argo data management

Argo data management

DOI: https://dx.doi.org/10.13155/40879

Argo Quality Control Manual For Biogeochemical Data

Version 1.1

1st September 2023

1

ARGO

part of the integrated global observation strategy

Argo Quality Control Manual for Biogeochemical Data

Authors: Catherine Schmechtig, Annie Wong, Tanya Maurer, Henry Bittig and Virginie Thierry

How to cite this document

Catherine Schmechtig, Annie Wong, Tanya Maurer, Henry Bittig and Virginie Thierry (2023).
Argo Quality Control Manual for Biogeochemical Data.
https://dx.doi.org/10.13155/40879

Table of contents

History

Date (dd/mmm/yyyy)	Comment
10/Nov/2015	Creation of the document by Virginie, Henry, Catherine, and Annie.
1/March/2016	Version 1.0
4/April/2019	Update version with pH, DOXY, NITRATE
1/Sep/2023	Added Section 5.3. "Recovering BGC data when float salinity is bad", and miscellaneous updates, by Annie Wong and update the document with the new test order and the data transmission to the GTS

Reference Documents

Reference N°	Title	Link
#RD1	Argo Quality Control Manual for CTD and Trajectory Data	http://dx.doi.org/10.13155/33951
#RD2	Argo Quality Control Manual for Biogeochemical Data	http://dx.doi.org/10.13155/40879
#RD3	Argo user manual	http://dx.doi.org/10.13155/29825
#RD4	BGC-Argo quality control manual for Chlorophyll-A concentration	http://dx.doi.org/10.13155/35385
#RD5	BGC-Argo quality control manual for particle backscattering	https://doi.org/10.13155/60262
#RD6	BGC-Argo quality control manual for radiometry	https://doi.org/10.13155/62466
#RD7	BGC-Argo quality control manual for Nitrate concentration	https://doi.org/10.13155/84370
#RD8	BGC-Argo quality control manual for CDOM	TBD
#RD9	BGC-Argo quality control manual for dissolved oxygen concentration	http://dx.doi.org/10.13155/46542
#RD10	Processing BGC-Argo pH data at the DAC level	https://doi.org/10.13155/57195

Preamble

During the ADMT16 meeting in 2015, it was decided to split the Argo quality control manual into two manuals:

- the Argo quality control manual for CTD and trajectory data (#RD1) and,
- the Argo quality control manual for biogeochemical data (this document, #RD2).

As there are many different groups of experts in charge of the assessment of different biogeochemical data set, the Argo quality control manual for biogeochemical data (this document) should be considered as the cover document of all biogeochemical data quality control manuals.

In this document, we present the generic tests applicable to biogeochemical data and we provide links to documentations specific to each biogeochemical parameter, thus reporting in detail all the quality control procedures (both generic and specific tests).

Users should be aware that although biogeochemical data are freely available at the Argo Global Data Assembly Centres (GDACs) along with their CTD data, the accuracy of these biogeochemical data in their raw state is not suitable for direct usage in scientific applications. Users are warned that the raw biogeochemical data should be treated with care, and that often, adjustments are needed before these data can be used for meaningful scientific applications.

Any user of these biogeochemical data who can develop specific adjustment methods to improve their accuracy is welcome to contact the BGC-ADMT on the developing methods.

1. Introduction

This document is the Argo quality control manual for biogeochemical data. It describes two levels of quality control:

- The first level is the real-time system that performs a set of agreed automatic checks.
  - Adjustment in real-time can also be performed and the real-time system can evaluate quality flags for adjusted fields
- The second level is the delayed-mode quality control system.

In core-Argo profile files, where <PARAM> = PRES, TEMP, PSAL, each <PARAM> has 5 qc and adjusted variables:

<PARAM>_QC, PROFILE_<PARAM>_QC, <PARAM>_ADJUSTED, <PARAM>_ADJUSTED_QC, and <PARAM>_ADJUSTED_ERROR.

In B-Argo profile files, <PARAM> can be classified into 3 groups:

(a). B-Argo <PARAM>: these are the ocean state biogeochemical parameters that will receive real-time qc tests, adjustments in real-time and in delayed-mode. They are stored in both the B-Argo profile files and the GDAC synthetic files (s- and sprof-).

(b). IB-Argo <PARAM>: these are the intermediate biogeochemical parameters that are only stored in the b-files. Some of them will receive real-time qc tests and adjustments.

(c). PRES: this is the stand-alone vertical axis that links the core-Argo and B-Argo profile files.

The following are some clarifications on what qc and adjusted variables are included in the B-Argo profile files for these 3 groups of parameters:

(a). B-Argo <PARAM>: all 5 qc and adjusted variables are mandatory for B-Argo PARAM in the B-Argo profile files.

(b). IB-Argo <PARAM>: <PARAM>_QC and PROFILE_<PARAM>_QC are mandatory for IB-Argo <PARAM>. <PARAM>_ADJUSTED, <PARAM>_ADJUSTED_QC and <PARAM>_ADJUSTED_ERROR are optional.

(c). PRES: the B-Argo profile files do not contain any qc or adjusted variables for PRES. They are in the core-Argo profile files.

In B-Argo profile files, biogeochemical parameters can receive adjustments at different times. Therefore the variable PARAMETER_DATA_MODE (N_PROF, N_PARAM) is added to B-Argo profile files to indicate the data mode of each <PARAM>. PARAMETER_DATA_MODE describes the data mode of the individual parameter:
R : real time data
D : delayed mode data
A : real time data with adjusted values

In B-Argo profile files, the variable PARAMETER_DATA_MODE for PRES is always 'R', as adjusted values for PRES are only stored in the core profile files. Thus, to access the 'best' existing version of a parameter (<PARAM>) data, except PRES, the user should:

1. Retrieve the data mode of the <PARAM> parameter (from DATA_MODE(N_PROF) in the core-files and from PARAMETER_DATA_MODE(N_PROF, N_PARAM) in the B-Argo profile files or the s-files),
2. Access the data:
    - If the data mode is ‘R’: In <PARAM>, <PARAM>_QC and PROFILE_<PARAM>_QC,
    - If the data mode is ‘A’ or ‘D’: In <PARAM>_ADJUSTED, <PARAM>_ADJUSTED_QC, PROFILE_<PARAM>_QC and <PARAM>_ADJUSTED_ERROR.

Note that the data mode of an IB-Argo parameter may depend on the DAC decision to include or not the adjusted fields for that IB-Argo parameter in the B-Argo profile file:

- If <PARAM>_ADJUSTED, <PARAM>_ADJUSTED_QC and <PARAM>_ADJUSTED_ERROR are present in the file, the data mode of the IB-Argo parameter can be ‘R’, ‘A’ or ‘D’,
- If not, the data mode of the IB-Argo parameter should always be ‘R’.

2. Real-time quality control for biogeochemical data and associated intermediate parameters

2.1. Introduction

Because of the requirement for delivering data to users within 24-48 hours of the float reaching the surface, the quality control procedures on the real-time data are limited and automatic.

At the present time, real-time tests are defined for the following biogeochemical parameters and the following intermediate parameters:

- DOXY, TEMP_DOXY, PPOX_DOXY,
- CHLA, CHLA_FLUORESCENCE
- BBP,
- DOWN_IRRADIANCExxx, DOWNWELLING_PAR
- NITRATE,
- pH.

QC of other biogeochemical parameters and other intermediate parameters should be set to ‘0’ (no QC was performed).

Each defined test has a unique number used to report the passed and failed tests list in the HISTORY_QCTEST variable in the Argo data files (#RD3).

The list of the defined tests is the following.

Test number	Test name	Type of parameter involved ('c', 'ic', 'b', 'ib')
1	Platform Identification	all
2	Impossible Date Test	all
3	Impossible Location Test	all
4	Position on Land Test	all
5	Impossible Speed Test	all
6	Global Range Test	all
7	Regional Range Test	'c' and 'ib'
8	Pressure Increasing Test	'c'
9	Spike Test	all
10	Top and Bottom Spike Test : removed	n/a
11	Gradient Test	'ib' and 'b'
12	Digit Rollover Test	'c' and 'ib'
13	Stuck Value Test	all
14	Density Inversion	'c'
15	Grey List	all
16	Gross salinity or temperature sensor drift	'c' and 'ib'
17	Visual QC	all
18	Frozen profile	'c' and 'ib'
19	Deepest pressure test	all
20	Questionable Argos position test	all
21	Near-surface unpumped CTD salinity test	'c'
22	Near-surface mixed air/water test	all
23	Real-time Quality Control Flag Scheme for float data deeper than 2000 dbar	all
24	specific test for RBRargo3|2K CTD data < 2000 dbar	'c'
25	MEDD Test	'c'
26	specific test for RBRargo3|2K TEMP_CNDC data	'c'
27 to 55	Not used	
56	pH specific test	'b'
57	DOXY specific test	'b'
58	CDOM specific test	'b'
59	NITRATE specific test	'b'
60	PAR specific test	'b'
61	IRRADIANCE specific test	'b'
62	BBP specific tests	'b'
63	CHLA specific tests	'b'

Table 1. Defined Tests

The test application order is as follow.

Order	Test nb	Test name	VP	NS	DP	TR
1	1	Platform Identification test				
2	2	Impossible Date test				
3	3	Impossible Location test				
4	4	Position on Land test				
5	20	Questionable Argos position test				
6	5	Impossible Speed test				
7	15	Grey list	x			x
8	19	Deepest pressure test	x			x
9	21	Near-surface unpumped CTD salinity test (NS)		x		x
10	22	Near-surface mixed air/water test (NS)		x		x
11	6	Global Range test	x			x
12	7	Regional Range test	x			x
13	8	Pressure increasing				
14	9	Spike test	x			
15	11	Gradient test	x			
16	25	MEDD				
17	12	Digit Rollover test	x			
18	13	Stuck value	x			
19	14	Density inversion				
20	16	Gross salinity or temperature sensor drift	x			
21	18	Frozen profile	x			
22	26	specific test for RBRargo3|2K TEMP_CNDC data				
23	24	specific test for RBRargo3|2K CTD data < 2000 dbar				
24	23	Real-time Quality Control Flag Scheme for float data deeper than 2000 dbar			x	
25	”appropriate test number”	BGC specific test	57, 59, 62, 63			57, 62
26	17	Visual QC				

Table 2. Global test application order,

The tests are applied or not based on the sampling scheme and the parameters, VP stands for Vertical Profiles, NS for Near Surface (the vertical sampling scheme of the profiles which start with “Near-surface sampling:”, for trajectory data it is selected from NS dedicated measurement codes), DP for Deep Profile (deep float have one of the following PLATFORM_TYPE: APEX_D, ARVOR_D, SOLO_D, NINJA_D, HM4000, HM6000), TR for TRajectory. “x” indicates whether there is a specific entry for the test associated to a BGC parameter.

2.2. Argo real-time quality control tests on vertical profiles

2.2.1. Common Argo real-time quality control tests on vertical profiles

This section lists the real-time tests that are common between CTD data and biogeochemical data with specific enties for BGC parameters. The same real-time test numbers for CTD data are used here. The common tests with no specific entry for BGC parameters are described in the Argo quality control manual (#RD1).

6. Global range test

This test applies a gross filter on observed values for DOXY, TEMP_DOXY, CHLA, BBP, NITRATE, IRRADIANCE, PAR and pH.

- DOXY in range [−5, 600] micromole/kg (#RD9)
- TEMP_DOXY in range [−2.5, 40.0] °C (#RD9)
- CHLA in range [−0.2, 100] mg/m^3 (#RD4)
- CHLA_FLUORESCENCE in range [−0.2, 100] RU (#RD4)
- CHLA_ADJUSTED in range [−0.1, 50] mg/m^3 (#RD4)
- NITRATE in range [−2, 50] micromole/kg (#RD7)
- PH_IN_SITU_TOTAL in range [7.0 , 8.8] (#RD10)
- PH_IN_SITU_TOTAL_ADJUSTED in range [7.3 , 8.5] (#RD10)
- DOWN_IRRADIANCExxx in W.m-2 .nm-1 (#RD6)
  - At 380nm in range [-1, 1.7]
  - At 412nm in range [-1, 2.9]
  - At 443nm in range [-1, 3.2]
  - At 490nm in range [-1, 3.4]
- DOWNWELLING_PAR in range [-1, 4672] µmolQuanta.m-2 .s-1 (#RD6)

Action: Values that fail this test should be flagged with a QC = '4' for DOXY, TEMP_DOXY, CHLA, NITRATE, IRRADIANCE, PAR and pH

7. Regional range test

This test applies to certain regions of the world where conditions can be further qualified. In this case, specific ranges for observations from the Mediterranean Sea and the Red Sea further restrict what are considered sensible values. The Red Sea is defined by the region 10N, 40E; 20N, 50E; 30N, 30E; 10N, 40E. The Mediterranean Sea is defined by the region 30N, 6W; 30N, 40E; 40N, 35E; 42N, 20E; 50N, 15E; 40N, 5W; 30N, 6W.

Red Sea

- TEMP_DOXY in range [21.7, 40.0] °C

Mediterranean Sea

- TEMP_DOXY in range [10.0, 40.0] °C

Action: If a value fails this test, it should be flagged as bad data (QC = ‘4’).

9. Spike test

Spike test for DOXY and TEMP_DOXY

The difference between sequential measurements, where one measurement is significantly different from adjacent ones, is a spike in both size and gradient. This test does not consider differences in depth but assumes a sampling that adequately reproduces changes in DOXY and TEMP_DOXY with depth.

Test value = | V2 − (V3 + V1)/2 | − | (V3 − V1) / 2 |

where V2 is the measurement being tested as a spike, and V1 and V3 are the values above and below.

For DOXY: The V2 value is flagged when

- the test value exceeds 50 micromol/kg for pressures less than 500 dbar, or
- the test value exceeds 25 micromol/kg for pressures greater than or equal to 500 dbar.

For TEMP_DOXY: The V2 value is flagged when

- the test value exceeds 6 °C for pressures less than 500 dbar, or
- the test value exceeds 2 °C for pressures greater than or equal to 500 dbar.

Action: Values considered as a spike should be flagged as bad data (QC = ’4’)

Spike test for NITRATE

We calculate the absolute difference between the nitrate concentration at a certain depth (V2) and a running median (5 values, V0, V1, V2, V3, V4) along the whole profile:

- TestValue = ABS[V2 - MEDIAN(V0,V1,V2,V3,V4)]

This test applies to certain regions of the world where conditions can be further qualified. In this case, specific ranges for observations from the Mediterranean Sea and the Red Sea further restrict what are considered sensible values.

The Red Sea is defined by the region 10N, 40E; 20N, 50E; 30N, 30E; 10N, 40E.

The Mediterranean Sea is defined by the region 30N, 6W; 30N, 40E; 40N, 35E; 42N, 20E; 50N, 15E; 40N, 5W; 30N, 6W.

Action:

In the Red Sea and the Mediterranean Sea:

IF TestValue > 1 micromole/kg THEN the test failed and V2 should be flagged with a QC = ‘4’ for NITRATE.

Other places:

IF TestValue > 5 micromole/kg THEN the test failed and V2 should be flagged with a QC = '4' for NITRATE.

Spike test for pH

Test value = |V2 – median(V0, V1, V2, V3, V4)|

where the test value represents the anomaly of the observed pH from the median of the surrounding data.

For pH, a data point is considered a spike and marked with quality flag '4' (bad data)

- if Test value > 0.04 * pH

Action: Values considered as a spike should be flagged as bad data (QC = '4').

11. Gradient test

Gradient test for DOXY and TEMP_DOXY

This test is failed when the difference between vertically adjacent measurements is too steep. The test does not consider differences in depth but assumes a sampling that adequately reproduces changes in DOXY and TEMP_DOXY with depth.

Test value = | V2 − (V3 + V1)/2 |

where V2 is the measurement being tested as a spike, and V1 and V3 are the values above and below.

For DOXY: the V2 value is flagged when

- the test value exceeds 50 micromol/kg for pressures less than 500 dbar, or
- test value exceeds 25 micromol/kg for pressures greater than or equal to 500 dbar.

For TEMP_DOXY: The V2 value is flagged when

- the test value exceeds 9 °C for pressures less than 500 dbar, or
- the test value exceeds 3 °C for pressures greater than or equal to 500 dbar.

Action: Values that fail this test should be flagged as bad data (QC = '4').

12. Digit rollover test

Only so many bits are allowed to store temperature values in a profiling float. This range is not always large enough to accommodate conditions that are encountered in the ocean. When the range is exceeded, stored values rollover to the lower end of the range. This rollover should be detected and compensated for when profiles are constructed from the data stream from the float.

This test is used to make sure the rollover is properly detected:

- TEMP_DOXY difference between adjacent pressures > 10 °C.

Action: When a rollover is detected, both values used to detect the jump in the data should be flagged with a QC = '4', other values of the profile with a QC = '3'.

13. Stuck value test

This test looks for all biogeochemical sensor outputs (i.e. ‘i’ and ‘b’ parameter measurements transmitted by the float) in a vertical profile being identical.

Action: Stuck values should be flagged as bad data (QC = ’4’).

15. Grey list

See Argo quality control manual (#RD1).

16. Gross temperature sensor drift

This test is implemented to detect a sudden and significant sensor drift. It calculates the average temperature from the deepest 100 dbar of a profile and the previous good profile. Only measurements with good QC are used.

Action: For TEMP_DOXY, if the difference between the two average values is more than 1°C, then all TEMP_DOXY values from the profile are flagged as probably bad data (QC = ‘3’).

17. Visual QC

See Argo quality control manual (#RD1).

18. Frozen profile test

This test is used to detect a float that reproduces the same profile (with very small deviations) over and over again.

Typically the differences between two profiles are of the order of 0.01 °C for temperature.

A). Derive TEMP_DOXY profiles by averaging the original profiles to get mean values for each profile in 50 dbar slabs (TEMP_DOXY_prof, TEMP_DOXY_previous_prof). This is necessary because the floats do not sample at the same level for each profile.

B). Obtain absolute values of the difference between the averaged temperature profiles as follows:

- delta_TEMP_DOXY = abs(TEMP_DOXY_prof - TEMP_DOXY_previous_prof)

C). Find the maximum, minimum and mean of the absolute values of the averaged differences between profiles for temperature:

- max(delta_TEMP_DOXY), min(delta_TEMP_DOXY), mean(delta_TEMP_DOXY)

D). To fail the test, require that:

- max(delta_TEMP_DOXY) < 0.3 °C
- min(delta_TEMP_DOXY) < 0.001°C
- mean(delta_TEMP_DOXY) < 0.02 °C

Action: If a profile fails this test, all measurements from the profile are flagged as bad data (QC = ‘4’). If a float fails this test on 5 consecutive cycles, it is inserted in the grey list.

19. Deepest pressure test

This test requires that a profile has pressures that are not greater than a pressure threshold, where PRESSURE_THRESHOLD = CONFIG_ProfilePressure_dbar + tolerance. See Argo quality control manual for details (#RD1).

Action: When PRES > PRESSURE_THRESHOLD, Measurements at those pressures should be flagged with QC = '3'.

2.2.2. Specific Argo real-time quality control tests on vertical profiles

2.2.2.1. For DOXY

DOXY specific tests are described in "BGC-Argo quality control Manual for dissolved oxygen concentration" (#RD9).

They all should be reported with the same test #57

2.2.2.2. For CHLA

63. CHLA specific Argo real-time quality control tests

CHLA specific tests are described in "BGC-Argo quality control manual for Chlorophyll-A concentration" (#RD4).

They all should be reported with the same test #63.

2.2.2.3. For BBP

62. BBP specific Argo real-time quality control tests

BBP specific tests are described in "BGC-Argo quality control manual for particle backscattering" (#RD5)

They all should be reported with the same test #62.

2.2.2.4. For IRRADIANCE

61. IRRADIANCE specific Argo real-time quality control tests

IRRADIANCE specific tests are described in "BGC-Argo quality control manual for radiometry" (#RD6)

They all should be reported with the same test #61.

2.2.2.5. For PAR

60. PAR specific Argo real-time quality control tests

PAR specific tests are described in "BGC-Argo quality control manual for radiometry" (#RD6)

They all should be reported with the same test #60.

2.2.2.6. For NITRATE

59. NITRATE specific Argo real-time quality control tests

NITRATE specific tests are described in "BGC-Argo quality control manual for nitrate concentration" (#RD7)

They all should be reported with the same test #59.

2.2.2.7. For CDOM

58. CDOM specific Argo real-time quality control tests

CDOM specific tests are described in "BGC-Argo quality control manual for CDOM" (#RD8)

They all should be reported with the same test #58.

2.2.2.8. For pH

56. pH specific Argo real-time quality control tests

pH specific tests are described in "BGC-Argo quality control manual for pH" (#RD8)

They all should be reported with the same test #56.

2.2.3. Scientific calibration information for each profile

If PARAMETER_DATA_MODE is 'R', there is no reason to fill the scientific calibration information, thus:

For PARAMs (B-Argo PARAMs and IB-Argo PARAMs) in 'R'-mode	
SCIENTIFIC_CALIB_COMMENT	FillValue
SCIENTIFIC_CALIB_EQUATION	FillValue
SCIENTIFIC_CALIB_COEFFICIENT	FillValue
SCIENTIFIC_CALIB_DATE	FillValue

A specific comment should however be set for PRES parameter

For PRES	
SCIENTIFIC_CALIB_COMMENT	'Adjusted values are provided in the core profile file'
SCIENTIFIC_CALIB_EQUATION	FillValue
SCIENTIFIC_CALIB_COEFFICIENT	FillValue
SCIENTIFIC_CALIB_DATE	FillValue

(see in Chapter 3 and Chapter 0 how to fill scientific calibration information when PARAMETER_DATA_MODE is 'A' or 'D' respectively).

2.3. Argo real-time quality control tests on trajectory file data

The following tests are applied in real-time to data stored in the trajectory files. These are data collected during the park-and-drift phase and the surface interval of a float's mission. Some trajectory file data are duplicates of data from the vertical profiles. For example, data from vertical levels with timing information from PROVOR/ARVOR floats are stored in the profile files (without their times) and duplicated in the trajectory files (with their associated times). These data should be duplicated with their associated QC values, which are set during the real-time quality control tests performed on the vertical profiles.

2.3.1. Common Argo real-time quality control tests on trajectory file data

6. Global range test

See 2.2.1.

7. Regional range test

See 2.2.1.

15. Grey list

See 2.2.1.

21. Near-surface unpumped DOXY test

Unpumped DOXY data from SBE63 oxygen sensor are of dubious quality because the flow rate and the water masses in front of the sensing foil are wrong when the pump is switched off. This test specifies that unpumped (or partially pumped) oxygen data returned by an SBE63 sensor should be flagged as “probably bad data” in real time.

Action: Unpumped (or partially pumped) oxygen data returned by an SBE63 sensor are flagged as “probably bad data” in real time. That is, DOXY_QC = ‘3’.

22. Near-surface mixed air/water test

Near-surface mixed air/water test for DOXY and TEMP_DOXY

Most near-surface profiles extend all the way to the sea surface. Therefore, the shallowest part of a near-surface profile will contain some mixed air/water measurements. This test identifies broadly the pressures at which this shallowest part of a near-surface profile takes place and specifies that data in that pressure range are “probably bad data”.

Action: Data from the shallowest part of a near-surface profile, which may contain mixed air/water measurements, are flagged as “probably bad data” (QC = ‘3’) in real-time.

(a). PROVOR/ARVOR (bin-averaged data)

For PROVOR/ARVOR floats that return bin-averaged data, if the first bin closest to the sea surface has PRES <= 1 dbar, then the temperature value from that first bin is suspected to contain averages of mixed air/water measurements and should be flagged as "probably bad data". That is, TEMP_DOXY_QC = '3'.

(b). PROVOR/ARVOR (spot-sampled data)

For PROVOR/ARVOR floats that return spot-sampled data, temperature observed at PRES <= 0.5 dbar should be flagged as "probably bad data". That is, TEMP_DOXY_QC = '3'.

(c). NOVA

For NOVA floats, if the first bin closest to the sea surface has PRES <=1 dbar, then the temperature value from that first bin is suspected to contain averages of mixed air/water measurements and should be flagged as "probably bad data". That is, TEMP_DOXY_QC = '3'.

(d). APEX

The shallowest part of near-surface profiles collected by APEX floats will contain pressure readings that are not necessarily monotonic with time. For APEX (ARGOS) floats with the NST firmware, when near-surface data have pressures shallower than 5 dbar, check difference in pressure between two successive measurements. If the difference is less than 0.5 dbar from the previous measurement, then data from that level and all levels after that should be flagged as "probably bad data". That is TEMP_DOXY_QC = '3'. For APEX (Iridium) floats equipped with the STS module, when near-surface data have pressures shallower than 5 dbar, check that pressure readings decrease monotonically with time. Data from the level when monotonicity stops and all levels after that should be flagged as "probably bad data". That is, TEMP_DOXY_QC = '3'.

2.3.2. Specific Argo real-time quality control tests on trajectory file data

2.3.2.1. For DOXY

DOXY specific tests for trajectory are described in "BGC-Argo quality control Manual for dissolved oxygen concentration" (#RD9).

2.3.2.2. For BBP

BBP specific tests for trajectory are described in "BGC-Argo quality control Manual for particle backscattering" (#RD5).

2.4. Argo real-time quality control tests on near-surface data

The near-surface data described in this section are specialised data that are collected with vertical sampling methods that are different from the primary CTD profiles. For most profiling floats, the CTD pump is normally switched off at around 5 dbar during ascent to avoid contamination of the conductivity cell. Several float types continue to sample up to the sea surface (with the pump off, or with the pump on closer to the sea surface) or carry auxiliary modules for high-resolution near-surface sampling. These specialised near-surface data are focused on the top 5 dbar of the ocean. They may extend deeper than 5 dbar so as to overlap with the primary CTD profiles for the purpose of cross-calibration. They are stored as additional profiles (N_PROF > 1) in the B-Argo profile files, and are identifiable by VERTICAL_SAMPLING_SCHEME = "Near-surface sampling". (Note: the full character string is “Near-surface sampling: averaged/discrete/mixed, pumped/unpumped [optional description]”. Please refer to Table 16 in the Argo Users Manual for details of the various vertical sampling schemes and their full character strings.) The following tests are applied in real-time to these specialised near-surface data.

Test 21. Near-surface unpumped DOXY test

See 2.3.1

Test 22. Near-surface mixed air/water test

Near-surface mixed air/water test for DOXY and TEMP_DOXY
See 2.3.1

2.5. Argo real-time quality control tests for deep float data

Some profiling floats equipped with an oxygen sensor have the capability to sample deeper than the original Argo profiling pressure target of 2000 dbar. However the accuracy of the raw data from these deep Argo floats below 2000 dbar is not yet well understood. Pilot studies of deep Argo data indicated possible pressure dependent salinity and oxygen bias, and the performance of the pressure sensor below 2000 dbar has not been validated by manufacturers. Therefore an interim real-time quality control flag scheme is implemented to warn users that the raw data from deep Argo floats below 2000 dbar may not be suitable for research applications requiring high data accuracy.

Test 23. Test for float data deeper than 2000 dbar

This test specifies that for float data deeper than 2000 dbar:

- real-time QC flag = ‘2’for temperature (TEMP_DOXY);

Action:
For profiles from deep Argo floats that extend deeper than 2000 dbar, real-time qc tests defined previously should be applied to the full-depth profiles. Float data deeper than 2000 dbar that have passed all previous real-time qc tests should be flagged in the following way TEMP_DOXY_QC = ‘2’.

2.6. Test application order between parameters

The following constraints should be observed to ensure that data useful or data needed for the quality control of a given parameter have been QCed prior to their use.

If the given parameter is available on a given float, then:

- Core QC precedes all B-Argo parameter QC
- DOXY QC precedes NITRATE QC
- DOXY QC precedes pH QC
- DOWN_IRRADIANCE and DOWNWELLING_PAR QC precedes CHLA QC

- BBP QC precedes CHLA QC
- CDOM QC precedes CHLA QC

2.7. Quality control flag application policy

The QC flag value assigned by a test cannot override a higher value from a previous test. Example: a QC flag ‘4’ (bad data) set by the Grey List Test cannot be decreased to QC flag ‘3’ (bad data that are potentially correctable) set by the Gross Temperature Drift Test.

A value with QC flag ‘4’ (bad data) is ignored by other quality control tests.

When a biogeochemical parameter is calculated from other intermediate (‘ib’ or ‘ic’) parameter or biogeochemical (‘b’) parameter data, its associated QC is set to the worse QC value of the input data.

For example, CHLA (‘b’ parameter) is calculated from FLUORESCENCE_CHLA (‘ib’ parameter). Then if FLUORESCENCE_CHLA_QC = ’4’ after failing the Stuck Value Test, the corresponding CHLA_QC is set to ‘4’.

3. Real-Time adjustment of biogeochemical data

3.1. Compulsory variables to be filled for data adjusted in Real-Time

When a B-Argo <PARAM> receives an adjustment in real-time, the following 5 mandatory _QC and _ADJUSTED variables must be filled in the BR profile file:

- <PARAM>_QC
- PROFILE_<PARAM>_QC
- <PARAM>_ADJUSTED
- <PARAM>_ADJUSTED_QC
- <PARAM>_ADJUSTED_ERROR (not mandatory but recommended)

As a consequence, the real-time quality control tests defined in the previous section are performed on the ADJUSTED parameters similarly to the raw parameters. PROFILE_<PARAM>_QC should be re-computed when <PARAM>_ADJUSTED_QC becomes available.

In addition, the SCIENTIFIC_CALIB section must be filled with adequate information on the real-time adjustment.

3.2. Scientific calibration information for each profile

When a biogeochemical parameter (B-Argo 'b' <PARAM>) has been through an adjustment procedure, its PARAMETER_DATA_MODE is set to ‘A’ which means “adjusted in real-time”. The PARAMETER_DATA_MODE of all intermediate parameters (IB-Argo 'i' <PARAM>) associated to this adjusted biogeochemical parameter are also set to ‘A’ when they have an _ADJUSTED field (but let to ‘R’ if not).

If PARAMETER_DATA_MODE is ‘A’, none of the scientific calibration information should be set to FillValue and every information should be filled.

As mentioned in §1, for IB-Argo <PARAM>, while <PARAM>_QC and PROFILE_<PARAM>_QC are mandatory, <PARAM>_ADJUSTED, <PARAM>_ADJUSTED_QC and <PARAM>_ADJUSTED_ERROR are optional (DAC decision).

For IB-Argo PARAMs with no corresponding _ADJUSTED field and for which the associated B-Argo PARAMs have been through adjustment in real-time	
SCIENTIFIC_CALIB_COMMENT	‘not applicable’
SCIENTIFIC_CALIB_EQUATION	‘not applicable’
SCIENTIFIC_CALIB_COEFFICIENT	‘not applicable’
SCIENTIFIC_CALIB_DATE	YYYYMMDDHHMISS(*)

For IB-Argo PARAMs with corresponding _ADJUSTED fields and for which the associated B-Argo PARAMs have been through adjustment in real-time	
SCIENTIFIC_CALIB_COMMENT	Content depends on <PARAM>. See #RD4 for CHLA,

	#RD5 for BBP #RD6 for Radiometry #RD7 for NITRATE, #RD8 for CDOM, #RD9 for DOXY, #RD10 for pH.
SCIENTIFIC_CALIB_EQUATION	Content depends on <PARAM>. See #RD4 for CHLA, #RD5 for BBP #RD6 for Radiometry #RD7 for NITRATE, #RD8 for CDOM, #RD9 for DOXY, #RD10 for pH.
SCIENTIFIC_CALIB_COEFFICIENT	Content depends on <PARAM>. See #RD4 for CHLA, #RD5 for BBP #RD6 for Radiometry #RD7 for NITRATE, #RD8 for CDOM, #RD9 for DOXY, #RD10 for pH.
SCIENTIFIC_CALIB_DATE	YYYYMMDDHHMISS(*)

For B-Argo PARAMs that have been through adjustment in Real-Time	
SCIENTIFIC_CALIB_COMMENT	Content depends on <PARAM>. See #RD4 for CHLA, #RD5 for BBP #RD6 for Radiometry #RD7 for NITRATE, #RD8 for CDOM, #RD9 for DOXY, #RD10 for pH.
SCIENTIFIC_CALIB_EQUATION	#RD4 for CHLA, #RD5 for BBP #RD6 for Radiometry

	#RD7 for NITRATE, #RD8 for CDOM, #RD9 for DOXY, #RD10 for pH.
SCIENTIFIC_CALIB_COEFFICIENT	Content depends on <PARAM>. See #RD4 for CHLA, #RD5 for BBP #RD6 for Radiometry #RD7 for NITRATE, #RD8 for CDOM, #RD9 for DOXY, #RD10 for pH.
SCIENTIFIC_CALIB_DATE	YYYYMMDDHHMISS(*)

(*): for a given calibration, the SCIENTIFIC_CALIB_DATE of an adjusted B-Argo parameter and of its associated IB-Argo parameters should be identical.

The three fields SCIENTIFIC_CALIB_COMMENT, _EQUATION, and _COEFFICIENT have netCDF dimensions (N_PROF, N_CALIB, N_PARAM, STRING256). This means that for each N_CALIB, each field is a 256-length character string. If character strings longer than 256-length are needed, the procedure should be separated and stored as multiple N_CALIB.

For a calibration that needs multiple N_CALIB:

- the SCIENTIFIC_CALIB_DATE should be identical for all N_CALIB,
- once the different fields are correctly filled, the remaining empty fields (unused) should be filled as follows:
  - ✓ SCIENTIFIC_CALIB_COMMENT: ‘No additional comment’,
  - ✓ SCIENTIFIC_CALIB_EQUATION: ‘No additional equation’,
  - ✓ SCIENTIFIC_CALIB_COEFFICIENT:’No additional coefficient’.

4. Real-time transfer of BGC-Argo data to the GTS

The value of assimilating Argo temperature and salinity data into operational ocean and coupled models, as well as global reanalysis systems, is now well understood (Penny et al. 2017; Johnson et al. 2022). As the use and development of coupled domain and biogeochemical (BGC) models expands, impacts from BGC parameter assimilation are also emerging (Verdy & Mazloff 2017; Cossarini et al. 2019; Fennel et al. 2019; Ford et al. 2020; Shu et al. 2022). A fundamental requirement for assimilation of Argo data by oceanic and climatic operational centers is timely delivery of real-time data to the Global Telecommunications System (GTS). Argo temperature and salinity data (TEMP and PSAL) and associated quality flags are delivered to the GTS in Binary Universal Form for the Representation of meteorological data (BUFR) format within 12 to 24 hours of satellite transmission. In order to further increase the access of BGC-Argo data to global operational centers, BUFR data streams should be initiated for BGC-Argo parameters as well.

Unlike Argo temperature and salinity data which, in their raw state, generally meet accuracy requirements needed for operational use, raw BGC-Argo data are typically not suitable for direct scientific use. This is stated explicitly in the preamble of this document and is due, in part, to biogeochemical sensors representing more emergent technology in comparison to CTDs. Many biogeochemical sensors are prone to drift significantly out of calibration prior to deployment.

Given these considerations, it is recommended that only real-time adjusted BGC-Argo data (<PARAMETER>_ADJUSTED with PARAMETER_DATA_MODE = ‘A’) be distributed on the GTS, unless explicitly stated otherwise. It is understood that this precludes the transfer and subsequent use of BGC-Argo data within the first few months of the float’s life once deployed. However, unlike core-Argo data which typically receive initial delayed-mode assessments after one year, most BGC-Argo data receive initial adjustments within three months of float deployment, and <PARAMETER>_ADJUSTED data-streams subsequently become available in real-time (see Bittig et al. 2019 for further description of BGC-Argo data-modes).

As BUFR format supports the inclusion of quality flags, all <PARAMETER>_ADJUSTED_QC values will be transmitted along with their associated <PARAMETER>_ADJUSTED data. Operational centers interested in accessing and/or assimilating Argo data from the GTS are advised to only select data with quality flag = ‘1’, representing ‘good data’ as is described in Argo reference table 2 (Argo QC Manual for CTD and Trajectory data, http://dx.doi.org/10.13155/33951). ‘Good data’ reflects parameter measurements that have passed all parameter-specific real-time tests.

Note to DACs: Similar to Argo temperature and salinity data, BGC-Argo data from vertical profiles that do not pass real-time quality control tests 1 & 2 should not be distributed to the GTS.

Table 3 shows the units, range and maximum expected precision for BGC-Argo parameters. These specifications are also listed in Reference Table 3 (parameter code table; http://www.argodatamgt.org/content/download/30910/209488/file/argo-parameters-list-core-and-b_20230203.xlsx). They are used to determine the BUFR format for each BGC-Argo parameter. The ranges listed are intended to fully cover all possible conditions (climatological and episodic) throughout the global ocean. Precision is sensor dependent, and “maximum expected precision” reflects the maximum precision that can be reached for some sensors.

BGC-Argo parameter	Descriptor	Units	Range	Precision	BUFR sequence
DOXY	Dissolved Oxygen	$\mu mol\ kg^{-1}$	[-5, 600]	0.001	3-06-044
NITRATE	Nitrate	$\mu mol\ kg^{-1}$	[-15, 65]	0.001	3 06 046
PH_IN_SITU_TOTAL	pH in situ	<unitless>	[7.0, 8.8]	0.0001	3 06 047
CHLA	Chlorophyll-A	$mg\ m^{-3}$	[-0.1, 50]	0.0001	3 06 045
BBP700	Particle Backscattering at 700 nm	m^{-1}	[-2.5e-5, 0.1]	0.0000001	3 06 048

Table 3: Units, range and maximum expected precision for each BGC-Argo parameter. Please refer to Argo Reference Table 3. These are used to determine the BURF format. Note that only the real-time adjusted values of these BGC-Argo parameters (<PARAMETER_ADJUSTED> with <PARAMETER_DATA_MODE> = 'A') are distributed on the GTS. Approved BUFR sequences listed are further described in WMO-No 306, V1.2 https://library.wmo.int/doc_num.php?explnum_id=11283

5. Delayed mode quality control of biogeochemical data

5.1. Editing raw qc and adjusted qc flags in delayed-mode

Delayed-mode operators should examine profile data for pointwise errors such as spikes and jumps, and edit and check the qc flags in <PARAM>_QC and <PARAM>_ADJUSTED_QC (when the adjustment is performed in Real Time). Here, <PARAM> refers to the biogeochemical parameters that have been through the delayed-mode process.

Examples where <PARAM>_QC, <PARAM>_ADJUSTED_QC should be edited in delayed-mode include:

- <PARAM>_QC/<PARAM>_ADJUSTED_QC should be changed to '4' for bad and uncorrectable data that are not detected by the real-time tests; and
- <PARAM>_QC/<PARAM>_ADJUSTED_QC should be changed to '1' or '2' for good data that are wrongly identified as probably bad by the real-time tests.

5.2. Compulsory variables to be filled in a BD profile file

This section lists the compulsory variables that must be filled in a BD profile file.

5.2.1. QC and ADJUSTED variables

Each B-Argo <PARAM> has 5 mandatory qc and adjusted variables in the B- profile file:

- <PARAM>_QC
- PROFILE_<PARAM>_QC
- <PARAM>_ADJUSTED
- <PARAM>_ADJUSTED_QC
- <PARAM>_ADJUSTED_ERROR

When a B-Argo <PARAM> has been through the delayed-mode process, the above 5 mandatory qc and adjusted variables must be filled in the BD profile file. PROFILE_<PARAM>_QC should be re-computed when <PARAM>_ADJUSTED_QC becomes available.

For IB-Argo <PARAM>, <PARAM>_QC and PROFILE_<PARAM>_QC are mandatory, but the 3 adjusted variables are optional in the B-Argo profile file:
<PARAM>_ADJUSTED, <PARAM>_ADJUSTED_QC, <PARAM>_ADJUSTED_ERROR.

If a data centre chooses to include these 3 adjusted variables for IB-Argo <PARAM> in the B-profile file, then these 3 adjusted variables must be filled when the IB-Argo <PARAM> has been through the delayed-mode process, and PROFILE_<PARAM>_QC should be re-computed with <PARAM>_ADJUSTED_QC.

Note that PRES in the B-Argo profile file does not carry any qc or adjusted variables. It is used as a stand-alone vertical index that links the core-Argo and B-Argo profile files. Users who want delayed-mode adjusted pressure values (PRES_ADJUSTED) should obtain them from the core-Argo profile files.

5.2.2. Scientific calibration information for each profile

It is compulsory to fill the scientific calibration section of a BD- profile file.

PARAMETER should contain every parameter recorded in STATION_PARAMETER (including PRES), even though not all STATION_PARAMETER have delayed-mode qc.

When a biogeochemical parameter ('b' parameter) has been through a delayed-mode procedure its PARAMETER_DATA_MODE is set to 'D'. The PARAMETER_DATA_MODE of all intermediate parameters ('i' parameters) associated to this adjusted biogeochemical parameter are also set to 'D' when they have an _ADJUSTED field (but set to 'R' if not).

If PARAMETER_DATA_MODE is 'D', none of the scientific calibration information should be set to FillValue and every information should be filled.

Here are the instructions on how to fill the scientific calibration section of a BD profile file.

For IB-Argo PARAMs with no corresponding _ADJUSTED field and for which the associated B-Argo PARAMs have been through delayed-mode qc	
SCIENTIFIC_CALIB_COMMENT	'not applicable'
SCIENTIFIC_CALIB_EQUATION	'not applicable'
SCIENTIFIC_CALIB_COEFFICIENT	'not applicable'
SCIENTIFIC_CALIB_DATE	YYYYMMDDHHMISS(*)

For IB-Argo PARAMs with corresponding _ADJUSTED fields and for which the associated B-Argo PARAMs have been through delayed-mode qc	
SCIENTIFIC_CALIB_COMMENT	Content depends on <PARAM>, See #RD4 for CHLA, #RD5 for BBP

	#RD6 for Radiometry #RD7 for NITRATE, #RD8 for CDOM, #RD9 for DOXY, #RD10 for pH.
SCIENTIFIC_CALIB_EQUATION	Content depends on <PARAM>, See #RD4 for CHLA, #RD5 for BBP #RD6 for Radiometry #RD7 for NITRATE, #RD8 for CDOM, #RD9 for DOXY, #RD10 for pH.
SCIENTIFIC_CALIB_COEFFICIENT	Content depends on <PARAM>, See #RD4 for CHLA, #RD5 for BBP #RD6 for Radiometry #RD7 for NITRATE, #RD8 for CDOM, #RD9 for DOXY, #RD10 for pH.
SCIENTIFIC_CALIB_DATE	YYYYMMDDHHMISS(*)

For PARAMs that have been through delayed-mode qc	
SCIENTIFIC_CALIB_COMMENT	Content depends on <PARAM>, See #RD4 for CHLA, #RD5 for BBP #RD6 for Radiometry #RD7 for NITRATE, #RD8 for CDOM, #RD9 for DOXY, #RD10 for pH.
SCIENTIFIC_CALIB_EQUATION	Content depends on <PARAM>, See #RD4 for CHLA, #RD5 for BBP #RD6 for Radiometry #RD7 for NITRATE,

	#RD8 for CDOM, #RD9 for DOXY, #RD10 for pH.
SCIENTIFIC_CALIB_COEFFICIENT	Content depends on <PARAM>, See #RD4 for CHLA, #RD5 for BBP #RD6 for Radiometry #RD7 for NITRATE, #RD8 for CDOM, #RD9 for DOXY, #RD10 for pH.
SCIENTIFIC_CALIB_DATE	YYYYMMDDHHMISS(*)

(*): for a given calibration, the SCIENTIFIC_CALIB_DATE of an adjusted B-Argo parameter and of its associated IB-Argo parameters should be identical.

The three fields SCIENTIFIC_CALIB_COMMENT, _EQUATION, and _COEFFICIENT have netCDF dimensions (N_PROF, N_CALIB, N_PARAM, STRING256). This means that for each N_CALIB, each field is a 256-length character string. If character strings longer than 256-length are needed, the procedure should be separated and stored as multiple N_CALIB.

For a calibration that needs multiple N_CALIB:

- the SCIENTIFIC_CALIB_DATE should be identical for all N_CALIB,
- once the different fields are correctly filled, the remaining empty fields (unused) should be filled as follows:
  - ✓ SCIENTIFIC_CALIB_COMMENT: ‘No additional comment’,
  - ✓ SCIENTIFIC_CALIB_EQUATION: ‘No additional equation’,
  - ✓ SCIENTIFIC_CALIB_COEFFICIENT:’No additional coefficient’.

5.2.3. Other variables in a BD profile file

Here are other variables in a BD profile file that need to be updated after delayed-mode qc.

- The variable DATA_STATE_INDICATOR should record '2C' or '2C+'.

- The variable DATE_UPDATE should record the date of last update of the netCDF file, in the format YYYYMMDDHHMISS.

- In both the core- and b- profile files, the variable DATA_MODE(N_PROF) is not related to a specific parameter. The value of DATA_MODE(N_PROF) is set to ‘D’ when delayed-mode adjusted values for one or more <PARAM> in each N_PROF become available. In b-Argo profile files, biogeochemical parameters can receive delayed-mode adjustments at different times. Therefore the variable PARAMETER_DATA_MODE(N_PROF, N_PARAM) is added to b-Argo profile files to indicate the data mode of each <PARAM> in each N_PROF.

The adjusted section (<PARAM>_ADJUSTED, <PARAM>_ADJUSTED_QC and <PARAM>_ADJUSTED_ERROR) for each <PARAM> in each N_PROF should then be filled independently according to its PARAMETER_DATA_MODE.

For example, in a b-Argo profile file with DOXY and NITRATE, it is possible that

PARAMETER_DATA_MODE = 'D' for DOXY, and

PARAMETER_DATA_MODE = 'R' for NITRATE.

In this case:

-the adjusted section for DOXY should be filled with delayed-mode adjusted values;

-the adjusted section for NITRATE should be filled with FillValues.

- A history record should be appended to the HISTORY section of the netCDF file to indicate that the netCDF file has been through the delayed-mode process. Please refer to the Argo User's Manual (§5 "Using the History section of the Argo netCDF Structure") on usage of the History section.

5.2.4. B-profile files naming convention

When one or more <PARAM> in a single-cycle B- profile file receive delayed-mode adjusted values, the file name changes from BR<WMO_ID>_xxx.nc to BD<WMO_ID>_xxx.nc.

5.3. Recovering BGC data when float salinity is bad

5.3.1. Introduction

When float salinity is bad, some BGC data can be recovered by using a salinity substitute. It is generally agreed that recovering BGC data in this manner is of value to the scientific community. The ADMT-22 meeting in December 2021 supported this action. This section intends to clarify some of the details for implementation.

5.3.2. In what situations should this be done?

This exercise should only be done when the CTD fails much earlier than the other BGC sensors on a float, thus compromising BGC data from multiple cycles.

This exercise should not be done when there are only occasionally bad salinity profiles, or spikes. For these cases, please follow the procedures of how to propagate PSAL_QC = '4' to BGC data QC flags, which are described in the specific manuals for each BGC parameter. (e.g. For DOXY, when PSAL_QC = '4', then DOXY_QC = '3'. See #RD9.)

This exercise should only be done after a float is dead and after the CTD delayed-mode operator has issued D- files for the entire float record.

The recovered BGC values should only be used to fill <B-PARAM>_ADJUSTED in 'D' mode in the BD- files. The raw values in <B-PARAM> should continue to be computed from the raw PSAL.

This exercise should not be used in any real-time adjustment setting to fill <B-PARAM>_ADJUSTED in 'A' mode in the BR- files.

This exercise should be considered as optional. Priority should continue to be on the good data. Recovery work should only be done when extra resources are available.

5.3.3. When to give up on float salinity and switch to a salinity subsitute?

In general, when float salinities from entire profiles are declared bad and unadjustable by the core DMQC operator over multiple cycles, the BGC DMQC operator can switch to a salinity substitute.

In some cases, some bad float salinities may still be usable for BGC data. If interactions between the core and BGC DMQC opeators are possible, they can work together to evaluate the special cases for BGC use.

5.3.4. What to use as a salinity substitute?

Argo does not endorse the use of any particular product as the salinity substitute. The decision on which product to use lies with the BGC DMQC operator. However, care should be taken to ensure that the product of choice has adequate spatial and temporal resolutions and is sufficiently accurate in its salinity estimate. An example of a suitable salinity substitute is ARMOR3D, which provides global data sets of temperature and salinity on a 1/4° horizontal regular grid at 50 depth levels from 0-5500m, with weekly temporal resolution. See https://catalogue.marine.copernicus.eu/documents/QUID/CMEMS-MOB-QUID-015-012.pdf

5.3.5. How to fill _ADJUSTED_QC and _ERROR in the BD- files?

All <B-PARAM>_ADJUSTED_QC rules remain the same, but <B-PARAM>_ADJUSTED_QC = '2' should be recommended in place of values that normally would have '1'.

<B-PARAM>_ADJUSTED_ERROR can be increased if deemed necessary by the BGC delayed-mode operator. For most cases, any error imposed on <B-PARAM>_ADJUSTED by using a salinity substitute will be within the stated parameter uncertainty. However, certain data products that are used as the salinity substitute may exhibit enhanced uncertainty in specific regions (e.g. near the coast or within marginal seas).

5.3.6. How to fill SCIENTIFIC_CALIB_COMMENT

SCIENTIFIC_CALIB_COMMENT must record the salinity substitute that is used.

It may also be useful to have a fixed char string at the beginning to distinguish these as recovered profiles. For example:

'**Recovered BGC profile.** Suboptimal float salinity used.'
'**Recovered BGC profile.** ARMOR3D product salinity used. DOI???'

Note that each SCIENTIFIC_CALIB_COMMENT is only a 256-length char string. To ensure there is sufficient space to record all the calibration information, you may want to keep the 'Recovered BGC profile ..." record at N_CALIB = 1, then other calibration information at N_CALIB > 1.

5.3.7. Should the salinity substitute be stored at the Argo GDACs?

The consensus is that the salinity substitute ***should not*** be stored in any Argo data files at the Argo GDACs.

There was a suggestion to store the salinity substitute in a supplemental file in the /aux directory at the Argo GDACs. This suggestion was rejected, because none of the reasons for doing so was valid.

- The /aux directory is used to store auxiliary data from Argo floats that do not fit in any of the Argo data files. But they are still data from Argo floats, not data from external products.
- There are proprietary and endorsement issues about storing data from external products at the Argo GDACs, even if they are in the /aux directory. Doing so will create a dangerous situation where people use these non-Argo data inappropriately, or misundertand them as Argo data.
- It is acceptable to use a salinity substitute in <B-PARAM>_ADJUSTED in 'D' mode, since these values are already heavily influenced by model data. However, we do not store other BGC calibration information in the /aux directory, such as outputs from CANYONB or LIR, so there is no reason why salinity substitutes should be an exception.
- There is the argument that without salinity, there is limited use for the recovered BGC data. Yes, that is the reality. Argo has had many instrumentation problems in the past that have led to significant data loss, and will continue to have these problems in the future, but at no time has there been a decision to pull in external products to fill those gaps. Users need to find their own salinity substitute that is suitable for their own application.

Therefore the decision is that salinity substitutes ***should not*** be stored anywhere at the Argo GDACs, including the /aux directory. Instead, BGC DMQC operators should manage the calibration information and data processing steps internally within their own workspace, and then record the details in SCIENTIFIC_CALIB_COMMENT.

5.3.8. Examples

Some DACs are now in the process of implementing this procedure for select floats. A more detailed description of select example(s) will be added to this section in the near future.

6. References

Bittig HC, Maurer TL, Plant JN, Schmechtig C, Wong APS, Claustre H, Trull TW, Udaya Bhaskar TVS, Boss E, Dall'Olmo G, Organelli E, Poteau A, Johnson KS, Hanstein C, Leymarie E, Le Reste S, Riser SC, Rupan AR, Taillandier V, Thierry V and Xing X (2019) A BGC-Argo Guide: Planning, Deployment, Data Handling and Usage. *Front. Mar. Sci.* 6:502. doi: 10.3389/fmars.2019.00502

Cossarini G., Mariotti L., Feudale L., Mignot A., Salon S., Taillandier V., Teruzzi A., D'Ortenzio F. Towards operational 3D-var assimilation of chlorophyll biogeochemical-argo float data into a biogeochemical model of the Mediterranean sea Ocean Model., 133 (2019), pp. 112-128, 10.1016/j.ocemod.2018.11.005

Fennel K, Gehlen M, Brasseur P, Brown CW, Ciavatta S, Cossarini G, Crise A, Edwards CA, Ford D, Friedrichs MAM, Gregoire M, Jones E, Kim H-C, Lamouroux J, Murtugudde R, Perruche C and the GODAE OceanView Marine Ecosystem Analysis and Prediction Task Team (2019) Advancing Marine Biogeochemical and Ecosystem Reanalyses and Forecasts as Tools for Monitoring and Managing Ecosystem Health. *Front. Mar. Sci.* 6:89. doi: 10.3389/fmars.2019.00089

Ford, David (2021) Assimilating synthetic Biogeochemical-Argo and ocean colour observations into a global ocean model to inform observing system design, Biogeosciences, 10.5194/bg-18-509-2021, **18**, 2, (509-534).

Johnson GC, Hosoda S, Jayne SR, Oke PR, Riser SC, Roemmich D, Suga T, Thierry V, Wijffels SE, Xu J. Argo-Two Decades: Global Oceanography, Revolutionized. Ann Rev Mar Sci. 2022 Jan 3;14:379-403. doi: 10.1146/annurev-marine-022521-102008. Epub 2021 Jun 8. PMID: 34102064.

Penny, S.G., Akella, S., Buehner, M., Chevallier, M., Counillon, F., Draper, C., Frolov, S., Fujii, Y., Karspeck, A., Kumar, A., Laloyaux, P., Mahfouf, J.-F., Martin, M.J., Peña, M., de Rosnay, P., Subramanian, A., Tardif, R., Wang, Y. and Wu, X. (2017). Coupled data assimilation for integrated earth system analysis and prediction: goals, challenges, and recommendations. Report WWRP 2017-3, World Weather Research Programmme, WMO, Geneva, Switzerland.

Shu C, Xiu P, Xing X, Qiu G, Ma W, Brewin RJW, Ciavatta S. (2022). Biogeochemical Model Optimization by Using Satellite-Derived Phytoplankton Functional Type Data and BGC-Argo Observations in the Northern South China Sea. *Remote Sensing*. 14(5):1297. https://doi.org/10.3390/rs14051297

Verdy, A., & Mazloff, M. R. (2017). A data assimilating model for estimating Southern Ocean biogeochemistry. *JGR:Oceans*, **122**(9), 69686988– 6988. https://doi.org/10.1002/2016JC012650