Thanks to the authors for providing an updated manuscript. Many of the author's responses have clarified points for me, and after gaining a better understanding of the research I find I still have many questions about the work. Therefore I recommend that there be a second major revision for the manuscript. I hope that my points help make the manuscript more accessible to an interested audience, as I still believe there can be a lot of value in using IE for ocean biogeochemistry. I am a bit disappointed that some of my concerns have not been addressed in this new manuscript. I have repeated those concerns within my review.

Major points:

Overall manuscript

There still needs to be much serious consideration on how best to frame this work for its intended audience. I appreciate the consideration about the mathematics behind information entropy and the mapping procedure; however, *Biogeosciences* reaches a wide variety of scientists interested in the Earth system with varying backgrounds in mathematics. Some of the references, like those to topology, can be easily misinterpreted or misunderstood by readers unfamiliar with the mathematics. Even if the terms are correct they can present a barrier to the audience.

I wonder if the title can more clearly reflect how the work uses stratification as a proxy to understand ocean oxygen and its evolution. Currently the manuscript sounds like a model intercomparison which it is not.

There are numerous errors in language use, generally in the use of gerunds and matching up singular/plural nouns and verbs. It makes parts of the manuscript quite difficult to understand. I cannot provide every example but I would strongly recommend the authors conduct a thorough proofreading for the manuscript.

Introduction

There is no reference to current reconstruction schemes of ocean oxygen, or indeed the fact that observations of oxygen are sparse in space and time. As it is part of both the abstract and the discussion it should be included in the introduction.

I appreciate the highlighting of the hypotheses and the questions in this new version, but the paper should be framed such that it references, in order:

1. The aim/objectives of the paper
2. The questions the paper aims to answer
3. The hypotheses for these questions

Right now the order is objective -> hypotheses -> questions, which makes it difficult to read. Additionally I do not think that the hypotheses need to be referenced as HYP1/2/3, as this just makes the manuscript more difficult to read and reference.

There is no literature review on the use of IE in oceanography. Any citations about its application exist much later and they should absolutely be included in the introduction to provide justification for this work. In addition, there needs to be a section in which the use of IE for physical properties like temperature is compared to its potential use for ocean biogeochemical properties. If space is an issue, this is something that could replace parts of the beginning discussion on ocean oxygen cycling which is a bit high-level, has very few references, and also mentions quite a few biological processes which are ignored for the rest of the manuscript.

Lines 113-115 seem like an additional objective in the manuscript. If this is the case it needs to be reformatted.

Generally an introduction ends with an outline of the following sections. In the current introduction there is too much emphasis on the methods (which should be in the immediately-following section). A brief outline of what is included in the methods can be in the section introducing the following sections.

Materials and methods

Which historical experiments are being used? ESMs have emissions-based historical forcings and concentration-based historical forcings. In either case the radiative forcing is not necessarily the same between models with different atmospheric components and atmospheric chemistry.

It would help to include which variables you take from the CMIP6 models. Also, please define how you calculate density from temperature and salinity and how it is implemented in Equation 4. Is it using EOS80 or TEOS10? Are all model outputs gridded so that 0-200m are full cells, or is there some interpolation involved?

I still find the names of these indices (Ind1, Ind2) to be very difficult to internalize, both in the methods section and afterwards. I would strongly suggest eliminating the shorthand entirely. For the definitions on lines 191-193 I would strongly suggest using an overbar to indicate time averages rather than angle brackets.

I am still unsure about the arguments about the robustness of the extreme metric. If a 32 year time series of, say, winter stratification is used, the extreme is effectively the 97th percentile. This definition is necessarily dependent on both the length of the time series used and the number of ensemble members used. Additionally, is there a reason why you use seasonal averages to calculate the extremes, but then divide by the total number of months within that season? I am not sure if I am missing something major here but this continues to confuse me.

Results

The manuscript mentions that the hindcast “not surprisingly” displays the best fit for IPV and O2. This makes me wonder about the usefulness of a fully-coupled ESM compared to other ocean BGC models forced by reanalyses. What are the benefits of using a fully-coupled ESM? If the representation of the PDO is the largest issue in reproducing IPV and O2 (although I’m not convinced this is true) then I would assume this work would be more successful using ocean-only runs. Coupled models exhibit a “signal to noise paradox” (e.g., Scaife and Smith

2018, Zhang et al 2021) which results in an apparent decrease in predictability that may be restricted to model world. How would this impact the interpretability of your results?

I am a bit confused about the argument about how low-frequency PDO modulation alongside higher-frequency variability can produce low predictability. By that argument there is information associated with the PDO, but it is just more difficult to isolate. Is this not a case where the timescales of interest need to be controlled for? You use interannual changes in seasonal values for IPV and oxygen, but if we know the PDO to operate on decadal timescales, is there any way that the IE method can be altered to take these longer timescales into account?

Please ensure that all figures follow inclusivity guidelines. In particular Figures 4, 5, 9, and 13 need to be revised so that readers with deuteranomaly (red-green colourblindness) can interpret the figures.

Additional points on figure legibility:

- Figure 1: Please indicate which panels illustrate biases within the figure itself and ensure that “RMSE = ” is included in each subplot.
- Figure 2: The years for the historical and future runs should be included in the figure, as well as some reference to IE so that the reader can differentiate the figure from later figures. Also, the colorbar direction seems counterintuitive to me, as I would normally associate green with “good” (i.e., lots of information) and this is actually for high entropy.
- Figure 3: Change the x-ticks for the historical run to green and the future scenario run to black. Subpanel b needs a similar label to differentiate the green and black dashed lines.
- Figures 4 & 5: the contours for ENSO and PDO are difficult to separate for some panels. Please use dashed lines for one of the indices so that the reader can more easily discern which regions are related to ENSO and PDO.
- Figure 6: The years in the caption are wrong and they should reflect the two periods used in the averaging rather than combine them into one period.

Discussion/Conclusions

Thanks for including a discussion. I think there still needs to be many more citations. For instance, the first page cites only Ito et al 2019. On line 534-535 you reference a wide range of mechanisms but there is no relevant literature included.

In the discussion, the manuscript references that reconstructions need to interpolate onto a regular grid. This is fleshed out more in the authors’ response. This is not true – there is no inherent requirement to interpolate observations *or* innovations onto a regular grid when creating a reconstruction. There is utility in doing this, particularly as reconstructions are often used to compare or benchmark models, but any further assertion should be removed from the manuscript.

Please include a concluding paragraph that restates the objectives and main findings of the paper.

Minor points:

Line 20: “may or may not differ among the two fields, with a strong model dependency” what does this mean?

Line 39: Please include citations to OMZ literature.

Line 59-60: The reference to IPV is difficult to understand, as it seems to reference an equation but yet has no equation in the vicinity. I would recommend either moving the definition to the methods or having a more first-order reference to IPV here.

Line 63: It is not solely large-scale ocean currents that determines the pathways of IPV and O2.

Line 97: “For example, given a time series,...” this example is unclear. What do you mean by information? Please explain in terms applicable to a general Earth Science audience.

Line 103-104: “It also allows to investigate the network of domains…” if this work is not done in the manuscript I would recommend deleting this.

Line 142: There is no need to reference a specific figure in another work.

Line 206: Please ensure all equations are given their own line and equation number.

Line 249: What sort of remapping is used?

Line 382: “as to be expected” Please refrain from using asides in the manuscript.

Line 411: It is not 2015 values that you are comparing but a longer averaging period, correct?

Line 515: I would say “with good accuracy” is not necessarily correct.

Line 613-614: Please rephrase the sentence about how IE can be connected to ocean reconstructions. I do not feel it is appropriate to imply futility in existing scientific research.

References

Scaife, A.A., Smith, D. A signal-to-noise paradox in climate science. *npj Clim Atmos Sci* **1**, 28 (2018). https://doi.org/10.1038/s41612-018-0038-4

Zhang, W., Kirtman, B., Siqueira, L. *et al.* Understanding the signal-to-noise paradox in decadal climate predictability from CMIP5 and an eddying global coupled model. *Clim Dyn* **56**, 2895–2913 (2021). https://doi.org/10.1007/s00382-020-05621-8