Plan for BGC DMQC at the MEDS DAC

Christopher Gordon
Fisheries & Oceans Canada

February 1, 2021

1 Introduction

As the `bgcArgoDMQC` package reaches an operational stage of development, it is time to begin performing DMQC on the DOXY floats where possible. Taking into consideration the current data mode state of profiles on the MEDS DAC and what our largest priorities are, I have a few possible "strategies" for doing so presented in section 3.

2 Current Status

The quality flags of TEMP and PSAL affect the quality flags of DOXY, so it is important for the Core DMQC process to be complete before performing the BGC DMQC process. The current data modes of profiles for the MEDS DAC are:

Table 1: Data mode for DOXY and Core variables

Data Mode	# Profiles	Percent
Both RT	390	8.5%
BGC RT, Core DM	825	18.0%
BGC A, Core RT	703	15.3%
BGC A, Core DM	2046	44.6%
Both DM	619	13.5%

Excluding any profiles less than 1 yr old shifts the ditribution slightly, mostly by reducing the number of Core profiles in RT mode:

Table 2: Data mode for DOXY and Core variables for profiles older than 1 yr

Data Mode	# Profiles	Percent
Both RT	240	9.0%
BGC RT, Core DM	725	19.0%
BGC A, Core RT	701	16.2%
BGC A, Core DM	2046	47.1%
Both DM	619	14.3%

Looking at the data mode report generated by Henry Bittig, we can see where we began putting DOXY floats into data mode "A", I assume using the DOXY audit generated by Josh Plant.
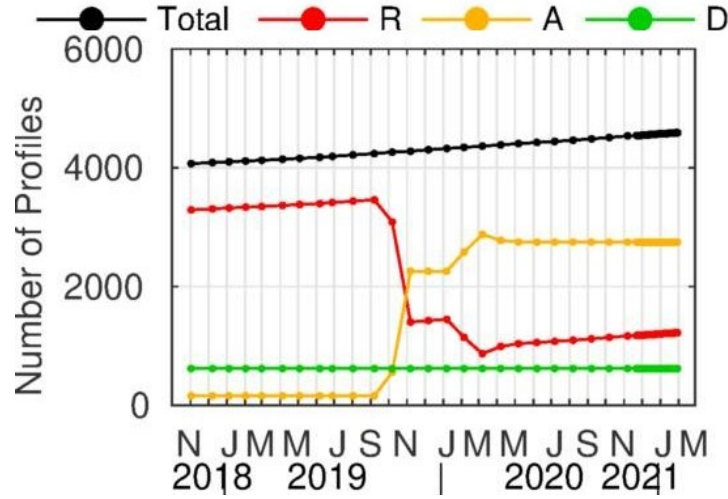

Figure 1: Data mode status of DOXY from report generated by Henry Bittig

3 Priorities and Timeline

Two plan options:

1 (a) Simply begin with the oldest profile, work on all possible profiles for that float, and then move to the oldest profile, and repeat.

1 (b) As above, but first prioritize floats that are no longer operational. This will move entire floats into delayed mode, and also will not merit any

change to RT processing[1]

2 (a) Prioritize floats that already have adjustments made to them.

2 (b) The opposite of above, prioritize floats that are in RT mode and *not* in A mode. Floats that already have adjusted data need DM processing less than those in RT mode, so more value is added by processing those floats first.

As for the timeline, I believe that things will move slowly at first, as I am likely to run into issues for the first few floats or profiles I process. The process will speed up as I add ways to optimize my workflow into the python package.

I think a good and achievable goal would be to be "caught up" and have all required (> 1 yr old) floats DMQC'ed by the next ADMT meeting.

---

[1]When DM files exist for a DOXY float, a temporal extrapolation of the gain is supposed to be applied in RT