General Comments:

I am re-reveiwing this manuscript in which the authors developed QC procedures to assess quality of oxygen data from ship-based CTD and BGC Argo float profiles. Then compared QC-ed data to reference samples (Winklers) to look at biases for CTD versus Argo profiles.

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.

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.

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.

Data availability section is numbered incorrectly. Data availability should contain links for the original CTD-DO, Winkler and ARGO datasets accessed. Code should also be made available.

Line Comments:

Line 51: AQC_FINAL_CTD.f ?

Line 52: sensor drift is due to fouling and electrolyte consumption

Line 82: is not isd

Line 92: BGC not BGH

Line 100-101: World Ocean Database (WOD)

Line 101: oxygen not Oxygen

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

Line 102: depository not depositary

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

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

Data Product Online

Line 192: should this be 2000 - 2014?

Line 243: What is PFL?

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

Line 262: What is PFL?

Line 378: AQC?

Line 430: remove KIO3 and just say standard reference

Line 437: program not programme

Line 538: which not what

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.

Figures:

Figure 7: What is PFL?

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

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

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

Figure 20: Panel J is missing

Figure 22: Panel D oxygen units are wrong

Supplemental Figures: y-label is offset in all e panels of supplemental 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.

I also found a number of issues with the data set provided with the manuscript.

- Read me file is empty.

- Incomplete list of netcdf files for OSD CTD data, ends in 2015.

- 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.

- 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.

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.

	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	2021	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
15												
16												

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

Additionally, the data associated with WOD_unique_IDs in the data set provided with the manuscript does not match the data associated with the WOD_unique_IDs when downloaded directly from WOD.
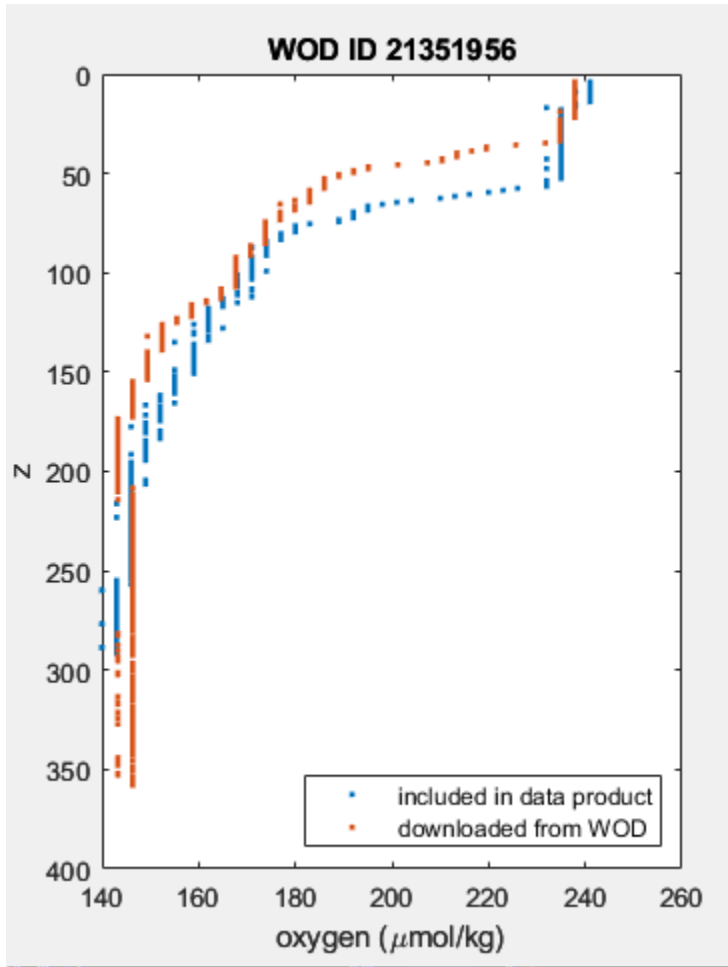

Figure 1 showing data labeled WOD ID 21351956 from the data set included in this manuscript (mat file) versus the data for WOD ID 21351956 downloaded directly as a nedcdf from WOD. Also note oxygen values can only be assigned to values every 3 umol (this issue seems to originate in the WOD data, not authors' data product).
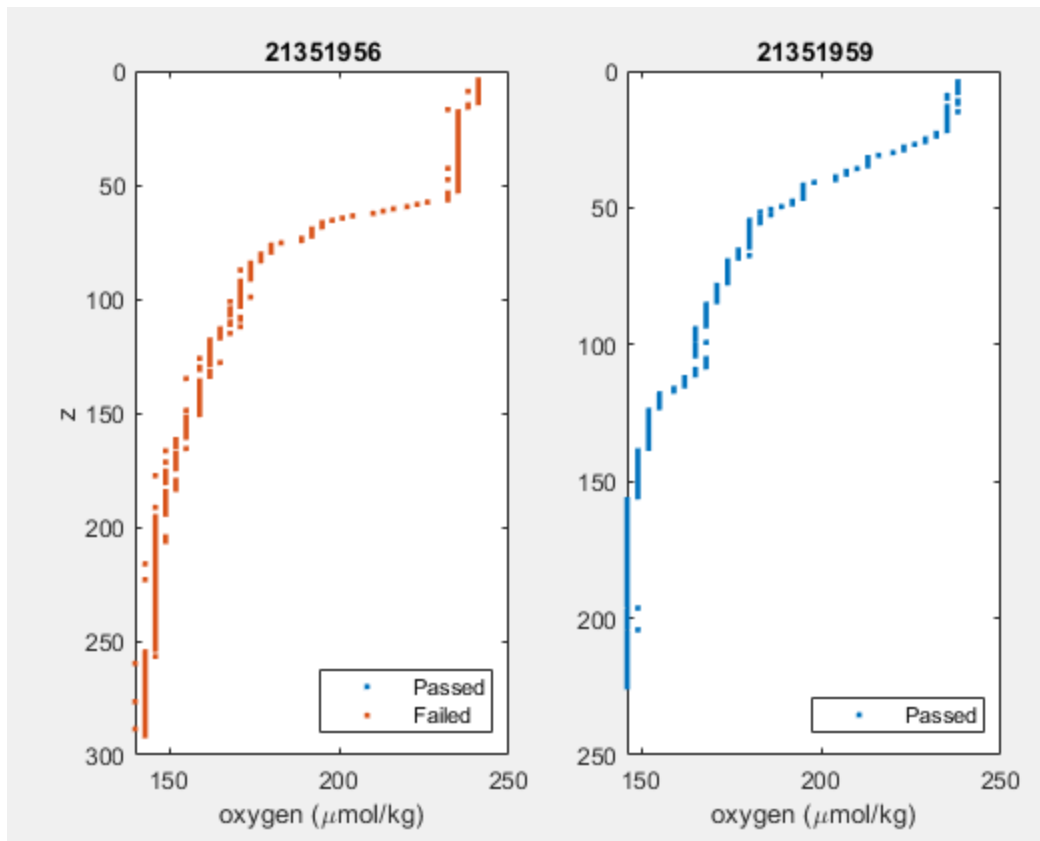
Figure 2 showing two profiles (titled with WOD ID) from the dataset included with the manuscript. Blue points passed QC checks while red points failed QC checks.

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.
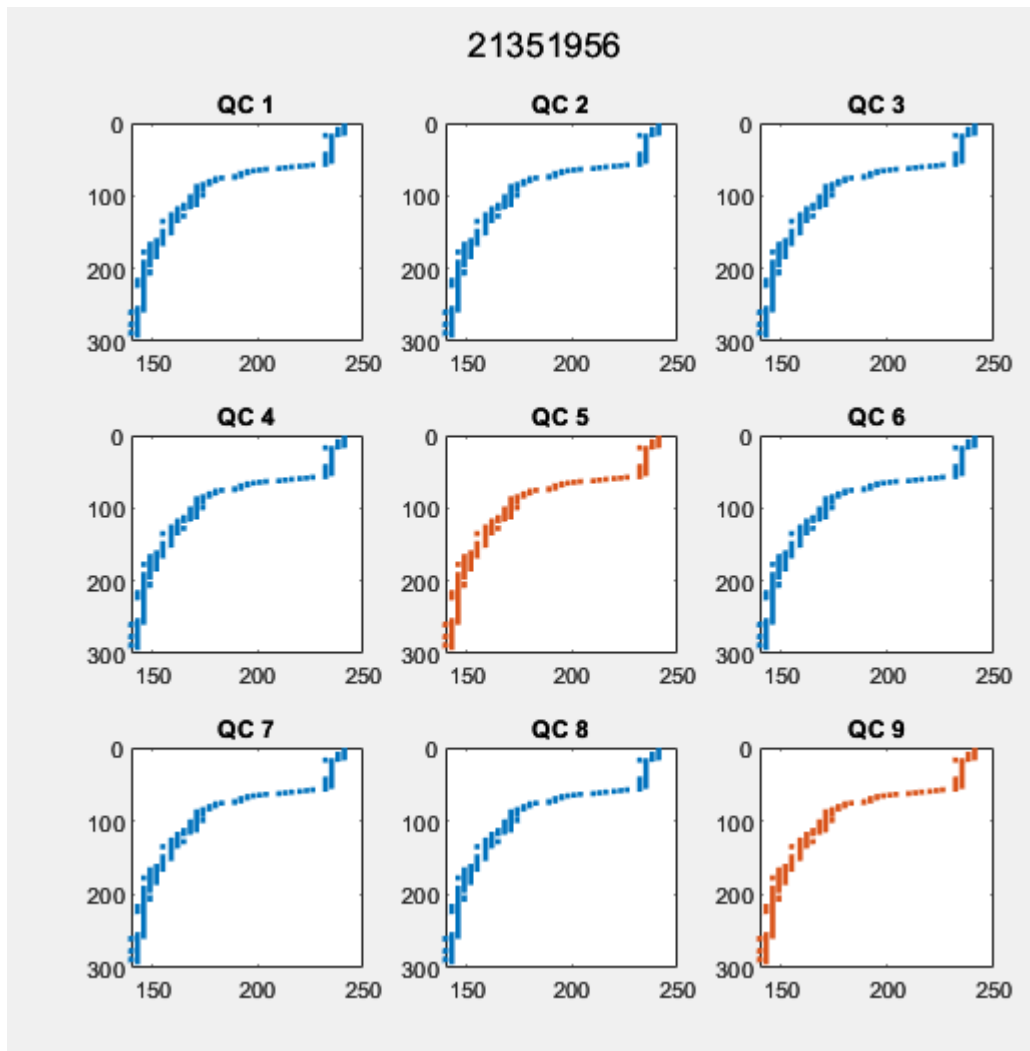
Figure 3 showing the results of the different QC tests for profile in left panel of Figure 2. Blue = passed QC test and red = failed QC test. Profile WOD ID 21351956 failed QC test 5.
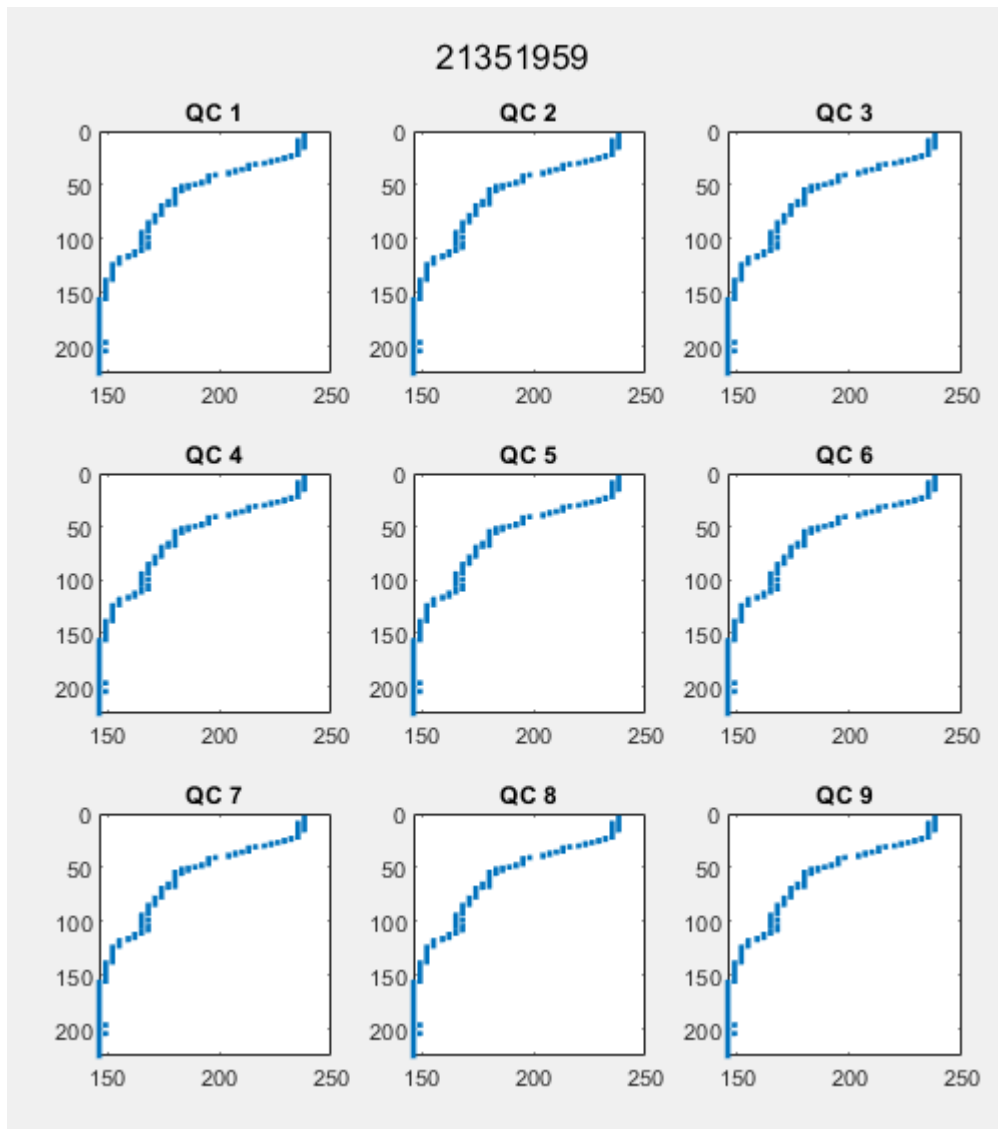
Figure 4 showing the results of the different QC tests for profile in right panel of Figure 2. Blue = passed QC test and red = failed QC test. Profile WOD ID 21351959 passed QC test 5 while Profile WOD ID 21351959 (Figure 3) did not.

Because of these issues, I do not have high confidence in the quality of this data product as is and recommend that this submission be rejected. I am willing to review the dataset again.