

Responses in bold

Anonymous Referee #1

Received and published: 10 February 2020

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 OMZ not be a more useful value when looking at impacts as this paper does?

Most of these choices were made to facilitate comparison with past intercomparison studies and previous IPCC reports. In particular, the recent SROCC report used a fixed definition of 0-100m when reporting changes in upper ocean nitrate concentrations. The O₂ concentration over 100-600m is a standard metric used to assess deoxygenation and the expansion of OMZs in observations and models. We propose to keep these definitions however provide additional metrics for the stratification analysis, where we will include changes in mixed layer depth and density differences between 0 and 1000m.

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.

Stippling will be added to maps in the revised manuscript to identify regions with strong model agreement.

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.

Providing percentage changes in maps is problematic for a number of variables. In certain regions concentrations can be very low and therefore percentage anomalies can explode with limited absolute change (e.g. NO₃/NPP in the subtropical gyres). We are happy to add the climatological means.

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'.

Although previous ocean biogeochemistry impact driver assessments have largely focussed on changes in mean state properties (e.g. Bopp et al., 2013) there has been a shift in the literature,

with greater recent recognition of the potential impact of changing temporal variability of drivers (Frölicher et al., 2018; Smale et al., 2019; Kroeker et al., 2020). We wanted to reflect this in the manuscript by including analysis of surface temperature and carbonate chemistry seasonality.

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?

We will add these and work on improving the clarity of this figure in the revised manuscript.

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.

We will work on improving the clarity of this figure in the revised manuscript.

Figure 8: Why do you see common changes in the Southern Ocean across ssp's? 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.

This was not explicitly assessed in the initial manuscript but the reviewer is likely correct that it reflects a lag between changes in atmospheric forcing and the benthic response. Stippling will be added in the revised manuscript to assess the significance of these benthic anomalies.

p2 l 68 Wm² is not a unit of warming.

This text will be corrected.

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

This text will be revised to clarify this.

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

These additional details will be added.

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 would be better to describe this as 'inter-model uncertainty'

We will change this to improve clarity.

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.

The text here was not sufficiently clear. We were merely trying to state that virtually the entire surface ocean warms. We take the reviewer's point and will revise the colour scales and associated text for Figure 2.

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)

We consciously chose not to include primary production in the initial manuscript because a thorough analysis of NPP changes is beyond the scope of this paper and requires a number of additional variables/processes to be assessed. However, on the recommendation of reviewers we will include the model NPP projections in the revised manuscript.

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

This will be added.

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 reinforces the suggestions made above for improving the MLD and euphotic zone analysis.

As stated above, this analysis will be extended.

p9 l 354 I 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.

As stated above, this analysis will be extended.

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

As discussed elsewhere in the manuscript, they are important, particularly for acidification projections. We will add mention of this to this paragraph in the revised manuscript.

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

Yes. Discussion on this will be included in the revised manuscript.

Typographic changes:

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

This will be corrected.