EDITOR

The two reviewers who assessed the first version of your work have evaluated the revised version. Both underline your efforts and find it greatly improved. They still have minor comments that I would like you to consider.

RESPONSE: Thank you for your comments and guidance through the review process. I have responded to the reviewer comments and made corresponding minor changes to the manuscript. I also made small edits where I found typographical errors and I removed three unnecessary sentences in the Methods section.

REVIEWER 1

This manuscript by Whitney has clarified all the major concerns and questions I raised in my previous review in a clear and direct way. I appreciated the effort from the author to strengthen the paper by adding additional climate model simulation analysis. The result section has been enriched including more data and further comparison, and discussion section has also been significantly improved. Although there are limitations from the global climate model and analysis, the results of this study provide a global perspective on global and coastal ocean warming with updated CMIP5 model and greenhouse gas emission scenarios considering large model ensembles. The data presented by this study is valuable for scientists and managers who are interested in warming pressure on coastal hypoxia and ecosystem dynamics.

RESPONSE: I appreciate the reviewer's comments.

There are a few remaining suggestions and concerns which are listed below.
(1) The rate of warming calculated in this study was based on a particular summer month (Feb/Aug) which I think should be clearly noted in the abstract and conclusion. Since climate models may project seasonal-varying warming rate, it's better to clarify it.

RESPONSE: I added "summer-month" in the Abstract for brevity and added the specific summer months in the Conclusions.

REVISIONS:
ABSTRACT: A global 40-year observational gridded climate data record and 21st century projections from the Community Earth System Model (CESM) under RCP8.5 forcing are analyzed for long-term linear trends in summer-month conditions, with a focus on warming-related pressures on coastal oxygen levels.
CONCLUSIONS: A global 40-year observational gridded climate data record (updated from Merchant et al., 2019) is analyzed for linear trends in temperature and oxygen conditions during summer months (August and February averages for northern and southern Hemispheres, respectively). CESM 21st century projections for the RCP8.5 scenario are analyzed in similar fashion for summer months with annual minimum oxygen conditions.

(2) This suggestion might require additional extra work. It is interesting to look at how many of eutrophication locations will become hypoxia at the end of 21st century based on Diaz et al. 2011 databased (L163) especially those at high latitude. This may enhance the highlight of this paper.

RESPONSE: This is a really good idea to explore. Extracting the projected oxygen trends for these locations would be straightforward, but determination of reaching the tipping point into hypoxia also would rely on the current near-bottom oxygen concentrations observed in each of these areas. For the latter step, I would want to check with published oxygen values in a variety of the eutrophic locations. That would be more work than I can complete in the short timeline for this round of revisions. I did not modify the paper in light of this idea, but there is clear value in conducting the suggested analysis in the future.

(3) Suggest to add temperature trend comparison among climate models in Table 3. The number of other CMIP climate models are not big enough and might be biased. Also it is better to list the names and a bit more detailed information of other climate models, like Table 1a in Bopp et al. 2013. Because certain climate models have bias and including the names of models included for analysis help the reader to understand the results.

RESPONSE: Adding the temperature trends to Table 3 would make the table dense with data. Particularly since the surface oxygen concentration trends and oxygen capacity trends should be added too if the temperature trends are added. For this part of the paper, keeping the focus on the vertical minimum oxygen trends is better for succinctly placing the CESM RCP8.5 projections in the context of other models. This variable requires the most from the biogeochemical models and is most directly connected to hypoxia. I added a sentence to that effect.

I analyzed all the CMIP5 and CMIP6 models with ocean biogeochemistry and the "o2min" variable currently available on the Earth System Grid Federation. The several models included do provide valuable context, but I agree that additional models would add to that context. I added a sentence about that.

Component model information for the "Other CMIP5" and "Other CMIP6" run sets is included in the Methods section and model names have been added to Table 3 for ready reference.

REVISIONS:
The contextual focus is on vertical-minimum oxygen concentrations because it is immediately linked with hypoxia and is most demanding of the biogeochemical models.

Increasing the number of runs in each set eventually may be practical if results from additional models with biogeochemistry become available on the Earth System Grid Federation.

The names "(HadGEM2-ES, IPSL-CM5A-LR, IPSL-CM5A-MR, MPI-ESM-LR, MPI-ESM-MR)" have been added to the Other CMIP5 entries and the names "(CanESM5, IPSL-CM6A-LR, MPI-ESM1-2-LR, MPI-ESM1-2-HR)" have been added to the Other CMIP6 entry in Table 3.

Other detailed comments:
L185-186: does it mean the lower rates are more trustable than higher values in Arctic region? Is it not very consistent with L189-199 that when only including points $P<0.1$ the observed global median SST trend is 0.22 deg/decade which is larger than including all points?

RESPONSE: There was a typographical error causing the problem. The sentence should have said P-values are higher in these locations.

REVISIONS: P-values are higher where calculated rates are lower in parts of the Atlantic, much of the South Pacific, and most of the Southern Ocean.

L226-227: the decline rate of oxygen capacity of global ocean is actually slower than coastal ocean. Please check it again.

RESPONSE: This faster and slower refer to the comparison to the Gilbert et al. (2010) rate. The sentence has been modified to clarify this point.

REVISIONS: The observed median coastal rate is half of the rate calculated for a global coastal band (within 30 km of the coast) for 1976-2000 (Gilbert et al., 2010). For the reasons mentioned above, a match between the studies is not expected. It is interesting that the current study has faster global rates and slower coastal rates than the Gilbert et al. (2010) study; the underlying reasons are not explored here.

L234-239: thank you for adding those sensitivity calculations on salinity.

RESPONSE: The sensitivity calculations on salinity helped support the observational results.

L275-276: actually, comparing to the 1990s instead of 2006 with 2090s should generate larger warming. Need reconsider this explanation.

RESPONSE: This latter point wasn't fully developed so I shortened it to a less specific statement. Going into further detail would be a detour from the main flow of the narrative.

REVISIONS: The smaller warming rate for the Bopp et al. (2013) results is likely connected to including multiple models and the different time period analyzed.

L316-321: the projected difference between coastal and hypoxia area in the decline of oxygen capacity is smaller than observed although the warming rate is similar, any explanation?

RESPONSE: I think this point relates to results that the projected rates for coastal points (-1.6 mmol m^{-3} per decade) and documented hypoxic areas (-1.4 mmol m^{-3} per decade) are closer to each other than the observed rates for coastal points (-1.3 mmol m^{-3} per decade) and documented hypoxic areas (-0.8 mmol m^{-3} per decade) are to each other. Observations provide less coverage because observed trends with $p<0.10$ cover <77% of the points, while projected trends have $p<0.10$ in most coastal and hypoxic areas. This is a contributor to the difference mentioned between observations and projections, but there may be other reasons that may be found with additional analysis beyond what is included here.

L352-353, L370-371, L471-473: there are also considerable coastal points showing increased minimum O2 in Figure 6b; increased in coastal hypoxic area it might be due to shifted timing in summer hypoxia with shifted biogeochemical cycle. Therefore, selecting one summer month (Feb/Aug) might be problematic (include Jun-Aug summer months might be better).

RESPONSE: The analysis of the model data accounts for the possibility of variability in the month of minimum oxygen conditions (as described in the Methods) to accommodate the possibility of biogeochemical cycle shifts. So the analysis of model projections already accounts for the reviewer concerns. The median month for the oxygen minimum is August in the northern hemisphere and February in the southern hemisphere. These months were used in the observational analysis. The fixed months were used for the analysis because the observational dataset is much more cumbersome due to the much larger data size connected with the finer grid and daily (rather than monthly) original data. So the observational analysis does have a limitation in this sense. The paper focuses most on the projections. If the paper focused primarily on observations, I would have invested the large additional time and data storage to allow for variation of the month to track the oxygen minimum month each year at each location.

L365-366: good point! But why there is little difference between surface oxygen capacity decline and vertical minimum oxygen in the hypoxic area (both -1.4mmol/m3)?

RESPONSE: The documented hypoxic areas have a smaller difference between surface and vertical-minimum trends partially because the depths tend to be shallower in these areas than in the coastal points as a whole. I added a sentence too this point. There likely are other reasons that could be teased out by a detailed analysis of the model biogeochemistry in these areas. Such analysis is beyond the scope of the current paper, but should be interesting for future study.

REVISIONS: The difference between rates for vertical-minimum and surface conditions is smallest in documented hypoxic areas where waters generally are shallower than in the coastal and global categories. The differences between trends in vertical-minimum oxygen concentrations and surface oxygen capacity are partially accounted for by the weakening of warming trends with depth (mentioned above) and the temperature-dependence of oxygen capacity.

L379-380: this might be resulted from the same climate model family CESM.

RESPONSE: I agree. This part of the comparison is for individual runs within the CESM RCP8.5 ensemble. An assessment of this intra-ensemble variability was recommended in the first round of revisions.

L430-431: in this case, nutrient concentration and local hydrodynamics will play a more important role.

RESPONSE: I added a sentence to address this comment.

REVISIONS: In these high-latitude areas, as in other locations, specific nutrient patterns and local hydrodynamics will play important roles in where hypoxia ultimately develops.

REVIEWER 2

The author has made a substantial effort to do more analysis and expand the discussion.

RESPONSE: I appreciate the reviewer’s comments.

I just have the following suggestion for part of the discussion:
Lines 485-500: I would ask for a little editing to these statements:
First, while there have been major nutrient management efforts in the Baltic, Chesapeake, Mississippi, load reductions have been somewhat minor, and perhaps localized (Baltic and Chesapeake). I think it would be worth pointing out that loads are still relatively high in these places compared to historic levels.

RESPONSE: I added a sentence conveying this point.

REVISIONS: Even with load reduction efforts, nitrogen loads in these areas and many other coastal systems remain much higher than historic levels prior to large increases in human population.

I am not sure that there has really been a load reduction realized in the Mississippi, and for example, Scavia et al. 2019 does not discuss one. P loads have gone up in the Mississippi according to USGS you cited, and NO23 is stable, though TN may have gone down slightly. The statement on Mississippi load may need to be more nuanced.

RESPONSE: I split the sentence into two sentences and specifically referred to the slight total nitrogen load decrease.

REVISIONS: Nutrient loads from the Mississippi watershed feed eutrophication in the Gulf of Mexico coastal hypoxic zone (Scavia et al. 2019; Giudice et al., 2020; Stackpoole et al., 2021USGS, 2022). Despite some managed reductions in the total nitrogen load (Stackpoole et al., 2021; USGS, 2022), hypoxic extent has not consistently decreased and remains well above the management goal (Rabalais and Turner, 2019).

For the Chesapeake, I would cite Zhang’s papers about nutrient loads as opposed to the Chesapeake Bay Program.

RESPONSE: I added references to Zhang et al. (2015) and Zhang et al. (2020) and kept the Chesapeake Bay Program (2022) reference because it is continually updated. Adding the Zhang et al. references gives additional scholarly weight to the statement.

REVISIONS: Hypoxia in the Chesapeake Bay has not been reduced despite extensive nitrogen management and somewhat decreased nitrogen loads (Zhang et al., 2015; Murphy et al., 2011; Zhang et al., 2020; Maryland Department of Natural Resources, 2021; Chesapeake Bay Program, 2022).