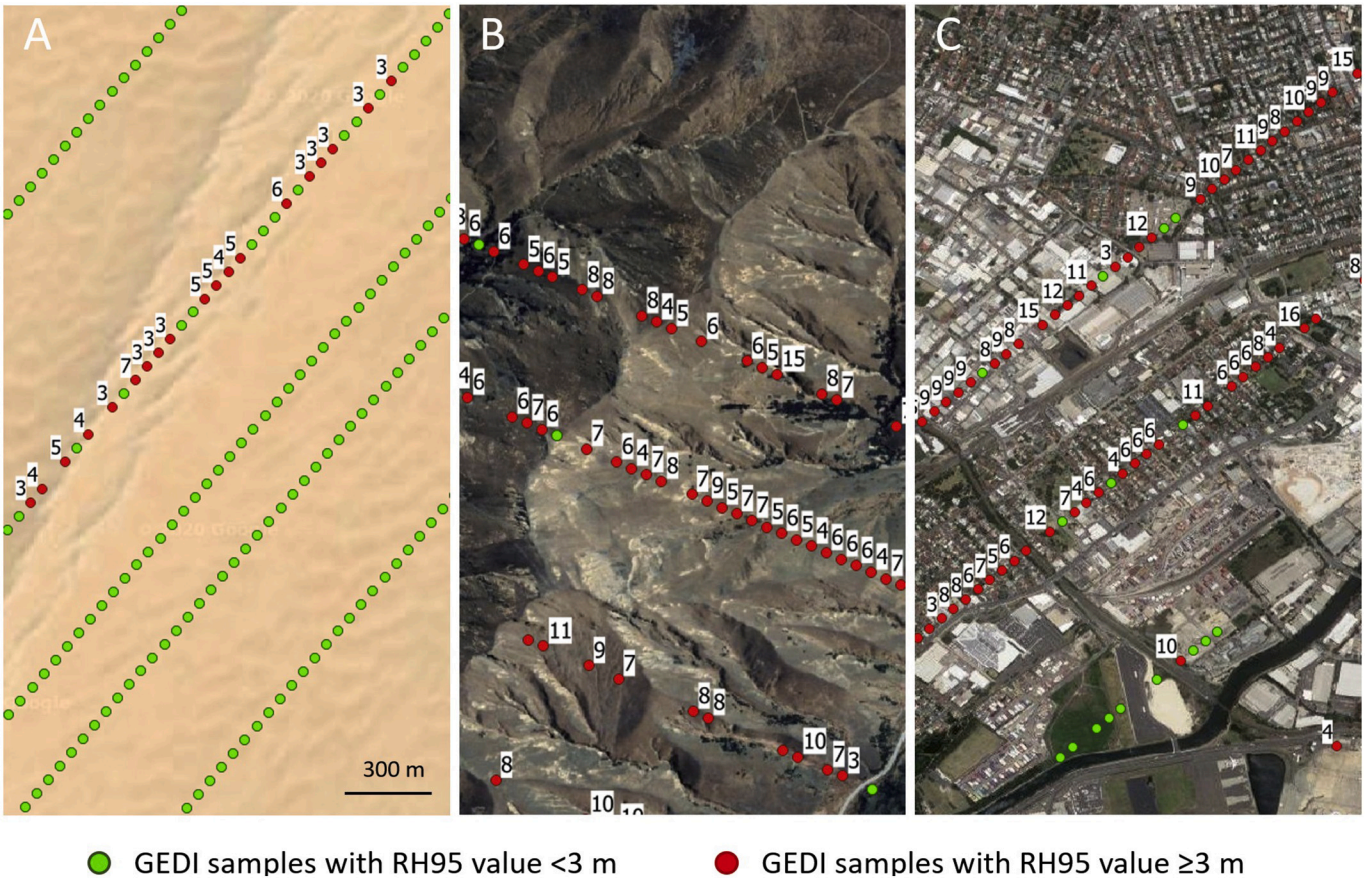


Fig. 4. Visual examples of differences between GEDI RH95 and forest canopy height. The RH95 values (meters) are shown for samples with value ≥3 m. The Google Earth imagery serves as the background. A – Sand dunes in Saudi Arabia (Lat/Long 52.33;20.85); B – Alpine grasslands, New Zealand (169.66;-44.57); C – Build-up area, Sydney, Australia (151.19;-33.93).

threshold. We did not validate the boreal forest height product prototype.

4. Results

Our 2019 global forest canopy height map demonstrates the potential of spaceborne lidar and optical data integration providing an overview of forest height at 30 m spatial resolution (Fig. 5 A). The full-resolution map is available at https://glad.umd.edu/dataset/gedi/. Map subsets demonstrate canopy height variation related to natural forest types, clearly separating tall Blue Ridge Mountain forests and lower canopy height forests of Piedmont Plateau in Maryland (Fig. 5 B); shorter flooded swamp forests and taller *terra firma* forests in the DRC (Fig. 5 C); shorter dry deciduous lowland forests and taller primary mountain evergreen forests in Cambodia (Fig. 5 D). Land management practices also can be identified by forest height: taller natural forests and shorter suburban tree cover in the USA; low canopy height of secondary forests affected by shifting cultivation in Africa; lower canopy height of rubber plantations compared to remaining tall natural forests in Southeast Asia.

The comparison of the Landsat-based forest height map with the GEDI RH95 metric from the validation set-aside data (Fig. 6 A) yielded RMSE = 6.6 m; MAE = 4.45 m, and $R^2 = 0.62$. The mean difference of −1.0 m indicates that the Landsat-based product underestimates forest height compared to validation data. The distribution of differences between Landsat-modeled and validation GEDI RH95 values for strata representing each integer height value from 3 to 40 m (Fig. 7 A) shows good correspondence between the map and validation data except for canopy height underestimation within short (<7 m) and tall (>30 m) forests.

The modeled forest height and GEDI RH95 metrics were converted to forest/non-forest classes using a set of tree height thresholds and accuracies of the resulting Landsat-based maps were quantified (Table 2). Accuracy metrics show that using a higher threshold (≥10 m) to define forests yielded higher overall accuracy with more balanced omission and commission errors (corresponding to user's and producer's accuracy metrics) compared to forest classes defined using lower height thresholds. The forest class defined using the ≥3 m threshold demonstrated substantial omission (reflected by the lower producer's accuracy) compared to GEDI validation data. The comparison of the Landsat-based model with set-aside GEDI data was done without applying any filters to the validation data. Thus, GEDI RH95 artifacts (i.e. forest height commission on slopes and within urban areas) affected the data intercomparison.

The comparison of our global map with the ALS validation data (Fig. 6 B) reveals higher uncertainty of the Landsat-modeled product, with RMSE = 9.07 m; MAE = 6.36 m, $R^2 = 0.61$. The mean difference of −3.8 m indicates underestimation of forest height by the model. The error distribution (Fig. 7 B) suggests that the mean underestimation was greater within short (3 m) and tall forests (above 21 m height). We should note that the comparison with the ALS data is affected by the disproportional sampling of tall forests by the airborne data. Specifically, in the DRC the ALS data were mostly collected over humid tropical forests and do not include savannas and dry miombo ecoregions, causing oversampling of tall forests in the aggregate reference dataset. The forest map validation (Table 2) yielded higher and balanced user's and producer's accuracies of the forest class compared to GEDI validation data. The lower overall accuracy compared to GEDI

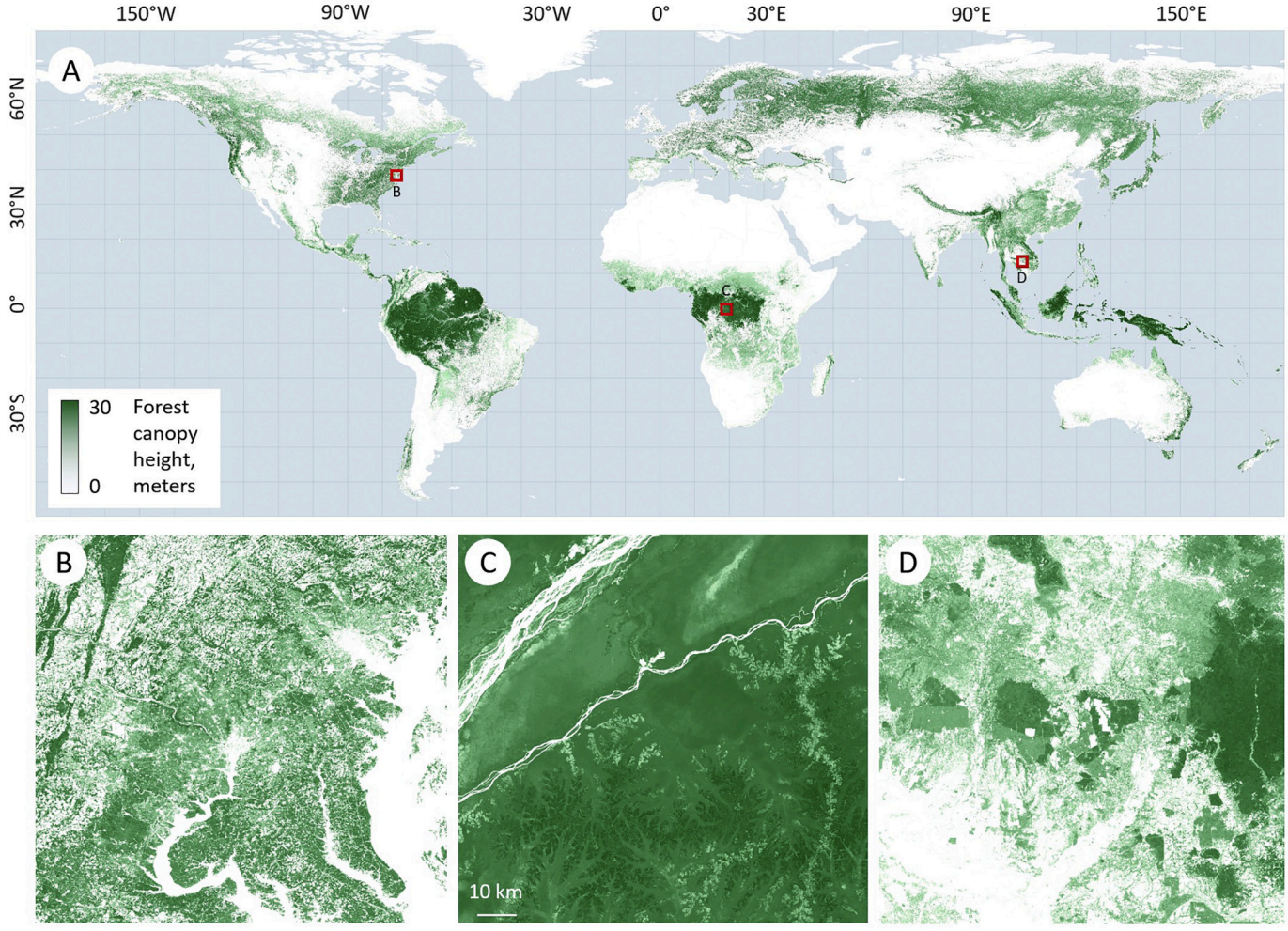


Fig. 5. Global forest cover height map for the year 2019 produced through the integration of GEDI data (April–October 2019) and multitemporal metrics derived from Landsat GLAD ARD (Potapov et al., 2020). A. Global map overview. Close-up examples illustrating different forest and land management types (B) in the USA (77°W;38.9°N); (C) in the Democratic Republic of the Congo (18.8°E;1°N) and (D) in Cambodia (105°E;13.2°N).

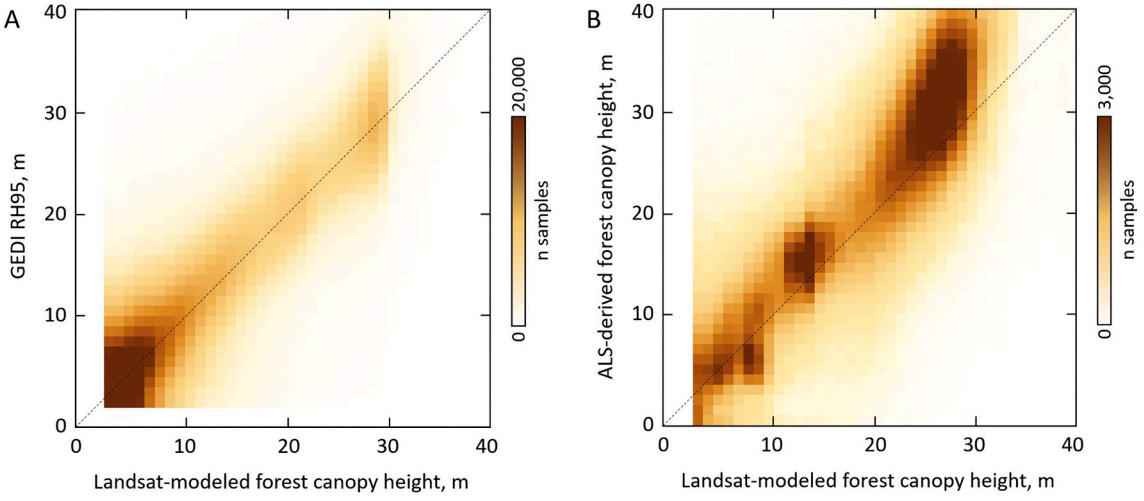


Fig. 6. A. Comparison of the GEDI RH95 for set-aside validation data set with Landsat-modeled forest canopy height product (between 52°N and 52°S); n = 3,631,862. B. Comparison of the ALS-based reference forest canopy height with Landsat-modeled forest canopy height product. The ALS data are collected in the USA, Mexico, DRC, and Australia; n = 1000,909.

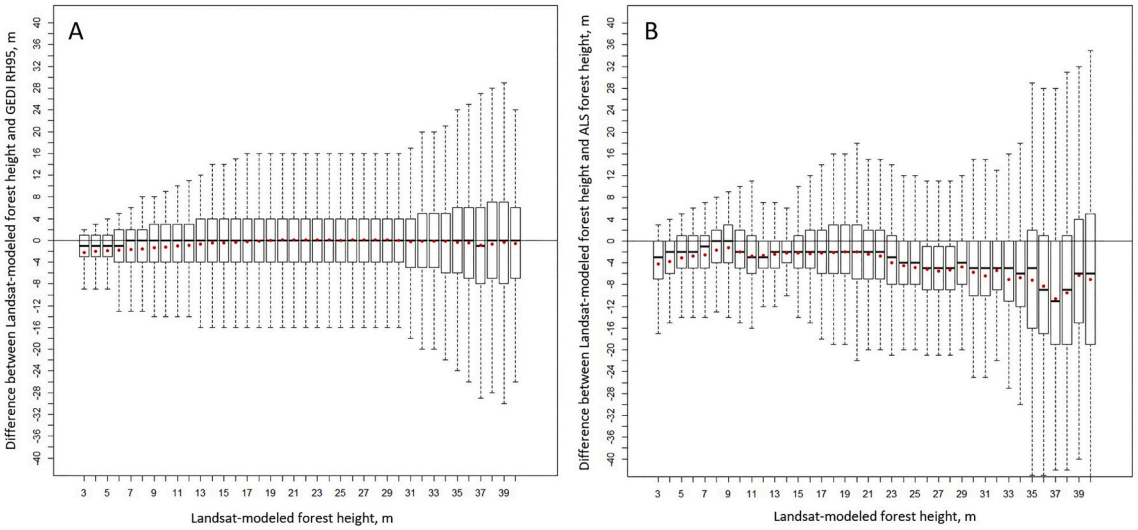


Fig. 7. Distribution of differences between Landsat-modeled forest height and (A) GEDI RH95 validation data and (B) ALS forest height validation data per strata defined by each integer value of Landsat-based forest height map between 3 and 40 m. Red dots represent strata means. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 2
Accuracy of binary forest maps derived from a Landsat-based forest height map validated with GEDI RH95 and ALS-based forest height validation data. Both Landsat-based map and reference data were converted into forest/non-forest classes using a set of tree height thresholds (3, 5 and 10 m). User's and producer's accuracies were computed for the forest class.

	Reference: GEDI RH95			Reference: ALS forest height		
Forest class threshold	Overall accuracy	User's accuracy	Producer's accuracy	Overall accuracy	User's accuracy	Producer's accuracy
≥3 m	83.0	93.0	62.6	87.6	97.4	87.9
≥5 m	87.8	89.0	66.7	88.1	96.1	88.7
≥10 m	92.9	85.5	71.4	90.1	95.1	90.0

validation data may be explained by the disproportionally high presence of forests in the ALS validation data.

One of the most important applications of the Landsat-based forest canopy height model is annual monitoring, enabling multidecadal analysis of forest dynamics. To illustrate the forest height monitoring, we applied the Landsat-based forest height model derived for the year 2019 to multitemporal metrics for each year, 2000 to 2019, for selected test regions. Presented examples show that annual time-series of canopy height may be used to detect stand-replacement dynamics, as well as forest degradation and enhancement (Figs. 8 and 9). An example from the Brazilian state of Mato Grosso (Fig. 8 A) illustrates that forest canopy height time-series capture not only forest clearing, but also degradation caused by selective logging and forest recovery after selective logging and surface fires. A data subset within selective logging operations on the island of Borneo (Fig. 8 B and 9 D) highlights forest height decrease as a result of selective tree removal followed by forest restoration. An example of intensive plantation forestry in Arkansas (Fig. 8 D and 9 B) reflects forest height change in a planting/harvesting cycle, while the clear-cut logging example from Oregon (Fig. 8 C and 9 A) reflects abrupt change in tree height following by slow regrowth. These examples illustrate the potential utility of annual maps of forest height in capturing forest loss and gain.

5. Discussion

Medium spatial resolution optical satellite data, such as Landsat, do not allow direct measurement of forest canopy height. However, this variable can be modeled by relating reference measurements to information on forest species composition, age, and growing conditions that can be derived from multitemporal spectral data. The model accuracy depends on the availability and quality of reference data, spatiotemporal consistency of optical data, and the geographic extent of model application. Due to the non-linear relation between surface reflectance and vegetation structure, non-parametric machine learning tools are usually employed for vegetation structure modeling (Hansen et al., 2016; Matasci et al., 2018; Potapov et al., 2019; Lang et al., 2019). Models calibrated within a limited geographic extent (i.e., at the national scale as demonstrated by Lang et al., 2019) may be expected to have higher accuracy compared to global or continental models. To compare with our tile-based model, we have tested a single global Landsat-based forest height model calibrated using the random subset of training data. The global model demonstrated higher uncertainty (RMSE 9.07 m and 10.43 m for tile-based and global models, respectively) and a lower correlation with the ALS reference data (R^2 0.61 and 0.53, respectively). We noticed that the global model underestimated tree cover in heterogeneous landscapes (i.e. agroforestry) and within broadleaf temperate forests (Fig. 10). Overall, tree models run over large areas performed less well than more locally calibrated models. GEDI affords the possibility of spatially disaggregating model runs, overcoming the limitations of decision trees when employing large training data sets and complex tree models across large geographic extents.

The availability of GEDI data per tile is the main limitation for the tile-based model calibration. The model accuracy may be compromised if training data do not adequately represent forest types and age cohorts. In regions with frequent cloud cover (i.e., Central Africa, Fig. 11), the limited high quality GEDI observations within a tile may disproportionally represent one forest type (in Fig. 11, primary humid tropical forests), while other land cover types (in Fig. 11, settlements and shifting cultivation complex) may be omitted. Using the data from neighboring tiles allowed us to collect representative training samples. We expect that a complete set of two years of GEDI data will permit a radius reduction for the neighboring tile search which will improve the local

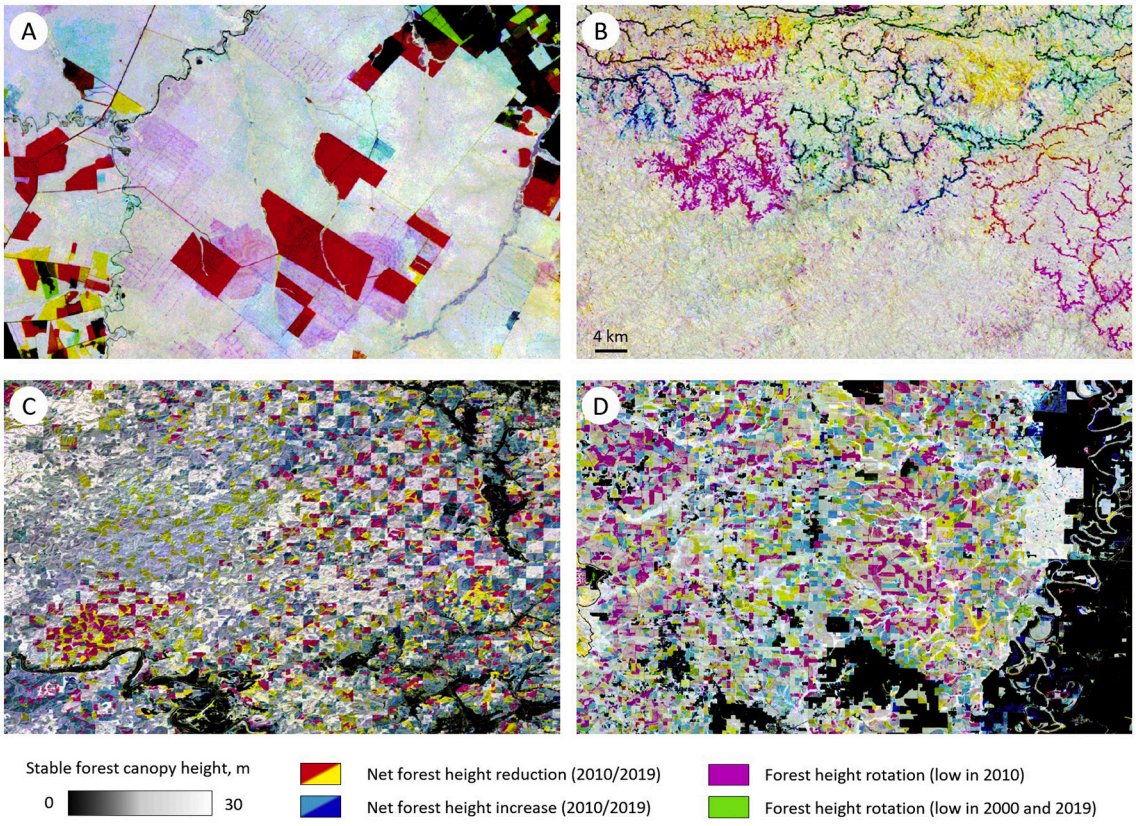


Fig. 8. Examples of forest canopy height time-series maps. Each map represents an RGB composite of forest canopy height layers from the years 2000 (red), 2010 (green), and 2019 (blue). The gray tones represent stable forest and non-forest areas and colors represent forest dynamics including unidirectional changes (forest height decrease or increase detected by 2010 or 2019) and rotation (when a land parcel experienced forest canopy height reduction in 2010 followed by increase by 2019, and the opposite). (A) Mato Grosso, Brazil (54.48°W;11.51°S); (B) Island of Borneo, border between Malaysia and Indonesia (113.44°E;1.27°N); (C) Oregon, USA (123.03°W;43.38°N); and (D) Arkansas, USA (91.29°W;33.18°N). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

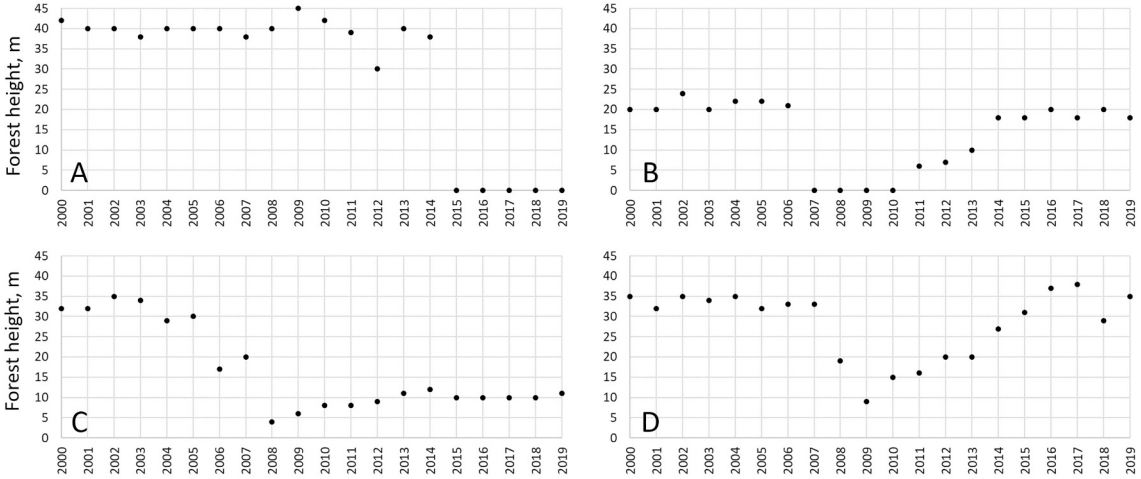


Fig. 9. Examples of forest canopy height time-series profiles for sample pixels representing different land use practices. (A) Clearcut in primary forest in Oregon, USA (123.7°W;43.69°N); (B) Pine plantation management in Arkansas, USA (91.71°W;33.4°N); (C) Oil palm plantation establishment after primary forest clearing on Borneo, Malaysia (113.38°E;2.51°N); (D) Selective logging in primary forest on Borneo, Malaysia (113.68°E;2.09°N).

model precision.

Several issues still exist in GEDI data that affect model calibration. Slopes affect tree cover height estimation by both spaceborne (Lefsky et al., 2007) and airborne (Neigh et al., 2014) lidar instruments. While GEDI footprint size reduces the effect of slopes on canopy height estimation (Dubayah et al., 2020a), additional correction of the footprint-level data may be required to reduce canopy height uncertainty. Properties of GEDI measurements such as the pulse width may also result in large differences between GEDI RH metrics and true canopy height, particularly in sparse low stature vegetation. Through visual data analysis, we found that forest canopy height is overestimated when using GEDI RH95 within temperate and subtropical mountain

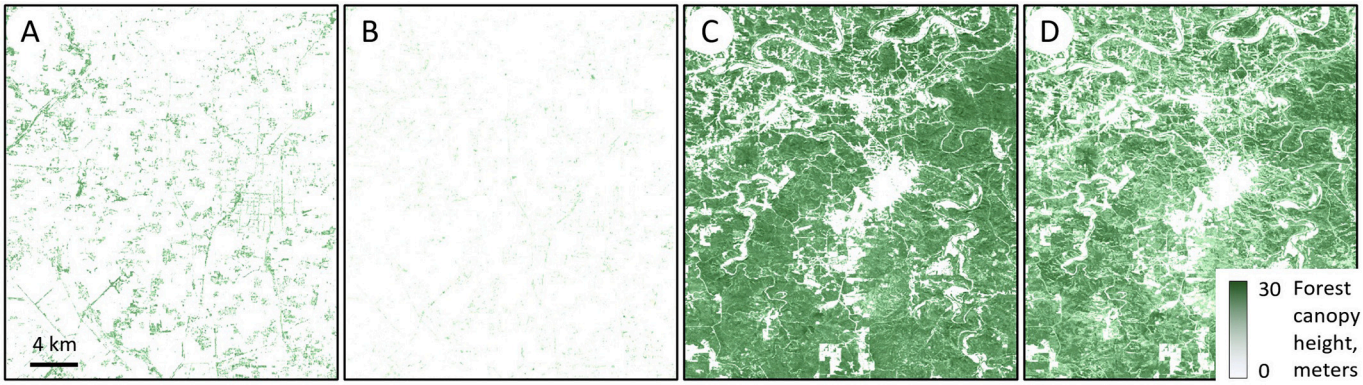


Fig. 10. Examples of forest canopy height maps created using tile-based (A, C) and global (B, D) models. (A, B) Agroforestry landscape in Shandong province, China (116.5°E;35.69°N); (C, D) Mixed forests in Missouri, USA (92.19°W;37.72°N).

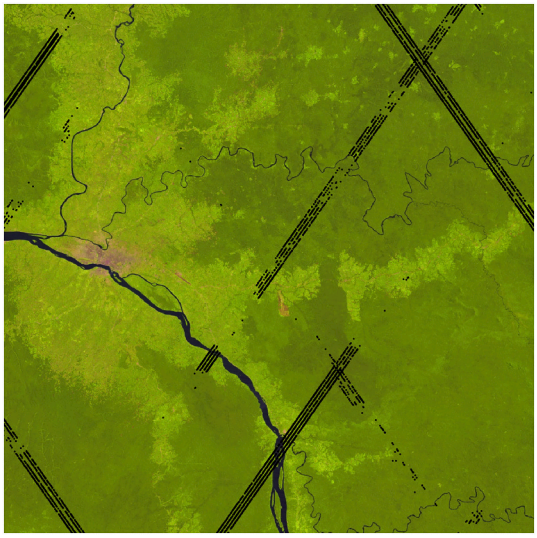


Fig. 11. Available high quality GEDI observations (black dots) over the 1 × 1° tile area in the DRC (tile center at 25.5°E;0.5°N). A Landsat clear-sky 2019 image composite serves as a background.

grasslands, e.g. in New Zealand and Lesotho, highlighting limitations of using a waveform energy quantile (GEDI RH95) as a direct proxy for ALS-based canopy height in areas of complex topography and sparse tree cover. The GEDI RH95 metric does not discriminate between the height of trees and man-made objects (buildings, power lines). Residual geolocation error of GEDI data affects calibration of the Landsat-based models and footprint level validations. Shifts of a few meters can result in several meters of height error. Subsequent versions of the GEDI data should achieve geolocation accuracies of <10 m (standard deviation) and include improvements to canopy height estimation and ground detection algorithms, an urban mask, and other refinements which will address some of these issues.

The observed underestimation of the Landsat-modeled forest height compared to the ALS data may be a consequence of several factors. The medium resolution of Landsat, which is the primary mapping data source in the current study, may limit the ability to correctly map the height of tallest forests (Hansen et al., 2016). Texture metrics derived from higher spatial resolution Sentinel-2 data were demonstrated to have better sensitivity to the taller canopy heights (Lang et al., 2019). The regression tree model employed in the current study uses the median as the output value for a terminal node. That approach has limitations in mapping shortest and tallest forests, especially if the training data contain errors. The higher sampling rate for non-forest training which we implemented to reduce the effect of GEDI RH95 errors may have affected the lower (3-5 m) forest height mapping.

Due to the limitations of optical data caused by atmospheric contamination and relatively low observation frequency, we may expect to observe high-frequency noise in the annual model outputs (Fig. 9 demonstrates the annual fluctuations of estimated forest height for test pixels). Data limitations precluded mapping forest height for parts of permanently cloudy regions in Central Africa. However, the approach prototyped by Potapov et al. (2019) that implements temporal filtering, gap-filling, and integration of the annual structure time-series with annual forest loss detection data has been shown to improve spatio-temporal consistency. Additionally, data from the future Landsat 9 satellite and integration of Sentinel-2 and Landsat time-series will increase observation frequency, which may reduce model uncertainty and eliminate the need for temporal filtering.

6. Conclusion

This study illustrates that integration of spaceborne optical (Landsat) and lidar (GEDI) data sets may enable global annual forest canopy height monitoring at a 30 m/pixel spatial resolution. As we move towards the operational version of the presented method, we are planning to improve model calibration by using additional and refined GEDI data as they become available. Subsequent reprocessing of GEDI waveforms should result in significantly improved geolocation and optimal algorithm settings, both of which should improve GEDI canopy height retrieval accuracies. We have demonstrated that a Landsat-based forest canopy height model calibrated using GEDI data from the recent years can be applied to historic Landsat imagery, enabling forest structure analysis from the 1980s to present, and providing a tool for operational monitoring of forest degradation and restoration. The next steps include (a) expanding forest height models beyond the GEDI data range using LVIS and ICESat-2 ATLAS data acquired in boreal forests for both training and validation, and (b) creating integrated annual forest canopy height, cover, and disturbance maps to monitor global carbon emissions from forestry operations and land conversions. Such operationally updated integrated forest structure and change maps will inform climate mitigation policy initiatives, including UNFCCC REDD+ (Reducing Emissions from Deforestation and Forest Degradation), the New York Declaration on Forests, and the Bonn Challenge land restoration initiatives (UNFCCC Secretariat, 2016; IUCN, 2018; NYDF Assessment Partners, 2019).

Declaration of Competing Interest

None.

Acknowledgments

Ralph Dubayah, John Armston, and Matthew C. Hansen are funded by the NASA contract number NNL15AA03C to the University of Maryland for the development and execution of the GEDI mission. This research was partially funded by the Global Forest Watch and the NASA SERVIR program (grant number 80NSSC20K0158).

References

Ahmed, O.S., Franklin, S.E., Wulder, M.A., White, J.C., 2015. Characterizing stand-level forest canopy cover and height using Landsat time series, samples of airborne LiDAR, and the random Forest algorithm. ISPRS J. Photogramm. Remote Sens. 101, 89–101.

Asner, G.P., Mascaro, J., Muller-Landau, H.C., Vieilledent, G., Vaudry, R., Rasamoelina, M., Hall, J.S., Van Breugel, M., 2012. A universal airborne LiDAR approach for tropical forest carbon mapping. Oecologia 168, 1147–1160.

Baccini, A., Laporte, N., Goetz, S.J., Sun, M., Dong, H., 2008. A first map of tropical Africa's above-ground biomass derived from satellite imagery. Environ. Res. Lett. 3 (4), 045011.

Balmford, A., Bruner, A., Cooper, P., Costanza, R., Farber, S., Green, R.E., Munro, K., 2002. Economic reasons for conserving wild nature. Science 297 (5583), 950–953.

Beck, J., Armston, J., Hofton, M., Luthcke, S., 2020. GLOBAL *Ecosystem Dynamics Investigation (GEDI) Level 02 User Guide*. U.S. Geological Survey, Earth Resources Observation and Science Center, Sioux Falls, South Dakota, USA.

Breiman, L., 1996. Bagging predictors. Mach. Learn. 24, 123–140.

Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J., 1984. Classification and Regression Trees. Wadsworth and Brooks/Cole, Monterey, California.

Chi, H., Sun, G., Huang, J., Guo, Z., Ni, W., Fu, A., 2015. National forest aboveground biomass mapping from ICESat/GLAS data and MODIS imagery in China. Remote Sens. 7 (5), 5534–5564.

Cook, B., Corp, L., Nelson, R., Middleton, E., Morton, D., McCorkel, J., Masek, J., Ranson, K., Ly, V., Montesano, P., Cook, B.D., Corp, L.A., Nelson, R.F., Middleton, E. M., Morton, D.C., McCorkel, J.T., Masek, J.G., Ranson, K.J., Ly, V., Montesano, P.M., 2013. NASA Goddard's LiDAR, hyperspectral and thermal (G-LiHT) airborne imager. Remote Sens. 5 (8), 4045–4066.

Coulter, L.L., Stow, D.A., Tsai, Y.-H., Ibanez, N., Shih, H., Kerr, A., Benza, M., Weeks, J. R., Mensah, F., 2016. Classification and assessment of land cover and land use change in southern Ghana using dense stacks of Landsat 7 ETM+ imagery. Remote Sens. Environ. 184, 396–409.

Crist, E., 1985. A TM tasseled cap equivalent transformation for reflectance factor data. Remote Sens. Environ. 301–306.

Dixon, R.K., Solomon, A.M., Brown, S., Houghton, R.A., Trexier, M.C., Wisniewski, J., 1994. Carbon pools and flux of global forest ecosystems. Science 263 (5144), 185–190.

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., 2020a. The global ecosystem dynamics investigation: high-resolution laser ranging of the Earth's forests and topography. Sci. Remote Sens. 1, 100002.

Dubayah, R., Hofton, M., Blair, J.B., Armston, J., Tang, H., Luthcke, S., 2020b. GEDI L2A elevation and height metrics data global footprint level V001 [data set]. *NASA EOSDIS Land Processes DAAC*. Accessed 2020-04-11 from. https://doi.org/10.5067/GEDI/GEDI02_A.001.

Dubayah, R., Tang, H., Armston, J., Luthcke, S., Hofton, M., Blair, J.B., 2020c. GEDI L2B Canopy Cover and Vertical Profile Metrics Data Global Footprint Level V001 [Data set]. *NASA EOSDIS Land Processes DAAC*. Accessed 2020-04-11 from. https://doi.org/10.5067/GEDI/GEDI02_B.001.

Duncanson, L.I., Niemann, K.O., Wulder, M.A., 2010. Integration of GLAS and Landsat TM data for aboveground biomass estimation. Can. J. Remote. Sens. 36 (2), 129–141.

Florczyk, A.J., Corbane, C., Ehrlich, D., Freire, S., Kemper, T., Maffenini, L., Sabo, F., 2019. GHSL Data Package 2019. Luxembourg. EUR 29788.

Gibson, L., Lee, T.M., Koh, L.P., Brook, B.W., Gardner, T.A., Barlow, J., Sodhi, N.S., 2011. Primary forests are irreplaceable for sustaining tropical biodiversity. Nature 478 (7369), 378–381.

Goetz, S., Dubayah, R., 2011. Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. Carbon Management 2, 231–244.

Griscom, B.W., Adams, J., Ellis, P.W., Houghton, R.A., Lomax, G., Miteva, D.A., Woodbury, P., 2017. Natural climate solutions. Proceedings of the National Academy of Sciences 114 (44), 11645–11650.

Hansen, M.C., Potapov, P.V., Moore, R., Hancher, M., Turubanova, S.A.A., Tyukavina, A., Thau, D., Stehman, S.V., Goetz, S.J., Loveland, T.R., Kommareddy, A., 2013. High-resolution global maps of 21st-century forest cover change. Science 342 (6160), 850–853.

Hansen, M.C., Potapov, P.V., Goetz, S.J., Turubanova, S., Tyukavina, A., Krylov, A., Kommareddy, A., Egorov, A., 2016. Mapping tree height distributions in sub-Saharan Africa using Landsat 7 and 8 data. Remote Sens. Environ. 185, 221–232.

Hofton, M., Blair, J.B., 2019. Algorithm Theoretical Basis Document (ATBD) for GEDI Transmit and Receive Waveform Processing for L1. Goddard Space Flight Center, Greenbelt, MD. and L2 Products.

Houghton, R.A., 1999. The annual net flux of carbon to the atmosphere from changes in land use 1850–1990. Tellus B 51 (2), 298–313.

Hudak, A.T., Lefsky, M.A., Cohen, W.B., Berterretche, M., 2002. Integration of lidar and Landsat ETM+ data for estimating and mapping forest canopy height. Remote Sens. Environ. 82 (2–3), 397–416.

IUCN, 2018. Getting Started with the Bonn Challenge.

Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E., 2008. Hole-filled seamless SRTM data V4, International Centre for Tropical Agriculture (CIAT), available from. http://srtm.csi.cgiar.org.

Lang, N., Schindler, K., Wegner, J.D., 2019. Country-wide high-resolution vegetation height mapping with Sentinel-2. Remote Sens. Environ. 233, 111347.

Lefsky, M.A., 2010. A global forest canopy height map from the moderate resolution imaging Spectroradiometer and the geoscience laser altimeter system. Geophys. Res. Lett. 37.

Lefsky, M.A., Cohen, W.B., Spies, T.A., 2001. An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. Can. J. For. Res. 31 (1), 78–87.

Lefsky, M.A., Keller, M., Pang, Y., De Camargo, P.B., Hunter, M.O., 2007. Revised method for forest canopy height estimation from Geoscience Laser Altimeter System waveforms. J. Appl. Remote. Sens. *1* (*1*), 013537.

Lovell, J.L., Jupp, D.L., Culvenor, D.S., Coops, N.C., 2003. Using airborne and ground-based ranging lidar to measure canopy structure in Australian forests. Can. J. Remote. Sens. 29 (5), 607–622.

Matasci, G., Hermosilla, T., Wulder, M.A., White, J.C., Coops, N.C., Hobart, G.W., Zald, H.S., 2018. Large-area mapping of Canadian boreal forest cover, height, biomass and other structural attributes using Landsat composites and lidar plots. Remote Sens. Environ. 209, 90–106.

Montesano, P.M., Rosette, J., Sun, G., North, P., Nelson, R.F., Dubayah, R.O., Ranson, K. J., Kharuk, V., 2015. The uncertainty of biomass estimates from modeled ICESat-2 returns across a boreal forest gradient. Remote Sens. Environ. 158, 95–109.

Neigh, C., Masek, J., Bourget, P., Cook, B., Huang, C., Rishmawi, K., Zhao, F., Neigh, C.S. R., Masek, J.G., Bourget, P., Cook, B., Huang, C., Rishmawi, K., Zhao, F., 2014. Deciphering the precision of stereo IKONOS canopy height models for US forests with G-LiHT airborne LiDAR. Remote Sens. 6, 1762–1782.

NYDF Assessment Partners, 2019. Protecting and Restoring Forests: A Story of Large Commitments Yet Limited Progress. New York Declaration on Forests Five-Year Assessment Report. Accessible at. forestdeclaration.org.

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Ciais, P., 2011. A large and persistent carbon sink in the world's forests. Science 333 (6045), 988–993.

Peterson, B., Nelson, K.J., 2014. Mapping forest height in Alaska using GLAS, Landsat composites, and airborne LiDAR. Remote Sens. 6 (12), 12409–12426.

Popescu, S.C., Zhou, T., Nelson, R., Neuenschwander, A., Sheridan, R., Narine, L., Walsh, K.M., 2018. Photon counting LiDAR: an adaptive ground and canopy height retrieval algorithm for ICESat-2 data. Remote Sens. Environ. 208, 154–170.

Potapov, P., Tyukavina, A., Turubanova, S., Talero, Y., Hernandez-Serna, A., Hansen, M. C., Saah, D., Tenneson, K., Poortinga, A., Aekakkararungroj, A., Chishtie, F., 2019. Annual continuous fields of woody vegetation structure in the lower Mekong region from 2000-2017 Landsat time-series. Remote Sens. Environ. 232, 111278.

Potapov, P., Hansen, M.C., Kommareddy, I., Kommareddy, A., Turubanova, S., Pickens, A., Adusei, B., Tyukavina, A., Ying, Q., 2020. Landsat analysis ready data for global land cover and land cover change mapping. Remote Sens. 12 (3), 426.

Pugh, T.A., Lindeskog, M., Smith, B., Poulter, B., Arneth, A., Haverd, V., Calle, L., 2019. Role of forest regrowth in global carbon sink dynamics. Proc. Natl. Acad. Sci. 116 (10), 4382–4387.

Sasaki, N., Putz, F.E., 2009. Critical need for new definitions of "forest" and "forest degradation" in global climate change agreements. Conserv. Lett. 2 (5), 226–232.

UNFCCC Secretariat, 2016. Key Decisions Relevant for Reducing Emissions from Deforestation and Forest Degradation in Developing Countries (REDD+) (UNFCCC Secretariat).

Simard, M., Pinto, N., Fisher, J.B., Baccini, A., 2011. Mapping forest canopy height globally with spaceborne lidar. Geophys. Res. Lett. 116, 4021.

TERN AusCover, 2020. Airborne LiDAR - Riegl Q560, point, raster and full waveform, Australian field sites. Obtained from the AusCover facility (http://www.auscover.org.au) of the Terrestrial Ecosystem Research Network (TERN, http://www.tern.org.au). Accessed 1/20/2020.

Tyukavina, A., Baccini, A., Hansen, M.C., Potapov, P.V., Stehman, S.V., Houghton, R.A., Krylov, A.M., Turubanova, S., Goetz, S.J., 2015. Aboveground carbon loss in natural and managed tropical forests from 2000 to 2012. Environ. Res. Lett. 10 (7), 074002.

UNFCCC [United Nations Framework Convention on Climate Change], 2006. *Report of the Conf. Parties serving as the meeting of the Parties to the Kyoto Protocol on its first session, held at Montreal (28 November–10 December 2005)* United Nations Office at Geneva, Switzerland.

White, J.C., Coops, N.C., Wulder, M.A., Vastaranta, M., Hilker, T., Tompalski, P., 2016. Remote sensing technologies for enhancing forest inventories: a review. Can. J. Remote. Sens. 42 (5), 619–641.

Xu, L., Saatchi, S.S., Shapiro, A., Meyer, V., Ferraz, A., Yang, Y., Bastin, J.-F., Banks, N., Boeckx, P., Verbeeck, H., Lewis, S.L., Muanza, E.T., Bongwele, E., Kayembe, F., Mbenza, D., Kalau, L., Mukendi, F., Ilunga, F., Ebuta, D., 2017. Spatial distribution of carbon stored in forests of the Democratic Republic of Congo. Sci. Rep. 7, 15030.