Dear Editor,

My co-authors and I wish to thank both reviewers for their valuable suggestions. A revised version of the paper accompanies this letter. We have addressed all of the reviewer suggestions as we detail below. I hope that this version attains the high standards of *Global Change Biology*. On the following pages we present a point-by-point response to each comment received.

Sincerely,

Eric Bastos Gorgens

Professor

**RESPONSE TO REVIEWER 1 COMMENTS**

**a/ around line 90, it would be useful to clarify what tree height metrics were derived. It seems that the authors were able to quantify the presence/absence of trees greater than a specified height per transect, but it would seem that they actually quantified the number of 1-m2 pixels exceeding the threshold height per transect. If the latter, it would be valuable to provide statistics of how many pixels are beyond 70 m (excluding non-tree maxima, as briefly alluded to line 91), and also discuss a possible link between number of spatial clusters of pixels and number of individual stems, which is the relevant ecological quantity. A relevant paper related to this discussion is Meyer et al. (2018).**

Answer: We included in the methodology section an improved explanation for the maximum tree height determination. We also included a supplementary figure to illustrate the maximum tree height determination per transect. Here is the added text: *Therefore we selected a single tallest tree per transect using an individual tree approach based on a local maximum filter. For each transect, the largest tree was inspected to exclude spurious maxima not related to tree structure. (Supplementary Figure 1).*

**b/ the results section could be made more quantitative. How many transects had trees > 70 m across Amazonia? What is the overall coverage of trees > 70 m across Amazonia (based on pixels above this height)? I appreciate that the authors do not wish to publish mean statistics for their dataset, as it may be part of a separate paper, but it would be great to plot the distribution of maximum tree height across the 906 landscapes. This is different from Figure 2, and it would be useful to compare raw and modelled results.**

Answer: We included a new figure 1 showing the histogram of the maximum tree height extracted from each transect. We included in new lines 477-482 more quantitative information about the maximum height occurrence. We also included supplementary figures 2 and 3 to improve the assessments of variable importance in our model. Supplementary figure 3 is a scatterplot comparing observed and predicted heights as requested by the reviewer.

**c/ The discussion of table 1 is exceedingly brief (lines 201-203) -- is there a natural way to select the most important variables? The authors choose to report all variable with an importance > 8% (n=4). Does it mean that precipitation seasonality can be safely ignored (importance 6.9%)?**

Answer: It was not our intention to report only the first four variables. We rewrote the relevant paragraph highlighting variable rank and how the variables are contributing to the rank. We also include Supplementary Figure 2 to bring alternative metrics from the random forests model to improve the discussion of the variable importance.

**d/ Figures could be improved. Figure 1 could show the location of all the transects, not only those where the presence of tall trees has been detected. It is difficult to interpret the bioregions as they are depicted in shades of grey. One solution would be to label the bioregion name onto the map itself (e.g., on the bottom-right panel). Also, it would be good to draw the rest of the continent, to aid the non-specialist reader. The legend does not have to be replicated within each sub-panel. Figure 2 would be much more impactful in colour. Here again, it would be important to map other features of South American geography (such as continental and country boundaries). The caption of Figure 3 is cryptic -- please explain more clearly what the lines represent. In short, lines 228-231 should belong in the caption of figure 3. Same comment as for Figure 3 applies to Figure 4. Same comment as for Figure 2 applies to Figure 5.**

Answer: The figures were completely revised following reviewer’s suggestions. We provided links to all the maps through the public repository Zenodo that we will make available should the paper be accepted.

**e/ Could the presence of tall trees detected in the 'pantepui' area North of Rio Negro and West of Rio Branco be due to the steep topography of this area? With 4 laser pulses per m2, it may be difficult to resolve ground location on hilly terrain, which could bias tree height retrieval. Actually, the same comment applies to the main identified region with tall Amazonian trees, around Rio Maicuru and Jarí. This region appears to be quite hilly relative to the rest of Amazonia. The paper should discuss this possible bias, but even If the authors are confident that terrain slopes do not contribute to inflating the occurrence of large trees, it would be also interesting to discuss whether rugged topography could offer the conditions for lower average wind speeds. This could be included near lines 327-335.**

Answer: We included a paragraph in the methodology section discussing the possibility of bias from the terrain model. Here is the new paragraph:

*For each transect we identified the returns from the ground and vegetation. We interpolated ground returns to produce a 1m horizontal resolution digital terrain model (DTM). Using the DTM, we calculated the heights above ground from vegetation returns. The uppermost vegetation heights were then employed to compute a 1 m horizontal resolution canopy height model (CHM). While errors in estimation of terrain height can affect CHM estimations, previous studies in tropical forests show that lidar surveys with at least 4 returns per m2 permit accurate DTM generation and tree height estimation even in complex terrain (Clark et al., 2004; Glenn et al., 2011; Leitold et al. 2015; Andrade et al., 2018).*

**line 184: this study does not include a survival analysis, and it may be better to refer to the occurrence/presence of the tallest trees, not their 'survival'.**

Answer: We changed text throughout to by replacing “survival” with “occurrence”.

**line 247: sentence 'The extraordinarily tall trees had a unique niche' is confusing. Why use the niche concept in this situation, as it is usually associated to individual species, why should this niche be 'unique'? Perhaps a better/simpler wording would be 'The extraordinarily tall trees were found in conditions characterised by a much smaller ...'**

Answer: Suggestion accepted and implemented.

**RESPONSE TO REVIEWER 2 COMMENTS**

**1. The study had two main questions (related to drivers of maximum height and favorable environmental conditions for the occurrence of giant trees) that need to be clearly presented with the background rationale in the introduction. The motivation for using (the 18) different environmental variables also needs to be clear in the introduction (how these variables relate to the availability of resources and disturbances and how these two main factors affect maximum height and occurrence of giant trees).**

Answer: In the new lines 27 to 49 we added two paragraphs to clarify the rationale used to select the environmental layers we included in the analysis.

**2. The methods need a more detailed description of the analyses used. The reader needs to understand the general idea of the analyses (and why they were chosen). All environmental variables must have the spatial resolution presented and those involving historical series must bring the duration. In my view, secondary forests and degraded areas should be excluded from the study a priori. I do not see how the maximum height of these anthropic areas can contribute to answer the study's questions, which are related to mature forests properties and dynamics (without recent anthropic influence).**

Answer: We enhanced the methodology section to detail the analyses used (lines 398-542) and provide the requested justification. We also included in Table 1 the spatial resolution and the time interval considered to generate each layer.

Because it is difficult to *accurately* classify second growth and degraded forests as opposed to intact old growth forest, we retained all areas for a first analysis. We then applied a filter to remove potential disturbed areas (133 transects) (see lines 638-645.) We found that “*The spatial distributions for maximum tree height persisted after removing these potential anthropogenic effects. Variable importance was similar and consistent (Supplementary Table 1)*.” Based on this analysis, we judge that inclusion of potential second growth and degraded areas had no effect on our conclusions.

**3. The results (and discussion) need to bring the perspective of combined environmental variables. What makes two regions with similar annual rainfall have different maximum heights? The final tree of the Random Forest Model (RFM) may help understanding these combinations of factors. I would also like to see a scatterplot with the relationship between predicted and observed values of the RFM.**

Answer: The final model of random forests is not a single tree, but a large number of trees. Therefore, it is extremely difficult to visualize. Rather than attempt to visualize the tree structure, we included supplementary figures 2 and 3 to illustrate the importance of the variables. We will make the R object available for the reader who wants to perform a deeper investigation. In supplementary figure 3 we provide the scatterplot requested by the reviewer.

**4. The discussion needs to be restructured to bring first a summary of the main findings of the study and to focus on the discussion of the results in the light of the hypotheses.**

Answer: The discussion section section now begins with a summary (lines 704-709). The entire discussion was rewritten reflecting the suggestions of both reviewers.

**5- The manuscript would also benefit from careful proofreading - there are numerous language errors (noticed even by a non-native English speaker) that need to be addressed.**

Answer: We have endeavored to improve the language. One of the co-authors, who is a native English speaker carefully edited the paper.

**Line 7: The author should make clear in the abstract the two questions addressed in the study.**

Answer: We included in the abstract the two questions addressed in this study. Here is the inserted text (lines 6-12): *We investigated how resources and disturbances shape the maximum height distribution across the Brazilian Amazon through the relations between the occurrence of giant trees and environmental factors. Common drivers of height development are fundamentally different from those influencing the occurrence of giant trees. We found that changes in wind and light availability drive giant tree distribution as much as precipitation and temperature, together shaping the forest structure of the Brazilian Amazon.*

**Line 6: Why including secondary forests in the study? How they contribute to understand the factors influencing maximum tree height on mature forests?**

Answer: Please see the reply to item 2 above.

**Line 10: "occurrence" is more appropriate since the author did not present data of survivorship/mortality of trees. Instead of "tree giant", "giant trees".**

Answer: We adopted the word “occurrence” throughout the paper rather than use “survivorship.” We also corrected the word order for “giant trees” as suggested.

**Line 18: composition of what? "species composition of the rainforest..".**

Answer: We rewrote the sentence highlighting the comparison of forest structure rather than species composition. Structure is more closely related to the aim of our paper.

**Lines 24-26: Are studies about the seasonality of forest productivity (Huete et al and Morton et al) a good base of comparison to the study of drivers of maximum tree height? I suggest the authors build on previous work more directly related to tree or forest canopy height. How the present study is going to fill a gap in the literature of trees/forest height in Amazonia?**

Answer: Following the reviewer’s suggestion we revised the text to refer to studies addressing forest canopy height.

**Lines 27-28: The authors could be more specific on what sort of environmental conditions beside climate they are referring as "other environmental conditions" limiting maximum tree height. Also, they should explain how climatic and these other factors can influence tree height.**

Answer: In the revised version of this paragraph, we no longer mention “other environmental conditions”. However, we moved the discussion of environmental conditions (such as climate, soil properties, and water)” to the following paragraph (new lines 27-48).

**Line 31: "local scale factors" seems more appropriate than "small-scale factors".**

Answer: We adopted the reviewer suggestion.

**Line 36: "Forest giant trees" instead of "Forest giants".**

Answer: The sentence was removed in this new version of the introduction.

**Line 38-40: Please provide a work as reference for hydraulic resistance.**

Answer: We included three references: Koch et al., 2004; Niklas, 2007; McDowell et al., 2008

**Lines 89-90: It would be helpful to have a figure as supporting information showing maximum height identification in a LiDAR transect.**

Answer: We included supplementary figure 1 to show the maximum height tree identification.

**Lines 167-169: Why use the Random Forest Model approach? Please include a brief explanation of this analysis and what are the advantages of using it for this study.**

Answer: We included in the new lines 399-530, the justification about why we used random forests, and also a brief explanation about how random forests works.

**Lines 171-172: The authors extracted values from environmental layers to the coordinates of the tallest trees in each LiDAR transect. Did the authors use a simple extraction for all layers or did they try to represent some heterogeneity around the tallest tree location. Which were the criteria to make this decision for the different layers?**

Answer: We included in the text that we did a simple extraction for all the layers. Here is the adjusted sentence: *Using the coordinates of the tallest tree within each lidar transect, we performed a simple extraction of the values for all the variable layers.*

**Line 187: Phillips et al. 2006 is missing in the list of references**

Answer: We included the citation in the reference list.

**Line 212: Table 1. I suggest including a column with the expected influence (positive or negative) of each variable on the maximum tree height. Also, inform the reader what is the spatial resolution of all variables used, what is the time resolution and the initial and final years of the time series used (when applicable). Include in the methods section, what is the depth of soil sampling for clay content.**

Answer: We included a new column in Table 1, following the reviewer’s suggestion.

**Line 213: Indicate the parameter in the Random Forest that inform the relative importance of each variable in the Random Forest Model.**

Answer: We specified in the new line 205 that the variable importance was measured by mean increase in accuracy. We also updated Table 1 considering the same importance metric used in the text (i.e. increase in accuracy).

**Because the random forest model is a kind of decision tree (or a collection of decision trees), there should be a hierarchical importance of variables from top to lower nodes. I missed the explanation of this hierarchical combination of variables in the results. For instance, does the region with the higher maximum tree heights present a lower number of clear days, high clay content and elevation above sea level and intermediary annual precipitation? Maybe a graphical output of the final tree (as supporting information) may help to explain that. A panel with maps showing the spatial distribution of the most important variables, together with the map of maximum heights prediction (Fig. 2), may also be a good way to show the combination of factors leading to different maximum heights across Amazonia.**

Answer: The final model from random forests is not a single tree, but a bunch of trees. It is not useful to visualize a single tree. Instead, we included supplementary figure 2 to display variable importance. We will provide the R object for the reader who wants to perform a deeper investigation. Supplementary figure 2 is a graphical summary of the variables included in the final random forests model showing two common measures of their relative importance.

**Lines 219-221: This seems a reverse logic in my opinion. Why use low values of FAPAR to exclude areas of degraded and anthropogenic regions (secondary forests?) instead of excluding these areas a priori? Secondary and degraded forests will present maximum heights lower than it would be potentially in mature and non-degraded forests.**

Answer: See reply to item 2 above.

**Line 244: Figure 3. I suggest order the graphs according the importance of the variable to predict maximum tree height, as presented in Table 1.**

Answer: We implemented the suggestion.

**Line 249: "driver of tall tree occurrence" instead of "...survival".**

Answer: Throughout the text, we replaced the word “survival” with “occurrence.”

**Lines 249-251: Why other variables showing higher probability values (up to ~0.5) in Figure 4 (e.g. Potential evap., Clear days, Days > 20mm) had lower relative contribution compared to elevation above sea level, which present maximum probability values of 0.25?**

Answer: We now explain that the marginal plots do not evaluate the explanatory contribution because the plot is created by estimating the maximum height by changing one variable at a time, keeping others constant at an average value. However, the marginal plots usefully complement the random forests and MaxEnt results. They help us to understand how the environmental variables influence the height estimation. Here is the text included to clarify the marginal plot contribution (line 414): *We visualized the relations of the environmental variables to maximum height using marginal plots, estimating the maximum height by one variable at a time, keeping other variables constant at an average value..*

**Line 258: Figure 4. What is the unit of the probabilities? I suggest order the plots according their relative importance in Maxent.**

Answer: We fixed all the figure 5 panels.

**Lines 266-267: Remind the reader what is your definition of tall trees (> 70m?). Be more specific about what part of the southern Amazon has tall trees. Data on LiDAR occurrence (Fig. 1) or the prediction model (Fig. 2) show tall trees predominantly in the northeastern of Brazilian Amazon (and in the north of subdivision II). Also, Malhi et al 2006 did not show a coincident spatial pattern of basal area as their map show higher values of basal area in forests of the northwestern basin.**

Answer: We implemented the suggestion in the new lines 390-394.

**Line 268-269: Different from what is expected according this pattern the authors showed maximum heights > 60m in areas with dry season longer than four months, in region II.**

Answer: This comment has been excluded from the discussion section.

**Lines 273-286: The message of this paragraph is not clear. How all these components (wood density, soil properties, climate) relate with the hypotheses and explain the results?**

Answer: This paragraph was removed from the discussion section.

**Lines 276-277: ter Steege et al. (2006) Nature, 443: 444-447 present a map of the average wood density of trees which also supports this statement.**

Answer: We included the reference.

**Line 294: The authors say "disturbance rates are far higher". Is it possible to provide a quantitative measure for disturbance differences (frequency or intensity/size) between western and eastern Amazon, according to Espirito-Santo et al 2014?**

Answer: We included a quantitative reference of 3 times higher, based on Espirito-Santo’s results. Here is the new text (lines 803-804): *The spatial distribution we observed also aligned with observed disturbance rates, that are three times higher in the Western compared to the the Eastern Amazon (Espírito-Santo et al., 2014).*

**Lines 300-307: The authors reported in the results a variation in u-wind speed of close to 0 to ~3 m/s. What is the relevance of 3 m/s wind speeds for tree fall? How the wind speed gradient reported here relates with windstorms as described in Negron-Juarez et al 2017? This is an important point to address since wind speed was an important variable to predict giant trees occurrence.**

Answer: As the reviewer is probably aware, accurate gridded data on maximum windspeeds for the region of study are not available. Although the wind speed values are low, they are mean wind speeds and the maxima would be substantially higher. Wind speeds generally follow a Weibull distribution, so the mean is related to the extremes. Predicting the extreme values accurately is difficult, but we assume that large spatial patterns in the means will reflect large spatial patterns in the extremes.

In addition, because the model is very sensitive to wind speed, we ran the model without wind speed and we discussed how the variable importance shifts in the absence of wind layers. All of this is now included in the methods and discussion sections.

**Lines 314-326: How the variation of these physiological, phenological or morphological traits help to explain the relationship between maximum tree height and the number of cloud-free days (or correlated environmental variables)?**

Answer: We reorganized the discussion section. In this new version those ideas are no longer considered.

**Lines 332-335: This needs to be focused on the results of the present study. I see no direct relationship of shifts in canopy chemical traits and maximum tree height variation in the regional scale. Be concise and make the links more clear to the reader.**

Answer: We reorganized the discussion section and chemical traits are no longer considered.

**Line 341: Soil water content was one of the worse predictors of maximum tree height. Is this a result of a poor soil water characterization?**

Answer: The soil water variable employed relates to water content at a fixed water potential. This potential may be far from the field reality and this may make it a poor predictor. Moreover, we admit the possibility that poor soil water characterization leads to the failure of this variable as a predictor. Ultimately, we need better data to evaluate the role of soil water retention on maximum tree height.