BIO Argo Data Management Meeting, Liverpool 14 and 15 October 2013

Forewords: This document summarizes the discussions and subsequent decisions taken during the ADMT14 meeting. Nevertheless, after the meeting, discussions between meeting attendees and other data managers have concerned the overall organization of Bio-Argo files and their complementary or independence with respect to the Argo core files. Therefore certain decisions taken during the meeting were somewhat nuanced thereafter. The conclusion section of the present report briefly summarize the decisions related to this new file organization.

We present the summary of the two half days dedicated to Bio-Argo.

Bio Argo Variables Naming (Justin Buck)

Justin Buck presented the proposal for naming of the Bio-Argo parameters, which closed the action 2. Note that a “family parameter” suffix is added to easily identify the new parameter (cf. DOXY added to the oxygen raw parameters TEMP_DOXY, BPHASE_DOXY, C1PHASE_DOXY…).

Naming Of Bio-Argo Configuration and technical parameters (Jean-Philippe Rannou and Thierry Carval)

Thierry Carval presented several suggestions for the bio parameters:

-It is proposed to add N_VALUES in the NetCDF file as a new dimension to handle the multiple optical wavelengths for the nitrate, the particle backscattering and the radiometry processing.

-Argo parameters used to be coded as “float”, radiometry parameters should be coded as “double”.

-New configuration parameters / metadata / technical data have to be defined to account for specificities of Bio floats

Ex: SUNA With SCOOP, ISUS in pump stream, distance between the bio sensor and the CTD…

-the units for bio parameters have to be clearly defined.

-The definition of a cycle must be updated

Processing CHLA at the DAC level: Action3 (Catherine Schmechtig)

The calculation of the Chlorophyll-A concentration from the Fluorescence measurements is presented. To follow the Argo requirements (i.e. store every raw parameters as well as derived parameters), the family suffix CHLA is adopted. FLUORESCENCE_CHLA is thus directly linked to the calculation of CHLA.

To consider the evolution of the sensors measuring Chlorophyll-A, it is recommended to find a place to store the sensor wavelengths of excitation and emission.

It is recommended to store the last calibration to use in the "SCIENTIFIC_CALIB_EQUATION".

Chla real-time QC: Actions 4, 6, 7 ,8 (Catherine Schmechtig)

Several quality control tests are presented:

- Range test (-0.1-50 mg/m3)
- Adjustment at depth (with identification of strong mixing situations and identification of a sensor drift)
- Gradient and spikes identification: after discussion, gradient test is not relevant for CHLA, as well as identification and flagging of positives spikes.
- Non-Photochemical Quenching correction

Regarding the tests considered as relevant, some recommendations are adopted:

- When adjustment at depth is required, CHLA is flagged 3 and CHLA_ADJUSTED is flagged 1.
- When quenching correction is required, CHLA is flagged 3 and CHLA_ADJUSTED is flagged 1.
- When strong mixing situations are identified, CHLA_ADJUSTED is calculated with the adjustment at depth corresponding to last non-mixing conditions.

CHLA delayed-mode QC: preliminary E-aims deliverable (Hervé Claustre)

For CHLA, DMQC the following recommendations are adopted. The satellite data are only used to identify potential issues in their float counter part. For that purpose:

- Find metrics to identify (anomalous) drift/changes in the deep values (1)
- Find metrics to compare surface float data with remote sensing CHLA (2)
- When appropriate, compare (1) and (2) to identify a mixing problem from a sensor drift

With respect to a reference database, the use of several methods to define biogeochemical provinces has to be tested. Such data aggregation (e.g. at the scale of a basin) is presently required given the scarcity of data.

Processing backscattering at the DAC level: Action 9

The calculation of the concentration of the Particle Backscattering from the measurements of Beta is presented. The contribution of pure seawater is removed to calculate the particle Backscattering. This contribution depends on Temperature and Salinity (the indication for taking into account this dependence will be integrated in the next version of the document: Processing Backscattering at the DAC level).

Processing the backscattering raises the issue of the N_VALUES as a new dimension of the NetCDF file. For example: for some floats of an European project (Eaims), there are two wavelengths of measurements 532nm and 700nm. (There are some floats under construction that will measure three wavelengths). However following post-meeting discussion, N_values was finally not retained for backscattering (see conclusions related to post-meeting discussions).

Backscattering real-time QC: Action 10 (Hervé Claustre)

Contrarily to chlorophyll-A, it is not recommended to correct for deep values (~0) in real time for BBP. Indeed from the analysis of BBP deep values of floats deployed in various areas, it appears that the intra-basin variability is lower than the inter-basin variability, suggesting that factory calibration provide coherent data even under very low particle concentration. Therefore, for real-time, it is considered that calibration factors are correct. Additionally the following recommendations are adopted:

- Range test: 5.10^{-5} to $10^{-1}m^{-1}$
- The negative spikes test is adopted.

Backscattering delayed-mode QC (Hervé Claustre)

For BBP, DMQC the following recommendations are adopted. The satellite data are only used to identify potential issues in their float counter part. For that purpose:

- Find metrics to identify (anomalous) drift/changes in the deep values (1)
- Find metrics to compare surface float data with remote sensing bbp (2)
- When appropriate, compare (1) and (2) to identify a mixing problem from a sensor drift

With respect to a reference database, the use of several methods to define biogeochemical provinces has to be tested. Such data aggregation (e.g. at the scale of a basin) is presently required given the scarcity of data.

Processing NO3 at the DAC level: Action:11 (Ken Johnson)

The calculation of the Nitrate concentration is presented (and the family suffix NITRATE adopted)

Processing the Nitrate concentrations at the DAC level raises also the issue of the N_VALUES as a new dimension of the NetCDF file (as wavelengths for the Nitrate spectra are different from the backscattering wavelengths, then it is assumed that N_VALUES corresponds to the sum of the number of measurements at different wavelengths from various sensors.

The document “Processing the nitrate at the DAC Level” needs to be read and commented by “Argo experts”.

NO3 real-time QC: Action 12,13 (Ken Johnson)

Two tests are proposed, they need to be implemented:

- Range test: 0-46 ± 5 µM kg^{-3}. Practically the values between -5 and 51 µM kg^{-3} are accepted
- Gradient / spike test

Additionally a third new test is proposed regarding the spectral quality which relies onto two sub-tests (1) the value of the residue of the absorbance fit (threshold 0.004) or (2) the value of the Absorbance at 240nm (threshold 0.8).

NO3 delayed-mode QC (Ken Johnson)

For the DMQC of the Nitrate, the WOA (World Ocean Atlas) is the best reference at the moment.

CONCLUSION AND PERSPECTIVES

The post meeting decision to have three files (Bio-File, core-File, merged-file) let to the following location of “derived” parameters and “raw” parameters. Parameters of geophysical interest will be merged with core Argo (P, T, S) into the merged file, while raw and intermediate parameters will be stored in the so-called Bio-File.

As a consequence and with respect to the spectral dimension

- For the parameters that have a spectral dimension (BBP, irradiance) the wavelength will be explicit in the parameter name and attribute that will be in the merged file

- For the parameters for which spectral measurements are only used for intermediate calculation to produce the final parameter value, the spectral dimension N_VALUE will be implemented in the B-file only.

Therefore:
For the processing of the CHLA at the DAC level, the proposition is:

-CHLA : in the Merged file

-FLUORESCENCE_CHLA: in the B-File

For the processing of the BBP at the DAC level, the proposition is:

-BBP700, BBP532: in the Merged File

-BETA_BACKSCATTERING700, BETA_BACKSCATTERING532: in the B-File

Furthermore, the attributes of the parameter will contain the wavelength of the measurements:

float BBP532(N_PROF, N_LEVELS);
BBP532:long_name = "Particle backscattering measured at 532nm";

float BBP700(N_PROF, N_LEVELS);
BBP700:long_name = "Particle backscattering measured at 700nm";

For the processing of the Nitrate at the DAC level, the proposition is:

-NITRATE: in the Merged File

-UV_INTENSITY_NITRATE: in the B-File

The dimension N_VALUES will be added in the B-file to handle the spectral dimension of the spectrum of the UV absorption for the NITRATE,

float UV_INTENSITY_NITRATE(N_PROF, N_LEVELS,N_VALUES);
UV_INTENSITY_NITRATE:long_name = "Intensity of Ultra violet flux from nitrate sensor";

Whereas different wavelengths for BBP, for example, will be handle defining two parameters in the merged file. (BBP532 and BBP700)