

1 Response to Reviewer 3

The manuscript aims to improve the quality of the BGC-Argo dataset and deliver more derived parameters for users and help delayed mode operators and Argo data assembly centers to identify remaining quality control issues to address them.

Qualifying data is always a huge and a time-consuming task. Moreover, as recalled by the authors, Argo data are collected over years and floats may be corrected individually. Then, to improve the consistency of the global dataset, it is important to perform an analysis at a global and/or at a basin scale. So, leading such a task is always a benefit for the Argo user's community.

By delivering, the “filtered” data and providing additional derived parameters such as dissolved inorganic carbon, total alkalinity (TA), and pCO₂, in a commonly used Argo format, the authors will be of great help for the users.

Definitely, it is also useful to remove from the data the remaining bottom hooks and the in-air data wrongly assigned to the data profiles, but the manuscript leaves the reader a bit disappointed while flying over the main quality issues in the Argo chemical data, like the oxygen response time and the potential remaining bias in Oxygen Minimum Zones, for example. Some path forward and some perspectives to address those issues would be a plus for this manuscript. Moreover, it could be stated directly in the title that the only variables that would be covered are the “Chemical” ones, while the Bio optics variables are poorly and sometimes wrongly served.

Finally, the huge effort on the visual data screening deserves recognition and an efficient path for the screened data to the Argo data system should be found.

Thank you for your careful and thoughtful comments. We have done our best to address all of your comments below. Regarding the oxygen response time and other biases, we now address this specifically in the discussion (lines 284-287):

"This dataset does not fix all known issues with oxygen sensor performance. For instance, in addition to the previously mentioned bias in deep oxygen concentrations (Bushinsky et al. 2025), slow oxygen optode response has been shown to induce biases in regions of the water column with strong oxygen gradients (Bittig et al., 2014, 2018). It is our hope that our work can be combined with other efforts by the DACs and outside groups to continue to improve BGC Argo data in the future. "

While it is definitely important to deal with the slow response bias and we did consider trying to include that correction here, it is simply too much to include in a single paper.

Specific comments

- Line 22: This sentence may be a bit reductive: “Optical measurements that provide information on particle density and fluorescence, light sensors, and many other types of sensors have been deployed over time as well”

Changed to: "Optical sensors that provide information on particle density, fluorescence, and light are also being deployed in increasing numbers."

The author would be well advised to specify in the title of this manuscript that this dataset will focus solely on the chemical parameters of Argo, rather than “treating” poorly bio-optics.

New title: BGC-Argo+ v1.0: Global biogeochemical Argo float dataset with secondary quality control of chemical measurements.

Original: BGC-Argo+: Global biogeochemical Argo float data with secondary quality control and derived parameters.

-Line 72 => For data reproducibility, considering that the Argo data set is constantly evolving the correct way to proceed is to get an archive corresponding to a certain snapshot of the dataset

For January 2025 Sprof snapshot, this would have been <https://doi.org/10.17882/42182#116312>

Instead of mentioning accessing the dataset on January 24, 2025.

[Added the correct citation.](#)

-In paragraph 2 related to data : it would have been useful to cite some references for the documentation to present the real time quality tests and delayed mode procedures that have been applied by the Argo Data team and lays the foundation of this manuscript.

For dissolved oxygen <https://doi.org/10.13155/46542>

For Nitrate <https://doi.org/10.13155/84370>

For pH <https://doi.org/10.13155/97828>

as well as, <https://doi.org/10.3389/fmars.2021.683207> (Maurer et al., 2021) describing the delayed mode procedures for all the chemical parameters.

As well as the Argo user's manual <https://doi.org/10.13155/29825> , the manual for the Sprof generation <http://dx.doi.org/10.13155/55637>

That will definitely be useful to understand the “jargon”

[References added to the methods where they were missing.](#)

-Line 92, Line 111 :

“filter out non-delayed mode data for most parameters”

It is not clear what this means, this needs to be precised

[Bullet changed to: "Download Sprof and Meta files, apply flags, remove non-delayed mode data for oxygen, nitrate, and pH, save to new "filtered" files."](#)

“We did not perform the same steps for pressure, temperature, or salinity, as our experience was that bad data in the "ADJUSTED" data fields for these physical parameters were already removed by the time any BGC data was put into Delayed Mode.”

Same as line 92, this needs to be precised, as it is admitted in the Argo community that BGC variables can be pushed in delayed mode before the CTD data.

[Clarified, see below.](#)

Do the authors mean that they only consider profiles with PRESSURE, TEMPERATURE, SALINITY , pH, Oxygen and NITRATE in Delayed Mode, if yes, they should say it like that, instead, this sentence is again confusing.

[Clarified, see below](#)

- Line 114, “as the vast majority of those did not appear to ever be set into Delayed Mode”, this sentence doesn't sound good. It looks like bio-optics variables are not within the scope of this manuscript, so why not saying it like that ? Are there any plans to account for bio-optics in BGC-Argo+ in the future and how ? Also reported in the table 1, it appears that 1/3 of radiometric data have been through delayed mode, so “the vast majority of those” expression may seem inappropriate.

[This paragraph now reads \(line 117+\):](#)

["For oxygen, nitrate, and pH data we only retained the Delayed Mode data, setting profiles in the other modes to NaN. Since one of the goals of this dataset is to make it easier for non-experts to work with this data, we decided to limit the dataset to data that had gone through the full DAC QC process. We did not, however, remove any pressure, temperature, or salinity data based on profile mode as there are sometimes BGC data that have been set to Delayed Mode before the physical measurements and we did not want to miss processing any Delayed Mode BGC data. Likewise we did not remove non-Delayed Mode data for other non-chemical parameters \(such as fluorescence, backscatter, or irradiance\) as they are not the focus of this manuscript and were not included in the outlier removal steps described below. "](#)

- Table 1 page 8

Regarding CHLA, these figures are wrong, presently in the Argo data set there are less than 200000 profiles for CHLA (June 2026). As CHLA is often associated to BBP, a more realistic number for CHLA profiles would be of 170000.

Thank you for catching that. We were over-counting profiles for BBP and CHLA. We had double-counted CHLA when both CHLA and CHLA_FLUORESCENCE were present in the files and had similar issues when multiple types of BBP were present. We now have 157,000 CHLA profiles, consistent with your numbers.

- Line 250, regarding the spikes, have the authors noticed some spikes simultaneously on pH and NITRATE ? Regarding the lags for both sensors, are they of the same nature ? while the Nitrate sensor could be mounted apart from the CTD, what can cause the lag for the pH ? Is there any time response related to the pH sensor ?

This is a good point. We have changed the sentence to (line 262):

"These seemed to coincide with a sharp transition in temperature or salinity, perhaps indicating a lag between the conditions experienced by the sensor and those used for processing the raw signal into nitrate concentrations or pH, though this is less likely for pH, which is typically mounted next to or in the same flow path as the CTD"

-Line 265 :

For the reader, for the users and for the Argo delayed mode operators, this analysis would have benefited from further examination, for example with illustrating how many bad surface data and bottom hook have been removed by the delayed mode process and how many of these are still in the data system and have been removed by this study (while it is potentially feasible mainly applying the test on DOXY and trying to catch if the DOXY_ADJUSTED_QC has been set to 4 or not) and the visual data screening performed by the authors is in line with the delayed mode operator data screening if any ?

This is a good suggestion, but we did not find it very easy to isolate only the oxygen points flagged as bad that correspond to bottom hooks compared to those flagged bad for other reasons. We will try to do this for a future version of the dataset and also will see if we can get internal diagnostics from one of the data centers to do this comparison.

This would have been also useful to highlight that the bad surface data are mainly related to old CTS4 PROVOR floats, when they were upgraded to take some in-air measurements.

We note this in line 142:

"This problem has been noted in the Argo O₂ quality control manual for PROVOR/ARVOR floats (one of the three main types of Argo floats) due to the optode being suspected of measuring a mix of air/water (Thierry et al., 2025b)"

Bottom oxygen hooks are also more visible for floats with a higher sampling frequency.

This may be true and would make sense, but we did not observe that ourselves and are therefore not including this in the manuscript.

-To cite the Argo dataset

Argo (2026). Argo float data and metadata from Global Data Assembly Centre (Argo GDAC). SEANOE. <https://doi.org/10.17882/42182>

Fixed in methods.

Technical corrections

About the data

File downloaded here

https://ftp.soest.hawaii.edu/bgc_argo_plus/outliers_removed/v0.1_2026_04/Individual_Floats/5901744_Sprof_BGCArgoPlus.nc

Some long names could be added to the new derived parameters to help when navigating through the files.

some are missing and some are certainly wrong for example, DOXY_SAT is certainly the oxygen saturation and not the practical salinity, so the long_name, standard_name, units, valid_min, valid_max are wrong.

```
double DOXY_SAT(N_PROF, N_LEVELS) ;
```

```
DOXY_SAT:_FillValue = NaN ;
```

```
DOXY_SAT:long_name = "Practical salinity" ;
```

```
DOXY_SAT:standard_name = "sea_water_salinity" ;
```

```
DOXY_SAT:units = "psu" ;
```

```
DOXY_SAT:valid_min = 2.f ;
```

```
DOXY_SAT:valid_max = 41.f ;
```

```
DOXY_SAT:C_format = "
```

```
DOXY_SAT:FORTTRAN_format = "F9.3" ;
```

```
DOXY_SAT:resolution = 0.001f ;
```

```
DOXY_SAT:coordinates = "JULD LATITUDE LONGITUDE PRES_ADJUSTED_BGCArgoPlus" ;
```

[Thanks for pointing this out. We have checked/changed attributes for all new variables.](#)

-Line 12 : Niskin bottles instead of “Nisken” bottles

[Corrected](#)