argo data management

Argo data management
January 3rd 2013

# Argo quality control manual
Version 2.8

1

ARGO
part of the integrated global observation strategy

# Table of contents

# History

| Date | Comment |
|---|---|
| 01/01/2002 | Creation of the document. |
| 28/03/2003 | Changed lower limit of temperature in Med Sea to 10.0 degrees C. |
| 08/06/2004 | Modified spike and gradient tests according to advice from Yasushi. Added inversion test. |
| 24/10/2003 | Real-time qc tests 15 and 16 proposed at Monterey data management meeting. Test 10 removed. |
| 07/10/2004 | 1. Real-time and delayed-mode manuals merged in "Argo quality control manual". 2. Frozen profile real time qc test 17 proposed at ADMT5 in Southampton. 3. Deepest pressure real time qc test 18 proposed at ADMT5 in Southampton. 4. Order list for the real time qc tests. 5. "Regional Global Parameter Test" renamed "Regional range test", test 7. 6. Grey list naming convention and format, test 15. 7. Real time qc on trajectories. |
| 23/11/2004 | §1: new introduction from Annie Wong. §4: delayed mode quality control manual from Annie Wong. |
| 26/11/2004 | §3: update of summary flow chart for salinity delayed mode qc. |
| 31/08/2005 | §2.2: update on test 17, visual qc. |
| 17/11/2005 | §2.3: added a section on Real-time salinity adjustment. §3.1: added usage of SURFACE PRESSURE from APEX floats. §3.3.5: added some more guidelines for PSAL_ADJUSTED_QC = '2'. §3.3.8: clarified that PROFILE_<PARAM>_QC should be recomputed when <PARAM>_ADJUSTED_QC becomes available. |
| 16/10/2006 | §3: updated Delayed-mode section based on DMQC-2 Workshop. |
| 20/11/2006 | §2.1.2: test 19: deepest pressure delta set to 10%. §2.1.2: test 14: density inversion test applied downward and upward. |
| 14/11/2007 | §2.1.2: test 6: minimum salinity set to 2 PSU instead of 0 PSU. §2.1.2: test 7: minimum salinity set to 2 PSU instead of 0 PSU. This change was decided during ADMT8 in Hobart. |
| 14/11/2007 | §3.3.1: use "known pressure drift" instead of delta P > 5dbar. This change was decided during AST8 meeting in Paris. |
| 14/11/2007 | §3.3.2: delayed-mode operators can edit real-time QC flags. This change was decided during ADMT8 in Hobart. |
| 14/11/2007 | §1.2.4: values with QC flag = '4' are ignored by quality control tests. This change was decided during ADMT8 in Hobart. |
| 14/11/2007 | §2.1.4: when salinity is calculated from the conductivity parameter, if temperature is flagged as bad then salinity is flagged as bad. This change was decided during ADMT8 in Hobart. |
| 21/01/2008 | §2.2: test 6: minimum salinity set to 2 PSU instead of 0 PSU. §2.2: test 7: minimum salinity set to 2 PSU instead of 0 PSU. |
| 04/11/2008 | §2.1.1: "Sigma0" specified in density inversion test. §2.3.1: added a section on Real-time pressure adjustment for APEX floats. §3.3.5: updated Delayed-mode section based on DMQC-3 Workshop. |
| 14/2/2009 | §3.1.1: added a section on Delayed-mode pressure adjustment for APEX floats. |
| 19/10/2009 | §2.1.2 & §2.2: test 6: minimum P set to −5 dbar. §3.1: added a section on Editing raw qc flags in delayed-mode. §3.2.2: updated delayed-mode treatment for APEX TNPDs, after DMQC-4. |
| 15/7/2010 | §2.4: added a section on Feedback from Statistical Test at Coriolis. |
| 05/11/2010 | §3.2.2: revised definition for TNPD after ADMT11 in Hamburg. |

| Date | Comment |
|---|---|
| 01/12/2011 | §2.1.2: added threshold of 0.03 kg m$^{-3}$ to RT Test 14, Density Inversion Test, for profile data, following ADMT12 in Seoul. Test to use potential density referenced to mid-point pressure between the two levels to be compared.<br>§2.2: Added RT Test 20, "Questionable Argos position test" from JAMSTEC, as a new real-time qc test for trajectory data, following 3$^{rd}$ Trajectory Workshop and ADMT12 in Seoul. |
| 03/01/2013 | §2.3.1: clarify real-time pressure adjustment when SP=0 |
| 03/01/2013 | §2.1.2: addition of real-time quality control for dissolved oxygen DOXY.<br>The real-time QC tests 6, 9, 11 and 13 are applied to DOXY. |

# Authors

Annie Wong, Robert Keeley, Thierry Carval and the Argo Data Management Team.

# 1.   Introduction

This document is the Argo quality control manual.

==The changes with the previous version of the manual are highlighted in yellow.==

The Argo data system has three levels of quality control.

- The first level is the real-time system that performs a set of agreed automatic checks on all float measurements. Real-time data with assigned quality flags are available to users within the 24-48 hrs timeframe.

- The second level of quality control is the delayed-mode system.

- The third level of quality control is regional scientific analyses of all float data with other available data. The procedures for regional analyses are still to be determined.

This document contains the description of the Argo real-time and delayed-mode procedures.

Please note that at the present time, quality control procedures exist only for the parameters JULD, LATITUDE, LONGITUDE, PRES, TEMP, PSAL ==and DOXY==.

# 2. Real-time quality controls

## 2.1. Argo Real-time Quality Control Test Procedures on vertical profiles

### 2.1.1. Introduction

Because of the requirement for delivering data to users within 24 hours of the float reaching the surface, the quality control procedures on the real-time data are limited and automatic. The test limits are briefly described here. More detail on the tests can be found in IOC Manuals and Guides #22 or at

http://www.meds-sdmm.dfo-mpo.gc.ca/ALPHAPRO/gtspp/qcmans/MG22/guide22_e.htm

Note that some of the test limits used here and the resulting flags are different from what is described in IOC Manuals and Guides #22.

If data from a float fail these tests, those data will not be distributed on the GTS. However, all of the data, including those having failed the tests, should be converted to the appropriate netCDF format and forwarded to the Global Argo Servers.

Presently, the TESAC code form is used to send the float data on the GTS (see http://www.meds-sdmm.dfo-mpo.gc.ca/meds/Prog_Int/J-COMM/J-COMM_e.htm). This code form only handles profile data and reports observations as a function of depth not pressure. It is recommended that the UNESCO routines be used to convert pressure to depth (Algorithms for computation of fundamental properties of seawater, N.P. Fofonoff and R.C. Millard Jr., UNESCO Technical Papers in Marine Science #44, 1983). If the position of a profile is deemed wrong, or the date is deemed wrong, or the platform identification is in error then none of the data should be sent on the GTS. For other failures, only the offending values need be removed from the TESAC message. The appropriate actions to take are noted with each test.

## 2.1.2. Quality control tests

**1. Platform identification**

Every centre handling float data and posting them to the GTS will need to prepare a metadata file for each float and in this is the WMO number that corresponds to each float ptt. There is no reason why, except because of a mistake, an unknown float ID should appear on the GTS.

Action: If the correspondence between the float ptt cannot be matched to the correct WMO number, none of the data from the profile should be distributed on the GTS.

**2. Impossible date test**

The test requires that the observation date and time from the float be sensible.

- Year greater than 1997

- Month in range 1 to 12

- Day in range expected for month

- Hour in range 0 to 23

- Minute in range 0 to 59

Action: If any one of the conditions is failed, the date should be flagged as bad data and none of the data from the profile should be distributed on the GTS.

**3. Impossible location test**

The test requires that the observation latitude and longitude from the float be sensible.

Action: If either latitude or longitude fails, the position should be flagged as bad data and none of the data from the float should go out on the GTS.

- Latitude in range −90 to 90

- Longitude in range −180 to 180

**4. Position on land test**

The test requires that the observation latitude and longitude from the float be located in an ocean.

Use can be made of any file that allows an automatic test to see if data are located on land. We suggest use of at least the 5-minute bathymetry file that is generally available. This is commonly called ETOPO5 / TerrainBase and can be downloaded from http://www.ngdc.noaa.gov/mgg/global/global.html.

Action: If the data cannot be located in an ocean, the position should be flagged as bad data and they should not be distributed on the GTS.

## 5. Impossible speed test

Drift speeds for floats can be generated given the positions and times of the floats when they are at the surface and between profiles. In all cases we would not expect the drift speed to exceed 3 m s$^{-1}$. If it does, it means either a position or time is bad data, or a float is mislabeled. Using the multiple positions that are normally available for a float while at the surface, it is often possible to isolate the one position or time that is in error.

Action: If an acceptable position and time can be used from the available suite, then the data can be sent to the GTS. Otherwise, flag the position, the time, or both as bad data and no data should be sent to the GTS.

## 6. Global range test

This test applies a gross filter on observed values for pressure, temperature, salinity and dissolved oxygen. It needs to accommodate all of the expected extremes encountered in the oceans.

- Pressure cannot be less than −5 dbar

- Temperature in range −2.5 to 40.0°C

- Salinity in range 2 to 41.0 PSU

- Dissolved oxygen in range −0.5 to 600 micromol/kg

Action: If a value fails, it should be flagged as bad data and only that value need be removed from distribution on the GTS. If temperature and salinity values at the same depth both fail, both values should be flagged as bad data and values for depth, temperature and salinity should be removed from the TESAC distributed on the GTS.

## 7. Regional range test

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

Action: Individual values that fail these ranges should be flagged as bad data and removed from the TESAC being distributed on the GTS. If both temperature and salinity values at the same depth both fail, then values for depth, temperature and salinity should be removed from the TESAC being distributed on the GTS.

Red Sea

- Temperature in range 21.7 to 40.0°C

- Salinity in range 2 to 41.0 PSU

Mediterranean Sea

- Temperature in range 10.0 to 40.0°C

- Salinity in range 2 to 40.0 PSU

## 8. Pressure increasing test

This test requires that the profile has pressures that are monotonically increasing (assuming the pressures are ordered from smallest to largest).

Action: If there is a region of constant pressure, all but the first of a consecutive set of constant pressures should be flagged as bad data. If there is a region where pressure reverses, all of the pressures in the reversed part of the profile should be flagged as bad data. All pressures flagged as bad data and all of the associated temperatures and salinities are removed from the TESAC distributed on the GTS.

## 9. Spike test

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 temperature, salinity and dissolved oxygen with depth. The algorithm is used on vertical profiles of temperature, salinity and dissolved oxygen.

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.

Temperature: The V2 value is flagged when

- the test value exceeds  6.0°C for pressures less than 500 dbar, or
- the test value exceeds  2.0°C for pressures greater than or equal to 500 dbar.

Salinity: The V2 value is flagged when

- the test value exceeds 0.9 PSU for pressures less than 500 dbar, or
- the test value exceeds 0.3 PSU for pressures greater than or equal to 500 dbar.

Dissolved oxygen: 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.

Action: Values that fail the spike test should be flagged as bad data and are removed from the TESAC distributed on the GTS. If temperature and salinity values at the same depth both fail, they should be flagged as bad data and the values for depth, temperature and salinity should be removed from the TESAC being distributed on the GTS.

## 10. Top and bottom spike test: obsolete

## 11. Gradient test

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 temperature, salinity and dissolved oxygen with depth. The algorithm is used on vertical profiles of temperature, salinity and dissolved oxygen.

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.

Temperature: The V2 value is flagged when

- the test value exceeds 9.0°C for pressures less than 500 dbar, or

- the test value exceeds 3.0°C for pressures greater than or equal to 500 dbar.

Salinity: The V2 value is flagged when

- the test value exceeds 1.5 PSU for pressures less than 500 dbar, or

- the test value exceeds 0.5 PSU for pressures greater than or equal to 500 dbar.

Dissolved oxygen: 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.

Action: Values that fail the test (i.e. value V2) should be flagged as bad data and are removed from the TESAC distributed on the GTS. If temperature and salinity values at the same depth both fail, both should be flagged as bad data and then values for depth, temperature and salinity should be removed from the TESAC distributed on the GTS.

## 12. Digit rollover test

Only so many bits are allowed to store temperature and salinity 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.

- Temperature difference between adjacent depths > 10°C
- Salinity difference between adjacent depths > 5 PSU

Action: Values that fail the test should be flagged as bad data and are removed from the TESAC distributed on the GTS. If temperature and salinity values at the same depth both fail, both values should be flagged as bad data and then values for depth, temperature and salinity should be removed from the TESAC distributed on the GTS.

## 13. Stuck value test

This test looks for all measurements of temperature, salinity and dissolved oxygen in a vertical profile being identical.

Action: If this occurs, all of the values of the affected variable should be flagged as bad data and are removed from the TESAC distributed on the GTS. If temperature and salinity are affected, all observed values are flagged as bad data and no report from this float should be sent to the GTS.

## 14. Density inversion

This test compares potential density between valid measurements in a profile, in both directions, i.e. from top to bottom, and from bottom to top. Values of temperature and salinity at the same pressure level $P_i$ should be used to compute potential density $\rho_i$ (or $\sigma_i = \rho_i - 1000$) kg m$^{-3}$, referenced to the mid-point between $P_i$ and the next valid pressure level $P_{i+1}$. A threshold of 0.03 kg m$^{-3}$ should be allowed for small density inversions.

Action: From top to bottom, if the potential density calculated at the greater pressure is less than that calculated at the lesser pressure by more than 0.03 kg m$^{-3}$, both the temperature and salinity values should be flagged as bad data. From bottom to top, if the potential density calculated at the lesser pressure is greater than that calculated at the greater pressure by more than 0.03 kg m$^{-3}$, both the temperature and salinity values should be flagged as bad data. Bad temperature and salinity values should be removed from the TESAC distributed on the GTS.


## 15. Grey list

This test is implemented to stop the real-time dissemination on the GTS of measurements from a sensor that is not working correctly.

The grey list contains the following 7 items:

- Float WMO Id

- Parameter: name of the grey listed parameter

- Start date: from that date, all measurements for this parameter are flagged as bad or probably bad

- End date: from that date, measurements are not flagged as bad or probably bad

- Flag: value of the flag to be applied to all measurements of the parameter

- Comment: comment from the PI on the problem

- DAC: data assembly center for this float

Example:

| Float WMO Id | Parameter | Start date | End date | Flag | Comment | DAC |
|---|---|---|---|---|---|---|
| 1900206 | PSAL | 20030925 | | 3 | | IF |


Each DAC manages a grey list, sent to the GDACs.

The merged grey list is available from the GDACs.

- Grey list format: ascii csv (comma separated values)

- Naming convention: xxx_greylist.csv
  xxx: DAC name (e.g.: aoml_greylist.csv, coriolis_greylist.csv, jma_greylist.csv)

- PLATFORM, PARAMETER, START_DATE, END_DATE, QC, COMMENT, DAC
  e.g. 4900228, TEMP, 20030909, , 3, , AO
  e.g  1900206, PSAL, 20030925, , 3, , IF

The decision to insert a float parameter in the grey list comes from the PI or the delayed-mode operator. A float parameter should be put in the grey list when sensor drift is too big to be corrected adequately in real-time, or when the sensor is judged to be not working correctly.

The grey list only concerns real-time files (R-files). When an anomalous float is dead and has been adjusted in delayed-mode, it should not appear in the grey list. When an anomalous float is active and has been partially adjusted in delayed-mode, it should remain in the grey list only if real-time adjustment is not adequate.

## 16. Gross salinity or temperature sensor drift

This test is implemented to detect a sudden and significant sensor drift.

It calculates the average salinity on the last 100 dbar on a profile and the previous good profile. Only measurements with good QC are used.

Action: if the difference between the two average values is more than 0.5 PSU then all measurements for this parameter are flagged as probably bad data (flag '3').

The same test is applied for temperature: if the difference between the two average values is more than 1°C then all measurements for this parameter are flagged as probably bad data (flag '3').

## 17. Visual QC

Subjective visual inspection of float values by an operator.

To avoid delays, this test is not mandatory before real-time distribution.

## 18. Frozen profile test

This test can 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 0.001 PSU for salinity and of the order of 0.01°C for temperature.

A). Derive temperature and salinity profiles by averaging the original profiles to get mean values for each profile in 50 dbar slabs (Tprof, T_previous_prof and Sprof, S_previous_prof). This is necessary because the floats do not sample at the same level for each profile.

B). Substract the two resulting profiles for temperature and salinity to get absolute difference profiles:

- deltaT = abs(Tprof − T_previous_prof)

- deltaS = abs(Sprof − S_previous_prof)

C). Derive the maximum, minimum and mean of the absolute differences for temperature and salinity:

- mean(deltaT), max(deltaT), min(deltaT)

- mean(deltaS), max(deltaS), min(deltaS)

D). To fail the test, require that:

- max(deltaT) <  0.3
- min(deltaT) < 0.001
- mean(deltaT) < 0.02
- max(deltaS) < 0.3
- min(deltaS) < 0.001
- mean(deltaS) < 0.004

Action: if a profile fails this test, all measurements for this profile are flagged as bad data (flag '4'). If the float fails the test on 5 consecutive cycles, it is inserted in the grey list.


## 19. Deepest pressure test

This test requires that the profile has pressures that are not higher than DEEPEST_PRESSURE plus 10%.

DEEPEST_PRESSURE value comes from the meta-data file of the float.

Action: If there is a region of incorrect pressures, all pressures and corresponding measurements should be flagged as bad data (flag '4'). All pressures flagged as bad data and all of the associated temperatures and salinities are removed from the TESAC distributed on the GTS.

## 2.1.3. Tests application order

The Argo real time QC tests are applied in the order described in the following table.

| Order | Test Number | Test Name |
|---|---|---|
| 1 | 19 | Deepest Pressure Test |
| 2 | 1 | Platform Identification |
| 3 | 2 | Impossible Date Test |
| 4 | 3 | Impossible Location Test |
| 5 | 4 | Position on Land Test |
| 6 | 5 | Impossible Speed Test |
| 7 | 6 | Global Range Test |
| 8 | 7 | Regional Range Test |
| 9 | 8 | Pressure Increasing Test |
| 10 | 9 | Spike Test |
| 11 | 10 | Top and Bottom Spike Test : removed |
| 12 | 11 | Gradient Test |
| 13 | 12 | Digit Rollover Test |
| 14 | 13 | Stuck Value Test |
| 15 | 14 | Density Inversion |
| 16 | 15 | Grey List |
| 17 | 16 | Gross salinity or temperature sensor drift |
| 18 | 18 | Frozen profile |
| 19 | 17 | Visual QC |

## 2.1.4. 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 Test 11 (gradient test) cannot be decreased to QC flag '3' (bad data that are potentially correctable) set by Test 15 (grey list).

A value with QC flag '4' (bad data) or '3' (bad data that are potentially correctable) is ignored by the quality control tests.

For floats where salinity (PSAL) is calculated from the temperature (TEMP) and conductivity (CNDC) parameters, if temperature is flagged '4' (or '3'), then salinity is flagged '4' (or '3').

## 2.2.  Argo Real-time Quality Control Test Procedures on trajectories

The following tests are applied in real-time on trajectory data.

### 1. Platform identification

Every centre handling float data and posting them to the GTS will need to prepare a metadata file for each float and in this is the WMO number that corresponds to each float ptt. There is no reason why, except because of a mistake, an unknown float ID should appear on the GTS.

Action: If the correspondence between the float ptt cannot be matched to the correct WMO number, none of the data from the profile should be distributed on the GTS.

### 2. Impossible date test

The test requires that the observation date and time from the float be sensible.

- Year greater than 1997
- Month in range 1 to 12
- Day in range expected for month
- Hour in range 0 to 23
- Minute in range 0 to 59

Action: If any one of the conditions is failed, the date should be flagged as bad data and none of the data from the profile should be distributed on the GTS.

### 3. Impossible location test

The test requires that the observation latitude and longitude from the float be sensible.

Action: If either latitude or longitude fails, the position should be flagged as bad data and none of the data from the float should go out on the GTS.

- Latitude in range −90 to 90
- Longitude in range −180 to 180

### 4. Position on land test

The test requires that the observation latitude and longitude from the float be located in an ocean.

Use can be made of any file that allows an automatic test to see if data are located on land. We suggest use of at least the 5-minute bathymetry file that is generally available. This is commonly called ETOPO5 / TerrainBase and can be downloaded from http://www.ngdc.noaa.gov/mgg/global/global.html

Action: If the data cannot be located in an ocean, the position should be flagged as bad data and they should not be distributed on the GTS.

5. Impossible speed test

Drift speeds for floats can be generated given the positions and times of the floats when they are at the surface and between profiles. In all cases we would not expect the drift speed to exceed 3 m s$^{-1}$. If it does, it means either a position or time is bad data, or a float is mislabeled. Using the multiple positions that are normally available for a float while at the surface, it is often possible to isolate the one position or time that is in error.

Action: If an acceptable position and time can be used from the available suite, then the data can be sent to the GTS. Otherwise, flag the position, the time, or both as bad data and no data should be sent to the GTS.

(Note: Floats that use the Argos system to obtain position data can use Test 20 instead.)


6. Global range test

This test applies a gross filter on observed values for pressure, temperature and salinity. It needs to accommodate all of the expected extremes encountered in the oceans.

- Pressure cannot be less than −5 dbar

- Temperature in range −2.5 to 40.0°C

- Salinity in range 2 to 41.0 PSU

Action: If a value fails, it should be flagged as bad data and only that value need be removed from distribution on the GTS. If temperature and salinity values at the same depth both fail, both values should be flagged as bad data and values for depth, temperature and salinity should be removed from the TESAC distributed on the GTS.


7. Regional range test

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

Action: Individual values that fail these ranges should be flagged as bad data and removed from the TESAC being distributed on the GTS. If both temperature and salinity values at the same depth both fail, then values for depth, temperature and salinity should be removed from the TESAC being distributed on the GTS.

Red Sea

- Temperature in range 21.7 to 40.0°C

- Salinity in range 2 to 41.0 PSU

Mediterranean Sea

- Temperature in range 10.0 to 40.0°C

- Salinity in range 2 to 40.0 PSU

### 20. Questionable Argos position test

For floats that use the Argos system to obtain position data, this test can be used in lieu of Test 5 (Impossible Speed Test).

This test identifies questionable Argos position data collected during surface drift of a float cycle by considering the float speed at the sea surface and Argos position errors. Details of the method can be found in Nakamura et al (2008), "Quality control method of Argo float position data", JAMSTEC Report of Research and Development, Vol. 7, 11-18. A brief description of the procedure is summarized here.

A). Collect all Argos positions during surface drift of a float cycle. The distance between two positions $A$ and $B$ is referred to as a segment. A segment is considered questionable

- if  (i) the float speed along the segment exceeds 3 m s$^{-1}$, **and**

- (ii) the length of the segment is longer than the critical error length, defined as

$$1.0 \times \sqrt{Er_A{}^2 + Er_B{}^2} \ ,$$

where $Er_A{}^2$ and $Er_B{}^2$ are the radii of position error of the Argos system (150m, 350m, and 1000m for Argos class 3, 2, and 1 respectively) at $A$ and $B$ respectively.

B). If a segment is not considered questionable, then both positions $A$ and $B$ are good.

C). If a segment is considered questionable, then:

- if the Argos class at $A$ and $B$ are different, then the position with the less accurate Argos class is flagged as '3';

- if the Argos class at $A$ and $B$ are the same, and there is one good position before and one good position after $A$ and $B$ (i.e. there are 4 positions for the check), then the position that gives the higher speed along the segment from the previous good position to the later good position is flagged as '3';

- if the Argos class at $A$ and $B$ are the same, and there is one good position either before or after $A$ and $B$ (i.e. there are 3 positions for the check), then the position that gives the higher speed along the segment either from the previous good position or to the later good position is flagged as '3';

- if the Argos class at $A$ and $B$ are the same, but there are no other good positions around $A$ and $B$ (i.e. there are 2 positions for the check), then both $A$ and $B$ are flagged as '3'.

## 2.3. Argo Real-time Adjustments on vertical profiles

### 2.3.1. Real-time pressure adjustment for non auto-correcting floats

There are many float types in Argo, each has its own way of treating the pressure measurements. For example, while PROVOR and SOLO floats internally correct for pressure offsets, APEX floats do not make any internal pressure correction. All Argo float types and their treatment of pressure are summarized in the "Surface Pressure Correction Table", which can be found at:

http://www.argodatamgt.org/Documentation

Floats that do not adjust their pressure data on board before data telemetry are referred to as *non auto-correcting* floats. Pressure measurements from non auto-correcting floats need to be treated in real-time whenever valid surface pressure values are available. The section below details the real-time pressure adjustment procedures for APEX floats, which is the largest group of non auto-correcting floats in Argo.

**Real-time pressure adjustment for APEX floats**

APEX floats return "raw" pressures, which are stored in the variable PRES in the Argo netCDF files. Pressure adjustment should be applied in real-time to all APEX floats by using SURFACE PRESSURE (SP) values returned by the APEX floats. The SP measurement is taken while the float is at the sea surface just before descent, and hence is different from the shallowest measured pressure in the vertical profile, which is taken on ascent and while the float is beneath the sea surface. These SP values are stored in the Argo technical files in the variable PRES_SurfaceOffsetTruncatedPlus5dbar_dBAR or PRES_SurfaceOffsetNotTruncated _dBAR, depending on the type of APEX controller used.

Subtract 5 dbar from the SP values in PRES_SurfaceOffsetTruncatedPlus5dbar _dBAR. SP values in PRES_SurfaceOffsetNotTruncated_dBAR are used as they are, without needing to subtract 5 dbar.

Then erroneous outliers in SP need to be removed. This is done in real-time in two steps:

(1). Discard SP values greater than 20 dbar or less than −20 dbar, then revert to the last valid SP.

(2). If the most recent SP value, SP(i), is different from the last valid SP by more than 5 dbar, that is, if abs[ SP(i) – last valid SP ] > 5 dbar, revert to the last valid SP.

When no valid SP value is available, no real-time pressure adjustment is available.

When there are valid SP values, real-time adjusted pressures will be recorded in the variable PRES_ADJUSTED, where PRES_ADJUSTED = PRES – SP.

Zero is a valid SP value. Therefore when SP = 0, PRES_ADJUSTED should be filled. In this case, PRES_ADJUSTED = PRES.

PRES should always record the *raw* data.

PRES_ADJUSTED_QC will be filled with the same values as PRES_QC.

PRES_ADJUSTED_ERROR and all variables in the SCIENTIFIC CALIBRATION section of the netCDF files will be filled with FillValue.

DATA_MODE will record 'A'.

There is no need to re-calculate salinity data in real-time by using the real-time adjusted pressure values. This is because the difference in salinity due to real-time pressure adjustment is small. Pressure adjustment of less than 20 dbar will result in salinity error of less than 0.01.

When the SP value exceeds 20 dbar (or −20 dbar) for more than 5 consecutive cycles, the float should be put on the grey list because of pressure error, after consultation with the PI.

When available, real-time adjusted values are distributed to the GTS, instead of the original values.

## 2.3.2. Real-time salinity adjustment

When delayed-mode salinity adjustment (see Section 3.3) becomes available for a float, real-time data assembly centres will extract the adjustment from the latest D*.nc file as an additive constant, and apply it to new salinity profiles. (If a better correction is available in real-time, DACs can use that instead.) In this manner, intermediate-quality salinity profiles will be available to users in real-time.

The values of this real-time adjustment will be recorded in PSAL_ADJUSTED.

PSAL_ADJUSTED_QC will be filled with the same values as PSAL_QC.

PSAL_ADJUSTED_ERROR and all variables in the SCIENTIFIC CALIBRATION section of the netCDF files will be filled with FillValue.

DATA_MODE will record 'A'.

When available, real-time adjusted values are distributed to the GTS, instead of the original values.

## 2.3.3. Real-time files with DATA_MODE = 'A'

When real-time files have DATA_MODE = 'A', it means real-time adjustments are available for one or more parameters. All PARAM_ADJUSTED variables should therefore be filled, where PARAM = PRES, TEMP, PSAL, CNDC, DOXY, etc …

PARAM_ADJUSTED = real-time adjusted values, or PARAM if no real-time adjustment is available.

PARAM_ADJUSTED_QC = PARAM_QC.

PARAM_ADJUSTED_ERROR = FillValue.

Users should be aware that even though the _ADJUSTED_ fields are filled in 'A' mode, the adjustments are applied in an automated manner in real-time and are not checked by delayed-mode operators.

## 2.4. Feedback from Statistical Test at Coriolis

At Coriolis, an objective analysis is performed on a daily basis on Argo temperature and salinity profiles that have been quality-controlled during the previous 3 weeks. As results of the comparison with climatology, anomalies on Argo profiles are detected by this objective analysis. For flag correction on those profiles, daily automatic feedbacks (in text files, by email) are sent to the appropriate DAC. The email message contains the list of Argo profiles highlighted by the objective analysis, and examined by a Coriolis operator, with the recommended flag correction listed at the end. The information is also available in a csv format file on the ftp site:

- ftp://ftp.ifremer.fr/ifremer/argo/etc/ObjectiveAnalysisWarning

# 3. Delayed-mode quality controls

## 3.1. Editing raw qc flags in delayed-mode

Delayed-mode operators should examine profile data for pointwise errors such as spikes and jumps, and edit the raw qc flags in PARAM_QC when they are set incorrectly. PARAM here refers to PRES, TEMP, CNDC, and PSAL. Cases where PARAM_QC should be edited in delayed-mode include:

(a). PARAM_QC should be changed to '4' for bad and un-correctable data that are not detected by the real-time tests; and

(b). PARAM_QC should be changed to '1' or '2' for good data that are wrongly identified as bad or probably bad by the real-time tests.

## 3.2. Delayed-mode procedures for pressure

Delayed-mode qc for PRES is done by subjective assessment of vertical profile plots of TEMP vs. PRES, and PSAL vs. PRES. This assessment should be done in relation to measurements from the same float, as well as in relation to nearby floats and historical data. The assessment should aim to identify: (a) erroneous data points that cannot be detected by the real-time qc tests, and (b) vertical profiles that have the wrong shape.

Bad data points identified by visual inspection from delayed-mode analysts are recorded with PRES_ADJUSTED_QC = '4'. For these bad data points, TEMP_ADJUSTED_QC and PSAL_ADJUSTED_QC should also be set to '4'. Please note that whenever PARAM_ADJUSTED_QC = '4',

- PARAM_ADJUSTED = FillValue;
- PARAM_ADJUSTED_ERROR = FillValue.

### 3.2.1. Delayed-mode pressure adjustment for APEX floats

Similar to the real-time procedure, pressures from APEX floats should be adjusted for offsets by using SURFACE PRESSURE (SP) values in delayed-mode. SP values are stored in the Argo technical files in PRES_SurfaceOffsetNotTruncated_dBAR or PRES_SurfaceOffsetTruncatedPlus5dbar_dBAR, depending on the type of APEX controller used. The SP time series is examined and treated in delayed-mode as follows:

(1). Subtract 5 dbar from PRES_SurfaceOffsetTruncatedPlus5dbar_dBAR. SP values in PRES_SurfaceOffsetNotTruncated_dBAR are used as they are, without needing to subtract 5 dbar.

(2). **Despike the SP time series to 1 dbar.** This is most effectively done by first removing the more conspicuous spikes that are bigger than 5 dbar (as in the real-time procedure), then the more subtle spikes that are between 1 to 5 dbar by comparing the SP values with those derived from a 5-point median filter. For standard Argo floats that sample every 10 days, a 5-point filter represents a filter window of 40 days (+/− 20 days from a profile), which is an appropriate time scale for retaining effects from the atmospheric seasonal cycle.

(3). Replace the spikes and any other missing SP values by interpolating between good neighbouring points. If missing values occur at the ends of the SP time series, extrapolate from the nearest good points.

The resulting SP time series should then be inspected visually to make sure there are no more erroneous points. Then the clean SP value from cycle *i+1* is used to adjust CTD pressures from cycle *i* by

$$\text{PRES\_ADJUSTED (cycle } i) = \text{PRES (cycle } i) - \text{SP (cycle } i+1).$$

The CTD profile and the associated SP is staggered by one cycle because the SP measurement is taken after the telemetry period, and therefore is stored in the memory and telemetered during the next cycle. The real-time procedure does not match SP value from cycle *i+1* with PRES from cycle *i*, because real-time adjustment cannot wait 10 days. However, in delayed-mode, it is important to match the CTD profile with the staggered telemetry of SP, because SP values can contain synoptic atmospheric variations, and because a missing CTD profile is often associated with an erroneous SP point. By this scheme, SP(1), which is taken before cycle 1 and therefore before the float has had its first full dive, is not used in delayed-mode.

Note that the real-time procedure does not adjust for pressure offsets that are greater than 20 dbar (or less than −20 dbar). This is because the real-time automatic procedure cannot determine whether SP values greater than 20 dbar (or less than −20 dbar) represent genuine sensor drift or erroneous measurements. Instead, in real-time, floats that return SP values greater than 20 dbar (or less than −20 dbar) for more than 5 consecutive cycles are grey-listed in consultation with the PI. In delayed-mode, operators can inspect the SP time series visually when severe pressure sensor drift occurs. Therefore there is no upper limit to the magnitude of delayed-mode pressure adjustment.

After adjustment, delayed-mode operators should check that PRES_ADJUSTED > 0. If PRES_ADJUSTED < 0, delayed-mode operators should check for decoding errors in SP or in the CTD pressures.

PRES should always record the *raw* data.

PRES_ADJUSTED_QC should be set appropriately. For example, floats that have had significant pressure adjustment should have PRES_ADJUSTED_QC = '2'.

PRES_ADJUSTED_ERROR = 2.4 dbar is the recommended error to quote, with 2.4 dbar being the manufacturer quoted accuracy of the pressure sensor.

Salinity should be re-calculated by using PRES_ADJUSTED, and recorded in PSAL_ADJUSTED. Salinity error due to pressure uncertainty is negligible, and can be ignored in the consideration of PSAL_ADJUSTED_ERROR.

Please use the SCIENTIFIC CALIBRATION section in the netCDF files to record details of the delayed-mode adjustment.

**Note to users:** The 1 dbar despiking threshold for SP assumes that spikes greater than 1 dbar represent noise in the SP measurement that should not be integrated into float pressures. After despiking to 1 dbar, the remaining SP values contain sea surface atmospheric pressure variations and variations due to other high-frequency surface processes. While sea surface atmospheric pressure variations affect the whole water column and therefore should be adjusted for, high-frequency surface processes do not affect the whole water column. Therefore users should be aware that PRES_ADJUSTED contains noise from high-frequency surface processes that are of the order < 1 dbar. In addition, other more subtle pressure errors such as those due to non-linear hysteresis and

other temperature- and pressure-dependent effects are not accounted for in PRES_ADJUSTED. Hence users should always heed the error bars quoted in PRES_ADJUSTED_ERROR.

### 3.2.2. Truncated negative pressure drift (TNPD) in APEX floats

APEX floats with Apf-5, Apf-7, or Apf-8 controllers that set negative SURFACE PRESSURE (SP) to zero (then add an artificial 5 dbar) present a challenge to delayed-mode qc because information from SP on any negative pressure offset is lost, thus making the pressure data unadjustable[1]. The problem with some of these APEX floats having unknown negative pressure error escalated with the discovery of the oil microleak defect in Druck pressure sensors. The Druck oil microleak defect manifests itself as increasingly negative offset at all pressures, and will eventually end the useful life of the float. For a detailed description of the Druck oil microleak defect, please refer to the article "A review of recent problems with float CTD units and Druck pressure sensors" by S. Riser in *Argonautics*, Number 11, September 2009.

During delayed-mode qc of pressure measurements from APEX floats with Apf-5, Apf-7, or Apf-8 controllers, operators should first remove erroneous SP values and any isolated spikes in the time series by following the procedure described in Section 3.2.1. The delayed-mode operator should then examine the resulting valid and despiked SP time series, and determine whether there are long periods of zero SP readings (after removing the artificial 5 dbar) that qualify as "Truncated Negative Pressure Drift", which has the following definition:

*Truncated Negative Pressure Drift (TNPD) refers to the part of a float's time series from which **valid and despiked** SP (after removing the artificial 5 dbar) reads continuously zero without reverting back to positive values or containing any occasional positive values.*

*The continuous valid zero-reading period needs to span at least 6 months, preferably longer. This captures the microleakers whose oil leak rates are fastest, and allows for seasonal variability from half of an annual cycle when surface pressure values may read just below zero.*

*For floats whose useful life is less than 6 months or when the continuous valid zero-reading period is shorter than 6 months, the qualifying time span is at the PI's discretion.*

Examples (a) to (d) illustrate some cases that should or should not be classified as TNPD. Please note that in all of the following schematic examples, SP represents valid and despiked values (after removing the artificial 5 dbar).

Example (a). 100% of the time series is TNPD.



---

[1] This feature was corrected in the Apf-9 and later versions of the controller.

Example (b). There are occasional valid positive SP readings in the first part of the time series, followed by a continuous zero-reading period that does not contain any occasional valid positive readings.
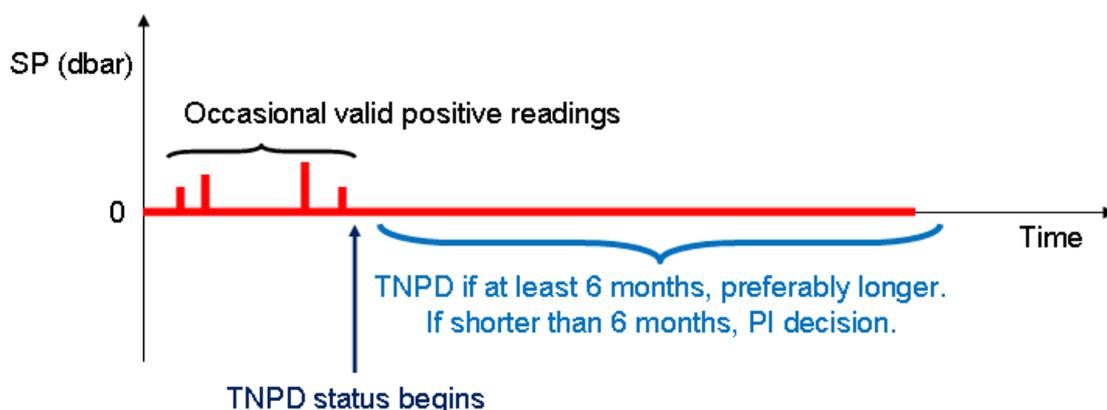


Example (c). The time series starts with continuous valid positive readings, then becomes continuously zero with no occasional valid positive readings.
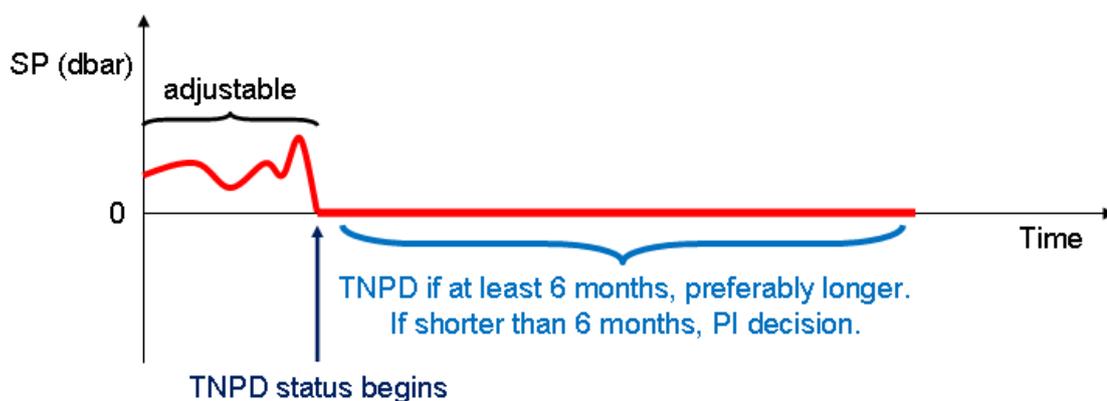


Example (d). The time series starts with a continuous zero-reading period, then reverts back to valid positive values. The initial zero-reading period does not qualify as TNPD. This is because pressure drifts are typically monotonic and therefore a reversal back to positive values indicates that the pressure sensor is not likely to have developed a negative drift.
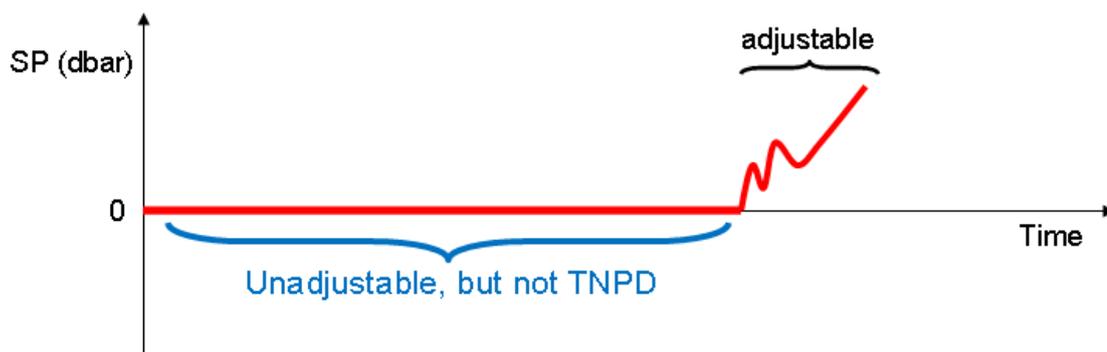
After determining which part of the time series qualifies as TNPD, the delayed-mode operator should then determine the probability of the TNPD data being affected by the Druck oil microleak problem. According to SeaBird, the date of manufacturing change at Druck that led to the oil microleak defect occurred sometime in mid-2006. The microleak failure rate jumped from 3% before 2006 to 30% in 2007. Any Druck pressure sensor with serial number greater than **2324175** falls into the group that has 30% likelihood of being affected by oil microleaks. Cross checking between the various APEX groups within Argo indicated that deployment of floats with Druck serial number greater than 2324175 occurred after October 2006. Since July 2009 SeaBird has begun screening Druck pressure sensors in order to identify those transducers that have microleaks.

One way to identify affected floats is by T/S analysis, since severe pressure error will lead to observable T/S anomalies. Anomalies associated with severe negative pressure drift include:

(a). Positive salinity drift; e.g. pressure error of −20 dbar will cause a positive salinity error of approximately 0.01 PSS-78. Statistical comparison methods that are used to determine conductivity sensor drift (e.g. WJO, BS, OW) can be used as diagnostic tools for these cases. Please refer to Section 3.4.1 for descriptions of these statistical comparison methods.

(b). Cold temperature anomaly whose size depends on vertical temperature gradient.

(c). Float-derived dynamic height anomalies significantly lower than satellite-derived sea level anomalies.

(d). Shoaling of isotherm depths independent of time/space migration of the float.

In addition to T/S analysis, delayed-mode operators should also observe when a float begins telemetering highly erratic data. This is a sign that it may be suffering from the Druck oil microleak problem, and that the pressure sensor may be about to fail completely. Note that symptoms of this failure are very similar to those of the Druck "snowflakes" problem, which affected floats that were manufactured in 2002 and 2003. Please refer to Appendix 4.3 for a brief description of the Druck "snowflakes" problem.

In light of these events, the following categories should be considered in assigning delayed-mode qc flags and error bars for data classified as TNPD.

1. <u>When float data do not show observable T/S anomalies that are consistent with increasingly negative pressure drift.</u> This means that the TNPD data may have unknown negative pressure error that is not severe. For these less severe cases, the adjusted variables should receive a delayed-mode qc flag of '2':

<div align="center">

PRES_ADJUSTED_QC = '2'

TEMP_ADJUSTED_QC = '2'

PSAL_ADJUSTED_QC = '2'.

</div>

(Note that TEMP_ADJUSTED_QC and PSAL_ADJUSTED_QC can change to '3' or '4' if TEMP and PSAL contain additional errors that are independent of the pressure error; e.g. pointwise temperature spike, conductivity cell drift, etc.)

For these less severe cases, two groups should be considered in assigning the pressure error bars in delayed-mode.

**(a). For TNPD data belonging to floats that used Druck pressure sensors with serial numbers less than 2324175, or were deployed before 1 October 2006 if the Druck serial numbers are unknown**, the likelihood of them being affected by the oil microleak problem is low, about 3%. Hence it is reasonable to cite the manufacturer quoted accuracy of 2.4 dbar as the pressure error for this group:

PRES_ADJUSTED_ERROR = 2.4 dbar.

**(b). For TNPD data belonging to floats that used Druck pressure sensors with serial numbers greater than 2324175, or were deployed after 1 October 2006 if the Druck serial numbers are unknown**, the likelihood of them being affected by the microleak disease is elevated, about 30%. For these suspicious data, an *upper bound* of the estimated error should be cited. Since a negative 20 dbar pressure error will cause a positive 0.01 salinity error, at which point T/S anomalies will become observable and data should be flagged as '4' as described in Category 2(b) below, 20 dbar has been chosen as the upper bound of the data error for this group:

PRES_ADJUSTED_ERROR = 20 dbar.

*Note that SeaBird will eventually provide a list of serial numbers that represents Druck sensors that have been screened as healthy. These healthy Druck sensors should be excluded from receiving the larger pressure error bar. Moreover, SeaBird has records that connect the Druck serial number to CTD number, and Teledyne WRC can make the connection to the float hull number.*

2. <u>When float data show observable T/S anomalies that are consistent with increasingly negative pressure drift after cycle-*n*.</u> This means that the TNPD data have unknown negative pressure error, and that the error becomes severe after cycle-*n*.

**(a). For the less severe part of the TNPD data before cycle-*n***, the adjusted variables should receive a dmqc flag of '2':

PRES_ADJUSTED_QC = '2'

TEMP_ADJUSTED_QC = '2'

PSAL_ADJUSTED_QC = '2',

while the pressure error should increase to 20 dbar:

PRES_ADJUSTED_ERROR = 20 dbar.

**(b). For the severe part of the TNPD data after cycle-*n*,** the adjusted variables should receive a dmqc flag of '4':

PRES_ADJUSTED_QC = '4'

TEMP_ADJUSTED_QC = '4'

PSAL_ADJUSTED_QC = '4'.

Please note that whenever PARAM_ADJUSTED_QC = '4', PARAM_ADJUSTED = FillValue, and PARAM_ADJUSTED_ERROR = FillValue.

**Note:** For the severe cases in Category 2(b), delayed-mode operators in consultation with float PIs should consider putting the real-time data on the grey list.

Example (e). A complex case belonging to Category 2, where T/S anomalies consistent with increasingly negative pressure drift are observed after cycle-*n*, part way through the TNPD portion of the time series.
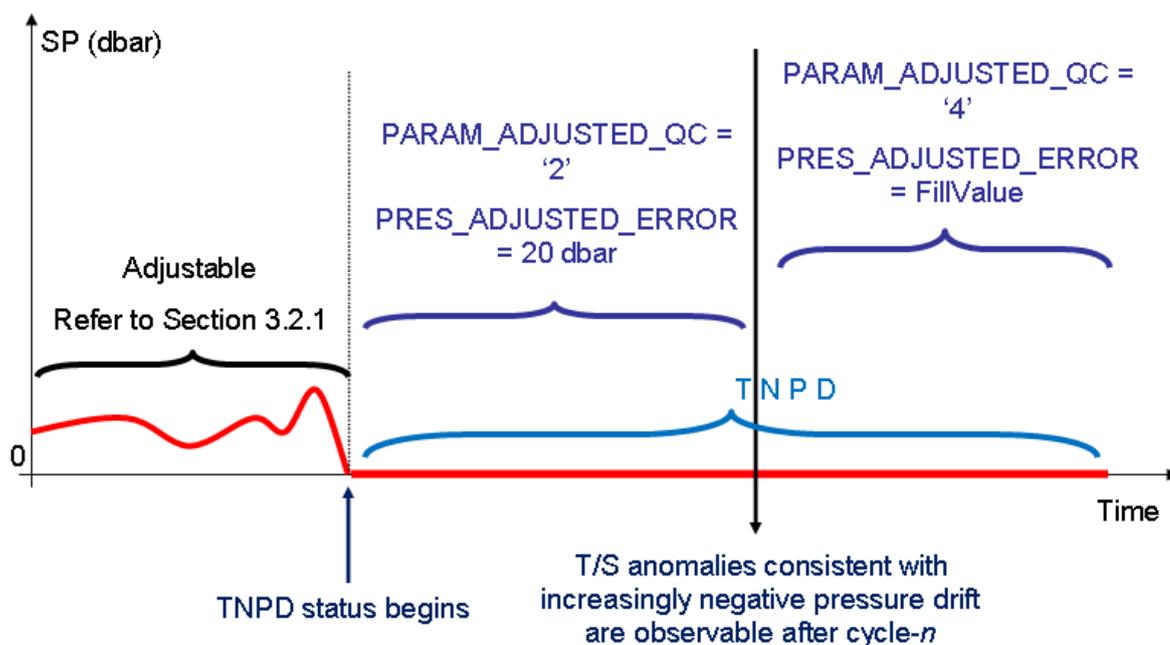


All TNPD data should receive a standard label in SCIENTIFIC_CALIB_COMMENT in the Argo single-cycle netcdf files, in the dimension corresponding to PRES. The standard label consists of the character string **"TNPD: APEX float that truncated negative pressure drift".** The delayed-mode operator may append to the end of this character string any other comments regarding PRES that he/she wishes to make.

For the portion of the time series that contains occasional valid positive SP readings (Example b), it is the PI's decision, based on the frequency of occurrence of the valid positive SP readings, on whether or not to adjust those profiles.

For the unadjustable but non-TNPD data (Example d, and cases in Example b where the PI decides to not adjust), any negative pressure offset is likely to be less than the manufacturer quoted accuracy of 2.4 dbar. For these unadjustable but non-TNPD data, if no additional error is found, then:

<div align="center">

PRES_ADJUSTED_QC = '1'

TEMP_ADJUSTED_QC = '1'

PSAL_ADJUSTED_QC = '1'

PRES_ADJUSTED_ERROR = 2.4 dbar.

</div>

**Summary flowchart for processing unadjustable APEX pressure data in delayed-mode**

```
┌─────────────────────────┐          ┌───────────────────────────────────┐
│   UNADJUSTABLE          │  ──────► │           Non-TNPD                  │
│   APEX PRESSURE DATA    │          │  PARAM_ADJUSTED_QC = '1'            │
│                         │          │  PRES_ADJUSTED_ERROR = 2.4 dbar     │
└─────────────────────────┘          └───────────────────────────────────┘
           │
           ▼
┌───────────────────────────────────┐
│              TNPD                  │
│     Insert TNPD comment string     │
└───────────────────────────────────┘
       │                        │
       ▼                        ▼
```

| Data do not show observable T/S anomalies that are consistent with increasingly negative pressure drift | Data show observable T/S anomalies consistent with increasingly negative pressure drift after cycle-$n$ |
|---|---|
| PARAM_ADJUSTED_QC = '2' | **Before cycle-$n$**<br>PARAM_ADJUSTED_QC = '2'<br>PRES_ADJUSTED_ERROR = 20 dbar |
| Druck pressure sensor with serial number <u>less than 2324175</u>, or float deployed before 1 October 2006<br><br>PRES_ADJUSTED_ERROR = 2.4 dbar | **After cycle-$n$**<br>PARAM_ADJUSTED_QC = '4'<br>PRES_ADJUSTED_ERROR = FillValue |
| Druck pressure sensor with serial number <u>greater than 2324175</u>, or float deployed after 1 October 2006<br><br>PRES_ADJUSTED_ERROR = 20 dbar | |

**Greylist real-time data**

## 3.3. Delayed-mode procedures for temperature

Delayed-mode qc for TEMP is done by subjective assessment of vertical profile plots of TEMP vs. PRES, and PSAL vs. TEMP. This assessment should be done in relation to measurements from the same float, as well as in relation to nearby floats and historical data. The assessment should aim to identify: (a) erroneous data points that cannot be detected by the real-time qc tests, and (b) vertical profiles that have the wrong shape.

Bad data points identified by visual inspection from delayed-mode analysts are recorded with TEMP_ADJUSTED_QC = '4'. Please note that whenever PARAM_ADJUSTED_QC = '4',

- PARAM_ADJUSTED = FillValue;

- PARAM_ADJUSTED_ERROR = FillValue.

TEMP_ADJUSTED, TEMP_ADJUSTED_ERROR, and TEMP_ADJUSTED_QC should be filled even when the data are good and no adjustment is needed. In these cases, TEMP_ADJUSTED_ERROR can be the manufacturer's quoted accuracy at deployment.

Please use the SCIENTIFIC CALIBRATION section in the netCDF files to record details of the delayed-mode adjustment.

## 3.4. Delayed-mode procedures for salinity

### 3.4.1. Introduction

Delayed-mode qc for PSAL described in this section are specifically for checking sensor drifts and offsets. Analysts should be aware that there are other instrument errors (e.g. conductivity cell thermal mass error, see Johnson *et al.* 2007; contact Gregory.C.Johnson@noaa.gov for the related adjustment software), and should attempt to identify and adjust them in delayed-mode. It is recommended that float salinity be adjusted for pressure offset and cell thermal mass error before sensor drift adjustment. If a measurement has been adjusted for more than one instrument error, analysts should attempt to propagate the uncertainties from all the adjustments.

The free-moving nature of profiling floats means that most float salinity measurements are without accompanying in-situ "ground truth" values for absolute calibration (such as those afforded by shipboard CTD measurements). Therefore Argo delayed-mode procedures for checking sensor drifts and offsets in salinity rely on reference datasets and statistical methods. However, since the ocean has inherent spatial and temporal variabilities, these drift and offset adjustments are subject to statistical uncertainties. **Users therefore should include the supplied error estimates in their usage of Argo delayed-mode salinity data.**

Three methods are available for detecting sensor drifts and offsets in float salinity, and for calculating adjustment estimates and related uncertainties:

1. Wong, Johnson, Owens (2003) estimates background salinity on a set of fixed standard isotherms, then calculates drifts and offsets by time-varying weighted least squares fits between vertically-interpolated float salinity and estimated background salinity. This method suits float data from open tropical and subtropical oceans. For the related software, please contact Annie Wong at awong@ocean.washington.edu.

2. Boehme and Send (2005) takes into account planetary vorticity in its estimates of background salinity, and chooses a set of desirable isotherms for calculations. This method suits float data from oceans with high spatial and temporal variabilities, where multiple water masses exist on the same isotherm, and where water mass distribution is affected by topographic barriers. For the related software, please contact Lars Boehme at lb284@st-andrews.ac.uk.

3. Owens and Wong (2009) improves the objective mapping scheme of WJO based on the method suggested by BS, and performs an optimal linear piecewise continuous fit in potential conductivity space. This method suits float data from the global ocean. For the related software, please contact Breck Owens at bowens@whoi.edu or Annie Wong at awong@ocean.washington.edu.

All three methods require an adequate reference database and an appropriate choice of spatial and temporal scales, as well as input of good/adjusted float pressure, temperature, position, and date of sampling. Therefore analysts should first check the reference database for adequacy and determine a set of appropriate spatial and temporal scales before using these methods. Operators should also ensure that other float measurements (PRES, TEMP, LATITUDE, LONGITUDE, JULD) are accurate or adjusted before they input them into the statistical tools for estimating reference salinity. See Sections 3.2 and 3.3 for delayed-mode procedures for PRES and TEMP.

### 3.4.2. Quality control and the semi-automatic part

The real-time qc procedures (described in Section 2) issue a set of qc flags that warns users of the quality of float salinity. These are found in the variable PSAL_QC. Float salinity with PSAL_QC = '4' are bad data that are in general unadjustable. However, delayed-mode operators can evaluate the quality and adjustability of these bad data if they have a reason to do so. Please refer to Section 4.1 for definitions of the Argo qc flags in real-time. The delayed-mode operators can edit <PARAM>_QC if they consider that data are flagged inappropriately.

In delayed-mode, float salinity values that have PSAL_QC = '1', '2' or '3' are further examined. Anomalies in the relative vertical salinity profile, such as measurement spikes and outliers that are not detected in real-time, are identified. Of these anomalies, those that will skew the least squares fit in the computation for drift and offset adjustments are excluded from the float series for evaluation of drifts and offsets. These measurements are considered unadjustable in delayed-mode.

Float salinity values that are considered adjustable in delayed-mode are assembled into time series, or float series. Sufficiently long float series are compared with statistical recommendations and associated uncertainties to check for sensor drifts and offsets. These statistical recommendations and associated uncertainties are obtained by the accepted methods listed in Section 3.4.1, in conjunction with appropriate reference datasets. These methods are semi-automatic and have quantified uncertainties.

Drifts and offsets can be identified in the trend of $\Delta S$ over time, where $\Delta S$ is the difference in salinity between float series and statistical recommendations. If $\Delta S = a + bt$ where t is time, then a is the offset and b is the drift rate. Note that these drifts and offsets can be sensor-related, or they can be due to real ocean events. PI evaluation is needed to distinguish between sensor errors and real ocean events.
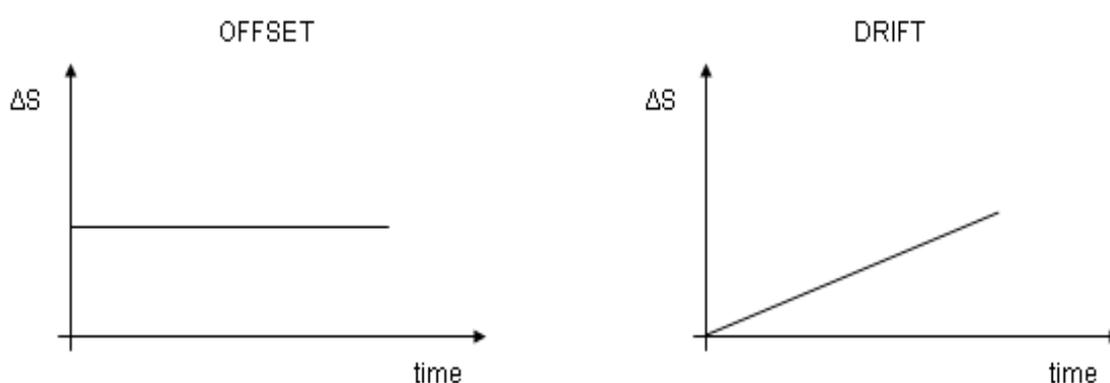


### 3.4.3. Splitting the float series and length of calibration window

If a float exhibits changing behaviour during its lifetime, the float series should be split into separate segments according to the different behaviours, so that one float series segment does not contaminate the other during the least squares fit process (e.g. the slowly-fouling segment does not contaminate the stable segment).

The following is a step-by-step guide on how to deal with float series with changing behaviours.

1). Identify different regimes in the float series. These can be:

- Stable measurements (no sensor drift), including constant offsets.
- Sensor drift with a constant drift rate.
- Transition phase where drift rate changes rapidly e.g. (a) 'elbow region' between stable measurements and constant drift; (b) initial biocide wash-off.
- Spikes.

2). Split the float series into discrete segments according to these different regimes or when there are too many missing cycles. Here is an example:
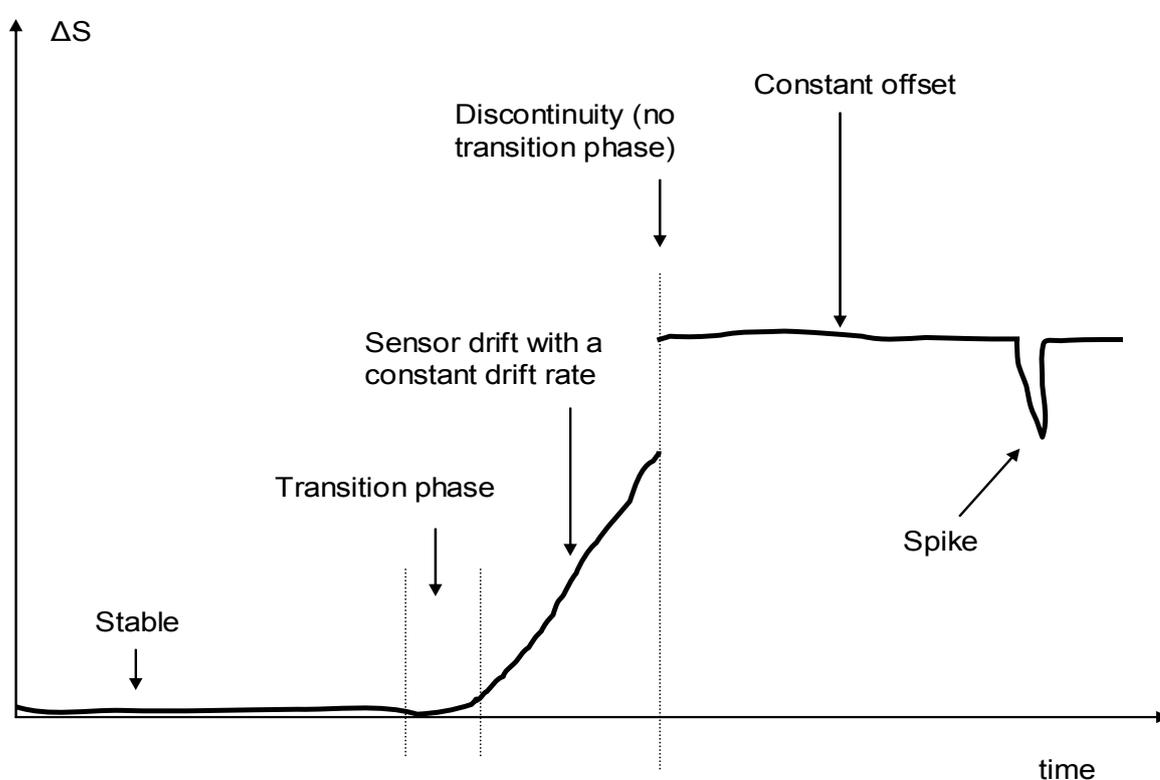


3). Choose length of sliding calibration window for each segment. These can be:

- Long window (+/− 6 months or greater) for the stable regime, or highly variable regimes where a long window is required to average over oceanographic variability to detect slow sensor drift, or period of constant drift rate.
- Short window (can be as short as +/− 10 days) for the transition phase.
- Zero length window for spikes. That is, adjust single profile.

4). Select temperature levels for exclusion from least squares fit (e.g. seasonal mixed layer, highly variable water masses).

5). Calculate proposed adjustment for each segment. The assembled proposed adjustments for the entire float series should be continuous and piecewise-linear within error bars, except where the delayed-mode operator believes there is a genuine discontinuity.

In general, the delayed-mode operator should aim to use as long a calibration window as possible, because a long calibration window (where the least squares fit is calculated over many cycles) will average over oceanographic noise and thus give a stable calibration. Hence splitting the float series into short segments is to be avoided (short segments mean short calibration windows, hence unstable calibrations).

## 3.4.4. The PI evaluation part

The PI (PI means Principal Investigator, or responsible persons assigned by the PI) should first check that the statistical recommendations are appropriate. This is because the semi-automatic methods cannot distinguish ocean features such as eddies, fronts, and water mass boundaries. Near such ocean features, semi-automatic statistical methods are likely to produce erroneous estimations. The associated uncertainties reflect the degree of local variability as well as the sparsity of reference data used in the statistical estimations. However, these associated uncertainties are sensitive to the choice of scales. Hence the PI also needs to determine that the associated uncertainties are realistic.

The PI then determines whether the proposed statistical adjustment is due to sensor malfunction or ocean variability. Care should be taken to not confuse real ocean events with sensor drifts and offsets. This is done by inspecting as long a float series as possible, and by evaluating other independent information. Some of the diagnostic tools are:

• Inspecting the trend of $\Delta S$ over time. Trends that reverse directions, or oscillate, are difficult to explain in terms of systematic sensor malfunction. These are often caused by the float sampling oceanographic features (e.g. eddies, fronts, etc.) that are not adequately described in the reference database.

• Visually checking the float trajectory with reference to oceanographic features such as eddies and rings that can introduce complications to the semi-automatic methods.

• Inspecting contour plots of float salinity anomaly time series. Systematic sensor malfunction should show up as salinity anomalies over several water masses.

• Using other independent oceanographic atlases to anticipate water mass changes that can occur along a float's path, and that can be misinterpreted as sensor malfunction.

• Inspecting residuals from objective maps.

• Cross-checking with nearby stable floats in cases of suspect sensor calibration offset.

If the PI is confident that sensor malfunction has occurred, then the recommended threshold for making an adjustment is when $\Delta S$ is greater than 2 times the error from the statistical methods, but the PI can provide an alternative estimate of uncertainty if they have a basis for doing so. Note that this guideline is to help the PI in deciding whether a slope or offset is statistically significant, and so should be used to evaluate the entire float segment being fitted, and not to single points.

In cases where the float series has been split into separate segments, the PI must ensure that the assembled adjustment for the entire float series is continuous within error bars, except where the PI believes there is a genuine discontinuity (see Step 5 in Section 3.4.3). This is to ensure that no artificial jump is introduced where the separate segments join. Adjustment continuity between separate float segments can be achieved by making adjustment in the transition phase even though the adjustment is below the 2 times error threshold limit.

In the following example, the float series experiences sensor drift after a stable period. The float series has been split for calibration. However, the float series has no discontinuity, so the final assembled adjustment should be continuous. Adjustment continuity is achieved by using model (a), not (b).
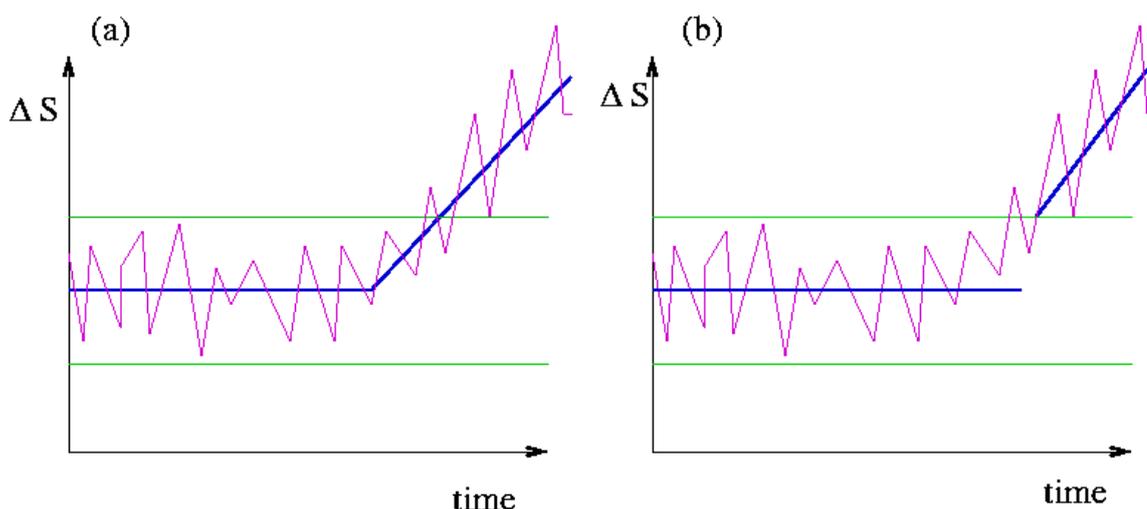


### 3.4.5. Assigning adjusted salinity, error estimates, and qc flags

After evaluating all available information, the PI then assigns adjusted salinity values, error estimates, and delayed-mode qc flags. In Argo netcdf files, these are found respectively in the variables PSAL_ADJUSTED, PSAL_ADJUSTED_ERROR, and PSAL_ADJUSTED_QC. Please refer to Section 4.1 for definitions of the Argo qc flags in delayed-mode.

Several Matlab-based graphical user interface softwares are available for interacting with Argo netcdf files. For examples of these softwares, please contact John Gilson (jgilson@ucsd.edu) or Paul Robbins (probbins@whoi.edu).

The following is a set of guidelines for assigning values to PSAL_ADJUSTED, PSAL_ADJUSTED_ERROR and PSAL_ADJUSTED_QC in Argo netcdf files.

For float salinity that are considered unadjustable in delayed-mode

For example, large spikes, or extreme behaviour where the relative vertical T-S shape does not match good data. These measurements are unadjustable.

- PSAL_ADJUSTED = FillValue;

- PSAL_ADJUSTED_ERROR = FillValue;

- PSAL_ADJUSTED_QC = '4'.

<u>For float salinity that are considered adjustable in delayed-mode</u>

These measurements have a relative vertical T-S shape that is close to good data. They are evaluated and adjusted for sensor drifts, offsets, and any other instrument errors.

i). When no adjustment is applied,

- PSAL_ADJUSTED = PSAL (original value);

- PSAL_ADJUSTED_ERROR = maximum [ statistical uncertainty, 0.01 ];

- PSAL_ADJUSTED_QC = '1', '2', or '3'.

ii). When an adjustment has been applied,

- PSAL_ADJUSTED = original value + adjustment recommended by statistical analyses, or adjustment provided by PI;

- PSAL_ADJUSTED_ERROR = maximum [ $(\sum \text{adjustment\_error}^2)^{1/2}$, 0.01 ], where "adjustment_error" is the uncertainty from each type of adjustment applied to PSAL. These can be statistical uncertainty from salinity drift adjustment, uncertainty from conductivity cell thermal mass adjustment, etc.

- PSAL_ADJUSTED_QC = '1', '2', or '3'.

iii). When LATITUDE, LONGITUDE, JULD are missing,

- Operators should fill the missing LATITUDE, LONGITUDE, JULD with interpolated x, y, and t wherever possible, and record POSITION_QC = '8', JULD_QC = '8'. The profile can then be evaluated and adjusted if necessary by using the interpolated x, y, t. The _ADJUSTED_ fields can then be filled accordingly.

<u>The following are some cases where PSAL_ADJUSTED_QC = '2' should be assigned</u>

- Adjustment is based on unsatisfactory reference database.

- Adjustment is based on a short calibration window (because of sensor behaviour transition, or end of sensor life) and therefore may not be stable.

- Evaluation is based on insufficient information.

- Sensor is unstable (e.g. magnitude of adjustment is too big, or sensor has undergone too many sensor behaviour changes) and therefore data are inherently of mediocre quality.

- When a float exhibits problems with its pressure measurements.

## 3.4.6. Summary flowchart

# Argo salinity sensor drift & offset QC procedures

Manual evaluation to detect anomalies on the relative profile, such as spikes, that are not detected in RT.
Remove anomalies that may skew the drift/offset correction.

Do not use in least squares fit.

Use in least squares fit.

PSAL_ADJUSTED
= FillValue
PSAL_ADJUSTED_ERROR
= FillValue
PSAL_ADJUSTED_QC = 4

Use accepted methods and reference database, split series and select appropriate length for the sliding window, to calculate recommended salinity adjustments.

**PI evaluation - consider long time series and other supporting information to determine whether sensor malfunction has happened.**

No sensor error has been detected, or sensor drift and/or offset is not significant

< max [ 2 x statistical uncertainty, instrument accuracy ]

Sensor drift and/or offset has been detected and is significant

> max [ 2 x statistical uncertainty, instrument accuracy ]

No adjustment needed.

Apply adjustment, or declare unadjustable.

PSAL_ADJUSTED = PSAL (original value)
PSAL_ADJUSTED_ERROR
= max [ ($\sum$adjustment_error$^2$)$^{1/2}$, 0.01 ],
or uncertainty provided by PI
PSAL_ADJUSTED_QC = 1, 2 or 3

PSAL_ADJUSTED
= value recommended by statistical analyses, or value provided by PI
PSAL_ADJUSTED_ERROR
= max [ ($\sum$adjustment_error$^2$)$^{1/2}$, 0.01 ],
or uncertainty provided by PI
PSAL_ADJUSTED_QC = 1, 2 or 3

**OR**

PSAL_ADJUSTED
= FillValue
PSAL_ADJUSTED_ERROR
= FillValue
PSAL_ADJUSTED_QC = 4

### 3.4.7. Timeframe for availability of delayed-mode salinity data

The statistical methods used in the Argo delayed-mode process for checking sensor drifts and offsets in salinity require the accumulation of a time series for reliable evaluation of the sensor trend. Timeframe for availability of delayed-mode salinity data is therefore dependent on the sensor trend. Some floats need a longer time series than others for stable calibration. Thus delayed-mode salinity data for the most recent profile may not be available until sufficient subsequent profiles have been accumulated. The *default* length of time series for evaluating sensor drift is 12 months (6 months before and 6 months after the profile). This means that in general, the timeframe of availability of drift-adjusted delayed-mode salinity data is 6+ months after a profile is sampled.

Users should also be aware that changes may be made to delayed-mode files at any time by DACs and delayed-mode operators. For example, delayed-mode files may be revised when new CTD or float data become available after the original delayed-mode assessment and adjustment. The date of latest adjustment of a parameter can be found in CALIBRATION_DATE.

Anytime an Argo file is updated for any reason, the DATE_UPDATE variable will reflect the date of the update. The "profile index file" on the GDACs contains the DATE_UPDATE information (along with other information) for every file on the GDACs and can be used to monitor updates. The profile index file is maintained in the top-level GDAC directory and is named "ar_index_global_prof.txt"; index files also exist for the meta-data and trajectory files.

### 3.4.8. References

Böhme, L. and U. Send, 2005: Objective analyses of hydrographic data for referencing profiling float salinities in highly variable environments. *Deep-Sea Research II*, **52/3-4**, 651-664.

Johnson, G.C., J.M. Toole, and N.G. Larson, 2007: Sensor corrections for Sea-Bird SBE-41CP and SBE-41 CTDs. *Journal of Atmospheric and Oceanic Technology*, **24**, 1117-1130.

Owens, W.B. and A.P.S. Wong, 2009: An improved calibration method for the drift of the conductivity sensor on autonomous CTD profiling floats by θ-S climatology. *Deep-Sea Research*, Part I: Oceanographic Research Papers, **56(3)**, 450-457.

Wong, A.P.S., G.C. Johnson, and W.B. Owens, 2003: Delayed-mode calibration of autonomous CTD profiling float salinity data by θ-S climatology. *Journal of Atmospheric and Oceanic Technology*, **20**, 308-318.

## 3.5.    Compulsory variables to be filled in a D file

This section lists the compulsory variables that must be filled in an Argo netCDF file that has been through the delayed-mode process.

### 3.5.1. Measurements for each profile

The following are compulsory measurement variables that must be filled in a D file:

- <PARAM>_ADJUSTED;

- <PARAM>_ADJUSTED_QC;

- <PARAM>_ADJUSTED_ERROR.

The variable PROFILE_<PARAM>_QC should be recomputed when <PARAM>_ADJUSTED_QC becomes available. See Section 4.2 for definitions.

Here, <PARAM> denotes all the measurement parameters that are reported in the netCDF file. Currently, <PARAM> = PRES, TEMP, PSAL are the fundamental measurement parameters that are reported in every Argo netCDF file and have approved delayed-mode qc procedures. See Sections 3.2, 3.3, 3.4 on how to fill their related _ADJUSTED_ variables.

For <PARAM> = CNDC, CNDC_ADJUSTED, CNDC_ADJUSTED_QC, and CNDC_ADJUSTED_ERROR can be their respective FillValues. If they are not their respective FillValues, then CNDC_ADJUSTED must be calculated to be consistent with PSAL_ADJUSTED, TEMP_ADJUSTED, and PRES_ADJUSTED. CNDC_ADJUSTED_QC must be consistent with PSAL_ADJUSTED_QC, and CNDC_ADJUSTED_ERROR must be consistent with PSAL_ADJUSTED_ERROR.

Some Argo netCDF files report DOXY. There is currently no approved method for delayed-mode qc on DOXY. Therefore, DOXY_ADJUSTED = original values recorded in DOXY, DOXY_ADJUSTED_QC = '0', DOXY_ADJUSTED_ERROR = FillValue, and PROFILE_DOXY_QC = ' ' (i.e. Blank, the FillValue for PROFILE_DOXY_QC).

### 3.5.2. Scientific calibration information for each profile

Within each single-profile Argo netcdf file is a scientific calibration section that records details of delayed-mode adjustments. It is compulsory to fill the variables in the scientific calibration section at the completion of delayed-mode qc.

In the scientific calibration section, every measurement parameter recorded in the netCDF file should be listed in the variable PARAMETER. For every measurement parameter listed in PARAMETER (PRES, TEMP, PSAL, CNDC, DOXY), there are four variables to record scientific calibration details:

- SCIENTIFIC_CALIB_EQUATION;

- SCIENTIFIC_CALIB_COEFFICIENT;

- SCIENTIFIC_CALIB_COMMENT;

- CALIBRATION_DATE.

In cases where no adjustment has been made, SCIENTIFIC_CALIB_EQUATION and SCIENTIFIC_CALIB_COEFFICIENT shall be filled by their respective FillValues.

SCIENTIFIC_CALIB_COMMENT shall contain wordings that describe the evaluation.

E.g. 1: "No adjustment is needed because no significant sensor drift has been detected."

E.g. 2: "No approved method for delayed-mode qc on DOXY is available."

In cases where adjustments have been made, examples of wordings for PSAL can be:

SCIENTIFIC_CALIB_EQUATION: "PSAL_ADJUSTED = PSAL + $\Delta$S, where $\Delta$S is calculated from a potential conductivity (ref to 0 dbar) multiplicative adjustment term r."

SCIENTIFIC_CALIB_COEFFICIENT: "r = 0.9994 ($\pm$ 0.0001), vertically averaged $\Delta$S = $-$ 0.025 ($\pm$ 0.003)."

SCIENTIFIC_CALIB_COMMENT: "Sensor drift detected. Adjusted float salinity to statistical recommendation as in WJO (2003), with WOD2001 as the reference database. Mapping scales used are 8/4, 4/2. Length of sliding calibration window is +/- 20 profiles"

The PI is free to use any wordings he/she prefers. Just be precise and informative.

Regardless of whether an adjustment has been made or not, the date of delayed-mode qc for each measurement parameter (PRES, TEMP, PSAL, CNDC, DOXY) should be recorded in CALIBRATION_DATE, in the format YYYYMMDDHHMISS.

## 3.5.3. Other variables in the netcdf file

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.

The variable DATA_MODE should record 'D'.

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.

Lastly, the name of the single-profile Argo netcdf file is changed from R*.nc to D*.nc.

# 4. Appendix

## 4.1. Reference Table 2: Argo quality control flag scale

This table describes the Argo qc flag scales. Please note that this table is used for all measured parameters. This table is named Reference Table 2 in the Argo User's Manual.

| n | Meaning | Real-time comment | Delayed-mode comment |
|---|---------|-------------------|----------------------|
| 0 | No QC was performed | No QC was performed | No QC was performed |
| 1 | Good data | All Argo real-time QC tests passed. | The adjusted value is statistically consistent and a statistical error estimate is supplied. |
| 2 | Probably good data | Probably good data | Probably good data |
| 3 | Probably bad data that are potentially correctable | Test 15 or Test 16 or Test 17 failed and all other real-time QC tests passed. These data are not to be used without scientific correction. A flag '3' may be assigned by an operator during additional visual QC for bad data that may be corrected in delayed-mode. | An adjustment has been applied, but the value may still be bad. |
| 4 | Bad data | Data have failed one or more of the real-time QC tests, excluding Test 16. A flag '4' may be assigned by an operator during additional visual QC for bad data that are uncorrectable. | Bad data. Not adjustable. Data replaced by FillValue. |
| 5 | Value changed | Value changed | Value changed |
| 6 | Not used | Not used | Not used |
| 7 | Not used | Not used | Not used |
| 8 | Interpolated value | Interpolated value | Interpolated value |
| 9 | Missing value | Missing value | Missing value |

## 4.2.    Reference Table 2a: profile quality flags

Please note that this table is used for all measured parameters. This table is named Reference Table 2a in the Argo User's Manual.

$N$ is defined as the percentage of levels with good data where:

- QC flag values of 1, 2, 5, or 8 are GOOD data

- QC flag values of 9 (missing) are NOT USED in the computation

- All other QC flag values are BAD data

The computation should be taken from <PARAM>_ADJUSTED_QC if available and from <PARAM>_QC otherwise.

| n | Meaning |
|---|---|
| " " | No QC was performed |
| A | $N$ = 100%; All profile levels contain good data |
| B | 75% <= $N$ < 100% |
| C | 50% <= $N$ < 75% |
| D | 25% <= $N$ < 50% |
| E | 0% < $N$ < 25% |
| F | $N$ = 0%; No profile levels have good data |

## 4.3.　Common instrument errors and failure modes

This section describes some common instrument errors and failure modes that will cause error in float measurements.

### 1. TBTO leakage

TBTO (tributyltinoxide) is a wide spectrum poison that is used to protect conductivity cells from biofouling. However, accidental leakage of TBTO onto the conductivity cell can occur. This will result in fresh salinity offsets in float series that usually gets washed off. Delayed-mode analysts should pay special attention to the shape of the salinity profiles at the beginning of the float series if TBTO leakage is suspected.

### 2. Pollution events

Any pollution on the conductivity cell will result in erroneously fresh salinity measurements. When pollution washes off, reversal of sensor drift trend can occur. Delayed-mode analysts need to be careful in splitting float series in such cases.

### 3. Ablation events

Any ablation of the conductivity cell, such as etching, scouring, or dissolution of the glass surface, will result in erroneously salty salinity measurements.

### 4. Conductivity cell geometry changes

The geometry of conductivity cells can change, thus causing electrodes to change distance. This will result in either an increase or decrease in salinity values.

### 5. Conductivity cell circuit changes

The circuit within the conductivity cell contains capacitors and resistors. Changes to any of these electrical components will affect electrical conductivity and thus will give erroneous (fresh or salty) salinity measurements. Electrical complication can result in sensor drifts that have varying drift rates (e.g. drift rates can change from slow and linear to exponential). Usually jumps in salinity measurements are an indication of electrical malfunction. If electrical complication is suspected, delayed-mode analysts should check the shape of the vertical salinity profiles for adjustability. Usually the vertical profiles after a measurement jump are wrong and so are uncorrectable.

### 6. Low voltage at end of float life, and "Energy Flu"

APEX floats often experience a sudden rapid decrease in available battery energy reserves. This premature exhaustion of battery, known as "Energy Flu", usually starts about 2 years after deployment. The sharp drop in battery voltage related to "Energy Flu", as well as the low voltage towards the end of a float's natural life, will produce low-of-correct salinity values. Towards the end of float life, low voltage will result in large drift, followed by death. "Energy Flu" will cause spikes that get worse and more frequent, also followed by death.

## 7. Druck pressure sensor "snowflakes" problem

About 4% of SBE41 CTDs that were manufactured in late 2002 through end of 2003 have experienced the Druck pressure sensor "snowflakes" problem. SeaBird has fixed this problem in 2004, so this feature is only included in this section for identifying the profiles that have been affected. The Druck pressure sensor "snowflakes" problem is due to internal electrical shorting by the growth of titanium oxide particles ('snowflakes') in the oil-filled cavity in the pressure sensor, causing the pressure sensor to report erratic pressure measurements, or going to full scale, i.e. either report PRES ~ 3000 dbar or −3000 dbar. These erratic pressure measurements will preferentially report deeper than correct. The firmware tries to adjust the piston according to the erroneous deeper pressures, causing the float to park shallower. The float will thus progressively become a surface drifter. Erroneous deeper pressures will also result in the firmware placing the pointer at the deeper nominal sampling levels in the lookup table, thus causing the float to take a sample everytime the firmware performs a lookup (every 6 seconds). The result is a series of measurements from very close-together depth levels. Progressively shallower profiles and close-together measurements are therefore two ways to identify whether the Druck pressure sensor has been contaminated. When the Druck pressure sensor has been contaminated, pressure measurements become suspect and should be considered bad. The corresponding temperature and salinity measurements are therefore also suspect and should be considered bad.

## 8. Druck pressure sensor "oil microleaks" problem

Another pathology in Druck pressure sensors is oil microleaks past the glass/metal seal. This oil leak leads to an internal volume loss, which then exhibits itself as an increasing negative offset at all pressures. At the early stages of microleak, float measurements are still correctable and usable. However, as more and more oil is leaked, the flexible titanium diaphragm will dip so far down the oil chamber that it will short the electrical parts, causing erratic behaviour in float measurements. This is the end stage of microleak, and the data at this point are bad and uncorrectable.

## 9. Incorrect pressure sensor coefficient

Incorrect scaling coefficient in the pressure sensor will give anomalous T-S curves at depth. The T-S relation will look acceptable, but at depth it will look as if the float is sampling an anomalous water mass relative to nearby floats. Delayed-mode analysts should try to re-scale pressure measurements (e.g. PRES' = PRES * X) to see whether the T-S curve can be recovered. Air bubbles in the pressure transducer can also cause erroneous pressure measurements that are visible as anomalous T-S curves.

## 10. Conductivity cell thermal mass error

Salinity reported immediately after a float has crossed a strong thermal gradient can be in error as a result of conductivity cell thermal mass. This error arises because the thermal inertia in the flow duct alters the temperature of water entering the conductivity cell, thus inducing a conductivity error. For details please refer to Johnson *et al.* 2007. A float that transits from cold to warm water can result in fresh error, and from warm to cold in salty error. These errors can exceed 0.01 (PSS-78) for strong thermal gradients, and sometimes result in unstable fresh spikes at the base of the mixed layer. This salinity error can be corrected if the ascent rate of the float is known. A correction algorithm is available from Greg Johnson at Gregory.C.Johnson@noaa.gov.

## 11. Abnormal "salty hooks" at base of profiles

In some APEX floats, a "salty hook" may be observed at the base of the profile. These are the deepest salinity measurements that are high of correct by about 0.005 (PSS-78) with respect to shallower samples. The "hook" appearance occurs when the two deepest measurements are reported at nearly identical pressures. The first measurement is the one taken at the end point of descent; the second measurement is the first deep sample taken during ascent from the pressure lookup table. These "salty hooks" are caused by asymmetry in Bernoulli flushing during ascent/descent. Higher salinity water in the conductivity cell carried from the surface or park level to deep profile level is not being flushed out completely before the deepest sample is taken, thus resulting in salinity that is high of correct. These "salty hooks" cannot be detected by the real-time tests, so delayed-mode operators are urged to examine carefully the base of profiles for these "hooks", and flag them appropriately in delayed-mode.

## 4.4. Criteria for CTD profiles to be retained in the reference database

The following criteria are used to select CTD data as reference for delayed-mode quality control of Argo salinity profiles in the open ocean.

1). Use only data that have passed all NODC quality control tests for observed level data.

2). Use all country codes.

3). Use only profiles that sampled deeper than 900 dbar.

4). Weed out all data points outside these ranges: $24 < S < 41$, $0.01 < P < 9999$, $0°C < T < 40°C$, except for WMO boxes with latitudes north of 60°N or south of 50°S, where

$-2.5°C < T < 40°C$.

5). For WMO boxes that contain more than 10,000 profiles, only select profiles that are post-1995.

6). Eliminate nearby duplicates.

7). Do objective residual analysis using previously qc'd reference data to identify anomalies. Then do visual inspection of anomalies.

8). Identify each reference profile with a unique ID, e.g. under the variable SOURCE.

It is recommended that in regions with adequate reference data, that delayed-mode qc for salinity should use CTD data only. If CTD data are too sparse, bottle data (BOT) may be included.

## 4.5. Criteria for Argo profiles to be retained in the reference database

The following criteria are used to select Argo data as reference for delayed-mode quality control of Argo salinity profiles in the open ocean.

1). No real-time data.

2). No floats that fail within 1 year of deployment.

3). No cycles within 6 months of end of record.

4). No cycles that have salinity drift adjustment ($\Delta S > .001$ PSS-78 in bottom data to distinguish from thermal lag adjustment at shallower levels).

5). No floats whose deepest sampling level is shallower than 800-dbar.

6). No cycles following ones that have salinity drift adjustment ($\Delta S > .001$ PSS-78 in bottom data).

7). No cycles where less than 90% of values (P, T, S) are good.

8). No cycles < 18 (first 6 months) to be used (due to the propensity of some floats to acquire TBTO contamination).

9). No cycles in the 6 months prior to salinity drift adjustment ($\Delta S > .001$ PSS-78 in bottom data).

## 4.6. Consistency checks for D files format at the GDACs

The following is a list that is used at the GDACs for checking D files format.

1). <PARAM>_ADJUSTED and <PARAM>_ADJUSTED_ERROR must contain data, except when <PARAM>_ADJUSTED_QC = '4' or '9'. Here, <PARAM> = PRES, TEMP, PSAL, and CNDC.

2). Where <PARAM> = DOXY, DOXY_ADJUSTED should be filled with the same values as DOXY. Furthermore, DOXY_ADJUSTED_QC should record '0', DOXY_ADJUSTED_ERROR = FillValue, and PROFILE_DOXY_QC = ' '.

3). If PRES_ADJUSTED_QC = '4', then TEMP_ADJUSTED_QC = '4' and PSAL_ADJUSTED_QC = '4'.

4). <PARAM>_ADJUSTED_QC cannot be '0', except when <PARAM> = DOXY.

5). POSITION_QC and JULD_QC cannot be '0'.

6). No variable should be filled with the netCDF value of IEEE NaN.

7). In the Scientific Calibration section, PARAMETER should have N_PARAM entries equal to the number of measurement parameters recorded in the netCDF file.

8). In the Scientific Calibration section, SCIENTIFIC_CALIB_COMMENT should have non-FillValue entries in every N_PARAM dimension.

9). In the Scientific Calbration section, CALIBRATION_DATE should have non-FillValue entries in every N_PARAM dimension, and should have format YYYYMMDDHHMISS (seconds must be 0 to 59).

10). DATE_UPDATE should be equal to or later than any CALIBRATION_DATE, HISTORY_DATE, DATE_CREATION, JULD, and JULD_LOCATION.

11). There should be at least one HISTORY record.

12). All dates must be after 1st Jan 1997, and before GDAC file time.

13). All dates must be 14 digit strings, in the format YYYYMMDDHHMISS (seconds must be 0 to 59).


14). Character strings should not contain the NULL character.