General comments:
This article explores the carbon allocation with different choice of scenarios, SSP1-1.9, SSP1-2.6, SSP2-4.5, SSP3-7.0 and SSP5-8.5, at three different global warming levels (2, 3 and 4 degree celsius). Authors comprehensively include a wide range of ESM outputs and design a quantitative analysis framework to calculate carbon fractions in different reservoirs. The current version of the manuscript matches the scope of ESD and the presentation of methodology is enough. However, the main finding from this manuscript is not clear to me. Authors also need to heavily revise their results and discussion sections to provide logical and robust analysis and cross validation and comparison to previous studies.

Specific comments:
I have the following major comments:

1. You have some discussions about the implication of your study on the future carbon management and relevant studies in the discussion section, which is good. But the same information in the introduction part is missing. It would be nice to see more introduction about how carbon allocation is important for relevant research. For example, 1) how the calculated parameter can be important to the next stage of model intercomparison, benchmarking and 2) if this parameter can be helpful to indicate the strategy of carbon management for the next stage.

2. Contents in Results and Discussion sections are stacked in a whole block and require more revisions to streamline your manuscript structure. Please summarize 2-3 subtitles and split your context and fill in these sub-sections.

3. Line 231: “In summary, fig. 3 shows that a model’s sensitivity to CO2 concentration significantly affects the total carbon allocation between the atmosphere, ocean and land at global warming levels, but is less impactful on the percentage allocation……the scenario has a much larger impact on the percentage carbon allocation at a given warming level than the ECS.” But as I found in fig. 3, the carbon allocation fraction after normalization (left pane) are quite similar to each other under different scenarios at least for GWLs at 2 and 3 degree celsius. To the opposite, certain models show very large discrepancy, e.g. EC-Earth3-CC compared to other models. Please explain how you get this conclusion?

4. My understanding is that the authors plan to use UKESM as one of the examples to help understand how different processes in ESMs can influence the calculated carbon fraction. But I only find qualitative speculation instead of quantitative analysis. For example, in Line 340, “The UKESM1’s higher AF at the year 2100 is likely due to the model limiting carbon uptake more than the other models. This could be Nitrogen limitation in the land surface or could be due to the model's higher ECS and thus warmer temperatures at 2100 than the multi-model mean.” I expect to see more analysis, figures or tables to list evidence and prove these statements. Otherwise, there’s no need to specifically highlight the result from one model and these conclusions from this manuscript are not robust.

5. In the discussion section, the manuscript lacks enough cross-validation or comparison against other similar published studies. There are published studies discussing carbon storage, residence time and feedbacks in land and ocean components under different future scenarios. Just to name a few here:

Friend, A. D., Lucht, W., Rademacher, T. T., Keribin, R., Betts, R., Cadule, P., et al. (2014). Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO2. Proceedings of the National Academy of Sciences, 111(9), 3280–3285. https://doi.org/10.1073/pnas.1222477110

Jiang, L., Yan, Y., Hararuk, O., Mikle, N., Xia, J., Shi, Z., et al. (2015). Scale-Dependent Performance of CMIP5 Earth System Models in Simulating Terrestrial Vegetation Carbon. Journal of Climate, 28(13), 5217–5232. https://doi.org/10.1175/JCLI-D-14-00270.1

Katavouta, A., & Williams, R. G. (2021). Ocean carbon cycle feedback in CMIP6 models: contributions from different basins. Biogeosciences, 18(10), 3189–3218. https://doi.org/10.5194/bg-18-3189-2021

6. Your key findings are not properly highlighted. To improve this draft, authors need to conclude a more solid and informative key finding, for example, “choice of forecast scenario impacts the carbon allocation at the same global warming levels more than model’s ECS/TCRE”. At the same time, provide more qualitative analysis to prove your key findings.

Technical corrections and minor comments:
Line 25: “and the land surface via primary production”. Here “primary production” can be replaced by “terrestrial carbon fixation”.
Line 27: “known as carbon allocation”. To avoid confusion with the “carbon allocation” widely used in terrestrial ecosystem modeling, I would suggest clarifying this point here, such as “known as carbon allocation in the Earth Systems (we simply use carbon allocation in the rest of the text)”.
Line 92: “land use emissions” contains how many different components? This LUE calculation may not contain the feedback from the settings of different ensembles.
Line 125: “can gives” shall be “can give”
Line 131: “Individual component models can be used by” can be clarified as “Same Individual component model can be used by”.
Line 137: Please clarify “All quoted values”. What are these values?
Line 148: “In addition, several models may share contributing component models” seems to be a repetition of the content in Line 131. Shall think about how to merge them.
Line 165: “These tools include quick ways to standardise, slice, re-grid, and apply statistical operators to datasets.” Can you provide a table or figure to summarize and explain the mathematical algorithms of the operators you applied in this paper through using ESMValTool for data pre-processing? I think this is necessary information to understand your methodology.
Line 193: “Figure 2 only shows the multi-model means, not single models.” It will be helpful to add the spread of carbon allocation fraction using the results from single models in figure 2.

Line 302: “Therefore, SSP3-7.0 can reaches” shall be “reach”.
Line 302: “Therefore, SSP3-7.0 can reaches the GWLs earlier than other scenarios at the same CO_2 concentration”. I’m not quite sure about this conclusion. If we take a look at figure 4, SSP3-7.0 is later than SSP5-8.5 to reach all 3 GWLs.
Line 315: “Higher CO_2 is causes” shall be “Higher CO_2 causes”
Line 323: “the rate at which surface waters and dissolved CO_2 is mixed downward will slow. This reduction is downward mixing reduces the overall absorption rate of CO_2 into the ocean” This statement is confusing. Please rephrase.

Figure 1: It’s better to clarify that your prescribed DCO_2 has accounted for the anthropogenic fossil fuel exploitation and the subsequent C emission from application.
Figure 4: “the historical observations from Raupach et al. (2014) & Watson et al. (2020),” It will be better if you can clarify in which year(s) these observations represent.

There are plenty of other typos and confusing statements in this draft and the authors shall be responsible to double check the whole document before resubmission.