The major comments from my previous review were about the data used, the assumption of a constant feedback parameter and the length/prose of the paper. While the authors have tried to address all, I am not yet satisfied with the solutions and I recommend another round of major revisions. I still believe the manuscript will become a valuable piece is the discussion of climate sensitivity, as it provides a comprehensive overview of modes of internal variability.

Major comments:

1. The authors still use HadCRUT4 as if it were a global average. They justified this by citing other papers that have done the same. The Nicholls paper seems to make the same mistake as the authors. However, the other paper does not. Liang et al. (2020) use HadCRUT4, but they do take into account that it is not a global average and that there is missing data, especially around the poles. They use a mask of the CMIP output so that the spatial coverage of the datasets is the same. This takes some work to implement, so I suggest the authors choose any of the four datasets with global coverage. Instead of comparing HadCRUT4 with CW14 (f.i. in Table 1, Figure 12), the authors can compare CW14 with NOAAGlobalTemp v5. Dropping HadCRUT4 from the manuscript completely also helps making the paper shorter.

   HadCRUT4 uses HadSST3 for sea surface temperatures, which further shows slower warming due to biases in ship measurements in comparison with HadSST4. As I understand it, even the incomplete (not infilled) provisional version of HadCRUT5 shows more warming than HadCRUT4. CW14 uses HadSST3 as well, potentially explaining why it warms more slowly compared to some other global averages.
2. The authors have developed an application of EM-GC with blended observations, but temperature output of CMIP6 to test whether EM-GC has predictive power for future temperatures. This is not quite what I intended with my comment, but I admit I wasn't clear before. I had hoped the authors would develop a pure model-based test of predictive power. The outcome of the blended result shows that EM-CG often underestimates ECS, but the authors claim in the body of the text that it is a very good predictor.
3. The authors now examine a time-varying feedback parameter, which varies with radiative forcing. They do not give justification for why they integrate a time-varying feedback like that. Global feedback is thought to change because of cloud feedbacks above a slow-changing ocean. A delay of a couple of decades between radiative forcing and the change of feedback is therefore expected. Disregarding the physics lead to a biased outcome, as the model is trying to fit the rise in the feedbacks too early, and it is only natural that would fail. Scaling with RF would mean that there is barely any feedback in the first half of the twentieth century, which is also unphysical.

   Different formulations for time-varying global feedback exist for simple models, such as (Armour et al., 2013; Geoffroy et al., 2013; Goodwin, 2018). All of these formulations have in common that the feedback only changes some time after radiative forcing, with different lags. I think the Goodwin approach is most suitable for adjustment into EM-GC. Goodwin, also using a data-driven model, shows that the upper range of climate sensitivity is extremely sensitive to the time-scale.

   I further believe that getting an optimal global constant by fitting, and then adjusting the

model to include time-varying feedbacks will tend to favour the former. Ideally, the fitting is done simultaneously.

4. I don't see how the authors determined the uncertainty around the carbon cycle. I cannot find a mention of 10% of Friedlingstein (which concluded that emission-driven simulations warm a tad more than concentration-driven simulations in CMIP5). Ten percent seems low, but this is not my expertise.
5. The paper is still quite long. In the minor comments I will make another set of suggestions to make the paper easier to understand. This will not be an exhaustive list. There are good guides on the internet for writing concisely, that have helped me become a better writer. For instance: https://writingcenter.gmu.edu/guides/writing-concisely.
   The EM-GC model does not model the carbon cycle explicitly, and discussion of the carbon cycle may also be an option to remove. I don't see the value of showing all SSPs in f.i. Figure 9. Consider dropping those with few CMIP6 models.

Geoffroy, O., Saint-Martin, D., Bellon, G., Voldoire, A., Olivié, D. J. L., & Tytéca, S. (2013). Transient climate response in a two-layer energy-balance model. Part II: Representation of the efficacy of deep-ocean heat uptake and validation for CMIP5 AOGCMs. Journal of Climate, 26(6), 1859–1876. https://doi.org/10.1175/JCLI-D-12-00196.1

Goodwin, P. (2018). On the Time Evolution of Climate Sensitivity and Future Warming. Earth's Future, 6(9), 1336–1348. https://doi.org/10.1029/2018EF000889

Armour, K. C., Bitz, C. M., & Roe, G. H. (2013). Time-Varying Climate Sensitivity from Regional Feedbacks. Journal of Climate, 26(13), 4518–4534. https://doi.org/10.1175/JCLI-D-12-00544.1

Minor comments:
79: Replace 'to designate future' with 'for the': future and scenarios are redundant
101: 'land-use change': check hyphens throughout the entire paper
131: remove 'of climate'
132: remove 'because', start new sentence at 'this'
142: consider removing 'Bony et al.' sentence, I don't see the use
150: due to this update, our model is
186: which update
202-205: long sentence
209: 'that is our primary data source', maybe replace with: 'which we use as default'
220: rung→panel
235-237: unnecessary sentence
240: reword: for this simulation, kappa =1.28, W/m^2/C fits the OHC data best
242: remove 'the' before 'IOD'
243: remove 'temporal variations in'
245: slight -> small
343: remove 'consequently'
347: remove 'multiplicative': factor is by definition multiplicative
348: split sentence after '2015'
354: remove 'thus'
367: remove sentence, already clear
379: remove 'scientific': what else?
408: consider replacing 'upon' with 'on' throughout: make it easy for your reviewers and readers to read your text

419: consider using the improved HadSST4, which removes biases in the ship measurements.
420: remove ‘variations in the strength’?
421: I’m not sure whether it’s appropriate to detrend using RF. Temperature lags RF quite a bit, especially in oceans.
433: remove everything between brackets
435: surely the numbers are altered. I cannot imagine that the feedback parameter isn’t dependant on AMOC in the fit.
438: consider using ‘use’ throughout instead of ‘utilising’
453: is this old factor still valid?
455: remove sentence ‘since … whole atmosphere’, redundant.
459: remove ‘temporal’
481: remove ‘however’
505: equal to→of
510: upon consideration of→by including
534: colouring seems to be off in figure S10
539: remove ‘the computation of’
552: remove sentence, redundant
Section 3.1: move methodology to methodology section 2.2.1 (the bit about blending)
Figure 8: what interval is plotted for each study?
772: changed word order, it seems like we’re coupling a two-box model to 2.6
793: Cox et al. based on CMIP5
834: remove ‘indicated on each plot’, redundant
834-835: remove sentence, the reader will know how to do a global average
858: I don’t think bimodality is clear here. There seems to be outliers, but not two roughly equal-sized groups of models. With so few models, passing any statistical test on bimodality would be tough. Drop it?
863: remove ‘apparent in figure 9’, redundant
918: remove ‘our’, redundant
Figure 12: choose bigger bin size: CMIP models displayed weirdly
934: three significant digits not justified, two better
Table 1: same
991-1003: you seem to be repeating the table, making the prose difficult to read, condense to half the size?
1015: since -> from / from … onwards
1023: I don’t think either of them studied the entire climate system. Instead, those studies were about the atmosphere.