Dear Editor:

Below please find our responses to reviews for our manuscript ’ Informing forest carbon inventories under the Paris Agreement using ground-based forest monitoring data’ (PPP-MT-2024-00903). We have addressed all points, as detailed below. The most important changes have been:

1. In order to give a better picture of the representativeness of the submitted data (requested by reviewer 1), we revised two of the figures (global map, histogram of submitted data) and added one new figure (a Whittaker-style diagram showing ForC sites and those with submitted data within the MAT-MAP climate space (new Fig. 3).
2. We have given some more specifics about the process (requested by reviewer 1), as well as specific recommendations (requested by reviewer 2).
3. We have revised the text to make it clearer and less verbose (requested by reviewer 1).

We believe these changes have signficantly improved the manuscript. Thank you for your consideration, and we look forward to receiving your decision.

Sincerely,

Kristina Anderson-Teixeira (on behalf of all coauthors)

# Response to Reviews

## Referee: 1

The manuscript presents a process developed for preparing data for submission to the Emission Factor Database (EFDB) and methods to assess the data in the Global Forest Carbon Database (ForC version 4.0) relevant to EFDB. The variables relevant to IPCC and EFDB were identified and listed. Results describe the databases and analyse some basic elements of the records they contain. Data contribution to ForC is mapped.

The aims of the study are clearly identified and very relevant to the community, and the idea itself of improving scientist’s contribution to international databases appreciable. Thus, there are good reasons to publish.

**Thank you.**

The results section is however disconcerting, partly because it is not clearly structured, and partly because it is quite unclear what point authors wanted to make out of the few graphs presented.

**We have reorganized and reworded to clarify the message and make the section less verbose. We have aimed to present an objective overview of the ForC data contributed to EFDB and the potential for future submissions to improve EFDB data volume and composition. We have also responded to the specific concerns (detailed below).**

The first paragraph is about records with deficiencies, a fraction being “deemed not applicable to EFDB…”. This is not informative at all in absence of a series of well-identified examples. Quite verbose but not that useful.

**We have added a sentence detailing some of the reasons for rejection: “Reasons included – but were not limited to – non-applicability to the IPCC methodology of the variable submitted (e.g., net ecosystem CO2 exchange, litter - OL layer; subsequently excluded from lists of relevant variables and counts of relevant records), inadequate information on the quality of data and thus on its replicability (e.g., confidence interval/uncertainty), and vaguely described disturbance histories (e.g.,”moderately”/ “severely” burned).” We also note that the authors of this study were not privy to the discussions of the review panel and do not have detailed information on all reasons for record rejection.**

**In addition, we have reworded and reorganized this paragraph to make it less verbose.**

Another example, the map (Fig 2) is fine as an illustration or proof of the global participation to the efforts of reporting, but it doesn’t help identify the gaps in the current reporting. Some vegetation types are well represented, others have a serious deficit. Identifying and quantifying this, for instance in terms of report per area, would have been a more eloquent and useful than a map. The imbalance in the data goes for the below/aboveground compartments too.

**We have added a graph to the map (Fig. 2) showing the proportional breakdown by FAO ecozone of (1) area with tree cover, (2) sites and records submitted to EFDB, (3) sites and records submitted to EFDB. We have also added a statement on relative representation by biome: “Relative to forested area, submitted data under-represented the tropics and over-represented temperate regions, reflecting the composition of ForC (Fig. 2).”**

**To assess the data distribution within in the global climate space, we have added a Whittaker-style diagram showing ForC sites and those with submitted data within the MAT-MAP climate space (new Fig. 3). We have also added a statement on climate representativity: “Submitted data were concentrated in climates that are relatively common over Earth’s land area and covered all forested climates with substantial land area except the colder boreal regions (boreal tundra woodland; Figs. 2, 3).”**

**Regarding imbalance among variable types, Table S3 provides a complete record of the numbers of records in ForC and submitted to EFDB. We feel that further analyzing by categories such as ecozone or age would be overkill, make the manuscript too long, and of little interest to most readers.**

Fig. 3 *(now 4)* isn’t conclusive either because ecozones have different sizes, so the number of reports on itself is not a sufficient to evaluate the coverage. The representativity of the data within an ecozone depends on the variability between records: if they all cover the same age classes, again, the representativity is bad.

**To evaluate the coverage, we have modified this figure to include bars representing the relative area (% of global) of each category and added some text summarizing the results/ citing this figure (2nd paragraph in Results).**

**Again, we feel that deeper analysis (e.g., age x ecozone) of the representativity would be overkill and of little interest to most readers.**

Consequently, the “Database needs” as discussed L297 are not convincingly demonstrated by the results which are limited to just frequencies.

**We believe that the changes to the results section and figures now convincingly demonstrate the claims in the discussion.**

Overall the discussion seems somewhat verbose.

**We have reviewed and revised the discussion, removing extra verbosity where possible.**

Some parts of the writing are a bit frustrating, as they suggest that authors have some relevant information that they do not share:

**We’d like to emphasize that all data, code, planning documents, and even our internal discussions are documented in the associated public GitHub repository (**[**https://github.com/forc-db/IPCC-EFDB-integration**](https://github.com/forc-db/IPCC-EFDB-integration)**), as explained in our data availability statement. We have done our best to present relevant information while keeping the manuscript concise and as interesting as possible. We appreciate the reviewer having pointed out places where more information is desirable, and have modified the text to include it.**

L156 “We identified ForC variables that were relevant to the IPCC…” Great, but how? The reasoning behind the identification is certainly of interest to scientists or practitioners.

**We have modified this statement to read, “Working in consultation with the EFDB Technical Support Unit and referencing IPCC guidance (IPCC, 2003, 2019; IPCC, 2006), we identified ForC variables that were relevant to the IPCC methodology and EFDB (Fig. 1, Notes S1).”**

L226 “To submit complete, reviewed ForC records into EFDB, we created R scripts to restructure ForC records and populate EFDB’s bulk import form.” What are those scripts? Why not share these scripts?

**These scripts are publicly available in the associated GitHub repository (**[**https://github.com/forc-db/IPCC-EFDB-integration**](https://github.com/forc-db/IPCC-EFDB-integration)**) and also archived in Zenodo. This is explained in our data availability statement, but that fact is easy to miss. To ensure that this is clear to readers, we have modified this statement to read: “To submit complete, reviewed ForC records into EFDB, we created R scripts to restructure ForC records and populate EFDB’s bulk import form (publicly available; see Data Availability Statement).**

L300. So far, only ~7% of the EFDB-relevant data in ForC have been submitted to EFDB. How do you get that number? What are those 7% and what consequences does this have? Providing concrete examples of mistakes to avoid would have been very appreciable, and made the text sound less theoretical.

**We now clarify what is meant by the 7% in the first sentence of the Results section: “…we had submitted to EFBD only a fraction (~7%) of potentially relevant ForC records (i.e., records for EFDB-relevant variables)…”**

**We have also made several changes to give more information on our process. For example, in the methods (section 3), we now refer readers to Methods S1 for a description of how we prioritized ForC records for review. We also added a description of our iterative process of submission, review, and improvement of our process: “Submissions were sent in six batches between April 2021 and June 2022, and feedback on each batch from the EFDB team was used to improve our system for subsequent submissions.” (methods, section 3)… “The iterative process of review by EFDB’s Technical Support Unit and review panel significantly improved our process, though it did not achieve a 100% acceptance rate.” (results, end of 1st paragraph)**

**We also revised the discussion paragraph on which this comment was focused, adding a sentence to help pu the 7% in context: “These are unevenly distributed across regions, forest types, and variables (Figs. 2-4, Table S3), reflecting the uneven composition of ForC (Anderson-Teixeira *et al.*, 2021) and the fact that our effort was focused on developing an accurate and efficient data submission system rather than optimizing the composition of submitted data.”**

**We now provide some concrete examples of mistakes to avoid in the first paragraph of the results (see above).**

The discussion section “Data reporting needs” should be named “Recommendations of good practices for reporting C stocks and stock changes”, or something more in line with the content itself.

**We have renamed this section as suggested, and have also renamed the other two to better describe the contents (“Potential for forest C databases to contribute to EFDB”, “Forest C data collection and analysis needs”).**

The summary lists up a point on human’s influence on earth’s climate, which is also heavily elaborated in the introduction. The emphasis on human’s influence on forests and climate are disproportionated and distracting (i.e. not appropriate here), since the manuscript is not dedicated to studying it as it solely focuses on scientists reporting C stocks evaluations. A simple statement that reporting C stocks is an obligation but still poses difficulties would be more honest and efficient. Removing the first paragraph of the introduction, which is totally disconnected, would allow to shorten a bit the overall too long introduction.

**We have removed most of the first paragraph, as we agreed that it was distracting. That said, we are retaining key statements on human’s influence in order to emphasize the relevance of this paper to the special issue “Anthromes and terrestrial carbon: from the deep past to net-zero”. This was motivated by a prior request from the editors.**

L239: “The 37% attenuation between records we reviewed and those submitted…” Why “attenuation”?, it seems to be a mere discrepancy, or deficit.

**This sentence has been reworded.**

In conclusion, although the aims and principle seem simple, the results are both vague and expeditive, resulting in a rather unconvincing article. Some efforts to make results more relevant, and some more details on key aspects would help.

**We believe that our responses to the comments above has greatly improved this manuscript. Thank you for your feedback.**

# Referee: 2

This manuscript is a timely and relevant paper about developing a standardized protocol to input forest carbon data into the IPCC emissions database. The authors identified two main problems: the lack of authors submitting their published forest carbon data into the IPCC Emission Factor Database, and some papers not fully reporting all information (pools, fluxes, methodology) necessary to submit in the EFDB). This team has developed a robust and standardized methodology for first aggregating forest carbon data into their already developed (but modified for this new methodology) ForC database, followed by inputting ForC data into the IPCC database. The lack of available, standardized forest carbon data is a real problem as the need to monitor ecosystem carbon stocks and fluxes becomes more crucial in the face of climate change– the methodology developed in this paper has high societal impact to better fill in data gaps in a globally important IPCC database.

**Thank you.**

This paper is high quality and well written with reliable findings. The authors used appropriate techniques for both the ForC database and the methodology for semi-automated data submission into the EFDB. Figure 1 is easy to follow and shows how detailed and robust the ForC to EFDB framework is.

**Thank you.**

I appreciate the authors recommendations in the discussion section (they are all relevant, warranted, and important) but I do think some of these recommendations could be more strongly worded. The authors have described large challenges with populating the EFDB, and have developed a robust solution (and identified other, easier solutions). I would like to see the authors clear and strong recommendations in this section. Some of the suggestions would require substantial increases in funding and infrastructure (line 328-333), which are huge challenges. I think that section has been well discussed in the discussion. However, the next paragraph discusses the lack of variables available because they have either not been calculated or just not presented in papers (although they may have been available to the authors) (lines 334-346). I think recommendations in this section could be stronger and more specific – it is very easy for authors to provide data they have already collected/calculated, or to complete relatively simple allometric calculations. Are there ways this could be required or incentivized (this question may be out of scope for this paper)? A more specific recommendation in the next paragraph (lines 348-362) would also be warranted.

**We appreciate this feedback. We have made some revisions so that the discussion gives more clear and strong recommendations:**

**1. In the section “Potential for forest C databases to contribute to EFDB”, we have added a couple sentences on the practicalities of additional database work: “While submitting records from ForC or other existing databases to EFDB is by far the most efficient approach to expanding forest C data in EFDB, it does require time-consuming database work. As the IPCC is not able to pay for data, such work will require support from research and government institutions, grant programs, and non-governmental organizations.”**

**2. In the section “Forest C data collection and analysis needs”, we have added a some details on how reporting of EFDB-relevant data might be increased: “We recommend that, when possible and relevant to their own study, researchers calculate and report all EFDB-relevant variables – a goal that could be facilitated by development of relevant analytical tools (sensu Chave *et al.*, 2014; Réjou-Méchain *et al.*, 2017; Gonzalez-Akre *et al.*, 2022) and encouraged by scientific journals.”**

**3.In the section “Recommendations of good practices for reporting forest C stocks and stock changes”, we have added a suggestion on how reporting might be improved: “Improved reporting – and potentially data submission to EFDB – could be encouraged by journal editors, reviewers, and funding agencies.”**

# References

**Anderson-Teixeira KJ, Herrmann V, Morgan RB, Bond-Lamberty B, Cook-Patton SC, Ferson AE, Muller-Landau HC, Wang MMH**. **2021**. [Carbon cycling in mature and regrowth forests globally](https://doi.org/10.1088/1748-9326/abed01). *Environmental Research Letters* **16**: 053009.

**Chave J, Réjou-Méchain M, Búrquez A, Chidumayo E, Colgan MS, Delitti WBC, Duque A, Eid T, Fearnside PM, Goodman RC, *et al.*** **2014**. [Improved allometric models to estimate the aboveground biomass of tropical trees](https://doi.org/10.1111/gcb.12629). *Global Change Biology* **20**: 3177–3190.

**Gonzalez-Akre E, Piponiot C, Lepore M, Herrmann V, Lutz JA, Baltzer JL, Dick CW, Gilbert GS, He F, Heym M, *et al.*** **2022**. [Allodb: An R package for biomass estimation at globally distributed extratropical forest plots](https://doi.org/10.1111/2041-210X.13756). *Methods in Ecology and Evolution* **13**: 330–338.

**IPCC**. **2003**. *Good Practice Guidance for Land Use, Land-Use Change and Forestry* (J Penman, M Gytarsky, T Hiraishi, T Krug, D Kruger, R Pipatti, L Buendia, K Miwa, T Ngara, K Tanabe, *et al.*, Eds.). Hayama, Japan: Institute for Global Environmental Strategies.

**IPCC**. **2006**. *2006 IPCC Guidelines for National Greenhouse Gas Inventories, Prepared by the National Greenhouse Gas Inventories Programme, Eggleston H.S., Buendia L., Miwa K., Ngara T. And Tanabe K. (eds).* Japan: IGES.

**IPCC**. **2019**. 2019 Refinement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories. In: Calvo Buendia E, Tanabe K, Baasansuren J, Fukuda M, Ngarize S, Osako A, Pyrozhenko Y, Shermanau P, Federici S, eds. Switzerland: IPCC.

**Réjou-Méchain M, Tanguy A, Piponiot C, Chave J, Hérault B**. **2017**. [Biomass: An r package for estimating above-ground biomass and its uncertainty in tropical forests](https://doi.org/10.1111/2041-210X.12753). *Methods in Ecology and Evolution* **8**: 1163–1167.