

Article

Spatially Continuous Mapping of Forest Canopy Height in Canada by Combining GEDI and ICESat-2 with PALSAR and Sentinel

Camile Sothe ^{1,*}, Alemu Gonsamo ¹, Ricardo B. Lourenço ¹, Werner A. Kurz ² and James Snider ³

¹ School of Earth, Environment & Society, McMaster University, Hamilton, ON L8S 4L8, Canada

² Canadian Forest Service, Natural Resources Canada, Victoria, BC V8Z 1M5, Canada

³ World Wildlife Fund Canada, Toronto, ON M5V 1S8, Canada

* Correspondence: sothec@mcmaster.ca

Abstract: Continuous large-scale mapping of forest canopy height is crucial for estimating and reporting forest carbon content, analyzing forest degradation and restoration, or to model ecosystem variables such as aboveground biomass. Over the last years, the spaceborne Light Detection and Ranging (LiDAR) sensor specifically designed to acquire forest structure information, Global Ecosystem Dynamics Investigation (GEDI), has been used to extract forest canopy height information over large areas. Yet, GEDI has no spatial coverage for most forested areas in Canada and other high latitude regions. On the other hand, the spaceborne LiDAR called Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2) provides a global coverage but was not specially developed to study forested ecosystems. Nonetheless, both spaceborne LiDAR sensors obtain point-based information, making spatially continuous forest canopy height estimation very challenging. This study compared the performance of both spaceborne LiDAR, GEDI and ICESat-2, combined with ALOS-2/PALSAR-2 and Sentinel-1 and -2 data to produce continuous canopy height maps in Canada for the year 2020. A set-aside dataset and airborne LiDAR (ALS) from a national LiDAR campaign were used for accuracy assessment. Both maps overestimated canopy height in relation to ALS data, but GEDI had a better performance than ICESat-2 with a mean difference (MD) of 0.9 m and 2.9 m, and a root mean square error (RMSE) of 4.2 m and 5.2 m, respectively. However, as both GEDI and ALS have no coverage in most of the hemi-boreal forests, ICESat-2 captures the tall canopy heights expected for these forests better than GEDI. PALSAR-2 HV polarization was the most important covariate to predict canopy height, showing the great potential of L-band in comparison to C-band from Sentinel-1 or optical data from Sentinel-2. The approach proposed here can be used operationally to produce annual canopy height maps for areas that lack GEDI and ICESat-2 coverage.

Keywords: spaceborne LiDAR; Sentinel-2; SAR; random forest

Citation: Sothe, C.; Gonsamo, A.; Lourenço, R.B.; Kurz, W.A.; Snider, J. Spatially Continuous Mapping of Forest Canopy Height in Canada by Combining GEDI and ICESat-2 with PALSAR and Sentinel. *Remote Sens.* **2022**, *14*, 5158. <https://doi.org/10.3390/rs14205158>

Academic Editor: Brian Alan Johnson

Received: 29 August 2022

Accepted: 13 October 2022

Published: 15 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

As one of the largest reservoirs of terrestrial carbon, forests hold approximately 45% of the world's active carbon [1,2], and sequester approximately 32% of anthropogenic emissions every year [3], which make them a key component of the global carbon cycle [4,5]. Accurate estimates of the distribution and total carbon storage in global forests are essential for modelling and monitoring climate change [6,7]. Particularly, aboveground biomass (AGB), that includes all live vegetation above the ground, is listed by the Intergovernmental Panel on Climate Change (IPCC) [8] as one of the most visible and dynamic terrestrial ecosystem carbon pools involving biomass, representing about 30% of the total terrestrial ecosystem carbon pool [9]. The knowledge of spatial distribution and global monitoring of AGB is essential to support the sustainable management of terrestrial ecosystems, improving understanding of the carbon cycle and, consequently, to reduce

carbon emissions [10]. Remote sensing provides a cost-efficient tool for mapping AGB across large regions, nationally and even globally [11].

Because there is no remote sensing sensor able to directly provide AGB information, studies rely on other variables that can be linked to AGB. Among them, canopy height, a variable that can be acquired from light detection and ranging (LiDAR) sensors, is an important predictor for AGB [12–17] as well as can help to monitor ecosystem response to climate, forest degradation, and land-use change as well as restoration [18–21]. Yet, LiDAR data over large, forested areas have been mostly acquired from aircraft, limiting their spatial coverage. Because of high cost, aerial LiDAR campaigns tend towards the acquisitions data over areas of high value forests rather than extensive nationwide or regional observations [22,23].

The new generation of space-based laser altimeters, Ice, Cloud, and Land Elevation Satellite-2 (ICESat-2) and Global Ecosystem Dynamics Investigation (GEDI), launched by NASA in 2018 can directly provide accurate information on vertical vegetation structure globally. When combined with optical or synthetic aperture radar (SAR) data, they allow for large-area estimates of forest structure, including canopy height [24–29], land cover [30], and AGB [31–33]. GEDI is the first spaceborne, full waveform LiDAR specifically created to provide vertical information of forest canopies [28]. Because the ISS orbit is limited between 51.6°N and S latitude, most of the global boreal forests, including in Canada, are not captured by GEDI.

On the other hand, ICESat-2 has no orbital or acquisition-related limitations [34] and provides good coverage in boreal forest areas. The Advanced Topographic Laser Altimeter System (ATLAS), onboard the ICESat-2, is a photon-counting LiDAR instrument with 11 m footprint. It is used to derive the ATL08, a product that provides canopy height percentiles along 100 m segments using returned photons classified as ground, noise, canopy, or top of canopy [35]. Several studies have explored the use of ATL08 to derive canopy height information [17,26,34,36–38], some of them comparing both ICESat-2 and GEDI to estimate canopy height [29,39] or AGB [19]. In general, these studies found GEDI more accurate than ICESat-2 to estimate canopy height. This is mainly related to the fact that ICESat-2, in addition of not being full waveform as GEDI, operates at the green region of the electromagnetic spectrum, with a limited ability to penetrate canopy cover when compared to GEDI.

So far, several studies have demonstrated the use of satellite LiDAR, in particular the Geoscience Laser Altimeter System (GLAS) data onboard the first ICESat mission (2003–2009), to provide large-scale forest height [40,41] and biomass maps [42–44]. Nevertheless, recent studies exploring new spaceborne LiDAR used simulated ATLAS or GEDI data [20,24,45,46]), or validate the data against airborne LiDAR (ALS) without providing wall-to-wall canopy height maps [29,34,36,39]. Because both ALS and spaceborne LiDAR such as GEDI and ICESat-2, are sampling missions, having gaps across tracks and from clouds, spatially continuous remote sensing data, such as from optical or SAR sensors, must be integrated to achieve a full coverage canopy height or AGB products. For example, Qi et al. [17] integrated GEDI data with height information from TanDEM-X InSAR data to obtain a wall-to-all map of AGB in two sites located in the United States and one site in Costa Rica. Francini et al. [31] integrated GEDI and Landsat data to map forest biomass changes due to disturbances in Italy.

Although there have been many studies focused on various forestry applications (e.g., forest AGB) using machine learning, only a small attempt has been made for forest height mapping using the new spaceborne LiDAR GEDI and ICESat-2 and continuous satellite data. Both optical and SAR data bring useful information to extrapolate forest height measurements from LiDAR data. While SAR beams directly interact with trunk and other plant biomass materials [47], providing important information about the physical scattering mechanisms [48], optical data capture both structural (i.e., leaf area index, LAI) and biochemical (e.g., foliage chlorophyll content) attributes that can be used as indicators of plant cover and growth [49].

As neither SAR nor optical data allow direct retrieval of vertical profiles inside forest canopies, they should be combined with other data sources such as LiDAR that provide this information. In this sense, Qi et al. [24] used GEDI to improve height estimates provided by TanDEM-X InSAR in sites located in the United States and Costa Rica. Li et al. [37] proposed integrating ICESat-2 data with Sentinel-1, Sentinel-2 and Landsat-8 data with machine and deep learning methods to map canopy height in Northeast China. Furthermore, in Northeast China, Xi et al. [50] used ATL08 product derived from ICESat-2, Sentinel-1, Sentinel-2 and topographic data, and machine learning methods to obtain canopy height maps for different forest types. At a global scale, Potapov et al. [27] used GEDI canopy height data together with Landsat imagery to create a 30 m spatial resolution global forest canopy height map for the year 2019. However, none of these studies used SAR L-band data from the Advanced Land Observing Satellite Phased Array type L-band SAR (ALOS-2/PALSAR-2). L-band has a deeper penetration through the forest canopy, providing more accurate information on vegetation vertical structure [51]. Emerging data such as ICESat-2 and GEDI require a deep understanding of how forest structure mediates the relationship of canopy height and L-band SAR backscatter. Combining these technologies can bring significant advances in large-scale mapping of forest canopy height.

Canada's forests, the one of the largest contiguous forest ecosystems on Earth, occupies an area of 400 million ha forming an east–west band across eight ecozones in a range of climatic, physiographic, and vegetation conditions [52]. National forest inventory programs are typically designed to produce long-term data for forest monitoring [53–55]. However, most forest inventories and forest mapping efforts are concentrated in managed forest areas of Canada, while northern areas with large tracts of noncommercial forest often lack detailed forest information [56]. In a recent effort, Castilla et al. [56] used optical and SAR data, in addition to the satellite LiDAR, GLAS, associated with field plots and airborne LiDAR data to estimate canopy height, AGB, among other variables in the Northwest Territories, Canada. At a national level, the existing studies used ALS and Landsat data to map forest structure variables, e.g., [57,58], however no study has tested the potential of combining the new satellite LiDAR GEDI and ICESat-2, and SAR data to produce continuous canopy height maps.

GEDI and ICESat-2 are particular in acquiring canopy height information in Canada—the former providing more accurate information of vegetation, yet not reaching full coverage in Canada, whilst the latter is not specifically designed to provide vegetation information but has a global coverage. This is the first study to explore the potential of both spaceborne LiDAR in this scenario. We proposed a straightforward methodology to integrate each ICESat-2 and GEDI, with SAR and optical data from Sentinel-1, Sentinel-2 and ALOS/PALSAR-2, and a machine learning method to provide wall-to-wall forest canopy height maps in forested areas of Canada. The results were validated in two different ways: (1) using a set-aside 30% of the GEDI/ICESat-2 observations serving as the primary validation data set, and (2) measuring the agreement of the produced canopy height maps with the national ALS data. The consistent and continuous monitoring of canopy height provided in this study is essential for estimating forest biomass and carbon stock dynamics, monitoring forest disturbance, and quantifying the effectiveness of forest management initiatives.

2. Materials and Methods

2.1. Study Area

Forested ecosystems of Canada comprise not only trees, but also wetlands and lakes, occupying an area of ~650 million ha [59] widespread across the country in 15 ecozones with different climatic, topographic, and vegetation types [60]. Here, we focused our prediction on treed vegetation and other wooded land in Canada, involving an area of 347 million ha [61]. Treed areas were based on Canada's National Forest Inventory—Treed Land in Canada data (<https://nfi.nfis.org/en/maps#>, accessed on 2 February 2022).

Canada's forests include both deciduous and coniferous tree species, dominated by white spruce (*Picea glauca*), black spruce (*Picea mariana*), balsam fir (*Abies balsamea*), jack pine (*Pinus banksiana*), trembling aspen (*Populus tremuloides*), and balsam poplar (*Populus balsamifera*) [57]. Most forested ecozones are part of the boreal zone, with areas belonging to the hemi-boreal subzone, a transition between temperate and boreal zones [58], extending through the southern part of the boreal zone and great part of the Pacific Maritime and Montane Cordillera ecozones. The main source of disturbances in Canada's forests are wildfire, insects and harvest [62,63]. Fires affected 1.6 million ha annually during the period 1985 to 2010 [64], and 2.25 million ha over the period 1970 to 2017 [65]. The area of forests affected by insects is much larger [5] but with much smaller impacts per hectare than fires or harvest, which affects 0.65 million ha annually [64].

2.2. ICESat-2 Data

Here, we used ATL08 data acquired from the National Snow and Ice Data Center (<https://nsidc.org/data/atl08>, accessed on 20 November 2021). ATL08 is a product derived from the photon-counting laser altimeter ATLAS, onboard ICESat-2 satellite. This laser operates at green region (532 nm) that is divided into three pairs of beams by slightly rotating the spacecraft [66]. The pairs of beams are composed of a strong beam and a weak beam with a strong-to-weak energy ration of approximately 4:1. ATLAS has a high repetition rate, sampling at approximately 0.7 m intervals along the track at a footprint size of approximately 12 m [67]. The ATL08 product provides relative height metrics at segments of 100 m (along-track) \times 12 m (across-track) [39] to ensure that a sufficient number of photons are available for canopy height estimation [37]. These data are composed of three pairs of ground tracks (gt1l-gt3r) with 90 m between the parts of each pair and 3.3 km between pairs.

For canopy height, we used the h_{canopy} (rh98) metric that provides the relative canopy height at 98th percentile of energy return height relative to the ground data. Instead of maximum height, rh98 is recommended to represent top canopy height because of the uncertainty of the signal-to-noise at the top of the canopy [39,68]. We collected data for the mid growing season between 15 June and 15 August 2020, selecting only strong power beams [39]. The parameter multiple scattered warning (msw_flag), derived from ATL09 was used to select segments with no scattering in the atmosphere. We only selected segments with msw_flag equals to 0 which indicates no observed scattering in the atmosphere. Only snow free segments were selected considering the snow flag parameter extracted from the NOAA daily snow cover product [69,70]. A snow_flag with a value of 0 indicates ice-free water, a value of 1 is considered snow-free land, a value of 2 indicates existence of snow, and 3 is presence of ice [68].

The sample dataset was composed of a geometry of points with h_{canopy} information represented by the centroids of the ATL08 segments (100 m \times 12 m). As our grid with covariates has 250 m \times 250 m, some points overlap the same grid tile. To avoid spatial autocorrelation and overfitting, we adopted a point thinning procedure to create a new sample dataset with the h_{canopy} average of points falling within the 250 m grid tile. Basically, this procedure takes the average of neighbor's points within a predefined distance (in this case, 250 m) and aggregates them into a new single. To compute the average, all points were used independent of the land cover type to reduce overestimation of canopy height in 250 m \times 250 m pixels that are not only covered by forest. This resulted in a final sample set of 208,554 points, which was randomly split into 70% for training and 30% for independent validation (Figure 1a).

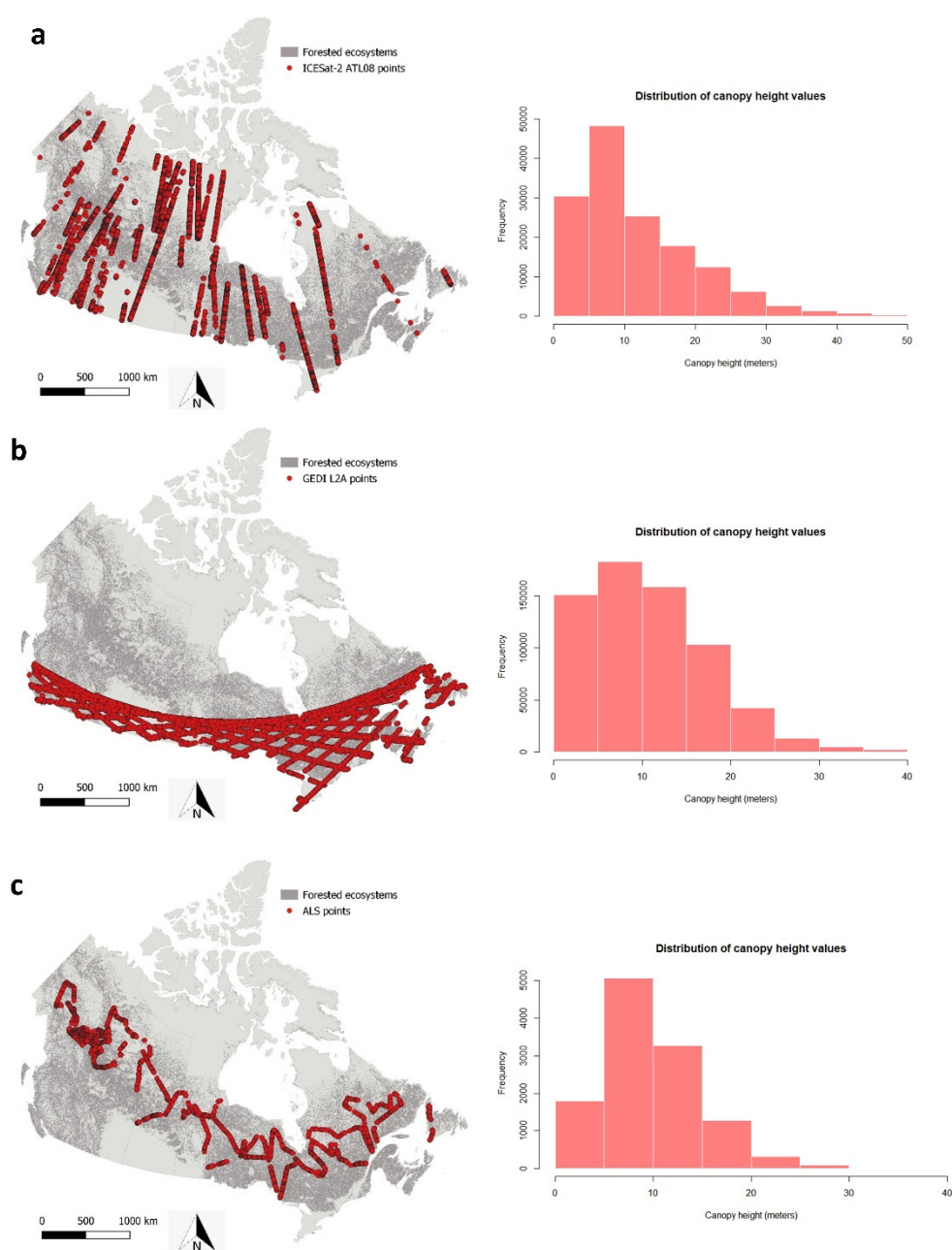


Figure 1. Distribution of LiDAR samples in Canada (**left**) and corresponding histogram with canopy height classes (**right**). (a) ICESat-2 ATL08. (b) GEDI L2A. (c) ALS.

2.3. GEDI Data

In this study, we used the GEDI L2A product data with relative height metrics acquired from NASA's Land Processes Distributed Active Archive Center (LPDAAC, https://e4ftl01.cr.usgs.gov/GEDI/GEDI02_A.001/, accessed on 20 November 2021). GEDI contains three lasers and full waveform recording LiDAR instrument operating at near infrared region (1064 nm). The lasers cover eight beam ground tracks (four coverage and four full power beams), each 600 m apart in the cross-track direction in a track of 4.2 km wide swath. Individual laser shots are spaced 60 m along track and have a circular footprint of approximately 25 m [38,39]. The data were collected for the same period as ICESat-2 (June to August 2020). We only selected full power beams as recommended by Liu et al. [39] and Potapov et al. [27]. For each 25 m footprint, we extracted the relative height metric

rh98 corresponding to 98th percentile of energy return height relative to the ground. The sample dataset was composed of a geometry of points with rh98 information represented by the centroids of the GEDI footprint (25 m).

We filtered the points based on the quality flag criteria, in which a value of 1 indicates the laser shot meets criteria based on energy, sensitivity, amplitude, and real-time surface tracking quality [39]. Since we used covariates at 250 m spatial resolution and GEDI provides data at 25 m footprint, we adopted the same point thinning procedure as ICESat-2 to aggregate points into a 250 m grid, in which the final point was the average of 2 to 8 points, including forest and other land cover types. Afterwards, the resulting 1,249,354 points were randomly split into 70% for training and 30% for independent validation (Figure 1b).

2.4. ALS Data

Acquiring canopy height measurements over large areas through field work is a difficult, expensive, time-consuming, and even dangerous activity, especially in forests with complex structure or situated in remote areas. When available, field measurements generally result in a limited number of sample plots and discontinued in time [71]. In this study, we had no access to field data that could be used to validate the canopy height maps. Therefore, we opted to use ALS data from a national LiDAR campaign to have consistent data across the country avoiding discrepancies upcoming from data collection in different fieldwork or LiDAR campaigns.

ALS data from boreal forest areas were acquired from a campaign conducted between June and August 2010 [22]. The data were collected in 34 flights using a discrete return sensor (Optech ALTM 3100, Teledyne Optech, Toronto, Canada) and resulted in a nominal pulse density of approximately 2.8 returns/m², covering more than 25,000 km of LiDAR transects. The data are available in the shapefile format with canopy height metrics ranging from 25th to 95th percentiles. The shapefiles contain points representing the centroid of each 25 × 25 m LiDAR plot [22]. Plots are provided in blocks of 10 × 10 LiDAR points, corresponding to the spatial extent of 1 km. To avoid having multiple points in the same 250 m × 250 m grid and oversample a pixel, we adopted the same procedure as ICESat-2 and GEDI, taking the average of point values within the same 250 m grid tile. To further validate spaceborne LiDAR data, we adopted the P95 that corresponds to 95th percentile of height, as rh98 is not available in this dataset.

Because of the time gap between ALS and spaceborne data (10 years), we removed the LiDAR points from areas affected by wildfires or harvest in this period in the last 37 years. We assume that after 37 years, tree height changes negligibly although other canopy structures may change. For this, we used Canada's forest change data available at https://opendata.nfis.org/mapserver/nfis-change_eng.html (accessed on 15 March 2022) [64,72,73]. This server provides raster files with forest areas affected by wildfires and harvest between the years 1985 and 2015, classified by change type and change year. After the abovementioned procedures, a total of 11,863 ALS points remained to be used as validation (Figure 1c).

2.5. Sentinel and ALOS-2/PALSAR-2 Data

Table 1 shows the covariates used in this study to estimate canopy height. Sentinel-1, Sentinel-2 and ALOS-2/PALSAR-2 data were acquired from the Google Earth Engine (GEE) platform. We used Sentinel-1 Ground Range Detected (GRD) (European Space Agency, Paris, France) scenes dual-polarization C-band (VV/VH), a collection processed using the Sentinel-1 Toolbox to generate a calibrated, ortho-corrected product. The data are provided in decibel (dB) and 10 m spatial resolution.

Table 1. Summary of covariates used in this study for canopy height modeling using the random forest algorithm.

Group	Band	Description	Spatial Resolution
Optical	S2_B4	ESA Sentinel 2 Multispectral Instrument Surface Reflectance- red and NIR bands (mar-apr, jun-jul, sep-oct)	10 m
	S2_B8		
	S1_VH	ESA Sentinel 1 SAR GRD: C-band Synthetic Aperture Radar Ground Range Detected, log scaling (mar-apr, jun-jul, sep-oct)	10 m
	S1_VV		
SAR	PALSAR_HV	Global ALOS PALSAR-2/PALSAR Yearly Mosaic, converted to decibels (DB)-L-band duo-polarization horizontal transmit/horizontal receive (HH) and horizontal transmit/vertical receive (HV) (annual)	25 m
	PAL-		
	SAR_HH		

From Sentinel-2 MultiSpectral Instrument (MSI) (European Space Agency, Paris, France), we used the Level-2A product which contains the surface reflectance data that were atmospherically corrected by the sen2cor model [74] and available on GEE. The Sentinel-2 image has 12 spectral bands with the spatial resolution ranging from 10 m to 20 m. As we wanted to simplify our approach and not overload the model with optical data, we only used the red and near infrared bands (bands 4 and 8). Information provided by the quality assessment band (QA60) to select only cloud free images.

Sentinel-1 from two polarizations (VV, VH), and Sentinel-2 satellite (two spectral bands, red, near-infrared), were obtained for three periods of the year 2020: March-April, June-July, September-October, in order to capture phenological variations during the growing season. If no cloud-free Sentinel-2 image was available for a certain period, we collected data from the previous year (2019). A median image composition was applied to each band of the Sentinel-1 and -2 images based on all the high-quality observations from the studied periods. Data from both satellites were resampled to 250 m using the pixel aggregate method to match the chosen resolution for the final maps.

The third dataset used is the yearly mosaic ALOS-2/PALSAR-2 L-band containing a co-polarized wave of HH and a cross-polarized wave of HV, with a spatial resolution of 25 m. Polarization data stored as 16-bit digital numbers (DN) were converted into backscatter gamma-naught (γ^0) value (unit: dB) for further analysis. We used Equation (1) in GEE [75]:

$$\gamma^0 = 10 \log^{10} (DN^2) - 83 \text{ dB} \quad (1)$$

The PALSAR-2 data were acquired for the selected year 2020 for all of Canada and resampled to 250 m spatial resolution.

2.6. Predicting Forest Canopy Height

To generate continuous canopy height maps, GEDI or ICESat-2 point sampling data were matched to contiguous data from the abovementioned optical and SAR sensors. For this, we overlaid the point sampling of each sensor with 14 predictor variables (covariates), Sentinel-1 (VV, VH) and Sentinel-2 (red, near-infrared) in three seasons and annual ALOS-2/PALSAR-2 (HH, HV), composing the regression matrix.

As mentioned above, we split the data of each sensor into training (70%), to train the model, and test (30%), used as independent data for accuracy assessment. For modelling, we used the random forest (RF) algorithm, a machine learning method proposed by Breiman [76] that operates as an ensemble of decision trees trained upon random subsets of the available labeled samples and predictors. In regression, each model grows with the number of trees for minimizing the prediction variance. We chose this method because of its ability to deal with multisource data and analyzing complex non-linear and possibly hierarchical interactions in large datasets [77,78], and for providing a ranking with the importance of each variable used in the regression, which allow us to make assumptions about the relationship between response and predictor variables.

To train the model, we defined the two RF parameters—the number of trees to build the “forest” (ntree) and the number of predictors considered for each node in the trees (mtry)—as follows: we chose 500 for ntree since no difference in accuracy was noticed between 500 and 1000, while the mtry parameter was set to 3, after testing values between 2 and 6 starting with the standard mtry recommended for regression, which is the number of variables divided by 3. The model fitting was executed using the ranger package [79] in R programming [80]. Ranger is an implementation of RF with internal parallelization that is particularly appropriate for high dimensional data [79].

For spatial prediction, the 14 covariates originally composing stacked raster files covering all Canada’s forests, were first converted to data frames so that each row corresponds to a 250 m cell grid and each column is a covariate. Due to the large size, the data were split into 400×400 km tiles and the trained model was applied tile-by-tile, where all tiles with predicted canopy height values were mosaicked at the end and exported in tiff format for the final map.

To evaluate the importance of predictor variables, we used the RF internal sampling procedure called out-of-bag (OOB) and Mean Squared Error (MSE). To rank the variables, RF omits one by one from the predictor variable list and, depending on how much this increases the OOB error, the variable is considered more or less important [81].

2.7. Accuracy Assessment

For accuracy assessment, we used common evaluation metrics such as r-square (R^2), a relative measure of fit, root mean square error (RMSE) to know how far apart the predicted heights are from the observed heights, and mean absolute error (MAE), that captures absolute difference between two height metrics. We also computed the mean difference (MD) or bias, which represents how much observed values systematically differ from predicted values, with negative values signifying underestimation and positive values signifying overestimation [38].

The evaluation metrics were computed assessed with two datasets: (1) a set-aside 30% of the GEDI/ICESat-2 observations served as the primary validation data set, and (2) measuring the agreement of the produced canopy height maps with the national ALS data.

3. Results

3.1. Forest Canopy Height Model Accuracy and Variable Importance

When the ICESat-2 model was evaluated using the set-aside dataset, we achieved an R^2 of 0.60, RMSE of 5.4 m and MAE of 3.4 m (Figure 2a), while a decrease in R^2 but also in RMSE was observed when ICESat-2 was validated against ALS (i.e., R^2 of 0.36 and RMSE of 5.2 m) (Figure 2b). For the GEDI set-aside dataset, we achieved an R^2 of 0.58, RMSE of 4.3 m and MAE of 2.9 m (Figure 2c). Again, when GEDI was validated against the ALS data, we obtained a R^2 of 0.37 and a slight decrease in the RMSE (i.e., 4.2 m) (Figure 2d). In both situations, the model tends to overestimate tall trees (>20 m) and underestimate the small ones (<5 m) (Figure 2). The MD between GEDI canopy height map and ALS data was 0.9 m, while for ICESat-2 it was 2.9 m, showing that both spaceborne LiDAR-derived products overestimate canopy height in relation to ALS with GEDI closer to ALS measurements.

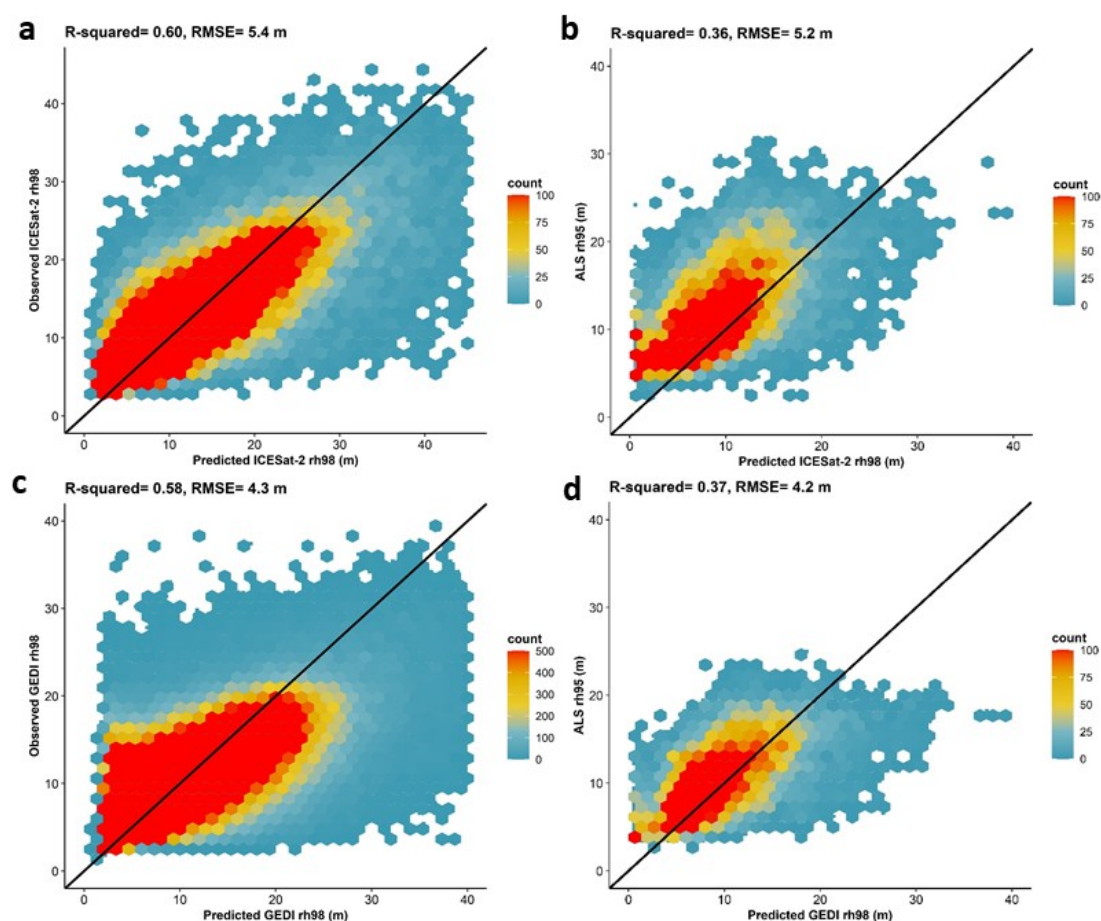


Figure 2. Relationship between the measured and estimated canopy height values using an independent validation assessment with GEDI and ICESat-2 sample set (a,c) and a random forest algorithm, and the resulting canopy height maps compared to ALS point data (b,d). Point density is indicated with a blue (low-density regions) to red (high-density regions) color gradient. The black line represents a 1:1 linear fit line. (a) Internal validation using ICESat-2 ATL08 modeled using 14 covariates and 145,987 points, and 62,566 points used for test. (b) Validation using 18,090 ALS points versus ICESat-2 canopy height map. (c) Internal validation using GEDI L2A modeled using 14 covariates and 874,548 points, and 374,806 points used in the test. (d) Validation using 18,090 ALS points versus GEDI canopy height map.

The ranking with variable importance provided by RF showed that both SAR bands from ALOS-2/PALSAR-2 were the most important information to predict canopy height using both ICESat-2 and GEDI data (Figure 3). In fact, ALOS-2/PALSAR-2 HV stood out over other covariates to predict canopy height representing alone almost 20% of relative importance. SAR data were also in second and third places of the ranking when using GEDI, whilst the two bands of Sentinel-2 from March–April period were important for ICESat-2. In both situations, cross polarized data from both SAR sensors PALSAR-2 (HV) and Sentinel-1 (VH) were more important than co-polarized (HH, VV).

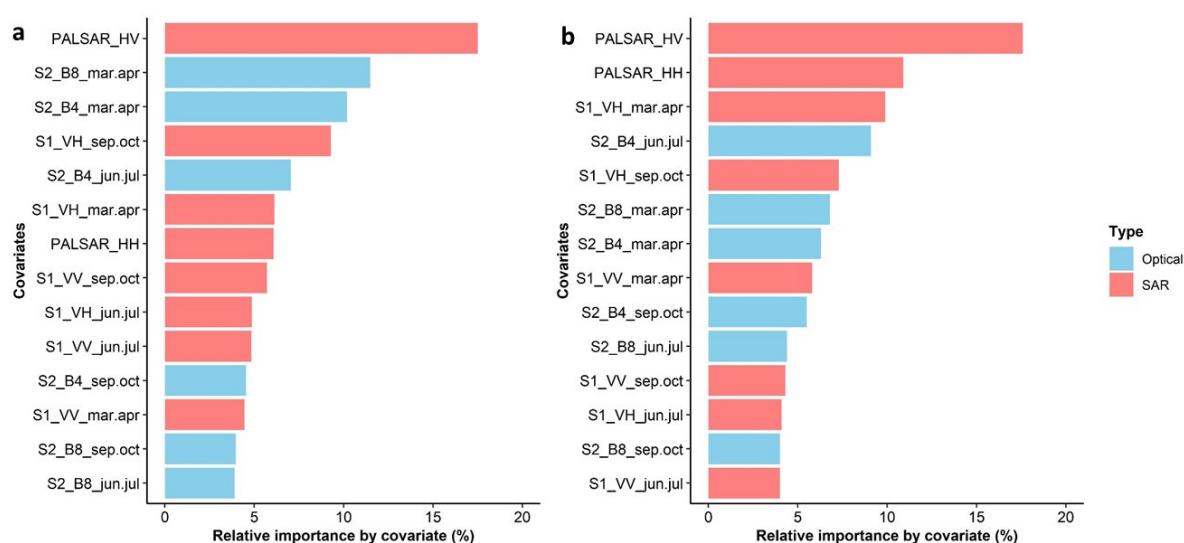


Figure 3. Ranking of variable importance generated by the random forest algorithm. (a) Ranking with relative importance of covariates when modeling canopy height using ICESat-2 data. (b) Ranking relative importance of covariates when modeling canopy height using GEDI data.

3.2. Spatial Distribution of Canopy Height

Figure 4 presents the maps with the spatial distribution of canopy height derived from ICESat-2 (Figure 4a) and GEDI (Figure 4c), and the difference between both (Figure 4e). In general, the maps agree in showing a decrease in canopy height with increasing latitudes, and larger canopy height values in forests of the Pacific Maritime and Montane Cordillera ecozones. The map derived from ICESat-2 shows higher canopy height compared to the GEDI map, as can be seen in the histograms of Figure 4b,d,f. For instance, we notice that in hemi-boreal forests, most trees are above 30 m in the ICESat-2 map, while GEDI falls in a range between 25 m and 30 m. The same trend is observed in the Atlantic Maritime ecozone and all the areas comprising hemi-boreal forests—ICESat-2 resulting in higher canopy height values than GEDI. This is consistent with the accuracy results observed from ICESat-2, with an overestimation of 2.9 m in relation to ALS data, while only 0.9 m observation was observed in the GEDI map (Figure 2). However, it is worth mentioning that fewer GEDI samples cover hemi-boreal forests while no ALS measurements cover those areas, where a forest canopy height taller than 30 m is expected. The map in Figure 4e shows that the difference between ICESat-2 and GEDI is more pronounced in areas with high canopy height values in the Pacific Coast with a general overestimation of ICESat-2 over GEDI. Although the difference between the two maps is small, with most histogram values around 0 (Figure 4f), there are some pixels that differ by more than 10 m.

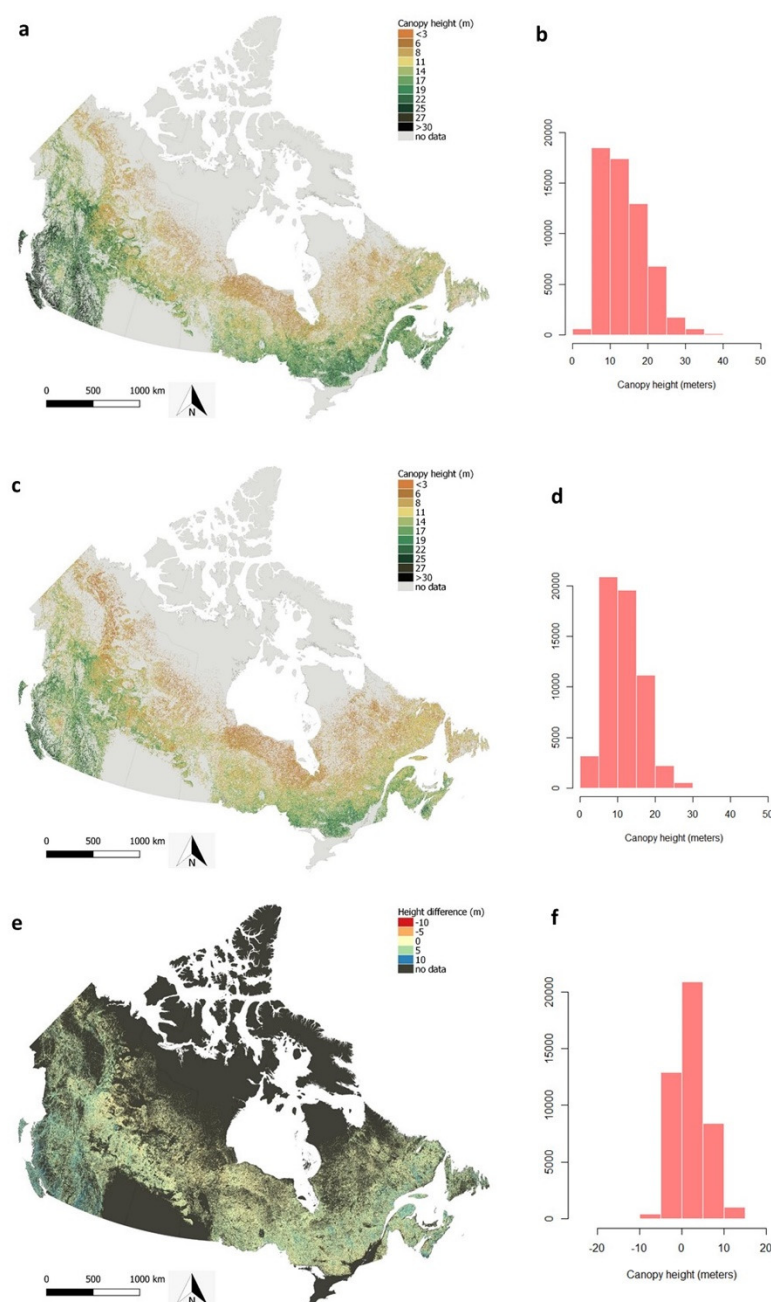


Figure 4. Predicted canopy height in forested ecosystems of Canada using spaceborne LiDAR associated with 14 covariates and a random forest algorithm. (a) Canopy height map from ICESat-2. (b) Histogram with distribution of canopy height values using ICESat-2. (c) Canopy height map for GEDI. (d) Histogram with distribution of canopy height values using GEDI. (e) Difference between ICESat-2 and GEDI canopy height map. (f) Histogram with distribution of the difference between ICESat-2 and GEDI canopy height values.

In Figure 5, we can see the histograms with the distribution of canopy height per ecozone using the GEDI map. As expected, higher canopy height values are predominantly located in the Pacific Maritime in British Columbia, whilst shorter trees are predominant in northern ecozones such as Hudson Plain, Taiga Plain and Taiga Cordillera.

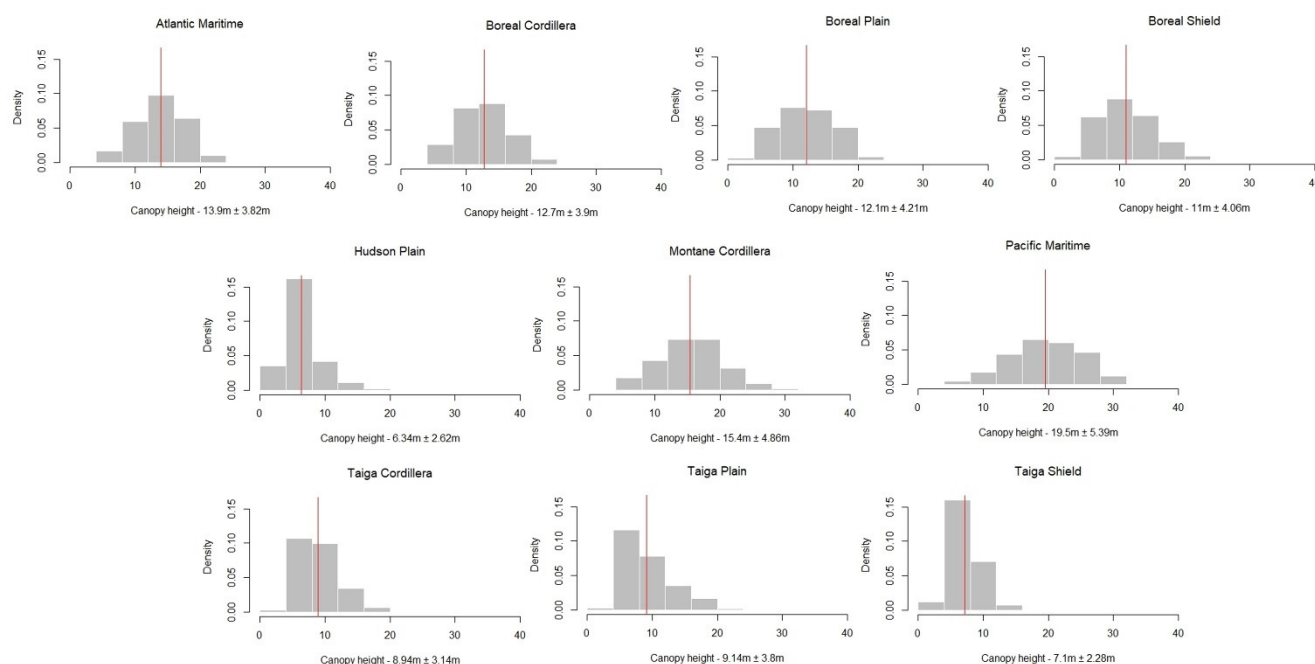


Figure 5. Distribution of canopy height values for Canada's ecozone according to canopy height map produced from GEDI data. Note: red line indicates canopy height average; mean \pm standard deviation is described in the bottom of each histogram.

In Figure 6 (zoom 1), the map subset derived from GEDI demonstrates canopy height variation related to natural forest types, clearly separating tall forests in the coastal area of British Columbia from shorter canopy height forests in mountains areas. Figure 6, zoom 2, shows shorter forests as expected for peatlands in the Hudson Bay Lowlands.

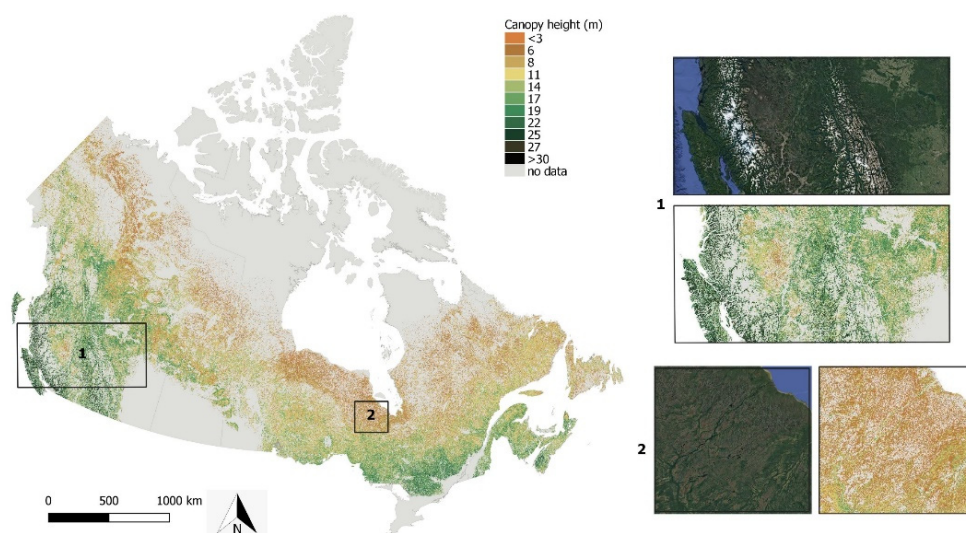


Figure 6. Forest canopy height map for the year 2020 produced through the integration of GEDI data (June–August 2020) and covariates derived from Sentinel and PALSAR-2. Close-up examples with google satellite images illustrating (1) forests in the Pacific Maritime and Montane Cordillera ecozones, (2) peatlands in the Hudson Plain ecozone.

4. Discussion

4.1. Comparing the Accuracy of ICESat-2 and GEDI to Predict Canopy Height

Using ICESat-2 and GEDI spaceborne LiDAR combined with Sentinel and ALOS-2/PALSAR-2 we were able to produce spatially continuous canopy height maps in forested ecosystems of Canada. Even with many areas beyond GEDI coverage, we achieve a RMSE of 4.2 m using this sensor, while a RMSE of 5.2 m was observed using ICESat-2 that reach full coverage in Canada. When producing a global canopy height map using GEDI and Landsat imagery, Potapov et al. [27] reported a RMSE of 9.07 m and a MD of -3.8 m when the model was compared to ALS data used for validation, indicating an underestimation of forest height. Our study shows a MD of less than 1 m when using GEDI model in relation to ALS data (0.9 m) and a RMSE more than 50% lower than the global study (4.2 m). ICESat-2 showed a MD of 2.9 m, which is a much greater overestimation than GEDI in relation to ALS data. When comparing GEDI and ICESat-2 for canopy height retrieval, Liu et al. [39] reported that GEDI outperforms ICESat-2, as ICESat-2 was more prone to bias for almost all forest types and cover conditions. When using all data pairs regardless of beam intensity and acquisition time, the authors achieved overall RMSEs of 7.01 m and 5.02 m for ICESat-2 and GEDI, respectively, while using only strong/power beam data acquired at night, resulted in a decrease in RMSEs to 3.93 m and 3.56 m. Neuenschwander et al. [36] used ATL08 from ICESat-2 to derive canopy height in boreal forests in Finland with an underestimation of 0.56 m with strong beam at night in summer. In this study, we used strong/power beam and summer season, but we included day and nighttime data to reach better coverage.

Our findings suggest that canopy height maps from both spaceborne LiDAR tended to overestimate smaller trees (<5 m) and underestimate taller trees (>20 m) and a MD suggesting overestimation (0.9 m from GEDI and 2.9 m from ICESat-2) in comparison to ALS data (Figure 2). In contrast, Neuenschwander et al. [36] and Malambo and Popescu [38] found that ICESat-2 tends to underestimate canopy height, yet they evaluated the sensor at the point level and not associated with other covariates and machine learning method as here. When measuring the agreement between multiple canopy and terrain height estimates from the ATL08 product and the corresponding height estimates from ALS in 12 sites across 6 biomes in the US, Malambo and Popescu [38] found the MD between ICESat-2 and ALS canopy height estimates of -1.71. In a global study using GEDI, Potapov et al. [27] reported that GEDI tended to underestimate shorter canopies (<3 m) and tall forests (>21 m height), while Liu et al. [39] reported that ICESat-2 tends to overestimate the canopy height of dwarf shrublands and underestimate the canopy height of forest. However, Liu et al. [39] argued that the ALS data used for accuracy assessment mainly sample tall forests, while in our study, the ALS data were acquired 10 years before the spaceborne LiDAR data and only involve boreal forest, not including tall canopies commonly found in Pacific Maritime and Montane Cordillera ecozones. We can summarize three main reasons that both sensors overestimate canopy height in relation to ALS: (a) forest growth and history of disturbance during the 10-years period, (b) ALS data do not cover the tallest forests of the country, (c) from ALS data, we used rh95 while for spaceborne LiDAR, rh98.

Our study covers 347 million ha, the R^2 we obtained using GEDI/ICESat-2 test data (i.e., 0.58 and 0.60) is within published range of previous works that was mostly conducted at landscape or smaller study areas [26,37,82]. ICESat-2 and GEDI reached similar R^2 using the set aside validation and ALS data. However, in both situations, a difference of 1 m in RMSE is observed (e.g., ~5.2 m ICESat-2, compared to ~4.2 m GEDI). In their study, Malambo and Popescu [38] recommended the use of metrics like MD (bias) and MAE, since ATL08 from ICESat-2 and ALS may have outlying estimates which could distort the association between the measurements and provide a distorted view of the agreement. In fact, although the validation using ALS data resulted in a R^2 much lower for both ICESat-2 and GEDI than when using set aside data, the RMSEs were also slightly lower. According to Potapov et al. [27], tree-based models applied over large areas performed less well

than locally calibrated models. At a smaller scale, promising results were shown by Li et al. [37] when using ICESat-2 data associated with Sentinel-1, Sentinel-2 and Landsat-8 data. Using an independent test dataset, they reached an R^2 of 0.78 and 0.68, and RMSE of 2.64 m and 2.93 m when using a deep learning and RF model, respectively.

4.2. Integrating Spaceborne LiDAR and Other Sensors for Continuous Canopy Height Mapping

Many of the existing studies evaluated ICESat-2 or GEDI at the plot level as they did not provide continuous canopy height maps (e.g., [37,38]) or using GEDI with optical data from Landsat and regression models to produce canopy height maps (e.g., [27]). Li et al. [37] is one of the few existing studies associating ICESat-2 with both optical and SAR to produce canopy height maps. In fact, they reported that backscattering coefficients from Sentinel-1 together with the red-edge related variables from Sentinel-2 positively contribute to the prediction of forest canopy height. In our study, PALSAR-2 was ranked as the best variable for both ICESat-2 and GEDI.

While optical sensors do not provide 3D information from forest, the short wavelength C-band data provided by Sentinel-1 only interacts with top canopy leaves. L-band is more sensitive to forest biomass and height estimation than C-band because its longer wavelength penetrates forest canopy capturing backscattering from branches and trunk [83,84]. Its high sensitivity to the forest vertical structure makes it advantageous in extracting forest canopy height. Yet, data from PALSAR-2 have not been sufficiently explored so far for canopy height retrieval and its potential is shown in this study. Cross polarized data from both SAR sensors, PALSAR-2 (HV) and Sentinel-1 (VH), were more important than co-polarized because they indicate the volumetric scattering from vegetation related to tree leaves and branches [85], while co-polarized as HH is mainly indicative of double bounce scattering associated with tree trunks, buildings or inundated vegetation. Upcoming satellite missions such as the NASA-ISRO Synthetic Aperture Radar (NISAR) will collect data in L and S frequencies which can greatly contribute to canopy height and AGB estimation.

4.3. Spatial Distribution of Canopy Height in Canada and Study Limitations

The spatial distribution of forest canopy height observed in this study is in line with existing maps covering Canadian territory. When estimating forest structure using ALS and Landsat, Matasci et al. [58] reported higher canopy height values in the Montane Cordillera and Pacific Maritime ecozones, with median canopy height values of approximately 20 m, which is similar to values of this study (Figure 5). They found the lowest canopy height values in the Taiga Shield West and Taiga Cordillera, while in addition to those areas, our study also shows lower canopy height values in Hudson Plain, and higher values in Atlantic Maritime ecozone. Our maps also show similarities with the Global Forest Height 2019 [27], with higher canopy height values in the hemi-boreal forest areas and southern parts of boreal forest. However, this study and none of the abovementioned studies directly compared the results with field measurements, it is not possible to say which one is more accurate in terms of canopy height estimates.

GEDI map showed a better result in terms of accuracy metrics. Nevertheless, as neither GEDI nor ALS (used to validate that models) have a good coverage in some ecozones as the Pacific Maritime, they may underestimate tree heights in those areas. For instance, Pacific Maritime showed an average canopy height of 19.5 m when using the map produced with GEDI data (Figure 5), whilst an average of 25 m was observed with ICESat-2 canopy height map, which is more realistic for that ecozone. This can also be related to lower spatial resolution adopted in this study (250 m). Within these ~6 ha of forest, we are likely to have a wide range of canopy heights as there will be stands of different ages and disturbance histories that were all aggregated into one pixel.

Considering that all covariates used here have a spatial resolution higher or equal to 30 m, the maps can be produced at finer scale and be used for local and regional scale analyses. Yet, Xi et al. [50] tested ICESat-2 associated with Sentinel data at 10, 30, and 250

m spatial resolution and obtained the best result at 250 m. Nevertheless, the canopy height maps produced here can be used as a predictor variable to estimate AGB or for other national scale analyses: to detect forest disturbances, to monitor changes in AGB stocks resulting from forestry operations and land conversions, to monitor the effects of forest degradation, to evaluate the progress of forest restoration projects, or to model key ecosystem variables such as primary production and biodiversity [27].

5. Conclusions

This study compared data from two new spaceborne LiDAR sensors, ICESat-2 and GEDI, to estimate canopy height in forested ecosystems of Canada. LiDAR data were associated with Sentinel-1 and -2 and ALOS-2/PALSAR-2 to produce continuous canopy height maps at 250 m resolution. In terms of RMSE, GEDI had an error of 4.2 m, which is 1 m lower than observed for ICESat-2. However, as GEDI has no full coverage in Canada, hemi-boreal and northern forests are underrepresented in the model. Visually, the map resulted from ICESat-2 represented very well the expected higher canopy heights in the Pacific Maritime and Montane Cordillera ecozones. The methodology applied here using both sensors overestimate canopy height in relation to national ALS data, with GEDI again performing better than ICESat-2, with a MD of 0.9 m in comparison to 2.9 of ICESat-2. Nevertheless, there is a 10 years-time gap between the data compiled from ALS and spaceborne sensors, and the fact that ALS only covers boreal forest areas, can be the reasons for overestimation. In fact, since no comparison with current ALS data or field measurements were made, it is difficult to state which sensor was more precise. For future studies, we recommend testing the potential of GEDI and ICESat-2 with new ALS data taken at the provincial level.

PALSAR-2 HV polarization was the most important covariate to predict canopy height, showing the great potential of L-band in comparison to C-band from Sentinel-1 or optical data. The fact that we used only a few and open access covariates, make the approach used here suitable for operational continuous canopy height mapping. To the best of our knowledge, there is no study associating GEDI or ICESat-2 with PALSAR-2, and this study highlighted the potential of this combination. PALSAR-2 is only available as annual mosaic which makes the approach proposed here better for canopy height monitoring at annual or longer observation periods.

Despite these caveats, this study shows that both GEDI and ICESat-2 can be used for large scale continuous mapping of canopy height in Canada or other high latitude regions. Canopy height is one of the main indicators of AGB that can be obtained with remote sensing observations, and the maps can be used to produce operationally updated integrated forest structure and change maps to inform climate mitigation policy initiatives. The next steps include producing annual forest canopy height estimates at higher resolution to quantify AGB and to monitor carbon stock changes from forestry operations, disturbances and land conversions, in addition to using more recent provincial ALS data to validate canopy height maps.

Author Contributions: Conceptualization, C.S. and A.G.; methodology, C.S. and A.G.; software, C.S. and R.B.L.; validation, C.S. and R.B.L.; formal analysis, C.S. and A.G.; investigation, C.S.; resources, A.G. and J.S.; data curation, C.S. and R.B.L.; writing—original draft preparation, C.S.; writing—review and editing, C.S., A.G., R.B.L., W.A.K. and J.S.; visualization, C.S.; supervision, A.G. and W.A.K.; project administration, C.S. and A.G.; funding acquisition, A.G. and J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by World Wildlife Fund Canada (WWF-Canada) and NSERC Alliance grants (application No: ALLRP 566310-21). A.G. acknowledges funding from NSERC Discovery Grant (RGPIN-2020-05,708) and the Canada Research Chairs Program.

Data Availability Statement: The data that support the findings of this study are available from the corresponding author, C.S., upon reasonable request.

Acknowledgments: This work was made possible by the facilities of the Shared Hierarchical Academic Research Computing Network (SHARCNET) and the Digital Research Alliance of Canada (Alliance de recherche numérique du Canada).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Houghton, R.A.; Hall, F.; Goetz, S.J. Importance of biomass in the global carbon cycle. *J. Geophys. Res.-Biogeosciences* **2009**, *114*, G00E03. <https://doi.org/10.1029/2009JG000935>.
2. Bonan, G.B. Forests and climate change: Forcings, feedbacks, and the climate benefits of forests. *Science* **2008**, *320*, 1444–1449.
3. Le Quéré, C.; Andrew, R.M.; Friedlingstein, P.; Sitch, S.; Pongratz, J.; Manning, A.C.; Korsbakken, J.I.; Peters, G.P.; Canadell, J.G.; Jackson, R.B.; et al. Global carbon budget 2017. *Earth Syst. Sci. Data* **2018**, *10*, 405–448. <https://doi.org/10.5194/essd-10-405-2018>.
4. Houghton, R.A. Aboveground forest biomass and the global carbon balance. *Glob. Chang. Biol.* **2005**, *11*, 945–958. <https://doi.org/10.1111/j.1365-2486.2005.00955.x>.
5. Kurz, W.A.; Shaw, C.H.; Boisvenue, C.; Stinson, G.; Metsaranta, J.; Leckie, D.; Dyk, A.; Smyth, C.; Neilson, E.T. Carbon in Canada's boreal forest—A synthesis. *Environ. Rev.* **2013**, *21*, 260–292. <https://doi.org/10.1139/er-2013-0041>.
6. Mitchard, E.T.A.; Feldpausch, T.R.; Brienen, R.J.W.; Lopez-Gonzalez, G.; Monteagudo, A.; Baker, T.R.; Lewis, S.L.; Lloyd, J.; Quesada, C.A.; Gloor, M.; et al. Markedly divergent estimates of Amazon forest carbon density from ground plots and satellites. *Glob. Ecol. Biogeogr.* **2014**, *23*, 935–946. <https://doi.org/10.1111/geb.12168>.
7. Stovall, A.E.L.; Fatoyinbo, T.; Thomas, N.M.; Armston, J.; Ebanega, M.O.; Simard, M.; Trettin, C.; Zogo, R.V.O.; Aken, I.A.; Debina, M.; et al. Comprehensive comparison of airborne and spaceborne SAR and LiDAR estimates of forest structure in the tallest mangrove forest on earth. *Sci. Remote Sens.* **2021**, *4*, 100034. <https://doi.org/10.1016/j.srs.2021.100034>.
8. Eggleston, H.S.; Buendia, L.; Miwa, K. Prepared by the National Greenhouse Gas Inventories Programme. In *2006 IPCC Guidelines for National Greenhouse Gas Inventories*; IGES: Kanagawa, Japan, 2006; pp. 27–43.
9. Kumar, L.; Mutanga, O. Remote sensing of above-ground biomass. *Remote Sens.* **2017**, *9*, 935. <https://doi.org/10.3390/rs9090935>.
10. Rodríguez-Veiga, P.; Wheeler, J.; Louis, V.; Tansey, K.; Balzter, H. Quantifying forest biomass carbon stocks from space. *Curr. For. Rep.* **2017**, *3*, 1–18. <https://doi.org/10.1007/s40725-017-0052-5>.
11. Brewer, C.K.; Monty, J.; Johnson, A.; Evans, D.; Fisk, H. *Forest Carbon Monitoring: A Review of Selected Remote Sensing and Carbon Measurement Tools for REDD+*. RSAC RPT1. United States Department of Agriculture Forest Service: Salt Lake City, UT, USA, 2012.
12. Dubayah, R.O.; Sheldon, S.; Clark, D.B.; Hofton, M.A.; Blair, J.B.; Hurtt, G.C.; Chazdo, R.L. Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica. *J. Geophys. Res. Biogeosci.* **2010**, *115*. <https://doi.org/10.1029/2009JG000933>.
13. Asner, G.P.; Mascaro, J. Mapping tropical forest carbon: Calibrating plot estimates to a simple lidar metric. *Remote Sens. Environ.* **2014**, *140*, 614–624. <https://doi.org/10.1016/j.rse.2013.09.023>.
14. Jucker, T.; Caspersen, J.; Chave, J.; Antin, C.; Barbier, N.; Bongers, F.; Dalponte, M.; van Ewijk, K.Y.; Forrester, D.I.; Haeni, M.; et al. Allometric equations for integrating remote sensing imagery into forest monitoring programmes. *Glob. Change Biol.* **2017**, *23*, 177–190.
15. Carreiras, J.M.; Quegan, S.; Le Toan, T.; Minh, D.H.T.; Saatchi, S.S.; Carvalhais, N.; Reichstein, M.; Scipal, K. Coverage of high biomass forests by the ESA BIOMASS mission under defense restrictions. *Remote Sens. Environ.* **2017**, *196*, 154–162. <https://doi.org/10.1016/j.rse.2017.05.003>.
16. Silva, C.A.; Saatchi, S.; Garcia, M.; Labriere, N.; Klauber, C.; Ferraz, A.; Meyer, V.; Jeffery, K.J.; Abernethy, K.; White, L.; et al. Comparison of small- and large-footprint lidar characterization of tropical forest aboveground structure and biomass: A case study from Central Gabon. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2018**, *11*, 3512–3526. <https://doi.org/10.1109/JSTARS.2018.2816962>.
17. Qi, W.; Saarela, S.; Armston, J.; Ståhl, G.; Dubayah, R.O. Forest biomass estimation over three distinct forest types using TanDEM-X InSAR data and simulated GEDI lidar data. *Remote Sens. Environ.* **2019**, *232*, 111283. <https://doi.org/10.1016/j.rse.2019.111283>.
18. Asner, G.P.; Mascaro, J.; Muller-Landau, H.-C.; Vieilledent, G.; Vaudry, R.; Rasamoelina, M.; Hall, J.S.; Van Breugel, M. A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia* **2012**, *168*, 1147–1160. <https://doi.org/10.1007/s00442-011-2165-z>.
19. Schneider, F.D.; Morsdorf, F.; Schmid, B.; Petchey, O.L.; Hueni, A.; Schimel, D.S.; Schaepman, M.E. Mapping functional diversity from remotely sensed morphological and physiological forest traits. *Nat. Commun.* **2017**, *8*, 1441. <https://doi.org/10.1038/s41467-017-01530-3>.
20. Duncanson, L.; Neuenschwander, A.; Hancock, S.; Thomas, N.; Fatoyinbo, T.; Simard, M.; Silva, C.A.; Armston, J.; Luthcke, S.B.; Hofton, M.; et al. Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County, California. *Remote Sens. Environ.* **2020**, *242*, 111779. <https://doi.org/10.1016/j.rse.2020.111779>.

21. Valbuena, R.; O'Connor, B.; Zellweger, F.; Simonson, W.; Vihervaara, P.; Maltamo, M.; Silva, C.; Almeida, D.; Danks, F.; Morsdorf, F. et al. Standardizing ecosystem morphological traits from 3d information sources. *Trends Ecol. Evolut.* **2020**, *35*, 656–667. <https://doi.org/10.1016/j.tree.2020.03.006>.
22. Wulder, M.A.; White, J.C.; Nelson, R.F.; Næsset, E.; Ørka, H.O.; Coops, N.C.; Hilker, T.; Bater, C.W.; Gobakken, T. Lidar sampling for large-area forest characterization: A review. *Remote Sens. Environ.* **2012**, *121*, 196–209. <https://doi.org/10.1016/j.rse.2012.02.001>.
23. Tompalski, P.; Coops, N.C.; White, J.C.; Goodbody, T.R.; Hennigar, C.R.; Wulder, M.A.; Socha, J.; Woods, M.E. Estimating Changes in Forest Attributes and Enhancing Growth Projections: A Review of Existing Approaches and Future Directions Using Airborne 3D Point Cloud Data. *Curr. For. Rep.* **2021**, *7*, 1–24. <https://doi.org/10.1007/s40725-021-00135-w>.
24. Qi, W.; Lee, S.-K.; Hancock, S.; Luthcke, S.; Tang, H.; Armston, J.; Dubayah, R. Improved forest height estimation by fusion of simulated GEDI Lidar data and TanDEM-X InSAR data. *Remote Sens. Environ.* **2019**, *221*, 621–634. <https://doi.org/10.1016/j.rse.2018.11.035>.
25. Healey, S.P.; Yang, Z.; Gorelick, N.; Ilyushchenko, S. Highly local model calibration with a new GEDI LiDAR asset on Google Earth Engine reduces Landsat forest height signal saturation. *Remote Sens.* **2020**, *12*, 2840. <https://doi.org/10.3390/rs12172840>.
26. Lin, X.; Xu, M.; Cao, C.; Dang, Y.; Bashir, B.; Xie, B.; Huang, Z. Estimates of Forest Canopy Height Using a Combination of ICESat-2/ATLAS Data and Stereo-Photogrammetry. *Remote Sens.* **2020**, *12*, 3649. <https://doi.org/10.3390/rs12213649>.
27. Potapov, P.; Li, X.; Hernandez-Serna, A.; Tyukavina, A.; Hansen, M.C.; Kommareddy, A.; Pickens, A.; Turubanova, S.; Tang, H.; Silva, C.E.; et al. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sens. Environ.* **2021**, *253*, 112165. <https://doi.org/10.1016/j.rse.2020.112165>.
28. Lang, N.; Kalischek, N.; Armston, J.; Schindler, K.; Dubayah, R.; Wegner, J.D. Global canopy height regression and uncertainty estimation from GEDI LIDAR waveforms with deep ensembles. *Remote Sens. Environ.* **2022**, *268*, 112760. <https://doi.org/10.1016/j.rse.2021.112760>.
29. Milenković, M.; Reiche, J.; Armston, J.; Neuenschwander, A.; Keersmaecker, W.D.; Herold, M.; Verbesselt, J. Assessing Amazon rainforest regrowth with GEDI and ICESat-2 data. *Sci. Remote Sens.* **2022**, *5*, 100051. <https://doi.org/10.1016/j.srs.2022.100051>.
30. Tang, H.; Armston, J.; Hancock, S.; Marselis, S.; Goetz, S.; Dubayah, R. Characterizing global forest canopy cover distribution using spaceborne lidar. *Remote Sens. Environ.* **2019**, *231*, 111262. <https://doi.org/10.1016/j.rse.2019.111262>.
31. Francini, S.; D'Amico, G.; Vangi, E.; Borghi, C.; Chirici, G. Integrating GEDI and Landsat: Spaceborne Lidar and Four Decades of Optical Imagery for the Analysis of Forest Disturbances and Biomass Changes in Italy. *Sensors* **2022**, *22*, 2015. <https://doi.org/10.3390/s22052015>.
32. Saarela, S.; Holm, S.; Healey, S.P.; Patterson, P.L.; Yang, Z.; Andersen, H.-E.; Dubayah, R.O.; Qi, W.; Duncanson, L.I.; Armston, J.D.; et al. Comparing frameworks for biomass prediction for the Global Ecosystem Dynamics Investigation. *Remote Sens. Environ.* **2022**, *278*, 113074. <https://doi.org/10.1016/j.rse.2022.113074>.
33. Sothe, C.; Gonsamo, A.; Arabian, J.; Kurz, W.A.; Finkelstein, S.A.; Snider, J. Large soil carbon storage in terrestrial ecosystems of Canada. *Glob. Biogeochem. Cycles* **2022**, *36*, e2021GB007213. <https://doi.org/10.1029/2021GB007213>.
34. Mulverhill, C.; Coops, N.C.; Hermosilla, T.; White, J.C.; Wulder, M.A. Evaluating ICESat-2 for monitoring, modeling, and update of large area forest canopy height products. *Remote Sens. Environ.* **2022**, *271*, 112919. <https://doi.org/10.1016/j.rse.2022.112919>.
35. Neuenschwander, A.; Pitts, K. The ATL08 land and vegetation product for the ICESat-2 mission. *Remote Sens. Environ.* **2019**, *221*, 247–259. <https://doi.org/10.1016/j.rse.2018.11.005>.
36. Neuenschwander, A.; Guenther, E.; White, J.C.; Duncanson, L.; Montesano, P. Validation of ICESat-2 terrain and canopy heights in boreal forests. *Remote Sens. Environ.* **2020**, *251*, 112110. <https://doi.org/10.1016/j.rse.2020.112110>.
37. Li, W.; Niu, Z.; Shang, R.; Qin, Y.; Wang, L.; Chen, H. 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.* **2020**, *92*, 102163. <https://doi.org/10.1016/j.jag.2020.102163>.
38. Malambo, L.; Popescu, S.C. Assessing the agreement of ICESat-2 terrain and canopy height with airborne lidar over US ecozones. *Remote Sens. Environ.* **2021**, *266*, 112711. <https://doi.org/10.1016/j.rse.2021.112711>.
39. Liu, A.; Cheng, X.; Che, Z. Performance evaluation of GEDI and ICESat-2 laser altimeter data for terrain and canopy height retrievals. *Remote Sens. Environ.* **2021**, *264*, 112571. <https://doi.org/10.1016/j.rse.2021.112571>.
40. Saatchi, S.S.; Harris, N.L.; Brown, S.; Lefsky, M.; Mitchard, E.T.; Salas, W.; Zutta, B.R.; Buermann, W.; Lewis, S.L.; Hagen, S.; et al. Benchmark map of forest carbon stocks in tropical regions across three continents. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 9899–9904. <https://doi.org/10.1073/pnas.1019576108>.
41. Simard, M.; Pinto, N.; Fisher, J.B.; Baccini, A. Mapping forest canopy height globally with spaceborne lidar. *Geophys. Res. Lett.* **2011**, *116*, 4021. <https://doi.org/10.1029/2011JG001708>.
42. Hu, T.; Su, Y.; Xue, B.; Liu, J.; Zhao, X.; Fang, J.; Guo, Q. Mapping Global Forest Aboveground Biomass with Spaceborne LiDAR, Optical Imagery, and Forest Inventory Data. *Remote Sens.* **2016**, *8*, 565. <https://doi.org/10.3390/rs8070565>.
43. Margolis, H.A.; Nelson, R.F.; Montesano, P.M.; Beaudoin, A.; Sun, G.; Andersen, H.-E.; Wulder, M.A. Combining satellite lidar, airborne lidar, and ground plots to estimate the amount and distribution of aboveground biomass in the boreal forest of North America. *Can. J. For. Res.* **2015**, *45*, 838–855. <https://doi.org/10.1139/cjfr-2015-0006>.

44. Nelson, R.; Margolis, H.; Montesano, P.; Sun, G.; Cook, B.; Corp, L.; Andersen, H.-E.; deJong, B.; Pellat, F.P.; Fickel, T.; et al. Lidar-based estimates of aboveground biomass in the continental US and Mexico using ground, airborne, and satellite observations. *Remote Sens. Environ.* **2017**, *188*, 127–140.
45. Narine, L.L.; Popescu, S.C.; Malambo, L. Using ICESat-2 to Estimate and Map Forest Aboveground Biomass: A First Example. *Remote Sens.* **2020**, *12*, 1824. <https://doi.org/10.3390/rs12111824>.
46. Silva, C.A.; Duncanson, L.; Hancock, S.; Neuenschwander, A.; Thomas, N.; Hofton, M.; Fatoyinbo, L.; Simard, M.; Marshak, C.Z.; Armston, J.; et al. Fusing simulated GEDI, ICESat-2 and NISAR data for regional aboveground biomass mapping. *Remote Sens. Environ.* **2021**, *253*, 112234. <https://doi.org/10.1016/j.rse.2020.112234>.
47. Joshi, N.; Mitchard, E.T.A.; Broly, M.; Schumacher, J.; Fernández-Landa, A.; Johannsen, V.K.; Marchamalo, M.; Fensholt, R. Understanding “saturation” of radar signals over forests. *Sci. Rep.* **2017**, *7*, 3505. <http://doi.org/10.1038/s41598-017-03469-3>.
48. Lee, J.-S.; Pottier, E. *Polarimetric Radar Imaging: From Basics to Applications*; CRC Press: Boca Raton, FL, USA, 2019.
49. Gonsamo, A.; Ciais, P.; Miralles, D.G.; Sitch, S.; Dorigo, W.; Lombardozzi, D.; Friedlingstein, P.; Nabel, J.E.M.S.; Goll, D.S.; O'Sullivan, M.; et al. Greening drylands despite warming consistent with carbon dioxide fertilization effect. *Glob. Chang. Biol.* **2021**, *27*, 3336–3349. <https://doi.org/10.1111/gcb.15658>.
50. Xi, Z.; Xu, H.; Xing, Y.; Gong, W.; Chen, G.; Yang, S. Forest Canopy Height Mapping by Synergizing ICESat-2, Sentinel-1, Sentinel-2 and Topographic Information Based on Machine Learning Methods. *Remote Sens.* **2022**, *14*, 364. <https://doi.org/10.3390/rs14020364>.
51. Pourshamsi, M.; Xia, J.; Yokoya, N.; Garcia, M.; Laval, M.; Pottier, E.; Balzter, H. Tropical forest canopy height estimation from combined polarimetric SAR and LiDAR using machine-learning. *ISPRS J. Photogramm. Remote Sens.* **2021**, *172*, 79–94. <https://doi.org/10.1016/j.isprsjprs.2020.11.008>.
52. Gonsamo, A.; Chen, J.M. Evaluation of the GLC2000 and NALC2005 land cover products for LAI retrieval over Canada. *Can. J. Remote Sens.* **2011**, *37*, 302–313.
53. Gillis, M.D.; Omule, A.Y.; Brierley, T. Monitoring Canada's forests: The national forest inventory. *For. Chron.* **2005**, *81*, 214–221.
54. Kangas, A.; Maltamo, M. Forest Inventory. In *Methodology and Applications*; Springer: Berlin/Heidelberg, Germany, 2006.
55. MacDicken, K.G. Global forest resources assessment 2015: What, why and how? *For. Ecol. Manag.* **2015**, *352*, 3–8. <https://doi.org/10.1016/j.foreco.2015.02.006>.
56. Castilla, G.; Hall, R.J.; Skakun, R.; Filiatrault, M.; Beaudoin, A.; Gartrell, M.; Smith, L.; Groenewegen, K.; Hopkinson, C.; van der Sluijs, J. The Multisource Vegetation Inventory (MVI): A Satellite-Based Forest Inventory for the Northwest Territories Taiga Plains. *Remote Sens.* **2022**, *14*, 1108. <https://doi.org/10.3390/rs14051108>.
57. Matasci, G.; Hermosilla, T.; Wulder, M.A.; White, J.C.; Coops, N.C.; Hobart, G.W.; Zald, H.S. Large-area mapping of Canadian boreal forest cover, height, biomass and other structural attributes using Landsat composites and lidar plots. *Remote Sens. Environ.* **2018**, *209*, 90–106. <https://doi.org/10.1016/j.rse.2017.12.020>.
58. Matasci, G.; Hermosilla, T.; Wulder, M.A.; White, J.C.; Coops, N.C.; Hobart, G.W.; Bolton, D.K.; Tompalski, P.; Bator, C.W. Three decades of forest structural dynamics over Canada's forested ecosystems using Landsat time-series and lidar plots. *Remote Sens. Environ.* **2018**, *216*, 697–714. <https://doi.org/10.1016/j.rse.2018.07.024>.
59. Wulder, M.A.; White, J.C.; Cranny, M.M.; Hall, R.J.; Luther, J.E.; Beaudoin, A.; Goodenough, D.G.; Dechka, J.A. Monitoring Canada's forests. Part 1: Completion of the EOSD land cover project. *Can. J. Remote Sens.* **2008**, *34*, 549–562. <https://doi.org/10.5589/m08-066>.
60. Brandt, J.P. The extent of the North American boreal zone. *Environ. Rev.* **2009**, *17*, 101–161. <https://doi.org/10.1139/A09-004>.
61. Natural Resources of Canada, 2020. The State of Canada's Forests. Annual Report 2020. Available online: <https://d1ied5g1xfpx8.cloudfront.net/pdfs/40219.pdf> (accessed on 25 March 2022).
62. Kurz, W.A.; Apps, M.J. 70-year retrospective analysis of carbon fluxes in the Canadian forest sector. *Ecol. Appl.* **1999**, *9*, 526–547. <https://doi.org/10.2307/2641142>.
63. Hanes, C.C.; Wang, X.; Jain, P.; Parisien, M.-A.; Little, J.M.; Flannigan, M.D. Fire-regime changes in Canada over the last half century. *Can. J. For. Res.* **2019**, *49*, 256–269. <https://doi.org/10.1139/cjfr-2018-0293>.
64. White, J.C.; Wulder, M.A.; Hermosilla, T.; Coops, N.C.; Hobart, G.W. A nationwide annual characterization of 25 years of forest disturbance and recovery for Canada using Landsat time series. *Remote Sens. Environ.* **2017**, *192*, 303–321. <https://doi.org/10.1016/j.rse.2017.03.035>.
65. Tymstra, C.; Stocks, B.J.; Cai, X.; Flannigan, M.D. Wildfire management in Canada: Review, challenges and opportunities. *Prog. Disaster Sci.* **2020**, *5*, 100045. <https://doi.org/10.1016/j.pdisas.2019.100045>.
66. Neumann, T.A.; Martino, A.J.; Markus, T.; Bae, S.; Bock, M.R.; Brenner, A.C.; Brunt, K.M.; Cavanaugh, J.; Fernandes, S.T.; Hancock, D.W.; et al. 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.* **2019**, *233*, 111325. <https://doi.org/10.1016/j.rse.2019.111325>.
67. Magruder, L.A.; Brunt, K.M.; Alonzo, M. Early ICESat-2 on-orbit geolocation validation using ground-based corner cube retro-reflectors. *Remote Sens.* **2020**, *12*, 3653. <https://doi.org/10.3390/rs12213653>.
68. Neuenschwander, A.; Magruder, L.; Guenther, E.; Hancock, S.; Purslow, M. Radiometric Assessment of ICESat-2 over Vegetated Surfaces. *Remote Sens.* **2022**, *14*, 787. <https://doi.org/10.3390/rs14030787>.
69. Robinson, D.A.; Estilow, T.W. NOAA CDR Program NOAA Climate Data Record (CDR) of Northern Hemisphere (NH) Snow Cover Extent (SCE), Version 1, NECI: Mansfield, MA, USA, 2012.

70. Estilow, T.; Young, A.; Robinson, D. A long-term northern hemisphere snow cover extent data record for climate studies and monitoring. *Earth Syst. Sci. Data* **2015**, *7*, 137–142. <https://doi.org/10.5194/essd-7-137-2015>.
71. Torres de Almeida, C.; Gerente, J.; Rodrigo dos Prazeres Campos, J.; Caruso Gomes Junior, F.; Providelo, L.A.; Marchiori, G.; Chen, X. Canopy Height Mapping by Sentinel 1 and 2 Satellite Images, Airborne LiDAR Data, and Machine Learning. *Remote Sens.* **2022**, *14*, 4112. <https://doi.org/10.3390/rs14164112>
72. Hermosilla, T.; Wulder, M.A.; White, J.C.; Coops, N.C.; Hobart, G.W.; Campbell, L.B. Mass data processing of time series Landsat imagery: Pixels to data products for forest monitoring. *Int. J. Digit. Earth* **2016**, *9*, 1035–1054.
73. Hermosilla, T.; Wulder, M.A.; White, J.C.; Coops, N.C.; Hobart, G.W. Updating Landsat time series of surface-reflectance composites and forest change products with new observations. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *63*, 104–111. <https://doi.org/10.1016/j.jag.2017.07.013>.
74. Louis, J.; Debaecker, V.; Pflug, B.; Main-Knorn, M.; Bieniarz, J.; Mueller-Wilm, U.; Cadau, E.; Gascon, F. SENTINEL-2 SEN2COR: L2A processor for users. In Proceedings of the ESA Living Planet Symposium, Prague, Czech, 9–13 May 2016.
75. Shimada, M.; Itoh, T.; Motooka, T.; Watanabe, M.; Tomohiro, S.; Thapa, R.; Lucas, R. New global forest/non-forest maps from ALOS PALSAR data (2007–2010). *Remote Sens. Environ.* **2014**, *155*, 13–31. <https://doi.org/10.1016/j.rse.2014.04.014>.
76. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. <https://doi.org/10.1023/A:1010933404324>.
77. Rodríguez-Veiga, P.; Quegan, S.; Carreiras, J.; Persson, H.J.; Fransson, J.E.S.; Hosillo, A.; Ziolkowski, D.; Stereńczak, K.; Lohberger, S.; Stängel, M.; et al. Forest biomass retrieval approaches from earth observation in different biomes. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *77*, 53–68. <https://doi.org/10.1016/j.jag.2018.12.008>.
78. Olden, J.D.; Lawler, J.J.; Poff, N.L. Machine learning methods without tears: A primer for ecologists. *Q. Rev. Biol.* **2008**, *83*, 171–193.
79. Wright, M.N.; Ziegler, A. Ranger: A fast implementation of random forests for high dimensional data in C++ and R. *J. Statist. Softw.* **2017**, *77*, 1–17. <https://doi.org/10.18637/jss.v077.i01>.
80. R Core Team. R: A Language and Environment for Statistical Computing R Foundation for Statistical Computing, Vienna, Austria, 2018. Available online: <https://www.r-project.org/> (accessed on 15 July 2021).
81. Liaw, A.; Wiener, M. Classification and regression by Random Forest. *R News* **2002**, *2*, 18–22.
82. Liu, Y.; Gong, W.; Xing, Y.; Hu, X.; Gong, J. Estimation of the forest stand mean height and aboveground biomass in Northeast China using SAR Sentinel-1B, multispectral Sentinel-2A, and DEM imagery. *ISPRS J. Photogramm. Remote Sens.* **2019**, *151*, 277–289.
83. Omar, H.; Misman, M.A.; Kassim, A.R. Synergetic of PALSAR-2 and Sentinel-1A SAR Polarimetry for Retrieving Aboveground Biomass in Dipterocarp Forest of Malaysia. *Appl. Sci.* **2017**, *7*, 675. <https://doi.org/10.3390/app7070675>.
84. Ji, Y.; Huang, J.; Ju, Y.; Guo, S.; Yue, C. Forest structure dependency analysis of L-band SAR backscatter. *PeerJ.* **2020**, *8*, e10055. <https://doi.org/10.7717/peerj.10055>.
85. Watanabe, M.; Shimada, M.; Rosenqvist, A.; Tadono, T.; Matsuoka, M.; Romshoo, S.A.; Ohta, K.; Furuta, R.; Nakamura, K.; Moriyama, T. Forest Structure Dependency of the Relation Between L-Band Sigma and Biophysical Parameters. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 3154–3165. <https://doi.org/10.1109/TGRS.2006.880632>.