Processing BGC-Argo pH data at the DAC level

Version 1.0
April 9th 2018
Argo data management
Processing BGC-Argo pH data at the DAC level

Authors: Kenneth S. Johnson, Joshua N. Plant, Tanya L. Maurer

How to cite this document
Table of contents

1 INTRODUCTION ........................................................................................................................................ 7
2 COMPUTING PH ......................................................................................................................................... 9
3 CONVERSION OF PH SCALES .................................................................................................................. 10
4 PH CALIBRATION FILE ............................................................................................................................. 11
5 SENSOR METADATA ................................................................................................................................. 12
6 REAL TIME QUALITY CONTROL FOR PH ............................................................................................... 13
  6.1 COMMON ARGO REAL-TIME QUALITY CONTROL TESTS ON PH VERTICAL PROFILES .............. 13
  6.2 SPECIFIC ARGO REAL TIME QUALITY CONTROL FOR PH VERTICAL PROFILES............... 14
  6.3 TEST APPLICATION ORDER ON PH VERTICAL PROFILES ......................................................... 14
  6.4 SCIENTIFIC CALIBRATION FOR EACH PROFILE ............................................................................ 14
7 ADJUSTMENT AND QUALITY CONTROL FOR PH IN DELAYED MODE AND REAL TIME ................. 15
  7.1 ADJUSTMENT METHOD ....................................................................................................................... 15
  7.2 SPECIFIC REAL TIME QUALITY CONTROL FOR ADJUSTED PH VERTICAL PROFILES.......... 16
  7.3 SCIENTIFIC CALIBRATION FOR EACH PROFILE ............................................................................ 17
8 QUALITY CONTROL FOR DATA ADJUSTED IN DELAYED MODE .................................................... 19
  8.1 DATA ADJUSTMENT PROCESS ........................................................................................................... 19
  8.2 PH DATA ADJUSTMENT SOFTWARE AND EXAMPLE .................................................................... 21
  8.3 EDITING RAW QC AND ADJUSTED QC IN DELAYED MODE ......................................................... 25
  8.4 COMPULSORY VARIABLES TO BE FILLED IN A BD PROFILE FILE ........................................... 25
    8.4.1 QC AND ADJUSTED VARIABLES ............................................................................................. 25
    8.4.2 SCIENTIFIC CALIBRATION FOR EACH PROFILE .................................................................. 26
    8.4.3 OTHER VARIABLES IN A BD PROFILE FILE ......................................................................... 27
    8.4.4 PROFILES NAMING CONVENTION ......................................................................................... 28
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>MATLAB CODE TO COMPUTE PH</td>
<td>28</td>
</tr>
<tr>
<td>9.1 CHECK VALUES</td>
<td>30</td>
</tr>
<tr>
<td>REFERENCES</td>
<td>32</td>
</tr>
</tbody>
</table>
History of the document

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Authors</th>
<th>Modification</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>September 2017</td>
<td>Kenneth S. Johnson, Joshua N. Plant, Tanya L. Maurer</td>
<td>Initial version</td>
</tr>
</tbody>
</table>

Reference Documents

<table>
<thead>
<tr>
<th>Reference N°</th>
<th>Title</th>
<th>Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>#RD1</td>
<td>Argo Quality Control Manual for CTD and Trajectory Data</td>
<td><a href="http://dx.doi.org/10.13155/33951">http://dx.doi.org/10.13155/33951</a></td>
</tr>
<tr>
<td>#RD2</td>
<td>Argo Quality Control Manual for Biogeochemical Data</td>
<td><a href="http://dx.doi.org/10.13155/40879">http://dx.doi.org/10.13155/40879</a></td>
</tr>
<tr>
<td>#RD3</td>
<td>Argo user manual</td>
<td><a href="http://dx.doi.org/10.13155/29825">http://dx.doi.org/10.13155/29825</a></td>
</tr>
</tbody>
</table>
Preamble:

This document was produced in response to action item 21 from the 5th Bio-Argo Data Management meeting in Tianjin, China (Oct., 2016).

“Few would [dispute] that pH is truly a master variable that permeates just about any area of study of living organisms and extends into the physical and chemical world in which living organisms come about, grow, develop, reproduce and die. There is arguably no other master variable that captures the complexities of the interactions between biological, physical and chemical aspects of the world to a similar extent” (Rengel, 2002)
1 Introduction

Seawater proton concentration is a master variable that controls the air-sea gas exchange of CO$_2$, the ability of organisms to produce calcium carbonate shells, and that tracks the production and respiration of organic carbon as CO$_2$ is removed or added to water by biological processes. The proton concentration in seawater [H$^+$] (mol kg-seawater$^{-1}$) is typically reported as the pH = -$\log_{10}$ [H$^+$]. The in situ proton concentration ranges from about 3 to 30 nmol kg-seawater$^{-1}$ (7.5<pH<8.5).

Seawater pH is measured from profiling floats using a Deep-Sea DuraFET pH sensor that is manufactured at MBARI (Johnson et al., 2016) or at Sea-Bird Scientific (Float pH). The core of each sensor is an Ion Sensitive Field Effect Transistor (ISFET), produced by Honeywell, that responds to proton activity and a Ag/AgCl reference electrode that responds to chloride ion activity. When properly calibrated, the sensor measures the activity of HCl, which is the product of the activities of the two ions aH$^+$ × ac$^{-}$ (Martz et al., 2010). These ion concentrations and activities are related through the activity coefficient ($\gamma$)

$$a_{H^+} \times a_{Cl^-} = m_{H^+} \times m_{Cl^-} \times \gamma_{H^+} \times \gamma_{Cl^-} = m_{H^+} \times m_{Cl^-} \times (\gamma_{\pm HCl})^2$$  (1)

where m signifies molality (mol kg-H$_2$O$^{-1}$). The activity coefficient is a thermodynamic property of neutral pairs of ions that can be determined by experiment. The mean activity coefficient of HCl in seawater ($\gamma_{\pm HCl} = (\gamma_{H^+} \times \gamma_{Cl^-})^{0.5}$) was measured by Khoo et al. (1977) as a function of temperature and salinity at atmospheric pressure.

The DuraFET sensor is operated in a mode where the potential between the reference electrode and the transistor source of the ISFET (VRS_PH with units of volt) follows the Nernst equation (Martz et al., 2010; Takeshita et al., 2014; Johnson et al., 2016)

$$VRS\_PH = k_{T,p} - R \times T$$

### Table 1 Mapping between variable names used in equations and the text, and Argo physical parameter names.

<table>
<thead>
<tr>
<th>Document Variable</th>
<th>Units</th>
<th>Argo Variable</th>
</tr>
</thead>
<tbody>
<tr>
<td>T</td>
<td>°K</td>
<td>TEMP + 273.15</td>
</tr>
<tr>
<td>t</td>
<td>°C</td>
<td>TEMP</td>
</tr>
<tr>
<td>S</td>
<td>Practical Salinity</td>
<td>PSAL</td>
</tr>
<tr>
<td>P</td>
<td>bar</td>
<td>PRES/10</td>
</tr>
<tr>
<td>pH</td>
<td>dimensionless</td>
<td>PH_IN_SITU_TOTAL or PH_IN_SITU_FREE (depends on pH scale)</td>
</tr>
</tbody>
</table>

In the text and equations of this document, a variety of concentration units are used to maintain consistency with underlying experimental determinations of required thermodynamic quantities. These units include measurements on the molality scale (mol of solute kg$^{-1}$ solvent, in this case pure water) and the informally named molinity scale (mol of solute kg-seawater$^{-1}$). There is not consistency in their usage through the seawater thermodynamics that underlie pH measurements so diligence is required to produce a final result which is accurate. The differences in scales are significant and correspond to variations in pH of 0.015. To further compound matters, the pH of seawater can be reported on several scales which are discussed in more detail below.

Additionally pressure and temperature are used to compute thermodynamic properties that have been fitted to empirical equations. These equations often use different units than reported by Argo floats (i.e. T °K vs T °C and pressure reported in bar vs dbar). Table 1 defines the mapping between Argo physical parameter names and those used in the following document. The aim of this document is to step the user through the calculation of pH from Argo profiling float data in a consistent manner, accounting for all of these nuances.
\[
\ln(10)/F \times \log_{10} (a_{H^+} \times a_{Cl^-}) \\
= k_{T,P} - R \times T \times \ln(10)/F \times \log_{10} \left( \gamma_{HCl, T,P} \right)^2 - R \times T \times \ln(10)/F \times \log_{10} (m_{H^+} \times m_{Cl^-})
\]

(2)

where \(k_{T,P}\) (volts) is the reference potential of the sensor at the in situ temperature and pressure, \(T\) is the temperature (°K), \(R\) is the gas constant \((8.31446 \text{ J mol}^{-1} \text{ °K}^{-1})\), and \(F\) is the Faraday constant \((96485 \text{ Coulombs mol}^{-1})\). In equation (2), \(m_{H^+}\) refers to the free protons only and not those bound by sulfate ion. The term \(RT \ln(10)/F\) is termed the Nernst slope. Laboratory experiments have demonstrated that the sensor response follows the Nernst slope with insignificant error from pH 2 to 10 (Takeshita et al., 2014). The mean activity coefficient of HCl is at the temperature and pressure of the measurement. The mapping between variable names used in equation (2), and following, and approved Argo physical parameter names is shown in Table 1.

Several different pH scales are in use for seawater measurements (Marion et al., 2011; Waters and Millero, 2012). As noted above, the Deep-Sea DuraFET responds to the free (unbound by other ions) proton activity and the molality of free protons can be computed from equation (2). However, it is more common in ocean sciences to report pH on the Total Proton scale, which includes binding of protons by the sulfate ion:

\[
\text{PH\_IN\_SITU\_TOTAL} = -\log_{10} \left( [H^+] + [HSO}_4^-] \right) = -\log_{10} \left( [H^+](1 + S_T / K_S) \right)
\]

(3)

where \(S_T\) is the total sulfate concentration (mol kg-seawater\(^{-1}\)) and \(K_S\) is the dissociation constant of bisulfate, which is defined as (Dickson, 1990)

\[
K_S = [H^+] \times S_T / [HSO}_4^-].
\]

(4)

The conversion of pH scales is discussed further in Section 3.

Also, note that there is considerable ambiguity in the literature as to whether the various pH scales are based on proton concentrations in units of molality \((m_{H^+}, \text{ mol kg-H}_2\text{O}^{-1})\) or units of molinity\(^1\) \(([H^+], \text{ mol kg-seawater}^{-1})\). The concentrations of a species \(X\) on these two scales are related by the equation

\[
m_X \text{ (mol kg H}_2\text{O}^{-1}) = [X] \text{ (mol kg seawater}^{-1}) / (1 - 0.001005 \times S)
\]

(5)

where the factor 0.001005 converts the practical salinity value \((S, \text{ distinguished from the sulfate concentration } S_T)\) to the approximate mass of salt (kg) in one kg of seawater (Dickson et al., 2007). The difference in concentration scales corresponds to a pH error of 0.015 at salinity 35, which is about double the sensor uncertainty (Johnson et al., 2017). Here the practice of Clayton and Byrne (1993), Dickson et al. (2007), and Liu et al. (2011) of using \([H^+]\) (mol kg-seawater\(^{-1}\)) for the proton concentration scale is adopted. This differs from the practice of Dickson (1993) or Waters and Millero (2012) to use \(m_{H^+}\) (mol kg H\(_2\)O\(^{-1}\)). A careful assessment back through the literature ensures that each of the computations outlined below uses the concentration scale appropriate to the original measurements of each thermodynamic property (e.g., Khoo et al., 1977; Dickson, 1993; Millero, 1983)

\(^1\)Note that molinity is not an officially (International Union of Pure and Applied Chemistry/IUPAC) recognized term, but is informally used to specify concentration as mol/kg-solution. These units are frequently used in ocean science and they are the base units used by Clayton and Byrne (1993) and Liu et al. (2011) for spectrophotometric pH measurements used to calibrate the sensors.
and pH measurement methods (Clayton and Byrne, 1993; Liu et al., 2011) used to calibrate the sensor (Johnson et al., 2016).

## 2 Computing pH

Computation of pH from the observed VRS_PH signal requires knowledge of the sensor reference potential at in situ temperature and pressure, the activity coefficient at in situ temperature, pressure, and salinity, and the chloride ion molality, which is estimated from salinity. The procedures used to compute the pH are described below. Seawater temperature (TEMP) is required in all of the calculations. As the sensor is in the pumped stream of the CTD, the CTD temperature (TEMP) is used. The MBARI DURA sensors, in any case, do not have an auxiliary temperature sensor, but the Sea-Bird SEAFET sensors do report an auxiliary pH temperature.

The reference potential of the sensor follows the relationship

\[ k_{T,P} = k_0 + k_2 \times t + f(P) \]  

where \( t \) is the temperature (°C) from the CTD and \( f(P) \) is a polynomial function of pressure. \( P \) has units of bar. The function \( f(P) \) is considered to be temperature independent over the oceanic range (-2 to 35 °C) (Johnson et al., 2016). \( k_{T,P} \) is somewhat equivalent to the Standard Potential of conventional electrochemical cell. However, the \( k_0, k_2, \) and \( f(P) \) values contain terms that are dependent on the FET design, semiconductor processing, mechanical design of the ISFET support, and electronic circuit design. As a result, \( k_{T,P} \) will be different for every device and each sensor reference potential must be individually calibrated as a function of temperature and pressure. However, once determined, these values are stable. So as not to imply that the \( k_{T,P} \) is the same for each sensor, it is referred to as a reference potential and it is designated with a \( k \), rather than the typical symbol \( E^\circ \) for an electrochemical cell. The values for \( k_0, k_2, \) and \( f(P) \) of each sensor are stored in the pH sensor calibration file with units such that the final computed value of \( k_{T,P} \) is in volts.

The mean activity coefficient of HCl in seawater (\( \gamma_{HCl} \)) is a thermodynamic property of seawater that is a function of temperature, salinity, and pressure. The values at various temperatures and atmospheric pressure, \( \gamma_{HCl}^{T,1} \), were reported by Khoo et al. (1977)

\[
\log_{10}(\gamma_{HCl}^{T,1}) = -A \times I^{0.5} / (1 + 1.394 \times I^{0.5}) + (0.08885 - 0.000111 \times t) \times I
\]

where the subscript \( T, 1 \) refers to temperature and a pressure of 1 bar. \( A \) is the Debye-Huckel constant and \( I \) is seawater ionic strength. Ionic strength of seawater (mol kg-seawater\(^{-1}\)) is a function of salinity (Dickson et al., 2007)

\[ I = 19.924 \times S / (1000 - 1.005 \times S). \]  

The Debye-Huckel constant is a function of pure water density and dielectric constant. The values reported in Table 2 of Khoo et al. (1977) were fitted to the empirical function

\[ A = 3.4286 \times 10^{-6} \times t^2 + 6.7524 \times 10^{-4} \times t + 0.49172143. \]  

The activity coefficient is also a function of pressure as required by the Gibbs-Duhem equation

\[
\log_{10}(\gamma_{HCl}^{T,P}) = \log_{10}(\gamma_{HCl}^{T,1}) + \Delta V_{Cl^-} \times P / (RT \ln 10) / 10
\]

where \( P \) is the pressure (bar) and \( \Delta V_{Cl^-} \) is the partial molar volume of HCl in seawater (Millero, 1983)

\[ \Delta V_{Cl^-} \quad (cm^3 \cdot mol^{-1}) = 17.85 + 0.1044 \times t - 0.001316 \times t^2. \]
The factor 10 on the right hand side of equation (10) is required for the combined conversion of units in cm$^3$ to m$^3$ and bar to pascal. Note that, by convention, the proton has a partial molar volume of zero. A compressibility term would normally be added to the partial molar volume, but it has no significant impact at the 2000 dbar maximum depth of a profiling float.

The chloride ion molality required in equation (2) can be computed from salinity (Dickson et al., 2007)

\[
m_{\text{Cl}^-} \text{ (mol kg H}_2\text{O}^{-1}) = \frac{0.99889}{35.453} \times S / 1.80655 / (1 - 0.001005 \times S)
\]

\[
= 0.015596 \times S / (1 - 0.001005 \times S).
\]

(12)

The molality of protons is then computed by rearranging equation (2) and substitution of equations (6) and (10)

\[
-\log_{10}(m_{\text{H}^+}) = (\text{VRS}_{\text{PH}} - k_0 - k_2 \times T - f(P)) \times F/(RT \ln(10)) + \log_{10}(m_{\text{Cl}^-}) + \log_{10}(\gamma_{\text{HCl}})^{1/2} + \Delta V_{\text{Cl}^-} \times P / (RT \ln 10).
\]

(13)

The proton concentration obtained in this manner is on the molality scale, mol per kg-H$_2$O. Seawater pH is reported on the mol per kg-seawater scale. This requires one additional conversion to obtain the pH of seawater on the Free proton scale

\[
\text{PH}_{\text{IN SITU FREE}} = -\log_{10}([\text{H}^+]) = -\log_{10}(m_{\text{H}^+}) - \log_{10}(1 - 0.001005 \times S)
\]

(14)

where [H$^+$] signifies mol kg-seawater$^{-1}$.

### 3 Conversion of pH scales

As noted above, the Deep-Sea DuraFET responds to the free ion concentration, but it is common practice to convert these measurements to the total proton concentration using equation (3). The conversion requires the concentration of bisulfate ion (mol kg-seawater$^{-1}$), which can be computed (Dickson et al., 2007)

\[
S_T = (0.1400 / 96.062) \times S / 1.80655 = 2.3049 \times 10^{-4} \times S.
\]

(15)

The bisulfate dissociation constant has been measured at atmospheric pressure (or summarized) in a number of studies (Khoo et al., 1977; Dickson et al., 1990; Campbell et al., 1993; Waters and Millero, 2012). Here, the equation reported by Dickson et al. (2007) is used for consistency with data sets such as GLODAPv2, but note that other alternatives are available

\[
K_S = \exp(-4276.1 / T + 141.328 - 23.093 \times \ln(T) + (-13856 / T + 771.54 + 114.723 \times \ln(T)) \times T^{-1.5} + 2698 / T \times T^{-1.5} + 1776 / T \times T^{-2} + \ln(1 - 0.001005 \times S)).
\]

(16)

The equilibrium constant is a function of pressure, as well. Correction for pressure requires the partial molar volume for the dissociation reaction. The partial molar volume of HSO$_4^-$ has been measured in pure water, but the value in seawater has not been measured. Its value was estimated by Millero (1983) using the pure water value and the assumption that the effect of sea salt was the same as that found for HCO$_3^-$. The partial molar volume in seawater is then

\[
\Delta V_{\text{HSO}_4} \text{ (cm}^3 \text{ mol}^{-1}) = -18.03 + 0.0466 \times t + 0.000316 \times t^2
\]

(17)
with a compressibility of

$$\kappa_{\text{HSO}_4} = \frac{(-4.53 + 0.09 \times t)}{1000}. \quad (18)$$

These terms are used to compute the ratio of $K_S$ at in situ pressure to the value and atmospheric pressure

$$\ln \left( \frac{K_S P}{K_S 1} \right) = \left( -\Delta V_{\text{HSO}_4} + 0.5 \times \kappa_{\text{HSO}_4} \times P \right) \times P / (R \times T \times 10). \quad (19)$$

The in situ equilibrium constant is then obtained from the ratio and the one atmosphere value. Finally, the pH on the Total Proton scale can be computed as

$$\text{PH} \_\text{IN} \_\text{SITU} \_\text{TOTAL} = \text{PH} \_\text{IN} \_\text{SITU} \_\text{FREE} - \log_{10} (1 + S_T / K_S 1). \quad (20)$$

The total proton scale is used throughout the remainder of this document.

## 4 pH calibration file

The pH calibration sheet (Figure 1) supplies values for $k_0$, $k_2$, and $f(P)$. $f(P)$ is a polynomial in pressure

$$f(P) = f_1 \times P \times 10 + f_2 \times (P \times 10)^2 + f_3 \times (P \times 10)^3 + f_4 \times (P \times 10)^4 + f_5 \times (P \times 10)^5 + f_6 \times (P \times 10)^6. \quad (21)$$

Sea-Bird Scientific typically only uses a fourth order polynomial, while MBARI calibrated sensors use the sixth order polynomial shown here. The $f_0$ coefficient on the Sea-Bird calibration sheet is not used.

Figure 1. Calibration sheet for a Sea-Bird Scientific pH sensor.
5 Sensor metadata

Sensor metadata are summarized in Table 2 and 3.

Table 2. Sensor metadata.

<table>
<thead>
<tr>
<th>SENSOR</th>
<th>TRANSISTOR PH</th>
</tr>
</thead>
<tbody>
<tr>
<td>SENSOR MAKER</td>
<td>MBARI or SBE</td>
</tr>
<tr>
<td>SENSOR MODEL</td>
<td>DURA or SEAFET</td>
</tr>
<tr>
<td>SENSOR SERIAL NO</td>
<td>To be filled</td>
</tr>
</tbody>
</table>

Table 3. Parameter metadata.

<table>
<thead>
<tr>
<th>PARAMETER</th>
<th>VRS PH</th>
</tr>
</thead>
<tbody>
<tr>
<td>PARAMETER UNITS</td>
<td>VOLT</td>
</tr>
<tr>
<td>PARAMETER_ACCURACY</td>
<td>0.000030</td>
</tr>
<tr>
<td>PARAMETER_RESOLUTION</td>
<td>0.00001</td>
</tr>
<tr>
<td>PARAMETER</td>
<td>PH_IN_SITU_TOTAL</td>
</tr>
<tr>
<td>PARAMETER UNITS</td>
<td>dimensionless</td>
</tr>
<tr>
<td>PARAMETER_ACCURACY</td>
<td>0.005</td>
</tr>
<tr>
<td>PARAMETER_RESOLUTION</td>
<td>0.0004</td>
</tr>
<tr>
<td>PARAMETER</td>
<td>PH_IN_SITU_FREE</td>
</tr>
<tr>
<td>PARAMETER UNITS</td>
<td>dimensionless</td>
</tr>
<tr>
<td>PARAMETER_ACCURACY</td>
<td>0.005</td>
</tr>
<tr>
<td>PARAMETER_RESOLUTION</td>
<td>0.0004</td>
</tr>
</tbody>
</table>
6 Real time quality control for pH

The two modes of real-time quality control applied to incoming pH data consist of automated screening to eliminate gross outliers, and data adjustments (drifts and offsets) to account for post-deployment sensor anomalies. Both modes are performed in-line of the data processing stream, although potential adjustment values must first be defined by the data manager after a period of visual inspection. Both modes are described in detail below.

6.1 Common Argo real-time quality control tests on pH vertical profiles

This section lists the real-time tests that are common between CTD data and biogeochemical data. The same real-time test numbers for CTD data are used here. See Argo quality control manual (#RD1, #RD2) The following tests are directly linked to pH, See Argo Quality Control Manual for Biogeochemical Data (#RD2).

6. Global range test

After computing PH_IN_SITU_TOTAL as described above, a simple range check is performed. If values do not fall in the range 7.3 to 8.5, the result is marked bad. These values with a preliminary QC assessment are then reported in the Argo BR profile files.

9. Spike test

The pH sensor can generate occasional spikes due to electrical noise and despiking is appropriate. The default Argo spike tests in core variables (Wong et al., 2018) is

\[
\text{Test value 1} = |V2 - (V3 + V1)/2| - |(V3 - V1)/2| \tag{22}
\]

where V2 is the measurement being tested as a spike, and V1 and V3 are the values above and below. This test does not work well for pH. The criteria depends strongly on the vertical gradient, making it regionally dependent. The test is also sensitive to the vertical resolution of the sensor data, which is platform dependent. We do not recommend this approach.
The spike test recommended for chlorophyll (Schmechtig et al., 2014) is more appropriate for pH. The value tested is

\[ \text{Test value 2} = |V2 - \text{median}(V0, V1, V2, V3, V4)| \] (23)

where the test value represents the anomaly of the observed pH from the median of the surrounding data. Schmechtig et al. (2014) recommend that the criteria used to determine if a data value is a spike be based on the statistical distribution of the 5 data points. Perhaps because pH has a smaller dynamic range than chlorophyll, we find that a constant value of 0.04 pH works well. A data point is considered a spike and marked with quality flag 4 (data bad) if Test value 2 > 0.04 pH.

### 6.2 Specific Argo Real Time quality control for pH vertical profiles

Currently, there is no QC specific test for the pH vertical profile, but if some happen in the future they should be described here and reported with the number “56”

### 6.3 Test application order on pH vertical profiles

The preferred test order is the range check first, then the spike test.

### 6.4 Scientific calibration for each profile

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

<table>
<thead>
<tr>
<th>For PARAMs (B-Argo PARAMs and I-Argo PARAMs) in ‘R’-mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCIENTIFIC_CALIB_COMMENT</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_EQUATION</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_COEFFICIENT</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_DATE</td>
</tr>
</tbody>
</table>

A specific comment should however be set for PRES parameter

<table>
<thead>
<tr>
<th>For PRES</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCIENTIFIC_CALIB_COMMENT</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_EQUATION</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_COEFFICIENT</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_DATE</td>
</tr>
</tbody>
</table>
(see in Chapter 7 and Chapter 8 how to fill scientific calibration information when PARAMETER_DATA_MODE is ‘A’ or ‘D’ respectively).

## 7 Adjustment and quality control for pH in Delayed Mode and Real Time

Data from Argo floats that pass through automatic quality control procedures and which are delivered to to the Argo GDAC, typically within 24 hours, are referred to as Real-Time (RT) data. If the float Principal Investigator applies further corrections to the data, which usually involves visual inspection of data relative to a reference data set, then the data are referred to as Delayed Mode (DM) data. DM corrections for core Argo pressure and salinity data are normally made within 6 to 12 months of collection. The adjustments that are applied to DM data after visual inspection may also be applied to RT data, as they are received, without visual inspection. These data are referred to as Adjusted Mode (AM) (Argo User’s Manual, 2017).

As noted by Johnson et al. (2017), most BGC-Argo data must receive a real-time adjustment to meet the Argo goals of delivering research-quality observations. Applications of uncorrected data are relatively limited, in comparison to the utility of corrected values. We recommend assigning an Argo quality flag of 3 (questionable, probably bad) to unadjusted RT pH data that passes the real-time tests outlined in Section 6.1. The Argo goals for research-quality data then require that the RT sensor data be adjusted in real-time, as noted above, to receive a quality flag of 1 (good data). The procedures for assessing the delayed mode corrections and any subsequent real-time corrections of pH are, therefore, included below in this document.

This requirement to adjust nearly all biogeochemical data to produce research-quality results will require that the DM process be initiated sooner than 6 months. It is therefore advised to perform the first delayed-mode quality control and adjustment after at least 5 cycles. Real-time adjustments then are made to carry these corrections forward.

### 7.1 Adjustment method

Johnson et al. (2017) described the notation for a set of nitrate corrections. We modified this notation slightly to include additional information on the parameter accuracy as it will change in time due to the frequency at which data are examined for drifts or offsets. The notation is the same for pH, nitrate, and oxygen.

The pH correction values are shown in Table 4 for float 9254/WMO # 5904395. The first line, which is optional, includes a correction for pH measurements made with the CTD pump on or off (offset = pH pump on – pH pump off). The offset may be produced by a streaming potential created by flowing seawater. It depends on sensor geometry and electronics, and is only occasionally present. Subsequent lines in the file show the float cycle (j) at which each adjustment is made, a possible gain correction, an offset correction (Oj), a drift correction (Dj, pH/year), and an accuracy estimate (ej). pH does not require a sensor gain correction and all values are set to 1. The gain term is only included for consistency with other parameters. For pH, the accuracy estimate is given in pH units on the total scale.

The net correction at each node (\(\Delta pH_j\)) is computed as

<table>
<thead>
<tr>
<th>Table 4. pH corrections for float 9254/WMO 5904395.</th>
</tr>
</thead>
<tbody>
<tr>
<td>9254SOOCN 05/23/17 15:06 pH, offset, 0.000</td>
</tr>
<tr>
<td>pH, 1.1,-0.003,0.000, 0.010</td>
</tr>
<tr>
<td>pH, 2.1,0.031,0.125, 0.010</td>
</tr>
<tr>
<td>pH, 9.1,-0.002,-0.027, 0.010</td>
</tr>
<tr>
<td>pH, 55.1,-0.000,0.000, 0.010</td>
</tr>
<tr>
<td>pH, 148.1,-0.001,0.000, 0.010</td>
</tr>
</tbody>
</table>
\[ \Delta p_{H_j} = \Delta p_{H_{j-1}} + O_j + D_{j-1} (T_j - T_{j-1}) \]  

(24)

The correction that is then applied between nodes at cycle (i), where \( i \geq j \) and \( i < j + 1 \) cycle, is computed as

\[ \Delta p_{H_i} = \Delta p_{H_j} + D_j (T_i - T_j) \]  

(25)

\[ \text{PH\_IN\_SITU\_TOTAL\_ADJUSTED} = \text{PH\_IN\_SITU\_TOTAL} + \Delta p_{H_i} \]  

(26)

Table 5 gives an example of thus obtained cumulative pH adjustments for float 5904395 (compare Table 4).

<table>
<thead>
<tr>
<th>Date</th>
<th>Cycle</th>
<th>Offset</th>
<th>Drift</th>
<th>Cumulative pH Adjustment</th>
<th>Raw - Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/21/2014</td>
<td>1</td>
<td>-0.003</td>
<td>0.000</td>
<td>-0.0030</td>
<td></td>
</tr>
<tr>
<td>4/27/2014</td>
<td>2</td>
<td>0.031</td>
<td>0.125</td>
<td>0.0280</td>
<td></td>
</tr>
<tr>
<td>6/3/2014</td>
<td>9</td>
<td>-0.002</td>
<td>-0.027</td>
<td>0.0387</td>
<td></td>
</tr>
<tr>
<td>2/2/2015</td>
<td>55</td>
<td>0.000</td>
<td>0.000</td>
<td>0.0206</td>
<td></td>
</tr>
<tr>
<td>5/15/2017</td>
<td>148</td>
<td>-0.001</td>
<td>0.000</td>
<td>0.0196</td>
<td></td>
</tr>
</tbody>
</table>

Note that adjustments in real time will essentially be past the last line (table 5) of the corrections obtained from delayed-mode, i.e., an extrapolation. To account for this, we propose to add 0.020 pH units per year since the last profile with delayed mode correction was made to the accuracy estimate:

\[ <\text{PARAM}>\_\text{ADJUSTED\_ERROR} = e_{\text{last}} + 0.020 \text{ year}^{-1} \times (T - T_{\text{last}}) \]  

(27)

where \( T \) is the time of the profile and \( T_{\text{last}} \) the time of the last profile with DM QC and adjustment, and \( e_{\text{last}} \) the last accuracy estimate of the DM correction (Table 4, last line).

### 7.2 Specific real time quality control for adjusted pH vertical profiles

pH sensors also return sensor diagnostic data. IB\_PH is the base current of the ISFET pH chip and IK\_PH is the counter electrode current. These are Argo intermediate parameters. A range check is performed on IB\_PH and IK\_PH (< +/- 150 nanoamps). If the currents exceed this range, then the QC value for PH\_IN\_SITU\_TOTAL\_ADJUSTED is set to 3 because the sensor is performing poorly, unless the QC value is already greater than this.
7.3 Scientific calibration for each profile

When a biogeochemical parameter (‘b’ parameter) 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 (‘i’ parameters) associated to this adjusted biogeochemical parameter are also set to ‘A’ when they have an ADJUSTED field (but otherwise left at ‘R’).

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 #RD2, for I-Argo <PARAM>, while <PARAM>_QC and PROFILE_<PARAM>_QC are mandatory, <PARAM>_ADJUSTED, <PARAM>_ADJUSTED_QC and <PARAM>_ADJUSTED_ERROR are optional (DAC decision).

<table>
<thead>
<tr>
<th>For I-Argo PARAMs with no corresponding _ADJUSTED field and for which the associated B-Argo PARAMs have been through adjustment in real-time</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCIENTIFIC_CALIB_COMMENT</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_EQUATION</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_COEFFICIENT</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_DATE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>For I-Argo PARAMs with corresponding _ADJUSTED fields and for which the associated B-Argo PARAMs have been through adjustment in real-time</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCIENTIFIC_CALIB_COMMENT</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_EQUATION</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_COEFFICIENT</td>
</tr>
</tbody>
</table>

As mentioned in #RD2, for I-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 B-Argo PARAMs that have been through adjustment in Real-Time

<table>
<thead>
<tr>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCIENTIFIC_CALIB COMMENT</td>
<td>Content depends on method. See #LIR, #MLR, #CANYON.</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB EQUATION</td>
<td>PH_IN_SITU_TOTAL_ADJUSTED = PH_IN_SITU_TOTAL + DeltapH_i</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB COEFFICIENT</td>
<td>DeltapH_i = Cumulative pH Adjustment ΔpH_i</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_DATE</td>
<td>YYYYMMDDHHMISS(*)</td>
</tr>
</tbody>
</table>

(*): for a given calibration, the SCIENTIFIC_CALIB_DATE of an adjusted B-Argo parameter and of its associated I-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 single 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’.
8 Quality control for data adjusted in delayed mode

8.1 Data adjustment process

This process will populate the PH_IN_SITU_TOTAL_ADJUSTED variable. Similar to Argo salinity adjustments (Owens and Wong, 2009), the pH adjustment process depends on having an accurate model for pH in waters below 1000 m depth, where temporal and spatial variability is minimal. Additionally, one should note that the accuracy and stability of the pH sensor is rapidly improving in time (Johnson et al., 2017). This does not yet remove the need for the QC assessment.

The background for the QC/adjustment process for pH is described in Johnson et al. (2016; 2017). Offsets or drifts in pH data appear to result from changes in the sensor $k_0$ constant. These changes in $k_0$ are assessed by comparing the in situ pH values observed at a depth range near the maximum depth that is consistently reached by a float with the values predicted at that depth using a variety of estimation methods. Three estimation methods are in use (Table 6). These include Multiple Linear Regression (MLR) equations tuned for specific regions (Williams et al., 2016), Locally Interpolated Regression (LIR) methods (Carter et al., 2018), which extend globally by encapsulating MLR equations in a geographically varying framework, and a neural network prediction system known as CANYON (Sauzéde et al., 2017), which also operates globally. All of these prediction methods utilize oxygen as an input variable for computing predicted pH. Therefore, it is strongly recommended to always deploy both pH and oxygen sensors together. Without an oxygen measurement, the correction schemes for pH are less robust.

The prediction systems in Table 6 were fitted to in situ pH values that were derived from shipboard measurements and then converted to in situ values using the CO2SYS software package (Lewis and Wallace, 1998; van Heuven et al., 2011). Careful attention must be paid to the underlying shipboard data that are used to produce the MLR, LIR and CANYON systems. Shipboard measurements of pH have evolved over time and there are some systematic differences in the results (Carter et al., 2018). The most accurate shipboard pH measurements are considered to be spectrophotometric determinations made using purified indicator dye (Liu et al., 2011). Measurements made using unpurified dye (generally before 2012) are systematically offset by small amounts that depend on each dye lot. The largest body of pH data comes from shipboard measurements of dissolved inorganic carbon (DIC) and seawater alkalinity ($A_T$), which are then used to compute pH with known seawater thermodynamics (Lewis and Wallace, 1998). These estimated pH values have a distinct, pH dependent offset from directly measured values that are made with either purified or raw dye (Carter et al., 2013, 2018; Williams et al., 2017). Equations that are fitted or trained with data using pH measurements by these various methods will differ in their predictions due to the underlying, systematic uncertainties in the measurements.

Here, we follow the general recommendation of Carter et al. (2018) to use pH estimates that are consistent with spectrophotometric measurements made using purified dye. We make this choice for the following reasons. First, the pH sensors are calibrated in the laboratory using spectrophotometric measurements with purified dyes. While the current generation of sensors often require in situ adjustments, which effectively shifts the calibration to the scheme used for in situ predictions, we believe drift problems will be resolved. At that point, sensor measurements will be directly related to the laboratory calibration method. The simplest, most precise reference measurement method for pH in the laboratory is direct measurement by spectrophotometric methods using purified dye. It will always be the preferred method for laboratory calibration when compared to the complexity and lower precision of pH based on DIC and $A_T$ measurements. To avoid any future need to change reference methods back to values consistent with purified dye determinations, we suggest that all adjustments be based on reference predictions that are based on observations that are consistent with purified dye measurements.
The MLR method, developed by Williams et al. (2016; 2017), used only recent pH measurements from two GO-SHIP cruises that utilized purified dye. There are two versions of the algorithm, specific to the Southern Ocean (Table 4). Algorithm 1 has been developed and tested for all regions of the Southern Ocean south of 50 degrees at depths from 1000 to 2100 m. Algorithm 2 was developed and tested for the Pacific basin in the latitude range 30° to 50°S and a similar depth range. These algorithms were intended primarily for float sensor validation.

The LIR method was developed using the GLODAPv2 (Olsen et al., 2016) data set, as well as a set of GO-SHIP cruises that were performed subsequent to the GLODAPv2 release (Carter et al., 2018). The GLODAPv2 pH data set is based primarily on values derived from DIC and AT measurements. Carter et al. (2018) adjusted these pH values, as well as measurements made with unpurified indicator dye to remove offsets from pH measured with purified dye. Note that the LIR algorithms have a flag to remove these corrections. The default setting (all pH values corrected for compatibility with purified dye) is used here for reasons discussed above. The LIR and MLR pH prediction in the Southern Ocean appear consistent to within 0.005 pH (Carter et al., 2018). Comparisons of the LIR predictions with the Hawaii Ocean Time-series pH values, which are not in the GLODAPv2 data set also indicate excellent agreement. Using a global test data set, Carter et al. (2018) report an uncertainty of 0.001±0.006 for LIR pH predictions at depths near 2000 m.

The CANYON method was also developed using the GLODAPv2 data set. The current version of CANYON (v2) does not attempt to reconcile the various pH measurements with purified dye observations. As a result, predictions with CANYON seem to be biased high by about 0.015 pH relative to MLR or LIR predictions at the same locations. It is expected that updated versions of CANYON will address this bias.

The expected error in float pH measurements is derived from the uncertainty in the reference data as well as sensor uncertainties. The expected error reported by Carter et al. (2018) for pH values predicted with LIR from depth, temperature, salinity and oxygen (Apparent Oxygen Utilization) at depths near 2000 m is 0.001±0.006, where the error terms are a bias and the root mean square error. Additional errors in float sensor pH may arise (Williams et al., 2017) because of uncertainty in the sensor pressure and temperature coefficients (0.0025 pH), and uncertainty in the carbon system thermodynamics used to convert laboratory pH values to in situ values (0.005). The pH sensor precision is near 0.003 (Williams et al., 2017). A combined error budget that recognizes both systematic and random errors suggests that sensor pH values corrected with the MLR or LIR equations to have an accuracy near 0.008. A conservative value of 0.01 for the error of corrected pH data is used in Table 4, unless the operator has reason to believe values are different. The error is roughly equivalent to 1 standard deviation. These error limits are substantiated by a comparison of profiling float pH sensor data with shipboard measurements made near the time of float deployment and then converted to in situ values (Johnson et al., 2017). The comparisons have a bias of 0.005 and a standard deviation of 0.007. If a float does not have an oxygen sensor, or the oxygen sensor has failed,
the MLR predictions have a larger error (Carter et al., 2018). In this case, the bias may be approximately doubled.

Finally, note that ocean pH is slowly decreasing due to the increase in atmospheric carbon dioxide. While this signal is largest near the surface, there are indications that it has penetrated to significant depths in the Southern Ocean and North Atlantic (Rios et al., 2015). Both the MLR and CANYON tools provide some capability to incorporate the ocean acidification signal. However, these should be used cautiously. Continued, high quality ship-based observations will be required to sustain confidence in the deep, reference pH values.

8.2 pH data adjustment software and example

The pH sensor correction process is done at MBARI using a software tool called SAGE (SOCCOM Assessment and Graphical Evaluation) that was developed to support the SOCCOM research project (http://soccom.princeton.edu). This tool allows the operator to select any of the prediction schemes described above as the metric against which to QC/adjust a float data set. As oxygen is an input to the prediction schemes, it must be corrected first as oxygen optodes are known to be affected by storage drift, resulting in underestimates of oxygen (Johnson et al., 2015). This underestimation is corrected when the sensor data is corrected for a gain error. A software tool called SAGEO2 is used for these corrections (Maurer et al., 2018).

The SAGE software allows the user to define a number of nodes where it appears that either an offset has occurred in the $k_0$ value or the $k_0$ has begun a linear drift in time, relative to the reference predictions (MLR, LIR or CANYON). The offsets and slopes at each of these nodes can be computed automatically using SAGE. The computed offsets and slopes are the values that minimize the sum of square difference between the observed pH and the reference pH. Values for offsets and slopes can also be entered manually. It is our general policy to add a final adjustment to the last cycle that is available, at the time a float is adjusted. The significance of adding each new node can be assessed by comparing the Akaike Information Criteria (AIC) statistic (Akaike, 1974) that is reported for each model. This is similar to the approach developed by Owens and Wong (200xx). As a float accumulates more cycles, adjustments may be modified.

The notation of the corrections is described in Section 7.1 Adjustment Method. Figure 2 shows the raw pH values and the adjusted pH values at 1500 dbar depth and at the 10 dbar for the delayed mode adjustment (i.e., all float cycles that are present). The float was launched at 40°S 150°W and has remained in that general vicinity. The MLR equation tuned for the Pacific Basin from 30°S to 60°S has, therefore, been selected as the reference data for this float.
Several screen grabs from the SAGE tool are shown in Figures 3 to 5 for float 9254. Figure 3 shows the unadjusted (raw) float data for PH_IN_SITU_TOTAL in the pressure range from 1480 to 1520 dbar. Inspection of the raw data suggests significant changes in the raw pH values, relative to the reference data at cycles 1, 2, 9, and 55. The last node (cycle 148) is manually set to have a slope of zero to avoid introducing artificial drift into the data. Note that 9254 is an early SOCCOM float. Significant changes to pH sensor preparation, outlined in Johnson et al. (2016; 2017), have substantially eliminated the drift seen in the first 6 months of sensor deployment.

Figure 4 shows a screen grab from SAGE of the adjusted data compared to the MLR. An updated AIC value is reported each time a new set of corrections is entered. If the operator is satisfied with the corrections, the values are saved in a .mat file which is used to create the Argo B profile file. The adjusted pH values are used to populate the PH_IN_SITU_TOTAL_ADJUSTED parameter.

To illustrate the sensitivity to different choices for the reference data set, Figure 5 shows the comparison to the LIR reference. There is a mean difference between the float data, which had been adjusted to the MLR, and the LIR prediction of 0.002 pH. This is a very acceptable agreement.
Figure 3  Screen grab from SAGE for float 9254/WMO 5904395 showing raw pH data near 1500 dbar. The MLR equation tuned for the Pacific Basin from 30°S to 60°S has been selected as the reference data for this float.
Figure 4 Screen grab from SAGE for float 9254/WMO 5904395 showing adjusted pH data near 1500 dbar. The MLR equation tuned for the Pacific Basin from 30°S to 60°S has been selected as the reference data for this float.
8.3 Editing raw QC and adjusted QC 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.

8.4 Compulsory variables to be filled in a BD profile file

8.4.1 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 recomputed when <PARAM>_ADJUSTED_QC becomes available.

For I-Argo <PARAM>, <PARAM>_QC and PROFILE_<PARAM>_QC are mandatory, but the 3 adjusted variables are optional in the B- profile file:

<PARAM>_ADJUSTED, <PARAM>_ADJUSTED_QC, <PARAM>_ADJUSTED_ERROR.

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

Note that PRES in the B- profile file does not carry any qc or adjusted variables. It is used as a stand-alone vertical index that links the core- and b- files. Users who want delayed-mode adjusted pressure values (PRES_ADJUSTED) should obtain them from the core- files.
8.4.2 Scientific calibration 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 let 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 indications on how to fill the scientific calibration section of a BD profile file.

<table>
<thead>
<tr>
<th>For I-Argo PARAMs with no corresponding _ADJUSTED field and for which the associated B-Argo PARAMs have been through delayed-mode qc</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCIENTIFIC_CALIB_COMMENT</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_EQUATION</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_COEFFICIENT</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_DATE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>For I-Argo PARAMs with corresponding _ADJUSTED fields and for which the associated B-Argo PARAMs have been through delayed-mode qc</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCIENTIFIC_CALIB_COMMENT</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_EQUATION</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_COEFFICIENT</td>
</tr>
<tr>
<td>SCIENTIFIC_CALIB_DATE</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>For PARAMs that have been through delayed-mode qc</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCIENTIFIC_CALIB_COMMENT</td>
</tr>
</tbody>
</table>
SCIENTIFIC_CALIB_EQUATION | PH_IN_SITU_TOTAL_ADJUSTED = PH_IN_SITU_TOTAL + ΔpH_i
--- | ---
SCIENTIFIC_CALIB_COEFFICIENT | ΔpH_i = Cumulative pH Adjustment ΔpH_i
SCIENTIFIC_CALIB_DATE | YYYYMMDDHHMISS(*)

(*) for a given calibration, the SCIENTIFIC_CALIB_DATE of an adjusted B-Argo parameter and of its associated I-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 single 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’.

### 8.4.3 Other variables in a BD profile file

Here are other variables in a B- 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 adjusted values for one or more <PARAM> in each N_PROF become available.

In b-Argo profile files, there are additional biogeochemical parameters which 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 their 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.

8.4.4 Profiles naming convention

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

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.

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

9 Matlab code to compute pH

The function phcalc was used to produce the pH data shown in Table 6.

function [phfree,phtot]= phcalc(Vrs, Press, Temp, Salt, k0, k2, Pcoefs)
%
% INPUTS:
%   Vrs     = Voltage between reference electrode and ISFET source
%   Press   = Pressures in decibars
%   Temp    = Temperature in degrees C
%   Salt    = Salinity (usually CTD salinity on the PSS)
%   k0      = Sensor reference potential (intercept at Temp = 0C)
%   k2      = linear temperature coefficient (slope)
%   Pcoefs  = sensor dependent pressure coefficients
%
% ************************************************************************
%  SET SOME CONSTANTS
% ************************************************************************
% Universal gas constant, (R), http://physics.nist.gov/cgi-bin/cuu/Value?r
R    = 8.31446; % J/(mol K)
F    = 96485; % Faraday constant Coulomb / mol
Tk   = 273.15 + Temp; % degrees Kelvin
ln10 = log(10); % natural log of 10
%
% ************************************************************************
% CALCULATE PHYSICAL AND THERMODYNAMIC DATA
% practices for ocean CO2 measurements.
% ************************************************************************
% IONIC STRENGTH OF SEAWATER (mol / kg H2O)
% Varified units by comparing to Dickson et al. 2007: Chap 5, p10 Table 2
% Dickson et al. 2007: Chap 5, p13 Eq 34  
IonS = 19.924 .* Salt ./ (1000 - 1.005 * Salt);

% MEAN SEAWATER SULFATE CONCENTRATION (mol / kg solution)  
% This wants to be mol/kg-seawater as KHSO4 is on that scale  
% Dickson et al. 2007: Chap 5, p10 Table 2  
Stotal = (0.14 / 96.062).* (Salt / 1.80655);

% MEAN SEAWATER CHLORIDE CONCENTRATION (mol / kg H2O)  
% this wants to be mol/kg H2O as activity is on mol/kg H2O scale  
% Dickson et al. 2007: Chap 5, p10 Table 2  
Cltotal = 0.99889 / 35.453 .* Salt / 1.80655;%(mol / kg solution)  
Cltotal = Cltotal ./ (1 - 0.001005 .* Salt); % (mol / kg H2O)

% BISULFIDE DISSCIATION CONSTANT AT T,S AND IONIC STRENGTH(mol/kg solution)  
% Dickson et al. 2007: Chap 5, p12 Eq 33  
Khs04 = exp((-4276.1 ./ Tk + 141.328 - 23.093 .* log(Tk)) + ...  
(-13856 ./ Tk + 324.57 - 47.986 .* log(Tk)) .* IonS .* 0.5 + ...  
(35474 ./ Tk - 771.54 + 114.723 .* log(Tk)) .* IonS - ...  
2698 ./ Tk .* IonS .* 1.5 + 1776 ./ Tk .* IonS .* 2 + ...  
log(1 - 0.001005 .* Salt));

% Millero 1983 Chemical Oceanography vol 8  
% partial molar volume and compressibility of HSO4 in seawater.  
deltaVHso4 = -18.03 + 0.0466 .* Temp + 0.000316 .* Temp .* 2;  
KappaHso4 = (-4.53 + 0.09 .* Temp) / 1000;

%%%%%% Press changed from dbar to bar here by / 10  
lnKhso4fac = (-deltaVHso4 + 0.5 .* KappaHso4 .* (Press / 10)) .* ...  
(Press / 10) ./ (R .* Tk);

% bisulfate association constant at T, S, P  
Khs04TPS = Khs04 .* exp(lnKhso4fac);

% GAMMA +/- HCl, activity coefficient of HCl at T/S, P=1  
% ADH is the Debye Huckel constant, calculated as a polynomial  
% fit to data in Khoo et al. 1977, doi:10.1021/ac50009a016  
% See Martz et al. 2010, DOI 10.4319/lom.2010.8.172, p175  
% Typo in paper 2nd term should be e-4 not e-6  
%  
% ADH = (3.4286e-6 .* Temp .* 2 + 6.7524e-4 .* Temp + 0.49172143);  

log10gammaHCl = -ADH .* sqrt(IonS) ./ (1 + 1.394 .* sqrt(IonS)) + ...  
(0.08885 - 0.000111 .* Temp) .* IonS;

% Millero 1983 partial molar volume of HCl in seawater  
deltaVHCl = 17.85 + 0.1044 .* Temp - 0.001316 .* Temp .* 2;

% effect of pressure on activity coefficient of HCl, divide by 2 because  
% its a mean activity coefficient, divide by 10 for units in the cm3 to F  
% conversion.  
log10gammaHCLtP = log10gammaHCl + deltaVHCl.*(Press./10)./(R.*Tk.*ln10)./2./10;

% Sensor reference potential
% 30
Argo data management Processing BGC-Argo pH data at the DAC level

% ************************************************************************
% k0T = k0 + k2 * Temp; % Temp in deg C
% CALCULATE PRESSURE CORRECTION (POLYNOMIAL FUNCTION OF PRESSURE)
% ALL SENSORS HAVE A PRESSURE RESPONSE WHICH IS DETERMINED IN THE LAB
% AND CONTAINED IN THE POLYNOMIAL Pcoeffs
pc = [flipud(Pcoeffs);0]; % Matlab wants descending powers & n+1 (add 0)
pcorr = polyval(pc,Press);
k0TP = k0T + pcorr;

% pH on free scale then corrected to get to pH total on mol/kg-seawater scale
% pHinSituFree = (Vrs - k0TP) / (R * Tk / F * ln10) + ...
% log(Cltotal) / ln10 + 2 * log10gammaHCltP
% this will be mol kg H2O need to convert to mol/kg-seawater
phfree = (Vrs - k0TP) ./ (R .* Tk ./ F .* ln10) + ...
  log(Cltotal) ./ ln10 + 2 * log10gammaHCltP; %mol/kg-H2O scale

% CONVERT TO mol/kg-seawater scale - JP 2/4/16
phfree = phfree - log10(1 - 0.001005 .* Salt); %mol/kg-seawater scale

% convert to total proton scale
phtot = phfree - log10(1 + Stotal ./ Khso4TPS);

% 9.1 Check Values

Table 7 contains a test data set that was computed with the Matlab function pHcalc.m (Section 8). It lists the sensor calibration coefficients and a vertical profile of PRESS/TEMP/PSAL values, the sensor VRS_PH, and the computed pH values at in situ conditions on the free and total proton scales (PH_IN_SITU_FREE, PH_IN_SITU_TOTAL).
Table 7. Check data for the calculation of pH. This data is from float 9254/WMO 5904395 and is a subsample of the values collected on cycle 8. Note that f(P) was fit to a 5th order polynomial.

### Calibration Data

<table>
<thead>
<tr>
<th>k0</th>
<th>k2</th>
<th>f1</th>
<th>f2</th>
<th>f3</th>
<th>f4</th>
<th>f5</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.372714</td>
<td>-0.001063</td>
<td>4.15927E-06</td>
<td>-2.53737E-08</td>
<td>2.87941E-11</td>
<td>-1.33263E-14</td>
<td>2.23781E-18</td>
</tr>
</tbody>
</table>

### Profile Data

<table>
<thead>
<tr>
<th>PRES</th>
<th>TEMP</th>
<th>PSAL</th>
<th>VRS_PH</th>
<th>PH_IN_SITU_FREE</th>
<th>PH_IN_SITU_TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>1598.78</td>
<td>2.9225</td>
<td>34.5006</td>
<td>-0.922691</td>
<td>7.9088</td>
<td>7.8673</td>
</tr>
<tr>
<td>1498.61</td>
<td>3.1302</td>
<td>34.4472</td>
<td>-0.922110</td>
<td>7.9112</td>
<td>7.8690</td>
</tr>
<tr>
<td>1398.77</td>
<td>3.3575</td>
<td>34.4010</td>
<td>-0.921226</td>
<td>7.9190</td>
<td>7.8761</td>
</tr>
<tr>
<td>1299.25</td>
<td>3.6987</td>
<td>34.3625</td>
<td>-0.920056</td>
<td>7.9307</td>
<td>7.8869</td>
</tr>
<tr>
<td>1198.84</td>
<td>4.2011</td>
<td>34.3332</td>
<td>-0.918651</td>
<td>7.9446</td>
<td>7.8994</td>
</tr>
<tr>
<td>1099.64</td>
<td>4.7072</td>
<td>34.3171</td>
<td>-0.917050</td>
<td>7.9614</td>
<td>7.9149</td>
</tr>
<tr>
<td>999.36</td>
<td>5.2709</td>
<td>34.3106</td>
<td>-0.915267</td>
<td>7.9800</td>
<td>7.9320</td>
</tr>
<tr>
<td>951.66</td>
<td>5.5806</td>
<td>34.3132</td>
<td>-0.914260</td>
<td>7.9908</td>
<td>7.9420</td>
</tr>
<tr>
<td>900.56</td>
<td>5.8022</td>
<td>34.3130</td>
<td>-0.913091</td>
<td>8.0049</td>
<td>7.9554</td>
</tr>
<tr>
<td>850.87</td>
<td>6.0924</td>
<td>34.3179</td>
<td>-0.912048</td>
<td>8.0157</td>
<td>7.9654</td>
</tr>
<tr>
<td>801.08</td>
<td>6.3237</td>
<td>34.3277</td>
<td>-0.911117</td>
<td>8.0249</td>
<td>7.9740</td>
</tr>
<tr>
<td>749.16</td>
<td>6.5121</td>
<td>34.3376</td>
<td>-0.910301</td>
<td>8.0319</td>
<td>7.9804</td>
</tr>
<tr>
<td>701.59</td>
<td>6.7193</td>
<td>34.3521</td>
<td>-0.909649</td>
<td>8.0359</td>
<td>7.9838</td>
</tr>
<tr>
<td>650.81</td>
<td>6.8914</td>
<td>34.3651</td>
<td>-0.908942</td>
<td>8.0406</td>
<td>7.9879</td>
</tr>
<tr>
<td>601.52</td>
<td>7.0795</td>
<td>34.3806</td>
<td>-0.908317</td>
<td>8.0436</td>
<td>7.9903</td>
</tr>
<tr>
<td>550.92</td>
<td>7.2701</td>
<td>34.3966</td>
<td>-0.907575</td>
<td>8.0483</td>
<td>7.9943</td>
</tr>
<tr>
<td>501.11</td>
<td>7.4840</td>
<td>34.4152</td>
<td>-0.906749</td>
<td>8.0542</td>
<td>7.9995</td>
</tr>
<tr>
<td>51.06</td>
<td>15.5472</td>
<td>34.5643</td>
<td>-0.892025</td>
<td>8.1754</td>
<td>8.0980</td>
</tr>
<tr>
<td>46.07</td>
<td>15.5544</td>
<td>34.5642</td>
<td>-0.891988</td>
<td>8.1759</td>
<td>8.0985</td>
</tr>
<tr>
<td>41.10</td>
<td>15.5561</td>
<td>34.5638</td>
<td>-0.891948</td>
<td>8.1766</td>
<td>8.0991</td>
</tr>
</tbody>
</table>
### 10 References


Dickson, A. G. 1990. Standard potential of the reaction: AgCl(s) + ½ H2(g) = Ag(s) + HCl(aq), and the standard acidity constant of the ion HSO4- in synthetic sea water from 273.15 to 318.15 K. Journal of Chemical Thermodynamics, 22, 113-127.


Takeshita, Yuichiro, Todd R. Martz, Kenneth S. Johnson and Andrew G. Dickson. 2014. Characterization of an Ion Sensitive Field Effect Transistor and chloride ion selective electrodes for pH measurements in seawater. Analytical Chemistry, dx.doi.org/10.1021/ac502631z.


