Tropical Forest Height Estimation Using L-and P-Band Polarimetric Synthetic Aperture Radar Backscatters with Machine Learning Models

REZAUL HASAN BHUIYAN
[August 2024]

SUPERVISORS:
dr. C. Paris (Claudia)
dr, M, Schlund (Michael)

Tropical Forest Height Estimation Using L-and P-Band Polarimetric Synthetic Aperture Radar Backscatters with Machine Learning Models

[REZAUL HASAN BHUIYAN]
Enschede, The Netherlands, [August 2024]

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Natural Resources Management

SUPERVISORS:

dr. C. Paris (Claudia)
dr, M, Schlund (Michael)

THESIS ASSESSMENT BOARD:
dr. T. Wang (Tiejun) (Chair)
M. Khodadadzadeh (Mahdi) (External Examiner)

ABSTRACT

Forest Canopy Height (CH) is an important biophysical parameter for effective forest management and conservation efforts. This study presents a Fully Convolutional Network (FCN)-based approach for estimating CH using L- and P-band polarimetric synthetic aperture radar (PolSAR) backscatters and their combinations. Specifically, a customized UNet architecture, tailored to the unique characteristics of SAR data, is employed to estimate CH in both heterogeneous and homogeneous forest sites, Lope and Mabounie respectively, located in Gabon. Results indicated that combinations of L- and P-band polarimetric backscatters led to CH predictions that were more accurate compared to single-band retrievals, with dual-band combinations producing Root Mean Square Error (RMSE) values of 4.03 m for Lope and 3.78 m for Mabounie. The estimation accuracies from the combinations of Synthetic Aperture Radar (SAR) bands were consistent across the two study areas, whereas the retrieval performance varied for individual bands. P-band-based retrievals were more accurate than L-band for the homogeneous Mabounie site (RMSE of 4.26 m vs. 4.63 m). However, for the heterogeneous Lope site, no significant RMSE difference was found between L- and P-band models. Upon comparison with other machine learning models, it was observed that the customized UNet model produced RMSE values three times lower than those of Random Forest (RF) and Light Gradient Boosting Machine (LGBM). These results are relevant in the context of upcoming long-wavelength SAR missions, such as the European Space Agency (ESA) BIOMASS and NASA-ISRO Synthetic Aperture Radar (NISAR), which could potentially be used for global forest canopy height mapping.

ACKNOWLEDGEMENTS

Where have all the flowers gone?
I dedicate this work to the children of Palestine, who had names and dreams.
.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Background

Climate change poses a global threat to ecosystems. Its adverse effects are manifesting more rapidly than predicted a decade ago (IPCC, 2022). Forests play a crucial role in regulating the global climate and offering essential ecosystem services such as carbon sequestration and biodiversity conservation. Despite their importance, forests are degrading at an alarming rate. For instance, Bourgoin et al., (2024) reported that anthropogenic intervention is responsible for an increase in tropical forest degradation by 200%. Besides, frequent climate extremes, such as heatwaves and droughts, have triggered alterations in forest composition and increased tree mortality rates (Hartmann et al., 2022). These natural and human-induced factors have caused forests to act as carbon sources in the atmosphere (Mitchard, 2018). Therefore, minimization of forest degradation can potentially reduce carbon dioxide emissions and increase terrestrial carbon sequestration.

Recognizing the importance of forest conservation and management, the United Nations Framework Convention on Climate Change (UNFCCC) introduced the Reducing Emissions from Deforestation and Forest Degradation (REDD+) program (UNEP, 2018). REDD+ encourages countries to adopt National Forest Monitoring Systems (NFMS) (UNEP, 2018). However, the Monitoring, Reporting, and Verification (MRV) components within NFMS require accurate, timely, and reliable approaches for measuring forest biophysical variables. Here, earth observation (EO) plays a significant role by providing direct, precise, and timely data at regional and global scales (Lang et al., 2023). Furthermore, EO methods have extended beyond spatial mapping and enabled detailed forest morphological information extraction. In contrast, In-situ forest data collection is often labor-intensive, time-consuming, and limited to small scales (Lu et al., 2016).

Canopy height (CH) is a crucial vegetation biophysical parameter for understanding ecosystem structure, carbon storage, and biodiversity in forests (Lang et al., 2023). The available EO datasets used for forest CH estimation include LiDAR (Dubayah & Drake, 2000), optical (Verrelst et al., 2015), and imaging radar (Moreira et al., 2013). LiDAR accurately estimates forest 3D structures (Duncanson et al., 2010), however, its extensive application is limited by high acquisition costs and cloud penetration inabilities, especially in tropical regions (Ge et al., 2022). While optical imagery primarily provides sensitivity to the chemistry of the target objects and is often subjected to saturation in dense tropical forests, radar is sensitive to their structure (Zhu et al., 2018). The potential of Synthetic Aperture Radar (SAR) in penetrating different forest layers and extracting detailed structural information makes it a well-suited dataset option for CH estimation (Ramachandran et al., 2023). Furthermore, polarimetric SAR (pol-SAR) backscatters, such as HH, HV, and VV, interact uniquely with different tree components. HH backscatter primarily interacts with trunks and branches, HV with the canopy structure, and VV with surface features (Wang et al., 1993). Furthermore, SAR techniques such as Interferometry (InSAR) and polarimetric Interferometry (PolInSAR) are traditionally used to estimate forest height (Chen et al., 2021). InSAR uses phase differences between at least two images, while PolInSAR combines polarimetric and interferometric data for detailed height information (Zhu et al., 2023). However, forest height retrievals using these techniques are influenced by various constraints associated with volume, temporal, and system decorrelation (Kugler et al., 2015; Lavalle & Hensley, 2012; Rizzoli et al., 2022). For example, parameters like window size and quantization compensation are assumed constant over a site but are data-dependent (Mahesh & Hänsch, 2023). This underscores the challenge posed by the lack of uniform processing parameters across diverse landscapes (Mahesh & Hänsch, 2023). Moreover, another significant constraint of this approach is the

scarcity of SAR data (Ge et al., 2022). Existing space-borne data sources such as Sentinel-1 (C-band), RADARSAT-2 (C-band), and ALOS-PALSAR2 (L-band) utilize repeat-pass acquisition mode. This mode is susceptible to temporal decoration and geometric distortions and produces a high bias in CH estimation due to coherence loss (Khati et al., 2018). Moreover, short wavelengths such as C- and X-bands fail to reach the ground by penetrating the canopy (Kuenzer et al., 2011). As a result, ground information remains unavailable for these wavelengths in dense, complex, and tall forest sites, such as tropical forests. Berninger et al. (2019) estimated CH using short wavelengths (C- & X-) PolInSAR in the Indonesian tropical forest site and found that C-band predicted CH with an R² of 0.62–0.63, overestimating heights of 0–15 m and underestimating heights above 15 m. And X-band-based models showed an R² of 0.58–0.66, with general underestimation. Both bands exhibited saturation effects for trees taller than 15 m. Another research conducted by Kugler et al. (2015) in the Sungai Wain tropical forest site in Indonesia used longer wavelengths such as L- and P-band PolInSAR to estimate forest CH. The study revealed that L-band Pol-InSAR was effective for estimating heights up to 40 m, while P-band PolInSAR was somewhat able to estimate heights up to 60 m, both with a correlation coefficient of 0.44. This implied the effectiveness of longer wavelengths on CH estimation. Polarimetric SAR (PolSAR) is another technique that has been used effectively in forest height estimation studies over the years (Pourshamsi et al., 2018, 2021). PolSAR usually provides information on physical scattering mechanisms (Lee & Pottier, 2017). Notably, PolSAR is unable to provide vertical information directly and requires reference height data to establish statistical regressions (Pourshamsi et al., 2021). Understanding the potential of the PolSAR technique and SAR wavelengths, Garestier et al. (2009) employed L-band and P-band polarimetric SAR (PolSAR) data in maritime pine forests in Nezer, France. They found that the P-band demonstrated a strong height correlation (R² of 0.93, RMSE error < 2 m), while the L-band was saturated above 6 m. The study also noted limitations related to the forest biome and indicated a need for further research on tropical forest sites using PolSAR data.

The growing importance of machine learning (ML) algorithms such as decision tree ensembles and support vector machines (SVM) has been observed over the years in forest CH studies owing to their ability to produce better accuracy compared to parametric regression-based models (Lu et al., 2016). They can also overcome the saturation effect of short wavelengths to some extent (Vafaei et al., 2018). Pourshamsi et al. (2021) explored the use of L-band PolSAR features combined with LiDAR measurements, employing ML models such as Random Forest (RF), Rotation Forest (RoF), Canonical Correlation Forest (CCF), and Support Vector Machine (SVM) for estimating CH in the tropical forests of Gabon. The study demonstrated good accuracy with an average R² of 0.70 and RMSE of 10 m. However, the study also underlined the potential underestimation of CH due to the limited penetration capabilities of L-band SAR and suspected that the coarse resolution of the SRTM DEM might have prevented the complete correction of topographic effects. Nevertheless, these pixel-based traditional ML models require hand-crafted features such as polarimetric ratios, indices, polarimetric decompositions, etc. Additionally, without incorporating external features such as Gray Level Co-occurrence Matrix (GLCM) parameters, they cannot incorporate spatial dependencies (Ge et al., 2022). Compared to traditional ML methods, Convolutional Neural Networks (CNNs) can more effectively capture spatial context information and recently have been employed in several forest CH estimations studies (Tolan et al., 2024; Li et al., 2023; Pascarella et al., 2023; Gazzea et al., 2023; Ge et al., 2022). Li et al. (2023) proposed a CNN-based framework for forest height estimation using L-band PolInSAR and PCGrad, achieving an RMSE of 10.15 m at the tropical forest site. However, the accuracy was inadequate and required further optimization and the use of advanced ML techniques like Fully Convolutional Neural Networks (FCN). Ge et al. (2022) proposed an improved semi-supervised UNet model based on FCN for forest height mapping using combined optical, SAR, and LiDAR imaging. Although the study achieved good accuracy with RMSEs ranging from 2.09 m to 2.02 m, such approaches are often constrained by acquisition costs,

cloud cover, and temporal resolution of different datasets. Mahesh & Hänsch (2023), measured forest height using complex coherence of TanDEM-X SAR data on UNet architecture. This study used rectified linear unit (ReLU) in the encoder and decoder blocks of UNet architecture. Applying the ReLU function to the real part of complex pixel values from an interferometric SLC image could cause a loss of information by setting all negative real parts to zero, potentially discarding significant amplitude details. Nonetheless, the results showed the potential of a UNet-based approach in achieving more accurate CH estimations.

To the best of my knowledge, no previous study has proposed a UNet model customized to the specific properties of the considered SAR data for CH mapping and compared its performances against standard ML algorithms. This study will focus on L- and P- bands PolSAR backscatters.

1.2. Problem Statement

Forest conservation and management require accurate and precise estimations of CH within the MRV system. With the upcoming ESA's BIOMASS and NISAR missions, there will be an abundance of SAR L- and P-band products, increasing the necessity for further research utilizing PolSAR backscatters and advanced ML techniques. Considering the research gaps addressed in the previous section, this study aims to propose a customized UNet framework to estimate CH using PolSAR backscatters of L-, P-, and L- and P- (dual-band) combinations. The proposed framework considered the SAR backscatter properties during the customization of the UNet convolutional layers and activation functions. As discussed in the literature review, longer wavelengths (L- and P-bands) can produce better accuracy in CH estimation (Patenaude et al., 2005). However, a comparative assessment of L- and P-band PolSAR backscatters needs to be carried out for CH estimation. Additionally, this study aims to address the limitations of using PolSAR bands individually. This will be done by analyzing the accuracy of CH retrieval using combinations of L- and P-band PolSAR backscatter compared to individual band retrievals. The performance will be evaluated on both heterogeneous and homogeneous forest sites in Lope and Mabounie, Gabon, respectively. Lastly, this study seeks to evaluate the performance of the proposed UNet model against existing ML models such as Random Forest (RF) and Light Gradient Boosting Machine (LGBM), utilizing dual-band PolSAR backscatter to estimate CH.

1.3. Objectives, Research Questions and Hypothesis

This study aims to propose a customized UNet framework to estimate forest CH using PolSAR backscatter data. Additionally, the models will evaluate the estimation accuracy performance of L-band, P-band, and dual-band PolSAR backscatter for both heterogeneous and homogeneous forest sites. This study also aims to compare the performance of the customized UNet model against two traditional decision-based ensemble tree ML models namely RF and LGBM.

Objective 1: To compare CH estimation accuracies for L- and P- bands PolSAR backscatters (HH, HV, and VV) and their combinations.

RQ1: For a heterogenous forest site, is there a significant difference between the accuracy of estimated CH from L-, P-, and dual-band PolSAR backscatters?

RQ2: For a homogenous forest site, is there a significant difference between the accuracy of estimated CH from L-, P-, and dual-band PolSAR backscatters?

Hypothesis: Dual-band PolSAR backscatters will produce significantly better accuracy than individual Bands.

Objective 2: To compare the accuracies of different ML algorithms for CH estimation using Dual-band PolSAR backscatters.

RQ1: For a heterogenous forest site, is there a significant difference between the accuracy of estimated CH from FCN (UNet) and traditional RF and LGBM models?

RQ2: For a homogenous forest site, is there a significant difference between the accuracy of estimated CH from FCN (UNet) and traditional RF and LGBM models?
Hypothesis: UNet will achieve significantly higher accuracies compared to RF and LGBM.

2. STUDY SITES AND DATA

The datasets utilized in this study were obtained as part of the AfriSAR campaign. The campaign was a joint venture by the European Space Agency (ESA) and the National Aeronautics and Space Administration (NASA)to gather airborne data from tropical forest sites in Gabon, including Lope, Mabounie, Mondah, and Rabi (Hajnsek et al., 2017). This study used L- and P-band polSAR data and Land, Vegetation, and Ice Sensor (LVIS) (LVIS)forest height data of Lope and Mabounie. Site selection was done based on the structural and species diversity of these forest areas where Lope is a heterogenous site and Mabounie is a homogenous forest site (Wang et al., 2022) Figure 1.
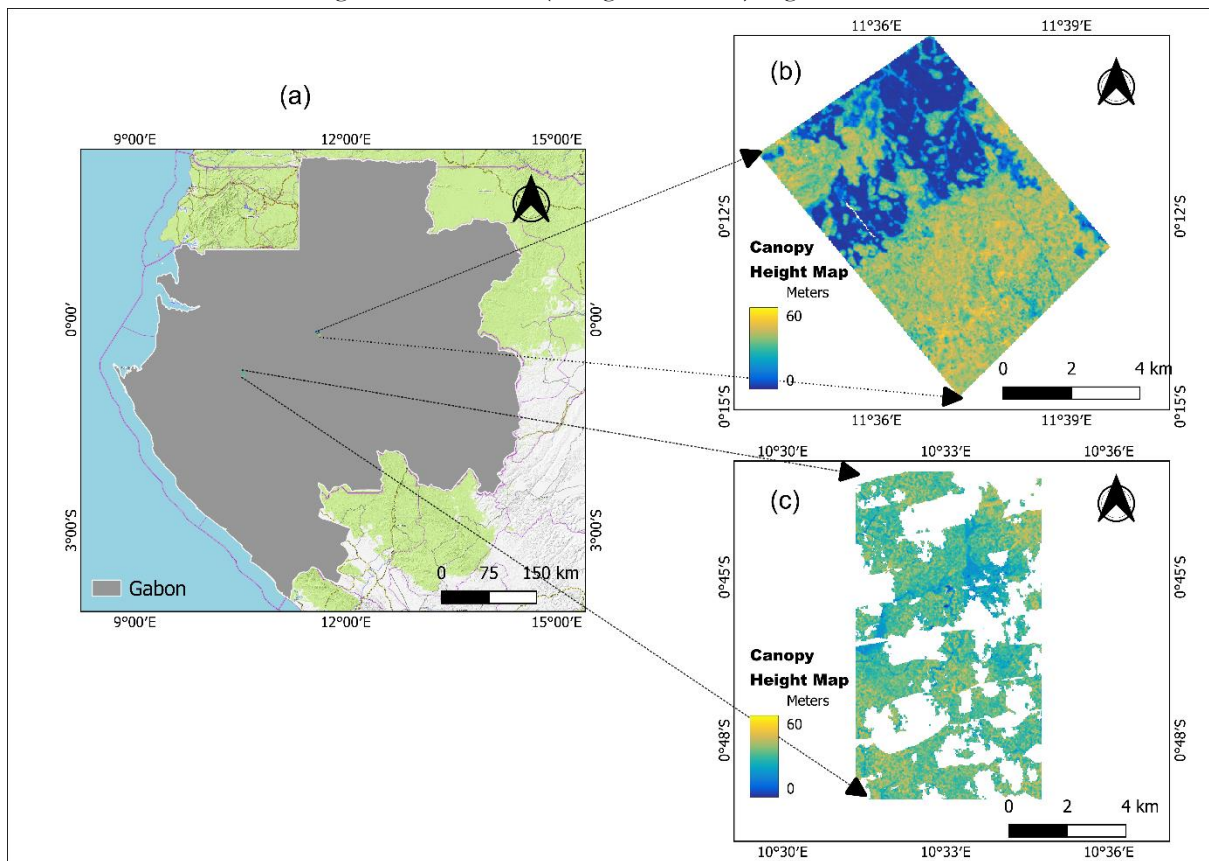
Figure 1. (a) Location of Lope and Mabounie sites in Gabon, (b) canopy height map of Lope derived from LiDAR (RH98), (c) canopy height map of Mabounie derived from LiDAR (RH98).

2.1. Lope

The first study site is Lope National Park, located at 0°30'00"S and 11°30'00"E in central Gabon (Figure 1 (b)) (Pourshamsi et al., 2021). This site covers an area of about 49,00 km², consisting of savannah and forest landscapes on hilly terrain (Marselis et al., 2018). The CH height distribution on the RH98 map shows three dominant forest regions (Figure 1(b) & Figure 2). The dark blue areas, covering 40% of the

site, consist of non-vegetated soils and sparse savannas up to 20 meters (Pourshamsi et al., 2018). Light blue areas, covering 20%, contain medium-height vegetation (20–40m) and are found at the edges between non-vegetated and forested regions. Green and light-yellow areas, covering 40%, represent dense forests (40–60m) with trees of various ages and species, although very tall trees above 50m are scarce (Pourshamsi et al., 2021). The forest is primarily composed of diverse tropical species such as the Custard Apple, Ethiopian Pepper, Frankincense, Okoume, Ochna, and Ironwood (Lewis & Labrière, 2016). The terrain is gently rolling, featuring an average slope of up to 25% (Pourshamsi et al., 2018).

2.2. Mabounie

The Mabounie site, located in central-western Gabon at 00°43' S and 10°31' E, covers about 43.3 km² (Figure 1(c)) (Wang et al., 2022). The landscape is mostly flat with a few mild slopes (Pardini et al., 2018). Unlike the Lope CH distribution, this site is a mature forest with old-growth tall trees, primarily 30-40 meters high, and very few exceed 45 meters (Figure 2) (Zhang et al., 2023a). The dominant tree species in Mabounie include Okoume, African Mahogany, Moabi, Ebony, Azobe, and Aiele (Marselis et al., 2018).
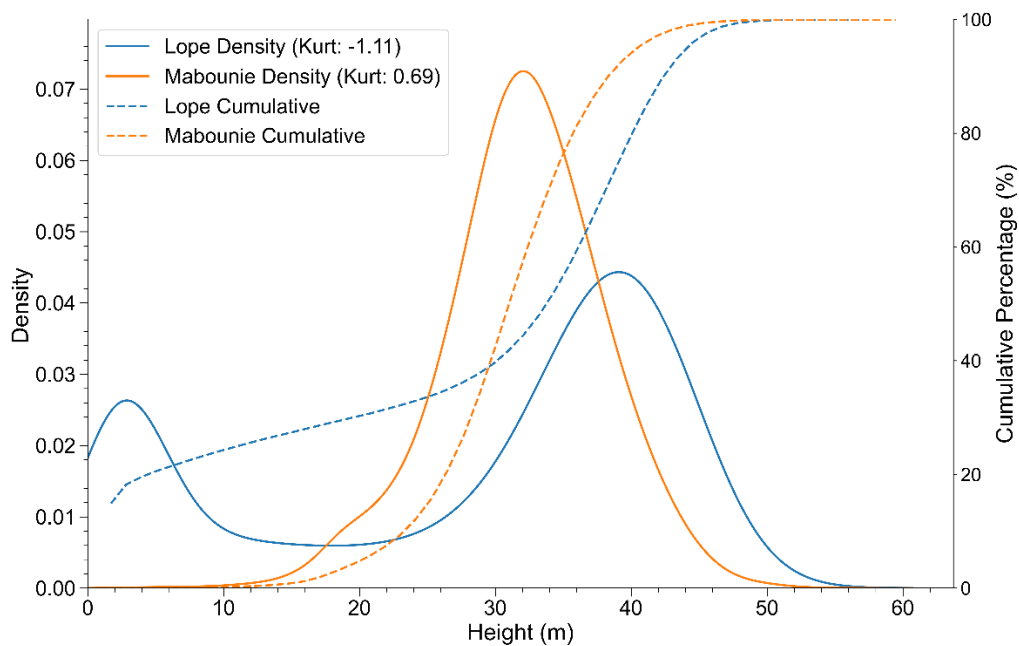
Figure 2 Density distribution and cumulative percentage of canopy height for Lope and Mabounie.

2.3. RADAR acquisition

The German Aerospace Centre (DLR) collected data for the AfriSAR campaign during the wet season in February 2016. This campaign acquired L-band (23 cm) data simultaneously with P-band (69 cm) data using along tracks (Liu et al., 2021). The F-SAR system was employed for SAR data collection. The key parameters of the F-SAR sensor are detailed in Table 1 (Hajnsek et al., 2017). The acquired raw data was processed to detect and geocode beta-naught values (Reigber et al., 2012). The data was geocoded using TanDEM-X DEM due to the closer temporal resolution to the SAR acquisition date (Hajnsek et al., 2017). This study used GTC products with an additional topographic correction, following the approach outlined by Hoekman & Reiche (2015), to detect gamma naught ($\gamma 0$). The terrain correction included an improved multi-model terrain correction approach: normalization for slope-induced variation and

scattering mechanisms correction (Hoekman & Reiche, 2015). It has been reported that the backscatter coefficient gamma-nought ($\gamma0$) is more suitable for volumetric backscatters such as vegetation (Small, 2011). Lastly, the SAR backscatter coefficient $\gamma0$ was converted into decibels using equation 1.

$$\gamma^0 \,[dB] = \, 10 \, * \, log10(\gamma^0 \,[linear])$$ (1)

Table 1 Key F-SAR Parameters

	Wavelengths	L-band	P-band
Parameters	Platform Flight Height	6069 m	6069 m
	Frequency	1325 MHz	435 MHz
	Polarization	VV, VH/HV, VV	VV, VH/HV, VV
	Bandwidth	150 MHz	50 MHz
	Range Resolution	1.92 m	3.84 m
	Azimuth Resolution	0.65 m	2.0 m
	Incident Angles	25-45°	25-55°

2.4.   Lidar Acquisition

NASA collected Lidar data using the Land, Vegetation, and Ice Sensor (LVIS) during the AfriSAR campaign. This lidar system was aimed at supporting the ESA's BIOMASS mission. The LVIS was mounted on the NASA Langley B200 aircraft and operated at an altitude of 7,315m (Pourshamsi et al., 2018). It produced 20 m wide data footprints (Table 2). The generated LVIS data products included Level 1B, which contains geolocated return energy waveforms, and Level 2, which provides elevation data for both the ground and canopy top, along with Relative Height (RH) measurements. LVIS height metrics are made publicly available by NASA and can be accessed at lvis.gsfc.nasa.gov. The standard data products including height metrics relative to the surface for points where 25%, 50%, 75%, 98%, and 100% of the waveform energy are detected, labeled as follows RH25, RH50, RH75, RH98, and RH100 (Lee et al., 2011). The 98% RH metric was used for its accuracy, as it showed similar ground and forest heights compared to small-footprint LiDAR (Schlund et al., 2019).

Table 2 Characteristics of the LVIS Sensor

Parameter	Specification
Looking Direction	Nadir
Nominal Flight Altitude:	7,315 m
Incidence Angle	0° to 6°
Resolution	18 × 25 m
Nominal LVIS Swath Width	1.5 km (200 mrad)
Nominal LVIS Footprint Diameter:	18 m (2.5 mrad)

3.   METHODOLOGY

In this study, HH, HV, and VV polarimetric backscatter images of L- and P-bands were used to estimate the CH of the Lope and Mabounie forest sites involving various ML algorithms. This study employed an FCN algorithm based on customized UNet architecture, specifically customized for SAR backscatter imageries to estimate CH. The study investigated the potential of L-band, P-band, and combined frequency backscatters using the proposed UNet architecture. Additionally, the performance of these bands was evaluated in both heterogeneous (Lope) and homogeneous (Mabounie) forests. The performance of the proposed UNet model was then statistically compared with existing pixel-based traditional ML models, namely RF and LGBM, which were selected based on their good performance in previous forest biomass and CH estimation studies (Ge et al., 2022; Morais et al., 2021; Pourshamsi et al., 2018, 2021; Wang et al., 2022).

The overall methodology is depicted in Figure 3. Section 3.1 elaborates on the image preprocessing and experimental setup. Section 3.2 represents the basic UNet model alongside the introduction of our customized UNet framework. In Section 3.3, pixel-based decision tree ensemble algorithms have been discussed. Model evaluation metrics are detailed in Section 3.4, followed by a brief discussion of evaluation metrics and statistical testing in Sections 3.5 & 3.6 respectively.

3.1.   Image processing and experimental setup

3.1.1.   Image preprocessing

Two co-polarized (HH and VV) backscatter images, along with cross-polarized HV backscatter images of L- and P-bands, were upsampled to 20 x 20 m pixel spacing using average parameter to align with the spatial resolution of the reference CH map. This resolution was chosen considering the footprint diameter of the LVIS data (Table 2). LiDAR image of Mabounie contained missing values due to non-adjacent flight lines. For consistency in analysis, these missing regions were assigned NaN (Not a Number) values. Noise in radar images is frequent due to inherent electronic noise, environmental factors, and various forms of interference (Argenti et al., 2013). Despite speckle filtering, the images can still contain noise that acts as an outlier (Singh et al., 2021). Besides, the heterogeneous nature of the Lope study site resulted in varying backscatter values across different regions, potentially causing outliers in the data distribution. Although normalization is commonly used for feature scaling; standardization, also known as Z-scaling, was used for this study due to its robustness in handling outliers. Standardization scales features by subtracting the mean (μ) and dividing by the standard deviation (σ) (Equation 2), facilitating a broader range of values (Ahsan et al., 2021). Later, HH, HV, and VV backscatter images were stacked together for each wavelength.

$$Z = \frac{(X\prime - \mu)}{\sigma} \tag{2}$$
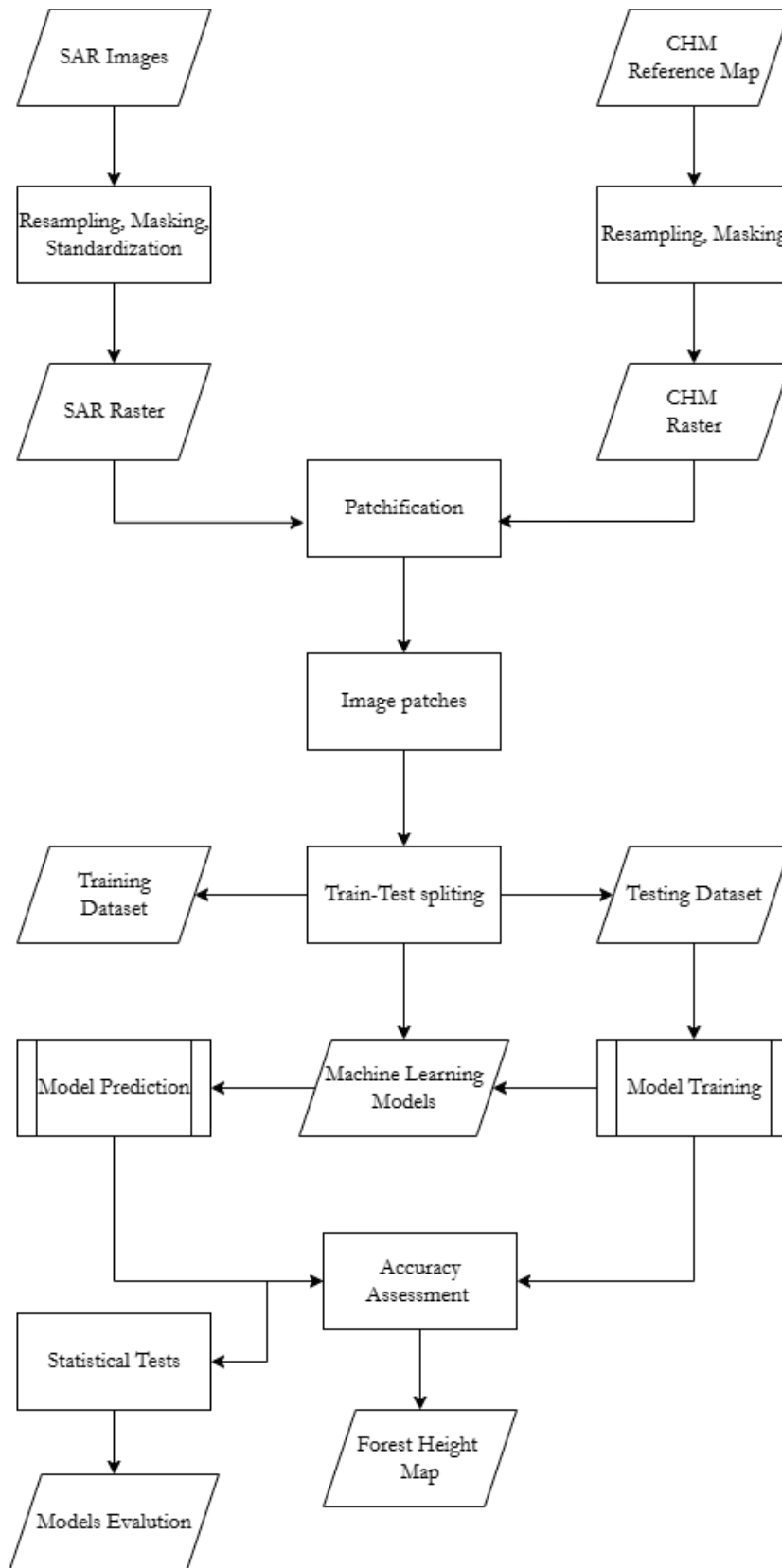
Figure 3 Workflow of the proposed method for estimating forest canopy height using polarimetric SAR backscatter images.

3.1.2. Experimental setup

Remote sensing images or tiles often contain a large array to cover areas on the ground. This makes them expensive to process with machine learning models, especially convolutional neural networks. Splitting an image into smaller patches can lower the computational cost compared to processing the whole image (Zoran & Weiss, 2011).

In deep learning applications for forestry, various patch sizes are commonly used: 256 x 256 (Li et al., 2023), 128 x 128 (Ge et al., 2022), 96 x 96 (Mahesh & Hänsch, 2023), 32x 32 (Gazzea et al., 2023), and 30 x 30 (Li et al., 2023). As indicated in the CH maps of the study area, one of our sites exhibits a leftward tilt, while another site has a substantial region of missing values. This limitation reduced the possibility of extracting large square patches. Because using larger patches would leave some parts of the study area out. Consequently, a small patch size of 32 x 32 px was selected for this study. From these, 289 non-overlapping patches were cropped from Lope and 180 from Mabounie. Since the FCN algorithm cannot process NaN values in the input SAR images, any patch containing NaN value was eliminated. After removing patches with NaN values, only 95 patches for Lope and 15 patches for Mabounie were left. The patches were then divided into a 70-30 split for training and testing, resulting in 65 patches for training 30 for testing from Lope, 12 patches for training, and 4 for testing from Mabounie (Figure 2). A validation set was not incorporated due to the limited number of training and testing images. Instead, the same testing set was used for both hyperparameter tuning and model evaluation. It should be noted that traditional data augmentation methods, such as flipping and rotation, were avoided in this study. These methods are not scientifically accepted for SAR imagery due to the significant impact of imaging direction on SAR images (Mittermayer et al., 2012). For instance, two images of the same area captured simultaneously have a higher possibility of differing from one another if one is taken from an ascending orbit and the other from a descending orbit. Consequently, simply rotating a target by 180 or 90 degrees does not yield a realistic representation of the original image.

The dataset dimensions for the UNet model are represented as (N×H×W×C). Here, N is the number of patches, H and W denote the height and width of each patch respectively, and C represents the number of channels. For wavelength comparison in the UNet model, the value of C was 3 for single-band data and 6 for dual-band data. Both the height and width of each patch (H and W) were set to 32.

Figure 4 Sampling design on Reference CH map of (a)Lope & (b) Mabounie

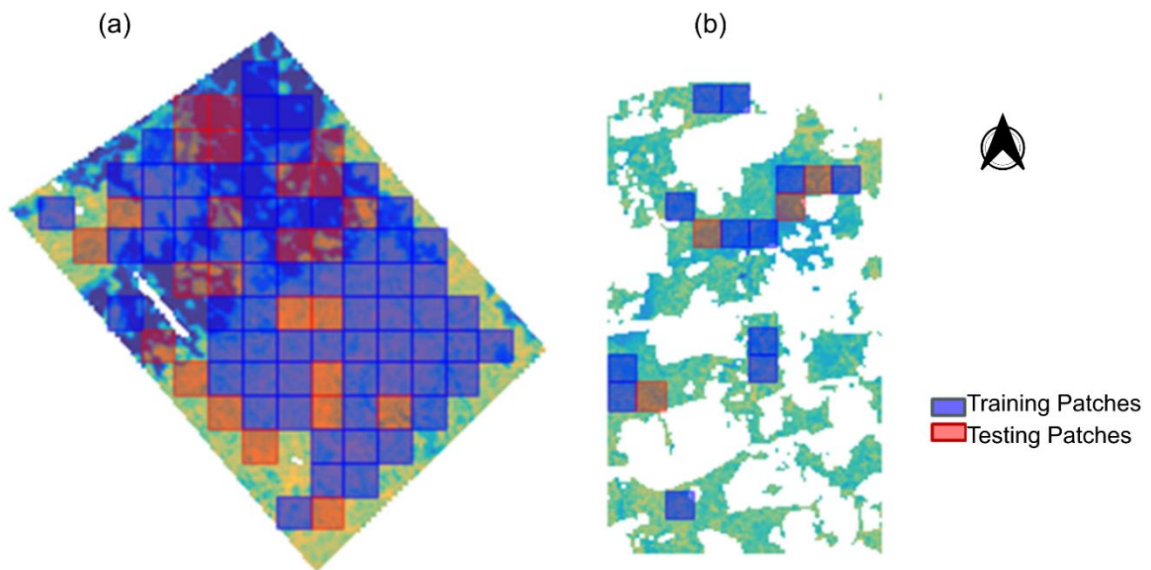For the pixel-based RF and LGBM models, eight GLCM features from HV polarization for each band were extracted using a 3x3 filter. This was done to introduce spatial context to these RF and LGBM models, thereby making the comparison with the UNet model fairer. The eight GLCM features include mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation. The GLCM features were selected based on previous CH and biomass estimation studies, as detailed in the works of (Jackson et al., 2023; Lourenço et al., 2021; Luo et al., 2023). L- and P-band SAR backscatter data were combined with 16 GLCM features, resulting in 23 channels (C= 23). The same training and testing patches used for the UNet model were also applied to the RF and LGBM models. However, pixels from the training and testing patches underwent stratified random sampling. Samples were collected across four different height classes for Lope ([0-20], [20-40], [40-50], [50-60]) and three classes for Mabounie ([0-25], [25-40], [40-60]). These strata were adopted considering the CH distribution of Lope and Mabounie (Figure 2). The aim was to ensure representation from each height strata. From each class, a minimum of 400 points were collected. The test set was then further split into an 80-20 test-validation ratio, with the validation set used for hyperparameter tuning. For both RF and LGBM, a grid search approach was adopted for hyperparameter tuning. The hyperparameter values that were used for tuning are listed in the following Table 3. The values were chosen based on the previous studies on forest CH estimations and biomass estimations using EO data. Through this grid search, 75 and 1620 combinations of hyperparameters were fitted for RF and LGBM respectively, to find the best combination.

Table 3 Grid Search Parameters for LGBM and RF Models

Model	No. Of estimators	Max depth	Learning rate	Number of leaves	Lambda l2
LGBM	[100, 200, 300, 500, 700]	[4, 8, 12]	[0.01, 0.05, 0.1]	[8, 16, 24]	[0, 0.1, 0.5, 1.0]
RF	[50,100, 200, 300, 500]	[None, 2,3,5,6]	NA	NA	NA

3.2.   UNet architecture and its components

The UNet architecture was introduced by Ronneberger et al. (2015) and is a variant of convolutional neural networks (ConvNets). It is widely used for RS image segmentation and regression tasks (Ge et al., 2022; Li et al., 2023; Pascarella et al., 2023; Zhang et al., 2023b). UNet can automatically learn and extract meaningful features from images through encoder and decoder blocks (Gazzea et al., 2023). This encoder and decoder path, along with the bridge, forms a U-shape. The convolution and max pooling blocks in the encoder and decoder path allow UNet to extract deeper features. For regression tasks, UNet is adapted by modifying the output layer to predict continuous values and using loss functions like Mean Squared Error (MSE). This adaptation makes UNet effective for tasks like depth estimation, providing detailed, pixel-wise predictions while preserving spatial information (Mahesh & Hänsch, 2023).

The encoder captures the context of the image by reducing its spatial dimensions and increasing the number of feature channels (Ge et al., 2022). Specifically, each block in the contracting path performs two 3x3 convolutions (unpadded), each followed by a ReLU activation and a 2x2 max pooling operation with a stride of 2 for downsampling (Ronneberger et al., 2015).

The decoder reconstructs the image's spatial dimensions while integrating high-resolution features from the encoder (Wang et al., 2023). It upsamples feature maps and uses a 2x2 convolution to halve the feature channels. Each step includes upsampling, concatenation with encoder features via skip connections, and two 3x3 convolutions with ReLU activation (Gazzea et al., 2023). Skip connections enhance upsampling by incorporating high-resolution details, improving segmentation accuracy. Skip connections in UNet transfer feature maps from the contracting to the expansive path, combining coarse, high-level information with fine, detailed data (Ronneberger et al., 2015). This maintains spatial information and

helps prevent the vanishing gradient problem by providing a direct path for gradient flow, ultimately improving pixel-level prediction accuracy (Ge et al., 2022).
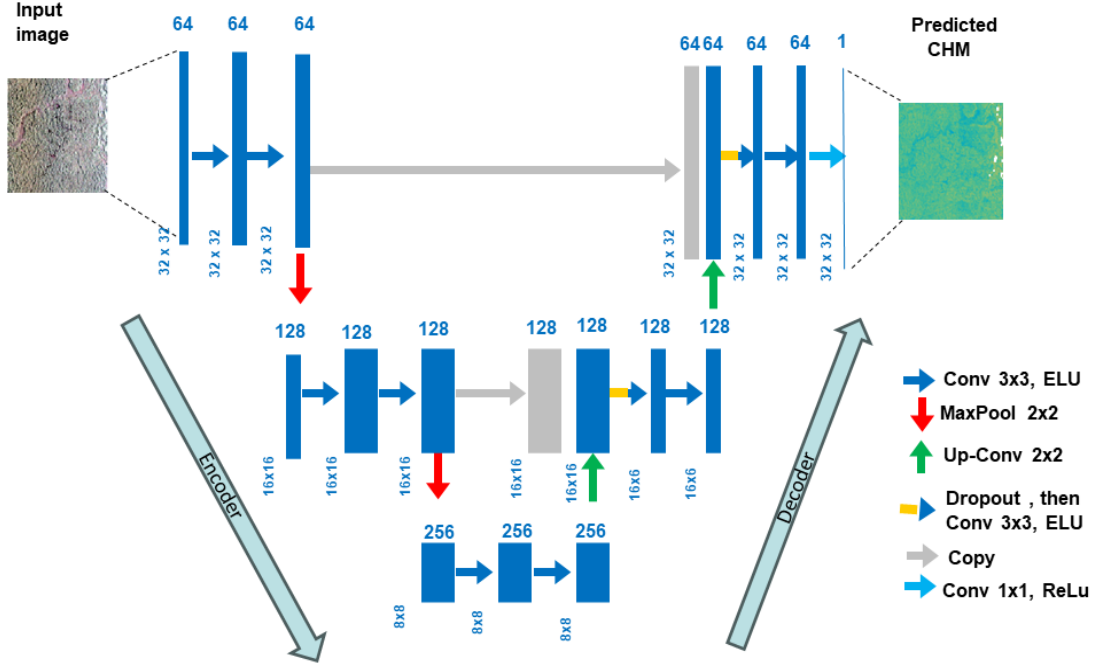
Figure 5 UNet architecture used for CH estimation

3.3.   UNet Modifications:

The base UNet model has performed well on both optical and SAR datasets (Gazzea et al., 2023; Ge et al., 2022; Mahesh & Hänsch, 2023; Wang et al., 2023). However, the unique characteristics of SAR backscatter images and the specific requirements of forest CH regression tasks necessitate further modifications to the UNet architecture (Figure 5). These adjustments include changes to activation and loss functions to handle the complexity and speckles of SAR data, as well as the integration of specific regularization and dropout techniques to improve regression performance.

The convolution block in both the encoder and decoder includes an activation function that helps analyze non-linearities, thus enabling the network to learn complex features. Nevertheless, the choice of activation function depends largely on the model input and objective, as it significantly affects the optimization, convergence, and overall performance of the model (Clevert et al., 2015). For this study, Exponential Linear Unit (ELU) was employed as the activation function, given the logarithmic negative values of SAR backscatter (dB). Using the standard ReLU, negative values scaled to zero, resulting in information loss. In contrast, the ELU activation function preserves negative values, improving learning efficiency and convergence. This helps mitigate the vanishing gradient problem and leads to faster and more accurate network training (Clevert et al., 2015). Additionally, the output activation function is set to ReLu to avoid negative forest height prediction as tree CH cannot be a negative value.

$$ELU(x) = \begin{cases} x & \text{for } x \geq 0 \\ \alpha(e^x - 1) & \text{for } x < 0 \end{cases} \tag{3}$$

With an input patch size of 32x32 and standard four layers, the UNet architecture first processes the input through several convolutional layers, gradually reducing the spatial dimensions via max-pooling operations

in the encoder path. This results in progressively deeper layers with reduced patch sizes: 16x16, 8x8, and 2x2. The bottleneck layer operates on the smallest patch size of 2x2, which can lead to artifacts because the patch size becomes very small at this stage, resulting in the loss of fine details and potential inaccuracies during reconstruction. To avoid this, the number of layers was reduced to 3, resulting in a minimum patch size of 8x8 (Figure 4). While reducing the number of layers decreases the number of parameters, this was mitigated by increasing the number of filters in each layer. The first layer had 64 filters, the second had 128, and the last layer had 256. These adjustments led to a total of 5,086,337 parameters, of which 5,081,729 are trainable and 4,608 are non-trainable. Furthermore, a dropout of 0.2 with L2 regularization was employed to reduce the overfitting of the models. The Adam optimizer was used with an initial learning rate of 0.0001 and a weight decay of 0.000001. The model was trained for 500 epochs, with a batch size of eight to achieve the results. The training and validation curves were found to be converged after 100 epochs (Appendix 1) for both sites and all the wavelengths. Although the loss function is not an integral part of the UNet model, it is crucial for convolutional network optimization. The ConvNets models adjust weight and biases depending on the loss. As mentioned previously, the reference CH map contains NaN values that impact gradient calculation. Also, SAR backscatters contain noises. To tackle this, the Huber loss function is employed to handle outliers and provide a robust measure of error (Clark et al., 2023). It computes the absolute difference between the true and predicted values and applies a threshold (δ) to differentiate between small and large errors. For small errors, it uses the Mean Squared Error (MSE), and for large errors, it uses a Mean Absolute Error (MAE) (equation 4). Additionally, the custom Huber loss function was designed to replace NaN values in the true and predicted outputs with zeros to avoid computational issues with the NumPy package.

$$L_\delta(y, \hat{y}) = \begin{cases} \frac{1}{2}(y - \hat{y})^2 & \text{for } |y-\hat{y}| \leq \delta \\ \delta \cdot |y - \hat{y}| - \frac{1}{2}\delta^2 & \text{for } |y-\hat{y}| > \delta \end{cases} \qquad (5)$$

Where:
- y is the true value.
- $\hat{y}$ is the predicted value.
- δ is the threshold parameter

3.4. Decision Tree Ensemble

Decision tree ensemble algorithms combine multiple decision trees to improve prediction accuracy (Pandey,2023). Each tree acts as an "if-then" rule system, splitting data based on input features until it reaches a prediction (leaf) (Saini & Ghosh, 2017). The ensemble technique uses voting techniques for the final prediction (Figure 4). ML models such as RF and LGBM are widely used in EO studies, as they analyze on pixel level (Ramachandran & Dikshit, 2022.).

3.4.1. Random Forest

RF is an ensemble learning technique that aggregates multiple decision trees, represented as, T {ξ1 (X), …, ξT (X)} where {X, Y} = {(x1, y1), …, (x$_n$, y$_n$)} denotes the collection of n training samples (Pourshamsi et al., 2021). Here, X signifies the D-dimensional feature vectors derived from HH, HV, and VV backscatter data of the L- and P-band. Hyperparameter tuning through grid search identified the best configurations as follows: for the Lope dataset, an unrestricted maximum tree depth with 500 trees provided the best results, capturing complex patterns effectively. For the Mabouinie dataset, a maximum depth of 6 with 500 trees prevented overfitting, ensuring the model remained generalizable. These settings optimize the model's performance by balancing flexibility and generalization.

3.4.2.   Light Gradient Boosting Machine (LGBM)

LGBM is an advanced ensemble-based decision tree method. Unlike traditional Gradient Boosting Machines (GBM), LightGBM uses a leaf-wise growth strategy rather than a level-wise approach (Tamiminia et al., 2021). This makes it faster and more efficient, especially with large datasets. In LGBM, trees are built sequentially, with each tree correcting the errors of its predecessors. The prediction $\hat{y}$ for an observation is updated iteratively (equation 6) (Pourshamsi et al., 2018):

$$\hat{y}(m) = \hat{y}(m-1) + v \cdot hm(x) \tag{7}$$

where $\hat{y}$ (m) is the updated prediction after the m-th iteration, $\hat{y}$ (m−1) is the prediction from the previous iteration, v is the learning rate, and hm(x) is the m-th tree's prediction of the residual.

LGBM is a leaf-wise gradient boosting method that enhances accuracy by splitting the tree based on the leaf with the highest loss reduction (Tamiminia et al., 2021). This approach can significantly improve model accuracy but also carries a risk of overfitting, necessitating careful tuning of parameters. After testing various hyperparameter settings, the optimal configurations were identified for two datasets. For the Lope dataset, the best parameters are regularization term (lambda_l2) of 1.0, learning rate of 0.1, maximum tree depth of 4, 100 estimators, and 8 leaves per tree. For the Mabounie dataset, the optimal parameters are the same regularization term (lambda_l2) of 1.0, lower learning rate of 0.05, maximum tree depth of 4, 100 estimators, and 8 leaves per tree. These settings balance the trade-off between accuracy and overfitting, tailored to the specific characteristics of each dataset.

3.5.   Evaluation metrics

Several metrics are used to assess model performance. R^2 (Coefficient of Determination) is commonly used and calculated as:

$$R^2 = 1 - \frac{\sum_{n=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{n=1}^{n}(y_i - \bar{y})^2} \tag{8}$$

In addition to R2, RMSE (Root Mean Squared Error) measures the average deviation of predicted values from true values:

$$RMSE = \sqrt{\frac{1}{n}\sum_{n-1}^{n}(y_i - \hat{y}_i)^2} \tag{9}$$

Percentage RMSE expresses RMSE as a percentage of the mean of the true values:

$$Percentage\ RMSE = \left(\frac{RMSE}{\bar{y}} \ x \ 100\right) \tag{10}$$

MAE (Mean Absolute Error) measures the average magnitude of the errors in predictions, without considering their direction:

$$MAE = \frac{1}{n}\sum_{n-1}^{n}|y_i - \hat{y}_i| \tag{11}$$

MAE provides a clearer picture of the average error magnitude, which can be easier to interpret compared to RMSE, particularly when dealing with outliers. Bias quantifies systematic errors in predictions, indicating potential underestimation or overestimation issues:

$$Bias = \frac{1}{n}\sum_{n-1}^{n}(y_i - \hat{y}_i) \tag{12}$$

A positive bias suggests underestimation, while a negative bias indicates overestimation. Where y_i are the true values $\hat{y}_i$ are the predicted values, and $\bar{y}$ is the mean. These metrics collectively provide a comprehensive evaluation of a model's predictive performance. While R^2 assesses how well the model predicts compared to the mean, RMSE and Bias offer insights into the accuracy and systematic errors of predictions respectively.

3.6. Statistical Analysis

Each model was iteratively run 15 times and the RMSE was recorded to compare model performances statistically. The recorded RMSE values of all the models can be found in (Appendix 2). One-way ANOVA (Analysis of Variance) on the RMSE values of the models was considered to test the two hypotheses in this study. However, one-way ANOVA can only inform about significant differences between the groups and cannot specify which group is significantly different. Therefore, post-hoc T-tests with Bonferroni Correction of significance level were employed to determine where the group differences lie. In the first hypothesis, the performance of the L-band, P-band, and dual-band (L- and P-band) on the UNet model was tested. In the second hypothesis, the prediction performance of UNet, RF, and LGBM on Dual-band PolSAR backscatter was compared. For the ANOVA test, the threshold value for p was set to 0.05. Then for post-hoc T-tests, the Bonferroni Correction adjusted the significance level to $\alpha = \frac{0.05}{k} = 0.02$, where k =3 is the number of comparisons made.

4. RESULTS

### 4.1. Forest CH retrieval using UNet on the Lope site

For the Lope study area, the trained UNet models for L-, P-, and combined (L- + P-) wavelengths were individually applied to the test data set, and estimated CH values were compared to reference CH values extracted from the LIDAR data. Figure 6 illustrates the model evaluation box and confusion plots for different wavelength predictions. The overall evaluation metrics of the models are reported in Table 4. For the L-band, the average error calculated in the predicted CH was 5.02 m (MAE). The R^2, RMSE, and RMSE% were 0.85, 6.45 m, and 25.52% respectively. CH values between 10-20 m had a noticeable variance, mostly overestimated (Figure 6a) in the predicted CH compared to actual CH. However, for CH above 50m, the L-band-derived CH underestimates the actual height with an overall positive bias of 12.5 m (Figure 6a). The overall L-band estimations demonstrated a high positive bias or underestimation of 3.80 m. The overall bias was found to be slightly reduced from 3.80 m to 2.95 m for the P-band compared to the L-band. The underestimation (3 m) for low vegetation was observed for P-band (Figure 6c), with the overall MAE standing at 4.09 m. The improved R^2 value indicated that the P-band enabled the model to explain 89% ($R^2 = 0.89$) of the variance in the CH prediction. Similarly, reduced RMSE and RMSE% values were reported at 5.39 m, and 21.34% respectively. P-band-derived estimation slightly reduced the bias for all the forest strata (Figure 6c). The dual-band combination (L- +P-) provided the best performance across all metrics compared to the individual L- and P-bands. It produced the highest R^2 value of 0.94, indicating the best fit of the model. The lowest RMSE, RMSE%, MAE, and Bias indicated the highest accuracy in predictions (Table 4). The RMSE% for the Dual-band was approximately 60% and 34% better than that of the L-band and P-band's RMSE%, respectively. The dual band reduced the overall bias. However, overestimation remained evident in 20-40 m strata. The underestimation for very high forest canopies (above 50m) was slightly reduced compared to L- and P-band (Figure 6e). Nevertheless, the minor differences in evaluation metrics required a statistical significance test for wavelength comparisons. Figures 7 (a-d) illustrate the CH maps estimated from L-band, P-band, dual-band, and LiDAR (RH98) data respectively. From visual interpretation, it is evident that the L-band tends to underestimate the forest CH, while the P-band overestimates it in some regions. To some extent, the dual-band approach balances the estimations, providing a more accurate representation.

Table 4 Performance Evaluation metrics of for L-band, P-band, and Dual-band on UNet-based Model for Lope site

Band and their combination	R²	RMSE (m)	RMSE%	MAE (m)	Bias (m)
L-	0.85	6.45	25.52	5.02	3.80
P-	0.89	5.39	21.34	4.09	2.95
Dual (L- + P-)	0.94	4.03	15.95	2.89	0.31
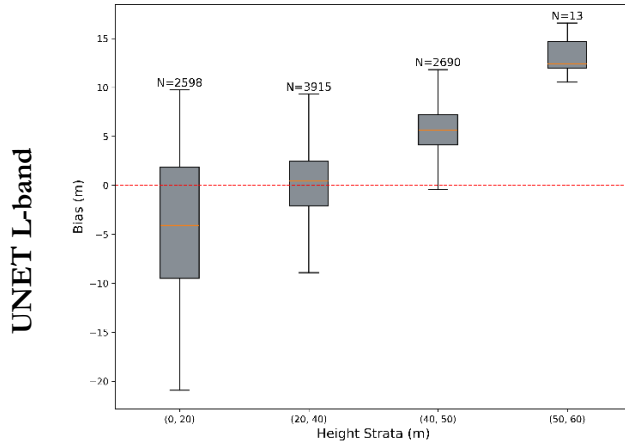
(a) CH bias of different height strata
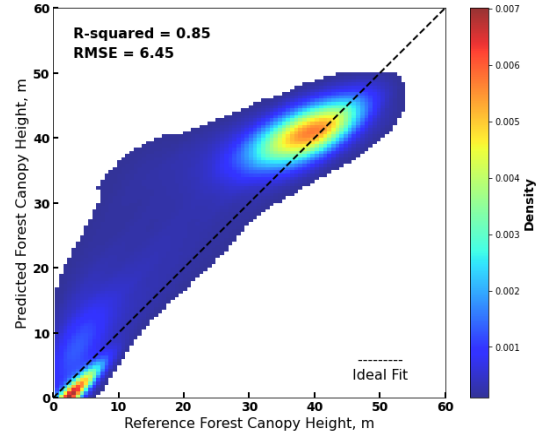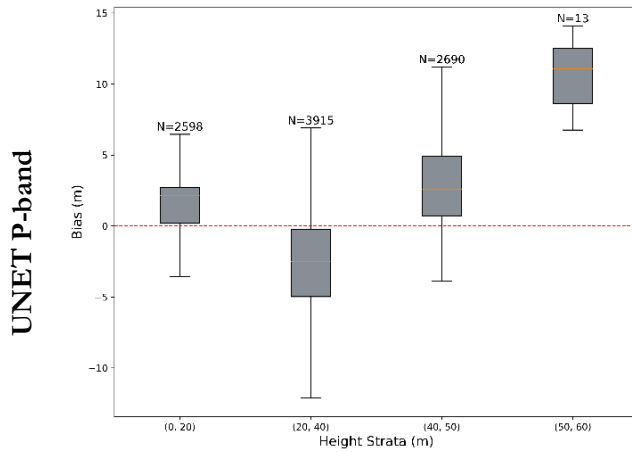
(b) Estimated vs reference CH

(c) CH bias of different height strata
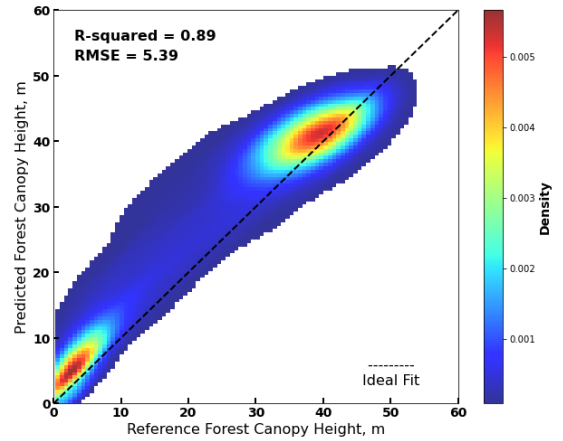
(d) Estimated vs reference CH

(e) CH bias of different height strata

(f) Estimated vs reference CH

Figure 6 **Model Evaluation for Lope site**- The box plots on the left illustrates the mean and standard deviation of the bias derived from the UNet-based CH prediction across different height strata. The confusion plots on the right compare UNet predictions to the LiDAR (RH98) map for test datasets from the same site, with a density colormap showing the observation distribution.
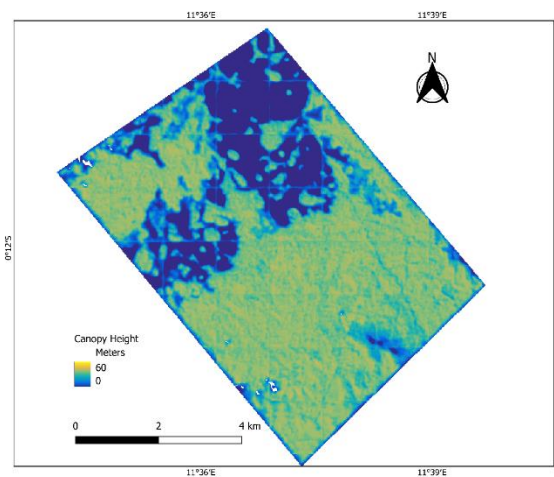
(a) CH map estimated from L-band
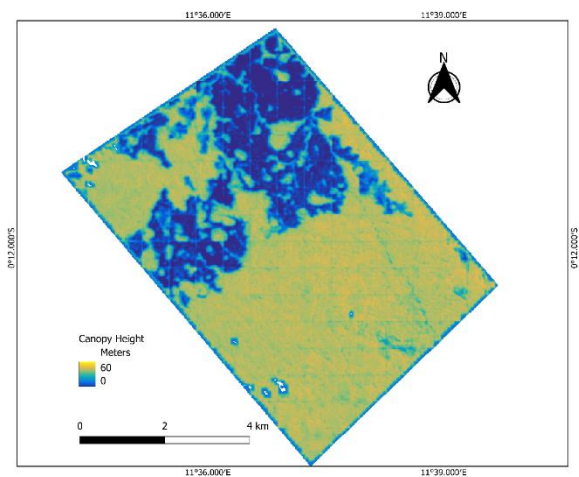
(b) CH map estimated from P-band
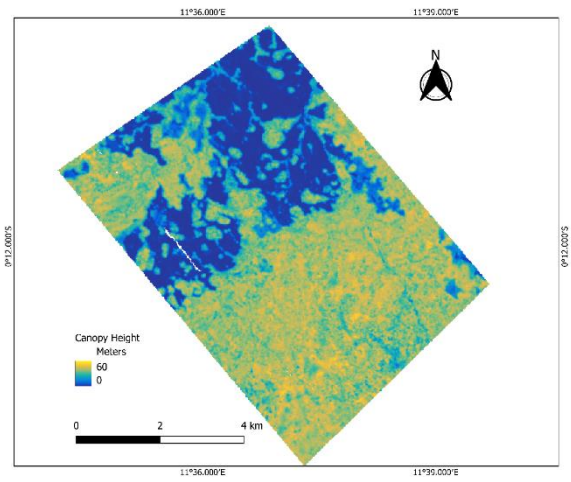
(c) CH map estimated from Dual-band   **(d) Reference CH map from LidAR (RH98)**

Figure 7 UNet Estimated CH Maps of the Lope Site Using Different Wavelengths

4.2.   Forest canopy height retrieval using UNet on Mabounie

The UNet was then tested on the Mabounie site using SAR L-, P-, and Dual-band. The performance metrics are reported in Table 5. The strata-wise bias plots and confusion plots for each band are shown in Figure 8. R^2 values indicated moderate model performances for Mabounie. However, the RMSE and RMSE% metrics suggested, models produced similar CH estimation accuracies as Lope. The Dual (L+P) band combination achieved the best performance, with the highest R^2 of 0.60, the lowest RMSE of 3.78 m and RMSE% of 11.89%, the lowest MAE of 1.69 m, and a small negative Bias of -0.38 m. The P-band also performed well, with an R^2 of 0.55 and an RMSE of 4.26 m, while the L-band had the lowest R^2 of 0.47 and the highest RMSE of 4.63 m. The L-band bias plot (Figure 8a) demonstrated the model's underestimation tendency for CH between 0-25 m. Additionally, for CH values above 30 m (Figure 8b), the model exhibited saturation with a mean positive bias of 15 m (8a). In contrast, the P-band showed a reduced underestimation compared to L-band (Figure 8e). The model prediction performance improved with the band combination (L-and P-) for 25–50 m strata. Further, the underestimation of tall CH (above 50 m) was minimized to some extent (Figure 8e). The estimated CH maps (Figure 9) illustrate the prediction differences between bands. The yellow forest CH areas of the reference map were predicted as Low strata (light blue) for the L-band. However, the estimation improved with P- and dual-band.

Table 5 Performance Evaluation metrics of for L-band, P-band, and Dual-band on UNet-based Model for Mabounie site

| Band combination	R^2	RMSE (m)	RMSE%	MAE (m)	Bias (m)
L-	0.47	4.63	14.73	1.95	1.35
P-	0.55	4.26	13.55	1.79	0.11
Dual (L+P)	0.60	3.78	11.89	1.69	-0.38
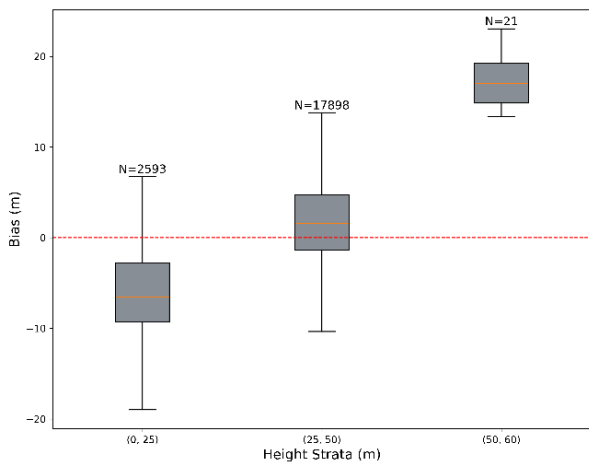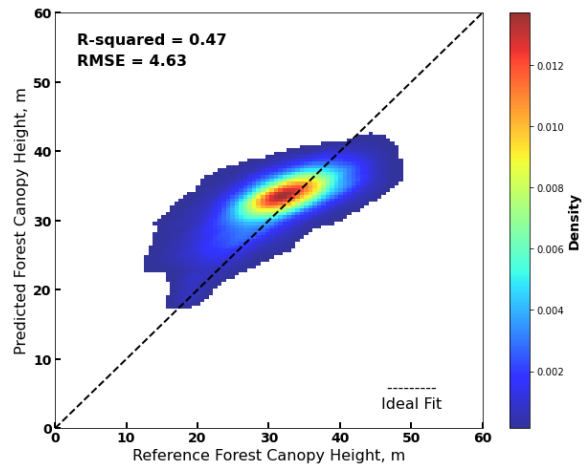
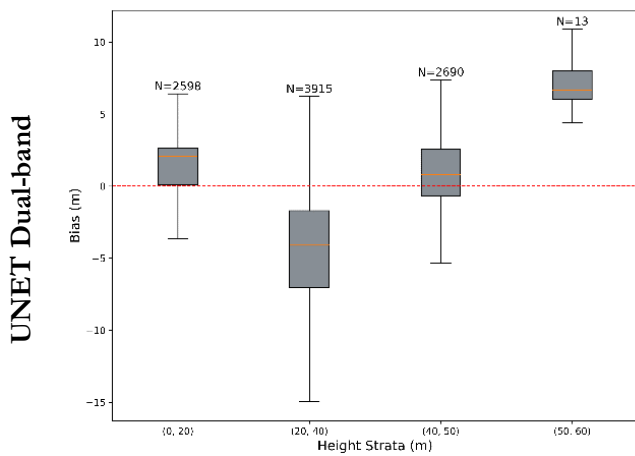**(a) CH bias of different height strata**
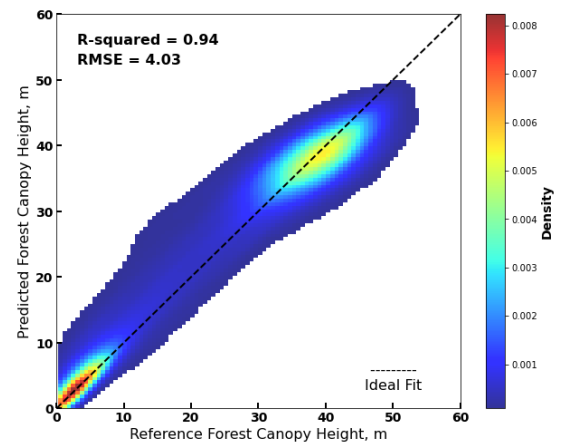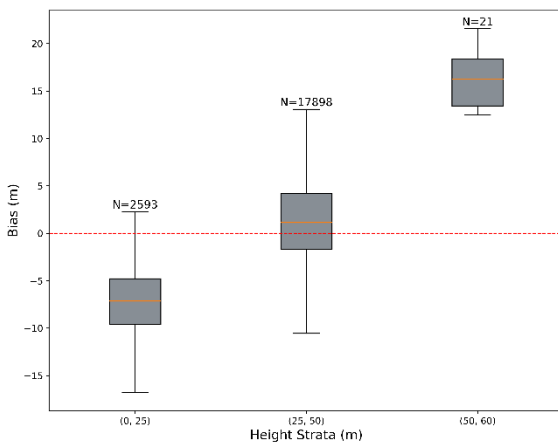
(b) Estimated vs reference CH

**(c) CH bias of different height strata**
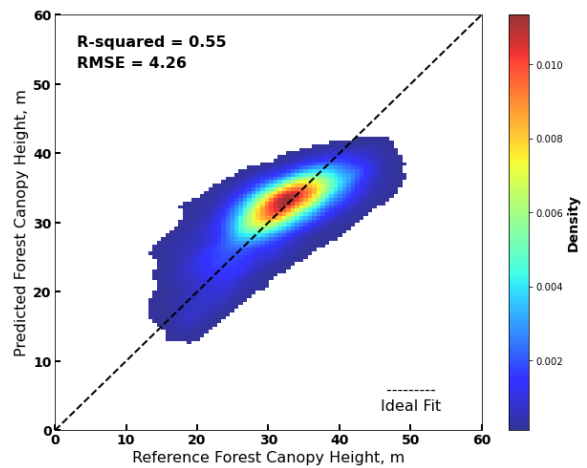
(d) Estimated vs reference CH
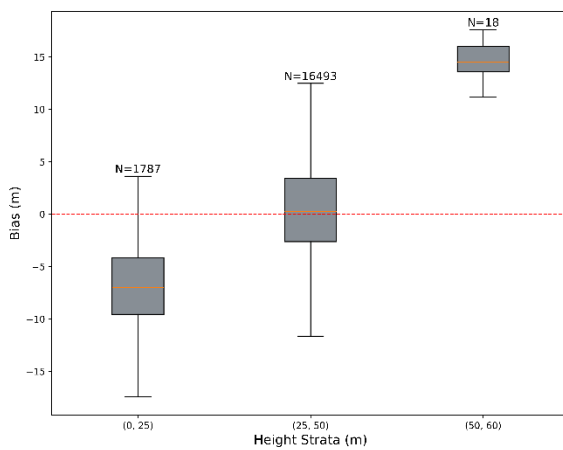
(e) CH bias of different height strata
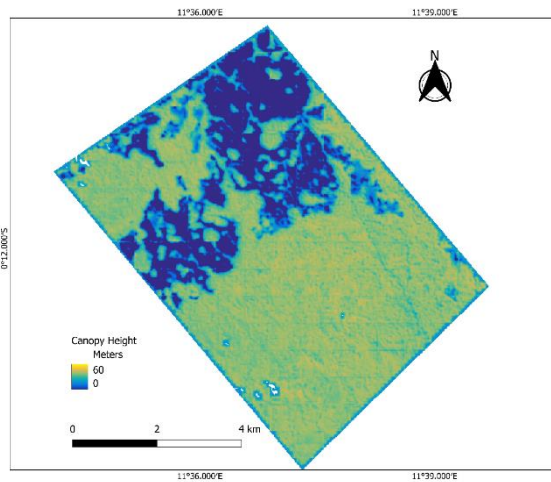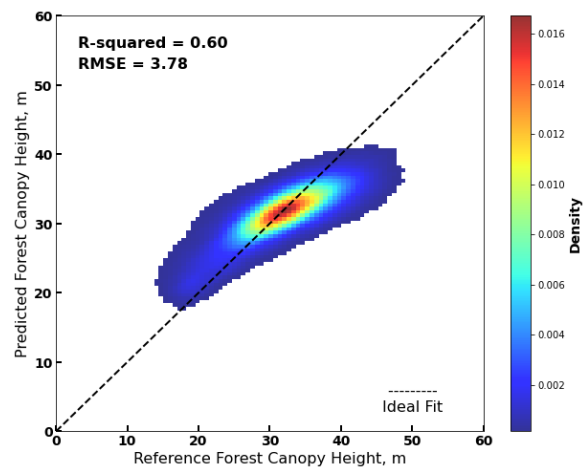
(f) Estimated vs reference CH

Figure 8 **Model Evaluation for Mabounie site**- The box plots on the left illustrates the mean and standard deviation of the bias derived from the UNet-based CH prediction across different height strata. The confusion plots on the right compare UNet predictions to the LiDAR (RH98) map for test datasets from the same site, with a density colormap showing the observation distribution.
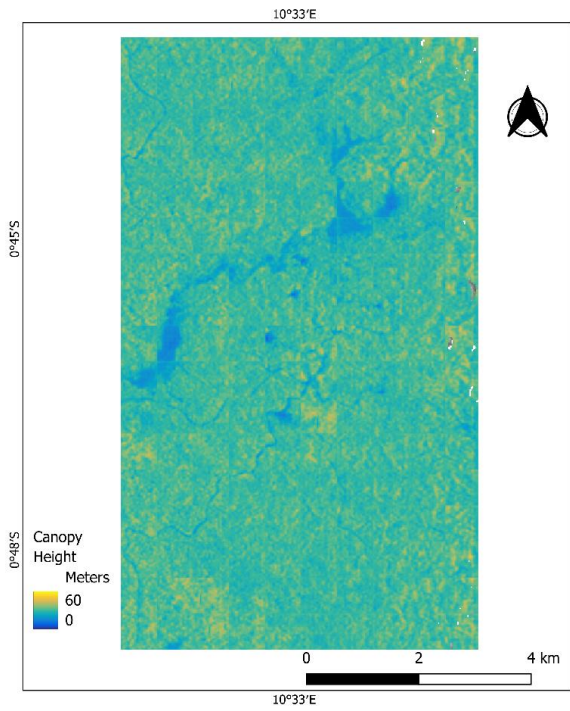
(a) CH map estimated from L-band
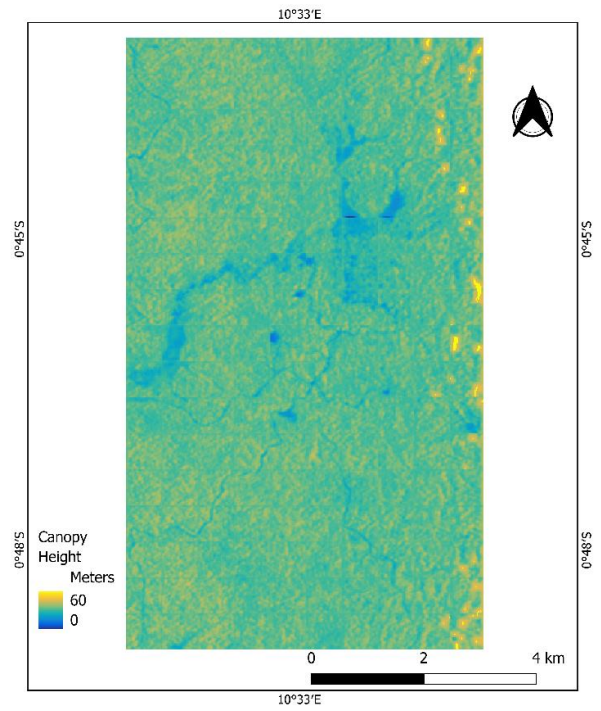
(b) CH map estimated from P-band
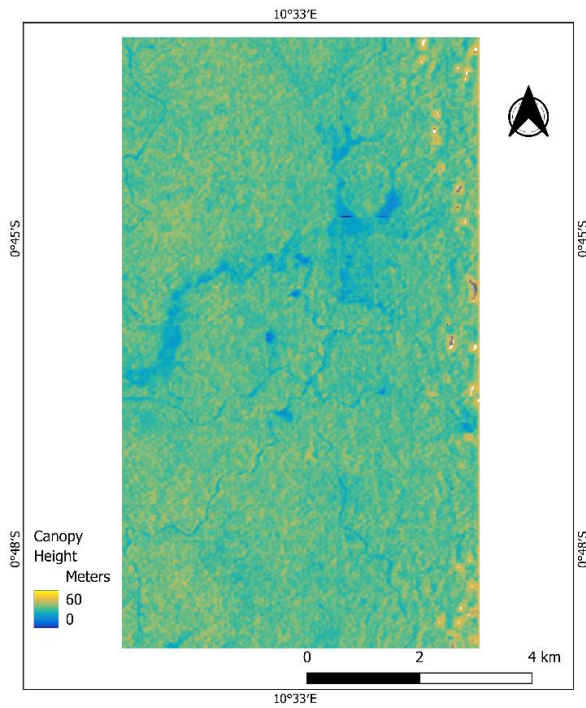
(c) CH map estimated from Dual band
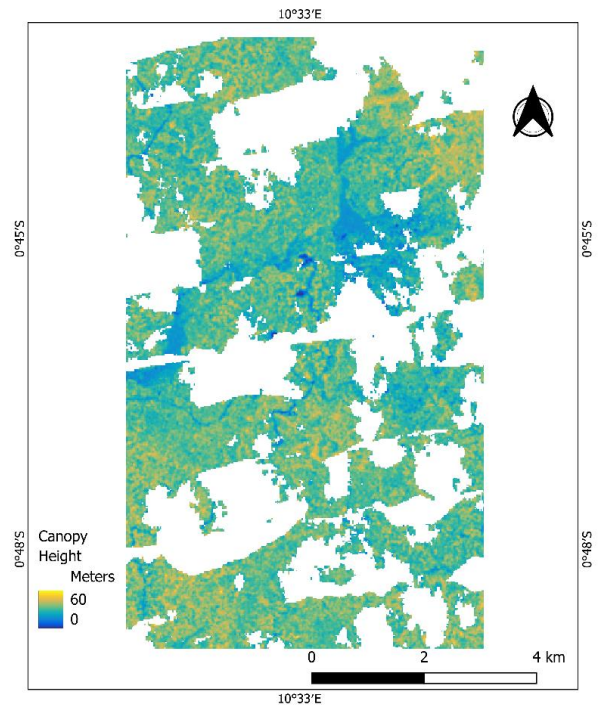
(d) Reference CH map from LiDAR (RH98)

Figure 9 UNet Estimated CH Maps of the Mabounie Site Using Different Wavelengths

4.3. Statistical significance test on wavelength prediction performances

A one-way ANOVA was used to analyze the RMSE values to compare the prediction performance of L-, P-, and combined SAR wavelengths on the UNet models (Chen et al., 2016) for both Lope and Mabounie sites. For the Lope site, the ANOVA test revealed that the F-statistic (45.06) was greater than the F critical value (3.22), thereby indicating the test's significance. In addition, the P-value (0.0035×10^{-8}) also rejected the null hypothesis and suggested that a statistically significant difference was present among the wavelengths' CH estimation performances. However, further evaluation was required to determine which wavelength's performance was statistically significant. Therefore, separate T-tests were performed with an adjusted α of 0.02 (Bonferroni correction). The post-hoc T-tests revealed that the RMSE of dual bands (4.25 ± 0.26) was significantly smaller compared to the RMSE of the P-band (5.64 ± 0.49, $P < 0.02$) and the L-band (6.50 ± 0.98, $P < 0.02$) and (Appendix 3a). Notably, this study found no significant difference between L-band and P-band estimation performances.

For Mabounie, the test results revealed that like Lope, Dual-band estimations produced better accuracy than individual bands. However, as shown in Appendix (3b), the CH estimation capabilities of individual L- and P-bands (P-value <0.02) were significantly different for this forest site.

4.4. Forest canopy height estimation for RF and LGBM on Lope

The confusion plots for RF and LGBM are presented in Figure 10. Overall, the accuracy of the estimated heights is moderate, with R^2 values of 0.72 and 0.75 and RMSE values of 10.04m and 9.52m for RF and LGBM, respectively (Table 6). The two algorithms performed quite similarly, with LGBM showing a slightly better RMSE than RF. Both models appeared to overcome underestimations of tall forest heights slightly (>50 m) (Figure 10). Additionally, LGBM had a lower MAE of 6.95 m compared to RF's 7.41 m, and both models had similar Bias values, with LGBM at -0.61 m and RF at -0.62 m. At Lope, it was evident from the estimated maps that both RF and LGBM overestimated some low CH regions compared to the reference Figure 10.

Table 6 Evaluation metrics for RF and LGBM for Dual-Band based CH Estimation of Lope site.

Models	R^2	RMSE (m)	RMSE%	MAE (m)	Bias (m)
RF	0.72	10.04	35.36	7.41	-0.62
LGBM	0.75	9.52	34.17	6.95	-0.61
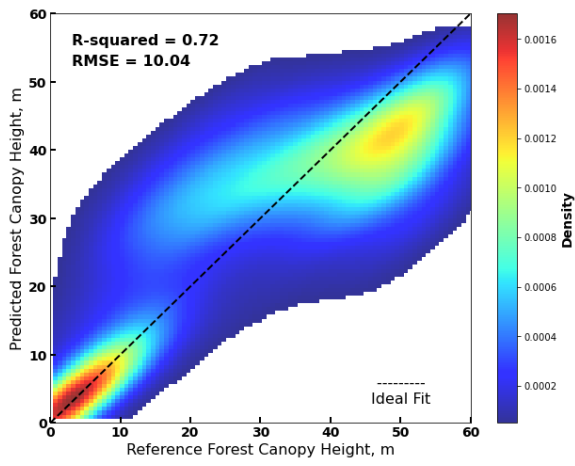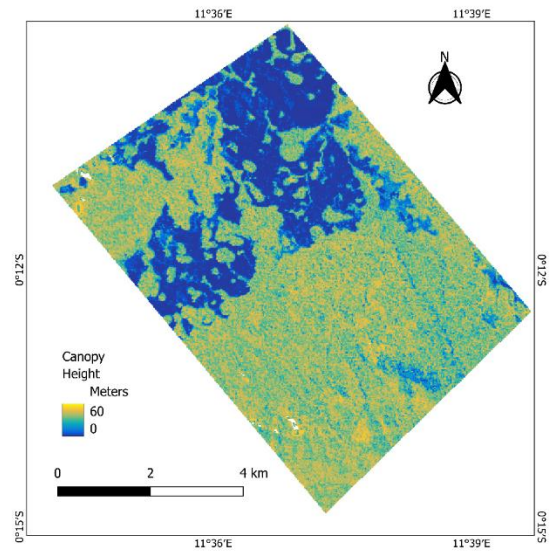
(a) Estimated vs reference CH
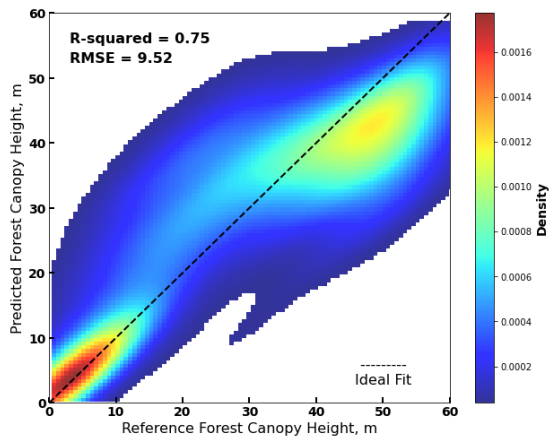
(b) Predicted CH map

(c) Estimated vs reference CH
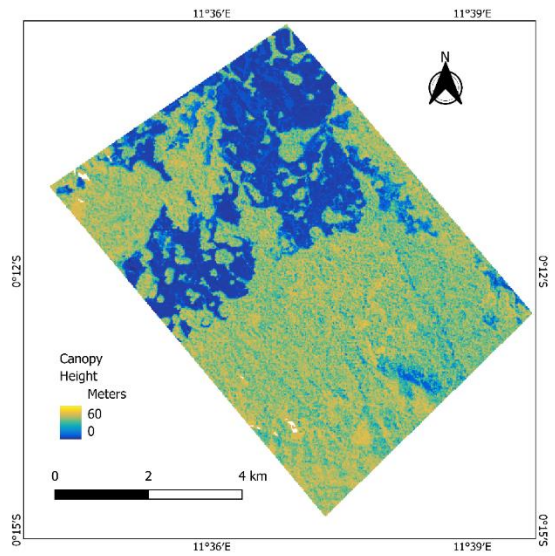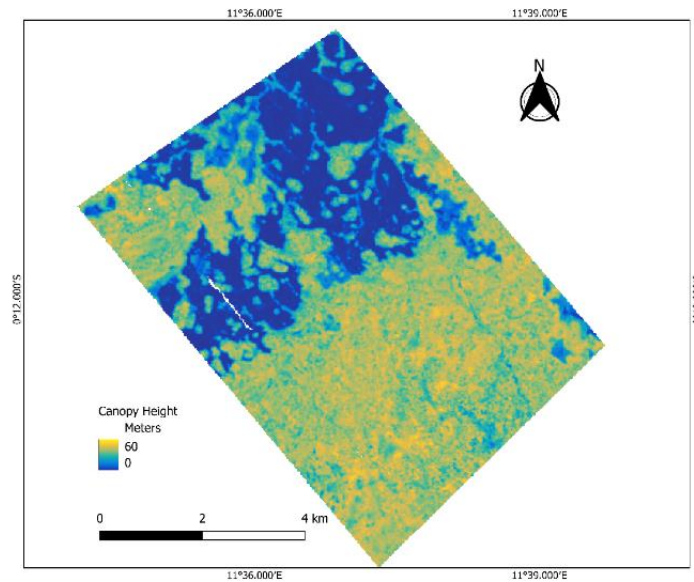
(d) Predicted CH map

Figure 10 The confusion plots on the left compare RF and LGBM (CH) predictions to the LiDAR (RH98) map for test datasets from the Lope site, with a density colormap showing the observation distribution. On the right are the estimated CH maps using RF and LGBM models and reference CH map from LiDAR RH98.

4.5.   Forest Canopy Height Estimation for RF and LGBM on Mabounie

The performance metrics for RF and LGBM algorithms on the Mabounie site are summarized in Table 7. with the corresponding confusion plots shown in the accompanying Figure 11. The right side of Figure 11 illustrates the map of estimated tree heights from the RF and LGBM algorithms. The R^2 values of 0.12 for RF and 0.15 for LGBM indicated poor performance in explaining the variance in the data (Table 9). However, the RMSE, RMSE%, MAE, and bias values suggested a similar performance of the models to those observed at the Lope site for RF and LGBM respectively. The low R^2 values resulted from the homogeneity of the CH distribution in the Mabounie forest site. As previously mentioned, most trees' heights ranged close to the forest's mean height, reducing the variability the models could capture. Notably, both models were inclined to highly saturate CH predictions, which led to underestimation for tall trees and overestimation for short trees. The overall prediction performance of the RF and LGBM models was similar to each other. This is also observed in the estimated CH map of both models in Figure (11).

Table 7 Evaluation metrics for RF and LGBM for Dual-Band based CH Estimation of Mabounie site

Models	R^2	RMSE (m)	RMSE%	MAE (m)	Bias (m)
RF	0.12	9.09	28.65	7.80	0.26
LGBM	0.15	8.94	28.19	6.95	0.14
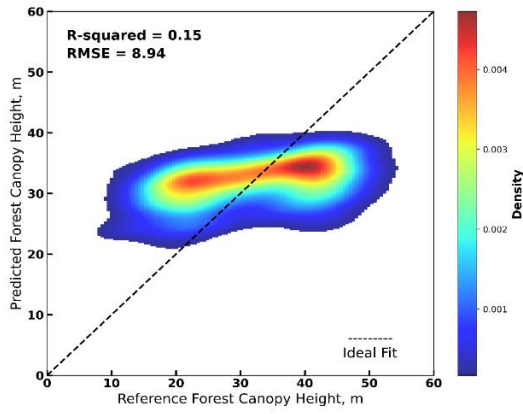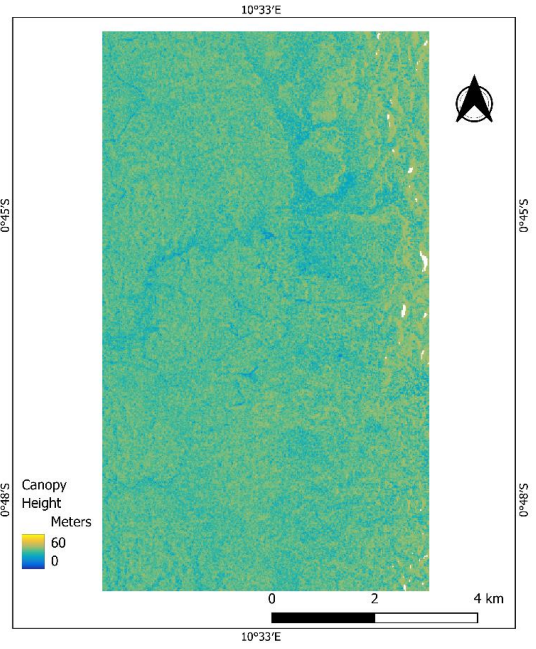
(a) Estimated vs reference CH
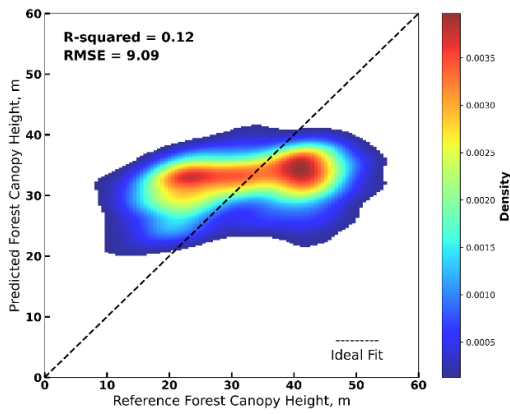
(b) Predicted CH map

(c) Estimated vs reference CH
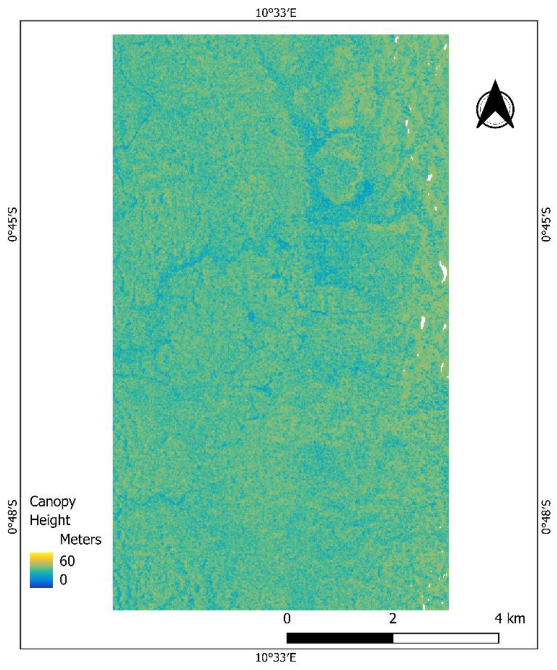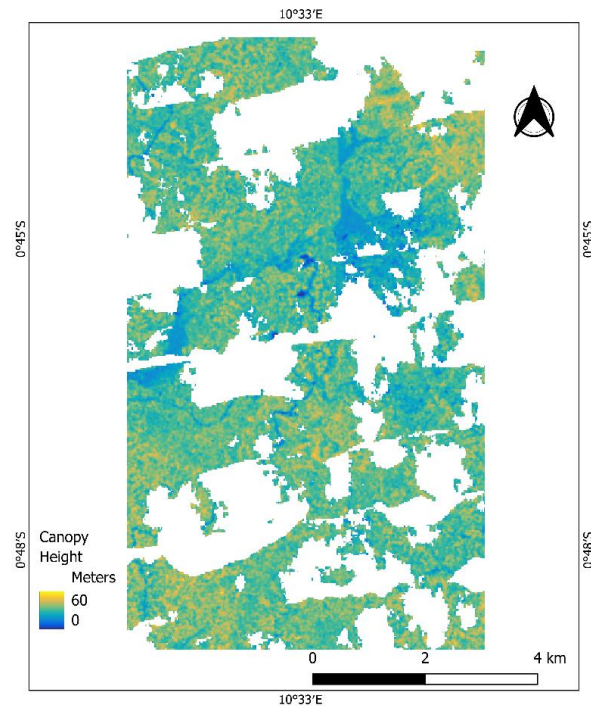
(d) Predicted CH map

Figure 11 The confusion plots on the left compare RF and LGBM (CH) predictions to the LiDAR (RH98) map for test datasets from the Mabounie site, with a density colormap showing the observation distribution. On the right are the estimated CH maps using RF and LGBM models.

4.6. Statistical significance test of model's prediction performance for Dual-band on Lope & Mabounie

The Lope site's performance metrics (Tables 4 & 6) and confusion plots (Figures 6f and 10(a & c)) indicated that the UNet model provided more accurate predictions than the baseline RF and LGBM models. The UNet model achieved a 34-35% improvement in RMSE% compared to the pixel-based ML models. The UNet-based model generated a much higher R^2 (0.94) value than both RF (0.72) and LGBM (0.75). In addition, the confusion plots illustrated that the UNet model had less prediction variability and uncertainty than RF and LGBM. However, to make a statistical comparative assessment of model performance, the RMSE values of the models underwent one-way ANOVA and a post hoc T-test. The result of the statistical significance test is summarized in Appendix (3c).

The statistical significance tests revealed a significant difference in model prediction accuracy with an F value of 45.06. The pairwise T-test confirmed the UNet's superiority in prediction accuracy compared to RF and LGBM, with p-values less than 0.02 (Bonferroni adjustment). Notably, the statistical test on the RMSE from the 15 iterative model runs showed no significant difference in the prediction capabilities of the RF and LGBM models. Besides, the RF and LGBM models resulted in higher prediction variability (SD = 0.98 and 0.49), while the UNet model had the lowest prediction variability (SD = 0.27), indicating more consistent predictions. Also, the statistical significance of the test on UNet, RF, and LGBM models on the Mabounie site is summarized in Appendix (3d). The results reiterated the Lope site's statistical evaluation with an F value of 45.06, where UNet outperformed the RF and LGBM.

5. DISCUSSION

5.1. Investigation of SAR Polarization and Wavelength Performance

This study investigated the possibility of estimating forest CH using different wavelengths and combinations, on the FCN-based UNet model. The findings of this study suggest that backscatter images of different polarizations can accurately estimate forest CH using UNet models with a single baseline image. The accuracies were consistent for homogenous and heterogenous forest sites for all bands. As expected, each polarization interacts distinctively with various parts of the forest structure; therefore, combining polarizations enables accurate estimation of forest CH (Green, 1998; Srivastava et al., 2009). While HV polarization is highly sensitive to forest canopy volume (Chen et al., 2021; Liu et al., 2021), backscatter of HH and VV polarizations are frequently observed in trunks and ground surface interactions (Chen et al., 2020; Sun & Simonett, 1988). Silveira et al. (2023) also highlighted that co-polarization penetrates deeper into forest layers and provides vertical information, while cross-polarization provides mostly canopy information due to shallow penetration. However, co-polarization (HH and VV) backscatters are subject to attenuation by canopy closure and canopy volume, and their effectiveness largely depends on the wavelength's penetration ability (Townsend, 2002).

When L- and P-bands were used individually, statistical significance tests indicated that their CH estimation performance differed between the Lope and Mabounie sites. This discrepancy could result from variations in forest structure, topography, and forest height distribution. Despite having similar dominant tree species, Mabounie is a homogeneous forest with mostly mature trees of uniform heights (30-35 m). In contrast, Lope's trees exhibited a wide range of heights (section 2.2). Additionally, Mabounie has plain topography, while Lope features hilly terrain (sections 2.1 and 2.2). As mentioned earlier, the effects of ground topography were corrected. This topographic correction approach improved the model output in the study by Schlund and Davidson (2018). So, it can be assumed that the terrain correction was sufficient to reduce the terrain effect for this study. Thus, the results indicated a different sensitivity of the two wavelengths to CH distribution. Our study suggested that the P-band performed better in CH estimation at the Mabounie site than the L-band. On the contrary, the performance of the L-band and P-band did not exhibit a statistically significant difference for Lope. This was because the L-band interacts mainly with the canopy and branches, while the P-band's longer wavelength penetrates deeper, capturing ground-stem interactions (Neumann et al., 2012; Sandberg et al., 2011). However, Gazzea et al. (2023) and our study both suggest that longer wavelengths like P-band can result in a negative bias or overestimation for low-height forest strata due to the intricate signals from the understory, trunks, and ground coupling. Similar performance of the L-band at Lope was also observed in a previous study by Liu et al. (2021). Liu et al. (2021) used tomograms obtained from multi-baseline interferometric data, which is different from polarimetric backscatter-based estimation, indicating that the observed consistency in CH accuracy might result from the penetration capabilities of the wavelengths used. The CH of forest stands in Lope exhibited a bimodal distribution, with most values concentrated in the ranges of 1-8 m and 35-48 m. This contrasts with the

s in Mabounie, which followed a more normal distribution. The distribution for Mabounie showed a high kurtosis of 0.69, indicating a more concentrated distribution. Approximately 80% of the CH in Mabounie fall between 25-40 m (Figure 2). So, it could be argued that the saturation and underestimation of CH for the L-band at the Mabounie site resulted from the homogeneity in the forest stands.

The results of this study also illustrated that the combination of L- and P-band could significantly improve model performance accuracy and reduce saturation effects for both forest sites, regardless of whether the forest structure is homogeneous or heterogeneous. While the P-band may better map CH due to increased penetration, it can also generate more complex signals from sub-canopy, trunks, and ground coupling (Gazzea et al., 2023). Consequently, the combination is complementary due to the different sensitivity to different forest layers of the two bands. However, no studies have been found that combine these bands specifically for CH estimation. Schlund and Davidson (2018) reported improved model performance by combining SAR wavelengths for AGB estimation in boreal and hemi-boreal forest sites. Notably, the underestimation of CH for heights above 50 m in Lope and 40 m in Mabounie could be considered insignificant when evaluating the wavelengths or model performances, as the frequency of occurrence for this CH range is less than 1% of the data (Figure 2). Interestingly, conventional RF and LGBM ML models were able to predict taller CH due to the stratification sampling technique, which is discussed further in section 5.2.

5.2.   Comparative Assessment of Model Performances

Statistical significance tests on the prediction performances of conventional RF and LGBM models, and customized UNet models indicated that UNet models provided better accuracy. It is often claimed that FCN-based UNet models can predict with higher accuracy than pixel-based RF and LGBM models because UNet accounts for spatial context (Ge et al., 2022) However, in this study, the spatial context was additionally introduced to the RF and LGBM using GLCM features (section 3.1.2), yet they still failed to achieve the same accuracy as the UNet model. Therefore, it can be assumed that other factors influenced the model prediction accuracy. The improved performance for CH estimation of the customized UNet framework can be attributed to the potential of UNet to capture relevant features from different dimensions using a 1x1 convolution block. This assumption is supported by findings from studies such as Schlund et al. (2018) and Soja et al. (2012), which highlighted the importance of combining dimensional channels to extract features. For instance, a simple HH/VV polarization ratio can reduce the topographic effect. In addition, the selection of suitable activation functions, such as the ELU function helps minimize the information and pattern loss. Besides, reducing the number of layers based on patch size helps to prevent edge artifacts on the patches (Pascarella et al., 2023).

Although UNet produced better overall accuracy than RF and LGBM, with RMSE three times lower, it failed to predict the taller CH (>50 m in Lope and >40 m in Mabounie), which comprised less than 1% of the dataset. This limitation arose because FCN models require large datasets for effective training (Lang et al., 2023), which were not available for this study. In contrast, RF and LGBM were able to predict this smaller portion of CH due to the stratification sampling technique and their inherent rule-based methods, which enabled these models to effectively split leaves for tall CH even with fewer training samples. Furthermore, due to the lack of sufficient variability in the CH distribution, RF and LGBM showed a saturation in the homogeneous forest site of Mabounie. Pourshamsi et al., (2021) also demonstrated similar saturation in the conventional RF, CCF, and SVM models using L-band PolSAR backscatter at the Lope site, due to insufficient variability in the training samples.

It is worth discussing that the UNet model is computationally expensive to train compared to RF and LGBM (Ge et al., 2022). In this study, while RF and LGBM were trained on only 16 GLCM features, the UNet architecture extracted a total of 1216 features from the encoder and decoder blocks using the six PolSAR backscatters. Therefore, to make a fairer comparison of the models, incorporating more parameters into RF and LGBM was necessary. Specifically, including cross-channel parameters might help overcome the shortcomings of conventional methods, presenting a potential direction for future research.

At the same time, introducing the explainability of UNet model predictions is essential to improve the model's transparency and transferability.

5.3. Uncertainties Associated with Reference Data

To validate ML model outputs, it is crucial to report uncertainties in the reference map. These uncertainties can propagate through the analysis and accumulate, potentially impacting model output. Fatoyinbo et al. (2021) reported significant variability in LVIS height metrics for Lope and Mabounie. In Lope, maximum canopy heights were estimated at 75.9 m (RH 98) with individual footprints up to 88.9 m, often overestimated due to complex topography (Fatoyinbo et al., 2021). In Mabounie, heights peaked at 75.26 meters (RH98), with sub-canopy heights from 8.9 m to 13.8 m. Further comparisons with ALS data showed a positive mean bias in LVIS estimates, with RMSE up to 4.2 m in Lope. This bias increases with canopy cover and slope, complicating accurate ground elevation and CH estimation (Hancock et al., 2017). However, for this study, upon upsampling to 20 m grid spacing the highest CH was found to be 58 m for Lope and 60 m for Mabounie. Additionally, dense cloud covers of the tropics caused gaps in the lidar data, and there were also missing values due to non-adjacent flight lines (Armston et al., 2020). This resulted in a shortage of sampling patches for ML models for the Mabounie site.

5.4. Comparison to Similar Works

Few studies have explored various methods and SAR data for forest height prediction in tropical forests (Li et al., 2023; Mahesh & Hänsch, 2023; Pourshamsi et al., 2021). The findings of our study are comparable to these studies. In the tropical region, Pourshamsi et al. (2021) reported forest CH estimation accuracies with L-band PolSAR backscatter and SRTM-derived variables, achieving an RMSE of 10m at the pixel level. In contrast, our study found an RMSE of 6.45m for the same site using L-band PolSAR backscatters alone. Additionally, while their results showed high saturation, our model demonstrated better performance in reducing the saturation of CH. L-band PolInSAR data and CNN in the same study area obtained an RMSE of 10.15m (Li et al., 2023). In another study in the tropical region, Mahesh and Hansch (2023) used X-band InSAR data with a UNet architecture, achieving an RMSE ranging from 5.78 m to 8.98 m, while our model demonstrated superior performance with an RMSE of 3.78 m to 4.63 m. The improved performance may result from the longer wavelengths used in this study. Moreover, to the best of our knowledge, our customized UNet models provided better accuracy than other CH estimation studies that combined SAR and optical imagery with FCN models in boreal and temperate regions (Astola et al., 2021; Ge et al., 2022). Recently Tolan et al. (2024), Lang et al., (2023), and Potapov et al. (2021) and produced global forest height maps from optical imagery using FCN models. The estimation accuracies of their models are within the same range as ours, with mean RMSE values between 4.5 and 7.5 m. It is noteworthy that the models from these studies were trained and tested on larger datasets and have higher spatial resolutions compared to the present study, providing evidence of their robustness and precision. To fairly compare these models with our own, it is essential to provide scientific evidence that generalizes this proof of concept across different biomes and enables forest height estimation at a global level. Nevertheless, the findings of this study are promising and encourage further research using UNet models and SAR data.

5.5. Implications & Future Research Directions

The findings from this study, as discussed in the previous section, have significant implications for future SAR missions such as BIOMASS and NISAR. An important note for these missions is that integrating L-band and P-band radar observables can significantly enhance forest CH estimations compared to using

each band individually. The complementary characteristics of these bands in estimating CH for tropical forests suggest that each band may provide more accurate estimations under varying biome conditions and forest structures, which needs further exploration. Furthermore, future research can be expanded by exploring the effectiveness of shorter SAR wavelengths, such as X- and C-band observables. Notably, the C-band of Sentinel-1, along with the upcoming P-band BIOMASS and L-band missions, offers the potential for global-scale forest height estimation. Therefore, enabling automation with UNet or other FCN frameworks can significantly foster ongoing forest conservation initiatives. While FCN models like UNet reduce efforts in producing handcrafted features for model prediction, they introduce complexities in model interpretability. Therefore, another future direction for this work could involve using polarimetric decomposition features as model inputs and employing Local Interpretable Model-agnostic Explanations (LIME) to better understand which scattering mechanisms and to what extent influence model estimations. Future studies could also explore transfer learning approaches to minimize data scarcity issues, potentially enhancing model robustness and applicability in diverse ecological settings.

6. CONCLUSION

This study demonstrates the effectiveness of using L- and P-band PolSAR backscatters in combination with a customized UNet architecture for accurate forest CH estimation. The dual-band approach outperformed single-band methods, providing more accurate CH predictions across both heterogeneous and homogeneous forest sites. The customized UNet model exhibited superior performance compared to traditional ML models (RF and LGBM), achieving higher R^2 and lower RMSE values.

It is worth noting the study's findings were limited to tropical regions and may not be generalizable to all biomes. The model also showed limitations in accurately predicting very tall canopy heights due to the scarcity of such samples in the training data. Thus, future research should incorporate more study areas or explore transfer learning to address these issues. Additionally, exploring other SAR wavelengths and features will help further refine and validate the proposed methodology. Also, incorporating explainability and interpretability into the model is crucial to enhance its transparency and utility for stakeholders.

Despite the limitations, the integration of this approach into national forest monitoring systems can significantly improve the accuracy of forest biophysical parameter estimations, supporting global efforts in forest conservation and climate change mitigation.

LIST OF REFERENCES

Ahsan, M. M., Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2023). A high-resolution canopy height model of the Earth. *Nature Ecology & Evolution*, *7*(11), 1778–1789.

Argenti, F., Lapini, A., Bianchi, T., & Alparone, L. (2013). A tutorial on speckle reduction in synthetic aperture radar images. *IEEE Geoscience and Remote Sensing Magazine*, *1*(3), 6–35.

Armston, J., Tang, H., Hancock, S., Marselis, S., Duncanson, L., Kellner, J. R., Hofton, M. A., Blair, J. B., Fatoyinbo, T., & Dubayah, R. O. (2020). AfriSAR: Gridded Forest Biomass and Canopy Metrics Derived from LVIS, Gabon, 2016. *ORNL DAAC*. https://doi.org/10.3334/ORNLDAAC/1775

Astola, H., Seitsonen, L., Halme, E., Molinier, M., & Lönnqvist, A. (2021). Deep neural networks with transfer learning for forest variable estimation using sentinel-2 imagery in boreal forest. *Remote Sensing*, *13*(12), 2392.

Berninger, A., Lohberger, S., Zhang, D., & Siegert, F. (2019). Canopy height and above-ground biomass retrieval in tropical forests using multi-pass X-and C-band Pol-InSAR data. *Remote Sensing*, *11*(18), 2105.

Bourgoin, C., Ceccherini, G., Girardello, M., Vancutsem, C., Avitabile, V., Beck, P., Beuchle, R., Blanc, L., Duveiller, G., & Migliavacca, M. (2024). Human degradation of tropical moist forests is greater than previously estimated. *Nature*, 1–7.

Chen, R. H., Pinto, N., Duan, X., Tabatabaeenejad, A., & Moghaddam, M. (2020). *Mapping tree canopy cover and canopy height with L-band SAR using LiDAR data and Random Forests*. 4136–4139.

Chen, W., Zheng, Q., Xiang, H., Chen, X., & Sakai, T. (2021). Forest canopy height estimation using polarimetric interferometric synthetic aperture radar (PolInSAR) technology based on full-polarized ALOS/PALSAR data. *Remote Sensing*, *13*(2), 174.

Clevert, D.-A., Unterthiner, T., & Hochreiter, S. (2015). Fast and accurate deep network learning by exponential linear units (elus). *arXiv Preprint arXiv:1511.07289*.

Dubayah, R. O., & Drake, J. B. (2000). Lidar remote sensing for forestry. *Journal of Forestry*, *98*(6), 44–46.

Duncanson, L., Niemann, K., & Wulder, M. (2010). Estimating forest canopy height and terrain relief from GLAS waveform metrics. *Remote Sensing of Environment*, *114*(1), 138–154.

Fatoyinbo, T., Armston, J., Simard, M., Saatchi, S., Denbina, M., Lavalle, M., Hofton, M., Tang, H., Marselis, S., Pinto, N., Hancock, S., Hawkins, B., Duncanson, L., Blair, B., Hansen, C., Lou, Y., Dubayah, R., Hensley, S., Silva, C., … Hibbard, K. (2021). The NASA AfriSAR campaign: Airborne SAR and lidar measurements of tropical forest structure and biomass in support of current and future space missions. *Remote Sensing of Environment*, *264*, 112533. https://doi.org/10.1016/j.rse.2021.112533

Garestier, F., Dubois-Fernandez, P. C., Guyon, D., & Le Toan, T. (2009). Forest biophysical parameter estimation using L-and P-band polarimetric SAR data. *IEEE Transactions on Geoscience and Remote Sensing*, *47*(10), 3379–3388.

Gazzea, M., Solheim, A., & Arghandeh, R. (2023). High-resolution mapping of forest structure from integrated SAR and optical images using an enhanced U-net method. *Science of Remote Sensing*, *8*, 100093.

Ge, S., Gu, H., Su, W., Praks, J., & Antropov, O. (2022). Improved semisupervised unet deep learning model for forest height mapping with satellite sar and optical data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 5776–5787.

Green, R. (1998). Relationships between polarimetric SAR backscatter and forest canopy and sub-canopy biophysical properties. *International Journal of Remote Sensing*, *19*(12), 2395–2412.

Hajnsek, I., Pardini, M., Jäger, M., Horn, R., Kim, J. S., Jörg, H., & Papathanassiou, K. (2017). *Technical assistance for the development of airborne SAR and geophysical measurements during the AfriSAR campaign*.

Hancock, S., Anderson, K., Disney, M., & Gaston, K. J. (2017). Measurement of fine-spatial-resolution 3D vegetation structure with airborne waveform lidar: Calibration and validation with voxelised terrestrial lidar. *Remote Sensing of Environment*, *188*, 37–50.

Hartmann, H., Bastos, A., Das, A. J., Esquivel-Muelbert, A., Hammond, W. M., Martínez-Vilalta, J., McDowell, N. G., Powers, J. S., Pugh, T. A., & Ruthrof, K. X. (2022). Climate change risks to global forest health: Emergence of unexpected events of elevated tree mortality worldwide. *Annual Review of Plant Biology*, *73*(1), 673–702.

IPCC. (2022). *Climate Change 2022 – Impacts, Adaptation and Vulnerability: Working Group II Contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. https://doi.org/10.1017/9781009325844

Jackson, C. M., Adam, E., Atif, I., & Mahboob, M. A. (2023). Feature extraction and classification of canopy gaps using glcm-and mlbp-based rotation-invariant feature descriptors derived from worldview-3 imagery. *Geomatics*, *3*(1), 250–265.

Khati, U., Singh, G., & Kumar, S. (2018). Potential of space-borne PolInSAR for forest canopy height estimation over India—A case study using fully polarimetric L-, C-, and X-band SAR data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(7), 2406–2416.

Kuenzer, C., Bluemel, A., Gebhardt, S., Quoc, T. V., & Dech, S. (2011). Remote sensing of mangrove ecosystems: A review. *Remote Sensing*, *3*(5), 878–928.

Kugler, F., Lee, S.-K., Hajnsek, I., & Papathanassiou, K. P. (2015). Forest height estimation by means of Pol-InSAR data inversion: The role of the vertical wavenumber. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(10), 5294–5311.

Lang, N., Jetz, W., Schindler, K., & Wegner, J. D. (2023). A high-resolution canopy height model of the Earth. *Nature Ecology & Evolution*, *7*(11), 1778–1789.

Lavalle, M., & Hensley, S. (2012). *Demonstration of repeat-pass POLINSAR using UAVSAR: The RMOG model*. 5876–5879.

Lee, J.-S., & Pottier, E. (2017). *Polarimetric radar imaging: From basics to applications*. CRC press.

Lee, S., Ni-Meister, W., Yang, W., & Chen, Q. (2011). Physically based vertical vegetation structure retrieval from ICESat data: Validation using LVIS in White Mountain National Forest, New Hampshire, USA. *Remote Sensing of Environment*, *115*(11), 2776–2785.

Li, D., Lu, H., Li, C., Mohaisen, L., & Jing, W. (2023). Forest Height Inversion by Convolutional Neural Networks Based on L-Band PolInSAR Data Without Prior Knowledge Dependency. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

Liu, X., Zhang, L., Yang, X., Liao, M., & Li, W. (2021). Retrieval of tropical forest height and above-ground biomass using airborne P-and L-band SAR tomography. *IEEE Geoscience and Remote Sensing Letters*, *19*, 1–5.

Lourenço, P., Godinho, S., Sousa, A., & Gonçalves, A. C. (2021). Estimating tree aboveground biomass using multispectral satellite-based data in Mediterranean agroforestry system using random forest algorithm. *Remote Sensing Applications: Society and Environment*, *23*, 100560.

Lu, D., Chen, Q., Wang, G., Liu, L., Li, G., & Moran, E. (2016). A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *International Journal of Digital Earth*, *9*(1), 63–105. https://doi.org/10.1080/17538947.2014.990526

Luo, Y., Qi, S., Liao, K., Zhang, S., Hu, B., & Tian, Y. (2023). Mapping the forest height by fusion of ICESat-2 and multi-source remote sensing imagery and topographic information: A case study in Jiangxi province, China. *Forests*, *14*(3), 454.

Mahesh, R. B., & Hänsch, R. (2023). *Deep Learning for Forest Canopy Height Estimation from SAR*. 5672–5675. In *IGARSS 2023-2023 IEEE International Geoscience and Remote Sensing Symposium* (pp. 5672-5675). IEEE https://doi.org/10.1109/IGARSS52108.2023.10281899

Marselis, S. M., Tang, H., Armston, J. D., Calders, K., Labrière, N., & Dubayah, R. (2018). Distinguishing vegetation types with airborne waveform lidar data in a tropical forest-savanna mosaic: A case study in Lopé National Park, Gabon. *Remote Sensing of Environment*, *216*, 626–634. https://doi.org/10.1016/j.rse.2018.07.023

Mitchard, E. T. A. (2018). The tropical forest carbon cycle and climate change. *Nature*, *559*(7715), Article 7715. https://doi.org/10.1038/s41586-018-0300-2

Mittermayer, J., Wollstadt, S., Prats-Iraola, P., López-Dekker, P., Krieger, G., & Moreira, A. (2012). Bidirectional SAR imaging mode. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(1), 601–614.

Morais, T. G., Teixeira, R. F., Figueiredo, M., & Domingos, T. (2021). The use of machine learning methods to estimate aboveground biomass of grasslands: A review. *Ecological Indicators*, *130*, 108081.

Moreira, A., Prats-Iraola, P., Younis, M., Krieger, G., Hajnsek, I., & Papathanassiou, K. P. (2013). A tutorial on synthetic aperture radar. *IEEE Geoscience and Remote Sensing Magazine*, *1*(1), 6–43.

Neumann, M., Saatchi, S. S., Ulander, L. M. H., & Fransson, J. E. S. (2012). Assessing Performance of L-and P-Band Polarimetric Interferometric SAR Data in Estimating Boreal Forest Above-Ground

Biomass. *IEEE Transactions on Geoscience and Remote Sensing*, *50*(3), 714–726. IEEE Transactions on

Geoscience and Remote Sensing. https://doi.org/10.1109/TGRS.2011.2176133

Pardini, M., Tello, M., Cazcarra-Bes, V., Papathanassiou, K. P., & Hajnsek, I. (2018). L-and P-band 3-D

SAR reflectivity profiles versus lidar waveforms: The AfriSAR case. *IEEE Journal of Selected Topics*

in Applied Earth Observations and Remote Sensing, *11*(10), 3386–3401.

Pascarella, A. E., Giacco, G., Rigiroli, M., Marrone, S., & Sansone, C. (2023). ReUse: REgressive Unet for

Carbon Storage and Above-Ground Biomass Estimation. *Journal of Imaging*, *9*(3), 61.

Potapov, P., Li, X., Hernandez-Serna, A., Tyukavina, A., Hansen, M. C., Kommareddy, A., Pickens, A.,

Turubanova, S., Tang, H., Silva, C. E., Armston, J., Dubayah, R., Blair, J. B., & Hofton, M. (2021).

Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote*

Sensing of Environment, *253*, 112165. https://doi.org/10.1016/j.rse.2020.112165

Pourshamsi, M., Garcia, M., Lavalle, M., & Balzter, H. (2018). A Machine-Learning Approach to

PolInSAR and LiDAR Data Fusion for Improved Tropical Forest Canopy Height Estimation

Using NASA AfriSAR Campaign Data. *IEEE Journal of Selected Topics in Applied Earth Observations*

and Remote Sensing, *11*(10), 3453–3463. IEEE Journal of Selected Topics in Applied Earth

Observations and Remote Sensing. https://doi.org/10.1109/JSTARS.2018.2868119

Pourshamsi, M., Xia, J., Yokoya, N., Garcia, M., Lavalle, M., Pottier, E., & Balzter, H. (2021). Tropical

forest canopy height estimation from combined polarimetric SAR and LiDAR using machine-

learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, *172*, 79–94.

https://doi.org/10.1016/j.isprsjprs.2020.11.008

Ramachandran, N., & Dikshit, O. (2022). *Forest Aboveground Biomass Estimation from Airborne L-Band SAR*

Data Using Machine Learning. In *IGARSS 2022-2022 IEEE International Geoscience and Remote Sensing*

Symposium (pp. 6403-6405). IEEE

Ramachandran, N., Saatchi, S., Tebaldini, S., d'Alessandro, M. M., & Dikshit, O. (2023). Mapping tropical

forest aboveground biomass using airborne SAR tomography. *Scientific Reports*, *13*(1), 6233.

Reigber, A., Scheiber, R., Jager, M., Prats-Iraola, P., Hajnsek, I., Jagdhuber, T., Papathanassiou, K. P.,

Nannini, M., Aguilera, E., & Baumgartner, S. (2012). Very-high-resolution airborne synthetic

aperture radar imaging: Signal processing and applications. *Proceedings of the IEEE*, *101*(3), 759–783.

Rizzoli, P., Dell'Amore, L., Bueso-Bello, J.-L., Gollin, N., Carcereri, D., & Martone, M. (2022). On the derivation of volume decorrelation from TanDEM-X bistatic coherence. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *15*, 3504–3518.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. *In Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18* (pp. 234-241). *Springer International Publishing.*

Saini, R., & Ghosh, S. K. (2017). *Ensemble classifiers in remote sensing: A review*. 1148–1152.

Sandberg, G., Ulander, L. M. H., Fransson, J. E. S., Holmgren, J., & Le Toan, T. (2011). L- and P-band backscatter intensity for biomass retrieval in hemiboreal forest. *Remote Sensing of Environment*, *115*(11), 2874–2886. https://doi.org/10.1016/j.rse.2010.03.018

Schlund, M., & Davidson, M. W. J. (2018). Aboveground Forest Biomass Estimation Combining L- and P-Band SAR Acquisitions. *Remote Sensing*, *10*(7), Article 7. https://doi.org/10.3390/rs10071151

Schlund, M., Erasmi, S., & Scipal, K. (2019). Comparison of aboveground biomass estimation from InSAR and LiDAR canopy height models in tropical forests. *IEEE Geoscience and Remote Sensing Letters*, *17*(3), 367–371.

Schlund, M., Scipal, K., & Quegan, S. (2018). Assessment of a Power Law Relationship Between P-Band SAR Backscatter and Aboveground Biomass and Its Implications for BIOMASS Mission Performance. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(10), 3538–3547. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. https://doi.org/10.1109/JSTARS.2018.2866868

Silveira, E. M. O., Radeloff, V. C., Martinuzzi, S., Martinez Pastur, G. J., Bono, J., Politi, N., Lizarraga, L., Rivera, L. O., Ciuffoli, L., Rosas, Y. M., Olah, A. M., Gavier-Pizarro, G. I., & Pidgeon, A. M. (2023). Nationwide native forest structure maps for Argentina based on forest inventory data, SAR Sentinel-1 and vegetation metrics from Sentinel-2 imagery. *Remote Sensing of Environment*, *285*, 113391. https://doi.org/10.1016/j.rse.2022.113391

Singh, P., Diwakar, M., Shankar, A., Shree, R., & Kumar, M. (2021). A Review on SAR Image and its Despeckling. *Archives of Computational Methods in Engineering*, *28*, 4633–4653.

Small, D. (2011). Flattening gamma: Radiometric terrain correction for SAR imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(8), 3081–3093.

Soja, M. J., Sandberg, G., & Ulander, L. M. (2012). Regression-based retrieval of boreal forest biomass in sloping terrain using P-band SAR backscatter intensity data. *IEEE Transactions on Geoscience and Remote Sensing*, *51*(5), 2646–2665.

Srivastava, H. S., Patel, P., Sharma, Y., & Navalgund, R. R. (2009). Multi-frequency and multi-polarized SAR response to thin vegetation and scattered trees. *Current Science*, 425–429.

Sun, G.-Q., & Simonett, D. S. (1988). Simulation of L-band and HH microwave backscattering from coniferous forest stands a comparison with SIR-B data. *International Journal of Remote Sensing*, *9*(5), 907–925.

Tamiminia, H., Salehi, B., Mahdianpari, M., Beier, C. M., Johnson, L., & Phoenix, D. B. (2021). A comparison of random forest and light gradient boosting machine for forest above-ground biomass estimation using a combination of landsat, alos palsar, and airborne lidar data. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *44*, 163–168.

Tolan, J., Yang, H.-I., Nosarzewski, B., Couairon, G., Vo, H. V., Brandt, J., Spore, J., Majumdar, S., Haziza, D., & Vamaraju, J. (2024). Very high-resolution canopy height maps from RGB imagery using self-supervised vision transformer and convolutional decoder trained on aerial lidar. *Remote Sensing of Environment*, *300*, 113888.

Townsend, P. (2002). Relationships between forest structure and the detection of flood inundation in forested wetlands using C-band SAR. *International Journal of Remote Sensing*, *23*(3), 443–460.

UNEP. (2018). Fundamentals on REDD+ Module 2: Understanding REDD+ and the UNFCCC. *In REDD+ Academy Learning Journals (Issue December)*. https://www.un-redd.net/documents/global programme-191/redd-academy-3509/redd-academy-learning-journals/english/17233 fundamentals-redd-module-2-understanding-redd-and-unfccc-1.html

Vafaei, S., Soosani, J., Adeli, K., Fadaei, H., Naghavi, H., Pham, T. D., & Tien Bui, D. (2018). Improving accuracy estimation of Forest Aboveground Biomass based on incorporation of ALOS-2

PALSAR-2 and Sentinel-2A imagery and machine learning: A case study of the Hyrcanian forest area (Iran). *Remote Sensing, 10*(2), 172.

Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J. P., Veroustraete, F., Clevers, J. G., & Moreno, J. (2015). Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties–A review. *ISPRS Journal of Photogrammetry and Remote Sensing, 108*, 273–290.

Wang, C., Hu, C., Shen, P., & Song, T. (2022). Evaluation of multilooking size on single-baseline PolInSAR forest height inversion. *Forests, 13*(7), 1031.

Wang, X., Hu, Z., Shi, S., Hou, M., Xu, L., & Zhang, X. (2023). A deep learning method for optimizing semantic segmentation accuracy of remote sensing images based on improved UNet. *Scientific Reports, 13*(1), 7600.

Wang, Y., Davis, F. W., & Melack, J. M. (1993). Simulated and observed backscatter at P-, L-, and C-bands from ponderosa pine stands. *IEEE Transactions on Geoscience and Remote Sensing, 31*(4), 871–879.

Zhang, B., Zhu, H., Xu, W., Xu, S., Chang, X., Song, W., & Zhu, J. (2023a). A Fourier–Legendre Polynomial Forest Height Inversion Model Based on a Single-Baseline Configuration. *Forests, 15*(1), 49.

Zhang, B., Zhu, H., Xu, W., Xu, S., Chang, X., Song, W., & Zhu, J. (2023b). A Fourier–Legendre Polynomial Forest Height Inversion Model Based on a Single-Baseline Configuration. *Forests, 15*(1), 49.

Zhu, J., Xie, Y., Fu, H., Wang, C., Wang, H., Liu, Z., & Xie, Q. (2023). Digital terrain, surface, and canopy height model generation with dual-baseline low-frequency InSAR over forest areas. *Journal of Geodesy, 97*(11), 100.

Zhu, L., Suomalainen, J., Liu, J., Hyyppä, J., Kaartinen, H., & Haggren, H. (2018). A review: Remote sensing sensors. *Multi-Purposeful Application of Geospatial Data, 19.*

Zoran, D., & Weiss, Y. (2011, November). From learning models of natural image patches to whole image restoration. *In 2011 international conference on computer vision (pp. 479-486). IEEE*

7. APPENDICES

Appendix 1: Represent the convergence graph of the training loss and validation loss of Lope (a)L- band ((b) P-band, (c) Dual-band and Mabounie (d) L- band, (e) P-band, (f) Dual-band

(a)

(b)

(c)

(d)

(e)

(f)
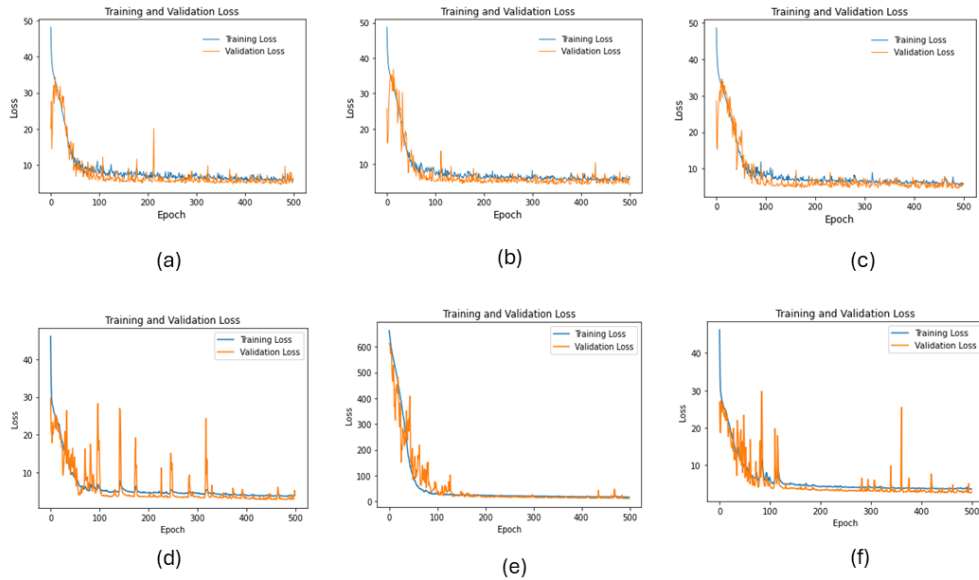
Appendix 2 (a) RMSE Values from Iterative UNet Model Runs for Lope

L- band	P-band	Dual band
7.8	8.94	4.53
10.2	8.91	3.78
7.56	8.03	4.42
9.09	8.94	3.78
9.33	8.89	3.74
9.12	8.08	4.53
9.99	9.21	3.78
9.09	9.09	4.42
8.99	9.04	3.78
9.09	9.09	3.74
9.88	9.09	3.74
10.22	7.55	3.74
8.98	9.01	3.74
9.09	9.13	3.74
9.09	8.88	3.68

(b) RMSE Values from Iterative UNet Model Runs for Mabounie

RF	LGBM	UNet	
7.8	8.94	3.78	
10	8.94	3.78	
7.56	8.03	4.42	
9.09	8.94	3.78	
9.09	8.94	3.74	
9.09	8.08	4.53	
9.09	9.09	3.78	
9.09	9.09	4.42	
9.09	9.09	3.78	
9.09	9.09	3.74	
10	9.09	3.75	
10.22	7.55	3.78	
7.8	8.94	3.74	
9.09	8.94	3.66	
9.09	8.94	3.68	

(c) RMSE Values from Iterative RF, UNet and LGBM Model Runs for Lope

RF	LGBM	UNet
9.77	9.09	4.53
10.04	6.56	3.78
9.17	9.51	4.42
9.42	9.51	3.78
10.04	9.52	3.74
9.77	9.52	4.53
10.04	9.52	3.78
9.17	9.09	4.42
9.42	9.09	3.78
10.04	10.04	3.68
9.42	10.04	3.75
10.04	9.09	3.74
9.77	9.13	3.79
6.57	6.56	3.74
10.04	9.13	3.68

(d) RMSE Values from Iterative RF, UNet and LGBM Model Runs for Mabounie

RF	LGBM	UNet
7.8	9.09	3.78
10	6.56	3.78
7.56	9.51	4.42
9.09	9.51	3.78
9.33	9.52	3.74
9.09	9.52	4.53

9.18	9.58	3.78
9.09	9.09	4.42
9.25	9.66	3.78
9.09	10.04	3.74
10	10.3	3.75
10.22	9.09	3.78
7.8	9.13	3.74
9.09	6.56	3.66
9.2	9.13	3.68

Appendix 3

(a) Statistical significance test on RMSE of L-, P- and Dual Band estimations for Lope site

Groups	Count	Sum	Average	Variance	Standard Deviation
L- band	15	97.57	6.50	0.97	0.98
P-band	15	84.64	5.64	0.25	0.50
Dual-band	15	63.81	4.25	0.07	0.27

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	38.68	2	19.34	45.06	0.00	3.22
Within Groups	18.03	42	0.43			
Total	56.71	44				

POST-HOC TEST

Groups	P-value (T-test)	Significant?
L- band v P- band	0.06	NO
P-band v Dual-band	0.00	YES
Dual-band v L-band	0.00	YES

ALPHA

Test	Alpha
ANOVA	0.05
Post-hoc test (Bonferroni corrected)	0.02

(b) Statistical significance test on RMSE of L-, P- and Dual Band estimations for Mabounie site

SUMMARY

Groups	Count	Sum	Average	Variance	Standard Deviation
L- band	15	97.57	6.50	0.97	0.98
P-band	15	84.64	5.64	0.25	0.50
Dual-band	15	63.81	4.25	0.07	0.27

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	38.68	2	19.34	45.06	0.00	3.22
Within Groups	18.03	42	0.43			
Total	56.71	44				

POST-HOC TEST

Groups	P-value (T-test)	Significant?
L- band v P- band	0.00	YES
P-band v Dual-band	0.00	YES
Dual-band v L- band	0.00	YES

ALPHA

Test	Alpha
ANOVA	0.05
Post-hoc test (Bonferroni corrected)	0.02

(c) Statistical significance test on RMSE of RF, LGBM, and UNet for Lope site

SUMMARY

Groups	Count	Sum	Average	Variance	Standard Deviation
RF	15	97.57	6.50	0.97	0.98
LGBM	15	84.64	5.64	0.25	0.50
UNet	15	63.81	4.25	0.07	0.27

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	38.68	2	19.34	45.06	0.00	3.22
Within Groups	18.03	42	0.43			
Total	56.71	44				

POST-HOC TEST

Groups	P-value (T-test)	Significant?

ALPHA

RF v LGBM	0.18	NO	
LGBM v UNet	0.00	YES	
UNet v RF	0.00	YES	

Test	Alpha
ANOVA	0.05
Post-hoc test (Bonferroni corrected)	0.02

(d) Statistical significance test on RMSE of RF, LGBM, and UNet for Mabounie site

SUMMARY

Groups	Count	Sum	Average	Variance	Standard Deviation
RF	15	97.57	6.50	0.97	0.98
LGBM	15	84.64	5.64	0.25	0.50
UNet	15	63.81	4.25	0.07	0.27

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	38.68	2	19.34	45.06	0.00	3.22
Within Groups	18.03	42	0.43			
Total	56.71	44				

POST-HOC TEST

Groups	P-value (T-test)	Significant?
RF v LGBM	0.33	NO
LGBM v UNet	0.00	YES
UNet v RF	0.00	YES

ALPHA

Test	Alpha
ANOVA	0.05
Post-hoc test (Bonferroni corrected)	0.02

ETHICAL CONSIDERATION

The AfriSAR datasets were openly provided by the European Space Agency (ESA). However, before this study, Dr. Michael Schlund conducted preprocessing on the raw data which has been used in his previous research works. This entitles him to the rights of the processed information and data. LVIS height metrics are made publicly available by NASA and can be accessed at lvis.gsfc.nasa.gov.

The conclusions drawn from this study are solely grounded in empirical evidence derived from these datasets. The outputs such as codes and images of this research will be made available for further use without any legal constraints. During the preparation of this work, the author used QuillBot and Chatgpt for paraphrasing and sentence structuring purposes, followed by a thorough review and editing process. The author takes full responsibility for the academic integrity and content of this work.