I thank the authors for their thorough revisions. I feel they have further improved an already excellent manuscript. I have just a few remaining concerns that I ask the authors to consider and address.

Response: Thank you again for helping improve our study. Please see below our responses.

Lines 88-89 – To make it clearer that the allometric results are compared to those of the machine learning approaches, please change this sentence to read “We compared the results of allometric upscaling to those of three machine learning techniques…”.

Response: Thank you. We clarified this part as suggested.

Lines 136-137 – I wouldn't necessarily expect slope and/or aspect to covary with elevation. A point at a given elevation could feasibly have any slope and be facing any direction. While, it is acceptable that you didn’t include these to variables as covariates (we can never include everything!), the justification you give is not logically sound. Please consider revising this explanation.

Response: Thank you. We removed our explanation about the slope and aspect in the main text.

Lines 140-141 – I’m still not satisfied by your justification for prioritizing one dataset over another when representing a given suite of covariates. More information is needed in the text. You say that you “favoured the database that is known to have higher quality data…”. How did you determine the relative quality of comparable datasets? For example, Is the Shangguan et al. soil dataset better than the new SoilGrids 2.0 dataset (https://soil.copernicus.org/preprints/soil-2020-65/soil-2020-65.pdf). If so, based on what standard? Both datasets appear to contain maps of all the same soil properties.

Response: We did not use SoilGrids 2.0 partly because this dataset does not have all the soil variables we need. SoilGrids 2.0 has bulk density, cation exchange capacity, coarse fragments, nitrogen, pH, organic carbon content, soil texture fractions. Shangguan et al. provides data for more soil variables. In addition to variables listed from SoilGrids 2.0, Shangguan et al. also has, for example, soil phosphorus, soil potassium, base saturation etc. We chose Shangguan et al. as it could allow us to test on more variables in a consistent way. In the final model, we have base saturation as one of the final predictors. SoilGrids 2.0 does not provide this variable. We do not have solid information on which dataset has better quality. Shangguan et al. dataset could provide us a more comprehensive set of soil variables in a consistent way. In the revision, we updated the format of some references to follow the journal style. We changed the reference of Shangguan et al. to Wang et al., 2014 to correct a mistake on first vs. last name. For the reason to choose Shangguan et al. dataset vs. SoilGrids 2.0. We added “For example, we used the GSES soil database (Wang et al., 2014) for the “central” estimate instead of SoilGrids 2.0 (Poggio et al., 2021) as the latter does not have the whole set of soil property variables needed in this study.”

Line 146 – Santoro et al now have a peer reviewed manuscript describing their product that is likely worth citing: https://essd.copernicus.org/preprints/essd-2020-148/

Response: We added the reference

Santoro, M., Cartus, O., Carvalhais, N., Rozendaal, D., Avitabilie, V., Araza, A., de Bruin, S., Herold, M., Quegan, S., Rodríguez Veiga, P., Balzter, H., Carreiras, J., Schepaschenko, D., Korets, M., Shimada, M., Itoh, T., Moreno Martínez, Á., Cavlovic, J., Cazzolla Gatti, R., da Conceição Bispo, P., Dewnath, N., Labrière, N., Liang, J., Lindsell, J., Mitchard, E. T. A., Morel, A., Pacheco Pascagaza, A. M., Ryan, C. M., Slik, F., Vaglio Laurin, G., Verbeeck, H., Wijaya, A., and Willcock, S.: The global

forest above-ground biomass pool for 2010 estimated from high-resolution satellite observations, Earth System Science Data Discussion, https://doi.org/10.5194/essd-2020-148, 2020.

Line 265 – Here you note a range of estimates based on the canopy cover threshold used. Since you have this data, it would seem more appropriate to compare the Spawn et al. estimate (188 Pg) to your 151 Pg (corresponding to a 0% canopy cover threshold) since Spawn et al. describe their estimate as encompassing all trees (i.e. they effectively use a 0% threshold). Thus comparing to 151 Pg rather than 142 Pg would yield a more direct comparison and allow you to say with greater confidence that the remaining difference is strictly due to differences in upscaling methodologies. By not making a direct comparison, you give the impression that difference is greater than it may be actually, which could be misleading. Note that this issue effects comparisons made in Table 1 and the abstract and related text in the discussion section

Response: We modified the comparison with Spawn et al. (188 Pg). Earlier we compared with 142 Pg from our study with a 15% tree cover threshold for forest definition, and we pointed to forest area and upscaling methodology as potential reasons for the differences. In this revision, we compare between Spawn et al. and our estimate with the 0% tree cover threshold for forest definition. We point out the difference is mostly likely due to the upscaling methodology, but we could not exclude out the slight differences in forest (woody tree) definition as Spawn et al. used modified aboveground biomass for Africa and tundra. The revised text read as "Compared to our estimation with the 0% tree cover threshold for forest definition (i.e., 151 Pg root biomass), the 24.5% higher estimation from Spawn et al. (2020) is most likely linked to the upscaling methodology, in addition to the slight difference in the definition of forest (woody) area especially in Africa and Tundra.". We also updated values in abstract and Table 1.

Line 367 – What do you mean by "After accounting for sparse forests"? Please make this more clear in the text.

Response: We rewrote this part without reference to sparse forests. Please see our response above.

Line 367 – Based on the description of this calculation given in the footnotes of table 1, the 32% estimate given here is calculated as a percent *increase*. However, here, you present it as though its a percent *decrease* by saying "32% smaller" which is inaccurate. To illustrate: 188*(1-0.32) = 128 not 142. Conversely 142*(1+0.32) = 188. In reality, the 188 Pg estimate is 32% *greater* than your 142 Pg estimate or your 142 Pg estimate 24.5% *less* than the 188 Pg estimate. (Further, if you instead compare to the 151 Pg estimate as I suggest above, your estimate (151 Pg) is 19.7% less than 188 Pg and 188 Pg is 24.5 greater than your estimate). Please make sure the language surrounding this comparison is accurate both here and in the abstract.

Response: Thanks a lot for pointing out this twist. We updated the language in abstract and the main text. Now we use "larger" to shift the base to estimations of our study.