Assessing Amazon rainforest regrowth with GEDI and ICESat-2 data

Milutin Milenković [a,*], Johannes Reiche [a], John Armston [b], Amy Neuenschwander [c], Wanda De Keersmaecker [a,d], Martin Herold [a,e], Jan Verbesselt [a]

[a] Laboratory of Geo-Information Science and Remote Sensing, Wageningen University & Research, Wageningen, the Netherlands
[b] Department of Geographical Sciences, University of Maryland, College Park, MD, USA
[c] Center for Space Research, University of Texas at Austin, Austin, TX, USA
[d] Unit Remote Sensing and Earth Observation Processes, Flemish Institute for Technological Research (VITO), Mol, Belgium
[e] Helmholtz GFZ German Research Centre for Geosciences, Section 1.4 Remote Sensing and Geoinformatics, Telegrafenberg, Potsdam, 14473, Germany

ARTICLE INFO

ABSTRACT

Two novel satellite LiDAR missions —GEDI and ICESat-2— are currently operational and combined provide near-global measurements of forest height and structure. Such data underpin a new era of large-area approaches for measuring forest height in regrowing forests of different ages and assessing associated regrowth rates. Two LiDAR missions further allow for comparing independently derived forest heights and regrowth rates. This study utilized both GEDI and ICESat-2 measurements to assess regrowth rates in regrowing forests of different ages for the Brazilian state Rondônia. We considered 19 data subgroups stratified by beam strength, light condition, beam sensitivity, and waveform processing algorithm to assess the retrieval uncertainty and identify data subgroups associated with the most reliable regrowth estimates. The quality assessment of GEDI and ICESat-2 forest heights over four 50 km long airborne LiDAR strips determined a root mean square error of 4.14 m (CV = 17%) and 5.91 m (CV = 19%) and a mean error of 0.04 m and −2.81 m, respectively. A linear calibration model between satellite- and airborne-LiDAR heights was then derived for each data subgroup and used to calibrate satellite heights. Forest regrowth rates were subsequently estimated for each satellite mission using a space-for-time imputation with forest heights' medians per stand age class. The total growth of GEDI and ICESat-2 median forest heights after 33 years was 20.17 m (SE = 1.3 m) and 20.13 m (SE = 2.8 m), respectively. However, when growth was approximated with different non-linear models, the total growth differed by up to 6%, and the average regrowth rate even by up to 23%. The study revealed that omitting either the calibration step or the removal of secondary-forest-border pixels would result in an underestimation of the regrowth rate by more than 20%. Furthermore, the ICESat-2 weak beams were found unreliable for regrowth retrieval. The study showed that the novel satellite LiDAR data and the proposed methods could assess median forest height growth over large areas. However, forest age errors should also be accounted for in the retrieval uncertainty before comparing the growth estimates across different regions.

Code and data necessary to reproduce the results are freely available on GitHub and Zenodo.

1. Introduction

The Amazon rainforest is undergoing dynamic changes, being the ecozone with the largest forest loss area globally and accommodating 13% of the total tropical forest regrowth area between 2000 and 2012 (Hansen et al., 2013). Forest regrowth plays an important role in global terrestrial carbon sink dynamics and needs to be accurately assessed to reduce the uncertainties in tropical carbon balance calculations (Baccini et al., 2017; Houghton 2003; Pan et al., 2011). Furthermore, spatially

explicit information about forest regrowth is required globally to model its effect on carbon sequestration accurately and better understand the impact of ongoing reforestation (Chazdon et al., 2016). Thus, studying forest regrowth as a part of the disturbance-recovery processes is the key to better understanding forest ecosystem functioning and ecological resilience (Poorter et al., 2016).

Traditionally, forest height growth is determined with repeated field measurements of permanent forest plots (Clark and Clark 2001), but a comprehensive spatio-temporal sampling for large-area monitoring is

both time- and cost-intensive. LiDAR (Light Detection and Ranging) remote sensing, on the other hand, collects systematic and rapid measurements of forest height across forest landscapes from which temporal change between repeat measurements can inform forest growth (Hopkinson et al., 2008; Hyyppä et al., 2008; Næsset and Gobakken 2005). For example, multi-temporal forest heights acquired with repeat-flight airborne LiDAR data were used to assess rainforest regrowth in Indonesia and boreal forest growth in Finland (Boehm et al., 2013; Yu et al., 2004). However, this approach to measure forest growth has not been implemented for the Amazon rainforest due to limited airborne lidar data. As multi-temporal LiDAR measurements are rare, several studies used a space-for-time substitution (Pickett 1989) to estimate forest regrowth, i.e., by combining single-year LiDAR forest heights with forest loss history (Becknell et al., 2018; Lefsky et al., 2005).

Although the above studies demonstrated that airborne LiDAR could assess forest regrowth, its capability to study broad-scale patterns is limited. Airborne LiDAR data provide accurate and dense estimates of forest height, but the acquisition costs are high, and data coverage is often limited to small regions. In contrast to airborne data, the ICESat (The Ice, Cloud, and Land Elevation Satellite, 2003–2009) mission provided satellite LiDAR data that enabled mapping forest heights with global coverage (Lefsky 2010; Simard et al., 2011). ICESat was the first, and until recently, the only satellite LiDAR used to assess forest regrowth (Dolan et al., 2009). ICESat, however, has not been operational since 2009, hampering the assessment of forest regrowth in recent years. Most recently, the GEDI (Global Ecosystem Dynamics Investigation, since 4/2019) and ICESat-2 (since 10/2018) missions provide a new generation of satellite LiDAR instruments producing high-resolution, contemporary forest heights with near-global coverage, i.e., GEDI between 51.6° and ICESat-2 between 88° North and South latitudes, respectively. Hence, these data enable studies of recent forest regrowth rates over large areas.

The GEDI mission utilizes a waveform LiDAR instrument to sample forest heights with footprints of 25 m in diameter, every 60 m along its ascending and descending orbital paths (Dubayah et al., 2020). Recent studies showed that GEDI is capable of mapping vertical forest structure (Marselis et al., 2020; Schneider et al., 2020). GEDI heights have also been extrapolated using multi-temporal Landsat metrics to derive a global forest height map with a 30 m pixel size (Potapov et al., 2021). The ICESat-2 mission carries a photon-counting laser altimeter that records single photons as separate returns within highly overlapping footprints (11 m diameter, 0.7 m along-track sampling) along transects parallel to its ascending and descending orbital paths (Neumann et al., 2019). ICESat-2 was designed primarily for monitoring ice, but several studies indicated the potential of ICESat-2 products for accessing forest height and aboveground biomass (Li et al., 2020; Narine et al., 2019; Neuenschwander and Magruder 2019).

Since April 2019, both GEDI and ICESat-2 have been providing forest height metrics on a near-global scale. In this early exploration phase, most of the studies have focused on the quality assessment of terrain and forest heights with airborne LiDAR data and the sensitivity of errors to different acquisition and environmental parameters. Liu et al. (2021), for example, evaluated the forest height retrieval from both GEDI and ICESat-2 data at sites with boreal-, temperate- and rain-forests across the United States. This study showed that ICESat-2 strong beams and GEDI power beams combined with their night acquisitions provided the lowest forest height root mean square errors (RMSEs) of 5.02 m and 3.56 m, respectively. The quality of ICESat-2 forest heights has been comprehensively assessed over a vast region of boreal forest in Finland, where the lowest forest height RMSE (2.45 m) was obtained again using the ICESat-2 strong beams, night, and summer acquisitions (Neuenschwander et al., 2020a). Adam et al. (2020) assessed the quality of GEDI forest heights over two temperate forest sites in Germany, reporting the median absolute error of 2.98 m and 3.17 m. The study also showed that the errors were sensitive to terrain slope, forest height, and the beam sensitivity parameter. In another quality assessment of GEDI forest

heights in a temperate forest in the Southwest of Spain, the lowest RMSEs (3.03 m and 3.07 m) were reported for two power beams (Quirós et al., 2021). Finally, in a study with simulated GEDI and ICESat-2 data, Duncanson et al. (2020) have shown that ICESat-2 relative height metrics underestimate their GEDI counterparts for a study area with temperate, mixed, and coniferous forests in Sonoma County, California. Until recently, only one study has focused on assessing regrowth rates from satellite LiDAR data (Guerra-Hernández and Pascual 2021). Guerra-Hernández and Pascual (2021) computed average heights increments between single-year GEDI and multi-temporal (2015–2019) national airborne LiDAR data to assess the growth dynamics of fast-growing tree plantations in the Northwest of Spain. A direct comparison between GEDI and ICESat-2 estimated forest heights and forest regrowth rates over the Amazon rainforest has yet to be studied.

In this study, we explore the utility of GEDI and ICESat-2 for assessing regrowth rates and uncertainties in the Amazon rainforest. Therefore, our two main objectives are:

- to define an appropriate methodology for assessing regrowth rates, and
- to derive and compare the forest regrowth rates from GEDI and ICESat-2 forest heights for the Rondônia state, Brazil.

To accomplish these objectives, we first calibrated and assessed the quality of GEDI and ICESat-2 forest heights with airborne LiDAR data at two calibration sites in the Amazon rainforest. Next, we derived regrowth rates using a space-for-time approach where single-year, spatially distributed GEDI and ICESat-2 heights of secondary forest were grouped according to their stand age class. The regrowth rates were derived from an auxiliary stand age map with 33 years of forest age history. We then analyzed the impact of ICESat-2 acquisition parameters such as beam strength, light conditions as well as GEDI beam sensitivity and waveform processing algorithm parameters on the retrieval of forest heights and regrowth rates. Finally, we examined different non-linear models to approximate the forest height distribution across the entire 33-year regrowth period.

2. Materials

2.1. Study area

The assessment of Forest regrowth was performed over the Brazilian federal state Rondônia (237,629 km^2), located in the central-west part of Brazil (Fig. 1). Rondônia has a tropical wet-and-dry climate with relatively uniform average monthly temperatures ranging from 24 °C to 27 °C, and monthly mean precipitations larger than 200 mm from November till March and lower than 20 mm from June till August (Harris et al., 2020). Rondônia is recognized as one of the environmental change hotspots (UNEP 2020), and approximately 29% of its area has undergone deforestation in the last 35 years (MapBiomas 2021). Deforestation started in the 1970s with a fishbone pattern around the main highway connecting the capitals of the Rondônia and Mato Grosso states (Pedlowski et al., 1997). Deforestation resulted in a vast pasture and agricultural area stretching diagonally across the state, and roughly 4% of the forest loss has regenerated to the secondary forest in the last 35 years (MapBiomas 2021). The remaining intact tropical forest is located mainly in the northern and the east-central parts.

2.2. Calibration sites

The satellite LiDAR forest heights were calibrated with airborne LiDAR forest heights available for two calibration sites in the Amazon rainforest. The first site is located close to the Amazon river at the border of the Tapajós National Forest (TNF), the Brazilian state Pará (Fig. 1). Highway BR-163 splits the TNF site such that the western portion contains intact forest and the eastern portion contains forests disturbed
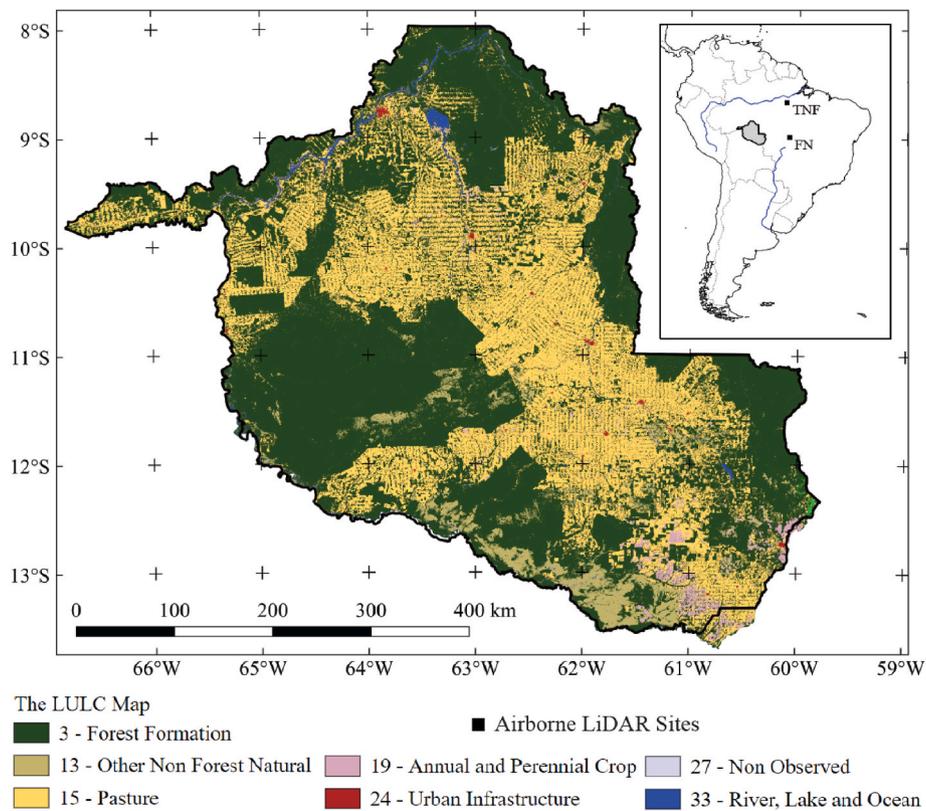
Fig. 1. The location and shape of the Rondônia state with the MapBiomas land use and land cover (LULC) map in the background. The two black squares in the overview figure show the locations of Tapajós National Forest (TNF) and Feliz Natal (FN) sites with airborne LiDAR data.

primarily by selective logging (Fig. 2a). The second site is located close to the southern border of the Amazon rainforest in the Feliz Natal (FN) municipality, the Mato Grosso state, Brazil (Fig. 1). The FN site includes mainly intact forests with a few clearcuts in the southwest (Fig. 2b). Each calibration site included two 50 km long and 200 m wide airborne LiDAR strips, providing a total calibration area of 40 km^2.

2.3. Stand age data

A 30 m stand age map of secondary forest updated annually for the 33 years between 1986 and 2018 was obtained from the benchmark datasets of the secondary forest in Brazil (Silva Junior et al., 2020). This freely available dataset is derived directly from annual, Landsat-based, land use and land cover (LULC) maps, collection 4.1, produced within the MapBiomas project (Souza et al., 2020). In addition to the LULC
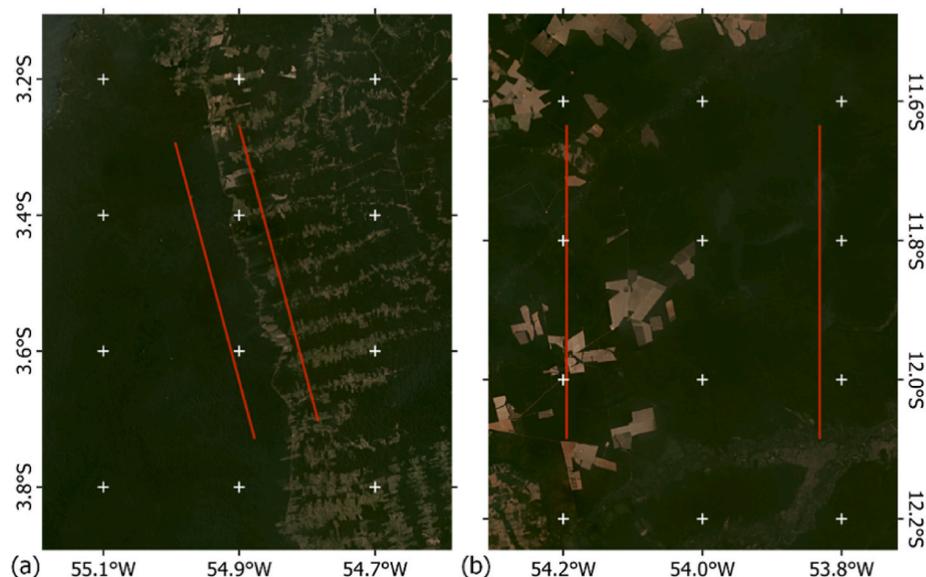
Fig. 2. The calibration sites (a) at the border of the Tapajós National Forest (TNF), Pará and (b) in the Feliz Natal (FN) municipality, Mato Grosso. The red polygons show the calibration area (40 km^2) covered with the airborne LiDAR survey. The background map is a Planet biannual mosaic from 2019.

maps, the stand age data considers the maximum water surface extent (Pekel et al., 2016) and masks out false positive pixels in wetland areas. The stand age values range from 0 to 33, where age class 1 refers to pixels that have been classified as forest cover in 2018 after an anthropic cover (e.g., pasture, agriculture, mining, etc.) in the previous year (Silva Junior et al., 2020). Other age class values show how many consecutive years a pixel stayed as forest cover after an anthropic cover. However, age class 0 refers to pixels with non-secondary forest cover such as old-growth forest, pasture, agriculture, mining, urban, water, etc.

To reduce the influence of satellite LiDAR geolocation errors, we removed secondary-forest pixels (stand age >0) that had at least one non-secondary pixel (stand age = 0) in their 3×3 neighborhood. Therefore, the resulting stand age map did not include those secondary-forest border pixels.

2.4. GEDI data

The GEDI Level 2A, version 2, data granules with footprint-level elevation and relative height (RH) metrics (Dubayah et al., 2021a) were downloaded for the period between April 18 – October 18, 2019, covering the first six months of GEDI data acquisition. These input data contained approximately 10.9 million geolocated footprints in our study area. These footprints, i.e., GEDI shots, are spatially organized along laser tracks corresponding to ascending and descending orbits of the International Space Station where the GEDI sensor is mounted. About each orbital ground track, the system samples waveforms every 60 m, along eight tracks separated by 600 m across-track distance. Four tracks are sampled with the full power laser beams, and four tracks are sampled with the coverage laser beams. The coverage beams use approximately half the energy of the power beams and thus, have a lower signal-to-noise ratio (SNR). The GEDI laser transmits 14 ns (4.2 m) long pulses of 1064 nm laser energy and records the waveform at a 1 ns rate (approximately 0.15 m). The mean 1-sigma horizontal geolocation error of GEDI footprints is currently reported to be 10.2 m with release 1, but the geolocation error is expected to be reduced in future releases (Beck et al., 2020).

For each GEDI shot, we selected latitude, longitude, and the RH98 metric. RH98 is the relative height between the 98th percentile of the returned laser energy and the peak of the ground return (Dubayah et al., 2020). RH98 is selected because it is a robust metric -used to estimate aboveground biomass by the GEDI science team (Dubayah et al., 2021b). Several additional acquisition parameters were selected to facilitate data filtering and stratification for the analysis. The acquisition parameters include beam sensitivity, quality/degradation flags, and selected algorithm. The beam sensitivity is a quality parameter of the derived RH metrics and ranges from 0 to 1, indicating the maximum canopy cover that can be penetrated considering the SNR of the waveform. This parameter was used to filter out GEDI shots with lower quality. The quality flag indicates GEDI land surface shots that meet several quality criteria, among which the minimum beam sensitivity measure of 0.90. The degradation flag is set for periods with low geolocation performance. The quality and degradation flags are used to remove low-quality GEDI shots from this analysis. The selected algorithm identifies one of seven possible algorithm setting groups selected automatically for each shot based on its geolocation and waveform complexity and then used for waveform processing. As selecting an algorithm setting group will directly affect the footprint-level metrics (including RH98 and beam sensitivity), the selected algorithm was used to stratify GEDI data and analyze how different processing strategies affect our results. Therefore, the RH98 values corresponding to the automatically-selected algorithm setting group will be referred to as the default RH98 values.

2.5. ICESat-2 data

The ICESat-2 ATL08, version 3, data granules with along-track terrain and forest heights (Neuenschwander et al., 2020b) were downloaded for the same six-month period as for the GEDI data. These input data contained approximately 6.7 million geolocated segments for our study area. Each segment represents 100 m of along-track photons used to derive terrain height and relative height canopy metrics. Like GEDI, the ATL08 segments are organized along a reference ground track corresponding to ascending and descending orbits of the ICESat-2 satellite. Per each reference ground track, ATL08 segments are written every 100 m along six tracks, three sampled with strong power beams and three with weak power beams. Each strong beam is paired with a weak beam forming three pairs in total, with 90 m and 3.3 km across-track within-in-pair and between-pair distance. The weak beams use one quarter the energy of the strong beams and thus have lower SNR. The ICESat-2 satellite utilizes the Advanced Topographic Laser Altimeter System (ATLAS), which is a green laser (532 nm wavelength) that transmits narrow pulses of 1.5 ns (approximately 0.4 m) length 10,000 times per second, which results in the along-track sampling of 0.7 m (Neumann et al., 2019). The footprint diameter for each ATLAS laser is estimated to be approximately 11 m on the ground (Magruder et al., 2020). Generally, it is expected to receive one to two signal photons per pulse interacting with land surfaces with the photon geolocation and vertical accuracy better than 5 m and 1 m, respectively (Neuenschwander and Magruder 2019).

For each ATL08 segment, we selected latitude, longitude, and the RH98 metric. The latter is the 98th relative height percentile of all the canopy photons in the corresponding segment, where the cumulative relative height is the difference of canopy photon height and the estimated terrain height across the 100 m segment. Like GEDI, we additionally utilized the following acquisition parameters: the ATLAS beam number and the night flag. The ATLAS beam number was used to stratify the data to the strong and weak beams, whereas the night flag was used to stratify the data to the day- and night-time acquisitions.

2.6. Airborne LiDAR data

Airborne LiDAR data were used to calibrate and assess the quality of GEDI and ICESat-2 forest heights at the TNF and FN calibration sites (Section 2.2). At each calibration site, two parallel, 50 km in length, airborne LiDAR strips with a swath width of 200 m were available. The data at the FN calibration site were situated in a north-south direction and were 40 km apart from one another. The data at the TNF calibration site were situated in an approximately north-south direction with a (clockwise-defined) azimuth of $-15°$ and were 11 km apart from one another. The lidar acquisition direction followed the border direction of TNF such that one TNF strip was placed inside the intact forest and another outside in the disturbed forest. The FN strips mainly covered the intact forest.

All four airborne LiDAR strips were acquired in early October 2018, using the Optech® ALTM 3100 Airborne Laser Terrain Mapper system. The surveys were performed through the Sustainable Landscapes project commissioned by the United States Forest Service in collaboration with the Brazilian Enterprise for Agricultural Research, and the data were archived through the Carbon Monitoring System project funded by NASA (Dos-Santos et al., 2019). The acquired data were provided as

georeferenced, ground-classified, and noise-filtered point clouds with an average density of more than 20 returns per m^2 and up to four returns recorded per laser pulse. For each return, its relative height to the ground was calculated using the software package Lastools (Isenburg 2019). The relative heights were then aggregated into 1 m gridded raster images using the 99th relative height percentile of all returns within the raster cell. This resulted in four 1 m rasters with airborne LiDAR forest heights that were then used for the comparison and calibration of GEDI and ICESat-2 forest heights.

3. Methods

Several steps were introduced to assess the forest heights and regrowth rates from GEDI and ICESat-2 data in the Amazon rainforest (Fig. 3). First, GEDI shots and ICESat-2 ATL08 segments were pre-processed such that any data outside secondary forest or covered low-quality forest heights were excluded from the analysis (Section 3.1). Next, the remaining GEDI shots and ICESat-2 segments were stratified into subgroups according to their acquisition parameters (Section 3.2). Next, for each GEDI and ICESat-2 subgroup, we analyzed the relation between satellite- and airborne-LiDAR forest heights at the calibration sites and derived corresponding calibration models (Section 3.3). Calibrated GEDI and ICESat-2 forest heights were used in subsequent calculations and comparisons. After the calibration step, we derived the distributions of median-aggregated, single-year GEDI and ICSat-2 heights per stand age class for the Rondônia state (Section 3.4). Due to an observed saturation of forest heights with stand age, we derived regrowth rates from the first two decades of data by applying a linear growth model (Section 3.5). At this stage, we analyzed regrowth retrieval uncertainty from 19 GEDI and ICESat-2 data subgroups. In the final step, when the most appropriate GEDI and ICESat-2 data subgroups were identified (Section 3.6), we investigated different non-linear models to approximate the regrowth over the entire 33-year period (Section 3.7).

Code, input, and intermediate datasets necessary to reproduce our results are made freely available on GitHub and Zenodo.

3.1. Pre-processing of GEDI and ICESat-2 data

Before parsing the data into different subgroups, two filters were applied to exclude non-secondary-forest and low-quality GEDI shots and ICESat-2 ATL08 segments. First, the secondary-forest GEDI shots and ICESat-2 ATL08 segments were selected based on their spatial over-lapping with the secondary forest pixels (stand age >0) in the Rondônia state. Second, the low-quality filter excluded erroneously large or small GEDI shots and ICESat-2 ATL08 segments, i.e., only the shots and segments with RH98 values larger than zero and smaller than 75 m were preserved for further analysis. These two datasets are referred to as GEDI ALL dataset and ICESat-2 ALL dataset (Table 1).

3.2. Stratification of GEDI and ICESat-2 data

GEDI shots were first stratified into four subgroups (Table 1). The first group contained GEDI shots that fulfilled additional quality requirements, i.e., all shots with compromised quality requirements (the quality flag = 0) or geolocation requirements (the degradation flag ≥0) were excluded. The GEDI shots of the first subgroup were further stratified into three additional groups applying beam sensitivity

Table 1
Overview of GEDI datasets used to understand the impact of different acquisition parameters on the forest heights and regrowth rates. The Parent Dataset is a dataset on which the filter in the Filter Applied column is applied.

Subgroup	Filters Applied	Parent Dataset
ALL	Stand Age Map and Erroneous Forest Height	–
QS90	Quality and Degradation flags	ALL
QS95	Beam Sensitivity ≥0.95	QS90
QS98	Beam Sensitivity ≥0.98	QS90
QS99	Beam Sensitivity ≥0.99	QS90
ALL-S2	Sensitivity derived with Algorithm Setting Group 2	ALL
QS90-S2	Quality and Degradation flags	ALL-S2
QS95-S2	Beam Sensitivity ≥0.95	QS90-S2
QS98-S2	Beam Sensitivity ≥0.98	QS90-S2
QS99-S2	Beam Sensitivity ≥0.99	QS90-S2
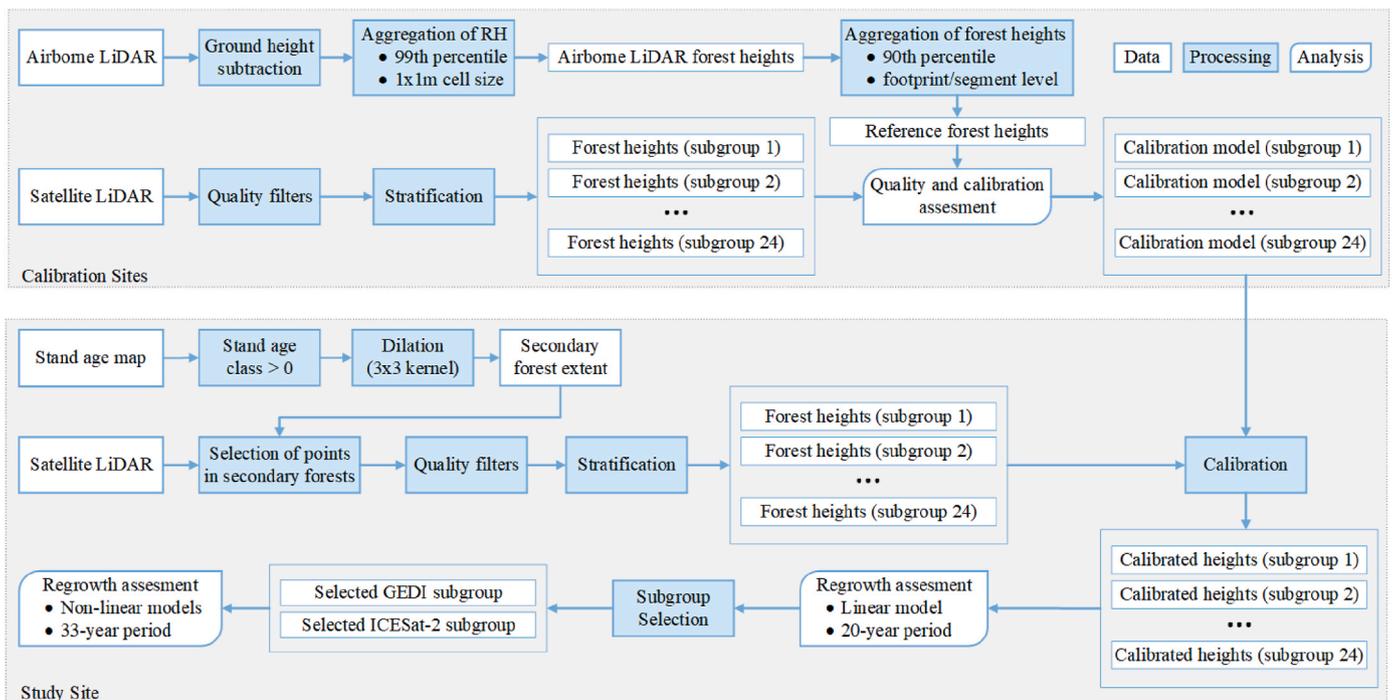
Fig. 3. An overview of the input data, processing, and analysis performed at the calibration and study sites. The term "Satellite LiDAR" refers to GEDI and ICESat-2 data, "RH" to relative heights, and "points" to GEDI shots or ICESat-2 ATL08 segments.

threshold values of 0.95, 0.98, and 0.99. As the quality flag includes a requirement on the sensitivity value, all the selected GEDI shots in those four subgroups had a beam sensitivity greater than 0.90. It should be noted that the data in the above GEDI subgroups were based on the automatically-selected algorithm setting group, i.e., subgroups contained GEDI variables with the default values.

The above GEDI subgroups were further parsed into five GEDI subgroups (Table 1) based on the algorithm selection group properties. The new subgroups were formed by keeping the default RH98 values but using the sensitivity derived with the algorithm setting group 2 (groups with "S2" as a suffix, Table 1). GEDI shots with the automatically-selected algorithm setting group 2 dominated (>70% of shots) for our study and calibration sites, whereas the remaining shots were with the automatically-selected algorithm setting group 1. That group has a higher waveform-end threshold, providing more lower forest heights than the algorithm setting group 2. Therefore, the S2 subgroups were introduced to ensure more balanced forest heights with the default RH98 values and consistent sensitivity values from the algorithm setting group 2.

The ICESat-2 ATL08 segments were stratified into eight groups (Table 2) based on their ATLAS beam number and the night flag parameter. Like GEDI, the ATLAS beam number and the night flag were used to identify points associated with either strong or weak beams and day versus night acquisition, respectively. Those categories were used to define eight ICESat-2 subgroups.

3.3. Calibration and quality assessment of GEDI and ICESat-2 forest heights

The calibration and quality of GEDI and ICESat-2 forest heights were assessed using the 1 m rasters with airborne LiDAR forest heights at the TNF and FN calibration sites. For each GEDI shot that falls inside one of the four airborne LiDAR strips, the reference forest height (H_{Ref}^{GEDI}) was calculated as the 90th percentile height of all the raster cells inside a circular buffer with 25 m diameter around the GEDI shot. This buffer diameter was selected to match the GEDI footprint diameter, and the 90th percentile height was already used in previous studies to assess the quality of GEDI forest heights and aggregate airborne LiDAR forest heights at the Landsat pixel level (Potapov et al. 2019, 2021). The 90th percentile height was also calculated for each ICESat-2 ATL08 segment ($H_{Ref}^{ICESat2}$), but using the raster cells inside an 11 m × 100 m rectangle centered at the ICESat-2 ATL08 segment center and oriented in the along-track direction with its longer side. The 100 m side was selected to match 100 m ATL08 segments, and the 11 m side was selected to match the estimated ICESat-2 footprint diameter (Magruder et al., 2020; Neuenschwander et al., 2020a). It should be noted that a geolocation error for either GEDI and ICESat-2 tracks was neither estimated nor corrected yet is potentially a significant source of uncertainty (Roy et al., 2021). This step was omitted to assess forest height error in the presence of geolocation error since such errors propagate through to the retrieval of the regrowth rates.

Table 2
Overview of ICESat-2 datasets used to understand the impact of different acquisition parameters on the forest heights and regrowth rates. The Parent Dataset is a dataset on which the filter in the Filter Applied column is applied.

Name	Filters Applied	Parent Dataset
ALL	Stand Age Map and Erroneous Forest Height	–
SN	Strong Beams and Nightlight Conditions	ALL
SD	Strong Beams and Daylight Conditions	ALL
WN	Weak Beams and Nightlight Conditions	ALL
WD	Weak Beams and Daylight Conditions	ALL
SND	Strong Beams and Nightlight and Daylight Conditions	ALL
WND	Weak Beams and Nightlight and Daylight Conditions	ALL
SWN	Strong and Weak Beams and Nightlight Conditions	ALL
SWD	Strong and Weak Beams and Daylight Conditions	ALL

Several quality assessment measures were calculated from the forest height differences ΔH_{GEDI} and $\Delta H_{ICESat2}$ ($\Delta H_s = H_s - H_{Ref}^S$, where s is used as a replacement for the sensor used, i.e., GEDI or ICESat-2). First, we calculated the mean error-as:

$$MeanE = \frac{\sum_1^{n_s} \Delta H_s}{n_s}$$

where n_s can be either the number of GEDI shots (n_{GEDI}) or the number of ICESat-2 ATL08 segments ($n_{ICESat2}$). Then, we calculated the maximum absolute error (MAE):

$$MAE = \max\{|\Delta H_s|\}$$

and the root mean squared error (RMSE):

$$RMSE = \sqrt{\frac{\sum_1^{n_s} \Delta H_s^2}{n_s}}$$

Finally, we calculated the mean of the reference (airborne LiDAR) forest heights ($\overline{H}_{Ref}^S$):

$$\overline{H}_{Ref}^S = \frac{\sum_1^{n_s} H_{Ref}^S}{n_s}$$

that was used to obtain the relative RMSE, i.e., CV(RMSE), by dividing RMSE by $\overline{H}_{Ref}^S$ and multiplying with 100 to express it in percentages.

To derive calibration models for satellite forest heights, we performed a regression analysis between the GEDI and ICESat-2 forest heights (H_{GEDI} and $H_{ICESat2}$) and their corresponding reference, i.e., airborne LiDAR, forest heights (H_{Ref}^{GEDI} and $H_{Ref}^{ICESat2}$). We calculated the coefficient of determination (R^2), slope (a), and intercept (b) of the linear calibration model to understand better the relation between satellite LiDAR and airborne LiDAR forest heights. The above calibration parameters were derived separately for each GEDI and ICESat-2 subgroup from Tables 1 and 2, respectively. Finally, a calibrated satellite height ($H_{s,g}'$) of a particular GEDI or ICESat-2 subgroup g was calculated as:

$$H_{s,g}' = \frac{H_{s,g} - b_g}{a_g}$$

where s is used as a replacement for the sensor used, i.e., GEDI or ICESat-2, and a_g and b_g are the calibration parameters for a GEDI or ICESat-2 subgroup g in Table 1 or Table 2, respectively. The calibrated satellite heights were then used in the subsequent analysis.

3.4. Aggregation of forest height per stand age

GEDI shots and ICESat-2 points were grouped according to their stand age class to derive the distribution of their forest heights across the 33 years of stand age. For each stand age *i*, the median GEDI and ICESat-2 forest height ($H_{Med,i}^S$) was calculated considering only GEDI shots and ICESat-2 points within this stand age class, respectively. As our stand age map covers the 1986–2018 period, the resulting distributions of GEDI and ICESat-2 forest heights spanned 33 years of regrowth ($i = 1, \ldots, 33$).

Initially, we calculated the distributions from the pre-processed GEDI and ICESat-2 data, i.e., from the two ALL subgroups in Tables 1 and 2, respectively. As the obtained distributions clearly showed a saturation of GEDI and ICESat-2 median forest height for stand age larger than 20 years, we decided to select the first 20 years of stand age (regrowth started between 1999 and 2018) to analyze the influence of the acquisition parameters on the retrieval of regrowth rates. However, different non-linear models were analyzed to approximate the entire 33-year regrowth period in a latter step.

3.5. Calculation of the regrowth rates

Regrowth rates were derived using a regression analysis of the median GEDI and ICESat-2 forest heights from the first 20 years of regrowth. The relationship between each response variable, i.e., median GEDI and ICESat forest height, was separately modeled using a linear regression model with stand age as the explanatory variable:

$$y^s = X\beta^s + \varepsilon^s$$

where s is used as a replacement for the sensor used, i.e., GEDI or ICESat-2, y^s is the vector of n observed median forest heights $H^S_{Med, i}$ where i is stand age ($i = 1, …, n$ and $n = 20$), X is the n x 2 design matrix, consisting of a n-dimensional vector with ones and a n-dimensional vector with stand age values i, β^s is the two-dimensional parameter vector with the regression intercept β^s_0 and the regression slope β^s_1 as its elements, and finally, ε^s is a n-dimensional vector of ε_i values, i.e., the error term. The parameter vector was estimated using the ordinary least square framework:

$$\widehat{\beta}^s = arg \min_\beta \varepsilon^2 = (X^T X)^{-1} X^T y^s$$

where ε is the vector norm of the error term. The standard error of the regression σ^s was calculated as:

$$\sigma^s = \sqrt{\frac{\widehat{\varepsilon}^T \widehat{\varepsilon}}{n-2}}$$

where $\widehat{\varepsilon}$ is the residual vector ($\widehat{\varepsilon} = y^s - \widehat{y}^s = y^s - X\widehat{\beta}^s$), and the standard errors (SEs) of the estimated parameters ($\sigma^s_{\widehat{\beta}_1}$ and $\sigma^s_{\widehat{\beta}_0}$) as:

$$\sigma^s_{\widehat{\beta}_0} = \sigma^s \sqrt{C_{11}} \quad \text{and} \quad \sigma^s_{\widehat{\beta}_1} = \sigma^s \sqrt{C_{22}}$$

where C_{11} and C_{22} are the diagonal elements of the cofactor matrix $(X^T X)^{-1}$. Finally, we calculated the coefficient of determination (R^2) as:

$$R^2 = \frac{\sum_{i=1}^{n}(\widehat{y}_i - \overline{y})}{\sum_{i=1}^{n}(y_i - \overline{y})}$$

where $\overline{y}$ is the mean of our response variable H^S_{Med}.

The regression slope $\widehat{\beta}^S_1$ and intercept $\widehat{\beta}^S_0$, their standard errors ($\sigma^s_{\widehat{\beta}_1}$ and $\sigma^s_{\widehat{\beta}_0}$) as well as the coefficient of determination R^2 were calculated for each GEDI and ICESat-2 subgroup from Tables 1 and 2. The regression slope $\widehat{\beta}^S_1$ was selected as an estimate of the average regrowth rate within the first 20 years of forest regrowth. The standard error $\sigma^s_{\widehat{\beta}_1}$ was selected as an estimate of the uncertainty associated with $\widehat{\beta}^S_1$. Finally, for each GEDI and ICESat-2 subgroup, we calculated another measure, the internal uncertainty, which is the absolute difference between their $\widehat{\beta}^S_1$ and $\widehat{\beta}^S_1$ from the best performing subgroup (one with the highest R^2).

3.6. Selection criteria for regrowth rates

Before comparing GEDI and ICESat-2 regrowth rates, we introduced several criteria to select the most appropriate GEDI and ICESat-2 subgroups for the comparison, i.e., those subgroups that lead to reliable regrowth measurements. The most appropriate subgroups had to fulfill the following six criteria:

I. **Erroneous Forest Heights**: Subgroups with no erroneous heights were considered to lead to more reliable calibration models and regrowth estimates.
II. **Quality of Forest Heights**: Subgroups with a smaller RMSE to the airborne LiDAR forest heights were considered to lead to more reliable calibration models and regrowth estimates.
III. **Sample Size**: Subgroups that lead to less reduction in sample size were considered more reliable, provided they had equally accurate forest heights.
IV. **R^2 of Regrowth Rate**: Subgroups with a higher R^2 were considered more reliable.
V. **Distribution of Forest Heights**: Subgroups with a higher R^2 and a height distribution different from the distribution of the best performing subgroup (one with the highest R^2) were considered less reliable.
VI. **Internal Uncertainty**: Subgroups with the internal uncertainty (regrowth and intercept differences) larger than their regrowth and intercept standard errors were considered less reliable. The regrowth and intercept differences were calculated relative to the regrowth rate and the intercept of the best performing subgroup, i.e., one with the highest R^2.

3.7. Non-linear regrowth models

Once the most appropriate GEDI and ICESat-2 subgroups were identified, we analyzed different non-linear models to best approximate the regrowth over the entire 33-year period. Three non-linear models were considered: spherical (SPH), exponential (EXP), and logarithmic (LOG). Each of these three models has a continuously increasing phase and a saturation phase, which was the behavior of our median GEDI and ICESat-2 forest heights over the 33 years. Furthermore, the LOG model has been used for approximating field data of Amazon rainforest growth (Freitas et al., 2019). Although field data typically consist of a much lower sampling size, spatial or temporal extent than satellite data, the LOG function is important to relate remote-sensing and field-data assessed regrowth.

The SPH model was defined as:

$$y(x) = \begin{cases} a + b\cdot\left(\frac{3}{2}\left(\frac{x}{c}\right) - \frac{1}{2}\cdot\left(\frac{x}{c}\right)^3 \right) & if \ x \leq c \\ a + b & if \ x > c \end{cases}$$

where y was the median GEDI or ICESat-2 forest height (H^S_{Med}) and x was the stand age. The coefficient a is the intercept, the coefficient b is the forest regrowth between the intercept and the saturation, and the coefficient c is the saturation year. A total height growth (THG) and an average regrowth rate (ARR) for the spherical model are then calculated as: $THG_{SPH} = a + b$ and $ARR_{SPH} = (a + b)/c$. The EXP model was defined as:

$$y(x) = a + b\cdot\left(1 - e^{-\frac{x}{c}} \right)$$

where the coefficients a and b have the same interpretation as their counterparts in the SPH model, whereas the saturation year is 3 c for the EXP model (Bivand et al., 2013, p.202). THG and ARR for the exponential model are then calculated as: $THG_{EXP} = a + b$ and $ARR_{EXP} = (a + b)/3c$. The LOG model was defined as:

$$y(x) = a + b\cdot\ln(x)$$

where ln is the natural logarithm, the coefficient a is the intercept, and the coefficient b controls the height stretch of the LOG model. It is noted

that the LOG model defines different forest height increments of b, $2b$, and $3b$, and consequently different regrowth rates of b/e, $2b/e^2$, and $3b/e^3$ for the first e (~ 2.7), e^2 (~ 7.4), and e^3 (~ 20.1) years of regrowth, respectively. THG and ARR for the logarithmic model are here calculated as: $THG_{LOG} = a + b \cdot \ln(33)$ and. $ARR_{LOG} = (a + 3b)/e^3$

The coefficients of the SPH, EXP, and LOG models were estimated similarly as in Section 3.5. For the SPH and EXP models, a $n \times 3$ design matrix ($n = 33$) with column elements $X_{i1} = dy(x_i)/da$, $X_{i2} = dy(x_i)/db$, and $X_{i3} = dy(x_i)/dc$, where ($i = 1, ..., n$), and the three-dimensional parameter vector $\beta = [a\ b\ c]^T$ were used. For the LOG model, the design matrix had only $X_{i1} = dy(x_i)/da$, and $X_{i2} = dy(x_i)/db$ columns as this model had only two unknown parameters. Finally, the model with the highest R^2 was selected as the most appropriate to approximate the entire 33-year regrowth period.

4. Results

4.1. Quality of GEDI forest heights and GEDI calibration models

Regression analysis between airborne LiDAR forest heights (ALFHs) and GEDI RH98 values showed that the coefficient of determination (R^2) constantly increased when GEDI shots with higher sensitivity were analyzed (Fig. 4). We also observed several GEDI shots with small (<8 m) RH98 values but with ALFHs larger than 12 m (blue rectangles in Fig. 4a and b). These GEDI shots were gradually filtered out with increasing the sensitivity threshold and were totally excluded when the sensitivity of 0.99 was applied (Fig. 4a–d).

The GEDI shots with RH98 < 8 m and ALFH >12 m were further examined with the Planet bi-annual image mosaic from the first half of 2020. Most of these GEDI shots fall within a former forest plot (the green polygon) deforested between airborne LiDAR and GEDI acquisitions (Fig. 5a), and thus those shots were excluded from the following quality assessment. We further excluded a few of the above GEDI shots that were found close to the forest border (Fig. 5b) because such border shots were

also excluded from our regrowth assessment by removing the border pixels from the forest age map. Finally, a few of the above GEDI shots were found inside the intact forest (Fig. 5c). Their erroneously small RH98 values may be possibly attributed to a combination of geolocation error and other causes such as waveform processing error or low sensitivity. Nevertheless, those shots were considered in the quality assessment since such shots may also be present in the GEDI subgroups with low sensitivity thresholds.

The beam sensitivity also affected all the quality assessment metrics (Table 3). For increasing sensitivities, RMSE and MAE values constantly decreased, whereas R^2 value constantly increased. The QS99-S2 was the subgroup with the lowest RMSE and the highest R^2 but its 0.99 sensitivity threshold left only 38% of GEDI shots compared to the QS98-S2 subgroup that had similar R^2. Furthermore, the QS98-S2 subgroup had the lowest mean error (MeanE) among all subgroups and a slightly larger RMSE than QS99-S2.

The results also showed that QS99-S2 did not have representation across the full range of heights (Fig. 6d). There are just a few QS99-S2 shots with RH98 and ALFH smaller than 5 m, and not a single shot in the 5–15 m height range. This was not the case for QS95-S2 and QS98-S2, whose shots covered the full range of heights (Fig. 6b and c). Furthermore, QS95-S2 and QS98-S2 were now free from the shots with erroneous RH98 heights (RH98 < 8 m and ALFH >12 m) located inside the deforested plot and close to the forest border.

4.2. Quality of ICESat-2 forest heights and ICESat-2 calibration models

The ICESat-2 subgroup with strong beams and day acquisitions (SD) had the highest R^2 (Fig. 7). There were a certain number of ICESat-2 ATL08 segments with RH98 values larger than 50 m and airborne LiDAR heights below 50 m (blue rectangles in Fig. 7a–c). Those errant ICESat-2 ATL08 segments were also present in the SND and SN subgroups (Fig. 7b and c) but not in the SD subgroup (Fig. 7d). This means that the errant segments were coming from a night acquisition.
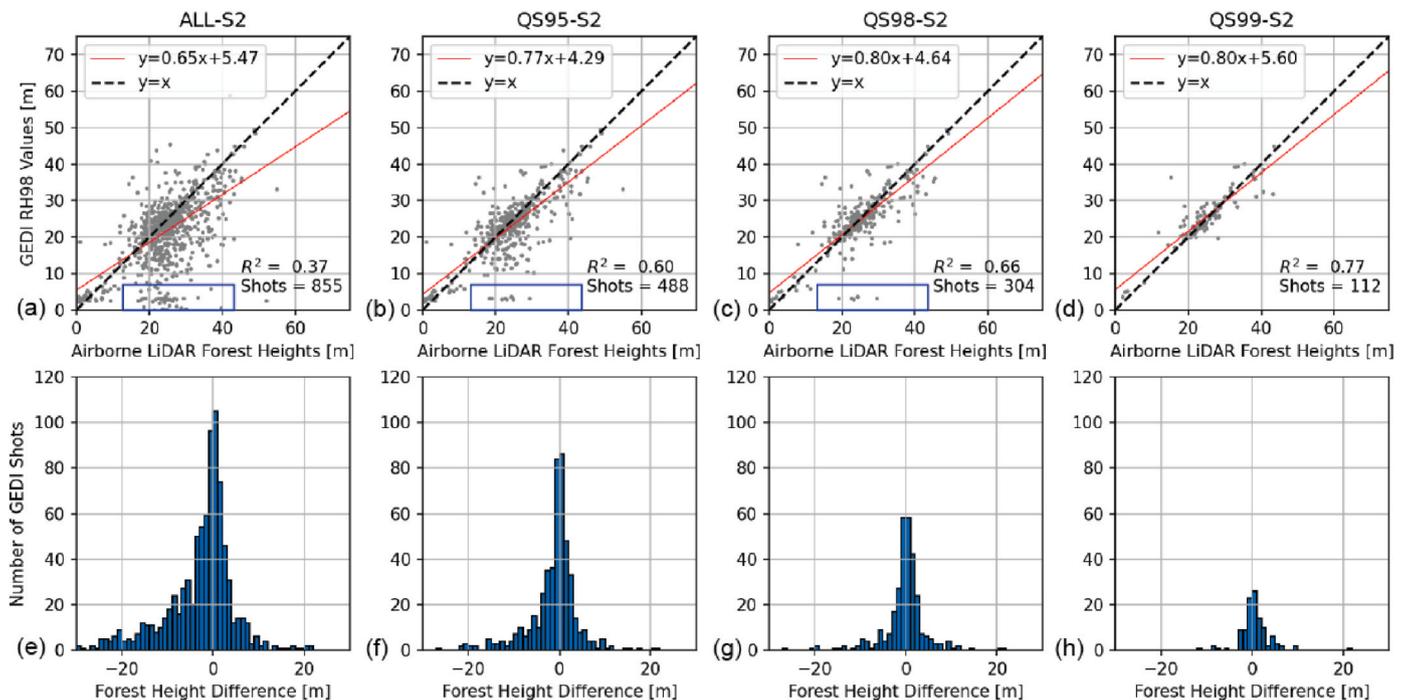
Fig. 4. The relation between GEDI RH98 (H_{GEDI}) and airborne LiDAR forest heights (H_{Ref}^{GEDI}) for different beam sensitivity thresholds, and the distributions of the corresponding forest height differences ($\Delta H_{GEDI} = H_{GEDI} - H_{Ref}^{GEDI}$). The blue rectangles in (a), (b), and (c) delineate GEDI shots with RH98 values smaller than 8 m and airborne LiDAR forest heights (ALFHs) larger than 12 m. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
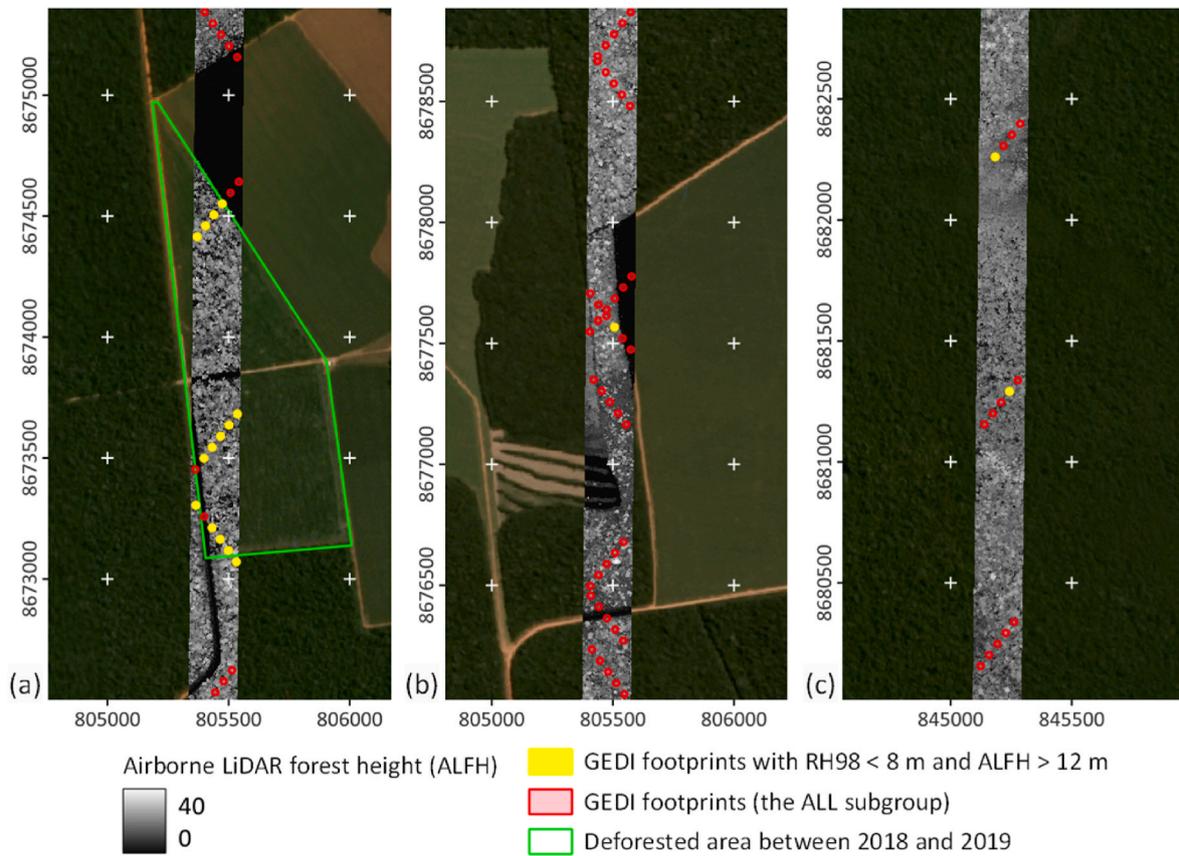
Fig. 5. Examples of GEDI footprints with erroneous RH98 heights caused by (a) a deforestation event between airborne LiDAR and GEDI acquisitions, or (b) geolocation error, or (c) a low sensitivity value (0.89 and 0.92). Planet Biannual mosaic (Dec.2019–June 2020) is used as the basemap. The coordinates are given in the UTM Zone 21S coordinate reference system.

Table 3
Quality assessment metrics for GEDI subgroups with different beam sensitivity thresholds and algorithm processing strategies.

Subgroup Name	Number of Shots	Number of Orbits	MAE [m]	$\overline{H}_{Ref}$ [m]	MeanE [m]	RMSE [m]	Slope [−]	Int. [m]	R^2
ALL	841	36	49.44	24.14	−2.62	7.93	0.65	5.85	0.38
QS90	496	22	30.99	23.94	−1.14	5.40	0.74	5.07	0.61
QS95	443	22	21.87	24.47	−0.91	4.91	0.75	5.25	0.61
QS98	240	20	21.12	24.90	−0.07	4.25	0.74	6.47	0.67
QS99	88	16	21.12	25.58	0.53	3.83	0.73	7.55	0.68
ALL-S2	841	36	49.44	24.14	−2.62	7.93	0.65	5.85	0.38
QS90-S2	496	22	30.99	23.94	−1.14	5.40	0.74	5.07	0.61
QS95-S2	478	22	21.87	23.91	−0.85	4.97	0.76	4.77	0.66
QS98-S2	298	20	21.12	24.64	0.04	4.14	0.80	5.08	0.74
QS99-S2	112	18	21.12	24.68	0.66	3.61	0.80	5.60	0.77

The errant ICESat-2 ATL08 segments were further examined by looking at the relation between the SN segments with their orbit number (Fig. 8b) and the number of canopy photons per segment (Fig. 8c). All the errant SN segments were coming from a single orbit, but not all the SN segments from that orbit were errant (Fig. 8b). However, the errant SN segments were associated with a large number of canopy photons (Fig. 8c). The remaining SN segments had a lower number of canopy photons. When a simple threshold of 140 or less canopy photons per segment was applied on the SN segments, almost all the erroneous ICESat-2 heights were filtered out (Fig. 8d). Therefore, the above canopy photon threshold was applied in all the calculations with the strong-beam night-acquisition ICESat-2 ATL08 segments.

After the canopy photon filter, the SD subgroup again had the lowest RMSE, but the ALL subgroup had the highest R^2 (Table 4). The remaining segments in the SN and SND subgroups had a slightly higher RMSE but included roughly 1.7 and 2.7 times more points than the SD

subgroup, respectively. The preprocessing of ICESat-2 ATL08 segments (the stand age and erroneous heights filters) left only one valid ICESat-2 weak-beam ATL08 segment for the analysis. Thus, there were not enough data to assess the performance of the weak-beam subgroups (WN, WD, WND, SWN, and SWD).

The results also showed that SN segments did not have representation for RH98 and ALFH heights below 20 m (Fig. 9c). However, this was not the case for ALL, SND, and SD, whose segments covered the full height range, though SD with only 27 segments in total (Fig. 9a, b, and d).

4.3. Distribution of GEDI and ICESat-2 forest heights across stand age

Distributions of GEDI and ICESat-2 median forest heights across stand age showed similar general behavior with a growth and a "saturation" phase (Fig. 10a and b). The growth phase covered approximately
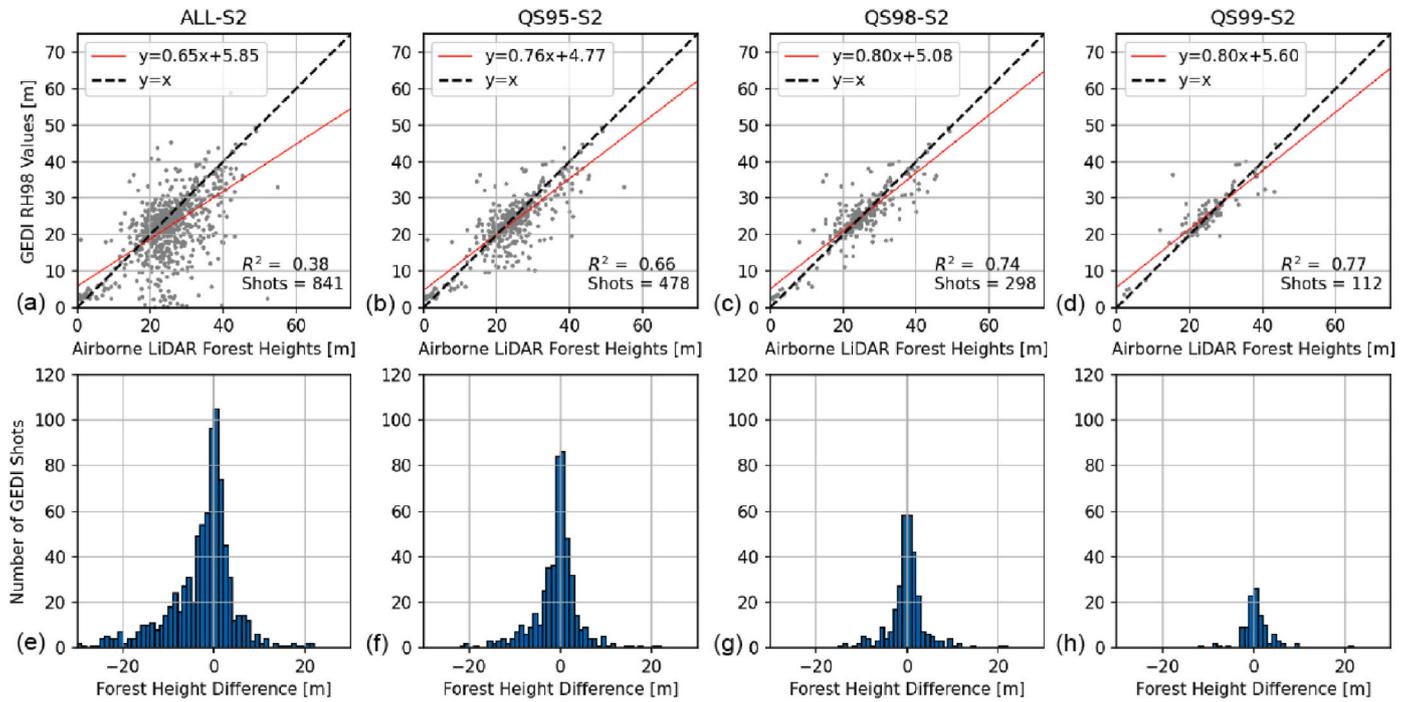
Fig. 6. The relation between GEDI RH98 (H_GEDI) and airborne LiDAR forest heights (H_{Ref}^{GEDI}) after excluding GEDI shots located inside the deforested plot and excluding GEDI shots close to the forest border.
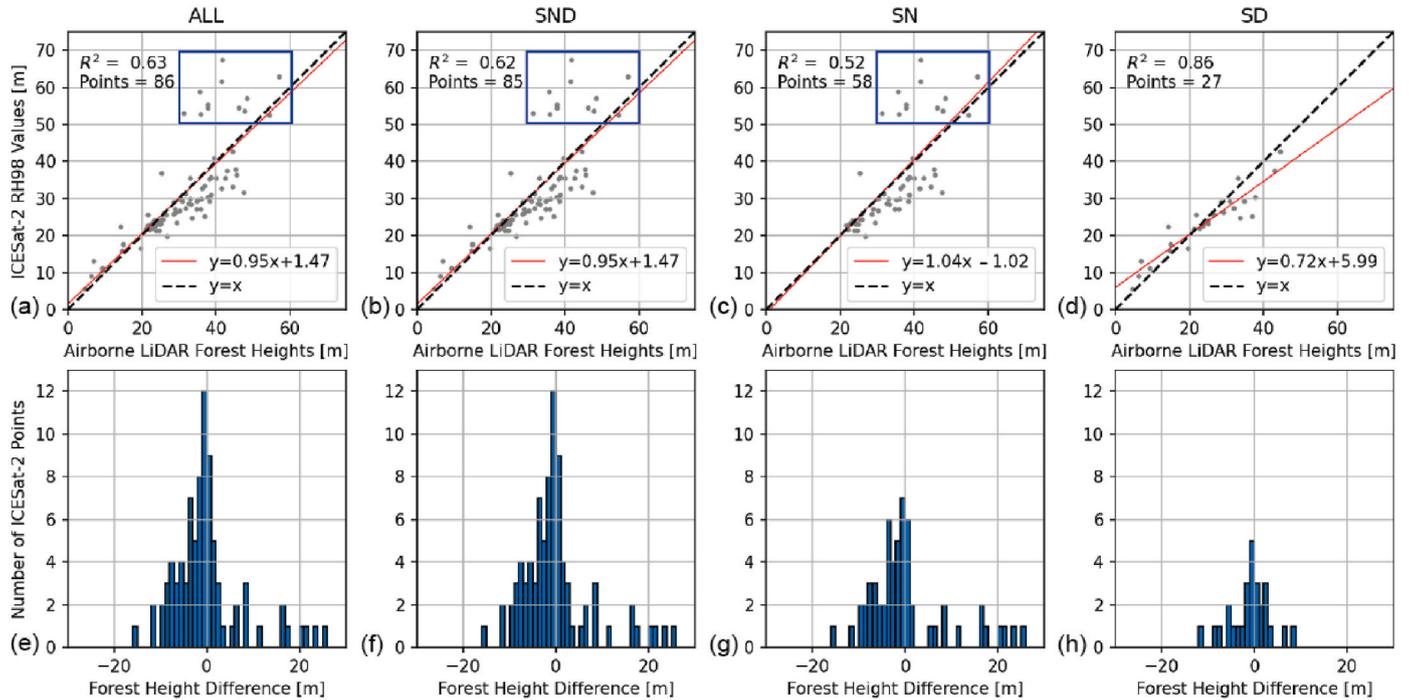
Fig. 7. The relation between ICESat-2 RH98 ($H_{ICESat2}$) and airborne LiDAR forest heights ($H_{Ref}^{ICESat2}$) for different beam strength and light conditions, and the distributions of the corresponding forest height differences ($\Delta H_{ICESat2} = H_{ICESat2} - H_{Ref}^{GEDI}$). The blue rectangles in (a), (b), and (c) delineate ICESat-2 ATL08 segments with erroneously large RH98 values. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
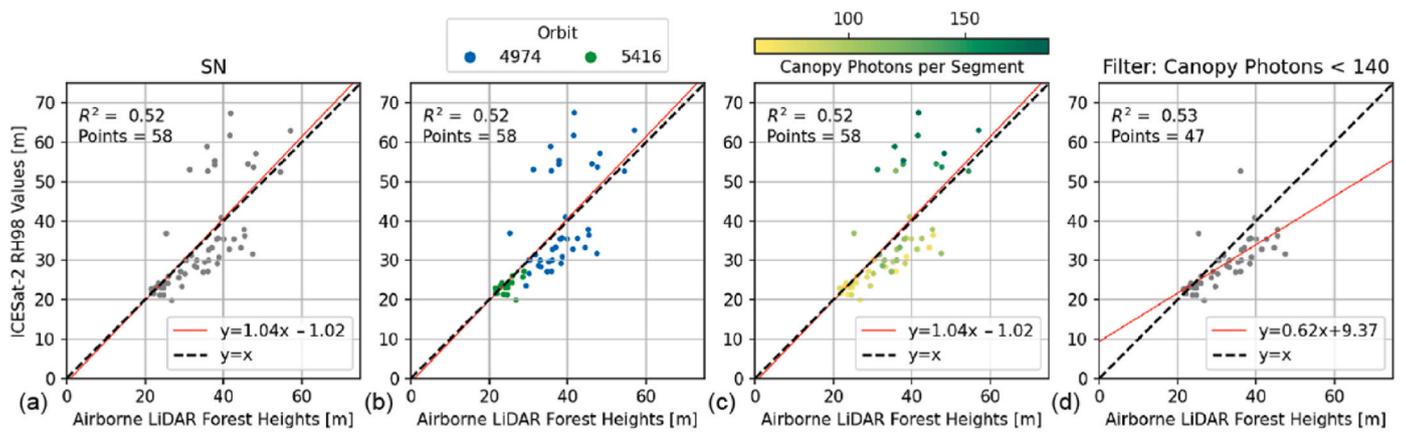
Fig. 8. The relation between the ICESat-2 SN points and their individual orbits and number of canopy photon per the corresponding ATL08 segment.

Table 4
Quality assessment metrics for ICESat-2 subgroups with different beam strength and light condition.

Subgroup Name	Number of Shots	Number of Orbits	MAE [m]	$\overline{H}_{Ref}$ [m]	MeanE [m]	RMSE [m]	Slope [−]	Int. [m]	R^2
ALL[a]	75	4	16.90	28.84	−2.10	5.38	0.69	6.87	0.74
SN[a]	47	2	16.90	31.69	−2.81	5.91	0.62	9.37	0.53
SD	27	2	11.76	24.59	−1.00	4.42	0.72	5.99	0.86
SND[a]	74	2	16.90	29.10	−2.15	5.41	0.68	7.16	0.72
WN	0	0	n/a	n/a	n/a	n/a	n/a	n/a	n/a
WD	1	1	1.03	9.39	1.03	1.03	n/a	n/a	n/a
WDN	1	1	1.03	9.39	1.03	1.03	n/a	n/a	n/a
SWN[a]	47	2	16.90	31.69	−2.81	5.91	0.62	9.37	0.53
SWD	28	2	11.76	24.05	−0.93	4.35	0.73	5.66	0.87

[a] The canopy photon filter applied, i.e., only segments with <140 canopy photons were preserved.
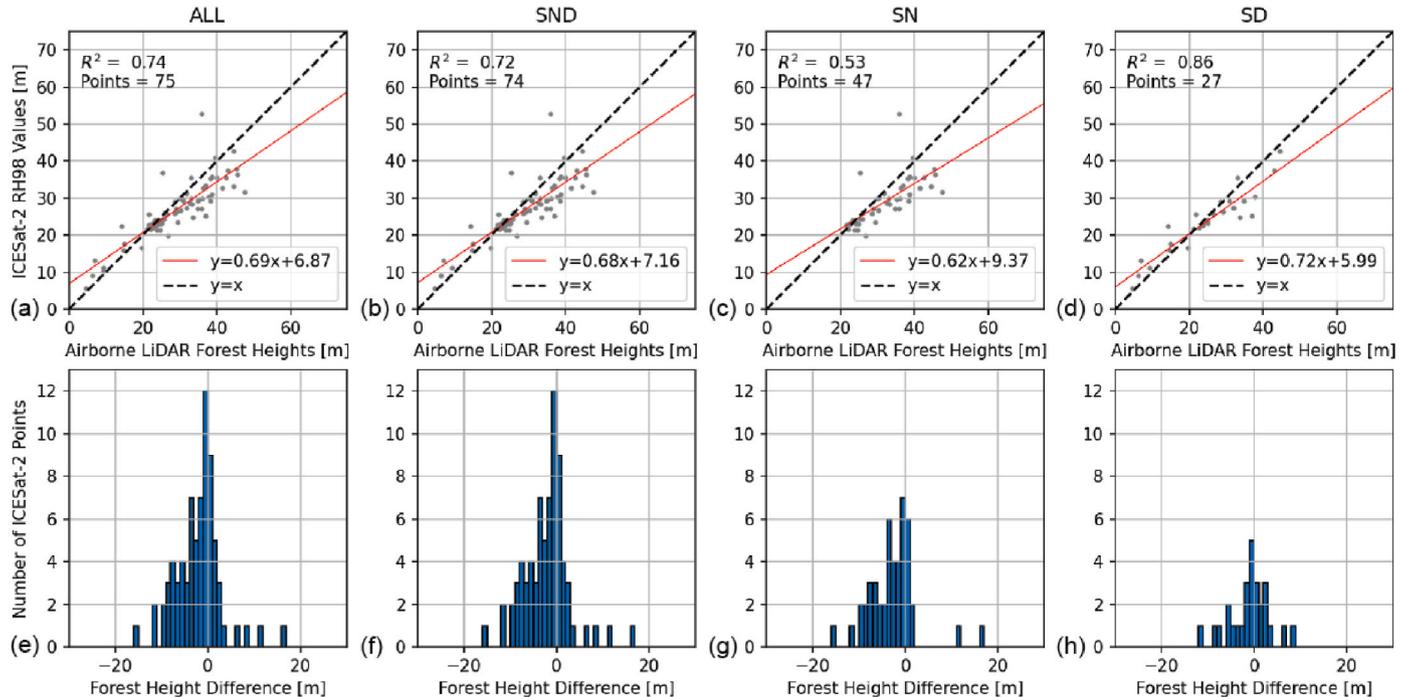
Fig. 9. The relation between ICESat-2 RH98 ($H_{ICESat2}$) and airborne LiDAR forest heights ($H_{Ref}^{ICESat2}$) after excluding segments with 140 or more canopy photons.
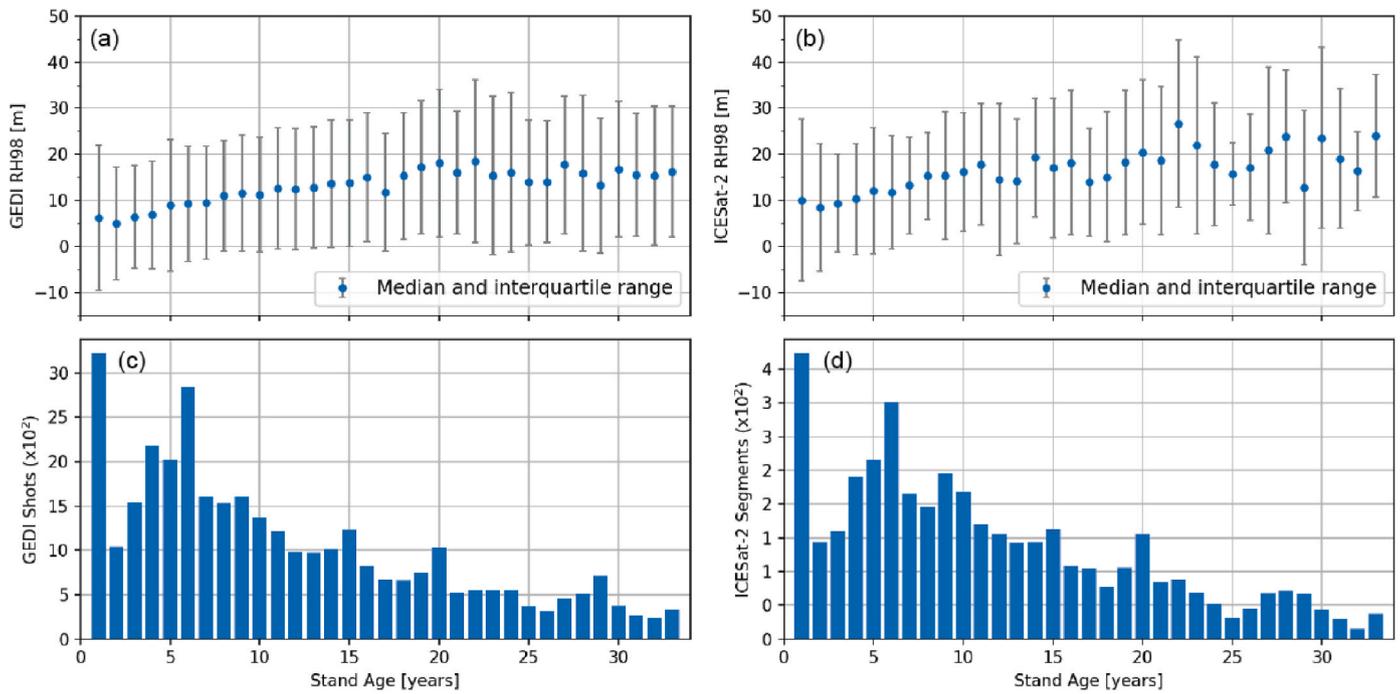
Fig. 10. Distributions of calibrated GEDI ALL forest heights (a) and GEDI shots (c) per stand age, as well as distributions of calibrated ICESat-2 ALL forest heights (b) and ICESat-2 ATL08 segments (d) per stand age. It is noted that the values on y-axis in (c) and (d) have to be multiplied by 10^2 to get the number of shots and segments per stand age class.

the first two decades of stand age and showed a quasi-liner increase in both GEDI and ICESat-2 median forest heights. The "saturation" phase covered approximately the last decade of stand age (regrowth started between 1986 and 1996) and showed that both GEDI and ICESat-2 median forest heights preserved a certain height level but with a larger height uncertainty compared to the growth phase. The distributions of GEDI shots and ICESat-2 segments were similar (Fig. 10c and d) but the number of GEDI shots was approximately an order of magnitude larger than the number of ICESat-2 segments. Furthermore, the number of GEDI shots and ICESat-2 segments in the "saturation" phase was much smaller than their numbers in the growth phase. The GEDI and ICESat-2 interquartile ranges showed that the calibration step introduced some negative heights, but the medians were positive for all age classes.

This first insight into median forest height distribution across stand age was derived from the GEDI ALL and ICESat-2 ALL subgroups. Before

modeling the median heights across all 33 years of regrowth, we first analyzed how different GEDI and ICESat-2 subgroups affect regrowth rates in the growth phase, i.e., in the first two decades of stand age. Then, together with the forest height quality assessment (Sections 4.1 and 4.2), the outcome of this analysis was used to identify the most appropriate GEDI and ICESat-2 subgroups to model and compare across the 33 years of regrowth.

4.3.1. Retrieval of GEDI regrowth rates

Linear regression of the first 20 years of forest age data showed that the S2 subgroups had $R^2 > 0.90$ and the highest R^2 among other subgroups with the same sensitivity threshold (Fig. 11a). Furthermore, for the sensitivities of 0.98 or larger, the S2 subgroups had notably larger R^2 than the other subgroups. Our comparison included ALL, QS90, QS95, QS98, and QS99 subgroups with not-calibrated heights (the red bars in
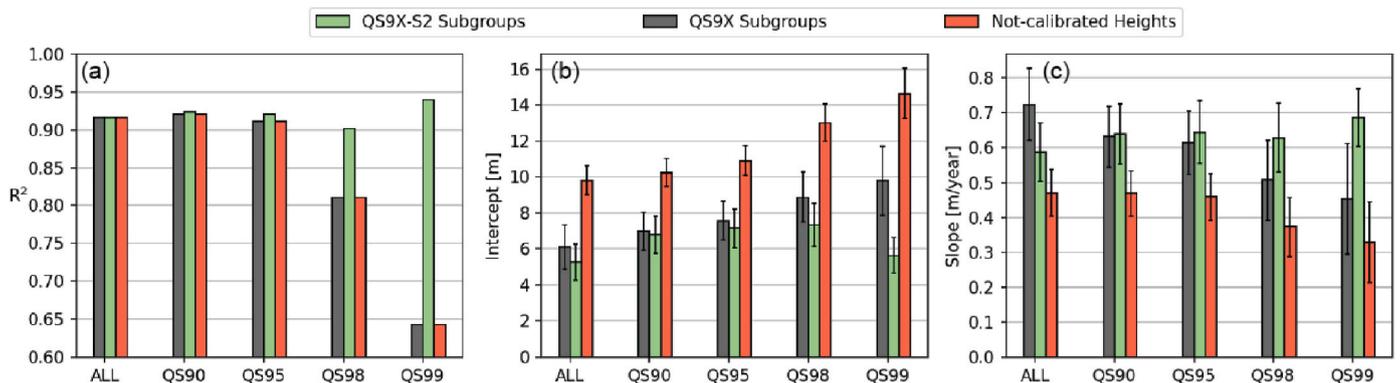
Fig. 11. Coefficient of determination (R^2), intercept and regrowth rate (slope) estimated with linear trends from the first 20 years of stand age and over different GEDI subgroups. Vertical error bars in the slope and intercept figures show the 2-sigma interval of the estimates. QS9X refers to QS90, QS95, QS98, QS99, and ALL. The red bars refer to the QS90, QS95, QS99, and ALL subgroups with not calibrated RH98 values. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)
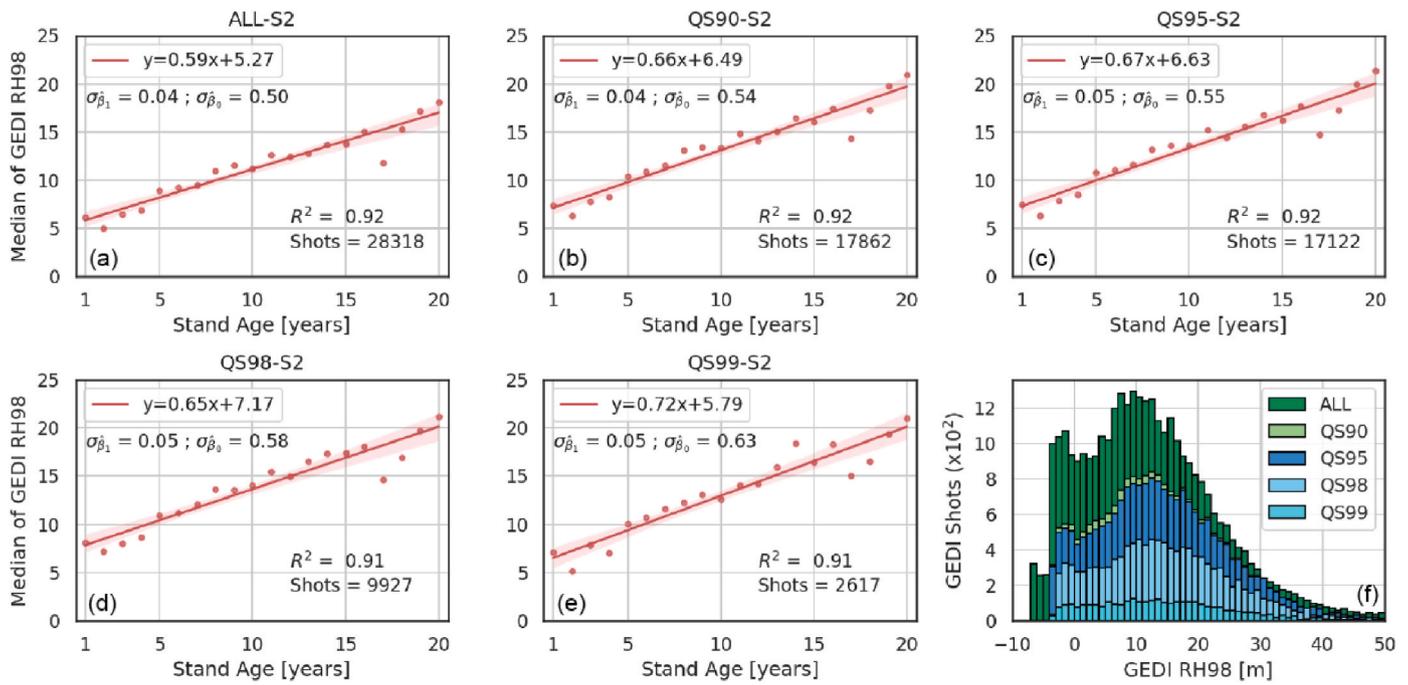
Fig. 12. Regrowth rates estimated from GEDI S2 subgroups with different beam sensitivity thresholds (a–e), and the corresponding distributions of RH98 values (f).

Fig. 11). However, the non-calibrated heights resulted in significantly larger intercepts and significantly lower regrowth rates than the remaining subgroups with calibrated heights. Their 2-sigma intervals did not, or just partly, overlap with one another. Furthermore, the subgroups with calibrated heights had intercepts closer to 5 m (the minimal forest height in the forest definition (FAO 2020)) than those with non-calibrated heights.

The height distributions of the S2 subgroups were all positively skewed (towards smaller forest heights), but the distribution of the QS99-S2 subgroup was platykurtic compared to the other (Fig. 12f). This was particularly visible in the 10–20 m height range, where the QS99-S2 distribution was without a notable mode. The distributions of QS90-S2, QS95-S2, and QS98-S2 were similar, but QS98-S2 had a slightly larger fraction of forest heights below 5 m. Furthermore, intercepts and regrowth rates of those three subgroups were more similar and with a higher overlap of their 2-sigma intervals than their counterparts in QS99-S2. The absolute difference between QS90-S2, QS95-S2, and QS98-S2 regrowth rates was 0.02 m/year, which was below the regrowth rates' standard errors (0.04–0.05 m/year). The absolute difference of their intercepts was 0.68 m, which was slightly above the intercepts' standard errors (0.54–0.58 m).

4.3.2. Retrieval of ICESat-2 regrowth rates

The retrieved regrowth rates differed for all ICESat-2 subgroups (Fig. 13). All the weak-beam subgroups (WN, WD, and WND) and the SD and SWD subgroups had small R^2 (0.01–0.50). The SN and SWN

subgroups had the highest R^2 (0.83 and 0.84, respectively) and very similar regrowth rates (0.90 m/year and 0.86 m/year) and intercepts (2.72 m and 2.80 m). However, 94% of the SWN segments were SN segments. The remaining SWN segments were from the WN subgroup that had a low R^2 (0.21).

We found much less weak-beam ICESat-2 ATL08 segments than the number of the strong-beam segments ($\sim$0.5 K and $\sim$3.3 K, respectively). For strong-beam subgroups, the number of day and night ICESat-2 segments were similar ($\sim$1.7 K and $\sim$1.6 K, respectively), whereas for the weak-beam subgroups, there were more day-than night-ICESat-2 segments ($\sim$0.4 K and $\sim$0.1 K, respectively).

The height distribution of all weak beams (WND) was platykurtic and shifted towards higher heights compared to the strong beams (Fig. 14a and b). Furthermore, the SN and SD distributions were different, with SN being more platykurtic and positively skewed (towards lower heights). The WN distribution was the most platykurtic and different from the other distributions.

The result also showed that segments with less than 10 photons were almost exclusively sampled with weak beams, whereas segments with 90 or more photons with strong beams (Fig. 14c and d). In the transition zone, the weak-beam segments continued to be in the majority up to 30 photons per segment after which the strong-beam segments dominate. For weak beams only, almost all night segments contained less than 50 canopy photons, whereas almost all day segments contained not more than 80 canopy photons. Based on these results, we do not have confidence in the weak beam regrowth rates.
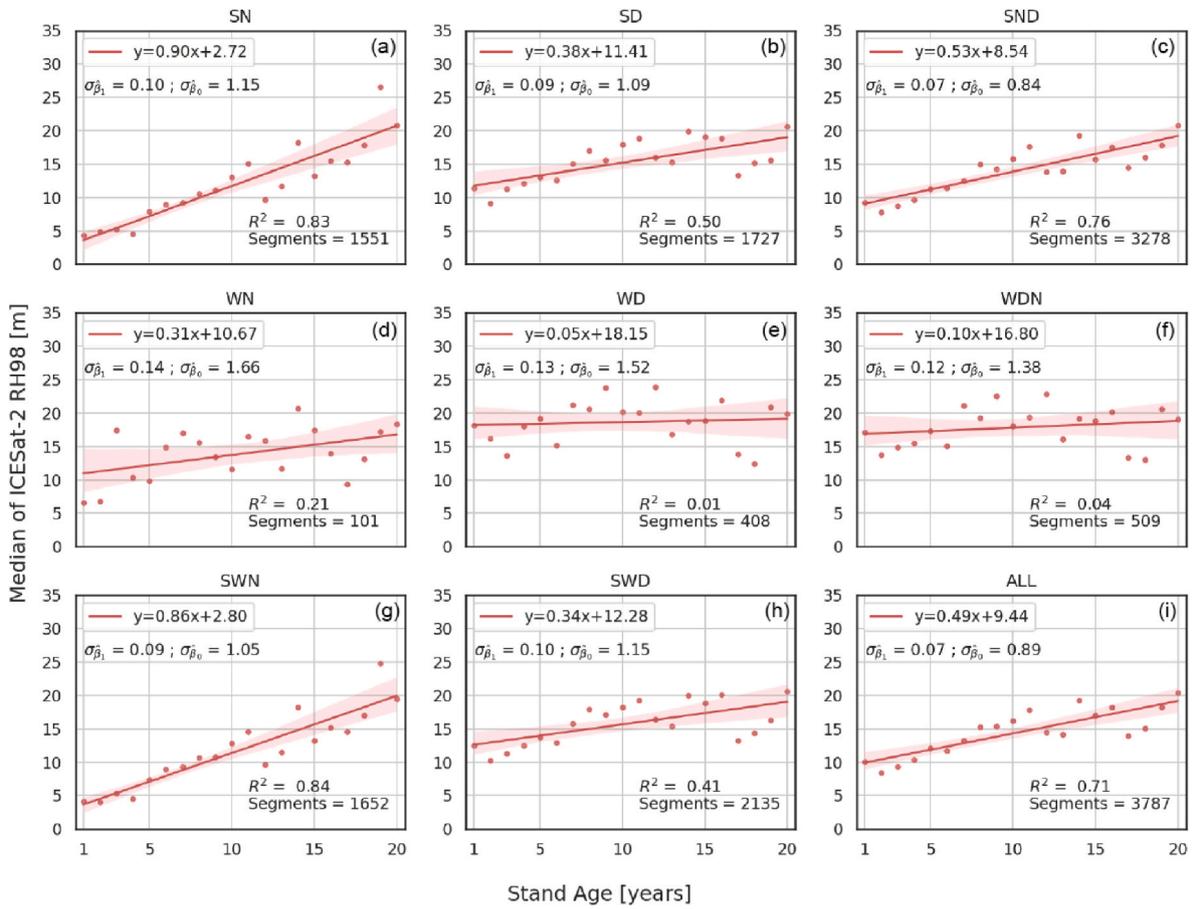
Fig. 13. Regrowth rates estimated from ICESat-2 subgroups with different beam strength and light conditions.
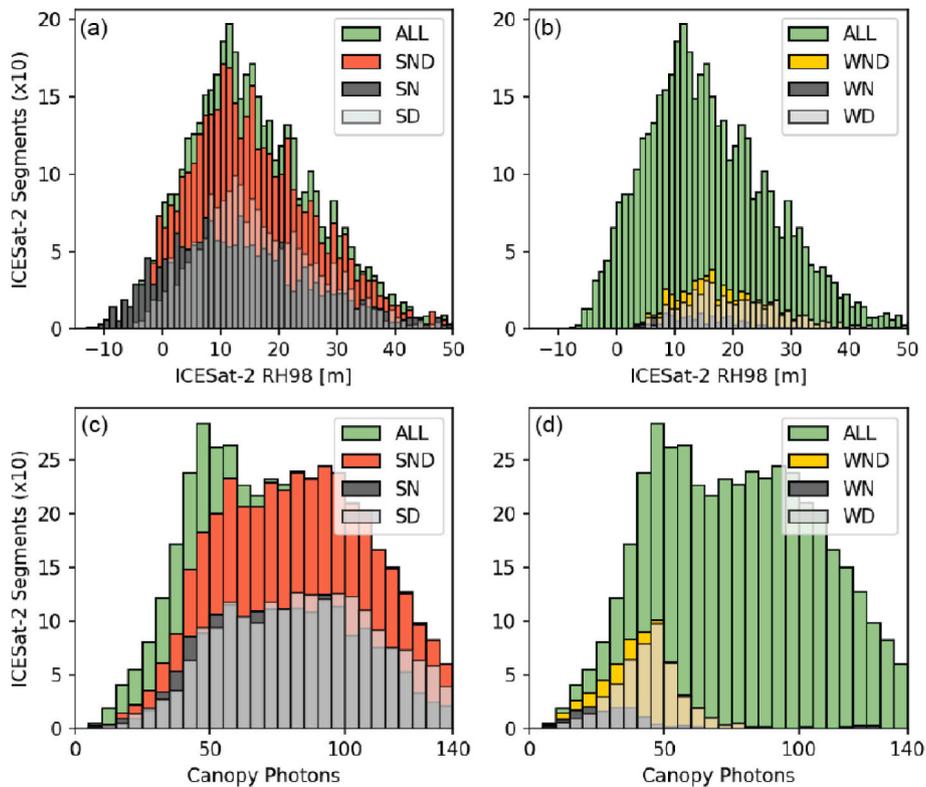
Fig. 14. The distributions of calibrated ICESat-2 RH98 values for strong-beam (a) and weak-beam (b) subgroups as well as the distributions of the canopy photons per ATL08 segment for strong-beam (c) and weak-beam (d) subgroups.

4.4. Selection of GEDI and ICESat-2 subgroups

The above analysis gave us the most confidence in the regrowth rates derived from the QS98-S2 and SN subgroups. Other GEDI and ICESat-2 subgroups failed or were not convincing in at least one of the criteria listed in Section 3.6. For example, QS99-S2 had the smallest RMSE and the highest R^2 compared to ALFH, but its height distribution, the intercept, and the regrowth rate were not consistent with the other S2 subgroups. Next, QS90-S2, QS95-S2, and QS98-S2 had similar intercepts and regrowth rates, but QS98-S2 had the smallest RMSE and the highest R^2 among them when ALFH was used as a reference. Furthermore, QS90-S2 showed that shots with such low sensitivities could contain erroneously small forest heights. The ICESat-2 weak-beam subgroups (WND, WN, and WD) failed on almost all the criteria. The SD subgroup had the smallest RMSE and the highest R^2 (whit ALFH as a reference) but notably smaller R^2 in the regrowth-rate regression analysis than the selected SN subgroup (0.50 and 0.83, respectively). Finally, SWN performed as good as SN, but we preferred SN because it did not contain non-descriptive heights from the WN segments.

4.5. Regrowth assessment with non-linear models

The spherical model best approximated the QS98-S2 and SN median height-distributions across the 33 years of regrowth, resulting in the highest R^2 among the three non-linear models (Fig. 15 and Table 5). The logarithmic model had the lowest R^2. The differences among THGs or ARRs were less sensitive to the selection of a satellited LiDAR dataset than to a non-linear model. For example, the difference among TGHs retrieved with the SPH model from the QS98-S2 and SN subgroups (20.17 m and 20.13 m, respectively) were smaller than the differences among TGHs retrieved from the QS98-S2 subgroup using the SPH, EXP, or LOG models (20.17 m, 21.06 m, 20.46 m, respectively). Nevertheless, the THG differences were below or close to the standard errors of the parameters. Finally, the SPH model showed a saturation after 25–26 years, whereas the EXP and LOG models showed no apparent saturation. The difference in saturation among the models also affected the corresponding ARRs.

Table 5

Coefficient of determination (R^2), coefficients ($\hat{a}$, $\hat{b}$, and $\hat{c}$) and their standard errors ($\sigma_{\hat{a}}$, $\sigma_{\hat{b}}$, and $\sigma_{\hat{c}}$) retrieved for spherical (SPH), exponential (EXP), and logarithmic (LOG) models fitted through the GEDI QS98-S2 and ICESat-2 SN subgroups. The average regrowth rates (ARRs) and the total height growths (THGs) were calculated as defined in Section 3.7.

Subgroup	Model	R^2	$\hat{a}$ [m]	$\hat{b}$ [m]	$\hat{c}$ [years]	$\sigma_{\hat{a}}$ [m]	$\sigma_{\hat{b}}$ [m]	$\sigma_{\hat{c}}$ [years]	THG [m]	ARR [m/yea]
QS98-S2	SPH	0.87	6.57	13.60	26.08	0.80	1.01	1.26	20.17	0.77
QS98-S2	EXP	0.86	4.73	16.33	10.76	1.38	1.22	2.47	21.06	0.65
QS98-S2	LOG	0.82	4.31	4.62	n/a	1.07	0.39	n/a	20.46	0.90
SN	SPH	0.70	2.16	17.97	25.26	1.75	2.23	1.75	20.13	0.80
SN	EXP	0.67	0.33	20.94	10.89	3.04	2.68	4.33	21.27	0.65
SN	LOG	0.63	−0.19	5.89	n/a	2.20	0.81	n/a	20.40	0.87
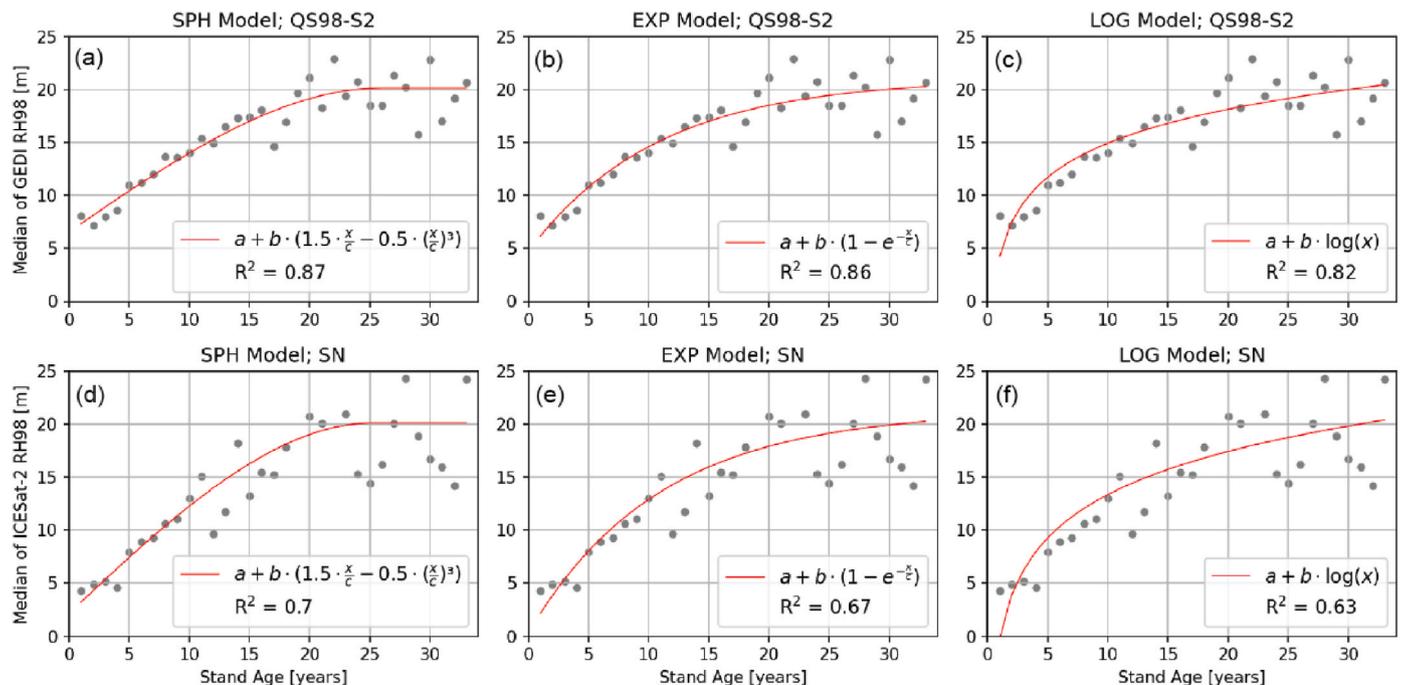
Fig. 15. Spherical (a) and (d), exponential (b) and (d), and logarithmic (c) and (f) models fitted over the 33 stand age period of the median GEDI and ICEsat-2 forest heights, respectively.

5. Discussion

5.1. Forest heights

5.1.1. GEDI forest heights and beam sensitivity

The GEDI beam sensitivity parameter demonstrated to be an important quality indicator of GEDI shots. We found the RMSE of forest heights decreased when increasing the sensitivity threshold from 0.90 towards 0.99 in the filtering of GEDI shots. GEDI shots with sensitivity 0.95 or larger (the QS95-S2, QS98-S2, and QS99-S2 subgroups) were free from erroneous GEDI heights. The QS99-S2 had the smallest RMSE and the highest R^2 among the subgroups. However, such filtering excluded heights in the 5–15 m range from the calibration (Fig. 6d) and then, in the regrowth assessment, left heights with a distribution different from the distributions from QS95-S2 and QS98-S2 heights (Fig. 12f). Those differences introduced inconsistencies between the QS99-S2 regrowth rate the QS95-S2 and QS98-S2 regrowth rates. Finally, we selected QS98-S2 because it had a smaller RMSE, a higher R^2, and was more restrictive on the sensitivity values than QS95-S2.

The GEDI Level 2A Version 2 algorithm setting group selection impacts the derived GEDI variables such as beam sensitivity and RH98. This was accounted for by introducing two different selection strategies for those two GEDI variables (Table 1). Our results showed that the strategy with using the default RH98 values together with the sensitivities from the algorithm setting group 2 (the S2 subgroups) outperformed the default selection strategy, i.e., one with all the default values (the QS90–QS99 subgroups). The QS90-QS99 subgroups performed similar to the S2 subgroups in the regrowth assessment, except for the sensitivities of 0.98 or larger. However, when airborne LiDAR heights were used as a reference, the QS90-QS99 subgroups had the largest RMSE and the lowest R^2 than their counterparts in the S2 subgroups.

The smallest RMSE of forest heights (3.61 m) was obtained with the QS99-S2 subgroup, whereas the selected QS98-S2 subgroup had a slightly larger RMSE (4.14 m). The RMSE observed for the rainforest were larger than the RMSEs (2.98–3.56 m) reported for forest sites in Germany, Spain, and United States (Adam et al., 2020; Liu et al., 2021; Quirós et al., 2021). However, the latter RMSEs correspond to GEDI data that were either corrected for the geolocation error (Liu et al., 2021; Quirós et al., 2021) or free from specific low-quality orbits (Adam et al., 2020). Nevertheless, it is encouraging to see that the RMSEs are generally similar even though the Amazon rainforest has the highest structural complexity that can notably attenuate laser energy and decrease the detection accuracy of the ground as well as the accuracy of relative height metrics.

5.1.2. ICESat-2 forest heights and canopy photons

Because ICESat-2 is a photon counting system, rather than a full-waveform system, only a few photons are detected by the ATLAS receiver for each outgoing shot. The number of detected photon events associated with each outgoing laser pulse is a function of the transmitted laser energy, surface reflectance, solar conditions, and scattering and attenuation in the atmosphere. The outgoing energy of the strong beams is four times the outgoing energy of the weak beams. Similarly, the detector array for the strong beams consists of a 4 × 4 array, where the weak beam array is a 2 × 2. The ATL08 data product is reported at the 100 m segment length along the orbit path. Based on the reported 70 cm distance between outgoing shots, each ATL08 segment should consist of approximately 140 shots in the along-track direction. Based on pre-launch simulations, the mean number of photons based on a Poisson

distribution for tropical forests is 0.6 (Neuenschwander and Magruder 2016) in the strong beam case and 0.15 in the weak beam case. A recent study shows that those numbers are much larger when calculated from the ICESat-2 data acquired over a region in the Amazon rainforest., i.e. 1.31 and 0.96 for the strong and the weak case, respectively (Neuenschwander et al., 2022). For this reason, we conservatively rejected ATL08 segments with more than 140 canopy photons due to the likelihood that segments with many canopy photons were mislabeled. By imposing this criterion, we removed the forest height outliers in the strong beam, night case. All rejected ATL08 segments were from the same granule, lending to the theory that the errant heights were likely caused by an external factor (i.e., low-lying clouds or fog).

Our results regarding the poor performance of the weak beam in forest height retrieval are confirmed by other studies. Although the forest heights extracted from the weak beam can often be considered valid, they always have a higher uncertainty (Neuenschwander et al., 2020a). For this reason, we felt we could not include them in the regrowth rates.

RMSE of ICESat-2 forest heights reported here are comparable with other ICESat-2 studies. The smallest RMSE of forest heights (4.42 m) was obtained with the SD subgroup, whereas the selected SND subgroup had a larger RMSE (5.41 m) and provided a larger sample size. Those values are similar to the RMSE value (5.02 m) reported for the study sites with diverse forest types (mainland temperate forest, Alaskan tundra, and Hawaiian rainforest) in the United States (Liu et al., 2021). However, our RMSEs are larger than one (2.45 m) reported for the test site with the boreal forest in Finland (Neuenschwander et al., 2020a). The study from Finland, however, applied a geolocation correction and was highly selective in the ATL08 segments analyzed.

For our two airborne LiDAR sites, the RMSEs of ICESat-2 forest heights are larger but still similar to the RMSEs of GEDI forest heights. The RMSE of the ICESat-2 SD subgroup (4.42 m) is similar to the RMSE values observed from the GEDI QS98-S2 subgroup (4.14 m). However, one important caveat to this result is a low number of ICESat-2 orbits (2) compared to the number of GEDI orbits (22) for the two sites. Nevertheless, ongoing global and pan-tropical calibration/validation initiatives of the GEDI and ICESat-2 Science Teams will be based on more data, which will give additional evidence of that relation.

5.2. Regrowth assessment

5.2.1. Regrowth rates

Approximation of median GEDI and ICESat-2 forest heights with the spherical model provided total growth estimates of 20.17 m and 20.13 m with standard errors of 1.3 m and 2.8 m. Their average regrowth rates were 0.77 m/year and 0.80 m/year, respectively. Those regrowth rates are comparable to values found in other LiDAR-based studies. For example, Becknell et al. (2018) reported regrowth rates of the Brazilian Atlantic forest between 0.9 m/year and 0.6 m/year on linearly modeled regrowth periods of 10 and 20 years, respectively. The regrowth rates in Becknell et al. (2018) are based on median airborne LiDAR heights and a Landsat-based disturbance history map for a 50 km^2 large study site. Then, Chave et al. (2020) reported a total growth of 18.04 m in the median forest height after 33 years of secondary forest regrowth in French Guiana. This total growth was calculated from a 4 km^2 area of the secondary forest site surveyed with airborne LiDAR.

The above regrowth rates were similar to field observations from 72 restoration sites in the neighboring Brazilian state Mato Grosso (Freitas et al., 2019). They modeled forest heights derived from the ten tallest individuals across 10 years of regrowth and derived two logarithmic models corresponding to different soil base saturation. They obtained

the intercepts of 1.27 m and 0.8 m as well as the scaling coefficients of 3.10 m and 3.16 m, which resulted in total 10-year growth of 8.40 m and 8.08 m, respectively. Those values were derived assuming the natural logarithmic model, but that was not explicitly stated in the study. Nevertheless, the observed differences between LiDAR-based and field-based regrowth rates are expected, considering the differences in the forest height definition and sampling size.

It should be noted that the above regrowth rates were underestimated when any of the processing steps, such as the removal of secondary-forest border pixels (Section 2.3) or the calibration of forest heights (Section 3.3), was omitted. For example, the QS98-S2 regrowth rate (from the linear approximation of the 20-year regrowth period) dropped from 0.65 to 0.50 m/year when the secondary-forest border pixels were not removed (Fig. 12d and Figure A1d, respectively). This was a reduction of more than 20% in the regrowth rate estimate. In addition, R^2 reduced from 0.91 to 0.85, and the regrowth intercept increased from 7.17 to 10.94 m. Next, when only the calibration step was omitted, the regrowth rate dropped to 0.50 m/year, the regrowth intercept increased to 10.25 m, whereas R^2 remained the same (Figure A2d). Other results related to those GEDI subgroups are not presented here to discourage any further usage of such subgroups and avoid potential misinterpretations of their results.

5.2.2. Regrowth intercepts

In addition to regrowth rates, our regrowth assessments resulted in GEDI and ICESat-2 intercepts of 6.57 m and 2.16 m, respectively. The ICESat-2 intercepts for each of the three non-linear models were lower than the GEDI intercepts, which might be due to different pulse width, imperfect calibration models, and a coarser spatial resolution of ATL08 segments compared to GEDI footprints. The intercepts theoretically refer to the median heights of the secondary forest with age 0, but age class 0 in our stand age map refers to the non-secondary forest cover. Therefore, it is more appropriate to look at the median forest height in the first year of regrowth (stand age = 1), which can be calculated by adding the corresponding ARR to the intercept. Those median heights refer to pixels that have been just classified as forest cover after an anthropic cover (e.g., pasture or agriculture) in the previous year (Silva Junior et al., 2020). Therefore, the median heights in the first year of regrowth are expected to be similar to the minimal forest height specified in the forest class definition. According to the FAO definition (FAO 2020), the minimal forest height is 5 m, which was overestimated by the SPH model and GEDI data (7.34 m).

5.2.3. Calibration of satellite LiDAR

The calibration analysis showed that GEDI and ICESat-2 forest heights tended to overestimate lower forest heights and underestimate higher forest heights, when compared to the airborne LiDAR forest heights (Fig. 16a). An underestimation of higher forest heights with ICESat-2 RH95 values was reported in a recent study where those

heights were compared with a nationwide forest height product over different ecozones in Canada (Mulverhill et al., 2022, Fig. 5). Such systematic distortions of satellite heights reduced the height range and introduced a "saturation" effect in a regrowth model based on non-calibrated heights (Fig. 16b). In addition, the systematic distortions led to larger regrowth intercepts, which in turn provided an underestimation of regrowth rates. As our calibration models were different for each subgroup, it was necessary to apply them and use the calibrated satellite heights in the comparison and parameter retrieval.

Our linear calibration models, based on the two calibration sites in the Amazon rainforest, provided regrowth intercepts close to 5 m, i.e., the minimal forest height according to the forest definition. Furthermore, the resulting average regrowth rates were similar to ones derived from the previous airborne LiDAR and field studies. This shows that our calibration models are acceptable, but due to a limited sample size, some of our models may be suboptimal. For example, our ICESat-2 SN calibration data did not have forest heights below 20 m, whereas the SD data had a small sample size (Fig. 9c and d). Next, due to an imperfect model and high data dispersion, some of our calibrated heights were negative (Fig. 10a and b). Truncating those heights to zero would only result in a larger intercept and a lower regrowth rate. Thus, we decided to use the medians of the calibrated heights because they were all positive real numbers. Finally, a larger sample size covering the whole calibration domain and a non-liner (or piecewise linear) modeling with constraints may lead to more accurate calibration models.

5.2.4. Median heights and regrowth uncertainty

Our approach with median GEDI and ICESat-2 forest heights is robust to outliers, but as a spatial aggregate, it also smooths out spatial variability of forest growth. Our study covers a wide range of regrowth types, from heavily disturbed to natural regenerations, which are here jointly considered in their type-independent median heights. Furthermore, it was observed that the forest height distribution per height class was skewed for young ages and that with the selection of the median heights might have affected the outcome of the model. Extensive, multi-date field reference data would help to better understand changes in height compared to satellite estimates. Such studies would also require other validation datasets such as large-area, multi-temporal airborne LiDAR data that overlap widely with the satellite LiDAR tracks. Therefore, it should be noted that our study presented an internal uncertainty of regrowth, i.e., regrowth precision, but future studies are necessary to assess the regrowth accuracy.

Our median heights also include errors coming from the stand age map that were not analyzed here, yet they can potentially be an important uncertainty source. The stand age map used here was derived from the forest class in the annual MapBiomas land use maps. Therefore, commission errors of the forest class will include satellite LiDAR heights from the non-secondary forest cover, such as pasture or even old-growth forest, into the calculation of median heights. Those commission errors
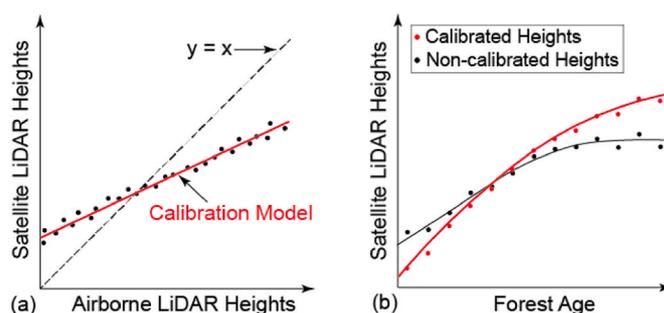
Fig. 16. Schematic representation of (a) a calibration model between satellite- and airborne- LiDAR heights, and (b) regrowth models derived using calibrated and non-calibrated satellite LiDAR heights.

propagate further into stand age classes by accumulating annual forest masks. Furthermore, the uncertainty of our calibrated heights ($\sigma^2_{H'}$) depends on the accuracy of calibration parameters (σ^2_a and σ^2_b), but also on the calibrated height itself (H'), which can be seen after applying a simple error propagation on the calibration equation and neglecting the covariance terms:

$$H' = \frac{H - a}{b} \rightarrow \sigma^2_{H'} = \frac{1}{b^2}\left(\sigma^2_H + \sigma^2_a + H'^2 \cdot \sigma^2_b\right)$$

where *a* and *b* are the calibration intercept and slope, whereas *H* is the observed satellite LiDAR height. Together with the stand age errors, this dependency of height error on its magnitude is why our median heights were more variable when going towards older stand age classes (Fig. 10a and b and Fig. 15). The omission errors of forest age class are not expected to affect the median heights significantly as they will only decrease the sample size. However, this is true only when those omissions happen randomly and do not significantly change the distribution of sampled heights. To assess how the stand age errors influence final regrowth rates, it would require a separate study design and methodology tailored according to that goal. Nevertheless, those advances are necessary to fully understand uncertainties associated with the regrowth rate retrieval.

5.3. Outlook

This study looked at methodological steps and uncertainty associated with the estimation of large-scale regrowth rates from contemporary satellite LiDAR data over the Amazon rainforest. The methodology presented here can be transferred to other regions to compare regrowth rates or look at different regrowth and reforestation types. Such spatial variabilities in regrowth rates are necessary, e.g., to identify regions with degraded ecosystem functioning, which is important for assessing their ecological resilience. The regrowth rate uncertainty presented here, together with future studies on the uncertainties caused by the errors in stand age maps, can help to determine the significance of the spatial variability in the regrowth rates.

Another step forward is to relate regrowth directly to aboveground biomass (AGB) gain, which is required to estimate and map the terrestrial carbon sink accurately. An important development in that direction is the GEDI footprint-level AGB density product, and its first version has been recently published by the GEDI Science Team (Dubayah et al., 2021b). Our results showed that the errors of ICESat-2 regrowth rate estimates from spatially aggregated median height estimates over large areas such as the state of Rondônia in Brazil are comparable to those from GEDI, which is encouraging for mapping AGB gain in the boreal domain that is not covered with GEDI data. Direct contributions to that end are the ongoing mapping of boreal-wide AGB with ICESat-2 forest heights supported through the NASA ABoVE program and the ICESat-2 Science Team.

6. Conclusions

This study on assessing Amazon rainforest regrowth from GEDI and ICESat-2 data revealed the importance of satellite LiDAR data and their ability to retrieve forest regrowth rates. Our results showed that erroneous forest heights were excluded when GEDI shots with a beam sensitivity of 0.95 or larger were selected. Furthermore, the strategy using RH98 values from the automatically-selected algorithm setting group and with the beam sensitivities from the algorithm setting group 2

outperformed other selection strategies. On the other hand, ICESat-2 ATL08 segments with more than 140 canopy photons per segment were excluded as they contained erroneously large forest heights.

The study revealed that calibration of satellite LiDAR heights with airborne LiDAR heights was necessary to avoid erroneously large regrowth intercepts and an underestimation of forest regrowth rate. RH98 values of both satellite missions often overestimated lower forest heights and underestimated higher forest heights, introducing a "saturation" effect of the median forest heights of 25 years or older secondary forests. Although a simple linear calibration model was able to correct those systematic distortions, it is expected that the calibration models can be further improved by including a larger calibration sample size, and refining approaches to quality filtering and selection of algorithm parameters. The study also showed that omitting either the calibration step or the removal of secondary-forest-border pixels would result in an underestimation of the regrowth rate by more than 20%.

The regrowth assessment showed that regrowth rates based on ICESat-2 subgroups were more variable than ones based on the GEDI subgroups. In addition, the ICESat-2 weak beams were found unreliable for regrowth retrieval. Therefore, we used only ICESat-2 strong night beams combined with the canopy photon filter. The GEDI shots with sensitivities larger than 0.95 and 0.98 provided similar regrowth rates with the highest R^2. However, the shots with sensitivities of 0.98 or larger were preferred because they had a smaller RMSE and a higher R^2 when the airborne LiDAR heights were used as a reference.

Finally, the study showed that the spherical model best approximated GEDI and ICESat-2 median forest height across the 33-year regrowth period. The second-best model (exponential) had a similar R^2, but yielded 18% and 23% lover ARR estimates for GEDI and ICESat-2, respectively, reviling how important can be a selection of a non-linear model in regrowth assessment. When the non-linear model was fixed, the differences between the GEDI and ICESat-2 estimates of the total height growth were within their standard error. However, GEDI estimates were preferred because of their larger sample size in regrowth and calibration analysis, better forest height accuracy, and lower internal uncertainty. Nevertheless, it is encouraging to see a general agreement between GEDI and ICESat-2 estimates, considering their differences in spatial and temporal resolution as well as the LiDAR principle.

This study demonstrated that novel satellite LiDAR data and the proposed methodology could assess median forest regrowth rates over large areas. However, the study also demonstrated the uncertainty associated with the retrieval of regrowth rates from GEDI and ICEsta-2 data. Errors on the age of secondary forest were not studied here but can potentially affect the retrieval of regrowth rates. Therefore, it would be necessary to study those errors further and understand how the retrieval uncertainty affects specific applications such as terrestrial carbon sink mapping, an ecological resilience assessment, or detection of forest degradation processes.

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.
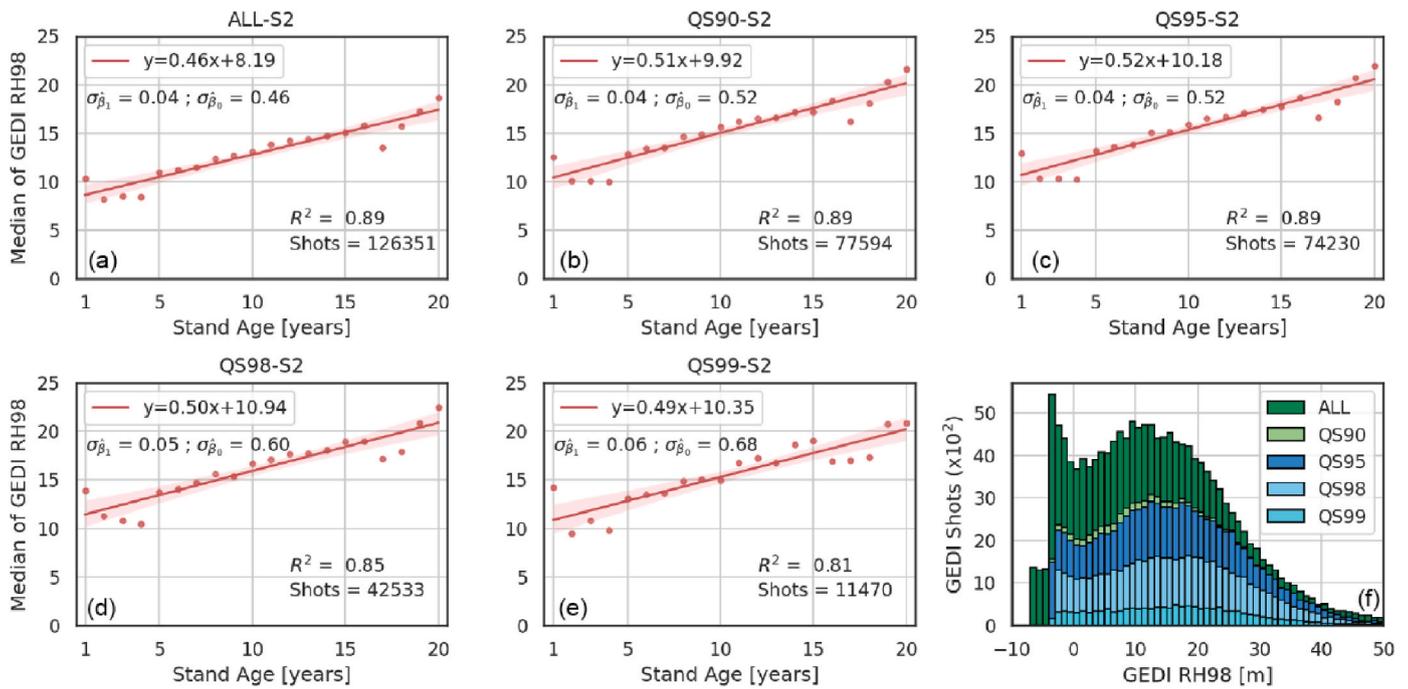
Fig. A.1. Regrowth rates assessed when the secondary-forest border pixels were not removed from the stand age map. The results were derived for the GEDI S2 subgroups with different beam sensitivity thresholds (a–e), and the corresponding distributions of RH98 values (f).
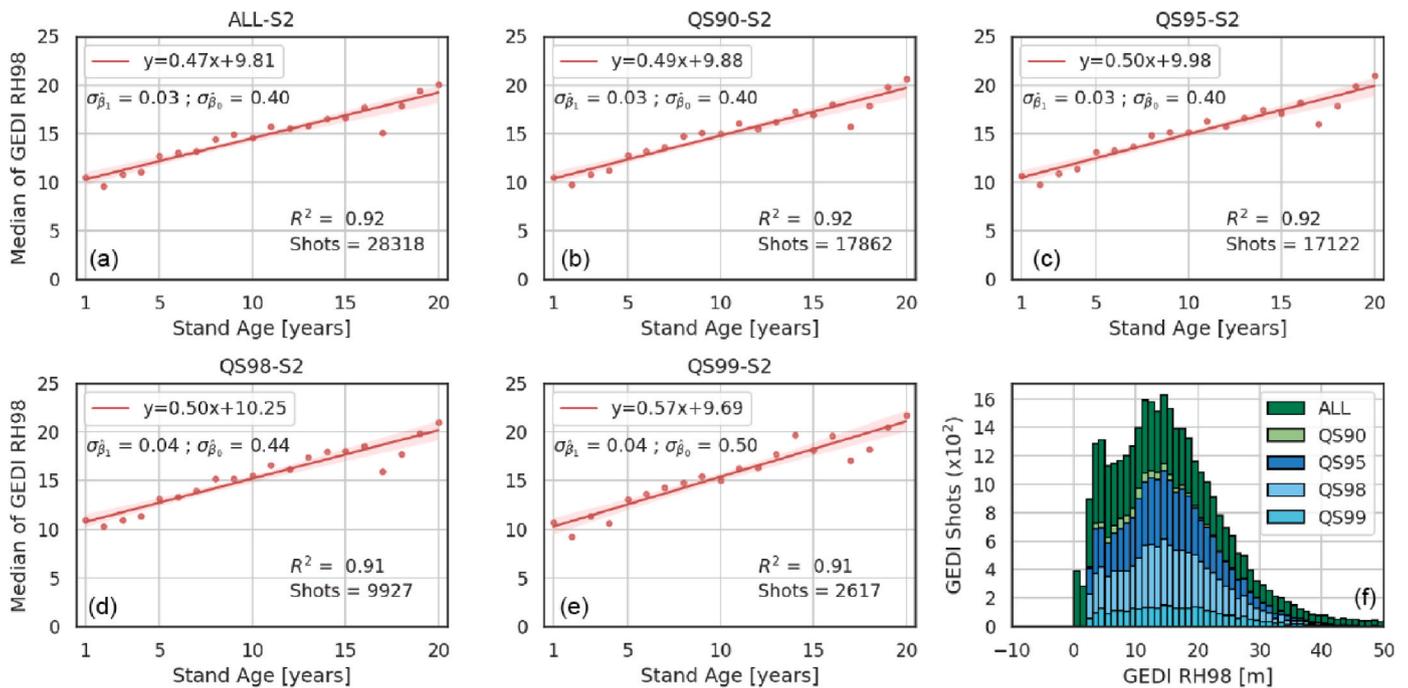
Fig. A.2. Regrowth rates assessed when the calibration of GEDI forest heights was not applied. The results were derived for the GEDI S2 subgroups with different beam sensitivity thresholds (a–e), and the corresponding distributions of RH98 values (f).

Acknowledgements

References

Adam, M., Urbazaev, M., Dubois, C., Schmullius, C., 2020. Accuracy assessment of GEDI terrain elevation and canopy height estimates in European temperate forests: influence of environmental and acquisition parameters. Rem. Sens. 12.

Baccini, A., Walker, W., Carvalho, L., Farina, M., Sulla-Menashe, D., Houghton, R.A., 2017. Tropical forests are a net carbon source based on aboveground measurements of gain and loss. Science 358, 230.

Beck, J., Luthcke, S.B., Hofton, M., Armston, J., 2020. In: Global Ecosystem Dynamics Investigation (GEDI) Level 1B User Guide. The USGS Earth Resources Observation and Science (EROS) Center: NASA's Land Processes Distributed Active Archive Center (LP DAAC), p. 13.

Becknell, J.M., Keller, M., Piotto, D., Longo, M., Nara dos-Santos, M., Scaranello, M.A., Bruno de Oliveira Cavalcante, R., Porder, S., 2018. Landscape-scale lidar analysis of aboveground biomass distribution in secondary Brazilian Atlantic Forest. Biotropica 50, 520–530.

Bivand, R.S., Pebesma, E., Gómez-Rubio, V., 2013. Interpolation and geostatistics. In: Bivand, R.S., Pebesma, E., Gómez-Rubio, V. (Eds.), Applied Spatial Data Analysis with R. Springer New York, New York, NY, pp. 213–261.

Boehm, H.V., Liesenberg, V., Limin, S.H., 2013. Multi-temporal airborne LiDAR-survey and field measurements of tropical peat swamp forest to monitor changes. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 6, 1524–1530.

Chave, J., Piponiot, C., Maréchaux, I., de Foresta, H., Larpin, D., Fischer, F.J., Derroire, G., Vincent, G., Hérault, B., 2020. Slow rate of secondary forest carbon accumulation in the Guianas compared with the rest of the Neotropics. Ecol. Appl. 30, e02004.

Chazdon, R.L., Broadbent, E.N., Rozendaal, D.M.A., Bongers, F., Zambrano, A.M.A., Aide, T.M., Balvanera, P., Becknell, J.M., Boukili, V., Brancalion, P.H.S., Craven, D., Almeida-Cortez, J.S., Cabral, G.A.L., de Jong, B., Denslow, J.S., Dent, D.H., DeWalt, S.J., Dupuy, J.M., Durán, S.M., Espírito-Santo, M.M., Fandino, M.C., César, R.G., Hall, J.S., Hernández-Stefanoni, J.L., Jakovac, C.C., Junqueira, A.B., Kennard, D., Letcher, S.G., Lohbeck, M., Martínez-Ramos, M., Massoca, P., Meave, J. A., Mesquita, R., Mora, F., Muñoz, R., Muscarella, R., Nunes, Y.R.F., Ochoa-Gaona, S., Orihuela-Belmonte, E., Peña-Claros, M., Pérez-García, E.A., Piotto, D., Powers, J.S., Rodríguez-Velazquez, J., Romero-Pérez, I.E., Ruíz, J., Saldarriaga, J.G., Sanchez-Azofeifa, A., Schwartz, N.B., Steininger, M.K., Swenson, N.G., Uriarte, M., van Breugel, M., van der Wal, H., Veloso, M.D.M., Vester, H., Vieira, I.C.G., Bentos, T.V., Williamson, G.B., Poorter, L., 2016. Carbon sequestration potential of second-growth forest regeneration in the Latin American tropics. Sci. Adv. 2, e1501639.

Clark, D.A., Clark, D.B., 2001. Getting to the canopy: tree height growth IN a neotropical rain forest. Ecology 82, 1460–1472.

Dolan, K., Masek, J.G., Huang, C., Sun, G., 2009. Regional forest growth rates measured by combining ICESat GLAS and Landsat data. J. Geophys. Res.: Biogeosciences 114.

Dos-Santos, M.N., Keller, M.M., Morton, D.C., 2019. In: LiDAR Surveys over Selected Forest Research Sites, Brazilian Amazon, 2008-2018. ORNL Distributed Active Archive Center.

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., Silva, C., 2020. 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., 2021a. GEDI L2A Elevation and Height Metrics Data Global Footprint Level V002 [April to October 2019, and 66.8 °W, 13.5 °S; 59.5 °W, 7.9 °S]. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/GEDI/GEDI02_A.002. (Accessed 14 June 2021).

Dubayah, R.O., Armston, J., Kellner, J.R., Duncanson, L., Healey, S.P., Patterson, P.L., Hancock, S., Tang, H., Hofton, M.A., Blair, J.B., Luthcke, S., 2021b. In: GEDI L4A Footprint Level Aboveground Biomass Density, Version 1. ORNL Distributed Active Archive Center.

Duncanson, L., Neuenschwander, A., Hancock, S., Thomas, N., Fatoyinbo, T., Simard, M., Silva, C.A., Armston, J., Luthcke, S.B., Hofton, M., Kellner, J.R., Dubayah, R., 2020. Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County, California. Remote. Sens. Environ. 242, 111779.

FAO, 2020. Global Forest Resources Assessment 2020: Terms and Definitions. The Food and Agriculture Organization (FAO), Rome, Italy.

Freitas, M.G., Rodrigues, S.B., Campos-Filho, E.M., do Carmo, G.H.P., da Veiga, J.M., Junqueira, R.G.P., Vieira, D.L.M., 2019. Evaluating the success of direct seeding for tropical forest restoration over ten years. For. Ecol. Manag. 438, 224–232.

Guerra-Hernández, J., Pascual, A., 2021. Using GEDI lidar data and airborne laser scanning to assess height growth dynamics in fast-growing species: a showcase in Spain. Forest Ecosyst. 8, 14.

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

Harris, I., Osborn, T.J., Jones, P., Lister, D., 2020. Version 4 of the CRU TS monthly high-resolution gridded multivariate climate dataset. Sci. Data 7, 109.

Hopkinson, C., Chasmer, L., Hall, R.J., 2008. The uncertainty in conifer plantation growth prediction from multi-temporal lidar datasets. Remote. Sens. Environ. 112, 1168–1180.

Houghton, R.A., 2003. Revised estimates of the annual net flux of carbon to the atmosphere from changes in land use and land management 1850–2000. Tellus B 55, 378–390.

Hyyppä, J., Hyyppä, H., Leckie, D., Gougeon, F., Yu, X., Maltamo, M., 2008. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. Int. J. Rem. Sens. 29, 1339–1366.

Isenburg, M., 2019. In: The LAStools Software Suite. https://rapidlasso.com/lastools/: rapidlasso GmbH.

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., Turner, D.P., Guzy, M., Cohen, W.B., 2005. Combining lidar estimates of aboveground biomass and Landsat estimates of stand age for spatially extensive validation of modeled forest productivity. Remote. Sens. Environ. 95, 549–558.

Li, W., Niu, Z., Shang, R., Qin, Y., Wang, L., Chen, H., 2020. High-resolution mapping of forest canopy height using machine learning by coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 data. Int. J. Appl. Earth Obs. Geoinf. 92, 102163.

Liu, A., Cheng, X., Chen, Z., 2021. Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals. Remote. Sens. Environ. 264, 112571.

Magruder, L.A., Brunt, K.M., Alonzo, M., 2020. Early ICESat-2 on-orbit geolocation validation using ground-based corner cube retro-reflectors. Rem. Sens. 12.

MapBiomas, 2021. In: Project MapBiomas - Collection 6.0 of Brazilian Land Cover & Use Map Series through the link: https://mapbiomas.org/en/statistics. (Accessed 2 November 2021).

Marselis, S.M., Abernethy, K., Alonso, A., Armston, J., Baker, T.R., Bastin, J.-F., Bogaert, J., Boyd, D.S., Boeckx, P., Burslem, D.F.R.P., Chazdon, R., Clark, D.B., Coomes, D., Duncanson, L., Hancock, S., Hill, R., Hopkinson, C., Kearsley, E., Kellner, J.R., Kenfack, D., Labrière, N., Lewis, S.L., Minor, D., Memiaghe, H., Monteagudo, A., Nilus, R., O'Brien, M., Phillips, O.L., Poulsen, J., Tang, H., Verbeeck, H., Dubayah, R., 2020. Evaluating the potential of full-waveform lidar for mapping pan-tropical tree species richness. Global Ecol. Biogeogr. 29, 1799–1816.

Mulverhill, C., Coops, N.C., Hermosilla, T., White, J.C., Wulder, M.A., 2022. Evaluating ICESat-2 for monitoring, modeling, and update of large area forest canopy height products. Remote. Sens. Environ. 271, 112919.

Næsset, E., Gobakken, T., 2005. Estimating forest growth using canopy metrics derived from airborne laser scanner data. Remote. Sens. Environ. 96, 453–465.

Narine, L.L., Popescu, S.C., Malambo, L., 2019. Synergy of ICESat-2 and Landsat for mapping forest aboveground biomass with deep learning. Rem. Sens. 11.

Neuenschwander, A., Guenther, E., White, J.C., Duncanson, L., Montesano, P., 2020a. Validation of ICESat-2 terrain and canopy heights in boreal forests. Remote. Sens. Environ. 251, 112110.

Neuenschwander, A., Magruder, L., Guenther, E., Hancock, S., Purslow, M., 2022. Radiometric assessment of ICESat-2 over vegetated surfaces. Rem. Sens. 14.

Neuenschwander, A.L., Magruder, L.A., 2016. The potential impact of vertical sampling uncertainty on ICESat-2/ATLAS terrain and canopy height retrievals for multiple ecosystems. Rem. Sens. 8.

Neuenschwander, A.L., Magruder, L.A., 2019. Canopy and terrain height retrievals with ICESat-2: a first look. Rem. Sens. 11.

Neuenschwander, A.L., Pitts, K.L., Jelley, B.P., Robbins, J., Klotz, B., Popescu, S.C., Nelson, R.F., Harding, D., Pederson, D., Sheridan, R., 2020b. In: ATLAS/ICESat-2 L3A Land and Vegetation Height, Version 3. [April to October 2019, and 66.8 °W, 13.5 °S; 59.5 °W, 7.9 °S]. Boulder, Colorado USA. NASA National Snow and Ice Data Center Distributed Active Archive Center.

Neumann, T.A., Martino, A.J., Markus, T., Bae, S., Bock, M.R., Brenner, A.C., Brunt, K.M., Cavanaugh, J., Fernandes, S.T., Harding, D.W., Harbeck, K., Lee, J., Kurtz, N.T., Luers, P.J., Luthcke, S.B., Magruder, L., Pennington, T.A., Ramos-Izquierdo, L., Rebold, T., Skoog, J., Thomas, T.C., 2019. The ice, cloud, and land elevation satellite – 2 mission: a global geolocated photon product derived from the advanced topographic laser altimeter system. Remote. Sens. Environ. 233, 111325.

Pan, Y., Birdsey, R.A., Fang, J., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., Ciais, P., Jackson, R.B., Pacala, S.W., McGuire, A.D., Piao, S., Rautiainen, A., Sitch, S., Hayes, D., 2011. A large and persistent carbon sink in the world's forests. Science 333, 988.

Pedlowski, M.A., Dale, V.H., Matricardi, E.A.T., da Silva Filho, E.P., 1997. Patterns and impacts of deforestation in Rondônia, Brazil. Landsc. Urban Plann. 38, 149–157.

Pekel, J.-F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. Nature 540, 418–422.

Pickett, S.T.A., 1989. Space-for-Time substitution as an alternative to long-term studies. In: Likens, G.E. (Ed.), Long-Term Studies in Ecology: Approaches and Alternatives. Springer New York, New York, NY, pp. 110–135.

Poorter, L., Bongers, F., Aide, T.M., Almeyda Zambrano, A.M., Balvanera, P., Becknell, J. M., Boukili, V., Brancalion, P.H.S., Broadbent, E.N., Chazdon, R.L., Craven, D., de Almeida-Cortez, J.S., Cabral, G.A.L., de Jong, B.H.J., Denslow, J.S., Dent, D.H., DeWalt, S.J., Dupuy, J.M., Durán, S.M., Espírito-Santo, M.M., Fandino, M.C., César, R.G., Hall, J.S., Hernández-Stefanoni, J.L., Jakovac, C.C., Junqueira, A.B., Kennard, D., Letcher, S.G., Licona, J.-C., Lohbeck, M., Marín-Spiotta, E., Martínez-Ramos, M., Massoca, P., Meave, J.A., Mesquita, R., Mora, F., Muñoz, R., Muscarella, R., Nunes, Y.R.F., Ochoa-Gaona, S., de Oliveira, A.A., Orihuela-Belmonte, E., Peña-Claros, M., Pérez-García, E.A., Piotto, D., Powers, J.S., Rodríguez-Velázquez, J., Romero-Pérez, I.E., Ruíz, J., Saldarriaga, J.G., Sanchez-Azofeifa, A., Schwartz, N.B., Steininger, M.K., Swenson, N.G., Toledo, M., Uriarte, M., van Breugel, M., van der Wal, H., Veloso, M.D.M., Vester, H.F.M., Vicentini, A., Vieira, I. C.G., Bentos, T.V., Williamson, G.B., Rozendaal, D.M.A., 2016. Biomass resilience of Neotropical secondary forests. Nature 530, 211–214.

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M.C., Kommareddy, A., Pickens, A., Turubanova, S., Tang, H., Silva, C.E., Armston, J., Dubayah, R., Blair, J. B., Hofton, M., 2021. Mapping global forest canopy height through integration of GEDI and Landsat data. Remote. Sens. Environ. 253, 112165.

Potapov, P., Tyukavina, A., Turubanova, S., Talero, Y., Hernandez-Serna, A., Hansen, M. C., Saah, D., Tenneson, K., Poortinga, A., Aekakkararungroj, A., Chishtie, F., Towashiraporn, P., Bhandari, B., Aung, K.S., Nguyen, Q.H., 2019. Annual continuous

fields of woody vegetation structure in the Lower Mekong region from 2000-2017 Landsat time-series. Remote. Sense. Environ. 232, 111278.

Quirós, E., Polo, M.E., Fragoso-Campón, L., 2021. GEDI elevation accuracy assessment: a case study of southwest Spain. IEEE J. Sel. Top. Appl. Earth Obs. Rem. Sens. 14, 5285–5299.

Roy, D.P., Kashongwe, H.B., Armston, J., 2021. The impact of geolocation uncertainty on GEDI tropical forest canopy height estimation and change monitoring. Sci. Remote. Sense. 4, 100024.

Schneider, F.D., Ferraz, A., Hancock, S., Duncanson, L.I., Dubayah, R.O., Pavlick, R.P., Schimel, D.S., 2020. Towards mapping the diversity of canopy structure from space with GEDI. Environ. Res. Lett. 15, 115006.

Silva Junior, C.H.L., Heinrich, V.H.A., Freire, A.T.G., Broggio, I.S., Rosan, T.M., Doblas, J., Anderson, L.O., Rousseau, G.X., Shimabukuro, Y.E., Silva, C.A., House, J. I., Aragão, L.E.O.C., 2020. Benchmark maps of 33 years of secondary forest age for Brazil. Sci. Data 7, 269.

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

Souza, C.M., Shimbo J, Z., Rosa, M.R., Parente, L.L., Alencar A, A., Rudorff, B.F.T., Hasenack, H., Matsumoto, M., Ferreira L, G., Souza-Filho, P.W.M., de Oliveira, S.W., Rocha, W.F., Fonseca, A.V., Marques, C.B., Diniz, C.G., Costa, D., Monteiro, D., Rosa, E.R., Vélez-Martin, E., Weber, E.J., Lenti, F.E.B., Paternost, F.F., Pareyn, F.G. C., Siqueira, J.V., Viera, J.L., Neto, L.C.F., Saraiva, M.M., Sales, M.H., Salgado, M.P. G., Vasconcelos, R., Galano, S., Mesquita, V.V., Azevedo, T., 2020. Reconstructing three decades of land use and land cover changes in Brazilian biomes with Landsat archive and earth engine. Rem. Sens. 12.

UNEP, 2020. Environmental change hotspots - rondônia. In: Division of Early Warning and Assessment (DEWA). United Nations Environment Programme (UNEP). https ://na.unep.net/atlas/webatlas.php?id=2287.

Yu, X., Hyyppä, J., Kaartinen, H., Maltamo, M., 2004. Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. Remote. Sense. Environ. 90, 451–462.