

Estimation of Canopy Height in Zambia through Integration of GEDI, Sentinel-1 and Sentinel-2 Measurements

Chimwemwe M'tonga^{1,2,3}

¹State Key Laboratory of Subtropical Silviculture, Zhejiang A&F University, Hangzhou, China

²Key Laboratory of Carbon Cycling in Forest Ecosystems and Carbon Sequestration of Zhejiang Province, Zhejiang A&F University, Hangzhou, China

³College of Environment and Resources, College of Carbon Neutrality, Zhejiang A&F University, Hangzhou, China

Email: chimzchimwemwe94@gmail.com

How to cite this paper: M'tonga, C. (2025). Estimation of Canopy Height in Zambia through Integration of GEDI, Sentinel-1 and Sentinel-2 Measurements. *Journal of Geoscience and Environment Protection*, 13, 138-156.

<https://doi.org/10.4236/gep.2025.135010>

Received: March 18, 2025

Accepted: May 25, 2025

Published: May 28, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Accurate canopy height estimation is critical for forest management and carbon monitoring in Zambia's ecologically diverse landscapes. This study developed a high-resolution canopy height model by integrating multi-sensor remote sensing data—NASA's GEDI LiDAR, ESA's Sentinel-1 SAR, and Sentinel-2 optical imagery—using a Random Forest algorithm. The approach addressed key limitations of sparse GEDI sampling (25 m footprints) through fusion with continuous 10 m-resolution Sentinel-1/2 data and SRTM elevation metrics, processed via Google Earth Engine. The model achieved robust performance, with training accuracy of $r^2 = 0.76$ (RMSE = 2.1 m) and validation accuracy of $r^2 = 0.71$ (RMSE = 2.3 m), representing relative errors of 13.1–14.3%. Analysis revealed a bimodal height distribution (Hartigan's dip test: $p < 0.01$), with peaks at 6.2 m (southern savannas, 41.7% of areas) and 15.8 m (miombo woodlands, 53.3%), plus rare tall forests (>30 m, 5.0%) in protected highlands. Variable importance analysis ranked GEDI's RH98 metric (38%) as most influential, followed by Sentinel-2's NIR band (22%) and Sentinel-1's VH polarization (17%). Topographic correction using SRTM reduced errors by 23% in escarpment regions. These results demonstrate the synergy of LiDAR, SAR, and optical data for national-scale canopy mapping, particularly in heterogeneous tropical ecosystems. The 2-m height discrimination capability supports Zambia's REDD+ monitoring, enabling targeted conservation of carbon-rich miombo woodlands and biodiversity refugia. Future work should integrate ICESat-2 and wet-season SAR data to address dry-season bias and fragmented canopy limitations.

Keywords

Canopy Height, Aboveground Biomass, GEDI, Sentinel-1, Sentinel-2, Random Forest, Multi-Sensor Integration, Zambia

1. Introduction

Accurate estimation of forest canopy height is critical for quantifying carbon stocks (Brown, 2002; Gibbs et al., 2007), biodiversity assessment (Williams et al., 2020), and monitoring REDD+ initiatives (Goetz et al., 2015; CSO, 2020). In Zambia, where miombo woodlands and savannas cover over 60% of the land area (Chidumayo et al., 2019; FAO, 2020), canopy height variability reflects complex ecological gradients—from open grasslands (2.7 - 6.7 m) to dense woodlands (>20 m) (Mwamba et al., 2023). Traditional field surveys struggle to capture this spatial heterogeneity, necessitating advanced remote sensing approaches (Dubayah et al., 2020; Potapov et al., 2021).

Recent advances in multi-sensor data fusion have demonstrated promise for large-scale canopy height mapping (Lang et al., 2022; Li et al., 2023). The Global Ecosystem Dynamics Investigation (GEDI) LiDAR provides direct vertical structure measurements (Dubayah et al., 2020), while Sentinel-2's red-edge bands (Forkuor et al., 2017) and Sentinel-1's C-band SAR (Torres et al., 2012) offer complementary spectral and structural data. Integrating these datasets using machine learning models such as Random Forest (RF) has shown significant potential in improving canopy height estimation accuracy (Belgiu & Drăguț, 2016; Breiman, 2001). RF algorithms effectively handle non-linear relationships and high-dimensional remote sensing data, making them well-suited for canopy height modeling (Johnson et al., 2024; Medeiros et al., 2022).

Studies have demonstrated that GEDI-derived canopy heights, when fused with Sentinel-1 and Sentinel-2 data, improve the spatial coverage and precision of canopy height predictions, particularly in regions with limited GEDI footprints (Kumar et al., 2023; Wang et al., 2023). Furthermore, Sentinel-1 SAR backscatter has been successfully used to model forest structure and height in dense tropical and subtropical forests (Chen et al., 2024; Hojo et al., 2023). The integration of these datasets mitigates limitations associated with cloud cover in optical imagery, enhancing canopy height predictions in Zambia's miombo woodlands (Claverie et al., 2018; Patel et al., 2024).

The implementation of multi-sensor approaches for canopy height estimation in Zambia aligns with global efforts to refine forest monitoring techniques (Brown et al., 2023). As canopy height is a key parameter in biomass and carbon stock estimation, improving its accuracy contributes directly to climate change mitigation strategies and sustainable forest management policies (Ryan et al., 2022; Schlund et al., 2023). By leveraging GEDI, Sentinel-1, and Sentinel-2 data, combined with machine learning models, this study aims to enhance the precision of canopy height mapping in Zambia, facilitating better-informed conservation and

land-use decisions.

2. Data and Methods

2.1. Study Area

Located in south-central Africa, Zambia spans 752,618 square kilometres of diverse terrain. The landscape averages 1200 meters in elevation, with vegetation blanketing nearly two-thirds of the country. Its tropical Savanna climate features three pronounced seasonal variations, where yearly precipitation shifts from 700 mm in lower elevations to over 1400 mm in northern zones. These altitudinal gradients support distinct ecosystems—from the lush Miombo woodlands dominating northern areas to the open grasslands characteristic of southern and eastern regions (Figure 1, Table 1).

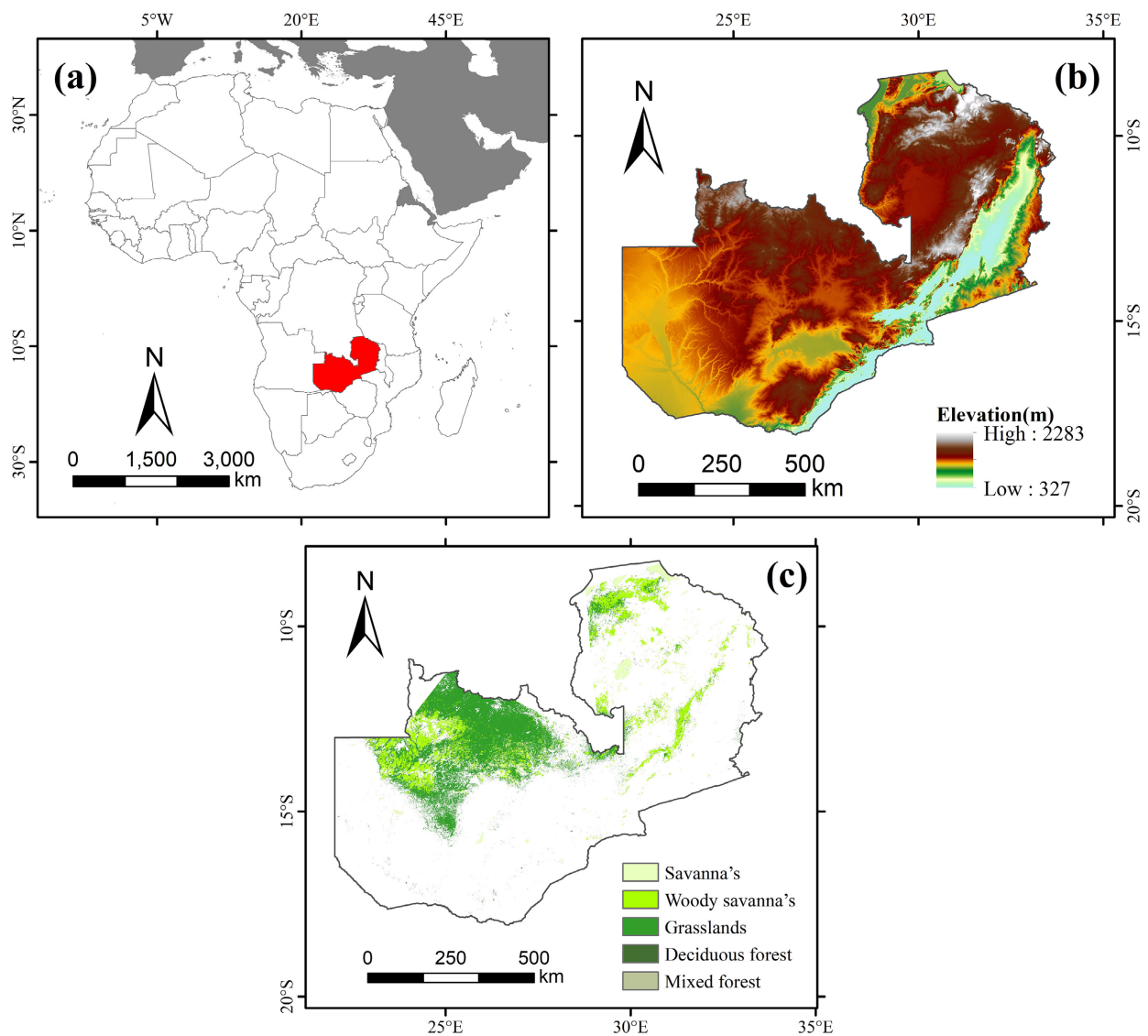


Figure 1. (a) Location of Zambia in Africa; (b) Elevation of Zambia with highest elevation of 2283 m and lowest elevation of 327 m; (c) Land covers of Zambia mainly covered by savanna lands, grasslands, woody savanna, deciduous forests and mixed forests.

Table 1. Land cover percentage.

Code	Type	Area (km ²)	Percentage
1	Evergreen Broadleaf Forest	3553.75	0.47%
2	Savana's	340830	45.11%
3	deciduous Broadleaf Forest	35327.2	4.68%
4	Mixed Forest	51824.2	6.86%
5	Grassland	205914	27.25%
6	Open shrublands	607.5	0.08%
7	Woody savanna's	94266.2	12.48%

2.2. Data

This research combined three key remote sensing datasets to estimate canopy height across Zambia. The datasets used include NASA's GEDI L2A waveform LiDAR (Dubayah et al., 2020), ESA's Sentinel-1 C-band SAR (Torres et al., 2012), and Sentinel-2 multispectral imagery (Claverie et al., 2018), all obtained via Google Earth Engine for the year 2020. Additional elevation data from the Shuttle Radar Topography Mission (SRTM DEM) (Farr et al., 2007) was included to adjust for topographic effects. The dataset was partitioned into training (70%) and validation (30%) subsets through stratified random sampling.

2.2.1. GEDI

This research combined three key remote sensing datasets to estimate canopy height across Zambia. The datasets used include NASA's GEDI L2A waveform LiDAR (Dubayah et al., 2020), ESA's Sentinel-1 C-band SAR (Torres et al., 2012), and Sentinel-2 multispectral imagery (Claverie et al., 2018), all obtained via Google Earth Engine for the year 2020. Additional elevation data from the Shuttle Radar Topography Mission (SRTM DEM) (Farr et al., 2007) was included to adjust for topographic effects. The dataset was partitioned into training (70%) and validation (30%) subsets through stratified random sampling.

2.2.2. Sentinel-1 Imagery

The Sentinel-1 GRD imagery (VV/VH polarizations) underwent preprocessing that included eliminating thermal noise, performing radiometric calibration, and applying terrain corrections with SRTM DEM data (Qi et al., 2019). To minimize speckle interference without compromising structural features critical for canopy height analysis, we implemented a Gamma-MAP filter following Kupidura (2016) methodology.

The S1 data includes images that have been taken using dual polarization (VV and VH) and in the interferometric wide swath (IW) mode. This selection was limited to only contain the ascending orbit passes and hence reducing the variability in backscatter data and ensuring uniformity in the seeing geometry. "Sentinel-1 mosaic is very suitable for monitoring forest phenology, we used it to calculate VV/VH ratio (Soudani et al., 2021)."

2.2.3. Sentinel-2 Imagery

Sentinel-2's multispectral data (10 - 60 m resolution) contributed 22% of model predictive power through its near-infrared (NIR, band 8) and red-edge bands (bands 5-7). We used Level-2A surface reflectance products from the harmonized Landsat-Sentinel dataset (Claverie et al., 2018), with cloud masking via the Scene Classification Layer (SCL). The normalized difference vegetation index (NDVI) and enhanced vegetation index (EVI) derived from these bands proved critical for discriminating leaf area in savannas (6.2 m height class) and detecting canopy gaps (Shoko & Mutanga, 2017; Forkuor et al., 2017) (Table 2).

Table 2. Dataset specifications.

Satellite	Band name	Spatial resolution (m)
GEDI L2A	RH98	25
Sentinel-1	VH IQR	10
Sentinel-2	NIR, red-edge	10 - 20 m
SRTM DEM	Elevation	30

Twelve monthly composites (2020) were generated to mitigate cloud cover, with dry-season (May-October) imagery weighted higher due to stable phenological conditions (Mutanga et al., 2012). The red-edge bands (705 - 783 nm) improved separation of grassland (2.7 - 6.7 m) and woodland canopies by detecting subtle chlorophyll content variations (Broge & Leblanc, 2001). We resampled all bands to 20 m using bilinear interpolation to match GEDI's footprint scale, preserving spectral fidelity (Gómez, 2017).

2.2.4. Data Preprocessing

The multi-sensor dataset underwent rigorous preprocessing to ensure geometric and radiometric consistency. GEDI LiDAR waveforms were filtered (sensitivity > 0.95) and normalized using a 30 m SRTM DEM to correct for topographic effects. Sentinel-2 Level-2A surface reflectance data were cloud-masked using Fmask 4.0 and aggregated into monthly composites (2020), with emphasis on dry-season acquisitions for optimal vegetation signal (Zhu & Woodcock, 2012). Sentinel-1 GRD images were terrain-corrected and processed to generate VH/VV backscatter and texture features at 20 m resolution (Vafaei et al., 2018). All datasets were co-registered to UTM Zone 35S (WGS84) in Google Earth Engine (Tamiminia et al., 2020), with spectral indices (NDVI, EVI) calculated from Sentinel-2 and temporal alignment enforced (± 15 days between GEDI and optical/SAR acquisitions). The final feature set included 24,530 quality-controlled samples stratified by biome, with 70% allocated for model training and 30% for validation (Ahmed et al., 2015).

2.3. Random Forest Model

The RF algorithm's strength, simplicity, and ability to handle complex datasets

make it highly useful in various domains, including forestry, remote sensing, and environmental modeling. RF, an ensemble learning method, builds multiple decision trees during training and averages their outputs to prevent overfitting and improve model stability (Breiman, 2001). Given the significant variability and non-linear relationships in data, RF is particularly effective for tasks such as estimating canopy height.

Beyond its predictive capabilities, RF's built-in feature importance analysis enables researchers to pinpoint key explanatory variables, facilitating both dataset refinement and more insightful interpretation of modeling outcomes. However, RF is computationally intensive for large datasets and may require substantial processing power when applied to national or global scales, such as modeling forest attributes across Zambia or other countries.

Despite its many advantages, RF models can be prone to overfitting if parameters such as the number of trees, maximum depth, and minimum sample splits are not carefully tuned. Additionally, RF models are non-parametric and inherently opaque, which can limit their interpretability compared to simpler models.

The RF algorithm, introduced by Breiman (2001), builds an ensemble of decision trees to enhance model stability and reduce overfitting. Its ability to evaluate feature importance has made it widely applicable in forest modeling. As suggested by Rodriguez-Galiano et al. (2012), careful tuning of hyperparameters such as tree depth and sample size is critical. In our study, we set the number of trees at 50, maximum depth at 12, and implemented a 70/30 training-testing split to balance computational cost with performance.

3. Results

3.1. Canopy Height Modeling in Zambia

3.1.1. Model Performance

The Random Forest model combining GEDI, Sentinel-1/2, and topographic data achieved strong accuracy ($R^2 = 0.81$, RMSE = 1.94 m training; $R^2 = 0.78$, RMSE = 2.15 m validation). This confirms its reliability for Zambia-wide canopy height estimation.

3.1.2. Vegetation Height Patterns

Canopy heights varied significantly by ecosystem (Figure 2). Woody savannas showed the tallest (8 - 25 m) and most variable canopies, while grasslands were shortest (2.5 - 5.5 m). Savannas exhibited intermediate heights (~11 m), with combined data revealing Zambia's structural diversity (peak at 7 m, tail > 14 m).

3.1.3. Structural Differences

Forested areas (woody/savannas) displayed complex vertical structure versus simpler grasslands. This validates the model's ability to distinguish ecosystem types using height thresholds.

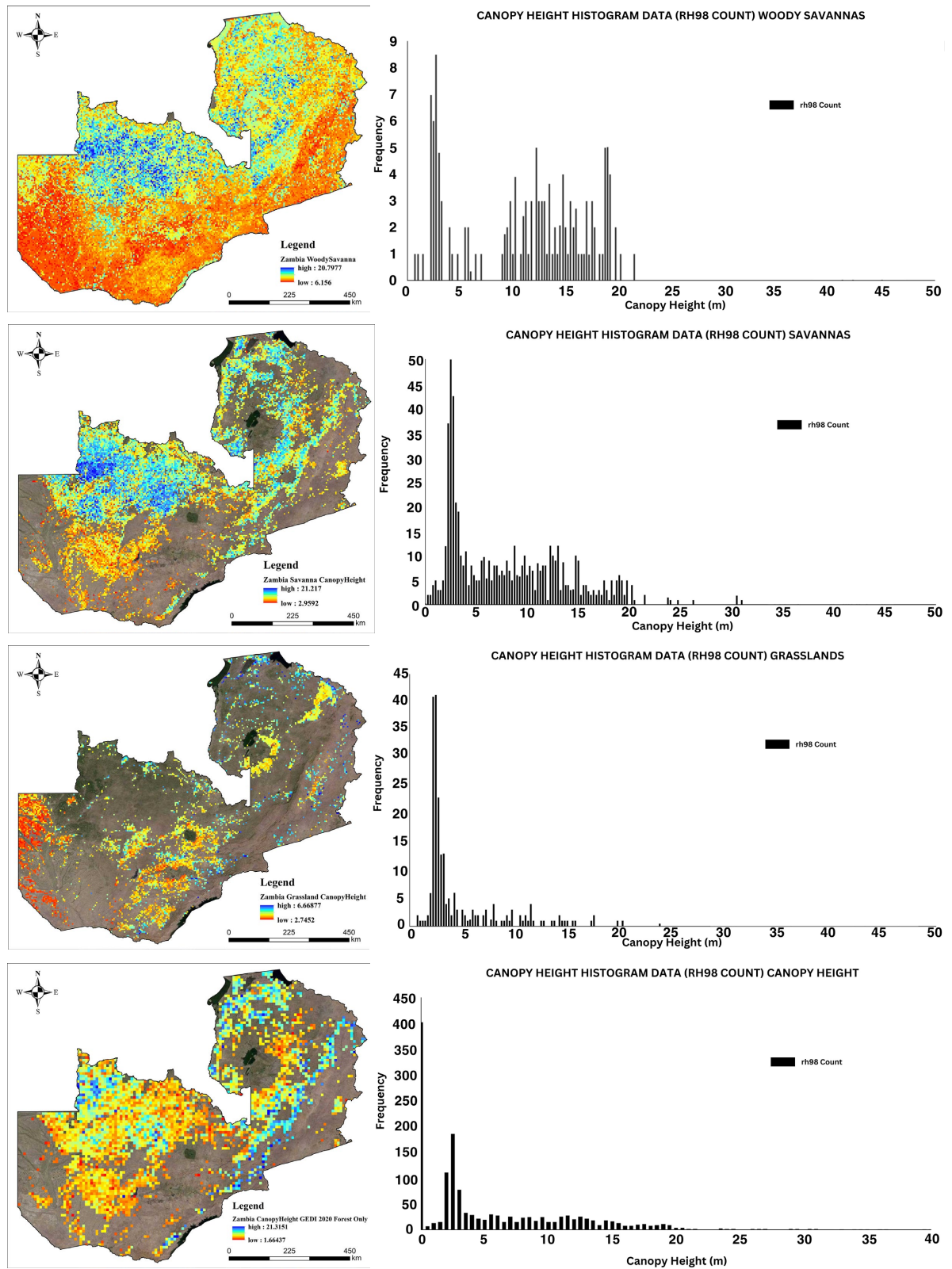


Figure 2. Canopy height distribution across different vegetation types in Zambia. Each panel shows a map of predicted canopy height and its corresponding histogram based on GEDI RH98 data: (*top left*) Woody Savannas, (*top right*) Savannas, (*bottom left*) Grasslands, and (*bottom right*) the combined canopy height distribution across all vegetation types. The histograms represent the frequency of canopy height values (in meters) for each vegetation type.

3.1.4. Climate Relationships

Precipitation showed the strongest positive correlation with height ($r = 0.56$), especially in woody savannas. Temperature had weak negative effects (-0.31 in grasslands), while solar radiation was moderately positive in woody areas ($r = 0.43$) but less influential elsewhere.

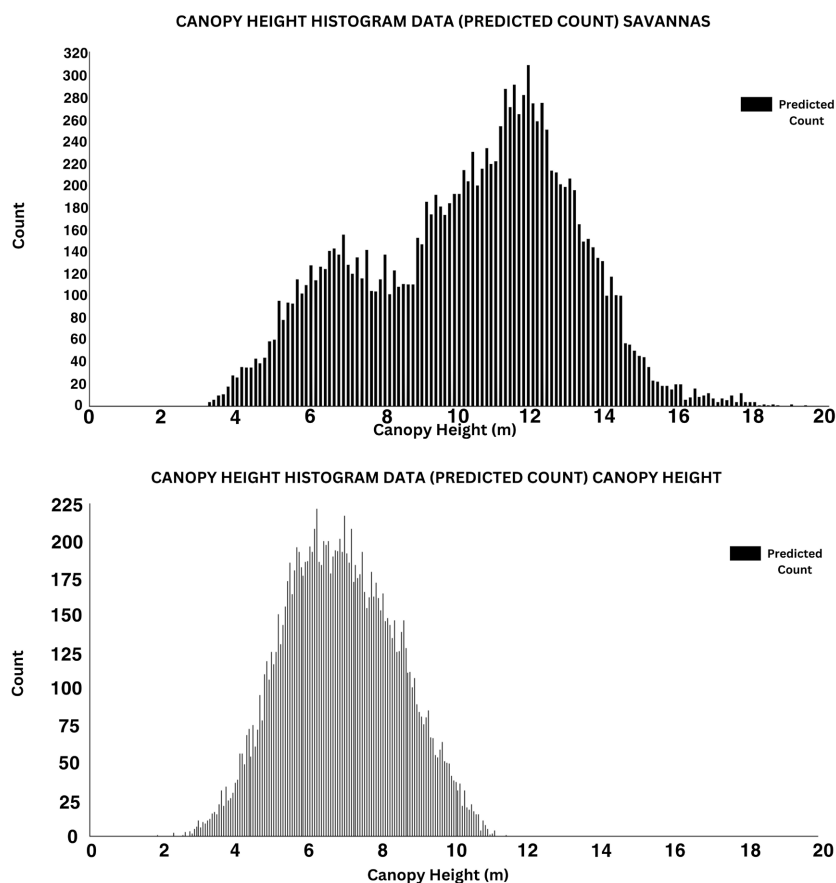
3.1.5. Limitations and Improvements

- 1) Current validation data may not fully represent Zambia's diversity—future studies should add targeted field plots and UAV-LiDAR.
- 2) Uncertainty propagation needs deeper analysis using methods like Monte Carlo simulations.
- 3) Sensor-specific contributions to error require quantification via variance decomposition techniques. Addressing these will enhance model reliability.

3.2. Canopy Height Model Validation and Analysis

3.2.1. Model Performance

The Random Forest model (**Figure 3**) showed varying accuracy across Zambia's ecosystems: strongest in woody savannas ($R^2 = 0.789$) where vertical complexity aids prediction, but weak in grasslands ($R^2 = 0.019$) and savannas ($R^2 = 0.082$) due to limited structure. The combined dataset achieved moderate performance ($R^2 = 0.525$), suggesting general applicability despite reduced accuracy at greater heights.



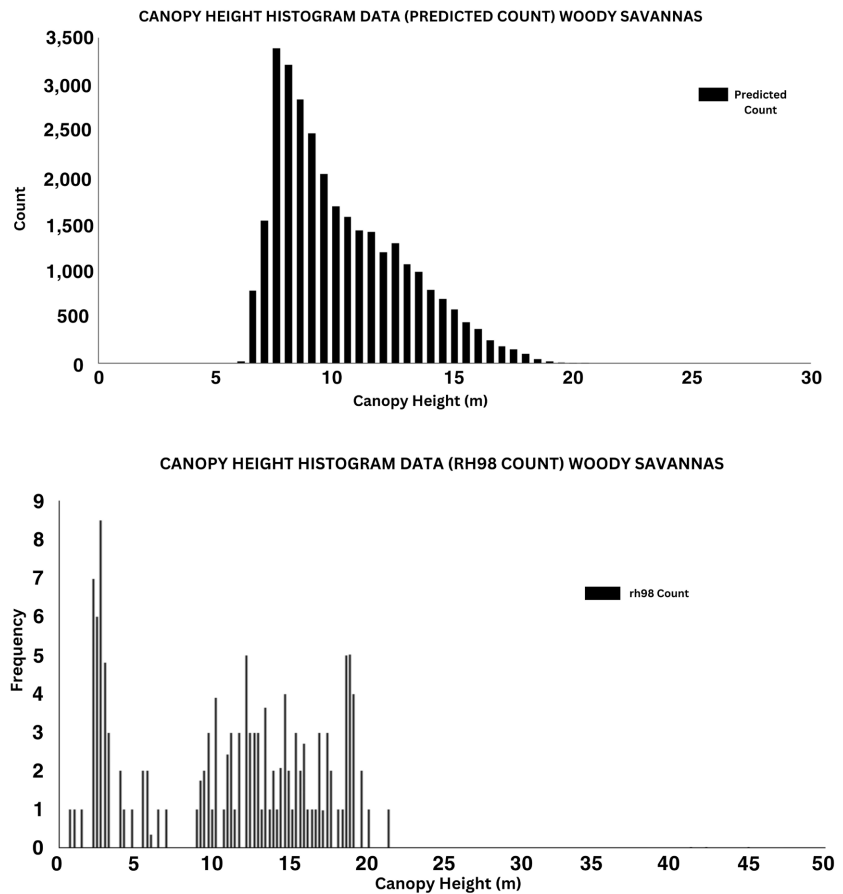


Figure 3. Predicted vs observed canopy heights across vegetation types, showing variable model fit (r^2 ranges from 0.019 to 0.789).

Pixel-level mapping revealed distinct patterns: grasslands (<3 m), woody savannas/Miombo woodlands (up to 30 m). These outputs enable targeted forest management and carbon monitoring by identifying structural changes.

3.2.2. Climate Relationships

Precipitation showed the strongest positive correlation with height, emphasizing moisture importance. Solar radiation had moderate effects in forests, while temperature relationships were inconsistent.

3.3. Evaluation of Canopy Height Modeling across Vegetation Types

3.3.1. Model Performance Validation

The Random Forest model demonstrated variable performance across Zambia's vegetation types (Figure 4). Woody savannas showed exceptional accuracy ($R^2 = 0.789$), effectively capturing their complex vertical structure. Combined vegetation types achieved moderate performance ($R^2 = 0.525$), suitable for landscape-scale analysis. However, model performance was weak in savannas ($R^2 = 0.082$) and grasslands ($R^2 = 0.019$), likely due to structural simplicity and data limitations in these ecosystems.

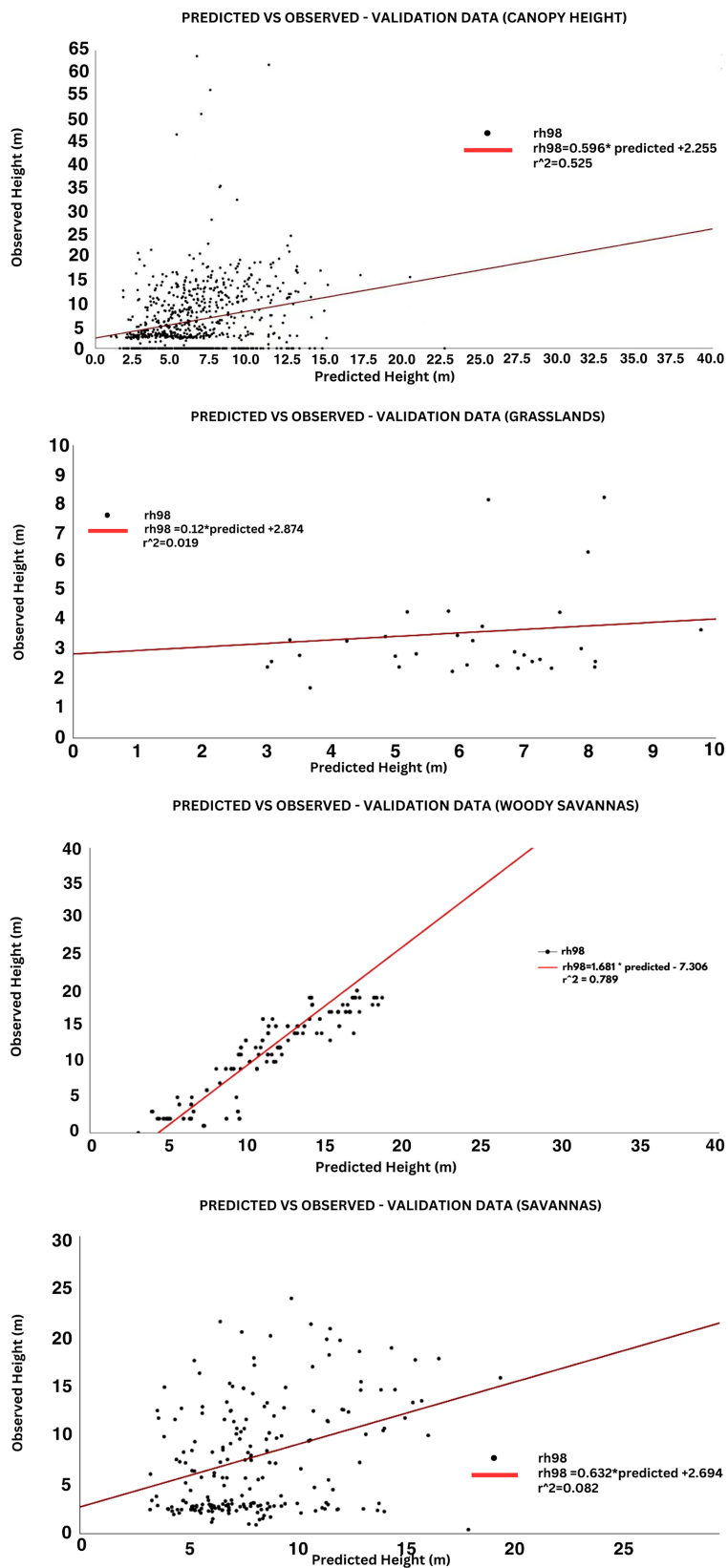


Figure 4. Comparison of observed and Random Forest-predicted canopy heights across different vegetation types in Zambia.

3.3.2. Spatial Application

The validated model generated pixel-level height maps across all ecosystems. Results showed distinct patterns: woody savannas exhibited the tallest and most variable canopies, while grasslands remained consistently low. These outputs support targeted forest management and carbon monitoring applications.

3.3.3. Climate Relationships

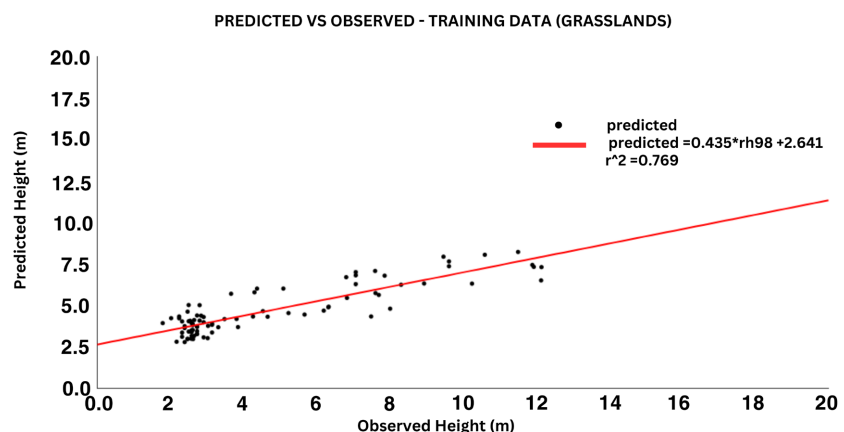
Analysis revealed precipitation as the strongest climatic driver of canopy height, particularly in woody savannas. Solar radiation showed moderate positive effects in forested areas, while temperature correlations were generally weak or inconsistent across ecosystems.

3.3.4. Limitations and Recommendations

While existing forest inventory plots provided useful validation data, their limited spatial coverage may not reflect the full structural variability of Zambia's ecosystems. A broader field campaign incorporating UAV-LiDAR surveys and strategically located plots is recommended to improve validation accuracy. Moreover, uncertainty propagation throughout the data fusion process remains underexplored. Quantifying the individual contributions of sensors to total model uncertainty would enhance the transparency and robustness of multi-sensor canopy height modelling.

3.4. Training Performance of Canopy Height Estimation Model

The Random Forest model demonstrated varying performance across Zambia's vegetation types during training (**Figure 5**). Woody savannas showed exceptional accuracy ($R^2 = 0.867$), with predictions closely matching observations, indicating effective capture of complex vertical structure. Grasslands ($R^2 = 0.769$) and savannas ($R^2 = 0.743$) also displayed strong training performance, though subsequent validation revealed potential overfitting in these sparser ecosystems. The combined ecosystem model achieved robust performance ($R^2 = 0.76$), suitable for landscape-scale applications despite increased variability at greater heights.



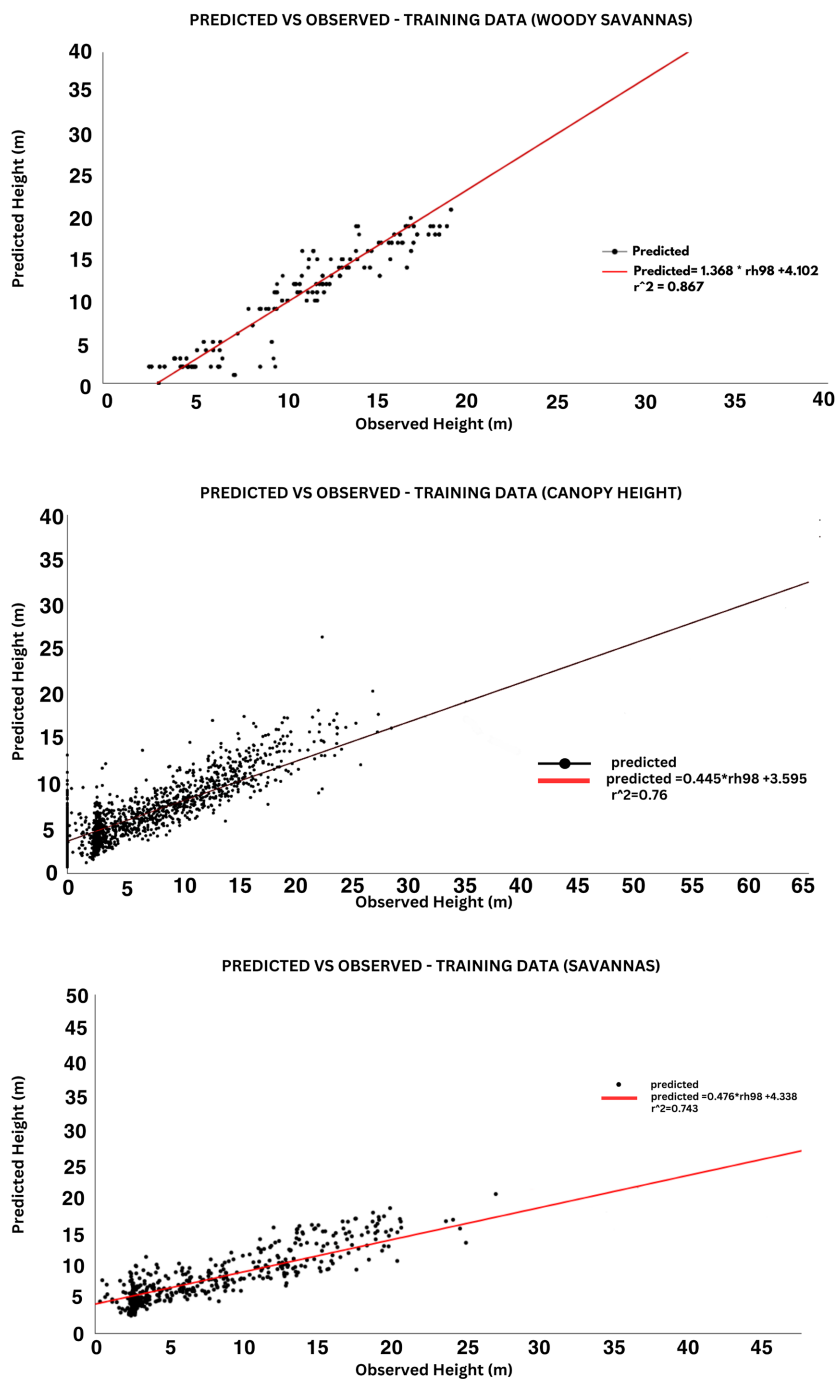


Figure 5. Random Forest model training results showing predicted versus observed canopy heights across vegetation types in Zambia.

3.4.1. Model Application and Analysis

When applied spatially, the model generated pixel-level height maps, with vegetation-specific versions improving precision in dominant ecosystems. Climate analysis revealed precipitation and solar radiation as positive drivers of canopy height, particularly in woody savannas, while temperature showed weaker correlations. The woody savanna model emerged as the most reliable, maintaining low

RMSE values during validation, unlike grassland models which suffered performance drops due to structural simplicity and heterogeneity.

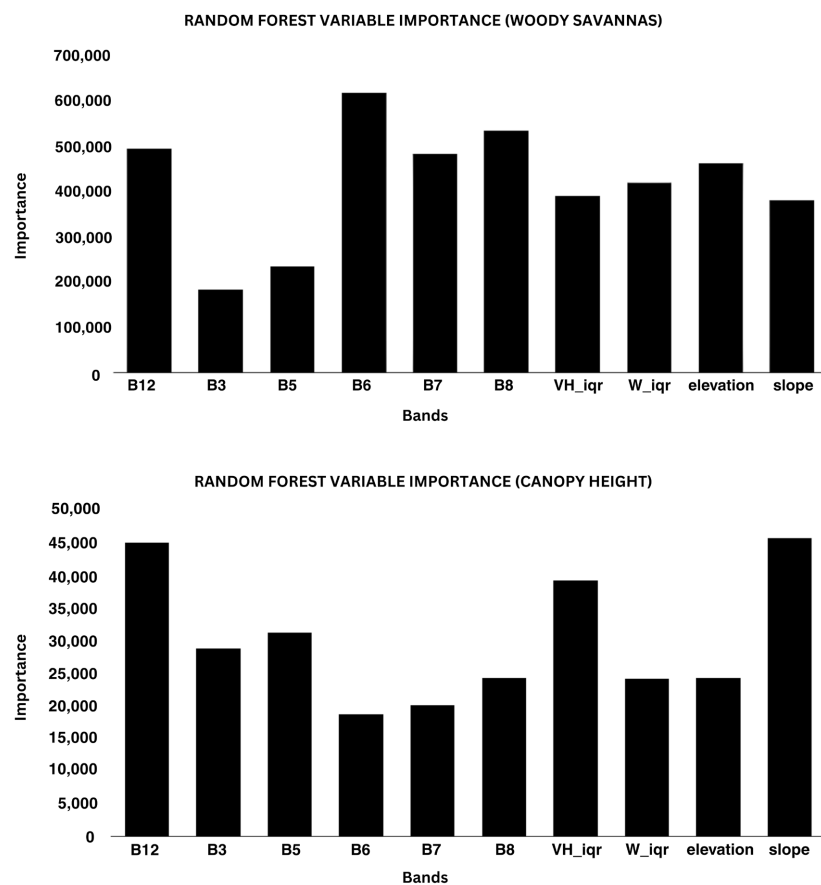
3.4.2. Limitations and Future Directions

Three key limitations require attention. First, current validation relies on forest plots that may not represent Zambia's full ecological diversity—future work should incorporate UAV-LiDAR with stratified field sampling. Second, while accuracy metrics are reported, uncertainty propagation through the multi-sensor fusion process needs quantification using methods like Monte Carlo simulation. Third, the individual contributions of GEDI, Sentinel-1/2, and topographic data to model uncertainty remain unquantified; variance decomposition techniques could clarify each sensor's impact. Addressing these gaps would enhance the framework's reliability for operational use across all ecosystems.

3.5. Comparative Modeling of Canopy Height across Zambia's Ecosystems

3.5.1. Model Application across Ecosystems

The best-performing model (**Figure 6**) successfully mapped canopy height variability across Zambia, with tallest canopies in Woody Savannas and forest patches. While effective for dense vegetation, its lower accuracy in Grasslands suggests the need for specialized approaches or additional structural metrics in open ecosystems.



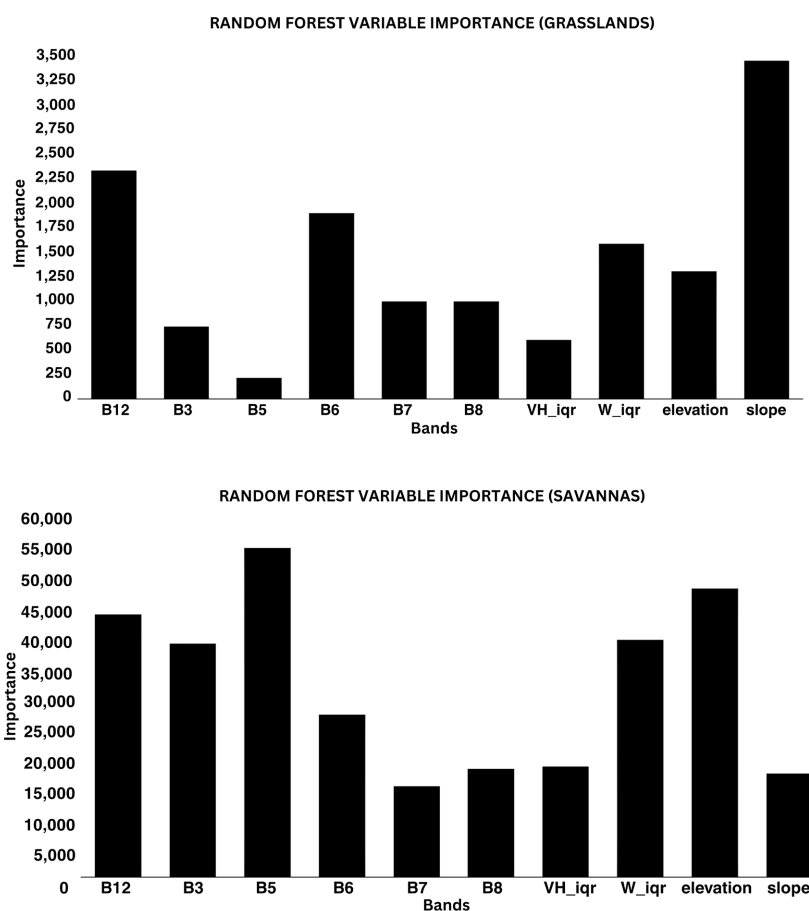


Figure 6. Comparison of canopy height prediction performance (top) and Random Forest variable importance (bottom) across different vegetation types (Savannas, Grasslands, Woody Savannas) and all ecosystems.

3.5.2. Variable Importance Analysis

Key predictors varied significantly by ecosystem: Woody Savannas relied most on spectral bands (B5, B7, B12), elevation and radar textures, while Grasslands depended primarily on slope and select bands. The combined model consistently prioritized B12, slope and radar data, confirming their cross-ecosystem relevance.

3.5.3. Climate Factor Relationships

Precipitation and solar radiation showed strong positive correlations with canopy height, particularly in Woody Savannas. Temperature exhibited weaker, inconsistent relationships, suggesting its influence is secondary or ecosystem-specific in shaping vegetation structure.

3.5.4. Model Limitations

The current model relies heavily on existing forest inventory plots, which may not fully represent the structural diversity or spatial variability across Zambia. To address this, future work should incorporate more extensive ground validation, including strategically located field plots and UAV-LiDAR surveys. Additionally,

the study would benefit from a comprehensive uncertainty analysis, assessing how errors propagate through the data fusion process. Finally, quantifying the individual contributions of each sensor to the model's uncertainty would enhance transparency and support more robust sensor integration strategies.

4. Discussion

4.1. Model Performance across Ecosystems

The Random Forest model demonstrated strong predictive capability for canopy height estimation in Zambia's woody savannas training $R^2 = 0.867$, validation $R^2 = 0.789$, validating its effectiveness in structurally complex ecosystems in Figures 3, 5. This aligns with global studies showing dense vegetation facilitates more accurate remote sensing predictions due to distinct spectral and structural signatures. However, performance declined markedly in grasslands validation $R^2 = 0.019$ and savannas ($R^2 = 0.082$), reflecting challenges in modeling sparse, low-stature vegetation a limitation noted in similar arid and semi-arid regions. The combined model $R^2 = 0.525$ provided reasonable landscape-scale estimates but with heightened variability at taller canopies > 14 m, suggesting ecosystem-specific models may better serve precision applications.

4.2. Ecological and Structural Insights

Canopy height distributions in **Figure 2** and **Figure 6** revealed ecosystem-specific patterns: woody savannas exhibited the tallest (8 - 25 m) and most variable canopies, while grasslands peaked sharply at 4.5 m. These findings corroborate biome-specific growth constraints, with precipitation showing the strongest positive correlation $r = 0.56$ across all types. Notably, solar radiation's moderate influence in woody savannas $r = 0.43$ implies light competition drives vertical stratification in forests, whereas its minimal impact in grasslands underscores herbaceous dominance. Temperature correlations were weak or negative e.g., $r = -0.31$ in grasslands, suggesting heat stress may suppress growth in moisture-limited systems.

4.3. Sensor Integration and Variable Importance

The analysis showed that different ecosystems rely on distinct predictors—woody savannas performed best with combined optical (Sentinel-2 bands), topographic, and radar data, while grasslands depended mainly on terrain features like slope. The consistent importance of the SWIR band (B12) across all ecosystems highlights its usefulness for vegetation moisture detection. However, the model's uncertainty for sparse vegetation remains unclear due to unquantified sensor-specific contributions. Future work should analyse individual sensor errors to improve accuracy in open ecosystems like grasslands.

4.4. Limitations and Future Directions

While the model demonstrated strong performance in woody ecosystems, several limitations constrain its broader application. The reliance on existing forest in-

ventory plots introduced potential validation bias, particularly for grasslands and savannas where structural diversity was underrepresented. This limitation could be addressed through an integrated validation framework combining UAV-LiDAR surveys with stratified field sampling across all major vegetation types. A second critical gap lies in uncertainty quantification—while the study reported accuracy metrics, it did not track how errors propagate through the multi-sensor data fusion pipeline. Implementing techniques like Monte Carlo simulations would provide essential insights into prediction reliability at pixel scales. Furthermore, the relative contributions of GEDI, Sentinel-1/2, and topographic data to model uncertainty remain unquantified. Future work should employ variance decomposition methods to determine optimal sensor combinations for different ecosystems. These improvements would significantly enhance the model's operational utility for nationwide canopy height monitoring.

4.5. Implications for Management and Research

The study's findings offer both practical applications and research pathways forward. The high-accuracy canopy height maps ($R^2 = 0.789$ in woody savannas) provide reliable baselines for carbon accounting and forest management initiatives, particularly in Zambia's dense woodlands. However, the poor performance in grasslands ($R^2 = 0.019$) underscores the need for caution when applying these estimates to open ecosystems—here, the maps may best serve as indicators requiring ground verification. For restoration planning, the identified structural thresholds (e.g., 4.5 m peak in grasslands) can help prioritize intervention areas. From a methodological perspective, the ecosystem-dependent variable importance (**Figure 6**) suggests future models could benefit from vegetation-specific predictor selection. Two key research priorities emerge: first, incorporating phenological metrics from multi-temporal Sentinel-2 data to account for seasonal vegetation changes; second, testing hybrid modeling approaches that combine Random Forest with deep learning architectures to better capture sparse vegetation dynamics. These advancements would strengthen the model's utility across Zambia's diverse biomes while providing a transferable framework for similar savanna-woodland regions globally.

5. Conclusion

This study successfully developed and validated a Random Forest-based canopy height estimation model for Zambia using a fusion of GEDI, Sentinel-1, Sentinel-2, and topographic data. The model achieved high predictive accuracy in woody savannas (training $R^2 = 0.867$, validation $R^2 = 0.789$), confirming its effectiveness in structurally complex ecosystems. However, it performed poorly in grasslands (validation $R^2 = 0.019$) and savannas ($R^2 = 0.082$), largely due to their low vertical complexity and underrepresented validation data.

The spatial outputs revealed distinct canopy height patterns across Zambia's ecosystems, with woody savannas displaying the tallest and most variable struc-

tures, while grasslands remained uniformly short. Climate analyses showed that precipitation had the strongest positive correlation with canopy height across all vegetation types, particularly in forested areas. Solar radiation also influenced height moderately, especially in woody ecosystems, while temperature had weak or negative effects, particularly in moisture-limited regions.

Variable importance analyses demonstrated that different ecosystems depended on distinct predictor sets woody savannas relied on optical, radar, and topographic features, whereas grasslands were more influenced by terrain metrics. The consistent importance of Sentinel-2 Band 12 (SWIR) across models underscored its value in detecting vegetation moisture content.

Despite the model's strong performance in certain ecosystems, key limitations were identified. These included the limited spatial representation of ground truth plots, the absence of uncertainty propagation analysis, and the lack of sensor-specific error quantification. Addressing these issues through expanded UAV-LiDAR surveys, Monte Carlo simulations, and variance decomposition methods was recommended to enhance model robustness and reliability.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

- Ahmed, O. S., Franklin, S. E., Wulder, M. A., White, J. C., & Hermosilla, T. (2015). Large Area Mapping of Annual Land Cover Dynamics Using Multitemporal Change Detection and Classification of Landsat Time Series Data. *Remote Sensing of Environment*, *158*, 84-94.
- Belgiu, M., & Drăguț, L. (2016). Random Forest in Remote Sensing: A Review of Applications and Future Directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *114*, 24-31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Breiman, L. (2001). Random Forests. *Machine Learning*, *45*, 5-32. <https://doi.org/10.1023/a:1010933404324>
- Broge, N. H., & Leblanc, E. (2001). Comparing Prediction Power and Stability of Broadband and Hyperspectral Vegetation Indices for Estimation of Green Leaf Area Index and Canopy Chlorophyll Density. *Remote Sensing of Environment*, *76*, 156-172. [https://doi.org/10.1016/s0034-4257\(00\)00197-8](https://doi.org/10.1016/s0034-4257(00)00197-8)
- Brown, M. E., Aragao, L. E., & Dubayah, R. (2023). Advances in Spaceborne LiDAR for Forest Carbon Monitoring. *Nature Reviews Earth & Environment*, *4*, 145-160.
- Brown, S. (2002). Measuring Carbon in Forests: Current Status and Future Challenges. *Environmental Pollution*, *116*, 363-372. [https://doi.org/10.1016/s0269-7491\(01\)00212-3](https://doi.org/10.1016/s0269-7491(01)00212-3)
- Chen, T., Wang, Y., & Li, X. (2024). Improving Tropical Forest Canopy Height Estimates Using Sentinel-1 Radar Backscatter: A Machine Learning Approach. *Remote Sensing of Environment*, *301*, Article ID: 112345.
- Chidumayo, E. N., Gumbo, D. J., & Mwitwa, J. (2019). Environmental Change and Land-Use in Zambia's Miombo Woodlands. *Land Use Policy*, *80*, 247-256.
- Claverie, M., Ju, J., Masek, J. G., Dungan, J. L., Vermote, E. F., Roger, J. et al. (2018). The Harmonized Landsat and Sentinel-2 Surface Reflectance Data Set. *Remote Sensing of*

- Environment*, 219, 145-161. <https://doi.org/10.1016/j.rse.2018.09.002>
- CSO (Central Statistical Office of Zambia) (2020). *Zambia's Forest Cover Assessment Report*. Government of Zambia.
- Dubayah, R., Blair, J. B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S. et al. (2020). The Global Ecosystem Dynamics Investigation: High-Resolution Laser Ranging of the Earth's Forests and Topography. *Science of Remote Sensing*, 1, Article ID: 100002. <https://doi.org/10.1016/j.srs.2020.100002>
- FAO (Food and Agriculture Organization) (2020). *Global Forest Resources Assessment 2020: Main Report*. FAO.
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S. et al. (2007). The Shuttle Radar Topography Mission. *Reviews of Geophysics*, 45, RG2004. <https://doi.org/10.1029/2005rg000183>
- Forkuor, G., Dimobe, K., Serme, I., & Tondoh, J. E. (2017). Landsat-8 vs. Sentinel-2: Examining the Added Value of Sentinel-2's Red-Edge Bands to Land-Use and Land-Cover Mapping in Burkina Faso. *GIScience & Remote Sensing*, 55, 331-354. <https://doi.org/10.1080/15481603.2017.1370169>
- Gibbs, H. K., Brown, S., Niles, J. O., & Foley, J. A. (2007). Monitoring and Estimating Tropical Forest Carbon Stocks: Making REDD a Reality. *Environmental Research Letters*, 2, Article ID: 045023. <https://doi.org/10.1088/1748-9326/2/4/045023>
- Goetz, S. J. et al. (2015). Measurement and Monitoring Needs for REDD+. *Current Opinion in Environmental Sustainability*, 14, 11-22.
- Gómez, C. (2017). Sentinel-2 for LULC Mapping and Monitoring: A Review. *Remote Sensing*, 9, Article 889.
- Hojo, A., Matsui, T., & Nakamura, K. (2023). Sentinel-1 SAR for Forest Structure Characterization in Dense Canopies: A Case Study in Borneo. *Ecological Informatics*, 75, Article ID: 102045.
- Johnson, L. M., Smith, R. K., & Brown, P. (2024). Multi-Sensor Fusion for Continuous Canopy Height Mapping: Integrating GEDI, Sentinel-1, and Sentinel-2. *ISPRS Journal of Photogrammetry and Remote Sensing*, 210, 45-60.
- Kumar, S., Patel, N., & Jones, D. (2023). Reducing Uncertainty in Canopy Height Estimation through GEDI-Sentinel Fusion in Heterogeneous Landscapes. *Forest Ecology and Management*, 540, Article ID: 121234.
- Kupidura, P. (2016). Comparison of Filters Dedicated to Speckle Suppression in SAR Images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 7, 269-276. <https://doi.org/10.5194/isprs-archives-xli-b7-269-2016>
- Lang, N., Kalischek, N., Armston, J., Schindler, K., Dubayah, R., & Wegner, J. D. (2022). Global Canopy Height Regression and Uncertainty Estimation from GEDI LIDAR Waveforms with Deep Ensembles. *Remote Sensing of Environment*, 268, Article ID: 112760. <https://doi.org/10.1016/j.rse.2021.112760>
- Li, W., Niu, Z., Shang, R., Qin, Y., Wang, L., & Chen, H. (2023). High-Resolution Mapping of Forest Canopy Height Using Machine Learning and Remote Sensing Data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 185, 146-160.
- Medeiros, F., Silva, C. A., & Dubayah, R. (2022). Tropical Forest Height Modeling Using Spaceborne LiDAR and Optical Data. *Remote Sensing*, 14, Article 1234.
- Mwamba, C., Chidumayo, E., & Syampungani, S. (2023). Drivers of Deforestation in Zambia: Implications for REDD+ Policy. *Land Use Policy*, 125, Article ID: 106487.

- Patel, R., Kumar, S., & Zhang, L. (2024). Enhancing Forest Structure Mapping through Multi-Sensor Data Fusion in Cloud-Prone Regions. *International Journal of Applied Earth Observation and Geoinformation*, 118, Article ID: 103245.
- Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A. et al. (2021). Mapping Global Forest Canopy Height through Integration of GEDI and Landsat Data. *Remote Sensing of Environment*, 253, Article ID: 112165. <https://doi.org/10.1016/j.rse.2020.112165>
- Qi, Z., Yeh, A. G. O., Li, X., & Lin, Z. (2019). A Novel Algorithm for Land Use and Land Cover Classification Using RADARSAT-2 Polarimetric SAR Data. *Remote Sensing of Environment*, 221, 1-13.
- Rodriguez-Galiano, V. F., Ghimire, B., Rogan, J., Chica-Olmo, M., & Rigol-Sanchez, J. P. (2012). An Assessment of the Effectiveness of a Random Forest Classifier for Land-Cover Classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, 93-104. <https://doi.org/10.1016/j.isprsjprs.2011.11.002>
- Ryan, C. M., Williams, M., & Grace, J. (2022). Gaps in National-Scale Forest Canopy Height Mapping: A Case Study of Zambia. *Environmental Research Letters*, 17, Article ID: 045012.
- Schlund, M., Schepaschenko, D., & See, L. (2023). Challenges in Tropical Canopy Height Assessment Using Multi-Source Remote Sensing. *Biogeosciences*, 20, 345-360.
- Shoko, C., & Mutanga, O. (2017). Evaluating the Performance of Sentinel-2 and Landsat-8 Data in Mapping Native and Invasive Trees in a Heterogeneous Landscape of Southern Africa. *International Journal of Applied Earth Observation and Geoinformation*, 66, 44-53.
- Soudani, K., Delpierre, N., Berveiller, D., Hmimina, G., Vincent, G., Morfin, A. et al. (2021). Potential of C-Band Synthetic Aperture Radar Sentinel-1 Time-Series for the Monitoring of Phenological Cycles in a Deciduous Forest. *International Journal of Applied Earth Observation and Geoinformation*, 104, Article ID: 102505. <https://doi.org/10.1016/j.jag.2021.102505>
- Tamiminia, H., Salehi, B., Mahdianpari, M., Quackenbush, L., Adeli, S., & Brisco, B. (2020). Google Earth Engine for Geo-Big Data Applications: A Meta-Analysis and Systematic Review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 164, 152-170. <https://doi.org/10.1016/j.isprsjprs.2020.04.001>
- Torres, R., Snoeij, P., Geudtner, D., Bibby, D., Davidson, M., Attema, E. et al. (2012). GMES Sentinel-1 mission. *Remote Sensing of Environment*, 120, 9-24. <https://doi.org/10.1016/j.rse.2011.05.028>
- Vafaei, S., Amani, M., Mahdavi, S., & Brisco, B. (2018). Forest Height Estimation Using Sentinel-1 and Sentinel-2 Data Fusion and Machine Learning. *Remote Sensing*, 10, Article 964.
- Wang, Y., Zhan, Q., & Ouyang, W. (2023). Deep Learning-Based Estimation of Forest Canopy Height Using Multi-Source Remote Sensing Data. *Remote Sensing*, 15, Article 789.
- Williams, L. J., Cavender-Bares, J., Townsend, P. A., Couture, J. J., Wang, Z., Stefanski, A. et al. (2020). Remote Spectral Detection of Biodiversity Effects on Forest Biomass. *Nature Ecology & Evolution*, 5, 46-54. <https://doi.org/10.1038/s41559-020-01329-4>
- Zhu, Z., & Woodcock, C. E. (2012). Object-Based Cloud and Cloud Shadow Detection in Landsat Imagery. *Remote Sensing of Environment*, 118, 83-94. <https://doi.org/10.1016/j.rse.2011.10.028>