

State-wide forest canopy height and aboveground biomass map for New York with 10 m resolution, integrating GEDI, Sentinel-1, and Sentinel-2 data



Haifa Tamiminia ^{a,*}, Bahram Salehi ^a, Masoud Mahdianpari ^{b,c}, Tristan Goulden ^d

^a Department of Environmental Resources Engineering, State University of New York College of Environmental Science and Forestry (ESF), NY 13210, USA

^b C-CORE, St. John's, NL A1B 3X5, Canada

^c Department of Electrical and Computer Engineering, Memorial University of Newfoundland, St. John's, NL A1B 3X5, Canada

^d National Ecological Observatory Network, Boulder, CO 80301, USA

ARTICLE INFO

Keywords:

Canopy height model
Forest aboveground biomass
Machine learning
GEDI
Spaceborne LiDAR
Large-scale
Google earth engine

ABSTRACT

Investigating the quantity of forest aboveground biomass (AGB) is crucial for understanding the role forests play in the global carbon cycle. The canopy height model (CHM) is a critical component in estimating AGB, as it provides a three-dimensional representation of the tree canopy. Traditional CHM estimation methods are time-consuming, labor-intensive, and expensive, particularly at large-scales. Remote sensing is a cost-effective and efficient alternative approach, providing valuable information over large areas in a timely manner. The Global Ecosystem Dynamics Investigation (GEDI) onboard the International Space Station is a space-based light detection and ranging (LiDAR) system designed to collect information on vertical structures of vegetation. One major problem with the collection of GEDI data is that it provides limited information over discrete ground samples, also known as footprints, and thus do not provide wall-to-wall gridded height products. The objective of this study was twofold: a) to integrate the GEDI LiDAR footprint heights with Sentinel-2 multispectral imagery to generate a 10 m wall-to-wall CHM map of New York State (NYS), USA for the year 2019 and b) to improve our previously generated AGB map (both accuracy and resolution) of NYS for the year 2019 by fusing Sentinel-2 multispectral, Sentinel-1 synthetic aperture radar (SAR), and the produced CHM. To generate the 10 m CHM map, the GEDI footprints height measurements were extrapolated using Sentinel-2 imagery and a random forest model. The CHM that was produced was assessed by using GEDI footprints that were not part of the training phase and were therefore independent (extrapolated). Comparing our 10 m CHM with the available global 30 m CHM map provided by Potapov et al. (2021) over NYS shows significant improvement not only in terms of spatial resolution, but also in terms of accuracy. The root mean square error (RMSE) of our 10 m CHM is 4.4 m while this value is 7.49 m for the 30 m CHM over NYS. Similarly, the R² value for the 10 m CHM map is 0.74, while that of the 30 m CHM is 0.46. Finally, the integration of produced 10 m CHM, Sentinel-1, and Sentinel-2 datasets were utilized to create a 10 m AGB map of NYS with the RMSE of 39.49 Mg/ha, and R² of 0.65. The results demonstrate the potential of integrating GEDI, Sentinel-1, and Sentinel-2 data for providing a valuable tool for large-scale mapping of forest canopy structure and biomass, which can help to inform forest management and carbon accounting efforts.

1. Introduction

Forests are one of the most critical natural resources which play a significant role in mitigating the effects of climate change by serving as a sink for carbon dioxide (CO₂) (Chavan and Rasal, 2012). However, the world has experienced an estimated loss of 178 million hectares of

forests since 1990 (FAO, 2020), making forest monitoring an important research and practice area. Accurate measurement of the vertical structure of forests is crucial for monitoring forest change, disturbance, and forest carbon stocks (Ahmed et al., 2015). An accurate estimation of forest above-ground biomass (AGB), a crucial indicator of carbon stocks, heavily relies on the forest's vertical structure. Therefore, monitoring

* Corresponding author.

E-mail address: htamimin@esf.edu (H. Tamiminia).

the vertical structure of forests is a key task of forest monitoring efforts. In addition, accurate and high resolution canopy height model (CHM) is of great interest for precise AGB estimation and forest characteristics modeling (Lisein et al., 2013). Thus, providing large-scale AGB and CHM maps is of paramount importance for sustainable management of natural resources (Ma et al., 2019). However, finding proper and sufficient samples that cover large areas might be challenging. These limitations can be overcome by remote sensing techniques which provide valuable information about Earth's surface at various spatiotemporal scales (Y. Li et al., 2019).

Recent advances in optical, synthetic aperture radar (SAR), and light detection and ranging (LiDAR) sensors have paved the way for large-scale CHM and AGB mapping (Hudak et al., 2002; Hussin et al., 2014; Jong et al., 2003; Kelsey and Neff, 2014; Mutanga et al., 2012; Urbazaev et al., 2018). Although AGB can indirectly be estimated using single optical or SAR images, incorporating direct measurement of canopy structure collected by lidar results in direct measurement of CHM leading to improved AGB and reduction of uncertainty in such products. The measurement of forest vertical structure has traditionally been conducted using airborne LiDAR data, also known as airborne laser scanning (ALS) (Potapov et al., 2021). While airborne LiDAR can provide accurate estimations of tree height, it is expensive and impractical for large-scale (e.g., State/ Nationwide) applications (Potapov et al., 2021). In recent years, spaceborne LiDAR sensors such as the Global Ecosystem Dynamics Investigation (GEDI) and Ice, Cloud, and land Elevation Satellite 2 (ICESat-2) have been developed to provide LiDAR measurements on a global scale which fills spatial gaps in current airborne LiDAR coverage.

Previous studies have shown the potential of using spaceborne LiDAR including GEDI in combination with optical and SAR remote sensing imagery to provide detailed estimates of CHM and AGB (Dorado-Roda et al., 2021; Duncanson et al., 2020; Francini et al., 2022; Liu et al., 2022; Qi et al., 2019; Saarela et al., 2018). Qi et al. in 2019 used the fusion of simulated GEDI LiDAR data and TanDEM-X InSAR data to improve forest height modeling. Liu et al. in 2020 integrated GEDI and ICESat-2 data for canopy height mapping of China's forests. Dorado-Roda et al. (2021) combined airborne LiDAR with GEDI data to estimate CHM and AGB for Mediterranean forests. The combination of simulated GEDI, ICESat-2, and the NASA-Indian Space Research Organization (ISRO) Synthetic Aperture Radar (NISAR) was used for biomass estimation by Duncanson et al. in 2020.

Among these spaceborne LiDAR datasets, GEDI is specifically focused on measuring vegetation canopy structure (Xi et al., 2022). GEDI is a full waveform LiDAR that produces profiles of forest canopies with 25 m diameter footprints. However, the sparse sample spacing of GEDI data (60 m along-track and about 600 m across-track) represents a point-based sampling approach, which is a limitation of this spaceborne LiDAR (Potapov et al., 2021). To address this limitation, the integration of GEDI with wall to wall optical/ SAR imagery can be used to spatially extrapolate GEDI samples (Potapov et al., 2021). This extrapolation technique has been used in several studies to generate large-scale CHM and AGB maps (Baccini et al., 2008; Chi et al., 2015; Simard et al., 2011). Nevertheless, the capability of the combination of GEDI data with other optical and SAR imagery for a more accurate and high-resolution CHM and AGB mapping has not been completely explored. For instance, Simard et al. (2011) and Chi et al. (2015) utilized the combination of Geoscience Laser Altimeter System (GLAS) onboard NASA's Ice, Cloud, and land Elevation Satellite (ICESat) and Moderate Resolution Imaging Spectro-radiometer (MODIS) imagery to create large-scale CHM maps with the spatial resolution of 1000 m and 500 m, respectively. Additionally, Potapov et al. in 2021 combined GEDI samples and Landsat imagery to create a global CHM of forest canopies at 30 m resolution (Potapov et al., 2021). Previous studies have operated at coarser spatial resolutions to overcome the sparse coverage of GEDI data. Thus, our study uniquely contributes by filling gaps in spatial coverage using finer spatial resolution and exploring the potential of GEDI, Sentinel-1, and

Sentinel-2 integration for more precise statewide CHM and AGB mapping.

The objective of this study is twofold. First, the integration of GEDI footprints with Sentinel-2 imagery is used to generate a 10 m wall-to-wall CHM of New York State, USA, hereafter simply New York, for the year 2019. Second, the combination of generated CHM, Sentinel-1, and Sentinel-2 data were utilized to create a 10 m AGB map of NYS for the year 2019. To develop this framework, the Google Earth Engine (GEE) platform was used to implement a random forest (RF) regression model (Gorelick et al., 2017). Specifically, this paper aims to address the following research questions: 1) investigate the potential of GEDI data for providing a high resolution CHM for large-scale applications, 2) explore the importance of combining footprint spaceborne LiDAR samples with optical, and SAR data for generating a gridded wall-to-wall statewide CHM, 3) and determine the effect of the CHM as an input predictor for accurate AGB mapping.

2. Study area and datasets

2.1. Study area

This study presents a framework for forest CHM and AGB estimation of the entirety of New York (Fig. 1). New York covers an area of 141,297 km², of which 61% is forested (DEC 2015). New York forest land is one of the most diverse resources in the region since it includes 94 tree species and 55 forest types (Albright, 2018). Among the dominant tree genera are *Acer* (maple), *Fagus* (beech), and *Betula* (birch) comprising the most common forest types, covering 53% of NYS forested land (DEC 2015).

2.2. Global ecosystem dynamics investigation (GEDI)

GEDI is the first spaceborne, full-waveform LiDAR onboard the International Space Station providing the first global, high-resolution observations of forest vertical structure (Qi and Dubayah, 2016). GEDI, launched in 2018, is designed to provide information on vertical structure of forests, forest canopy height, and surface elevation. This spaceborne LiDAR provides profiles of forest canopies with 25 m diameter footprints separated by 60 m along-track and about 600 m across-track (Qi and Dubayah, 2016). GEDI comprises three lasers that produce eight parallel tracks of observations. Each laser fires 242 times per second and illuminates a 25 m spot (footprint) on the surface over which 3D structure is measured.

In this study, GEDI's Level 2 A Geolocated Elevation and Height Metrics Product in GEE was used for CHM prediction. GEDI Level 2 data provides geolocated, calibrated, and validated waveform data of individual GEDI footprints, as well as derived metrics of vegetation vertical structure, such as canopy height, canopy cover, and vertical distribution of vegetation layers (Dubayah et al., 2020). GEDI Level 2 data is produced by processing GEDI Level 1B data, which includes raw waveform data and ancillary data such as GPS, attitude, and instrument status information (Dubayah et al., 2020). This data is composed of 100 relative height (RH) percentiles representing the percentile of energy return height relative to the ground (Potapov et al., 2021). To generate the 10 m CHM, the RH95th percentiles were collected between June and September of 2019. In this study, the RH95th percentile was chosen due to its sensitivity to capture the upper boundaries of the canopy (Potapov et al., 2021). In addition, it minimizes the outliers and represents a more robust metric of dominant canopy heights. Thus, the RH95 percentile contributes to a precise and reliable characterization of forest vertical structure. Fig. 1 shows GEDI footprint samples of New York in 2019 (June–September). The colors are categorized based on height.

To ensure the reliability and accuracy of the analyses, a quality mask was applied to the GEDI data to filter out data points with poor quality or degradation. The quality mask was implemented through a custom function named qualityMask, which selectively retained only the data

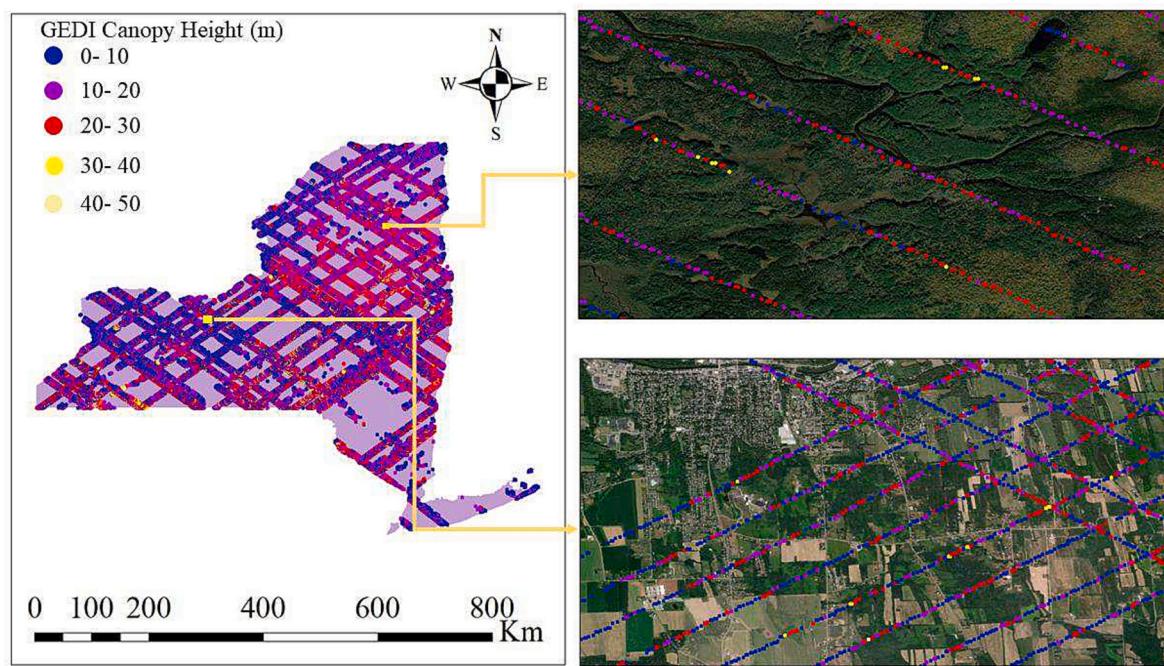


Fig. 1. Visual example of spaceborne GEDI footprints for New York from June-2019 to September-2019, shown at different geographic scales.

points that met specific filtering conditions. Specifically, the function first extracted the ‘quality_flag’ band from the GEDI image and retained the pixels where the ‘quality_flag’ value was equal to 1, indicating good data quality. Additionally, the function checked the ‘degrade_flag’ band and retained the pixels with a value of 0, indicating data points without degradation. The combination of these filtering conditions ensured that only high-quality and reliable GEDI data was used for subsequent analyses. The application of the quality mask plays a critical role in enhancing the validity and robustness of the findings presented in this paper, enabling accurate assessments of forest structure and biomass based on the GEDI spaceborne LiDAR data.

2.3. Sentinel-2 imagery

The Sentinel-2 Multi-Spectral Instrument (MSI) is onboard the Sentinel-2A and -2B satellites orbiting the Earth at 786 km altitude (Claverie et al., 2018). The sensor has a 20.6° field of view corresponding to an image swath width of approximately 290 km (Claverie et al., 2018). Sentinel-2 carries a multispectral imager with 13 spectral bands covering the visible, near infrared, and shortwave infrared regions of the electromagnetic spectrum. The spatial resolution of Sentinel-2 imagery varies depending on the spectral band, with a spatial resolution ranging from 10 m to 60 m: 10 m for the visible (Red, Green, Blue) and the near-infrared (NIR) bands, 20 m for the red edge, narrow NIR and shortwave infrared (SWIR) bands, and 60 m for the atmospheric bands (e.g., coastal aerosol, water vapor, and SWIR-Cirrus).

In this paper, first, Sentinel-2 images between June and September of 2019 were collected using GEE platform (Tamiminia et al., 2020). Second, a cloud mask was applied to produce cloud free images. Then, visible, NIR, SWIR-1, SWIR-2, and red-edge bands were extracted. Bands with 20 m spatial resolution were resampled to 10 m using bicubic interpolation and re-projected to NAD83 Conus Albers EPSG: 5070 coordinate system. The decision to resample bands from 20 m to 10 m spatial resolution was made with the intention of harmonizing the dataset for consistency and compatibility with other data sources. Several studies have demonstrated the practice of downsampling 20 m Sentinel-2 bands to 10 m to leverage the finer spatial resolution and capture more detailed information (Du et al., 2016; Q. Wang et al., 2016; Li et al., 2017; Kaplan and Avdan, 2018; J. Wang et al., 2019).

Specifically, Du et al., 2016 pointed out that upsampling of Sentinel-2 10 m bands to 20 m neglects the detailed information available at 10 m resolution. The bands were used to calculate some vegetation indices that are important for vegetation monitoring. Normalized difference vegetation index (NDVI), normalized burn ratio (NBR), Normalized Difference Red Edge (NDRE), normalized difference moisture index (NDMI), and tasseled cap greenness were calculated using spectral bands. Vegetation indices can provide valuable information for forest mapping (Zheng et al., 2007). Finally, these vegetation indices in addition to spectral bands were stacked and used as input predictors to implement the RF regression model.

2.4. Sentinel-1 imagery

SAR data provides complement textural data for environmental applications including forest AGB estimation (Saatchi, 2019). We used Sentinel-1 dual polarization C-band data with 10 m resolution for our AGB estimation. Collection of images in June and September of 2019 in vertical transmit/vertical receive (VV) and vertical transmit/horizontal receive (VH) polarizations was used to estimate AGB. GEE provides ground range detected (GRD), pre-processed with Sentinel-1 Toolbox to create calibrated, ortho-corrected products (Tamiminia et al., 2020). The pre-processing consists of three steps: thermal noise removal, radiometric calibration, and terrain correction. Terrain correction was performed using Shuttle Radar Topography Mission (SRTM) 30 m DEM and the final terrain-corrected values were converted to decibels using log scaling and a smoothing speckle filter was applied. In addition to VV and VH bands, we calculated span and band ratios (Eqs. (1) and (2)) (Lee and Pottier, 2009).

$$span_{s1} = VH^2 + VV^2 \quad (1)$$

$$ratio_{s1} = VV/VH \quad (2)$$

2.5. The European Space Agency (ESA) world cover

The European Space Agency (ESA) provides a global land cover map at 10 m resolution using Sentinel-1 and Sentinel-2 datasets (Zanaga et al., 2022). The ESA WorldCover 2020 map consists of 11 land cover

classes including tree cover, shrubland, grassland, build-up, bare/sparse vegetation, snow-ice, permanent water bodies, herbaceous wetland, mangroves, and moss-lichen. These classes are defined by the United Nations (UN) Food and Agriculture Organization (FAO) land cover classification system. We used the ESA land cover map to mask out water bodies from our CHM and AGB maps. In addition, GEDI height datasets are not able to discriminate between trees and buildings (Potapov et al., 2021). To address this limitation, the build-up class was employed to identify and mask out non-forest areas from forested regions. The build-up class helps us to mitigate the potential of overestimation of tree height in non-forest areas.

3. Methods

3.1. Method overview

We have extensively used different machine learning algorithms for AGB estimation (Tamiminia et al., 2021; Tamiminia et al., 2021a; Tamiminia et al., 2022). Since the RF model has less hyperparameters in comparison to gradient boosting machines and it is less sensitive to hyperparameter tuning, it provides more accurate results for biomass estimation (Tamiminia, 2021; Y. Li et al., 2019; M. Li et al., 2013). Thus, in this paper, we utilized an RF regression algorithm to create a 10 m CHM and AGB map of New York for 2019. Our state-wide CHM and AGB mapping approach comprises two main steps (CHM and AGB mapping) as depicted in Fig. 2. First, GEDI footprints were extrapolated to a 10 m pixel resolution raster using an RF model and Sentinel-2 imagery. Then, the trained RF regression model was used to generate a 10 m CHM of New York for the year 2019. Second, the produced CHM was used along with the combination of Sentinel-1 and Sentinel-2 data to generate the AGB map of New York at 10 m resolution. In this step, airborne LiDAR derived AGB rasters were utilized as training/testing datasets to train and evaluate the RF regression model to produce a 10 m AGB map of New York.

3.2. Canopy height model (CHM) extrapolation

In this study, a 10 m CHM of New York was generated for the year 2019. The process involved three main steps, as follows:

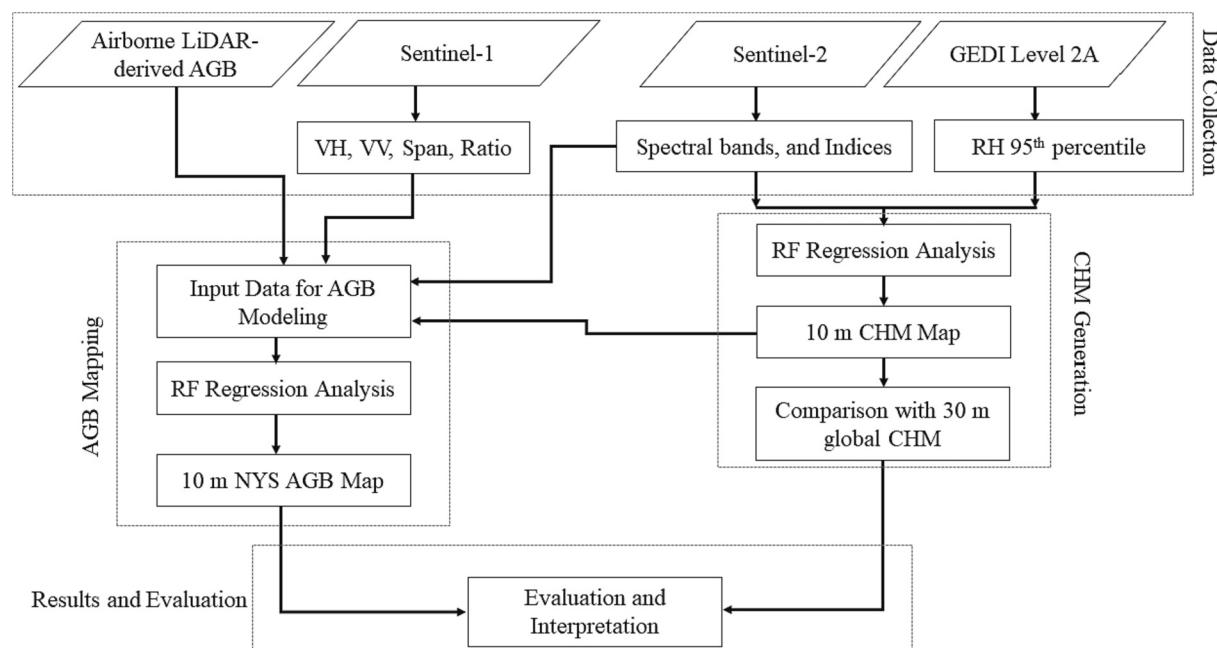


Fig. 2. An overview of the proposed method for CHM and AGB modeling using GEDI, Sentinel-1, and Sentinel-2 data and an RF method.

a. Extraction of GEDI footprint height measurements: The first step involved the extraction of point-based height measurements from GEDI Level 2 A data. These point-based height measurements provided a baseline for the estimation of CHM across New York. A total of 1,574,441 samples were randomly split to 70/30 training/testing samples. The RH 95th percentile of GEDI height was used as a dependent variable for the RF regression model.

The RH95th percentile of GEDI Level 2 A is a metric that represents the height of the top of the canopy (Potapov et al., 2021). It is a robust metric that is less affected by the presence of multiple layers of vegetation and vegetation gaps than other height metrics. This makes it particularly useful for forest canopy height modeling.

b. Extrapolation of GEDI height measurements using Sentinel-2 imagery: The extracted GEDI height measurements were then extrapolated using spectral bands and indices of Sentinel-2 imagery as independent variables for training the RF model to generate a high-resolution CHM. Given the 25 m footprint of GEDI samples, it was necessary to derive a single independent value for each GEDI height by computing the average pixel values of the 10 m Sentinel-2 predictors. In instances where a Sentinel-2 pixel overlaps with two GEDI footprints, a threshold of 70% was established. If the GEDI footprint covers less than 70%, the average height values is calculated. Conversely, if the GEDI footprint covers more than 70%, the dominant height is assigned to the Sentinel-2 pixel.

The RF model was built using the RF built-in function in the GEE platform. The trained RF regression model was utilized to generate the 10 m CHM of New York for the year 2019. This extrapolation helped to fill the gaps in the GEDI height measurements and enabled the creation of a synoptic CHM map.

c. Validation of CHM using GEDI height measurements: Finally, the generated CHM was validated using the original GEDI testing height measurements. This validation process helped to ensure that the generated CHM map was accurate and reliable, and that it accurately represented the height of the forest canopies across New York. In

addition, the 10 m CHM was compared to the 30 m global forest CHM provided by Potapov et al.

3.3. Aboveground biomass (AGB) extrapolation

LiDAR data has become a widely used resource for generating accurate AGB maps in the field of remote sensing (Hirata et al., 2018; Hudak et al., 2020). Following this trend, our study also leveraged airborne LiDAR data for four pilot areas to create AGB maps in New York.

The process of generating AGB maps for these pilot areas involved a two-stage approach. First, we utilized an RF regression model (Tamiminia et al., 2022) to produce airborne LiDAR-derived AGB rasters for each of the four pilot areas. The procedure of generating airborne LiDAR AGB rasters was done by Johnson et al. (2022). This step provided us with essential AGB estimates, which served as the foundation for the subsequent stages. The second stage was focused on the creation of a comprehensive training and testing dataset from the generated airborne LiDAR AGB maps for the four pilot areas (as illustrated in Fig. 3). To ensure the dataset's representativeness, we employed a stratified random sampling method. This approach involved sorting the AGB samples into 2 Mg ha⁻¹ bins and then randomly selecting 500 pixels from each bin. In cases where a bin contained fewer than 500 pixels, we randomly selected half of the pixels for the training and testing split (Hudak et al., 2020). This strategic sampling technique was chosen to capture the full distribution of AGB values within the LiDAR-derived AGB maps accurately. This stratified random sampling approach was chosen to ensure that the training/testing dataset accurately represented the distribution of AGB values within the LiDAR-derived AGB maps. The use of this approach was crucial in ensuring that the AGB maps generated were accurate and representative of the distribution of AGB values in the study area.

3.4. RF model

In our study, we adopted a grid search approach to optimize the

parameters of our machine learning model. Specifically, we focused on tuning several critical parameters, namely 'numberOfTrees', 'minLeafPopulation', 'bagFraction', and 'seed'. The 'numberOfTrees' represents the number of decision trees in the ensemble, and through the grid search, we explored different values to determine the ideal number of trees that balances model complexity and predictive performance. Similarly, 'minLeafPopulation' refers to the minimum number of samples required to form a leaf node in each tree. We experimented with various values to find the optimal setting that prevents overfitting while capturing meaningful patterns in the data. The 'bagFraction' indicates the proportion of the training dataset used to train each individual tree, and we searched for the best value to enhance model diversity and generalization. Lastly, 'seed' is a random number seed used to ensure reproducibility, and we evaluated multiple seeds to select the one that provides the most stable and reliable results. To carry out the grid search, we defined a range of potential values for each parameter based on prior knowledge and literature review. The grid search then exhaustively tested all possible combinations of these values and evaluated the model's performance using cross-validation techniques. For each combination, we utilized the RMSE as the evaluation metric, aiming to minimize the RMSE to achieve the most accurate predictions. By employing the grid search method, we aimed to optimize the model's hyperparameters and enhance its ability to generalize well to new, unseen data, making it a robust and effective tool for our regression task. The winning parameters are as follows: numberOfTrees: 150, minLeafPopulation: 3, bagFraction: 1, and seed: 123.

4. Results

4.1. CHM mapping

This section describes the result of the generated 10 m CHM of New York using the GEDI RH 95th percentile height footprints and Sentinel-2 imagery and the RF model. The 10 m CHM was then compared with a global 30 m resolution CHM map generated by Potapov et al. (2021) using Landsat imagery and GEDI data. Both our 10 m CHM and global

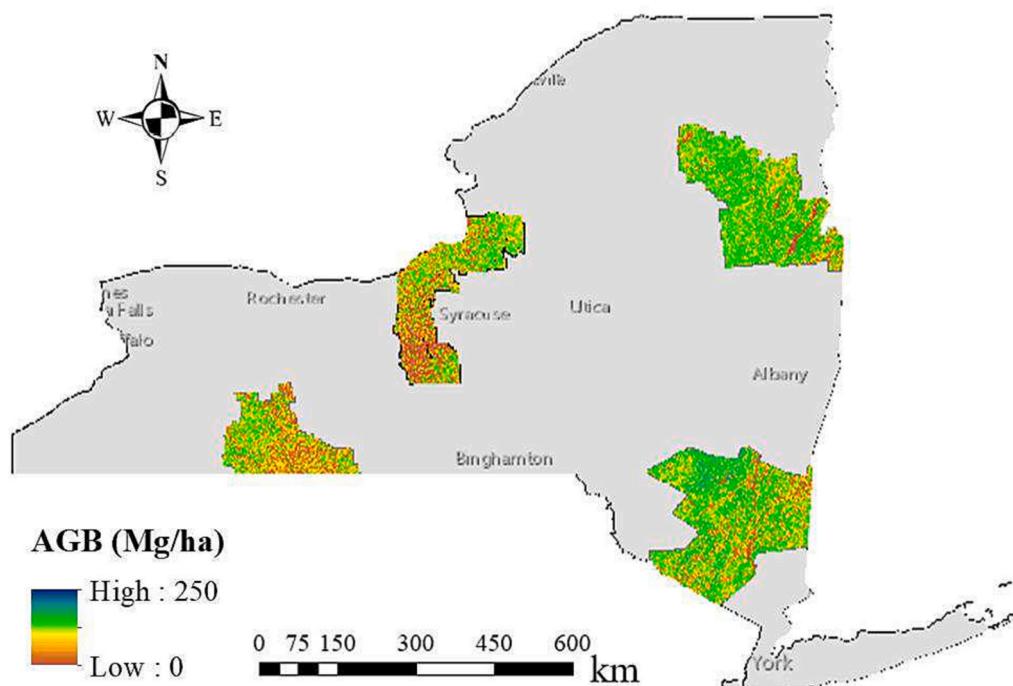


Fig. 3. Airborne LiDAR-derived AGB maps at 4 pilot areas: (Three counties-2014, Warren Washington Essex-2015, Allegany and Steuben-2016, CayugaOswego-2018).

30 m CHM were validated against the independent test samples of GEDI footprints (that were not participated in the training phase) and accuracy metrics of mean bias error (MBE) and root mean square error (RMSE) as well as R² were calculated. Table 1 lists the mean bias error (MBE), root mean square error (RMSE), and R² of both CHM maps for the year 2019. The MBE for the 10 m CHM was -0.05 m, indicating that the predicted values were slightly more than the observed values on average (a sign of a slight overestimation). The RMSE for the 10 m CHM was 4.40 m, indicating the average difference between the predicted and observed values was 4.40 m. The R² value for the 10 m CHM was 0.74, indicating a moderate correlation between the predicted and observed values. On the other hand, the MBE for the 30 m CHM was 1.05 m, indicating that the predicted values were higher than the observed values on average. The RMSE for the 30 m CHM was 7.49 m, indicating a larger difference between the predicted and observed values. The R² value for the 30 m CHM was 0.46, indicating a weak correlation between the predicted and observed values. These results put emphasis that the 10 m CHM is more accurate than the 30 m CHM for forest canopy height modeling.

Fig. 4 shows the 10 m and 30 m state-wide CHM of New York and a visual comparison of the generated 10 m New York CHM, the 30 m global CHM, and the original Sentinel-2 imagery. The comparison with the 30 m global CHM demonstrates the advantage of using finer spatial resolution imagery for height mapping. In comparison to the 10 m resolution CHM map generated in the current study, the 30 m global CHM map produced by Potapov et al. (2021) exhibits distinct differences in the spatial distribution of canopy height. Fig. 4 (b) displays a noticeably higher concentration of yellowish areas (heights from 0 to 5 m) in the north, east, and southern regions of New York. This has happened because Potapov et al. (2021) assigned every height below 3 m to zero.

According to Fig. 4 the visual comparison of close-up samples highlights the differences in the spatial resolution and accuracy of the CHM maps and provides a better understanding of the impact of the finer resolution on the results.

Fig. 4 (A) depicts a clear-cut in the northern part of Adirondack Park, which is easily recognizable in the Sentinel-2 imagery. The clear-cut pattern is accurately captured in the 10 m CHM, while the global 30 m CHM only provides a general indication of the disturbance. Fig. 4 (B) demonstrates a linear forest cut in the southern part of Huntington in Adirondack Park. A similar pattern can be seen in that a 10 m CHM can capture more detailed information about the cut and more accurate information on the CHM variation in the area. On the other hand, the 30 m global CHM maps the whole area to a region with an elevation of 0 to 5 m. Fig. 4 (C) illustrates an example of a high-density canopy structure. According to Fig. 4 (C) the Sentinel-2 imagery shows some high-density forests which are mapped correctly using the 30% GEDI validation set in the 10 m CHM with a height of 25 to 35 m. The global 30 m CHM estimates the region with lower canopy height between 20 and 25 m. The comparison results revealed that the 10 m CHM generated in this study showed higher accuracy and finer resolution compared to the global 30 m CHM map. The finer resolution of the 10 m CHM allowed for a more detailed and precise representation of the forest canopy structure. The results of the 10 m CHM of New York provide insight into the effectiveness of using Sentinel-2 imagery and an RF model for accurate height estimation.

In order to investigate the capability of different spectral bands and vegetation indices, variable importance scores were calculated. A Gini-

based approach was used in Python to compare the importance of each variable for the RF model. This method used the calculation of Gini impurity when the decision is required to select a predictor to split a node (Menze et al., 2009). Fig. 5 illustrates the Sentinel-2 predictors with corresponding normalized coefficients that contributed to the RF regression model. These top five vegetation indices and spectral bands are highly related with biomass and chlorophyll content of vegetation, resulting in improving the performance of the regression models.

4.2. AGB mapping

Fig. 6 shows the quantitative results of the RF model for AGB estimation of New York in 2019 using the integration of GEDI-derived 10 m CHM, Sentinel-1, and Sentinel-2 data. According to Fig. 6, among satellite imagery, Sentinel-1 SAR data provides the highest RMSE and MBE (58.66 Mg/ha and 9.47 Mg/ha) and the lowest R² (0.47). Landsat 8 with 30 m resolution presents RMSE, MBE, and R² of 54.53 Mg/ha, 5.76 Mg/ha, and 0.56, respectively, while Sentinel-2 data with 10 m resolution provides RMSE, MBE, and R² of 43.15 Mg/ha, -0.38 Mg/ha, and 0.59. The result demonstrates the importance of spatial resolution and spectral information (e.g., red-edge bands and indices calculated from Sentinel-2 imagery) in the performance of RF model for AGB estimation.

Moreover, the combination of CHM with Sentinel-1 and Sentinel-2 improves the result of AGB estimation. To see model performance of Sentinel-1 and Sentinel-2 adding in combination with CHM, a comparison of GEDI-derived CHM along with Sentinel-1 and Sentinel-2 data was conducted. The results show that adding CHM can improve the performance of the model. The combination of Sentinel-2 and CHM decreased the RMSE and MBE by about 2.12 Mg/ha and 0.209 Mg/ha, respectively and increased the R² by 0.04. Adding CHM to Sentinel-1 predictors also decreased RMSE and MBE by about 3.14 Mg/ha and 6.358 Mg/ha and increased the R² about 0.08. Finally, the combination of GEDI-derived 10 m CHM, Sentinel-1, and Sentinel-2 data indicates promising results (RMSE: 39.49 Mg/ha, MBE: -0.486, and R²: 0.65) for AGB estimation using only spaceborne sensors. The results put emphasis on the importance of CHM and the vertical structure of the trees for a more accurate AGB estimation.

Fig. 7 shows the 10 m and 30 m state-wide AGB of New York and a visual comparison of the generated 10 m New York CHM, the 30 m AGB, and the original Sentinel-2 imagery. The comparison with the 30 m Landsat AGB demonstrates the advantage of using finer spatial resolution imagery for height mapping. In comparison to the 10 m resolution AGB map generated in the current study, the 30 m AGB map generated in our recent paper (Tamiminia et al., 2022) exhibits distinct differences in the spatial distribution of biomass. According to Fig. 7 the visual comparison of close-up samples highlights the differences in the spatial resolution and accuracy of the AGB maps and provides a better understanding of the impact of the finer resolution and red-edge bands on the results.

Fig. 7 (A) depicts a clear-cut in the northern part of Adirondack Park, which is easily recognizable in the Sentinel-2 imagery. The clear-cut pattern is accurately captured in the 10 m AGB, while the Landsat 30 m AGB only provides a general indication of the disturbance. Fig. 7 (B) demonstrates a linear forest cut in the southern part of Huntington in Adirondack park. A similar pattern can be seen in that a 10 m AGB can capture more detailed information about the cut and more accurate information on the AGB variation in the area.

Fig. 7 (C) illustrates an example of a high-density biomass area. The Landsat 30 m AGB estimates the region with lower AGB values. The comparison results revealed that the 10 m AGB generated in this study showed higher accuracy and finer resolution compared to the 30 m AGB map. The finer resolution of the 10 m AGB allowed for more detailed and precise representation of the forest AGB structure. The results of the 10 m AGB of New York provide insight into the effectiveness of using Sentinel-2 imagery and an RF model for accurate height estimation.

Table 1

Results of RF model for CHM estimation of New York for 2019 using the integration of GEDI and Sentinel-2 data and the result of 30 m CHM using Landsat imagery by (Potapov et al., 2021).

| RF model | RMSE (m) | MBE (m) | R ² |
|----------|----------|---------|----------------|
| CHM 10 m | 4.40 | -0.05 | 0.74 |
| CHM 30 m | 7.49 | 1.05 | 0.46 |

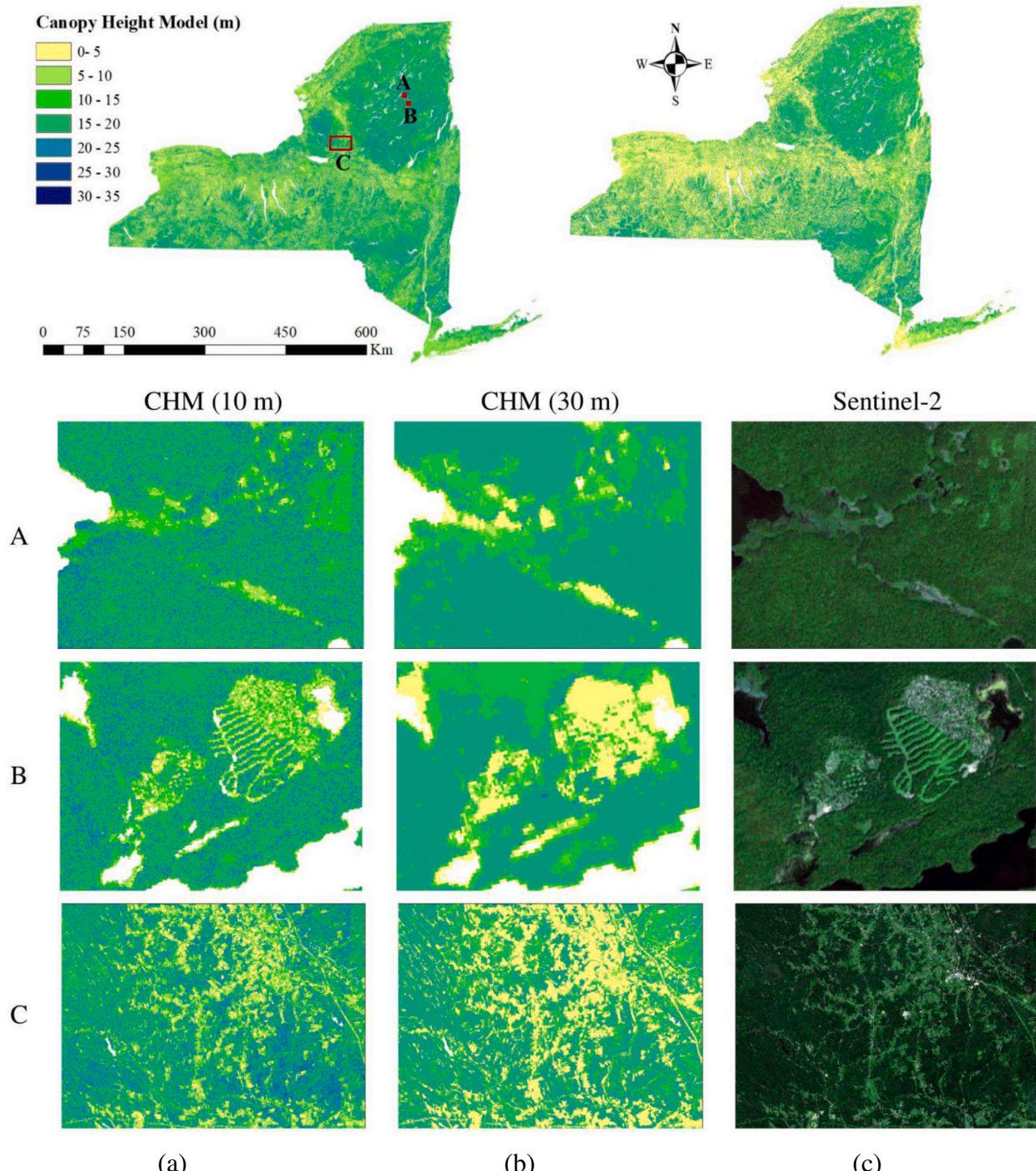


Fig. 4. Three close-up examples (insets A-C) of New York CHM for the year 2019 using GEDI point samples and an RF model; (a) a 10 m CHM using Sentinel-2 imagery; (b) a 30 m CHM using Landsat imagery by (Potapov et al., 2021); (c) Sentinel-2 imagery.

4.3. Map accuracy assessment

To evaluate the maps quantitatively, we adopted a map accuracy assessment protocol developed by Riemann et al. (2010). This protocol is designed to systematically measure the accuracy of the generated maps by comparing them to reference data or ground truth information. The map accuracy assessment process consists of comparing the predicted values from the generated maps to the actual or observed values at specific locations or pixels. In this study, the accuracy assessments

were carried out for multiple locations within the study area, which was divided into 216,500 hexagonal cells. Each hexagonal cell represents a specific spatial unit in the study area. By aggregating the accuracy assessments within the hexagonal cells, the study obtains a comprehensive evaluation of the model's performance across the entire study area. This aggregated evaluation provides valuable insights into the overall accuracy and reliability of the generated maps.

Fig. 8 demonstrates the mapped differences in modeled area estimates at the 216,500 ha scale with respect to plot-based confidence

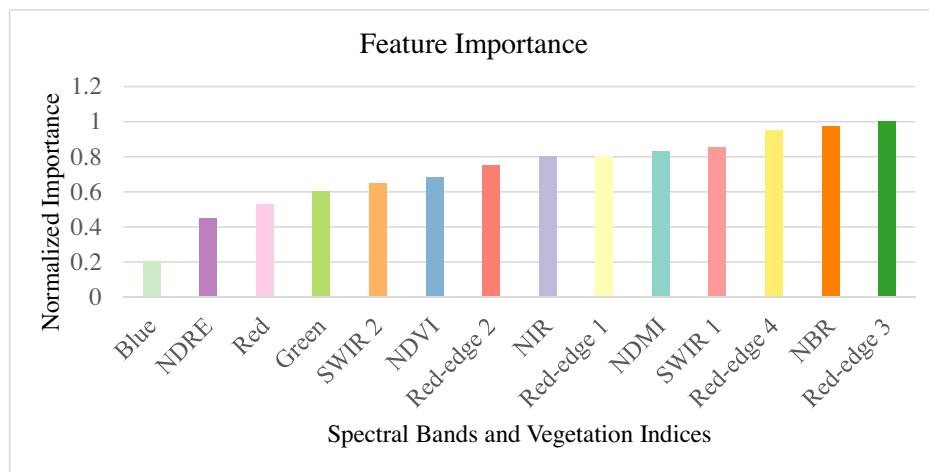


Fig. 5. Normalized feature importance of input predictors for the RF regression model using Sentinel-2 predictors.

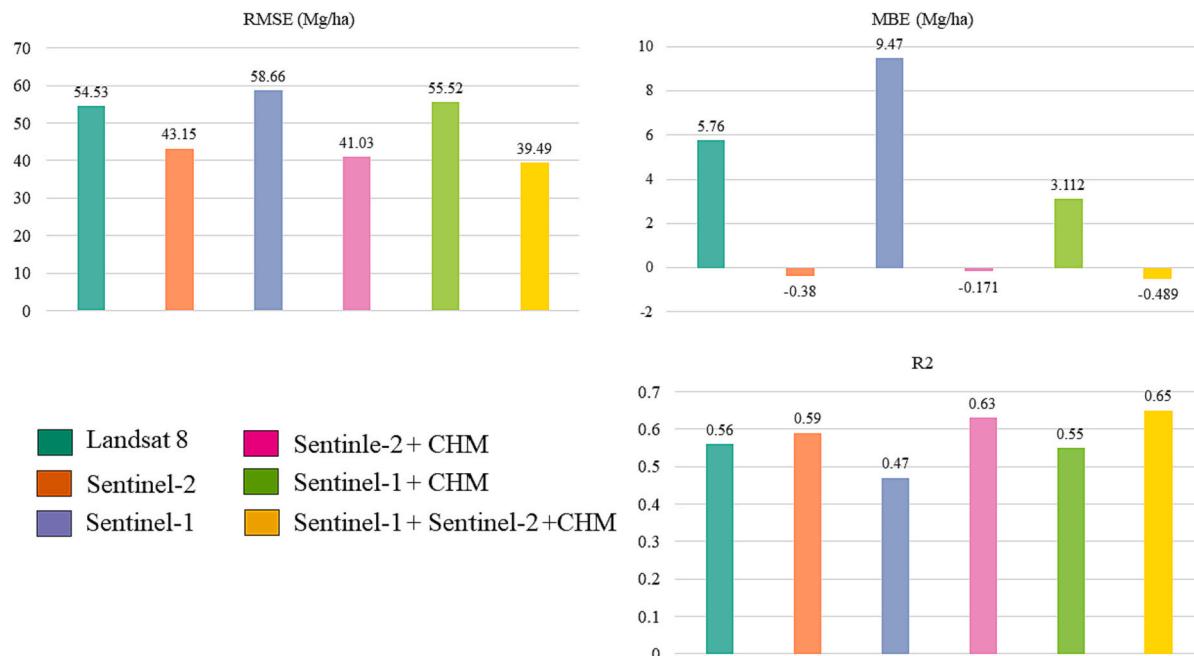


Fig. 6. Results of RF model for AGB estimation of New York for 2019 using the combination of GEDI-derived CHM, Sentinel-1, and Sentinel-2 data.

intervals. The mean biomass values fall within the 95% CI associated with the plot means of 86% of the time. According to Fig. 8, the largest errors occur outside the airborne LiDAR AGB pilot areas. It should be noted that training/testing datasets are within 4 pilot areas; however, the proposed model predicts the AGB for the entire NYS.

5. Discussion

5.1. Potential of GEDI data for large-scale CHM mapping

The GEDI mission is a laser-based technology that is capable of producing high-resolution and large-scale canopy height maps (Leite et al., 2022). These measurements can be used to determine the height of the vegetation canopy and create a CHM. The GEDI data have several advantages over other methods for CHM mapping, making it a valuable tool for large-scale CHM mapping (Potapov et al., 2021). One of the advantages of GEDI data is that it provides profiles of forest canopies with 25 m diameter footprints (Leite et al., 2022). This fine spatial resolution enables the creation of CHMs that accurately represent the

topography of the landscape and the height of the vegetation canopy.

Another advantage of GEDI data is its ability to penetrate the forest canopy. Unlike passive sensor satellite-based methods, which are often unable to penetrate the dense forest canopy, the GEDI laser instrument is able to provide height measurements even in heavily forested areas (Dhargay et al., 2022; C. Liu and Wang, 2022). This allows for the creation of CHMs that accurately represent the height of the vegetation canopy, even in dense forests. Its fine spatial resolution, ability to penetrate the forest canopy, and information on the vertical structure of the vegetation canopy make it a valuable tool for a variety of applications, especially CHM mapping on a large scale.

Several papers have utilized spaceborne LiDAR technology along with MODIS or Landsat imagery for large-scale CHM and AGB mapping (Baccini et al., 2008; Chi et al., 2015; Potapov et al., 2021; Simard et al., 2011). Our results align with the outcomes of the mentioned research, indicating the potential of spaceborne LiDAR for large-scale CHM mapping. All studies agree on the high correlation existing between spaceborne LiDAR and other optical imagery including MODIS, Landsat, and Sentinel-2 (Baccini et al., 2008; Chi et al., 2015; Potapov et al.,

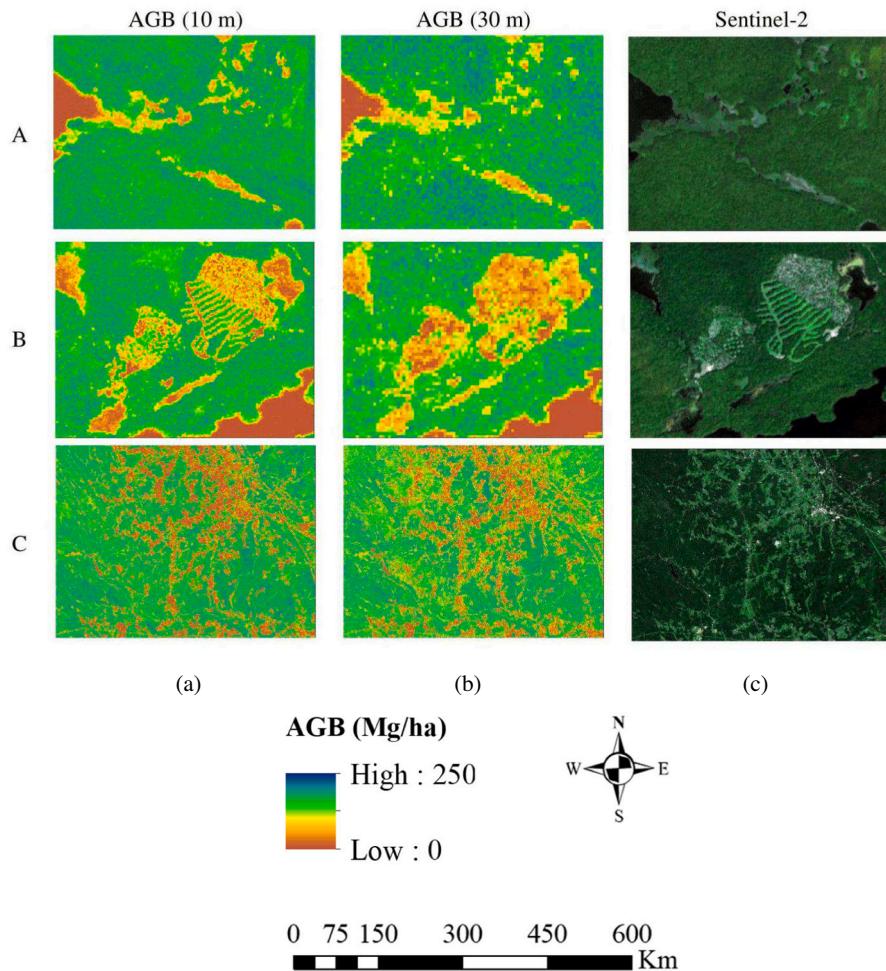


Fig. 7. Three close-up examples (insets A-C) of New York AGB for the year 2019 using the combination of GEDI-derived CHM, Sentinel-1, Sentinel-2, and an RF model; (a) a 10 m AGB; (b) a 30 m AGB using Landsat imagery; (c) Sentinel-2 imagery.

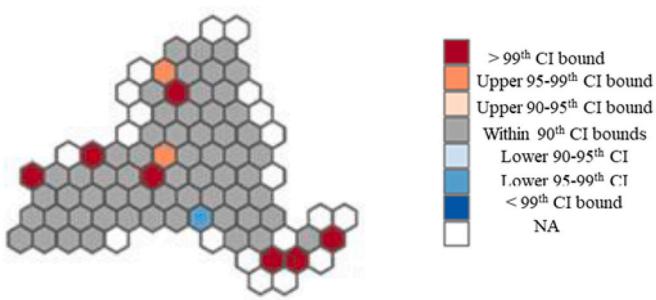


Fig. 8. Mapped differences in modeled area estimates at the 216,500 ha scale with respect to plot-based confidence intervals.

2021; Simard et al., 2011). Simard et al. (2011) leveraged GLAS data and incorporated MODIS imagery, achieving an RMSE of 6.1 m and R² of 0.5. In a more recent study, Potapov et al. (2021) utilized GEDI data for the same purpose, reporting an RMSE of 6.6 m and an improved R² of 0.62. Notably, our study generates a 10 m statewide CHM map with the RMSE of 4.4 m and R² of 0.74.

While it is true that ALS data provides higher accuracy in AGB estimation, it has certain limitations in terms of coverage and cost. ALS data is collected on a smaller scale and may not be available for large areas or remote regions. In contrast, GEDI spaceborne LiDAR offers a significant

advantage in terms of global coverage, making it a valuable resource for AGB mapping in areas where ALS data might be scarce or unavailable.

The cost associated with operating ALS instruments and collecting data can be substantial, especially when considering extensive study areas. GEDI, as a spaceborne sensor, provides a cost-effective alternative. Moreover, the issue of aligning and matching ALS data with other datasets becomes a concern, particularly when dealing with multiple data sources collected at different time periods. GEDI data, on the other hand, provides a consistent dataset from 2019 onwards, eliminating the need to worry about temporal alignment and ensuring a standardized approach to AGB estimation.

5.2. Importance of CHM for AGB mapping

Canopy height is an important characteristic of forests, as it has direct implications for a variety of ecological processes, including the carbon cycle and energy balance (Potapov et al., 2021). In forest ecology, AGB is a critical metric that describes the amount of organic matter in a forest ecosystem (Fernández-Guisuraga et al., 2022). AGB is a function of both tree species and forest structure and is a crucial indicator of the health and productivity of a forest. Canopy height is a critical factor in determining AGB, as taller trees tend to have greater biomass. For this reason, the use of accurate CHM data is becoming increasingly common in AGB mapping.

The use of CHM data in AGB mapping is particularly important in tropical forests, where the dense canopy structure and high levels of

vegetation make it difficult to obtain accurate AGB estimates using other remote sensing techniques (Zhang et al., 2021). However, in this study, the integration of CHM with other optical and SAR data such as Sentinel-1 and Sentinel-2 were used for temperate forests in New York. The use of a 10 m GEDI-derived CHM map improved the RMSE by 3.66 Mg/ha and R² by 0.06.

Therefore, CHM is an important tool for AGB mapping, as it provides valuable information about the structure of forest ecosystems. The integration of CHM data with other remote sensing data can help to improve the accuracy of AGB estimates, particularly in dense and complex forest environments.

5.3. Integration of GEDI, Sentinel-1, and Sentinel-2 for CHM and AGB

GEDI data provides fine-resolution height information, which is critical for CHM mapping (Mandal et al., 2020). SAR signals are sensitive to the geometric structure of trees (e.g., size, shape, orientation and distribution of leaves, stalks and fruits) and dielectric properties of the crop canopy, as well as soil characteristics (e.g., roughness and water content) (Tamiminia et al., 2017). The data from Sentinel-1 can be used to detect and map large-scale forest structures, such as clear-cuts and forest edges, which are important for CHM mapping. On the other hand, Sentinel-2 provides data on the spectral characteristics of forests, which can be used to map different forest types and identify areas of forest disturbance. This information can be used to estimate the above ground amount of biomass in the vegetation canopy by analyzing the spectral signature of the vegetation. Sentinel-2 provides fine-resolution data in multiple spectral bands, making it an important tool for mapping AGB.

The MBE, RMSE, and R² values reflect the effect of finer spatial resolution in creating the CHM. Sentinel-2 provides imagery with higher spatiotemporal resolution (10 m ~ 20 m, revisit cycle within 6 days) compared to Landsat data (W. Li et al., 2020). In addition, Sentinel-2 contains three red-edge bands which can improve the results of CHM mapping. Red-edge bands are important indicators of vegetation growth; thus, they can provide valuable information about canopy height (Hua and Zhao, 2021). Several studies have proven the advantage of red-edge bands for vegetation mapping (W. Li et al., 2020; Bayle et al., 2019; Hua and Zhao, 2021). A recent study showed that the Sentinel-2 multi-spectral and texture metrics are very promising co-variables in the country-level high-resolution mapping of vegetation height when coupled with the airborne LiDAR-derived vegetation height data (W. Li et al., 2020).

The results indicate that the combination of S1 and S2 with CHM (Sentinel-1 + Sentinel-2 + CHM) model has the best performance in terms of RMSE and MBE (39.49 Mg/ha and -0.489) with the highest R² (0.65) value. Thus, the integration of GEDI, Sentinel-1, and Sentinel-2 data offers a powerful tool for creating a fine resolution CHM and AGB on a large scale. These three sensors can provide complementary textural, spectral, and vertical information that can be used to create highly accurate CHM and AGB maps. In addition, integrating GEDI data with Sentinel-1 and Sentinel-2 data can help to correct for potential biases in the CHM and AGB mapping products, improving their accuracy and reliability. The findings contribute to the field of remote sensing and provide valuable information for understanding the role of forests in the global carbon cycle.

6. Conclusion

This study demonstrates the potential of integrating GEDI, Sentinel-1, and Sentinel-2 data for generating 10 m CHM and AGB maps at a large extent, in this case, New York for the year 2019. Our results show that the combination of these datasets improves the accuracy of CHM and AGB estimates compared to using them individually. To obtain the 10 m CHM, the RH 95th percentile measurements of GEDI footprints were extrapolated using Sentinel-2 imagery and an RF regression model. The comparison of generated 10 m CHM (MBE: -0.05 m, RMSE: 4.4 m, and

R²: 0.74) and the global 30 m resolution CHM (MBE: 1.05 m, RMSE: 7.49 m, and R²: 0.46) demonstrated the promising capability of GEDI spaceborne LiDAR and Sentinel-2 data for accurate state-wide CHM mapping. The visual comparison of the 10 m CHM, the 30 m CHM, and the original Sentinel-2 image further highlights the advantage of using high-resolution remote sensing data for generating CHMs. Sentinel-2 satellite imagery provides finer spatiotemporal resolution compared to Landsat data, and its three red-edge bands play an important role in mapping vegetation growth and canopy height. In addition, we also highlighted the importance of CHM data for accurately estimating AGB. Thus, the created 10 m CHM along with Sentinel-1 backscatter information and Sentinel-2 spectral indices, was used to produce an AGB map of New York in 2019. Airborne LiDAR-derived AGB rasters were utilized as training/testing data to train and evaluate the performance of the RF model. Quantitative and visual analysis of the 10 m State-wide AGB map (MBE: 0.38 Mg/ha, RMSE: 43.15 Mg/ha, and R²: 0.59) with our previously Landsat-based 30 m AGB map (MBE: 5.76 Mg/ha, RMSE: 54.53 Mg/ha, and R²: 0.56) showed a remarkable improvement in capturing all ranges of biomass from low to high. Our methodology can be applied to other regions to improve the accuracy and spatial resolution of CHM and AGB estimates. We hope that our findings will encourage further exploration of the potential of combining remote sensing datasets for forest mapping and carbon accounting.

CRediT authorship contribution statement

Haifa Tamiminia: Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **Bahram Salehi:** Funding acquisition. **Masoud Mahdianpari:** Data curation. **Tristan Goulden:** Writing – review & editing.

Data availability

Data will be made available on request.

Acknowledgment

This project was supported by the Climate & Applied Forest Research Institute at SUNY ESF with funding from the NYS Department of Environmental Conservation and as well as by a NIFA (National Institute of Food and Agriculture) McIntire-Stennis grant. We would also like to thank Dr. Colin M. Beier, Lucas K. Johnson and Michael J. Mahoney for providing the airborne LiDAR-derived AGB rasters.

References

- Ahmed, Oumer S., Franklin, Steven E., Wulder, Michael A., White, Joanne C., 2015. Characterizing stand-level Forest canopy cover and height using Landsat time series, samples of airborne LiDAR, and the random Forest algorithm. *ISPRS J. Photogramm. Remote Sens.* 101, 89–101.
- Albright, Thomas A., 2018. Forests of New York, 2017. In: Resource Update FS-170. Newtown Square, PA: US Department of Agriculture, Forest Service, Northern Research Station, 4, p. 170, 1–4.
- Baccini, A., Laporte, N., Goetz, S.J., Sun, M., Dong, H., 2008. A first map of tropical Africa's above-ground biomass derived from satellite imagery. *Environ. Res. Lett.* 3 (4), 045011.
- Bayle, Arthur, Carlson, Bradley Z., Thierion, Vincent, Isenmann, Marc, Choler, Philippe, 2019. Improved mapping of mountain Shrublands using the Sentinel-2 red-edge band. *Remote Sens.* 11 (23), 2807.
- Chavan, Balbhim, Rasal, Ganesh, 2012. Total sequestered carbon stock of *Mangifera Indica*. *J. Environ. Earth Sci.* 2 (1).
- Chi, Hong, Sun, Guoqing, Huang, Jinliang, Guo, Zhifeng, Ni, Wenjian, Anmin, Fu., 2015. National forest aboveground biomass mapping from ICESat/GLAS data and MODIS imagery in China. *Remote Sens.* 7 (5), 5534–5564.
- Claverie, Martin, Junchang, Ju, Masek, Jeffrey G., Dungan, Jennifer L., Vermote, Eric F., Roger, Jean-Claude, Skakun, Sergii V., Justice, Christopher, 2018. The harmonized Landsat and Sentinel-2 surface reflectance data set. *Remote Sens. Environ.* 219, 145–161.
- Dhargay, Sonam, Lyell, Christopher S., Brown, Tegan P., Inbar, Assaf, Sheridan, Gary J., Lane, Patrick N.J., 2022. Performance of GEDI space-borne LiDAR for quantifying

- structural variation in the temperate forests of South-Eastern Australia. *Remote Sens.* 14 (15), 3615.
- Dorado-Roda, Iván, Pascual, Adrián, Godinho, Sergio, Silva, Carlos A., Botequim, Brigitte, Rodríguez-Gonzálvez, Pablo, González-Ferreiro, Eduardo, Guerra-Hernández, Juan, 2021. Assessing the accuracy of GEDI data for canopy height and aboveground biomass estimates in Mediterranean forests. *Remote Sens.* 13 (12), 2279.
- Du, Yun, Zhang, Yihang, Ling, Feng, Wang, Qunming, Li, Wenbo, Li, Xiaodong, 2016. Water Bodies' mapping from Sentinel-2 imagery with modified normalized difference water index at 10-m spatial resolution produced by sharpening the SWIR band. *Remote Sens.* 8 (4), 354.
- Dubayah, Ralph, Hofton, Michelle, Blair, James, Armston, John, Tang, Hao, Luthcke, Scott, 2020. GEDI L2A Elevation and Height Metrics Data Global Footprint Level V001. NASA EOSDIS Land Processes DAAC. https://doi.org/10.5067/GEDI/GEDI02_A.001.
- Duncanson, Laura, Neuenschwander, Amy, Hancock, Steven, Thomas, Nathan, Fatoyinbo, Temilola, Simard, Marc, Silva, Carlos A., Armston, John, Luthcke, Scott B., Hofton, Michelle, 2020. Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County, California. *Remote Sens. Environ.* 242, 111779 <https://doi.org/10.1016/j.rse.2020.111779>.
- FAO, 2020. Global Forest Resources Assessment 2020 – Key Findings. <https://doi.org/10.4060/ca8753en>.
- Fernández-Guisuraga, José Manuel, Suárez-Seaone, Susana, Fernandes, Paulo M., Fernández-García, Víctor, Fernández-Manso, Alfonso, Quintano, Carmen, Calvo, Leonor, 2022. Pre-fire aboveground biomass, estimated from LiDAR, spectral and field inventory data, as a major driver of burn severity in maritime pine (*Pinus Pinaster*) ecosystems. *For. Ecosyst.* 9, 100022.
- Francini, Saverio, D'Amico, Giovanni, Vangi, Elia, Borghi, Costanza, Chirici, Gherardo, 2022. Integrating GEDI and Landsat: Spaceborne Lidar and four decades of optical imagery for the analysis of forest disturbances and biomass changes in Italy. *Sensors* 22 (5), 2015.
- Gorelick, Noel, Hancher, Matt, Dixon, Mike, Ilyushchenko, Simon, Thau, David, Moore, Rebecca, 2017. Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>.
- Hirata, Yasumasa, Furuya, Naoyuki, Saito, Hideki, Pak, Chealy, Leng, Chivin, Sokhi, Heng, Ma, Vuthy, Kajisa, Tsuyoshi, Ota, Tetsuji, Mizoue, Nobuya, 2018. Object-based mapping of aboveground biomass in tropical forests using LiDAR and very-high-spatial-resolution satellite data. *Remote Sens.* 10 (3), 438. <https://doi.org/10.3390/rs10030438>.
- Hua, Yiyi, Zhao, Xuesheng, 2021. Multi-model estimation of Forest canopy closure by using red edge bands based on Sentinel-2 images. *Forests* 12 (12), 1768.
- Hudak, Andrew T., Lefsky, Michael A., Cohen, Warren B., Berterretche, Mercedes, 2002. Integration of Lidar and Landsat ETM+ data for estimating and mapping Forest canopy height. *Remote Sens. Environ.* 82 (2–3), 397–416.
- Hudak, Andrew T., Fekety, Patrick A., Kane, Van R., Kennedy, Robert E., Filippelli, Steven K., Falkowski, Michael J., Tinkham, Wade T., Smith, Alistair M.S., Crookston, Nicholas L., Domke, Grant M., 2020. A carbon monitoring system for mapping regional, annual aboveground biomass across the Northwestern USA. *Environ. Res. Lett.* 15 (9), 095003 <https://doi.org/10.1088/1748-9326/ab93f9>.
- Hussin, Younis Ali, Gilani, Hammad, Louise van Leeuwen, M.S.R., Murthy, Rachna Shah, Baral, Srijana, Tsendbazar, Nandin-Erdene, Shrestha, Saurav, Shah, Shyam Kumar, Qamer, Faisal Mueen, 2014. Evaluation of object-based image analysis techniques on very high-resolution satellite image for biomass estimation in a watershed of hilly Forest of Nepal. *Appl. Geomat.* 6 (1), 59–68.
- Johnson, Lucas K., Mahoney, Michael J., Bevilacqua, Eddie, Stehman, Stephen V., Domke, Grant M., Beier, Colin M., 2022. Fine-Resolution Landscape-Scale Biomass Mapping Using a Spatiotemporal Patchwork of LiDAR Coverages. *International Journal of Applied Earth Observation and Geoinformation* 114, 103059.
- Jong, De, Steven, M., Pebesma, Edzer J., Lacaze, Bernard, 2003. Above-ground biomass assessment of Mediterranean forests using airborne imaging spectrometry: the DAIS-Peyne experiment. *Int. J. Remote Sens.* 24 (7), 1505–1520.
- Kaplan, Gordana, Avdan, Ugur, 2018. Sentinel-2 Pan sharpening—comparative analysis. In: Proceedings. 2. MDPI, p. 345. <https://www.mdpi.com/2504-3900/2/7/345>.
- Kelsey, Katharine C., Neff, Jason C., 2014. Estimates of aboveground biomass from texture analysis of Landsat imagery. *Remote Sens.* 6 (7), 6407–6422.
- Lee, Jong-Sen, Pottier, Eric, 2009. Polarimetric Radar Imaging: From Basics to Applications. CRC Press.
- Leite, Rodrigo Vieira, Silva, Carlos Alberto, Broadbent, Eben North, Do Amaral, Cibele Hummel, Liesenberg, Veraldo, De Almeida, Danilo Roberti Alves, Mohan, Midhun, Godinho, Sérgio, Cardil, Adrian, Hamamura, Caio, 2022. Large scale multi-layer fuel load characterization in tropical savanna using GEDI Spaceborne Lidar data. *Remote Sens. Environ.* 268, 112764.
- Li, Manqi, Im, Jungho, Beier, Colin, 2013. Machine learning approaches for Forest classification and change analysis using multi-temporal Landsat TM images over Huntington wildlife Forest. *GISci. Remote Sens.* 50 (4), 361–384. <https://doi.org/10.1080/15481603.2013.819161>.
- Li, Zhongbin, Zhang, Hankui K., Roy, David P., Yan, Lin, Huang, Haiyan, Li, Jian, 2017. Landsat 15-m panchromatic-assisted downscaling (LPAD) of the 30-m reflective wavelength bands to Sentinel-2 20-m resolution. *Remote Sens.* 9 (7), 755.
- Li, Yingchang, Li, Chao, Li, Mingyang, Liu, Zhenzhen, 2019. Influence of variable selection and Forest type on Forest aboveground biomass estimation using machine learning algorithms. *Forests* 10 (12), 1073. <https://doi.org/10.3390/f10121073>.
- Li, Wang, Niu, Zheng, Shang, Rong, Qin, Yuchu, Wang, Li, Chen, Hanyue, 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.
- Lisein, Jonathan, Pierrot-Deseilligny, Marc, Bonnet, Stéphanie, Lejeune, Philippe, 2013. A photogrammetric workflow for the creation of a Forest canopy height model from small unmanned aerial system imagery. *Forests* 4 (4), 922–944.
- Liu, C., Wang, S., 2022. Estimating tree canopy height in densely forest-covered mountainous areas using gedi spaceborne full-waveform data. In: ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences, 1.
- Liu, Xiaoqiang, Yanjun, Su, Tianyu, Hu, Yang, Qiuli, Liu, Bingbing, Deng, Yufei, Tang, Hao, Tang, Zhiyao, Fang, Jingyun, Guo, Qinghua, 2022. Neural network guided interpolation for mapping canopy height of China's forests by integrating GEDI and ICESat-2 data. *Remote Sens. Environ.* 269, 112844.
- Ma, Lei, Liu, Yu, Zhang, Xueliang, Ye, Yuanxin, Yin, Gaofei, Johnson, Brian Alan, 2019. Deep learning in remote sensing applications: a Meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* 152, 166–177.
- Mandal, Dipankar, Kumar, Vineet, Ratha, Debanshu, Dey, Subhadip, Bhattacharya, Avik, Lopez-Sánchez, Juan M., McNairn, Heather, Rao, Yalamanchili S., 2020. Dual Polarimetric radar vegetation index for crop growth monitoring using Sentinel-1 SAR data. *Remote Sens. Environ.* 247, 111954.
- Menze, B.H., Kelm, B.M., Masuch, R., Himmelreich, U., Bachert, P., Petrich, W., Hamprecht, F.A., 2009. A comparison of random forest and its Gini importance with standard chemometric methods for the feature selection and classification of spectral data. *BMC bioinformatics* 10, 1–16.
- Mutanga, Onisimo, Adam, Elhadi, Cho, Moses Azong, 2012. High density biomass estimation for wetland vegetation using WorldView-2 imagery and random forest regression algorithm. *Int. J. Appl. Earth Obs. Geoinf.* 18, 399–406.
- Potapov, Peter, Li, Xinyuan, Hernandez-Serna, Andres, Tyukavina, Alexandra, Hansen, Matthew C., Kammareddy, Anil, Pickens, Amy, Turubanova, Svetlana, Tang, Hao, Silva, Carlos Edibalido, 2021. Mapping global Forest canopy height through integration of GEDI and Landsat data. *Remote Sens. Environ.* 253, 112165.
- Qi, Wenlu, Dubayah, Ralph O., 2016. Combining Tandem-X InSAR and simulated GEDI Lidar observations for Forest structure mapping. *Remote Sens. Environ.* 187, 253–266.
- Qi, Wenlu, Lee, Seung-Kuk, Hancock, Steven, Luthcke, Scott, Tang, Hao, Armston, John, Dubayah, Ralph, 2019. Improved Forest height estimation by fusion of simulated GEDI Lidar data and TanDEM-X InSAR data. *Remote Sens. Environ.* 221, 621–634.
- Riemann, Rachel, Wilson, Barry Tyler, Lister, Andrew, Parks, Sarah, 2010. An effective assessment protocol for continuous geospatial datasets of Forest characteristics using USFS Forest inventory and analysis (FIA) data. *Remote Sens. Environ.* 114 (10), 2337–2352.
- Saarela, Svetlana, Holm, Sören, Healey, Sean P., Andersen, Hans-Erik, Petersson, Hans, Prentius, Wilmer, Patterson, Paul L., Næsset, Erik, Gregoire, Timothy G., Ståhl, Göran, 2018. Generalized hierarchical model-based estimation for aboveground biomass assessment using GEDI and Landsat data. *Remote Sens.* 10 (11), 1832.
- Saatchi, Sassan, 2019. SAR methods for mapping and monitoring forest biomass. In: Book: Flores A, Herndon K, Thapa R, Cherrington E, Eds. SAR Handbook: Comprehensive Methodologies for Forest Monitoring and Biomass Estimation, pp. 207–246.
- Simard, Marc, Pinto, Naiara, Fisher, Joshua B., Baccini, Alessandro, 2011. Mapping forest canopy height globally with spaceborne lidar. *J. Geophys. Res. Biogeosci.* 116 (G4).
- Tamiminia, 2021. A comparison of random forest and light gradient boosting machine for forest above-ground biomass estimation using a combination of landsat, Alos Palsar, and Airborne Lidar Data. *Int. Arch. Photogramm. XLIV-M-3-2021*, 163–168.
- Tamiminia, Haifa, Homayouni, Saeid, Hansen, Matthews, Erik, Gregoire, Timothy G., Safari, Abdoreza, 2017. A particle swarm optimized kernel-based clustering method for crop mapping from multi-temporal Polarimetric L-band SAR observations. *Int. J. Appl. Earth Obs. Geoinf.* 58, 201–212.
- Tamiminia, Haifa, Salehi, Bahram, Mahdianpari, Masoud, Quackenbush, Lindi, Adeli, Sarina, Brisco, Brian, 2020. Google earth engine for geo-big data applications: a meta-analysis and systematic review. *ISPRS J. Photogramm. Remote Sens.* 164, 152–170.
- Tamiminia, Haifa, Salehi, Bahram, Mahdianpari, Masoud, Beier, Colin M., Klimkowski, Daniel J., Volk, Timothy A., 2021. Comparison of machine and deep learning methods to estimate shrub willow biomass from UAS imagery. *Can. J. Remote. Sens.* 47 (2), 209–227.
- Tamiminia, Haifa, Salehi, Bahram, Mahdianpari, Masoud, Beier, Colin M., Johnson, Lucas, Phoenix, Daniel B., 2021a. A comparison of decision tree-based models for forest above-ground biomass estimation using a combination of airborne lidar and landsat data. In: ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences, 3.
- Tamiminia, Haifa, Salehi, Bahram, Mahdianpari, Masoud, Beier, Colin M., Johnson, Lucas, Phoenix, Daniel B., Mahoney, Michael, 2022. Decision tree-based machine learning models for above-ground biomass estimation using multi-source remote sensing data and object-based image analysis. *Geocarto Int.* 1–29.
- Urbazaev, Mikhail, Thiel, Christian, Cremer, Felix, Dubayah, Ralph, Migliavacca, Mirco, Reichstein, Markus, Schmullius, Christiane, 2018. Estimation of Forest aboveground biomass and uncertainties by integration of field measurements, airborne LiDAR, and SAR and optical satellite data in Mexico. *Carb. Balance Manag.* 13 (1), 1–20. <https://doi.org/10.1186/s13021-018-0093-5>.
- Wang, Qunming, Shi, Wenzhong, Li, Zhongbin, Atkinson, Peter M., 2016. Fusion of Sentinel-2 images. *Remote Sens. Environ.* 187, 241–252.
- Wang, Jing, Huang, Bo, Zhang, Hankui K., Ma, Peifeng, 2019. Sentinel-2A image fusion using a machine learning approach. *IEEE Trans. Geosci. Remote Sens.* 57 (12), 9589–9601.
- Xi, Yanbiao, Tian, Qinqiu, Zhang, Wenmin, Zhang, Zhichao, Tong, Xiaoye, Brandt, Martin, Fensholt, Rasmus, 2022. Quantifying understory vegetation density

- using multi-temporal Sentinel-2 and GEDI LiDAR data. *GISci. Remote Sens.* 59 (1), 2068–2083.
- Zanaga, Daniele, Van De Kerchove, Ruben, Daems, D., De Keersmaecker, W., Brockmann, Carsten, Kirches, G., Wevers, Jan, Cartus, O., Santoro, M., Fritz, S., 2022. *ESA WorldCover 10 m 2021 V200*.
- Zhang, He, Bauters, Marijn, Boeckx, Pascal, Van Oost, Kristof, 2021. Mapping Canopy Heights in dense tropical forests using low-cost UAV-derived photogrammetric point clouds and machine learning approaches. *Remote Sens.* 13 (18), 3777.
- Zheng, G., Chen, J.M., Tian, Q.J., Ju, W.M., Xia, X.Q., 2007. Combining remote sensing imagery and forest age inventory for biomass mapping. *Journal of Environmental Management* 85 (3), 616–623.