Delayed-mode quality control of oxygen, nitrate and pH data on SOCCOM biogeochemical profiling floats

Tanya Lea Maurer[1], Josh N. Plant[1], and Kenneth S. Johnson[2]

[1]MBARI
[2]Monterey Bay Aquarium Res. Inst.

November 22, 2022

Abstract

The Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) project has deployed 194 profiling floats equipped with biogeochemical (BGC) sensors, making it one of the largest contributors to global BGC-Argo. Post-deployment quality control of float-based oxygen, nitrate, and pH data is a crucial step in the processing and dissemination of such data, as in-situ chemical sensors remain in early stages of development. In-situ calibration of chemical sensors on profiling floats using atmospheric reanalysis and empirical algorithms have been shown to bring accuracy to within 3 µmol O2 kg-1, 0.007 pH units, and 0.5 µmol NO3- kg-1. Routine quality control efforts utilizing these methods can be conducted manually through visual inspection of data to assess sensor drifts and offsets, but more automated processes are preferred to support the growing number of BGC floats and reduce subjectivity among delayed-mode operators. Here we present a methodology and accompanying software designed to easily visualize float data against select reference datasets and assess quality control adjustments within a quantitative framework. The software is intended for global use and has been used successfully in the post-deployment calibration and quality control of over 250 BGC floats, including all within the SOCCOM array. Results from validation of the proposed methodology are also presented which can provide a metric for tracking data adjustment quality through time.

Delayed-mode quality control of oxygen, nitrate and pH data on SOCCOM biogeochemical profiling floats

Tanya L. Maurer[1], Joshua N. Plant[2], and Kenneth S. Johnson[3]

[1, 2, 3] Monterey Bay Aquarium Research Institute, 7700 Sandholdt Road, Moss Landing, CA 95039.

Key Points:

- A methodology and related software for visualizing biogeochemical data against references aids in correcting data for shifts in calibration
- Described methods bring data accuracies to within the range required for climate studies and remain applicable over the lifetime of a float
- A standardized approach to quality control supports cross-platform data management and routine monitoring of fleet-wide sensor performance

Abstract

The Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) project has deployed 194 profiling floats equipped with biogeochemical (BGC) sensors, making it one of the largest contributors to global BGC-Argo. Post-deployment quality control of float-based oxygen, nitrate, and pH data is a crucial step in the processing and dissemination of such data, as in-situ chemical sensors remain in early stages of development. In-situ calibration of chemical sensors on profiling floats using atmospheric reanalysis and empirical algorithms have been shown to bring accuracy to within 3 µmol O_2 kg^{-1}, 0.007 pH units, and 0.5 µmol NO_3^- kg^{-1}. Routine quality control efforts utilizing these methods can be conducted manually through visual inspection of data to assess sensor drifts and offsets, but more automated processes are preferred to support the growing number of BGC floats and reduce subjectivity among delayed-mode operators. Here we present a methodology and accompanying software designed to easily visualize float data against select reference datasets and assess quality control adjustments within a quantitative framework. The software is intended for global use and has been used successfully in the post-deployment calibration and quality control of over 250 BGC floats, including all within the SOCCOM array. Results from validation of the proposed methodology are also presented which can provide a metric for tracking data adjustment quality through time.

Plain Language Summary

The amount of chemical oceanographic data available to researchers is rapidly increasing thanks to robotic drifting floats such as those deployed through the Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) project. Because these floats live the entirety of their life at sea, ensuring that the sensors are working as expected and that the quality of the data returned is fit for scientific use must be done remotely. This paper describes the approaches and accompanying software used to assess performance of oxygen, nitrate and pH sensors on profiling floats and correct for any shifts in sensor performance through the life of the float. An independent validation of the proposed methods is also presented which provides an added level of confidence to the described methods and overall quality of the dataset.

1 Introduction

The Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) project has finished its sixth year reaching a total of 194 biogeochemical (BGC)-Argo profiling floats deployed throughout the Southern Ocean (Fig. 1). Funded by the US National Science Foundation (NSF) Office of Polar Programs, this novel basin-scale network of biogeochemical sensors has filled one of the largest observational gaps in the global ocean. Due to the success of the current program, the SOCCOM project has been renewed for an additional four years, with the goal of deploying 120 more BGC profiling floats south of 30S. In addition, the NSF has funded the Global Ocean Biogeochemistry (GO-BGC) Array, which will extend the current BGC-Argo program considerably through the deployment of an additional 500 floats throughout the global ocean. Emerging data from floats within the SOCCOM array have already expanded our understanding of the Southern Ocean's role in the global carbon cycle and have improved the capability of ocean models to predict future change (Bushinsky et al., 2019a; Gray et al., 2018; Russell et al., 2018; Swart et al., 2019; Verdy & Mazloff, 2017; Williams et al., 2018). Key to these advancements has been the underlying quality of the supporting dataset which relies on pre-deployment sensor calibration and post-deployment quality control (QC), bringing sensor accuracies to within the narrow range required for climate studies (Johnson et al., 2017).

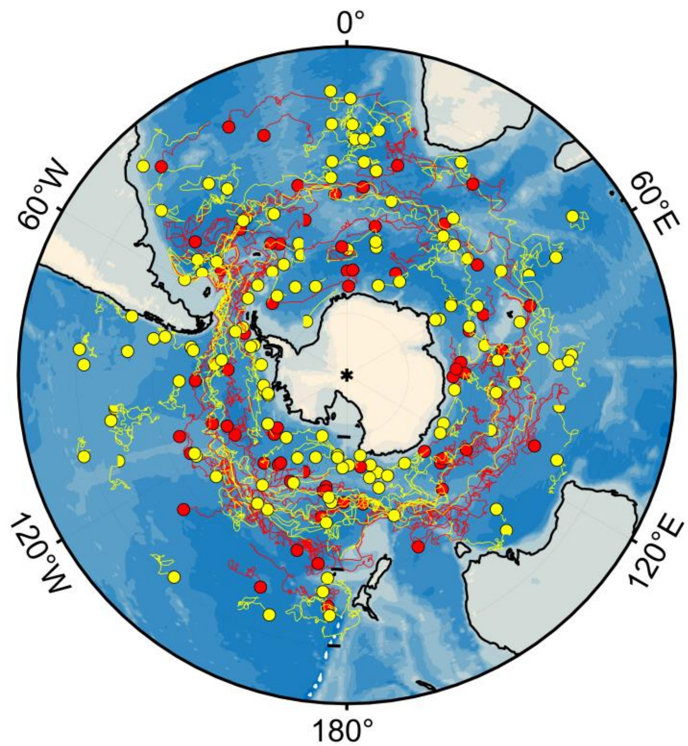

Figure 1. Current location (circles) and associated trajectories (lines) from floats within the SOCCOM array, as of December, 2020. Both operational (yellow) and inactive (red) floats are shown. Historical data from inactive floats remains a valuable part of the SOCCOM dataset.

Operational procedures for post-deployment processing of CTD data from the Argo array are well established. A number of real-time checks constitute the first level of quality control, many of which have been adopted for BGC data as well (Schmechtig et al., 2016). Salinity profiles from Argo floats are also subject to various delayed-mode assessments that typically apply interpolation methods to relate float data to a climatology (Gaillard et al., 2009; Guinehut et al., 2009; Owens & Wong, 2009; Wong et al., 2003). Argo salinity data have been estimated to be accurate to 0.01 PSU, after delayed-mode adjustments, and temperature and pressure data are generally thought of as acceptable for use in data assimilation and other direct applications prior to receiving any delayed-mode assessment (Wong et al., 2020).

In contrast, in situ chemical sensors for measuring oxygen, nitrate and pH on BGC-Argo floats represent newer technologies that require significantly more quality control. Generally, the scientific use of raw, unadjusted BGC-Argo float data is not recommended. The real-time and delayed-mode adjustment processes greatly improve the quality of the BGC sensor data and result in a data set that is suited for research in a variety of applications. Various delayed-mode methods for BGC sensor recalibration and quality control for oxygen, pH and nitrate have been suggested (Bittig et al., 2018a; Johnson et al. 2013; 2015; 2017; Takeshita et al., 2013; Williams et al., 2016) but integrating the suite of methodologies into a coherent framework that can be used operationally across a fleet has proven challenging. Producing science-quality biogeochemical data requires consistent and traceable correction methods that can be adopted globally across all data centers.

In this paper we present the methodology developed as part of the SOCCOM program to assess oxygen, nitrate, and pH sensor gain, drifts and offsets in delayed-mode. The two accompanying MATLAB tools, SAGE (SOCCOM Assessment and Graphical Evaluation) and SAGE-O_2, are also described. The magnitude of required adjustments within the SOCCOM array and an independent validation of described methods are also discussed.

2 SOCCOM float array

The SOCCOM array of profiling floats includes both Teledyne/Webb Research (TWR) APEX and Sea-Bird Scientific (SBE) Navis floats. All SOCCOM floats utilize Iridium two-way satellite communication and are outfitted with ice-avoidance software developed at the University of Washington (UW) (Riser et al., 2018; Wong & Riser, 2011). For profiles taken while under ice, geographic coordinates cannot be obtained so latitude and longitude are estimated through linear interpolation. All SOCCOM floats are programmed to perform the nominal Argo mission of 10-day profile frequency from a maximum depth of 2000m with an interim park depth of 1000m.

The SOCCOM floats carry a suite of biogeochemical sensors, with sensor models varying slightly between the two platforms (Table 1). The ISUS nitrate (Johnson & Coletti, 2003) and Deep-Sea DuraFET pH (Johnson et al., 2016) sensors used on APEX floats are primarily built and calibrated at the Monterey Bay Aquarium Research Institute (MBARI). pH sensors from SBE are also deployed on APEX floats. These receive pressure and temperature calibrations at SBE, and a final pH calibration at MBARI. All other sensors listed in Table 1 receive factory-calibration direct from the manufacturer. Both sensor categories (MBARI-calibrated or manufacturer-calibrated) can suffer from shifts in laboratory calibration leading to changes in performance that manifest as sensor offsets or drifts in the field.

Parameter	Navis sensor model	APEX sensor model
T, S, P	SBE 41N	SBE 41CP
Oxygen	SBE 63 Optode	Aanderaa Optode (3830 or 4330)
Nitrate	SUNA[2]	ISUS[1]
pH	Deep-Sea DuraFET from Sea-Bird	Deep-Sea DuraFET
Bio-optics	WET Labs MCOMS (chl-a fluorometer, 700nm backscatter, FDOM)	WET Labs ECO-FLBB AP2 (chl-a fluorometer, 700nm backscatter)

[1]In-Situ Ultraviolet Spectrophotometer
[2]Submersible Ultraviolet Nitrate Analyzers

Table 1. Sensor models used on Sea-Bird Navis and MBARI/UW-built Teledyne-Webb APEX floats in SOCCOM.

Automatic QC procedures are applied in real-time to flag grossly erroneous data within the SOCCOM array. These tests roughly follow the Argo real-time tests for BGC data as outlined in Schmechtig et al. (2016). Fig. 2 shows SOCCOM float tracks colored by data quality. Points along a float track marked in purple represent profiles where >50% of the data has been marked "bad" by one of the automated QC tests. Of the three parameters shown, pH sensor data have the highest number of "bad" quality flags, at 35.59% of the data.

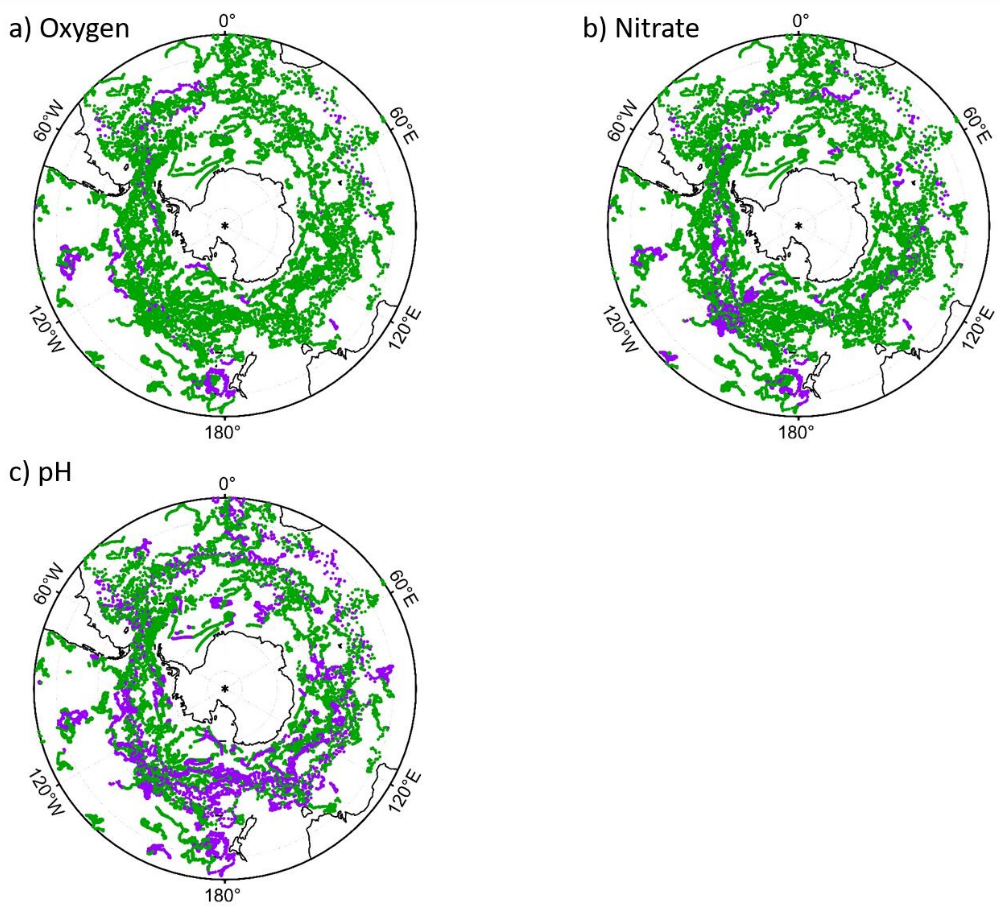

Figure 2. SOCCOM float tracks colored by data quality for (a) oxygen, (b) nitrate, and (c) pH. Points along a float track marked in purple (green) represent profiles where >50% of the data has been marked "bad" ("good") by one of the automated QC tests.

After passing the real-time quality checks, oxygen, nitrate and pH data that are considered adjustable can be brought up to the accuracy level required for global biogeochemical studies through relatively simple correction procedures (Johnson et al., 2018b; Thierry et. al, 2018b). This represents the second level of quality control (Bittig et al., 2019). In the next sections, the delayed-mode procedures and accompanying software tools used in the adjustment of oxygen, nitrate and pH data on a SOCCOM float are presented.

3 Adjustment of oxygen data

3.1 Gain adjustments for optodes

The delayed-mode correction procedure for biogeochemical data on a SOCCOM float begins with oxygen. This is because the deep reference fields used in nitrate and pH quality control (described in Section 4) are generated from empirical algorithms that require accurate oxygen measurements (along with other core variables and position information) as input parameters. Takeshita et al. (2013) have shown that the raw oxygen data from floats can be in error by as much as 20% of surface water oxygen saturation due to storage drift. Following Johnson et al. (2015), oxygen concentrations can be corrected using a multiplicative gain factor, G, to reduce the effects of storage drift and improve the accuracy of the sensor (for additional information on optode storage drift see Bittig et al. (2018a) and D'Asaro & McNeil (2013)):

$$[O_2]_{corr} = G \times [O_2]_{raw} \qquad (2)$$

There is some evidence in the literature that a slope correction on oxygen concentration could potentially be improved by the inclusion of an intercept, especially in regions of near-zero oxygen levels (Bittig & Kortzinger, 2015; Bushinsky et al., 2016; Drucker & Riser, 2016; Nicholson & Feen, 2017). However, such corrections appear to be small (<1 µmol kg^{-1}), based on an assessment of 20 floats in the Arabian Sea and Bay of Bengal (Johnson et al., 2019) and are thus not implemented within the SOCCOM program.

SAGE-O_2 is the MATLAB Graphical User Interface (GUI) developed at MBARI to assist in deriving oxygen optode gain corrections by comparing oxygen data from a float to various reference datasets, including measurements of oxygen partial pressure in the atmosphere. An image of the interface, including the plot display window and user-controlled sidebar is shown in Fig. 3 for SOCCOM float 9752 (WMO 5904694) in the Southwest Pacific, east of New Zealand. The top panel of the interface displays a time series of float data (blue) in comparison to the user-selected reference (red). Details related to the calculation of the gain factor, G, over the lifetime of a float, as implemented through the software, are described further below.

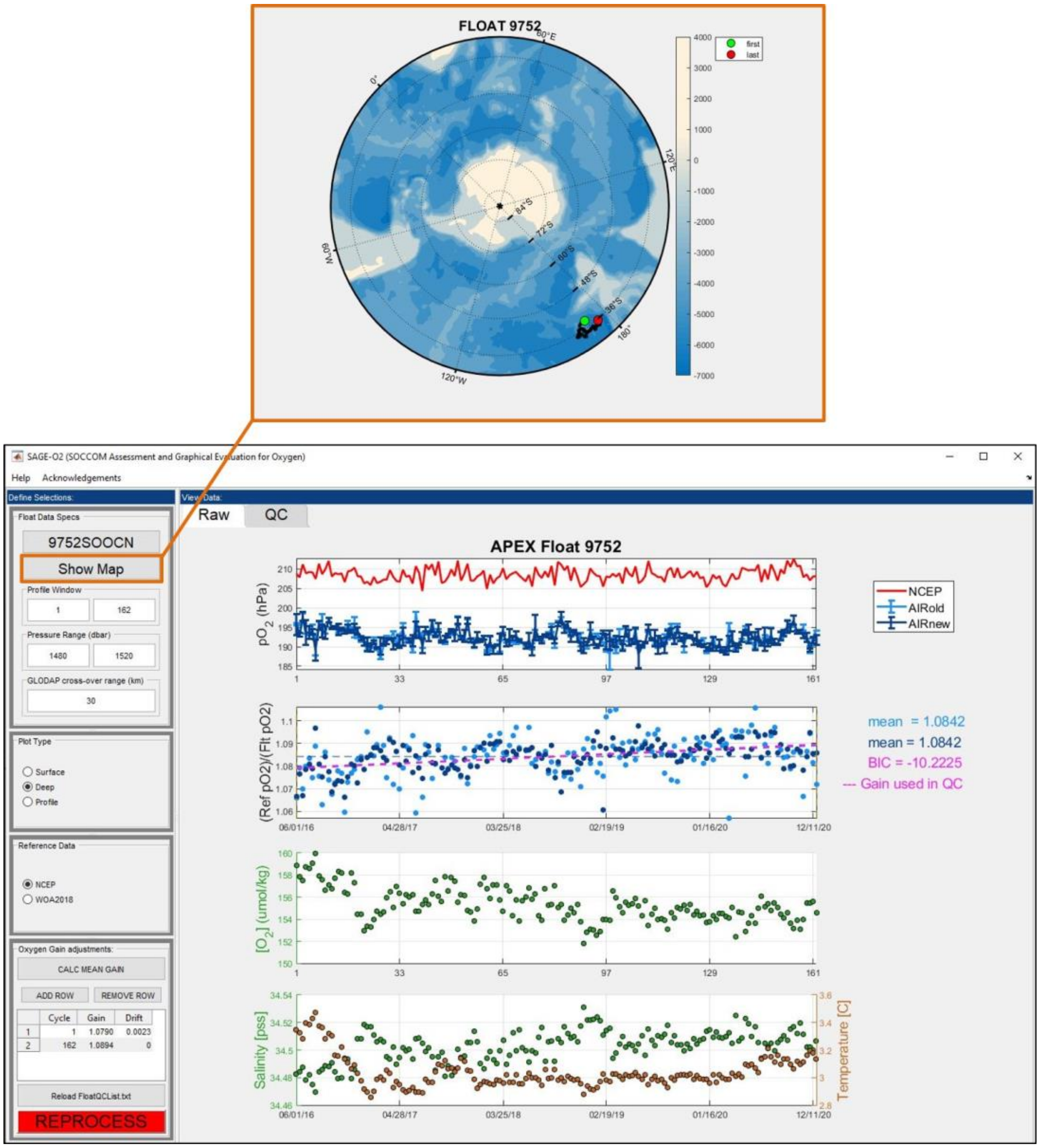

Figure 3. SAGE-O_2 software interface showing results of the calibration for sample float 9752, WMO 5904694. The map display functionality is also indicated.

3.1.1 Gain computation using in-air oxygen with NCEP/NCAR Reanalysis

In-air calibration of oxygen optodes onboard profiling floats has been shown to bring accuracy to within 1% and is currently the operational standard (Johnson et al., 2015). For floats with in-air measurement capabilities, an estimate of atmospheric pressure must be available to compute the local oxygen partial pressure. The product referenced for oxygen gain computation within the SAGE-O_2 software is NCEP/NCAR Reanalysis-1 six-hourly surface pressure (Kalnay et al., 1996). This is a Gaussian gridded product with units of Pascals, which are converted to hectopascals (millibar equivalent) prior to proceeding. The NCEP atmospheric surface pressure (P_{NCEP}) values are interpolated to the time and location of the float's surfacing. Values are then converted to oxygen partial pressure based on the assumption that the atmosphere is 100%

saturated with water vapor at the sea surface (equation (3)). The water vapor pressure (p_{H_2O}, in hPa) is calculated using equation (4), where T represents temperature in degrees Celsius (Aanderaa Instruments, 2017).

$$p_{O_2} = \left(P_{NCEP} - p_{H_2O}\right) \times 0.20946 \tag{3}$$

$$p_{H_2O} = e^{\left[52.57 - \left(\frac{6690.9}{T+273.15}\right) - 4.681 \times ln(T+273.15)\right]} \tag{4}$$

The sensor gain that is estimated from air oxygen for each individual profile, *i*, is then computed using equation (5), as outlined in Johnson et al. (2015):

$$G_i = p_{O_2 NCEP} / p_{O_2 FLOAT} \tag{5}$$

where $p_{O_2 NCEP}$ follows from equation (2) and $p_{O_2 FLOAT}$ is the partial pressure of oxygen computed from the float (reported in millibars). The overall gain factor, G, used to correct all in water oxygen observations is then the mean of the n individual G_i values.

Mean gain values over the float's life are displayed within the SAGE-O_2 interface in blue to the right of the plot panels (Fig. 3). Note that at the start of the SOCCOM program, APEX floats were programmed to take a single in-air oxygen reading with each surfacing that was associated with the telemetry phase of the cycle. A subsequent upgrade to the mission programming was initialized such that the optodes on APEX floats take a sequence of in-air measurements at each surfacing at the end of ascent (4 subsurface measurements followed by 8 measurements in air after inflation of the air-bladder). Therefore, the majority of APEX floats in the SOCCOM program have 2 sets of in-air measurements: one associated with the telemetry phase (light blue in the GUI interface), and another larger set associated with the in-air measurement series (dark blue in the interface). Both of these are plotted in the GUI for comparison. Average gain between the two sets differs by less than 0.1 % fleet-wide.

In the future the SAGE-O_2 software may be upgraded to utilize the now real-time NCEP/DOE-R2 reanalysis product. Additional reanalysis products from other centers are also available, including the European Centre for Medium-Range Weather Forecasts (ECMWF) ERA5 reanalysis. The ERA5 product utilizes a more state-of-the-art (4D-variational) data assimilation system but its data latency (3 month lag for quality assured updates) may limit timely delayed-mode QC operations. The absolute uncertainty in reanalysis surface pressure fields from different products can be difficult to fully quantify although a comparison of NCEP and ECMWF operational models by Salstein et al. (2008) found that rms differences between surface pressure and shipboard observational stations were between 2 and 5 hPa in Southern latitudes with minimal difference between the two products, especially in more recent years. Surface pressure uncertainties of this magnitude roughly translate to less than 0.5% change in corrected O_2 measurements on individual floats.

3.1.2 Gain computation using shipboard bottle data

The SBE63 optodes onboard SOCCOM Navis floats are plumbed in line with the pumped CTD flow stream and are thus not fully exposed to ambient air during surfacing. In-situ calibration of these floats thus relies on comparison to high-quality Winkler titrations from shipboard samples taken at the time of float deployment. The Winkler oxygen are generated primarily on GO-SHIP cruises or by research groups that regularly participate in GO-SHIP cruises and they are considered to be of a quality consistent with GO-SHIP measurements (Hood et al., 2010 state a target accuracy of 2σ less than 0.5% of the largest oxygen concentration found in the ocean). Comparisons of the float and bottle data can be viewed through the software (Fig. 4). We focus on the upper 50m near the surface where oxygen is close to 100% saturated and the vertical gradients are small. A comparison of average gain values derived using shipboard Winkler measurements versus in-air samples for 97 SOCCOM APEX floats shows a mean difference (float minus bottle) of -0.31% (standard deviation of 2.2%). This suggests that there is no large systematic bias for Navis floats when optodes are calibrated using bottle data.

In addition to providing an alternative approach to in-situ optode calibration, comparison to shipboard data offers a simple and independent means for validating gain values derived from other methods. The gain correction for the float shown in Fig. 3 was performed using in-air measurement data as described in Section 3.1.1. Fig. 4 shows data from this float in profile view. Pressure is along the x-axis for all plot panels. The top two panels show mean float data (solid blue line) along with GLODAPv2 profile data that are within a 30km radius, and the computed residuals. The bottom two panels show the float's first and second profiles (blue) along with shipboard Winkler and CTD oxygen data (circles), and computed residuals. Note that the 'QC' tab is selected, thus all float data in the display have been adjusted using the computed gain shown in Fig. 3. If the 'Raw' tab was chosen, the float profile would have no adjustments applied. The small positive bias shown in reference to the bottle data is due to temporal mismatch between the shipboard data and float measurements within high-gradient regions of the profile. The mean residual (bottle-float) is 1.245 μmol kg^{-1}. The mean residual against all GLODAPv2 data within 30 km is -0.060 μmol kg^{-1}, although the range is larger than the hydrocast data due to the larger time range included in the matchup criteria.

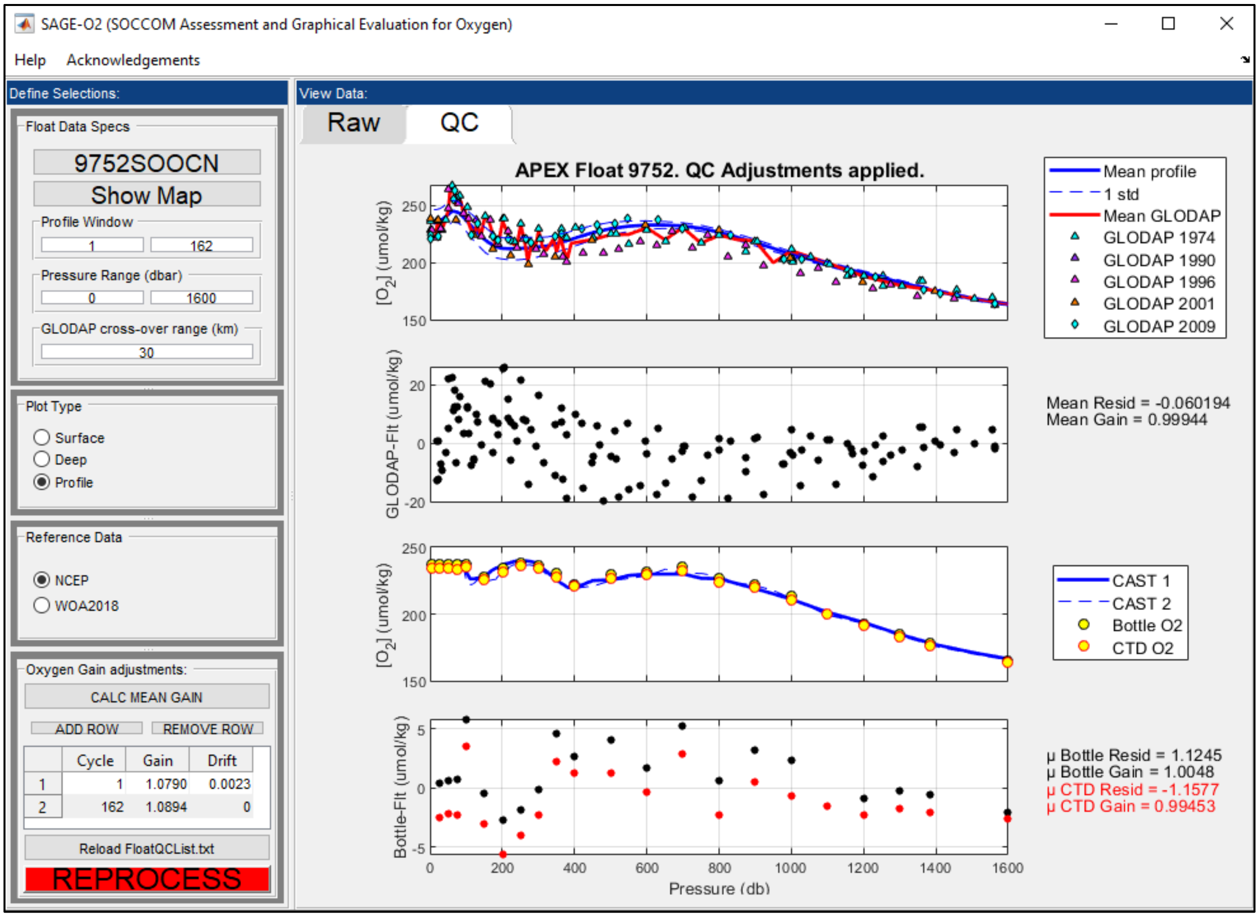

Figure 4. A comparison of adjusted oxygen data to GLODAPv2 (top two panels) and shipboard hydrocast matchups (lower two panels), as viewed through the SAGE-O_2 interface for UW/MBARI float 9752 (WMO 5904694).

3.1.3 Gain computation using World Ocean Atlas climatology (WOA)

For floats incapable of taking in-air oxygen measurements, and when shipboard reference data are not yet available, a preliminary optode gain correction factor can be derived within the SAGE-O_2 GUI using WOA percent oxygen saturation in surface water. This method follows Takeshita et al. (2013), which suggest an accuracy of 1-3% for sensors calibrated against WOA values. Percent saturation from the float is calculated following equation (7) below, where the solubility of oxygen (O_{2Sol}) is computed as a function of temperature and salinity following Garcia & Gordon (1992) equation 8 (omitting erroneous term [$A3*T_s^2$]) and using solubility constants from Benson and Krause (1984) (see equation 8 and Table 1 in Garcia & Gordon, 1992). Individual gain values, G_i, are then computed using equation (8), where $\%Sat_{WOA}$ and $\%Sat_{Float}$ represent the mean WOA and mean float percent saturation values for the upper 25m of the profile, respectively.

$$\%Sat = [O_2]/[O_{2Sol}] \times 100 \qquad (7)$$

$$G_i = \%Sat_{WOA}/\%Sat_{Float} \qquad (8)$$

The overall gain factor, G, is calculated as the mean of the individual gain values (G_i) computed for each cycle. A comparison of gain factors computed using WOA percent saturation versus NCEP reanalysis air pressure as reference for 95 floats with in-air measurement capabilities shows a bias between the methods of 1.4% with a standard deviation near 2%. The largest differences occur in floats near seasonal sea ice or very close to the coast where WOA reference climatology data is limited and/or seasonally biased. Note that for many floats within the global BGC Argo array, this method is the most accessible option for data managers and should be applied wherever possible as a first-order correction.

3.2 Drift in optode gain

The effects of pre-deployment storage drift are readily apparent across the majority of optodes used on profiling floats. Oxygen data from all Aanderaa and Sea-Bird optodes onboard SOCCOM floats require gain correction, with a fleet-wide mean gain correction of 7.0 ± 4.6 (1σ) %. While an optode's stability once deployed is substantially smaller, it is less predictable (Bittig & Kortzinger, 2015; Bittig & Kortzinger, 2017; Bushinsky et al., 2016; Johnson et al., 2015). Bittig et al. (2018a) provides a thorough review on this topic, and suggests that individual optodes may exhibit significant post-deployment drift of up to +/- 0.6%/yr. If not accounted for, such drift could lead to significant biases in certain biogeochemical analyses such as air-sea fluxes.

Characterizing the amount of optode drift is possible within the SAGE-O_2 software through comparison against reference values over time. This method was recently put into practice for select floats within the SOCCOM fleet. The software allows the user to auto-calculate the drift relative to a reference such as NCEP. The computed offset (initial gain), *b*, and slope (drift), *m*, are calculated using a model I regression of computed gain on each cycle against cycle time. The gain value applied at each cycle (following equation 1) then becomes:

$$G_{i=1:k} = b + m(\Delta T) \tag{9}$$

where ΔT is the time, in years, elapsed since the first cycle (or time at which the drift began). If the chosen ending node at cycle *k* is not the final cycle reported from the float upon assessment, a drift assessment on the subsequent segment (cycles i=k:n) is automatically performed. The slope of the second segment, m_2, is found by first subtracting the recomputed gain at the end of the first segment (G_k) from individual gains, g_i, of segment 2, and then regressing segment 2 through the origin. This can be expressed as

$$m_2 = \frac{\sum_{i=k}^{n}(g_i - G_k) * x_i}{\sum_{i=k}^{n} x_i{}^2} \tag{10}$$

where x represents the time elapsed since the ending cycle of segment 2. This method results in drifting gains that remain continuous throughout segments. However, note that drift assessment within the GUI (and especially multi-segment drifts) should be limited to advanced users. It is recommended that drift assessment be performed only after a sufficient amount of data has been received (optimally at least 2 years). Care must be taken in order to prevent correcting for an apparent drift that has been influenced by a seasonal cycle.

Within the GUI there are two methods to test whether or not a computed drift over the lifetime of a float is statistically robust. Upon auto-computation of the drift, a two-tailed T-test is performed to assess whether the calculated slope is significantly different than zero at the 95% confidence interval (results are returned on screen). Additionally, on the right-side panel in the interface, the GUI reports the computed Bayesian Information Criteria (BIC) (Schwarz, 1978) following Equation 11 below, where *SSR* represents the sum of squared residuals of the model, *K* is the number of model parameters, and *n* represents the number of data points. The BIC weighs the number of predictors within a model against the goodness-of-fit, allowing the user to prevent over-fitting of the data (the model with the lowest BIC is always preferred).

$$BIC = \log\left(\frac{SSR}{n}\right) + \frac{K \log n}{n} \qquad (11)$$

In the SOCCOM array, of the 126 floats currently considered candidates for optode drift correction, 32 exhibited significant drift rates. Both positive and negative drift rates were observed, with a mean of -0.07%/yr, a standard deviation of 0.65%/yr and a total range of -1.1 to 1.2 %/yr.

The drift correction proposed here relies on the existence of air oxygen measurements relative to the NCEP atmospheric reference. However it does not address the root cause of sensor drift behavior which is somewhat unsatisfying. Bittig et al. (2018a) show how inadequate temperature calibration of the oxygen optode can oftentimes account for in-situ drift rates apparent in a float's optode time series. They describe a correction method (Equation 23 of referenced publication) that can simultaneously correct for inadequate temperature calibration and any seawater carryover on the sensor during sampling while in air. The supplementary material to their paper highlights the results of applying the method to UW/MBARI float 9313 (WMO 5904474); the strong oxygen-temperature response exhibited by this float is shown to bias the sensor gain time series and application of the correction method effectively removes the apparent drift in sensor gain. However, recent testing demonstrates that the results of this correction are not consistent across the SOCCOM array. Fig. 5 plots computed drift in optode gain against the residual drift in optode gain after temperature compensation with Equation 23 from Bittig et al. (2018a) is applied for 82 SOCCOM floats that have been operational for at least 2 years. The Model II regression (shown in red) gives an offset of 0 which suggests that the Bittig et al. (2018a) correction is robust and does not add spurious drift. The slope of the Model II regression is 0.797 (different than 1 at the 99% significance level) suggesting that across the SOCCOM array, the correction reduces the apparent drift in gain by 20.3%. For certain floats, the Bittig et al. (2018a) correction tends to underestimate the magnitude of the true drift of the optode, thus, additional drift correction may be warranted. The mean difference in gain drift before versus after the correction is -0.021%/year and the standard deviation of the differences is 0.31 %/yr. These results highlight the fact that the optode-temperature response is unique to each sensor. This result is in accordance with findings of Johnson et al. (2017) who show that only 20% of the change in gain over time can be accounted for by temperature changes observed by a float. Such corrections should therefore not be applied systemically across the whole fleet, but rather integrated on a float-by-float basis in delayed-mode with statistical indexing to weigh the benefit of added complexity of the correction, similar to what is currently being done to assess the need for drift corrections. These methods may be integrated into the GUI framework in a similar manner in a future revision.

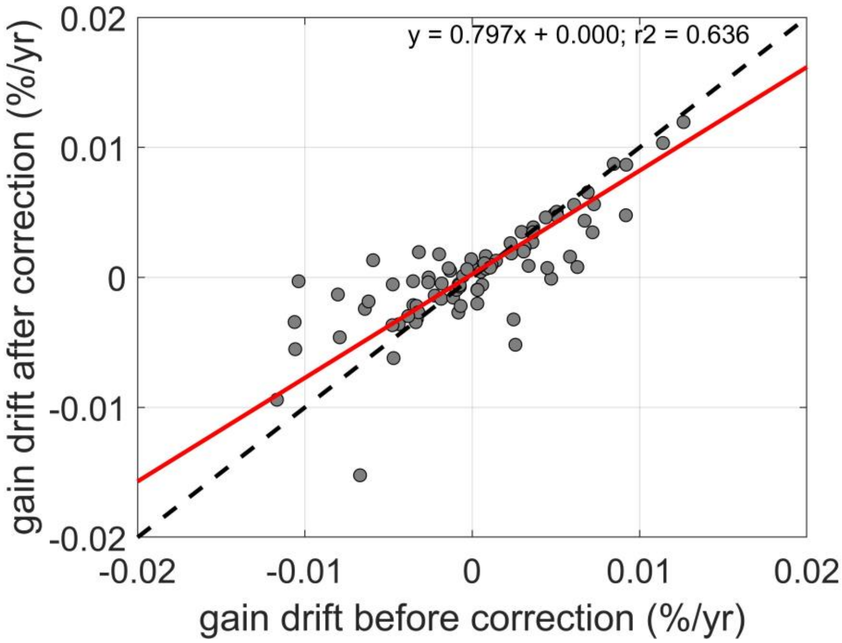

Figure 5. Comparison of post-deployment optode drift before and after application of Bittig et al. (2018a) Eqn 23. Analysis includes 82 SOCCOM floats. Dashed line depicts the 1:1 relationship; red line is the Model II regression.

4 Adjustment of nitrate and pH data

Adjustment of nitrate and pH data are performed after oxygen data has been corrected. Similar to oxygen optodes, nitrate and pH sensors on profiling floats often suffer from initial calibration shifts that must be corrected prior to scientific use. Such inaccuracies can manifest as offsets and/or drifts throughout the data series. As described in Johnson et al. (2017), pH offsets and drifts can be attributed to changes to the sensor reference potential (k_0) over time, while those apparent in nitrate usually result from changes in light throughput due to aging or fouled optical components. Therefore, adjustments to pH and nitrate are applied as offsets to k_0 and nitrate concentration [μmol kg^{-1}], respectively.

The general adjustment process for pH and nitrate is based on evidence that the offsets and drifts are constant throughout an entire profile (Johnson et al., 2013; 2017). Corrections then involve comparison of raw float data to select reference fields at depths below 1000m where spatial and temporal variability in ocean chemistry is minimal. The corrections determined at depth are then applied to the entire profile. This process is similar to the protocol used to correct Argo salinity data (Owens & Wong, 2009). Fig. 6 below shows the SAGE GUI interface where such comparisons can easily be performed. Upon selecting a float, default view specifications are loaded into the GUI, including a profile window encompassing the entirety of the float's life-span, and a pressure range of 1480 to 1520 m where adjustment assessment is performed. Float (blue) and reference (red) data within selected time and pressure ranges are plotted in the top panels, and the anomaly series (float minus reference) is plotted below in green. Global Data Analysis Project v2 (GLODAPv2; Olsen et al., 2020) crossover data is also shown in the upper panel plots as a climatological reference, but only to assess the consistency of adjusted data. As in SAGE-O_2, the search distance for GLODAPv2 data from each profile can be set in the GUI.

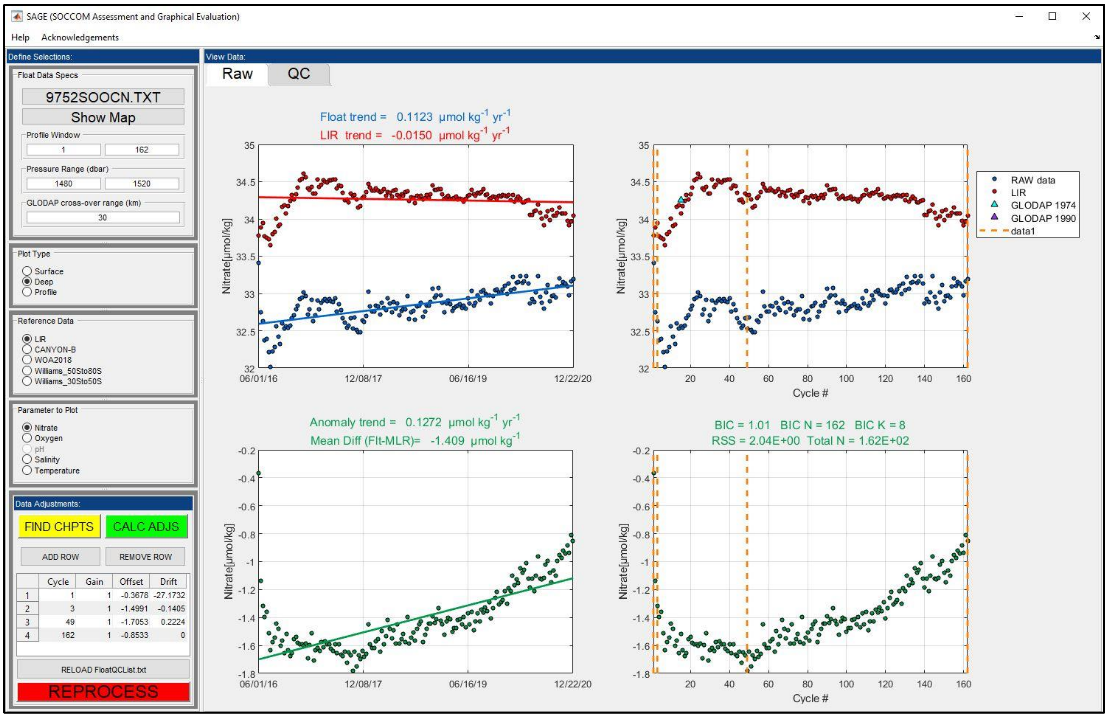

Figure 6. SAGE GUI software interface showing raw nitrate data (blue) from MBARI/UW float 9752 (WMO 5904694).

Similar to conductivity sensors (Owens & Wong, 2009), drifts and offsets occurring in data from nitrate and pH sensors often vary linearly over long time periods, and calibration jumps in the time series are not uncommon. Oftentimes the largest drift rates occur over the first few cycles in a float's life as can be seen in the nitrate anomalies shown in Fig. 6. Nitrate and pH anomalies from a float data series are thus best modeled as discontinuous piecewise linear fits, where both drifts and offsets change independently between segments that are bounded on either side by defined cycle breakpoints. In the Fig. 7 schematic, the correction, ΔANOM, at each cycle breakpoint, j, is calculated as

$$ANOM_j = O_j \qquad (12)$$

and the data correction for any subsequent cycle, i, within the same segment becomes

$$ANOM_i = O_j + D_j(T_i - T_j) \qquad (13)$$

where O and D represent the offset (in µmol kg^{-1}) and drift (in µmol kg^{-1} per year), respectively, of the linear least squares fit to the anomaly data series between cycles located at breakpoints j and j+1 (not including the latter bounding breakpoint), and T represents time (in years). For nitrate data, this modeled correction (represented by gray lines in Fig. 7) is then subtracted from the original data series. For pH data, the modeled correction is applied as an offset to the reference potential (k_0) of the sensor as described in Johnson et al. (2017). A matrix of correction factors (as shown in the lower left corner of Fig. 6) is stored in a float-specific text file along with any derived oxygen corrections for use in reprocessing applications. This method constitutes a delayed-mode correction approach that can be revisited and characterized at periodic intervals throughout the float's life.

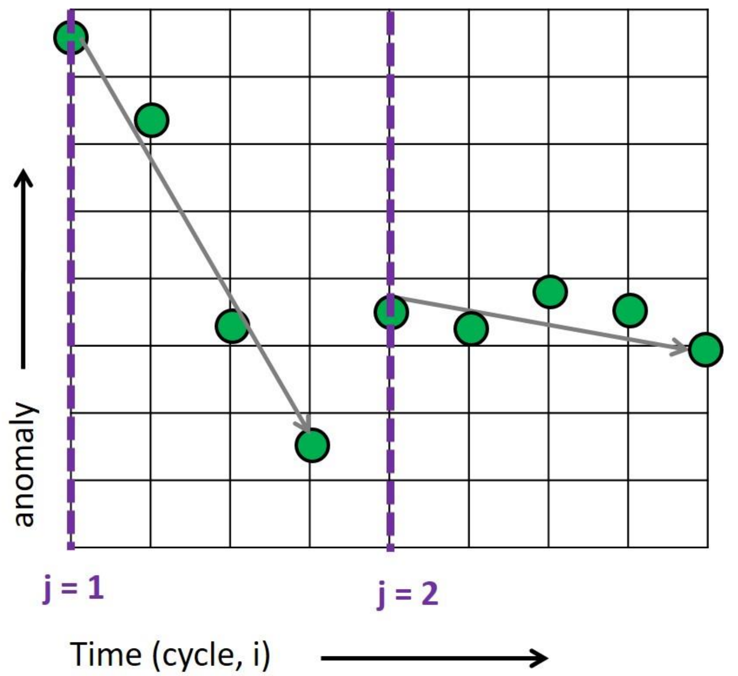

Figure 7. Qualitative schematic showing the adjustment model of a theoretical sensor anomaly series. The two series breakpoints, identified in purple, occur at cycles 1 and 5. Gray lines represent the least-squares fit (adjustment model) to the elements (green dots) within each segment.

4.1 Reference models for pH and nitrate

Multiple options are available for use in the estimation of deep pH and nitrate reference fields for comparison against float data. These include World Ocean Atlas climatological fields as well as empirical algorithms derived from high-quality shipboard data acquired from GO-SHIP cruises (Bittig et al., 2018b; Carter et al., 2018; Williams et al., 2016). While the algorithms provide estimated fields rather than direct measurements, their performance has been extensively validated. The set of multiple linear regression models (MLRs) by Williams et al. (2016) were the first of such reference algorithms available in the Southern Ocean and were utilized in the quality control of SOCCOM nitrate and pH float data during the early years of the program. Nitrate and pH estimates produced using the Williams method rely on MLR equations specific to two latitudinal bands around the Southern Ocean. Predictor variables include pressure, salinity, temperature, and oxygen. A key distinction between the Williams MLRs and the other methods available for use within the SAGE software is the lack of global extent in the Williams MLRs. In addition, this method is limited in depth space to the range of 1000 to 2100 m. While this fully encompasses the depth nominally used in quality control for the majority of SOCCOM floats, sometimes shallower reference depths are required, for example when a float is under-ballasted and cannot reach 1000 m. Nonetheless, the Williams MLR algorithms perform very well when used within their specific range limits. Williams et al. (2016) states root mean square errors (RMSE) of 0.3 µmol kg^{-1} and 0.004 total pH units for deep (1500m) nitrate and pH estimates, respectively. Additionally, Johnson et al. (2017) show linear regressions between first nitrate and pH profiles from SOCCOM floats, adjusted to the Williams MLRs at depth, and shipboard bottle data taken at the time of deployment to be near unity, with midrange differences (bottle minus float) of -0.1 µmol kg^{-1} and 0.006 pH units, respectively. These findings are important as they validate the method as an acceptable reference option for other float programs in the Southern Ocean.

However, as increasing numbers of BGC floats are being deployed outside of the Southern Ocean, an alternative reference algorithm with full global extent is now the operational standard. This allows for a consistent procedure, homogenous across float arrays. The current default choices for estimating nitrate and pH for comparison against SOCCOM float data are the locally interpolated nitrate regression (LINR) and the locally interpolated pH regression (LIPHR) (or LIRs, collectively) (Carter et al., 2018). The LIR algorithms were developed from a series of MLRs trained using GLODAPv2, resulting in a separate set of coefficients for each 5 degree latitude and longitude grid box and 33 different depth surfaces. The derived coefficients at each grid point then get interpolated onto a float's location for use in generating a final nitrate or pH estimate. As described in Carter et al. (2018), there are 16 possible groupings of predictor variables available to use in producing a final estimate. For SOCCOM assessments, LIR regression #7 is used with depth, salinity, temperature, and dissolved oxygen as input parameters, in addition to the profile latitude and longitude. The RMSE of the residuals between LIPHR and LINR estimates within 1000 and 2000m using predictor set #7 and the test observations used for algorithm validation were 0.006 pH units and 0.47 µmol kg^{-1}, respectively (Carter et al., 2018).

A third optional reference algorithm is the CArbonate system and Nutrient concentration from hYdrological properties and Oxygen using a Neural-network, Bayesian approach (CANYON-B, Bittig et al., 2018b). This is a neural network mapping performed in a Bayesian framework, that is, informed by an ensemble of model components at each stage rather than fixed values. This model is a revised version of an earlier individual neural-network approach, CANYON, originally developed by Sauzéde et al. (2017). In their publication, Bittig et al. (2018b) compare the performance of CANYON-B with LIR for various parameters, including nitrate and pH, against a post-GLODAPv2 validation dataset. The authors stress that, while both methods perform similarly well in a bulk statistical sense, local estimates can still be quite different. Fig. 8 compares differences between pH and NO_3^- estimates for the SOCCOM array using CANYON-B and LIR algorithms. The mean (standard deviation) of the differences at the depth that QC is performed within SAGE are -0.001 (0.006) for pH and -0.053 (0.278) µmol kg^{-1} for NO_3^-. Larger differences near the surface are largely due to greater uncertainties of the LIR algorithms at these depths (although, as noted by Bittig et al., (2018b), estimates from all algorithms show some level of enhanced uncertainty toward the surface due to difficulty in accurately capturing seasonal variability and effects of air-sea gas exchange). The enhanced skill in near-surface depths exhibited by CANYON-B, relative to LIR, can serve as an independent validation to the calibration approach. Surface data from floats corrected at depth using LIR frequently align well with CANYON-B estimates at the surface, in a qualitative sense. However, it should be noted that pH estimates generated by CANYON-B are intended to be in line with pH calculated from DIC and TA, rather than pH that has been spectrophotometrically measured, whereas pH measurements using the LIPHR method are the opposite. While the LIPHR algorithm has a flag to apply a linear adjustment that will subsequently produce estimates consistent with calculated pH, this method should not be used for calibrating a pH measurement from a float, as ISFET pH is consistent with spectrophotometric pH measurements (Takeshita et al., 2020). The differences shown in Fig. 8 were performed after a linear transformation was applied to CANYON-B estimates following Carter et al. (2018) (Equation 1) to bring estimates back into alignment with spectrophotometrically measured pH.

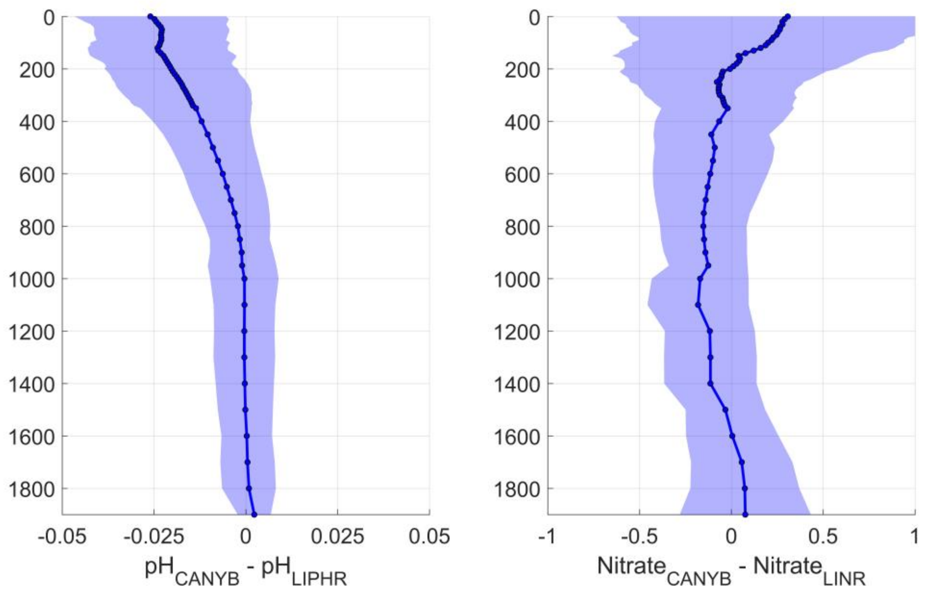

Figure 8. Fleet-wide differences of computed pH (left) and nitrate (right) using CANYON-B and LIR algorithms. Data were binned at 10, 50 and 100m pressure intervals for 0-350, 350-1000, and 1000-2000 db, respectively. The blue line represents the mean difference and the shaded areas represent +/- 1 standard deviation.

A final note should be made regarding the use of pH estimates that are based on measurements made over a large time span. Ocean pH is decreasing due to increasing atmospheric carbon dioxide concentrations and these effects are sometimes detectable at the depth range used for pH sensor adjustment (Rios et al., 2015). Each of the algorithms described here has been trained on shipboard data that may exhibit this effect. While the LIPHR algorithm does include a flag for optional application of an ocean acidification adjustment, this is a static adjustment and does not account for geographic differences in ocean acidification rates, nor does it account for changes in global ocean acidification rates over time. This highlights the need for such reference equations to be periodically updated, utilizing recent training datasets to provide more accurate algorithm coefficients.

4.2 Computation of correction factors using automated change-point detection

In the initial version of the SAGE software, the user manually chose the location of each breakpoint (node). The inherent subjectivity in this approach in addition to the increasing time investment required by the operator to complete a full adjustment assessment of the SOCCOM array proved less than optimal. In the current software version, both the optimal number and location of each breakpoint can be assigned automatically through an automated multi-step process. First, the binary segmentation method of change-point detection is applied using the MATLAB function, ischange, which begins by splitting the data series for variable y, of length n, into two segments separated by a change-point, j (Killick et al., 2012). The location of j along the time series is then iteratively shifted until a minimization of the left side of the following equation is reached:

$$C\left(y_{1:j}\right) + C\left(y_{(j+1):n}\right) < C\left(y_{1:n}\right) \tag{14}$$

where C represents the cost function

$$C(x) = nVar(x) \tag{15}$$

where n is the number of data points in the segmented data series, x, and Var is the variance. This process is then repeated, further splitting up the segments to find the optimal location of an increasing number of changepoints. Next, in order to statistically determine the best number of changepoints of the various groupings tested, a modified BIC is calculated for each model, following

$$BIC = \log\left(\frac{SSR}{n} + \alpha^2\right) + \frac{K \log n}{n} \tag{16}$$

where the α term is used as a threshold on the mean residual, driven by the basic precision of the sensor. In SOCCOM processing operations, α=0.5 (0.005) is used for nitrate (pH) data. If α is omitted, equivalent to assuming the sensor has no inherent noise, the changepoint algorithm will often find an excessive number of change points, which is inconsistent with known sensor behavior. The location and number of changepoints from the model with the lowest BIC value is then used to derive offsets and drifts as described in Section 4.

A key concern in the move from a manually-assigned to an automated definition of breakpoints in the sensor QC process was to maintain the final quality of the adjusted SOCCOM dataset. Thus, prior to operational implementation of the automated method, a quality assessment was performed using two adjusted datasets, one done manually by a trained biogeochemical float quality control operator and the other performed automatically using the changepoint detection method described above. Fig. 9 (a-d) shows that the use of automated changepoint detection in the SOCCOM QC process results in a fewer number of change-points, on average, and an overall better model of the anomaly time series, in a statistical sense (lower BIC value), than the previously employed manual correction method.

However, the absolute difference in BIC between models is small in most cases (mean differences of 0.658 and 1.165 for nitrate and pH, respectively) with the automated method showing progressively better performance as model complexity increases (Fig. 9, e-f). It is generally accepted that when comparing candidate models, a difference in computed BIC less than 2 is relatively inconsequential, meaning that the two models are statistically similar and minimal (if any) improvement can be attained by choosing one over the other (Fabozzi et al., 2014; Kass & Raferty, 1995). When taken in this context, results from this comparison suggest that the initial manual method of change-point detection for QC across the SOCCOM fleet was not of poor quality, and that the move to automated changepoint detection sustains such quality while concurrently reducing the time required to perform an objective fleet-wide assessment.

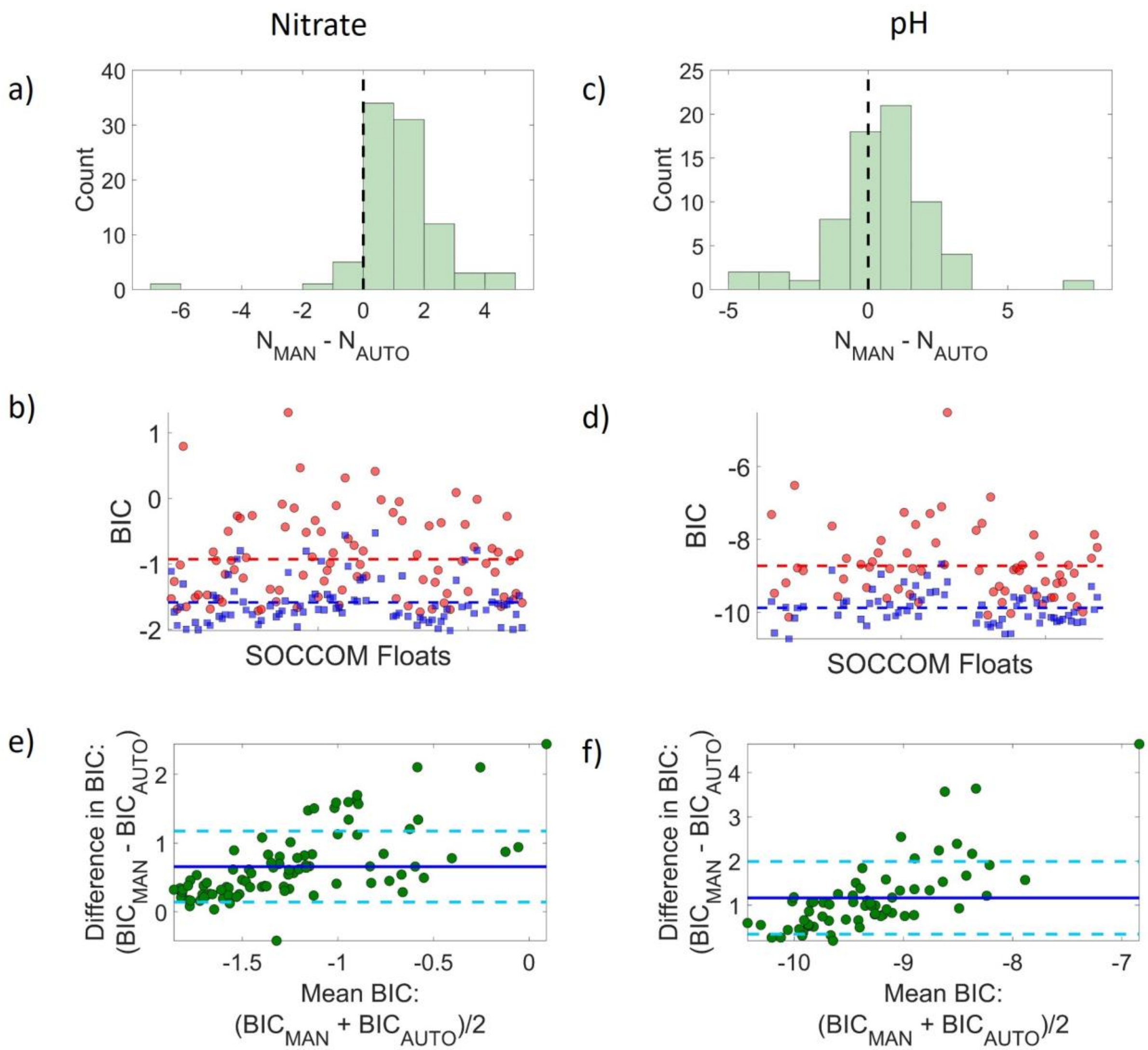

Figure 9. (a,b): Histograms showing differences in number of changepoints identified by the manual (N_{MAN}) versus automated (N_{AUTO}) method for nitrate and pH sensor QC. (c,d): Comparison of computed Bayesian Information Criterion (BIC) for manual (red circles) and automated (blue squares) changepoint identification in nitrate and pH QC. Dashed lines represent mean BIC values for each method. (e,f): Difference in computed BIC (manual versus auto) against mean BIC value for each float, for nitrate and pH. Solid and dashed lines represent mean difference +/- 1 standard deviation, respectively. 120 SOCCOM floats were used in each analysis.

4.3 pH and nitrate adjustments in the SOCCOM array

The magnitude of a required sensor adjustment, as derived from the methods described in the previous sections, represents the degree to which sensor performance has changed since laboratory calibration. A summary of the adjustments required over time across a full array of sensors can unveil any systematic biases and subsequently help identify key areas for which to focus future development efforts. While the adjustment methods described in this paper improve data accuracy, reducing the magnitude of required adjustments to a sensor is the optimal goal. As described in Section 4, the coefficients to the linear fits of each segmented anomaly series are included within a single float-specific correction matrix that is used in the data adjustment process. The offset associated with the first segment exemplifies sensor performance upon deployment. As each segment is treated independently, the value of any subsequent offset can provide information on sensor health over time when viewed relative to the first offset.

The distributions of the first and second offsets required for nitrate and pH data in the SOCCOM array are shown in Fig. 10. The positive skew of the nitrate first offset distribution demonstrates that the majority of SOCCOM nitrate sensors are biased high upon deployment while the opposite is true for pH sensors within the array. The magnitude of the bias is 0.91 μmol kg^{-1} for nitrate, and -0.032 for pH (Table 2). Distributions of the second offsets (relative to the first) show reduced spread across both sensor types and an elimination of bias in pH sensor data. This behavior is not surprising; oftentimes the largest anomaly is observed on the first cycle as the sensor re-conditions to an aqueous environment. Continued exposure to seawater at 1500m helps to stabilize the sensors, particularly the pH sensor. The optics of the nitrate sensor are more sensitive to perturbations so jumps in the data series are more often observed. This is exemplified by the fact that a small bias (negative) remains in the distribution of second nitrate offset, showing that a second offset is almost always required to bring nitrate data in line with climatology.

Also notable in the distributions is that there is a small subset of floats receiving relatively large first offset corrections for nitrate and pH sensor data. Currently there is no operational threshold in place for maximum allowable adjustment. Floats requiring larger than normal nitrate or pH adjustments are analyzed on a case-by-case basis and may be grey-listed as bad or questionable by the delayed-mode operator upon review of laboratory calibration and sensor diagnostics. These large offsets may be the result of changes in optical alignment or sensor contamination during transport.

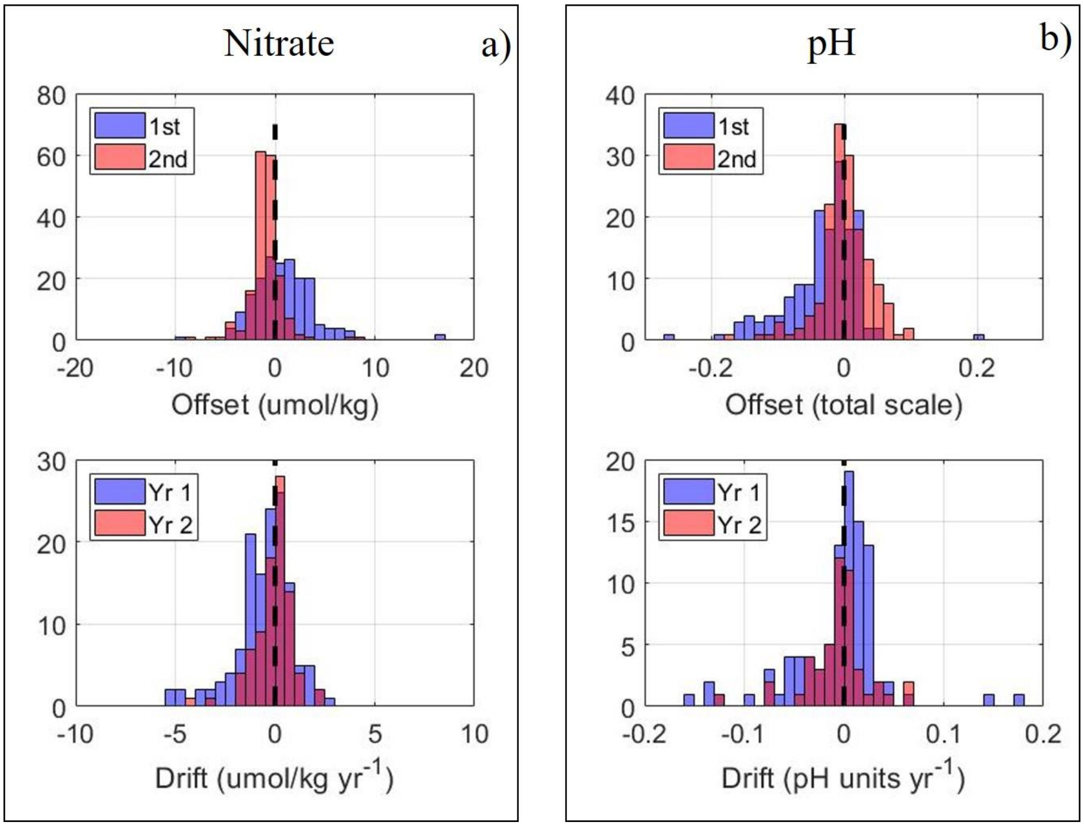

Figure 10. Histograms of first and second offsets (top), and first-year and second-year drift rates (bottom) for nitrate (a) and pH (b) data. Offsets were computed as float data minus reference data at a nominal calibration depth of 1500m; the second offset is relative to the first. Drift rates were computed using a Model I regression on the anomaly time series.

a)	Nitrate 1st offset (μmol kg^{-1})	Nitrate 2nd offset (μmol kg^{-1})	Nitrate 1st-year drift (μmol kg^{-1} yr^{-1})	Nitrate 2nd-year drift (μmol kg^{-1} yr^{-1})
Median	0.72	-0.95	-0.17	0.08
Mean	0.91	-0.95	-0.51	-0.09
Std dev	3.12	1.75	1.52	0.96

b)	pH 1st offset	pH 2nd offset	pH 1st-year drift (yr^{-1})	pH 2nd-year drift (yr^{-1})
Median	-0.020	0.002	0.000	-0.002
Mean	-0.032	0.001	-0.017	-0.005
Std dev	0.059	0.040	0.060	0.032

Table 2. Data adjustment summary statistics for nitrate (a) and pH (b).

First year and second year sensor drifts for nitrate and pH are also shown in Fig. 10 (lower histograms). These were computed as the slope of a Model I regression over the first and second year of data for each float. This ensured a uniform time frame for drift comparison across the array (as the length of each segment within a float's adjustment matrix can vary). While drift in the second year is not completely eliminated, there is an 80% (70%) reduction in mean drift rate across the array for nitrate (pH) sensors from year 1 to year 2. The reduction in sensor drift from year 1 to year 2 is not a uniform rate of change. This can be seen in Fig. 11 that shows percentiles across the array of computed anomalies at each profile relative to that of the first profile (top) and percentiles of the rate of change in anomaly by profile (center). By the second year, around 25% of nitrate anomalies have drifted beyond 2 µmol kg^{-1} of their initial value with the majority of sensors drifting negative (measuring low relative to reference fields) and the largest proportion of drift occurring within the first five cycles (red line). pH sensors see both positive and negative drift rates, with close to 50% of the data drifting beyond 0.03 pH units of their initial value. However, similar to nitrate sensors, pH sensors are also relatively stable beyond the first few cycles. Because both nitrate and pH sensors exhibit the largest rates of in situ drift within the first 2 months since deployment, it is recommended that initial QC assessment be performed only after the first 5 cycles have been returned from a float.

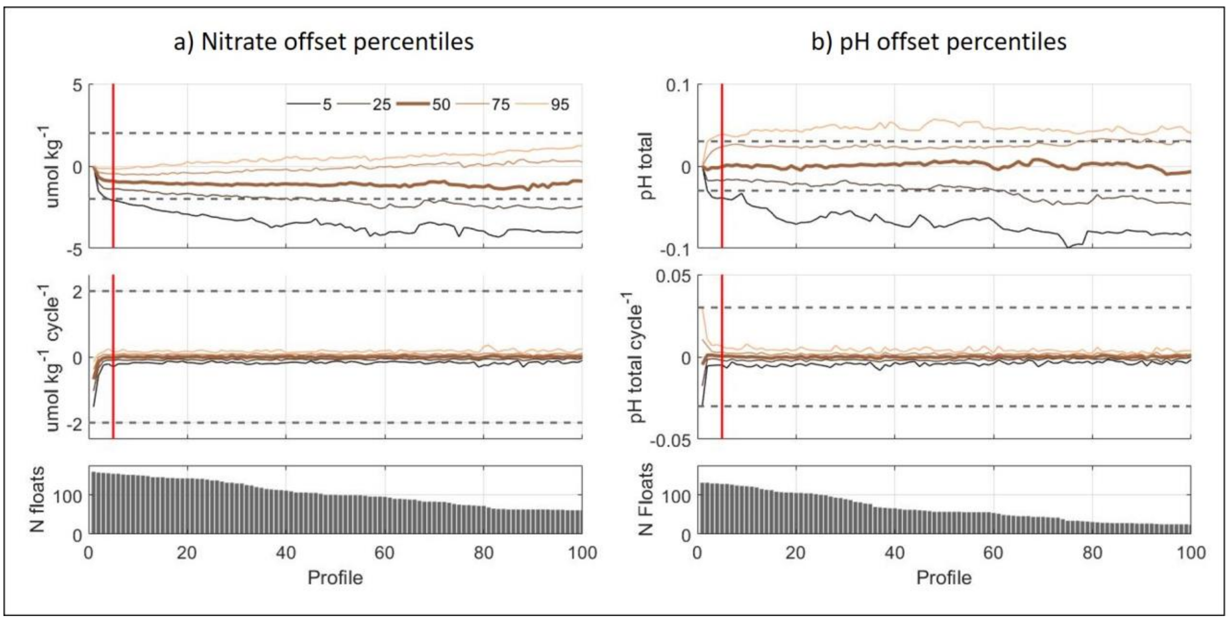

Figure 11. Nitrate (a) and pH (b) offset percentiles. Offsets are computed as the anomaly (float – reference) at each profile across the array. Top panels display offsets relative to profile one; center panels display the rate-of-change (first derivative) in offset from profile to profile (SOCCOM floats cycle at 10-day intervals); lower panels show the number of floats at each profile number.

While we see sensor stability improving with time since deployment for individual sensors, it is also important to understand if adjustment requirements across the array are improving over each subsequent deployment year. Fig. 12 shows box plots of the first offsets required for nitrate (left) and pH (right) data grouped by deployment year. Median offsets for nitrate seem to be more or less randomly distributed around zero. For pH, this is not the case. Median values remain negative over all deployment years which suggests a systematic negative bias for this sensor. pH sensor offset statistics also show a more dramatic change over time, in both the location of central tendency and degree of dispersion. These shifts in offset statistics are likely linked to changes in sensor design or laboratory calibration procedure. For example, significant improvements were seen in 2016 and 2018 in conjunction with the move to a thicker ISFET covering, and the switch from silver to platinum wire connections on the ISFET electrode, respectively. Beginning in 2016 the offset distributions are centered closer to zero than in previous years, and the 2018 distribution has a much tighter interquartile range, indicating more consistent sensor behavior.

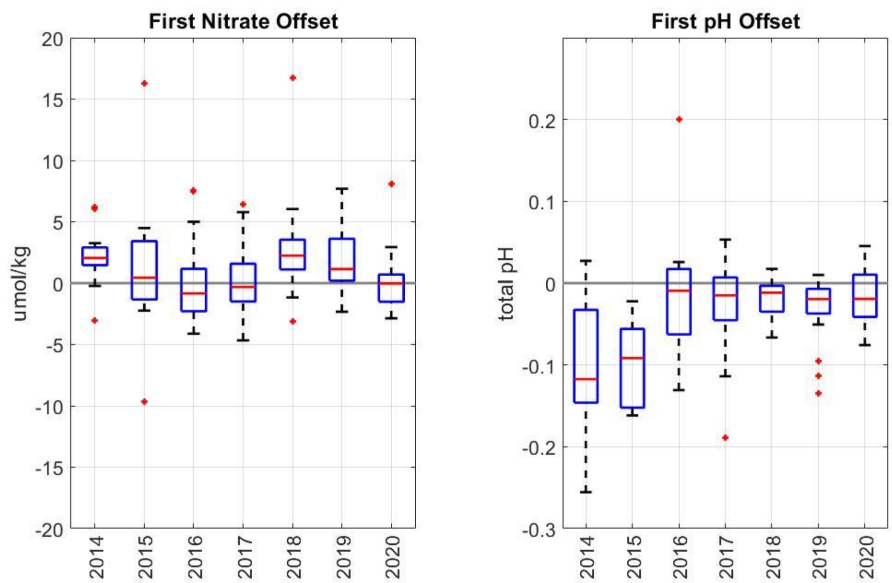

Figure 12. Boxplot summaries of first nitrate (left) and first pH (right) offsets, grouped by deployment year. Red lines represent the median, box boundaries represent the interquartile range (Q3 – Q1), whiskers are the outer range of data, excluding outliers (red stars) which are defined as data points that are larger than Q3+1.5*(Q3-Q1) or smaller than Q1-1.5*(Q3-Q1).

5 Validating SOCCOM float data adjustments

In this section, we discuss a system for validating our calibration methods. This involves comparison of post-corrected float data to data from both high-quality shipboard bottle casts taken alongside each SOCCOM float at the time of deployment, and nearby stations within the GLODAPv2 dataset (Olson et al., 2020). While shipboard data can also be useful for assessing initial offsets along a profile, it is not essential to float calibration and is typically reserved as an independent validation of the employed correction methods.

5.1 The use of shipboard bottle data

With the exception of oxygen calibration on Navis floats, the methods described in the previous sections for adjusting chemical data from a float do not depend on the existence of shipboard reference data collected alongside a float's deployment. This is advantageous in that any shipboard data taken at the time of deployment can be used to validate the applied in situ calibration methods. The SOCCOM program has required shipboard data collection alongside float deployment wherever possible to support the building of a robust validation dataset.

However, because it is not essential to sensor quality control, shipboard data collection may be reduced to select cruises in the future.

Comparisons of SOCCOM quality-controlled float data against shipboard data taken near the time of deployment are shown in Figs. 13 and 14. All float data have been interpolated onto the pressure axis of the hydrocast data. A portion of the error in the differences can be attributed to spatial and temporal changes in hydrography between the float profile and bottle samples. Float deployments typically occur as the ship begins heading away from a sampling station after the CTD rosette cast has been performed. This is done to reduce the chances of the ship running into the float. An additional lag time exists between deployment and when the float completes its first profile. Float-to-bottle matchups in the SOCCOM array are on average 23 hours and 8 km apart in time and space because of this. Nonetheless, the float to bottle matchups show very good agreement. The slope of the Model II regression for each parameter is indistinguishable from the 1:1 line. The median bottle-minus-float difference for oxygen, nitrate and pH are 0.35 µmol kg^{-1}, -0.12 µmol kg^{-1}, and 0.002 total pH units, respectively (Table 3). These values are very close to the accuracies reported in Johnson et al. (2017). Oxygen shows the largest improvement; this can likely be attributed to the implementation of the optode drift correction which was not yet accounted for at the time of the Johnson et al. (2017) publication.

Additionally, an independent analysis by Mignot et al. (2019) of quality controlled BGC-Argo float data in the Mediterranean Sea shows similar results, stating accuracies for oxygen and nitrate data (when compared to shipboard measurements) of 2.9 and 0.46 µmol kg^{-1}, respectively. Maximum depths reached by floats in the Mignot et al. (2019) analysis was 1000 m, as opposed to 2000 m on SOCCOM floats. The upper water column, therefore, made up a larger relative proportion of their float-to-bottle dataset; spatio-temporal mismatch due to greater oceanic variability at these depths likely accounts for the slightly larger biases observed.

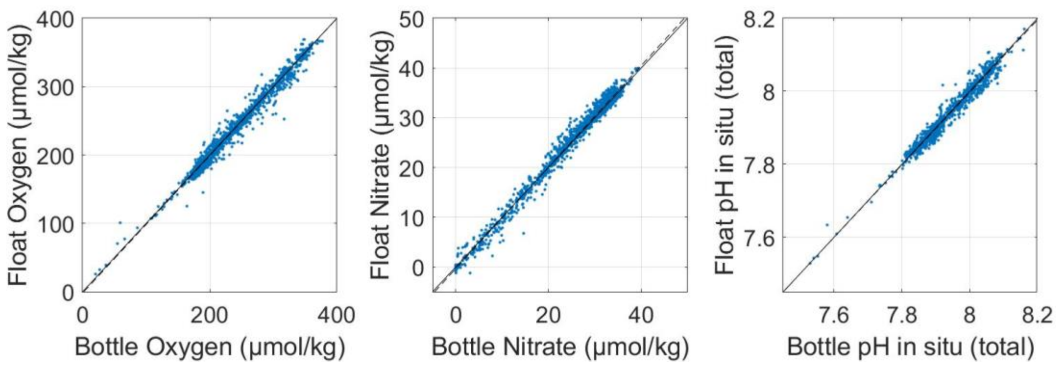

Figure 13. Scatter plots of float oxygen (left), nitrate (middle) and pH (right) data versus shipboard bottle data. The solid and dashed lines represent the 1:1 and Model II least squares fit, respectively.

	N observations	mean	median	standard deviation	maximum	minimum
Oxygen (μmol kg^{-1})	2366	0.94	0.35	6.84	64.50	-41.82
Nitrate (μmol kg^{-1})	2240	-0.22	-0.12	1.00	7.89	-5.14
pH (in-situ total)	1145	0.002	0.002	0.015	0.061	-0.096

Table 3. Bottle – minus – float matchup summary statistics for oxygen, nitrate and pH.

Float-bottle matchups in pressure space provide a validation of the assumption that sensor offsets are constant with depth (Johnson et al., 2013; 2016; 2017). Fig. 14 shows the bottle-minus-float differences for all oxygen, pH and nitrate matchups, plotted against pressure. The blue lines represent binned averages. There are no large trends in the oxygen or nitrate values with depth, confirming the assumptions in our calibration method. For pH, the pressure-binned distribution of mean differences show a negative bias of 5 millipH at depth. This bias changes sign toward the surface. Johnson et al. (2016) show a similar trend in comparison to discrete data (Fig. 6 in their publication, note trend is reversed as their plot represents float-minus-discrete) which they attribute to an incomplete understanding of carbonate-system thermodynamics at high pressures. While the magnitude of this bias is within the limits of stated uncertainty in the pH correction method (see section 4.2), the depth-dependent nature of the pH bias, as evident in the data, should be researched further.

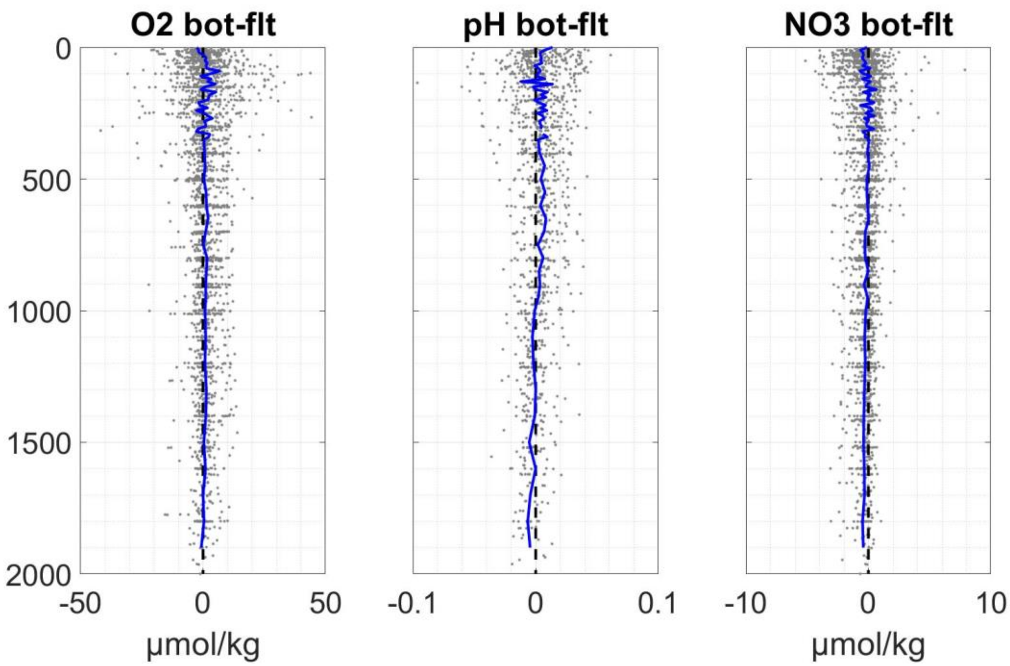

Figure 14. Scatterplots of bottle minus float matchups for oxygen (left), pH (center) and nitrate (right) data, plotted in depth space. Blue lines represent the mean of data within depth bins.

5.2 Comparisons to GLODAPv2

As described in the previous section, SOCCOM data quality validation is performed primarily in reference to shipboard hydrographic data taken at the time of deployment and is thus limited in scope to the initial profile returned from each float. Since in-situ drift is often observed in nitrate and pH (and to a lesser degree, oxygen) sensors onboard SOCCOM floats, a logical question is whether or not the quality of the applied adjustments remains stable throughout the duration of a float's life. For nitrate and pH, degradation in the quality of the adjustment over time could come from either a reduction in accuracy in one of the input parameters to the reference models (namely, temperature, salinity or oxygen), or a reduction in the accuracy of the reference algorithm itself due to gradual changes in deep ocean conditions that challenge the validity of the empirical relationships. The first possibility poses less of a threat, as temperature and salinity data on Argo floats are quite stable and require minimal adjustment. And, although drift is observed in some optodes onboard SOCCOM floats (see Section 3.2), comparison to a stable atmospheric reference provides a robust means for correction. The potential for degradation in data adjustment quality through time due to changes in the pressure or temperature coefficients of the sensor is more of a concern. If such changes in calibration occurred, then corrections derived at depth as the sensor aged would not be accurate near the surface.

The impacts from the issues described in the preceding paragraph can be assessed for the current SOCCOM dataset through an independent comparison of SOCCOM quality-controlled data at different stages of a float's life with hydrographic data from nearby stations in the GLODAPv2 dataset (Olson et al., 2020). Fig. 15 shows histograms of GLODAPv2 minus float data for oxygen, nitrate and pH crossovers within 20km distance of GLODAPv2 station data with no temporal restrictions; only data below 300 dbar were used to minimize discrepancies due to seasonal variability in the upper water column. The upper panels in the figure include comparisons from floats older than 6 months of age, and the lower panel includes data from floats greater than two years of age. A 4 μmol kg^{-1} and 0.02 pH bias between float and GLODAPv2 data can be observed for oxygen and pH data, respectively. The consistency of the biases for young (< 6 months) and old (> 2 years) floats are thus more likely a result of temporal differences between mean GLODAPv2 data used in the analysis and the corrected SOCCOM dataset. The mean age difference between the two datasets is 17.8 years. A 4 μmol/kg decrease in oxygen over nearly two decades (0.2 μmol/kg/y) is consistent with reported rates of oxygen change in the Southern Ocean that are based on shipboard data (Helm et al., 2011). Additionally, the observed rate of change in pH across this time frame (0.001 pH/y) is consistent with expected and observed rates of ocean pH decrease due to increasing atmospheric CO_2 (ocean acidification; Rios et al., 2015; Williams et al., 2018). Further, as is shown in Johnson et al. (2017), both the oxygen and pH biases increase with the mean age difference between the GLODAPv2 station time and the profiling float measurement time. This lends support to our hypothesis that the biases for oxygen and pH seen in Fig. 15 are the result of dynamic ocean change in the Southern Ocean in response to global climatic shifts (Bronselaer et al., 2020). This provides strong evidence that the quality control methods continue to be accurate over the lifetime of the float.

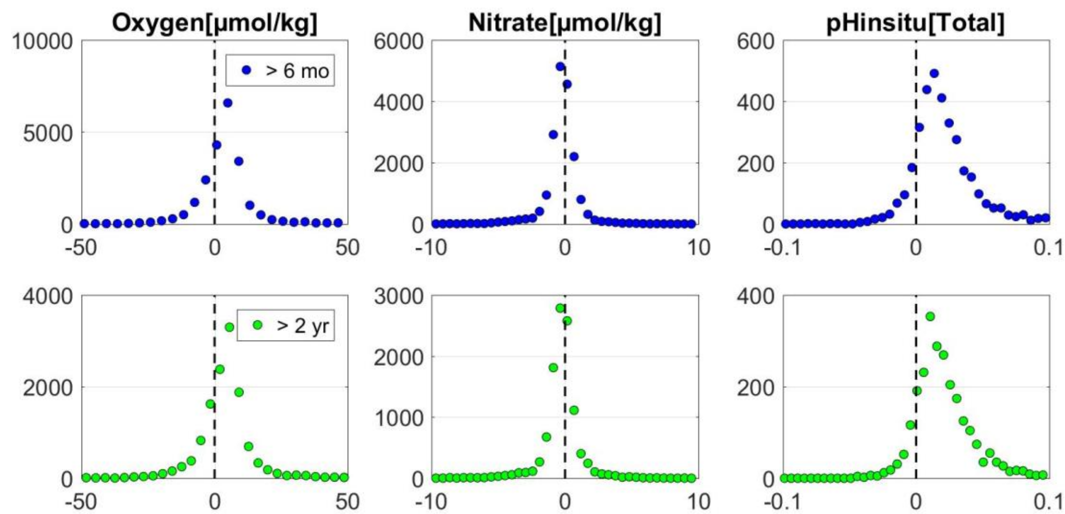

Figure 15. Histograms of GLODAPv2 minus quality-controlled oxygen (left), nitrate (center) and pH (right) float data. Upper panels include data from floats older than six months of age; lower panels include data from floats older than two years of age. Matchups were restricted to data that was within 20 km of GLODAPv2 reference stations.

6 Discussion and Conclusions

In this paper, we presented a coherent framework for applying delayed-mode adjustment procedures to oxygen, nitrate and pH data from SOCCOM biogeochemical profiling floats. The software GUIs presented, SAGE (SOCCOM Assessment and Graphical Evaluation) and SAGE-O_2, provide a robust way to visualize and assess the quality of these data. These software are open source and available through GitHub (https://github.com/SOCCOM-BGCArgo/ARGO_PROCESSING). The tools are intended to be used periodically throughout a float's life to reexamine sensor performance in delayed-mode. Adjustments derived using the software can then be applied to existing data and propagated forward in real-time until the next delayed-mode assessment is completed. A notable aspect of the procedure is in the relationship between the oxygen adjustment and that of nitrate and pH. The collective use of both SAGE-O_2 and SAGE offers a clear pathway to adjusted data for oxygen optodes, nitrate and pH sensors, all of which commonly coexist on biogeochemical profiling float platforms.

The successful expansion of the BGC-Argo program on a global scale, as described by Roemmich et al. (2019), depends partially on the implementation of standardized data adjustment methods across float platforms. The SAGE tools have already been adopted for use by other Argo data centers and are helping to increase the level of high-quality biogeochemical profiling float data available to users around the world. Although these software were developed specifically for the SOCCOM program, output files can be transformed to Argo NetCDF format via a separate processing pathway. Structuring the tools in this way has allowed for flexibility in adaptation across data centers. Additionally, this flexibility means that applications are not limited to Argo float data. The SAGE tools have the potential for use in post-deployment calibration of nitrate and pH data from other platforms such as gliders as well (Takeshita et al., 2020). As Bushinsky et al. (2019b) describe, sustaining multiple types of observational platforms in the ocean can increase our ability to resolve key processes at different spatial and temporal scales and in regions particularly susceptible to the effects of global change such as coral reef habitats and coastal upwelling zones. Ensuring that biogeochemical data is comparable across platforms is therefore essential.

Furthermore, along with performing repeated, standardized QC procedures it is important to run validation analysis, as described in Section 5, with regularity. This provides a metric for tracking improvements to sensor accuracy over time and testing the effects of processing upgrades or changes in QC methodology on the quality of the dataset. While data from biogeochemical sensors onboard profiling floats are revolutionizing capabilities in global ocean carbon research and modeling (Ford, 2020), the operational limitations of the sensors and the measurements they provide cannot be overlooked. Characterizing the uncertainties associated with such measurements helps to identify gaps in our understanding and guide future research and development. It is our hope that the calibration methods applied within the SOCCOM program, as outlined above, will serve as a global model for profiling float quality control, but also that the validation that follows will help to constrain the scientific questions that can be asked and provide inspiration for future research in both chemical sensor development and quality control.

Acknowledgments, Samples, and Data

All float data were collected and made freely available by the Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) project funded by the National Science Foundation, Division of Polar Programs (NSF PLR-1425989, with extension NSF OPP-1936222), supplemented by NASA, and by the International Argo Program and the NOAA programs that contribute to it. The Argo Program is part of the Global Ocean Observing System (https://doi.org/10.17882/42182, https://www.ocean-ops.org/board?t=argo). Additionally, all data analysis and supporting work performed at the Monterey Bay Aquarium Research Institute was supported in part by the David and Lucile Packard Foundation. The comparison in Section 4.2 was done using SOCCOM data archives https://doi.org/10.6075/J0QJ7FJP and https://doi.org/10.6075/J01G0JKT. Float data used for analysis in all other sections can be found at https://doi.org/10.6075/J0B27ST5. Raw msg files returned from SOCCOM floats are also freely available at ftp://ftp.mbari.org/pub/SOCCOM/RawFloatData/combined/. Shipboard data used in validation of SOCCOM float data is available through CCHDO (https://cchdo.ucsd.edu/search?q=soccom). World Ocean Atlas data used in this study can be found at https://www.ncei.noaa.gov/data/oceans/woa/WOA18/. GLODAPv2 data used in this study can be found at https://www.glodap.info/index.php/merged-and-adjusted-data-product/. NCEP Reanalysis data was provided by the NOAA/OAR/ESRL PSL, Boulder, Colorado, USA, from their Web site at https://psl.noaa.gov/. We would like to thank all people who contributed to the production, calibration and assembly of all sensors and floats used within the SOCCOM program, as well as all personnel involved in their transport and deployment at sea. We gratefully acknowledge contributions provided by Yui Takeshita and Annie Wong during the initial revision of the manuscript. We would also like to thank MBARI intern, Tatiana Ellis, for her contributions to the testing of automated change-point-detection within the SAGE GUI framework. MATLAB code for the SAGE and SAGE-O_2 software tools is freely available at https://github.com/SOCCOM-BGCArgo/ARGO_PROCESSING/ (release 2.0).

References

Aanderaa Instruments. (2017), TD269 Operating Manual, oxygen optode 4330, 4831, 4835, Seventh Ed.

Benson, B. & Krause, D. (1984), The concentration and isotopic fractionation of oxygen dissolved in freshwater and seawater in equilibrium with the atmosphere. *Limnology and Oceanography,* 29: 620-632.

Bianchi, D., Dunne, J. P., Sarmiento, J. L., & Galbraith, E. D. (2012), Data-based estimates of suboxia, denitrification, and N_2O production in the ocean and their sensitivities to dissolved O_2. *Global Biogeochemical Cycles*, 26, GB2009, doi:10.1029/2011GB004209

Bittig, H. C., Fiedler, B., Scholz, R., Krahmann, G., & Körtzinger, A. (2014), Time response of oxygen optodes on profiling platforms and its dependence on flow speed and temperature. *Limnology and Oceanography: Methods*, 12, doi: 10.4319/lom.2014.12.617.

Bittig, H.C., & Körtzinger, A. (2015), Tackling Oxygen Optode Drift: Near-Surface and In-Air Oxygen Optode Measurements on a Float Provide an Accurate in Situ Reference. *Journal of Atmospheric and Oceanic Technology*, 32, 1536–1543, https://doi.org/10.1175/JTECH-D-14-00162.1

Bittig, H. C., & Körtzinger, A. (2017), Technical note: update on response times, in-air measurements, and in situ drift for oxygen optodes on profiling platforms. *Ocean Science.* 13, 1–11. doi: 10.5194/os-13-1-2017

Bittig, H. C., Körtzinger, A., Neill, C., van Ooijen, E., Plant, J. N., Hahn, J., Johnson, K., S., Yang, B., & Emerson, S. R. (2018a), Oxygen optode sensors: principle, characterization, calibration and application in the ocean. *Frontiers in Marine Science*, 4:429. doi: 10.3389/fmars.2017.00429

Bittig H.C., Steinhoff, T., Claustre, H., Fiedler, B., Williams, N. L, Sauzède, R., Körtzinger, A., & Gattuso, J.P. (2018b), An Alternative to Static Climatologies: Robust Estimation of Open Ocean CO_2 Variables and Nutrient Concentrations From T, S, and O_2 Data Using Bayesian Neural Networks. *Frontiers in Marine Science*. 5:328. doi: 10.3389/fmars.2018.00328

Bittig, H.C., Maurer, T.L., Plant, J.N., Schmechtig, C., Wong, A.P.S., Claustre, H., Trull, T.W., Udaya Bhaskar, T.V.S., Boss, E., Dall'Olmo, G., Organelli, E., Poteau, A., Johnson, K.S., Hanstein, C., Leymarie, E., Le Reste, S., Riser, S.C., Rupan, A.R., Taillandier, V., Thierry, V. & Xing, X. (2019), A BGC-Argo Guide: Planning, Deployment, Data Handling and Usage. *Frontiers in Marine Science,* 6:502. doi: 10.3389/fmars.2019.00502

Bronselaer, B., Winton, M., Griffies, S.M., Stouffer, R.J., Hurlin, W.J., Sergienko, O.V., Rodgers, K., & Russell, J. (2018), Change in future climate due to Antarctic meltwater. *Nature,* 564, 53–58. https://doi.org/10.1038/s41586-018-0712-z

Bushinsky, S. M., Emerson, S. R., Riser, S. C., & Swift, D. D. (2016), Accurate oxygen measurements on modified Argo floats using in situ air calibrations. *Limnology and Oceanography Methods*, 14: 491-505. doi:10.1002/lom3.10107

Bushinsky, S. M., Landschützer, P., Rödenbeck, C., Gray, A. R., Baker, D., Mazloff, M. R., et al. (2019a), Reassessing Southern Ocean air-sea CO_2 flux estimates with the addition of biogeochemical float observations. *Global Biogeochemical Cycles*. 33, 1370–1388. https://doi.org/10.1029/2019GB006176

Bushinsky, S.M., Takeshita, Y., & Williams, N.L. (2019b), Observing Changes in Ocean Carbonate Chemistry: Our Autonomous Future. *Current Climate Change Reports.* 5, 207–220 https://doi.org/10.1007/s40641-019-00129-8

Carter, B.R., Feely, R.A., Williams, N.L., Dickson, A.G., Fong, M.B., & Takeshita, Y. (2018), Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate. *Limnology and Oceanography Methods*, 16: 119-131. doi:10.1002/lom3.10232

D'Asaro, E. A., & McNeil, C. (2013), Calibration and Stability of Oxygen Sensors on Autonomous Floats. *Journal of Atmospheric and Oceanic Technology*, 30, pp. 1896-1906

Drucker, R., & Riser, S. C. (2016), In situ phase-domain calibration of oxygen Optodes on profiling floats. *Methods in Oceanography*, 17, 296-318, https://doi.org/10.1016/j.mio.2016.09.007.

Fabozzi, F. J., Focardi, S. M., Rachev, S. T., & Arshanapalli, B. G. (2014), The Basics of Financial Econometrics: Tools, Concepts, and Asset Management Applications. Appendix E: Model Selection Criterion: AIC and BIC. Hoboken, New Jersey: Wiley & Sons, Inc.

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

Gaillard, F., Autret, E., Thierry, V., Galaup, P., Coatanoan, C., & Loubrieu, T. (2009), Quality control of large Argo datasets. *Journal of Atmospheric and Oceanic Technology*, 26, 337-351, doi:10.1175/ 2008JTECHO552.1.

Garcia, H. E. & Gordon, L. I. (1992), Oxygen solubility in seawater: Better fitting equations. *Limnology and Oceanography*, 37, doi: 10.4319/lo.1992.37.6.1307.

Gordon, C., Fennel, K., Richards, C., Shay, L. K., & Brewster, J. K. (2020), Can ocean community production and respiration be determined by measuring high-frequency oxygen profiles from autonomous floats? *Biogeosciences*, 17, 4119-4134, doi: 10.5194/bg-17-4119-2020

Gray, A., Johnson, K. S., Bushinsky, S. M., Riser, S. C., Russell, J. L., Talley, L. D., Wanninkhof, R., Williams, N. L., & Sarmiento, J. L. (2018), Autonomous biogeochemical floats detect significant carbon dioxide outgassing in the high-latitude Southern Ocean. *Geophysical Research Letters*, 45, 9049–9057. https://doi.org/10.1029/2018GL078013

Guinehut, S., Coatanoan, C., Dhomps, A. L., Le Traon, P. Y., & Larnicol, G. (2009), On the use of satellite altimeter data in Argo quality control. *Journal of Atmospheric and Oceanic Technology*, 26, 395-402. https://doi.org/10.1175/2008JTECHO648.1

Helm, K. P., Bindoff, N. L., & Church, J. A. (2011), Observed decreases in oxygen content of the global ocean, *Geophysical Research Letters*, 38, L23602, doi:10.1029/2011GL049513.

Hood, E. M., Sabine, C. L., & Sloyan, B. M. eds. (2010), The GO-SHIP Repeat Hydrography Manual: A Collection of Expert Reports and Guidelines. IOCCP Report Number 14, ICPO Publication Series Number 134. Available online at http://www.go-ship.org/HydroMan.html.

Johnson, K. S., Coletti, L., Jannasch, H., Sakamoto, C., Swift, D. & Riser, S. (2013), Long-term nitrate measurements in the ocean using the in situ ultraviolet spectrophotometer: Sensor integration into the apex profiling float, *Journal of Atmospheric and Oceanic Technology*, 30, 1854– 1866, doi:10.1175/JTECH-D-12-00221.1.

Johnson, K. S., Plant, J. N., Riser, S. C., & Gilbert, D. (2015), Air oxygen calibration of oxygen optodes on a profiling float array. *Journal of Atmospheric and Oceanic Technology*. 32, 2160-2172.

Johnson, K. S., Jannasch, H. W., Coletti, L. J., Elrod, V. A., Martz, T. R., Takeshita, Y., Carlson, R. J., & Connery J. J. (2016), Deep-Sea DuraFET: A pressure tolerant pH sensor designed for global sensor networks, *Analytical Chemistry*. doi:10.1021/acs.analchem.5b04653.

Johnson, K. S., Plant, J. N., Coletti, L. J., Jannasch, H. W., Sakamoto, C. M., Riser, S. C., Swift, D. D., Williams, N. L., Boss, E., Haëntjens, N., Talley, L. D., & Sarmiento, J. L. (2017), Biogeochemical sensor performance in the SOCCOm profiling float array. *Journal of Geophysical Research Oceans*, 122, 6416–6436, doi:10.1002/2017JC012838.

Johnson, K. S., Pasqueron De Fommervault, O., Serra, R., D'Ortenzio, F., Schmechtig, C., Claustre, H., & Poteau, A. (2018a), Processing Bio-Argo nitrate concentration at the DAC Level. Argo data management. https://doi.org/10.13155/46121

Johnson K. S., Plant J. N., & Maurer T. L., (2018b), Processing BGC-Argo pH data at the DAC level. Argo data management. https://doi.org/10.13155/57195

Johnson, K. S., Riser, S. C., & Ravichandran, M. (2019), Oxygen variability controls denitrification in the Bay of Bengal oxygen minimum zone. *Geophysical Research Letters*, 46, 804– 811. https://doi.org/10.1029/2018GL079881

Kalnay, E. M., and Coauthors. (1996), The NCEP/NCAR 40-Year Reanalysis Project. Bull. *American Meteorological Society*, 77, 437–470, doi:10.1175/1520-0477(1996)077,0437:TNYRP.2.0.CO;2.

Kass, R. E., & Raferty, A. E. (1995), Bayes Factors. *Journal of the American Statistical Association*, 90: 773-795.

Killick R., Fearnhead, P., & Eckley, I.A. (2012), Optimal detection of changepoints with a linear computational cost. *Journal of the American Statistical Association*, 107: 1590-1598.

Mignot, A., D'Ortenzio, F., Taillandier, V., Cossarini, G., & Salon, S. (2019), Quantifying observational errors in Biogeochemical-Argo oxygen, nitrate, and chlorophyll *a* concentrations. *Geophysical Research Letters*, 46, 4330– 4337. https://doi.org/10.1029/2018GL080541

Nicholson, D. P., & Feen, M. L. (2017), Air calibration of an oxygen optode on an underwater glider. *Limnology and Oceanography Methods*, 15: 495-502. doi:10.1002/lom3.10177

Olsen, A., Lange, N., Key, R. M., Tanhua, T., Bittig, H. C., Kozyr, A., Àlvarez, M., Azetsu-Scott, K., Becker, S., Brown, P. J., Carter, B. R., Cotrim da Cunha, L., Feely, R. A., van Heuven, S., Hoppema, M., Ishii, M., Jeansson, E., Jutterström, S., Landa, C. S., Lauvset, S. K., Michaelis, P., Murata, A., Pérez, F. F., Pfeil, B., Schirnick, C., Steinfeldt, R., Suzuki, T., Tilbrook, B., Velo, A., Wanninkhof, R., & Woosley, R. J. (2020), GLODAPv2.2020 – the second update of GLODAPv2. doi:10.5194/essd-2020-165.

Owens, B. W., & Wong, A. P. S. (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 – Ocean Research Papers*, 56, 450-457. doi:10.1016/j.dsr.2008.09.008.

Ríos, A. F., Resplandy, L., García-Ibáñez, M. I., Fajar, N. M., Velo, A., Padin, X. A., et al. (2015), Decadal acidification in the water masses of the Atlantic Ocean. *Proceedings of the National Academy of Sciences*, 112(32), 9950–9955. https://doi.org/10.1073/pnas.1504613112

Riser, S. C., Swift, D., & Drucker, R. (2018), Profiling floats in SOCCOM: Technical capabilities for studying the Southern Ocean. *Journal of Geophysical Research: Oceans*,123, 4055–4073. https://doi.org/10.1002/2017JC013419

Roemmich, D., Alford, M. H., Claustre, H., Johnson, K., King, B., Moum, J., Oke, P., Owens, W. B., Pouliquen, S., Purkey, S., Scanderbeg, M., Suga, T., Wijffels, S., Zilberman, N., Bakker, D., Baringer, M., Belbeoch, M., Bittig, H. C., Boss, E., Calil, P., Carse, F., Carval, T., Chai, F., Conchubhair, D. O., d'Ortenzio, F., Dall'Olmo, G., Desbruyeres, D., Fennel, K., Fer, I., Ferrari, R., Forget, G., Freeland, H., Fujiki, T., Gehlen, M., Greenan, B., Hallberg, R., Hibiya, T., Hosoda, S., Jayne, S., Jochum, M., Johnson, G. C., Kang, K., Kolodziejczyk, N., Körtzinger, A., Le Traon, P.-Y., Lenn, Y.-D., Maze, G., Mork, K. A., Morris, T., Nagai, T., Nash, J., Garabato, A. N., Olsen, A., Pattabhi, R. R., Prakash, S., Riser, S., Schmechtig, C., Schmid, C., Shroyer, E., Sterl, A., Sutton, P., Talley, L., Tanhua, T., Thierry, V., Thomalla, S., Toole, J., Troisi, A., Trull, T. W., Turton, J., Velez-Belchi, P. J., Walczowski, W., Wang, H., Wanninkhof, R., Waterhouse, A. F., Waterman, S., Watson, A., Wilson, C., Wong, A. P. S., Xu, J., & Yasuda, I. (2019). On the Future of Argo: A Global, Full-Depth, Multi-Disciplinary Array, *Frontiers in Marine Science*, 6, 439, https://doi.org/10.3389/fmars.2019.00439

Russell, J. L., Kamenkovich, I., Bitz, C., Ferrari, R., Gille, S. T., Goodman, P. J., Hallberg, R., Johnson, K., Khazmutdinova, K., Marinov, I., Mazloff, M., Riser, S., Sarmiento, J. L., Speer, K., Talley, L. D. & Wannikhof, R. (2018), Metrics for the Evaluation of the Southern Ocean in Coupled Climate Models and Earth System Models. *Journal of Geophysical Research: Oceans*, 123, 3120-3143. https://doi.org/10.1002/2017JC013461

Salstein, D. A., Ponte, R. M., & Cady-Pereira, K. (2018), Uncertainties in atmospheric surface pressure fields from global analyses, J. Geophys. Res., 113, D14107, doi:10.1029/2007JD009531, 2008.

Sampson, D. (2021), GUI Layout Toolbox (https://www.mathworks.com/matlabcentral/fileexchange/47982-gui-layout-toolbox), MATLAB Central File Exchange.

Sauzède R., Bittig, H.C., Claustre, H., Pasqueron de Fommervault, O., Gattuso, J-P, Legendre, L., & Johnson, K.S. (2017), Estimates of Water-Column Nutrient Concentrations and Carbonate System Parameters in the Global Ocean: A Novel Approach Based on Neural Networks. *Frontiers in Marine Science,* 4:128. doi: 10.3389/fmars.2017.00128

Schmechtig, C., Thierry, V., & The Bio Argo Team. (2016), Argo quality control manual for biogeochemical data. https://doi.org/10.13155/40879

Swart, S., Gille, S. T., Delille, B., Josey, S., Mazloff, M., Newman, L., Thompson, A.F., Thomson, J., Ward, B., du Plessis, M.D., Kent, E.C., Girton, J., Gregor, L., Heil, P., Hyder, P., Pezzi, L.P., de Souza, R.B., Tamsitt, V., Weller, R.A., & Zappa, C.J. (2019), Constraining Southern Ocean Air-Sea-Ice Fluxes Through Enhanced Observations. *Frontiers in Marine Science*, 6:421. doi: 10.3389/fmars.2019.00421

Takeshita, Y., Martz, T. R., Johnson, K. S., Plant, J. N., Gilbert, D., Riser, S. C., Neill, C., & Tilbrook, B. (2013), A climatology-based quality control procedure for profiling float oxygen data. *Journal of Geophysical Research Oceans*, 118, 5640-5650. doi:10.1002/jgrc.20399

Takeshita, Y., Jones, B. D., Johnson, K. S., Chavez, F. P., Rudnick, D. L., Blum, M., Conner, K., Jensen, S., Long, J. S., Maughan, T., Mertz, K. L., Sherman, J. T., & Warren, J. (2020), Accurate pH and O2 measurements from Spray underwater gliders. *Journal of Atmospheric and Oceanic Technology*, https://doi.org/10.1175/JTECH-D-20-0095.1

Thierry, V, Bittig, H., Gilbert, D., Kobayashi, T., Kanako, S., & Schmid, C. (2018), Processing Argo oxygen data at the DAC level. https://doi.org/10.13155/39795

Thierry V., Bittig., The Argo-BGC Team (2018b), Argo quality control manual for dissolved oxygen concentration. https://doi.org/10.13155/46542

Verdy, A., & Mazloff, M. R. (2017), A data assimilating model for estimating Southern Ocean biogeochemistry. *Journal of Geophysical Research Oceans*, 122, 6968–6988, doi:10.1002/2016JC012650.

Williams, N. L., Juranek, L. W., Johnson, K. S., Feely, R. A., Riser, S. C., Talley, L. D., Russell, J. L., Sarmiento, J. L. & Wanninkhof, R. (2016), Empirical algorithms to estimate water column pH in the Southern Ocean. Geophys. Res. Lett., 43, 3415–3422, doi:10.1002/2016GL068539.

Williams, N. L., Juranek, L. W., Feely, R. A., Russell, J. L., Johnson, K. S., & Hales, B. (2018), Assessment of the carbonate chemistry seasonal cycles in the Southern Ocean from persistent observational platforms. *Journal of Geophysical Research: Oceans*, 123, 4833–4852. https://doi.org/10.1029/2017JC012917

Wong, A. P. S., Johnson, G. C., & Owens, W. B. (2003), Delayed-mode calibration of autonomous CTD profiling float salinity data by θ-S climatology. *Journal of Atmospheric and Oceanic Technology*, 20, 308-318. https://doi.org/10.1175/1520-0426(2003)020,0308: DMCOAC.2.0.CO;2.

Wong, A. P. S. & Riser, S. C. (2011), Profiling Float Observations of the Upper Ocean under Sea Ice off the Wilkes Land Coast of Antarctica. *Journal of Physical Oceanography*, 41(6), 1102– 1115.

Wong, A.P.S., S.E. Wijffels, S.C. Riser, S. Pouliquen, S. Hosoda, D. Roemmich, J. Gilson, G.C. Johnson, K. Martini, D.J. Murphy, M. Scanderbeg, T.V.S.U. Bhaskar, J.J.H. Buck, F. Merceur, T. Carval, G. Maze, C. Cabanes, X. André, N. Poffa, I. Yashayaev, P.M. Barker, S. Guinehut, M. Belbéoch, M. Ignaszewski, M.O.N. Baringer, C. Schmid, J.M. Lyman, K.E. McTaggart, S.G. Purkey, N. Zilberman, M.B. Alkire, D. Swift, W.B. Owens, S.R. Jayne, C. Hersh, P. Robbins, D. West-Mack, F. Bahr, S. Yoshida, P.J.H. Sutton, R. Cancouët, C. Coatanoan, D. Dobbler, A. Garcia Juan, J. Gourrion, N. Kolodziejczyk, V. Bernard, B. Bourlès, H. Claustre, F. D'Ortenzio, S. Le Reste, P.-Y. Le Traon, J.-P. Rannou, C. Saout-Grit, S. Speich, V. Thierry, N. Verbrugge, I.M. Angel-Benavides, B. Klein, G. Notarstefano, P.-M. Poulain, P. Vélez-Belchí, T. Suga, K. Ando, N. Iwasaska, T. Kobayashi, S. Masuda, E. Oka, K. Sato, T. Nakamura, K. Sato, Y. Takatsuki, T. Yoshida, R. Cowley, J.L. Lovell, P.R. Oke, E.M. van Wijk, F. Carse, M. Donnelly, W.J. Gould, K. Gowers, B.A. King, S.G. Loch, M. Mowat, J. Turton, E.P. Rama Rao, M. Ravichandran, H.J. Freeland, I. Gaboury, D. Gilbert, B.J.W. Greenan, M. Ouellet, T. Ross, A. Tran, M. Dong, Z. Liu, J. Xu, K. Kang, H. Jo, S.-D. Kim & H.-M. Park (2020), Argo Data 1999–2019: Two Million Temperature-Salinity Profiles and Subsurface Velocity Observations from a Global Array of Profiling Floats, *Frontiers in Marine Science*, **7**, https://doi.org/10.3389/fmars.2020.00700