

General comments

This manuscript describes a new high resolution, global map of root biomass—crucially, one generated by a different approach than most previous studies, that produces a very different (lower) result globally. This is an interesting and important topic, and the resulting dataset will be a valuable resource for a wide range of scientists. The text is generally clear and well written, methods clearly described, and uncertainty and cross validation quantification comprehensive. I have not tried to rerun the code posted on Figshare but applaud its availability, which is crucial for scientific transparency and reproducibility. From scanning through, it looks clear and complete.

Response: Thank you.

There are a few problems. I echo the previous reviewer's point about sampling depths—this should be more clearly described. There also is a key reference and comparison dataset that isn't cited but almost certainly should be: Spawn, S. A., Sullivan, C. C., Lark, T. J., and Gibbs, H. K.: Harmonized global maps of above and belowground biomass carbon density in the year 2010, *Sci Data*, 7, 112, 2020. <http://dx.doi.org/10.1038/s41597-020-0444-4>. Finally, the text has a few unclear or awkward points (see short list below).

In summary, this is a strong and interesting manuscript documenting a valuable global dataset, and I think it will generate much interest in understanding the discrepancies between these and previous results. It needs minor to moderate revisions for clarity and to include the recent Spawn et al. paper and dataset.

Response: Thank you. Please see our responses to previous reviewer's point about the sampling depth. Thank you for referring to this nice recent study. We added the study of Spawn et al 2020 to the abstract, Table 1, and the discussion. Please check below our point-by-point responses to your comments.

Specific comments

Line 31: perhaps "a key role"

Response: We modified the text as suggested.

37: here and elsewhere, I'm puzzled the exclusion of recent Spawn et al. 2020

Response: We added the Spawn et al. 2020 to Table 1, abstract and the discussion. In the discussion, we stated "Spawn et al. (2020) estimated root biomass from shoot biomass and the correlation between root/total biomass and temperature. Among previous studies, the recent study from Spawn et al. (2020) shows the smallest difference from our study. After accounting for sparse forests, the 32% smaller estimation from our study is likely linked to differences in the definition of forests and upscaling methodology" (around Line 365 of the tracked version manuscript).

65-66: I agree this (increasing with time) is interesting, but you never return to this point in the discussion...why do you think this occurred?

Response: We added "likely associated with improved methods in excavating roots that reduce under-sampling" as a possible reason for this increasing with time

180-181: this sentence is awkward and either out of place, or not well connected to the material around it. Rework, probably starting a new paragraph for readability

Response: We deleted “A model with an overall good performance will not guarantee a good prediction on woody plants with higher biomass” to make this part more readable.

188-: the cross-validation step is crucial and I feel like this is a little light on the details. For example, the continental cross-validation is only described in the Figure 9 caption I think; should be here as well.

Response: In the main texts around Line 270 (tracked manuscript), we explained cross validation and the model performance. We added a supplementary Table 8 to compare the performance of the random forest, the allometric fitting and another two machine learning algorithms. We also added “By continents, the performance of RF is worst in Africa ($R^2 = 0.6$; MAE = 44 kg) partly due to limited observations (Supplementary Figure 9)”. Through these changes, we improved the coverage of the cross-validation.

262: probably start new paragraph here

Response: We started a new paragraph as suggested.

Table 1: see #2 above

Response: We added Spawn et al. 2020 to Table 1.

333-: interesting!

Response: Thank you

347-348: this is a good and succinct point; include in abstract?

Response: Thanks a lot for this nice suggestion. We use this example to show, if we use the biome-level mean R:S from observations to estimate root biomass, we would overestimate the total root biomass. Because in observations, sampling is biased towards small trees. It is one of the several reasons that non-spatially explicit method might overestimate the root biomass. After a second thought, we decided not to put it into the abstract to reduce the risk of mis-interpretation of the number 233 Pg, as it requires long text to put the context.

Supplementary I. 96-97: this is crucial detail and should be included in the F7 and F8 captions as well.

Response: Thank you. We added this information to the captions of F7 and F8