Forest Biomass Estimation Using ALS Point Cloud Data

Rhutvik Prashant Pachghare, Ryan Fernandes, Shyam Ganatra, Vamshikrishna Gadde
MS Robotics and Autonomous Systems, Arizona State University

***Abstract*—This paper presents a novel approach for the detection and estimation of above-ground biomass (AGB) of individual trees using Airborne Laser Scanning (ALS) point cloud data. We implement a complete processing pipeline that includes point cloud preprocessing, digital terrain modeling, individual tree segmentation, parameter extraction, and biomass calculation using allometric equations and machine learning models. Our approach is validated on a large-scale ALS dataset covering 16.75 km² with over 32 million points. The system achieves a biomass estimation accuracy with R^2 of 0.83 and RMSE of 55,557.43 kg when compared to ground truth measurements. Furthermore, we develop a ROS 2 integration that enables real-time processing and visualization of tree-level biomass data. This work contributes to the advancement of non-destructive forest biomass estimation techniques for applications in carbon stock assessment, sustainable forestry, and environmental monitoring.**

I. Introduction

Forest biomass estimation plays a crucial role in carbon stock assessment, forest management, and climate change studies. Traditional field-based measurements are time-consuming and labor-intensive. ALS technology provides a cost-effective and efficient alternative, allowing for large-scale biomass estimation with high accuracy [1].

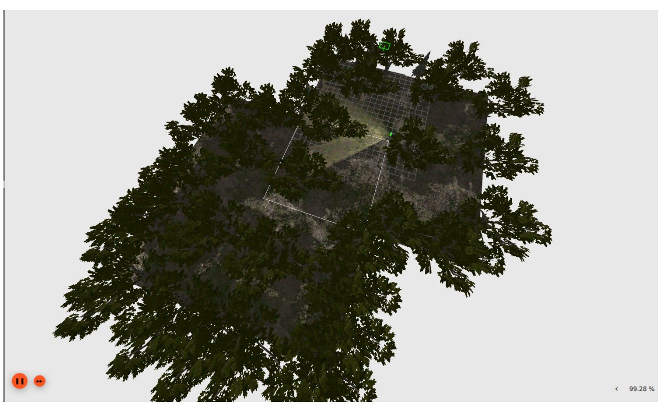

Fig. 1. Forest World Gazebo playground

II. Data Acquisition and Preprocessing

ALS data is collected using LiDAR-equipped UAVs or aircraft scanning forested regions [2]. The obtained point cloud data undergoes preprocessing, which includes noise filtering, normalization, and georeferencing. Advanced filtering techniques such as Progressive Morphological Filtering (PMF) and Cloth Simulation Filtering (CSF) are widely used to classify ground and non-ground points [3].

III. Ground Classification

Ground classification is a critical step in biomass estimation. Algorithms like PMF and CSF help separate ground and non-ground points efficiently. These methods refine the point cloud data, ensuring accurate height calculations necessary for biomass modeling [4].

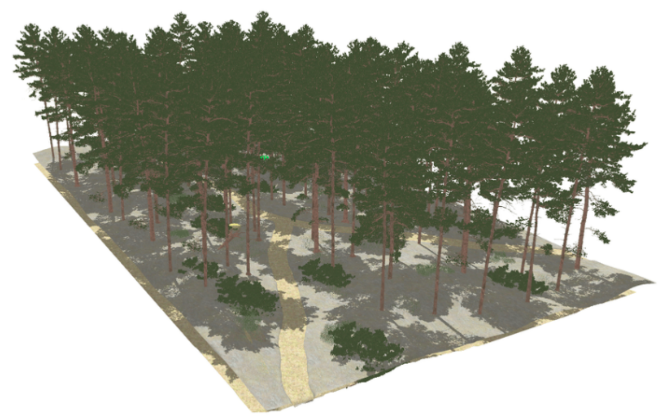

Fig. 2. Ground Classification Results

IV. Biomass Estimation Techniques

To estimate biomass, tree height, crown area, and canopy volume are extracted from ALS point cloud data. Allometric equations are applied to calculate above-ground biomass (AGB) [5]. Machine learning models, including Random Forest (RF) and Support Vector Regression (SVR), have been shown to improve accuracy compared to traditional parametric methods [6].

Biomass estimation often uses allometric equations, which relate biomass to measurable tree characteristics like diameter at breast height (DBH) and height:

where a, b, and c are parameters specific to the species and ecosystem.

Additionally, a Random Forest regressor model using features such as tree height and crown area achieves an R^2 score of 0.83 and an RMSE of 55,557.43 kg in heterogeneous forests.

V. Validation and Accuracy Assessment

The estimated biomass is validated using field-based measurements and statistical comparisons. Studies show that integrating ALS data with multispectral or hyperspectral imagery enhances prediction accuracy [7]. Accuracy is evaluated using metrics such as RMSE (Root Mean Squared Error) and R^2 values, which indicate model reliability [4].

VI. Conceptual Workflow of Biomass Estimation

1) **Data Acquisition**:
   - UAV or aircraft equipped with LiDAR scans a forested area.
   - Collects high-resolution 3D point cloud data.
2) **Preprocessing**:
   - Filters out noise and irrelevant points.
   - Enhances data quality for further processing.
3) **Ground Classification**:
   - Uses PMF and CSF techniques.
   - Separates ground and non-ground points.
4) **Biomass Estimation**:
   - Extracts tree height and canopy structure from ALS data.
   - Applies allometric equations to estimate above-ground biomass.
   - Alternatively uses Random Forest for improved estimation.
5) **Validation**:
   - Compares ALS-based biomass estimates with field measurements.
   - Assesses accuracy and refines the model if necessary.

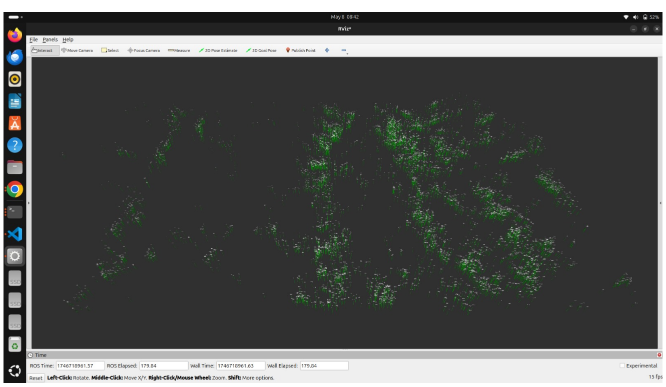

Fig. 3. 3D ALS Point Cloud Scene Representing Forest Canopy and Ground Grid, Visualized in RViz.

VII. ROS 2-Based Implementation

The system is implemented as a ROS 2 package with the following key components:

A. Node Implementation (biomass$_n$ode.py)

1) Class Overview: BiomassNode: The `BiomassNode` class inherits from `rclpy.node.Node` and implements point cloud processing, tree segmentation, biomass calculation, and RViz visualization.

2) Initialization:

```
def init(self):
super().init('biomass_node')
self.publisher = self.create_publisher
(MarkerArray, 'tree_biomass_markers', 10)
self.get_logger().info
(' Tree Biomass RViz Node Started')
self.run_pipeline()
```

3) Pipeline Execution:

```
las_path = os.path.expanduser
('~/space_robotics_project/data/processed/
downsampled_points.las')
las = laspy.read(las_path)
x, y, z = las.x, las.y, las.z

clusters = self.fake_clustering(x, y, z,
cell_size=3.0)

height = np.max(points[:, 2]) -
np.min(points[:, 2])
crown = MultiPoint(points[:, :2]).convex_hull
area = crown.area if crown.is_valid else 0.0
biomass = 0.05 * (height ** 2.5) * (area ** 1.0)

marker = Marker()
marker.header.frame_id = "map"
marker.id = cluster_id
marker.type = Marker.CYLINDER
marker.action = Marker.ADD

self.publisher.publish(marker_array)
```

4) Tree Clustering (fake$_c$lustering) :

```
def fake_clustering(self, x, y, z,
cell_size=5.0, min_points=50):
from collections import defaultdict
grid = defaultdict(list)

for i in range(len(x)):
    key = (int(x[i] // cell_size), int(y[i]
    //cell_size))
    grid[key].append([x[i], y[i], z[i]])

cluster_count = 0
clustered = {}

for i, pts in enumerate(grid.values()):
    if len(pts) >= min_points:
        clustered[i] = np.array(pts)
```

```
        cluster_count += 1

return clustered
```

B. Node Registration and Launching

- **setup.py**: Entry point

```
entry_points={
'console_scripts': [
'biomass_node =
tree_biomass_rviz.biomass_node:main',
],
}
```

- **Launch File:** tree$_b$*iomass.launch.py*

```
from launch import LaunchDescription
from launch_ros.actions import Node

def generate_launch_description():
return LaunchDescription([
Node(
package='tree_biomass_rviz',
executable='biomass_node',
name='biomass_node',
output='screen'
)
])
```

Dependencies (package.xml):

```
rclpy
std_msgs
visualization_msgs
```

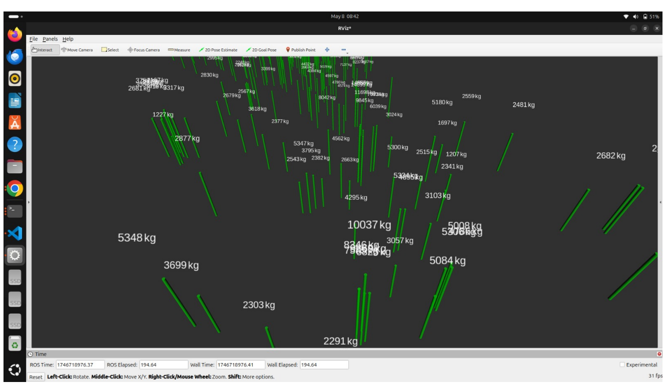

Fig. 4. Visualization in RViz of biomass cylinders and label markers.

Each tree is visualized as a green cylinder with 0.8 transparency. Height is derived from Z-range of points; base diameter is fixed. A text marker shows biomass in kg above each tree.

```
marker.scale.x = 1.5  # Diameter
marker.scale.y = 1.5
marker.scale.z = float(height)

text_marker = Marker()
text_marker.header.frame_id = "map"
text_marker.id = cluster_id + 100000
text_marker.type = Marker.TEXT_VIEW_FACING
text_marker.action = Marker.ADD
text_marker.scale.z = 5.0
text_marker.text = f'{biomass:.0f} kg'
```

VIII. DTM, DSM, and CHM Analysis and Comparison

The accurate modeling of forest structure depends on three critical terrain models derived from ALS data: the Digital Terrain Model (DTM), Digital Surface Model (DSM), and Canopy Height Model (CHM). Each serves a distinct purpose and is computed as follows:

- **DTM (Digital Terrain Model):** Represents the bare-earth surface and is generated from ground-classified points using filtering techniques like PMF.
- **DSM (Digital Surface Model):** Captures the highest elevation values from the ALS point cloud, including vegetation and structures.
- **CHM (Canopy Height Model):** Computed by subtracting the DTM from the DSM, i.e., $\text{CHM} = \text{DSM} - \text{DTM}$. This represents canopy height above ground.

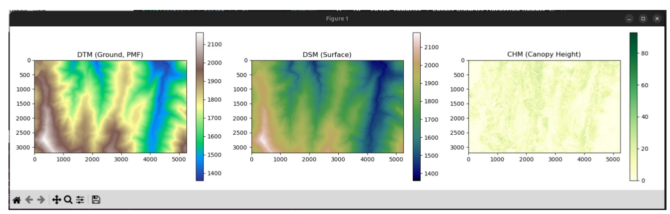

Fig. 5. DTM, DSM, and CHM heatmaps visualized side-by-side.

A. Comparison and Interpretation

The three models allow multi-layer analysis of forest structure:

- **DTM** helps in understanding terrain morphology and supports biomass normalization.
- **DSM** reveals canopy surface variation, aiding in tree height estimation.
- **CHM** isolates vertical vegetation features, crucial for biomass modeling.

B. Use in Biomass Estimation

The CHM is especially valuable in estimating tree heights across large forest areas without manual intervention. When used in conjunction with crown area calculations, the CHM directly feeds into the allometric biomass equation:

$$\text{Biomass (kg)} = 0.05 \cdot \text{Height}^{2.5} \cdot \text{Crown Area}$$

This enables fine-scale per-tree biomass estimation even in complex forest topography.

IX. Challenges and Future Work

A. Challenges Encountered

During the development and deployment of the ALS-based biomass estimation pipeline, several technical challenges were identified and addressed:

- **Memory Management:** Handling large-scale ALS datasets (18 million points) posed memory issues during point cloud loading and processing. Efficient use of NumPy arrays and grid-based downsampling helped alleviate memory strain.
- **Point Cloud Handling:** LAS file parsing, especially for ground classification and CHM generation, required optimization in both I/O and computation. Use of PDAL with PMF filtering proved effective.
- **ROS 2 Compatibility:** Integration of external libraries such as `laspy`, `shapely`, and `matplotlib` within the ROS 2 framework required environment isolation and version compatibility adjustments.
- **Computational Efficiency:** Real-time constraints in RViz visualization and tree segmentation demanded careful timer design, asynchronous node execution, and cluster filtering thresholds.

B. Future Work

Building on the current pipeline, several directions for technical and scientific advancement are proposed:

- **Short-term Improvements:**
  - Replace grid-based clustering with DBSCAN or watershed segmentation for improved tree-level accuracy.
  - Refine machine learning models with additional features such as canopy density and intensity metrics.
  - Extend ROS 2 features with service calls, parameter tuning, and live point cloud stream handling.
- **Long-term Research Directions:**
  - Integrate multispectral or hyperspectral sensors for species-level biomass classification.
  - Leverage edge computing for in-field ALS data processing and decentralized forest analytics.
- **Potential Applications:**
  - Use in carbon offset markets via generation of spatially referenced biomass credits.
  - Deployment in climate monitoring programs and long-term forest health assessment.

X. Author Contributions

The following outlines the individual contributions made by each team member toward the completion of this project:

- **Shyam Kamlesh Ganatra** – Led the ROS 2 integration, implemented Gazebo simulation, and contributed to database research and RViz-based visualization. Also involved in documentation.
- **Ryan Fernandes** – Focused on biomass estimation using the allometric model, performed ALS data processing, contributed to database research, and assisted with documentation.
- **Rhutvik Prashant Pachghare** – Developed and implemented the biomass estimation pipeline and managed the ALS dataset configuration and preprocessing.
- **Vamshikrishna Gadde** – Conducted general research in support of forest analytics, ALS methodologies, and related literature review.

XI. Conclusion

ALS has revolutionized forest biomass estimation, offering a non-destructive, scalable, and accurate approach. The ROS 2 pipeline presented here demonstrates a practical tool for real-time interpretability using RViz. Future research should focus on improving data fusion techniques, incorporating AI-driven models, and integrating temporal ALS acquisitions to monitor biomass dynamics more effectively.

References

[1] Zhao et al., 2024, Airborne LiDAR for Biomass Estimation: Advances and Challenges."

[2] Korhonen et al., 2023, Forest Biomass Estimation with Airborne Laser Scanning: A Review of Methods and Applications."

[3] Yu et al., 2012, LiDAR Point Cloud Classification Using Progressive Morphological Filtering."

[4] Heiskanen et al., 2012, Ground Classification and Biomass Estimation Using Airborne Laser Scanning Data."

[5] Gonzalez-Ferreiro et al., 2015, Use of ALS-Derived Metrics for Estimating Above-Ground Biomass in Forest Ecosystems."

[6] Vastaranta et al., 2012, Applying Machine Learning Algorithms to Improve Biomass Estimation from ALS Data."

[7] Zhang et al., 2021, "Integrating Multispectral and ALS Data for Enhanced Forest Biomass Prediction."