

Looking at Figure 3, why does CC exhibit noticeable drift in several variables, such as ocean NPP (panel j) as well as the variables shown in panels d–g, whereas this drift is absent in CC2? Could the authors comment on the reason for this difference? Is the improved stability in CC2 primarily the result of model tuning, or does it arise from changes in the model physics?

Version CC was not properly tuned. We had not run the simulation long enough to obtain soil carbon in equilibrium and that caused the land flux to be drifting and thus the ocean as well. Additionally the ocean NPP was drifting because surface ocean nitrogen concentration, the main limiting nutrient for phytoplankton growth over most of the ocean, exhibited a negative drift of 0.2 $\mu\text{molN/kg/century}$ in CC. Drifts in nitrogen concentration, and the resulting drift in NPP, were a result of both insufficient tuning, which caused too much sequestration of nitrogen in the deep ocean, and differences in spin-up periods for the physical climate and the ocean carbon cycle. In both CC and CC2, the ocean carbon cycle is not initialized from observations until thousands of years after the physical model is spun-up (3000 in CC; Ito et al., 2020, vs. 7000 in CC2). Once initialized, the carbon cycle was spun-up for 700 years in CC, and 300 years in CC2. Century scale spin-ups do not guarantee equilibrated solutions, as ocean carbon cycle models may need thousands of years to spin-up, depending on how “close” the initial conditions are to the steady-state solution (Li and Primeau, 2008). Since CC2 involved extensive re-tuning to bring nutrients and NPP closer to observations, and the model is initialized from these same observations, it follows that the model would require less time to reach a steady-state solution. We also emphasize that these are inter-dependant issues; the spin-up period needed to reach a quasi-steady state is dependent on both initial conditions and tuning of ocean biogeochemical parameters (Seferian et al., 2016). CC2 likely takes less time to reach steady-state than CC because the model solution has been tuned to be closer to observations.

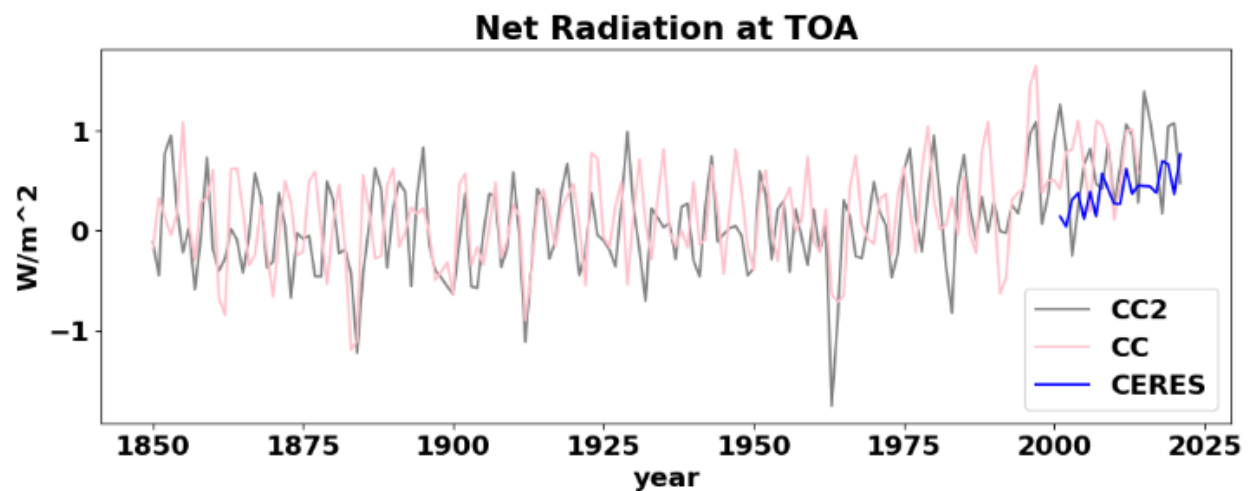
Moreover, the drift in ocean NPP in CC does not appear to affect the model's mean climate state, such as the global mean TOA radiation balance or global mean surface temperature. Does this suggest that production-related variables are largely decoupled from the climate system's energy budget? In contrast, all variables in CC2 appear relatively stable, making it difficult to assess whether drift in variables such as NPP would eventually influence other aspects of the system. A brief discussion of this difference would be helpful.

We are not clear on this question, hence the answer may be non-specific. We first think it is worth clarifying that the ocean carbon cycle only affects the climate system's energy budget through its effects on atmospheric CO₂. The ocean carbon cycle and its effects on the atmospheric CO₂ includes both the biological pump and the physical pump. While the biological pump is sensitive to NPP, which higher NPP leading to higher organic carbon export and

increased capacity for the ocean to adsorb atmospheric CO₂, the physical pump is determined by ocean surface temperatures, and vertical mixing and transport of seawater between the surface and interior ocean. So even if NPP is reduced compared to observations and has a negative drift, the physical pump can still be responsible for the majority of the exchange between ocean and atmospheric CO₂ in the pre-industrial climate in which the ocean carbon cycle is spun-up.

In Figure 4, is there a reason why the net TOA radiation is not shown, as it is in Figure 3a for the pre-industrial control simulation?

Figure 4 now also shows net TOA as shown below, where validation data is from CERES.



Validation data: <https://ceres-tool.larc.nasa.gov/ord-tool/jsp/EBAFTOA421Selection.jsp>

Finally, I think the overall quality of the figures could be improved. It would help to adopt a more consistent style across all figures, including fonts, legends, line styles, panel labels, and other formatting details, to improve readability.

We have redone Figure 4 to ensure that it uses consistent fonts and line styles. Also, we have added a panel with the Net Radiation TOA, as discussed before. We also redid Figures 6 and 7 to have the same styles.

References

Xingwen Li, X. and Primeau, F.W. 2008. A fast Newton–Krylov solver for seasonally varying global ocean biogeochemistry models, *Ocean Modelling*, Volume 23, Issues 1–2, Pages 13-20, <https://doi.org/10.1016/j.ocemod.2008.03.001>.

Séférián, R., Gehlen, M., Bopp, L., Resplandy, L., Orr, J. C., Marti, O., Dunne, J. P., Christian, J. R., Doney, S. C., Ilyina, T., Lindsay, K., Halloran, P. R., Heinze, C., Segschneider, J., Tjiputra, J., Aumont, O., and Romanou, A., 2016: Inconsistent strategies to spin up models in CMIP5:

implications for ocean biogeochemical model performance assessment, *Geosci. Model Dev.*, 9, 1827–1851, <https://doi.org/10.5194/gmd-9-1827-2016>.