Dear Reviewers,

We are very grateful to you for your extraordinarily detailed and insightful review of our manuscript. We really appreciate the time you invested. We find your comments invaluable in helping us to improve the quality of the submitted manuscript. We believe that we could properly address all issues raised in your review. Please find below our point-by-point response. Your questions/comments are written in black; our response is given in blue; our revised text in the manuscript is indicated in orange.

Response to Reviewer #1:

Quality control is a crucial step in handling observational datasets. However, the quality control routine alone does not suffice for a data paper. The revised manuscript still lacks a detailed analysis on how the quality-controlled data influences deoxygenation or the oxygen budget at regional and global scales.

Re: Thanks. To address the issue, we have included a new section to show the impacts of QC and bias adjustment on estimating ocean oxygen mean state and its changes (annual cycle and long-term changes) after applying a previously published mapping approach (Cheng et al. 2017, 2020). Although we think it deserves further study to fully evaluate the application of this mapping approach in oxygen reconstruction, we feel adding this section strengthens the study by illustrating how the quality-controlled data influences deoxygenation or the oxygen budget at regional and global scales, as you are suggesting.
We copy the section here for your reference:

7. Impact of quality control and bias adjustment on estimating oxygen changes

Applying the QC and bias adjustment to historical *in situ* oxygen data is expected to impact the derived ocean oxygen changes on various spatial/temporal scales. To illustrate this impact, we implemented the new Auto-QC system for all oxygen data and adjusted the Argo data based on the approach described in Section 6. Based on these data, we applied the mapping method (Ensemble Optimal Interpolation approach with a Dynamic Ensemble from climate model simulations, EnOI-DE) proposed by Cheng et al. (2017, 2020) to spatially interpolate oxygen data, yielding a spatially complete gridded global ocean oxygen dataset. Because of the limited spatial coverage of oxygen data, we combine each successive three years of data to derive oxygen fields for each calendar year. Respectively, the oxygen time series are based on these fields. The reconstruction is only done for the upper 2000 m because of the insufficient in situ data in the abyssal layers. The resultant oxygen field is denoted as “after QC/adjustment”. To show the impact of QC and adjustment on the oxygen changes estimate, we also applied the same method to the data without QC (e.g. with only several crude QC checks applied to remove most likely erroneous values, including overall range checks, solubility check, and spike check) and without Argo adjustments. The resultant field is denoted as "before QC/adjustment".

The long-term mean states (e.g., the climatology, reconstructed using all data between 1990-2022 based on EnOI-DE approach) of the upper 1000 m oxygen before and after QC/adjustment are very similar (**Figs. 23a, b**). One reason is the EnOI-DE method (as any mapping approach) has a smoothing effect, so the erroneous data is less visible behind high spatial variability. This indicates the robust large-scale pattern, where the oceans in the low latitudes have lower oxygen concentrations than in the higher latitudes because of the water temperature and ocean circulation difference. The Eastern Pacific and North Indian Oceans show even lower oxygen levels because of the subsurface oxygen minimum zone. The difference between oxygen climatologies calculated before and after QC/adjustment ranges from -15~15 μmol kg^{-1} but differs at different locations (**Fig. 23c**). The zonal mean difference is smaller (-3~1 μmol kg^{-1}) because of the error cancellation at each latitude (**Fig. 23d**).

The QC/adjustment also impacts the annual cycle (including both phase and magnitude) of the global mean oxygen changes (**Fig. 23e**). Examples for the layers 0 – 100 m (representing the upper seasonal change layer), 100 – 600 m (representing the main thermocline) and 0 – 2000 m (showing the ocean oxygen inventory) are shown in **Fig. 23e**. For 0 – 100 m, the mean oxygen level shifts from negative to positive in November after QC/adjustment but in September before QC/adjustment. The magnitude of the annual cycle, if simply defined as the difference between the maximum and minimum of the 12-month climatology time series, is 1.45 μmol kg^{-1} but slightly reduced after QC/adjustment (1.22 μmol kg^{-1}). The magnitude of the 100 – 600 m and 0 – 2000 m annual cycle has also been reduced after QC/adjustment (1.18, 0.55 μmol kg^{-1} before QC/adjustment and 0.79, 0.48 μmol kg^{-1} for 100 – 600 m and 0 – 2000 m, respectively, **Fig. 23e**).

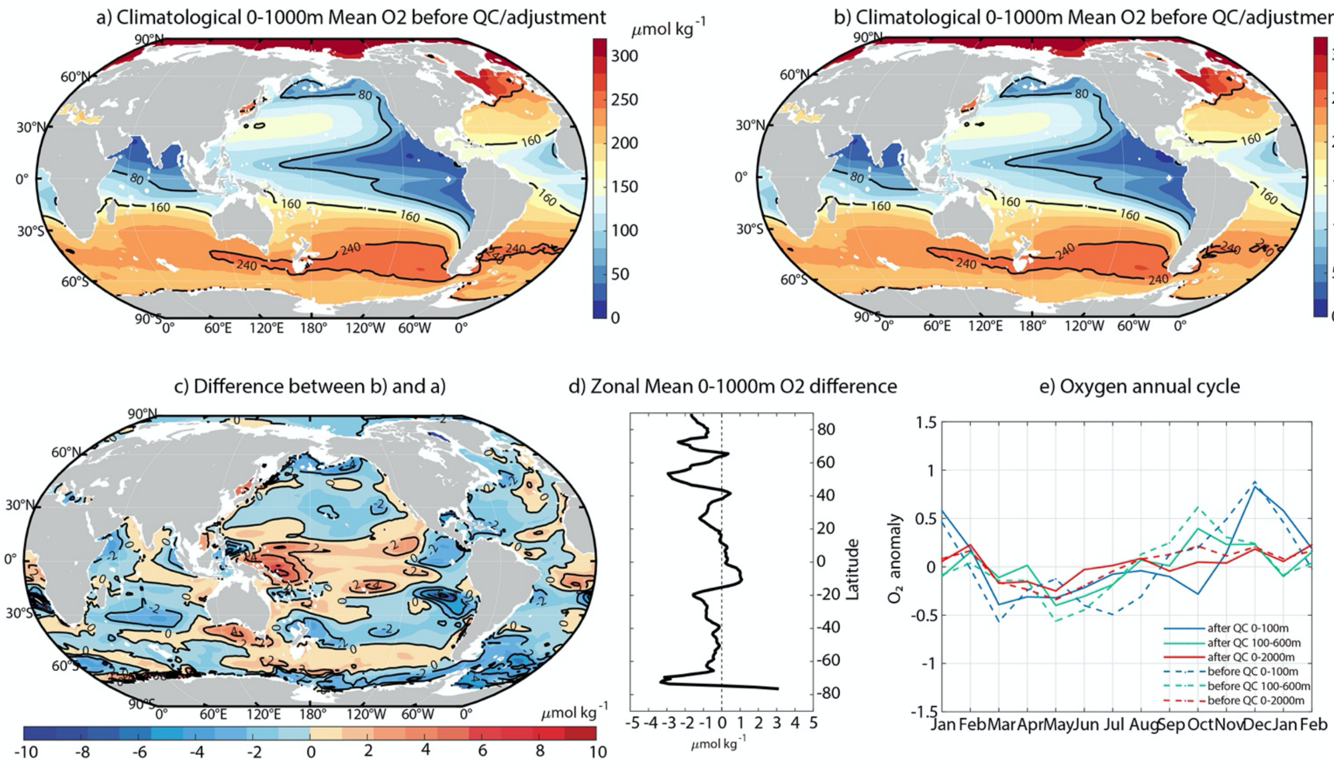


Figure 23. The climatological upper 1000 m oxygen field before (a) and after (b) QC/adjustment, with their spatial difference shown in (c) and zonal mean differences in (d).

The annual cycle (relative to the climatological annual mean level) before (dashed line) and after (solid line) QC/adjustment are compared in (e) for different vertical layers. The climatology field is reconstructed by combining all data within 1990-2022 with EnOI-DE mapping method (Cheng et al. 2017, 2020).

The QC and adjustment also impact the estimates of long-term oxygen changes, for example the global deoxygenation estimates for 0 – 100 m, 100 – 600 m and 0 – 2000 m layers depicted in **Fig. 24**. After QC/adjustment, the standard deviation of the time series is decreased from 1.71 (0 – 100m), 2.37 (100 – 600m), 1.60 (0 – 2000m) to 1.62 (0 – 100m), 2.24 (100 – 600m), 1.44 (0 – 2000m) µmol kg^{-1}, showing a reduced variability in global oxygen time series after QC/adjustment. This indicates a reduction of noise, which is mainly attributed to both QC and Argo adjustment. For example, before QC/adjustment, there was a big global 0-100m deoxygenation of ~ 3 µmol kg^{-1} from 1995 to 1996, which is likely non-physical and spurious. Such change disappeared after QC/adjustment (**Fig. 24**). The linear rate of deoxygenation differs for the two tests as well: -0.77 ± 0.43 (0 – 100m), -1.45 ± 0.30 (100 – 600m), -0.95 ± 0.30 (0 – 2000m) µmol kg^{-1} dec^{-1} before QC/adjustment and -0.90 ± 0.38 (0 – 100m), -1.37 ± 0.40 (100 – 600m), -0.84 ± 0.41 (0 – 2000m) µmol kg^{-1} dec^{-1} after QC/adjustment. The linear trend is calculated by the ordinary least square regression with a 90% confidence interval shown (accounting for the reduction in degree of freedom). The deoxygenation rates are reduced after QC/adjustment for both 100 – 600m and 0 – 2000m, mainly because of the Argo adjustment, which shifted the oxygen level in the past decade by ~0.76 µmol kg^{-1} for 100 – 600 m average and ~0.82 µmol kg^{-1} for 0 – 2000 m average within 2015-2023 (**Fig. 24**).

By means of these tests we demonstrate that QC and bias adjustment can impact the estimation of the oxygen changes at various temporal-spatial scales, highlighting the need for careful oxygen data processing before application. However, we note here that the validity of the mapping approach on oxygen reconstruction has not been thoroughly evaluated, which deserves a separate study.

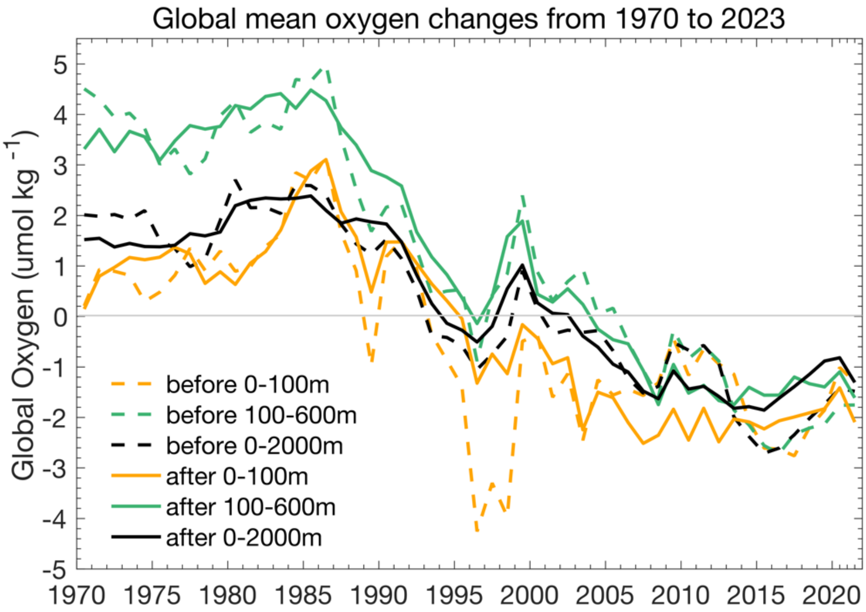


Figure 24. The reconstructed global averaged oxygen time series before (dashed line) and after (solid line) QC/adjustment from 1970 to 2023 for the layers 0 – 100 m, 100 – 600 m and 0 – 2000 m. Here, we combine each successive three years of data to estimate the oxygen changes. The anomalies are calculated relative to the climatology shown in Fig. 23.

Response to Reviewer #2:

General Comments:

Section 3.4, Stuck value is a profile with a standard deviation less than a certain threshold (e.g., 3 umol/kg for CTD data). Is this threshold applied to the whole profile? or in depth increments. Main text suggests whole profile while caption in supplemental figure suggests this was applied to 100-m data bins.

Re: Your understanding of our stuck value check is correct. The standard deviation is calculated for each profile using data from all observed levels, and the decision to flag it is made by comparing the standard deviation with the threshold value.

The indicated bin size has nothing to do with this quality check. Figure S5 caption just provided information about the bin size for the two-dimensional plot (e.g. percentage of outliers versus depth and year, plates b) and e)). To avoid confusion, we deleted information about the bin size from the figure caption.

The stuck value check section has been improved, which is pasted here, we hope it is clearer now: "Malfunctioning of sensors often results in stuck values when the same oxygen concentration is reported for all or most of the observed levels. To identify such profiles, we calculated oxygen standard deviations for each oxygen profile to build histograms (Fig. 6) for each instrumentation type. Only profiles with at least seven oxygen levels are considered. Unlike the OSD and Argo data, for which the frequency of profiles drops for low standard deviation values, the CTD profiles are characterized by a distinct peak for the lowest standard deviation values (Fig. 6c). Accordingly, based on the histograms (Fig. 6b, c), we set the thresholds of 3 μmol kg^{-1} and 1 μmol kg^{-1} for CTD and Argo profiles, respectively. No lowest value thresholds are applied for OSD profiles, as stuck values are only characteristics of the electronic sensors. The geographical distribution of profiles failing this check is given in Fig. S5 a, d. The check is applied only to the CTD and Argo sensor data and reveals a high percentage of outliers for CTD profiles, especially after 2000 (Fig. S5b). Argo profiles which fail the check are not numerous and are located mostly in the Northern Hemisphere (Fig. S5d)."

Unclear to me how over half the CTD measurements would fail the local range test. It seems like the local range may be poorly defined if this is the case.

Re: Thanks for the concern, and we offer more explanations here:

1) After numerous trials of our codes and applying the quality control procedure to a number of quite different sets of oxygen data, we are confident that the local climatological thresholds

have been calculated correctly. Please note that the same thresholds are applied to all subsets of oxygen data analyzed in the manuscript: Winkler, CTD, and Argo.

2) Many CTD profiles are rejected through stuck value check because oxygen values at all levels are identical or "almost identical". We have checked many of these CTD profiles and have confirmed that they are what we got from the WOD. **We communicated with the WOD/NCEI/NOAA group; we confirm that the CTD data issue is not an issue of the WOD archive. Instead, it is the issue of the data when they are submitted to WOD. WOD keeps the data as they are flowing in; WOD doesn't change or alter the data.**
3) The quality checks in our automated procedure are independent of each other. In case of CTD data a significant fraction of oxygen profiles is flagged not only because of failing the stuck value check (65.5%) but also because the same profiles also fail the local climatological range check (61.5%). Please see the outlier statistics in Table 2. Therefore, other checks also confirm that these CTD profiles are likely erroneous.

Furthermore, we'd also like to introduce our additional modification in our revised manuscript: Based on a thorough discussion among the co-authors of the manuscript and with several colleagues working with Argo oxygen data, we decided to present biases for distinct sensor models and DACs. This is a deviation from the previous version of the manuscript, where bias assessment was made only for distinct DACs, ignoring different oxygen sensor models implemented on BGC Argo floats. Respectively, the former Section 6.3 ("Residual Oxygen Biases for Argo profiles from distinct DACs") is now Section "6.3 Residual Oxygen Biases for distinct oxygen sensor".

In this section, we included Table 3, which gives the names of different sensors and shows the respective number of Argo profiles equipped with each sensor. Figure 18 has also been changed. Now, it shows the yearly number of Argo profiles (for different sensors and for different DACs) that have collocations with the reference data.

Fig. 19 shows overall oxygen biases estimated for distinct sensors and DACs. Fig. 20 aims to demonstrate the relative stability of the diagnosed offsets over the observation period. The recommended overall bias corrections are summarized in Table 4.

The former Fig. 19 is now represented by Fig.21, where residual biases are shown for eight DACs before and after application of our bias corrections.

In addition to the abovementioned changes, we added a new section: "7. Impact of quality control and bias adjustment on estimating oxygen changes". This was done following the suggestion of the other reviewer to show the impact of our QC and offset adjustment on estimating oxygen changes from the annual cycle to multi-decadal trends/variability.

The new parts of the text are highlighted in cyan in the manuscript.

Line Comments:

In Section 6 you introduce the bias for both CTD-DO and Argo instrumentation so Section 7 should really be section 6.4 as it continues the Section 6 topic.

Re: We followed your suggestion: former Section 7 is now Section "6.4 Residual Oxygen Biases for CTD oxygen sensors".

Data availability section is numbered incorrectly.
Re: Following the other reviewer's advice, we added a new Section7: "Impact of quality control and bias adjustments on estimating oxygen changes". In accordance with the changes explained above, the "Conclusion and Discussion" section has now become number 8, and the "Data Availability section" is number 9.

Data availability should contain links for the original CTD-DO, Winkler and ARGO datasets accessed.
Re: Added:" The quality control procedure described above was applied to the OSD and CTD oxygen profiles between 1920 and 2023 from the World Ocean Database (https://www.ncei.noaa.gov/access/world-ocean-database-select/dbsearch.html) and to the oxygen profiles from the BGC Argo floats (https://www.seanoe.org/data/00311/42182/).".

Code should also be made available.
Re: A new section, "10 Code availability," has been added. Here, we provided the FORTRAN code, which performs automated quality control on sample input observed data. We hope the open access to codes will help users, and we are happy to resolve any issues users may encounter in the future.

Line Comments:

Line 51: AQC_FINAL_CTD.f ?
Re: This sequence was deleted

Line 52: sensor drift is due to fouling and electrolyte consumption
Re: we change the wording according to your suggestion

Line 82: is not isd

Re: “isd” changed to “is”

Line 92: BGC not BGH

Re: BGH changed to BGC

Line 100-101: World Ocean Database (WOD)

Re: the abbreviation WOD was introduced after “World Ocean Database”

Line 101: oxygen not Oxygen

Re: Oxygen changed to oxygen

Line 101: Argo Global Data Assembly Center? not Argo Global Assembly Center

Re: thank you for noting this: the word “Data” is inserted

Line 102: depository not depositary

Re: the word was corrected

Line 125-128: Write out DAC names before using abbreviations

Re: we introduced the respective changes: the full name is now provided before each name abbreviation.

Line 134: Bushnell et al. (2015) not Bushnell et al., (2015)

Re: corrected: comma deleted

Line 192: should this be 2000 – 2014?

Re: yes: 2000-2014 is correct, we changed “200-2014” to “2000-2014”

Line 243: What is PFL?

Re: The abbreviation PFL goes back to the World Ocean Data Base and means “Profiling Floats” The PFL oxygen profiles from the World Ocean Dataset were used at the early stage of the manuscript preparation. Later, we switched to using Argo oxygen profiles obtained from DACs instead of PFL. **We apologize for this inconsistency!** The abbreviation PFL was also used at two other places in the text: in Section 3.4 and the figure caption to Fig. 7. In all cases, the abbreviation PFL was substituted by the word “Argo”. The respective change was also introduced in the Fig.7.

Line 245: I think you mean to refer to supplemental figure 5.

Re: Yes, indeed: Fig. S5 is meant. We introduced respective corrections.

Line 262: What is PFL?

Re: See our response to Line 243. PFL was changed to Argo

Line 378: AQC?

Re: We apologize for overlooking that. AQC means "automated quality control". The respective change has been done.

Line 430: remove KIO3 and just say standard reference

Re: we removed KIO3

Line 437: program not programme

Re: we changed programme to program.

Line 568-570: I disagree and think that CTD oxygen data are typically submitted uncalibrated. From the WOD documentation: "Note that in many cases the dissolved oxygen and chlorophyll data are uncalibrated and not of high quality. Information on whether these variables are calibrated is not usually supplied by the data submitter (see Chapter 3.)"

The fact that CTD-DO data is often uncalibrated should be made clear throughout the document.

Re: We have introduced the respective changes in the text. We now cite the statement (pointed to in your review) by Boyer et al., 2018 (page 20). We removed our previous statement about the majority of CTD data being calibrated. Further, we contacted the WOD team again to clarify where the high percentage of unrealistic CTD oxygen values comes from. They had checked and confirmed first that it is not the WOD's issue; instead, the erroneous profiles are the data submitted by the data submitters

With respect to the data you found below, Tim Boyer from WOD said "*I would need some examples to see what the reviewer found, but I expect the integer oxygen data found by the reviewer are oxygen data which have been converted from ml/l to umol/kg. Number of significant figures don't change on conversion so a value of X.XX ml/l with 3 significant figures would be YYY umol/kg with 3 significant figures. We write out data in csv file to the number of significant figures given. One of the defects of netcdf is you cant easily do that - so the full number is written out, the significant figures is another field which must be applied by the user.*".

We will continue working with the WOD group to resolve the data issues. This will be a major international effort that consumes time, so it is good that our paper identifies this potential issue.

Figures:

Figure 7: What is PFL?

Re: See our comments to Line 243 and 262. PFL has been changed to Argo.

Figure 10: b and c are flipped in the caption.

Re: We have changed the labels in Figure caption.

Figure 11: Percent not percen and the caption shouldn't be split above and below the figure.

Re: “percen” changed ‘percent’. The caption is not splitted now.

Figure 18: What are AROD FT and ARO FT referring to?

Re: Fig. 18 has been changed in the revised version of the manuscript. The figure now shows the yearly number of Argo profiles according to the sensor model and DAC. Table 3 now gives the names of sensor models and shows the number of profiles available for each sensor model. AROD_FT and ARO_FT are optoid oxygen sensors—this is indicated now in Table 3.

Figure 20: Panel J is missing

Re: Thank you for noting this! This figure is not shown in the revised version of the manuscript.

Figure 22: Panel d) oxygen units are wrong

Re: we have changed the units for the color bar

Supplemental Figures:

y-label is offset in all e) panels of supplemental figures.

Re: Thank you for noting that we have corrected all supplementary figures.

Review of data product:

- Appreciate multiple file formats.

In my review of this dataset I looked further into the WOD CTD dataset to investigate why there was such a high percentage of CTD measurements that failed the stuck value test. Ultimately, I agree with the authors’ conclusion that the WOD archive suffers from major quality issues. When looking at the original oxygen data downloaded from the WOD database (at least the files I looked at from Nov 1-30 2021), it seems like there is an issue with the resolution for the oxygen concentrations. When downloading *csv files from WOD, oxygen concentrations are reported in integer values that are spaced by 3 umol/kg. When downloading the NetCDFs for this same time period directly from the WOD, all oxygen values are spaced by 3.0489 umol/kg I was expecting oxygen concentrations to be reported to at least one decimal place. This seems to be an issue with the WOD oxygen product. Have the authors reached out to managers of the WOD product to see if this is an artifact of how the data is being downloaded or something else they may be able to help troubleshoot? I suspect that the poor quality CTD data is failing the QC stuck value test but it isn’t really a stuck value, seems to be a bigger more systematic WOD issue.

Re: Thanks for the confirmation of the CTD data issue! That is an important independent confirmation.

-I also found a number of issues with the data set provided with the manuscript.
Re: we apologize for overlooking some issues in the dataset; we have carefully checked the dataset now to make sure everything is correct now.

Read me file is empty.
Re: Now the file is available, and users can download it.

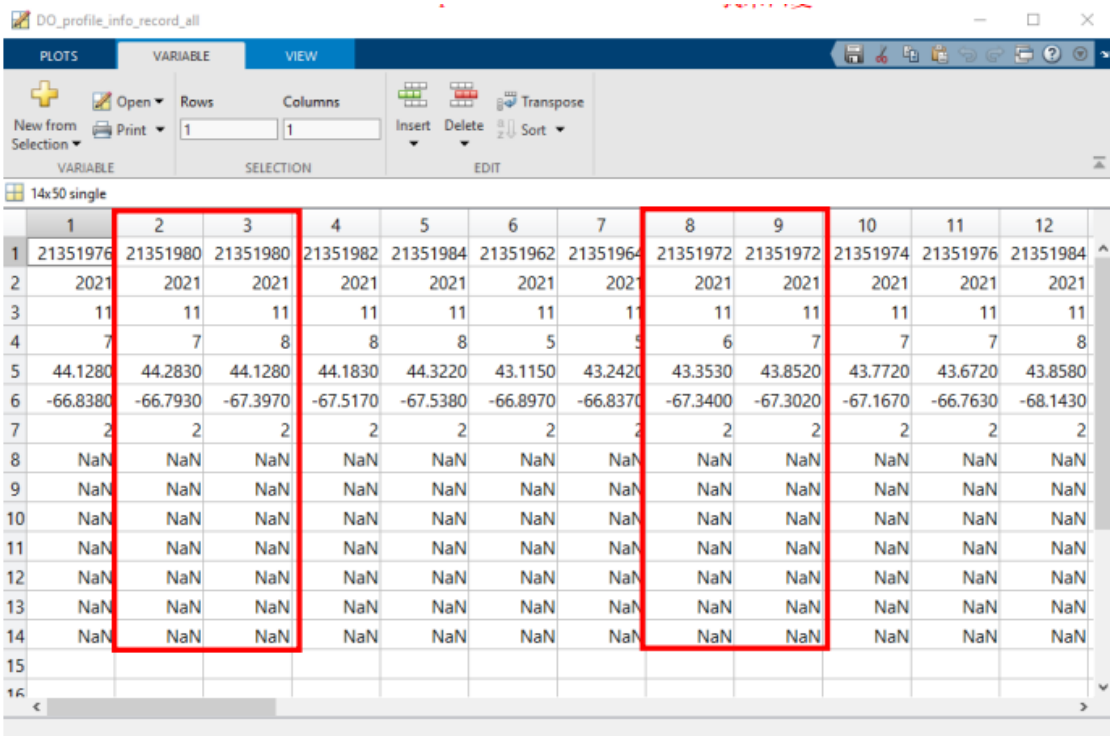
- Incomplete list of netcdf files for OSD CTD data, ends in 2015.
Re: We corrected the list and re-uploaded all files to the data centre (DOI: http://dx.doi.org/10.12157/IOCAS.20231208.001)
As an alternative, we also provided another link (http://www.ocean.iap.ac.cn/ftp/cheng/IAP_oxygen_profile_dataset/) to assess this dataset.

-NetCDF metadata for OSD_CTD files refer to high-resolution CTD data (original data) as raw data. Raw data is in instrument units (e.g., volts, counts). These data have been processed to scientific units using CTD calibration coefficients and some rare cases further calibrated using discrete samples. They should not be referred to as raw.
Re: We agree. To avoid confusion, we decided to remove the description of 'raw data' and just keep it as 'original data' (i.e., downloaded directly from GDAC or WOD). We have also updated all corresponding netCDF metadata descriptions ('long_name').

-NetCDFs metadata list QC 9 checks in this order: 'Geographical Location Check, Crude range check, Maximum oxygen solubility check, Stucked value check, Spike check, Multiple extrema check, Oxygen Vertical Gradient check, Local Climatological range check, Excessive flagged level percentage check' while the table in the main manuscript lists 10 QC check in a different order. Also Missing Global Oxygen on T surfaces QC check in metadata in NetCDF files. Not sure which levels of QC are actually being referenced in the NetCDF files.
Re: Thanks for your efforts to go deep into the dataset. Sorry for the mistake. The QC system includes 10 QC checks in total, as indicated in Table 2 of the manuscript. The corrected order of these 10 checks is: Location check, Global Oxygen Range at depth levels; Global Oxygen Range on T surfaces; Maximum oxygen solubility check; Stuck value check; Multiple extrema check; Spike check; Local climatological oxygen range check; Local climatological oxygen vertical gradient range check; Excessive flagged level percentage check (same as the Table 2 in the manuscript). We note however that the checks are independent of each other. We have corrected all these mistakes and ensure the consistency between the netCDF files and MAT files.

-Another issue in the *mat file is the referencing of the original WOD data files which are indexed by unique WOD ids. For reference I looked at the mat file for CAS_Oxygen_CTD_OSD_2021_11. While investigating the *mat files I found that the variable DO_profile_info_record_all should contain a unique WOD_unique_ID for each profile. WOD_unique_ID numbers are not unique and the same WOD_unique_ID numbers refer to multiple profiles at different lat/lon coordinates. I assume these WOD_unique_ID numbers are supposed to line up with WOD_unique_ID numbers in the DO_profile_info_str_all variable. They do not. I don't know if this issue persists in the netcdfs provided with this manuscript because there was no accompanying netcdf for this mat file.

Re: We apologize again for this mistake. We double-checked the 'WOD_unique_id' and found the numbers in the 'DO_profile_info_str_all' variable are correct, and the numbers in 'DO_profile_info_record_all' are wrong. This is because of a bug in our data formatting codes. Therefore, we debugged and the 'DO_profile_info_record_all' fields are correct (Fig. A1). Both mat and netCDF files are updated and corrected. Now, the WOD_unique_ID numbers are unique across all the data files.



	1	2	3	4	5	6	7	8	9	10	11	12
1	21351976	21351980	21351980	21351982	21351984	21351962	21351964	21351972	21351972	21351974	21351976	21351984
2	2021	2021	2021	2021	2021	2021	202	2021	2021	2021	2021	2021
3	11	11	11	11	11	11	11	11	11	11	11	11
4	7	7	8	8	8	5	5	6	7	7	7	8
5	44.1280	44.2830	44.1280	44.1830	44.3220	43.1150	43.2420	43.3530	43.8520	43.7720	43.6720	43.8580
6	-66.8380	-66.7930	-67.3970	-67.5170	-67.5380	-66.8970	-66.8370	-67.3400	-67.3020	-67.1670	-66.7630	-68.1430
7	2	2	2	2	2	2	2	2	2	2	2	2
8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

Screenshot showing example of repeating WOD_unique_IDs (Row 1) for different profiles lat/lons (Rows 5 and 6).

变量 - DO_profile_info_record_all

DO_profile_info_record_all

14x50 double

	1	2	3	4	5	6	7	8	9	10	11
1	21351977	21351979	21351981	21351982	21351983	21351962	21351963	21351972	21351973	21351974	2135197
2	2021	2021	2021	2021	2021	2021	2021	2021	2021	2021	202
3	11	11	11	11	11	11	11	11	11	11	1
4	7	7	8	8	8	5	5	6	7	7	
5	44.1280	44.2830	44.1280	44.1830	44.3220	43.1150	43.2420	43.3530	43.8520	43.7720	43.672
6	-66.8380	-66.7930	-67.3970	-67.5170	-67.5380	-66.8970	-66.8370	-67.3400	-67.3020	-67.1670	-66.763
7	2	2	2	2	2	2	2	2	2	2	
8	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Na
9	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Na
10	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Na
11	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Na
12	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Na
13	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Na
14	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	Na
15											
16											
17											
18											
19											
20											
21											
22											
23											
24											
25											
26											

Fig. A1. The same screenshot as the above figure, but it is for the revised data version. The CAS_Oxygen_CTD_OSD_2021_11.mat are used here.

Additionally, the data associated with WOD_unique_IDs in the dataset provided with the manuscript does not match the data associated with the WOD_unique_IDs when downloaded directly from WOD (Figure 1).

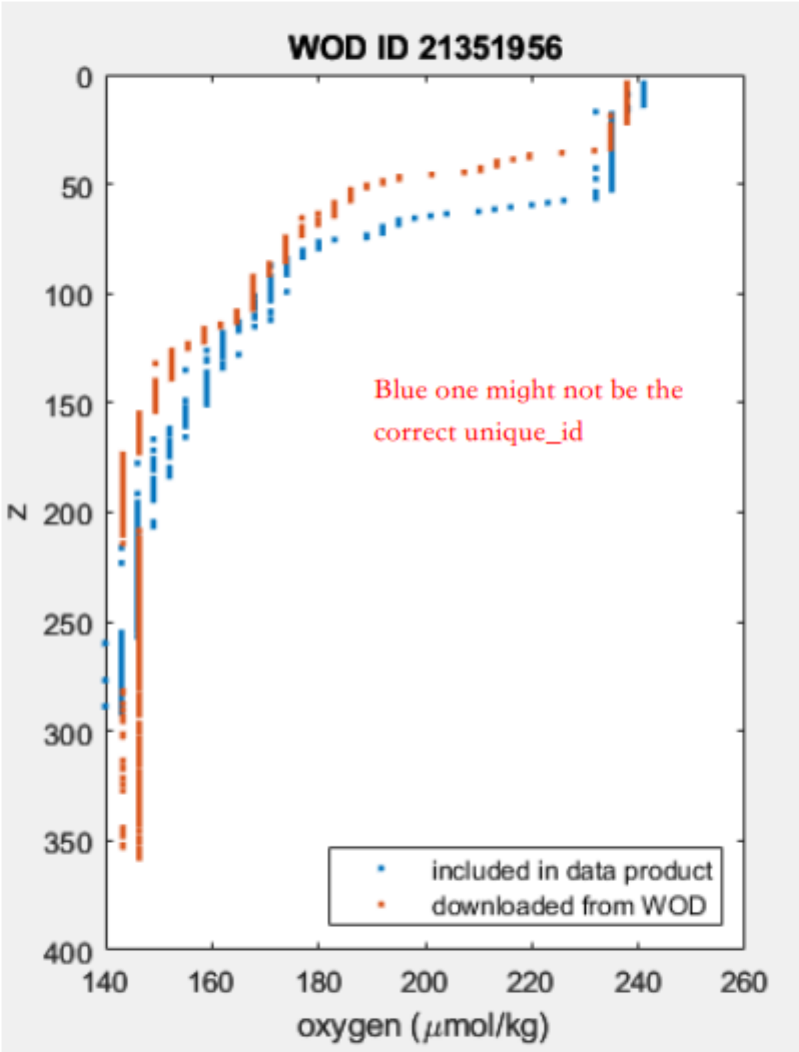


Figure 1 showing data labeled WOD ID 21351956 from the data set included in this manuscript (mat file) versus the data for WOD ID 21351956 downloaded directly as a nedcdf from WOD. Also note oxygen values can only be assigned to values every 3 umol (this issue seems to originate in the WOD data, not authors' data product).

Thanks for pointing to this error. We found this happened due to a bug in our codes (as discussed in the last paragraph, the ID in our previous dataset is not unique). We have corrected this error, and now, data associated with WOD_unique_IDs in our dataset matches the data downloaded directly from WOD (Fig. A2). All the profiles have been double-checked. Interestingly, we helped WOD/NCEI group to identify an issue in WOD data: the data downloaded as WOD native ASCII format and as NetCDF format are actually not consistent because of the different data processing procedures in WOD (some data will be rounded off in one version but not in the other, because of the unit change process). We (in this paper) now use the WOD native ASCII format version of WOD data, and we are now working with the WOD group to resolve this issue.

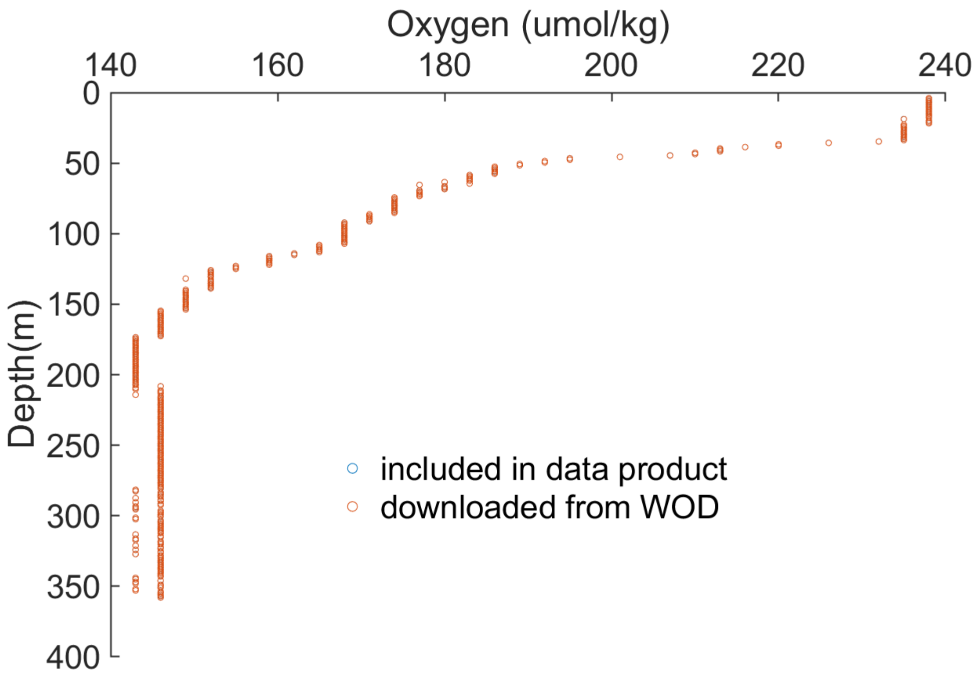


Fig. A2. WOD ID 21351956 from the revised data set included in this manuscript versus the data for WOD ID 21351956 downloaded directly as a nedcdf from WOD.

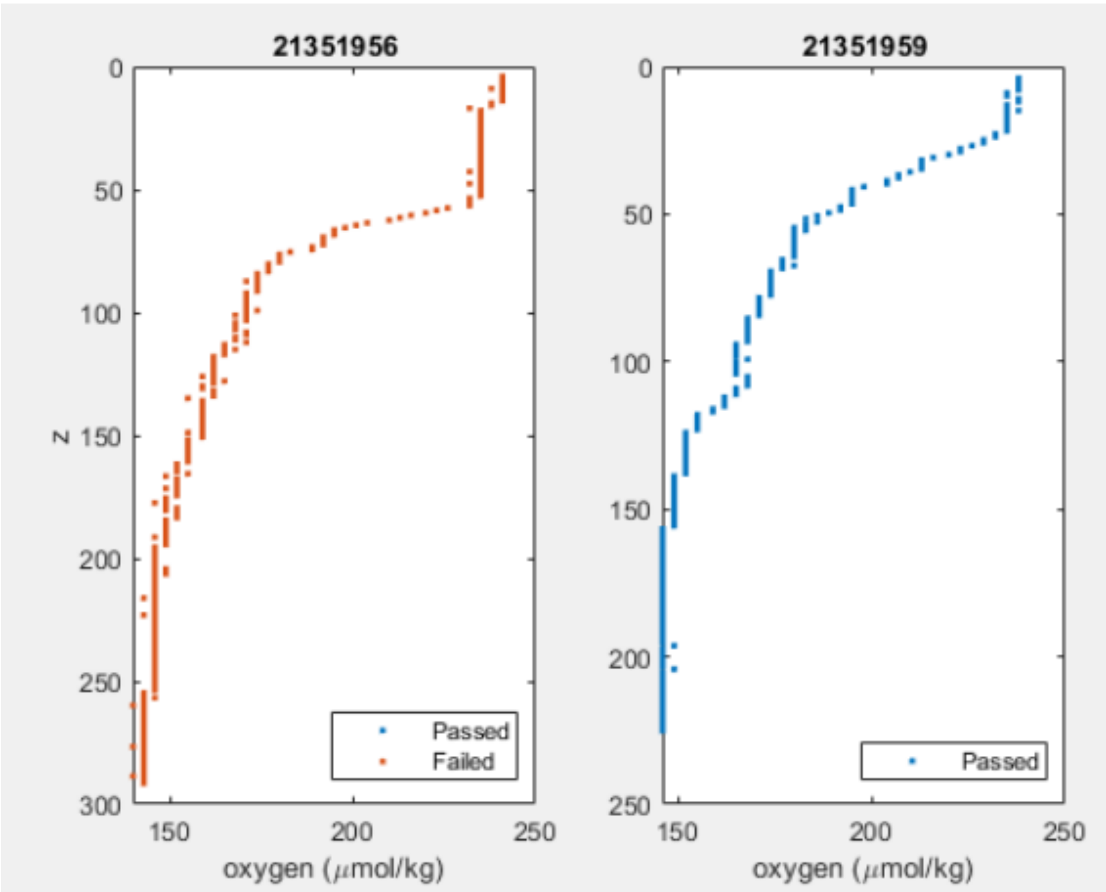


Figure 2 showing two profiles (titled with WOD ID) from the dataset included with the manuscript. Blue points passed QC checks while red points failed QC checks.

It was also unclear to me why profiles that are similar in quality and shape result in different QC outcomes after going through the authors' QC pipeline. Unclear why profile in left panel of Figure 2 would fail the authors' QC tests and the profile on the right would pass. After looking at the individual QC tests for these profiles, it appears that the profile in the left panel of Figure 2 failed QC test 5. According to the mat file metadata QC test 5 is spike value while according to the manuscript QC test 5 is the stuck value test.

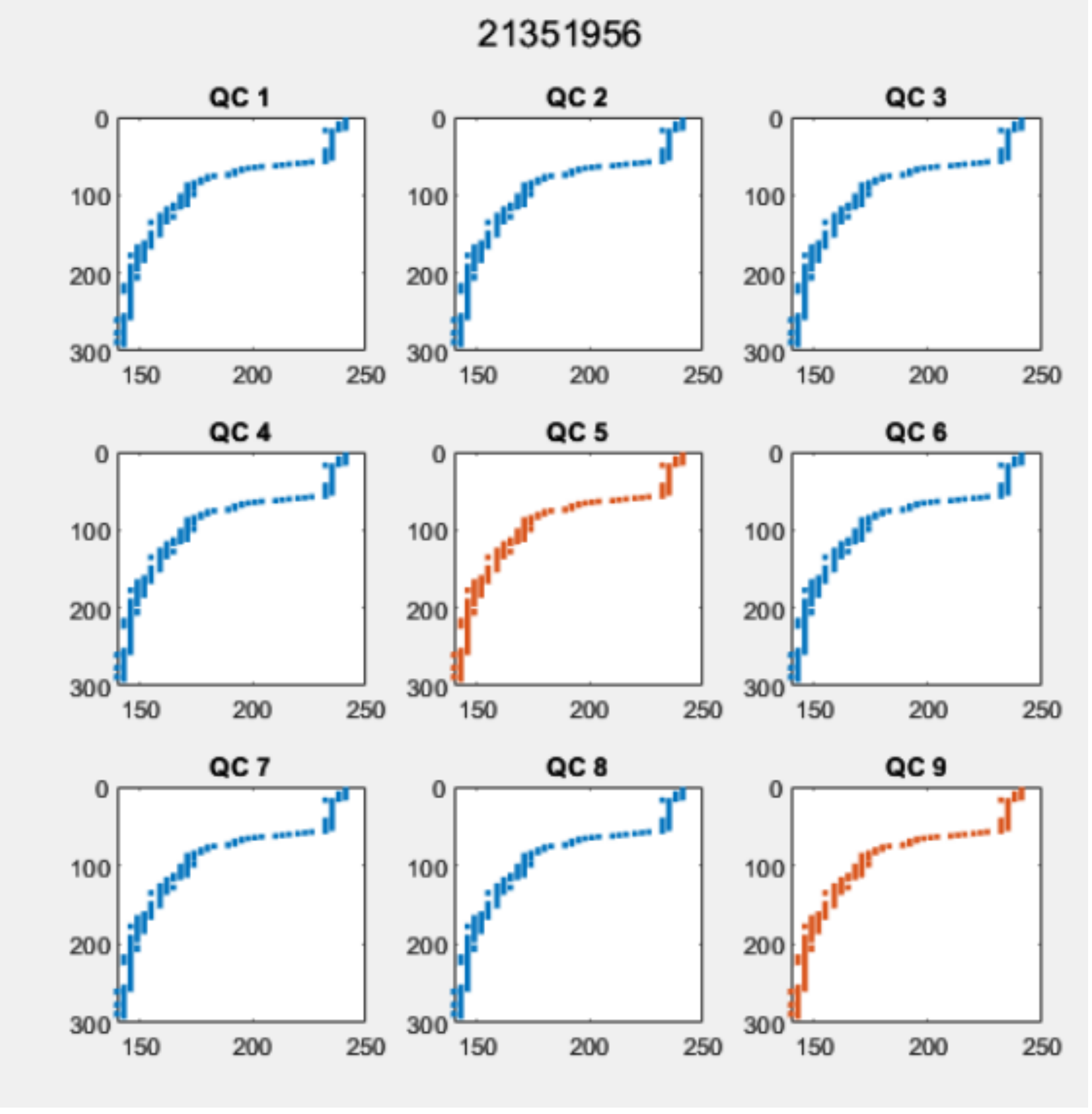


Figure 3 showing the results of the different QC tests for profile in left panel of Figure 2. Blue = passed QC test and red = failed QC test. Profile WOD ID 21351956 failed QC test 5

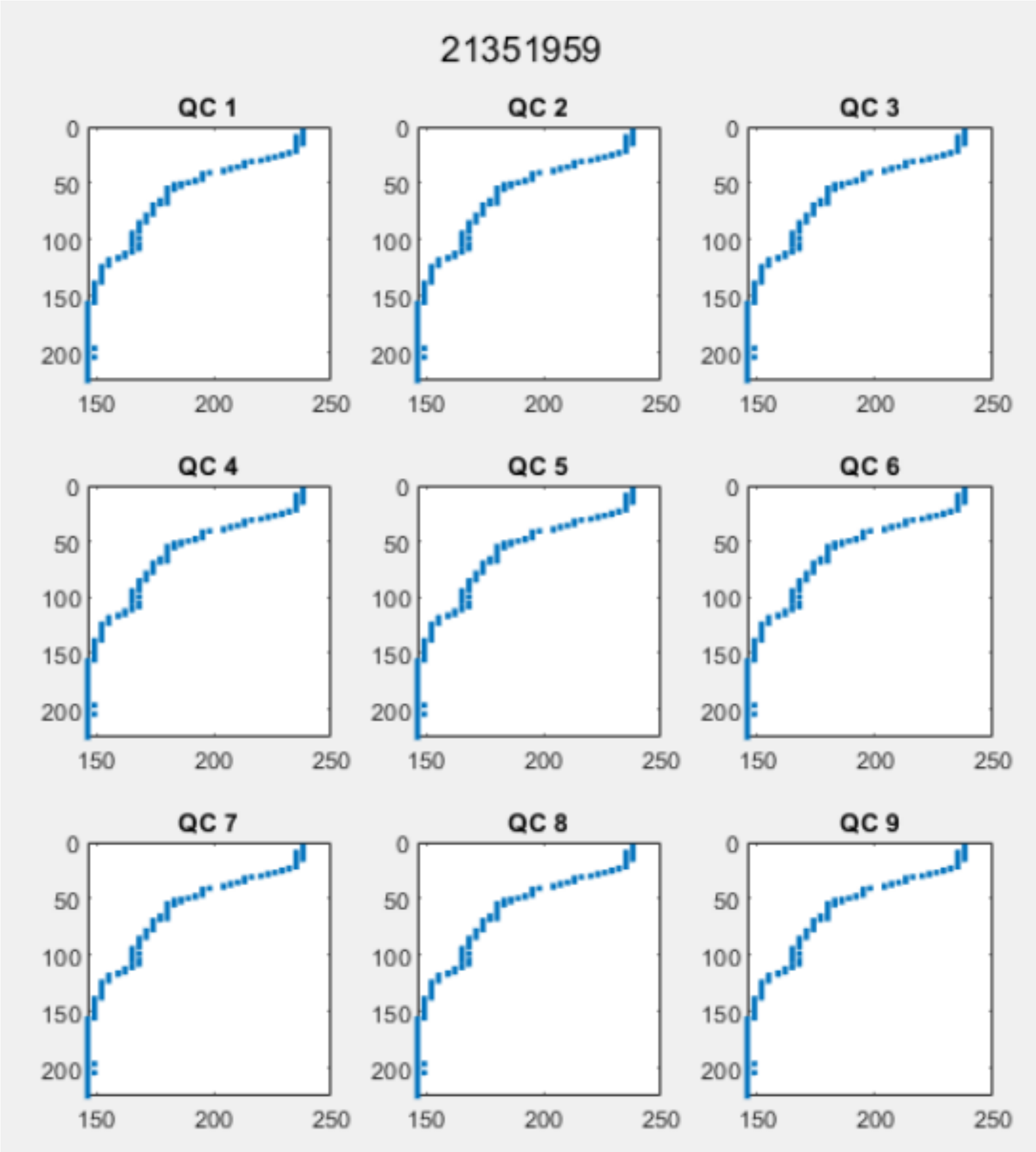


Figure 4 showing the results of the different QC tests for profile in right panel of Figure 2. Blue = passed QC test and red = failed QC test. Profile WOD ID 21351959 passed QC test 5 while Profile WOD ID 21351959 (Figure 3) did not.

Re: Thank you again for your careful test. Again, because of the bug in the data formatting code, the QC records are not correct in the previous dataset. Now, all corrected. We double-checked and reformatted the data files to ensure that all QC records were correct across all data files. Additionally, according to the latest metadata description, the name of the QC test 5 has been changed from 'spike check' to 'stuck value check' (see earlier reply).
Indeed, Profile WOD ID 21351956 and Profile WOD ID 21351959 pass all QC tests in our QC program (see Figs. A3-A5).

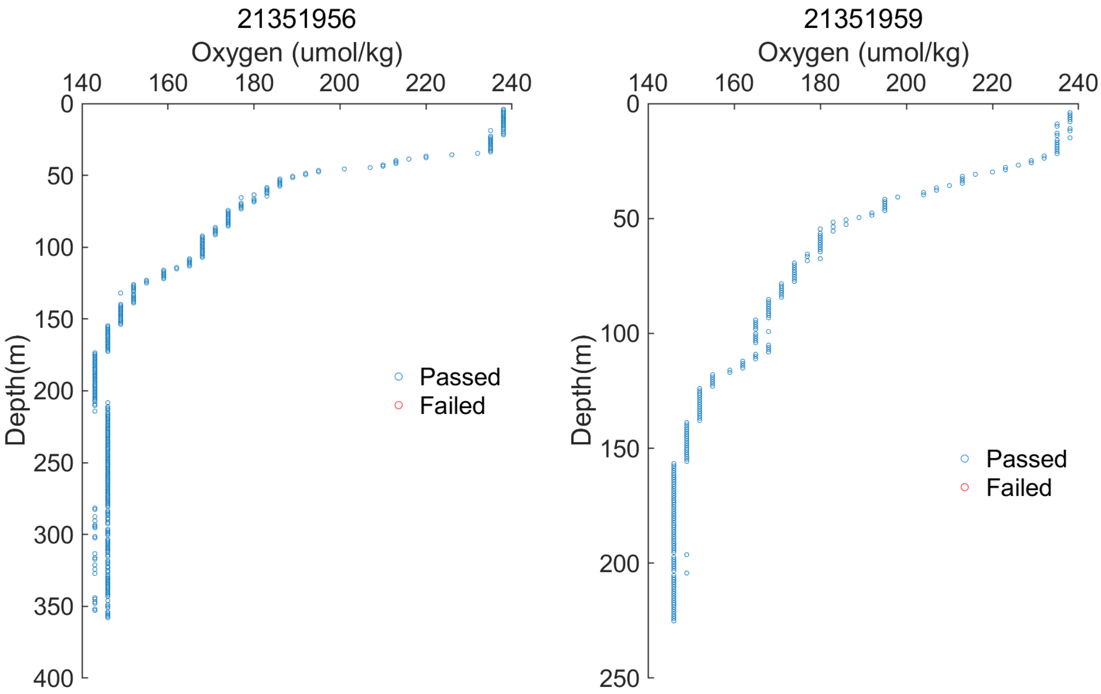


Fig. A3. The same as Fig. 1 (WOD_id=21351956 and WOD_id=21351959), but is for the revised data version. Here, blue points passed QC checks while red points failed QC checks.

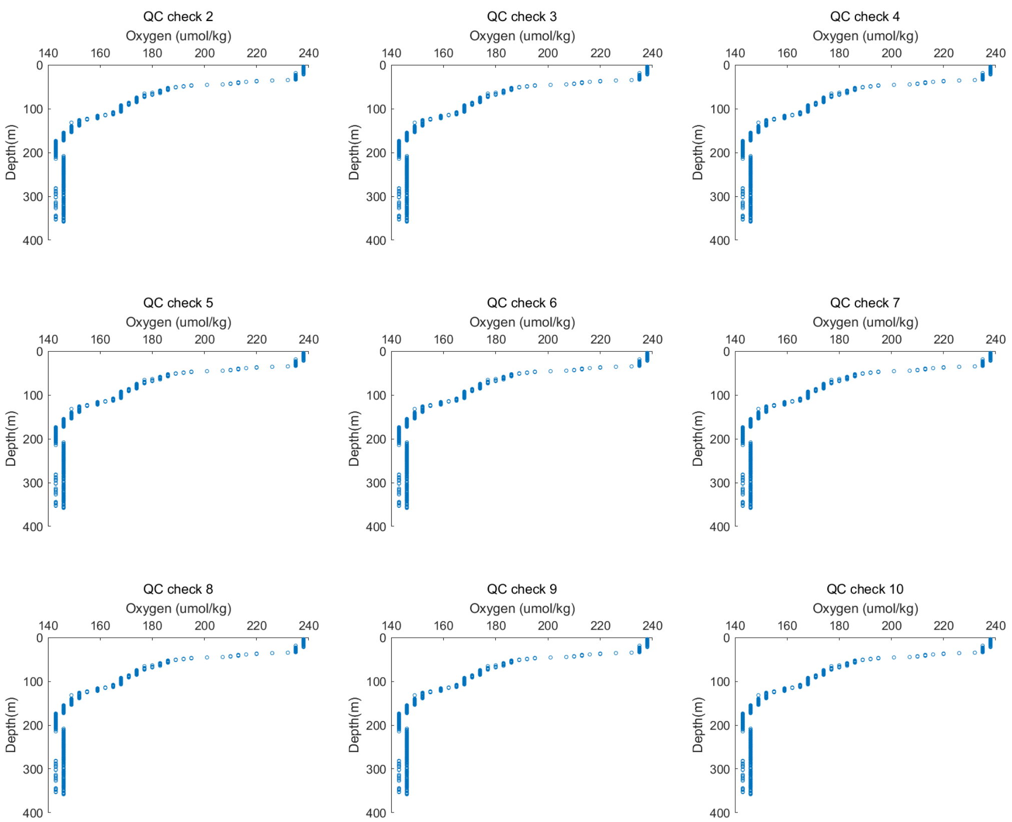


Fig. A4. The results of the different QC tests for Profile WOD ID 21351956. Blue = passed QC test, and red = failed QC test.

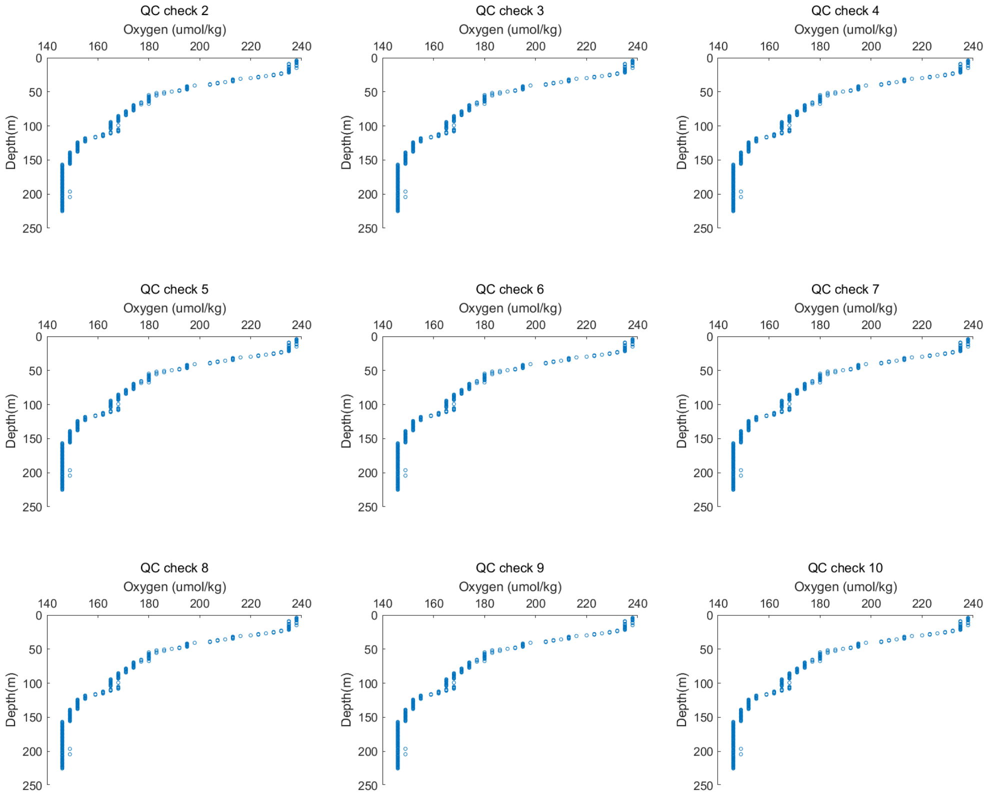


Fig. A5. The results of the different QC tests for Profile WOD ID 21351959. Blue = passed QC test, and red = failed QC test.