Padrón et al report an analysis of drivers of the terrestrial carbon sink in the CMIP6 ensemble scenario SSP126 where warming is limited to 2 oC. This is a useful study of the latest CMIP model results in a policy relevant scenario, showing that terrestrial carbon sink projections by 2100 (cumulative NBP) in the ensemble vary from 56 to 207 Pg C, mean 144 and standard deviation 47 Pg C. Using linear regression Padrón et al partition this variability among sensitivity to CO2, temperature (T), soil moisture (SM), and differences in baseline temperature and soil moisture. Their methods show that the greatest proportion of this variance is explained by sensitivity to T and SM combined, with sensitivity to CO2 as the second most important driver of variability. Based on these results, they conclude that the gamma feedback (climate) is greater than the beta feedback (physiological) under this policy-relevant scenario and thus climate sensitivities require the greatest attention. They also show compensating drivers of cumulative NBP variability such that reduction of uncertainty in response to one driver would not greatly reduce overall NBP variability.

Overall this is a well written and executed study. The analysis of the relatively low-warming SSP126 scenario is timely and to my knowledge has not been done before. I have several comments and criticisms that I hope will help to make the analysis and conclusions more robust and impactful. First, I think that for a number of reasons the method has low-biased the estimation of the impact of CO2 sensitivity on NBP variability. Second, I encourage a little more quantification and thought into exactly what is quantified and communicated.
First: underestimation of the impact of CO2 sensitivity on NBP variability. CO2 sensitivity is estimated as the sensitivity of GPP to CO2 in the 1 % per year increases in CO2 simulations (1pctCO2-bgc) in which CO2 ranges from 350 to 800 ppm. This method assumes 1) the CO2 sensitivity of NBP is the same as that for GPP, 2) that CO2 sensitivity is linear across the range 350 to 800 ppm, and 3) that there are no interactions between CO2 sensitivity and either T or SM. It is likely that all three of these assumptions will low-bias the estimate of the impact of model CO2 sensitivity on cross-model NBP variability.

- While GPP sensitivity to CO2 is likely the main driver of NBP sensitivity to CO2, as asserted in the current ms, the assumption ignores potential changes in turnover rates that can also occur in response to CO2, which can be substantial. Using cross-model GPP sensitivity to CO2 will result in a lower correlation with NBP variability than using NBP sensitivity to CO2. Further, for T and SM sensitivity, NBP is used, biasing results in favor of T and SM sensitivity. Comparing the sensitivities of GPP to CO2 to NBP to T and SM is not a like-for-like comparison. Sensitivity of NBP to CO2 should be estimated and used in the regression analysis.

- The CO2 response over 350 to 800 ppm is likely not linear in these models, it almost certainly is not at the leaf scale which drives model CO2 responses. The SPP126 simulations max out at 446 ppm. There is likely saturation in the CO2 response for many models somewhere between 450 and 800 ppm. CO2 sensitivities should be estimated over the range of CO2 concentrations that preserve linearity over the range 350 – 446 (i.e. concentrations can be higher but responses must be linear over the range). A supplemental figure showing NBP against CO2 for the 1pctCO2-bgc simulations would be useful.

- There are interactions between CO2 and T and SM. Interactions with T are likely the most important for this discussion. At high T it is well known that CO2 can alleviate some of the reductions in photosynthesis due to interactive effects on photorespiration. This could alleviate GPP reductions in high T years that I’m not sure would be removed by detrending NBP. I’m not sure there is an easy way to account for this, and that is OK. But some acknowledgment of this effect and some attempt to quantify it would help make results more robust.

Second: encourage more thought into exactly what is quantified and communicated. I suggest quantifying statement in the abstract, cumulative NBP variability etc. Also, as well as putting these numbers in the context of current annual emissions, I think it might also be useful to present them as a proportion of the assumed emissions in the SSP126 scenario (if someone has calculated those). Why is the proportion of variance in NBP variability to CO2 sensitivity not quantified on ln 354? I encourage the authors to think about what is best to present given this is a study of the global carbon cycle. Most figures are presented in the units per meter squared. When aggregating to broad zonal regions I suggest it is more informative to present results as the absolute sum across the whole
area – this would make it easier to relate the regions and sensitivities directly to the
global aggregate numbers. Finally, have differences in grid-square area been taken into
account when presenting the global aggregate drivers of NBP variability?

Technical Comments

Title: suggest switching “controls” for “drivers” as control suggests some degree of intention.

While I see some of the benefits of the narrative style with the methods spread throughout the results (e.g. Ins 157-164, 179-189, 263-283, etc), I think it is more practical to have the methods all in one place where they can be found easily and assessed side by side.

Fig 6: can probably go into the supplement.

Fig 7: A little hard to read, I think the sensitivities could be presented more clearly if they were presented as in Fig 8. I recognise that would necessitate removal of uncertainties from the figure and I appreciate the effort made to quantify uncertainty but is clarity of communication is the trade off. Fig 8 as is could go to the supplement.
Fig 8: Suggest adding a dot for the actual cumulative NBP. Also I really think this would be better off expressed in global sums rather than per meter squared. The white dot could be a little larger.

Ln 30: There are several commentaries explaining why Wang et al 2020 is not a reliable analysis.

Ln 37: Are there other disturbances that release C directly back to the atmosphere?

Ln 61: Note the editor’s note for Keenan 2021

Ln 69: can delete “consider it important to instead”