Response to comments from RC1

On behalf of all authors, we would like to thank the reviewer Dr. Toste Tanhua for their thorough reading of the manuscript and their constructive remarks and suggestions. Your comments provided valuable insights to refine and clarify the manuscript. We have taken into consideration all suggestions.

In the following, we try to address some issues raised, all comments and suggestions were implemented in the new version as best as possible. For a full overview , please refer to the undated manuscript version.

Below, the reviewers' comments are given in *italic blue* and responses in normal black font.

The manuscript from Belgacem et al. is setting out to produce a data product focusing on consistent ocean oxygen values in the western Mediterranean Sea, with some important restrictions, such as only considering data collected on Italian vessels, and only considering the oxygen data from the CTD sonde. It is a well needed attempt to move toward a consistent set of ocean data, fit to determine variability and trends in oxygen in this particular area.

- We appreciate your recognition of the value of our effort to establish a consistent oxygen dataset for the WMED region. We have clarified in the manuscript the rationale behind restricting the dataset and focusing on CTD sensor data, emphasizing data quality and methodological consistency

The abstract promises "The quality assurance process involves calibration of CTD measurements against Winkler analyses and the comparison of deep observations with reference datasets, using the crossover analysis". However, I see only weak connections in the manuscript between Winkler and sonde (CTD) data. The supplementary information is very short, only 2 tables. These are of interest and I suggest including this information in the main manuscript rather than having a SI. Maybe the information from table S1 could be included in tables 1, for instance. There are a lot of details on how the analysis of Winkler and CTD data is done that needs to be discussed in the manuscript. Table S1 has no information on any post-calibration of the CTD sensor: Was a "post-calibration" needed? If so, how was this done, and how large was the post-calibration?There needs to be a discussion on the difference of this analysis, that uses CTD sensor data, vs. the Winkler data.

- We have added a dedicated section in the revised manuscript describing the post-calibration assessment. Moreover, we expanded the discussion on the differences between CTD sensor data and Winkler titrations, highlighting the strengths and limitations of each and their complementary roles in our quality control process.

It took me a while to understand that the manuscript is about the CTDO data, and not about the discrete bottle data, as is the case for CARIMED and GLODAP. Also, and importantly, if there was an adjustment applied to the CTDO data, was the same adjustment applied to the Winkler data for that cruise?

- We have clarified early in the manuscript that our focus is on CTD oxygen sensor (CTDO2) profiles rather than discrete bottle data. Because Winkler titrations were limited in number, we did not apply the same adjustment to Winkler data; instead, Winkler measurements served as independent calibration points for the CTD sensors. This distinction and rationale are now explicitly stated.

And where is this documented, and where are the bottle data for those cruises? Actually, where are the bottle data for all of these cruises? This connection is particularly important for any

future effort to make a data product of oxygen data in the Mediterranean Sea that is more universal, but based on bottle data.

- Metadata including information about bottle data for each cruise is now provided in Cruise report link the Supplementary Material, the data is not yet available. We acknowledge the importance of integrating bottle data for a more universal Mediterranean oxygen product and have noted that gathering accessible Winkler datasets is a key future goal.

The link to the temporary data set did not work, so I could not inspect the product. This in itself is a disqualifier for publication, that can only happen once the data is on a recognized repository.

- We apologize for the inconvenience. The link is now active and points to the updated dataset currently under review at PANGAEA: https://doi.pangaea.de/10.1594/PANGAEA.974725. The dataset will be publicly available soon.

Minor issues:

Abstract: Quality assurance is the process of checking the data while doing the measurements, so that you can react and correct any issues you notice. What you have done here is quality control, not quality assurance.

- Done: We revised the terminology to use “quality control” consistently

Line 38. Is that the end of the sentence? Add period.

Line 39: Oxygen is not leading to high productivity, but is the result of high (primary) productivity.

- Done: Corrected the sentence accordingly.

Line 41: OMZs are certainly important, but the OMZ in the Med is very weak, i.e. most areas of the (open) Med is well oxygenated. I suggest to note this in the text. OMZs cannot become “more frequent”, as they the term OMZ refer to an area, not an event. OMZ can become larger/smaller, more, or less, intense, but not more/less frequent. Low oxygen events can, however, become more frequent.

- Done: We clarified the limited extent and intensity of Mediterranean OMZs and rephrased to reflect that low oxygen events, rather than OMZ frequency, may increase.

Table 1: Maybe add a column on which oxygen sensor was used, and if there was a post-calibration needed after comparison with Winkler data.

- We included information about post-calibration.

Lines 132-135: As Far as I know, the CARIMED data product is not published, or final. How can you be sure that the oxygen has “reference cruise” quality by simply stating that they were part of a non-ready product? Also, did you make a check on consistency between Winkler data and CTD data for the reference cruises as well, noting that CARIMED is only bottle data`

- Done: We clarified the reference cruises were evaluated using Langdon et al. (2010) methodology with certified reference materials (OSIL KIO3) and note the CARIMED product is preliminary.

Would there be another, more objective, way of assigning reference cruises, such as using cruises that used CRMs or certified standards?

- To our knowledge he references cruises used the Langdon (2010) methodprecisely they used OSIL 0.01N KIO3 to standardize the thiosulfate; this is a reference material about CARIMED

Lines 181-219, roughly. Most paragraphs is this section is only one sentence long. It makes it awkward to read. Consider modifying. Much of this information is probably better in a table.

Line 227: Why is this an issue for recent cruises?

- Added explanation

Line 232: It seems very strange to average each cruise to one single profile. I cannot see how that can work, except for cruises with limited geographic extent. Why not make an average profile for each of your sub-regions instead, that might work out.

- Done: We now provide regional mean profiles in addition to cruise averages to better represent spatial variability.

Line 269: Section crossovers. I did not understand if you made a cross-over analysis of all cruises vs. each other, or only the cruises in the product vs. the reference cruises. The text is not clear on that.

- Done: The text now clearly states that crossover analysis was performed between all cruises and reference cruises, with a detailed explanation of the methodology.

Section 4: I would strongly suggest to use coherent definitions, and standardized language for the description of the adjustments to each cruise. I would suggest to make sure the ms is consistent in using "offset" – this is the difference between a cruise and reference cruises: "correction" – this is the reciprocal of the offset and indicate what would need to be done to get consistency based on the determined off-sets – "Adjustment"- this is what you did to the cruise based on different lines of evidence.b Reading the text on each cruise, it is not always clear what you did to the cruise due to fluffy language. I would recommend not to to fancy on language, but be consistent and clear.

- Done: Terminology was harmonized for clarity and consistency.

There are many many figures to the section describing the individual cruises. Maybe having one or two cruises with many panels would be much easier to read .

- Done: We reduced the number of panels per figure and focused on two representative cruises for detailed display.