 doi:10.20944/preprints202210.0190.v1

Article

Estimation of Above Ground Biomass of Mangrove Forests from Terrestrial LiDAR Data using Supervised Machine Learning Algorithms

Yeshwanth Kumar Adimoolam[1], Nithin D. Pillai[2], Gnanappazham Lakshmanan[3]* Deepak Mishra[4] and Vinay Kumar Dadhwal[5]

[1] CYENS Center of Excellence, Cyprus,[2]GFZ Helmholtz Centre, Potsdam, German, [3] Indian Institute of Space Science and Technology, Valiamala, Trivandrum, India, [4] National Institute of Advanced Studies, Bangaluru, India

* Correspondence: gnanam@iist.ac.in

Abstract: Accurately quantifying the above-ground volume (AGV) and thus above-ground biomass (AGB) of forest stands is an important aspect in the conservation of mangrove ecosystem owing to their ecological and economic benefits. However, the number of studies focusing on quantifying mangrove forests' biomass has been relatively low due to their marshy terrain, making exploratory studies challenging. In recent times, the use of LiDAR technologies in forest inventory studies has become increasingly popular, due to the reliability of LiDAR as a highly accurate means of 3D spatial data acquisition. In this study, we propose an end-to-end methodology for estimating AGV of mangrove forest stands from terrestrial LiDAR data. Many of the recent studies on this topic effectively employ machine learning algorithms such as multi layer perceptron, random forests, etc. for filtering foliage in the point cloud data of single trees. This study further extends that approach by incorporating the impact of class imbalance of forest point cloud data in a weighted random forest classifier. For the task of segmentation of wood/foliage points in a single tree point cloud, this approach yielded an average increase of 15.27% in the balanced accuracy score, 0.20 in Cohen's Kappa score, 15.27% in the ROC AUC score and 5.45% in the F1 score. For the task of AGV estimation of a single tree, this approach resulted in an average coefficient of determination of 0.93 with respect to the ground truth volumes. Also, the machine learning classifier and geometric features used in this study were invariant to tree species and hence could be generalised for the classification of point clouds of other tree species as well. This study aims to incorporate the detection and AGB estimation of pneumatophores (inverted root like structures above the ground or aerial roots produced by some of the mangrove species that are special adaptation for gaseous exchange in marshy environment) in the inventory of mangrove forest stands which is first of its kind using LiDAR point cloud data. Above Ground Biomass estimation of mangroves using our approach based on TLS data has the mean bias of 4.4kg and RMS variation of 25.85kg when compared with the conventional allometric methods of AGB estimation. For the task of counting pneumatophores in a plot-level point cloud, a breadth-first graph-search segmentation based approach is also proposed as part of the pipeline to estimate the contribution of pneumatophores to the overall AGB of mangrove forest stands. The proposed breadth-first searching method yielded an average coefficient of determination of 0.94. The contribution of pneumatophores to the AGB of mangrove forest plots was estimated to increase the AGB density by 0.196 kg/m^2 that could also aid future mangrove forest inventory studies in modeling the underlying root network and estimating the below-ground biomass of mangrove trees.

Keywords: Above-ground biomass; mangroves; pneumatophores; terrestrial LiDAR; machine learning; random forest

 doi:10.20944/preprints202210.0190.v1

1 Introduction

Forest's biomass is the principal indicator of the carbon stock and potential of carbon sequestration of a terrestrial ecosystem [1]. Owing to the significance of biomass in understanding the patterns of global climate change, the United Nations Framework Convention on Climate Change (UNFCCC) proposed biomass as an Essential Climate Variable (ECV) in order to bridge the information gaps in the knowledge base on the global climate system [2]. Forest's biomass can be broadly classified as above-ground and below-ground biomass. Since the study of below-ground biomass often requires destructive uprooting of trees, most studies focus on the quantification of above-ground biomass (AGB) of a forest. Conventional remote sensing methods often use either high resolution optical sensors or SAR sensors to study the AGB of a forest [3].

In a forest, the collective above-ground wood volume and merchantable stem volume of trees are the main attributes of forest inventory that are of interest to most forest managers [4]. Tree heights are also measured, but typically only for a subset of trees in the plot, since height measurements are more tedious and costly to perform than measuring diameters [5,6]. However, carrying out such field measurements can prove to be greatly time consuming and infeasible. Furthermore, field measurements can only be made in forest plots that can be physically accessed by field personnel. In the last 4 to 5 decades such difficulties are overcome with the help of remote sensing data that significantly improves the efficiency and spatial coverage of forest inventory. Optical and SAR remote sensing data have been employed extensively to study the structural and biophysical parameters of forests using various regression methods [3,7,8]. However, most recently published studies on estimation of AGB from remotely sensed data exhibit noticeable variability in the accuracy of the biomass estimate depending on the forest environment and type of remote sensing data used [9]. Adding to this, optical remote sensing data can only cover the canopy of forests and SAR can penetrate only crown height whereas the sub-canopy vegetation contributing greatly to the total AGB of the forest is rarely sensed by them.

In the past years, LiDAR technology has showed great potential for forest inventory applications and forest attribute studies at very high resolutions [10–12]. Also, among the presently available LiDAR data acquisition techniques, Terrestrial Laser Scanning (TLS) appears to deliver the most detailed and precise characterisation of forest structures. This can be especially true in complex forest stands such as that of mangroves, where the sub-canopy vegetation is characterised by a high degree of branching, varying diameters, overlapping crowns. Unlike, other terrestrial forests, mangroves have special adaptations like, pneumatophores and stilt roots (inverted root like structures above the ground or aerial roots produced by some of the mangrove species that are special adaptation for gaseous exchange in marshy environment) that also have significant contribution in the AGB estimation of mangroves [13] and LiDAR techniques will be able to meet this requirement very efficiently with more accuracy. Upon considering these aspects, it can be reasonably inferred that using a combination of terrestrial LiDAR, remote sensing imagery and field measurements will certainly yield better results in quantification of above-ground biomass as opposed to conventional methods employed in forest inventory studies.

In order to make meaningful decisions for the approaches and techniques used in this study, the prevailing literature relevant to this study were studied extensively and several key observations were noted. Point cloud classification algorithms can be broadly grouped into two categories: (a) direct methods that operate on the point clouds in their raw form and (b) indirect methods which involve transforming the input point cloud into other rasterized representations such as voxels or projected images. Among the numerous methods reviewed for point cloud classification, it was noted that the methods which processed point clouds directly performed better both in terms of space and time complexity [14–19]. The view-based [20,21] and volumetric methods [22–24] both involve abstraction of the raw point cloud data into rasterized forms such as multi-

 doi:10.20944/preprints202210.0190.v1

view stereo rasters or voxel grids. Although both these approaches offer the advantage of easy integration of point cloud data into existing 2D computer vision algorithms, the main disadvantage of these methods is that the abstraction of point cloud data into intermediate rasterized forms leads to loss of the rich intrinsic geometry of point clouds. This could potentially introduce errors in the classification results. Also, these approaches were found to be less efficient in both space and time complexity compared to directly extracting features from point clouds in their raw form. Owing to these factors, it was inferred that a point-wise feature extraction and geometric approaches were found to yield the best results for classification of point clouds and in the present study we apply this approach to classify point clouds of foliage and wood.

When estimating the above-ground biomass of forests from terrestrial laser scans, the foliage points in the forest stand point cloud tend to cause gross overestimation of the estimated AGB value. Therefore, it is often necessary to filter out the foliage points from the wood points of the trees and estimate AGB only using the wood points. The presently available foliage filtering methods could be grouped into either those using radiometric features for classification [25] or those using geometric features [26]. Methods relying on radiometric features of points were found to be largely dependent on sensor-specific characteristics, which rendered such methods unable to generalise well over data from different sensors. Therefore, methods that used purely geometric features were found to be more robust in this regard. Tao et al. (2015) [26] proposed a geometric approach for foliage filtering which used only the spatial coordinates of the points as inputs to cluster and segment the foliage and wood points. Ferrara et al. (2018) [27] presented an automated method in which the points were partitioned into voxels and a DBSCAN clustering was carried out using features computed at a voxel level.

Random forest algorithm being widely used for the classification of point cloud data obtained using any type of platform including space borne, aerial borne, terrestrial and mobile through the pixel based or voxel based approaches owing to its ability to reduce over fitting and variance thus improving the accuracy. Xue et al (2020)[28] have applied improved Cloth Simulation Filter (ICSF) and weekly correlated Random Forest Classifier to classify point clouds of urban features namely buildings, vegetation and other man-made features with 4% improved accuracy. Zhu et al, (2018)[29] have used geometric and radiometric features based Random Forest methods to classify the foliage and wood point clouds of a mixed forest environment of Bavarian Forest National Park, Germany. Similarly, geometric features estimated through fast KD tree to classify leaf, trunk and ground using RF and xgboost classifiers with a maximum of about 0.9 kappa coefficient for the forests of Mongolian oak plantation [19]. RF classifier was applied on the clump covariance eigenvalues by constructing three dimensional and two dimensional features and recursive elimination process to identify the photsynthetic, non-photosynthetic and ground components of Washington Park USA ([30]. They could achieve an overall accuracy of 89% also leading to estimation of Leaf Area Index. In general RF algorithms outperformed the classification of point clouds based on satellite data [31], Aerial plat forms [28], UAVs and mobile [32] as well as relating and validating ICESat -2 based height parameters against height derived from airborne point clouds [31].

Vicari et al. (2019) [33] presented another automated approach which combined unsupervised clustering of point geometric features and shortest path analysis to identify points belonging to the main stem of the tree. Moorthy et al. (2019) [34] proposed a supervised foliage filtering method that used point-wise features computed from radially bounded nearest neighbours at multiple spatial scales. Wang et al. (2020) [35] developed LeWoS, a fully automatic geometric tool for foliage filtering of tree laser scans. They used a recursive point cloud segmentation and regularisation approach. From the extensive review of the specified methods for foliage filtering, it was observed that computing geometric features for each point at multiple local neighbourhood scales made the classifier invariant to point clouds with highly varying point densities. Owing to these

 doi:10.20944/preprints202210.0190.v1

observations, the point-wise geometric features computed at multiple scales were used for foliage filtration in this study.

To measure the various forest inventory properties from laser scans of forest stands, one of the preliminary steps would be the instance segmentation of each tree from the LiDAR data. This is required because most of the methods for estimation of biophysical properties of trees often require point clouds of single trees as inputs. The existing methods that were relevant to this task were studied and the following observations were made. Of the available methods for individual tree segmentation from laser scans of forests, the region-growing methods [36–38] required that the forest stand be relatively simple with well-spaced near-vertical trees for good performance. They may not perform well on terrestrial laser scans of mangrove forest stands which are characterised by a more complex branching structure and overlapping tree crowns. The graph-cut approach for individual tree segmentation proposed by Yang et al. (2016) [39], despite performing better than the region-growing approaches, still did not perform satisfactorily on point clouds of complex forest stands with many misclassifications of branches of one tree as that of another. Zhang et al. (2019) [40] proposed a segment-based approach which involved thinning of dense TLS point clouds using the curvature points followed by a connected components segmentation to delineate the individual trees. This method appeared to perform better in comparison to the other methods on point clouds of complex forest stands such as those of mangrove forests.

In order to estimate the biophysical parameters of individual trees from their LiDAR point clouds, an important step is to reconstruct the polygonal 3D model of the tree. It is from this polygonal representation of the tree, called a quantitative structure model (QSM) that the structural and biophysical parameters of the tree can be reliably measured. Boudon et al. (2014) [41] developed an algorithm called "PlantScan3D" originally to generate QSMs for studying plant development in the context of agronomy and biology. In this method, a branch structure graph (BSG) was constructed from neighbouring points. Then, cylinders were fitted to each part of the skeleton to reconstruct the volume of the tree. Landes et al. (2015) [42] developed an algorithm called "TreeArchitecture" in which the skeletonization method of Cao et al. (2010) [43] was used followed by a Delaunay triangulation to create raw skeletons of trees. This skeleton is then extended at its extremities followed by cylinder fitting to each segment of the skeleton to reconstruct the tree model. Hackenberg et al. (2015) [44] developed a method called "SimpleTree" which involved fitting cylinders to segments of skeleton nodes extracted by cutting spheres with the point cloud. This method also allows the estimation of biophysical parameters of a tree in a hierarchical order. Of the different methods reviewed for reconstruction of QSMs from point clouds of trees, the "SimpleTree" method [44] was chosen in this study to produce most realistic QSMs and more accurate volume estimates for each tree.

Based on the above literature survey, we set the following objectives for the present study:

1. To compare the Standard Random Forest (RF) and Weighted RF algorithms to segregate foliage and woody 3D point cloud of terrestrial laser scans collected from the mangroves of Maharashtra, India.
2. To accurately estimate the above-ground volume and biomass of mangroves as an individual tree and as a plot by generating a quantitative structure model of the trunks from the classified 3D point cloud.
3. To explore the spatial search algorithm in quantifying the contribution of pneumatophores present in the plot for the estimation of AGB estimation.

 doi:10.20944/preprints202210.0190.v1

2 Study area and Field Data Collection

2.1 Study Area

Maharashtra, the third-largest state in India accounts for a coastal length of 720 km running from North to South [45]. The mangroves of Maharashtra coast (is popularly known as Konkan coast) known for its extensive plant diversity. Most of the mangrove forests are spread across the districts of Mumbai City and Mumbai Suburban, Raigad, Ratnagiri, Sindhudurg, and Thane (FSI 2019) accounting to 320 sq.km. The region was experiencing a constant increase in environmental stress due to various anthropogenic activities during 1980s to 2010s [46]. However, later about 134sq.km of increase in forest cover was noticed between 2009 and 2019. Hence, there is a need to assess and monitor the stock of the forest cover for the sustainable conservation and management that motivate the present study on estimating the Above Ground Biomass of the mangroves. The area chosen for this study is bounded by the latitudes 19.029080° N to 19.164466° N and the longitudes 72.922204° E to 73.020012° E. This study area is also characterised by extensive adjoining urban cover where the main contributing rivers are the Ulhas and the Vaitarna rivers (Figure 1).

The major creeks in the study area from north to the south include Vasai, an estuarine creek in the northern boundary of Salsette Island; Manori and Malad creeks, located in the western waterfront of Mumbai; Thane creek, a 26 km long stretch located in the midst of the bustling cities of Mumbai and Thane; Panvel creek, a tide-dominated creek that opens up in Thane creek; Karanja creek, situated 30 km from Mumbai and Dharamtar creek, which falls along the central west coast of India. The climate, in general, is warm and humid and receives average annual precipitation 2142mm in monsoon season. The study area is dominated by *Avicennia marina*. Other mangrove species in the region include *Avicennia officinalis, Sonneratia apetala, Aegiceras corniculatum, Bruguiera cylindrica, Rhizophora mucronata, Rhizophora apiculata, Excoecaria agallocha* and associated species like *Acanthus ilicifolius, Salvodora persica* (Mugade and Sapkale 2014) [47]. The data used for this study were collected from the mangrove forests of Thane Creek, Mumbai (Figure 1).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2022 doi:10.20944/preprints202210.0190.v1

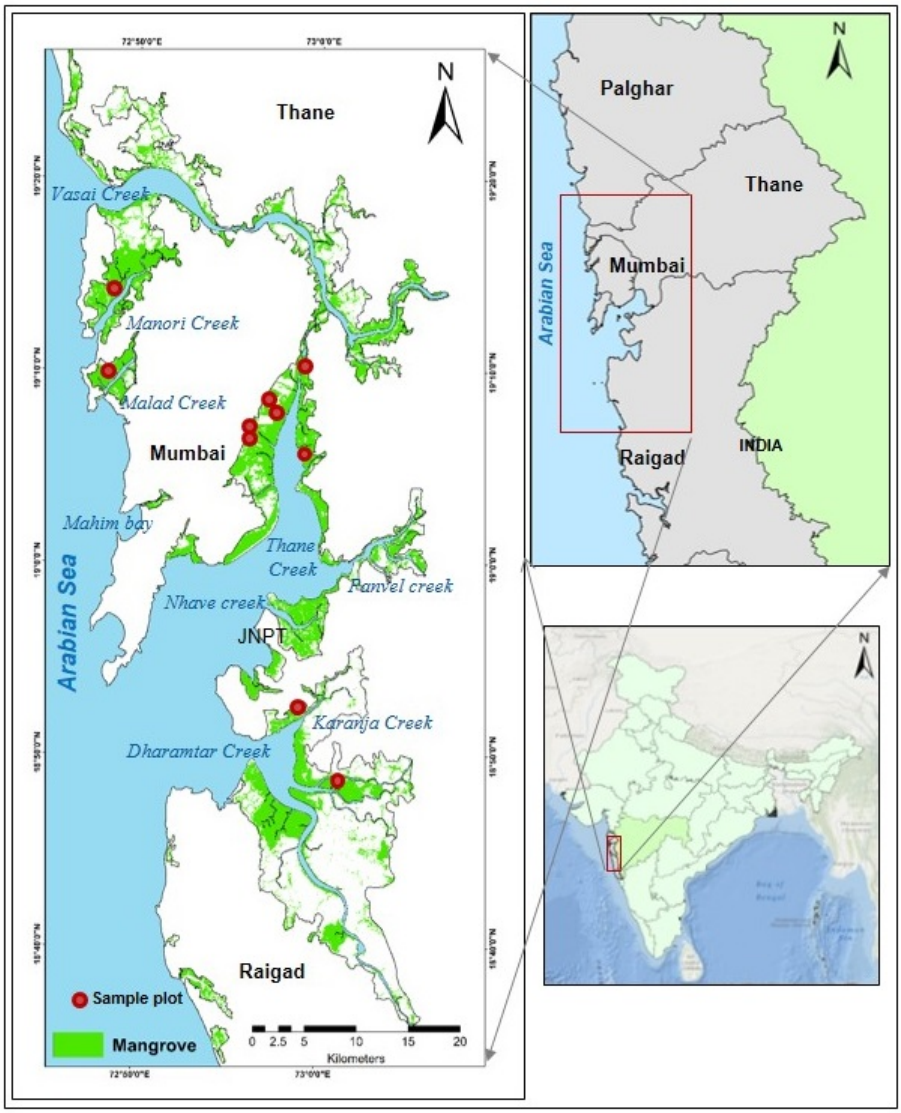


Figure 1. Study Area showing the mangroves of Mumbai, its surroundings and the sample locations of TLS data collection

 doi:10.20944/preprints202210.0190.v1

2.2 Field data collection

Figure 2. Laser scanning using FARO Focus S 350 Terrestrial Laser Scanner.

The various sources of the data used in this study and their modes of acquisition are detailed in this subsection. The primary data used in this study was acquired from the field surveys and terrestrial laser scans of the mangrove forest locations. Tree level and plot level laser scans of the mangrove forests are acquired using FARO Focus S 350 terrestrial laser scanner (Figure 2). FARO sensor has a scanning range of 0.6 to 350m, with measurement speed of 976,000 points per second and field of view of 300° vertical, 360° horizontal at the wavelength of 1550nm and was chosen for this study owing to its compact size and ease of use, which made it ideal for traversing the difficult terrain in mangrove forest plots. Apart from the TLS scanner, a laser distance meter was used to measure the tree heights and a handheld GPS receiver was used to acquire the coordinates of the various plot centres and individual trees. Mangroves are under very marshy environment that are hard to conduct field survey, install the instruments and controls within the muddy soil background. This limits the number of sample plots of size 33 x 33m to ten (this size is chosen to meet the resolution of open source multispectral data of Landsat for future upscaling of biomass inventory) (Figure 1). Eventhough only ten plots are surveyed, an average of 80 individuals present in each plot were measured for further modelling and analysis similar to survey carried out in the mangroves of Everglades by Feliciano et al (2014). Two modes of data were acquired in this study:

i Plot level terrestrial laser scans of ten forest plots of size 33m x 33m.
ii Laser scans of individual trees.

Ten plots of size 33m x 33m ($\approx$ 0.1 ha) were laid out at the locations given in Figure 1. At each plot, the following data were collected:

i Terrestrial laser scans of the plot from 9 scan stations.
ii Diameter at Breast Height (DBH) of each tree in the plot.
iii Wood core specimen from sample trees in the plot.
iv Total tree count in each plot.

All of the plots chosen for the study were dominated by *Avicennia marina* with associations of *Acanthus ilicifolius*. However for the current study, individual tree scans

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2022 doi:10.20944/preprints202210.0190.v1

(a) **(b)**

Figure 3. Multi-station Terrestrial Laser Scanner Survey Layouts of (a) a plot and (b) a single tree

were taken from three girth classes namely Large (girth greater than 90cm) Medium (girth equal or greater than 40.4 cm), and Small (girth equal or greater than 29.5 cm) measured at breast height to estimate bio-volume of the above ground part of the tree contributing significantly to AGB. Also, one plot is used to segment the trees and estimate plot level Above Ground tree volume.

The plot was scanned from nine scan stations following the layout (Figure 3a). Five scan stations were placed within the plot boundary and four other scan stations were placed outside the plot to ensure comprehensive coverage of the entire plot and minimise occlusion. Six white spherical targets were placed inside every plot in such a way that at least four of the six targets were visible from all the scan stations. In case of individual trees, each tree was scanned from three positions(Figure 3b). Four spherical targets were placed around the tree such that all of them visible from each scan station. These scans were later registered, exported and preprocessed to a single point cloud for both the plot-level and tree-level scans. In addition to the laser scans of each tree, wood core samples, diameter at breast height and tree height were also collected for each tree during the survey. Pnematophores were counted and heights were measured in five 1mx1m square subplots within the plot.

3 Methodology

The above-ground biomass (AGB) estimation was carried out through a set of three stages using terrestrial LiDAR scanned data for both individual trees and forest plots (Figure 4).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2022 doi:10.20944/preprints202210.0190.v1

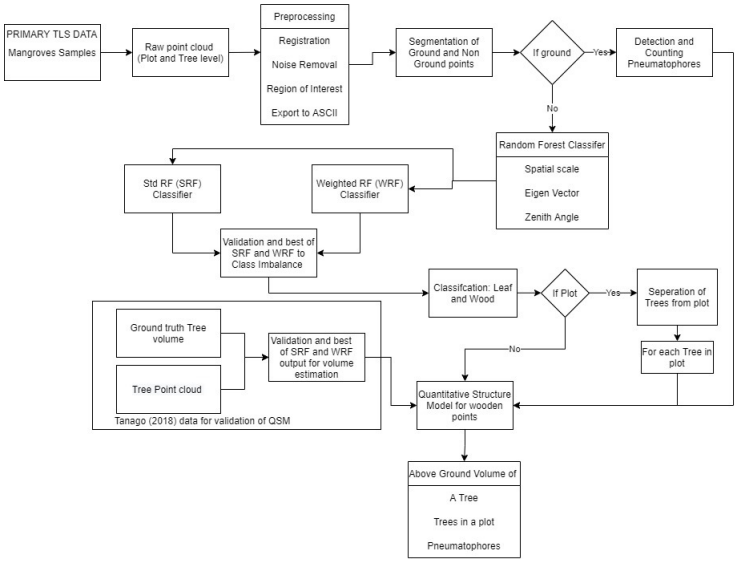


Figure 4. Methodology of the steps involved in estimating above ground volume of a tree, groups of trees in a plot and pneumatophores.

3.1 Stage 1: Preprocessing of the TLS Scans

Upon acquisition of the TLS scans of individual trees and plots as detailed in Section 2.2, the point cloud data are subjected to a series of preprocessing operations.

The scans from each of these stations were registered to form the complete point cloud of the tree or plot. For this purpose, all the scans within a single survey project were registered together using automatic target-based registration provided by the FARO SCENE software. This resulted in a complete point cloud of the entire scene that was visible to the sensor from the multiple scan stations during the survey. This point cloud was then converted to ASCII (.xyz) format.

The point cloud after registration was found to contain a lot of noise points due to ghosting of moving branches and leaves during scanning. It also possessed a very high point density that made manipulation of the data memory intensive. For these reasons, the point cloud was first spatially downsampled by voxelisation, with a minimum spacing between points in the point cloud of 1 cm. After downsampling, the noisy points in the point cloud were removed using a statistical outlier removal tool. The tool computes the average distance of each point to its nearest 'k' neighbours and removes points that are farther than the sum of the average distance and a multiple of the standard deviation. A 'k' value of 6 and a standard deviation multiplier threshold of 1.0 were chosen. These operations were carried out in the open-source CloudCompare [48] software.

After down sampling and noise removal, the resulting point cloud consisted of several background objects and regions beyond the plot boundaries. This was because the TLS was allowed to record points along a 360° horizontal field of view during the survey. To only extract the plot or individual tree of interest from this point cloud, the region of interest in the point cloud was visually identified and manually extracted. The resulting individual tree or forest plot point clouds were then exported in a suitable ASCII format (.xyz). These point clouds are the final results of the preprocessing pipeline and form the inputs to the subsequent pipelines for estimation of above-ground biomass (AGB).

3.2 Stage 2: Segmentation and Volume Estimation

The preprocessed terretrial LiDAR point cloud data was subjected to

i Segregation of ground and non-ground points

 doi:10.20944/preprints202210.0190.v1

ii Segmenting non-ground points into either wood or foliage points
iii Generation of the Quantitative Structure Model (QSM) and then
iv Estimation of the above-ground volume of wooden structure of individual trees and plots

The first step of our approach is to separate and remove the ground points and non-ground points from the preprocessed point cloud. Although, in the case of mangrove species the ground points of terrestrial LiDAR scans will contain pneumatophores (Figure 2, the contribution of pneumatophores to AGB at an individual tree level varies depending on the species. This study considers the tree level estimation for the species *Avicennia marina* for which the contribution of pneumatophores to AGB is negligible. Hence, the ground points of the tree point cloud are removed along with the pneumatophores for the scope of this study at the tree-level analyses, however, we made an attempt to estimate the contribution of pneumatophores at plot level discussed at the end of this research. To remove the ground points from the point cloud, the cloth simulation filter [49] plugin in CloudCompare was used. Cloth Simulation filter is commonly used to generate the approximation of surface at the ground surface of objects or features on the ground when point clouds are absent or very less on the inverted point clouds and here used to separate the ground points from the above ground points.

After the removal of ground points, the resulting point cloud was broadly grouped into two categories i.e. those belonging to photosynthetic parts of the tree (leaves) and those belonging to the non-photosynthetic parts of the tree (stem, branches and other wooden components). Out of them, the contribution of leaves to AGB is considered to be minimal and hence ignored for the scope of this study and thus, the points belonging to the wooden components of the tree are considered most essential in estimating the AGB, contributing to more than 97% of AGB of mangroves [50]. To perform this, a supervised random forest classifier was used to classify the points into two classes (leaf and wood) using the features computed at multiple spatial scales as described by Moorthy et al. (2020) [34]. Five spatial scales of 0.1m, 0.25m, 0.5m, 0.75m and 1m were chosen to compute the features. At each of these spatial scales, the eigenvalues $(\lambda 1, \lambda 2, \lambda 3)$ and the zenith angles $(\theta 1, \theta 2, \theta 3)$ of the corresponding eigenvectors of the covariance matrix for each point were computed. The covariance matrix of the N 3D points in a neighbourhood can be computed from Equation 1 and the azimuth (θ) of the eigenvectors, V can be calculated as in Equation 2.

$$\Sigma = \frac{\sum_{i=1}^{N}(X_i - \bar{X})(X_i - \bar{X})^T}{N} \tag{1}$$

$$\theta = arctan(\frac{V_2}{V_1}) \tag{2}$$

where, V_1 and V_2 are the first and second elements of the eigenvector respectively. This led to 6 features being computed for each scale which resulted in a total of 30 features across multiple scales for each point in the point cloud. The first three features among these are plotted in Figure 6 to depict the separability achieved from these features.

These features were then used to classify the point cloud using a supervised random forest (RF) classifier. As there existed a class imbalance between the foliage and wood point clouds as discussed later in the results section 4.2 we used weighted RF classifier also to handle the class imbalance. The weighted random forest classifier is essentially the 'balanced' variant of the standard random forest classifier in the scikit-learn [51] python library. This classifier assigns weights to each class inversely proportional to the corresponding class' frequency in the input dataset. The procedure followed for the leaf and wood points separation [34] is detailed in Algorithm 1. To validate the classification results we used the standard classification metrics such as Balanced Accuracy Score, Cohen's Kappa Score, F1 Score and Area under Receiver Operator Characteristics (ROC-

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2022 doi:10.20944/preprints202210.0190.v1

AUC) Curve and it was made by performing 10-fold cross-validation on 818328 points of 10 mangrove trees of large, medium and small diameters which are manually labelled.

Algorithm 1: Leaf Wood Classification

```
Structure tree point cloud P, using kdtree data structure.
Define a set of spatial neighbourhood scales, S.
  S = {0.1m, 0.25m, 0.5m, 0.75m, 1.0m}
for s ∈ S do
    for each point p ∈ P do
        Identify nearest neighbours n, within s distance from p.
        Compute covariance matrix of n.
        Compute the eigenvalues and zenith angles of corresponding
          eigenvectors of the covariance matrix.
Train a supervised weighted RF classifier on a training set with these features.
Classify the point clouds of individual trees using the trained classifier.
Output (O) is a point cloud with each point assigned a label as either wood or
  leaf.
```

3.3 Stage 3: Development of Quantitative Structure Model

After segmenting the tree point cloud, it is now necessary to generate a Quantitative Structure Model (QSM) from the wooden points of the point cloud. The SimpleTree [44] plugin in the Computree [52] software platform was used to generate the QSMs from the wood points of the tree point clouds. In this method, the wooden components of a tree are modelled as a hierarchical collection of cylinders. The output of this procedure is a polygonal mesh file accompanied by an ASCII file containing information such as cylinder descriptions, the volume of the tree, number of branches in each level of the branching hierarchy, etc. From this output, the above-ground volume of the tree is retrieved which in turn is used to calculate the above-ground biomass of the tree as a product of its estimated volume and wood density. The wood density of the tree can be obtained by measuring the specific gravity of the core wood samples of that tree, by a simple oven drying experimental setup. Current study assumed a uniform density for the tree to calculate AGB as studied by Fajardo [53]. As there was no permission to get the above-ground volume by destructive sampling of mangrove trees and parts, we validated the processes of the volume estimation using a benchmark dataset of ground truth tree volume obtained by destructive methods from the LUCID repository [54].

3.4 Above-Ground Biomass Estimation of Forest Plots

Analogous to the estimation of AGB from individual tree point clouds, the first step in estimating the AGB of forest plots is to separate the ground and non-ground points in the forest plot point cloud. The same procedure for ground segmentation of a single tree point cloud is followed here. However, in contrast to the case of individual trees, the ground points in the plot level point clouds contain a large mass of points returned from pneumatophores. Although the contribution to AGB by pneumatophores was ignored in the case of individual trees, their contribution to the AGB of forest plots is not negligible. Due to this, a parallel workflow is carried out to estimate the contribution to AGB of forest plots by the pneumatophores in the ground points. Each tree in the plot-level point cloud is classified into wood and foliage components using the procedure followed in the case of a single tree point cloud described earlier by deriving geometric features (Section 3.2).

After the removal of foliage points from the plot level point cloud, it is necessary to detect and segment the individual trees from the plot level point cloud. This is so

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2022 doi:10.20944/preprints202210.0190.v1

that each tree point cloud in the plot could be used as an input for generating QSMs which in turn could be used to estimate the AGB of plots. To segment the individual trees from plot-level LiDAR scans, the connected components segmentation tool of the open-source software platform, CloudCompare [48] was used. This tool is a version of the classic connected components segmentation algorithm described in [55], extended to work with 3D voxel grids. To perform this segmentation, the point cloud data is first voxelized into a uniform grid. From the voxelized point cloud, a 3D binary occupancy grid is constructed. Once the 3D binary grid is constructed, the connected component segmentation is carried out as detailed in Algorithm 2. The result of this algorithm stored each tree as an ASCII (.xyz) point cloud file with a tree ID assigned to each file.

Algorithm 2: Connected Component Segmentation

The point cloud is converted to a 3D binary occupancy grid.
Each cell in the binary grid is categorised as either background or foreground.
Only foreground cells are considered for segmentation.
Select first cell in 3D grid. Current cell = C.
Initialise current label, *curr_label* to 1.
Initialise empty queue, Q.
if $Cell, C \in Foreground\ \&\ label(C) = None$ **then**
    Set label(C) = *curr_label*.
    Add C to Q.
    Go to step 13.
else
    Repeat steps 6-12 for the next cell in the 3D grid.
Pop cell from queue, Q.
Add all neighbours of the current cell (based on 26 connectivity) to set N.
for $eachneighbour, n \in N$ **do**
    if $n \in foreground\ \&\ label(n) = None$ **then**
        Set label(n) = *curr_label*.
        Add n to Q.
Repeat steps 13-20 until queue (Q) becomes empty.
Choose next cell in 3D grid as C & increment current label by 1. *curr_label* += 1
Start again from step 5.

QSM for each tree was constructed on the segmented point clouds using the SimpleTree [44] plugin in the Computree [52] software. The QSM of each tree was stored as a polygonal mesh file with its corresponding tree ID and its above-ground volume and in turn AGB were estimated using the average density of *A marina* from field measurements as 513 kg/m ^{3}. It is estimated using samples of stem cores collected from 10 trees of A. marina of varying DBH. The TSL based estimation of AGB is compared against the conventional method of allometric equations[56]. Even though, the allometric equations are developed for matured trees with more than 10cm diameter at breast height (DBH) they are widely used for rapid estimation where destructive measurements are either restricted or to reduce cost interms of time and labor [57]. However, most of the satellite remote sensing based biomass models are developed using allometry based biomass estimates [58].

In the case of plot-level point clouds, it is also necessary to estimate the overall contribution of the pnematophores to the above-ground biomass of the forest plot. As the pneumatophores lie just above the ground surface for a few centimeters height specific to each species, first, a set of hierarchical layers based on elevation were created from the ground points containing the pneumatophores. These ground points were one of the results of the first segmentation to separate the entire point cloud into ground and non-ground point clouds. Then, the layer having the median elevation was extracted based on the species present, *Avicennia marina* for this study and exported as a separate

 doi:10.20944/preprints202210.0190.v1

ASCII file. This median layer of the ground points was then converted to an undirected graph. In this graph of median layer ground points, an unlabelled point was selected at random as the root node, a label was assigned to this root node and a breadth-first search was conducted to identify all points connected to the current root node. Based on field observations, a threshold of 3cm for the horizontal distance between the root node (based on the species *Avicennia marina*) and the visited node was chosen as the connectivity criteria. If a visited node satisfied the connectivity criteria, then it was assigned the same label as the root node. After all the nodes are visited, a new root node from the remaining unlabelled points was chosen and the process was repeated until all points in the input point cloud were assigned a label. The result of this procedure was a point cloud where points belonging to each pneumatophore of the plot were assigned a unique label. To validate the count of pneumatophores by this method, five test patches of size 2m x 2m were extracted from across several plots and the pneumatophores in these patches were manually counted. This manually counted number of pneumatophores was used as a reference to validate the estimated count of pneumatophores. Due to the minimum number of point clouds for each pneumatophore, construction of 3D structure was not attempted. Instead, once each pneumatophore was identified and counted in a plot with reasonable accuracy, it was modelled into a conical shape to estimate its volume based on the height estimated through segmented pneumatophores and the average diameter of 1cm based on in situ sample observations. And its biomass is estimated using the uniform wood density 513kg/m^2 for *Avicennia marina*) as mentioned in Sec 3.3.

4 Results

This section presents the results of the various algorithms sequentially implemented in this study towards the estimation of the tree and plot-level AGB and discussion about the accuracy measures of each estimation from the point clouds data.

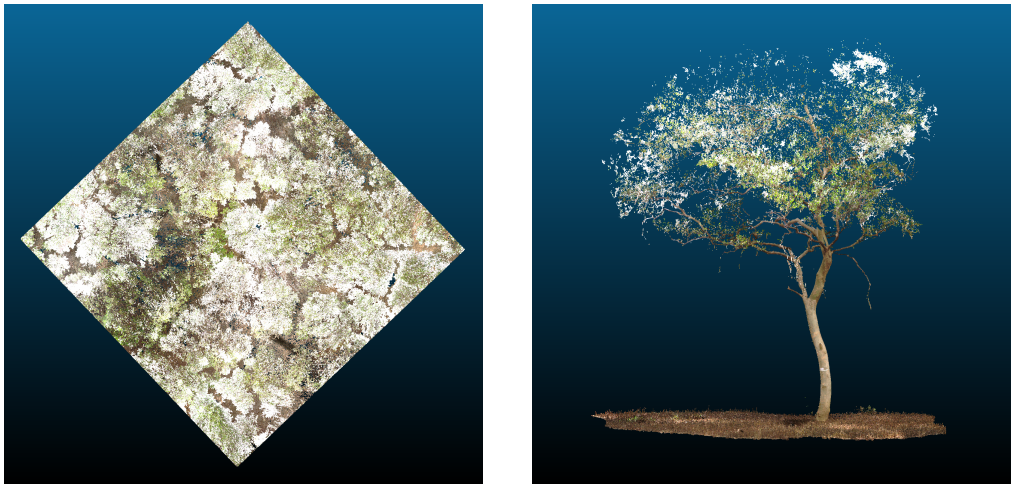
4.1 Preprocessed Point Clouds

The TLS scan data acquired from nine scan positions for the plot level scans and three positions for tree level scans were registered using the FARO SCENE software using the target spheres as control points. Due to large memory occupied by integration of nine scans of the plot, the integrated point clouds were down sampled in to half and then used for further processing. For the tree level processing, point clouds that are falling within the projected extent of the canopy on the ground space was extracted and used further (Figure 5). Canopy crowns of the individual mangrove trees represented using the RGB value obtained by the TLS within the plot could be seen from the top view of the plot level point cloud (Figure 5a) along with open ground between them too. While the side view of a single tree of large DBH displayed with RGB shows clear distinction between point clouds of trunk and foliage (Figure 5b).

4.2 Segmentation of Tree Point clouds into Photosynthetic and Non-photosynthetic Components

To separate the points corresponding to the photosynthetic and non-photosynthetic components of a tree from each other, a Random Forest (RF) classifier was implemented. To classify the points, geometric features were computed for each point at multiple spatial scales. The chosen geometric features were the eigenvalues and zenith angles of eigenvectors of the covariance matrix of the nearest neighbours of a point [34]. As the class imbalance between foliage and wood point clouds was evident from the scatter plot of the first three features (Figure 6), the performance of the standard RF classifier was tested against a weighted RF classifier on the labelled point cloud of mangrove tree and compared using standard classification metrics. The comparison was made

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2022 doi:10.20944/preprints202210.0190.v1



(a) (b)

Figure 5. The preprocessed point cloud obtained from TLS scans of (a) a plot of size 33mx33m and (b) a single *A marina* tree of large DBH.

by performing 10-fold cross-validation on our labelled dataset of mangroves. Upon comparison, it was observed that there was an average increase of 20% in the balanced accuracy score, 0.3 in Cohen's Kappa score, 20% in the ROC AUC score and 21% in the F1 score when using the weighted RF classifier over the standard RF classifier (Figure 7). The labeled point clouds and the output of the weighted random forest classifier of a mangrove tree, *A. marina* is depicted in Figure 8.

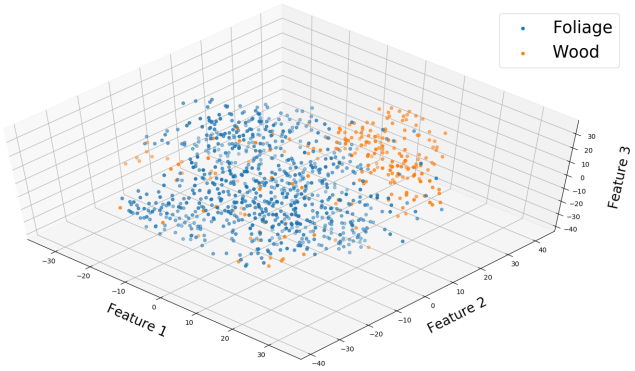


Figure 6. A 3D scatter plot (tSNE) of first three features out of the 30 features derived for the Random Forest Classification shows the clear separability and class imbalance between foliage and wood points.

From these results, it is apparent that the weighted RF classifier performs consistently better than the standard RF classifier in all cases of foliage vs wood classification. The final classification outputs of the weighted random forest classifier on the *Avicennia marina* trees of various diameter classes are depicted on left three images of Figure 10.

4.3 Estimation of Above-Ground Biomass of a Single Tree

After filtering the foliage points from a tree point cloud, the remaining wood points in the point cloud were used to generate a QSM. The estimated volumes of the QSM generated for the benchmark point cloud data of ten trees of Tanago et al, [54] using both the standard and weighted RF classifiers are illustrated along with their performance

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2022 doi:10.20944/preprints202210.0190.v1

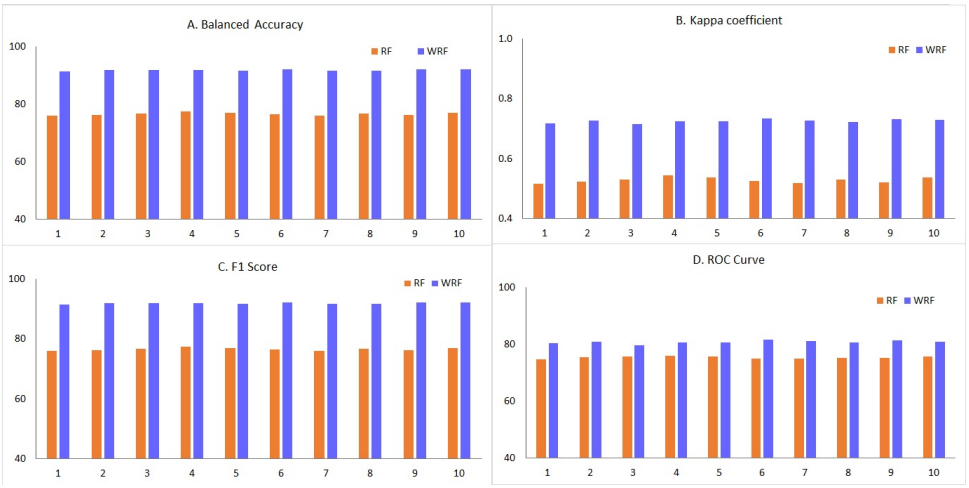


Figure 7. Balanced Accuracy(%), Cohen's Kappa coefficient, ROC AUC(%) and F1 Scores(%) of Standard RF Classifier (orange colored bar) vs Weighted RF Classifier (blue colored bar).

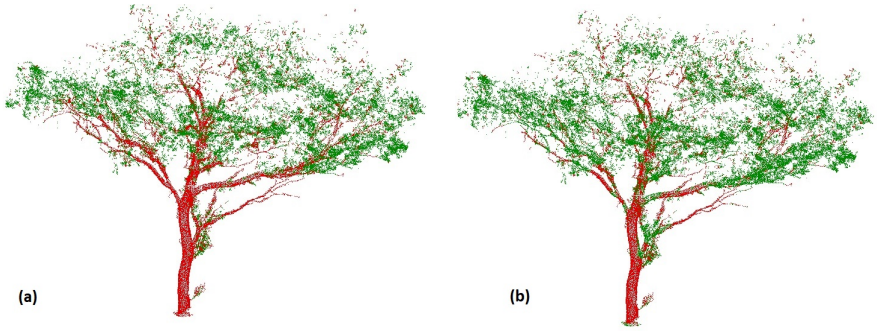


Figure 8. Validation of Segmentation Results of Weighted RF Classifier (a) Labelled Point Cloud and (b) Prediction of the Weighted Random Forest Classifier of *A. marina*

for tree volume estimation in Figure 9. From the results of the QSM reconstruction, the total above-ground volume was obtained, which in turn was used to calculate the above-ground biomass.

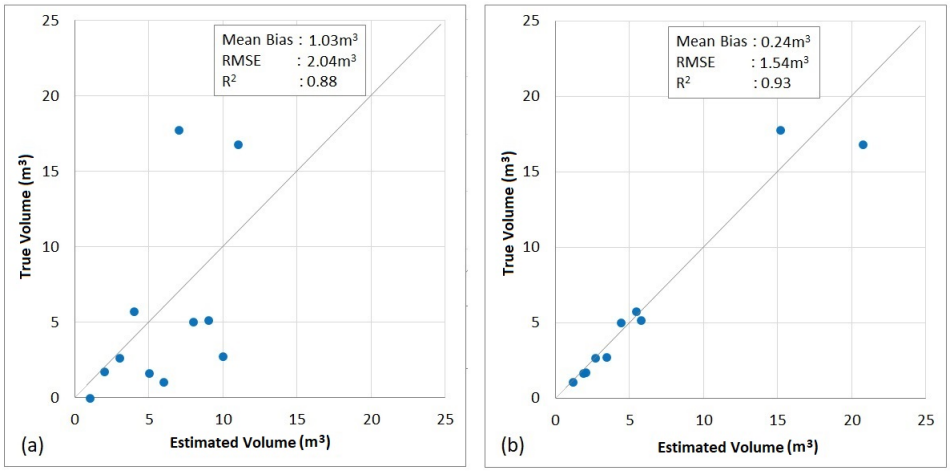


Figure 9. Performance of (a) Standard RF (SRF) and (b) Weighted RF (WRF) Classifiers output in estimating tree volume using data from Gonzalez de Tanago *et al* [54]

From these results, it is evident that the tree volumes estimated using the classification results of the weighted RF classifier correlates better with the ground truth volumes

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2022 doi:10.20944/preprints202210.0190.v1

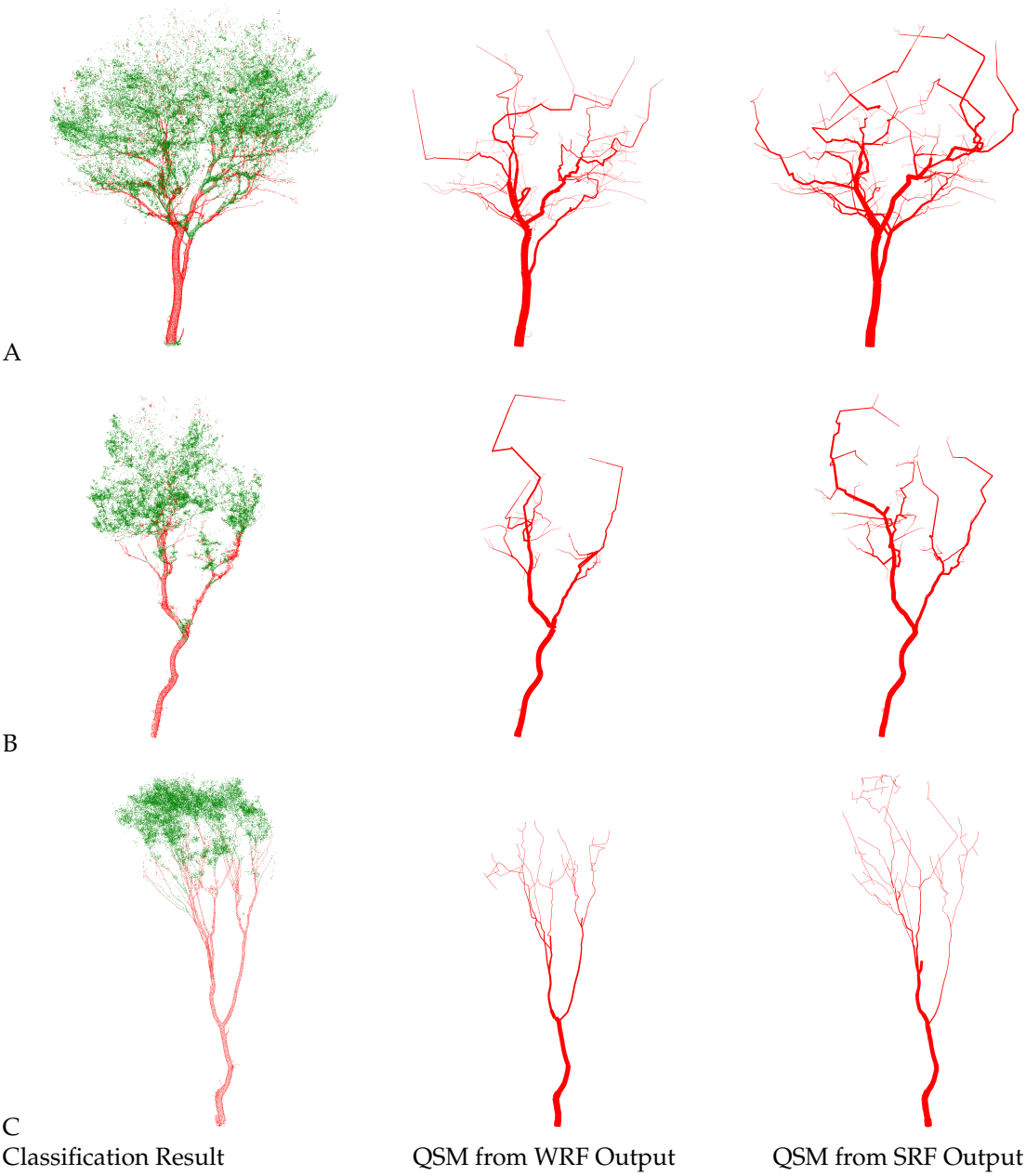
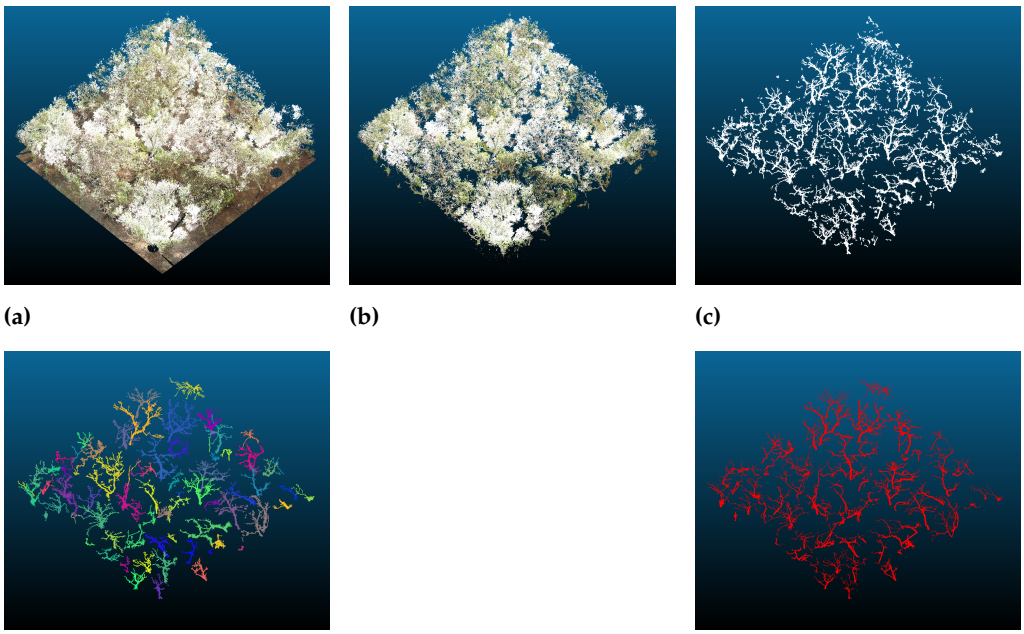


Figure 10. QSMs of *Avicennia marina* trees generated from Outputs of Weighted (WRF) and Standard (SRF) RF Classifiers. (A) Large (B) Medium and (C) Small diameter classes. Overestimation of branches is reduced in the case of the QSM from WRF outputs.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2022 doi:10.20944/preprints202210.0190.v1

than those from the classification results of the standard RF classifier. Based on the better accuracy obtained for the test data we applied both classifiers on *A. marina* of three different DBH. The QSMs generated from both the SRF and WRF outputs have been depicted in Figure 10. From these figures and visual interpretation, it is observed that the presence of misclassified foliage points in the output of the SRF classifier results in the overestimation of the branches during QSM reconstruction. When the misclassification was reduced by addressing the issue of class imbalance using WRF classifier, the overestimation of branches in the QSMs generated was also reduced and was more representative of the actual tree structure when viewed in three dimensional mode as well.



(a) **(b)** **(c)**

(d) **(e)**

Figure 11. Outputs of the Sequence of Steps followed for the Estimation of Above-Ground Biomass of an Entire Forest Plot. (a) Preprocessed Plot-Level Point Cloud. (b) Ground Segmented Plot-Level Point Cloud. (c) Foliage Filtered Plot-Level Point Cloud. (d) Individual Tree Segmentation of the Foliage Filtered Plot-Level Point Cloud. (e) QSM Reconstruction of Each Tree in the Plot-Level Point Cloud.

4.4 Results of plot level analysis of Forest Point Clouds

Similar to the process carried out for the point clouds of individual trees, classification was carried out for foliage and wood separation but here for the entire plot and then segmented into individual trees of the plot. Once each tree in the plot level point cloud was segmented, the QSM for each tree was constructed and the structural parameters namely tree height and diameter at breast height (DBH) were estimated. The point cloud results of the various processes carried to estimate the above-ground biomass of forest plots have been depicted in Figures 12a - 12e. The structural parameter estimation based on TLS point clouds was comparable with height measured in the field however, it is varied for DBH measurement (Figure 13). The huge variation in DBH was due to the omission of thin trees during field survey below 7 cm and varied number of samples. DBH ranges from 8 to 29 cm (average of 16.4 ±7.4 cm) and tree height from 5.7 to 9.5 m (average of 7.3 ±1.3m) from of field measurements while they range from 2.6 to 9.8cm (7.3 ±1.5cm) and 3.5 to 24.1 m (average of 11.4 ±4.8m) respectively from TLS based estimation. As such there are no allometric formulae available to estimate tree volume using from height and DBH rather than biomass for mangroves, our comparison

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2022 doi:10.20944/preprints202210.0190.v1

is restricted to the tree height and biomass estimation only. From the QSM output of

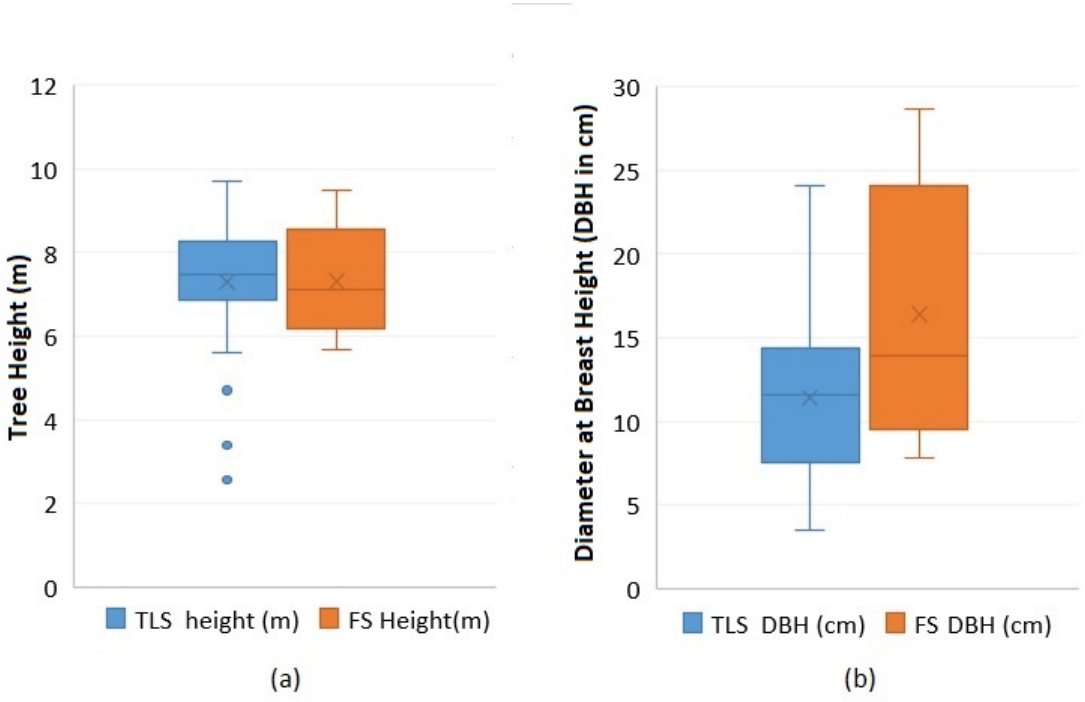


Figure 12. Comparison of (a) Diameter at Breast Height(DBH) and (b) Height estimated for the trees of sample plot using TLS derived tree structure and the field survey measurements

each tree, it was then possible to obtain the above-ground volume of all trees in the plot (0.01 m^3 - 0.53 m^3), which in turn was used to compute the above-ground biomass of the entire forest plot with an average wood density of 513 kg/m3 resulting into an average of AGB of 44.11kg/tree with a range of 0.14 - 253.89kg/tree using allometric methods and 46.16 kg/tree with a range of 0.85 - 276.9 kg/tree using TLS based estimation(Figure 13). The average density of AGB using allometric methods and TLS based method was estimated as 1.74kg/m^2 and 1.82kg/m^2.

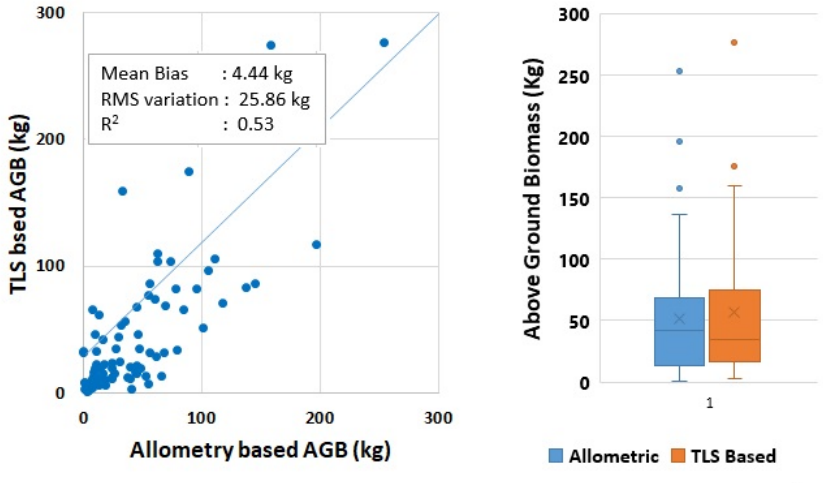


Figure 13. Comparison of above ground biomass estimated for the sample plot using allometric equation and proposed TLS derived AGB (a) scatter plot and Deviation metrics and (b) box plot chart showing their distribution

 doi:10.20944/preprints202210.0190.v1

Furthermore, from the ground points of the plot level point cloud, the contribution of pneumatophores to plot-level AGB was also calculated. This was done by performing a breadth-first graph search based segmentation of the ground points of the plot-level point cloud, to segment each pneumatophore shown in multiple colors (Figure 14) and obtain their total count from the plot level point cloud. The results and performance metrics of the estimation of pneumatophores in a forest plot have been illustrated in Figure 15 showing a overestimation of count that might have been resulted due to the point clouds acquired from closely spaced pneumatophores. Because of the presence of very dense pneumatophores and very few points for each pneumatophores, construction of the QSM from the point clouds was not yielding meaningfull result. Hence, each pneumatophore was constructed into a conical shape with a base diameter of 1cm and the height of each pneumatophore stand measured using segmentation process (Figure 14. c). The average volume of constructed pneumatophores is estimated as $0.401 m^3 / m^2$ of the sample plots contributing to an average of 213kg /plot of 33 x 33m size (resulting into increase in the AGB density of mangroves from $1.82 kg/m^2$ to $2.01 kg/m^2$).

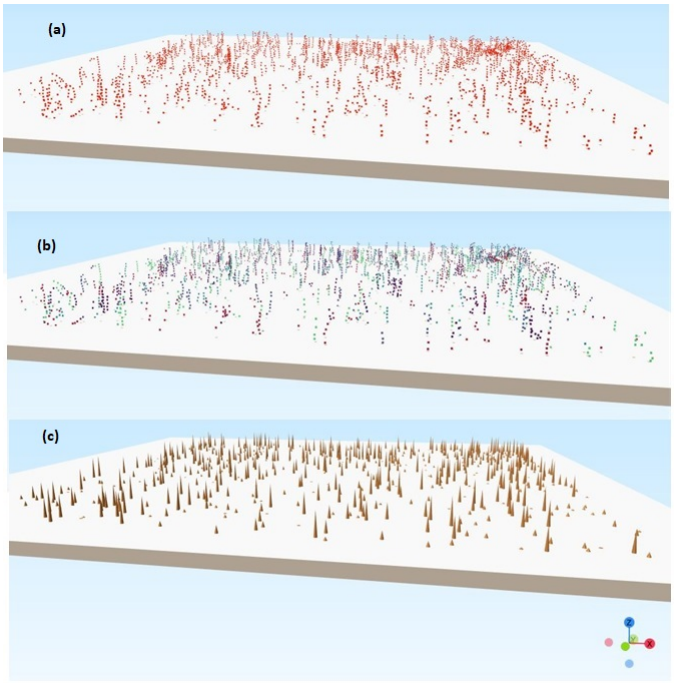


Figure 14. Detection of Individual Pneumatophores from a section of forest Plot (a) Point Clouds of Ground Points containing Pneumatophores (b) Individually Segmented Pneumatophores and (c) Constructed pneumatophores

The present analysis of TLS point clouds obtained from the sample locations of Mumbai mangroves resulted in the estimation of volume by constructing the Quantitative Structure Model of (i) *Avicennia marina* trees having large, medium and small diameters, (ii) plots having a number of *A. marina* trees and (iii) pneumatophores and thus, used to estimate the biomass of respective categories by incorporating density values computed from field samples.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 13 October 2022 doi:10.20944/preprints202210.0190.v1

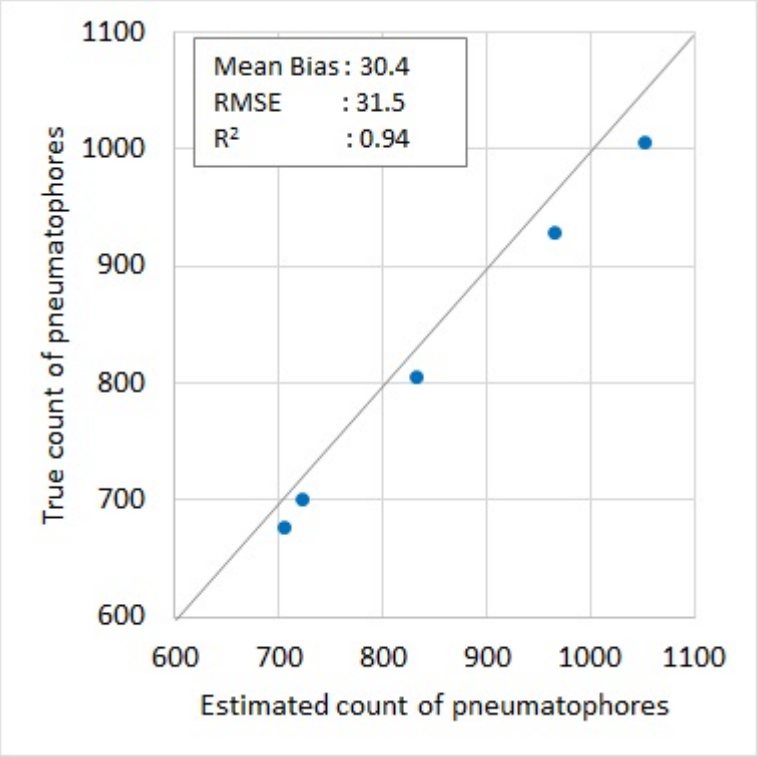


Figure 15. Performance of the Individual Pneumatophore Segmentation Method.

5 Discussion

The main objective of this study was to propose a processing pipeline to estimate the above-ground volumes of individual trees and forest plots from their terrestrial LiDAR data and then estimate the biomass. In this process we aimed at collecting sufficient field samples to estimate the mangrove biomass using lidar scans. In a similar study [59], 5 plots of 50 x 50m size of Everglades mangroves were surveyed using TLS and estimated the volume and biomass followed by validation using destructive methods. In our study, we chose ten plots distributed across the study area with the size of 33 x 33m with an average of 80 number of trees. Also for precise estimation, individual trees of large, medium and small DBH were also acquired.

Preprocessing for noise removal was carried out using statistical outlier method followed by ground segmentation was carried out using CLS filter. The resultant non-ground point clouds were first classified into foliage and wooden components using standard and weighter RF algorithms and followed by the construction of tree structure upon the wooden components by generating polygonal mesh file. The present state-of-the-art QSM reconstruction algorithms only accept the non-photosynthetic components of a tree point cloud that can form near cylindrical or spherical surfaces [44]. Hence, the foliage points in the tree point clouds are filtered from the tree point cloud to improve the accuracy of the tree volume estimates. It has to be noted that the foliage part of the tree was not considered in estimating the volume and biomass because, the mass to volume ratio of the foliage part of Avicennia marina is significantly lower when compared to the woody part of the tree [49]. The random forest classifier used in this study for foliage filtering performed noticeably better when adding a weight factor to each of the input classes [50] when Airborne Laser Scanner (ALS) data was analysed to classify the land cover features like ground, vegetation, building and water. This is due to the issue of foliage vs. wood class imbalance in point clouds of trees and forest plots i.e., the number of foliage points is several times higher than the wood points. This larger number of foliage points in the point cloud can be attributed to ghosting effects caused by the

 doi:10.20944/preprints202210.0190.v1

movement of leaves during the scanning process due to large angle of incidence while using TLS point clouds that were removed using the intensity values of point clouds[60].

Above ground volume estimation using the proposed approach resulted with a root mean square error of 2.037m^3 and 1.543m^3 using the standard and weighted RF classifiers respectively that could have resulted frthe substantial improvement (20%) in the classification accuracy of weighted RF. The structural parameter estimation based on TLS point clouds was comparable with height measured in the field however, it is varied for DBH measurement. Significant variation in DBH was noticed because thin trees (below 7 cm DBH) were omitted during field survey. As the mangroves of India including our study area are prohibited for biophysical parameter estimation using destructive methods, we could compare the AGB estimation of our study against the estimation using allometric equation as we could see from other research by Kargar et al (2020)[61]. In their study, the mangroves of Federated States of Micronesia have been surveyed using TLS for the detection of stems and roots using integrated RF and SVM methods and constructing tree structure using alpha shapes and convex hull. They have also compared the volume of TLS based estimation against allometric equations resulting a consistency of 85 % between them. While comparing of AGB using allometric equation and TLS approach of our study shows a mean bias of 4.4kg and RMS variation of 25.86 kg respectively for the plot. Also it shows a significant correlation between them with the value of 0.53. With the support of concerned authorities, if true biomass are estimated by destructive sampling, the established potential of the techniques could be proved for mangroves also.

Segmentation of individual pneumatophores of the forest plot was successfully carried out from TLS point clouds from the study. This could help in the quantification of pneumatophores and in turn their contribution to above-ground biomass in inventory studies of mangrove forests [13]. Based on the observations and inferences drawn in this study, further research can be carried out to explore the potential of deep learning algorithms to achieve better results in the leaf vs. wood classification of point clouds, to develop a robust QSM reconstruction including the presence of foliage points in the point cloud.

Further, the contribution of pneumatophores, especially for the species with large pneumatophores like Sonneratia apetalla would be most significant component of above ground biomass [13]. This can be better estimated by developing a methodology by combining the field observations and geometric features to capture the shape of the larger pneumatophores from the data instead of assigning an arbitrary shape to them.

Indian mangroves have been widely studied for the estimation of biophysical parameters such as height, biomass, basal area and Leaf Area Index (LAI) using varied remote sensing methods from multispectral data based regression to hyperspectral and microwave data which are purely based on the radiometric and scattering properties of the canopy cover. For example, Strong relationship between vegetation indices and basal area was noticed for the mangroves of Coringa by Satyanarayana et al. (2001) [62]; Bhumika Vaghela et al (2021) [63]regressed allometry based biomass against Sentinel 1A Synthetic Aperture Radar (SAR) to model the biomass of Gulf of Kutch mangroves; Spectral indices of EO1 Hyperion hyperspectral data to model the AGB of Bhitarkanika mangroves ([64] and [65]); George *et al* (2018) [66] estimated AGB and LAI of Andaman mangroves using Hyperion data using vegetation indices. Later, the integration of radiometric properties optical data and geometric properties like height from stereo photogrammetry were found to improve the accuracy [67]. Recent advancements in laser technology enable the accurate and precise field survey that pave way to upscale the satellite and aerial remote sensing in vegetation studies. One hectare of Savanna was mapped using laser scans at spatial resolutions greater than 2 m and 4 m obtained from three different platforms- terrestrial, mobile (MLS)and UAV (ULS) and no significant difference was found but their intergration could give better results [68]. Also the study was suggesting TLS to optimize the acquisition parameters of ULS and MLS. The use

 doi:10.20944/preprints202210.0190.v1

of TLS sampling can expand the options for the calibration and validation of multiple spaceborne LiDAR, SAR, and optical missions in studies over larger area. The development of LiDAR techniques enabled the assessment of three dimensional structure of the tree cover to replicate the actual tree structure for modelling the biophysical parameters to close to real. With the currently available space borne LiDAR sensors namely the Advanced Topographic Laser Altimeter System (ATLAS) of ICESat – 2 and Global Ecosystem Dynamics Investigation (GEDI) of Japanese Experiment Module – Exposed Facility(JEM-EF) and still evolving LiDAR sensor technology (Multi-footprint Observation Lidar and Imager (MOLI) of JEM-EF to be launched in 2022) there are plenty of opportunities to overcome the uncertainties of radiometry based modelling of biophysical parameters such as volume or biomass. Field based instruments like Terrestrial laser scanner also reduce the laborious manual process involved in the measurements of parameters at the same time would help the researchers to upscale the biophysical characterisation of forest types from field based modelling approach to space based modelling (Coops et al, 2021).

6 Conclusions

The primary objectives of this study were to develop a pipeline for the estimation of above-ground biomass (AGB) from the terrestrial LiDAR point clouds of mangrove forests at tree level and plot level. These objectives were successfully met during this study by classifying foliage and woody points using SRF and WRF and construction of tree structure of woody points into tree trunk and pneumatophores using 3D polygonal shapes. The accuracy of classifying the point clouds of foliage and wood could be significantly improved using weighted RF algorithm in comparison with standard RF by assigning weights as balanced variants to overcome the error due to class imbalance. The mangroves of the plots having varying height and overlapping canopy have very noisy point clouds and segmentation of individual trees and QSM construction was challenging part of the study, however, we could get the near real tree distribution using visualization tools. The pipeline developed in this study achieved satisfactory performance on the benchmark dataset also used for validation in this study. The methodology proposed in this study could be generalised to predict the AGB of forest plots of any tree species. This study also proposes a method to estimate the contribution of pneumatophores to AGB, which are specially adapted aerial roots characteristic to mangrove forests that are generally neglected in AGB estimation using point cloud data. The proposed pipeline will be a pre-runner and would form a baseline ground truth information for the effective utilisation of the recent space borne laser data like GEDI and MOLI for upscaling the biophysical characterization of inaccessible and ecologically important mangrove ecosystem in a larger extent.

Author Contributions: GL and YKA Conceptualised the study; YKA and NP collected field data, carried out preliminary and formal data analysis; GL, DVK and DM devised the methodology and flow of Analysis; YKA and DM worked on programming; result analysis was performed by YKA and GL; YKA prepared the original draft and was edited by all authors and reviewed by GL and DVK. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Mangrove Foundation, Mumbai, India and supported by Indian Institute of Space Science and Technology, Trivandrum, India.

Acknowledgments: Authors would like acknowledge the support from Mangrove Foundation, Mumbai, India to carry out the TLS Survey in the mangroves of Maharashtra. Special thanks to Dr. Sruthi Krishna Moorthy Parvathi for her guidance via prompt mails in implementing the RF classification approach proposed in their paper [34].

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

Abbreviations

The following abbreviations are used in this manuscript:

AGB	Above-Ground Biomass
LiDAR	Light Detection and Ranging
UNFCCC	United Nations Framework Convention on Climate Change
ECV	Essential Climate Variable
COP	Committee of Parties
REDD	Reduction of Emissions due to Deforestation and Forest Degradation
SAR	Synthetic Aperture Radar
TLS	Terrestrial Laser Scanner/Scanning
DBSCAN	Density-Based Spatial Clustering of Applications with Noise
QSM	Quantitative Structure Model
BSG	Branch Structure Graph
SOR	Statistical Outlier Removal
ASCII	American Standard Code for Information Interchange
DBH	Diameter at Breast Height
tSNE	t-distributed Stochastic Neighbour Embedding
RF	Random Forest
LUCID	Land Use, Carbon and Emission Data
ROC AUC	Receiver Operator Characteristics - Area Under Curve
WRF	Weighted Random Forest
SRF	Standard Random Forest

References

1. Le Toan, T.; Quegan, S.; Davidson, M.W.; Balzter, H.; Paillou, P.; Papathanassiou, K.; Plummer, S.; Rocca, F.; Saatchi, S.; Shugart, H.; Ulander, L. The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle. *Remote Sensing of Environment* **2011**, *115*, 2850–2860. doi:10.1016/j.rse.2011.03.020.
2. Ciais, P.; Moore, B.; Steffen, W.; Hood, M.; Quegan, S.; Cihlar, J.; Raupach, M.; Tschirley, J.; Inoue, G.; Doney, S.; others. IGOS-P Integrated Global Carbon Observation Theme. *A Strategy to Realise a Coordinated System of Integrated Global Carbon Cycle Observations, Produced by IGBP (http://www. igbp. kve. se), 53pp* **2004**.
3. Franklin, S.E. *Remote sensing for sustainable forest management*; CRC press, 2001.
4. Bouvier, M.; Durrieu, S.; Fournier, R.A.; Renaud, J.P. Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data. *Remote Sensing of Environment* **2015**, *156*, 322–334. doi:10.1016/j.rse.2014.10.004.
5. Avery, T.E.; Burkhart, H.E. *Forest measurements*; Waveland Press, 2015.
6. Kangas, A.; Maltamo, M. *Forest inventory: methodology and applications*; Vol. 10, Springer Science & Business Media, 2006.
7. Le Toan, T.; Beaudoin, A.; Riom, J.; Guyon, D. Relating forest biomass to SAR data. *IEEE Transactions on Geoscience and Remote Sensing* **1992**, *30*, 403–411.
8. Leboeuf, A.; Fournier, R.A.; Luther, J.E.; Beaudoin, A.; Guindon, L. Forest attribute estimation of northeastern Canadian forests using QuickBird imagery and a shadow fraction method. *Forest Ecology and Management* **2012**, *266*, 66–74. doi:10.1016/j.foreco.2011.11.008.
9. Zolkos, S.G.; Goetz, S.J.; Dubayah, R. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sensing of Environment* **2013**, *128*, 289–298. doi:10.1016/j.rse.2012.10.017.
10. van Leeuwen, M.; Nieuwenhuis, M. Retrieval of forest structural parameters using LiDAR remote sensing. *European Journal of Forest Research* **2010**, *129*, 749–770. doi:10.1007/s10342-010-0381-4.
11. Næsset, E. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research* **2004**, *19*, 164–179.
12. Nelson, R.; Krabill, W.; Tonelli, J. Estimating forest biomass and volume using airborne laser data. *Remote Sensing of Environment* **1988**, *24*, 247–267. doi:10.1016/0034-4257(88)90028-4.
13. Kauffman, J.B.; Donato, D.C.; others. *Protocols for the measurement, monitoring and reporting of structure, biomass and carbon stocks in mangrove forests*; Citeseer, 2012.

14. Sedlacek, D.; Zara, J. Graph cut based point-cloud segmentation for polygonal reconstruction. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 2009, Vol. 5876 LNCS, pp. 218–227. doi: 10.1007/978-3-642-10520-3_20.
15. Niemeyer, J.; Rottensteiner, F.; Soergel, U. Classification of urban LiDAR data using conditional random field and random forests. Joint Urban Remote Sensing Event 2013. IEEE, 2013, pp. 139–142.
16. Qi, C.R.; Su, H.; Mo, K.; Guibas, L.J. PointNet: Deep learning on point sets for 3D classification and segmentation. Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, 2017, Vol. 2017-Janua, pp. 77–85, [1612.00593]. doi: 10.1109/CVPR.2017.16.
17. Qi, C.; Yi, L.; Su, H.; Guibas, L. PointNet++: Deep Hierarchical Feature Learning on. *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems* **2017**, pp. 5105–5114.
18. Li, Y.; Bu, R.; Sun, M.; Wu, W.; Di, X.; Chen, B. PointCNN: Convolution on X-transformed points. Advances in Neural Information Processing Systems, 2018, Vol. 2018-Decem, pp. 820–830, [arXiv:1801.07791v5].
19. Wang, Y.; Sun, Y.; Liu, Z.; Sarma, S.E.; Bronstein, M.M.; Solomon, J.M. Dynamic graph Cnn for learning on point clouds. *ACM Transactions on Graphics* **2019**, *38*, [arXiv:1801.07829v2]. doi:10.1145/3326362.
20. Pang, G.; Neumann, U. 3D point cloud object detection with multi-view convolutional neural network. *Proceedings - International Conference on Pattern Recognition* **2016**, *0*, 585–590. doi: 10.1109/ICPR.2016.7899697.
21. Xiaozhi, C.; Huimin, M.; Ji, W.; Bo, L.; Tian, X. Multi-View 3D Object Detection Network for Autonomous Driving | Spotlight 4-2B - YouTube. *ComputerVisionFoundation Videos* **2017**, pp. 1907–1915.
22. Jing, H.; You, S. Point cloud labeling using 3D Convolutional Neural Network. *Proceedings - International Conference on Pattern Recognition* **2016**, *0*, 2670–2675. doi: 10.1109/ICPR.2016.7900038.
23. Xu, Y.; Tuttas, S.; Hoegner, L.; Stilla, U. Voxel-based segmentation of 3D point clouds from construction sites using a probabilistic connectivity model. *Pattern Recognition Letters* **2018**, *102*, 67–74. doi:10.1016/j.patrec.2017.12.016.
24. Li, B. 3D fully convolutional network for vehicle detection in point cloud. *IEEE International Conference on Intelligent Robots and Systems* **2017**, *2017-September*, 1513–1518. doi: 10.1109/IROS.2017.8205955.
25. Béland, M.; Baldocchi, D.D.; Widlowski, J.L.; Fournier, R.A.; Verstraete, M.M. On seeing the wood from the leaves and the role of voxel size in determining leaf area distribution of forests with terrestrial LiDAR. *Agricultural and Forest Meteorology* **2014**, *184*, 82–97. doi: 10.1016/j.agrformet.2013.09.005.
26. Tao, S.; Guo, Q.; Xu, S.; Su, Y.; Li, Y.; Wu, F. A geometric method for wood-leaf separation using terrestrial and simulated lidar data. *Photogrammetric Engineering and Remote Sensing* **2015**, *81*, 767–776. doi:10.14358/PERS.81.10.767.
27. Ferrara, R.; Virdis, S.G.; Ventura, A.; Ghisu, T.; Duce, P.; Pellizzaro, G. An automated approach for wood-leaf separation from terrestrial LIDAR point clouds using the density based clustering algorithm DBSCAN. *Agricultural and Forest Meteorology* **2018**, *262*, 434–444. doi:10.1016/j.agrformet.2018.04.008.
28. Xue, D.; Cheng, Y.; Shi, X.; Fei, Y.; Wen, P. An Improved Random Forest Model Applied to Point Cloud Classification. IOP Conference Series: Materials Science and Engineering. IOP Publishing, 2020, Vol. 768, p. 072037.
29. Zhu, X.; Skidmore, A.K.; Darvishzadeh, R.; Niemann, K.O.; Liu, J.; Shi, Y.; Wang, T. Foliar and woody materials discriminated using terrestrial LiDAR in a mixed natural forest. *International journal of applied earth observation and geoinformation* **2018**, *64*, 43–50.
30. Ma, Z.; Pang, Y.; Li, Z.; Lu, H.; Liu, L.; Chen, B. Fine classification of near-ground point cloud based on terrestrial laser scanning and detection of forest fallen wood. *J. Remote Sens* **2019**, *23*, 743–755.
31. Li, W.; Niu, Z.; Shang, R.; Qin, Y.; Wang, L.; Chen, H. High-resolution mapping of forest canopy height using machine learning by coupling ICESat-2 LiDAR with Sentinel-1, Sentinel-2 and Landsat-8 data. *International Journal of Applied Earth Observation and Geoinformation* **2020**, *92*, 102163.

32. Zeybek, M. Classification of UAV point clouds by random forest machine learning algorithm. *Turkish Journal of Engineering* **2021**, *5*, 48–57.
33. Vicari, M.B.; Disney, M.; Wilkes, P.; Burt, A.; Calders, K.; Woodgate, W. Leaf and wood classification framework for terrestrial LiDAR point clouds. *Methods in Ecology and Evolution* **2019**, *10*, 680–694. doi:10.1111/2041-210X.13144.
34. Krishna Moorthy, S.M.; Calders, K.; Vicari, M.B.; Verbeeck, H. Improved Supervised Learning-Based Approach for Leaf and Wood Classification From LiDAR Point Clouds of Forests. *IEEE Transactions on Geoscience and Remote Sensing* **2019**, *58*, 3057–3070. doi: 10.1109/tgrs.2019.2947198.
35. Wang, D.; Momo Takoudjou, S.; Casella, E. LeWoS: A universal leaf-wood classification method to facilitate the 3D modelling of large tropical trees using terrestrial LiDAR. *Methods in Ecology and Evolution* **2020**, *11*, 376–389. doi:10.1111/2041-210X.13342.
36. Li, W.; Guo, Q.; Jakubowski, M.K.; Kelly, M. A new method for segmenting individual trees from the lidar point cloud. *Photogrammetric Engineering and Remote Sensing* **2012**, *78*, 75–84. doi:10.14358/PERS.78.1.75.
37. Dalponte, M.; Coomes, D.A. Tree-centric mapping of forest carbon density from airborne laser scanning and hyperspectral data. *Methods in Ecology and Evolution* **2016**, *7*, 1236–1245. doi:10.1111/2041-210X.12575.
38. Silva, C.A.; Hudak, A.T.; Vierling, L.A.; Loudermilk, E.L.; O'Brien, J.J.; Hiers, J.K.; Jack, S.B.; Gonzalez-Benecke, C.; Lee, H.; Falkowski, M.J.; Khosravipour, A. Imputation of Individual Longleaf Pine (Pinus palustris Mill.) Tree Attributes from Field and LiDAR Data. *Canadian Journal of Remote Sensing* **2016**, *42*, 554–573. doi:10.1080/07038992.2016.1196582.
39. Yang, B.; Dai, W.; Dong, Z.; Liu, Y. Automatic forest mapping at individual tree levels from terrestrial laser scanning point clouds with a hierarchical minimum cut method. *Remote Sensing* **2016**, *8*. doi:10.3390/rs8050372.
40. Zhang, W.; Wan, P.; Wang, T.; Cai, S.; Chen, Y.; Jin, X.; Yan, G. A novel approach for the detection of standing tree stems from plot-level terrestrial laser scanning data. *Remote Sensing* **2019**, *11*, 1–20. doi:10.3390/rs11020211.
41. Boudon, F.; Preuksakarn, C.; Ferraro, P.; Diener, J.; Nacry, P.; Nikinmaa, E.; Godin, C. Quantitative assessment of automatic reconstructions of branching systems obtained from laser scanning. *Annals of Botany* **2014**, *114*, 853–862. doi:10.1093/aob/mcu062.
42. Landes, T.; Saudreau, M.; Najjar, G.; Kastendeuch, P.; Guillemin, S.; Colin, J.; Luhahe, R. 3D tree architecture modeling from laser scanning for urban microclimate study. 9th International Conference on Urban Climate, 12th Symposium on the Urban Environment, 2015, number November, p. 6.
43. Cao, J.; Tagliasacchiy, A.; Olsony, M.; Zhangy, H.; Su, Z. Point cloud skeletons via Laplacian-based contraction. SMI 2010 - International Conference on Shape Modeling and Applications, Proceedings, 2010, pp. 187–197. doi:10.1109/SMI.2010.25.
44. Hackenberg, J.; Spiecker, H.; Calders, K.; Disney, M.; Raumonen, P. SimpleTree - An efficient open source tool to build tree models from TLS clouds. *Forests* **2015**, *6*, 4245–4294. doi: 10.3390/f6114245.
45. Jagtap, T.G.; Untawale, A.G.; Inamdar, S.N. Study of mangrove environment of Maharashtra coast using remote sensing data. *Indian Journal of Marine Sciences* **1994**, *23*, 90–93. doi: http://nopr.niscair.res.in/handle/123456789/37548.
46. Kulkarni, V.A.; Jagtap, T.G.; Mhalsekar, N.M.; Naik, A.N. Biological and environmental characteristics of mangrove habitats from Manori creek, West Coast, India. *Environmental monitoring and assessment* **2010**, *168*, 587–596.
47. Mugade, N.; Sapkale, J. A Review of Mangrove Conservation Studies in Maharashtra, India. *International Journal of Engineering and Technical Research (IJETR) ,ER Publications* **2014**, *ISSN 321-0869*, 2321–869.
48. GPL Software. CloudCompare (version 2.11).
49. Zhang, W.; Qi, J.; Wan, P.; Wang, H.; Xie, D.; Wang, X.; Yan, G. An easy-to-use airborne LiDAR data filtering method based on cloth simulation. *Remote Sensing* **2016**, *8*. doi: 10.3390/rs8060501.
50. Tran, P.; Gritcan, I.; Cusens, J.; Alfaro, A.C.; Leuzinger, S. Biomass and nutrient composition of temperate mangroves (Avicennia marina var. australasica) in New Zealand. *New Zealand Journal of Marine and Freshwater Research* **2017**, *51*, 427–442. doi:10.1080/00288330.2016.1260604.
51. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher,

M.; Perrot, M.; Duchesnay, E. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research* **2011**, *12*, 2825–2830.
52. Krebs M., Alexandre P.. Computree (version 4.0).
53. Fajardo, A. Insights into intraspecific wood density variation and its relationship to growth, height and elevation in a treeline species. *Plant Biology* **2018**, *20*, 456–464.
54. Gonzalez de Tanago, J.; Lau, A.; Bartholomeus, H.; Herold, M.; Avitabile, V.; Raumonen, P.; Martius, C.; Goodman, R.C.; Disney, M.; Manuri, S.; Burt, A.; Calders, K. Estimation of above-ground biomass of large tropical trees with terrestrial LiDAR. *Methods in Ecology and Evolution* **2018**, *9*, 223–234. doi:10.1111/2041-210X.12904.
55. Lumia, R.; Shapiro, L.; Zuniga, O. A new connected components algorithm for virtual memory computers. *Computer Vision, Graphics and Image Processing* **1983**, *22*, 287–300. doi:10.1016/0734-189X(83)90071-3.
56. Komiyama, A.; Poungparn, S.; Kato, S. Common allometric equations for estimating the tree weight of mangroves. *Journal of tropical ecology* **2005**, *21*, 471–477.
57. Han, S.H.; Park, B.B. Comparison of Allometric Equation and Destructive Measurement of Carbon Storage of Naturally Regenerated Understory in a Pinus rigida Plantation in South Korea. *Forests* **2020**, *11*. doi:10.3390/f11040425.
58. Li, Y.; Li, M.; Li, C.; Liu, Z. Forest aboveground biomass estimation using Landsat 8 and Sentinel-1A data with machine learning algorithms. *Scientific reports* **2020**, *10*, 1–12.
59. Feliciano, E.A.; Wdowinski, S.; Potts, M.D. Assessing mangrove above-ground biomass and structure using terrestrial laser scanning: A case study in the Everglades National Park. *Wetlands* **2014**, *34*, 955–968.
60. Balduzzi, M.A.; Van der Zande, D.; Stuckens, J.; Verstraeten, W.W.; Coppin, P. The properties of terrestrial laser system intensity for measuring leaf geometries: a case study with conference pear trees (Pyrus Communis). *Sensors* **2011**, *11*, 1657–1681.
61. Rouzbeh Kargar, A.; MacKenzie, R.A.; Apwong, M.; Hughes, E.; van Aardt, J. Stem and root assessment in mangrove forests using a low-cost, rapid-scan terrestrial laser scanner. *Wetlands Ecology and Management* **2020**, *28*, 883–900.
62. Satyanarayana, B.; Thierry, B.; Seen, D.L.; Raman, A.; Muthusankar, G. Remote sensing in mangrove research–relationship between vegetation indices and dendrometric parameters: A Case for Coringa, East Coast of India. 22nd Asian conference on remote sensing, 2001, pp. 5–9.
63. Vaghela, B.; Chirakkal, S.; Putrevu, D.; Solanki, H. Modelling above ground biomass of Indian mangrove forest using dual-pol SAR data. *Remote Sensing Applications: Society and Environment* **2021**, *21*, 100457. doi:https://doi.org/10.1016/j.rsase.2020.100457.
64. Anand, A.; Pandey, P.C.; Petropoulos, G.P.; Pavlides, A.; Srivastava, P.K.; Sharma, J.K.; Malhi, R.K.M. Use of hyperion for mangrove forest carbon stock assessment in Bhitarkanika forest reserve: A contribution towards blue carbon initiative. *Remote Sensing* **2020**, *12*, 597.
65. Prasad KA, G.L. Estimation of Above Ground Biomass using High Resolution Multispectral Worldview 2 image. *Indian Cartographer* **2018**, *38*, 569–579.
66. George, R.; Padalia, H.; Sinha, S.; Kumar, A.S. Evaluation of the use of hyperspectral vegetation indices for estimating mangrove leaf area index in middle Andaman Island, India. *Remote Sensing Letters* **2018**, *9*, 1099–1108.
67. Almeida, A.; Gonçalves, F.; Silva, G.; Souza, R.; Treuhaft, R.; Santos, W.; Loureiro, D.; Fernandes, M. Estimating Structure and Biomass of a Secondary Atlantic Forest in Brazil Using Fourier Transforms of Vertical Profiles Derived from UAV Photogrammetry Point Clouds. *Remote Sensing* **2020**, *12*.
68. Levick, S.R.; Whiteside, T.; Loewensteiner, D.A.; Rudge, M.; Bartolo, R. Leveraging TLS as a calibration and validation tool for MLS and ULS mapping of savanna structure and biomass at landscape-scales. *Remote Sensing* **2021**, *13*, 257.