BGC-Argo Data Management Team
“task team”
Visio conference
September 9, 2019
Report of the visio conference of 9 September 2019

All the DACs concerned by BGC-Argo (Aoml, Bodc, Coriolis, Csio, Csiro, Incois, Jma, Meds) are represented.

Participants: Josh Plant, Matt Donnelly, Violetta Paba, Henry Bittig, Thierry Carval, Catherine Schmehchtig, Xiaogang Xing, Liu Zenghong, Christina Schallenberg, Udaya Bhaskar, Kensaku Kobayashi, Anh Tran

Guests: Hervé Claustre, Ken Johnson, Tom Trull

Excused: Tanya Maurer, Annie Wong

A tour is organized to present everybody, a bit complicated with technical issues.

BGC-Argo workshop:

- The BGC-Argo workshop will take place in Villefranche sur Mer, 14 October Afternoon and 15 October.

- Catherine Schmehchtig explained that on the 14th October morning a meeting of the Work Package task 4.2 of the european project EARISE take place, dedicated to the organization of the BGC-Argo quality control (data flow, procedure definitions ...) at the European level. Every BGC-Argo people are more than welcome to attend.

- During the slots allotted to DAC, DAC should emphasize the issues that prevent them from populating ADJUSTED fields and how the BGC-ADMT task team could help reaching this goal. Ken Johnson and Josh Plant insisted on the crucial need to remove/mark data that are obviously bad from the data system, even using a rough QC control procedures.

- The DAC Germany slot will be shorten/remove to give time for Matt Donnelly to present results regarding tests on generic programs (developed in Matlab at BODC, M. Cazaly, V. Paba, M. Donnelly) not developed specifically to sit in the BODC system, but as a genuine toolbox (library). The testing of this toolbox was attempted on various floats from different DACs. They found that the parsing of constants, coefficients and other information in metadata fields varied so much that it took up a significant part of the development activity to be able to parse them and use them to recompute the derived BGC Argo variables from the intermediate parameters. In some cases, this metadata was completely missing, and this reduced the number of useable floats in the testing from around a dozen to 2 or 3.

- In the agenda, there is a slot dedicated to R, A, D, PARAMETER_DATA_MODE definition for BGC and recommendations on how these definitions should be reported in the documentation to warn users. Tanya prepared some slides, based on last year discussion and Henry Bittig’s paper. As the time slot is 45 minutes, we add a discussion on different strategies to estimate DM on a bunch of floats in the same area (Tom’s suggestion).

How to organize us so that we can help each other’s

- As, each institution that runs a DAC has different internal working practices, different software versions, and different ancillary tools in addition to Argo community tools, it is not sustainable for entire software stacks to be developed and shared. It results in the receiving DAC having very limited
ability to contribute to the development of such a software stack, makes them dependent on another
DAC to progress the capability, reduces the chance errors will be picked up in individual pieces of code,
and potentially creates a single point-of-failure. There are potentially also security issues to consider
when a DAC is considering deploying someone else’s software on their own system. Matt’s suggestion
is that we aim to develop a series of community toolboxes that are infrastructure agnostic (not
designed with any one DAC in-mind), which can be shared, cross-tested, and developed collaboratively
as deemed appropriate. For instance, we could have float type specific file readers that are decoupled
from further processing, a full BGC derived variable toolbox and an NRT Automated QC toolbox in the
DAC context. In a detailed QC context, adjusted-mode and delayed-mode QC toolboxes for all
variables, plus options for file readers/writers could be developed which could work with local
infrastructure.

-It should be interesting to share routines (QC) that works directly on Bfile and insert/update QC in the
Bfile. (Anh’s suggestion)

-Regarding DOXY DM procedures, it is not clear which methods should be used to calculate the
adjustment (SAGE-O2, Bittig methods?) (Anh’s remark). Henry will clarify that point. More generally,
the BGC-ADMT task team encourage experts to give stronger recommendations to help in choosing
DM procedures.

-Uday Bhaskar explained that, he already used tools like SAGE-O2 to calculate adjustment necessary
for DOXY fields, but he would also be interested in getting routines that apply this correction in BRfile
and prepare BDfile.

-There is a suggestion to set up workshop to train each other’s on several methods, although it is
complicated to organize and expensive.

Terms of reference

The terms of reference of the BGC-ADMT task team is changed to explain how the co-chairs are chosen
(generic definition).