Interactive comment on "Twenty-first century ocean warming, acidification, deoxygenation, and upper ocean nutrient decline from CMIP6 model projections" *by* Lester Kwiatkowski et al.

Anonymous Referee #1

This is a well written, valuable paper, which pulls together initial ocean biogeochemical results from the CMIP6 models and places them in context with the CMIP5 results. This will no doubt provide a useful set of figures from the upcoming IPCC assessment, and will be a useful resource for people looking for the headline CMIP6 biogeochemistry results.

I can see no major issues with the manuscript, but have a number of suggestions which hopefully can improve the clarity of the analysis and results. I will start with the more substantial comments.

I appreciate why for practical reasons fixed depths have been used for the nitrate,

O2 and stratification analyses, but I worry that it is oversimplifying things and leading to artefacts which are not obvious as such in the results. For example, assuming stratification to be represented well by the density difference between 0 and 100m will not hold up in many areas (e.g. the Arctic), why not use the mixed layer depth outputted by the models? Similarly, taking a fixed definition of the euphotic zone depth as being 0-100m will be appropriate in some places and not others. Finally, is an average O2 concentration from 0-600m really a good way to understand OMZ volume? Would the column thickness and depth of the the OMZ not be a more useful value when looking at impacts as this paper does?

The results are typically presented very clearly, but do not attempt to distinguish between significant and non-significant results. In the map-based analysis, I would strongly suggest using an approach like that now routine in much of atmospheric science to highlight model agreement by adding stippling to the maps.

A number of the figures (e.g. figure 2) show an absolute anomaly from a climatological value. This is hard to interpret. Please display either as a percentage change (preferentially in my view), or also display the climatological value.

Section 3.8 (seasonality in ocean acidification parameters) is interesting, and the original paper looking at this is really nice, but I would question whether it is a useful section to have in a paper on impacts. Maybe I'm, wrong I would not describe this as an 'impact driver'.

Minor comments (in the order presented in the manuscript):

Figure 1: - It would be useful to plot the CO2 and radiative forcing time-series here so that the reader can visualise the differences between the RCPs and SSPs. - It is really hard to see the CMIP5 results on this. Larger dots and more transparent plumes might help?

Figure 5: This is not very clear. I can not even make out where the O2 < -30 areas

are on my printed copy. I can also not distinguish SST AND NO3 from SST OR NO3, which is pretty key given that the figure is about compound drivers.

Figure 8: Why do you see common changes in the Southern Ocean across ssps? Is it that they are still responding to a common historical period, or is it that it is dominated by large internal variability in one or two models. I think stippling for 'significance' would really help here.

p2 l 68 Wm2 is not a unit of warming.

p5 l191 'best available' what does 'best' mean?

p5 l203 'were vertically regridded' - on to what grid (this could be important for the benthic work)

p7 l261 (and subsequently) 'model structural uncertainty' - my understanding is that this describes only differences in model component design, but actually what you are describing here is all differences between models (e.g. including parameter uncertainty). I think it woudl be better to describe this as 'inter-model uncertainty'

P8 l1 'near global relatively uniform' - it is hard to tell if this is true or just a function of the wide colour bar chosen to allow the two scenarios to be on the same color range. Given that these figures are included to allow the reader to interpret the spatial nature of the changes, and figure 2 already allows the reader to understand the relative changes between scenarios, I would avoid using a common color scale where it masks the detail.

p8 l317 Why is primary production not mentioned when it comes to explaining the O2 changes? Surely it is very important (as seen in fig 3)

p8 l 320 'a subset of the CMIP6 models' - what is this subset, please state.

p9 l350 'regions of enhanced stratification are typically projected to experience reductions in euphotic NO3...' I don't think this is clear from a comparison of figure 2 and

4. The fact that this is clearly what we would expect reenforces the suggestions made above for improving the MLD and euphotoc zone analysis.

p9 l 354 IO would not simply attribute the Arctic behaviour to sea-ice. Arctic stratification is highly complex, and I suspect your simple stratification metric is not representing it well.

p10 1st paragraph - are GHG concentration and pathway differences between the MIPs not important also?

P11 Is vertical resolution potentially an important souce of uncertainty in your global averaged benthis numbers? high vertical resolution == more shelf sea.

Typographic changes:

p2 l66 'Global sea surface...' should be 'Globally averaged sea surface...'?