

Article

Remote Sensing of Leaf Area Index from LiDAR Height Percentile Metrics and Comparison with MODIS Product in a Selectively Logged Tropical Forest Area in Eastern Amazonia

Yonghua Qu ^{1,2,3,*}, Ahmed Shaker ^{3,*}, Carlos Alberto Silva ^{4,5}, Carine Klauberg ⁵ and Ekena Rangel Pinagé ⁶

- State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and Institute of Remote Sensing and Digital Earth of Chinese Academy of Sciences, Beijing 100875, China
- ² Beijing Engineering Research Center for Global Land Remote Sensing Products, Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
- ³ Department of Civil Engineering, Ryerson University, Toronto, ON M5B2K3, Canada
- ⁴ Biosciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20707, USA; carlos_engflorestal@outlook.com
- ⁵ US Forest Service (USDA), Rocky Mountain Research Station, RMRS, 1221 South Main Street, Moscow, ID 83843, USA; carine_klauberg@hotmail.com
- ⁶ Climate Change Cluster, University of Technology Sydney, Ultimo, NSW 2007, Australia; Ekena.RangelPinage@student.uts.edu.au
- * Correspondence: qyh@bnu.eud.cn (Y.Q.); ahmed.shaker@ryerson.ca (A.S.); Tel.: +8610-58802041 (Y.Q.); +1-416-9795000 (ext. 6458) (A.S.)

Received: 30 April 2018; Accepted: 13 June 2018; Published: 17 June 2018



Abstract: Leaf area index (LAI) is an important parameter to describe the capacity of forests to intercept light and thus affects the microclimate and photosynthetic capacity of canopies. In general, tropical forests have a higher leaf area index and it is a challenge to estimate LAI in a forest with a very dense canopy. In this study, it is assumed that the traditional Light Detection and Ranging (LiDAR)-derived fractional vegetation cover (fCover) has weak relationship with leaf area index in a dense forest. We propose a partial least squares (PLS) regression model using the height percentile metrics derived from airborne LiDAR data to estimate the LAI of a dense forest. Ground inventory and airborne LiDAR data collected in a selectively logged tropical forest area in Eastern Amazonia are used to map LAI from the plot level to the landscape scale. The results indicate that the fCover, derived from the first return or the last return, has no significant correlations with the ground-based LAI. The PLS model evaluated by the leave-one-out validation shows that the estimated LAI is significantly correlated with the ground-based LAI with an R² of 0.58 and a root mean square error (RMSE) of 1.13. A data comparison indicates that the Moderate Resolution Imaging Spectrometer (MODIS) LAI underestimates the landscape-level LAI by about 22%. The MODIS quality control data show that in the selected tile, the cloud state is not the primary factor affecting the MODIS LAI performance; rather, the LAI from the main radiative transfer (RT) algorithm contributes much to the underestimation of the LAI in the tropical forest. In addition, the results show that the LiDAR-based LAI has a better response to the logging activities than the MODIS-based LAI, and that the leaf area reduction caused by logging is about 13%. In contrast, the MODIS-based LAI exhibits no apparent spatial correlation with the LiDAR-based LAI. It is suggested that the main algorithm of MODIS should be improved with regard to tropical forests. The significance of this study is the proposal of a framework to produce ground-based LAI using forest inventory data and determine the plot-level LAI at the airborne and satellite scale using LiDAR data.



Keywords: leaf area index; LiDAR; height percentile metrics; partial least square; tropical forest; MODIS

1. Introduction

Tropical forests comprise nearly 50% of the global forest cover [1]. Thus, understanding the growth stage of forests is an important basis for understanding their ecological responses to climate change [2]. The foliage is the exchange surface between the photosynthetic active component of the plant and the atmosphere, and controls the light, thermal, and moisture conditions in the canopy. A simple measure of the amount of foliage is the leaf area index (LAI), which is half of the foliage area per unit of ground area [3]; LAI can be unitless or in the units of m^2/m^2 ; in this study, we use the unitless format.

Selective logging is an important harvesting method in forested areas throughout the humid tropics [4]. Logging results in many changes in the forest structure and contributes substantially to gross carbon fluxes in the Brazilian Amazon and in other tropical regions [5]. Selective logging can be monitored by determining the changes in the forest gap fraction [4] or LAI [6]. Therefore, measurements of LAI in selective logging areas are important for monitoring the impact of forest management and for understanding the carbon cycle dynamics in tropical forests.

LAI can be directly or indirectly measured. Since direct methods are labor intensive, indirect methods, which usually involve the use of optical sensors, are often preferred due to their high efficiency [7]. Measured LAI from indirect are influenced by the leaf clumping effects, or the clumping index [8]. When the clumping index is not accounted for, the output of indirect methods, such as LAI-2000 canopy analyzer (Li-Cor Biosciences, Lincoln, NE, USA) and hemi-spherical imaging method, is called effective leaf area index. Thus, in this paper, when we refer to indirect measurements of LAI, we mean effective LAI.

The most commonly used methods for estimating LAI across landscapes rely on the relationships between the LAI and various manipulations of spectral information from aircraft- or satellite-based imagery [9–11]. Throughout the past decade, Light Detection and Ranging (LiDAR) data have provided an alternative approach for estimating LAI across landscapes [12–15].

In previous studies, numerous LiDAR-derived metrics have been proposed to link the LiDAR data with ground-measured LAI; for example, fractional vegetation cover indices [16–19], height and density metrics [20], and the combination of these metrics have been used [21]. An analysis of the success of the use of LiDAR for the estimation of forest LAI shows that there are few reports on estimating LAI in very dense forests. Commonly LiDAR metrics have been used as suitable surrogates to approximate LAI of canopies with a relatively low LAI (e.g., LAI = 3–5). However, in forests with high leaf density, e.g., an LAI higher than 5.0, the shadowing effect becomes more significant [22]. Clark et al. [23] reported that the maximum LAI of tropical forest is 12.0 with a mean value of 6.0 on a regional scale. For dense forest cover, which is common in the tropical forests of Eastern Amazonia, improved methods are required to estimate forest LAI from LiDAR data.

The canopy leaf amount based on area (m²) or weight (leaf mass per area (LMA) has a high correlation with the canopy height [24,25]. For example, Cavaleri et al. [26] measured the height and LMA of all species encountered along 45 vertical canopy transects across a Costa Rican tropical rain forest. Their results showed a remarkably linear relationship between canopy height and LMA throughout the entire vertical canopy profile. The relationship between canopy height and forest LAI is the basis for estimating LAI using height-related metrics from LiDAR data, e.g., mean height, top-of-canopy height, and height percentiles.

Unlike a single height variable, LiDAR-derived height percentiles are a series of statistical descriptors of the canopy's vertical profile. As a result, the height percentiles offer the potential to reveal multi-layer information on the canopy rather than information for only the top; these metrics

have been used to estimate a variety of forest structure characteristics [21,27]. However, due to the collinearity between the height percentiles, a careful design of the regression model is required before using these metrics as explanatory variables, to eliminate or minimize the collinearity of the predictors [28,29].

Global LAI data from optical sensors are available from the Moderate-Resolution Imaging Spectroradiometer (MODIS) satellites. Examples include the MODIS global LAI products at four- or eight-day temporal resolution [30–32], which have been used to evaluate the effects of climate change and the responses of tropical forests [33,34]. However, as a passive sensor, MODIS is affected by cloudy conditions and dense forest cover, which is particularly problematic when studying tropical forests [35]. A comparison between the LAI from LiDAR and MODIS sensors will highlight the potentials and limitations of the two types of LAI products for determining the effects of logging activities on forest canopies.

The first objective of this study is to estimate the LAI of a dense tropical forest in Eastern Amazonia using height percentile metrics derived from airborne LiDAR data. It is assumed that there is no significant relationship between the traditional LiDAR fractional and LAI in a dense forest. A partial least squares (PLS) regression model [36], which is a linear learning model that uses variables to capture the maximal covariance between the input and output data, is developed using ground-based plot LAI data and LiDAR metrics. Our second objective is to compare LiDAR-derived and the MODIS-derived LAI estimates at the landscape scale and determine their relationship and potential limitations. The quality control (QC) data of the MODIS LAI are used to determine the factors affecting the performance of the MODIS products. In addition, the damage caused by selective logging is analyzed using both types of LAI data.

In this study, we first generate the plot-based LAI from forest inventory datasets that are used as training data to link with the LiDAR metrics. Next, a PLS model is established and evaluated using the plot-based LAI and LiDAR-based height percentiles. Finally, we generate the landscape-level LAI and determine the relationship between the satellite-based and LiDAR-based estimations of the LAI at the pixel scale and at the unit scale used for the selective logging.

2. Materials and Methods

2.1. Study Area

The study area is located at Fazenda Cauaxi in the Paragominas Municipality of Pará State, Brazil (Figure 1a–c). The average annual temperature is 25 °C and the average annual precipitation is 2200 mm [37]; the area is classified as a moist tropical forest [38]. The terrain is characterized with flat to steep slopes, and the soils in the region are predominately classified as dystrophic yellow latosols according to the Brazilian soil classification system [39].

This site covers an area of $3 \times 4 \text{ km}^2$, divided into 12 logging plots of approximately 100 hectares. Ten of them have been logged through reduced-impact logging (RIL) from 2006 to 2013 and the remaining two are still unlogged. In this study area, there are 167 trees species; the three most dominant species are Jatereu (*Lecythis idatimon* Aubl.), Abiurana (*Pouteria gongrijpii* Eyma), and Matamata branco (*Eschweilera coriacea* (DC.) S.A. Mori).

2.2. Field Data

The forest inventory was collected in 2014 in 88 plots of 50×50 m each and distributed in the 12 logging plots (Figure 1(d,d1,d2)). At each plot, a sub-plot (strip) along one side of the plot with dimensions of 5×50 m (250 m²) was also demarcated [40]. According to the diameter at breast height (DBH), two sampling methods were used for collecting the survey data: at the plot level, all trees with a DBH equal to or greater than 35 cm were measured, whereas trees with a DBH equal to or greater than 35 cm were measured (5 × 50 m). The ground data surveying was conducted from 18 February to 25 April 2014 [41].

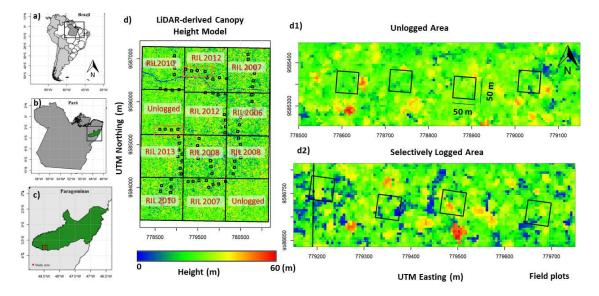


Figure 1. The location of the experiment in Brazil (a-c) and the plot distribution along the transect overlaid on the Light Detection and Ranging (LiDAR)-derived canopy height model (d). The enlarged images of the plots in the unlogged and selectively logged areas are shown in subplots (d1,d2) respectively.

The ground-based LAI was generated from the forest inventory data using the above-ground biomass (AGB) allocation method. This method is based on the assumption that the AGB of the forest is comprised of a woody component and a leaf component and that there is an empirical ratio between these parts. We summarize the process in the following flowchart (Figure 2).

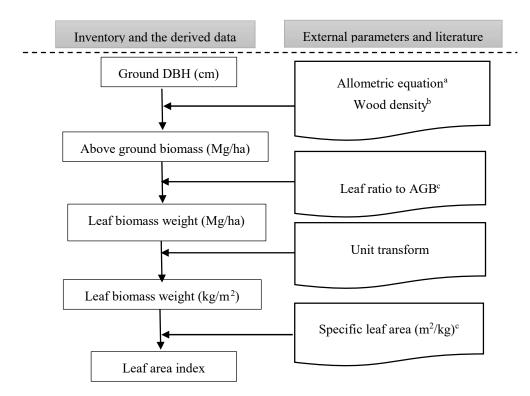


Figure 2. Workflow for the generation of the plot-based leaf area index (LAI) from ground measured diameter at breast height (DBH). The superscripts a, b, c denote the references [42–44] respectively.

For the calculation of the AGB in this area, we used the allometric equation (Equation (1)) derived from [42] which was used by Silva et al. [40] in the same research area.

AGB (kg) = exp
$$\left(-1.803 - 0.976E + 0.976\ln(\rho) + 2.673\ln(DBH) - 0.0299\ln^2(DBH)\right)$$
, (1)

where AGB (kg) is the live aboveground biomass of the trees in kg; DBH (cm) is the diameter at breast height (1.30 m), ρ is the wood density, which is derived from a published database [43], and E is a measure of the environmental stress and it is set as a constant of 0.103815 [40]. The total live AGB of the plots and sub-plots is obtained by aggregating the individual tree biomass values and converting to Mg/ha.

The next step is to determine the ratio of the leaf biomass to the AGB for the tropical forest. McWilliam et al. [44] collected data in central Amazon tropical forest (Manaus, Brazil) and analyzed data from published historical literature [45]. Their results showed that the mean ratio of the leaf biomass to the AGB was 2.5% for a range of the AGB from 190 Mg/ha to 405 Mg/ha. Therefore, we used the mean value of 2.5% as the ratio of the leaf AGB for most trees. However, this value is not a constant during all growth stages of trees. According to the research result of Reich et al. [46], trees with a lower AGB have a larger fraction of leaves than those with a higher AGB. In this study, the lower AGB threshold is defined as the mean AGB minus the standard deviation. The calculated mean AGB is 232 Mg/ha and the standard deviation is 84 Mg/ha (cf. Section 3.1); therefore, the threshold is approximated to 150 Mg/ha. When the AGB is lower than the threshold, the leaf biomass ratio (LBR) of 3.7%, which is the maximum value in [44], is used as the ratio to calculate the leaf biomass. Finally, the following equation is used to calculate the leaf biomass from the total AGB:

$$\begin{aligned}
\text{Lea} f_{biomass} &= \alpha \cdot \text{forest}_{AGB} \\
\alpha &= \begin{cases} 0.025, \text{forest}_{AGB} > 150 \\ 0.037, \text{forest}_{AGB} \le 150 \end{cases}
\end{aligned} \tag{2}$$

The specific leaf area (SLA) is defined as the ratio of the leaf area to the dry mass of the leaf (m^2/kg) [47] and is used to calculate the leaf area from the leaf weight. For tropical forest leaves, the SLA ranges from 8.0 to 9.0 m²/kg [44,48]. Here, we used the value of 9.0 m²/kg derived from measured results by [44] because there are similar tree species in our research area.

Using the workflow presented in Figure 2 and Equations (1) and (2), the ground-based LAI is calculated from the inventory DBH variable by using the proxy variable AGB.

2.3. Remote Sensing Data

2.3.1. LiDAR Data Collection and Processing

The LiDAR data were collected on 26 December and 27 December 2014 using the ALTM300 sensor. The flight altitude was 850 m, and the pulse density was 37.5 pulses/m². The LiDAR acquisition covered most of the field inventory area [49]. Herein, even though the plot corners were geolocated with differential GNSS to within <1 m in most cases [50], we chose to optimize plot location to reflect canopy conditions according to Silva et al. [40]. After geometrically registering the two datasets, it was found that three of the 88 plots were located outside the LiDAR coverage area; thus, only the inventory data of 85 plots were used to link the ground- and LiDAR-based observations.

The original LiDAR data had a very high sampling frequency of 37.5 pulses/m² which leads the mean point density of 57 point/m². To improve the data processing efficiency, we thinned the original LiDAR data to lower densities using LASThin utility of LASTool software [51]. The reduction algorithm places a uniform grid over the point cloud and within each grid keeps only random points up to the desired density. We set LASThin utility to select 2 points per meter in *x* and *y* direction (grid size of 0.5 m) and it produces LiDAR point density of 4 point/m². Though the estimated forest LAI is affected by point density, study result from Luo et al. [52] showed that when point density varied from 4 to 8 point/m², RMSE of estimated LAI varied from 0.45 to 0.47 and R² varied from 0.74 to 0.78. Singh et al. [53] reached the similar conclusion while they studied the influence of LiDAR point density on the forest biomass. As a result, we assume that the bias of estimated LAI caused by different point density is ignorable.

The digital terrain model (DTM) generated from the same LiDAR data was resampled to a resolution of 1×1 m [41]. The next step in the LiDAR data preprocessing was the classification of the cloud points as ground and forest (non-ground) using the LASGround utility in the LASTool software [51]. The input parameter for the offset, which is the threshold to separate ground from vegetation in LASGround, was set to 1.5 m. Finally, the LiDAR data were processed using FUSION software [54] to extract a series of LiDAR metrics, as described in Table 1. The grid size of the LiDAR data was set to 50×50 m, which corresponds to the size of the ground plot. The metrics of 85 grids at the locations of the ground plots were extracted as training data to develop the relationship between the ground-based LAI and LiDAR data.

Table 1. Metrics deri	ved from the LiDA	R data.
-----------------------	-------------------	---------

Metrics	Definition
H_Px	Height percentiles, where <i>x</i> is 10th to 90th at the interval of 10, and additional percentiles of 1st, 5th, 25th, 75th, 95th, and 99th are included
fCover_last	Canopy fractional cover using last return data
fCover_first	Canopy fractional cover using first return data

The canopy fractional cover (fCover) from LiDAR data is calculated as follows:

$$fCover = \frac{N_c}{N_c + N_g} \tag{3}$$

where N_c and N_g are the point counts returned from the canopy and the ground, respectively.

The LiDAR return points are labeled by the return number according to their return order of a pulse. For fCover_last, the last return points of the canopy and ground are used in Equation (3), and for fCover_first, the first return points are used. The height percentiles are calculated using only the first returns.

The metrics are used to link the LiDAR observations with the ground-based LAI. We expected that the variables derived from the metrics with high correlations with the ground-based LAI could be used as predictor variables to estimate LAI.

2.3.2. MODIS Data Collection and Processing

The MODIS (version 6) LAI product (MCD15A3H) is a four-day composite dataset with 500-m pixel size [55]. It is estimated using an algorithm that compares the MODIS surface reflectance with the reflectance in a Look-Up-Table (LUT), which is simulated from a three-dimensional radiative transfer (RT) model. The LUT is parameterized by varying biophysical parameters such as LAI and land cover and the estimated LAI is based on finding the "best" matches between the MODIS data and the RT model. When this method fails to find the "best" LAI, a backup algorithm based on biome-specific empirical relations between the normalized difference vegetation index (NDVI) and LAI is utilized. The information on the algorithm and the cloud state of the reflectance data is recorded in the quality control (QC) layer, which is distributed together with the LAI product [30,55].

According to the MODIS reflectance data quality, the LAI retrieved via the RT inversion algorithm is separated into two retrieval classes. We refer to such pixel retrievals as Main_Best and Main_Sat. The former is the solution of low variance, which has the best possible result and the latter is the solution when the LAI is considered saturated. The LAI values derived from the backup algorithm also are classified into two types depending on the reasons why the main algorithm failed due to geometry problems (we refer to this as Backup_Geo) or problems other than geometry (we refer to this as Backup_Oth).

Based on the date of the LiDAR data acquisition, we selected the MODIS LAI dataset on the day-of-year (DOY) 361 in this area (tile H13V09). We converted the MODIS LAI product from the original sinusoidal projection to the UTM/WGS84 projection using the MODIS reprojection tool (MRT). To overlay the MODIS data with the LiDAR and ground inventory data, the projection of all the spatial data are unified into UTM zone 22S, datum SIRGAS 2000 using the QGIS software [56].

2.4. Modeling LAI from LiDAR Data and PLS Regression

PLS regression is a bilinear calibration method that uses data compression by reducing the number of measured collinear variables (represented here by the LiDAR height percentiles) to a few non-correlated principal components (PCs) [57]. The process of projecting the original height percentile variables to PCs was conducted to maximize the covariance of the response variable (LAI) and independent variables (height percentile metrics). This process is similar to another multi-variable analysis method, principal component analysis (PCA), which creates PCs to explain the observed variability of the predictor variables but without considering the response variable. PLS takes the response variable into account and, therefore, models are produced to fit the response variable with fewer components. The detailed mathematical basis of the PLS algorithm is described in Rosipal and Krämer [36].

In general, PLS regression uses component projection successively to find latent structures. For our analysis, the number of latent components is based on the estimated prediction uncertainty of the regression model due to two sources of information. The first one is the proportion of variance explained by the components, and the second one is the mean squared error (MSE) from a k-fold cross validation (CV) [58].

Visual inspections of the variance-explained plots and the validation MSE plots are used to find the optimal number of PCs. In most cases, this procedure reduces the number of LiDAR-derived height percentile variables to a few independent PCs. This provides a means to determine the absolute and relative importance of each height percentile for predicting the LAI. The MATLAB function PLSREGRESS was used to obtain a PLS estimate [59]. The final model had the following form:

$$y = b_0 + \sum_{c=1}^n b_c \times t_c \tag{4}$$

where t_c (c = 1, ..., n) are the scores of the original x variable (height percentile). The scores are calculated on the basis of mean-centered data. For a linear regression of t versus y, the PLS-regression coefficients [60] $b_0, b_1, ..., b_n$ are obtained and Equation (4) is used to predict the LAI once the LiDAR height percentile metrics are obtained.

The PLS model is validated using the leave-one-out CV (LOOCV), which is a special case of the k-fold CV, in which the number of folds equals the number of samples. This means that the function is trained 85 separate times using all the data except for the sample that is left out and a prediction is made for that point. The root means square error (RMSE) and coefficient of determination (R²) are computed and used to evaluate the performance of the PLS model.

2.5. Comparing LiDAR and MODIS-Derived LAI Estimates

After the PLS model is validated as described in Section 2.4, the model is applied at the landscape level to compute the LAI for the entire study area. When the LiDAR LAI is overlaid on the MODIS LAI, there are about 60–80 LiDAR cells in one MODIS pixel. The number varied depending on the LiDAR pixel locations relative to the grid boundaries of the MODIS pixels (Figure 3). To guarantee the representation of the LiDAR LAI data in the MODIS pixel, we select the MODIS pixel containing at least 70 LiDAR LAI grid cells. We calculate the mean and standard deviation (STD) of the LiDAR LAI

in each MODIS pixel using the Zonal Statistics Plugin in QGIS software [56]. The mean LiDAR LAI in one MODIS pixel is the indicator of the LAI at the landscape scale and the STD is the description of the LAI's spatial variability. The comparison of the LiDAR and MODIS LAI is conducted by inspecting the bias of the MODIS LAI with the LiDAR LAI plus/minus its STD. The QC layer information of the MODIS LAI is used to further understand the factors related to the bias.

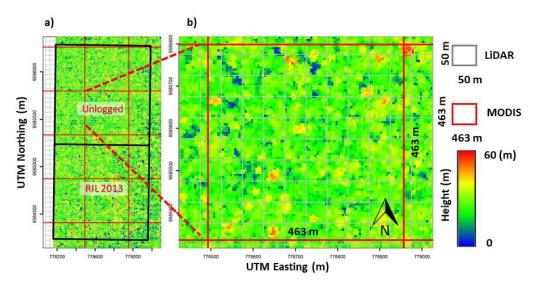


Figure 3. Illustration of the LiDAR leaf area index (LAI) (50×50 m) grids overlaid on the Moderate Resolution Imaging Spectrometer (MODIS) LAI map (463×463 m).

3. Results and Discussion

3.1. Leaf Area Index at the Plot Level Calculated from Inventory Data

The plot AGB ranges from 61 to 527 Mg/ha and the mean value is 232 (\pm 84) Mg/ha. Note that for deriving Equation (2), which is used to calculate the ground-based LAI from the AGB, we use the AGB data from published work which has a maximum AGB of 405 Mg/ha. Although the calculated maximum AGB values in this research area are larger than 405 Mg/ha, it is evident in Figure 4a that only a few AGB values occur in that range (about 2%). This result corroborates the rationale behind using the leaf ratio value in Equation (2).

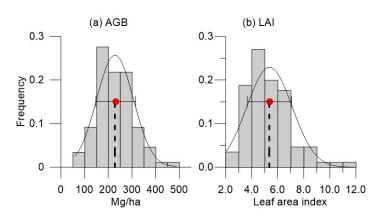


Figure 4. Frequency distribution of ground plot variables, (**a**) aboveground biomass (AGB, Mg/ha) and (**b**) leaf area index (LAI). The frequency distribution is fitted using a Gaussian model (black curve). The mean values and standard deviations are plotted as the vertical dashed lines and horizontal lines, respectively.

The ground-based LAI calculated using the workflow shown in Figure 2 ranges from 2.05 to 11.86 with a mean value of 5.36 (\pm 1.7). The frequency distribution is illustrated in Figure 4b. Since the LAI values will be used as training data in the PLS model (cf. Equation (4)), it is critical to check their validity before they are used to model the LiDAR metrics. We used published LAI data that were measured near our research area or came from areas with a similar forest type to evaluate the plot-level LAI calculated from the forest inventory data.

Metcalfe et al. [61] measured the LAI of a tropical forest at an experimental site located in the Caxiuanã National Forest, Pará State, Brazil (1°43'3.5"S, 51°27'36"W), which is the nearest location where a historic LAI dataset was recorded. They reported LAI values derived from hemispherical images ranged from 4.90 to 5.80 with the mean of 5.30 (reproduced from Figure 1 in [61]). The measured values are very close to the LAI values in this study (mean LAI 5.36). Another dataset for a tropical forest (northern Costa Rica, Central America) was destructively measured by Clark et al. [23], who reported LAI values in a similar range (1.20–12.94) to our results (2.05–11.86) but a slightly higher average LAI value of 6.0. The consistency or approximation of our calculated LAI with the ground-measured results indicates that our results are rational for representing the LAI in the research area and can be used as training data for modeling the LiDAR LAI.

3.2. Exploratory Correlation of LiDAR Metrics with Plot-Based LAI

We determined whether the LiDAR-derived canopy cover metrics could be used to estimate the LAI for a closed canopy. The statistics of the fCover are listed in Table 2.

	fCover_First (%)	fCover_Last (%)
Minimum	80.4	69.4
Maximum	99.9	97.5
Mean	96.5	89.6
Standard deviation	4.2	6.5

Table 2. Statistics of the LiDAR-derived fractional cover.

The results show that in the research area, the mean values of the two fractional covers are 96.5% and 89.6% respectively and this result is comparable with other published fractional cover data. Stark et al. [62] observed similar high fractional cover in an Amazon tropical forest in Brazil. They reported a fractional cover of 96% for a mean gap fraction of 4% in a one-hectare plot.

The scatter plot between fCover and LAI (Figure 5) provides a visual illustration of the relationship between the plot-based LAI and the LiDAR canopy cover. As expected, the canopy cover values calculated from the first and last return points are at the top of the scatter plot, i.e., most of the values are large (the minimum values were 80% and 68% respectively) or saturated (about 51% of the plots had a cover value larger than 98%). This result indicates that the fractional cover values from the first or last return points are not good proxies to estimate LAI in this research area.

This conclusion differs from previous studies [12,17,63,64]. However, a detailed analysis of the published studies indicates that the success of the LiDAR-derived fractional cover or related variables, e.g., penetration ratio, was attributed to the case of lower LAI values. In most cases, there were much larger gaps within or between the forest canopies in the previous studies. It was reported that the maximum ground-measured LAI was relatively low (<3.5). Our result showed that as the LAI increased, the number of LiDAR returns that penetrated the forest was lower and this illustrates the difficulty in detecting the effective fractional cover in tropical forests. In this case, alternative modeling methods or complex vegetation indices for LiDAR data are needed. For example, Heiskanen et al. [16] reported that they successfully estimated the LAI in a tropical forest for maximum LAI values of 7.0 using Solberg's cover index (SCI) [18]. Nevertheless, much care should be taken in a situation when, as stated by Olivas et al. [65], fCover approaches ~80% because it is difficult to detect large changes in the LAI for such high cover.

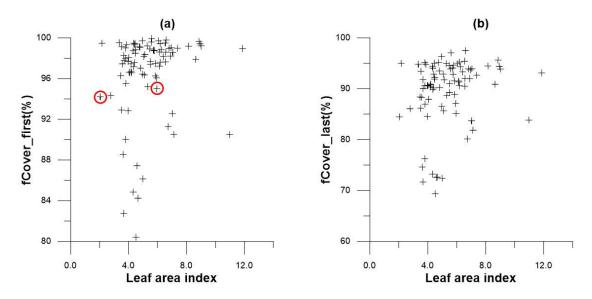


Figure 5. Scatter plots of ground LAI and LiDAR-derived fractional cover. The subplots (**a**,**b**) show the fractional cover calculated from the first and last return points respectively. The points in the red circles are used as examples in Figure 6.

Compared to the fCover metrics, the height percentile metrics are more effective for describing the structure of a tropical forest. We use two points with different LAI values (2.05 and 5.96) but similar fCover (94.2% and 95.0%) in Figure 5a (red circle) to illustrate the difficulty in estimating LAI from the LiDAR fCover variables. The LiDAR-derived height distribution are shown in Figure 6. Although the fCover values are similar for the points (Figure 5a), much difference of the tree height percentiles is observed in Figure 6. This result indicates that, as the fCover is a two-dimensional (2-D) variable that is used to represent the three-dimensional (3-D) tree canopy, in order to determine the variation of the leaf amount along the vertical profile, variables related to the canopy's internal structure, such as multi-height percentile variables, might be more suitable for estimating the LAI for very dense canopies.

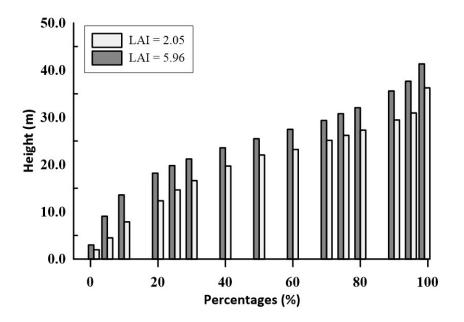


Figure 6. LiDAR-derived height cumulative percentile of the two example points of LAI = 2.05 and LAI = 5.96, as indicated in Figure 5a

We calculate the correlation coefficients between the plot-based LAI and the fifteen height percentiles and the plot is illustrated in Figure 7. Two points can be deduced from the results. First, all the height percentile variables have a positive correlation with the canopy LAI. This is in agreement with the well-known fact that trees with a larger height have more leaves to support their photosynthetic activities [25,66]. However, we want to emphasize the second point revealed in this figure. The coefficients range from 0.22 to 0.69 and the strongest correlations are observed for the 60th height percentile. Although all the percentile variables have a positive correlation with the LAI, none of them explain more than 50% of the variability of the LAI if only one height percentile variable is used as a predictor in a linear regression model (maximum $R^2 = 0.48$). Thus, it is difficult to estimate LAI using an ordinary linear regression method.

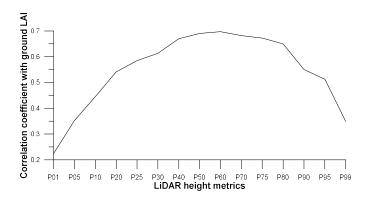


Figure 7. Correlation coefficients between plot-based LAI with LiDAR height metrics. The labels of the *x*-axis ticks, P01, P05, ..., P99 indicate the percentile values of the canopy height at the 1st, 5th, ..., 99th percentiles.

Applying a multiple linear or stepwise regression may partly solve the above-mentioned problem by using multiple LiDAR-derived metrics as predictors of the LAI. However, another problem is that the predictors in the regression model exhibit collinearity, which can be manifested by their correlation coefficients as illustrated in Figure 8. This result shows that the largest coefficients are near the diagonal line, i.e., adjacent height percentile variables have stronger correlations.

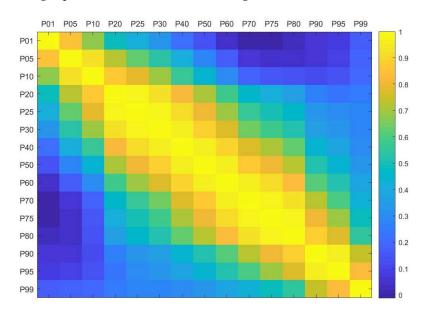


Figure 8. Color plot of Pearson's correlation coefficient between the height percentile metrics (*x* and *y* labels).

The basis of the PLS model is to maximize the information of multiple variables by projecting the original variables (*x* or predictors and *y* or response) into fewer "newer" components with minimum correlation [60]. The performance of the projection process can be indicated by the MSE of the new components relative to the original variables. In theory, more predictors provide more explanatory power; however, using redundant predictors results in overfitting. The optimum number of predictors can be determined by visually checking the plot of the explained variance by the model vs. the number of the PLS components and the MSE of the predictors and response variables vs. the number of the PLS components.

The Matlab PLSREGRESS function produces the explained variance vs. the number of predictors and we plot it in Figure 9a. The results show that 62.5% of the variance of the model is explained when only the first component is used as a predictor; this number increases to 67% when five components are incorporated into the PLS model. When more predictors are used, the rate of increase in the explained variance decreases. However, this does not imply that five components are the optimum number. The best option to select predictors is to combine the results of the explained variance and the predictive error of the candidate models. Figure 9b shows that the lowest error for the response variable (LAI) and a relatively low error for the predictor variables (height percentiles) occur when one predictor is used. After that point, there is no apparent improvement in the response variable although the MSE of the predictor variables continues to decrease. This indicates that the contribution of the second and subsequent components may result in an overfitting of the model of the LAI. Therefore, based on the analysis of the PLS predictors, we select one component as the predictor of the final PLS model.

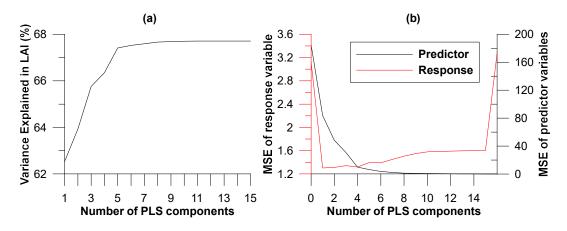


Figure 9. Percent of the explained variance (**a**) and the mean squared error (MSE) vs. the number of partial least squares (PLS) components (**b**).

The major difference between the PLS model and an ordinary linear regression model is that the former uses the transformed variables rather than the original LiDAR metrics, i.e., any of the selected predictors is the linear combination of the 15 height percentile metrics. The weight of the linear combination corresponds to the contribution of the original LiDAR metrics to the predictor.

Since the first predictor of the PLS explains nearly 62.5% variance of the plot LAI (Figure 9a), the weight of the first predictor which is one of the outputs of Matlab PLSREGRESS function, is plotted in Figure 10 to better understand the contribution of the individual LiDAR height percentile metrics. We can see that the contribution of the first predictor is higher for the height percentiles that are larger than the 10th percentile. Recalling the correlation coefficients in Figure 7, it can be inferred that the PLS uses all the information of the LiDAR metrics that have a relatively high correlation with the plot LAI. In this step, we find that all the metrics have a relatively high contribution except for the lower percentile heights, which are near the ground, i.e., the height of the 1st and 5th percentile.

height metrics higher than the 5th percentile, P20 and P95 have the highest weight. This shows that the information from the upper layer (P95) and the low-middle layer (P20) have similar and high contributions to the estimated LAI.

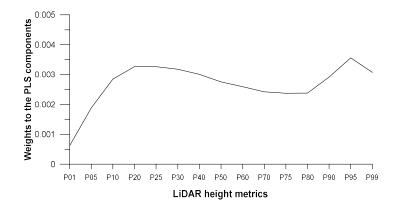


Figure 10. The weight (contribution) of the LiDAR-derived height percentile metrics to the first PLS component. The *x*-axis labels represent the height percentile metrics as indicated in the *x*-axis of Figure 7.

The PLS model is then applied on the plot level and the 85 plot-level LAI values are calculated from the LiDAR height percentile metrics. The LOOCV result is shown in Figure 11. A Pearson's correlation analysis is performed to determine the relationship between the predicted LiDAR LAI values and the ground-based values. There is a very strong, positive correlation between the LiDAR LAI and ground LAI values (N = 85, r = 0.72, p < 0.0001) as indicated by the scatter plot (Figure 11a). The PLS model of the LAI explains 58% of the variation in the ground-based LAI with an RMSE of 1.13.

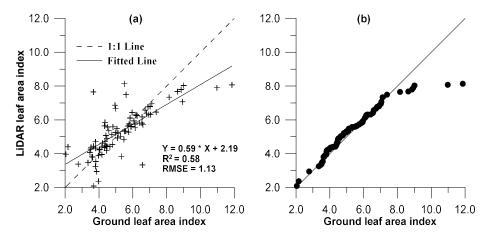


Figure 11. Leave-one-out cross-validation result of the PLS model. The scatter plot (**a**) shows the paired comparison and the Q-Q plot (**b**) shows the similarity in the distribution of the two LAI datasets.

The accuracy of the estimated LAI is comparable with that of a previous study conducted in a tropical forest. Tang et al. [67] estimated the LAI of a tropical forest at La Selva, Costa Rica using airborne LiDAR data and an inverted geometric optical and radiative transfer (GORT) model [68]. This study was conducted in a similar forest environment where the maximum LAI was 10.0. The reported R^2 and RMSE values ranged from 0.42 to 0.63 and 1.36 to 1.91 respectively.

In our study, when the ground-based LAI is lower than 8.0, the points in the scatter plot are near the 1:1 line, which indicates that there is a low bias between the LiDAR-based LAI and the ground-based LAI, as shown in the Q-Q plot in Figure 11b. A Q-Q plot is a probability plot that

compares the probability distribution of two datasets by plotting their quantiles against each other. Each point plotted on the Q-Q plot represents the same quantile in each data set. If the two distributions are similar, the points will lie on or near the line y = x. As suggested by the Q-Q plot, it can be concluded that the LAI datasets have similar distributions when the LAI values are lower than 8.0.

For LAI values that are higher than 8.0, the ground-based LAI values tend to be underestimated by the PLS model. This can be partly explained by the fact that the LiDAR signal is less sensitive to very high LAI values in areas of dense leaf layers. Although the active LiDAR sensor is superior to passive remote sensing for determining the LAI in dense forests, signal saturation still occurs. However, the saturation occurs at higher LAI values than in reported passive optical methods, in which saturation occurs when the LAI > 4.0 or 6.0, depending on the measuring instruments [65] or the forest structure [1,69,70].

3.4. Comparison of LiDAR and MODIS LAI

There are 43 MCD15A3H pixels that meet the criterion that at least 70 LiDAR cells fall in one MODIS pixel. In the corresponding MODIS pixels, we compare the LiDAR- and MODIS-based LAI in Figure 12. The average value of the landscape-level LiDAR-based LAI (6.26) is very close to the ground-measured result reported by Williams et al. [71] in the Eastern Amazon forest. They reported that at the regional scale, the LAI at 13 sites varied from 5.2 to 7.0 with a mean of 6.2. In general, the LiDAR-based LAI values have a higher mean than the MODIS product and the bias is 1.38. This means that MODIS underestimates the tropical forest LAI by about 22% compared to the LiDAR-derived LAI. Previous studies also noted the underestimation of the MODIS-based LAI compared to ground measurements in tropical forests [72–74]. For example, the LAI values from MODIS were 17% lower than the LAI ground measurements in a forest with high LAI values [73]. In a conifer forest, Jensen et al. [75] observed that the MODIS LAI product resulted in an average underestimation of 10%–50%, depending on the algorithm.

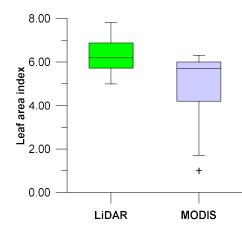


Figure 12. Boxplots of LiDAR landscape-level LAI and the MODIS MCD15A3H product. The central line in the box indicates the median and the bottom and top edges indicate the 25th and 75th percentiles. Outliers are plotted using the "+" symbol.

In the boxplots, the different positions of the median values also suggest that the MODIS observations deviate considerably from its mean value (middle of the upper and lower box boundaries) whereas the LiDAR data provides a nearly normal distribution. The asymmetric distribution of the MODIS LAI values indicates a skew towards the upper boundary (75% quantile), which indicates that there are many MODIS observations with high LAI values. This may be explained by the fact that in dense canopies, the MODIS surface reflectance was saturated and was, therefore, insensitive to changes in the LAI [30].

To further address the specific factors affecting the performance of the MODIS LAI product, we plot the 43 landscape-level pixels of the LiDAR LAI and the corresponding pixels of the MODIS LAI in Figure 13. The standard deviations of the LiDAR LAI data are also plotted as the vertical bars.

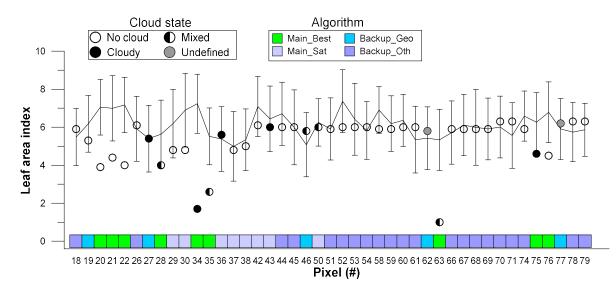


Figure 13. Landscape-level LiDAR LAI values of the 43 pixels and the corresponding MODIS LAI values. The vertical bars indicate one standard deviation of the LiDAR LAI values.

The results in Figure 13 indicate that most pixels are cloud-free and only 5 pixels contain clouds. Although cloud-free conditions prevailed, when grouping the MODIS LAI values by the algorithm, we observe a relative low retrieval ratio [30], i.e., only 9 out of 43 pixels are fall into the Main_Best category (green labels). In addition, the LAI values of those pixels are relatively low compared to those produced by the backup algorithms (Backup_Geo and Backup_Oth) or the solution that indicates a saturated reflectance (Main_Sat). The reason for the higher LAI values obtained from the backup and the Main_Sat algorithms is related to the reflectance saturation of a dense forest, which has been well documented in previous studies [9,76]. We will focus on the underestimation of the Main_Best algorithm because it is assumed to provide the best quality LAI data.

Pinto-Júnior et al. [74] stated that in the southern Amazon Basin, the presence of clouds increased the reflectance values and consequently reduced the mean LAI. However, this is not the case in our study. The results of this study suggest that in the dry season [4] of the Amazon tropical forest, the cloud condition is not the primary factor affecting the MODIS LAI products. The main algorithm of the MODIS LAI product uses the reflectance in the LUT simulated from a 3-D RT model to match the MODIS surface reflectance [30]. The low retrieval ratio of the main algorithm can be explained by the inability of simulating a range of observed reflectance values over the Amazon tropical forest [77]. The lower LAI values of the main algorithm may be attributed to the fact that the MODIS product is predominantly correlated with the overstory LAI rather than the total forest LAI [73]. This result partly supports the findings of [75], who reported that the MODIS main algorithm was rarely used to retrieve LAI values greater than 4.0 when they evaluated the MODIS LAI values of a conifer forest.

The findings of this study are also consistent with the results of Cohen et al. [78], who validated the MODIS LAI in a tropical forest using the Bigfoot field data (TAPA site, part of the Large-Scale Biosphere-Atmosphere Experiment in Amazonia) [78]. Their result showed that the MODIS LAI product was very unstable in a tropical forest area and the mean generally varied between 3.0 and 6.0. Although they validated the Collection 4 MODIS LAI, which differs from the version (Collection 6) used in this study work, the main conclusion still holds for the comparison of the MODIS LAI with the LiDAR LAI data.

We then explore the spatial distribution of the two types of landscape-level LAI products by mapping them at two scales, i.e., first, the LiDAR-based LAI is scaled up to the same resolution as the MODIS data (Figure 14a,b), and then the data are aggregated at the 100-ha scale and overlaid with the work units (Figure 14c,d). We expected to find a spatial correlation between the LiDAR-and MODIS-based LAI values and wanted to investigate the impact of the logging operation on the landscape-level LAI values.

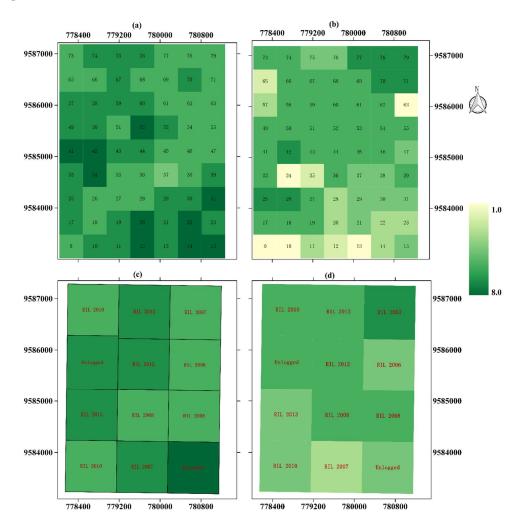


Figure 14. LiDAR-based LAI and MODIS-based LAI values in resolution of 463 m (**a**,**b**) and aggregated into work-unit scale (**c**,**d**). The left column shows the LiDAR LAI (**a**,**c**) and the right column the MODIS LAI (**b**,**d**). The pixel indices are indicated in subplot (**a**,**b**). The years of the reduced-impact logging (RIL) or the unlogged status are indicated in the labels of the subplots in (**c**,**d**) in red text.

A visual analysis of the 463-m resolution maps (Figure 14a,b) does not reveal any spatial patterns, which indicates that there is no apparent spatial correlation between the two datasets. Specifically, the MODIS-based LAI exhibits more homogeneity than the LiDAR-based LAI and the homogeneous patches also have relatively higher LAI values. Yet, this homogeneity does not reflect the ground-based spatial patterns shown in the LiDAR-based LAI values (Figure 14a) but rather results from the reflectance saturation that affects the MODIS algorithm. This result is also manifested by the skewness in the boxplot (Figure 12) and the larger values produced by the backup algorithm to the right part to the pixel index 43 in Figure 13.

In the aggregated 100-ha scale LAI maps with the work units, the maximum value of the LiDAR LAI in the unlogged area is 7.31 (right bottom in Figure 14c) and the average LiDAR LAI value for the two unlogged blocks (6.58 and 7.31) is 6.94 (Figure 15). There were no LAI measurements in logged areas to compare our results with, but some gap fraction data were available. As there is an exponential function relationship between LAI and gap fraction, we converted the calculated LAI into gap fraction (GF) according to the Beer-Lambert law and using the extinction coefficient of 0.5 [79,80],

$$GF = \exp(-0.5LAI) \tag{5}$$

The calculated gap fraction values of the two unlogged blocks are 0.037 and 0.026 respectively, and the average gap fraction of two unlogged blocks is 0.032. This result is almost identical to the result of Pereira et al. [4], who reported a gap fraction of 0.031 in unlogged blocks at the landscape scale while they investigated the impact of ground logging activities in a tropical forest in an area (3°43.878'S, 48°17.438'W) very close to our research area.

The logged blocks have an average lower LAI value of 6.04 (Figure 15); compared with the unlogged blocks, this amounts to a 13.0% reduction in the LAI caused by the RIL. This value is similar to the result of Pereira et al. [4], who reported a reduction in the LAI of 13.2% after three years of RIL but it is lower than the value of 36% when RIL is conducted within one year (reproduced from Table 3 in [4]). In the same area, this result is very close to the reduction in the AGB of 15.5% reported by Longo et al. [50], who compared the aboveground carbon density of intact and RIL forests. This result is also consistent with the findings of Aragão et al. [81], who suggested that selective logging practices in Eastern Amazonia did not result in significant changes in the LAI values.

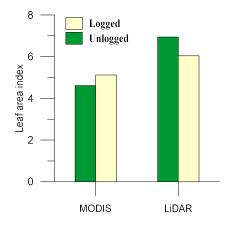


Figure 15. Comparison of MODIS and LiDAR LAI values in logged and unlogged areas.

In contrast, the maximum value of the MODIS LAI occurs in the logged area in 2007 (right upper in Figure 14d), which is irrational taking into account the logging activity although some regrowth of the forest had occurred. Furthermore, compared to the unlogged area, the average LAI in the logged areas exhibit an unexpected higher value (Figure 15). Considering the data quality analysis mentioned above (Section 3.4), it can be concluded that the MODIS LAI product needs be improved to account for the forest characteristics in the Amazon basin. For example, the LUT should be based on a more detailed land cover classification, or more complex RT model with higher accuracy under saturated conditions should be used [9].

3.6. Factors Related to the Uncertainty of the Results

There was a lag of about 8–10 months between the ground inventory (February to April) and the LiDAR (December) collection dates; therefore, when the LiDAR-derived metrics were used to link them with the ground-based LAI, we assumed that the LAI at the study site had remained stable.

Although we could not fully validate this assumption for this research area in the absence of a time series of ground-measured LAI values, there is evidence supporting that there is no strong pattern of seasonal variation in the LAI in the Amazonian tropical forest [48]. There is also some evidence from ground camera images from the PhenoCam Network in Rio Tapajos, Santarém, Brazil, where a similar forest type was observed from 2009 to 2011 [82]. Here, we present three pictures from January to December in 2011 to determine if seasonal dynamics of the forest were apparent. Greenness is one of the proxy variable to estimate LAI from RGB images [83]. similar to [83], we calculate the greenness index (GI), which is the ratio of digital number of green band to the sum of red, green and blue bands, to represent the forest structure in daily image, and take the averaged GI from month images as month indictor (Figure 16). In the selected 3 months, the value of GI varied from 0.335 to 0.338, and this result shows that there is no evidence of leaf dynamics in this area (Figure 16). Based on these results, the uncertainty caused by the date lag is negligible.



Figure 16. Three images obtained in different seasons of an Amazon tropical forest near the study area.

Two parameters can affect the ground-based LAI values, i.e., the LBR and the SLA when the ground-based LAIs are calculated from the AGB. In fact, it is difficult to measure the above parameters at the landscape scale, even though some researchers demonstrated that the relationship between leaf area and plant biomass was non-linear and varied depending on the carbon partitioning [84]. For example, Metcalfe et al. [61] suggested that the SLA exhibited significant change over time and the annual SLA ranged from 8.5 to 13 in a tropical forest. Kalacska et al. [1] also suggested that in order to calculate LAI from litter collected in traps, the SLA must be calculated for each species rather than using the mean value of multiple species. To minimize the uncertainty of the LBR and SLA, we calculated the leaf biomass using a piecewise linear equation to simulate the nonlinear feature of the relationship between the plant biomass and the leaf biomass (Equation (2)). The SLA values were obtained from the literature while accounting for the similarity of the dominant plant species. Although a careful effort was made to select these parameters, the results presented may be improved when newer ground data are obtained in this research area. However, the SLA parameter only affects the magnitude of the LiDAR-based LAI and the correlation between the ground-based LAI and LiDAR-based height percentiles still holds even if new SLA values are available. In contrast, changes in the LBR will change the result with regard to magnitude and correlation because its effect on the ground-based LAI is nonlinear, as shown in Equation (2).

The performance of the PLS model will be influenced by the data used as a validation dataset. In this study, to maximize the use of the ground LAI, we calibrated and validated the PLS model using the same dataset but with a leave-one-out cross validation (LOOCV) method. The ideal method is to separate the original dataset into two independent subsets, i.e., one for calibration and another for validation. However, it would require a larger dataset than we had available for this study.

4. Conclusions

In this study, we proposed a framework to estimate the LAI in a selectively logged tropical forest at the plot scale and landscape scale. The ground-based LAI values were computed from forest

inventory data measured in the field and a PLS regression method was proposed to determine the ground-based LAI and the plot and landscape scales. Finally, the LiDAR-based LAI was compared with the MODIS-based LAI for a number of logging units.

The proposed PLS method uses the LiDAR-based height percentiles and differs from an approach using only one LiDAR metric. The results show that, by utilizing multi-height percentile metrics, which are related to the leaf density of the canopy layers in the vertical profile, the new method has the ability to obtain canopy information from multiple layers rather than only from the top of the canopy. The plot-level LAI obtained by the PLS model is significantly correlated with the ground-based LAI with $R^2 = 0.58$ and RMSE = 1.13. The weight of the PLS model suggests that all the selected height percentile metrics are positively correlated with the LAI but have different explanatory power. The first component of the PLS model explains 62.5% of the variance of the LAI and its major contribution comes from the low-middle to upper layers of the canopy.

Compared with the LiDAR-based LAI, the MODIS-based LAI values underestimated the tropical forest LAI by about 22%. A quality analysis of the selected tile reveals that clouds are not an important factor affecting the performance of the MODIS LAI but the main algorithm, which produces significantly lower LAI values, may be the main source of uncertainty. The LAI values from the backup algorithm have relatively higher values but are a result of the reflectance saturation and skewed towards the upper boundary of the results. The MODIS- and LiDAR-based LAI values show no apparent spatial correlation at the landscape scale. The LiDAR-based LAI estimates indicate that the RIL activity causes a decrease in the LAI of about 13% compared to the intact forest, but this is not captured by the MODIS-based LAI estimates.

It should be noted that the work of comparison with the MODIS-based LAI and the LiDAR-based LAI is not a substitution for the MODIS product validation in this area. A strict validation campaign involves more aspects than were considered in this study, such as a ground sampling strategy, transferring function, and upscaling [76,77]. Furthermore, multi-temporal LAI datasets may be useful to fully understand the LAI dynamics across seasons. However, this validation scheme is beyond the scope of this study.

To our knowledge, this was the first effort to compute LAI estimates from a ground-based inventory dataset in a selectively logged tropical forest. This can overcome the lack of LAI ground data because forest inventory data are much more abundant than LAI ground data. The significance of this effort is also reflected by the development of a robust method for estimating LAI in a tropical forest using LiDAR active remote sensing. The proposed workflow may be of interest to researchers in the field of forest management who want to develop LAI datasets from the ground to the landscape scale, but it is advised that the parameters discussed in Section 3.6 should be calibrated locally.

Supplementary Materials: The following are available online at http://www.mdpi.com/2072-4292/10/6/970/s1. Matlab script to model ground LiDAR LAI using partial least square (PLS) regression method; LiDAR metrics and the ground plot LAI which were used to develop the PLS model.

Author Contributions: Y.Q. and A.S. conceived and designed the framework of this paper; Y.Q. analyzed the data and wrote the paper; C.A.S. contributed the co-registration of ground plot data with LiDAR data and mapping some of the figures; C.K. and E.R.P. contributed to the improvement of the manuscript.

Funding: This work was supported in part by the National Natural Science Foundation of China (41671333/41531174) and National key research and development program of China (2016YFC0500103).

Acknowledgments: LiDAR and forest inventory data were acquired with support from USAID and the US Department of State with the technical assistance of the Brazilian Corporation for Agricultural Research (EMBRAPA) and the US Forest Service Office of International Programs via the Sustainable Landscapes Brazil Project. The MODIS LAI product and MRT tools were supported by the funding from the NASA Earth Science Data and Information Systems (ESDIS) Project. Many thanks to Prof. Jiquan Chen of Michigan State University for his help in proof-reading on trees' Latin name.

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

References

- Kalacska, M.; Calvo-Alvarado, J.C.; Sanchez-Azofeifa, G.A. Calibration and assessment of seasonal changes in leaf area index of a tropical dry forest in different stages of succession. *Tree Physiol.* 2005, 25, 733–744. [CrossRef] [PubMed]
- 2. Restrepo-Coupe, N.; da Rocha, H.R.; Hutyra, L.R.; da Araujo, A.C.; Borma, L.S.; Christoffersen, B.; Cabral, O.M.R.; de Camargo, P.B.; Cardoso, F.L.; da Costa, A.C.L.; et al. What drives the seasonality of photosynthesis across the Amazon basin? A cross-site analysis of eddy flux tower measurements from the Brasil flux network. *Agric. For. Meteorol.* **2013**, *182–183*, 128–144. [CrossRef]
- 3. Chen, J.M.; Black, T.A. Defining leaf area index for non-flat leaves. *Plant Cell Environ.* **1992**, *15*, 421–429. [CrossRef]
- 4. Pereira, R.; Zweede, J.; Asner, G.P.; Keller, M. Forest canopy damage and recovery in reduced-impact and conventional selective logging in eastern Para, Brazil. *For. Ecol. Manag.* **2002**, *168*, 77–89. [CrossRef]
- 5. Asner, G.P.; Knapp, D.E.; Broadbent, E.N.; Oliveira, P.J.C.; Keller, M.; Silva, J.N. Selective Logging in the Brazilian Amazon. *Science* 2005, *310*, 480–482. [CrossRef] [PubMed]
- Pfeifer, M.; Kor, L.; Nilus, R.; Turner, E.; Cusack, J.; Lysenko, I.; Khoo, M.; Chey, V.K.; Chung, A.C.; Ewers, R.M. Mapping the structure of Borneo's tropical forests across a degradation gradient. *Remote Sens. Environ.* 2016, 176, 84–97. [CrossRef]
- Jonckheere, I.; Fleck, S.; Nackaerts, K.; Muys, B.; Coppin, P.; Weiss, M.; Baret, F. Review of methods for in situ leaf area index determination: Part I. Theories, sensors and hemispherical photography. *Agric. For. Meteorol.* 2004, 121, 19–35. [CrossRef]
- 8. Chen, J.M.; Cihlar, J. Plant canopy gap-size analysis theory for improving optical measurements of leaf-area index. *Appl. Opt.* **1995**, *34*, 6211–6222. [CrossRef] [PubMed]
- 9. Shabanov, N.V.; Huang, D.; Yang, W.; Tan, B.; Knyazikhin, Y.; Myneni, R.B.; Ahl, D.E.; Gower, S.T.; Huete, A.R.; Aragao, L.E.O.C.; et al. Analysis and optimization of the MODIS leaf area index algorithm retrievals over broadleaf forests. *IEEE T Geosci. Remote Sens.* **2005**, *43*, 1855–1865. [CrossRef]
- 10. Qu, Y.; Han, W.; Ma, M. Retrieval of a Temporal High-Resolution Leaf Area Index (LAI) by Combining MODIS LAI and ASTER Reflectance Data. *Remote Sens.* **2015**, *7*, 195–210. [CrossRef]
- 11. Chen, J.M.; Cihlar, J. Retrieving leaf area index of boreal conifer forests using Landsat TM images. *Remote Sens. Environ.* **1996**, *55*, 153–162. [CrossRef]
- 12. Riaño, D.; Valladares, F.; Condés, S.; Chuvieco, E. Estimation of leaf area index and covered ground from airborne laser scanner (Lidar) in two contrasting forests. *Agric. For. Meteorol.* **2004**, *124*, 269–275. [CrossRef]
- 13. Farid, A.; Goodrich, D.C.; Bryant, R.; Sorooshian, S. Using airborne lidar to predict Leaf Area Index in cottonwood trees and refine riparian water-use estimates. *J. Arid Environ.* **2008**, *72*, 1–15. [CrossRef]
- 14. Hayduk, E.A. Using LiDAR Data to Estimate Effective Leaf Area Index, Determine Biometrics and Visualize Canopy Structure in a Central Oregon Forest with Complex Terrain. Master's Thesis, The Evergreen State College, Olympia, WA, USA, 2012.
- 15. You, H.; Wang, T.; Skidmore, K.A.; Xing, Y. Quantifying the Effects of Normalisation of Airborne LiDAR Intensity on Coniferous Forest Leaf Area Index Estimations. *Remote Sens.* **2017**, *9*, 163. [CrossRef]
- 16. Heiskanen, J.; Korhonen, L.; Hietanen, J.; Pellikka, P.K.E. Use of airborne lidar for estimating canopy gap fraction and leaf area index of tropical montane forests. *Int. J. Remote Sens.* **2015**, *36*, 2569–2583. [CrossRef]
- 17. Zhao, K.; Popescu, S. Lidar-based mapping of leaf area index and its use for validating GLOBCARBON satellite LAI product in a temperate forest of the southern USA. *Remote Sens. Environ.* **2009**, *113*, 1628–1645. [CrossRef]
- 18. Solberg, S.; Brunner, A.; Hanssen, K.H.; Lange, H.; Næsset, E.; Rautiainen, M.; Stenberg, P. Mapping LAI in a Norway spruce forest using airborne laser scanning. *Remote Sens. Environ.* **2009**, *113*, 2317–2327. [CrossRef]
- 19. Richardson, J.J.; Moskal, L.M.; Kim, S.-H. Modeling approaches to estimate effective leaf area index from aerial discrete-return LIDAR. *Agric. For. Meteorol.* **2009**, *149*, 1152–1160. [CrossRef]
- 20. Pope, G.; Treitz, P. Leaf Area Index (LAI) Estimation in Boreal Mixedwood Forest of Ontario, Canada Using Light Detection and Ranging (LiDAR) and WorldView-2 Imagery. *Remote Sens.* 2013, *5*, 5040–5063. [CrossRef]
- 21. Jensen, J.L.R.; Humes, K.S.; Vierling, L.A.; Hudak, A.T. Discrete return lidar-based prediction of leaf area index in two conifer forests. *Remote Sens. Environ.* **2008**, *112*, 3947–3957. [CrossRef]
- 22. Xia, S.; Wang, C.; Pan, F.; Xi, X.; Zeng, H.; Liu, H. Detecting Stems in Dense and Homogeneous Forest Using Single-Scan TLS. *Forests* **2015**, *6*, 3923–3945. [CrossRef]

- Clark, D.B.; Olivas, P.C.; Oberbauer, S.F.; Clark, D.A.; Ryan, M.G. First direct landscape-scale measurement of tropical rain forest Leaf Area Index, a key driver of global primary productivity. *Ecol. Lett.* 2008, 11, 163–172. [CrossRef] [PubMed]
- 24. McDowell, N.; Barnard, H.; Bond, B.; Hinckley, T.; Hubbard, R.; Ishii, H.; Köstner, B.; Magnani, F.; Marshall, J.; Meinzer, F.; et al. The relationship between tree height and leaf area: Sapwood area ratio. *Oecologia* **2002**, *132*, 12–20. [CrossRef] [PubMed]
- 25. Yuan, Y.; Wang, X.; Yin, F.; Zhan, J. Examination of the Quantitative Relationship between Vegetation Canopy Height and LAI. *Adv. Meteorol.* **2013**, 2013. [CrossRef]
- 26. Cavaleri, M.A.; Oberbauer, S.F.; Clark, D.B.; Clark, D.A.; Ryan, M.G. Height is more important than light in determining leaf morphology in a tropical forest. *Ecology* **2010**, *91*, 1730–1739. [CrossRef] [PubMed]
- 27. Donoghue, D.N.M.; Watt, P.J.; Cox, N.J.; Wilson, J. Remote sensing of species mixtures in conifer plantations using LiDAR height and intensity data. *Remote Sens. Environ.* **2007**, *110*, 509–522. [CrossRef]
- 28. Kiers, H.A.L.; Smilde, A.K. A comparison of various methods for multivariate regression with highly collinear variables. *Stat. Methods Appl.* **2007**, *16*, 193–228. [CrossRef]
- 29. Tang, J.; Zhang, J.; Wu, Z.; Liu, Z.; Chai, T.; Yu, W. Modeling collinear data using double-layer GA-based selective ensemble kernel partial least squares algorithm. *Neurocomputing* **2017**, *219*, 248–262. [CrossRef]
- 30. Myneni, R.B.; Hoffman, S.; Knyazikhin, Y.; Privette, J.L.; Glassy, J.; Tian, Y.; Wang, Y.; Song, X.; Zhang, Y.; Smith, G.R. Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. *Remote Sens. Environ.* **2002**, *83*, 214–231. [CrossRef]
- 31. Knyazikhin, Y.; Glassy, J.; Privette, J.L.; Tian, Y.; Lotsch, A.; Zhang, Y.; Wang, Y.; Morisette, J.T.; Votava, P.; Myneni, R.B. MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15) Algorithm Theoretical Basis Document; Theoretical Basis Document; NASA Goddard Space Flight Center: Greenbelt, MD, USA, 1999; Volume 20771.
- 32. Daac, N.L. Version 6. Leaf Area Index and Fractional Photosynthetically Active Radiation, NASA EOSDIS Land Processes. Available online: https://lpdaac.usgs.gov (accessed on 8 March 2018).
- 33. Saleska, S.R.; Didan, K.; Huete, A.R.; Da Rocha, H.R. Amazon forests green-up during 2005 drought. *Science* **2007**, *318*, 612. [CrossRef] [PubMed]
- Myneni, R.B.; Yang, W.; Nemani, R.R.; Huete, A.R.; Dickinson, R.E.; Knyazikhin, Y.; Didan, K.; Fu, R.; Negron Juarez, R.I.; Saatchi, S.S.; et al. Large seasonal swings in leaf area of Amazon rainforests. *Proc. Natl. Acad. Sci. USA* 2007, 104, 4820–4823. [CrossRef] [PubMed]
- 35. Sanchez-Azofeifa, A.; Antonio Guzmán, J.; Campos, C.A.; Castro, S.; Garcia-Millan, V.; Nightingale, J.; Rankine, C. Twenty-first century remote sensing technologies are revolutionizing the study of tropical forests. *Biotropica* **2017**, *49*, 604–619. [CrossRef]
- Rosipal, R.; Krämer, N. Overview and Recent Advances in Partial Least Squares. In Subspace, Latent Structure and Feature Selection. SLSFS 2005; Saunders, C., Grobelnik, M., Gunn, S., Shawe-Taylor, J., Eds.; Springer: Berlin/Heidelberg, Germay, 2006; pp. 34–51.
- Costa, M.H.; Foley, J.A. A comparison of precipitation datasets for the Amazon Basin. *Geophys. Res. Lett.* 1998, 25, 155–158. [CrossRef]
- 38. IBGE Mapa de Vegetação do Brasil. Escala 1:5.000.000. Available online: ftp://geoftp.ibge.gov.br/ informacoes_ambientais/vegetacao/mapas/brasil/vegetacao.pdf (accessed on 3 March 2018).
- 39. IBGE Pedologia—1:250.000. Available online: https://www.ibge.gov.br/geociencias-novoportal/ informacoes-ambientais/pedologia.html (accessed on 3 March 2018).
- Silva, A.C.; Hudak, T.A.; Vierling, A.L.; Klauberg, C.; Garcia, M.; Ferraz, A.; Keller, M.; Eitel, J.; Saatchi, S. Impacts of Airborne Lidar Pulse Density on Estimating Biomass Stocks and Changes in a Selectively Logged Tropical Forest. *Remote Sens.* 2017, 9, 1086. [CrossRef]
- 41. dos-Santos, M.N.; Keller, M.M. *CMS: Forest Inventory and Biophysical Measurements, Para, Brazil, 2012–2014;* ORNL DAAC: Oak Ridge, TN, USA, 2016. [CrossRef]
- 42. Chave, J.; Réjou-Méchain, M.; Búrquez, A.; Chidumayo, E.; Colgan, M.S.; Delitti, W.B.C.; Duque, A.; Eid, T.; Fearnside, P.M.; Goodman, R.C.; et al. Improved allometric models to estimate the aboveground biomass of tropical trees. *Glob. Chang. Biol.* **2014**, *20*, 3177–3190. [CrossRef] [PubMed]
- 43. Chave, J.; Coomes, D.; Jansen, S.; Lewis, S.L.; Swenson, N.G.; Zanne, A.E. Towards a worldwide wood economics spectrum. *Ecol. Lett.* **2009**, *12*, 351–366. [CrossRef] [PubMed]

- 44. McWilliam, A.L.C.; Roberts, J.M.; Cabral, O.M.R.; Leitao, M.V.B.R.; de Costa, A.C.L.; Maitelli, G.T.; Zamparoni, C.A.G.P. Leaf Area Index and Above-Ground Biomass of terra firme Rain Forest and Adjacent Clearings in Amazonia. *Funct. Ecol.* **1993**, *7*, 310–317. [CrossRef]
- 45. Fittkau, E.J.; Klinge, H. On Biomass and Trophic Structure of the Central Amazonian Rain Forest Ecosystem. *Biotropica* **1973**, *5*, 2–14. [CrossRef]
- Reich, P.B.; Luo, Y.; Bradford, J.B.; Poorter, H.; Perry, C.H.; Oleksyn, J. Temperature drives global patterns in forest biomass distribution in leaves, stems, and roots. *Proc. Natl. Acad. Sci. USA* 2014, *111*, 13721–13726. [CrossRef] [PubMed]
- Vile, D.; Garnier, É.; Shipley, B.; Laurent, G.; Navas, M.-L.; Roumet, C.; Lavorel, S.; DÍAz, S.; Hodgson, J.G.; Lloret, F.; Midgley, G.F.; et al. Specific Leaf Area and Dry Matter Content Estimate Thickness in Laminar Leaves. *Ann. Bot.-Lond.* 2005, *96*, 1129–1136. [CrossRef] [PubMed]
- 48. Malhado, A.C.M.; Costa, M.H.; de Lima, F.Z.; Portilho, K.C.; Figueiredo, D.N. Seasonal leaf dynamics in an Amazonian tropical forest. *For. Ecol. Manag.* **2009**, *258*, 1161–1165. [CrossRef]
- 49. dos-Santos, M.N.; Keller, M.M. *CMS: LiDAR Data for Forested Areas in Paragominas, Para, Brazil, 2012–2014;* ORNL DAAC: Oak Ridge, TN, USA, 2016. [CrossRef]
- 50. Longo, M.; Keller, M.; dos-Santos, M.N.; Leitold, V.; Pinagé, E.R.; Baccini, A.; Saatchi, S.; Nogueira, E.M.; Batistella, M.; Morton, D.C. Aboveground biomass variability across intact and degraded forests in the Brazilian Amazon. *Glob. Biogeochem. Cycles* **2016**, *30*, 1639–1660. [CrossRef]
- 51. Isenburg, M. Rapidlasso Gmbh—Fast Tools to Catch Reality. Available online: https://rapidlasso.com/lastools/ (accessed on 8 March 2018).
- 52. Luo, S.; Chen, J.M.; Wang, C.; Xi, X.; Zeng, H.; Peng, D.; Li, D. Effects of LiDAR point density, sampling size and height threshold on estimation accuracy of crop biophysical parameters. *Opt. Express* **2016**, *24*, 11578–11593. [CrossRef] [PubMed]
- 53. Singh, K.K.; Chen, G.; McCarter, J.B.; Meentemeyer, R.K. Effects of LiDAR point density and landscape context on estimates of urban forest biomass. *ISPRS J. Photogramm.* **2015**, *101*, 310–322. [CrossRef]
- 54. McGauchey, R.J. *FUSION/LDV: Software for LiDAR Data Analysis and Visualization;* Forest Service Pacific Northwest Research Station USDA: Seattle, WA, USA, 2016; Available online: http://forsys.cfr.washington.edu/fusion/FUSION_manual.pdf (accessed on 8 March 2018).
- 55. Myneni, R.; Knyazikhin, Y.; Park, T. MCD15A3H MODIS/Terra+Aqua Leaf Area Index/FPAR 4-day L4 Global 500m SIN Grid V006. 2015. Available online: http://dx.doi.org/10.5067/MODIS/MCD15A3H.006 (accessed on 1 March 2018).
- 56. QGIS Development Team, QGIS Geographic Information System. Available online: https://qgis.org (accessed on 1 June 2018).
- 57. Hansen, P.M.; Schjoerring, J.K. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens. Environ.* **2003**, *86*, 542–553. [CrossRef]
- 58. Geisser, S. The Predictive Sample Reuse Method with Applications. J. Am. Stat. Assoc. 1975, 70, 320–328. [CrossRef]
- 59. *MATLAB and Statistics and Machine Learning Toolbox Release* 2017*a*; The MathWorks, Inc.: Natick, MA, USA; Available online: http://www.walkingrandomly.com/?p=4767 (accessed on 1 June 2018).
- Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemometr. Intell. Lab.* 2001, 58, 109–130. [CrossRef]
- 61. Metcalfe, D.B.; Lobo-do-Vale, R.; Chaves, M.M.; Maroco, J.P.; Aragão, C.; Luiz, E.O.; Malhi, Y.; Da Costa, A.L.; Braga, A.P.; Gonçalves, P.L.; et al. Impacts of experimentally imposed drought on leaf respiration and morphology in an Amazon rain forest. *Funct. Ecol.* **2010**, *24*, 524–533. [CrossRef]
- 62. Stark, S.C.; Leitold, V.; Wu, J.L.; Hunter, M.O.; de Castilho, C.V.; Costa, F.R.C.; McMahon, S.M.; Parker, G.G.; Shimabukuro, M.T.; Lefsky, M.A.; et al. Amazon forest carbon dynamics predicted by profiles of canopy leaf area and light environment. *Ecol. Lett.* **2012**, *15*, 1406–1414. [CrossRef] [PubMed]
- 63. Solberg, S. Mapping gap fraction, LAI and defoliation using various ALS penetration variables. *Int. J. Remote Sens.* **2010**, *31*, 1227–1244. [CrossRef]
- 64. Alonzo, M.; Bookhagen, B.; McFadden, J.P.; Sun, A.; Roberts, D.A. Mapping urban forest leaf area index with airborne lidar using penetration metrics and allometry. *Remote Sens. Environ.* **2015**, *162*, 141–153. [CrossRef]

- 65. Olivas, P.C.; Oberbauer, S.F.; Clark, D.B.; Clark, D.A.; Ryan, M.G.; O'Brien, J.J.; Ordoñez, H. Comparison of direct and indirect methods for assessing leaf area index across a tropical rain forest landscape. *Agric. For. Meteorol.* **2013**, *177*, 110–116. [CrossRef]
- 66. Byrne, K.A.; Kiely, G.; Leahy, P. CO₂ fluxes in adjacent new and permanent temperate grasslands. *Agric. For. Meteorol.* **2005**, *135*, 82–92. [CrossRef]
- 67. Tang, H.; Dubayah, R.; Swatantran, A.; Hofton, M.; Sheldon, S.; Clark, D.B.; Blair, B. Retrieval of vertical LAI profiles over tropical rain forests using waveform lidar at La Selva, Costa Rica. *Remote Sens. Environ.* **2012**, 124, 242–250. [CrossRef]
- 68. Ni-Meister, W.; Jupp, D.L.B.; Dubayah, R. Modeling lidar waveforms in heterogeneous and discrete canopies. *IEEE T Geosci. Remote Sens.* **2001**, *39*, 1943–1958. [CrossRef]
- 69. Mason, E.G.; Diepstraten, M.; Pinjuv, G.L.; Lasserre, J.-P. Comparison of direct and indirect leaf area index measurements of Pinus radiata D. Don. *Agric. For. Meteorol.* **2012**, *166–167*, 113–119. [CrossRef]
- 70. Cutini, A.; Matteucci, G.; Mugnozza, G.S. Estimation of leaf area index with the Li-Cor LAI 2000 in deciduous forests. *For. Ecol. Manag.* **1998**, *105*, 55–65. [CrossRef]
- 71. Williams, M.; Shimabukuro, Y.E.; Herbert, D.A.; Pardi Lacruz, S.; Renno, C.; Rastetter, E.B. Heterogeneity of Soils and Vegetation in an Eastern Amazonian Rain Forest: Implications for Scaling Up Biomass and Production. *Ecosystems* **2002**, *5*, 692–704. [CrossRef]
- 72. Doughty, C.E.; Goulden, M.L. Seasonal patterns of tropical forest leaf area index and CO₂ exchange. *J. Geophys. Res. Biogeosci.* 2008, 113. [CrossRef]
- 73. Biudes, M.S.; Machado, N.G.; Danelichen, V.H.d.M.; Souza, M.C.; Vourlitis, G.L.; Nogueira, J.d.S. Ground and remote sensing-based measurements of leaf area index in a transitional forest and seasonal flooded forest in Brazil. *Int. J. Biometeorol.* **2014**, *58*, 1181–1193. [CrossRef] [PubMed]
- 74. Pinto-Júnior, O.B.; Sanches, L.; de Almeida Lobo, F.; Brandão, A.A.; de Souza Nogueira, J. Leaf area index of a tropical semi-deciduous forest of the southern Amazon Basin. *Int. J. Biometeorol.* 2011, 55, 109–118. [CrossRef] [PubMed]
- Jensen, J.L.R.; Humes, K.S.; Hudak, A.T.; Vierling, L.A.; Delmelle, E. Evaluation of the MODIS LAI product using independent lidar-derived LAI: A case study in mixed conifer forest. *Remote Sens. Environ.* 2011, 115, 3625–3639. [CrossRef]
- 76. Aragao, L.E.O.C.; Shimabukuro, Y.E.; Espirito-Santo, F.D.B.; Williams, M. Spatial validation of the collection 4 MODIS LAI product in eastern Amazonia. *IEEE T Geosci. Remote Sens.* **2005**, *43*, 2526–2534. [CrossRef]
- 77. Wenze, Y.; Tan, B.; Dong, H.; Rautiainen, M.; Shabanov, N.V.; Wang, Y.; Privette, J.L.; Huemmrich, K.F.; Fensholt, R.; Sandholt, I.; et al. MODIS leaf area index products: From validation to algorithm improvement. *IEEE T Geosci. Remote* **2006**, *44*, 1885–1898. [CrossRef]
- 78. Cohen, W.B.; Maiersperger, T.K.; Turner, D.P.; Ritts, W.D.; Pflugmacher, D.; Kennedy, R.E.; Kirschbaum, A.; Running, S.W.; Costa, M.; Gower, S.T. MODIS land cover and LAI collection 4 product quality across nine sites in the western hemisphere. *IEEE T Geosci. Remote Sens.* 2006, 44, 1843–1857. [CrossRef]
- 79. Hirose, T. Development of the Monsi–Saeki Theory on Canopy Structure and Function. *Ann. Bot.-Lond.* 2005, 95, 483–494. [CrossRef] [PubMed]
- 80. Goudriaan, J. The bare bones of leaf-angle distribution in radiation models for canopy photosynthesis and energy exchange. *Agric. For. Meteorol.* **1988**, *43*, 155–169. [CrossRef]
- 81. Aragão, L.E.O.C.; Shimabukuro, Y.E.; Espírito Santo, F.D.B.; Williams, M. Landscape pattern and spatial variability of leaf area index in Eastern Amazonia. *For. Ecol. Manag.* **2005**, *211*, 240–256. [CrossRef]
- 82. PhenoCam an Ecosystem Phenology Web Camera Network. Available online: https://phenocam.sr.unh. edu/webcam/ (accessed on 8 March 2018).
- 83. Liu, J.; Pattey, E. Retrieval of leaf area index from top-of-canopy digital photography over agricultural crops. *Agric. For. Meteorol.* **2010**, *150*, 1485–1490. [CrossRef]
- 84. Weraduwage, S.M.; Chen, J.; Anozie, F.C.; Morales, A.; Weise, S.E.; Sharkey, T.D. The relationship between leaf area growth and biomass accumulation in Arabidopsis thaliana. *Front. Plant Sci.* **2015**, *6*, 167. [CrossRef] [PubMed]



© 2018 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 (http://creativecommons.org/licenses/by/4.0/).