Response to Anonymous Referee #2

Authors quantify carbon-concentration and carbon-climate feedback for negative emissions for an idealized scenario and compare the magnitude of these feedbacks for the positive emissions part of an idealized scenario. The manuscript is relatively well written and in principle it all makes sense. However, I would suggest improving the manuscript in the following ways.

Please include equations in the main text that should clarify your methodology (where you subtract the effect of zero emissions run on quantities considered during the rampdown phase). If a picture is worth 500 words, an equation is worth at least 200 words.
In the absence of the equations, it is difficult to understand your methodology

- *We thank the reviewer for their comments and positive feedback. We agree that including equations for our proposed approach would make our methodology easier to follow. We will include the equations for our proposed approach in the main text, along with the carbon cycle feedback framework in our supplement.*

Please introduce your sign notation in the beginning and then use it consistently throughout the manuscript. Recall that carbon-concentration feedback is negative from the atmosphere's perspective because it reduces atmospheric CO2 If you use the term "when carbon is gained" then please clarify explicitly which component is gaining carbon - land/ocean or the atmosphere.

Near lines 308-313, I was confused with the sign notation even more because it seems, as you interpret it, sign notation reverses during the ramp-down phase. This needs to be better explained because I am unable to understand why sign notation reversal is needed. If carbon-concentration feedback is negative from an atmosphere's perspective (let's say a value of -1.0 Pg C/ppm) this implies that an increase in atmospheric CO2 concentration will be reduced from its initial amount due to this negative feedback. The corollary of this is that if atmospheric CO2 is reducing then the change in CO2 is negative (say -2 ppm) which when multiplied by -1.0 Pg C/ppm yields +2.0 Pg C implying 2 Pg C is added to the atmosphere. All this makes sense in my mind. So why is reversal of sign notation needed?

- *We thank the reviewer for highlighting this – we will address both comments here. We agree that the concentration-carbon feedback remains a negative feedback under negative emissions. Likewise, the climate-carbon feedback remains a positive feedback. What changes here is* ***the meaning of the sign of the feedback parameters****. The concentration-carbon feedback parameter is calculated by rearranging equations 3.3 and 3.4 for the land and ocean respectively. Hence, the equations are then expressed as:*

$$[3.3.1] \quad \beta_L = \frac{\Delta C_L}{\Delta C_A} \qquad [3.4.1] \quad \beta_O = \frac{\Delta C_o}{\Delta C_A}$$

- *Under positive emissions, CO_2 concentration increases (positive denominator) and both the land and ocean gain carbon (positive numerator) resulting in a positive feedback parameter for both land and ocean. Under positive emissions, a positive β_L and β_O is associated with land and ocean carbon gain. Under negative emissions, when both numerator and denominator decrease, β_L and β_O remain positive as the reviewer correctly stated, but now, a positive feedback parameter is associated with land and ocean carbon loss.*
- *We also agree that the phrase "carbon is gained" is unclear, and we have clarified all instances of "carbon is gained" and "carbon is lost" with where it is gained or lost from.*
- *Lastly, we will add a paragraph introducing the sign convention to the carbon cycle feedback framework in our supplement (along with the two equations above for β and those for γ), which we will be moving to Section 2: Methodology in the main text. We will also clarify the meaning of the sign of the feedback parameters under positive and negative emissions, and hence, the reversal of the sign convention.*

I would also like to note that feedback parameters are most "realistic" or "relevant" when found using FULL and BGC runs. The real world operates like a fully-coupled simulation. For finding feedback parameters in addition to FULL we need a BGC or RAD simulation. Since the carbon-concentration feedback is the dominant feedback perhaps it makes more sense to use the BGC simulation.

- *We agree that feedback parameters are more realistic calculated using the FULL-BGC approach. We include feedback parameters calculated from this approach in Table 1 (shown in parentheses) as well as the relevant land, ocean, vegetation and soil carbon changes in Figure S5 of the supplement.*
- *We find that the feedback parameters from the FULL-BCG approach are qualitatively consistent with those from the RAD approach: the magnitude of the climate-carbon feedback parameters calculated from both approaches is smaller under negative emissions than under positive emissions. Moreover, we find that results are easier to interpret in our proposed approach with feedback parameters computed using the RAD approach rather than the FULL-BGC approach. In the latter case, we would need to subtract the difference between FULL and BGC zero emissions simulations from the difference between the FULL and BGC CDR-reversibility ramp-down phases. This double difference would make it difficult to make sense of the resulting feedback parameters.*

Finally, my last major comment is that when in the real world we do ramp down emissions then, at that point in time, the land and ocean C cycles won't be in equilibrium with the atmospheric CO2. There will be inertia in the real system, and the response of land and ocean at the time will be affected by this inertia. So is the purpose of attempting to correct the feedback parameters for this inertia on the ramp-down side only to compare them with their ramp-up counterparts?

- *Yes, that is correct. Since the feedback parameters under negative emissions include land and ocean responses to both the negative emissions and prior positive emissions, we* *isolate the response to the negative emissions alone by using zero emissions simulations, bringing the feedback parameters closer to the "true" sensitivity of land and ocean carbon under negative emissions.*

Minor comments

1. I realize the purpose of Figure 1 is to clarify things but for me text for easier to follow. Perhaps you can try to improve Figure 1.

We will improve Figure 1 or reconsider including it in the main text.

2. Line 107, "generates permafrost". Please reword this sentence. I think it is incorrect to say "generate permafrost". Permafrost is a state which results from sub-zero temperatures.

Done.

3. Lines 145-152 need equations to clarify the methodology used.

As previously mentioned, we will move the carbon cycle feedback framework (including all equations) in Section A of the supplement to Section 2: Methodology of the main text to clarify Lines 145-152.

4. Line 172. "This temperature change is driven by biophysical responses to increasing CO2". Please add another sentence of explanation at the end of this sentence for completeness.

We thank the reviewer for highlighting this. The temperature change in the BGC mode is driven by changes in evaporative fluxes. We have included a sentence explaining this.

5. Please put a zero line in Figures 3c,d,e,f, and Figures 4a,b.

We will add the zero line to those figures as well as Figures 6 and 7c-f

6. Lines 258-261 read " … except in the vegetation carbon pool where the width of the hysteresis increases throughout the simulation (figure 5(c)). The land and ocean carbon pools in the RAD mode also exhibit hysteresis (figure 6). The hysteresis in the land carbon pool is dominated by the soil carbon pool (figure 5(d)), and the width of the hysteresis appears to increase throughout the simulation for all carbon pools except the vegetation carbon, which shows nearly constant hysteresis".

I am confused here. Please reword clearly. Hysteresis is defined as the difference in paths going up and down. Isn't hysteresis zero at the point of turn? With this in mind please reword the above sentences.

*We agree with this definition and agree that hysteresis should be 0 at the point of turn. We have changed the phrase "… the width of the hysteresis appears to increase **throughout the simulation** for all carbon pools …" to "… the width of the hysteresis appears to increase **throughout the ramp-down phase** for all carbon pools …" to make this clearer.*

7. Lines 273-274 read "The ocean holds only 70PgC less than at preindustrial, but unlike the land carbon pool, a miniscule amount of ocean carbon is regained in the rampdown phase (figure 5d)".

But Figure 5d is the soil C figure. Please refer to the correct figure.

We thank the reviewer for pointing out this error. We corrected the figure reference.

Line 308 reads "For positive emissions, feedback parameters are positive (negative) for a gain (loss) of carbon". Please consider not using sentences that use pair of parentheses to note two points. This can get very confusing. Also, please clarify whether the gain or loss is by which component – land/ocean or the atmosphere.

Done.

9. Line 309 reads " … resulting in a negative denominator (see supplementary equations 3.3 – 3.6)".
There is no denominator in these equations. I think I know what's implied here but it may not be obvious to other readers.

10. Lines 308 – 313. Please use equations here because the sign convention is becoming confusing.

We will address both comments 9 and 10 here. As mentioned in our response to the second major comment above, we will add the two equations above for β *(along with those for* γ*) to our carbon cycle feedback framework in our supplement as a visual reference and to help readers understand the sign convention.*

11. Comparison of Figure 5a and S4a shows there's more hysteresis in BGC run than in the FULL run. Can this be explained? Isn't this a good reason to use the FULL simulation to find feedback parameters on the ramp-up and ramp-down portions?

Although we find the difference in hysteresis between the FULL and BGC modes very interesting, we do not explore or explain this further because reversibility is beyond the scope of our study. In addition, assuming you mean using the FULL-BGC approach for computing the climate-carbon feedback in the two phases, conclusions drawn from feedback parameters computed from the FULL-BGC and RAD approaches are consistent in our study (see response to last major comment).

12. What does "All" means in Figure 7a legend?

The "ALL" label refers to the fact that all three modes (fully coupled, biogeochemically coupled and radiatively coupled) are initialized from the same simulation: the ramp-up phase of the CDR-reversibility simulation. We have clarified this in the figure caption.

13. Zero emissions runs were initialized from the end of ramp-up. What does BGC and RAD mean for these runs? Do the RAD and BGC runs in Figure 7, see and not see temperature change, respectively, relative to end of the ramp-up or relative to the preindustrial state? Please clarify.

As the RAD and BGC runs are initialized from the end of the ramp-up phase, they see and do not see temperature change relative to the end of the ramp-up. Therefore, in the RAD run, the land and ocean see changes in temperature, but see CO_2 concentration fixed at four times the preindustrial CO_2 concentration (~1120ppm). In the BGC run, the land and ocean see changes in CO_2 concentration, but the radiation code remains fixed at four times the preindustrial CO_2 concentration. We will clarify this in the text.

14. Lines 375-377 read "Under negative emissions, the magnitudes of b[eta] and g[amma] from our novel approach are larger compared to those from the "CDRreversibility" simulation WHEN RAMPING UP (CORRECT?), implying greater carbon loss due to the concentration-carbon feedback and greater carbon gain due to the climatecarbon feedback under negative emissions".
"Greater carbon loss" and "greater carbon gain" for what component – land/ocean or atmosphere?

Here, we are referring to magnitudes of carbon cycle feedback parameters under negative emissions i.e., in the ramp-down phase. Therefore, we are comparing the feedback parameters from the ramp-down phase of the CDR-reversibility simulation to those from our "ramp-down – zeroemit" approach (the ramp-down of the CDR-reversibility simulation minus zero emissions simulation). We also include an example of this with values of feedback parameters in Lines 377-381 that read, "For example, a decrease in atmospheric CO2 of one ppm would result in the loss of 0.68 PgC of land carbon in the standard approach and 0.80 PgC of land carbon in our approach due to the concentration-carbon feedback whereas, cooling by one 380 degree, would result in land carbon gain of 56.4 PgC in the standard approach and almost three times as much (157.1 PgC) in our approach due to the climate-carbon feedback." Finally, "greater carbon loss" and "greater carbon gain" refers to the land and ocean components; we have clarified this in all instances in the manuscript.

15. Lines 383-384 read "… due to the concentration-carbon feedback, carbon pools take up carbon in the ramp-up phase, continue to take up carbon in the early ramp-down phase."
Actually, it's the other way around. Carbon pools don't behave according to the feedbacks but rather feedbacks are derived from the behavior of the C pools. Please consider rewording.

16. Next two sentences …
"Due to the climate-carbon feedback, carbon pools lose carbon in the ramp-up phase,

continue to lose carbon in the ramp-down phase, then switch into carbon sinks"

"… suggesting that land and ocean carbon changes due to carbon cycle feedbacks …"
Here too, please consider rewording.

We thank the reviewer for their comments. We will reword both sentences.

17. Lines 404-405 read " … we subtract the zero emissions simulations from the "CDR reversibility"
simulations …".
Please use equations to show how.

As mentioned in the first major comment, we will provide the equations in the main text along with the carbon cycle feedback framework in our supplement.

18. Lines 427-428 read "… concentration-carbon feedback parameter is more positive (Table S2)".
Please clarify if this is from the land's perspective. Please use a single notation consistently.

*This is from the land's perspective. We say, "Models without a nitrogen cycle exhibit greater **land carbon gain** under positive emissions relative to other CMIP5 and CMIP6 models, that is, the concentration-carbon feedback parameter is more positive (Table S2)." [Lines 426-427].*

19. Lines 428-429 read … "They [i.e. land models with N cycle] also exhibit greater carbon loss under positive emissions, that is, the climate-carbon feedback parameter is more negative".

This seems incorrect. Note that land models with N cycle typically have a smaller absolute magnitude of carbon-climate feedback because increase in temperature promotes vegetation growth due to enhanced N mineralization which somewhat compensates for increased soil C respiratory losses.

*We thank the reviewer for their comment. Here, we are referring to models without a N cycle. Lines 427-429 read, "Models without a nitrogen cycle exhibit greater land carbon gain under positive emissions relative to other CMIP5 and CMIP6 models, that is, the concentration-carbon feedback parameter is more positive (**Table S2**). They also exhibit greater carbon loss under positive emissions, that is, the climate-carbon feedback parameter is more negative."*

20. Lines 433 – 435 read "With the consideration of nitrogen limitation, the already weakened CO_2 fertilization effect under declining CO_2 concentrations would be further constrained, exacerbating the carbon loss due to the concentration-carbon feedback".
This seems like a bit of speculation. Why would this be? It could be the other way around too. If increasing CO_2 causes C:N ratios to increase and constrain photosynthesis, more than the case when the N cycle is not represented, then decreasing CO_2 should lower C:N ratio and help vegetation photosynthesize a bit more (compared to when the N cycle is

not represented).
Of course, overall photosynthesis will still be reducing since CO2 is going down but off the top of my head it's difficult for me to imagine the effect of N cycle when CO2 is reducing.

Perhaps is prudent to not speculate.

We thank the reviewer for bringing this to our attention. We had only considered the impact of nitrogen limitation on the CO_2 *fertilization effect. We will also consider C:N ratios and instead of making a definitive statement of the effect of the N cycle, we will suggest both as potential responses.*

21. Finally, what is the CDR-reversibility simulation? Does this refer to both the ramp-up and ramp-down portions or just the ramp-down portion? Note that the ramp-up portion already has a standard experiment name i.e. 1pctCO2. Please clarify this in the beginning and then use the correct terminology throughout the rest of the manuscript

This is correct. The CDR-reversibility simulation is the combination of the ramp-up and ramp-down simulation. We say, "To explore how the magnitude of carbon cycle feedbacks under positive emissions differs from that under negative emissions, we ran the "CDR-reversibility" simulation from the Carbon Dioxide Removal Model Intercomparison Project (CDRMIP) (Keller et al., 2018). Starting from a preindustrial equilibrium state, atmospheric CO_2 *concentration was prescribed to increase at 1% per year until quadrupling, then decline back to preindustrial levels at the same rate." [Lines 119-122]. We also chose to refer to it as the CDR-reversibility ramp-up phase as opposed to "1pctCO2"as we find it clearer which portion of the CDR-reversibility we are referring to. We will make sure to use consistent terminology throughout the manuscript.*