Minutes of the 6th BGC-Argo meeting 27, 28 November 2017, Hamburg

Catherine Schmechtig

Hervé Claustre / Ken Johnson. Introduction/ general objectives of this meeting

This is an introduction to the 6th BGC Argo meeting : presenting the BGC array, the national BGC meeting, the number of peer-reviewed papers related to BGC-Argo as well as issues that have to be addressed during the meeting : the file format, the reprocessing of the data… The special issue on best practices in observing ocean of Frontiers in marine science is presented, Henry Bittig will lead the BGC Argo contribution.

Udaya Bhaskar. DAC: BGC Argo data management-India

From 2012, 51 BGC-floats were deployed by India and 40 are still active.

An Indian working group was formed to work on the data, suggest calibration protocols and QC protocols. This group wrote some recommendations on profiling time (to avoid grazing and Non photo chemical quenching) and on in situ observations that should be collected to calibrate BGC float data.

An Indian-Australian collaboration was set up.

Regarding DOXY QC, some issues occurred in OMZ area, with LOCODOX tools or MBARI S/W tool a neural network approach is used to check and correct the data.

BGC data are used in modelling studies, validation of remote sensing data, Tuna fish studies.

For the future, the objectives are to actively collaborate for float deployment , to get in situ data as well as to improve conjointly the quality control.

Matt Donnelly. DAC: BGC Argo data management-UK

The UK operates or has operated APEX, NAVIS and PROVOR floats with a range of sensor combinations totaling 38 deployed, and another 21 planned. These data are not currently being delivered in v3.1, with the exception of 13 PROVOR floats that Coriolis are processing on BODC's behalf. BODC has made progress towards delivering all data from BGC floats as a result of progress with the conversion to v3.1 and the implementation of the Coriolis processing stream at BODC to augment the BODC Argo System. An improved funding situation and expanded team should accelerate progress in the coming year. BODC is seeking to collaborate with other DACs on the development of a shared APF11 decoder.

Katsunari Sato. DAC: BGC Argo data management-Japan

Japan has deployed 79 Biogeochemical-Argo floats since 2005. About 62 floats of them are equipped with only oxygen sensor, 3 floats are equipped with Chl-a and TURBIDITY and 14 floats are equipped with oxygen sensor and Chl-a and/or BBP sensor. In addition to oxygen data, the submission to GDAC of Chl-a and BBP has begun but not put through RTQC. We will introduce RTQC according to BGC QC manual. We will also send oxygen data on the GTS in BUFR format.

Rebecca Cowley DAC: BGC Argo data management-Australia

Currently have 21 live BGC floats, mostly oxygen sensors.

• Many deployed as part of the Australia-India Joint Indian Ocean Bio-Argo Project.

• All Bio floats are processed in Real Time

• Bio-Data, including both raw and derived parameters, is delivered to GDACs in format version 3.1 BR files

Data QC:

• DOXY data and TEMP_DOXY data goes through RTQC procedures but is not calibrated in RT

• Other Bio data is not QCd in our routine RT processing, but final parameters are derived from the raw variables in RT. Web plots are available :http://wa-shiny.imos.csiro.au:3838/bioargo/

• DMQC for pH awaiting SOCCOM QC procedures.

• Nitrate QC is underway using Johnson [2016] protocols.

• Regionalized RT overview plots of all profiles of all float instruments for CSIRO BioArgo floats are available: https://research.csiro.au/iobioargo/

Annie Wong / Tanya Maurer. DAC: BGC Argo data management-US

Report from biogeochemical float data management in the US – Annie Wong. In the US, efforts to transfer biogeochemical float data to the Argo GDACs come from 3 groups: MBARI, UW, AOML.

MBARI

- Scientific processing of SOCCOM and pre-SOCCOM float data
- Real-time production of BR- files for SOCCOM floats
- 'A' and 'D' mode adjustments of bgc data
- Synthesis and analysis projects

UW

- Satellite data telemetry
- Real-time production of intermediate msg files for AOML
- Delayed-mode production of D- and BR- files
- UW historical oxygen floats

AOML

- Real-time production of tech, meta, core traj and core R- files
- Other historical oxygen floats (WHOI, remaining UW)

Accomplishments to date include BR-file production for bgc floats from the SOCCOM project and the pre-SOCCOM UW-MBARI equivalent project. SOCCOM data transfer to the Argo GDACs (via AOML) is now carried out in near-real-time by MBARI, at a frequency of once per day. (Since the time of the meeting, transfer frequency has increased to twice per day for delivery of new incoming cycles).

Tasks that still need to be done include the production of Btraj files to store in-air Optode data, and the conversion of the remaining legacy DOXY floats to V3.1.

Anh Tran / Catherine Schmechtig. DAC: BGC Argo data management-Canada

Canada deployed 5 new DOXY floats equipped with SBE63 sensors in 2017. We currently have 10 active floats out of 19 DOXY floats with SBE63 sensors ever deployed. Some DM QC was tried on DOXY obtained with an Aanderaa optode. Most of the netcdf files have been converted to format V3.1.

In the framework of the Greenedge experiment, 3 CTS5 were deployed in 2016 and 2 in 2017. Most of them were equipped with CHLA, BBP, CDOM, DOXY, NITRATE sensors and are able to wintering.

Thierry Carval . DAC: BGC Argo data management-France

The matlab processing chain for Coriolis floats is freely available on the web (http://doi.org/10.17882/45589). In 2017, data and metadata from these floats have been distributed on Argo GDAC. They feature version 3.1 core and bio profiles, core and bio trajectories, metadata and technical data. In November 2017, 40 781 BGC-Argo profile files from 315 floats were available on Coriolis DAC and 131 308 BGC Argo profiles from 863 floats were available on Argo GDAC. This is a strong increase compared to 2016 :+65% more profiles and +54% more floats.

Xiaogang Xing. DAC : BGC Argo data management -China

-CSIO data receiving and processing system begins to integrate the BGC data processing functions, and plan to finish all integration in the end of 2018.

-Six Argo floats (equipped with Optode) in the Southern Ocean were contributed to Argo project and global database.

Henry Bittig / Josh Plant. End-user utility of BGC-Argo files and how far does BGC-Argo go down the processing chain(L0->L1->L2->L3)

Biogeochemical-Argo files have the task to keep all relevant data from the floats that are needed for parameter calculations, corrections, and re-processing, as well as to provide calibrated observations in an accessible form. The v3.1 m-files contain raw and adjusted c- and b-parameters and are the highest level of processing in the current Argo data system. However, structure can be significantly different among different float types depending on the detail of data transmitted. For example, Apex/Navis floats transmit sequentially sampled BGC observations (~10 – 15 s delay) on one single pressure level, i.e., one single N_PROF, and Provor floats transmit pressure levels where the BGC sample was actually obtained, i.e., one N_PROF per BGC sensor. Additionally, the complexity of certain format structures can lead to parameter misalignment and proportionally large regions of wasted space (fill values). These issues hinder the use of data across different float types / DACs within the global network and need to be addressed. As a result of the discussions, Annie, Josh, Thierry (by e-mail) and Henry came up with an initial proposal to create simplified bgc profile files, starting from the GDAC m-files, that are coherent between different float types

/ acquisition strategies and allow simplified access to biogeochemical float observations. This proposal will be discussed on the argo-bio list and implemented thereafter.

Emmanuel Boss / Giorgio Dall Olmo / Antoine Poteau / Andrew Barnard. BBP Status

We discussed:

-The recent findings of Poteau et al., 2017, regarding the inconsistency of different sensors at depth. Andrew Barnard explained the bulk of the discrepancy as a due to a problem at WETLabs with assignment of appropriate calibration constants. He has compiled a list of corrected slope factors. It was agreed that the DACs will coordinate their application.

-We have agreed that from now on we will use the manufacturer dark values rather than the on-float dark values. We continue to recommend measurements of dark on float (and their inclusion in metadata) so we can learn more about potential bias between the two.

Xiaogang Xing / Emmanuel Boss / Josh Plant. Chla status + reprocessing discussion

We have agreed that,

1) Keep the factory-calibration-based measurement as raw data: CHL_FAC = SLOPE_FAC * (DC - DARK_FAC)

2) Compute the in situ OFFSET for each fluorometer: first compute the minimum value of CHL_FAC for the first five deep profiles (maximum depth > 900 dbar). Then, compute the OFFSET as the median of these five values.

3) Include Roesler_factor=2 into the DAC level processsing: CHL_ADJ = (CHL_FAC - OFFSET)/ROESLER_FACTOR

4) Using solar zenith angle to identify daytime and nighttime profiles: NIGHT profile: ZENITH >= 90; DAY profile: not(NIGHT)

5) No NPQ correction for NIGHT profiles, QC_FLAG = 2, and finish the processing.

6) For DAY profiles, CHL_ADJ_0 = CHL_ADJ, and then Using 3 or 5-point median filter (5-point median for NAVIS and PROVOR, 3-point for APEX) to smooth CHL_ADJ_0 profile and get baseline (CHL_ADJ_1) and SPIKE profile: SPIKE = CHL_ADJ_0 - CHL_ADJ_1

7) MLD criterion uses 0.03 kg m-3 as the difference of potential density from 10 m.

8) Xing12 is applied on CHL_ADJ_1 for NPQ correction, only for DAY profiles (using MLD obtained above), getting CHL_ADJ_2

9) adding back the spike: CHL_ADJ = CHL_ADJ_2 + SPIKE, and QC FLAG = 5, and finish the processing.

10) In region where NPQ is not applied, CHL_ADJ = CHL_ADJ_0, and QC FLAG =2, and finish the processing.

Some others tests will be done this year to test different quenching corrections.

Catherine Schmechtig and All. Discussion 1

-The PARAMETER_DATA_MODE is to be added to the merge profile index to improve accessibility to QCed float data:

Parameters from different sensors can receive DMQC at different times and by different people depending on their expertise. The BGC file name only records whether any single one of the parameters has received DMQC (BD... and MD...) or not (BR... and MR...). To add transparency and allow selection of profiles with a particular parameter in a particular parameter data mode by the user (e.g., all profiles with NITRATE in 'A' or 'D' mode), it was decided to add the parameter data mode after the parameter to the b- and m-file index. The order or parameter data mode follows the order of parameter, e.g., the merge file index will change from:

```
file,date,latitude,longitude,ocean,profiler_type,institution,parameters,date_update
coriolis/4901802/profiles/MR4901802_092.nc,20161015153058,68.980,-61.858,A,835,IF,PRES TEMP PSAL DOXY
NITRATE,20171009140541
coriolis/5902295/profiles/MD5902295_020.nc,20130121004100,-13.772,162.326,P,841,IF,PRES TEMP PSAL
DOXY,20170322113910
```

to:

```
file,date,latitude,longitude,ocean,profiler_type,institution,parameters,parameter_data_modes,date_update
coriolis/4901802/profiles/MR4901802_092.nc,20161015153058,68.980,-61.858,A,835,IF,PRES TEMP PSAL DOXY
NITRATE,RRRRR,20171009140541
coriolis/5902295/profiles/MD5902295_020.nc,20130121004100,-13.772,162.326,P,841,IF,PRES TEMP PSAL
DOXY,DDDR,20170322113910
```

-Deepest pressure test (Test 19) will only be applied to PRES and no longer to i-and b-parameters. If there were no other QC tests specified, test 19 is the only test on i-and b-parameters. The solution is to remove this test application to i-and b-parameters and keep their QC=0.

-For a parameter to pass to mode 'A' (i.e., adjusted in real-time), the calculation for the adjustment must involve the parameter itself (e.g., with an offset or slope). If a different parameter used for the calculations is in mode 'A' (e.g., PSAL_ADJUSTED), this does not transitions onto the parameter itself and does not put it into mode 'A'. The <PARAM> field is always calculated with other parameters in 'R' mode (e.g., PSAL). <PARAM>_ADJUSTED is only populated with a "real" parameter adjustment as defined above. A calculation without a "real" parameter adjustment but involving other adjusted parameters (e.g., PSAL_ADJUSTED) is not performed/not recorded in the BGC-Argofiles.

-Test 8: Pressure increasing test: High resolution, unbinned (BGC-) data can have a non-significant fraction of data with the same PRES or with pressure inversions. For optical and BGC data (that are outside the CTD's pumped path), any data are considered good data. If their PRES_QC is set to 4, they will be excluded from analysis. The following Proposition was not adopted: This test is not applied to N_PROFs that are accompagnied by BGC-Argo measurements (i.e., N_PROF>1 without T/S near surface sampling, N_PROF>2 with T/S nearsurface sampling).

Henry Bittig / Tanya Maurer. DOXY

The cookbook and QC manuals were updated to

- reflect the calculations described in the cookbook cases in the scientific description. (The cookbook case calculations did not change.)

- give more detailed recommendations on DOXY RT and DM adjustments as well as PARAMETER_ADJUSTED_ERROR estimates

- facilitate implementation of RT adjustments following DM

The following items were discussed and decided on during the meeting:

- Default RT adjustment: Oxygen observations are almost exclusively done using oxygen optodes. Their uncorrected data, however, is systematically biased low. To remove this systematic bias, it was decided to adjust their data by default in one of two ways:

1) If the oxygen optode is incapable of obtaining in-air measurements, the adjustment procedure will consist of using surface data from a climatology (e.g., WOA13). After the 5th profile, an average gain is calculated from the shallowest float O_2 data converted to PPOX_DOXY and the WOA13 seasonal climatological oxygen data converted to PPOX_DOXY and interpolated to the float locations. (*Discussion* : *MBARI uses %sat when performing comparisons to WOA, per Takeshita et al, 2013, while LOV wants to keep PPOX_DOXY because it's a well-defined quantity and unambiguous static O2 conversion to PPOX_DOXY can be used)*.

2) If the oxygen optode is capable of obtaining in-air measurements, the adjustment procedure will consist of comparing in-air optode measurements, as pO_2, to surface pO_2 values from NCEP reanalysis at each cycle. After the 5th profile, an average gain is then computed over all cycles, as described in Johnson et al. (2015) for RT processing. For DM and higher precision it is recommended to correct for the "carry-over" effect as described in Bittig and Körtzinger (2015).

The gain obtained from either procedure (1) or (2) is then applied for incoming profiles in RT. To indicate the automated and only first-order nature of this bias correction, the ADJUSTED_ERROR will be populated with a still large uncertainty.

- Specification of oxygen near-surface vertical sampling schemes

For CTD data, there can be two distinct vertical sampling schemes, the primary profile/vertical sampling scheme (N_PROF=1) with CTD data acquired with the CTD pump on and a near-surface sampling (N_PROF>1) for CTD data acquired near the surface where the CTD pump is off to avoid ingestion of surface films/bubbles. This distinction is important for PSAL, as pumping (or not) significantly affects PSAL data quality.

If oxygen data follow the same vertical sampling scheme(s) as CTD data, they are stored in the same N_PROF(s) as the TEMP and PSAL data.

If oxygen data follow an independent vertical sampling scheme, their data are not split into two, a profile and near-surface sampling, but put into one single vertical sampling scheme (N_PROF>1).

For oxygen data, the quality reduction in PSAL (used for calculations) when unpumped is not substantial enough to justify a separate treatment of oxygen data acquired while the CTD pump was on and of oxygen data acquired while the CTD pump was off.

-Conclusions about how to store surface oxygen data in the Btraj files

A decision was reached to abolish MC=1100, and move the assignment of surface data in the Btraj files back to the "true" MCs in the following manner:

X – 10 = in-water samples, part of end of profile, shallower than nominal 10 dbar

X + 10 = in-water samples, part of surface sequence (guidance in RT: before air-bladder inflation / before max. buoyancy)

X + 11 = in-air samples, part of surface sequence (guidance in RT: after air-bladder inflation / after max. buoyancy)

X – 1 = individual surface observations

Data to include should all be in PPOX_DOXY.

Users should be warned that the distinction between x-10, x+1-, x+11 is known definitively for some floats (e.g. some newer Apf9i APEX with Optode), but is only a best guess estimate for other floats (e.g. PROVORs).

The X + 10 / X + 11 codes apply only for X = 600 (AET), 700 (TST) and 800 (TET), i.e., when the float is at the surface.

Ken Johnson. NO3 (Sensor updates, LINR/Canyon, RT & DM QC principles)

Ken encouraged Argo scientists creating off-line, quality controlled data sets to also apply those corrections to Argo data sets in D mode to ensure that Argo data is of science quality

Carole Sakamoto's paper about pressure correction for the computation of nitrate concentrations has been published in 2017. The reference should be added in the draft documentation « processing nitrate at the DAC level » and this documentation should be published on the web site.

Using SAGE tool, presented by Tanya Maurer, it is possible to calculate offset, gain and drift coefficients to adjust Nitrate concentration. MBARI is able to fill nc files in near real time with good quality NO3 concentrations compared to bottle data acquired. As these correction coefficients are obtained with an operator intervention, there was a discussion on how the adjusted values should be considered. The consensus was obtained as follows:

-When the coefficients are applied to past profiles (backward direction), the NITRATE PARAMETER_DATA_MODE should be set to D and files should be « BD ».

-When the coefficients are applied as a first guess correction to future profiles (forward direction), the NITRATE PARAMETER_DATA_MODE should be set to A and files should be « BR ».

- The same guidelines would apply to other BGC parameters, eg pH, O2, CHL. The overarching data-mode identifier for the Bfile ('R' or 'D') will be set to the highest level <PARAMETER>_DATA_MODE described within that cycle's file.

-It is decided that BD files can be provided at the DAC even if core D files are not already there.

-A real time adjustment should be written in the documentation for DAC that cannot be as responsive as MBARI.

Tanya Maurer. SAGE

MBARI oversees the data processing, quality control, and overall management of all incoming BGC Argo float data from the SOCCOM array. With support from UW and AOML, we serve this science-quality BGC data (with adjusted-parameters filled) to the Argo GDACs in near-real time. To assist with real-time and delayed-mode data quality adjustments for oxygen, pH, and nitrate, MBARI has developed and implemented two MATLAB-based GUIs, SAGE (SOCCOM Assessment and Graphical Evaluation) and SAGE-O2. These software tools were designed to assist with visualizing float data against select reference datasets, including shipboard analyses and ocean climatologies, and to assess quality control adjustments within a quantitative framework. Gains, drifts and offsets are derived within the GUIs and are then propagated back into the processing stream and applied to incoming float data. The software has been used successfully in the post-deployment calibration and quality control of over 100 BGC floats and is currently being adapted for use by other Argo Data Assembly Centers. However, while these software tools can assist other DACs in the derivation of parameter-specific adjustments, the procedures for storing this information and propagating the adjustments back into the data stream/repositories will be DAC-specific and may require additional effort.

Poteau Antoine / Xiaogang Xing. Radiometry QC

<u>Real Time QC on radiometry.</u>

-A range test is proposed so far on radiometry data.

- The Minimum value is 0
- The Maximum value is calculated with a model of solar irradiance (Gregg and Carder) just below the surface and a function of light attenuation in the water.
- The Maximum range will depend on the position, time, depth and the wavelength of the measurement.

  -If Values < 0 => QC=3
  -For Depth in 0-10m, If values > Max_range => QC=2 (data could be affected by wave focusing)
  -For Depth > 10m, if Values > Max_range => QC=3

<u>Delayed Mode QC on radiometry,</u>

The temperature has an effect on the measurements, mainly on the estimation of the dark currents. To estimate this effect and be able to apply a correction, a night profile is still recommended.

Ken Johnson / Henry Bittig. pH presentation and discussion

pH measurements are calibrated in the laboratory with spectrophotometric pH measurements using purified dye. After floats are deployed, drift is assessed by comparing these calibrated measurements with predictions from the LIpHR system of regression equations or the CANYON Neural Network. These predictive systems were tuned to reproduce in situ pH observations computed from the GLODAPv2 data set. pH values in GLODAP data set may be measured spectrophotometrically with purified or unpurified dye, or computed using carbon system thermodynamics from dissolved inorganic carbon and total alkalinity measurements. These 3 methods do not yield equivalent pH values, as discussed by Carter et al. (Limnol. Oceanogr.: Methods, accepted). There was some discussion regarding which of the GLODAPv2 pH data sets should be used for adjustment of pH sensor data. The LIpHR system (Carter et al., accepted) provides estimates (as an option) that are consistent either with purified dye measurements or with DIC/AT derived values. The use of purified dye measurements as the target measurement would be directly consistent with the sensor calibration procedure. There are, however, more pH values in GLODAPv2 that are derived from DIC and AT values. These derived pH values are significantly less precise than direct measurements due to limitations of the DIC and AT values. Further, they are highly dependent on the carbonate system equilibrium constants used to compute pH, which remain in a state of flux. Accuracy of direct spectrophotometric pH relies on the appropriate characterization of the dye's pHs value. Further, they are dependent on the carbonate system equilibrium constants used to convert pH between laboratory and in-situ conditions, which remain in a state of flux. The discussion centered around whether spectrophotometric pH are not only precise but also the most accurate pH available, and ISFET pH observations should therefore be adjusted to the best other (direct) pH observations available, vs. pH data are very likely to be used in a carbonate system context, e.g., to calculate pCO_2, and therefore be adjusted to fit within carbonate system calculations without need for additional corrections (some users might be unaware of). The **interim** solution is to **use purified dye measurements as the target reference value** and to coordinate with other observations (e.g., GLODAP collection of ship-based data).

The opinion of Ken Johnson was that it would be extremely unfortunate if BGC-Argo were to decide not to use purified dye measurements as the target reference value. In essence BGC-Argo would be creating its own pH scale, inconsistent with the rest of the oceanographic community.

The opinion of Henry Bittig was that – given that one BGC-Argo science aim is to provide parameters and/or a description of the inorganic carbon system – it would be extremely unfortunate if BGC-Argo were to decide to use purified dye measurements as the target reference value. In essence, BGC-Argo would use a pH scale that is not compatible with measurements of other carbon system parameters by the rest of the oceanographic community.

Action item: Ken/Henry/Siv: Keep link with GLODAP community and report back if there are updates.

Catherine Schmechtig and ALL. Discussion 2, Open Questions

- Two documents for processing and QC ?

- Pros : Processing for DOXY is more than 200 pages, merging processing and QC will not be manageable.
- Cons : the QC is often really close to the processing. It is difficult to find the documentation on the ADMT web site.

Proposition : keep two documents and gather documentations links in a table on the ADMT website