Participants: Henry Bittig, Thierry Carval, Hervé Claustre, Matt Donnelly, Tetsuichi Fujiki, Chris Gordon, Ken Johnson, Robert Keeley, Tanya Maurer, Yuka Okunaka, Josh Plant, Raphaëlle Sauzede, Christina Schallenberg, Catherine Schmechtig, Anh Tran, Kamila Walicka, Annie Wong.

Excused: Xiaogang Xing, Violetta Paba

The meeting agenda is:

- Standardization of Quality Flags
- Progress on parameter decision trees
- Formalizing a BGC reference database
- Propagating more adjusted data to the GDAC
- Additional ADMT topics: Optode time response, Combined trajectory files, Fine-tuning Sprof files, Meta data

1. Standardization of Quality Flags

The problem is that different DACS are interpreting and applying Argo Quality flags differently for BGC. The example of DOXY is given: At MBARI, the unadjusted data RT flags is set to 3 meaning “caution: do not use data without correction” while at Coriolis, it is set to 1 meaning “This data has passed all real-time tests”. The proposition from MBARI is to set QF=3 for RT unadjusted DOXY, to prevent users to misuse the data.

Anh reports that, now the bufr format (compare to the old TESAC format) allows to transmit the data as well as the QF, so there is not particular recommendations on quality flags settings regarding GTS transmission (mainly dedicated to meteorological uses: MET OFFICE).

If we flag the DOXY initially at 3, is it still needed to run the real time tests? (Anh’s question). There is a key difference between the QC=3 (probably bad) and the QC=4 (bad) reflecting out of range or outliers, so the real time tests are still needed.

Thierry recalls that QC=0 is for parameter that has never been through any quality control and has no specific entry for common tests like range, spikes ...

Catherine is in favor of setting up a Real Time QC test for all BGC sensors that are known to have a calibration issue (pH, NITRATE, DOXY). For DOXY, it could be based on a comparison with the WOA climatology (See Josh’s Audit). It is also mentioned that for BBP and radiometry, the sensors don’t show any obvious need for an adjustment (confirmed by Hervé), so maybe it is not worth to set their unadjusted raw QC to 3.

Tanya is afraid that implementation of a real time test might be slow. So, she is in favor of populated the adjusted fields (using a gain value in RT).

Annie thinks that it is a good idea to move on populating the ADJUSTED field and its associated QC field, but there is still the need to populate the raw flags. Even if the adjustment is done in RT or in
Near RT, there is still the need to go back and populate the raw QC with a 3, if you suspect the data to be out of calibration. There are two suggestions:

1. Put every doxy sensors into the greylist (QC=3) prior the first adjustment
2. Set up a Real Time quality tests, comparing to a gain and set a limit, if the limit is reached it changes the QF of unadjusted raw data to QC=3.

Henry doesn’t think that doing better with adjusting the data is a good reason to say that the raw data are bad. Instead of asking DACS to put a QC to 3, we should ask them to correct the data of the 6 percent (which is the first guess), that will remove biais and make the data consistent with the climatology. The problem is to adjust the data, not to flag them. Assigning an ADJUSTED_ERROR should solve the issue.

According to ken, the 6 percents can be up to 20 percents depending on the sensors (Rinko, Aanderaa, Seabird), the storage...

Tanya and Matt don’t agree with henry’s reasoning, as the raw data are still existing in the system and they should carry a flag that makes sense without considering that the adjusted fields are populated or not.

In DM operations for core data, Annie reports that the ADJUSTED_QC are reported to the raw QC, pointwise, mainly with visual control on spikes. Moreover, if the DM reports an adjustment that is over a threshold, raw qc should be reassigned to 3. This should be equivalent to setting the raw DOXY QC to 3 and it keeps the consistency between core argo and BGC argo. Annie reports that regarding the core data, the quality flags bring also some idea of the data accuracy and not only the quality.

Matt reports that we can not presume what will be the user behaviour so we have to minimize the risk of misuse by being clear about the meaning of Argo quality control and consistent in how we use this meaning throughout Argo. He also explains that we should clarify what is our definition of “Science-Ready” in Dmode. He suggests to make a video (5 minutes) on quality control - R/D files, R/A/D mode, and QC flags to help users. This final idea is well received.

A consensus is reached on a DOXY_QC =3

The decision is still pending for radiometry and BBP while these are really different sensors (Hervé and Catherine).

Actions:
In DOXY documentation:
- Set up a real time QC test to put DOXY_QC=3

All DACs:
- Set up an unadjusted raw DOXY_QC=3 (pH, NITRATE)
- Move on populating ADJUSTED fields (see Josh’s audit)

All:
- Investigate the video on quality control (where will we put it ?)
2. Progress on parameter decision trees

**CHLA**

Christina presents the flow chart that has been discussed recently for CHLA.

- Global range should be done early to exclude negative values with new limits to reflect the Roessler factor
- Dark estimation should be done on 5/10 profiles?
- Spikes (negative or positive) are no longer marked as bad (QC=4), the spike test is used for the quenching area and to prevent the use of negative spikes in the dark estimation
- CHLA will mostly be QC=3 to reflect the factory dark is different from the dark in-situ
- The decision tree should also address issues with PSAL and TEMP (QC=4) impact on CHLA (for example with the MLD estimation)

The question of whether 5 or 10 profiles should be used to estimate the DARK value is raised. Moving to 10 profiles instead of 5 would be an issue for MBARI floats since they cycle every 10 days. This would mean that no stable dark estimation would occur before 100 days (so maybe a solution would be to refer to time instead of cycle). Several examples of MCOMS sensors (aoml, csiro) illustrate that an exponential decrease of the DARK value is observed just after the deployment of the float. Hervé reports that it should be an issue relative to the MCOMS sensor as such an exponential decrease has never been observed on ECO sensors (and this decrease doesn’t appear on an illustration with aoml FLBB sensors).

Regarding the filtering of the spikes for the CHLA_ADJUSTED, Catherine says that maybe this decision should be addressed to AST with illustrating examples as it is directly related to science (phytoplankton aggregates, quenching on spikes ...).

**Actions:**

Christina:

- Prepare some tests and some metrics to circulate through the different DACS on the new CHLA_RTQC and ADJUSTMENT method

All DACS:

- For CHLA, run the tests and calculate the metrics to present at the ADMT so that we can come to an agreement at the ADMT.

**NITRATE**

Catherine reports about the decision tree for NITRATE RTQC and ADJUSTMENT. The decision tree was discussed between the LOV and MBARI, but the draft document is available to circulate among DACs that are interested in. The document is almost ready and is a high request from Sylvie Pouliquen as it is needed for other DACs (other than Aoml and Coriolis) with NITRATE sensors. Tanya mentions that there is still an issue on the assignment of the NITRATE_ADJUSTED_QC after the adjustment in Real Time (1 for aoml, 2 for Coriolis). A large part of the sensors at the Coriolis DAC are deployed in the Mediterranean Sea and are adjusted automatically on the WOA, while it is known not be well representative in this area. The proposition could be to assign the NITRATE_ADJUSTED_QC to 1 everywhere except in the Mediterranean Sea and Red Sea, if performed with a WOA (like “a regional adjustment”).
DOXY

Tanya presents some slides on decision tree for DOXY, that started on a side meeting at the ADMT, some more work that should be included is

- Include RT tests for O2 & flagging scheme
- Include RT adjustment gain estimation: 1) WOA, or 2) last DM gain assessment, ie SAGE-O2
- Organize into single diagram
- Decision tree for DM

Henry wonders if the presented decision trees are only a “visual representation” of what is already in the documentation.

Catherine and Tanya say that they should also work on how every parameter interact with other parameters (example PSAL on DOXY). This was also mentioned in Christina’s slides on how should we report PSAL and TEMP bad in the estimation of the MLD for the RTQC of the CHLA. Catherine reports that the first objective is to help DACS (for RT) or experts (for DM), to take decision.

Actions:

DOXY experts :
- Work on DOXY decision trees for RT and DM

Catherine :
- Work on decision trees for BBP and Radiometry and circulate.

3. Formalizing a BGC reference database

The question initially posed by Lynne Talley is “is there a part of the ADMT that deal with the shipboard data sets that can be used for calibration and/or validation?”. That is the case for Core Argo, but not yet the case for BGC. The requirements for such a database will be :

- A centralized group/place to receive data
- Data quality controlled in line with BGC-ARGO needs
- Formatted so as to be easily used

GLODAP should be a logical partnership :

- Easy to use products (and timely updates, release of merged datasets)
- Pathway from CCHDO and other centers
- No formal processing
- Take care of BGC + tracer parameters but still need support for HPLC and POC
- Big programs (WOCE, CLIVAR) are only a small percentage of the total cruises (datasets also come directly from PI)

Robert Key presents some insights for GLODAP :
Merging Bottle data and optical data was done only for a couple of cruises at the beginning of the SOCCOM project while at least when you include optical data, the size of the product doubles, for only a couple of points. All optical data are now going to NASA. It is not a big issue to use the GLODAP building tool to merge bottle data with optical data at the NASA but you need a person familiar with optical data first to identify the cruises and second to provide quality control on the data. (cf EOS, nico ??? ).

The decision that should be also taken is whether to use the GLODAP dataset or the WOA datasets which are not the same products at all. With WOA, BGC people would not be able to calibrate DOXY and NITRATE better than a couple of percents (units issues, looser QC) while the GLODAP dataset is smaller than the WOA dataset. So the question is, not to say this product is better than the other, but which one will fit our problem/issue/question.

There is also a distinction between WOA gridded data (used for example in Josh’s audit) and the WOA discrete dataset.

SEABASS is also an option for biooptics (https://seabass.gsfc.nasa.gov/wiki/System_Description, The NASA Ocean Biology Processing Group (OBPG) maintains a local repository of in situ oceanographic and atmospheric data to support their regular scientific analyses). Tanya mentions that it is not completely clear what data are going there. According to Josh, the data format is not always consistent (HPLC data) making the data less easy to extract/use.

Catherine presents what was done with deployment data at the LOV to illustrate that setting up a database would be very useful as it is a tremendous effort for every teams to keep these data available and usable (server maintenance, embargo gestion, access gestion, formatting). She also rapidly presents how they begin to use WMS (Web Map Service) to get direct access to the pigments database of the LOV (on going work).

Matt comments that from a DAC point of view there should be clarity about data that should be used, what should be developed, whether there is a clear pathway to access enough data. It is a tricky exercise to decide if it is what I need. Is it a decision that should come from Argo to decide what the best practices for the ship based observations are? Ie GOSHIP already has standards. Matt is afraid of a multiplication of the standards.

Tanya answers that it is not to explain for example to GOSHIP how to proceed with ship based observations but for others cruises also and to set up a list of information, metadata, QF... that are needed to use as reference database for BGC-Argo (not a new standard but the minimum information needed to get the data usable).

This may require some fundings, someone needs also to be identified. (Ken)

Anh reports that even for core Argo it is sometimes difficult to get reference data, so this task should be a common task and not from different groups.

Actions :
- Set up an agenda item at the next ADMT (finding key speakers)
- Specify the needs (formatting, quality, processing, metadata ) in a BP document

4. Propagating more adjusted data to the GDAC
Josh presents that a strong effort has been made in 2020 to push more adjusted data in the system with now more than 60% of the DOXY are ADJUSTED whether in DM or in RT. There is still some work to do. Josh presents the audit which aims to flag bad data. This audit represents a huge work and it needs the data manager to agree or disagree with the profiles assessment. Josh mentions that he only receives a few feedbacks from data managers and wonder if this audit is useful. Everybody agrees on the utility of Josh’s audit.

Ken also recalls that the agreement at the ADMT is to put on a grey list all the floats that have been in Josh’s audit for one year, with the risk to prevent good data to be in the system, so a feedback from DAC is requested.

Henry wonders how josh managed to reach the PI of the float. Josh sent the ftp link through the bio-argo list and expects the DAC to reach the PI (ftp://ftp.mbari.org/pub/BGC_argo_audits/DOXY/).

Matt explains that BODC really tends to go back to the PI expertise in the core context.

For meds, Anh has been using Josh’s coefficient. There is a new member in the team that will work on DM DOXY QC.

Catherine reports that for Coriolis, she tried to get in touch with several PI with no success and then apply the correction.

There is a clear need for the system is to get an official contact person to take decision for BGC parameters. Annie reports that this issue has been tackled at the last ADMT and Mathieu Belbeoch is setting up a list for BGC DM operators (in DM files, DM operator names should also be reported now).

Catherine reports that one request from Ken Johnson at the last ADMT, was to push quality controlled data in the system, especially those which have been scrutinized, unbiased and filtered. This was mainly for BBP and CHLA. In a side meeting, few data managers agreed on pushing published data into the system. Catherine presents the work made by Coriolis on floats dead prior to 1st January 2019. Because of an issue with Xing2018 method, CHLA data have not been DModed. All the routines to transform R files in D files are available on github (https://github.com/qjutard/). This work will be refined with Giorgio’s Dall’Olmo work on the uncertainty estimation work.

5. Additional ADMT topics:

Optode time response

Tanya presents some analysis on sensitivity of optode time response correction to sensor sampling interval, including some recent work done by Yui Takeshita. Some remaining questions are how error fields should be updated to include uncertainties in this correction (reflecting, the vertical resolution of the different platforms) and how to document that correction in the SCIENTIFIC_CALIB_xxx.

Combined trajectory files

The combined (Core and BGC) trajectory files was approved at the AST. In air measurements for the optode is / will be stored in this file as well as the timing information.

S-profiles

There are some on going discussion for the S-profiles:

- Create a S-profile, even with no good BGC-data to have a complete record of the float cycles
• Replace QC=8 (interpolated) with 4 or 3 if bounded by 4 or 3
• How to construct the pressure axis with CTD if no accompanying BGC data

Metadata

We should work on the severe discrepancies between metadata at the different DAC, particularly in the PREDEPLOYMENT_CALIB fields, So DACS are encouraged to ask questions or for help.

Matt will do a more complete review of the metadata at the next ADMT. Associated to that, Matt reports that the Argo manual reference tables are migrating to the NERC vocabulary server (work in progress).