Title: Response to reviews

Re: "Carbon cycling in mature and regrowth forests globally" by Anderson-Teixeira, Kristina; Herrmann, Valentine; Banbury Morgan , Rebecca ; Bond-Lamberty, Benjamin; Cook-Patton, Susan; Ferson , Abigail; Muller-Landau, Helene; Wang, Maria Article reference: ERL-109898

REFEREE REPORT(S):

Referee: 2

COMMENTS TO THE AUTHOR(S)

The manuscript provides an important contribution to carbon cycle research. It provides an update to a previously published data base of forest carbon stocks and fluxes (ForC) and analyses the dataset across forests biomes and in relation to stand age. This is a valuable dataset for assessing the current state of the carbon cycle and comparing across biomes. The dataset presents carbon cycle modelers with an opportunity to rigorously compare their models with observations. Figures 2-5, 8-11 are excellent summaries of the dataset. Yet I have some concerns and clarifications that must be addressed.

1. It is not clear what the various sources of data are. Perhaps this is because the authors have already described data sources in previous documentation of the dataset. However, I think some summary is needed for readers new to the dataset. I need to know what I am looking at when I see the various numbers in Figures 2-5, 8-11.

This just requires adding some explanation.

   a. For example, is Fluxnet the source of GPP and Reco? The authors state that "ForC amalgamates numerous intermediary data sets" (line 119) and cite Luyssaert et al (2007). Is that the source of GPP and Reco, or have the fluxes been updated to newer Fluxnet data products? Similarly for the soil respiration data. Is this the data from Bond-Lamberty and Thomson (2010) or has it been updated?

This just requires adding some explanation.

   b. I need information on how Figures 2-5, 8-11 were prepared. Carbon stocks and fluxes such as NPP, litterfall, etc. are probably plot data. There is a detailed discussion of how the data were filtered to remove, for example, disturbance effects (lines 140-149). The GPP, Reco, and soil respiration databases seem to me to be different and may not be collocated with the plot data. Was disturbance filtering applied to the GPP, Reco, and soil respiration databases, too?

This just requires adding some explanation.

   c. If Luyssaert et al (2007) is a source of data, is there circularity in the comparison with latitudinal trends found in other datasets, for which Luyssaert et al is cited (lines 317-318)?

This just requires some re-wording. We are building on/ updating results of Luyssaert.

2. Figures 2-5, 8-11 are excellent summaries of the dataset. Yet there is much unexplained in the figure and the reported data values. No guidance is given on how to use the data, and especially how to resolve discrepancies in the data.

a. For example, Figure 2 lists foliage and woody aboveground biomass in tropical broadleaf forests. These do not sum to aboveground biomass (3.53 + 125.42 = 128.95 Mg C/ha vs 146.69). Nor do coarse root biomass and fine root biomass sum to root biomass (23.15 + 9.29 = 32.44 vs. 21.86). Total biomass (147.96) is not the sum of aboveground (146.69) and root (21.86).

This just requires adding some explanation.

b. Figure 4 for temperate conifer forests has large discrepancies in carbon stocks: e.g., 315.61 Mg C/ha for woody aboveground biomass, but only 167.46 for aboveground biomass and 226.58 for total biomass. The authors have a detailed discussion about lack of closure for root biomass, but do not discuss aboveground biomass. Presumably, this is because the data fall within the closure criteria, but what are we to make of the inconsistent data?

This just requires adding some explanation.

I understand that the data comes from many different sources and so may not always be compatible and that they have large standard deviations, but how are we to interpret and use the various data entries? For example, what has higher confidence: the component variables or the aggregate variables? What is the best estimate for aboveground biomass or root biomass? The authors take great pride in that component variables sum to within one standard deviation of the aggregate variables in all but one instance (lines 225-226, 301-302, 309-310). This is somewhat remarkable! But users of the dataset need guidance on how to interpret and use the data.

This just requires adding some explanation. Note that R2 and R3 have very different interpretations of the "closure"

3. There is a general sloppiness to the manuscript that makes me think the authors did not proofread the manuscript carefully.

This is true. Need to fix all points below. None difficult.

a. Line 58: The authors say GPP is estimated to be >69 Gt C/year, and cite Badgley et al. (2019). This statement is factually correct, but is quite misleading. Badgley et al. "estimate global annual terrestrial photosynthesis to be 147 Pg C/year (95% credible interval 131-163 Pg C/year)". The value of 69 Gt C is less than one-half that reported by Badgley. What are the authors trying to say? Such a large discrepancy does not instill confidence in the other numbers reported in the manuscript.

b. Line 81: "rare exceptions that span regions or continents" is repeated twice.

c. Figure 1: Explain the gray scale for "forest cover". It seems like this refers to gradations of the various biomes, but the figure caption does not provide an explanation.

d. Line 162: Clarify what is meant by "the minimum diameter breast height (DBH) threshold for tree census was <10cm"?

e. Table 1: (i) The definition of GPP refers to NEE, but this in never defined. The preceding entry defines NEP not NEE. (ii) The description of the biome labels is inaccurate. The table uses "TrB" but the note refers to "Tr". The table uses "BoN", but the note refers to "B".

f. Line 210: I am confused by the statement that of the 39762 records in ForC v3.0, 11923 were included in this study. Previously (lines 135-149), the authors described creating "ForC-simplified" with 17349

records. Clarify the difference between ForC-simplified and ForC v3.0.

g. Line 212: The authors refer to ForC, where in the previous sentence they referred to ForC v3.0, and previously used ForC-simplified. Clarify what database is used in the analyses.

4. This is not a review article in the traditional sense that a review is a critical assessment of recent papers in the field of carbon cycle research or identifies future research priorities. Instead, the paper documents a database and shows the utility of the database for carbon cycle research. My comments on the review aspect of the manuscript are:

a. Yes, there is a need for this database; but

b. No, this is not a critical and authoritative review of the carbon cycle; and

c. The authors reference many other publications and datasets, but they do not critically evaluate their own dataset or other datasets.

This requires careful consideration and response

Referee: 3

COMMENTS TO THE AUTHOR(S)

This papers uses a database recently created and updated by the author team – For C – to understand how forest carbon stocks and pools varies across broad biome classifications (e.g. boreal, temperature, and tropical) and stand age. The find that the rate of C cycling is faster in warmer climates, and that many C fluxes and pools increase with stand age (at least up to 100 years of age).

There were several things I really liked about this paper. First, I applaud the author's ambition in creating (and maintaining) this database, which has already been used in multiple high-profile papers. It is a novel idea to curate all the C fluxes and pools together in one virtual location, and the carbon cycle budgets illustrated in Figures 2-5, and 8-11, will likely function as useful "reality checks" against which both empirical and modeling results can be assessed. I also appreciated the focus on understanding how carbon cycles varies with stand age, as this is an important unknown that limits understanding of the usefulness of reforestation as a natural climate solution (among other unknowns). Overall I found the manuscript to be clearly written.

Thank you

However, I also found that many aspects of the paper gave me considerable pause.

First, the paper covers an awful lot of ground. It strikes me that each of the three research questions (bottom of page 4) could easily motivate an entire paper on their own. By attempting to address all three in one (relatively short) manuscript, it is not possible to discuss in any detail the mechanisms and processes that determine the results. There are no process-oriented hypotheses or frameworks against which the results are evaluated.

(*add some broad-brush hypotheses/ frameworks*)

For example, concerning expectations about how various C cycles and pools vary with stand age? There is a fairly extensive literature on this topic ... Odum's classic paper on the topic ("The Strategy of Ecosystem Development, 1969) has been cited thousands of times. But this literature is not referenced or cited in the

current manuscript. Odum's hypothesis suggests that NEP (arguably the most important flux, at least from a climate mitigation perspective) should increase initially with stand age and then eventually decrease as forests continue to mature. However, this framework/literature is not referenced, and the way the results are presented make it difficult to understand whether those expectations were borne out in the data. On that note, I didn't understand why the authors chose to show flux trends explicitly as a function of age up to 100 years (Figure 6), but then group all the forests >100 years old into a single "bar). Why not plot the mature forests explicitly on the flux versus age axis? This would allow a clearer assessment of whether flux trends with age are really linear. It would also help the reader understand better one of the most striking results from this manuscript: that NEP of mature forests is indistinguishable across biomes. The authors describe this as unsurprising, but I think it is a bit unexpected, especially given the results of some of the more synthetic work from FLUXNET (e.g. Luyssaert et al. 2007). The authors explain this result in a couple of sentences (Page 25, lines 22-26), speculating that the result is driver by "moderate disturbances" or "disequilibrium of Rsoil relative to C inputs). A deeper dive into the results, combined with some mechanistic grounding, might reveal to what extent this result represents a climate-age interaction that is predicted from the existing theory.

Regarding reference to Odum, ..., *add citation to Odum*

Regarding the question of showing flux trends as a function of age beyond 100 years, this unfortunately doesn't make sense within the context of the database because tropical forests can rarely be aged beyond 100 years (if that). This is because tropical trees rarely form annual rings, which are used to age (older) extratropical forests. It would not make sense to treat tropical and extra-tropical forests differently.

Regarding the NEP result for mature forests, ... *(Note that R4 also found this surprising. I think an Odum-esque diagram can help clarify.)*

Finally, I wondered about the interacting effects of changing climate (rising CO2, warming temperatures) and stand age, especially in determining trends in the pools. Mature forests will have experienced a much wider range of climate conditions than younger stands. How does this complicate the comparison of live biomass across forest of different age?

This is a good point, but unlikely to have much influence at this relatively coarse scale of analysis. ...

Second, I had some questions about the representativeness of the dataset. While I appreciate that the authors choose to use distinct geographic areas as the unit of analysis (avoiding some issues of pseudo-replication from many observations from a single site), I still wondered about the extent to which the distinct geographic areas were representative of the climate space within each biome. For example, if mean annual climate versus presentation for all boreal, temperate and tropical locations are shown in a scatterplot (for example, using reanalysis data), and then mean annual temperature and precipation of the observations are shown on top, how much of the "climate space" is covered by the dataset?

Issue #57

Third, I found the presentation of the results made it difficult to see clearly the major differences in C fluxes and pools across biomes and age classes. The illustrated C budgets (the majority of the figures, 8 in total) are visually very appealing, but the reader has to do a lot of flipping back and forth to see how any

particular flux or pool varies across biomes and age class. Figure 6 is more synthetic, but each panel is very small and the differences from one group to the next are hard to see. It is also difficult to compare results for young forests (as scatterplots) with the box plots for the mature forests. My advice is to move some of the budgets to SI, and include in the main manuscript more figures that clearly illustrate the most interesting trends with biome, and to allow an expansion of the results in Figure 6 (for example, by first showing scatterplots of all forests, young and old, as a function of stand age), and then perhaps another that is a box plot comparison of mature versus old forests in each biome.

- *move young forest C budgets to SI?*
- *remove maps from age trend figures?*

Regarding the suggestion of showing age of old forests as a scatterplot, this is not possible give unknown ages of most mature tropical forests, as explained above.

Regarding the proposed figure rearrangement, ... *issue #58*

A few other comments:

Page 6, first paragraph: I wondered about the extent to which filtering the data for "managed" affected the results. In the Eastern US, for example, its difficult to find any forests on public land that aren't managed to some extent (for example, through periodic selective harvests), and many of them regenerated from "planted" stands back in the 1930s. I would be curious to know if including "managed" forests substantially altered results.

The binary age classification (two ages, less than 100 years or greater than 100 years) was difficult to accept, as forests that are in the 80-100 year age range are often considered to be mature (at least in the temperate zone). I realize that even with a dataset as rich as For C, data availability will limit stratification into too many bins. Nonetheless, at least for some measurements, I wonder if it's possible to consider a greater number of age classes (for example, young, maturing, and mature).

No changes here; just respond....

We prefer not to separate forests of known age into categories, but rather to represent age as a continuous variable. ...

The consistency check (e.g. do component fluxes/pools sum to within one standard deviation of the aggregate flux or pool) seems like it is destined to provide a favorable assessment of the degree of closure so long as the data within each grouping represent a wide range of natural climatic and soil variability. As long as the aggregate variable has a large standard deviation, the results are likely guaranteed to be "consistent." It seems this metric would benefit from some simulations (perhaps with artificially generated data) to understand exactly how poor closure needs to be at the site-level to generate an inconsistent results when aggregated across sites.

(No! Too much work!)

I'm inclined to place less–not more–emphasis on representativeness: Issue #60

Referee: 4

COMMENTS TO THE AUTHOR(S) General: This paper presents the findings of a rather comprehensive modeling analysis of C fluxes and stocks in the world's major forest biomes. An important strength of this work is its reliance on a fairly large empirical data set to calculate a comprehensive suite of fluxes and stocks to close (or come close) C budgets in these systems. Another unique component of this work is the contrasts of young vs mature forests in each biome. An important contribution of this work is its highlighting of gaps in data (e.g., deadwood) and non-random distribution of empirical data (though this, of course is well-known) and how these factors influence C accounting efforts. Overall, the paper is well written with beautiful figures and the analyses and data sets are sound. I think this work makes an important and timely contribution to fields of C cycle science, forest ecology, and ecosystem ecology and is likely to be of interest to many ERL readers. I have no fundamental concerns with this manuscript. However, in addition to the detailed comments/suggestions below, I think it would greatly benefit the paper if the authors could include some discussion of how the forest ecosystems they characterize here compare to the forest ecosystems that actually exist. The data here represent generally interior forest ecosystems, which of course are incredibly important. But, work over the last 5 years or so highlights the extent to which forest fragmentation influences a large proportion of the world's forests (Haddad et al., 2015). Fragmentation and the creation of edges has been shown to have important implications for C stocks and fluxes with regional and global implications (e.g., (Chaplin-Kramer et al., 2015; Remy et al., 2016; Reinmann & Hutyra, 2017; Smith et al., 2019; Ordway & Asner, 2020; Reinmann et al., 2020: FULL REFS BELOW). Logging and other forms of management also influence a large proportion of the world's forests. I am not suggesting this be included in modeling efforts here, but in placing this work in the broader context of C stocks and fluxes of the world's forests I think it would do the scientific community a great service to more explicitly recognize what is being modeled and perhaps the proportion of the world's forests these data might represent... even if discussed in a qualitative sense.

add discussion of the more intact stands studied here, vs what actually exists. straightforward.

Introduction:

L81: "exceptions" in sentence twice.

text edit

L97: Typo "Since the its most"

text edit

Methods:

L144: If there is a non-trivial proportion of the world's forests (especially in certain biomes) is plantation or planted forests, does removing such plots from the dataset bias the results of a global modeling product?

straightforward–just a bit of text

Results:

L243-245: I think this is per unit forest area, correct? If so, I think it would be helpful to specify that here. Also, the reader should be referred to Figs 2-5, not just 5, right?

straightforward–just a bit of text edit.

L270-271: The wording of the sentence "There were sufficient data to model. . . ." Seems a bit awkward. Should it read ". . . WHICH were also significant. . . "?

straightforward–just a bit of text edit.

Discussion: You might consider reiterating in the first paragraph of the Discussion section that your findings indicate that Temperate Broadleaf Forests are actually the most productive forest biome in terms of NEP. I think this is a surprising finding (we normally think of Tropical forests as being more productive) that warrants attention.

It strikes me that both R3 and R4 find this surprising. I don't. An Odum-esque conceptual diagram + some discussion should help to explain this.

L297: The fact that ForC does not include soil C is important. It is not a flaw in the model or the approach here, but throughout you discuss C stocks, which many would interpret as being inclusive of soil C. I think the authors should consider clearly indicating in the Methods section, and perhaps reiterating in the Results and Discussion, that stocks here are defined as litter layer, biomass, and necromass, but excludes C in soil.

straightforward–just a bit of text edit.

L322-323: As you point out in the Results section, while there are no statistically sig differences in NEP across biomes, there are large differences in the means. Is this an artifact of the data sources used (i.e. distribution and number of sites with empirical data)? Can you speak a little more to this point in this section?

straightforward–just a bit of text edit.

L382: You might consider changing text to "increases with age AT LEAST up to the 100-yr threshold examined here" so that it does not come across as suggesting that NEP only increases for the first hundred years of stand development, which we know is not true.

straightforward–just a bit of text edit.

L445-450: I come back to the data in Figs 2-5 and text in the results section (L243-245), which indicates no sig differences in NEP, but highest means in temperate forests. Of course, the high biomass in tropical forests makes them critical to protect from a C storage perspective, but if NEP (i.e. rates of C sequestration) are highest, at least as a mean, in temperate broadleaf forests how should those ecosystems factor into conservation priorities. Related to this point, is the high NEP of temperate forests driven by the relatively young nature of temperate broadleaf forests in the eastern U.S., where a lot of data exist?

straightforward–just a bit of text edit.

References mentioned above:

Chaplin-Kramer R., Ramler I., Sharp R., Haddad N. M., Gerber J. S., West P. C., . . . King H. (2015). Degradation in carbon stocks near tropical forest edges. Nature Communications, 6, 1–6.
https://nam02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.1038%2Fncomms10158&data=04%7C01%7Cteixeirak%40si.edu%7Cd27c53d6ec434c8ae42108d8a03017cf%7C989b5e2a14e44efe93b78cdd5fc5d11c%7C0%7C0%7C637435475707545147%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=maYY3mHl%2F%2FX0jgkAOK5Kc9K1TpaoTVkFM52HsOxeMjU%3D&reserved=0.

Haddad N. M., Brudvig L. a., Clobert J., Davies K. F., Gonzalez A., Holt W. M., ... Townshend J. R. (2015). Habitat fragmentation and its lasting impact on Earth's ecosystems. Science Advances, (March), e1500052. https://nam02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.4028%2Fwww.scientific.net%2FAMM.315.108&data=04%7C01%7Cteixeirak%40si.edu%7Cd27c53d6ec434c8ae42108d8a03017cf%7C989b5e2a14e44efe93b78cdd5fc5d11c%7C0%7C0%7C637435475707545147%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=1L35xf2RdBj6Y9PjsYhSY5tct4RhnJLE8Jd50QJdzBM%3D&reserved=0.

Ordway E. M. & Asner G. P. (2020). Carbon declines along tropical forest edges correspond to heterogeneous effects on canopy structure and function. Proceedings of the National Academy of Sciences of the United States of America, 117(14), 7863–7870. https://nam02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.1073%2Fpnas.1914420117&data=04%7C01%7Cteixeirak%40si.edu%7Cd27c53d6ec434c8ae42108d8a03017cf%7C989b5e2a14e44efe93b78cdd5fc5d11c%7C0%7C0%7C637435475707545147%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=DnTMc4LCIoRcHMLUEAcpvNOn%2FeJLO3%2Bh8pXVYSUb%2BI4%3D&reserved=0.

Reinmann A. B. & Hutyra L. R. (2017). Edge effects enhance carbon uptake and its vulnerability to climate change in temperate broadleaf forests. Proceedings of the National Academy of Sciences, 114(1), 107–112. https://nam02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.1073%2Fpnas.1612369114&data=04%7C01%7Cteixeirak%40si.edu%7Cd27c53d6ec434c8ae42108d8a03017cf%7C989b5e2a14e44efe93b78cdd5fc5d11c%7C0%7C0%7C637435475707545147%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=GncJJ%2FLa2f96c0tjkl5RZZv2JLJFoAoQnOTm0b%2BYyNo%3D&reserved=0.

Reinmann A. B., Smith I. A., Thompson J. R. & Hutyra L. R. (2020). Urbanization and fragmentation mediate temperate forest carbon cycle response to climate. Environmental Research Letters, 15, 114036.

Remy E., Wuyts K., Boeckx P., Ginzburg S., Gundersen P., Demey A., ... Verheyen K. (2016). Strong gradients in nitrogen and carbon stocks at temperate forest edges. Forest Ecology and Management, 376(2016), 45–58. https://nam02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.1016%2Fj.foreco.2016.05.040&data=04%7C01%7Cteixeirak%40si.edu%7Cd27c53d6ec434c8ae42108d8a03017cf%7C989b5e2a14e44efe93b78cdd5fc5d11c%7C0%7C0%7C637435475707545147%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=u2ylmC4K4VIxHWamrCnk%2BNfu%2FsZDUkSPfABgfN%2BCzT4%3D&reserved=0.

Smith I. A., Hutyra L. R., Reinmann A. B. & Thompson J. R. (2019). Evidence for Edge Enhancements of Soil Respiration in Temperate Forests Geophysical Research Letters. Geophysical Research Letters, 46, 1–10. https://nam02.safelinks.protection.outlook.com/?url=https%3A%2F%2Fdoi.org%2F10.1029%2F2019GL082459&data=04%7C01%7Cteixeirak%40si.edu%7Cd27c53d6ec434c8ae42108d8a03017cf%7C989b5e2a14e44efe93b78cdd5fc5d11c%7C0%7C0%7C637435475707545147%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C1000&sdata=92jjRZSUzQrNX6SGlV%2BZRfHlUN%2B6AL96%2FJIl4PIK3a8%3D&reserved=0.

References