REPORT #2

General Comments:
This manuscript assesses key components of the terrestrial carbon cycle—GPP, NPP, NEP, Ra, and Rh—over South America, with a focus on the Amazon and savanna biomes. Using output from 18 Earth System Models, the authors compare carbon fluxes across models and against literature benchmarks (e.g., CARDAMOM), and also examine differences between 'dry' and 'wet' years. The topic is relevant and the multi-model perspective is potentially valuable, particularly for carbon cycle assessments in tropical ecosystems. However, the analysis and presentation are currently too limited and underdeveloped. Significant restructuring and clarification would be required to reach publishable quality. While I hope the comments provided are useful for future revisions, I unfortunately recommend rejection in the current form.
[Response] Thank you for taking the time to thoroughly review our manuscript and provide important feedback. We will undertake significant restructuring and clarifications in a revised version of the manuscript, to be resubmitted.
Detailed comments:
Structure and framing of the manuscript:
It's not entirely clear what the main 'results' of this paper are and often the framing changes throughout. For instance, I would suggest the novelty of this paper comes from the attempted assessment of ***how well models are able to simulate carbon components in South America & its biomes, and how the models compare against one another for the region and in wet/dry years.*** Yet the abstract doesn't mention the outcome of such an assessment but rather stating the patterns of the temporal evolution (which could be argued, isn't really anything new).
[Response] Thank you for your comments on how to make our work more relevant. We agree. As you noted, our focus is indeed the South America and its major biomes. The topics you have mentioned will be addressed in a revised version of the manuscript, to be resubmitted, is special in the Introduction and Abstract.
On a similar note, the introduction lacks a **justification for the research presented**, what are the research gaps and why did the authors feel that this work should be done? What is the novelty? The paragraph starting at L101 under Earth System Models is trying to argue this I think, but this should be earlier on in the manuscript and argumentation improved. The several lines on the 2024 burning seems somewhat unnecessary, more context should be added (eg. This shows vulnerability and fast changes that can occur in terrestrial ecosystems and their carbon cycles etc) or **it should be removed.**
[Response] Thank you for your comments and suggestions. We will enhance the justifications for our research earlier in the text, in the Introduction, and add more context to the text mentioning the 2024 fires, which were mentioned to provide a recent example on the importance and potential impacts of the carbon fluxes from the study region on atmospheric composition

globally.
Results/Discussions and Conclusions could be better organised. By the conclusion, several points have been repeated numerous times.
[Response] Thank you for your comment on this section of the manuscript. We will revise the text and remove repetitions.

Models: **More information is needed on the models presented throughout this study and likely some further analysis is required on the following points:**
18 ESMs are used according to data availability. However, the models selected greatly vary in their set up and functionality, especially related to carbon dynamics and terrestrial processes, which will be influencing the results you show and discuss. For example, the EC-Earth3-Veg model used is a configuration that does not include ocean biogeochemistry and therefore does not have a fully coupled carbon model activated whilst other models such as CESM2 does have a fully coupled carbon model on. On the other hand, I believe that in the French IPSL-CM6A-LR model, ORCHIDEE does not have the dynamic vegetation scheme activated, which limits PFT responses and hence carbon dynamics. Whilst it is completely understandable to not have descriptions of each of the ESMs, Table 1 should be expanded to have more details of differences in the models or something similar added to SI. I also think there needs to be a discussion on how these differences impact the results added.
[Response] Thank you for the suggestion. We will expand and include more information about the models in the supplementary material. We will add more details about each Earth System Model, and this text will provide further substantiation for Table 1.
More details should be added about what CARDAMOM is and how it works. Looking at late figures in the manuscript (e.g. Fig.8), it looks like CARDAMOM disagrees with other estimates in literature. As a reader I therefore have no indication about how ‘good’ CARAMOMs outputs are. Therefore, a couple of sentences about CARDAMOMS performance or sources for CARDAMOM evaluations would help.
[Response] Thank you very much for the suggestion. We plan to include a section providing further details on CARDAMON, as well as on the other research sources shown in the manuscript figures.
Simulation details: More details on the simulations that generated the ESM output in this study would be beneficial under the Earth System Models/Methods section. For example, what data was used to drive land surface change in the models? This would help later in the discussions of land use change impacts on NBP – do the model simulations include the land use changes discussed from L396 onwards? How?
[Response] Thank you very much for the comments. We will add more details on the simulations as you have suggested.

Biome definition:
How are the biomes presented in Figure 2 defined? Did the authors select the regions based on observations/literature/models? Figure 2 looks like perhaps it is from CARDAMOM - this should be added to the methods and figure caption.
[Response] Thank you for your observation. Both, the Amazon and the South American Savannas, were defined based on published literature (e.g. Olson et al (2001), Eva et al (2005), Cardoso et al (2017) and Stier et al. (2020)), and official geographical surveys from the Brazilian Institute of Geography and Statistics, cited in the text. We will revise and improve the explanation on the biome regions definition in section 2.3, “Study area”.
The next related question is, how are these regions implemented in the analysis. Presumably each model will have slightly **different simulated regions for the 2 biomes and also different PFTs defined within them which then influences the carbon fluxes** and stocks calculated in the results. Some spatial plots of the models either in the results or as an SI figure are needed to see how spatially different the vegetation in the models are and how well they align with Fig 2. For instance, dominant PFT in each gridcell or vegetation biomass.
[Response] Thank you for your observation. For the development of this research, all model outputs and study region definition were first interpolated on a common regular grid with 1°x1° lat/lon spatial resolution, aligning all datasets in space and allowing us to filter for each biome in separate. We will improve the text to add more details on how the datasets were harmonised before analysis.

NEP: The sentence defining NEP/NEE (L78) is worded in a way that suggests NEP is always related to forests, when perhaps the authors meant just the net flux of CO2 from a given ecosystem? It should also be made clear how NEP was calculated in the analysis, i.e was it calculated by looking at the net ecosystem flux for given spatial definitions (e.g masks such as in Fig.2) or is it calculated only from gridcells that contain ‘forests’ in the model outputs (for instance where tree cover >60% or similar). How is NEP calculated for the whole of south America (Fig.3)?
[Response] We sincerely appreciate the observations. NEP was calculated by subtracting Rh from NPP at each grid cell of the ESMs. Additional details on NEP and NEE will be provided in the manuscript.
L152: In your selection of Wet/Dry years, the author writes that comparing observational years with simulation meteorological conditions, the results were mostly inconsistent. However, this is what is expected with ESMs. Models have their own internal variability (natural, chaotic fluctuations in the climate system that occur even under the same external forcings) and therefore timing of specific events, such as wet and dry years is often don’t align with observations. It is therefore good practice not to compare observations with specific years in most ESMs. I would change the lines 152-155 to explain this and frame your approach

accordingly.
[Response] We greatly appreciate the comments. This explanation will be added to clarify the selection of Wet and Dry years.

Figures: In general, the size of figures should be increased relative the legend text size. The horizontal line around zero (yellow dashed line) should be explained in the captions, changed colour or removed.
[Response] Thank you very much for the comments. In the revised manuscript, we will enhance the size and resolution of the figures, and we will also change the color of the line near zero.
Fig 6: Where does the big range in carbon pools come from? Again, this could maybe be explained by spatial vegetation plots.
[Response] Thank you for the comment. In addition to the differences in areal extent of vegetated land, the lines in Fig. 6 may also reflect differences in vegetation types and treatment of disturbances, resulting in more/less biomass and carbon stocks. As you have suggested, vegetation maps can help understand these substantial range in the models results. We will discuss this point in a new version of the manuscript, and perhaps add supplemental maps of vegetation types.

Wet/Dry years: Why did the authors choose not to do the wet/dry years analysis also for the whole of South America as they did with the temporal fluxes?
[Response] Thank you for the question. We plan to incorporate this analysis in the revised version of the manuscript.
Section 4.3: One could argue that the model results do not match well with the reference data in all cases. For example, in Figure 10. GPP and Rh reference data for the Amazonian do not match with the ensemble of models, yet the authors indicate otherwise. Perhaps a better explanation is needed here.
[Response] Thank you for your comment. We agree. Currently, the text do not provide details on the agreement between models and observations for all variables, and need improvements. In the revised version of the manuscript, we will address your observation.

Technical comments:
L45 lacks a comma after factors. I would also avoid the use of ‘probably’ here.
[Response] Thank you for the suggestion, we will correct it.
L393: EC-Earth3-Veg is listed as a model whose coupled surface model is CLM which is not the

case.
[Response] Thank you for the comment. We agree, and will fix that in the manuscript revision.