



# Canopy Height Mapping for Plantations in Nigeria Using GEDI, Landsat, and Sentinel-2

Angela Tsao <sup>1,\*</sup> , Ikenna Nzewi <sup>2</sup>, Ayodeji Jayeoba <sup>2</sup>, Uzoma Ayogu <sup>2</sup> and David B. Lobell <sup>1</sup>

<sup>1</sup> Department of Earth System Science, Center on Food Security and the Environment, Stanford University, Stanford, CA 94305, USA

<sup>2</sup> Releaf, Lagos Victoria Island Suite, Eti-Osa, Lagos 101233, Nigeria; ikenna.nzewi@releaf.ng (I.N.); dejijaye@releaf.ng (A.J.)

\* Correspondence: angtsao@stanford.edu

**Abstract:** Canopy height data from the Global Ecosystem Dynamics Investigation (GEDI) mission has powered the development of global forest height products, but these data and products have not been validated in non-forest tree plantation settings. In this study, we collected field observations of the canopy heights throughout oil palm plantations in Nigeria and evaluated the performance of existing global canopy height map (CHM) products as well as a local model trained on the GEDI and various Landsat and Sentinel-2 feature combinations. We found that existing CHMs fared poorly in the region, with mean absolute errors (MAE) of 4.2–6.2 m. However, the locally trained models performed well (MAE = 2.5 m), indicating that using the GEDI and optical satellite data can still be effective, even in a region with relatively sparse GEDI coverage. In addition to improved overall performance, the local model was especially effective at reducing errors for short (<5 m) trees, where the global products struggle to capture the canopy height.

**Keywords:** GEDI; plantation; LiDAR; Landsat; canopy height; tree crops



**Citation:** Tsao, A.; Nzewi, I.; Jayeoba, A.; Ayogu, U.; Lobell, D.B. Canopy Height Mapping for Plantations in Nigeria Using GEDI, Landsat, and Sentinel-2. *Remote Sens.* **2023**, *15*, 5162. <https://doi.org/10.3390/rs15215162>

Academic Editor: Jan Altman

Received: 18 August 2023

Revised: 19 October 2023

Accepted: 26 October 2023

Published: 29 October 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Trees have been widely studied for their important role in sequestering carbon, bearing crops, and filtering air [1], with quantitative estimates of their value typically derived through combinations of forest inventories and remote sensing [2–4]. However, much of this work has focused on forests, with the extent and value of trees outside of forests (TOF) relatively less understood [5,6]. Due to varying definitions of what constitutes a forest, ranging from tree cover thresholds to land use constraints, it has even been difficult to define what type of land cover constitutes TOF [5,7]. Plantations, colloquially defined as areas in which trees were planted by humans, are one specific case of TOF that is differentiated by the central role of human management. Mapping the structural characteristics of plantations is particularly valuable for monitoring outcomes such as yield or carbon storage and for optimizing disaggregated supply chains [8,9].

Although large-scale plantations are common, trees are also ubiquitous in smallholder systems, with estimates indicating that more than 30% of smallholder farms contain trees, either interspersed with crops or as small, monocrop plantations [10]. Based on this lower-bound estimate, trees on farms account for 17% of the annual gross income for the farms where they are grown [10], providing many smallholders with a resilient source of annual income as well as goods for direct consumption [11]. Unlike annual agricultural crops, tree crops in plantations have structural characteristics that allow them to accumulate and store carbon over a multi-decade life cycle [12]. Additionally, as opposed to natural forests, plantations often receive more regular, costly management, as the production of a primary crop provides a vital source of household income and nutrition. This aspect of production may enable better implementation of carbon management practices and other efforts to increase the environmental benefits associated with tree crops [13]. Measuring

the height and vertical structure of plantations will help with monitoring these benefits and quantifying their value at scale.

Measurement of the canopy height is a key goal in many tree monitoring efforts, as tree height is a key input into many downstream tasks including the estimation of biomass, age, and carbon accumulation. This interest in canopy height maps (CHMs) has led to significant advances both in novel data sources and modeling approaches. The Global Ecosystem Dynamics Investigation (GEDI) mission was launched in 2019 with the scientific goal to measure forest biomass, biodiversity, and quality [14]. The spaceborne light detection and ranging (LiDAR) captures the vertical structure of vegetation with active, full-waveform LiDAR beams at a 25 m radius footprint spatial resolution (the size of an active LiDAR collection) spaced approximately every 60 m along mission tracks with an across-track width of 4.2 km. This mission builds upon a number of disaggregated efforts to use airborne LiDAR scanning (ALS) to capture similar insights into forest metrics [15–17] and older LiDAR satellite missions with canopy height products, such as Ice, Cloud, and Land Elevation Satellite-2 [18]. The GEDI is powerful for its near-global scope for data collection and has a higher sampling density compared with its predecessors. However, across this huge spatial extent, it is a sparse dataset with footprints spaced out along mission tracks [14]. This spatial sampling means that large regions of the world are well represented, but any individual location is unlikely to be observed. Although the GEDI science team also produces gridded products that cover all regions, the 1 km resolution of these products is too coarse to capture small, irregularly shaped fields typical of smallholder plantations. Rather, wall-to-wall data such as optical imagery provide features with the appropriate coverage to observe plantations but lack a mapping of the image to a GEDI metric like the relative height.

Recent efforts have focused on combining data from the GEDI with optical datasets to develop wall-to-wall maps of forest metrics such as the canopy height at a finer spatial resolution [19,20]. These fusion strategies leverage the GEDI as reference labels to train a model to predict the canopy height from wall-to-wall optical imagery. Prior to the release of the GEDI, similar methods combined optical datasets with GEDI predecessors such as regional ALS [17]. The GLAD Global Forest Canopy Height 2019 Map [20] trains a collection of localized ensemble machine learning models on GEDI relative height metrics to predict the canopy height from spatiotemporal features aggregated from time series Landsat data. A unique model is trained for every 1 degree by 1 degree cell on a larger coordinate grid based on training data collected from cells within a 12 degree radius. Predictions are made independently for each cell on the coordinate grid and then stitched back together into a global-scale map with a 30 m resolution [20]. In a different approach, the ETH Global Sentinel-2 Canopy Height 2020 Map [19] generated predictions at a 10 m resolution using convolutional neural networks trained on Sentinel-2 image tiles and GEDI reference heights. Five CNNs with identical structures but different initializations were trained, and each was applied for global-scale inference. The final global canopy height map was generated by randomly selecting one of the five models for inference for each 100 km × 100 km tile Sentinel-2 image across a global grid [19]. Both the GLAD and ETH maps are examples of spatially contiguous products that track the global status of trees, providing critical inputs for various downstream applications such as monitoring nature-based solutions.

However, these tree height models are explicitly designed for forest land and have not been systematically tested on non-forest tree cover. Due to modeling constraints, canopy height models have shown a tendency for underpredicting at the very upper end (>95% for height) and overpredicting plots below the median height of a training set [20,21]. Much emphasis has been placed on improving the prediction for very tall forests, including a study finding that highly localized training works well for reducing “signal saturation” in tall regions [22]. Regional deep learning methods have also been developed and applied globally as a strategy to reduce saturation for tall canopies [19,23,24]. The tallest trees are important when predicting heights for forests, because tall forests have higher species richness in several taxonomic groups and are a priority for conservation

efforts [25]. However, less focus has been placed on improving performance where trees are short. Previous reasoning has deferred to traditional definitions of woody matter as “having the potential to reach 2–5 m height”, and some methods such as the GLAD CHM actively filter all pixels with a predicted height <3 m as 0 m tall non-forest land [20]. For typical old-growth forests, this definition is valid, but CHMs could suffer significant shortcomings when predicting heights for non-forest tree cover, such as that in plantations. In plantations, the shortest trees are important as they represent young or newly planted trees with significant growth potential. For crops like oil palm, height and age are associated heavily with yield, and the shortest trees may be expected to rapidly increase production as they mature [26]. Height maps can then inform the local need for equipment for harvesting or post-processing. Given that stresses from climate change will affect the suitable growing zones for oil palm, it is particularly important to be able to map the emergence and growth of palm plantations [27].

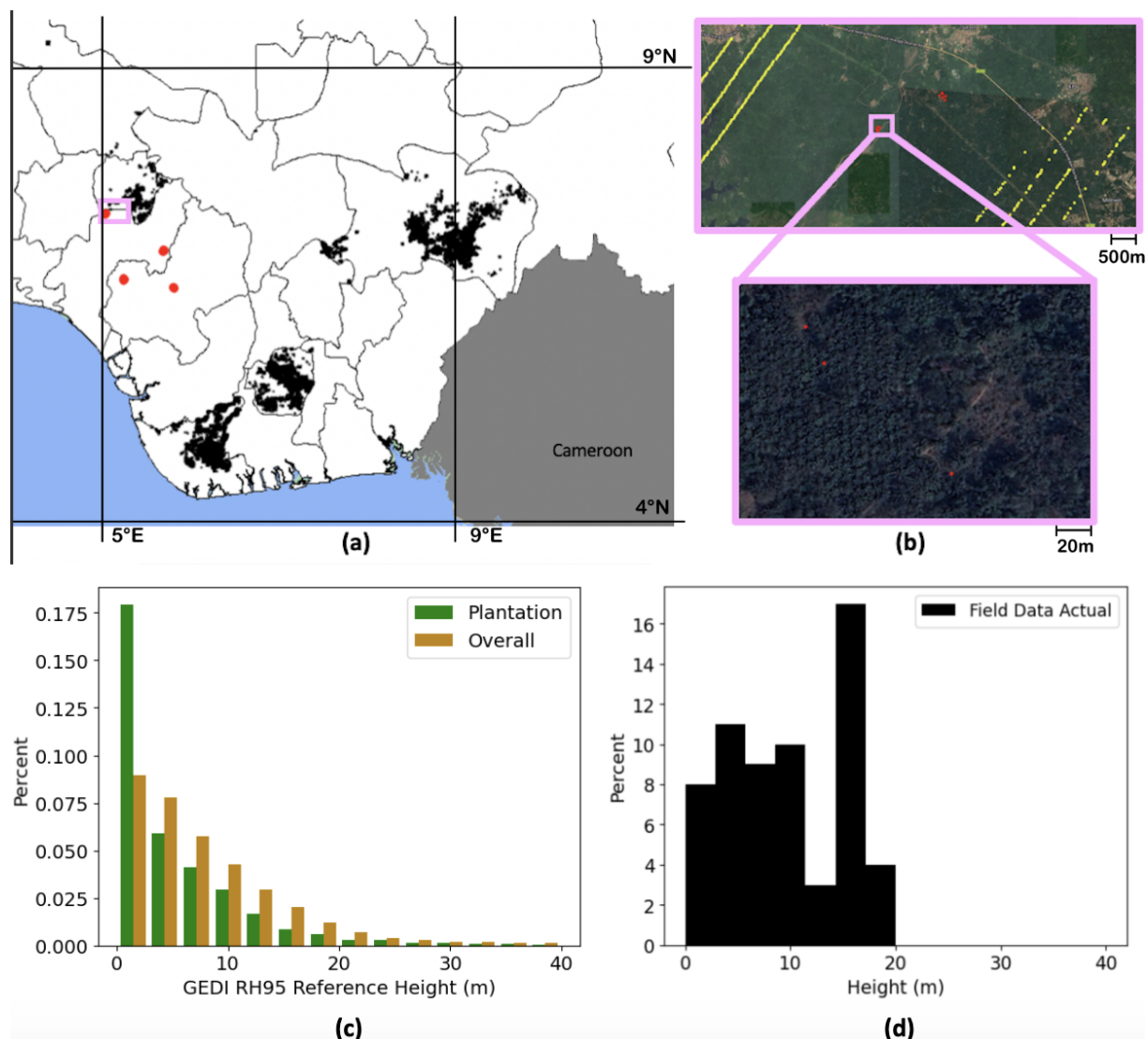
Footprint-level GEDI data also hold higher geolocation uncertainty in landscapes with heterogeneity in canopy cover, particularly where small stands of trees, sparse trees, or forest edges are the predominant type of tree cover [28]. These characteristics are particularly common for smallholder plantations, and thus it is uncertain how suitably GEDI-based height estimation methods can perform for disaggregated plantations. Studies comparing GEDI measurements directly to airborne LiDAR scanning measurements generally find high agreement between the two data sources but with normally distributed errors. This uncertainty demands the availability of expanded, verified ground truth canopy height data from field teams with standardized data collection methodologies. Given the GEDI’s initial design and calibration emphasis on tall, dense forests, some studies have explored how well the data and current methods work in border settings such as savannas. In savannas, the GEDI accurately characterized the heights for trees taller than 3 m but failed to accurately predict the heights for sparse woody vegetation shorter than that 3 m threshold [29]. Despite this consideration of natural TOF, GEDI height prediction methods have not yet been validated in plantation settings, where human management and cultivation causes a distribution shift in tree cover readings from remote sensors. To fill this gap, we focused in this study on smallholder palm plantations in Nigeria (the largest producer of palm oil in Africa [27]) and aimed to evaluate existing global canopy height products using high-quality ground measurements collected in plantations. We then tested how a model trained with country-level data performed against these existing products, comparing the ability of different types of features and sensors to predict the canopy height.

## 2. Materials and Methods

### 2.1. Plantation Mask

We investigated and collected data over palm-growing southern Nigeria, which is the largest oil palm producer in Africa. Through a community partner providing sustainable processing in the palm oil chain, and as a precursor to both the field work and GEDI analysis, we developed a non-exhaustive map of oil palm locations in the region (Figure 1a). Based on the partner’s capabilities, these efforts were focused on five main states with socioeconomic relevance and varying plantation qualities: Bayelsa, Benue, Edo, Ekiti, and Imo. One of the biggest challenges for developing the plantation mask was differentiating oil palm and other vegetation, especially coconut trees. Field agents collected reference data on the leaf structure, canopy density, shape, and interspacing in confirmed oil palm plantations and coconut locations. High-resolution satellite imagery was then preprocessed using ENVI to correct for atmospheric distortions, radiometric calibration, and geometric registration. The normalized difference vegetation index (NDVI) was calculated using the same software to assess the vegetation density. Pixel-level classification was performed with random forest using ground observations as training, and then the boundaries of the oil palm fields were mapped by hand in QGIS. Historical land use records were referenced to minimize potential misclassifications and ensure high accuracy. Additionally, we took into account field knowledge and local expertise from stakeholders to validate the delineated

boundaries. Field agents also visited a sample of locations to collect ground truth data and verification. In total, 45,459 fields were identified on this map.



**Figure 1.** Visualization of the study region in Nigeria. (a) Map of oil palm plantation mask (black polygons) and field sample locations (red dots). (b) An example of field sampling locations and sampled GEDI shots (yellow dots) overlaid on aerial imagery. Panel shows field-level imagery. Observe that GEDI tracks may not always intersect with field sites or points of interest. (c) Histograms of GEDI RH95 values for random samples across the plantation extent vs. across Nigeria as a whole. (d) Histogram of field-captured height for test set ( $n = 62$ ).

Using this non-exhaustive dataset of oil palm plantations in Nigeria, we analyzed the characteristic differences between the plantation regions and the overall landscape (Figure 1c). As the overall landscape included non-forest areas, it contained a large number of low-height points. However, many of these “low-height” returns may have been non-tree land cover, such as grassland, urban area, or cropland.

## 2.2. Field Data

In addition to the regional plantation mask, we collected ground truth height metrics at 62 different points in the Edo region of Nigeria (Figure 1a,b). These data were held out as a test set to confirm the model’s performance on unseen data given the strategy of using the GEDI as a label for supervised learning. Field sites were sampled from plantations where

the property managers or owners had high certainty about the ages of the plantations. The sites were chosen to collect heights for trees across a wide range of age classes in order to obtain a diversity of tree heights. Furthermore, to ensure that various types of smallholder settings were captured, we included a mix of organized, larger-scale operations and scattered palm agroforestry. The samples were collected with the same field agents across two site visits in January 2023 and June 2023 using the GLOBE Observer smartphone-based app. Data collection using GLOBE Observer has been found to correspond well with airborne LiDAR-based tree height measurements, generally falling within a 10% error range [30]. The latitude and longitude coordinates for each location were collected via smartphone with an error range of 5–20 m. A representative tree from each field was chosen from a central location within the field. The centers of the fields were used to avoid field edges and mitigate the impact of geolocation errors.

### 2.3. GEDI Data

We sampled the GEDI footprints over areas with at least 10% tree cover according to the Hansen Forest Change dataset [31] for each year between 2019 and 2021 of GEDI availability. Sampling was random across this spatial extent. The relative height at the 95th percentile (RH95) for the point was taken based on the value of the footprint's root group, which selected the optimal processing algorithm for the waveform from a pool of six L2A processing algorithms [32]. The points were also filtered for quality and lack of degradation, and points with RH95 values greater than 40 m were treated as outliers and discarded. Sampling prioritized more recent footprints, and thus for experimental settings where more than 50,000 footprints could be sampled from a single year of GEDI measurements, we sampled the 2021 footprints only.

### 2.4. Multispectral Satellite Data

We explored both Landsat and Sentinel-2 as potential predictors of the canopy height, with all image processing and sampling performed in Google Earth Engine (GEE) [33]. For Landsat, surface reflectance data (GEE asset name: "LANDSAT/LC08/C02/T1\_L2") were extracted at each of the sampled GEDI footprints. We masked out cloudy pixels by creating a bitmask from the "QAPixels" property and applying it over the "cloud\_shadows" and "Cloud" properties within the dataset. The cloud confidence and cloud shadow confidence were identified at bits 3 and 5 from the quality band. We derived spatiotemporally aggregated band and index metrics from a sliding 3 year time series of Landsat data, starting with the year of GEDI sample collection. The same procedure was applied to spatiotemporally aggregate features from the cloud-masked Sentinel-2 data (GEE asset name: "COPERNICUS/S2"), with the GEDI footprints sampled to a higher resolution scale of 10 m per pixel. We masked out pixels based on the cloud probability ( $p$ ) value, cloud displacement index ( $cdi$ ), and cirrus band ( $c$ ), masking the low-to-mid atmospheric clouds where

$$(p > 0.65 \ \& \ cdi < -0.5) \ \text{or} \ c > 0.01, \quad (1)$$

accounting for both low-to-mid atmospheric clouds and higher atmosphere cirrus clouds. For both Landsat and Sentinel-2, the composite metrics included various percentiles and the mean and standard deviation for the surface reflectance band values (Red, Green, Blue, Near Infrared (NIR), Short-Wave Infrared 1 (SWIR1), Short-Wave Infrared 2 (SWIR2), and Panchromatic) and computed indices (normalized difference vegetation index (NDVI) [34], normalized difference water index (NDWI) [35], and normalized difference moisture index (NDMI) [36]). The indices are computed by the following definitions:

$$NDVI = (NIR - Red) / (NIR + Red), \quad (2)$$

$$NDWI = (Green - NIR) / (Green + NIR), \quad (3)$$

$$NDMI = (NIR - SWIR)/(NIR + SWIR). \quad (4)$$

A higher NDVI and NDWI are both correlated with higher levels of vegetation, while the NDMI measures the moisture content in vegetation. These metrics are thus well recognized for applications involving trees and woody vegetation. The NDVI in particular is a sensitive indicator for the canopy structure in sparse canopies [37]. The NDVI and NDMI had strong correlations with an ALS-captured dataset in central Spain, showing strong potential to predict the canopy height from spectral indices [38].

### 2.5. Random Forest Model and Feature Ablation

A well-established model for predicting the biomass and height from multispectral satellite data is the random forest (RF) model [39–42], which implicitly incorporates ensembling to improve on the performance of a single classifier [43]. Python’s scikit-learn package was used to implement the model, used here as a regressor. The hyperparameters were set to the default scikit-learn settings (number of trees = 100, criterion = squared error, minimum sample-split = 2, minimum samples per leaf = 1, and no max depth). For Landsat, we trained our base RF model on an 80-20 train-validation split given 71,621 sampled footprints from 2021 and corresponding composite metrics for the 3 year range from 2021 to 2023. For Sentinel-2, a separate RF model was trained on a different random sample of GEDI footprints and extracted composite metrics, with the same feature set across the same spectral bands and indices. The mean absolute error (MAE) and root mean square error (RMSE) metrics were computed for both Landsat and Sentinel by using the corresponding model to predict the heights for the 62 field points. We also measured the errors for the important subset of short trees on plantations, computing the error for trees <5 m. This definition included <3 m tall woody matter that would have been explicitly ignored by other models, as well as the shortest defined tree pixels.

Table 1 shows the range of features included for our model and different experimentation settings for which features were included to train the best-performing model. RF models may suffer from the curse of dimensionality, which means that feature selection can often improve problems with overfitting [42]. As such, it is relevant to test which features have the greatest importance for model performance. We found that inclusion of all features resulted in the most reliable model, with no significant difference between the 1 year and 3 year composited statistics (Figure 2).

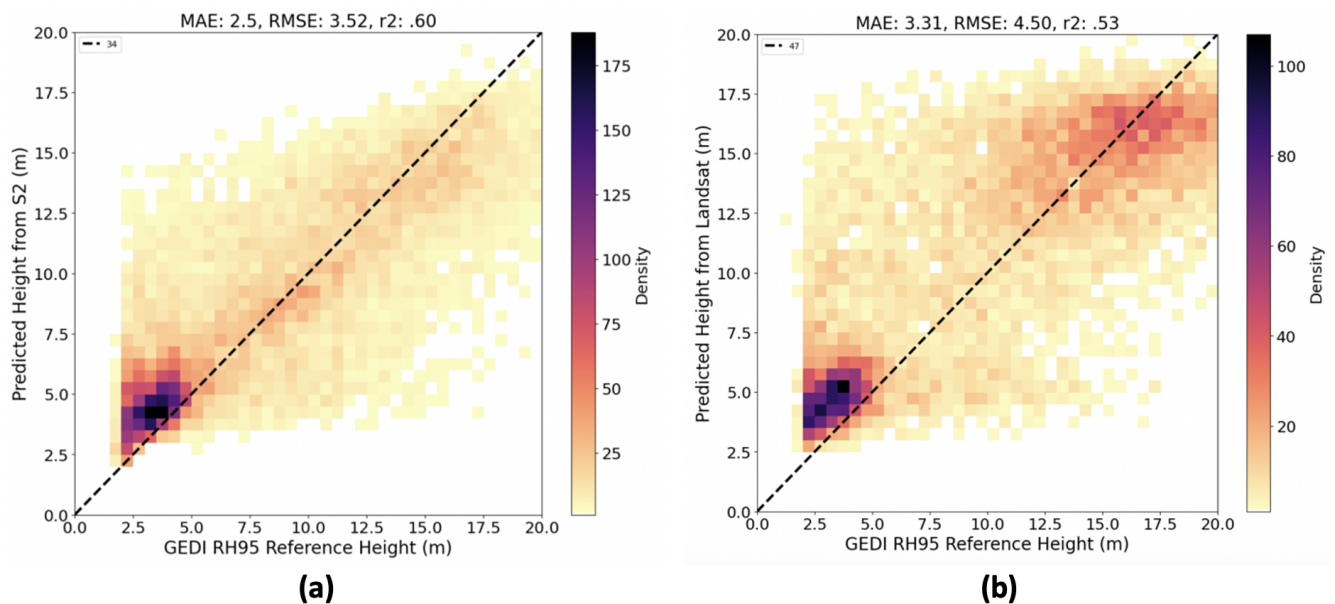
**Table 1.** Feature combinations tested for different models.

<i>1 year composite</i>
All percentiles: Surface Reflectance Band Values (SR) + indices
25/50/75/mean/std: SR + indices
<i>3 year sliding composite</i>
All percentiles: SR + indices
25/50/75/mean/std: SR + indices
25/50/75/mean/std: SR only
25/50/75/mean/std: RGB and NIR only
25/50/75/mean/std: indices only

### 2.6. Local Calibration via Spatially Radiating Sampling

Random forest models perform best when given sufficient training samples to overcome complexity in high-dimensional feature spaces, as well as when these training data match the distribution of the test set [42]. One application of canopy height estimation with RF found that the localized models performed well but could not transfer to other regions or vegetation types. However, increasing the spatial extent of the training data allowed for

a generalizable model to be trained but only within a spatial extent of 50 km<sup>2</sup> [40]. For local calibration, we sampled the GEDI footprints from concentric circles with increasingly large radii, which were always centered at each of the 62 field sites. We took the union of each of these 62 buffered  $r$ -radius circles as the complete sampling region for a given sampling radius. The radii used were 0, 5, 10, and 20. At smaller radii (<10 km), the sampling was exhaustive, but not all available footprints were used when the sampling radius was large. (The total number of samples was restricted to 8000 samples per 2.5 coordinate degree  $\times$  2.5 coordinate degree grid cell along the latitude and longitude axes).



**Figure 2.** Heatmap for relationship between model's predicted height and reference height for (a) Landsat model and (b) Sentinel-2 model, both trained with GEDI-labeled points retrieved from known plantations.

### 3. Results

When testing the GLAD CHM and higher-resolution ETH CHM on the 62 plantation field points, we observed RMSE values of 5.09 m and 7.54 m (Table 2). In comparison with their global errors, the GLAD CHM and ETH CHM actually performed slightly better in terms of absolute error over our test set than on the global scale. However, the plantations in our dataset have predominantly shorter trees compared with old forest settings, and therefore the errors were large relative to the heights of the trees. The best hyperparameter combination of our random forest model included all spectral bands and indices. A comparison against the GLAD CHM [20] and higher-resolution ETH CHM [19] showed our model significantly outperforming the global, forest-focused models.

**Table 2.** Landsat-based model results on field data with different feature inputs. Feature ablation demonstrates that best performance was achieved by including the combination of all surface reflectance bands and computed indices (NDVI, NDWI, and NDMI) in prediction (results on ground-verified field dataset for 3 year Landsat composite with standard RF settings).

Metric	Our Model (30 m)	GLAD CHM (30 m)		ETH CHM (10 m)
		<3 m Points Excluded	<3 m Points Included	
MAE	2.52	4.57	4.18	6.19
RMSE	3.37	5.62	5.09	7.54
R <sup>2</sup>	0.63	−0.59	0.13	−0.91
MAE for Trees <5 m	0.9	1.28	2.77	5.53

We evaluated the performance of Sentinel-2 and Landsat, finding that the models were comparable in validation performance for held-out GEDI samples, but a Landsat model performed somewhat better on the field points (Figure 2). One possible explanation is that Sentinel-2's features are more sensitive to geolocation errors in the field data compared with Landsat's features. As a different experimental setting, we contrasted the model test performance under the different feature combinations mentioned in Table 1. The combination of surface reflectance bands and indices together resulted in the best height estimates, with all other settings leading to both severe underprediction of tall and medium heights as well as overprediction of short heights (Table 3). Notably, the surface reflectance bands on their own were more useful in prediction than just the computed indices.

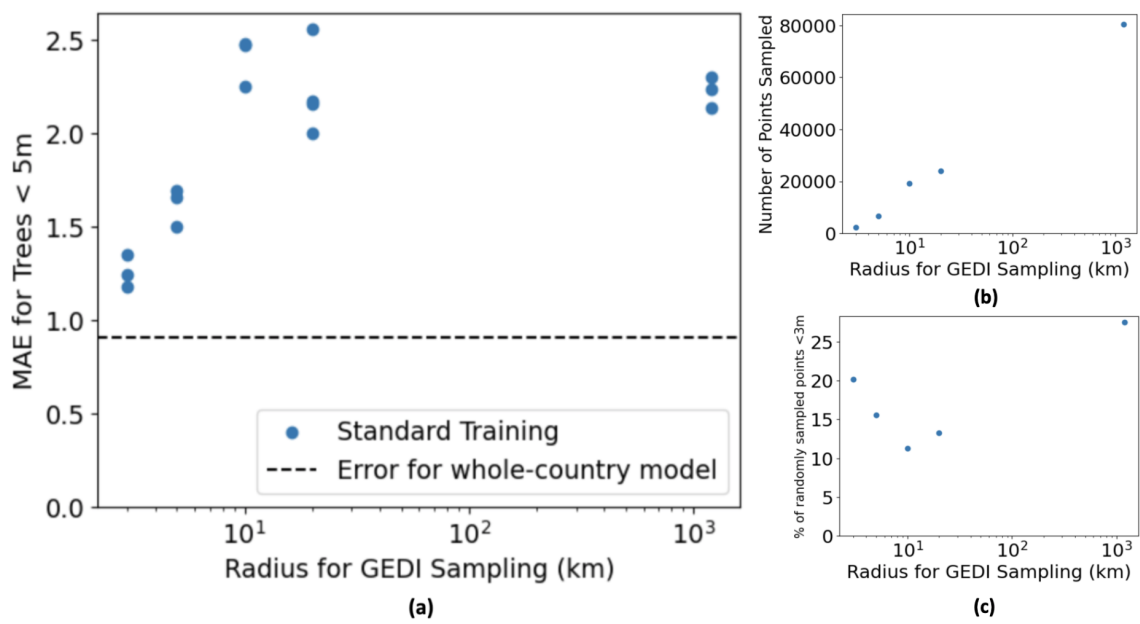
**Table 3.** Landsat-based model results for field data with different feature inputs. Feature ablation demonstrates that best performance was achieved by including the combination of all surface reflectance bands and computed indices (NDVI, NDWI, and NDMI) in prediction (results on ground-verified field dataset for 3 year Landsat composite with standard RF settings).

Metric	SR Bands Only	Indices Only	RGB + NIR Only	All Features (SR Bands + Indices)
MAE	2.76	3.26	3.22	2.67
RMSE	3.61	4.21	4.01	3.54
R <sup>2</sup>	0.56	0.41	0.46	0.58
MAE for trees < 5	1.32	2.4	2.2	0.91

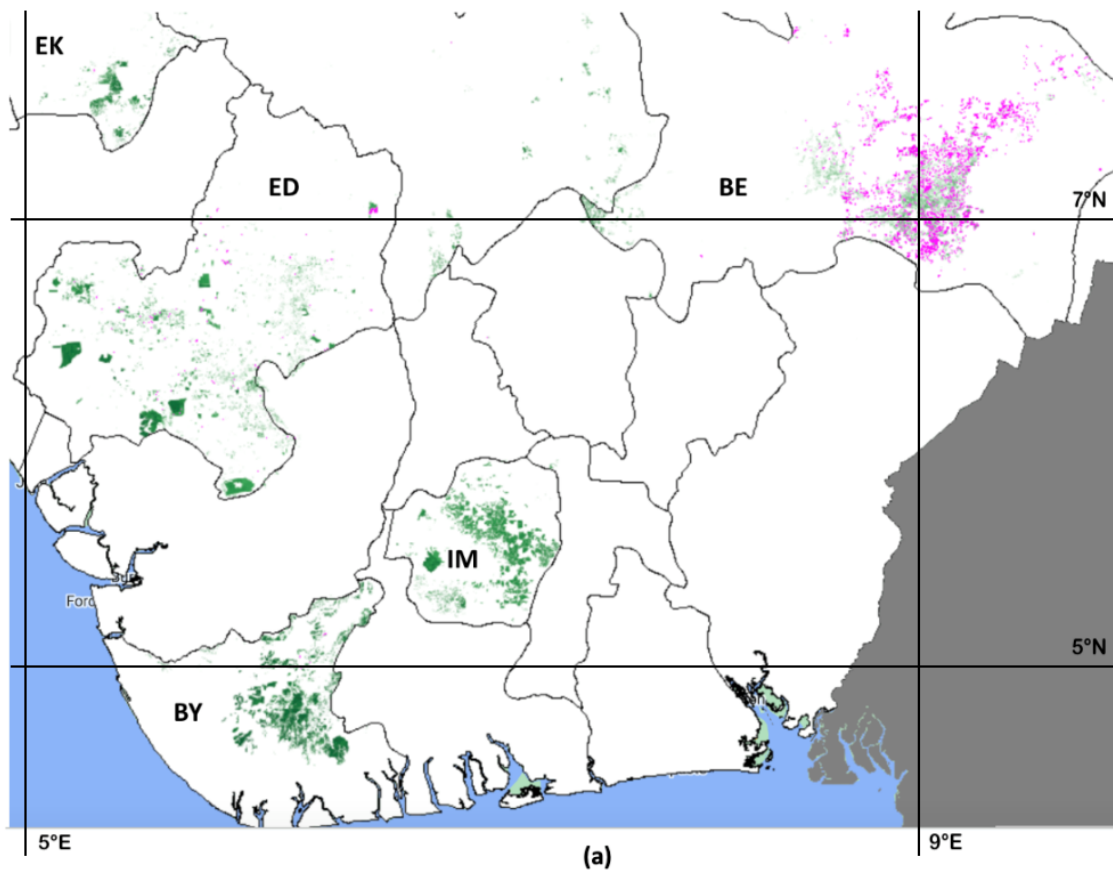
Given past work demonstrating that local calibration can improve performance for outlier short and tall points [22], we tested models trained on samples drawn at different spatial scales (Figure 3). We found that local calibration did not influence the overall errors on the test data but showed a local trend of reducing errors on short trees as the sampling grew more localized. The opposite pattern emerged if shorter trees were treated as unidentifiable points, such as the GLAD methodology of setting all points below a threshold (3 m) to a 0 m height. In this case, the error for the short trees worsened under highly local sampling (Figure 3). We also tested the comparison of training a model on points from the whole country versus points from just the regions selected within our plantation mask. However, this subsetting of the training data resulted in a slight decrease in performance on the test data. It is possible that our plantation mask did not fully represent all plantations in the region or even those in our ground-validated test set. Spectral band- and index-based representations of plantations and forests are highly diverse even within a class. The best way for a model like random forest to learn multiple complex relationships between features and labels is through being trained on a representative dataset; otherwise, the shift in distribution between the test points and training points could cause poor model performance. As such, increasing the spatial extent of training could lead to a training distribution that has more similar examples to the test set.

Finally, we generated height predictions across the known oil palm plantations in the country (Figure 4). We observed that different states had different plantation age distributions and that most short plantations can be found in the Benue state. This indicates new planting and development of young plantations.

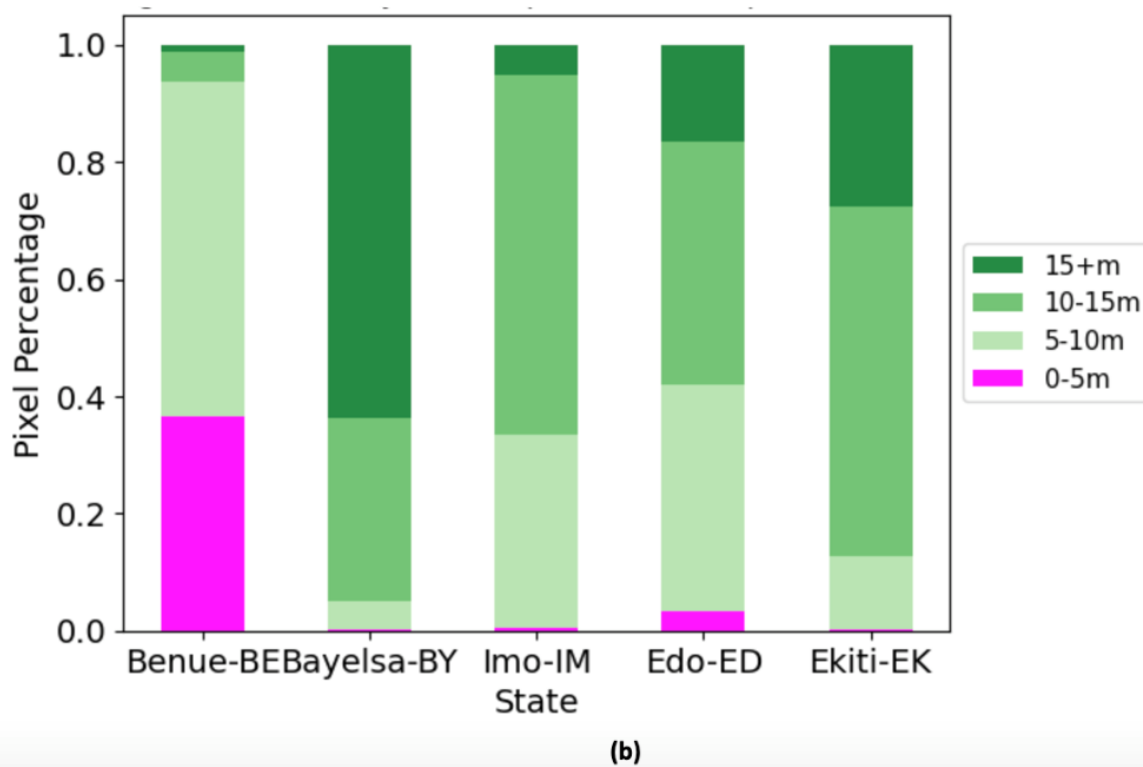




**Figure 3.** Short tree errors when increasing the sampling radius of the training data. (a) Within a certain threshold of a sampling radius, localized calibration reduced error for “short” (<5 m) trees under a standard training regime. (b) More points in total were able to be sampled because the spatial extent of sampling increased, and hyperlocal GEDI may not have been abundant. (c) We observed this alongside a local distribution shift under a certain sampling radius. Including more data points from wider plantation-like regions during training resulted in higher errors.



**Figure 4.** Cont.



**Figure 4.** Height predictions using best model for known oil palm plantations in Nigeria. (a) “Short” plantations are mapped in magenta and dominate the Benue state. Other states see less recent planting activity. (b) Stacked bar chart depicts frequency by height class for the 5 states with the most plantation pixels.

#### 4. Discussion

Following the findings from [20], it can be expected that national- or local-scale models tend to have higher accuracies than global-scale products. Our results hold true to this pattern. However, our model is significant for the degree by which it outperformed the standard canopy height product. This suggests that calibration on sparser sites could support more accurate prediction in non-forest tree cover. Our model, trained at the country-level scale of Nigeria, has a smaller training extent compared with both the GLAD and ETH models. The GLAD model trained tree ensembles for each individual 1 degree cell based on training data sampled from neighbors within a 12 degree radius. The “local” scale proposed in the GLAD CHM resulted in land with a very wide range of management practices, different biomes, and climate zones, even though the sampling regime attempted to train a more localized model [20]. The ETH model trained its CNN on satellite images sampled globally, but the mapping of texture features to a canopy height may not follow the same relationship across tree cover with heterogeneous appearances [19].

When comparing their model’s test accuracy on ALS field data to a GEDI RH95 validation set, the authors of [20] found that higher performance on ALS could be explained by a skew toward tall, dense forest regions sampled in the airborne LiDAR study’s test distribution. They suggested that it might be easier to predict height for these tall forests [20]. This stands alongside studies demonstrating that the GEDI holds the highest uncertainty in areas where canopy cover is patchy and heterogeneous [28]. Our experimentation conducted a critical evaluation of the performance of canopy height maps in sparse small-holder plantation settings and tested potential ways to improve performance for models making these predictions. The plantation mask provided greater confidence that a sampled GEDI footprint would be over a region with at least some tree cover, although it could be interspersed with other types of non-tree land cover. We also observed that there were not many tall (>30 m) samples in either the plantations or the larger region.

Previous work has cited signal saturation as a significant reason for the decline in performance at the very short or tall “extremes” of the height spectrum [22]. We suggest that the reported superior performance of the ETH CHM might come at the expense of performance in shorter regions. Because the ETH CHM sought to improve predictions for tall forests (which are less frequent globally compared with low- or zero-vegetation regions), sample loss was inversely weighted to the frequency of occurrence of the sample height [19]. Through this weighting, the models’ precision for a wide range of low-to-no-vegetation short samples may have been negatively impacted in the process of improving the representation of tall forests. While the GLAD CHM did not explicitly focus on improving performance at tall heights, it likewise suffered from lower performance in short regions due to a methodology that explicitly discarded short or young trees from consideration [20]. Another potential challenge with using ETH’s methodology over managed plantation settings is that management practices may generate a stronger texture signal compared with how much tree height affects the texture [19]. Furthermore, the relationship between height and textural changes may appear differently in forests and managed plantations, and distribution shifts between the train and test distributions may cause the methods to perform worse on managed trees.

Within a certain sampling radius, the strategy of local calibration to create more tailored height prediction models was effective at reducing errors (Figure 3). This suggests that around plantation regions (where our ground-verified fields are located), there must be a minimum importance paid to shorter trees to accurately estimate the height of managed tree fields. As the radius of sampling increased, we included more points that were drawn from a different distribution than the field points (Figure 3b). These points comparatively affected the model more, suggesting some trade-off between the quantity of data and how different the train and test distributions became following an influx of new training data. An inflection point was present where the sampling radius changes occurred and where sub-five-meter points initially became less common as the sampling radius increased from zero meters and then started to become more common again. This U-shaped curve may have happened in part because trees less than five meters tall are more likely to be found in managed settings than natural forests, whereas in a very wide sampling regime, non-tree land cover types are also included. A different study predicting canopy heights with RF also found that local prediction error trends were thresholded as the spatial extent of sampling reached 50 km [40]. As more non-plantation tree cover was sampled, the sub-five-meter trees may have become harder to predict because the samples were becoming more different than the test points. But as even more distant, diverse locations were sampled overall, the new data points were useful because they provided more heterogeneity. Some characteristics of non-plantation land cover types may have been similar to those of young plantations, leading to the improved height estimation under national-scale sampling.

In Landsat-based models, signal saturation at both the low and high ends of the height range can practically be reduced through highly localized model calibration [22]. However, GEDI coverage is sparsest around the equator due to the path of ISS orbital crossings [44]. As such, there may be a trade-off between the improved saturation offered by localization and the reduced bias from reducing the number of total training samples. This proved to be a significant challenge in mapping the heights of oil palm plantations in Nigeria, as the region has particularly low GEDI coverage. We found that there were insufficient data at the hyper-local level (3 km buffered sampling), where only 2444 footprints in a 3 km range of our field sites were retrievable across 3 years of GEDI data collection. Despite having significantly less training data, the three-kilometer model still performed reasonably well compared with the larger models. However, further expanding the localization range to just 5 km (6666 points) resulted in a model that performed with comparable error metrics to training with a national-level sampling regime. Although localization was effective for reducing signal saturation for the shortest plantations (<5 m), it did not appear to affect the overall errors (Figure 3).

## 5. Conclusions

We evaluated existing GEDI canopy height products against a set of tree height data collected in the field with a standardized procedure. We found that existing global products based on GEDI and optical satellite data cannot be reliably applied to the task of mapping heights for planted trees, as plantations have distinct characteristics from other types of tree cover. We also found that locally trained models performed much better, with significantly improved absolute and percent errors. Given the relatively sparse coverage of the GEDI at tropical latitudes, the degree to which training should be localized requires balancing the trade-off between larger sample sizes and potential distribution shifts as one includes a broader swath of training data. By comparing different feature sets, we also showed that including multiple spectral bands and indices allowed for better height predictions in Nigerian plantations. Robust wall-to-wall maps of the heights and locations of planted trees are necessary to better understand the value of managed trees, to correct historical estimates of carbon sequestered in nature-based solutions, and to understand patterns of growth for targeted sustainable development interventions. We found that models trained at a relatively local scale can perform well even in regions with low GEDI coverage. In general, predicting heights over plantations requires additional effort to improve performance for short trees, as shorter (typically younger) trees are abundant and play a vital economic and environmental role. Further improvements to the task of plantation mapping may continue to emphasize these short trees. Additionally, subsequent data collection from a renewed GEDI or other spaceborne LiDAR missions may allow for higher data availability and further improved performance in locally calibrated models, altering the trade-off between data availability and distributional representativeness.

**Author Contributions:** Conceptualization, A.T., I.N., A.J., U.A. and D.B.L.; methodology, A.T. and D.B.L.; software, A.T.; validation, A.T.; formal analysis, A.T.; investigation, A.T.; resources, I.N., A.J., U.A. and D.B.L.; data curation, I.N., A.J. and U.A.; writing—original draft preparation, A.T.; writing—review and editing, A.T., I.N. and D.B.L.; visualization, A.T. and D.B.L.; supervision, D.B.L.; project administration, D.B.L.; funding acquisition, I.N., A.J., U.A. and D.B.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** The height map for oil palm plantations is available as an ImageCollection in Google Earth Engine, asset name (“projects/ee-peregriney/assets/releaf-fieldpoints-height-predfig-finalic”).

**Acknowledgments:** Thanks to Muyideen Dauda for leading field data collection.

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

## Abbreviations

The following abbreviations are used in this manuscript:

GEDI	Global Ecosystem Dynamics Investigation
TOF	Trees outside forests
RF	Random forest
MAE	Mean absolute error
RMSE	Root mean square error

## References

1. Cavender-Bares, J.M.; Nelson, E.; Meireles, J.E.; Lasky, J.R.; Miteva, D.A.; Nowak, D.J.; Pearse, W.D.; Helmus, M.R.; Zanne, A.E.; Fagan, W.F.; et al. The hidden value of trees: Quantifying the ecosystem services of tree lineages and their major threats across the contiguous US. *PLoS Sustain. Transform.* **2022**, *1*, e0000010. [[CrossRef](#)]
2. Le Toan, T.; Quegan, S.; Davidson, M.; Balzter, H.; Paillou, P.; Papathanassiou, K.; Plummer, S.; Rocca, F.; Saatchi, S.; Shugart, H.; et al. The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle. *Remote Sens. Environ.* **2011**, *115*, 2850–2860. [[CrossRef](#)]

3. Hunter, M.O.; Keller, M.; Victoria, D.; Morton, D.C. Tree height and tropical forest biomass estimation. *Biogeosciences* **2013**, *10*, 8385–8399. [[CrossRef](#)]
4. Crowther, T.W.; Glick, H.B.; Covey, K.R.; Bettigole, C.; Maynard, D.S.; Thomas, S.M.; Smith, J.R.; Hintler, G.; Duguid, M.C.; Amatulli, G.; et al. Mapping tree density at a global scale. *Nature* **2015**, *525*, 201–205. [[CrossRef](#)] [[PubMed](#)]
5. Zomer, R.J.; Neufeldt, H.; Xu, J.; Ahrends, A.; Bossio, D.; Trabucco, A.; van Noordwijk, M.; Wang, M. Global Tree Cover and Biomass Carbon on Agricultural Land: The contribution of agroforestry to global and national carbon budgets. *Sci. Rep.* **2016**, *6*, 29987. [[CrossRef](#)] [[PubMed](#)]
6. Thomas, N.; Baltezar, P.; Lagomasino, D.; Stovall, A.; Iqbal, Z.; Fatoyinbo, L. Trees outside forests are an underestimated resource in a country with low forest cover. *Sci. Rep.* **2021**, *11*, 7919. [[CrossRef](#)] [[PubMed](#)]
7. Lund, H. When Is a Forest Not a Forest? *J. For.* **2002**, *100*, 21–28. Available online: <http://xxx.lanl.gov/abs/https://academic.oup.com/jof/article-pdf/100/8/21/22555273/jof0021.pdf> (accessed on 25 October 2023).
8. Fang, S.; Xue, J.; Tang, L. Biomass production and carbon sequestration potential in poplar plantations with different management patterns. *J. Environ. Manag.* **2007**, *85*, 672–679. [[CrossRef](#)]
9. Zhou, X.; Zhu, H.; Wen, Y.; Goodale, U.M.; Li, X.; You, Y.; Ye, D.; Liang, H. Effects of understory management on trade-offs and synergies between biomass carbon stock, plant diversity and timber production in eucalyptus plantations. *For. Ecol. Manag.* **2018**, *410*, 164–173. [[CrossRef](#)]
10. Miller, D.C.; Muñoz-Mora, J.C.; Christiaensen, L., Do Trees on Farms Matter in African Agriculture? In *Agriculture in Africa: Telling Myths from Facts*; World Bank eLibrary: Washington, DC, USA, 2017; pp. 115–121. [[CrossRef](#)]
11. Khatun, K.; Maguire-Rajpaul, V.A.; Asante, E.A.; McDermott, C.L. From Agroforestry to Agroindustry: Smallholder Access to Benefits From Oil Palm in Ghana and the Implications for Sustainability Certification. *Front. Sustain. Food Syst.* **2020**, *4*, 29. [[CrossRef](#)]
12. Sharma, S.; Rana, V.S.; Prasad, H.; Lakra, J.; Sharma, U. Appraisal of Carbon Capture, Storage, and Utilization Through Fruit Crops. *Front. Environ. Sci.* **2021**, *9*, 700768. [[CrossRef](#)]
13. Diao, J.; Liu, J.; Zhu, Z.; Wei, X.; Li, M. Active forest management accelerates carbon storage in plantation forests in Lishui, southern China. *For. Ecosyst.* **2022**, *9*, 100004. [[CrossRef](#)]
14. Dubayah, R.; Blair, J.B.; Goetz, S.; Fatoyinbo, L.; Hansen, M.; Healey, S.; Hofton, M.; Hurtt, G.; Kellner, J.; Luthcke, S.; et al. The Global Ecosystem Dynamics Investigation: High-resolution laser ranging of the Earth’s forests and topography. *Sci. Remote Sens.* **2020**, *1*, 100002. [[CrossRef](#)]
15. Luther, J.E.; Fournier, R.A.; van Lier, O.R.; Bujold, M. Extending ALS-Based Mapping of Forest Attributes with Medium Resolution Satellite and Environmental Data. *Remote Sens.* **2019**, *11*, 1092. [[CrossRef](#)]
16. Coomes, D.A.; Šafka, D.; Shepherd, J.; Dalponte, M.; Holdaway, R. Airborne laser scanning of natural forests in New Zealand reveals the influences of wind on forest carbon. *For. Ecosyst.* **2018**, *5*, 10. [[CrossRef](#)]
17. Wilkes, P.; Jones, S.D.; Suarez, L.; Mellor, A.; Woodgate, W.; Soto-Berelov, M.; Haywood, A.; Skidmore, A.K. Mapping Forest Canopy Height Across Large Areas by Upscaling ALS Estimates with Freely Available Satellite Data. *Remote Sens.* **2015**, *7*, 12563–12587. [[CrossRef](#)]
18. Neuenschwander, A.; Pitts, K. The ATL08 land and vegetation product for the ICESat-2 Mission. *Remote Sens. Environ.* **2019**, *221*, 247–259. [[CrossRef](#)]
19. Lang, N.; Jetz, W.; Schindler, K.; Wegner, J.D. A High-Resolution Canopy Height Model of the Earth. *arXiv* **2022**, arXiv:cs.CV/2204.08322.
20. Potapov, P.; Li, X.; Hernandez-Serna, A.; Tyukavina, A.; Hansen, M.C.; Kommareddy, A.; Pickens, A.; Turubanova, S.; Tang, H.; Silva, C.E.; et al. Mapping global forest canopy height through integration of GEDI and Landsat data. *Remote Sens. Environ.* **2021**, *253*, 112165. [[CrossRef](#)]
21. Wang, H.; Seaborn, T.; Wang, Z.; Caudill, C.C.; Link, T.E. Modeling tree canopy height using machine learning over mixed vegetation landscapes. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *101*, 102353. [[CrossRef](#)]
22. Healey, S.P.; Yang, Z.; Gorelick, N.; Ilyushchenko, S. Highly Local Model Calibration with a New GEDI LiDAR Asset on Google Earth Engine Reduces Landsat Forest Height Signal Saturation. *Remote Sens.* **2020**, *12*, 2840. [[CrossRef](#)]
23. Becker, A.; Russo, S.; Puliti, S.; Lang, N.; Schindler, K.; Wegner, J.D. Country-wide retrieval of forest structure from optical and SAR satellite imagery with deep ensembles. *ISPRS J. Photogramm. Remote Sens.* **2023**, *195*, 269–286. [[CrossRef](#)]
24. Lang, N.; Schindler, K.; Wegner, J.D. Country-wide high-resolution vegetation height mapping with Sentinel-2. *Remote Sens. Environ.* **2019**, *233*, 111347. [[CrossRef](#)]
25. Huang, Q.; Xu, J.; Wong, J.P.; Radeloff, V.C.; Songer, M. Prioritizing global tall forests toward the 30-by-30 goal. *Conserv. Biol.* **2023**, e14135. [[CrossRef](#)] [[PubMed](#)]
26. Danylo, O.; Pirker, J.; Lemoine, G.; Ceccherini, G.; See, L.; McCallum, I.; Hadi.; Kraxner, F.; Achard, F.; Fritz, S. A map of the extent and year of detection of oil palm plantations in Indonesia, Malaysia and Thailand. *Sci. Data* **2021**, *8*, 96. [[CrossRef](#)] [[PubMed](#)]
27. Paterson, R.R.M.; Chidi, N.I. Climate Refuges in Nigeria for Oil Palm in Response to Future Climate and Fusarium Wilt Stresses. *Plants* **2023**, *12*, 764. [[CrossRef](#)] [[PubMed](#)]
28. Roy, D.P.; Kashongwe, H.B.; Armston, J. The impact of geolocation uncertainty on GEDI tropical forest canopy height estimation and change monitoring. *Sci. Remote Sens.* **2021**, *4*, 100024. [[CrossRef](#)]

29. Li, X.; Wessels, K.; Armston, J.; Hancock, S.; Mathieu, R.; Main, R.; Naidoo, L.; Erasmus, B.; Scholes, R. First validation of GEDI canopy heights in African savannas. *Remote Sens. Environ.* **2023**, *285*, 113402. [[CrossRef](#)]
30. Enterkine, J.; Campbell, B.A.; Kohl, H.; Glenn, N.F.; Weaver, K.; Overoye, D.; Danke, D. The potential of citizen science data to complement satellite and airborne lidar tree height measurements: Lessons from The GLOBE Program. *Environ. Res. Lett.* **2022**, *17*, 075003. [[CrossRef](#)]
31. Hansen, M.C.; Potapov, P.V.; Moore, R.; Hancher, M.; Turubanova, S.A.; Tyukavina, A.; Thau, D.; Stehman, S.V.; Goetz, S.J.; Loveland, T.R.; et al. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science* **2013**, *342*, 850–853. [[CrossRef](#)]
32. Beck, J.; Wirt, B.; Armston, J.; Hofton, M.; Luthcke, S.; Tang, H. GLOBAL Ecosystem Dynamics Investigation (GEDI) Level 2 User Guide. 2021. Available online: [https://lpdaac.usgs.gov/documents/986/GEDI02\\_UserGuide\\_V2.pdf](https://lpdaac.usgs.gov/documents/986/GEDI02_UserGuide_V2.pdf) (accessed on 25 October 2023).
33. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [[CrossRef](#)]
34. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, D.W. *Monitoring Vegetation Systems in the Great Plains with ERTS*; NASA: Washington, DC, USA, 1973.
35. McFEETERS, S.K. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *Int. J. Remote Sens.* **1996**, *17*, 1425–1432. [[CrossRef](#)]
36. Gao, B. NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sens. Environ.* **1996**, *58*, 257–266. [[CrossRef](#)]
37. Gamon, J.A.; Field, C.B.; Goulden, M.L.; Griffin, K.L.; Hartley, A.E.; Joel, G.; Penuelas, J.; Valentini, R. Relationships between NDVI, Canopy Structure, and Photosynthesis in Three Californian Vegetation Types. *Ecol. Appl.* **1995**, *5*, 28–41. [[CrossRef](#)]
38. Pascual, C.; García-Abril, A.; Cohen, W.B.; Martín-Fernández, S. Relationship between LiDAR-derived forest canopy height and Landsat images. *Int. J. Remote Sens.* **2010**, *31*, 1261–1280.
39. Nandy, S.; Srinet, R.; Padalia, H. Mapping Forest Height and Aboveground Biomass by Integrating ICESat-2, Sentinel-1 and Sentinel-2 Data Using Random Forest Algorithm in Northwest Himalayan Foothills of India. *Geophys. Res. Lett.* **2021**, *48*, e2021GL093799.
40. Jin, S.; Su, Y.; Gao, S.; Hu, T.; Liu, J.; Guo, Q. The Transferability of Random Forest in Canopy Height Estimation from Multi-Source Remote Sensing Data. *Remote Sens.* **2018**, *10*, 1183. [[CrossRef](#)]
41. Fayad, I.; Baghdadi, N.; Bailly, J.S.; Barbier, N.; Gond, V.; Hajj, M.E.; Fabre, F.; Bourguine, B. Canopy Height Estimation in French Guiana with LiDAR ICESat/GLAS Data Using Principal Component Analysis and Random Forest Regressions. *Remote Sens.* **2014**, *6*, 11883–11914. [[CrossRef](#)]
42. Belgiu, M.; Drăguț, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 24–31. [[CrossRef](#)]
43. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
44. Dubayah, R.; Armston, J.; Healey, S.P.; Bruening, J.M.; Patterson, P.L.; Kellner, J.R.; Duncanson, L.; Saarela, S.; Ståhl, G.; Yang, Z.; et al. GEDI launches a new era of biomass inference from space. *Environ. Res. Lett.* **2022**, *17*, 095001. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.