To: Associate Editor, Jochen Schöngart

Dear Editor,

Thank you for handling this manuscript. We are pleased to see that the feedback from the reviewers was overall positive, and that they both suggested improvements to the manuscript, which we took onboard.

Both reviewers pointed out issues with the calibration of our new model (henceforth referred to as LCA model). Our methodology was not clearly stated. The LCA model was not calibrated from Lidar data but from ground data at 4 sites, and we edited the manuscript to avoid confusion about it. We developed the local models based on MCH to confirm the optimal height threshold for segmentation as indicated by Figure 2 and Figure 3. MCH-inferred AGB values are now just used as a test for validation of our height threshold in Figure 3. We also added sections comparing the LCA model to a similar model based on MCH calibrated from the same 4 sites, as suggested by the reviewers.

Comments made by both reviewers were addressed in the authors' comments as part of the interactive discussion process, and are presented here again, with additional information about changes that were made in the manuscript.

Please note that all references to changes in manuscript correspond to the line numbers of the revised manuscript with track changes.

We believe that these changes and the ones described below improved the clarity of our paper, and that it is now acceptable for publication in your journal.

Sincerely,

Victoria Meyer, on behalf of all co-authors.

Response to Anonymous Referee #1

Response to General comments:

Thank you for reviewing our manuscript. We greatly appreciate your comments and did our best to address the issues you brought up. Your comments highlighted the fact that our methodology was not clearly stated. The LCA model was calibrated using inventory data from the four sites referred to as "calibration sites" in the manuscript. Based on both reviews of the paper, we decided to remove Figure 5b and moved the paragraph explaining how AGB_{Lidar} (renamed AGB_{Local} for clarity) was calculated to the Supplementary Information (S.2), to make the paper more straight forward and focused on LCA. AGB_{Local} values are now just used as a test for validation of our height threshold in Figure 3.

Comment: "In the methods section, it is unclear whether they are predicting AGB_Lidar and AGB_LCA from an equation that already exists or whether they are doing a regression analysis to find values for parameters 'a' and 'b' in Eqs. 1-3. If it's the former, show the actual values for 'a' and 'b'."

Response: The form of Equation 1 (now Equation S4) is a commonly used model form to estimate AGB from Lidar locally (see Asner and Mascaro, 2014). For each site (or group of sites for Manaus, Tapajos and Cotriguaçu), we performed a regression based on that form and obtained coefficients a and b, presented in Table S1 (SI, ls.50-51: "All coefficients are presented in Table S1").
We decided to move this section to the Supplementary Information, as it is not central to the paper and is just used to obtain Figure 3a in this new version of the paper.
Coefficients a and b for Equation 2) and 3) (now Eq 1 and 2) are presented in Table 3. We added a sentence that makes a clear reference to the coefficients in that table. Also, we moved the section presenting the form of the LCA models from the Methods to the Results section, for clarity (ls.358-364).
Changes to manuscript: ls.334-335 "The coefficients of the models, as well as their respective coefficients of correlation, RMSE and bias from all training data and cross-validation are reported in Table 3."

Comment: "Either way, it doesn't seem necessary to predict AGB from MCH other than to compare AGB estimates from LCA to those from MCH (eg, show improvement in new method)."
Response: Based on both reviewers' comments, we removed the part of the analysis that compared AGBLCA to the locally estimated AGB_{Lidar}. As a result, Figure 5b was removed. Instead, we are now comparing AGB estimations from LCA and MCH based on the same methodology: in both cases, models were fitted using the field AGB_{inv} of the four calibration plots. This is presented in the Methods (ls.218-240), in the Results (ls. 345-379) and in the Discussion (ls.563-569).
Changes to manuscript: see ls. 218-240, ls. 345-379 and ls. 563-569. Figure 5

Comment: "In section 2.3 the authors say they have only 4 calibration sites (instead of 9 in the abstract)."
Response: We realize that the abstract was misleading. We added a sentence stating that the model was calibrated using 4 sites. We also removed the word "nine" in the title of the paper. Changes to manuscript: ls.45-46: "…and ground inventory data in nine undisturbed old growth Neotropical forests, of which four had plots large enough (1ha) to calibrate our model."

Comment: "So, is AGB in the other five sites predicted by Eq 1 (MCH)?"
Response: AGB_{LCA} in the other sites was estimated using the same LCA model calibrated from the 4 calibration sites (Eq 2). AGB_{MCH} was calculated using the MCH model presented in Table S3.

Comment: "I suggest the authors remove AGB_Lidar estimates and focus on relating LCA metrics to AGB determined from ground inventories."
Response: Thank you for your suggestion. We removed figure 5b and removed the paragraphs related to AGB_{Lidar} in Section 2.2. The information on AGB_{Lidar} (renamed as AGB_{Local}) are now provided in the Supplementary Information (S.2). AGB_{Local} is now only used to provide additional information on the choice of the height threshold in Figure 3. (nb: equation numbers have changed).
We edited the text to emphasize the role of the calibration plots and show that AGB_{Local} was just used as an additional/confirmation step.
Changes to manuscript: ls.200-206: "We determined the optimal minimum canopy height threshold calculating the coefficient of correlation between AGB_{inv} and LCA at the four calibration sites. (…).. We also estimated AGB from Lidar data locally (AGB_{Local}) using a commonly used model fit relating MCH to AGB_{inv} in each site, to further examine the variations of LCA and AGB in all nine sites (see S.2, Table S1)."

Comment: "Furthermore, I suggest trying to optimize AGB estimates from LiDAR by, for example, estimating AGB with both LCA and MCH."
Response: We tested different model forms for Equation 2 and 3 (now Equations 1 and 2), including models using both LCA and MCH as predictors. Using MCH in addition to LCA did not improve the performance of the model. This is stated in the sentence ls.234-237 "We tested different models to infer AGB_{inv} from LCA, henceforth called AGB_{LCA}, at the four calibration sites, and explored if adding more parameters, such as mean wood density of a site, mean wood density of large trees (DBH ≥50 cm), mean canopy height or top percentiles of canopy height improved the predicting power of the moded." We added:
Changes to manuscript: ls.311-331"Adding more parameters did not improve the performance of the model, except when using WD as a normalizing factor. The two models we retained are therefore of the form of Eq. (1) and Eq. (2)"

Responses to specific comments:
Comment: How is the LCA method weighted by WD if there isn't ground data at 5 sites?
Response: Ground data are available in all sites except Cotriguaçu, but plot size was too small to be used in the LCA model calibration process. However, wood density estimation does not depend on plot size, and wood density information was used from all sites to obtain a site-averaged wood density (see Table 1). A sentence was added to highlight this point:

Changes to manuscript: ls.138-145: “For this reason, all plots smaller than 1 ha were excluded from the LCA analysis but were used in estimating average wood density for each site, which does not depend on plot size. Stand averaged wood density was calculated based on the wood density of all trees present in a site, determined using the commonly used global wood density database, and is reported in Table 1 (Chave et al., 2009; Zanne et al., 2009). For Cotriguaçu, we used stand averaged wood density given by Fearnside, (1997) for a region covering the site.”

Comment: Line104: what do you mean by ‘unique’?
Response: by “unique”, we mean one model that would work across sites in the Neotropics.
Changes to manuscript: l.112: We modified the sentence accordingly to “single”.

Comment: Line 166: What model? Line 167: what data?
Response: The text was edited to clarify this sentence.
Changes to manuscript: SI, ls.55-58 “For the remaining sites of the Central Amazon (Cotriguaçu, Manaus and Tapajós), we developed a model based on existing data in Manaus and Tapajós from a previous study, derived from airborne and spaceborne Lidar (see Lefsky et al., 2007).” Note that this section is now part of the Supplementary Information, as explained above.

Comment: Lines 203-4: This indicates that AGB_LCA is being tested against AGB_Lidar, where LiDAR is being treated as the reference. AGB_Lidar is only an estimate.
Response: This is correct. The goal here was to test AGB_{LCA} against locally derived AGB_{Lidar}. Based on both reviewers’ comments, we realized that this step was not necessary and was removed from the paper.
Changes to manuscript: Figure 5b and any text related to this graph were removed from the paper.

Comment: Lines 205-6: Here you say that these results were compared to ‘a traditional model relying on MCH to estimate AGB’. Isn’t AGB_Lidar the model relying on MCH to estimate AGB?
Response: Thank you for highlighting this point. Here, we refer to a single model based on MCH from all the calibration sites, the same way that the LCA model was calibrated. This way, we can compare the LCA model to a MCH model. We realize that this sentence is confusing and edited the manuscript to clarify it: as stated above, AGB_{Lidar} is now only used to obtain Figure 3b and is no longer compared to AGB_{LCA}. Instead, we added a new section in the methods, results and discussions comparing AGB_{LCA} and AGB_{MCH} (based on a model calibrated on the same 4 calibration sites). Please report to our response to earlier comment.

Comment: Section 2.5: Is it possible to apply the same methods to logged areas, since you may not know which areas have been harvested or not – or have before and after pictures?
Response: We agree that we need before-after data to detect logging. In the example we are showing, we do have before and after logging Lidar data. Details are provided in Anderson et al., 2014.
We added a sentence to emphasize on the need for this type of datasets.
Changes to manuscript: ls.246-247: “provided that Lidar data are available from pre and post-logging.”.

Comment: Line 269: Where did wood volume data come from?
Response: we edited the manuscript to clarify this point:
Changes to manuscript: ls.307-309: "Since AGB depends on DBH, H and WD (see Chave et al., 2014), average wood volume can be computed approximately as the ratio of AGB divided by the average wood density".

Comment: Lines 315-6: In what way does Antimary not represent Peruvian Amazon and Amazon-Andes gradients?
Response: We added the following sentence to be more specific :
Changes to manuscript: ls.418-421: "However, this site does not represent forests in the western Amazon or the Amazon-Andes gradients with relatively lower wood density (Baker et al. 2004) and more fertile volcanic soils impacting the forest structure and dynamics (Quesada et al., 2011)."

Comment: Line 323: by how much does it explain the variation?
Response: Overall 78% is explained (R2=0.78).
Changes to manuscript: l.428: "and explained 78% of the variation".

Comment: Section 4.3: Would be helpful to refer to tables and figures
Response: Thank you for the suggestion. We added references to table 2 and figure 3.
Changes to manuscript: references l.465 and l.468.

Comment: Lines 344-6: This sentence is unclear to me, but it sounds like it supports my point that using AGB_Lidar as a reference is circular and not proving anything
Response: This sentence was not clear and was removed from the manuscript. Moreover, we are now comparing AGB from LCA and MCH in a separate section of the results and discussion to avoid any confusion.

Comment: Line 374: Change 'only' to 'primarily' or something similar.
Response: "only" was removed.

Comment: Line 391: Change 'Any' to 'Most'
Response: We changed "'Any' to 'Most'.

Comment: Lines 423-5: Maybe the relationship is not linear at the high end of LCA
Response: It is indeed a possibility. We added this suggestion to the manuscript.
Changes to manuscript: ls. 589-591: "It is also possible that the relationship between AGB and LCA is not linear for very high AGB values. This could be tested in the future with a larger number of sites with very high biomass."

Comment: Line 467: If the relationship remains unique across forest types, is it not then broadly applicable?
Response: Yes, this is an important point of the paper. We added two sentences highlighting this fact.
Changes to manuscript:
- in the Discussion:

ls.538-539: “Our model can therefore potentially be applied to a wide range of forest types, provided that there is information about wood density of the study area in the literature.”
- in the Conclusion:
ls.640-641: “. This linear relationship remains unique across different forest types, making the LCA model broadly applicable.”

Comment: Fig 3: Clever way to find the optimal H threshold
Response: Thank you for this positive comment.

Comment: Fig 4b: This doesn’t look like a perfectly fit.
Response: With a R2 of 0.78, RMSE of 46 and no bias, we consider the fit to be good. These number are provided in Table 3. R2 was added to Figure 4b to emphasize this point.
Changes to manuscript: R2 was added to Figure 4b to emphasize this point.

Comment: Fig 5b: All calibration sites are above the 1:1 line. Why are Nouragues and Choco below the line?
Response: Based on your comments and that of Reviewer 2, we removed this figure. The fact that some plots were above/below the line was likely due to the fact that AGB_{Lidar} was estimated locally for different sites and included some error. We are now simply comparing the LCA and MCH methods based on the inventory data only (Figure 5, attached here as Fig.1).

Comment: Fig 7: It would be helpful to see the actual data, not just regression lines.
Response: The point of this figure is to clearly see where the lines cross the y axis. For Fig 7a), we are just showing where the LCA model crosses the y axis, with different wood density from the different sites. Each line represents the model curve with various wood density values. To see the actual data from the calibration sites, see Figure 4b.
For fig 7b, actual data could be added, but just showing the lines gives the figure a clean look, considering that the information we are looking for here is the intercept of each line.

Response to Anonymous Referee #2

Thank you for taking the time to review our paper. We did our best to address all your comments in the hope this will improve the quality of the manuscript.

Comment: For this method to be useful, it must either (1) outperform existing methods, (2) perform similarly to existing methods but at lower computational cost or (3) open up new applications not allowed by existing methods.

Response : Our study does open up new applications compared with existing methods. We demonstrate that our method performs similarly to another method relying on information from all trees within a plot (MCH). The point of our paper is not to say that the LCA method is better than the MCH method, but rather to show that information on large trees is enough to estimate biomass. Our findings confirm what has been shown in several studies focusing on ground data (Bastin et al, Slik et al…) and shows for the first time that relying on large trees from a remote sensing perspective allows to estimate AGB. It opens up new applications both for field inventory and remote sensing applications. In the discussion (section 4.8), we talk about how methods focusing on large trees could help future space missions, such as BIOMASS and GEDI, to accurately estimate biomass and open up new applications. LCA also gives information on the presence of large trees in a study area, which other metrics such as MCH cannot do. It is an important point, considering that large trees are often the most affected by natural disturbance and targeted by logging companies.

Changes to manuscript: ls.455-457: "LCA provides information on the presence of large trees in a study area, which other metrics such as MCH cannot do. It is an important point, considering that large trees are often the most affected by natural disturbance and targeted by logging companies."

ls.564-565: "The comparison of LCA and MCH metrics showed that both performed similarly in estimating AGB, highlighting the importance of large canopy trees to estimate biomass."

ls.645-647: "The results of our study may encourage further research in the use of Lidar data for detecting the distribution of larger trees in tropical forests for ecological and conservation studies."

Comment: The paper is framed around comparing the new LCA method against the existing MCH method, but a clear comparison of the two against ground-based validation data is not presented.

Response: Thank you for pointing this out. We added a short paragraph in the method section, as well as a new section in the Results and in the Discussion, comparing the performance of LCA and MCH methods. This is presented in the Methods (ls.218-240), in the Results (ls. 345-379) and in the Discussion (ls.563-569).

To avoid any confusion, we moved the MCH local estimations of AGB from the main Lidar data paragraph to the Supplementary information (S.2). AGB_{Lidar} was also renamed LCA_{Local} for clarity.

Changes to manuscript: see ls. 218-240, ls. 345-379 and ls. 563-569. Figure 5 (attached here as Fig. 1)

We chose to keep Table S3 in the Supplementary Information for clarity, but we added a figure comparing AGB estimations using the 2 methods (Figure 5).

Comment: Is LCA quicker to calculate than MCH? It would be useful to present a comparison of the computational time taken to calculate LCA versus MCH.

Response: LCA is not quicker to calculate than MCH, but it is not significantly slower either (below 1s for both methods). Also, the strength of LCA lies in the structural information it provides, not in its computational time. Thus, we chose not to add a detailed comparison of computational time.

Comment: The application to detect the impacts of selective logging is potentially very important.

Response: We agree. We emphasized this point in the Discussion:

Changes to manuscript: ls.609-611: "LCA could become an important tool to detect forest degradation, in particular selective logging, considering that large trees are targeted by logging companies."

Comment: My main suggestion to improve this paper are to concentrate on testing the relative performance of LCA and MCH approaches at estimating biomass when validated against inventory data (even if LCA performs worse, this is still a very useful result for method development),

Response: Thank you for your suggestion. As mentioned above, we added a paragraph in the method section, as well as two new sections (results and discussion) and a figure comparing the two methods, showing that they perform very similarly. We also show how they differ in terms of AGB estimations in different sites.

Comment: and comparing the performance of the two approaches when applied to detect the impacts of selective logging.

Response: We compared the performance of the 2 approaches when applied to selective logging detection. The MCH model showed a loss of biomass of 19 Mg ha^{-1}, compared to 15 with LCA and 9 from a previous study based on rh25. We added this information in the results and the discussion.
Changes to manuscript: ls.393-394: "As a comparison, the MCH model led to an estimated biomass loss of 19 Mg ha^{-1}."
ls.607-609: "The higher biomass loss estimation from the MCH model (19 Mg ha^{-1}) again shows how different metrics can lead to different results. Here, three methods based on three different Lidar metrics yielded results that differed by more than twofold.".

Comment: I agree with reviewer 1 in that I don't see much value in testing the performance of LCA against biomass estimates using MCH.

Response: Thank you for your suggestion. We removed Figure 5b. Performance comparison of LCA and MCH model at the calibration sites is now based on Figure 5a. The models applied to the nine sites are now Figure 5b, following your other suggestion to focus on the comparison of LCA and MCH methods.

Specific comments:

Comment: Line 205 – How was bias calculated?

Response: We added the definition of bias to the manuscript:
Changes to manuscript: ls.214-215: "bias (mean difference between the expected values of AGB and the observed values of AGB)".

Comment: Line 262 – What are the other models apart from a power law fit?

Response: For both LCA and MCH models, we tested linear models and power laws, which are

the 2 common fits. We modified the sentence to avoid any confusion:
Changes to manuscript: ls.302-303: "with a better coefficient of correlation and RMSE than a power law fit"

Comment: Line 262 – 263 – Are RMSE values and r squared values here from cross-validation or from the training data? Line 263 – Just present the bias from cross-validation.

Response: R^2 and RMSE are from training data.
We removed the bias from the training data and present the bias from cross-validation.
Changes to manuscript: l.304: "$bias_{cross\_val} = 0.16$ Mg"
ls.334-336: **"**coefficients of correlation, RMSE and bias from training data and cross-validation are reported in Table 3."

Comment: Line 271 – How feasible is it to scale by wood density in the absence of inventory data? Presumably errors would be larger if modelled estimates of wood density were used.

Response: We agree. If there is no information in the literature from previous studies, modelled WD could be used, but would indeed give greater errors. This is now covered in the Discussion.
Changes to manuscript: ls.558-561: "In the absence of information on wood density from the literature, modelled wood density could potentially be used, but would give greater errors. These errors should be taken into account when reporting on the uncertainty of the results."

Comment: Lines 287-301 – It would be useful to also see how MCH performs at detecting this loss of biomass.

Response: The MCH model (Table S3) gives a biomass loss of 19mg/ha, more than twice what was reported in Andersen et al., 2014. These results were added to the results section and the discussion section 4.6.:
Changes to manuscript: ls.393-394: "As a comparison, the MCH model led to an estimated biomass loss of 19 Mg ha^{-1}."
ls.607-609: "The higher biomass loss estimation from the MCH model (19 Mg ha^{-1}) again shows how different metrics can lead to different results. Here, three methods based on three different Lidar metrics yielded results that differed by more than twofold.".

Comment: Lines 376-377 – This is a very nice approach to identify how much biomass is missed by LCA.

Response: Thank you for this positive comment.

Comment: Figure S2 - Given that the minimum cluster size didn't have a major effect on the AGB estimates, I would be interested in seeing a comparison of the performance of the LCA metric just following masking versus the LCA metric following removal of segments below the threshold cluster size. How computationally costly are these last steps?

Response: This is a good point. For a reference image of 1000x1000m pixels, the full process takes less than one second. Just using masking may be slightly faster, but the computational cost

is not an issue here. Just using masking gives similar results as when using LCA, because the pixels removed by the full process represent a small fraction of the area covered by large trees (1.73% on average). (R2=0.78, RMSE=45.7, bias=0.55)
These isolated pixels either represent single branches reaching above 27m or the tip of a tree whose crown is mainly below 27m. Therefore, these pixels have no meaning in terms of our LCA metric and do not represent large trees. This is why we chose to remove them. The goal of our study is to show that large trees are sufficient to estimate AGB. We clarified this point in the manuscript:
Changes to manuscript: ls.450-454: “Clusters smaller than 100 m^2 add only a small fraction (1.7% on average) to LCA values across sites. Including these clusters in LCA would not impact the performance of the model (similar R^2, RMSE and bias) and would allow to skip the final steps of the LCA retrieval (see Fig. S2). However, since these pixels either represent single branches reaching above 27m or the tip of a tree crown, they have no meaning in terms of our LCA metric and do not represent large trees.”.

Comment: Technical comments: Inconsistent approach to using capitals in section headings. Line 209 – => Detecting changes of selective logging. Line 385 - => LCA as an AGB estimator

Response: Thank you for pointing this out. We removed the capital letters accordingly.

Additional changes

We made some additional minor edits to the paper to clarify some sentences. Please refer to the track changes of the revised manuscript, notably:

- Paragraph ls.485-503.
- Figure 6: “2012” was replaced by “2011”.
- The word “nine” was removed from the title to be more consistent with the content of the manuscript.

Canopy Area of Large Trees Explains Aboveground Biomass Variations across Neotropical Forest Landscapes

Deleted: Nine

Victoria Meyer[1,2], Sassan Saatchi[1], David B. Clark[3], Michael Keller[4,5], Grégoire Vincent[6], António Ferraz[1], Fernando Espírito-Santo[1,7], Marcus V.N. d'Oliveira[5], Dahlia Kaki[1] and Jérôme Chave[2]

[1] *Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA. USA*
[2] *Laboratoire Evolution et Diversité Biologique UMR 5174, CNRS Université Paul Sabatier, Toulouse, France*
[3] *Department of Biology, University of Missouri, St. Louis, Missouri, U.S.A.*
[4] *USDA Forest Service, International Institute of Tropical Forestry, San Juan, Puerto Rico*
[5] *EMBRAPA Acre, Rio Branco, Brazil*
[6] *IRD, UMR AMAP, Montpellier, 34000 France*
[7] *Lancaster Environmental Centre, Lancaster University, Lancaster, United Kingdom, LA1 4YQ*

Correspondence to:

Victoria Meyer
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA. 91109 USA
Email: victoria.meyer@jpl.nasa.com

Abstract

Large tropical trees store significant amounts of carbon in woody components and their distribution plays an important role in forest carbon stocks and dynamics. Here, we explore the properties of a new Lidar derived index, large tree canopy area (LCA) defined as the area occupied by canopy above a reference height. We hypothesize that this simple measure of forest structure representing the crown area of large canopy trees could consistently explain the landscape variations of forest volume and aboveground biomass (AGB) across a range of climate and edaphic conditions. To test this hypothesis, we assembled a unique dataset of high-resolution airborne Light Detection and Ranging (Lidar) and ground inventory data in nine undisturbed old growth Neotropical forests, of which four had plots large enough (1ha) to calibrate our model. We found that the LCA for trees greater than 27 m (~25–30 m) in height and at least 100 m^2 crown size in a unit area (1 ha), explains more than 75 % of total forest volume variations, irrespective of the forest biogeographic conditions. When weighted by average wood density of the stand, LCA can be used as an unbiased estimator of AGB across sites (R^2 = 0.78, RMSE = 46.02 Mg ha^{-1}, bias = -0.63 Mg ha^{-1}). Unlike other Lidar derived metrics with complex nonlinear relations to biomass, the relationship between LCA and AGB is linear and remains unique across forest types. A comparison with tree inventories across the study sites indicates that LCA correlates best with the crown area (or basal area) of trees with diameter greater than 50 cm. The spatial invariance of the LCA–AGB relationship across the Neotropics suggests a remarkable regularity of forest structure across the landscape and a new technique for systematic monitoring of large trees for their contribution to AGB and changes associated with selective logging, tree mortality, and other types of tropical forest disturbance and dynamics.

Deleted: 76

Deleted: . Results show that this

Deleted: relationship remains

Deleted: >

Keywords

Lidar, biomass, tropical forest, large trees, crown area, wood density

1 Introduction

In humid tropical forests, tree canopies contribute disproportionately to the exchange of water and carbon with the atmosphere through photosynthesis (Goldstein et al., 1998; Santiago et al., 2004). From a physical standpoint, canopies are rough interfaces formed by crowns of emergent and large trees, regularly disturbed by wind thrusts and gap dynamics. This structurally complex boundary layer is challenging for scaling of biogeochemical fluxes and modeling of vegetation dynamics (Baldocchi et al., 2003). Large canopy trees are among the first to be impacted by storms or heavy precipitation (Espírito-Santo et al., 2010), drought stress (Nepstad et al., 2007; Saatchi et al., 2013; Phillips et al., 2009), and fragmentation (Laurance et al., 2000), potentially leading to tree death and formation of large canopy gaps (Denslow, 1980; Espírito-Santo et al., 2014). Several studies suggest that forest canopies can show fractal properties that tend to evolve from a non-equilibrium state towards a self-organized critical state, involving gap formation and recovery (Pascual and Guichard, 2005; Solé and Manrubia, 1995), with crowns preferentially growing towards more sunlit parts of the canopy (Strigul et al., 2008).

Over the past decade, stand level canopy metrics have been increasingly derived using small footprint airborne Lidar systems (ALS), a widely used remote sensing technique to study the structure of forests (Kellner and Asner, 2009; Lefsky et al., 2002). Lidar derived mean top canopy height (MCH) is a good predictor of tropical forest aboveground carbon content and its spatial variability (Jubanski et al., 2013), but it does not provide information on the presence of large trees that are important when monitoring changes of forest biomass from logging and other small scale disturbance (Bastin et al., 2015). Moreover, different forests with the same MCH may differ in their stem density, notably of large trees, and in stand mean wood density, two aspects that are important in constructing a robust model to infer AGB from lidar data (Asner et al., 2012; Mascaro et al., 2011). Ground observations suggest that stem density, basal area, height and crown size of large tropical trees may all be good indicators of forest AGB (Clark and Clark, 1996; Goodman et al., 2014). This implies that including information on crown area of individual large trees should improve carbon stock assessments, as confirmed in temperate and boreal regions (eg. Packalen et al., 2015; Popescu et al., 2003; Vauhkonen et al., 2011, 2014). In tropical forests, identifying and delineating crowns of large trees is a difficult and time consuming process due to the layered structure of the forest canopy and overlapping crowns (Zhou et al., 2010, but see Ferraz et al., 2016).

Here, we explore how the fractional area occupied by crowns of large trees in a forest stand can be used as a reliable indicator of forest biomass across a wide range of forest structure, climate and edaphic geographic variations. We define large tree canopy area (LCA) as a metric capturing the cluster of crowns of large trees within a forest patch using height and crown area measured by high resolution airborne Lidar measurements. Precisely, LCA is the number of pixels in the canopy height model above a reference height, and excluding the pixel clusters smaller than a reference area. Since this metric quantifies the proportional presence of large trees, it can be used to estimate AGB and monitor changes associated with the disturbance of large trees from mortality events and selective logging. We first explore the properties of LCA across a range of landscapes in the Neotropics. Next, we hypothesize that LCA is a good predictive metric of the spatial variations of AGB over a wide range of old growth forests.

To this end, we assembled a collection of airborne Lidar measurements and ground inventory data at nine sites in old growth Neotropical forests. The Lidar data provide variations in canopy height and distribution of large trees that allow us to address the following questions: 1) is there a single definition of LCA at the landscape scale across different sites? 2) does LCA metric capture variations of AGB?

Deleted: unique

2 Materials and Methods

2.1 Study sites

We studied the canopy structure at nine old growth lowland Neotropical forest sites that span a broad range of climatic and edaphic conditions (Fig. S1, Table 1). All sites are located in low elevation areas (less than 500 m above sea level) but have small scale surface topography that may influence the distribution of crown formations and gaps. These forests are for the most part undisturbed *terra firme* forests. Tapajós, Antimary and Cotriguaçu get the least rainfall, with approximately 2000mm yr^{-1}, while La Selva and Chocó both receive more than 4000 mm yr^{-1} (Table 1).

Permanent forest inventory plots were available for all sites except Cotriguaçu (Table 1). Sites where tree level inventory data were available were used to estimate the stand level aboveground biomass, thereafter referred to as AGB_{inv}: BCI (50 plots of 1 ha each), Chocó (42 plots of 0.25 ha each), La Selva (11 plots of 1 ha each), Manaus (10 plots of 0.25 ha each), Nouragues (7 plots of 1 ha each) and Tapajós (10 plots of 0.25 ha each). In these plots, all trees with a diameter at breast height (DBH) ≥10 cm have been mapped, measured and identified to the species. Trees with irregularities or buttresses were measured higher on the bole. Total tree height measurements were available for a subset of these trees. The method for calculating AGB_{inv} from forest inventories is reported in S.1 of the supplementary information. Four sites (BCI, La Selva, Nouragues and Paracou) with 1 ha inventory plots, were used as “calibration sites” to compare the LCA metric and AGB. Sites with smaller plots were not used as calibration of LCA because of the probability of crowns of large trees extending outside the plot boundary and the introduction of uncertainty in estimating LCA from edge effects (Meyer et al., 2013; Packalen et al., 2015). For this reason, all plots smaller than 1 ha were excluded from the LCA analysis but were used in estimating average wood density for each site, which does not depend on plot size. Stand averaged wood density was calculated based on the wood density of all trees present in a site, determined using the commonly used global wood density database, and is reported in Table 1 (Chave et al., 2009; Zanne et al., 2009). For Cotriguaçu, we used stand averaged wood density given by Fearnside, (1997) for a region covering the site. Additional plot level data (AGB_{inv} and mean wood density) were provided for Antimary (50 plots of 0.25 ha each), Nouragues (27 plots of 1 ha each) and Paracou (85 plots of 1 ha each).

2.2 Lidar data

Lidar sensors scan the vegetation vertical structure and return a three dimensional point cloud derived from the time it took each pulse to return to the instrument. The Lidar datasets acquired over the study sites come from discrete return Lidar instruments and were gridded horizontally at a 1m resolution using the echoes classified as either vegetation or ground. They yield three products: digital surface model (DSM) corresponding to the top canopy elevation, digital terrain model (DTM) corresponding to the ground elevation, and canopy height model (CHM), which is the height difference between the DSM and the DTM. DTMs were interpolated from a Delaunay

Moved (insertion) [1]

Deleted: were available were used

Deleted: , and are here referred to as “calibration sites” (BCI, La Selva, Nouragues and Paracou). Smaller plots have a higher probability of having the

Deleted: e

Deleted: , which can introduce

Deleted: because of edge effect

Moved up [1]: The four sites where 1 ha plots were available were used to compare the LCA metric and AGB, and are here referred to as “calibration sites” (BCI, La Selva, Nouragues and Paracou). Smaller plots have a higher probability of having the crown of large trees extend outside the plot boundary, which can introduce uncertainty in estimates of LCA because of edge effect (Meyer et al., 2013; Packalen et al., 2015). For this reason, all plots smaller than 1 ha were excluded from this analysis.

triangulation or comparable interpolation methods, after outliers have been removed. DSMs were created using the highest return within a cell. Lidar data over Paracou were acquired in last return mode, causing a bias of 50 cm on the CHM (Vincent et al., 2012). This bias is not addressed in this study because our height increment for the determination of optimal height thresholding is larger (1m) (see Sect. 4.3). Data were acquired between 2009 and 2013, using relatively similar sensors and acquisition configurations (Table 2). The potential differences between the Lidar datasets and their impact on the results are addressed in the Discussion.

For each site, we selected a 1x1 km (100 ha) area of old growth forest, oriented north-south, without any human disturbance to the extent possible. Topography derived from Lidar data within the selected 1 km^2 subset images provides information on landscape variations that may impact the forest structure. Data visualization was done using ENVI version 4.8 (Exelis).

Deleted: ... [1]

2.3 Computing Large Canopy Area (LCA)

At each study site, we extracted the area of canopy that relates to total area of the canopy height model above a standard height (h) threshold, or LCA(h), and explored how this metric scales along two axes. First, we varied the threshold height h with increments of 1m, between 5m and 50m, in 100 m by 100 m subareas (100 subareas for each site). Second, to denoise the data, we excluded the clusters with less than a set number of 1m^2 pixels (50, 100, 150 or 200). We then prioritized the crown area of large trees, and filtered out pixels that could be related to outliers or to single branches. This method thus quantifies the area of large crowns covering a plot or larger landscape unit area, as a percentage of covered area.

LCA maps were produced at 1 ha resolution. Pixel clustering was based on the similarity of the four nearest neighbors (similar results were obtained with an eight neighbor model, results not shown here). Figure S2 summarizes the steps taken to go from the Lidar canopy height model to the final LCA map. Processing was conducted using the IDL software (Interface Description Language, Exelis).

We determined the optimal minimum canopy height threshold calculating the coefficient of correlation between AGB_{inv} and LCA at the four calibration sites. This step allowed us to examine if optimal height thresholds differed from one site to the other. The goal was to find a single optimal height threshold and crown size that could be applied for LCA retrieval across closed canopy Neotropical forests. We also estimated AGB from Lidar data locally (AGB_{Local}) using a commonly used model fit relating MCH to AGB_{inv} in each site, to further examine the variations of LCA and AGB in all nine sites (see S.2, Table S1).

2.4 Relating LCA to biomass

We tested different models to infer AGB_{inv} from LCA, henceforth called AGB_{LCA}, at the four calibration sites, and explored if adding more parameters, such as mean wood density of a site, mean wood density of large trees (DBH ≥50 cm), mean canopy height or top percentiles of canopy height improved the predicting power of the model. We evaluated our results by applying a jackknife validation to our regression models, based on 1000 iterations of bootstrapping. The coefficients of correlation (R^2), root mean square error (RMSE) and bias (mean difference between the expected values of AGB and the observed values of AGB) are reported for the models providing the best results. The analysis was performed using the R statistical software (R Core Team, 2014).

We compared the new approach based on LCA to a similar approach based on MCH, which relies on information on all pixels of an area of interest. In both cases, models were calibrated by

Deleted: AGB_{Lidar} and LCA

Deleted: We also performed the same analysis using AGB_{inv} and LCA at the four calibration sites.

Deleted: Adding more parameters did not improve the performance of the model, except when using WD as a normalizing factor. The two models we retained are therefore of the formThe two models we retained are of the form of Eq. (12) and Eq. (23): ... [2]

Deleted: We also compared AGB as derived from LCA (AGB_{LCA}) to the Lidar derived aboveground biomass (AGB_{Lidar}) in the nine 1km^2 images. The coefficients of models for equation 1 and 2, as well as cCoefficients of correlation (R^2), root mean square error (RMSE) and bias (mean difference between the expected values of AGB and the observed values of AGB) are reported in Table 3.. We finally compared these results to a traditional model relying on MCH to estimate AGB.

using field data from the four calibration sites and their respective mean wood density. This comparison is meant to investigate if a metric based on large trees only (LCA) can estimate AGB similarly to a metric that uses information about 100% of the canopy (MCH).

2.5 Detecting changes of selecting logging

Deleted: 5
Deleted: C
Deleted: S
Deleted: L

Forest degradation due to selective logging is difficult to detect with conventional remote sensing techniques due to small scale and minor impacts on the forest canopy and biomass compared to severe forest disturbances (e.g. fires, storms, or clearing). However, selective logging targets large trees (Pearson et al., 2014) and thus may be detectable using LCA, provided that lidar data are available from pre and post-logging. Here, we use the Antimary study site that was selectively logged after the 2010 Lidar acquisition to examine the use of LCA for detecting logging impacts on the forest canopy and AGB. We apply the large tree segmentation approach on both the 2010 image and on a 2011 post-logging Lidar image (see Andersen et al., 2014 for details) to quantify the logging impacts in terms of the distribution of large trees removed from the forest and the loss of aboveground biomass.

Deleted: data

3 Results

3.1 Intersite comparison of landscapes and MCH

Topographic variation within the 1 km^2 images ranged from about 4 m elevation gain in flat area of Tapajós to steep elevation gain of up to about 100 m in Cotriguaçu and Chocó (Fig. S3). Top canopy height reached up to 60m, but varies across sites, with Chocó having the lowest MCH (24.1 m) and Nouragues the highest (29.7 m). Forest height in Manaus was more homogeneous than in the other sites, with a standard deviation of 6.8 m for MCH, versus 10.3 m in Paracou.

We found no relationship between topography and canopy height, which suggests that variability in forest structure may be due to other ecological and edaphic factors in each site.

3.2 Large canopy area index

The choice of the canopy height threshold impacted LCA more than the minimum number of pixels per cluster (Table S2). The difference due to the choice of the minimal cluster size threshold was on average 1.4 %, calculated as the mean of the difference between the smallest grain (50 pixels) and the largest one (200 pixels) across sites and height thresholds. Based on this analysis, we chose to define LCA using a minimum cluster size of 100 pixels (100 m^2 for crown area) in the remainder of this study. This corresponds to an area of at least 10 m x10 m or a circle of approximately 11m in diameter, consistent with the average crown diameter of large trees of the region (Bohlman and O'Brien, 2006; Figueiredo et al., 2016; Clark, unpublished results).

In contrast, the canopy height thresholds markedly impacted the magnitude of LCA among sites (Fig. 1 and Fig. 2, Table S2). As the height threshold increased, intra-site variation of LCA(h) became apparent, showing differences of LCA associated with differences of forest structure (Fig. 1). Tapajós and Nouragues stood out with more area of large trees at the height threshold of 30 m (LCA_{30m} = 51 and 48 %, respectively) , while Antimary and Chocó showed much lower LCA at this height threshold (LCA_{30m} = 21 %) (Table S2). The steepest slopes of the LCA(h) function corresponded to the highest sensitivity of LCA to height thresholds and the inflection in LCA was found between 24m in Antimary and 30m in Nouragues (Fig. 2). The average height of the steepest slope was about 27 m, a value that was used as the optimal threshold across all sites.

Regressing AGB_{inv} and LCA at the calibration sites (Fig. 3b) showed the best relationships corresponded to height thresholds between 27m (Nouragues and Paracou) and 28m (BCI and La Selva), with maximum coefficients of correlation ranging between 0.5 and 0.8. The same analysis repeated using AGB_{Local} and LCA in the nine sites also confirmed the earlier results that the highest coefficients of correlation between the two metrics occurred between 23 m (Chocó) and 30 m (Tapajós) height thresholds (Fig. 3a), explaining more than 75 % of AGB variation in each site. Based on these results, we defined LCA as the cumulative area of clusters of the canopy height model greater than 27 m height, as the mean of optimal height threshold with highest R^2 across sites, with clusters covering areas larger than 100 m^2.

3.3 Variation of AGB derived from LCA

AGB_{inv} was found to depend linearly on LCA (Eq. 1), with a better coefficient of correlation and RMSE than a power law fit ($R^2_{linear} = 0.59$, $RMSE_{linear} = 62.53$ Mg ha^{-1}, vs. $R^2_{power} = 0.54$, $RMSE_{power} = 65.38$). Although this model was unbiased ($bias_{cross\_val} = 0.16$ Mg), there were clear differences among study sites (Fig. 4a, Table 3). These differences were largely explained by landscape scale differences in wood density, an important factor representing the influence of species composition on the spatial variation of AGB. Since AGB depends on DBH, H and WD (see Chave et al., 2014), average wood volume can be computed approximately as the ratio of AGB divided by the average wood density (Fig. 4b). The linear relationship between LCA and wood volume yielded an estimate of the average total volume of forests independently of the site characteristics, through Vol = a LCA + b. Adding more parameters did not improve the

Deleted: Regressing AGB_{Lidar} and LCA showed that the highest coefficients of correlation between the two metrics occurred between 23 m (Chocó) and 30 m (Tapajós) height thresholds (Fig. 3a), explaining more than 75 % of AGB variation in each site. The same analysis repeated using

Deleted: also confirmed the earlier results

Deleted: ing

Deleted: are found to be

Deleted: 2

Deleted: other models, such as

Deleted: bias = 0.0 Mg,

Deleted: 05

Deleted: To explore the contribution of wood density across the study sites,

Deleted: we

Deleted: the average wood volume

Deleted: as

Deleted: (Table 3)

performance of the model, except when using WD as a normalizing factor. The two models we retained are therefore of the form of Eq. (1) and Eq. (2):

$$AGB_{LCA} = a\,LCA + b \qquad (1)$$

$$AGB_{LCA} = (a\,LCA + b) \times WD \qquad (2)$$

where here WD is the mean wood density of a site. The coefficients of the models, as well as their respective coefficients of correlation, RMSE and bias from training data and cross-validation are reported in Table 3.

For AGB estimation, the model based on LCA weighted by WD gives the best result by bringing R^2 up to 0.78 and RMSE down to 46.02 Mg ha^{-1} (Fig. 4b, Fig. 4c, Table 3, Eq. (2)), with AGB_{inv} and AGB_{LCA} falling around a one-to-one line in Fig. 4c. At all sites, RMSE values are between 20.87 and 42.22 Mg, except Nouragues, where RMSE remains large (71.21 Mg) due to high biomass and several outliers from the linear relation. The relationship between LCA and other metrics derived from ground data, such as Lorey's height or basal area, are presented in S.3 and Table S4.

3.4 LCA vs. MCH approach

Finally, we compared these results to AGB estimated using a similar approach based on MCH (AGB_{MCH}) for the calibration plots (Fig. 5a), and we also compared AGB_{LCA} to AGB_{MCH} in all nine sites, using LCA and MCH of the 1km^2 images (Fig. 5b). Both methods perform similarly (R^2_{MCH} = 0.80, $RMSE_{MCH}$ = 42.52 Mg ha^{-1}, $bias_{cross\_val}$=-0.21 Mg ha^{-1}, Table S3), showing that relying on a fraction of the Lidar information performs as well as using a metric depending on information from all pixels. However, Fig. 5 also shows that the

Deleted: ,

Deleted: 5

Deleted: 3

Deleted: 5a

Deleted:

Deleted: b

Deleted: ,

Deleted: applied the model from Eq. (3) to all 1km^2 areas and

Deleted: the derived

Deleted: AGBLidar

Deleted: (see Sect. 2.2)

Deleted: , for which local models based on MCH were used

Deleted: cb

Deleted: . Global RMSE was found to be 34.72 Mg and RMSE per site varied between 20.79 Mg at BCI and 49.58 Mg at Manaus

Deleted: Our ground calibrated LCA model defined by Eq. (3) had a similar performance as the MCH based AGB model (R^2_{MCH} = 0.79, $RMSE_{MCH}$ = 44.2 Mg, Table S3). These findings

Deleted: gives comparable results

Deleted: s

Deleted: , such as MCH

Deleted: ure 5b and 5c

LCA method tends to overestimate AGB compared to the MCH method (bias=9.66 Mg ha^{-1}), especially in La Selva, BCI, Cotriguaçu and Manaus.

3.5 AGB changes from logging

The impacts of logging on the distribution of large trees and changes of AGB was detected by simply deriving the LCA index from pre and post-logging Lidar data acquired in 2010 and 2011 respectively in Antimary (Fig. 6). Difference in LCA between the two dates (2010–2011) (Fig. 6a) at 1 ha grid cell captured the areas of largest changes in the few months following logging (logging took place between June and November 2011, Lidar data were collected in late November 2011). The LCA approach was able to detect approximately a 17 % decrease in LCA, from a mean LCA of 34.8 % in 2010 to 29.2 % in 2011.

The changes were also captured in the frequency distribution of large canopy trees before and after logging (Fig. 6b) and the differences in the spatial distribution (Fig. 6c and 6d).

These changes in LCA correspond to a biomass loss of 15.2 Mg ha^{-1} when integrated in equation (2) and were of the same magnitude of the planned selectively logging removal rate (12–18 Mg ha^{-1} or 10–15 m^3 ha^{-1} of timber volume) (Andersen et al., 2014). As a comparison, the MCH model led to an estimated biomass loss of 19 Mg ha^{-1}. Difference in the Lidar index (ΔLCA) at the native resolution of 1 m (Fig. 6e) was able to capture both the location of all large trees removed from the forest stand and partial regeneration and gap filling that occurred in the forest between the two dates.

4 Discussion

4.1 Inter-site Comparisons

Deleted: , highlighting the importance of large canopy trees to estimate biomass. The relationship between LCA and other metrics derived from ground data, such as Lorey's height or basal area, are presented in Table S4.

Deleted: 4

Deleted: 2

Deleted: 3

Cross-site studies on the structure of tropical forests have led to significant advances in our understanding of tropical forest ecology (Gentry 1993; Phillips et al., 1998; ter Steege et al., 2006). They have also yielded important insights on new techniques to predict carbon stocks across regions (eg. Asner and Mascaro, 2014). Comparison of sites in terms of MCH derived for the study sites confirms that there is a strong regional variation of AGB with respect to canopy height, and that East Amazonian sites tend to have much taller trees than Central and Western Amazonia sites. This was already apparent in the canopy height maps produced by the GLAS sensor (Lefsky, 2010; Saatchi et al., 2011; Simard et al., 2011). Comparing sites in terms of LCA showed a similar pattern of larger trees, being relatively more present in eastern Amazonia, notably in the French Guiana sites and Tapajos. Our most southwestern site was Antimary, in the state of Acre (Brazilian Amazon). However, this site does not represent forests in the western Amazon or the Amazon-Andes gradients with relatively lower wood density (Baker et al. 2004) and more fertile volcanic soils impacting the forest structure and dynamics (Quesada et al., 2011). The site in Chocó is also unique in its characteristics because of extremely wet condition and potential disturbance (e.g., selective logging). Additional lidar and ground measurements will allow validating the performance of the LCA in representing the AGB variations in the western Amazon region.

Deleted: areas in

Deleted:

Deleted: the Peruvian Amazon and western

Deleted: unknown

Deleted: history

Deleted: would be needed in western Amazonia

Deleted: to further

Deleted: the patterns observed in this study.

4.2 Physical Interpretation of LCA

In this study, we introduced a simple structural metric that captures the proportion of area covered by large trees over the landscape (> 1 ha) and explained 78% of the variation in average forest volume and biomass when weighted by wood density in four sites of old growth Neotropical forests. LCA cannot separate the crown areas of individual trees. However, it is

Deleted: nine

adapted for large scale monitoring of forest volume and biomass change, as it is a robust and readily accessible metric. For individual tree separation, complex and more computationally intensive approaches are available (Ferraz et al., 2016).

In estimating LCA from Lidar data, we examined the spatial clustering properties of LCA and found that the minimum cluster size was less important than the threshold of canopy height, as long as the analysis focused on the relative covered area instead of on the density of large trees. We found that using the percentage of the area covered by large canopy trees is an efficient way of overcoming the problem of individual crown segmentation in Lidar data. LCA is related to how trees reaching the forest canopy (above a certain height) fill the space and how this characteristic may follow a spatially invariant scaling across tropical forests (West et al., 2009). Clusters smaller than 100 m^2 add only a small fraction (1.7% on average) to LCA values across sites. Including these clusters in LCA would not impact the performance of the model (similar R^2, RMSE and bias) and would allow to skip the final steps of the LCA retrieval (see Fig. S2). However, since these pixels either represent single branches reaching above 27m or the tip of a tree crown, they have no meaning in terms of our LCA metric and do not represent large trees. LCA provides information on the presence of large trees in a study area, which other metrics such as MCH cannot do. It is an important point, considering that large trees are often the most affected by natural disturbance and targeted by logging companies.

4.3 Correlation between LCA and AGB

The distribution of R^2 between LCA and AGB for (Fig. 3) is such that the maximum difference in R^2 between a threshold of 25m and 30m is approximately 0.1, a negligible value. Hence, AGB retrieval by LCA is relatively insensitive to the height threshold. For most sites, except

Antimary, we found a height threshold such that LCA explains about 80–90 % of the variation of AGB or total volume of the forests for each site (60–70 % when compared with ground plots) (Fig. 3). Using a height threshold of 27 m for all sites reduced the R^2 by 0.04 on average (max = 0.08) compared to the optimal height threshold for each site.

Deleted: Hence, the difference between the R^2 of Lidar and ground plots is due to the relative correlation between MCH used in Lidar derived biomass and LCA. Differences in Lidar characteristics for each site and differences in timing of Lidar observations and ground plots further amplify this problem. Finally, a limit to how much LCA can explain variation in AGB relates to forest structure and the AGB of small trees.

Potential differences in MCH among sites are due to footprint size, scan angle and return density (Disney et al., 2010; Hirata, 2004; Hopkinson, 2007) (Table 2). However, these effects are generally smaller than the 1m increment that we used to determine the optimal height thresholds of LCA. As a result, LCA estimation, and therefore AGB inferred from LCA, should depend little on instrument, acquisition and processing (Table 2). This is an important finding given the increasing variety of airborne Lidar sensors, and also given the pre and post-processing methods available for monitoring tropical forest structure and aboveground biomass. However, determining whether the 27m threshold holds for LCA calculation across in the tropics would require a validation at more study studies across continents.

Deleted: Our ground calibrated LCA model defined by Eq. (3) had a similar performance as the MCH based AGB model (R^2_{MCH} = 0.79, $RMSE_{MCH}$ = 44.2 Mg, Table S3).

4.4 LCA Relation to Ground Measurements

The relation between LCA derived from Lidar and the ground measurements can be further investigated by converting the 27 m height threshold into equivalent DBH values, using a height–diameter relationship. In the absence of a local DBH–height relation at each site, we made use of the following equation (Chave et al., 2014):

$$ln(H) = 0.893 - E + 0.760 \times ln(D) - 0.0340 \times (ln(D))^2 \quad (3)$$

Deleted: 4

where E is a measure of environmental stress for each site that potentially impacts the tree allometry. The corresponding DBH values fall around 35–55 cm, except for Chocó, where the best coefficient of correlation is reached with a DBH threshold of 29 cm (Fig. S4). The average minimum DBH to assign for the definition of large trees that represent variations of AGB is below 50 cm. By choosing a DBH threshold of 50 cm for old-growth undisturbed forests, the LCA model for estimating biomass can have an approximate analog in inventory data. This comparison suggests that the LCA model can also be adjusted with the average wood density of trees lager than 50 m, allowing a much faster ground data collection of calibrating LCA model for different sites (S.4).

A limit to how much LCA can explain variations in AGB relates to forest structure and the AGB of small trees. The lower range of biomass estimation for the LCA model, associated with the intercept for LCA equal to zero, ranged between 122 Mg ha^{-1} in La Selva and 192 Mg ha^{-1} in Paracou (Fig. 7a). This lower range identified with the intercept of the LCA–AGB linear model can be interpreted as the AGB associated with all trees smaller than 27 m height (approximately all trees with DBH <50 cm). Note that the differences between sites are due to differences in their mean wood density and not the volume of trees (see Eq. (2) and Fig. 4). Similarly, the contribution of small trees to the total biomass in the ground inventory ranges between around 100 and 200 Mg ha^{-1}, except in Paracou (261 Mg ha^{-1}) (Fig. 7b). AGB estimation based on LCA in these sites cannot go under 100 Mg ha^{-1} or over 500 Mg ha^{-1}. This is not a limitation of the model because LCA is designed to provide AGB estimates for forests reaching at least 27 m in mean canopy height, and such forests generally exceed 100 Mg ha^{-1} in AGB. Also, the upper threshold of 500 Mg ha^{-1} is consistent with upper values found globally at 1 ha scale (Brienen et al., 2015; Slik et al., 2013). A recalibration of the method should be envisaged in secondary and highly degraded forests.

Deleted: DBH estimation suggests that using a minimal DBH threshold of about 50 cm for large trees for old growth neo-tropical forests better represents the total AGB variations.

Deleted: Finally, a

Deleted: of

Deleted:

Deleted: only

Deleted: 3

4.5 LCA as AGB Estimator

The correlation of LCA to AGB_{inv} suggests that a Lidar based approach can lead to the estimation of AGB at the landscape scale and give useful information on the presence of large canopy trees and their distribution, extending the analysis of large trees in plot level inventory based studies (Bastin et al., 2015; Slik et al., 2013).

Therefore, LCA can explain the variations of total forest volume without any ancillary data about the forest or the landscape. Most bias in conversion of LCA to AGB, however, can be corrected across landscapes and sites by scaling the LCA–AGB relationship with average wood density at the landscape scale. Our model can therefore potentially be applied to a wide range of forest types, provided that there is information about wood density of the study area in the literature.

Wood density has been shown to be a key element of allometric models of AGB estimation (Baker et al., 2004; Brown et al., 1989; Chave et al., 2004; Nogueira et al., 2007). If wood density is assumed to be constant across DBH classes, the mean wood density at the plot scale can readily be used to scale LCA to biomass. However, if the wood density of large trees is smaller or larger than the average wood density, (e.g. in BCI and Chocó: S.4, Fig. S5), the use of mean wood density to scale LCA may introduce a slight bias in biomass estimation. A difference in mean wood density of 0.1 g cm^{-3} would introduce a bias of ±10 % in the biomass estimation when using our model. We found that using mean wood density of large trees or basal area weighted wood density instead can give slightly better results and could circumvent the differences in size distribution of the wood density (S.4). Instead we could rely on the wood density of large trees only. This would make the collection of ground data easier and cost effective for biomass estimation, because trees ≥50 cm DBH only represent 5–10 % of the stems of a plot (S.4, Fig. S6). Focusing on the wood density of dominant or hyper dominant species could also be an alternative approach for future use of Lidar derived LCA for large scale biomass estimation (Fauset et al., 2015; ter Steege et al., 2013). In the absence of information on wood density from the literature, modelled wood density could potentially be used, but would give greater errors. These errors should be taken into account when reporting on the uncertainty of the results.

4.6 LCA and MCH

The comparison of LCA and MCH metrics showed that both performed similarly in estimating AGB, highlighting the importance of large canopy trees to estimate biomass. The differences between the two methods in estimating AGB show that two methods can have similar performance in terms of R^2 and RMSE and nonetheless lead to different estimations, with LCA giving higher AGB estimations in some sites. The choice of a metric is therefore crucial to estimate AGB, especially when estimating the changes in biomass (see Section 4.7).

Both MCH and LCA–AGB models performed relatively poorly in high biomass plots of the Nouragues study area, by underestimating biomass values greater than 500 Mg ha^{-1} (Fig. 4 and 5). To explain the underestimation, we performed three tests: 1. We examined the differences in the ground estimated biomass values with and without tree height and found no significant impact in reducing the effect of underestimation. 2. We tested the hypothesis that the height threshold used for LCA estimation across sites was not suitable for the Nouragues study site and dismissed the hypothesis because 27 m was found to be the optimum threshold for Nouragues plots. 3. We examined the errors in the Lidar estimation of forest height and found that except for an extremely high AGB_{inv} of 617 Mg ha^{-1}, the four other high biomass outliers are all located in the 6 ha Pararé plot located on a very steep topography. The Lidar digital terrain model

Deleted: models

Deleted: 6

Deleted:

Deleted:

(DTM) of this area shows an average within plots elevation range of 90 m. Ground detection on steep terrain can be erroneous, depending on the Lidar point density and the view angle, causing large area interpolation errors for DTM development and significant error in canopy height measurements (Leitold et al., 2015). Other factors that may affect the underestimation of AGB by LCA or MCH in the Nouragues site may be due to the presence of forest patches with clusters of large trees and overlapping crown areas. It is also possible that the relationship between AGB and LCA is not linear for very high AGB values. This could be tested in the future with a larger number of sites with very high biomass.

4.7 LCA and forest degradation

Although LCA and MCH may perform similarly in capturing the forest biomass variations and changes, the use of LCA in detecting forest degradation and logging is more straightforward because of its relation to large trees. The LCA approach was able to accurately detect changes in forests after logging by locating where the large trees are extracted. Our estimate of biomass change from the LCA approach was higher than the biomass loss of 9.1 Mg ha^{-1} reported by another study using the 25^{th} percentile height above ground as the Lidar metric for biomass estimation (Andersen et al. 2014). It can be expected that relying on the 25^{th} percentile height metric for biomass estimation would place more emphasis on the lower part of the canopy (understory) that is either less damaged or has gone through some level of regeneration after logging. Models based on LCA or MCH, on the other hand, may be more realistic for estimating AGB changes because they capture the changes in large trees and upper forest canopy structure that contain most of the biomass and are directly impacted by logging and biomass removal.

The higher biomass loss estimation from the MCH model (19 Mg ha^{-1}) again shows how different metrics can lead to different results. Here, three methods based on three different Lidar metrics yielded results that differed by more than twofold. LCA could become an important tool to detect forest degradation, in particular selective logging, considering that large trees are targeted by logging companies.

4.8 Future Applications of LCA

Deleted: 7

LCA definition in our study relies on the high resolution information on forest height, allowing for the detection of crown area of large canopy trees. Can a similar measure be derived from large footprint Lidar observations such as the future NASA spaceborne Lidar mission GEDI (Global Ecosystem Dynamic Investigation)? GEDI will not provide spatially continuous data on forest height, but its footprint size (~ 25 m) and dense sampling may be adequate to develop statistical indicators of large trees over the landscape.

Similarly, future spaceborne radar missions could also provide useful information to retrieve large canopy areas. The synthetic aperture radar (SAR) tomographical observations of the European Space Agency (ESA) BIOMASS mission will provide wall-to-wall imagery of canopy profile that could be converted to LCA over the landscape (Le Toan et al., 2011). Preliminary research based on airborne TomoSAR measurements has already shown that backscatter power at about 30 m above the ground, with sensitivity to the distribution of large trees, explained the variation of AGB over Nouragues and Paracou plots better than the backscatter power related to the lower part of the canopy (0–15 m) (Minh et al., 2016; Rocca et al., 2014). Future research on exploring the use of an equivalent radar index product from BIOMASS height or tomography measurements at a height threshold (e.g. 27 m) may provide a potential algorithm to map the area of large trees and estimate forest volume and biomass changes across the landscape.

5 Conclusions

We introduce LCA as a new Lidar derived index to capture the variations of large trees and total volume and biomass across landscapes that remain spatially and regionally invariant. The importance of LCA is in its relevance to the structure and ecological characteristics of large trees in filling the canopy space and their unique contribution in determining the total volume and biomass of forests. Unlike other Lidar derived metrics, LCA is linearly related to total aboveground biomass after being weighted by average wood density. This linear relationship remains unique across different forest types, making the LCA model broadly applicable. The comparison of LCA index with ground plots suggests that DBH >50 cm is a more reliable threshold to quantify the number and distribution of large trees in undisturbed old growth tropical forests and in capturing the variations of the total aboveground biomass across landscapes and regions. The results of our study may encourage further research in the use of Lidar data for detecting the distribution of larger trees in tropical forests for ecological and conservation studies.

Author contribution

V. Meyer and S. Saatchi developed the model and designed the study. V. Meyer developed the model code and performed the analysis. J. Chave, G. Vincent, M. Keller, F. Espírito-Santo, D.

Clark and M. d'Oliveira provided inventory data and derived metrics necessary to run the experiments. A. Ferraz contributed to the data processing. D. Kaki performed a preliminary analysis of the data. V. Meyer prepared the manuscript with contributions from all co-authors. The authors declare that they have no conflict of interest.

Acknowledgements

The work described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. This work has benefited from "*Investissement d'Avenir*" grants managed by the French *Agence Nationale de la Recherche* (CEBA, ref. ANR-10-LABX-25-01 and TULIP, ref. ANR-10-LABX-0041; ANAEE-France: ANR-11-INBS-0001) and from CNES (TOSCA project; PI T Le Toan). Field and Lidar data from the Brazilian sites were acquired by the Sustainable Landscapes Brazil project supported by the Brazilian Agricultural Research Corporation (EMBRAPA), the US Forest Service, and USAID, and the US Department of State. La Selva field work was supported by the U.S. National Science Foundation LTREB Program NSF LTREB 1357177. Data in Chocó are available as part of the Reducing Emissions from Deforestation and forest Degradation (REDD) project. FES was supported by Natural Environment Research Council (NERC) grants ('BIO-RED' NE/N012542/1 and 'AFIRE' NE/P004512/1) and Newton Fund ('The UK Academies/FAPESP Proc. N°: 2015/50392-8 Fellowship and Research Mobility'). The AGB data for Paracou were made available courtesy of CIRAD (B. Hérault).

Data accessibility

The BCI lidar and forest inventory dataset used in this research are publically available from the Office of Bioinformatics, Smithsonian Tropical Research Institute. All relevant data are within the paper and its Supporting Information files.

References

Andersen, H. E., Reutebuch, S. E., McGaughey, R. J., d'Oliveira, M. V. and Keller, M.: Monitoring selective logging in western Amazonia with repeat lidar flights. Remote Sens. Environ., 151, 157-165, 2014.

Asner, G. P., Mascaro, J., Muller-Landau, H. C., Vieilledent, G., Vaudry, R., Rasamoelina, M., Hall, J. S. and van Breugel, M.: A universal airborne Lidar approach for tropical forest carbon mapping. Oecologia, 168(4), 1147-1160, 2012.

Asner, G. P. and Mascaro, J.: Mapping tropical forest carbon: Calibrating plot estimates to a simple Lidar metric. Remote Sens. Environ. 140, 614-624, 2014.

Baker, T. R., Phillips, O. L., Malhi, Y., Almeida, S., Arroyo, L., Di Fiore, A., Erwin, T., Killeen, T. J., Laurance, S. G., Laurance, W. F. and Lewis, S. L.: Variation in wood density determines spatial patterns in Amazonian forest biomass. Glob. Change Biol., 10(5), 545-562. doi: 10.1111/j.1365-2486.2004.00751.x, 2004.

Baldocchi, D. D.: Assessing the eddy covariance technique for evaluating carbon dioxide exchange rates of ecosystems: past, present and future. Glob. Change Biol., 9(4), 479-492, 2003.

Basset, Y., Cizek, L., Cuénoud, P., Didham, R. K., Guilhaumon, F., Missa, O., Novotny, V., Ødegaard, F., Roslin, T., Schmidl, J. and Tishechkin, A. K.: Arthropod diversity in a tropical forest. Science, 338(6113), 1481-1484, 2012.

Bastin, J.-F., Barbier, N., Réjou-Méchain, M., Fayolle, A., Gourlet-Fleury, S., Maniatis, D., de Haulleville, T., Baya, F., Beeckman, H., Beina, D. and Couteron, P.: Seeing Central African forests through their largest trees. Sci. Rep.-UK, 5, 13156, 2015.

Bioredd.org/ accessed 4.13.2016

Bohlman, S., and O'Brien, S.: Allometry, adult stature and regeneration requirement of 65 tree species on Barro Colorado Island, Panama. J. Trop. Ecol., 22(02), 123-136, 2006.

Brienen, R. J. W., Phillips, O. L., Feldpausch T. R., Gloor E., Baker, T. R., Lloyd, J. and Lopez-Gonzalez G.: Long-Term Decline of the Amazon Carbon Sink. Nature, 519, 344. http://dx.doi.org/10.1038/nature14283, 2015.

Brown, S., Gillespie, A. J., and Lugo, A. E.: Biomass estimation methods for tropical forests with applications to forest inventory data. Forest Sci., 35(4), 881-902, 1989.

Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S., and Perez, R.: Error propagation and scaling for tropical forest biomass estimates, Philos. T. R. Soc. B, 359, 409–420, 2004.

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M. S., Delitti, W. B., and Vieilledent, G.: Improved allometric models to estimate the aboveground biomass of tropical trees. Glob. Change Biol., 20(10), 3177-3190, 2014.

Clark D. B. and Clark D. A.: Abundance, growth and mortality of very large trees in neotropical lowland rain forest. Forest Ecol. and Manag., 80, 235–244, 1996.

Clark, D. B. and Clark, D. A.: Landscape-scale variation in forest structure and biomass in a tropical rain forest. Forest Ecol. and Manag., 137, 185–198, 2000.

Condit, R.: Tropical Forest Census Plots. Springer Verlag and R.G. Landes Company. Berlin and Georgetown, TX, 1998.

d'Oliveira, M. V. N., Reutebuch, S. E., McGaughey, R. J. and Andersen, H. E. : Estimating forest biomass and identifying low-intensity logging areas using airborne scanning lidar in Antimary State Forest, Acre State, Western Brazilian Amazon. Remote Sens. Environ., 124, 479-491, 2012.

Denslow, J. S. : Gap portioning among tropical rainforest trees. Biotropica, 12, 47–55, 1980.

Disney, M. I., Kalogirou, V., Lewis, P., Prieto-Blanco, A., Hancock, S., and Pfeifer, M.: Simulating the impact of discrete-return Lidar system and survey characteristics over young conifer and broadleaf forests. Remote Sens. Environ., 114(7), 1546-1560, 2010.

ENVI/IDL, Exelis Visual Information Solutions, Boulder, Colorado.

Espírito-Santo, F. D. B., Keller, M., Braswell, B., Nelson, B. W., Frolking, S., and Vicente, G.: Storm intensity and old-growth forest disturbances in the Amazon region. Geophys. Res. Lett., 37(11), 2010.

Espírito-Santo, F. D. B., Keller, M. M., Linder, E., Oliveira, R. C. Junior, Pereira, C. and Oliveira, C. G.: Gap formation and carbon cycling in the Brazilian Amazon: measurement using high-resolution optical remote sensing and studies in large forest plots. Plant Ecol. Divers., 7, 305–318, 2014.

Fauset, S., Johnson, M. O., Gloor, M., Baker, T. R., Monteagudo, A., Brienen, R. J., Feldpausch, T. R., Lopez-Gonzalez, G., Malhi, Y., Ter Steege, H. and Pitman, N. C.: Hyperdominance in Amazonian forest carbon cycling. Nat. Commun., 6, 2015.

Fearnside, P. M.: Wood density for estimating forest biomass in Brazilian Amazonia. Forest Ecol. and Manag., 90(1), 59-87, 1997.

Ferraz, A., Saatchi, S., Mallet, C., and Meyer, V.: Lidar detection of individual tree size in tropical forests. Remote Sens. Environ., 183, 318-333, 2016.

Figueiredo, E. O., d'Oliveira, M. V. N., Braz, E. M., de Almeida Papa, D. and Fearnside, P. M.: LIDAR-based estimation of bole biomass for precision management of an Amazonian forest: Comparisons of ground-based and remotely sensed estimates. Remote Sens. Environ., 187, 281-293, 2016.

Gentry, A. H.: Four neotropical rainforests. Yale University Press, 1993.

Goldstein, G., Andrade, J. L., Meinzer, F. C., Holbrook, N. M., Cavelier, J., Jackson, P., and Celis, A.: Stem water storage and diurnal patterns of water use in tropical forest canopy trees. Plant Cell Environ., 21(4), 397-406, 1998.

Goodman, R. C., Phillips, O. L., and Baker, T. R.: The importance of crown dimensions to improve tropical tree biomass estimates. Ecol. Appl., 24(4), 680-698, 2014.

Gourlet-Fleury, S., Guehl, J.-M. and Laroussinie, O.: Ecology and management of a neotropical rainforest. Lessons drawn from Paracou, a long-term experimental research site in French Guiana. Elsevier, Amsterdam, 2004.

Hirata, Y.: The effects of footprint size and sampling density in airborne laser scanning to extract individual trees in mountainous terrain. Proc. ISPRS WG VIII/2 "Laser-scanners for forestry and landscape assessment", Vol. XXXVI, Part 8/W2, 3–6 October 2004, Freiburg, Germany, 2004.

Hopkinson, C.: The influence of flying altitude, beam divergence, and pulse repetition frequency on laser pulse return intensity and canopy frequency distribution. Can. J. Remote Sens., 33(4), 312−324, 2007.

Hubbell, S. P., Foster, R. B., O'Brien, S. T., Harms, K. E., Condit, R., Wechsler, B., Wright, S. J. and De Lao, S. L.: Light gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. Science, 283, 554-557, 1999.

Jubanski, J., Ballhorn, U., Kronseder, K., Franke, J., and Siegert, F.: Detection of large above-ground biomass variability in lowland forest ecosystems by airborne Lidar. Biogeosciences, 10(6), 3917-3930, 2013.

Kellner, J. R., and Asner, G. P.: Convergent structural responses of tropical forests to diverse disturbance regimes. Ecol. Lett., 12(9), 887-897, 2009.

Laurance, W. F., Delamônica, P., Laurance, S. G., Vasconcelos, H. L., and Lovejoy, T. E.: Conservation: rainforest fragmentation kills big trees. Nature, 404(6780), 836-836. doi:10.1038/35009032, 2000.

Le Toan, T., Quegan, S., Davidson, M. W. J., Balzter, H., Paillou, P., Papathanassiou, K., Plummer, S., Rocca, F., Saatchi, S., Shugart, H. and Ulander, L.: The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle. Remote Sens. Environ., 115(11), 2850-2860, 2011.

Lefsky, M. A., Cohen, W. B., Parker, G. G., and Harding, D. J.: Lidar remote sensing for ecosystem studies, BioScience, 52, 19–30, 2002.

Lefsky, M. A.: A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. Geophys. Res. Lett., 37(15), 2010.

Lefsky, M. A., Keller, M., Pang, Y., De Camargo, P. B., and Hunter, M. O.: Revised method for forest canopy height estimation from Geoscience Laser Altimeter System waveforms. J. Appl. Remote Sens., 1(1), 013537, 2007.

Leitold, V., Keller, M., Morton, D. C., Cook, B. D., and Shimabukuro, Y. E.: Airborne Lidar-based estimates of tropical forest structure in complex terrain: opportunities and trade-offs for REDD+. Carbon Balance Management, 10(1), 3, 2015.

Mascaro, J., Detto, M., Asner, G. P., and Muller-Landau, H. C.: Evaluating uncertainty in mapping forest carbon with airborne Lidar. Remote Sens. Environ., 115, 3770-3774, 2011.

Meyer, V., Saatchi, S. S., Chave, J., Dalling, J. W., Bohlman, S., Fricker, G. A., Robinson, C., Neumann, M., and Hubbell, S.: Detecting tropical forest biomass dynamics from repeated airborne Lidar measurements. Biogeosciences, 10(8), 5421-5438, 2013.

Minh, D. H. T., Le Toan, T., Rocca, F., Tebaldini, S., Villard, L., Réjou-Méchain, M., Phillips, O. L., Feldpausch, T.R., Dubois-Fernandez, P., Scipal, K. and Chave, J.: SAR tomography for the retrieval of forest biomass and height: Cross-validation at two tropical forest sites in French Guiana. Remote Sens. Environ., 175, 138-147, 2016.

Nepstad, D. C., Tohver. I. M., Ray D., Moutinho, P., and Cardinot, G.: Mortality of large trees and lianas following experimental drought in an Amazon forest. Ecology 88, 2259–2269, 2007.

Nogueira, E. M., Fearnside, P. M., Nelson, B. W., and França, M. B.: Wood density in forests of Brazil's ‘arc of deforestation’: Implications for biomass and flux of carbon from land-use change in Amazonia. Forest Ecol. and Manag., 248(3), 119-135, 2007.

Packalen, P., Strunk, J. L., Pitkänen, J. A., Temesgen, H., and Maltamo, M.: Edge-tree correction for predicting forest inventory attributes using area-based approach with airborne laser scanning. IEEE J. Sel. Top. Appl., 8(3), 1274-1280, 2015.

Pascual, M., and Guichard, F.: Criticality and disturbance in spatial ecological systems. Trends Ecol. Evol., 20(2), 88-95, 2005.

Pearson, T. R., Brown, S., and Casarim, F. M.: Carbon emissions from tropical forest degradation caused by logging. Environ. Res. Lett., 9(3), 034017, 2014.

Phillips, O. L., Malhi, Y., Higuchi, N., Laurance, W. F., Núnez, P. V., Vásquez, R. M., Laurance, S. G., Ferreira, L. V., Stern, M., Brown, S. and Grace, J.: Changes in the carbon balance of tropical forests: evidence from long-term plots. Science, 282(5388), 439-442, 1998.

Phillips, O. L., Aragão, L. E., Lewis, S. L., Fisher, J. B., Lloyd, J., López-González, G., Malhi, Y., Monteagudo, A., Peacock, J., Quesada, C. A. and Van Der Heijden, G.: Drought sensitivity of the Amazon rainforest. Science, 323(5919), 1344-1347, 2009.

Popescu, S. C., Wynne, R. H., and Nelson, R. F.: Measuring individual tree crown diameter with Lidar and assessing its influence on estimating forest volume and biomass. Can. J. Remote Sens., 29(5), 564-577, 2003.

Quesada, C. A., Lloyd, J., Anderson, L. O., Fyllas, N. M., Schwarz, M., and Czimczik, C. I.: Soils of Amazonia with particular reference to the RAINFOR sites, Biogeosciences, 8, 1415-1440, https://doi.org/10.5194/bg-8-1415-2011, 2011.

R Core Team, 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Réjou-Méchain, M., Tymen, B., Blanc, L., Fauset, S., Feldpausch, T. R., Monteagudo, A., Phillips, O. L., Richard, H. and Chave, J.: Using repeated small-footprint Lidar acquisitions to infer spatial and temporal variations of a high-biomass Neotropical forest. Remote Sens. Environ., 169, 93-101, 2015.

Rocca, F., Dinh, H. T. M., Le Toan, T., Villard, L., Tebaldini, S., d'Alessandro, M. M., and Scipal, K.: Biomass tomography: A new opportunity to observe the earth forests. Int. Geosci. Remote Se., 1421-1424 , 2014.

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E.T., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S. and Petrova, S.: Benchmark map of forest carbon stocks in tropical regions across three continents. P. Natl Acad. Sci. USA, 108(24), 9899-9904, 2011.

Saatchi, S. S., Asefi-Najafabady, S., Malhi, Y., Aragão, L. E., Anderson, L. O., Myneni, R. B., and Nemani, R.: Persistent effects of a severe drought on Amazonian forest canopy. P. Natl Acad. Sci. USA, 110(2), 565-570, 2013.

Santiago, L. S., Goldstein, G., Meinzer, F. C., Fisher, J. B., Machado, K., Woodruff, D., and Jones, T.: Leaf photosynthetic traits scale with hydraulic conductivity and wood density in Panamanian forest canopy trees. Oecologia, 140(4), 543-550, 2004.

Simard, M., Pinto, N., Fisher, J. B., and Baccini, A.: Mapping forest canopy height globally with spaceborne lidar, Journal of Geophysical Research - Biogeosciences, 116, G04021, doi:10.1029/2011JG001708, 2011.

Slik, J. W., Paoli, G., McGuire, K., Amaral, I., Barroso, J., Bastian, M., Blanc, L., Bongers, F., Boundja, P., Clark, C. and Collins, M. : Large trees drive forest aboveground biomass variation in moist lowland forests across the tropics. Global Ecol. and Biogeogr., 22(12), 1261-1271, 2013.

Solé, R. V., and Manrubia, S. C.: Are rainforests self-organized in a critical state?. J. Theor. Biol., 173(1), 31-40, 1995.

Strigul, N., Pristinski, D., Purves, D., Dushoff, J., and Pacala, S.: Scaling from trees to forests: tractable macroscopic equations for forest dynamics. Ecol. Monogr., 78(4), 523-545, 2008.

Deleted:

Ter Steege, H., Pitman, N. C., Phillips, O. L., Chave, J., Sabatier, D., Duque, A., Molino, J. F., Prévost, M. F., Spichiger, R., Castellanos, H. and Von Hildebrand, P.: Continental-scale patterns of canopy tree composition and function across Amazonia. Nature, 443(7110), 444-447, 2006.

Ter Steege, H., Pitman, N.C., Sabatier, D., Baraloto, C., Salomão, R. P., Guevara, J.E., Phillips, O. L., Castilho, C. V., Magnusson, W. E., Molino, J. F. and Monteagudo, A. :Hyperdominance in the Amazonian tree flora. Science, 342(6156), 1243092, 2013.

Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K. M. and Lien, V.: Comparative testing of single-tree detection algorithms under different types of forest. Forestry, 85(1), 27-40, 2011.

Vauhkonen, J., Næsset, E., and Gobakken, T.: Deriving airborne laser scanning based computational canopy volume for forest biomass and allometry studies. ISPRS J. Photogramm., 96, 57-66, 2014.

Vincent, G., Sabatier, D., Blanc, L., Chave, J., Weissenbacher, E., Pélissier, R., Fonty, E., Molino, J. F. and Couteron, P.: Accuracy of small footprint airborne Lidar in its predictions of tropical moist forest stand structure. Remote Sens. Environ., 125, 23-33, 2012.

West, G.B., Enquist, B. J. and Brown, J. H. : A general quantitative theory of forest structure and dynamics. P. Natl Acad. Sci. USA, 106, 7040–7045, 2009.

Zhou, J., Proisy, C., Descombes, X., Hedhli, I., Barbier, N., Zerubia, J., Gastellu-Etchegorry, J. P. and Couteron, P.: Tree crown detection in high resolution optical and Lidar images of tropical forest. P. Soc. Photo-Opt. Ins., 7824. SPIE, 2010.
, 2010.

Table 1. Information on forest inventory plots. * indicates that a site has been used for the calibration of the LCA model. Sources: Antimary and Cotriguaçu: Fearnside, 1997; d'Oliveira et al., 2012, BCI: Center for Tropical Forest Science (CTFS) (Condit, 1998; Hubbell et al., 1999, 2005), Chocó: (bioredd.org), La Selva: Carbono project (Clark and Clark, 2000), Manaus and Tapajós: Espírito-Santo (unpublished results), Nouragues: Réjou-Méchain et al., 2015, Paracou: Gourlet-Fleury et al., 2004; Vincent et al., 2012.

Site	Data	Plots Size (ha)	N plots	Year	Mean WD ($g\ cm^{-3}$)	Mean AGB ($Mg\ ha^{-1}$)	Annual rainfall ($mm\ yr^{-1}$)
Antimary (Brazil)	Plot level	0.25	50	2010	0.61	234	2000
BCI * (Panama)	Tree level	1	50	2010	0.56	235	2600
Chocó (Colombia)	Tree level	0.25	42	2013	0.60	224	10000
Cotriguaçu (Brazil)	Not available	-	-	-	0.60	-	2000
La Selva * (Costa Rica)	Tree level	1	11	2009	0.45	178	4000
Manaus (Brazil)	Tree level	0.25	10	2014	0.66	263	2200
Nouragues * (French Guiana)	Plot level Tree level	1 1	33 7/33	2012	0.66	424	3000
Paracou * (French Guiana)	Plot level	1	85	2009-10	0.71	353	3000
Tapajós (Brazil)	Tree level	0.25	10	2014	0.62	238	1900

Table 2. Information on Lidar data and locations of the 9 research sites.

Site (1km^2 images)	Sensor	Year	Returns m^{-2}	Flight Altitude (m)	Scanning angle (°)	Frequency (kHz)	NW corner lat	NW corner lon
Antimary	Optech ALTM3100EA	2010-2011	10-15	500	11	70	9°17'47.26"S	68°17'15.06"W
BCI	Optech ALTM3100EA	2009	8	1000	35	70	9°9'28.56"N	79°51'18.9"W
Chocó	Optech ALTM3033	2013	4	1000	20	33	3°57'5.71"N	76°49'10.31"W
Cotriguaçu	Optech ALTM3100EA	2011	10-15	850	11	60	9°27'8.87"S	58°51'51.22"W
La Selva	Optech ALTM3100EA	2009	4	1500	20	70	10°25'37.97"N	84°1'8.76"W
Manaus	Optech ALTM3100EA	2012	10-15	850 (max)	11	60	2°56'38.48"S	59°56'12.57"W
Nouragues	Riegl LMS-Q560	2012	12	400	45	200	4°3'10.0"N	52°42'19.95"W
Paracou	Riegl LMS-280i	2009	4	120-220	30	24	5°15'47.73"N	52°56'26.96"W
Tapajós	Optech ALTM3100EA	2011	10-15	850 (max)	11	60	2°50'53.41"S	54°57'44.53"W

Table 3. Coefficients, R^2, RMSE and bias for the models used to estimate AGB_{LCA} without and with wood density as a weighting factor (m_LCA) and m_LCA_wd, respectively).

Model	Equation	a	b	R^2	RMSE	Bias	R^2 cross-val	RMSE cross-val	Bias cross-val
m_LCA	AGB = aLCA + b (Eq. (2))	3.56	136.91	0.59	62.53	0.0	0.58	63.26	0.16
m_LCA_wd	AGB = (aLCA+b) × WD (Eq. (3))	4.47	270.27	0.78	46.02	-0.76	0.77	46.47	-0.63

Figure 1. Segmentation of the 1 km × 1 km images in each site using five canopy height thresholds. A minimum of 100 contiguous pixels was used as a segmentation threshold in all cases.

Figure 2. LCA in function of height thresholds in the nine study sites. The steepest slopes are between 24 m (Antimary) and 30 m (Nouragues), with an average of 27 m across sites. Steepness of slope was obtained by calculating the derivative of the sigmoid models charactering each site.

Figure 3. Distribution of R^2 between tree height thresholds used to determine LCA and AGB_{Local} in the nine 1 ha subareas (a) and distribution of R^2 between tree height thresholds and AGB_{inv} in 1 ha inventory plots of the four calibration sites (b). All optimal thresholds are between 23 m and 30 m. The average maximal height threshold is 27 m.

Figure 4. Relationship between AGB_{inv} and LCA (a), AGB_{inv} normalized by averaged wood (b), and AGB_{inv} vs. AGB_{LCA} estimated with LCA_wd model (c). The black line represents the 1-to-1 line. Normalizing AGB by averaged wood density brings the data from different sites closer to a common fit.

Figure 5. AGB_{MCH} vs. AGB_{LCA} in the plots of the four calibration sites (a), and AGB_{MCH} vs. AGB_{LCA} in the 1km^2 images of the nine sites (b). The black line represents the 1-to-1 line.

Figure 6. Detection of changes of forest structure from selective logging in the Antimary study area showing a) the difference between pre- and post- logging (2010–2011) Lidar derived LCA at 1 ha grid cells over the entire study area, b) the histogram of LCA for the two Lidar datasets showing the mean difference and the reduction of medium and large LCA areas from selective logging, c) 2010 Lidar LCA segmentation at 1 m resolution over a sample area in the north of the study site, d) same LCA segmentation for 2011 Lidar data, and e) difference of the two segmented areas showing the extent of the logging impact on large trees in addition to natural changes of forest structure from changes in canopy gaps from tree falls and tree growth.

Figure 7. Relationship between LCA and AGB_{LCA} (a) and relationship between AGB_{inv} of large trees (>50 cm DBH) and total AGB_{inv} (b). In both cases, the intercepts represent the contribution of small trees to total AGB. Note that Manaus and Nouragues overlap because they have the same mean wood density, as well as Chocó and Cotriguaçu.

Deleted: AGB_{Lidar}

Deleted: density

Deleted: and

Deleted: AGB

Deleted: density

Deleted: .

Deleted: AGB_{inv}

Deleted: density

Deleted: estimated with LCA_wd model

Deleted: .

Formatted: Superscript

Deleted: AGB_{Lidar} density from the 1km^2 images vs. AGB_{LCA} estimated with LCA_wd model (b).

Figure 1

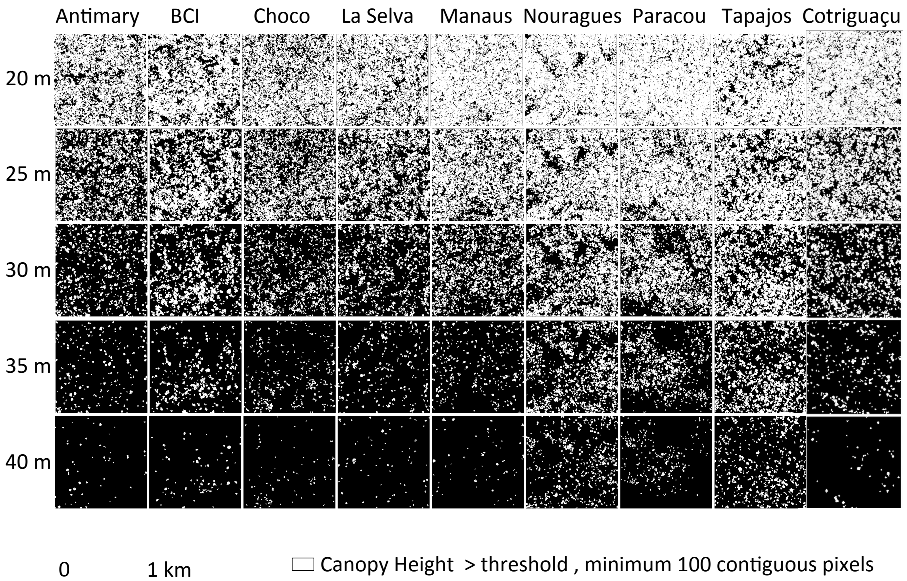

Antimary
BCI
Choco
La Selva
Manaus
Nouragues
Paracou
Tapajos
Cotriguaçu
20 m
25 m
30 m
35 m
40 m
0
1 km
Canopy Height > threshold , minimum 100 contiguous pixels

Figure 2

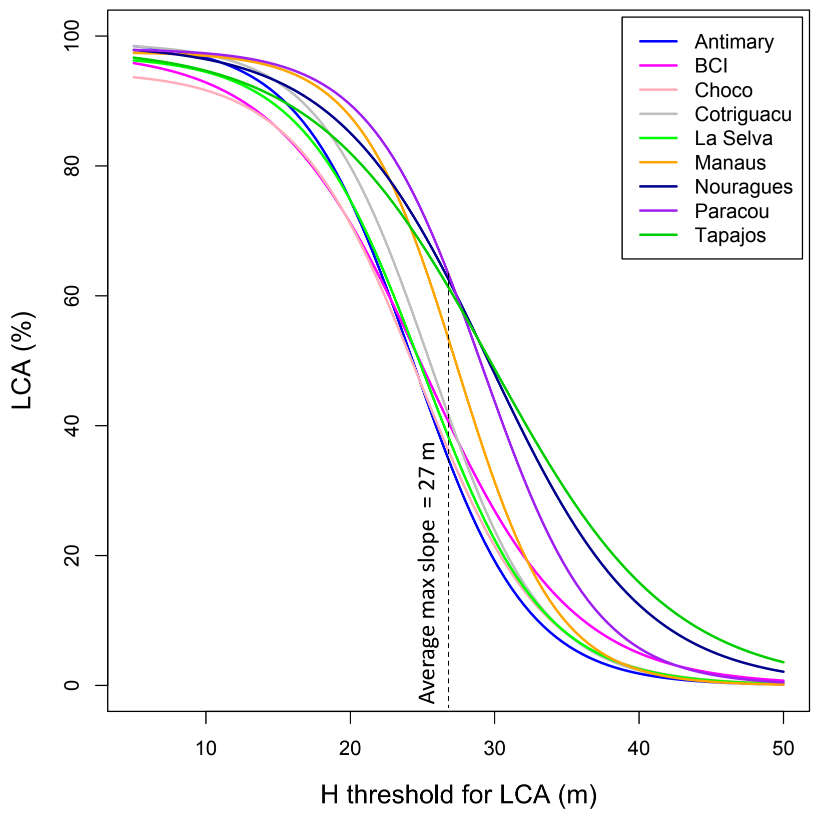
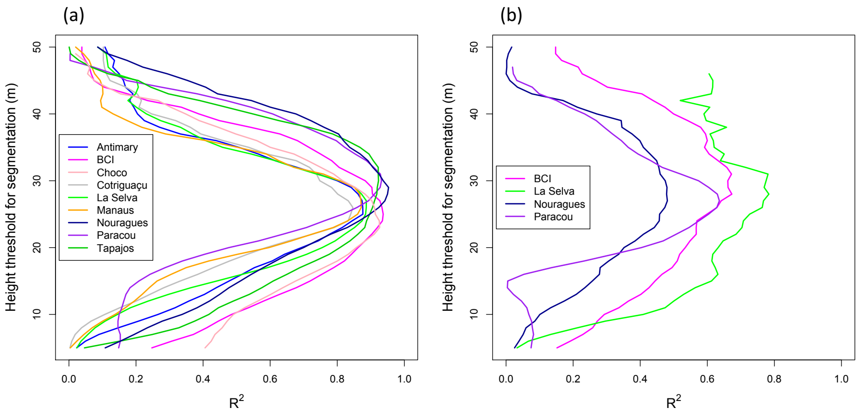
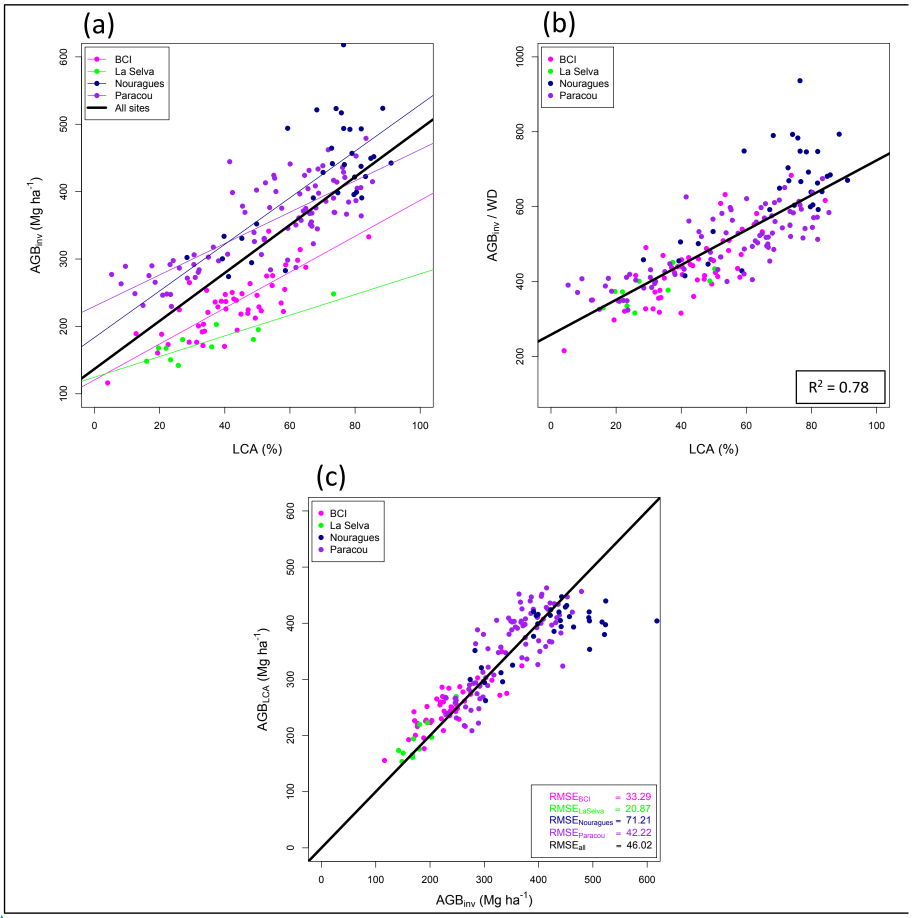

Antimary
BCI
Choco
Cotriguacu
La Selva
Manaus
Nouragues
Paracou
Tapajos
Average max slope = 27 m
LCA (%)
H threshold for LCA (m)
0
20
40
60
80
100
10
20
30
40
50

Figure 3

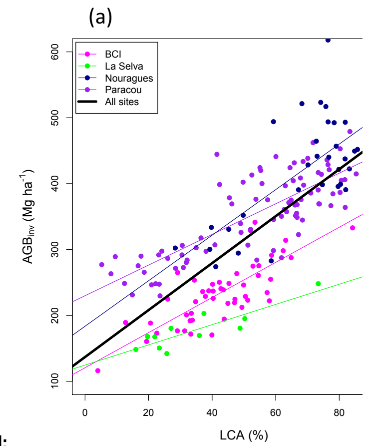


(a)
(b)
Height threshold for segmentation (m)
R^2
Antimary
BCI
Choco
Cotriguaçu
La Selva
Manaus
Nouragues
Paracou
Tapajos
BCI
La Selva
Nouragues
Paracou
0.0
0.2
0.4
0.6
0.8
1.0
10
20
30
40
50

Figure 4

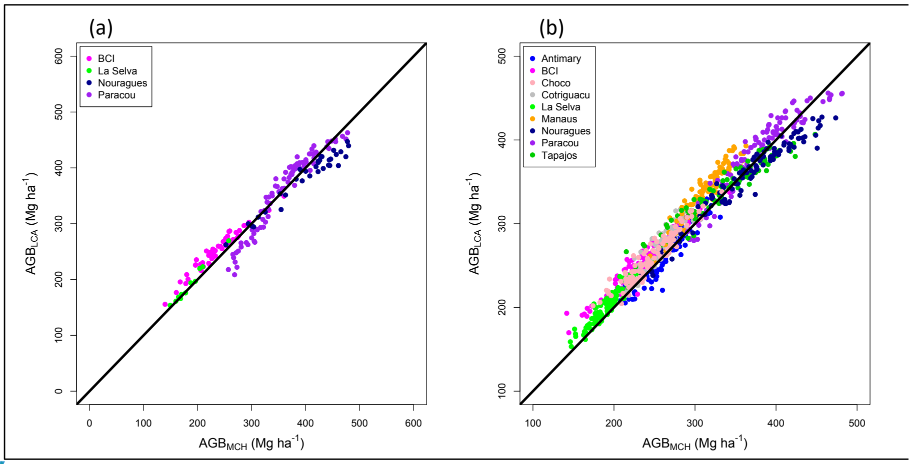


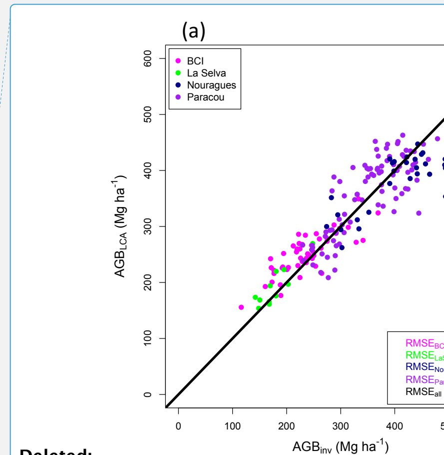
Formatted: Font color: Black



Deleted:

Figure 5

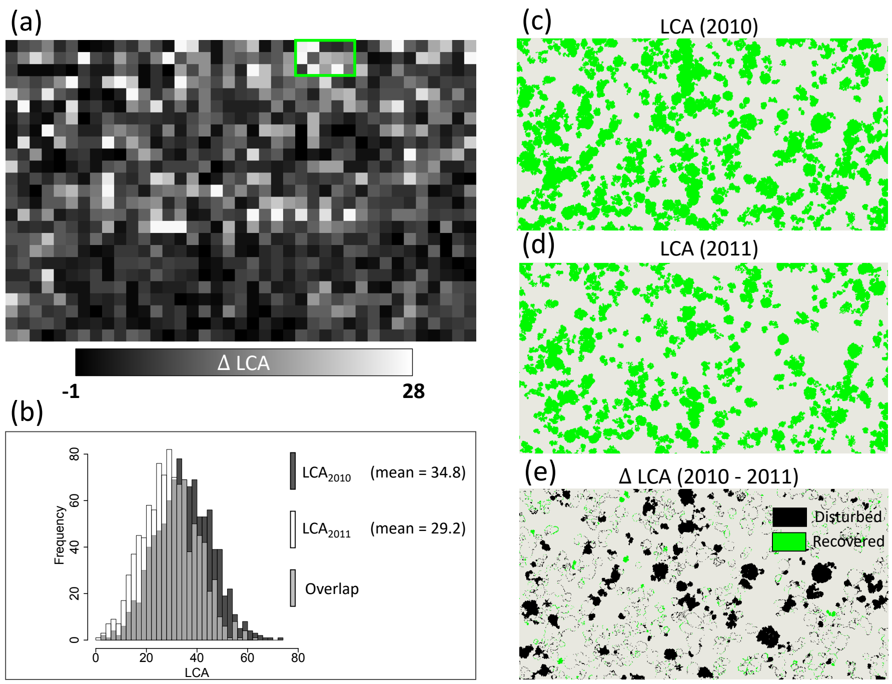




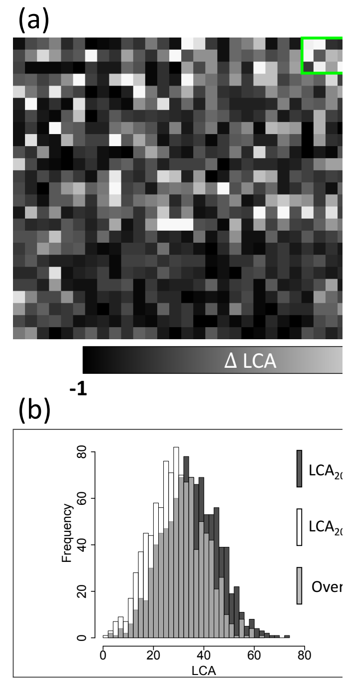
Deleted:

Formatted: Font color: Black

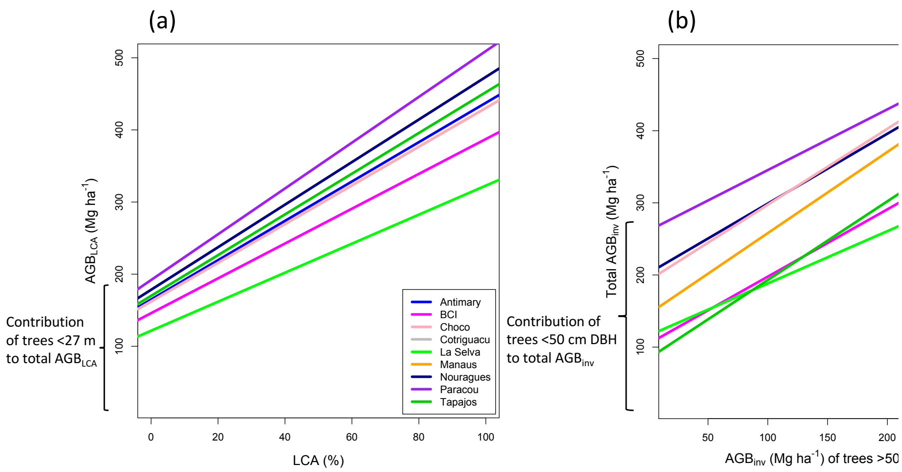
Figure 6



Deleted:



1029 Figure 7



1030