Reviewer comments

Author responses (**Bold is new text in manuscript**)

RC2

Varney et al. present results on soil carbon changes in CMIP5 and CMIP6 models under different levels of climate change, and quantify the contribution of productivity and turnover controls on these changes in soil carbon. They find that the spread in soil carbon responses across CMIP6 models was less than that across CMIP5 models, suggesting a potential reduction of uncertainty in 21st century soil carbon projections. The study shows that there are still differences in the relative contributions of controls (e.g., NPP and tau) on soil carbon changes across the models. They also illustrate a linear relationship between the change in carbon from NPP and turnover time, which they connect to the concept of false priming and demonstrate that this relationship is tighter across the CMIP6 models. In all, this is an interesting study with nicely summarized figures. Some findings could be discussed in greater detail with appropriate caveats, and I include specific comments below.

We are thankful to the reviewer for the helpful comments which we feel have now improved the paper.

Main comments:

It is an important result that the study finds such differences across the two CMIP generations, and I think more discussion on this point could be helpful. Could the authors provide further details on a few of the soil C models that showed large differences between the CMIP generations? How did these soil C model representations change?

We agree that the comparison between two CMIP generations is useful to try and interrupt new additions within ESMs. The response to this comment has been incorporated into related responses below (see ***Comments 2*** and ***Comments 4***).

In particular, did the soil C models change with regards to the number of pools from CMIP5 to CMIP6? I was under the impression that most were based on the the Century or DayCent models, which were developed decades ago. It'd be great if the authors could provide more details on the models here. Otherwise, the last sentence of the abstract about the "inclusion of more soil carbon models with multiple pools in CMIP6" does not have sufficient support currently.

We have included a new table in the manuscript (which has been adapted from Tables 1 and 2 from Varney et al. 2022), which includes information on the related soil carbon components of the ESMs in both CMIP5 and CMIP6. Specifically, the no. of dead soil carbon pools within the ESMs **(see below)**.

Earth System Model	Nitrogen cycle	No. of live carbon pools	No. of dead carbon pools	References
BNU-ESM	No	-	-	Ji et al. (2014); Dai et al. (2003)
CanESM2	No	3	2	Arora et al. (2009); Arora and Boer (2010)
GFDL-ESM2G	No	8	2	Dunne et al. (2012, 2013); Shevliakova et al. (2009)
GISS-E2-R	No	7	5	Schmidt et al. (2014); Yue and Unger (2015)
HadGEM2-ES	No	3	4	Jones et al. (2011); Best et al. (2011): Clark et al. (2011)
IPSL-CM5A-LR	No	-	7	Dufresne et al. (2013); Krinner et al. (2005)
MIROC-ESM	No	4	2	Watanabe et al. (2011); Ito and Oikawa (2002); Sato et al. (2007)
MPI-ESM-LR	No	4	2	Raddatz et al. (2007); Knorr (2000)
NorESM1-M	Yes	13	7	Bentsen et al. (2013); Iversen et al. (2013); Lawrence et al. (2011)
ACCESS-ESM1.5	Yes	3	6	Ziehn et al. (2020); Haverd et al. (2018);
BCC-CSM2-MR	No	3	8	Wu et al. (2019); Ji et al. (2008)
CanESM5	No	3	2	Swart et al. (2019); Melton et al. (2020); Seiler et al. (2021)
CESM2	Yes	22	7	Danabasoglu et al. (2020); Lawrence et al. (2019)
CNRM-ESM2-1	No	6	7	Séférian et al. (2019); Delire et al. (2020)
GFDL-ESM4	No	6	4	Dunne et al. (2020); Zhao et al. (2018)
IPSL-CM6A-LR	No	8	3	Boucher et al. (2020); Cheruy et al. (2020); Guimberteau et al. (2018)
MIROC-ES2L	Yes	3	6	Hajima et al. (2020); Ito and Oikawa (2002)
MPI-ESM1.2-LR	Yes	3	18	Mauritsen et al. (2019); Goll et al. (2017); Goll et al. (2015)
NorESM2-LM	Yes	22	7	Seland et al. (2020); Lawrence et al. (2019)
UKESM1-0-LL	Yes	3	4	Sellar et al. (2020); Wiltshire et al. (2021)

<u>Table NEW</u>: "***The CMIP5 and CMIP6 Earth system models included in this study and the relevant features of associated land carbon cycle components: simulation of interactive nitrogen, number of live carbon pools and the number of dead soil carbon pools (Varney et. al 2022; Arora et al. 2013; 2020).***"

Secondly, we have included extra text in the Methods describing the key differences between the CMIP5 and CMIP6 generations.

Ln 57:

"~~Specific soil carbon related updates within ESMs from CMIP5 to CMIP6 are included in Varney et al. 2022 within the 'Earth system models' section of the Methods, and more general model updates are presented within the 'Model descriptions' section of the Arora et al. 2020 Appendix.~~

The use of CMIP allows for comparison between ESMs in the different ensemble generations. Table 1 presents key soil carbon related information from the ESMs for both CMIP6 and CMIP5 (adapted from Tables 1 and 2 in Varney et al. 2022). The Table can be used to identify key ESM updates between CMIP5 and CMIP6, such as: the simulation of interactive nitrogen in CMIP6 (ACCESS-ESM1.5, CESM2, MIROC-ES2L, MPI-ESM1.2-LR, NorESM2-LM and UKESM1-0-LL) compared to CMIP5 (NorESM1-M) and the number of soil carbon pools (dead carbon pools). The ESMs where CMIP5 and CMIP6 generations are included in our analysis are: CanESM2 and CanESM5, GFDL-ESM2G and GDFL-ESM4, IPSL-CM5A-LR and IPSL-CM6A-LR, MIROC-ESM and MIROC-ES2L, MPI-ESM-LR and MPI-ESM1.2-LR, NorESM1-M and NorESM2-LM, and HadGEM2-ES and UKESM1-0-LL, respectively, where in these cases direct comparisons can be made. It is noted that some Land Surface Models

within ESMs share similarities (e.g. CESM2 and NorESM2-LM both use the Community Land Model version 5; Arora et al. 2020)."

Ln 60:

"***Within ESMs, specific soil carbon processes are modelled using soil biogeochemical models which are used to simulate the flow and storage of carbon within the soil. Since early models, both the litter and soil are simulated using separate carbon pools, which are used to represent differing sensitivities of carbon to decomposition and allocation into pools is often dependent on the molecular structure of the litter and the long-term stability (Exbrayat et al., 2013). Early examples of soil carbon models are the grass and agroecosystems dynamic model (CENTURY; Parton et al. (1988)) and the Rothamsted carbon model (ROTH-C; Jenkinson et al. (1991)). Updated variants of these models are still widely used to represent soil carbon decomposition in modern ESMs within CMIP (Arora et al., 2020; Todd-Brown et al., 2018). Table 1 presents the number of soil carbon pools (dead carbon pools) within both CMIP5 and CMIP6 ESMs, which can be used to compare between the ESMs.***"

1. Exbrayat, J.-F., Pitman, A., Zhang, Q., Abramowitz, G., and Wang, Y.-P. (2013). Examining soil carbon uncertainty in a global model: response of microbial de- composition to temperature, moisture and nutrient limitation. *Biogeosciences*, 10(11):7095–7108.
2. Parton, W. J., Stewart, J. W., and Cole, C. V. (1988). Dynamics of c, n, p and s in grassland soils: a model. *Biogeochemistry*, 5(1):109–131.
3. Jenkinson, D. S., Adams, D., and Wild, A. (1991). Model estimates of co2 emissions from soil in response to global warming. *Nature*, 351(6324):304.
4. Todd-Brown, K., Zheng, B., and Crowther, T. W. (2018). Field-warmed soil carbon changes imply high 21st-century modelling uncertainty. *Biogeosciences*, 15(12):3659–3671.

The abstract has changed to be more appropriate for the findings of the study, then more specific details included in Discussion.

Ln 8:

"*We show* ***that the concept of 'false priming' is likely to be contributing to this emergent relationship*** ~~*this emergent relationship is the result of `false priming'*~~*, which leads to a decrease in the effective soil carbon turnover time as a direct result of NPP increase and occurs when the rate of increase of NPP is relatively fast compared to the slower timescales of a multipool soil carbon model.* ***This finding suggests that the structure of soil carbon models within ESMs in CMIP6*** ~~*The inclusion of more soil carbon models with multiple pools in CMIP6 compared to CMIP5, therefore*~~ ~~*seems to have*~~ ***has likely*** *contributed towards the reduction in the overall model spread in future soil carbon projections* ***since CMIP5***."

How were the particular ESMs included in the analysis chosen? Line 55 says that this was due to data availability, but it is surprising that certain of the CMIP6 models (e.g., CESM2) did not have the necessary data for CMIP5 as well. Either way, it seems that 5-6 models (CanESM, IPSL, MIROC, MPI, NorESM, and HadGEM/UKESM) are included for both the CMIP6 and CMIP5 model output. These models would be good candidates to explore the earlier point above, regarding changes to the soil C modules and the number of pools in the CMIP5 and CMIP6 representations.

The study used all ESMs which had data available for all required variables (cSoil, cLitter, NPP, Rh) on the ESGF website (https://esgf-node.llnl.gov/projects/cmip6/, where all nodes where checked). We have checked the data source again and additional data is still not available. We agree that the ESMs with CMIP5 and CMIP6 model generations provide good candidates to explore common model updates. As above, the study now includes details regarding changes in the number of C pools between CMIP5 and CMIP6 generations.

How many of the models did not report a separate litter carbon pool? (These could be briefly listed on line 88.) It seems that the analysis could be more consistent by using only the soil carbon pool across the models. If not, additional rationale can be provided.

Most ESMs do include a separate litter carbon pool variable. Those that do not in this analysis is only UKESM1-0-LL in CMIP6, plus HadGEM2-ES and GISS-E2-R in CMIP5. The use of combining soil and litter carbon pools to consider 'total soil carbon' in ESMs is a common method used within literature (Todd-Brown et al. 2013, 2014; Koven et al. 2015; Varney et al. 2022). The models which do not report the litter pool do have the representation of litter within the model, but not report a separate variable in the CMIP simulations. The sentence has been updated as suggested.
Ln 88:
"*For models that do not report a separate litter carbon pool (cLitter), soil carbon is taken to be simply the cSoil variable* ***(UKESM1-0-LL in CMIP6, GISS-E2-R and HadGEM2-ES in CMIP5)****.*"

1. Todd-Brown, K. E. O., Randerson, J. T., Post, W. M., Hoffman, F. M., Tarnocai, C., Schuur, E. A. G., and Allison, S. D.: Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations, Biogeosciences, 10, 1717–1736, https://doi.org/10.5194/bg-10-1717-2013, 2013.
2. Todd-Brown, K. E. O., Randerson, J. T., Hopkins, F., Arora, V., Hajima, T., Jones, C., Shevliakova, E., Tjiputra, J., Volodin, E., Wu, T., Zhang, Q., and Allison, S. D.: Changes in soil organic carbon storage predicted by Earth system models during the 21st century, Biogeosciences, 11, 2341–2356, https://doi.org/10.5194/bg-11-2341-2014, 2014.
3. Koven, C. D., Chambers, J. Q., Georgiou, K., Knox, R., Negron-Juarez, R., Riley, W. J., Arora, V. K., Brovkin, V., Friedlingstein, P., and Jones, C. D.: Controls on terrestrial carbon feedbacks by productivity versus turnover in the CMIP5 Earth System Models, Biogeosciences, 12, 5211–5228, https://doi.org/10.5194/bg-12-5211-2015, 2015.
4. Varney, R. M., Chadburn, S. E., Burke, E. J., and Cox, P. M.: Evaluation of soil carbon simulation in CMIP6 Earth system models, Biogeosciences, 19, 4671–4704, https://doi.org/10.5194/bg-19-4671-2022, 2022.

The finding of a turning point from increasing to decreasing soil carbon in Fig. 2 (lines 155-165) is really interesting, and the authors mention that this suggests a potential limit to the Cs increase. Why do you think this could be? And why does this turning point appear later or not at all in some models? Do any geographic regions contribute more to this turning point? Some more discussion here would be great, as this is an interesting finding.

The result of a turning point in future soil carbon simulations has been expanded upon and extra discussion has been included.
Ln 163:
"~~*This finding suggests a potential limit to* ΔC_s *increase and a reduced likelihood of a carbon sink under more extreme levels of climate change.*~~
Generationally related ESMs between CMIP5 and CMIP6 allow us to highlight some key changes between the CMIP generations and to suggest potential model updates which may have contributed to the change. For example, within CMIP5 the models HadGEM2-ES and MPI-ESM-LR predicted the greatest increases in soil carbon within the ensemble. Conversely, within CMIP6 the updated versions UKESM1-0-LL and MPI-ESM1-2-LR predict reduced increases (Fig. 1). This is likely

due to the inclusion of interactive nitrogen within simulations, which could limit carbon sequestration through limiting the magnitude of CO_2 fertilisation (Wiltshire et al. 2021). The projected 'turning point' seen in UKESM1-0-LL within CMIP6, suggests a saturation of the CO_2 fertilisation effect, compared with no saturation of increased respiration with warming. This is again most likely due to the more widespread inclusion of nutrient limitations on CO_2 fertilisation in CMIP6. This finding suggests a potential limit to ΔC_s increase and a reduced likelihood of a carbon sink under more extreme levels of climate change."

1. Wiltshire, A. J., Burke, E. J., Chadburn, S. E., Jones, C. D., Cox, P. M., Davies- Barnard, T., Friedlingstein, P., Harper, A. B., Liddicoat, S., Sitch, S., et al. (2021). Jules-cn: a coupled terrestrial carbon–nitrogen scheme (jules vn5. 1). *Geoscientific Model Development*, 14(4):2161–2186.

The immediate response in respiration in Fig 10b looks surprising, especially the abruptness and shape of the tau curve. I guess this may be because the Cs1 and Cs2 pools both equilibrate almost instantaneously, with intrinsic turnover times of 1 and 10 years. However, it is difficult to see any of these details associated with the short-term response, because the x-axis spans 500 years. It would be helpful to focus on the first 100 or so years following the perturbation, as in Fig. 10a.

Figure 10b demonstrates how the 3-box model responds to a step change of NPP increase. The aim here is to demonstrate the subsequent effects of soil carbon, Rh and turnover changes clearly as a response to the increased input and allowing an equilibrium to be reached. Fig. 10b is a complementary figure to Fig. 10a, which shows the short-term transient response in the same 3-box model. More discussion has been added to the text discussing the figure.
Ln 314:
*"Fig. 10(b) demonstrates that false priming is a transient effect associated with ~~a disequilibrium in~~ the distribution of soil carbon between the 3 pools**, which emerges from the differences in the mass-weighted and flux-weighted responses**. **The fast soil carbon pool reaches equilibrium before the slower soil carbon pool resulting in** the* ***… (Fig. 10(b)).***"

The authors could consider adding some discussion/conclusions on how their results on effective turnover times may connect with radiocarbon-based insights from soil carbon ages (e.g., He et al. Science 2016; Shi et al. Nat Geosci 2020) in data and ESMs.

Discussion connecting with these stated studies has now been included.
Ln 329:
"***A more apparent false priming affect within CMIP6 could suggest an improved representation of the slower components of soil carbon since CMIP5, commonly by including more dead carbon pools within the ESM (Table 1). Based on observational radiocarbon estimates it has been found that CMIP5 ESMs underestimate carbon age within the soil (He at al. 2016; Shi et al. 2020), suggesting that ESMs underestimate the amount of carbon in the slow carbon pools. It has been shown that representing soil carbon ages more in line with radiocarbon estimates leads to a reduced potential for soil carbon sequestration in the future (He et. al 2016), which agrees qualitatively with the projected 21st century soil carbon changes as predicted by CMIP6 compared to CMIP5 found in this study.***"

1. He, Y., Trumbore, S. E., Torn, M. S., Harden, J. W., Vaughn, L. J., Allison, S. D., & Randerson, J. T. Radiocarbon constraints imply reduced carbon uptake by soils during the 21st century. *Science*, *353*(6306), 1419-1424 (2016).

2. Shi, Z., Allison, S.D., He, Y. *et al.* The age distribution of global soil carbon inferred from radiocarbon measurements. *Nat. Geosci.* **13**, 555–559 (2020).

The last point in the conclusions (#6) reads as if false priming itself is a mechanism that affects soil carbon storage. However, it is in fact an effective bulk quantity that results from differences in mass-weighted and flux-weighted responses when there are multiple soil components with different residence times (as is the case in most models and in soil itself). It can thus be a useful quantity to further probe and diagnose model responses in response to perturbations. The authors may want to clarify and refine this last point.

This last point has been changed.
Line 352:
"~~It is recommended that the full extent of false priming on future soil carbon is understood, where if increased carbon inputs to soil carbon pools preferentially enters fast soil carbon pools, this could limit the maximum increase in soil carbon storage in the future.~~"
"***Our study highlights the significant role that false priming can play under transient changes in atmospheric CO_2 and climate. We advise caution in the interpretation of changes in the effective soil carbon turnover time in terms of climate affects alone. Idealised C4MIP simulations, which can be used to separate the effects of CO_2 and climate on the effective soil carbon turnover time, are very useful to assess the role of false priming in models. Understanding these factors will be key to predicting soil carbon changes over the next 100 years.***"

Minor comments:

Lines 18-24: This background jumps around a bit, giving a case study for warming and then saying 'therefore' with a statement about elevated CO2 importance. Consider reorganizing this intro paragraph. Also, there could be a discussion somewhere here regarding the uncertainty resulting from underestimation of soil C ages in most ESMs (e.g., He et al. Science 2016). This is particularly relevant for the discussion of uncertainties resulting from increasing NPP and elevated CO2.

The paragraph has been edited to be clearer.
Ln 18:
"*However, the long-term response of soil carbon is uncertain due to large stocks which are known to be particularly sensitive to changes in CO_2 and the subsequent global warming (Cox et al. 2000). For example, permafrost thaw under climate change has the potential to release significant amounts of carbon into the atmosphere over a short period of time with increased warming, representing a significant feedback within the climate system (Schuur et al. 2022; Hugelius et al. 2020; Burke et al. 2017). Therefore, quantifying the future response of soil carbon* ***under future changes to climate*** *~~to increased CO2~~ is vital in determining the long-term potential land carbon storage.*"

Line 30: There is only reference to Crowther et al. 2016 here, but just a note that there was conflicting evidence in a follow-up to that paper by van Gestel et al. 2018.

Follow up paper is noted and included. The Crowther et al. 2016 study was used as they attempted to quantify the carbon loss from global soils under global warming. However, we note here that Van Gestel et al. (2018) stated a limiting factor of this study is the number of field experiments within northern latitude regions, potentially leading to an underestimation of soil carbon loss. Additional citations have been added here ***(van Gestel et al. 2018)***.

1. van Gestel, N., Shi, Z., van Groenigen, K. *et al.* Predicting soil carbon loss with warming.*Nature* **554**, E4–E5 (2018).

Line 70: 'subtracted' instead of 'taken away'

Sentence changed.

Line 120: can 'be' expanded

Sentence changed.

Line 314: Can you elaborate on what you mean by "offsets about 40% of the increase in soil carbon that would arise from the NPP increase alone" here?

This sentence has been removed to avoid confusion as is not vital to the presentation and understanding of false priming.
Ln 314:
"**.** *~~and for this set of parameters offsets about 40% of the increase in soil carbon that would arise from the NPP increase alone~~*."

Line 315: I'm not sure that it is a 'disequilibrium' per se, but rather an apparent quantity that emerges from differences in mass-weighted and flux-weighted responses.

This sentence has been changed as suggested.
Ln 315:
"*Fig. 10(b) demonstrates that false priming is a transient effect associated with ~~a disequilibrium in~~ the distribution of soil carbon between the 3 pools***, which emerges from the differences in the mass-weighted and flux-weighted responses***."