We are very grateful for all the comments and suggestions from Referee #2. We appreciate he acknowledged the importance of our collected dataset and its rarity due to limited amount of long-term measurement campaigns in the Antarctic coastal region. We are grateful for overall positive feedback to our methodology and presentation of results, but even more so for all the questions and suggestions of improvements.

Our detailed responses to them are as follow:

Referee #2: Given that these datasets represent ocean profiles, it would have been nice to have a figure showing variability of properties with depth along profiles. Figure 5 does present some aspect of this variability, however.

In the initial stages of our data analysis we actually did look closely into all the profiles in full, but we later realized that in this scale we are missing the point of the biggest variability of all the measured properties that were noted in the surface layers. This is why we have developed Figure 5 to truly show this phenomenon, and we skipped a step of showing the reader information on profiles in their totality, to make our paper as economical as possible. However after reading your comment we have realized that maybe this information is still worth presenting, therefore we have developed another figure (Figure 1 in the attachment) that now we are planning to put it into our revised manuscript. In your opinion this is a good idea?

We will have a joined response to three following comments:

Referee #2: I have suggested the authors add additional detail regarding data processing (which they state follows YSI protocols with a link to the manual), as I am unfamiliar with the processing pipeline for YSI instrument and would therefore as a reader benefit from knowing more about it.

and

Referee #2: Inspection of the data (in Ocean Data View) was facilitated by the data being accessible in tabular format in Pangea. Plotting the physical data shows few (<<< 1%) unrealistic density values (a parameter not included in the dataset but derived using temperature and salinity) due to low salinity values. Similarly there is some scatter in
other biogeochemical parameters, which the authors discuss. Typically hydrographic data would first be processed using manufacturer software, and then potentially subject to further QA/QC via, for example, QARTOD (https://ioos.noaa.gov/project/qartod/). I am unfamiliar with the requirements of this publication as to data quality / status (e.g. raw v. L2). Adding additional details regarding processing as detailed above, here and in the Pangeo repository, would however allow for better usage of the data in the future (e.g. indicate to researchers the data should be passed through their preferred qa/qc pipeline).

Referee #2: Section 3.2: Additional details on data processing should be included, as the description of the data centers primarily on the collection and sensor calibration. For example: what software was used to download / record the data? Was it recorded in a YSI proprietary format, and later converted in some software? Did the profiles go through any QA/QC or interpolation / binning, as is common for seabird data processing?

Thank you very much for these questions, they certainly need some more explanation and will benefit our article. The measurement data was initially recorded to YSI proprietary format, handled by the manufacturer software embedded in all of the sensors. Through this software a real-time data filtering using basic rolling filter is performed that was done using default and recommended settings of the manufacturer. To be precise, let us quote on this YSI manual:

“As a sonde takes measurements, it compares new readings to those taken in the previous 2-30 seconds (depending on the selected option). If the new reading is not significantly different than past measurements, then it merely factors into the rolling average with older data points to create a smooth curve. If the new reading is significantly different than past measurements, then it restarts the rolling average of data points.”

The default mode provide optimum data filtering with up to 40 seconds of filtering on the sensors.

Additionally YSI sensors perform adaptive filtering and outlier rejection, again as per YSI manual:

“The drawback to a basic rolling filter is that response time to an impulse event is delayed, and the more entries in the average summation, the longer the delay for the result to converge on the true value. To correct this, the filter algorithm monitors the new data arriving and compares it to the current averaged result, looking for indication of an impulse event. When new data deviate from the average by more than a predetermined tolerance, the number of data entries within the rolling average is reduced to a minimum count and the remaining values are flushed with the new data. The result is a more accurate capture of the impulse event data, entirely eliminating the inherent delay caused by the rolling average.

Every time a newly acquired data value is added, the rolling average entries are scanned for outlier data. Although such data has already been determined to fall within the tolerances defined above, the remaining worst offenders are removed from the rolling average calculation. This outlier rejection allows for smoother continuous data results.”

Also automatically, through YSI software all the derived values were calculated from direct measurements (meaning salinity from conductivity and temperature, depth from pressure, pH from electric potential difference, quantities of turbidity, ODO, fDOM, chlorophyll A and phycoerythrin using linear regression from optical measurement results).

The gathered dataset was downloaded using YSI KorExo program from which it was
downloaded into csv format that was later analyzed using Matlab. Using Matlab data quality check was performed. We have analyzed all the distribution of all the property values and extracted questionable values based on one of the following reasons:

- Notes from the measurement crew indicated malfunctions or some difficulties during measurements
- In majority of measurements in sites with depth smaller than 100 m, sonde after reaching the bottom showed unrealistic values from all of the sensors which was caused by the contact with the seafloor. This was best observed through rapid spikes in turbidity values, so in all these profiles we have cut out all the measurements, from all the sensors from before these disturbances till the end of each of affected profiles.
- Extreme and outlier data was scrutinized individually:
  - Continuous abnormal values of a particular sensor during measurement day were extracted indicating sensor malfunction or decalibration
  - Incidental extreme values recorded within otherwise reasonable datasets were extracted indicating some momentary disturbance

Despite this procedure, our data did not go through any standard QA/QC procedure so we will make that clear in our revised manuscript.

To be sure the above information is given to our readers we will firstly add another column to Table 2, describing our sensors in which we will add information from what direct measurement given values was derived. More importantly, both in Pangea and our revised manuscript we will add above information about used software and data management procedure. In your opinion will this be a sufficient amount of information in this matter?

We will have a joined response to two of the following comments:

Referee #2: The authors discuss issues with a small subset of the data which they trace back to calibration issues, and I've suggested the authors include additional details if available with the idea that such details may aid other researchers studying similar harsh high-latitude systems.

and

Referee #2: L116: how often did negative values show up? Add a % of dataset for relevant parameters. I suspect it is small, which would further demonstrate the value / robustness of the rest of your dataset L153: You repeat the fact that there are negative values a few times (see L 116). You may consider consolidating that discussion when addressing the underlying reason for the negative number (say, in L 116, where you introduce it in the context of observations), and then simply note that negative values exist in the plots which is discussed earlier. Also, you mention that negative values are due to methodological and calibration issues. Do you have specific insight / recommendations into what would have corrected this issue? It may be worth including here, as it could be helpful for other scientists to know whether, for example, a seasonal calibration is needed, whether a combination of conditions (extreme cold and turbidity) reduces the accuracy of the instrument, etc. You cite the YSI manual in a number of sections, it may be good to detail some of the content here to give context to how calibration is done and what parts of this process may have been impacted in your case.

This is are some very good points and we appreciate them greatly, because we truly are almost certain of the main issue that caused the miscalibration of optical sensors of fDOM and Total Algae (measuring chlorophyll A and phycoerythrin content). YSI Exo manual outlies couple of procedures of calibration for the optical sensors, dependent on each sensor. These are either 1- 2- or 3-point calibrations, and obviously the more points of calibration the better the results of it. Unfortunately we have been able to calibrate fDOM
and Total Algae sensors only using 1-point procedure, using deionized water as our 0-fluorescence standard and this proved to be an insufficient method. Therefore, we will put this information into our revised manuscript and recommend future researchers to use a more robust method of calibration.

Unfortunately the % of negative values of the properties is not small, in fact it is 77.82% for chlorophyll A, 70.87% for phycoerythrin and 60.45% but looking at the histograms (Figure 2 in the attachment) of their distributions and vertical profiles in the above figure we are convinced that their relative distribution is significant. We believe that the highest pick in the histograms describe the actual state of 0 for each of the properties. The only question here is with the fDOM values distribution but after further analysis it is revealed that during our measurement we have found that in majority of the samplings the fDOM quantities were very low, suggesting lack of dissolved organic matter in the water, but when its values rose, they rose rapidly, showing sensors great sensitivity to its presence. This would explain the two spikes in fDOM values histogram with lower one showing instances of lack of dissolved organic matter in water and the second one, its variable presence.

We truly appreciate Referee #2 in depth questioning on this matter and we will put this information in our revised manuscript.

Also, as per the Referee #2 suggestion, we will consolidate the information about our negative results in the section describing the measurement and data handling procedure and we will make suggested changes to a further text.

Referee #2: L 17: unrealistic instead of impossible

We will change that in the revised version of the article.

Referee #2: L 20: I would shorten the discussion of GMW and make it a bit more clear. For example, it's the export of freshwater that changes the ocean's chemical composition, not GMW itself (which is the resulting water mass). I suggest: "When freshwater from glaciers is introduced to the marine environments, it mixes with ambient ocean water masses leading to the formation of new glacially modified water (GMW; Straneo 2012). Freshwater export has in this way been shown to influence properties of the coastal ocean, with impacts on the hydrodynamics and thermodynamics..."

Thank you for that comment and suggested solution, we will surely use it.

Referee #2: L30: I would change to: "While the majority of studies examining the influence of glacial meltwater on the marine ecosystem have been performed in the Northern Hemisphere, its importance for the functioning of coastal Antarctic waters has long been hypothesized (Dierssen et al. 2002)"

(https://www.pnas.org/doi/10.1073/pnas.032206999).

Again, thank you very much for the suggestion and reference, we will surely use it.

Referee #2: L41: West Antarctica

Of course, thank you for noticing, will be fixed.

Referee #2: L41 / L93: One thing that isn't obvious in your description is how challenging of an environment you sampled in. I suspect the presence of ice (sea ice, bergs and bergy bits), had an impact on operations. You mention it in L110, but it would be good highlight early in the description of the place that this is a remote and harsh environment.
Yes, it was a challenging environment to work in. We will add an appropriate paragraph describing it in this section. Thank you for that suggestion.

Referee #2: L175: a detail, but the link does not work as is in pdf (fine if I copy and paste into browser).

Sorry for that, of course we will fix it.

Referee #2: L179: I would remove the mention of negative values here, as you’ve discussed it several times prior in the text, and focus on the big picture value of your measurements. Instead, I would use sentence 1, skip 2, modify sentence 3 to highlight details of the scope of the measurements, and finish as you do. You could otherwise add a sentence as you do in the abstract, after you've summarized the strengths and value of your data, stating that while absolute values of parameters showed some issues due to calibration, the relative distribution and seasonality is still insightful, as it is one of the few existing, long term multi parameter time series in polar regions broadly.

Thank you for that suggested improvement, we will do so.

Referee #2: Figure 2: Excellent idea to have a visual of observation platform and sensors, as it is a unique environment to sample in

Thank you for that comment.

Referee #2: Table 1: I would add details in the caption to give context to the metadata, even if some of the details appear in the text. For example: depth was measured in this way, with depth >100 indicating that... While all stations are to some extent influenced by glacial input, distance from glacial front was measured only for those stations located within designated glacial coves...

We will add this information to Table 1 caption

Referee #2: L265: Snazelle should be cited as “Snazelle, T.T., 2015, Evaluation of Xylem EXO water-quality sondes and sensors: U.S. Geological Survey Open-File Report 2015-1063, 28 p., http://dx.doi.org/ofr20151063.” as per report (i.e. including U.S. Geological Survey Open-File Report 2015-1063).

Of course we will correct this in revised manuscript.

Please also note the supplement to this comment: https://essd.copernicus.org/preprints/essd-2022-320/essd-2022-320-AC2-supplement.pdf