Journal of Applied and Natural Science
17(1), 52 - 64 (2025)
ISSN : 0974-9411 (Print), 2231-5209 (Online)
journals.ansfoundation.org

Research Article

Integration of ICESat/GLAS data and random forest to estimate canopy height and biomass in Central Indian Forest

Alekhya VVL Padma
AMNEX Infotechnologies Pvt. Ltd. Ahmedabad-380001 (Gujarat) India
Mohammad Suhail*
Centre of Applied Remote Sensing and GIS Applications, Samarkand State University named after Sharof Rashidov, Samarkand, 140104, Uzbekistan

*Corresponding author. Email: netgeo.suhail@gmail.com

Article Info
https://doi.org/10.31018/jans.v17i1.5860
Received: June 10, 2024
Revised: January 25, 2025
Accepted: January 30, 2025

How to Cite
Padma,A. VVL and Suhail, M. (2025). Integration of ICESat/GLAS data and random forest to estimate canopy height and biomass in Central Indian Forest. *Journal of Applied and Natural Science*, 17(1), 52 - 64. https://doi.org/10.31018/jans.v17i1.5860

Abstract
Satellite lidar systems, such as ICESat, GEDI, and ICESat-2, have revolutionized global above-ground biomass (AGB) estimation by providing precise forest height data. These missions highlight the transformative potential of spaceborne lidar in advancing biomass assessment and forest monitoring. The present research effectively utilized spaceborne ICESat-1 LiDAR to measure forest canopy height and estimate above-ground biomass (AGB) in Madhya Pradesh's Central Indian Forest. Data from LiDAR, radar, optical, and digital elevation models were integrated with ancillary climate variables and field-based forest inventory during 2009-10. Further, this approach was validated against new data obtained from ICESAT-2 (October 2018 onwards), GEDI (March 2019 onwards), and other spaceborne LiDAR sensors by numerous researchers globally. Lorey's height method established a relationship between GLAS-derived AGB and other variables, creating a spatial–height map using a K-Nearest Neighbour-based random forest approach. Estimated forest canopy height ranged from 2.16 m to 17.63 m with an RMSE of ± 2.57 m. Spatial AGB was estimated for prominent forest types, with R^2 values ranging from 0.62 to 0.71 ($p < 0.01$). The total AGB over Madhya Pradesh's forests were estimated at 315.77 Mt with an RMSE of ±19.22 t ha^$^{-1}$. Relative errors ranged from 33% to 45% over different forest types, suggesting newer missions like ICESat/GLAS-2 are needed for more precise estimates in future research.

Keywords: Above Ground Biomass, ICESat/GLAS, k- Nearest Neighbour, Madhya Pradesh, Random Forest, Tree height

INTRODUCTION

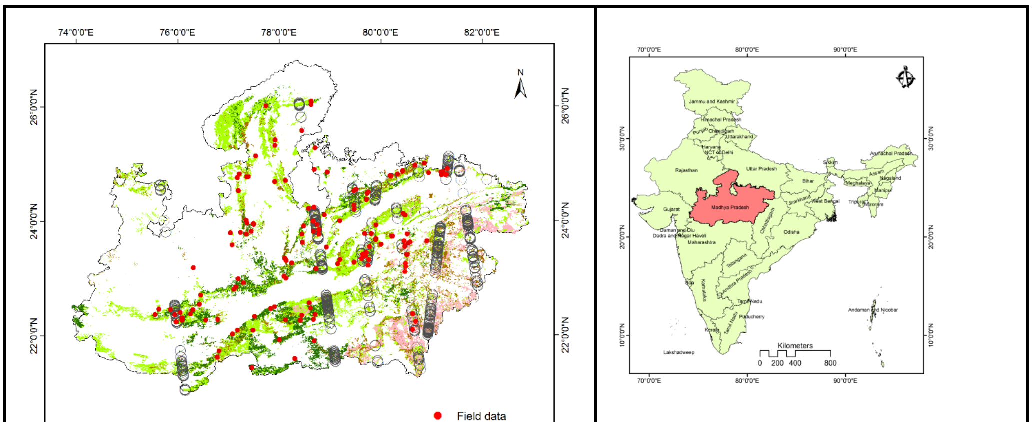
Due to growing worries about global climate change, above-ground biomass (AGB) assessment has garnered more attention recently, particularly in tropical forests. A crucial factor in determining how well a forest stores carbon is biomass, which serves as a substitute for potential carbon emissions from deforestation or conversion to non-forest land use/cover (Mora *et al.,* 2013). According to Ahern *et al.* (1998), data from remote sensing is the most effective means of estimating biomass over a wide area with the help of small-scale field measurements. It has also been demonstrated that it is an important method for monitoring international carbon emissions protocols, such as the Kyoto Protocol. However, spatial AGB has been estimated using satellite remote sensing techniques except for the highest levels of biomass (over 400 t ha^{-1}), which is difficult to quantify due to the saturation of remote sensing signals at both optical and microwave sensors. Previous research work has indicated that the intermediate biomass levels (about 250 t ha^{-1}) have been studied effectively using remote sensing data (Madugundu *et al.,* 2008; Thumaty *et al.,* 2015). FAO conducted a benchmark study on forest mapping to estimate biomass and forest change in Tropical forests (Brown, 1997), which was later extended to study biomass in three continents by Saatchi *et al.* (2011). Forest biomass has also been estimated using airborne SAR tomography (Ramachandran *et al.,* 2023). LiDAR – Light Detection and Ranging system has demonstrated an effective way of extracting forest height as it works on the principle of active remote sensing and relies upon a laser scanning approach (Lefsky and Harding *et al.,* 1999; Lefsky *et al.,* 2005; Lefsky and Harding *et al.,* 2005; Chen and Wang *et al.,* 2023; Lu and Jiang, 2024). Accurate estimations of forest biomass across many biomes have been made possible by this system while exploiting the comprehensive knowledge of the vertical structure of forest vegetation (Nelson *et al.,* 2009). The first spaceborne LiDAR sensor intended for continuous worldwide Earth observation was the Geoscience Laser Altimeter System (GLAS), which was installed on the Ice, Cloud, and Land Elevation Satellite (ICESat) (Zwally *et al*., 2002). This full waveform spaceborne LiDAR system offers a large footprint for adequate information to determine forest height. The large-scale mapping of forest heights has been made possible by the comparatively large footprint and spatial coverage of the GLAS data (Lefsky and Harding *et al.,* 2005). The system used a 1064 nm laser wavelength and collected full-waveform altimetry data between 2003 and 2009. The LiDAR footprints stretch out along the track by 172 meters and have an estimated diameter of 70 meters (Schutz *et al.,* 2005). Several characteristics related to forest properties have been retrieved from full waveform GLAS data to measure forest height and AGB. Lefsky *et al.* (2005) presented a technique to calculate the maximum canopy height using the terrain index—which is obtained from an extra data set—and waveform extent, which refers to the interval between signal start and signal wrap-up. This approach was further enhanced with a leading-trailing edge technique to eliminate the need for an extra data set (Lefsky *et al.,* 2007, Lefsky, 2009 and 2010). To estimate forest biomass, spectral information from optical and microwave remote sensing pictures and vertical forest structures from GLAS data have recently been integrated (Fararoda *et al.,* 2021; Chen and Sun *et al.,* 2023). In one of the first studies using GLAS-LiDAR data, Boudreau *et al.* (2008) estimated AGB for forests across Quebec, Canada. The study exploits an innovative approach to integrate spaceborne LiDAR, Landsat ETM+ land cover, SRTM digital elevation model (DEM), and ground inventory plots. Using a regression tree-based model and observations from MODIS imagery and GLAS points, Baccini *et al.* (2008) created the very first map of AGB in tropical Africa region, in which approximately 82% of the variance in AGB was explained by the model with an RMSE of 50.5 tons per hectare. They also generated carbon dioxide emissions and density maps for tropical deforestation (Baccini, *et al.,* 2012). Another study by Simard *et al.* (2011) also estimated worldwide forest height at 1km spatial resolution grid. The study used waveform parameters from GLAS data in combination with climate parameters, SRTM-DEM data, and MODIS tree cover. It demonstrated adequate scopes for future research by using a regression tree model and obtained a mean RMSE of 4.4m. Sun *et al.* (2011) used a synergy of radar imagery and waveform LiDAR to estimate forest biomass accurately. Michard *et al.* (2012) studied AGB over densely forested landscapes and showed that the AGB can be precisely and rather accurately estimated using a combination of GLAS LiDAR data, terrain-corrected L-band radar data, and field measurements. Although attempts have been made to collect vertical details of structures from GLAS data to create global forest height and AGB maps, their regional validity is still debatable. In particular, validation sampling across the nation was haphazard, which is very true in the case of Indian forests (Simard *et al.,* 2011; Goparaju *et al.,* 2021). Studies on Indian forests include Reddy & Jha *et al.* (2015), Reddy & Rajashekar *et al.* (2015), Reddy *et al.* (2016), Yadav *et al.* (2021), Pasha *et al.* (2023) and others.

The present study focused on the spatial assessment of AGB by creating a forest height map using a regionalized model, which also combines optical and microwave GLAS data with field observations of climatic and elevation data. The objective was to estimate the spatial biomass above ground utilizing height information derived from GLAS data and a random forest approach based on K-Nearest Neighbours (Tomppo *et al.,* 2004). The main goals were : (1) estimate canopy height using a multi-linear regression model with field-based measurements and GLAS-derived parameters; (2) establish the relationship to derive canopy heights for all GLAS-covered areas; (3) application of RF-based k-NN imputation to derive a spatial estimate of canopy height over Madhya Pradesh, while considering other parameters derived by using remote sensing data (climate variables, optical, and microwave data); and (4) validate the results by using field-derived relationships and finally obtain the spatial height map to above-ground biomass.

MATERIALS AND METHODS

Study area

One of India's states, Madhya Pradesh, is the second-biggest Indian state by area, covering 308245 km² or 9.38% of the country's total land area. It is the sixth most populous state, with more than 72 million people. The state is located between latitudes 21° 17'N and 26° 52'N, and longitudes 74° 08'E and 82° 49'E. The annual rainfall decreases from southeast and east to northwest and west, ranging from 800 mm to around 1800 mm. The average annual temperature ranges between 22.5°C and 25°C. It borders the states of Rajasthan in the northwest, Maharashtra in the south, Gujarat in the west, Uttar Pradesh in the northeast, and Chhattisgarh in the east (Fig. 1). The 17th cycle of forest cover mapping assessed all lands exceeding 1 hectare in the area with a tree canopy density greater than 10 percent. This comprehensive evaluation included traditional forests, trees in orchards, bamboo groves, palm stands, and other similar vegetation types located on recorded forests and other government lands, as well as private, community, or institutional properties. According to the assessment, out of the total forest cover of approximately 77,493 km² (constituting 25.14 percent of the



Study area – red dots represent the ground sample plots and grey circles represents GLAS footprints across Madhya Pradesh in four different vegetation types.

Fig. 1. *Study area – State of Madhya Pradesh*

state's total land area), 6,665 km² was classified as dense forest, 34,209 km² as moderately dense forest, and 36,619 km² as open forest cover (ISFR, 2021).

Further, The methodology used in the present research has also been rigorously validated against new data obtained from ICESAT-2 (October 2018 onwards), GEDI (March 2019 onwards), and other spaceborne LiDAR sensors by numerous researchers globally, including Thumaty *et al.* (2015); Shufan and Chun, (2022); Rodda *et al.,* (2023); Mora *et al.,* (2013); Lu and Jiang, (2024); Huang *et al.,* (2023); Fararoda *et al.,* (2021); Chen and Wang *et al.,* (2024), among others. A precise validation detail is available in the supplementary file (Table ST1 and Fig. SF1). These studies have collectively reinforced the robustness and accuracy of the methodology, establishing its reliability for diverse geospatial and ecological applications. The subsequent section provides a detailed, step-by-step explanation of the working methodology, offering a comprehensive overview of the processes, techniques, and analytical approaches employed. This systematic breakdown ensures clarity and facilitates a deeper understanding of the methodology's structure, enabling its application across various research and practical contexts.

Field sample

The majority of the forests in the research area (51%) were dry and mixed deciduous forests, followed by sal forests (9%), teak (30%), and scrub, thorn, and grassland (9%) (ISFR, 2019). Across the study region, the researcher carried out a comprehensive field inventory of individual trees during 2009–2010, involving detailed measurements taken at the plot level in the field (Fig. 2; Dadhwal *et al.* 2012).To estimate the spatial AGB for the forests in MP, a total of 370 plots in the field were selected. At every 0.1 ha plot, the diameter at breast height (DBH > 5 cm) was measured. Sixty out of 370 plots are selected and positioned beneath the GLAS footprints to estimate the spatial height map in this study. The basal area and tree height were estimated at each plot. Lorey's average height was determined by applying these two factors (Rajab *et al.* 2017). The mean height of comparable trees weighted by their basal area is expressed as Lorey's mean height (Hz) (Lorey,1878).

$$H_z = \frac{\Sigma(g*h)}{\Sigma g} \qquad \text{Eq. 1}$$

Where H is the height of the matching tree in each plot and g is the basal area of each tree in the area of the plot. The largest trees in a stand are given more weight by the basal area weighting of tree heights, which also typically indicates the height of the tallest tree in the stand. After determining Lorey's mean height and ground AGB for 370 plots, spatial estimation of AGB was generated. Using the species-specific allometric equation, the AGB was calculated for each tree.

GLAS data

The data obtained by the Geoscience Laser Altimeter System (GLAS) on board ICESat was used to generate spaceborne LiDAR top canopy height in Madhya Pradesh State. To optimize coverage of ICESat, the GLAS mission began operating with a 91-day repeat orbit (with one daily sub-cycle) at specific periods of the year (Abdalati *et al.,* 2010). Within its ellipsoidal footprints, GLAS waveform data at intervals of about 170 meters provide information on vegetation cover and land elevation. Ballhorn et al., 2011 successfully demonstrated the use of GLAS data in topography and Forest Biomass estimates. There are 15 different level-1 and level-2 deliverables (GLA01 to GLA15) that contain the GLAS products. This study utilized release-33 of GLAS

Fig. 2. *Measurement of tree girth and height (Field survey) – State of Madhya Pradesh*

laser altimetry data obtained from the National Snow and Ice Data Center. It was collected by the satellite from March 2009 to October 2009. A total of 1245 GLAS data (GLA01 and GLA14) footprints have been used to cover the representative forest area of Madhya Pradesh state. This GLAS level-2 land surface altimetry (GLA14) product, which comprises waveform parameters such as the signal starting and echo energy peaks, as well as footprint localization (Zwally *et al.*, 2002), is considered to be best suitable for the present study, while its footprints were assumed a circular diameter of approximately 70 meters. Moreover, GLA14 offers the footprints' geolocation and depicts the surface characteristics, among others. Record number and shot number are used to merge these two datasets, and then normalization, smoothing, and Gaussian fit to the waveforms are applied to extract variables (Harding *et al.,* 2001; Lefsky *et al.,* 2005; Lefsky *et al.,* 2007; Lee *et al.,* 2011). Waveform start, waveform end, waveform extent, number of Gaussian fits, and distance from the ground return peak to the waveform centroid i.e., Height of Median Energy - HOME, or H25, H50, H75, and H100, were the parameters that were extracted from the waveform. The height quartiles at each percentage point where return energy is focused are marked by 25, 50, 75, and 100 numerals. The flowchart of the methodology to extract the above parameters is given (Fig. 3).

Ancillary satellite data analysis

Remote sensing

For the present study, a set of selected satellite instrument data, both active and passive, have been used. Landcover data was obtained from the Indian Remote Sensing (IRS) – P6 Advance Wide Field Sensor (AWiFS) instrument, which can capture image data in 56 m spatial resolution with 4 multispectral bands. From this data, the Normalized Difference Vegetation Index (NDVI) and Normalized Difference Moisture Index (NDMI)) were extracted to understand vegetation cover in the region. Further, Phased Array type L – band Synthetic Aperture Radar data from Advance Land Observing Satellite (ALOS PALSAR) was also utilized with limited capacity to understand the backscatter coefficient, and subsequently vegetation information in the study area. From this dataset, a ratio between HH/HV, and HV/HH polarization was obtained for forest types classification.

Climate data

To understand the climate regime and cause-and-effect relationship in the study region, several climatic parameters were studied. Ancillary gridded data for this purpose was obtained from WorldClim – Global Climate Data with a spatial resolution of about 1 km^2 (Fick and Hijmans, 2017). To understand the biological process and development stage of forest cover, the monthly rainfall and temperature values are used to calculate the bioclimatic variables. Both seasonality and yearly trends were studied from these variables. Out of 19 bioclimatic variables, the following seven variables were utilized as hereunder (Fig. 4):

Variable -1: Annual Mean Temperature (ºC)

Variable -2: Mean Diurnal Range (ºC), i.e. (mean of monthly (max temp - min temp))

Variable -3: Iso-thermality (Variable 2/Variable 7) (*100)

Variable -4: Temperature Seasonality (standard deviation *100)

Variable -7: Temperature Annual Range (ºC)

Variable -12: Annual Precipitation (mm)

Variable -15: Precipitation Seasonality (coefficient of variation) (mm)

In a geographic information system environment, all of the variables mentioned above were resampled to 56 m using the nearest neighbour resampling technique. The forest canopy height, which relates to bioclimatic historical records, spectral bands and vegetation indices, elevation, and slope, was extracted by combining all ancillary variables along with GLAS data.

Downscaling up to the study area

The study uses a two-step approach to produce spatial AGB estimates across the territory of Madhya Pradesh, India. Firstly, it estimates the spatial tree height map, and then, in the second step, it converts the spatial tree height map into AGB by utilizing a statistical relationship with the samples that were collected from field observations.

Spatial mapping of tree heights – Random forest and K-NN algorithms

To generate the spatial height of trees, a relationship was established between GLAS footprints and selected 60 sampled plots. The flow chart of the methodology is given in Fig. 5. With a multiple linear regression model (eq.1), Lorey's height, as observed from field samples, and GLAS-derived parameters were utilized to get the final output of a spatial map for tree height (Alekhya *et al.* 2015).

Lorey's height (m) = 1.214 + (-0.204*HOME) + (0.117*H25) + (1.018*H75) Eq. 2

Using the K-Nearest Neighbour method, Lorey's height at each GLAS footprint as shown in Fig. 1 was then used to train a random forest (RF) model along with all ancillary datasets. The procedure to carry out this modeling was done using the rattle package in the R program (Fig. 6 a & b). Using random subsampling within the provided dataset, the RF data mining technique allows the deployment of several decision trees (Breiman, 2001; Liaw and Wiener, 2002). Additionally, the RF was incorporated into R software (The R Core Team, 2015) using the "yaImpute" and rattle package (Crookston and Finley, 2008). Following variable assignment and processing, the model was then applied in a spatial domain using 'yaImpute,' which uses k-nearest Neighbour imputation. With this approach, the best determination of the correlation coefficient was obtained along with the estimated tree canopy height (m) in each 56 m x 56 m grid in Madhya Pradesh Forest zones (Fig. 7).

K-fold cross-validation

To assess the accuracy of machine learning methods on a small sample of data, resampling is done using a process called cross-validation. Training and test sets of data are separated in a model for machine learning, and the training set uses the test set to evaluate the outcome. An individual data sample's desired number of groups to be divided into is indicated by the procedure's sole parameter, k. Using this method, the set of observations is divided into k folds, or groups, at random that are roughly similar in size. The algorithm fits

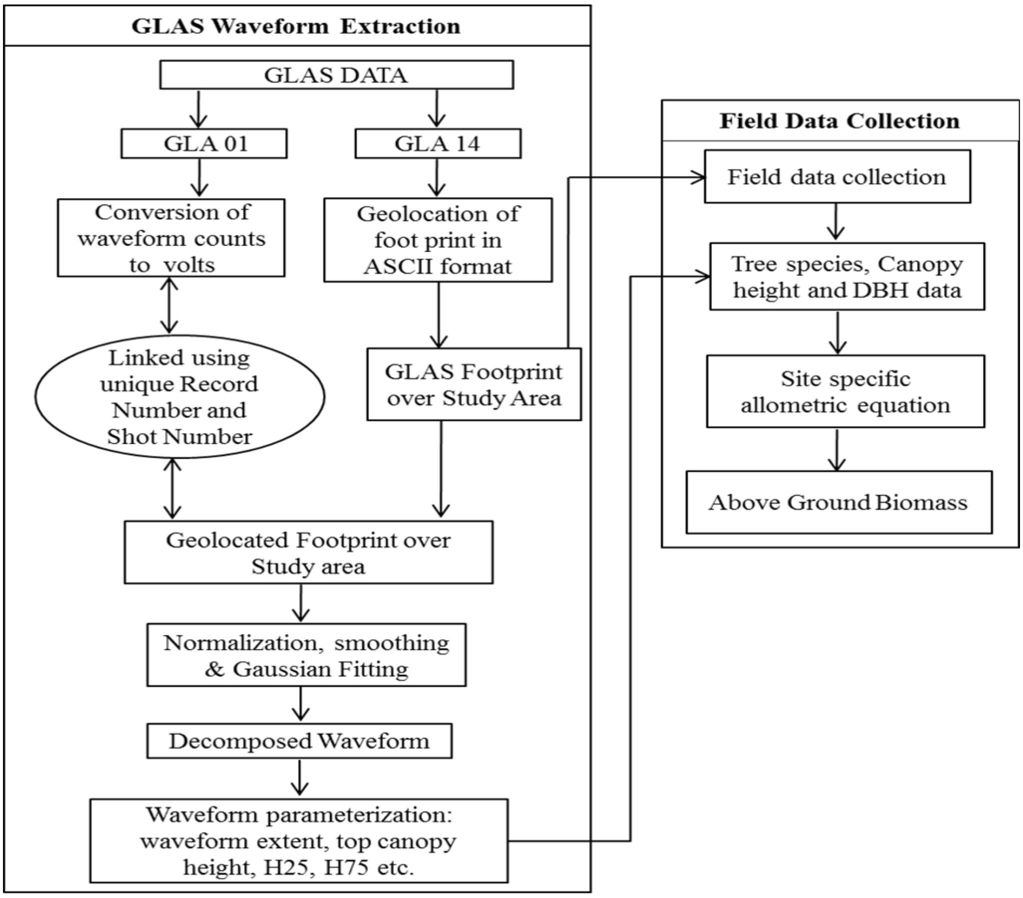


Fig. 3. *Methodology for extracting the GLAS waveform parameters – State of Madhya Pradesh*

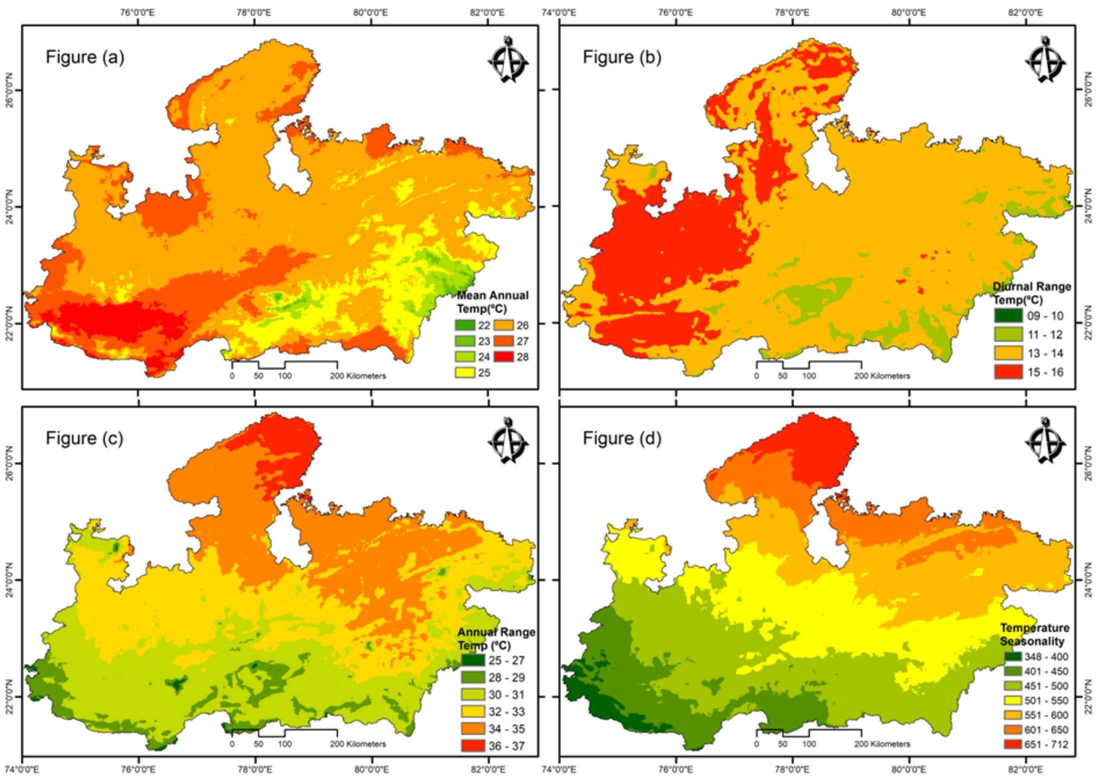


Fig. 4. *Bio-Climatic variables – State of Madhya Pradesh. Source: Prepared by Authors from WorldClim database (Fick and Hijmans, 2017) – Fig. (a) Annual Mean Temperature (ºC); Figure (b) Mean Diurnal Range (ºC); Figure (c) Temperature Annual Range (ºC); and Figure (d) Temperature seasonality*

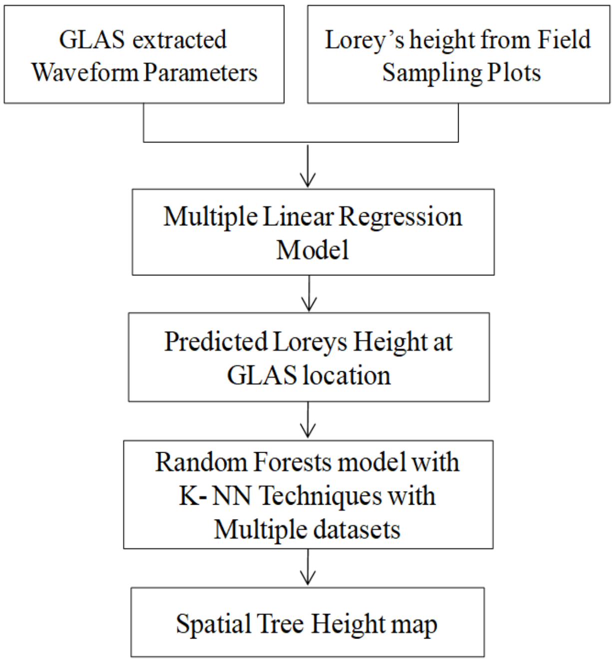


Fig. 5. *Flow chart of spatial tree height map preparation – State of Madhya Pradesh*

the remaining k − 1 folds, using the initial fold as a validation set. The model's reference may correspond to a specific number for k, such as k=10, which becomes a 10-fold cross-validation, instead of k. When the model is used to generate predictions on data that was not used for training, it needs a small sample size to estimate how it should perform generally (Lu *et al.,* 2012). It is widely used because, when compared to other techniques, such as a simple train/test split, it usually produces a less biased or optimistic forecast of model performance and is simple to comprehend. In the present study, the model was validated using the above method, and the model's RMSE was computed for the observed and predicted tree heights in a K-fold cross-validation (k=10). RMSE can be written as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - y_i)^2} \qquad \text{Eq.3}$$

Where yi is the expected Lorey's height at site "i," and xi is the measured Lorey's height at plot "i," and n is the number of plots (i=1, 2..., n). Moreover, Fig. 8 shows that the distribution, scatter plots, and their respective RMSE errors in estimates between observed and GLAS-derived AGB (t ha^{-1}) across four forest types, i.e., dry deciduous (DD), moist deciduous (DM), Sal and Teak in the state of Madhya Pradesh.

Spatial biomass estimation

Using established relationships between Lorey's predicted canopy height and field observed biomass, spatial biomass was calculated from spatial measures of tree height. Lorey's estimated tree height from GLAS and observed AGB were compared using a power regression model over 60 field measurements in 0.1 ha for every forest type (Table 1). The equations were then utilized to calculate the spatial AGB of forests in Madhya Pradesh. Table shows a strong relationship exists among the four classified forest types i.e., dry deciduous (DD), moist deciduous (DM), Sal, and Teak. The coefficient of determination ranges from 0.62 to

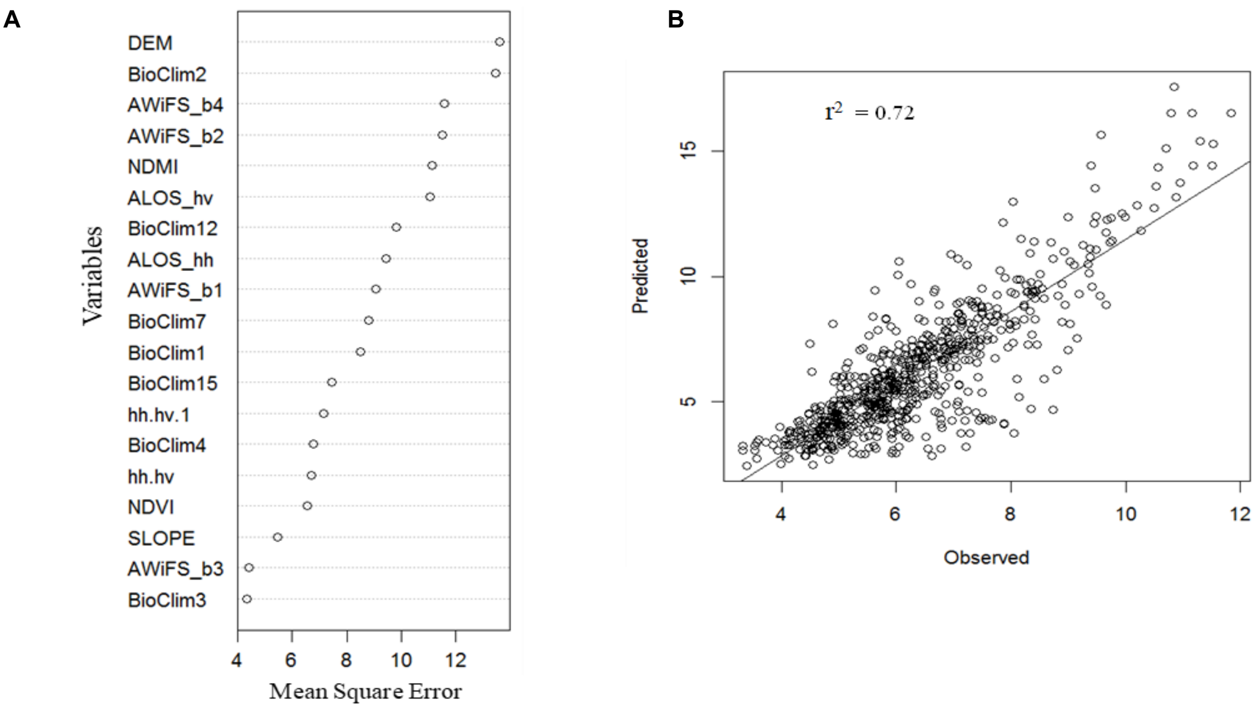


Fig. 6. (a). *Relative importance of predictor variables for tree heights estimation using decision tree approach and RF algorithm; (b) Correlation coefficient between observed versus predicted tree height using RF algorithm*

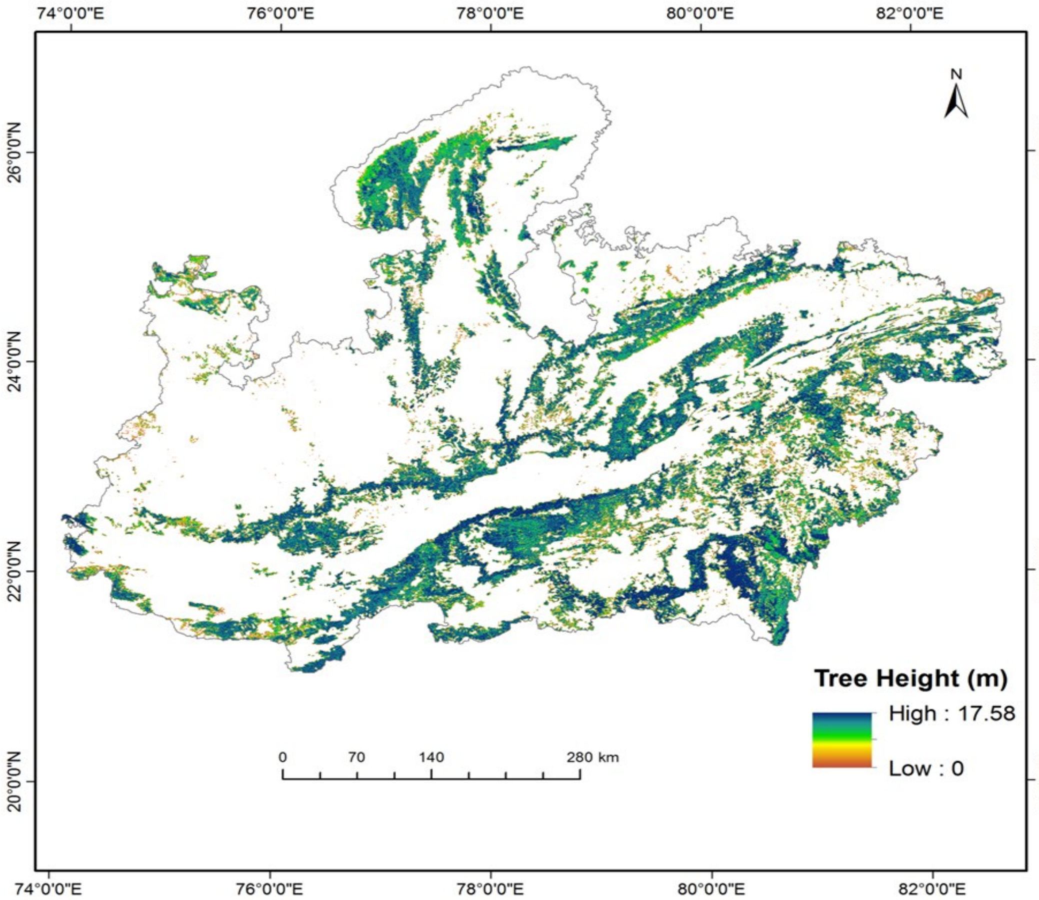


Fig. 7. *Spatial map of tree height (estimated) derived from GLAS data using RF algorithm – State of Madhya Pradesh*

0.71 with a 95% confidence interval. Moreover, the final estimates of AGB derived from GLAS data were presented in Fig. 9, along with the normalized difference vegetation index (NDVI) generated from AWiFS data.

RESULTS AND DISCUSSION

The spatial height of trees was measured using two alternative models, i.e., the multiple linear regression model and a random forest using k-NN imputation (section 2.5.1). Furthermore, the spatial estimation of forest cover biomass was estimated based on the correlations between Lorey's predicted height of the canopy and field observed biomass. It shows that the R^2 between the measured (Lorey's height) and observed (GLAS derived) height of forest cover was 0.74 with a 95% confidence interval (Alekhya *et al.* 2015). Thumaty *et al.* (2015) by using ALOS data showed that the average AGB for Madhya Pradesh has to be 58 t/ha, while the model demonstrated that the observed versus predicted AGB had a strong correlation ($R^2 = 0.7$), indicating the reliability of the spatial analysis. Further, supplementary datasets and GLAS footprints are trained in a random forest model for spatial assessment of tree height. It demonstrates that the BIO2 climatic parameter, spectral bands, indices of vegetation, height factors, and ALOS PALSAR L-band backscatter coefficient (HH, HV) have a significant relationship and are accountable for the predictor value (tree height). On the other hand, variables also contributed to the predictor value in the RF model, as is evident in Fig. 6 (a&b). Using these variables, the RF model developed a spatial tree height map with a height ranging from 2.16 m to 17.63 m while the RMSE of ± 2.57 m was found in the analysis (Fig. 7). The correlation study was performed between observed Lorey's heights from field data and anticipated Lorey's heights from GLAS data, with R^2=0.7 (Fig. 8). Finally, the predicted tree height map at 56 m resolution was compared to the available worldwide forest height map at 1 km resolution (Simard *et al.,* 2011). The GLAS dataset classifies tree heights into three categories: <5m, 5-10m, and >10m.

Fig. 9 depicts the spatial distribution of AGB including 60 field survey plots, which are utilized to estimate spatial biomass map alongside an estimated tree height map based on GLAS footprints (see section 2.6). Most of the AGB plots were observed in dry deciduous forest regions followed by teak, moist deciduous, and Sal Forest areas (Table 1). The plot-level mean AGB of teak, dry deciduous, Sal, and moist deciduous forests was estimated as 58.52, 43.70, 43.67, and 40.11 ton ha^{-1} respectively. Further, the regression model was estimated between Lorey's height (observed) and AGB (estimated) on various vegetation types for all 60 plots (Table 1). The analysis shows that Sal woods had the highest association between GLAS-derived Lorey's tree height and AGB with R^2 = 0.71 and p<0.01, which is followed by moist deciduous forests with R^2 = 0.68 and p<0.01, Teak Forest with R^2 = 0.65 and p<0.01, and dry deciduous forests with R^2 = 0.62 and p<0.01, respectively. The geographical estimates of AGB from the GLAS datasets using the equations (Table 1) ranged from 8.56 to 252.04 tons per hectare across different vegetation types (Fig. 9). The plot level AGB estimated for the research area's forest types ranged from 3.17 to 198.41 tons per hectare, while GLAS data estimated AGB ranged from 8.56 to 252.04 tons per hectare. The average estimated AGB from GLAS for four vegetation types has to be 57.10 tons per hectare with RMSE ±14.65 tons per hectare for Sal Forest, 42.73 tons per hectare with RMSE ±24.46 tons per hectare for Teak, 41.56 tons per hectare with RMSE ±18.06 tons per hectare for moist deciduous, and 35.17 tons per hectare with RMSE ±19.73 tons per hectare for dry deciduous, respectively. Additionally, the estimated AGB based on GLAS data together with ancillary variables and the estimated field biomass for different forest types across the study region are fairly similar to each other and based on the F test (F=1.42 and p -value = 0.68), were not statistically different. Using a decision-tree-based technique and an RF algorithm, the study region's total AGB stock was assessed to be 315.77 metric tons, with a RMSE of ± 19.22 tons per hectare, while it ranged from 238.88 to 392.66 metric tons. One notable study on the region found 367.4 metric tons of AGB for Madhya Pradesh's deciduous woods using ALOS-PALSAR L-band data from 2010 (Thumaty *et al.,* 2015). Chaturvedi *et al.* (2011) estimated that the carbon stock in India's tropical dry forest ranged between 15.6 and 151 tons per hectare. Salunkhe *et al.* (2014) investigated above-ground biomass estimation exclusively in selected plots or districts in Madhya Pradesh. The average above-ground biomass of dry deciduous and mixed deciduous forests across all locations was 31.8 and 20.7 tons per hectare, respectively. Huang *et al.* (2023) compared six models using machine learning technique in Yunnan–Guizhou Plateau in southwest China. The result of this study shows that the estimate for coniferous was better than that of mixed forest. For coniferous forests, the R^2 value was 0.63 with an RMSE of 43.23 Mg/ha, while for mixed forests, the R^2 was 0.56 with an RMSE of 47.79 Mg/ha, reflecting moderate predictive accuracy for both forest types. In a recent study, Nandy *et al.* (2021) adopted a similar approach to estimate AGB in Doon valley, Uttarakhand by using ICESat -2 data. The study revealed a strong correlation between field-measured canopy height and ICESat-2 data (R^2 = 0.89, RMSE = 1.11 m). Predicted canopy heights ranged from 15.32 to 31.02 m. The Random Forest model demonstrated

Table 1. Regression model estimates between Lorey's height (observed) and AGB (estimated) on various vegetation types

Vegetation type	No. of plots	Equation	R^2	RMSE	p-value
Dry Deciduous	27	$Y = 3.6192x^{1.231}$	0.62	19.73	<0.01
Moist Deciduous	13	$Y = 3.6018x^{1.2791}$	0.68	18.06	<0.01
Sal	3	$Y = 2.1135x^{1.6677}$	0.71	14.65	<0.01
Teak	17	$Y = 2.0786x^{1.5783}$	0.65	24.46	<0.01

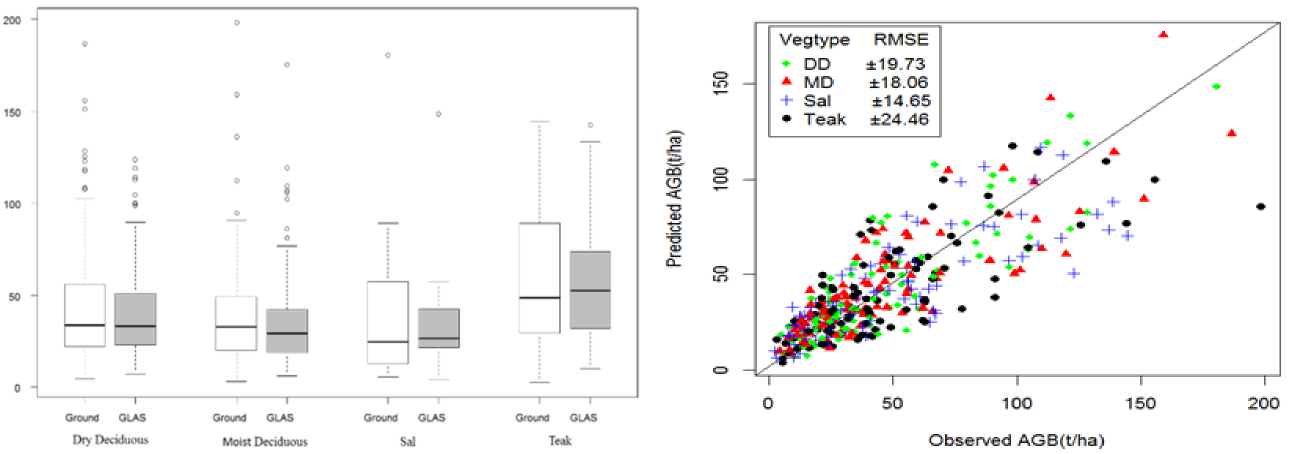


Fig. 8. *Distribution and Scatter plots between observed and GLAS derived AGB (t ha^{-1}) across four vegetation types (Dry Deciduous (DD), Moist Deciduous (MD), Sal, Teak and their respective RMSE – State of Madhya Pradesh*

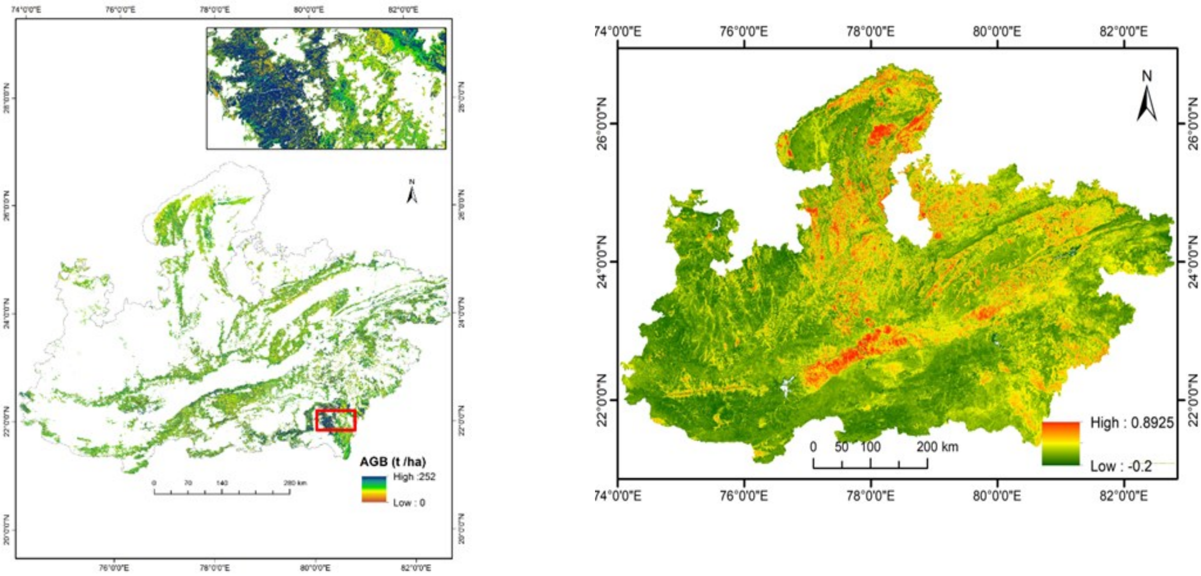


Fig. 9. *Estimated spatial map of above ground biomass (AGB) and NDVI derived from GLAS and AWIFS data respective – State of Madhya Pradesh*

high accuracy (R^2 = 0.84, RMSE = 1.15 m). The forest AGB model, integrating canopy height and spectral data, achieved an R^2 of 0.83 and an RMSE of 19.98 Mg/ha, with AGB predictions ranging from 234.91 to 547.56 Mg/ha and a mean of 426.41 Mg/ha. Nonetheless, the estimated above-ground biomass from GLAS data in the present study was 57.10 tons per hectare (R^2 = 0.62, RMSE = ±19.73) for the dry deciduous forest of MP. This study integrates ICESat-1-derived forest canopy height with spectral variables to enhance AGB predictions, aligning with the primary science objective of the ICESat mission to estimate vegetation canopy height for biomass assessment. Previous and current research, such as Rodda *et al.* (2023), Mora *et al.* (2013), Huang *et al.* (2023), Fararoda *et al.* (2021), Li *et al.* (2021), Urbazaev *et al.* (2018), Nandy *et al.,* (2021), and among others, supports the effectiveness of multi-sensor integration with machine learning algorithms in improving forest biomass estimates, particularly forest structure, with higher accuracy than traditional field inventory approaches.

Conclusion

In this present study, the effective measurement of forest canopy height and the estimation of above-ground biomass (AGB) were achieved by utilising spaceborne LiDAR, specifically the Geoscience Laser Altimeter System (GLAS). The study focused on the vast expanse of the State of Madhya Pradesh, situated in the Central Indian Forests region. The integration of data from various sources, including LIDAR (ICESat/GLAS), Radar (ALOS-PALSAR), Optical (IRS-P6 AWIFS), and digital elevation model (SRTM), along with ancillary climate variables and field-based forest inventory, facilitated a comprehensive analysis. The Lorey's height method, employed in this research, established a relationship with GLAS-derived AGB, further combined with climatic factors, elevations, and land cover to map the spatial extent of forest areas in the studied region. The development of a spatial height map through a K-Nearest Neighbour-based random forest approach demonstrated variability in estimated forest canopy height, ranging from 2.16 m to 17.63 m, with a root mean square error (RMSE) of ± 2.57 m. Additionally, spatial AGB was estimated using height-biomass models for different forest types. The findings revealed that the total AGB over forests in Madhya Pradesh was estimated to be 315.77 Mt with an RMSE of ±19.22 t ha-1. The relative error across different forest types ranged from 33% to 45%. The study highlighted the significance of accurate AGB assessment, especially in the context of global climate change concerns. Forest biomass, a key factor in carbon storage, is crucial in mitigating potential carbon emissions from deforestation or land-use changes. The methodology employed in this research, integrating spaceborne LiDAR with ancillary data from various sensors, proved effective in estimating forest canopy height and AGB over a large and diverse landscape. The accuracy of the estimates was further validated through the establishment of height-biomass models and cross-validation techniques. The research contributes to the broader understanding of the applicability of spaceborne LiDAR in assessing forest structure and biomass, with implications for global efforts in carbon monitoring and sustainable forest management. The spatially explicit maps generated through this study provide valuable information for policymakers, conservationists, and researchers involved in forestry and environmental management. Therefore, future research should include a comparative analysis using forest biomass and canopy cover estimates.

Conflict of interest

The authors declare that they have no conflict of interest.

REFERENCES

1. Abdalati, W., Zwally, H. J., Bindschadler, R., Csatho, B., Farrell, S. L., Fricker, H. A., &. . . Webb, C. (2010). The ICESat-2 Laser Altimetry Mission. *Proceedings of the IEEE, 98*(5), 735–751. https://doi.org/10.1109/jproc.2009.2034765
2. Ahern, F., Janetos, A. C., & Langham, E. (1998). *Global Observation of Forest Cover: One Component of CEOS' Integrated Global Observing Strategy*. https://doi.org/10.4095/219613
3. Alekhya, V. V. L. P., Reddy. R. S., Jha, C. S., Rajashekar, G. & Prasada Rao, P. V. V. (2015). Estimation of vertical forest structure and above-ground biomass using ICESat/GLAS data. *International Journal of Geomatics and Geosciences,* Vol. 6(2), 1530-1538.
4. Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., . . . Houghton, R. A. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change*, 2(3), 182–185. https://doi.org/10.1038/nclimate1354
5. 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. *Environmental Research Letters, 3*(4), 045011. https://doi.org/10.1088/1748-9326/3/4/045011
6. Ballhorn, U., Jubanski, J., & Siegert, F. (2011). ICESat/GLAS Data as a Measurement tool for peatland topography and peat swamp forest biomass in Kalimantan, Indonesia. *Remote Sensing*, 3(9), 1957–1982. https://doi.org/10.3390/rs3091957
7. Boudreau, J., Nelson, R., Margolis, H., Beaudoin, A., Guindon, L., & Kimes, D. (2008). Regional aboveground forest biomass using airborne and spaceborne LiDAR in Québec. *Remote Sensing of Environment*, 112(10), 3876–3890. https://doi.org/10.1016/j.rse.2008.06.003
8. Breiman, L. (2001). Random Forests. *Machine Learning,* Vol. 45, 5–32. https://doi.org/10.1023/A:1010933404324
9. Brown, S. (1997). Estimating biomass and biomass change of Tropical forests – A primer. Forestry Paper – 134, FAO, Rome.
10. Chaturvedi, R., Raghubanshi, A., & Singh, J. (2011). Carbon density and accumulation in woody species of tropical dry forest in India. *Forest Ecology and Management*, 262(8), 1576–1588. https://doi.org/10.1016/j.foreco.2011.07.006
11. Chen, A., Wang, X., Zhang, M., Guo, J., Xing, X., Yang, D., . . . Yang, X. (2023). Fusion of LiDAR and Multispectral Data for Aboveground Biomass Estimation in Mountain Grassland. *Remote Sensing, 15*(2), 405. https://doi.org/10.3390/rs15020405
12. Chen, Z., Sun, Z., Zhang, H., Zhang, H., & Qiu, H. (2023). Aboveground Forest Biomass Estimation Using Tent Mapping Atom Search Optimized Backpropagation Neural Network with Landsat 8 and Sentinel-1A Data. *Remote Sensing, 15*(24), 5653. https://doi.org/10.3390/rs15245653
13. Crookston, N. L., & Finley, A. O. (2008). yaImpute: An R Package for kNN Imputation. *Journal of Statistical Software*, 23(10). https://doi.org/10.18637/jss.v023.i10
14. Dadhwal, V. K., Kushwaha, S. P. S., Singh, S., Patel, N. R., Nayak, R. K., Patil, P., . . . Ali, M. M. (2012). Recent results from EO studies on Indian carbon cycle assessment. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVIII-8/W20, 3–9. https://doi.org/10.5194/isprsarchives-xxxviii-8-w20-3-2011
15. Fararoda, R., Reddy, R. S., Rajashekar, G., Chand, T. K., Jha, C. & Dadhwal, V. (2021). Improving forest above-ground biomass estimates over Indian forests using multi-source data sets with machine learning algorithm. *Ecological Informatics, 65*, 101392. https://doi.org/10.1016/j.ecoinf.2021.101392
16. Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1□km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302–4315. https://doi.org/10.1002/joc.5086
17. ISFR. (2021). India State Forest Report – 2021. Forest Survey of India, Ministry of Environment and Forests, Dehradun, India.
18. Goparaju, L., Madugundu, R., Ahmad, F., Sinha, D. & Singh, J. S. (2021). Temporal dynamics of above-ground biomass of Kaimoor Wildlife Sanctuary, Uttar Pradesh, India: conjunctive use of field and Landsat data. *Proceedings of the Indian National Science Academy, 87*(3), 499–513. https://doi.org/10.1007/s43538-021-00046-1
19. Harding, D., Lefsky, M., Parker, G., & Blair, J. (2001). Laser altimeter canopy height profiles: methods and validation for closed-canopy, broadleaf forests. *Remote Sensing of Environment*, 76(3), 283–297. https://doi.org/10.1016/s0034-4257(00)00210-8
20. Huang, X., Cheng, F., Wang, J., Duan, P., & Wang, J. (2023). Forest Canopy Height Extraction Method Based on ICESat-2/ATLAS Data. *IEEE Transactions on Geoscience and Remote Sensing, 61*, 1–14. https://doi.org/10.1109/tgrs.2023.3240848
21. Lee, S., Ni-Meister, W., Yang, W., & Chen, Q. (2011). Physically based vertical vegetation structure retrieval from ICESat data: Validation using LVIS in White Mountain National Forest, New Hampshire, USA. *Remote Sensing of Environment, 115*(11), 2776–2785. https://doi.org/10.1016/j.rse.2010.08.026
22. Lefsky, M. A. (2009). Biomass accumulation rates of Amazonian secondary forest and biomass of old-growth forests from Landsat time series and the Geoscience Laser Altimeter System. *Journal of Applied Remote Sensing*, 3(1), 033505. https://doi.org/10.1117/1.3082116
23. Lefsky, M. A. (2010). A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. *Geophysical Research Letters, 37*(15). https://doi.org/10.1029/2010gl043622
24. Lefsky, M. A., Harding, D. J., Keller, M., Cohen, W. B., Carabajal, C. C., Del Bom Espirito□Santo, F., . . . de Oliveira, R. (2005). Estimates of forest canopy height and aboveground biomass using ICESat. *Geophysical Research Letters*, 32(22). https://doi.org/10.1029/2005gl023971

25. Lefsky, M. A., Harding, D., Cohen, W., Parker, G., & Shugart, H. (1999). Surface Lidar Remote Sensing of Basal Area and Biomass in Deciduous Forests of Eastern Maryland, USA. *Remote Sensing of Environment*, 67(1), 83–98. https://doi.org/10.1016/s0034-4257(98)00071-6
26. Lefsky, M. A., Hudak, A. T., Cohen, W. B., & Acker, S. (2005). Geographic variability in lidar predictions of forest stand structure in the Pacific Northwest. *Remote Sensing of Environment*, 95(4), 532–548. https://doi.org/10.1016/j.rse.2005.01.010
27. Lefsky, M. A., Keller, M., Pang, deCamargo, & Hunter. (2007). Revised method for forest canopy height estimation from Geoscience Laser Altimeter System waveforms. Journal of Applied Remote Sensing, 1(1), 013537. https://doi.org/10.1117/1.2795724
28. Liaw, A., Wiener, M. (2002). Classification and regression by random forest. The R Journal, R News 2, 18-22. Retrieved from https://journal.r-project.org/articles/RN-2002-022/
29. Li, Z. Bi, S. Hao, S. Cui, Y. (2021). Aboveground biomass estimation in forests with random forest and Monte Carlo-based uncertainty analysis. *Ecological Indicators,* 142 (2022) 109246. https://doi.org/10.1016/j.ecolind.2022.109246
30. Lorey T. (1878). Die mittlere Bestandeshöhe. *Allgemeine Forst- und Jagdzeitung, 54*: 149-155. (in German).
31. Lu, D., & Jiang, X. (2024). A brief overview and perspective of using airborne Lidar data for forest biomass estimation. *International Journal of Image and Data Fusion,* 1–24. https://doi.org/10.1080/19479832.2024.2309615
32. Lu, D., Chen, Q., Wang, G., Moran, E., Batistella, M., Zhang, M., . . . Saah, D. (2012). Aboveground Forest Biomass Estimation with Landsat and LiDAR Data and Uncertainty Analysis of the Estimates. *International Journal of Forestry Research, 2012,* 1–16. https://doi.org/10.1155/2012/436537
33. Madugundu, R., Nizalapur, V., & Jha, C. S. (2008). Estimation of LAI and above-ground biomass in deciduous forests: Western Ghats of Karnataka, India. *International Journal of Applied Earth Observation and Geoinformation, 10*(2), 211–219. https://doi.org/10.1016/j.jag.2007.11.004
34. Mora, B., Wulder, M., White, J., & Hobart, G. (2013). Modeling Stand Height, Volume, and Biomass from Very High Spatial Resolution Satellite Imagery and Samples of Airborne LiDAR. *Remote Sensing, 5*(5), 2308–2326. https://doi.org/10.3390/rs5052308
35. Nandy, S. Srinet, R. Padalia, H. (2021). Mapping forest height and aboveground biomass by integrating ICESat-2, Sentinel-1 and Sentinel-2 data using Random forest algorithm in northwest Himalayan foothills of India. *Geophysical Research Letters*, *48*, e2021GL093799. https://doi.org/10.1029/2021GL093799
36. Nelson, R., Ranson, K., Sun, G., Kimes, D., Kharuk, V., & Montesano, P. (2009). Estimating Siberian timber volume using MODIS and ICESat/GLAS. *Remote Sensing of Environment, 113*(3), 691–701. https://doi.org/10.1016/j.rse.2008.11.010
37. Pasha, S. V., Dadhwal, V. K., & Reddy, C. S. (2023). Rubber expansion and age-class mapping in the state of Tripura (India) 1990–2021 using multi-year and multi-sensor data. *Environmental Monitoring and Assessment, 195*(2). https://doi.org/10.1007/s10661-023-10942-2
38. Rajab Pourrahmati, M., Baghdadi, N., Darvishsefat, A. A., Namiranian, M., Gond, V., Bailly, J. S., & Zargham, N. (2017). Mapping Lorey's height over Hyrcanian forests of Iran using synergy of ICESat/GLAS and optical images. *European Journal of Remote Sensing, 51*(1), 100–115. https://doi.org/10.1080/22797254.2017.1405717
39. Ramachandran, N., Saatchi, S., Tebaldini, S., d'Alessandro, M. M., & Dikshit, O. (2023). Mapping tropical forest aboveground biomass using airborne SAR tomography. *Scientific Reports, 13*(1). https://doi.org/10.1038/s41598-023-33311-y
40. Reddy, C. S., Jha, C. S., Dadhwal, V. K., Hari Krishna, P., Vazeed Pasha, S., Satish, K. V., . . . Diwakar, P. G. (2015). Quantification and monitoring of deforestation in India over eight decades (1930–2013). *Biodiversity and Conservation, 25*(1), 93–116. https://doi.org/10.1007/s10531-015-1033-2
41. Reddy, C. S., Rajashekar, G., Hari Krishna, P., Jha, C., & Dadhwal, V. (2015). Multi-source and multi-date mapping of deforestation in Central India (1935-2010) and its implication on standing phytomass carbon pool. *Ecological Indicators, 57*, 219–227. https://doi.org/10.1016/j.ecolind.2015.05.003
42. Reddy, C. S., Rakesh, F., Jha, C., Athira, K., Singh, S., Alekhya, V. P., . . . Dadhwal, V. (2016). Geospatial assessment of long-term changes in carbon stocks and fluxes in forests of India (1930–2013). *Global and Planetary Change, 143*, 50–65. https://doi.org/10.1016/j.gloplacha.2016.05.011
43. Rodda, S. R., Nidamanuri, R. R., Fararoda, R., Mayamanikandan, T., & Rajashekar, G. (2023). Evaluation of Height Metrics and Above-Ground Biomass Density from GEDI and ICESat-2 Over Indian Tropical Dry Forests using Airborne LiDAR Data. *Journal of the Indian Society of Remote Sensing.* https://doi.org/10.1007/s12524-023-01693-1
44. Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., . . . Morel, A. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences, 108*(24), 9899–9904. https://doi.org/10.1073/pnas.1019576108
45. Salunkhe, O., Khare, P. K., Sahu, T. R., Singh, S. (2014). Above Ground Biomass and Carbon Stocking in Tropical Deciduous Forests of State of Madhya Pradesh, India. *Taiwania 59*(4): 353-359. https://taiwania.ntu.edu.tw/pdf/tai.2014.59.353.pdf
46. Schutz, B. E., Zwally, H. J., Shuman, C. A., Hancock, D., & DiMarzio, J. P. (2005). Overview of the ICESat Mission. *Geophysical Research Letters, 32*(21). https://doi.org/10.1029/2005gl024009
47. Shufan, W., & Chun, L. (2022). Urban forest canopy height modeling using spaceborne laser ICESAT-2 LIDAR. IOP Conference Series: *Earth and Environmental Science, 1004*(1), 012023. https://doi.org/10.1088/1755-1315/1004/1/012023
48. Simard, M., Pinto, N., Fisher, J. B., & Baccini, A. (2011). Mapping forest canopy height globally with spaceborne lidar. *Journal of Geophysical Research, 116*(G4). https://doi.org/10.1029/2011jg001708
49. Sun, G., Ranson, K. J., Guo, Z., Zhang, Z., Montesano,

P., & Kimes, D. (2011). Forest biomass mapping from lidar and radar synergies. *Remote Sensing of Environment, 115*(11), 2906–2916. https://doi.org/10.1016/j.rse.2011.03.021

50. The R Core Team. (2015). R: A language and environment for statistical computing (Reference Index). R Foundation for Statistical Computing, Vienna, Austria. https://cran.r-hub.io/doc/manuals/fullrefman.pdf
51. Thumaty, K. C., Fararoda, R., Middinti, S., Gopalakrishnan, R., Jha, C. S., & Dadhwal, V. K. (2015). Estimation of Above Ground Biomass for Central Indian Deciduous Forests Using ALOS PALSAR L-Band Data. *Journal of the Indian Society of Remote Sensing, 44*(1), 31–39. https://doi.org/10.1007/s12524-015-0462-4
52. Tomppo, E., Halme, M. (2004). Using coarse scale forest variables as ancillary information and weighting of variables in k-NN estimation: a genetic algorithm approach. *Remote Sensing of Environment, 92*(1), 1–20. https://doi.org/10.1016/j.rse.2004.04.003
53. Urbazaev, M. Thiel, C. Cremer, F. Dubayah, R. Migliavacca, M. Reichstein, M. Schmullius, C. (2018). Estimation of forest aboveground biomass and uncertainties by integration of field measurements, airborne LiDAR, and SAR and optical satellite data in Mexico. *Carbon Balance Manage (2018) 13:5*. https://doi.org/10.1186/s13021-018-0093-5
54. Wang, Y., Peng, Y., Hu, X., & Zhang, P. (2023). Fine-Resolution Forest Height Estimation by Integrating ICESat-2 and Landsat 8 OLI Data with a Spatial Downscaling Method for Aboveground Biomass Quantification. *Forests, 14*(7), 1414. https://doi.org/10.3390/f14071414
55. Yadav, S., Padalia, H., Sinha, S. K., Srinet, R., & Chauhan, P. (2021). Above-ground biomass estimation of Indian tropical forests using X band Pol-InSAR and Random Forest. *Remote Sensing Applications: Society and Environment, 21*, 100462. https://doi.org/10.1016/j.rsase.2020.100462
56. Zwally, H., Schutz, B., Abdalati, W., Abshire, J., Bentley, C., Brenner, A., . . . Thomas, R. (2002). ICESat's laser measurements of polar ice, atmosphere, ocean, and land. *Journal of Geodynamics, 34*(3–4), 405–445. https://doi.org/10.1016/s0264-3707(02)00042-x