Interactive comment on "Differences in carbon cycle and temperature projections from emission and concentration-driven earth system model simulations" *by* P. Shao et al.

P. Shao et al.

shaopu0608@gmail.com

We appreciate Dr. Jones's comments that help us to clarify our presentations and motivate us to do additional analyses with new insights.

General comments

But I find some of the logic and order of the analysis does not make sense and there are some gross errors in some of the results due to mistakes with units of carbon fluxes and stores. I have highlighted below some areas where the paper needs to address these issues before it can be accepted for publication. I have also listed a handful of comments which I hope are useful in more minor ways.

Reply: (a) We didn't imply causality and hence the order of the discussion in the original manuscript was fine. We have added material to increase clarity. (b) The units of carbon fluxes and stores were correct in the original manuscript. Additional explanations have been added to increase clarity. (c) We agreed on the comment on climate-carbon feedback metrics, and we have made relevant revisions. (d) We have made other revisions to address other minor comments.

Major comments

1. *On page 997 you try to explain the differences in [CO_2] between the C-driven and E-driven runs by looking at the differences in land and ocean carbon uptake. This is the wrong way round. It is the differences in [CO_2] which drive the differences in the fluxes, not the other way round. The reason why E-driven runs have different [CO_2] from C-driven is because the ESM fluxes differ from the IAM which created the scenario. In the IAM the simple model MAGICC is used to map from emissions to*

concentration. If the ESM uptake differs from MAGICC then the E-driven run will have different [CO_2] from the C-driven run. This difference between E-driven and C-driven runs then CAUSES (is not caused by) the different land/ocean fluxes in the ESM between the two simulations. You cannot use the land/ocean differences to explain the [CO_2] differences.

Reply: We didn't imply causality in our discussion, and hence the order of the discussion in the original manuscript was fine. We have added material to increase clarity: (a) include the [CO_2] mass balance equation, and emphasize we discuss different terms of this equation; (b) following Dr. Jones' comment, briefly discuss the adjustment process (i.e., emission increases atmospheric [CO_2], which then forces the adjustment of land and ocean fluxes to satisfy the mass balance); and (c) mention that, while Jones et al. (2013) evaluated the emission differences between C-driven simulations and integrated assessment models (IAMs), we focus on the differences of various [CO_2] mass balance terms between E-driven and C-driven simulations.

2. *In figure 1 you have mixed up your units of carbon leading to very wrong and misleading numbers in panel 1d. Panel 1a presents the differences in atmospheric CO_2 - it is the same as in Friedlingstein et al. (2014) figure 2a, but subtracting the C-driven [CO_2] from the E-driven results. Here you choose to convert the units to Pg(CO_2) rather than PgC. In panels 1b and 1c you use units of PgC for land/ocean equivalents. You therefore cannot simply add these together. When you combine them to get panel 1d you therefore have mis-matched units and hence VERY odd results. Did you not wonder why these numebrs are so large and different from other studies? e.g. Jones et al (2013) compare the timeseries and cumulative totals of emissions between C-driven runs and the IAMs and do not see numbers this big (Jones et al figure 5b shows most ESMs are within 100 PgC of the IAM and all are within about 500 PgC). If you convert your panel 1a to PgC before calculating the diagnosed emissions you should get consistent numbers. i.e. the 2100 values in your figure 1d should give the same values as the red bars/black dots in Jones et al. figure 5b.*

Reply: (a) The units of carbon fluxes and stores in Figure 1 were correct in the original manuscript. In fact, to avoid possible confusion, we specifically mentioned in the figure caption that "All ordinate units are converted to Pg CO_2." To further increase clarity, we have added additional explanations (including the use of the [CO_2] mass balance equation). (b) After unit conversion (from PgCO_2 to PgC by multiplying a factor of 12/44), the C-driven values by 2100 in our Fig.1d are indeed consistent with Fig. 5b of Jones et al. (2013). (c) We have also compared the emission differences between C-driven simulations and integrated assessment models (IAMs) in Jones et al. (2013) versus those between E-driven and C-driven simulations in our study.

3. *I found the discussion on the climate feedback metrics a bit over simplistic. It is well known (in fact for over a century) that radiative forcing scales with ln(CO_2) and not CO_2 itself. This is precisely why transient climate sensitivity ("TCR") in climate models is measured using an exponential (1%/yr) rise in CO_2 - so that the forcing and hence T response is linear in time. So your finding is of course not surprising that alpha-prime is more constant in time than alpha. If the only point of alpha was a diagnosis of the climate response to CO_2 then this would indeed make more sense. But from the C4MIP feedback framework alpha is required to be a linearization so that the subsequent feedback framework fits. The whole of section 4 of Friedlingstein et al 2006 relies on this linearization, and alpha is then a vital component of the gain factor, eg, in their eqn 7. You cannot simply redefine one term in this framework. If you want to derive an improved framework it would need to be consistent right through all the metrics. Gregory et al (2009, J. Clim., 22, 5232-5250) is a good start at looking at the implications of linearizing the responses about different points.*
I also did not understand why you suggest using log(CO_2) in the gamma term. You do not discuss this at all in the text, so strange to come out of the blue in the conclusion section. Gamma is defined as the change in land/ocean carbon storage per degree of temperature change. It does not involve [CO_2] at all.

Reply: Following this excellent comment, we have (a) deleted the sentence on the γ term, (b) emphasized the use of α and α' as a diagnosis of the climate response to CO_2 change, and (c) emphasized the need of using (linear) α, rather than (nonlinear) α', in the carbon-climate feedback formalism of Friedlingstein et al. (2006).

The minor comments

1. *There are many biases and errors in the carbon cycle simulations of CMIP5 ESMs. Do you have a reason why you single out the seasonal cycle of CO_2 (in abstract and conclusion) as being particularly urgent to address? Anav et al (2013., J. Clim., 26, 6801–6843) look across many variables for example and see big errors in terrestrial carbon stores. It's not clear the seasonal cycle of the fluxes is a more urgent area to fix than this for example. As you say, the CO_2 seasonal cycle is driven mainly by land fluxes. There is very large uptake (GPP) and release (respiration) fluxes which have large seasonality. The cycle of the net flux is a fine balance between these. So I agree the seasonal cycle of CO_2 is at least one good metric of model performance, but it is very hard to know the origins (and hence implications) of any discrepancies. Too large/small GPP might have very different implications from too large/small respiration or simply a is-match in their phasing. Models may have reasonable components and a very poor net cycle, or may get a good net cycle which hides large but compensating errors in the processes. So I would recommend focus on a much wider set of metrics is required.*

Reply: Following this excellent comment, we have added sentences to clarify: (a) processes that dominate the seasonality of [CO_2], such as terrestrial productivity and respiration, have been analyzed before, including in our recent studies (Shao et al., 2013a,b); (b) observational data on the carbon cycle (e.g., terrestrial productivity and respiration) have much larger uncertainties than the [CO_2] observations; (c) evaluation of the seasonality of [CO_2] suggested in our study would be complementary to the other evaluations of the carbon cycle, and (d) delete the word "urgent".

Furthermore, to better connect the [CO_2] seasonality discussion to other parts of our study, we have analyzed the correlation of the [CO_2] seasonality with α and α', and added sentences to summarize the results.

2. *Your description of the experimental design (E-driven vs C-driven, p.993) is good. You might also site Box 6.4 of Ciais et al (IPCC AR5 WG1 Ch.6) which shows this diagrammatically.*

Reply: Ciais et al. (2013) has been cited in the revised manuscript.

3. *p. 997, line 1. When you refer to differences of the E-driven run from the C-driven don't use the word "bias". This implies the E-driven run is wrong and the C-driven is correct. We don't know which is right/wrong/better/worse so just say they are different from each other.*

Reply: The "highest bias" has been replaced by "biggest difference".

4. *P.999/ figure 3a. Why does BNU have a significant T difference at the start of the runs?*

Reply: Thanks. It's because we computed 11-year running means with the end-points utilizing reflective (symmetric) conditions. The temperature difference in January 1850 is indeed very close to zero and is just –0.1 K by December 1850. We have re-plotted Figure 3 without using any smoothed beginning and end points. For instance, the first value would be at 1855, representing the 11-year average from 1850-1860. We have revised the figure caption accordingly.

5. *P.1000. You can define an alpha for runs driven by many climate forcings, but is it meaningful? The temperature in these runs is only partially due to CO_2. Idealized simulations with CO_2 as the only forcing are better to calculate the sensitivity.*

Reply: We have added sentences to clarify: (a) the b term in Eq. (1) represents the temperature change due to non-CO_2 forcings (e.g., other greenhouse gases, aerosols); (b) Friedlingstein et al. (2006) also used this equation in their Fig. 2a; and (c) α in our analysis may not be exactly the same as that in idealized simulations with CO_2 as the only forcing. To help quantify point (c), we have also calculated the correlation coefficient associated with the regression in Eq. (1).

6. *p.1002, lines 4-10. The seasonal cycle of temperature is driven by insolation/earth's orbit. I wouldn't expect the seasonal cycle of CO_2 to affect it. If it were important then you should look at N. hemi and S. hemi CO_2 separately and not a global mean.*

Reply: Agree. Our point was to show that the seasonal cycle of CO_2 does not affect the seasonal cycle of temperature. We have revised the sentences to make it clearer.

7. *Figure 1 - Does INM really have 1200 PgC different land carbon between these two simulations? I haven't looked at the data, but this seems huge. Is it a real difference or some diagnostic issue? You mention they don't impose LUC, but this is the same for both runs isn't it?*

Reply: The unit in Figure 1 is $PgCO_2$, and the results are correct. For instance, the land accumulated differences by 2100 in INMCM4 between E-driven and C-driven experiments is about –295 PgC (= 197 – 492). To increase clarity, we have: (a) added PgC (in addition to $PgCO_2$) in some of the sentences; (b) provided the gross primary production and respiration values from both simulations in INMCM4; and (c) mentioned that, while the accumulated land uptake in the E-driven simulation of INMCM4 is consistent with those from other models, the C-driven value of accumulative net carbon uptake is almost twice as large as the second highest value from BNU-ESM among the eight models.

8. *table 1. Can you define how you calculate the correlation with HadCRUT - is it based on annual temperatures? or smoothed to give decadal/longer trends? If the former then is this meaningful given we don't expect the models to agree in their phase of internal variability, if the latter then is this meaningful? the MAGNITUDE of*

response is not captured so simply getting a good correlation is only part of the requirement.

Reply: It is based on annual temperature, and we have clarified it in the caption. We agree that correlation is just one of the metrics, and the median correlation is about 0.82 in Table 1, rather than much higher, partly because of the phase differences of internal variability mentioned in the above comment.

9. - *[CO_2] differs by 2100 by -19 to +207 ppm. OK.*

- The diagnosed emissions therefore differ by »1000 PgC. No - this is an error due to mixed up units of carbon/CO_2.

Reply: The unit is PgCO_2 (not PgC), and our results are correct. Also see our Reply to Major Comment #2.

- E-driven results have a wider temperature spread than C-driven. OK. But this is not new in itself - see Friedlingstein 2014 Figure 2b. This was also noted in the AR4 projections chapter.

Reply: Motivated by this comment and similar comments from other reviewers, we have: (a) analyzed the change of the global, land, and ocean temperature ranges, separately; and (b) tested the relationship between model [CO_2] differences (i.e., the differences between simulated [CO_2] from E-driven simulations and prescribed [CO_2] in C-driven simulations) for the present-day (2005) versus for the year 2100 as presented in Friedlingstein et al. (2014). Based on these results, we have added materials in different parts of the revised manuscript. (1) The global temperature range increase by 49% (as computed in the original manuscript) is consistent with the finding of Friedlingstein et al. (2014), although our eight models and their eleven models share just six models. This further demonstrates the robustness of this conclusion. (2) The land and ocean temperature ranges increase by nearly the same: 42.3% over land and 41.8% over ocean, but both values are less than that for global temperature (48.5%). (3) In contrast to the statistically significant correlation (R)

between model differences in 2005 versus in 2100, with R^2 = 0.62 in Figure 3 of Friedlingstein et al. (2014), we found R^2 = 0.35 in our calculation, which is not significant at the 0.05 level. Furthermore, R^2 becomes much smaller (0.10) for model differences in 1980 versus in 2100. This implies that, for the eight ESMs analyzed in our study, model [CO_2] biases during the historical period may not be a good predictor of the model differences between simulated and prescribed [CO_2] in 2100.

- carbon cycle feedback framework should use ln(CO_2) for its temperature sensitivity. No - because then the whole linearization breaks down. While I acknowledge the linearization is not perfect it has proven useful. If you only want to look at the climate sensitivity of the model to CO_2 then I agree the C4MIP definition of alpha is not optimal. But within the feedback framework it is still OK.

- use ln[CO_2] in the gamma term instead of [CO_2]. No - [CO_2] is not in gamma in the first place?

Reply: As mentioned in our Reply to Major Comment #3, we have (a) deleted the sentence on the γ term, (b) emphasized the use of α and α' as a diagnosis of the climate response to CO_2 change, and (c) emphasized the need of using (linear) α, rather than (nonlinear) α', in the carbon-climate feedback formalism of Friedlingstein et al. (2006).

- seasonal cycle of CO_2 should be used as a benchmark. I agree, but only within a basket of other metrics too. We need to understand why models differ and ensure that they get the right answer for the right reasons.

Reply: Agree. We have added sentences to clarify: (a) evaluation of the seasonality of [CO_2] suggested in our study would be complementary to the other evaluations of the carbon cycle, and (c) delete the word "urgent." Also see our Reply to Minor Comment #1.

- ESMs and IAMs should be more consistent. They are already fairly consistent. Jones

et al. show this (figure 5). They agree well for low scenarios, and less well for higher scenarios. But the agreement is much better than you show in your figure 1d due to your units error.

Reply: Our units are correct (see our Reply to Major comment #2). We have added sentences to clarify the complementary results in Jones et al. (2013) and our study: (a) our results show the relatively large [CO_2] differences between E-driven and C-driven simulations in 2100; (b) Jones et al. (2013) in their Figure 5 showed that the differences between the diagnosed fossil-fuel emissions in C-driven simulations and the emissions estimated from IAMs are smaller in magnitude than those between E-driven and C-driven simulations shown in Figure 1d; and (c) E-driven ESMs (rather than both E-driven and C-driven ESMs) and IAMs should be more consistent.

References

Ciais, P., Sabine, C., Bala, G., Bopp, L., Brovkin, V., Canadell, J., Chhabra, A., DeFries, R., Galloway, J., Heimann, M., Jones, C., Le Quéré, C., Myneni, R. B., Piao, S., and Thornton, P.: Carbon and Other Biogeochemical Cycles. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S.K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 465–570, doi:10.1017/CBO9781107415324.014, 2013.

Friedlingstein, P., Cox, P., Betts, R., Bopp, L., Von Bloh, W., Brovkin, V., Cadule, P., Doney, S., Eby, M., Fung, I., Bala, G., John, J., Jones, C., Joos, F., Kato, T., Kawamiya, M., Knorr, W., Lindsay, K., Matthews, H. D., Raddatz, T., Rayner, P., Reick, C., Roeckner, E., Schnitzler, K. G., Schnur, R., Strassmann, K., Weaver, A. J., Yoshikawa, C., and Zeng, N.: Climate-carbon cycle feedback analysis: Results from the C^4MIP model intercomparison, J. Climate, 19, 3337−3353, doi:10.1175/JCLI3800.1, 2006.

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., and Knutti, R.: Uncertainties in CMIP5 Climate Projections due to Carbon Cycle Feedbacks, J. Climate, 27, 511-526, 10.1175/JCLI-D-12-00579.1, 2014.

Jones, C. D., Robertson, E., Arora, V., Friedlingstein, P., Shevliakova, E., Bopp, L., Brovkin, V., Hajima, T., Kato, E., Kawamiya, M., Liddicoat, S., Lindsay, K., Reick, C. H., Roelandt, C., Segschneider, J., and Tjiputra, J.: Twenty-First-Century compatible CO_2 emissions and airborne fraction simulated by CMIP5 Earth System Models under four Representative Concentration Pathways, J. Climate, 26, 4398−4413, doi:10.1175/JCLI-D-12-00554.1, 2013.

Shao, P., Zeng, X.-B., Sakaguchi, K., Monson, R. K., and Zeng, X.-D.: Terrestrial carbon cycle - climate relations in eight CMIP5 earth system models, J. Climate, doi: 10.1175/JCLI-D-12-00831.1, 26, 8744–8764, 2013a.

Shao, P., Zeng, X.-B., Moore, D. J. P., and Zeng, X.-D.: Soil microbial respiration from observations and Earth System Models, Environ. Res. Lett., 8, doi:10.1088/1748-9326/8/3/034034, 2013b.