Review of "Carbon-concentration and carbon-climate feedbacks in CMIP6 models, and their comparison to CMIP5 models" by Arora et al.

General comments:

This manuscript provides an update on carbon cycle feedback parameters in Earth system models based on simulations conducted under the CMIP6 framework. This is a useful manuscript that demonstrates how feedback metrics have evolved from CMIP5 to CMIP6 models and provides explanations for these differences. However, the manuscript is very long, and inconsistent in its quality. A few sections are merely descriptive and do not offer much insight into the reasons for inter-model differences, whereas other parts are missing important details (see specific comments). Also, I am not convinced of the usefulness for this manuscript of the analysis that decomposes feedback parameters into various terms (section 4.4). The decomposition has meaning to experts in the respective land and ocean biogeochemistry disciplines, but in my view has little value to readers not familiar with the concepts of one or either discipline. I suggest that the authors consider removing these analyses from the manuscript and preparing separate disciplinary submissions. This would free room to strengthen other aspects of the manuscript. For instance, the abstract concludes that the approach of calculating the feedback parameters from COU and BGC simulations is "most relevant", yet there is no discussion of this point in the main manuscript. Also, an interesting result is that despite the fact that several models now include N limitation, the model-mean land carbon uptake has increased relative to CMIP5. This is an important result, which in my view warrants more attention in the results and conclusion sections.

Specific comments:

Abstract, l. 48-49: How are they different?

Abstract, l. 50-51: This conclusion is not supported by the discussion in the manuscript. There is discussion on this point in the Appendix that should be elevated to the main manuscript if this point is to be kept in the abstract.

Abstract, l. 51-55: Report interesting findings rather than methodological approaches.

p. 4, l. 97: "Offers several benefits": which specifically?

p. 4, l. 116: "Comparison is useful". I don't think comparison between A13 and F06 is particularly useful. The two studies differ with regard to a number of key assumptions that make resulting feedback metrics hard to compare. In addition to different scenarios these include emissions versus concentration-driven runs and different methods to compute gamma.

p. 6, l. 142: Section 2 title: I suggest to rename the title to "Feedbacks *metrics* in the coupled climate-carbon system". The current title suggests a section on processes, which it is not.

p. 8, l. 183: Even though c' does not appear in Eq. 1a I suggest to mention that c' is the same for RAD as well.

p. 9, l. 192-193: Unclear what is meant by "evolve over time" (from one year to another or from CMIP5 to CMIP6?).

p. 10, l. 225: Include reference to Zickfeld et al. (2011). The paper provides a detailed analysis of non-linearities in the coupled climate-carbon cycle system.

p. 11, l. 234-235: First part of the sentence is repetitive.

p. 11, l. 237: How is it different, i.e. is it larger or smaller?

p. 13, l. 276-277, "explicitly considering…": contradictory. Needs clarification.

p. 17, l. 361-362: Need to say that changes in biological carbon inventory are assumed to be small.

p. 19, l. 407: How is the function f defined?

p. 20, l. 415-416, "Do slightly affect": Can this be quantified?

p. 21, l. 450-452. This statement is confusing. After reading it I thought that the additional figure panels show CMIP6 results for the subset of models that was used in CMIP5, but from the figure captions I gather that those panels show the CMIP5 results from A13. Please clarify.

p. 22, L. 470-472: Mention that the CMIP6 model ensemble includes some high climate sensitivity models.

p. 23, l. 481-482: "fitting a polynomial"? Justify why you chose to do this. Fitting procedure needs to be described in the Methods section.

p. 23, l. 489-490: It is not intuitive why temperature in the RAD simulation is sensitive to inclusion of NorESM2-LM whereas land and ocean carbon fluxes are not. Please explain.

p. 24, l. 512, "has not meaningfully declined". The bottom panels in Fig. 2 suggest that it has actually increased.

p. 25, l. 523-524: How about changes in ocean circulation?

p. 25, l. 526-540: I don't find this pargarph particularly useful as it merely describes what is evident in the figure. I suggest to either include an explanation for inter-model differences, or delete the paragraph. The figure could then be moved to the Appendix.

p. 26, l. 543: Which simulation – RAD? The factor is lower (~two) for the COU and BGC simulations.

p. 26, l. 552-553: Need to clarify that DC' refers to a change in a reservoir. As such, DC'_A is not the atmospheric growth rate (PgC/yr) but the change in atmospheric carbon burden (PgC).

p. 26, l. 555: Which equation/section of the Appendix? Could also refer to Eq. (18).

p. 27, l. 564: Which equation/section of the Appendix?

p. 28, l. 588: It should be emphasized that the difference between models with and without representation of the N cycle is much smaller than in CMIP5.

p. 28, l. 595-596: It would he helpful to have a brief explanation of the increase in CO2 fertilization effect in CanESM5.

p. 28, l. 602-603: I suggest to remove quantitative information from this and the subsequent paragraph (not needed).

p. 29, l. 611 – p. 30 l. 628: The discussion would be easier to follow if differences in land models were first discussed, followed by a discussion of differences in ocean models.

p. 28, l. 592 – p. 30 l. 628: It would be worth emphasizing (here and in the conclusions) that with implementation of N limitation in several models land carbon uptake increased in CMIP6 relative to CMIP5.

p. 30, l. 641-646: Avoid repeating information from the figure legend in the text.

p. 31, l. 650-652: This is not immediately evident from the figure (e.g. beta_L calculated with RAD-COU differs from that calculated with other approaches for CMIP6 models). What measure was used to quantify the sensitivity?

p. 31, l. 656-658: It could again be noted that difference between models with and without N limitation is smaller than for CMIP5.

p. 31, l. 663: It is worth mentioning in my view that the spread in feedback parameters for models with and without N cycle has widened compared to CMIP5.

p. 32, l. 680, "existing studies": Provide references.

p. 33, l. 703-707: Avoid repeating information from the figure legend in the text.

p. 32, l. 680-681: The preferred use of COU and BGC over other approaches to calculate the feedback parameters is a conclusion highlighted in the abstract, yet the discussion is relegated to the Appendix. If the conclusion is to be kept in the abstract, the text in A2 should be elevated to the main manuscript.

p. 34, l. 722-724: Suggest to delete quantitative information in parenthesis (not needed).

p. 37, l. 772-773, "This is one of the few times…": Include references.

p. 38, l. 798-801: Unclear why this needs to be stated upfront.

p. 46, l. 961: climate response to cumulative *carbon* emissions

p. 47, l. 977-978: Is the increase in the mean value of the TCRE since CMIP5 due to changes in TCR, diagnosed emissions or both?

p. 47, l. 986-987, "representation of the nitrogen cycle is helpful in reducing this uncertainty": unclear what results this statement is based on.

p. 47, l. 993-p. 48, l. 1003: Several studies have explored the decomposition of TCRE into various terms and their contribution to TCRE uncertainty. Given that a comprehensive discussion of this literature is out of scope here I suggest to delete this paragraph that is based on a single study.

p. 50, l. 1038, "… a reduced spread across land models". I don't think this is a correct characterization of the results. The CMIP6 models including N limitation have a smaller spread than the models without N limitation (Fig. 6) but the overall spread is not reduced compared to CMIP5.

p. 50, l. 1052-1055: Again, the manuscript lacks discussion supporting this conclusion.

p. 59, Fig. 5 caption: Equation references need to be corrected.

p. 62, Fig. 8: The upper panel is reproduced in Fig. 9, so this figure could be cut.

p. 65-66: Figs. 11 and 12 could be combined. Also, the figure caption needs to draw attention to the different vertical scale used in the panels.

<u>Editorial comments</u>

p. 3, l. 61-63: Style of sentence could be more fluid

p. 4, l. 83-84: Delete text in parenthesis

p. 4, l. 97: Delete “of course”

p. 15, l. 333: Delete “some”.

p. 32, l. 680: Delete “and” after “7”.

p. 34, l. 714: Sentence unclear.

p. 36, l. 748-749, “While …. Feedback over land (β_L),”: Delete

p. 48, l. 1011-1014: Sentence is convoluted. Delete “that allows….”

p. 49, l. 1027: Replace “but” with “and”.

p. 49, l. 1029: Delete “very”.