Synergetic of PALSAR-2 and Sentinel-1A SAR Polarimetry for Retrieving Aboveground Biomass in Dipterocarp Forest of Malaysia

Hamdan Omar *, Muhamad Afizzul Misman [ID] and Abd Rahman Kassim

Geoinformation Programme, Division of Forestry and Environment, Forest Research Institute Malaysia, Selangor 52109, Malaysia; afizzul@frim.gov.my (M.A.M.); rahmank@frim.frim.gov.my (A.R.K.)
* Correspondence: hamdanomar@frim.gov.my; Tel.: +60-3-6279-7200

Abstract: Space borne synthetic aperture radar (SAR) data have become one of the primary sources for aboveground biomass (AGB) estimation of forests. However, studies have indicated that limitations occur when a single sensor system is employed, especially in tropical forests. Hence, there is potential for improving estimates if two or more different sensor systems are used. Studies on integrating multiple sensor systems for estimation of AGB over Malaysia's tropical forests are scarce. This study investigated the use of PALSAR-2 L-band and Sentinel-1A C-band SAR polarizations to estimates the AGB over 5.25 million ha of the lowland, hill, and upper hill forests in Peninsular Malaysia. Polarized images, i.e., HH–HV from PALSAR-2 and VV–VH from Sentinel-1A have been utilized to produce several variables for predictions of the AGB. Simple linear and multiple linear regression analysis was performed to identify the best predictor. The study concluded that although limitations exist in the estimates, the combination of all polarizations from both PALSAR-2 and Sentiel-1A SAR data able to increase the accuracy and reduced the root means square error (RMSE) up to 14 Mg ha^{-1} compared to the estimation resulted from single polarization. A spatially distributed map of AGB reported the total AGB within the study area was about 1.82 trillion Mg of the year 2016.

Keywords: aboveground biomass; tropical forest; microwave sensor system

1. Introduction

Aboveground biomass (AGB) includes all vegetation above the ground (i.e., stems, branches, bark, seeds, flowers, and foliage of live plants) and approximately 50% of its composition is carbon [1]. AGB usually measures in metric tons of dry matter per hectare (e.g., t ha^{-1} or Mg ha^{-1}) or in metric tons of carbon per hectare (e.g., t C ha^{-1} or Mg C ha^{-1}). The United Nations Framework Convention on Climate Change (UNFCCC) identified it as an Essential Climate Variable (ECV). Therefore, accurate information on biomass stock in world forests is necessary to reduce uncertainties and to fill the knowledge gaps of the climate system [2]. Further strong impetus to improve methods for measuring global biomass comes from the reduction of emissions due to deforestation and forest degradation (REDD) mechanism, which was introduced in the UNFCCC Committee of the Parties (COP-13) Bali Action Plan. REDD which is now popular with REDD+ (with additional elements of carbon stock enhancement and biodiversity conservation) is dedicated to the developing countries around the world including Malaysia. Its implementation relies fundamentally on systems to assess available carbon stock and monitor changes due to loss of biomass from deforestation and forest degradation [3], which are amalgamated in a system called monitoring, reporting, and verification (MRV).

Remote sensing has been recognized as one of the primary spatial inputs for this process [4–6]. Satellite remote sensing technologies are currently widely tested and suggested as a tool in REDD+

MRV. Along with scientific programs and field tests, there is also a debate as to the overall feasibility and cost–benefit ratio of remote sensing approaches, depending on the wide range of ecosystem and land use conditions as well as the range of approaches to carbon credit accounting [7].

In many parts of the world, especially in tropical region, the frequent cloud conditions often restrain the acquisition of high-quality remotely sensed data by optical sensors. The acquisition of cloud-free, wall-to-wall optical satellite images in tropical countries is almost impossible [8]. Thus, SAR data become the only feasible way of acquiring remotely sensed data within a given timeframe because the SAR systems are independent of cloud coverage, weather, and light conditions. Due to this unique feature compared with optical sensor data, the SAR data have been used extensively in many fields, including forest-cover identification and mapping, discrimination of forest from other land covers, and forest biomass estimation.

Previous studies demonstrated that the L-band polarimetry backscatter tends to saturate at certain levels of biomass, and hence limits the accuracy of estimates [9–11]. However, the saturation level varies with the type and structure of the forest. It was demonstrated that the sensitivity of SAR polarimetry is depending on the structure, density, and tree elements (i.e., trunk/stem, branches, and leaves) of the forests [12]. Other than these issues, several other inter-related issues can affect the biomass estimations using remotely sensed data. These issues can be generalized into three major groups, which are (i) the natural conditions of the forest, (ii) the forest management system being practiced, and (iii) the technical issues related to the remote sensing system being used [13].

Short wavelength SAR sensors on board several satellites such as the Earth Resources Satellite (ERS-1), Radarsat, and Environmental Satellite (Envisat) have been used to quantify forest biomass. A number of studies have been conducted in relatively homogeneous or young forests, but the signal tends to saturate at low biomass (100–200 Mg ha^{-1}) [14–16]. However, L-band SAR has shown better potential in retrieving the biomass of forests, including those in the tropics [17–21]. Recently, there has been rising interest in integrating data from several SAR sensors and SAR with optical sensors to improve the accuracy of biomass estimates [22,23].

In Malaysia, there are limited studies on the applications of SAR for estimating biomass. Out of many studies conducted worldwide, very few have been done in Malaysia [11,24,25]. This indicates that the potential, limitations, and advances of L-band SAR in estimating tropical forest in Malaysia are not extensively explored. Methods of applying this SAR system are also scarcely exploited. Therefore, the objective of this study is to explore the synergy of SAR sensors, i.e., PALSAR-2 L-band and Sentinel-1A C-band for estimation of AGB in inland dry dipterocarp forest in Peninsular Malaysia. This study highlights and discusses advantages and limitations of this technique.

2. Materials and Methods

2.1. The Study Area

The study area comprised lowland, hill, and upper hill dipterocarp forests, which are categorized based on land altitude, i.e., <300, 300–750, and 750–1200 m, respectively. These forests are major, occupying about 5,257,395 ha or about 89% out of the total forested land (i.e., about 5.9 million ha) in Peninsular Malaysia. These forests occur within the entire Peninsular Malaysia, which has an extent between 1–7° latitude and 99–105° longitude. These forests embrace all the well-drained primary forests of the plains, undulating land, and foothills and hill terrain up to about 750 m altitude. Trees from the family of Dipterocarpaceae are dominant species, which make the forests major timber production areas in Peninsular Malaysia. Almost the entire area (i.e., 4.9 million ha) is categorized as Reserve Forest which is meant for production and protection. About 1.98 million ha have been allocated for protection forests in the form of national parks, wildlife sanctuaries, and nature reserves [26]. The most common tree species found in this forest come from the genera such as Shorea, Hopea, Dipterocarpus, Dryobalanops, Neobalacarpus, Anisoptera, and Vatica. The remaining forested

land is comprised of peat swamp, mangrove, and montane forests. Figure 1 shows the distribution of major forest types in Peninsular Malaysia.
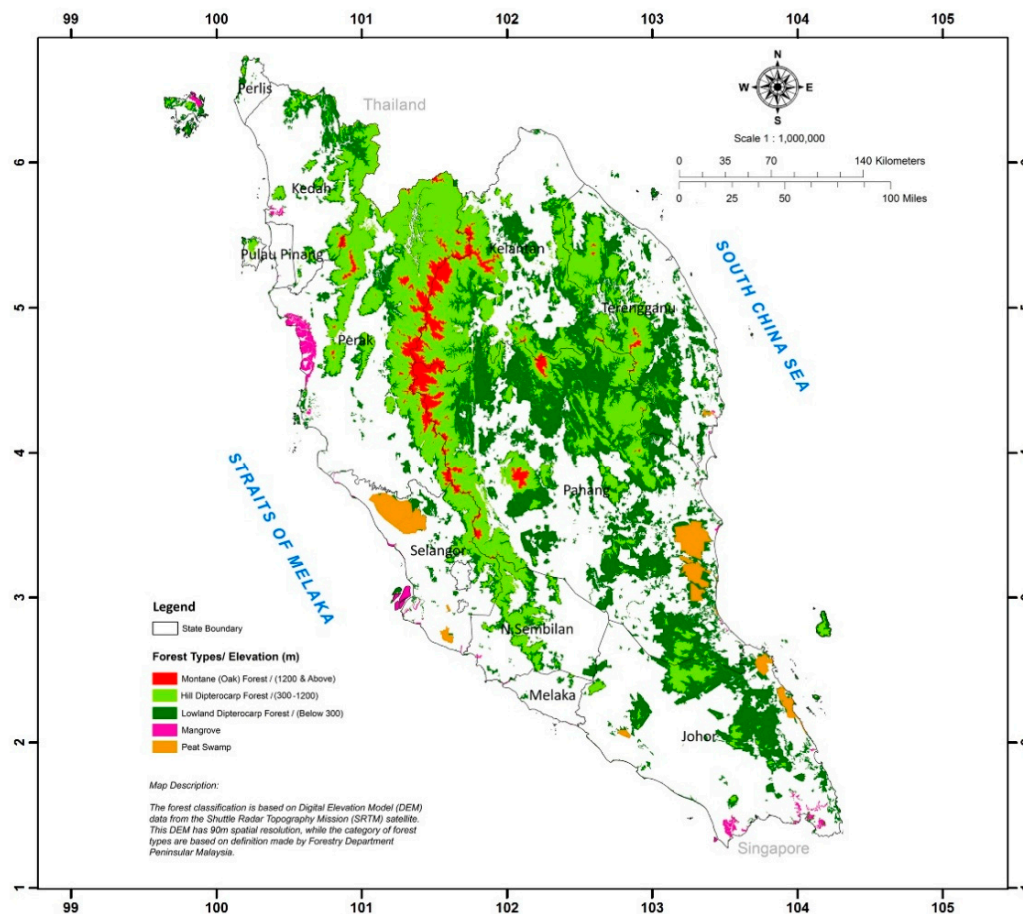

Figure 1. Forest types in Peninsular Malaysia.

2.2. Satellite Datasets

2.2.1. Satellite Images Acquisition

The satellite datasets that have been used in this study came from two satellites, which are; (i) Advanced Land Observing Satellite 2 (ALOS-2) that carries Phased Array type L-band Synthetic Aperture Radar-2 (PALSAR-2) on-board and (ii) Sentinel-1A that carries C-band imaging SAR sensor. Table 1 summarizes the properties of the data and Figure 2 shows both datasets that have been processed and used in this study.

ALOS-2 is the successor of the ALOS, but the structure of the new satellite is quite different from its predecessor. ALOS was launched in January 2006 and brought the Phased Array type L-band Synthetic Aperture Radar (PALSAR) on-board. After five years of observations, it stopped transmitting in April 2011. ALOS-2 was then launched on 24 May 2014, which carried the PALSAR-2 sensor. PALSAR-2 is currently operating and producing L-band SAR data, that has similar (with some advancements) characteristics with PALSAR. The data were acquired from Earth Observation Research Center (EORC) under Japan Aerospace Exploration Agency (JAXA). The data was acquired under the Kyoto and Carbon (K&C) Initiave, a research agreement between Forest Research Institute Malaysia (FRIM) and JAXA, whereby FRIM has special permission to access the PALSAR-2 product at all imaging modes and resolutions. JAXA also provides free access of PALSAR-2 mosaic product at 25-meter resolution for public which is available at http://www.eorc.jaxa.jp/ALOS/en/PALSAR_fnf/data/index.htm.

The SENTINEL-1 mission is the European Radar Observatory for the Copernicus joint initiative of the European Commission (EC) and the European Space Agency (ESA). The SENTINEL-1 was launched on 3 April 2014 and its mission operating in four exclusive imaging modes with different resolution (down to 5 m) and coverage (up to 400 km). It provides dual polarization capability, very short revisit times, and rapid product delivery. The Sentine-1A data was acquired in Level-1 Ground Range Detected (GRD) format so that radar cross-section of both distributed and point targets can be easily derived from the data. The data is available at https://scihub.copernicus.eu/dhus/#/home and free to download.

Another satellite data that was used in this study was the digital elevation model acquired from the Shuttle Radar Topography Mission (SRTM). This data was used to classify the forest into specified elevation categories, according to the type of forests. It was also used for radiometric terrain correction on both PALSAR-2 and Sentinel-1A images. These data are available at the US Geological Survey's EROS Data Center for download at http://srtm.usgs.gov/index.html.

Table 1. Summary of satellite images used in this study.

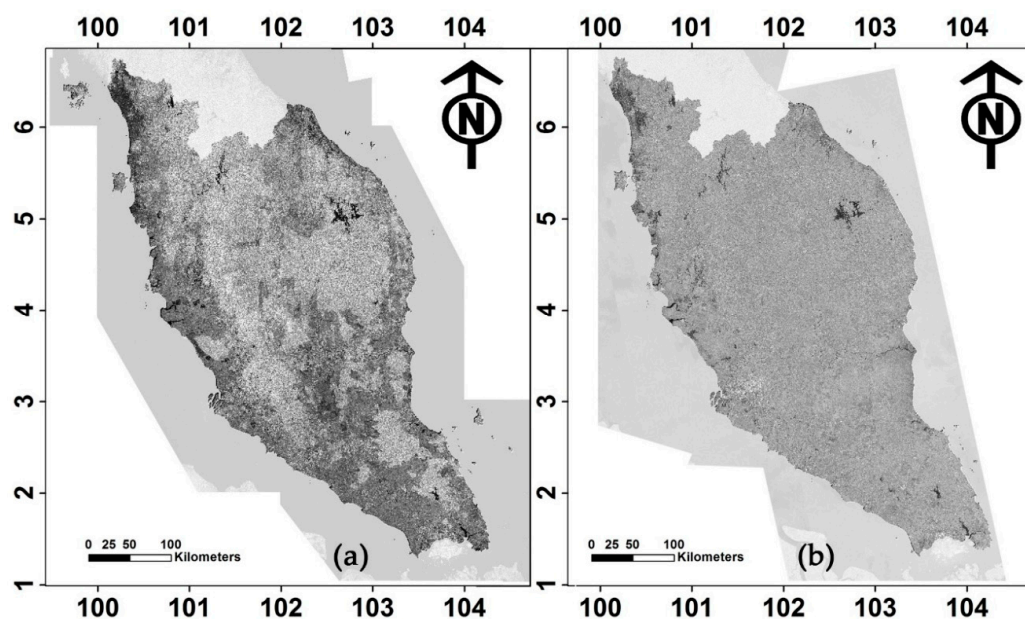| Sensor	Wavelength	Date of Acquisition	Mode/Polarization	No. of Scene	Ground Resolution (m)	Incidence Angle (°)
PALSAR-2	C-band (5.405 GHz)	Between March and June 2016	Fine Beam Dual (FBD)/HH, HV	52	6	29.1–46.0
Sentinel-1A	L-band (1.270 GHz)	November 2016	Interferometric Wide swath (IW) VV, VH	8	9	37

Figure 2. Mosaics of (**a**) PALSAR-2 and (**b**) Sentinel-1A over Peninsular Malaysia, displayed HV and VH polarizations, respectively.

2.2.2. Satellite Image Pre-Processing

PALSAR-2 images that were used in this study came in Level-1.5 product, which means the range and multilook azimuth compressed data is represented by amplitude data. The range coordinates were also converted from slant range to ground range, and map projection was performed.

Level-1 GRD consist of Sentinel-1A products consist of focused SAR data that has been detected, multi-looked and projected to ground range using an Earth ellipsoid model such as WGS84. The ellipsoid projection of the GRD products is corrected using the terrain height specified in the product general annotation. The terrain height used varies in azimuth but is constant in range.

Ground range coordinates are the slant range coordinates projected onto the ellipsoid of the Earth. Pixel values represent detected magnitude. Phase information is lost. The resulting product has approximately square resolution pixels and square pixel spacing with reduced speckle at a cost of reduced geometric resolution. In addition, the GRD products have thermal noise removed to improve the quality of the detected image.

Other than these processes, they are two important pre-processing stages, namely, speckle suppression and radiometric terrain correction. Spatial domain Lee Sigma filter with a kernel size of 7 × 7 pixels was used to remove speckle effect on the images. Digital Elevation Model (DEM) acquired from SRTM was used to minimize terrain shadowing effect on the images. The presence of this effect on SAR imagery was because the signal strengths were dependent on two variables, which were incidence angle and surface roughness or topography of terrain. If slope is facing the SAR transmitter, the signal will become stronger than the other side of the slope. Semi-empirical method was used for radiometric terrain correction [27] and the processes were performed in ENVI/IDL (Harris Corporation, Melbourne, Australia) following the same approach as found in Canty et al. [28]. This process was necessary to normalize both sides of the slopes and it minimized errors towards the end of AGB prediction.

2.2.3. Satellite Image Calibration

The objective of SAR calibration is to provide imagery in which the pixel values can be directly related to the radar backscatter of the scene. The PALSAR-2 image that was used in this study was built on a 16-bit data type and all pixels have digital numbers (DN) ranging from 0 to 65,535. These DNs however do not represent the radar signal of features or objects on the ground. Therefore, the DNs have to be converted to backscatter (i.e., the returned radar signals) known as Normalized Radar Cross Section (NRCS) and represented as σ^0 in decibels (dB). The equation that was used for the calculation of NRCS for PALSAR are slightly different from other sensors in that the usual sine term has already been included in the DN values. Thus, for the products stored at Level 1.5 and above, the equation for NRSC of any of the polarization component can be obtained by the following formula with single calibration factor (CF), which can be expressed as follows [29].

$$\sigma^0_{dB} = 10 \cdot log_{10}\left(DN^2\right) - 83 \qquad (1)$$

The Sentinel-1A product uses radiometric calibration look-up table (LUT) to do the calibration. This was performed on Sentinel Application Platform (SNAP) tool, a software that was designed specifically for Sentinel-1 products processing and it available for free at http://step.esa.int/main/download/. The essential conversion of amplitude to *DN* and from *DN* to sigma nought were done automatically on SNAP and once the sigma nought values was obtained, the computation of backscatter (σ^0_{dB}) can be performed as

$$\sigma^0_{dB} = 10 \cdot log_{10}\sigma^0 \qquad (2)$$

2.3. Forest-Non-Forest Classification

Forest-non-forest (FNF) classification was performed on PALSAR-2 polarizations images to delineate the forests from other land cover. This process is critical to define the boundary of forests and to ensure that the estimated AGB did not include other types of vegetation. The reason was that the forests are often confused with rubber, teak, and other timber tree plantations, which are common in Peninsular Malaysia and they appear almost identical on both HH and HV polarizations. Instead of

using only the original backscatter HH and HV polarizations, an attempt has been also made to derive other image variables derived from PALSAR HH and HV images. Image variables, namely (i) simple band ratio (HH/HV), (HV/HH), (ii) average (HH + HV/2), and (iii) square root of products ($\sqrt{}$(HH × HV)) were produced.

The incorporation of texture measure also can improve classification of spatially distributed pixels on an image. Gray-level co-occurrence matrix (GLCM) uses a gray-tone spatial dependence matrix to calculate texture values. This is a matrix of relative frequencies with which pixel values occur in two neighboring processing windows separated by a specified distance and direction. For this purpose, texture has been defined as repeating pattern of local variations in image intensity which is too fine to be distinguished as separate class at the observed resolution. Thus, a connected set of pixels satisfying given gray-level properties which occur repeatedly in an image region constitute a textured region. A mean-type GLCM was applied to the original HH and HV polarizations to produce textured images with clearer definitions of the objects on the images [30].

These inputs were used for the FNF classification and the Maximum Likelihood Classifier algorithms with nearest neighbor technique was applied. The forests were then further classified into several forest types by using the DEM from SRTM.

2.4. Forest Survey Data

The sampling design in this study modified from the standard operating procedure (SOP) that was developed by Winrock International [31], which follows the Intergovernmental Panel on Climate Change (IPCC) standards [1]. A cluster comprises of four plots and the design is shown in Figure 3. The plot was designed in circular with smaller nests inside. The biggest nest measures 20 m in radius, followed by the smaller nests, measuring 12 and 4 m. The sizes of trees were measured according to the nest sizes, which is summarized in Table 2. Depending on the nest size, it indicates that not all stands were measured in a single plot. In additional to these nests, there is another small nest measuring 2 m in radius, which is used to count saplings (i.e., trees measuring <10 cm in diameter at breast height (dbh) and >1.3 m in height). Clustering of plots at each sampling unit was recommended for natural forest areas and areas that have been selectively logged. The sampling system was designed in such a way to make the data collection process easier and faster, but reliable and representative for a particular forest stratum.
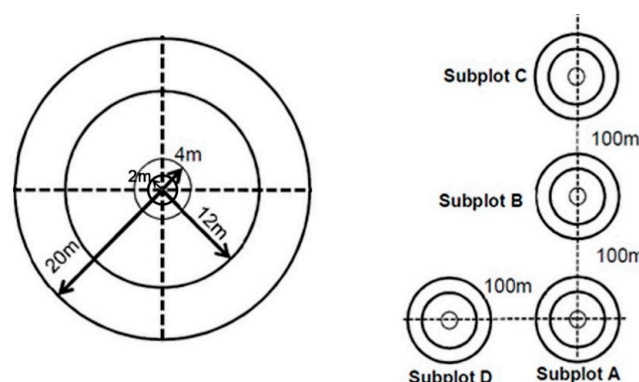

Figure 3. Layout of the sampling plot and sampling design of a cluster of sample.

A forest ecosystem normally has five terrestrial carbon pools, which are; (i) aboveground living biomass, (ii) belowground living biomass, (iii) deadwood, (iv) non-tree vegetation and litters, and (v) soil. However, one of the most significant carbon pools is aboveground biomass as it the easiest and the most practical pool to assess, while being representative to an ecosystem. Aboveground biomass comprises all the living components of a tree, including stems, branches, and leaves. Allometric

functions are the best way that AGB can be estimated. In this study, a published allometric function for dry inland forest in Asia region was used to estimate the AGB of living trees [32].

$$AGB = [\exp(-1.803 - 0.976E + 0.976\ln(\rho) + 2.673\ln(D) - 0.0299[\ln(D)]^2] \qquad (3)$$

AGB denotes aboveground biomass (kg/tree), E represents bioclimatic variable, ρ is wood specific gravity/wood density, and D is dbh.

Table 2. Summary of living trees measurement in a plot.

Nest Radius (m)	Size	Diameter at Breast Height, dbh (cm)
2	Sapling	<5 cm (dbh) & >1.3 m (height)
4	Small	10.0–19.9
12	Medium	20.0–39.9
20	Large	≥40.0

A total number of 332 plots have been surveyed between years 2014 and 2016 and were used as sample plots information for this study. The forest survey was conducted in a number of field trips that cover mainly the central parts of Peninsular Malaysia. The States include Terengganu, Pahang, Johor, Negeri Sembilan, Selangor, Perak, Kelantan, and Perlis in the north. In each plot, every tree which meets the dbh size in the nest radius was inventoried. Species of every stand being inventoried was also recorded. Position (coordinate) of each plot was recorded at the center by using hand-held Global Positioning System (GPS) (Trimble Inc., Sunnyvale, CA, USA). The locations of all plots were post-processed by using base position data from the Department of Survey and Mapping Malaysia to ensure the accuracy of the position acquired. Locations of the sample plots are shown in Figure 4 and a summary of the sample plots is given in Table 3.
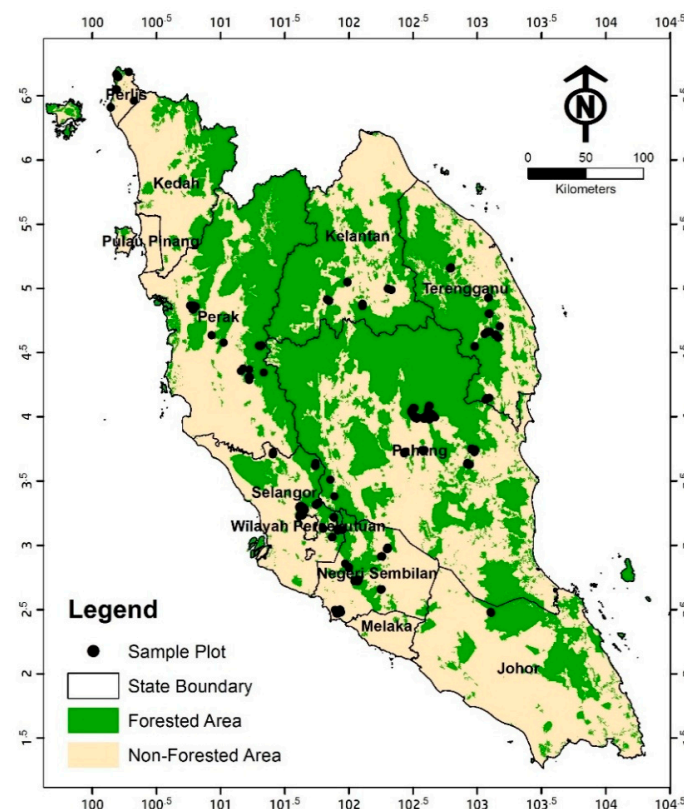

Figure 4. Distribution of ground sample plots within the study area.

Table 3. Summary of distribution of sample plots within the study area.

State	No. of Plot			Total
	Lowland	Hill	Upper Hill	
Perlis	8	2	0	10
Terengganu	48	18	4	70
Pahang	46	18	8	72
Johor	6	0	0	6
Negeri Sembilan	40	14	0	54
Selangor	42	18	8	68
Perak	20	8	4	32
Kelantan	12	8	-	20
Total	222	86	24	332

2.5. Correlation Analysis

The backscatter values from both PALSAR-2 and Sentinel-1A were extracted from the images, which represented HH, HV, VV, and VH polarizations. The AGB at the sample plots on the ground was then correlated with the corresponding backscatter of these polarizations by using linear regression method. Instead of using the single polarization as a variable, several other variables have been derived by manipulating these single polarizations. This manipulation was performed to produce variety of image variables and that to examine the roles of polarization in estimating AGB. Table 4 lists the variables that have been derived from the individual PALSAR-2 and Sentinel-1A and combination of polarizations from both sensors. These variables act as predictors to the AGB at the sample plots.

Table 4. Variables that were derived from PALSAR-2 and Sentinel-1A.

Variable	PALSAR-2	Sentinel-1A
Single polarization	HH HV	VV VH
Polarization multiplicative	$HH \times HV$	$VV \times VH$
Simple polarization ratio	$HH - HV$ $HV - HH$	$VV - VH$ $VH - VV$
Polarization averaging	$(HH + HV)/2$	$(VV + VH)/2$
Square root of multiplicative	$(HH \times HV)^{1/2}$	$(VV \times VH)^{1/2}$
Combination of polarizations	$HH - VV$ $HV - VH$ $(HH + HV)/(VV + VH)$ $(VV - VH)/(VH - HV)$ $(HH + HV + VV + VH)/4$ $(HH \times HV)/(VV \times VH)$ $(HH \times HV \times VV \times VH)^{1/4}$	

All polarizations are in sigma nought (σ^0, dB).

Simple linear regression method was used to investigate the relationship between the AGB and the image variables. Multiple linear regressions were also applied to the variables to observe whether the combination of polarizations from both PALSAR-2 and Sentinel-1A able to improve strength of the correlations. Manipulation of polarization of individual PALSAR-2 and Sentinel-1A as well as combination of both sensors were tested and multiple variable equations have been produced. The relationship between backscatter and AGB is represented in a common linear function as $y = ax + b$, x and y denote image variables and AGB, respectively and a and b are the equation coefficients. The strength of the relationship was measured by the derived coefficient of correlation (R^2). The greater R^2 indicates a stronger relationship between two variables; the value R^2 of 0 means no correlation and 1 is

a perfect correlation. In this case, the prediction equation with the highest R^2 was selected to estimate AGB within the entire study area.

Studies [9,11,33] have demonstrated that PALSAR polarization data actually has a logarithmic relationship with AGB. Therefore, instead of employing linear regression only, the study also attempted to correlate the AGB with the polarizations in a non-linear form. However, this method was applied only on the individual polarizations, i.e., HH, HV, VV, and VH. Similar to simple linear regression, the estimation models used AGB as independent variable to observe the sensitivity of the backscatter to the AGB. The relationship between backscatter and AGB is commonly represented in exponential an function as $y = a \times e^{(xb)}$, where x and y denote image variables and AGB, respectively and a and b are the equation coefficients.

2.6. Validation Approach

The study used K-fold cross validation method to evaluate the performance of the best prediction model derived from PALSAR-2, Sentinel-1A, and combination of both data. This method provided better indication on the prediction performance than the common residual method. Residual method does not provide an indication as to how well the model makes new predictions over new sample data, but this method does. In this study, 10-fold cross validation method [34] was used where all sample plots data were randomly grouped into 10 groups. One group was used as a testing set while the other nine groups were used in developing the model. The root mean square error (RMSE) was calculated using the testing set. This process was iterated 10 times where each group was used as a testing set once. Then, the average of all RMSEs was calculated to get the overall RMSE of that model.

3. Results and Discussion

3.1. Satellite Datasets

The images that were used for analysis have been calibrated, geometrically and radiometrically corrected, and topographically normalized. Examples of the images that went through all the pre-processing are depicted in Figure 5. The topographic normalization outcome is also shown in Figure 6. The study found that these processes are necessary and must be done on any SAR images before further analyses are carried out.
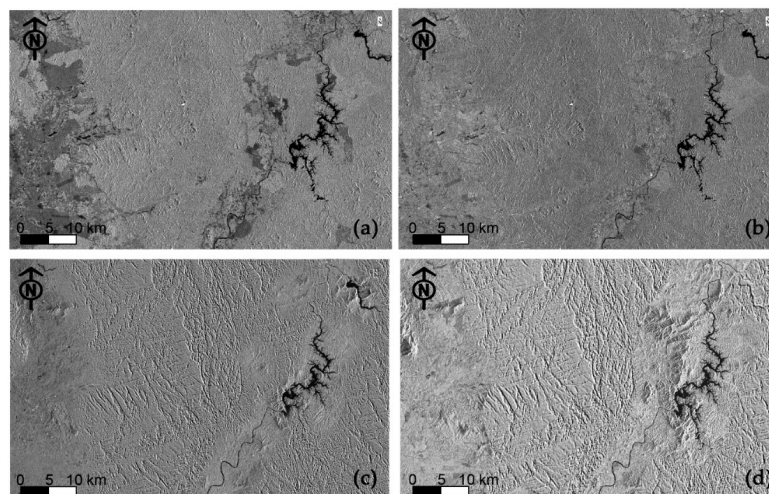

Figure 5. Corrected images for all polarizations of (**a**) PALSAR-2 HV; (**b**) PALSAR-2 HH; (**c**) Sentinel-1A VH; and (**d**) Sentinel-1A VV, displayed in backscatter values.
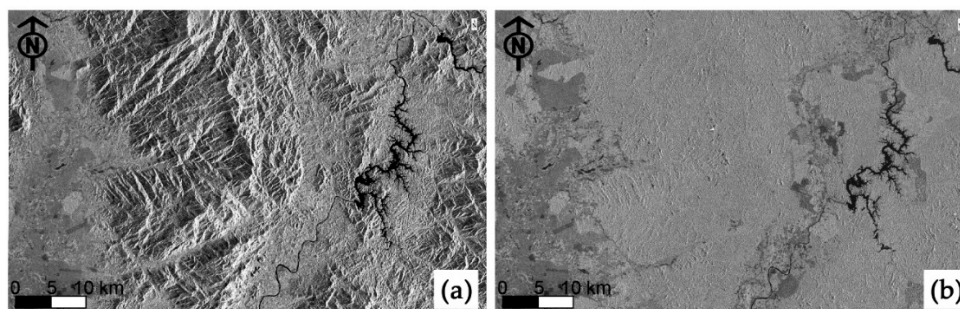
Figure 6. Topographic effect on SAR images (**a**), which has been normalized (**b**).

3.2. Forest-Non-Forest Classification

The classification that was carried out over the HH and HV polarizations and all the manipulations found that PALSAR-2 images have different capability in defining forests. The study demonstrated that the most effective polarization for FNF classification was the HV. However, the HH polarization was found effective on delineating plantation areas, such as rubber and teak, because the orientation of the plantations is systematic and homogenous, which can be interpreted well by the HH polarization. The classification was made based on the major forest types found in Peninsular Malaysia and the results are summarized in Table 5. However, this study concentrated only in lowland, hill, and upper hill dipterocarp forests. The breakdowns of these forest types are summarized in Table 6. The classification results were compared with the land use map for the year 2014 that was produced by the Department Agriculture Peninsular Malaysia and the classification accuracy was attained at 91.3% with a kappa coefficient of 0.88. The remaining 8.7% belonged to errors due to misclassification of secondary forest and rubber plantation as defined on the land use map. The results were reliable because the classification interested only in distinguishing forests from other land covers.

Table 5. Extents of forests in Peninsular Malaysia.

Forest Type	Extent (ha)
Inland	5,525,034
Peat swamp	264,578
Mangrove	106,198
Total	5,895,810

Table 6. Forest types within the study area.

Forest Type	Lowland Dipterocarp	Hill Dipterocarp	Upper Hill Dipterocarp	Total (ha)
Extents (ha)	2,704,816	2,004,991	547,588	5,257,395
Percentage (%)	51.5	38.1	10.4	100

3.3. Forest Survey Results

Aboveground biomass within all the 332 sample plots have been estimated at plot level. In general, the average AGB was 399.42 Mg ha^{-1} within the range between 35.57 and 615.50 Mg ha^{-1} and the standard deviation of 127.82 Mg ha^{-1}. AGB of small trees (dbh 10–19.9 cm) contributes only about 15% of the total AGB. However, trees under this category were plenty in terms of number. Figure 7 shows the relationship between the number of trees and AGB in a hectare of the forest. The AGB is actually stored in the huge trees measuring dbh from 40 cm and above. Although the number of huge trees is low, the amount of AGB within these trees is large.
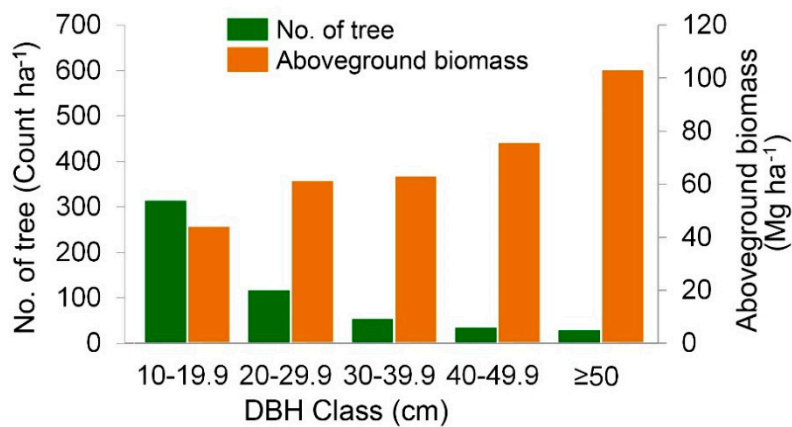
Figure 7. Relationship between tree size, number of trees, and AGB in a hectare of dipterocarp forest.

3.4. Correlation Analysis

Backscatter values from all polarizations have been extracted at all sample plots and the distribution is depicted in Figure 8. The boxplot indicates that PALSAR-2 basically had stronger backscatter over the sample plots at both polarizations as compared to Sentinel-1A polarizations. Higher variation of PALSAR-2 HV polarization indicates the capability in discriminating AGB level. On the other hand, for Sentinel-1 data, VV polarization is more sensitive to the forest as compared to VH. These backscatter values were used in the correlation analysis.
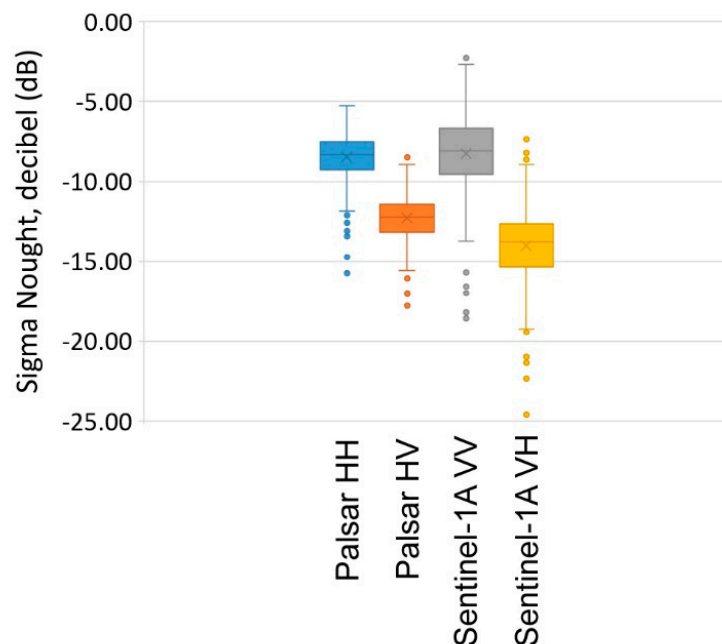
Figure 8. Backscatter values from all polarizations at sample plots of PALSAR-2 (HH and HV) and Sentinel-1 (VV and VH).

3.4.1. Simple Linear Regression

The results indicated that all variables showed weak linear relationships with AGB even though the correlations were significant at 95% confidence level. Table 7 summarized the correlation strength of all variables derived from the polarizations of PALSAR-2 and Sentinel-1A. The corresponding scatter plots listed in the table are shown in Figure 9. The results proved that the manipulation of

polarizations from a single sensor slightly improve the correlation strength. Further improvement was attained when the polarizations are combined together, from one either single sensor or integration of both sensors.

Table 7. Summary of the AGB prediction equations produced from simple linear regression.

Sensor	Variable	Scatter Plot	Prediction Equation	R^2
PALSAR-2	HH	a	28.59HH + 641.68	0.119
	HV	b	47.21HV + 978.87	0.276
	HH × HV	c	−1.98(HH × HV) + 608.64	0.223
	HH − HV	d	−6.5969(HH − HV) + 369.08	0.005
	HV − HH	e	6.5969(HV − HH) + 369.08	0.005
	(HH + HV)/2	f	43.49((HH + HV)/2) + 850.56	0.219
	$(HH \times HV)^{1/2}$	g	$-40.72((HH \times HV)^{1/2}) + 813.88$	0.201
Sentinel-1A	VV	h	15.56VV + 527.84	0.091
	VH	i	8.48VH + 518.33	0.023
	VV × VH	j	−0.69(VV × VH) + 482.13	0.090
	VV − VH	k	13.78(VV − VH) + 319.88	0.041
	VH − VV	l	−13.328(VH − VV) + 321.76	0.038
	(VV + VH)/2	m	14.13((VV + VH)/2) + 556.84	0.062
	$(VV \times VH)^{1/2}$	n	$-14.46((VV \times VH)^{1/2}) + 553.95$	0.073
Combination	HH − VV	o	−4.11(HH − VV) + 398.50	0.004
	HV − VH	p	6.48(HV − VH) + 388.08	0.015
	(HH + HV)/(VV + VH)	q	−31.53((HH + HV)/(VV + VH)) + 429.95	0.0001
	(VV − VH)/(VH − HV)	r	−590.85((VV − VH)/(VH − HV)) + 269.13	0.099
	(HH + HV + VV + VH)/4	s	36.98((HH + HV + VV + VH)/4) + 797.17	0.176
	(HH × HV)/(VV × VH)	t	−0.91((HH × HV)/(VV × VH)) + 400.39	0.003
	$(HH \times HV \times VV \times VH)^{1/4}$	u	$-35.30((HH \times HV \times VV \times VH)^{1/4}) + 764.90$	0.177

All polarizations are in sigma nought (σ^0, dB). All correlations are significant at $p < 0.05$.

3.4.2. Multiple Linear Regression

Synergy of the prediction has been obtained when the variables were integrated into an empirical prediction equation derived from multiple line regression. This method was applied to the single PALSAR-2, Sentinel-1A polarization, and also to the variables from the combination of both PALSAR-2 and Sentinel-1A. The best three models have been produced as summarized in Table 8. Evidently the combination of PALSAR-2 L-band and Sentinel-1A able to strengthen the relationship between AGB and the polarization, thus improving the accuracy of estimates. Both datasets have complemented to each other that eliminated the effects of backscattering diffusion.

Table 8. The best correlations derived from multiple regression from a single sensor and combination of sensors.

Sensor	Prediction Equation	R^2
PALSAR-2	$146.90HH + 169.78HV - 7.03(HH \times HV) + 416.96(HH \times HV)^{1/2} + 227.07$	0.342
Sentinel-1A	$-17.040VH - 2.344(VV \times VH) + 24.327(HH \times HV)^{1/2} + 181.918$	0.138
Combination	$-10.877VH - 13.292(HH \times HV)^{1/2} + 139.702HH + 162.287HV - 6.526(HH \times HV) + 394.502(HH \times HV)^{1/2} + 238.524$	0.356

All polarizations are in sigma nought (σ^0, dB). All correlations are significant at $p < 0.05$.
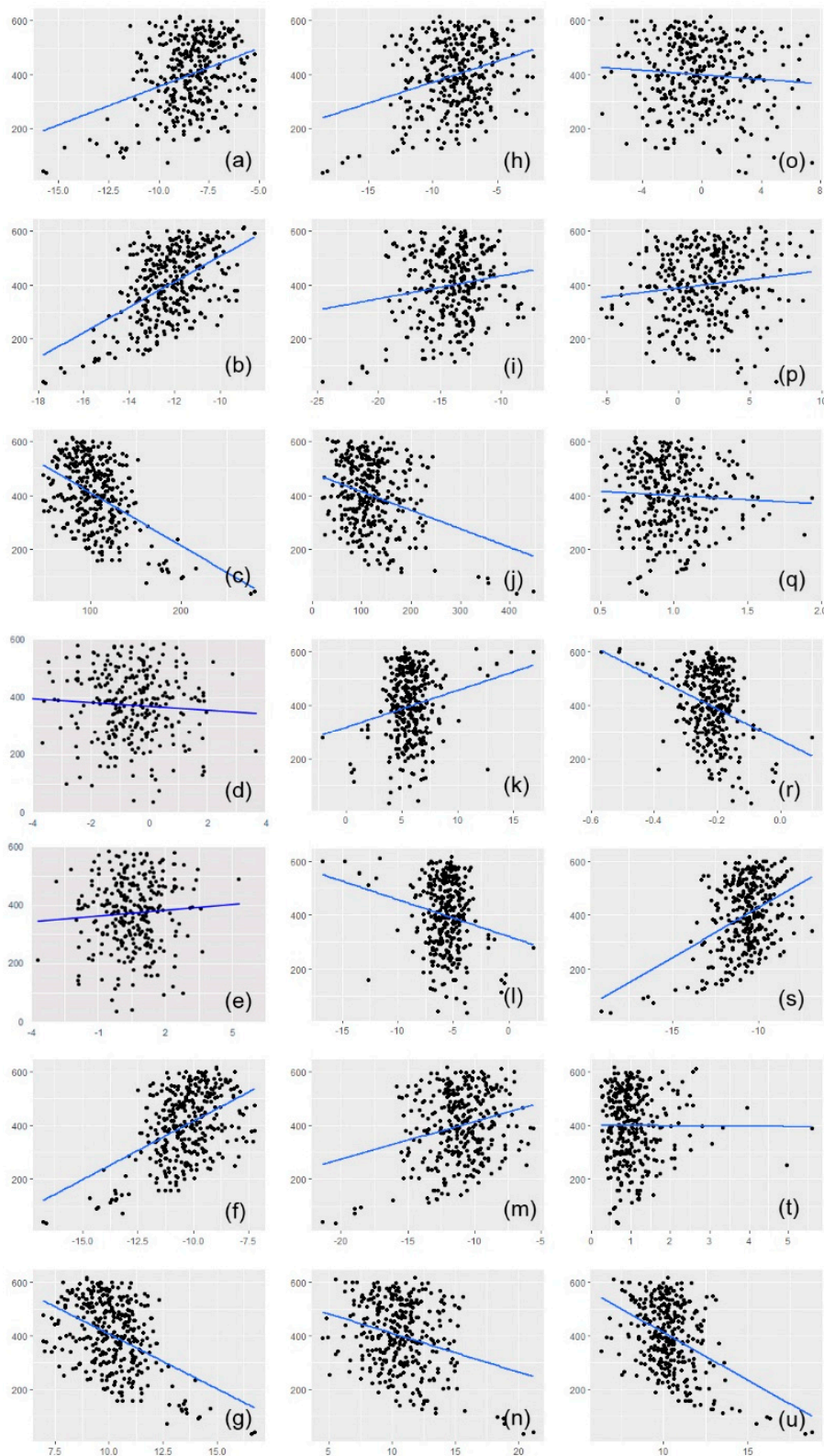
Figure 9. Scatter plots of simple linear correlations between AGB (*y*-axis) and image variables (*x*-axis).

3.4.3. Non-Linear Regression

Referring to the correlations listed in Table 9 and depicted in Figure 10, the backscatter of PALSAR-2 HV polarization gave better R^2 as compared to the HV as well as Sentinel-1A VV and VH.

The HV backscatter ranged from -1 to -20 dB and the saturation point concentrated -12 dB. Rapid increment occurred, especially at lower biomass level (i.e., up to 200 Mg ha^{-1}), and then decreased towards higher AGB. The trend line became almost constant when the AGB exceeded 200 Mg ha^{-1}. It was obvious that the estimation uncertainties are larger at AGB > 200 Mg ha^{-1}. The results were even worse for HH polarization.

Table 9. Summary of non-linear correlation between AGB and individual polarization.

Sensor	Polarization	Prediction Equation	R^2
PALSAR-2	HH	$y = 1043 \times e^{0.1215x}$	0.2058
	HV	$y = 3114.2 \times e^{0.173x}$	0.3502
Sentinel-1A	VV	$y = 644.1 \times e^{0.0664x}$	0.1596
	VH	$y = 718.72 \times e^{0.0469x}$	0.0749

All correlations are significant at $p < 0.05$.
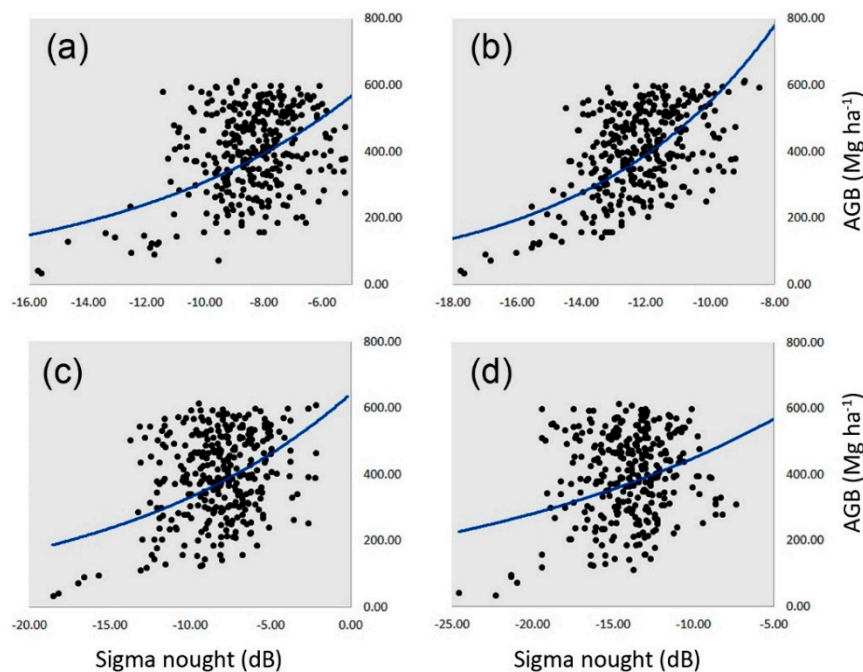
Figure 10. Scatter plots of non-linear correlations between AGB and single polarization of (**a**) PALSAR-2 HH; (**b**) PALSAR-2 HV; (**c**) Sentinel-1A VV; and (**d**) Sentinel-1A VH.

It has been reported that, at a given polarization and incidence angle, the saturated backscatter value for forest was within a small range of backscatter [11]. The dynamic range is determined primarily by the backscatter at low levels of AGB. It increases with decreasing frequency and it is higher at HV compared to HH polarization. Similarly, Sentiel-1A polarizations saturated quickly at AGB lower than 100 Mg ha^{-1}. The Sentinel-1A VV and VH backscatter ranged from -18 to -3 dB and -24 to -8 saturated at -8 dB and -14 dB, respectively. This was because the C-band wavelength is shorter than the L-band, and thus not very sensitive to the AGB at high level. Figure 11 illustrates how L- and C-bands interact with the forest canopy structure that influence the strength of the backscatter at high biomass forest.
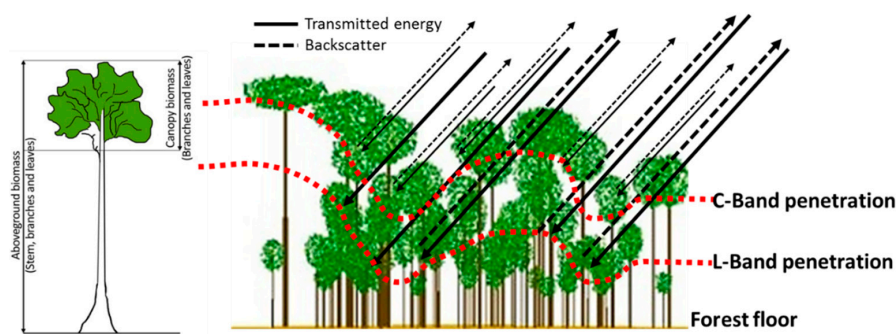
Figure 11. Common interaction of SAR L- and C-bands with the forest structure.

The accuracy was also mostly influenced by the tree density, soil surface roughness, soil moisture, tree sizes, and the layering effects of the SAR itself [35]. An experiment has found that the backscattering intensity interacted only with trees of dbh larger than 15 cm. These stands are actually dominating the higher canopies, which gave the best response to backscatter in HV polarization of L-band [13]. Other factors—such as orientation of the forest, polarimetry, incidence angle, and crown structure—also play important role in the estimated biomass [36,37].

3.4.4. The Combination Effects

Referring to Figures 8 and 11, the responses of PALSAR L-band and Sentinel-1A C-band towards AGB are different in terms of strength and variation. L-band observations penetrate more into the forest canopy and between branches and spaces, compared to the C-band, which only interacts with top canopy layers before it is scattered back or extinct. In tropical forests, volume and volume-surface scattering dominated the HH in while volumetric scattering due to dense vegetation cover dominated the HV by [38]. It is also likely that forest has a higher amount of canopy variability influencing scattering due to significant surface roughness as observed by the cross-pol (HV) term. Since structure influences the cross-pol term, forest areas that undergo selective harvesting are theoretically observable by PALSAR-2 HV, but not by Sentinel-1A, unless there is excessive timber extraction from the forest. SAR sensors can receive a relatively higher amount of surface scattering in low-density forest rather than a majority of scattering from trunks and trees or branches and crowns. Therefore, stand density, basal area, and AGB influence these relationships; although variability remains low regardless of height of the stands.

Since most of the sample plots were located in dense and mature forest (of AGB $\geq$ 200 Mg ha^{-1}), the variation of backscatter from both PALSAR-2 and Sentinel-1A polarizations are within the saturation threshold. Except for a number of sample plots that were located inside the secondary and logged forests, which contained relatively lower AGB than that inside the dense forest. Consequently, these factors have influenced the scatterplots in the correlations. The presence of C-band in the combination has complement the L-band at lower part of AGB (<200 Mg ha^{-1}) forest thus produce a better prediction overall. Taking the best linear correlations from PALSAR-2 HV and Sentinel-1A VV, with R^2 0.276 and 0.091 respectively, the R^2 increased to 0.356 when combined. This has increased the explained variance by about an average of 17.25%. Although a single PALSAR-2 HV polarization from the non-linear correlation can predict the AGB with an R^2 of 0.3502, the multiple linear correlations remain stronger even with an increase of explained variance by about 0.58%. In addition, the combination of L-band polarizations, as well as L- and C-band fusion has proven to be successful, thus confirming the hypothesis of the study.

3.5. Estimated AGB and Mapping

The study demonstrated that the combination of polarizations from PALSAR-2 L-band and Sentinel-1A C-band provided advantages for AGB estimation in dipterocarp forest with relatively

high AGB. The equation that was derived from the multi linear regression resulted from the best combination of PALSAR-2 and Sentinel-1A polarization, which gives the highest correlation value of 0.356, was used to estimate the AGB within the entire study area. The equation is expressed as

$$\text{AGB (Mg ha}^{-1}) = -10.877\text{VH} - 13.292(\text{HH} \times \text{HV})^{1/2} + 139.702\text{HH} + 162.287\text{HV} - 6.526(\text{HH} \times \text{HV}) + 394.502(\text{HH} \times \text{HV})^{1/2} + 238.524 \tag{4}$$

By using this equation, AGB within the entire study area has been retrieved and mapped. Figure 12 shows the spatial distribution of AGB in the study area. From the map, the total AGB in about 5.25 million ha of the study area was estimated at 1,821,214,202 Mg over the year 2016. Figure 13 summarizes the distribution of AGB in the study area, represented by histogram of frequency of pixel occurrences. The distribution was found to be normal throughout the entire study area. Further classification was made to the AGB distribution, reported in intervals as shown in Figure 14. More than half of the study area comprised AGB within the range of 300–400 Mg ha^{-1}. The highest AGB was concentrated mainly in the northern part of Pahang and southern part of Kelantan, where the largest National Park in Peninsular Malaysia is located. High density of AGB occurred also in the northern part of Perak where Royal Belum State Park is located. These forests are virgin and have existed for millions years and are still intact now. Variations are found scattered in other areas where there was a mix between natural virgin and logged over forests. Low density of AGB appeared near the edges of forest areas, which mostly interacted with other land use activities nearby.
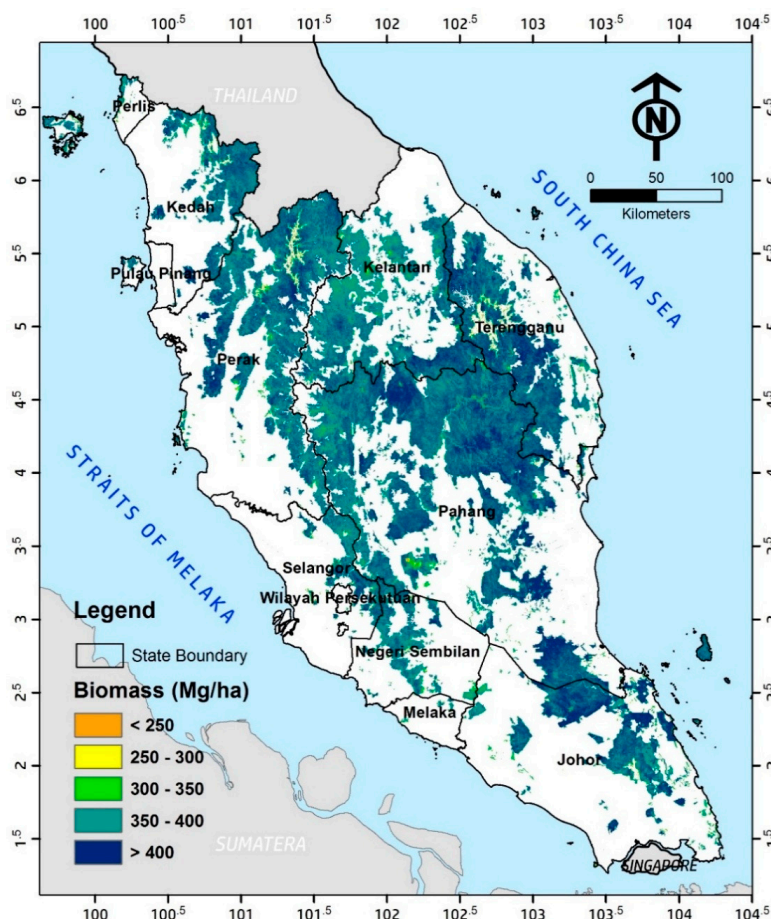

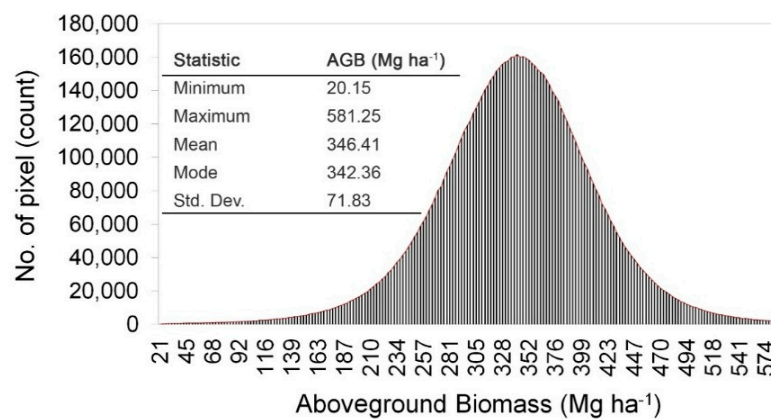**Figure 12.** Spatially distributed map of AGB within the study area.

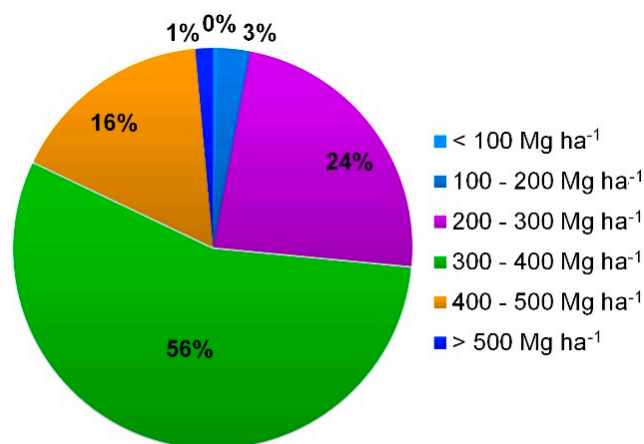**Figure 13.** Distribution of AGB within the study area.

Figure 14. Overall breakdowns of AGB density in the study area.

3.6. Validation of the Estimates

Overall, the RMSE, resulted from 10-fold cross validation method, for the best model for PALSAR-2, Sentinel-1A, and combination of PALSAR-2 and Sentinel-1A are 99.10 Mg ha^{-1}, 111.18 Mg ha^{-1}, and 98.41 Mg ha^{-1}, respectively. The best prediction model, which was produced from the combination of polarization from both PALSAR-2 and Sentinel-1A, gave the lowest error. However, the variation of RMSE between them was considered small, which means that even if the prediction were carried out from a single PALSAR-2 or Sentinel-1A, the error will be almost at the same level. Figure 15 shows the relationship between the reference AGB calculated from ground data and predicted AGB estimated using Equation (4). The error was observed to occur at around 400 Mg ha^{-1} and the distribution shows that the prediction model slightly overestimated the AGB within the study area.
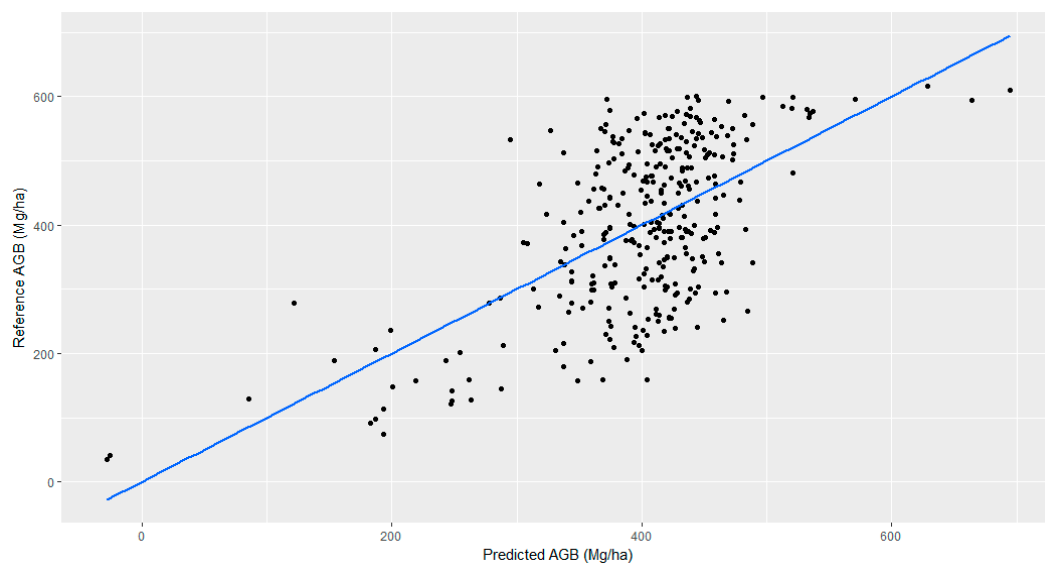
Figure 15. Relationship between predicted and reference AGB.

4. Conclusions

The study has successfully quantified the AGB over the lowland, hill, and upper hill dipterocarp forests in Peninsular Malaysia. The total AGB was estimated at about 1.82 trillion Mg over the year 2016. The extent of forested area—i.e., 5,895,810 ha—was also identified from the L-band PALSAR-2 data. The study confirmed that the synergetic of PALSAR-2 and Sentinel-1A produced better estimates than the single sensor. Although there were limitations found, the study provided an alternative for AGB retrieval that can be utilized in a practical manner to assist in the management and protection of forested areas. The study, to some extent, can also provide a significance contribution towards the MRV in the REDD+ implementation. One of greatest advantages of using the PALSAR-2 and Sentinel-1A data is the free access policy to the datasets. Free-cloud cover and rapid acquisition made them more valuable, especially for this kind of study in Malaysia.

Author Contributions: H.O. led the project, contributed the main idea of this article and conducted mapping, and prepared most of the research design and findings. M.A.M. conceived the statistical analysis and led the ground survey team. A.R.K. conducted an overall assessment over the scientific merit of this study and contributed in analysis of the ground data.

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

References

1. Intergovernmental Panel on Climate Change (IPCC). *IPCC Guidelines for National Greenhouse Gas Inventories*; Prepared by the National Greenhouse Gas Inventories Programme; Eggleston, H.S., Buendia, L., Miwa, K., Ngara, T., Tanabe, K., Eds.; Japan Institute for Global Environmental Strategies: Hayama, Japan, 2006.
2. Sessa, R.; Dolman, H. (Eds.) *Terrestrial Essential Climate Variables for Climate Change Assessment, Mitigation and Adaptation*; FAO GTOS-52; FAO: Rome, Italy, 2008.
3. Le Toan, T.; Quegan, S.; Davidson, M.W.J.; 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]

4.   Pedro, R.V.; Wheeler, J.; Louis, V.; Tansey, K.; Balzter, H. Quantifying forest biomass carbon stocks from space. *Curr. For. Rep.* **2017**, *3*, 1–18.

5.   Gibbs, H.K.; Brown, S.; O'Niles, J.; Foley, J.A. Monitoring and estimating tropical forest carbon stocks: Making REDD a reality. *Environ. Res. Lett.* **2007**, *2*, 045023. [CrossRef]

6.   Angelsen, A.; Brown, S.; Loisel, C.; Peskett, C.; Streck, C.; Zarin, D. *Reducing Emission from Deforestation and Degradation (REDD): An Options Assessment Report*; A Report Prepared for the Government of Norway; Meridian Institute: Washington, DC, USA, 2009; p. 100.

7.   Holmgren, P. *Role of Satellite Remote Sensing in REDD*; UN-REDD Programme; MRV Working Paper 1; UN FAO: Rome, Italy, 2008.

8.   Asner, G.P. Cloud cover in Landsat observations of the Brazilian Amazon. *Int. J. Remote Sens.* **2001**, *22*, 3855–3862. [CrossRef]

9.   Lucas, R.; Armston, J.; Fairfax, R.; Fesham, R.; Accad, A.; Carreiras, J.; Kelley, J.; Bunting, P.; Clewley, D.; Bray, S.; et al. An evaluation of the ALOS PALSAR L-band backscatter—Aboveground biomass relationship Queensland, Australia: Impacts of surface moisture condition and vegetation structure. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2010**, *3*, 576–593. [CrossRef]

10. Avtar, R.; Suzuki, R.; Takeuchi, W.; Sawada, H. PALSAR 50 m mosaic data based national level biomass estimation in Cambodia for implementation of REDD+ mechanism. *PLoS ONE* **2013**, *8*, e74807. [CrossRef] [PubMed]

11. Hamdan, O.; Mohd, H.I.; Khali Aziz, H.; Norizah, K.; Helmi Zulhaidi, M.S. Determining L-band saturation level for aboveground biomass assessment of dipterocarp forests in Peninsular Malaysia. *J. Trop. For. Sci.* **2015**, *27*, 388–399.

12. Imhoff, M.L. A theoretical analysis of the effect of forest structure on synthetic aperture radar backscatter and the remote sensing of biomass. *IEEE Trans. Geosci. Remote Sens.* **1995**, *33*, 341–352. [CrossRef]

13. Hamdan, O.; Mohd, H.I.; Khali Aziz, H. Combination of SPOT-5 and ALOS PALSAR images in estimating aboveground biomass of lowland Dipterocarp forest. In *IOP Conference Series Earth and Environmental Science*; Institute of Physics: London, UK, 2014; p. 6.

14. Chenli, W.; Zheng, N.; Xiaoping, G.; Zhixing, G.; Pifu, C. Tropical forest plantation biomass estimation using RADARSAT-SAR and TM data of South China. In Proceedings of the Fourth International Symposium on Multispectral Image Processing and Pattern Recognition (MIPPR), Wuhan, China, 31 October–2 November 2005; Liangpei, Z., Jianqing, Z., Mingsheng, L., Eds.; Volume 6043, pp. 61–69.

15. Patenaude, G.; Hill, R.A.; Milne, R.; Gaveau, D.L.A.; Briggs, B.B.J.; Dawson, T.P. Quantifying forest aboveground carbon content using lidar remote sensing. *Remote Sens. Environ.* **2004**, *93*, 368–380. [CrossRef]

16. Le Toan, T.; Quegan, S.; Woodward, I.; Lomas, M.; Delbart, N.; Picard, C. Relating radar remote sensing of biomass to modeling of forest carbon budgets. *Clim. Chang.* **2004**, *76*, 379–402. [CrossRef]

17. Proisy, C.; Mougin, E.; Fromard, F.; Karam, M.A. Interpretation of polarimetric radar signatures of mangrove forests. *Remote Sens. Environ.* **2000**, *71*, 56–66. [CrossRef]

18. Simard, M.; DeGrandi, G.; Saatchi, S.; Mayaux, P. Mapping tropical coastal vegetation using JERS-1 and ERS-1 radar data with a decision tree classifier. *Int. J. Remote Sens.* **2002**, *23*, 1461–1474. [CrossRef]

19. Lucas, R.; Moghaddam, M.; Cronin, N. Microwave scattering from mixed-species forests, Queensland, Australia. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 2142–2159. [CrossRef]

20. Hamdan, O.; Khali, A.H.; Abd, R.K. Remotely sensed L-Band SAR data for tropical forest biomass estimation. *J. Trop. For. Sci.* **2011**, *23*, 318–327.

21. Simard, M.; Hensley, S.; Lavalle, M.; Dubayah, R.; Pinto, N.; Hofton, M. An empirical assessment of temporal decorrelation using the uninhabited aerial vehicle synthetic aperture radar over forested landscapes. *Remote Sens.* **2012**, *4*, 975–986. [CrossRef]

22. Englhart, S.; Keuck, V.; Siegert, F. Aboveground biomass retrieval in tropical forests—The potential of combined X- and L-band SAR data use. *Remote Sens. Environ.* **2011**, *115*, 1260–1271. [CrossRef]

23. Reiche, J.; Lucas, R.; Anthea, L.; Mitchell, J.V.; Dirk, H.; Hoekman, J.H.; Kellndorfer, J.M.; Rosenqvist, A.; Eric, A.L.; Woodcock, C.E.; et al. Combining satellite data for better tropical forest monitoring. *Nat. Clim. Chang.* **2016**, *6*, 120–122. [CrossRef]

24. Morel, A.C.; Saatchi, S.; Malhi, Y.; Berry, N.J.; Banin, L.; Burslem, D.; Nilus, R.; Ong, R.C. Estimating aboveground biomass in forest and oil palm plantation in Sabah, Malaysian Borneo using ALOS PALSAR data. *For. Ecol. Manag.* **2011**, *262*, 1786–1798. [CrossRef]

25. Morel, A.C.; Fisher, J.B.; Malhi, Y. Evaluating the potential to monitor aboveground biomass in forest and oil palm in Sabah, Malaysia, for 2000–2008 with Landsat ETM+ and ALOS-PALSAR. *Int. J. Remote Sens.* **2012**, *33*, 3614–3639. [CrossRef]

26. Forestry Department Peninsular Malaysia. *Annual Report 2014*; Forestry Department Peninsular Malaysia: Kuala Lumpur, Malaysia, 2015.

27. Hoekman, D.H.; Reicheb, J. Multi-model radiometric slope correction of SAR images of complex terrain using a two-stage semi-empirical approach. *Remote Sens. Environ.* **2015**, *156*, 1–10. [CrossRef]

28. Canty, M.J. *Image Analysis, Classification and Change Detection in Remote Sensing: With Algorithms for ENVI/IDL and Python*, 3rd ed.; CRC Press: Boca Raton, FL, USA, 2014.

29. Shimada, M. Model-based Polarimetric SAR calibration method using forest and surface-scattering targets. *IEEE Trans. Geosci. Remote Sens.* **2011**, *49*, 1712–1733. [CrossRef]

30. Sarker, M.; Rahman, L.; Janet, N.; Baharin, A.; Busu, I.; Alias, A.R. Potential of texture measurements of two-date dual polarization PALSAR data for the improvement of forest biomass estimation. *ISPRS J. Photogramm. Remote Sens.* **2012**, *69*, 146–166. [CrossRef]

31. Walker, S.M.; Pearson, T.R.H.; Casarim, F.M.; Harris, N.; Petrova, S.; Grais, A.; Swails, E.; Netzer, M.; Goslee, K.M.; Brown, S. *Standard Operating Procedures for Terrestrial Carbon Measurement: Version 2014*; Winrock International: Little Rock, AR, USA, 2012.

32. Chave, J.; Maxime, R.E.; Alberto, B.; Chidumayo, E.; Matthew, S.C.; Welington, B.C.D.; Alvaro, D.; Tron, E.; Philip, M.F.; Rosa, C.G.; et al. Improved allometric models to estimate the aboveground biomass of tropical trees. *Glob. Chang. Biol.* **2014**, *20*, 3177–3190. [CrossRef] [PubMed]

33. Liesenberg, V.; Gloaguen, R. Evaluating SAR polarization modes at L-band for forest classification purposes in Eastern Amazon, Brazil. *Int. J. Appl. Earth Obs. Geoinf.* **2013**, *21*, 122–135. [CrossRef]

34. McLachlan, G.J.; Do, K.A.; Christophe, A. *Analyzing Microarray Gene Expression Data*; Wiley: Hoboken, NY, USA, 2004.

35. Quinones, M.J.; Hoekman, D.H. Exploration of factors limiting biomass estimation by polarimetric radar in tropical forests. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 86–104. [CrossRef]

36. Watanabe, M.; Shimada, M.; Rosenqvist, A.; Tadono, T.; Matsuoka, M.; Romshoo, S.A.; Ohta, K.; Furuta, R.; Nakamura, K.; Moriyama, T. Forest structure dependency of the relation between L-band σ^0 and biophysical parameters. *IEEE Trans. Geosci. Remote Sens.* **2006**, *44*, 3154–3165. [CrossRef]

37. Guo, Z.; Ni, W.; Sun, G. Analysis of the effect of crown structure changes on backscattering coefficient using modelling and SAR data. *IEEE Trans. Geosci. Remote Sens.* **2009**, *4*, 386–389.

38. Nathan, T.; Lindsay, L.; William, S.; Meng, Z. Regional mapping of plantation extent using multisensor imagery. *Remote Sens.* **2016**, *8*, 236.