

We thank both reviewers for their comments, which helped to improve the revised manuscript.

Reviewer(s)' Comments to Author:

Reviewer: 1

Abstract: The abstract begins with a reference to the “vertical expansion of the intermediate-depth low-oxygen zones”, a topic developed in the earlier Stramma et al. (2008) paper. The present manuscript does not really discuss “vertical expansion” and I suggest modifying this introductory abstract sentence so that it focuses on the present paper.

Thanks, good remark. The first sentence focusses now on the oxygen decrease instead of the vertical expansion.

Methods: I am not a specialist with the statistical methodology (a limitation discussed in advance by email with the editor). However, many readers of this paper are likely to also be non-specialists in this regard. In order to make this paper more comprehensible to the broader global change community, a simple non-technical summary paragraph of how the data were chosen and assembled to get a single point representing 1 year per depth zone and area, and the limitations of the dataset and methodology, would be valuable and would prevent mis-interpretation. Although I read the referenced methods in Schmidt et al. (2017), I still have questions (and most general readers would probably not refer back to that reference).

We added a more non-technical summary paragraph in the Data and methods chapter on page 5. With regard to limitation of the method also the text in the Discussion and Summary chapter was modified and more details added that are easier to comprehend the steps taken in this manuscript and in the references. See also our answer below.

Some methodological questions that could be addressed in a summary paragraph include: Were only individual bottle sample data used or were data from continuous electronic sensors from the last few decades included (vertical profiles from the many oceanographic instruments that include CTD sensors)? If electronic profiles were used, how were they condensed into 1 value for that depth interval? Were any data from Argo floats used since temporal and geographic coverage is expanding (although extended geographical oxygen coverage is only just beginning with the Biogeochemical Argo floats)? If these various electronic profilers were not used, a major source of modern data was ignored, and this limitation should be acknowledged. The discussion could include suggestions about how these continuous profiles might be applied to future trend calculations.

In the additional paragraph in the Data and methods section it is now described that bottle and CTD data are used and that float data were not included as measurements from our own floats showed drifts probably due to biological activity at the sensors which would lead to erroneous trends. The possibility to use float data in future once drifts can be removed is mentioned in the Discussion and Summary section.

The main focus of this manuscript is to compare the extended time period with the Stramma et al. (2008) results, this paper is even today highly cited. Therefore, we used the similar methods to make the results comparable (mentioned now in the additional paragraph in the data and Methods chapter).

CTD data are available in 1 dbar steps. We used only the CTD values 5 dbar apart to reduce the amount of data, as also the vertical gridding was made on 5 dbar steps. This information is now

included in the supplement text. Using an objective mapping scheme described in the manuscript the available data were mapped on 5 dbar steps and then the mean value and standard deviation for the two depth intervals 50 to 300 m and 300 to 700 m was computed. This is now explained in the text.

How many individual data points (or specific CTD casts) were included for each year? This is especially important for interpreting the earliest records. Could this be included in a supplementary table?

The data points and CTD profiles vary for each year in each area. A table with the numbers of data points would not show how large the bias of the computed mean oxygen value is. However, we mention now the possible influence of the fewer data point within a selected depth layer in the additional paragraph in “Data and methods”.

The authors do discuss how they accounted for chemical measurement improvements over this time period. However, since the 1950's, there have also been substantial improvements in location (satellite-based) and depth (electronic sensor) data. Was there any consideration for those sampling changes and uncertainty of those parameters in earlier data? That should at least be mentioned.

We mention now in the Data and methods paragraph the uncertainty caused by less accurate depth measurements from bottle data with CTD-depth measurements. As all measurements within each area independent from the geographical location were used, the better satellite derived location should not influence the results presented here.

The authors mention standard vertical depth levels (line 126). What are the depths and how were they used? As someone who deals with the small-scale vertical and temporal variability of oxygen through some of these depth strata, I remain confused about how that single value was derived, what it represents, and how robust that calculated value is. Is this supposed to be an indicator of the concentration of oxygen in the OMZ? Small changes in depth (and location and time) within their indicated depth zones can have quite different oxygen concentrations, as we now know from continuous profiling sensors and short-term replicated profiles. Since many of the long-term trends are slight, small changes in this value (for example from bottles offset by < 10 m or taken a few hrs apart) could have large consequences for the calculated trends. In the discussion, the authors should include more about the uncertainty associated with these issues, how to account for that, and how to improve future predictions.

In the revised version more information is presented with regard to uncertainties. The reviewer is correct that time of measurements and vertical resolution play a role for uncertainties of the trends. Never-the-less on the larger scale, due to internal waves, depth uncertainties and other factors we can assume these errors to be noise the no systematic bias over time. Thus, any trends derived will be less certain though not biased in decline or increase of the parameter analyzed. Furthermore as mentioned above we use the same methods as in Stramma et al. 2008 to make the results comparable. With CTD oxygen measurements on 1 dbar steps the uncertainties between years will be much less than in comparison to bottle oxygen measurements in earlier years. Nutrient trend computations could be improved if future nutrient measurements are made on defined standard depth levels to make the results better comparable. This is now explained in the Data and method section as well as in the supplementary file.

Results and Figures

Tables 1 and 2. It would be helpful to have more space between items or some other font indicator (bold?). It took several readings to notice the year listing since this blended in with the other numbers.

The years in Tables 1 and 2 are now bold, hence it should be well recognizable what are the trends and what are the time periods.

Fig. 2 and 3. The plots should be labelled A, B, C etc. There should be bigger tick marks for the decadal divisions on the x axis so those are visually differentiated from the climate index modal divisions. The caption should include a sentence noting that the X's in the plots represent an annual value (one X per year). There should also be a sentence to point out that the Y axes for oxygen (and nitrate in Fig. 4) differ in each plot.

In Ocean Science subsets of figures are labelled a, b, c (not A, B, C)

Bigger tick marks now show the decadal divisions on the x-axis.

The (x) is included in the figure legends and a sentence added to point out that the y-axes changes.

Fig 5. First line of the caption should delete "oxygen" and its units since this figure shows different variables in each graph.

Thanks, "oxygen" was replaced by "parameter" and the units were deleted in the first line of Fig. 5 caption.

It would be very interesting to see all the overall trend lines for a particular variable (oxygen in the two depth zones, nitrate) plotted on a single graph (one graph for each variable) with a single Y axis range. This would allow a visual comparison of the geographic variability in the strength of these trend lines. This could be a supplemental figure.

As supplementary figures the trends for both oxygen layers and the nitrate trends are presented.

Discussion This paper provides strong support for the trend of increasing deoxygenation worldwide at mid ocean depths. Many interesting possibilities are discussed to explain the trends in the different ocean areas and the interactions of different variables with regional physical oceanography, the climate modes, and each other. However, this also becomes confusing with details. It would be helpful to have some summary conclusions, perhaps a numbered list by geographic area. I would also like to see more discussion about the broader implications of these trends especially for ocean biology and human impacts. The authors could also bring in some mention of how their results could contribute in the future to some of the major climate change discussions and documents (IPCC report, UN Ocean Decade, etc.)

We included near the end of the manuscript a paragraph listing the observations by ocean basins. The implication of oxygen trends for biology and human impacts quite large and a lot of literature exist. All aspects of oxygen trends are discussed in the different chapters of the IUCN report (2019) and this is now mentioned and referenced at the end of the summary.

Minor wording comments: Line 47: sinking, transport, and subduction into the deep

ocean (the biological pump includes more processes than sinking) Line 73: extent Line 74: biologically Line 97: despite the fact that the low Line 434: nutrients Line 441: agrees with Line 443: sentence? Line 452: adding Line 463: trend-line; Line 478: areas. Hence, Line 484: appear Line 485: ocean; therefore, Line 495: inverse Line 496: observed; however, Line 519: fisheries

Thanks for pointing to the correct writing, all proposed changes were done in the modified manuscript.

We thank both reviewers for their comments, which helped to improve the revised manuscript.

Reviewer(s)' Comments to Author:

Reviewer: 2

The manuscript describes changes in dissolved oxygen and nutrients which might have various and important impacts on the ocean ecosystem. As the authors note, the manuscript confirms conclusions done in the previous work by Stramma et al (2008), which describes trends in the same selected geographical areas.

1) My main concerns regarding the significance of the presented results relates to the poor observational basis. In the concluding remarks the authors themselves acknowledge this fact and point to the necessity for further verification in the future.

The visual inspection of the calculated trends shows high variability of the yearly parameter concentration values imposed on a much weaker climatological signal. Some of the trends are not significant at 95 percent level. The others, even formally significant, leave the impression that the removal of just few data points would lead to significantly different trend estimates. Here, increasing the number of investigated areas of similar size, or increasing the size of the investigated areas could help to confirm (or not to confirm!) the robustness of the presented results.

The main focus of this manuscript is to compare the extended time period with the Stramma et al. (2008) results, this paper is even today highly cited. Therefore, we used similar methods to make the results fully comparable (mentioned now in the additional paragraph in the data and Methods chapter).

The data base for the manuscript was similar low as for the investigation of Stramma et al. (2008). Hence there is in this manuscript and in the earlier paper a larger uncertainty possible. For the 2008 paper later measurements in literature confirmed the decreasing oxygen trends. One focus of this manuscript is to compare the trends for the longer time period with the shorter one in the 2008-paper. For additional areas we would need an additional area with measurements covering a large part of the 1950 to present period, we are not aware of, as in the tropics there are very limited areas with a larger data base. Due to changing oxygen on the geographical locations an extension of the areas would include more uncertainties. Hence, we stayed with the same areas as in Stramma et al. (2008), but we describe these limitations more specific in the Chapters Data and methods, Discussion and Summary and in a supplementary file.

2) Using optimally averaged yearly values which in turn are used to estimate trends is justified, as the averaging procedure acts as a smoother. However, I would appreciate the elaboration about the possible errors related to the mapping scheme, and, because

of the data paucity, even these averaged data might be linked to relatively large errors.

The reviewer is correct, this procedure smoothes the dataset to a degree. We added more information to the supplemental methods discussing the possible errors. In general, this mapping scheme is quite conservative and is more likely to show no trend than a trend for two reasons. First with a non-smoothed data set an oversampled anomalous year has significant impact on the whole timeseries, second with lower overall smoothed data points the uncertainty of a trend analysis is larger. With those points in mind, one can assume that the results, if statistically significant are a robust find.

3) The authors do provide measurement precisions for the modern measurements and offset estimates for the older data. The reference is done to the paper by Tahua et al., 2010.

Even for temperature which is easier to measure compared to other parameters, systematic instrumental errors pose a big problem in estimating the ocean heat content changes. Possible instrumental biases in oxygen and nutrient data is even a bigger issue, which was treated, for instance, during the WOCE time. Here, references to Johnson et al., 2001, and to Gouretski and Jancke, 2001 could be added to the reference list. From the manuscript text it is not clear, whether the original data were corrected for systematic biases, or not. More discussion of possible bias impact is needed.

The offsets derived from an inter-cruise comparison described by Gouretski and Jancke (2001) is now included as well as the initial standard deviations of cross-over differences from Johnson et al. 2001. These papers are now used to describe a possible bias from the measurements. For heat content the systematic bias is one directional and addressed mainly expandable thermographs, such data is not used in this analysis. In the case of our data, we have evidence as referenced that the bias is more likely to be noise than systematic, thus not impacting any trend analysis.

4) Several areas show a certain shift in oxygen concentration for the data after 2000 (Fig.2a,b,d,f; Fig.3 a,e) Could these change in concentration be due to unexplored biases?

We think that these changes in oxygen concentration are related to the climate signals changing in all three oceans shortly before the year 2000. However, we included a discussion related to these changes in the supplementary text.

5) it is interesting to know, how sensitive the calculated trends are to the thickness of the layers (the fixed thickness of 250 and 400 meters are used for the presented results).

As mentioned above, one focus of this manuscript is a comparison with the Stramma et al. (2008) paper, therefore the layer 300 to 700 m was selected again. Depending on the parameter gradients at the boundaries of the layer, the trends will be different. However, as the oxygen trends for the 50 to 300 m layer and the 300 to 700 m are all negative (except for the 50 to 300 m layer of area C due to a local effect) the result of oxygen decrease is not related to the depth layer chosen. This is now described in the Discussion and Summary chapter.

6) Can the yearly parameter values be seasonally biased considering the poorness of the data basis??

As the annual data were computed independent of the season due to the pooriness of the data basis, a seasonal influence might be possible. However, as the seasonal cycle in the tropics is weaker than in most subtropical and subpolar regions (Louanchi and Najjar, 2000) we mention it near the end of the Introduction and now also in the supplementary text.

Minor comments:

1) I do not see much use in presenting mean parameter values for the investigated areas without providing the number of available original profiles and standard deviations (Table 3)

The mean parameter values for a depth layer and year are used to compute the mean parameter values for the time period covered and here the number of years and the mean standard deviation are presented. The standard deviation is often large, as it is related to the variability of the annual mean parameter value and the strength of the trend during the measurement period, which is now explained in the text.

As for each year a different amount of data is available, computing the standard deviation for each year would provide only information with regard to variability within a year, but not for the mean values presented in table 3. The fact that the mean parameter is derived from the annual values is now mentioned in Table 3.

2) Please, indicate in the figure captions, that crosses denote the annual parameter values used to calculate trends

Thanks, it is now mentioned that the annual parameter values are used to calculate trends.

Line 13: change indicates to indicate

As proposed changed to 'indicate'

Lines 125-126: probably the word "interpolated" is missing after the word "profiles"
(Line 126)

Thank you, word "interpolated" added

Line 267: change nutrient to nutrients

As proposed changed to 'nutrients'