RECOMMENDATIONS FOR ENHANCEMENT OF pH QC METHODS

Ref.: D4.7_V2.0

Date: 25/07/2022

Euro-Argo Research Infrastructure Sustainability and Enhancement Project (EA RISE Project) - 824131

This project has received funding from the European Union's Horizon 2020
research and innovation programme under grant agreement no 824131.
Call INFRADEV-03-2018-2019: Individual support to ESFRI
and other world-class research infrastructures

Document Reference

Project	**Euro-Argo RISE - 824131**
Deliverable number	**D4.7**
Deliverable title	Recommendations for enhancement of pH QC Methods
Description	**Report**
Work Package number	**4**
Work Package title	**Biogeochemical Extension**
Lead Institute	**GEOMAR**
Lead authors	**Arne Körtzinger, Tobias Steinhoff, Cathy Wimart-Rousseau, Katharina Seelmann**
Contributors	**Fabrizio D'Ortenzio, Catherine Schmechtig**
Submission date	**25.07.2022**
Due date	**31.03.2022**
Comments	**Further delay due to Covid-related issues**
Accepted by	**Fabrizio D'Ortenzio**

Document History

Version	Issue Date	Author	Comments
Ver. 1.0	04.07.2022	Arne Körtzinger	
Ver. 2.0	25.07.2022	Arne Körtzinger	Final version

EXECUTIVE SUMMARY

BGC-Argo achieved access to the marine CO_2 system through the addition of the SeaFET pH sensor to the suite of BGC sensors. This in principle opens the possibility to detect variability and long-term changes in interior ocean inorganic carbon storage and quantify the ocean sink for atmospheric CO_2. Due to the huge inorganic carbon background in the ocean, detection of the anthropogenic changes requires exorbitantly high measurement accuracy (typically about 1 ‰) for marine CO_2 system variables. From a sensor perspective, the marine CO_2 system is therefore a particularly challenging case. This puts high emphasis on the quality control of float-based pH measurements. By applying established Argo cookbook routines to the pH measurements from a pH/O_2 float pilot study in the subpolar North Atlantic Ocean, we detected significant issues with the achievable pH accuracy. We note that for our floats the two proposed referencing methods (LIPHR and CANON-B) rather strongly disagree by about 0.02 pH units. In our analysis we applied additional and completely independent QC methods which are based on crossovers of float data with data from discrete hydrocasts and Ship-of-Opportunity observations. Our results are drastically different between apparently well-functioning pH sensors and a production series of pH sensors with major manufacturing issues. Results for the latter therefore have to be viewed with caution. Results for the former, however, show that the minimum accuracy requirement of 0.01 pH units (equivalent to the minimum pCO_2 accuracy for inclusion into SOCAT of 10 µatm) is not systematically met in the upper ocean. This is particulary important in the context of converting surface ocean pH into pCO_2 for the purpose of calculating of the air-sea CO_2 flux. Our results point at the possibility of inacceptably high pH biases in the surface ocean which call for an additional independent reference in the surface ocean. We show evidence that the existing Ship-of-Opportunity observatory could potentially serve this purpose in certain regions. Given its large synergy potential we therefore propose a systematic coupling of the SOOP and BGC-Argo observatories. In addition, we stress the need of improving both pH sensor performance characteristics and manufacturing quality.

Disclaimer:

This document represents the situation at the time of data evaluation and writing of the report which is primarily based on data from 2021. Unfortunately, all pH sensors of the 5 pH/O_2 floats deployed in 2021 are affected by the manufacturing problem that leads to deteriorated data quality and early senor failure. This has both reduced the amount and compromised the quality of the data coming from the 5 floats. Our results are therefore based largely on 3 earlier pH/O_2 floats with much higher endurance and data quality.

However, as pH floats and SOOP line are still in operation more data are coming in and the database is growing daily. This will allow us to improve the statistics of our analyses and hence the robustness of the results. Therefore the results presented here are based on the status quo and are not necessarily the final word on these matters. We therefore point out that further analyses will and need to be carried out.

TABLE OF CONTENT

1 Background

The emerging Biogeochemical Argo Programme (BGC-Argo) holds the potential to revolutionize Marine Biogeochemisty in a way similar to how the classical Argo revolutionized Physical Oceanography. One particular area of interest for BGC-Argo is the possibility to observe properties of the marine CO_2 system. This in principle would allow quantification of the air-sea flux of CO_2 and monitoring of temporal changes in the ocean's inorganic carbon pool. The latter could provide insight into changes in the natural components of the marine carbon cycle, i.e. primary production, organic matter decomposition, vertical particle flux attenuation etc. It would also allow insight into the anthropogenically perturbed part of the marine carbon cycle, both in terms of the passive uptake and storage of anthropogenic CO_2 as well as in terms of monitoring and verification of carbon dioxide removal techniques.

1.1 Chemistry of the Marine Carbon Dioxide System

A series of chemical equilibria govern the dissolution of CO_2 in water and its reaction with water. This gives rises to a suite a different chemical species which are connected through the corresponding equilibrium reactions. These are CO_2 in aquatic solution, carbonic acid, bicarbonate and carbonate (i.e., $CO_2(aq)$, H_2CO_3, HCO_3^-, CO_3^{2-}). Unfortunately, the concentrations of these individual species of the carbon dioxide system in solution can not be measured directly. There are, however, four parameters that can be measured at high accuracy. These are used, together with ancillary information, to obtain a complete description of the CO_2 system in seawater (Dickson et al., 2007).

These four parameters are :

- Total dissolved inorganic carbon (DIC)
- Total alkalinity (TA)
- Fugacity/partial pressure of CO_2 in gas phase in chemical equilibrium with seawater (fCO_2/pCO_2)
- Total hydrogen ion concentration (pH)

1.2 Marine Carbon Dioxide System Parameter of Choice for BGC-Argo

It is possible, in theory, to obtain a complete description of the marine CO_2 system in a sample of sea water at a particular temperature and pressure provided that the following information is known (Dickson et al., 2007):

- the solubility constant for CO_2 in sea water (K_0),
- the equilibrium constants for each of the acid-base pairs that are assumed to exist in the solution,
- the total concentrations of all the non-CO_2 acid-base pairs,
- the values of at least two of the CO_2 related parameters: DIC, TA , f/pCO_2, pH

From a scientific perspective it would thus be highly desirable to always measure two of the four parameters of the marine CO_2 system. Since the other pieces of required information listed above are mostly available this would allow a full chemical characterization of the marine CO_2 system as the basis for in-depth interpretation of the results obtained. In practice, however, this is limited by the availability of suitable sensors for the Argo float platform.

Early tests with state-of-the-art pCO_2 sensors demonstrated both the potential and the significant obstacles (high power demand, slow response time, long-term drift etc.) of float-based pCO_2 observations (Fiedler et al., 2013). Hopes that pCO_2 optode technology might mature in a way the oxygen optode technology did, which would potentially allow in-air pCO_2 measurements to be used for in-situ offset and drift correction during a float's lifetime similiar how this has been implemented for oxygen (Bittig and Körtzinger, 2015 ; 2017), did not come true.

The successful testing of the Honeywell Durafet ion-sensitive field-effect transistor for seawater pH by Martz et al. (2010) has made the ISFET the most promising sensor technology to access the marine CO_2 system from autonomous platforms such as floats and gliders. Significant engineering has been put into this sensor technology to become the Deep-Sea DuraFET pH sensor that has successfully been integrated into state-of-the-art Argo floats (Johnson et al., 2016).

It should be noted, that currently the SeaFET pH is the only suitable, tested and commercially available sensor for float-based observations of the marine CO_2 system. While pH is rightly being viewed as a master variable in marine chemistry and potentially allows direct observations of the phenomenon of ocean acidification, its measurement does not directly allow observation of other, arguably even more desired quantities such as the air-sea CO_2 flux or changes in the inorganic carbon inventory of the ocean. For quantification of these, pH observations need to be combined with data for a second marine CO_2 system variable. In the absence of a suitable sensor for $p\mathrm{CO}_2$, DIC or TA a reasonable work-around is to predict TA using algorithms that employ variables such as T, S, pressure (P), and O_2 which are also measured on floats. One such option for this alkalinity estimation is LIAR (Locally Interpolated Alkalinity Regression; Carter et al., 2016), which uses data from the Global Ocean Data Analysis Project (GLODAPv2) data set (Olsen et al., 2016). Generally, TA predictions have a typical average uncertainty on the order of 6 µmol kg^{-1} (e.g., Williams et al., 2017). As the data coverage of the GLODAPv2 data product is far from perfect both in terms of spatial and annual variability (seasonal cycle) the robustness of TA algorithm based on it (or other data products) by necessity is globally not uniform. Regional and/or seasonal biases are therefore to be expected which have the potential to compromise the accuracy of DIC and $p\mathrm{CO}_2$ calculated from measured pH and predicted TA.

To overcome this limitation, we should continue as a community to develop a suitable sensor for reliable float-based observations of a second marine CO_2 system variable allowing for complete analytical characterization of the system from float measurements alone.

It should be noted that significant efforts are underway to also advance the optode technology for pH. While earlier versions of the pH optode had significant issues with, for example, drift and response time, recent advances (e.g., AquapHOx platform with pH optode, PyroScience GmbH, Aachen, Germany) are promising which may warrent dedicated field testing.

2 Concepts for Correction of Float pH

Similar to other chemical sensors such as oxygen optodes, even well-calibrated pH sensors require a number of corrections to achieve adequate data accuracy. These, among other steps, involve initial offset correction and long-term drift correction. The general philosophy for pH correction employs empirical algorithms based on high-quality GO-SHIP hydrographic measurements to estimate water column pH (Williams et al., 2017). This yields a single calibration point at depth (typically 1480-1520 dbar) which is then applied as a uniform offset correction in the pH sensor reference potential across the entire water column. In lack of a second calibration point (e.g., at the surface), this correction has to therefore assume that the at-depth correction applies throughout the water column, i.e. across the entire gradients in T, S, and P.

2.1 Cookbook Procedure for Correction of Float pH (Johnson et al., 2017)

Internal sensor algorithms convert the measured potential into pH on the total proton scale using laboratory based calibration coefficents. This pH undergoes automatic real time checks which include a global range (7.3-8.5) and a spike check to flag out-of-range data and outliers as bad data. This real-time data (RT) typically does not yield research-quality pH. Therefore almost all float pH data must receive an adjustment which is calculated as part of the delayed-mode quality control (and can be implemented as a – if necessary regularly updated – real-time adjustment). The following corrections are discussed in the context of float pH quality control:

- Pump offset: This pH offset is frequently observed when the pump of the CTD flow-path into which the pH sensor is integrated is switched between off and on during profiling. As some float profiling routines change from unpumped to pumped pH measurement mode at some point during the profiling (e.g., Apex float) a jump in pH can be observed that requires correction. This offset is explained by a streaming potential created by the flowing seawater that depends on sensor geometry and the electronics which makes it specific for each sensor.
- Gain correction: According to current understanding the SeaFET pH does not require a gain correction. It is therefore typically set to 1.
- Offset correction: This correction is most critical as it corrects for the typically significant (negative) offset that is observed with float-based pH measurements despite the effort that routinely goes into individual sensor calibration with the manufacturer. The correction is only possible against an external reference and involves a criticial assumption. Several variants of this correction have been proposed and a we have also developed our own Matlab algorithm for this purpose. These variants have in common that they employ high-quality hydrographic data sets with discrete pH measurements (or pH calculated from discrete measurements of two other CO_2 system variables) as reference. Second, all of them establish the correction on the basis of comparisons made at depths below 1000 m (typically 1480-1520 dbar) where spatial and temporal variations are expected to be minimal and where any temporal change can be neglected.
- Drift correction: This correction removes any observed temporal drift in the pH data as it emerges from the evolvement of the offset correction over a float's lifetime.

2.2 Additional Quality Control Measures for Float pH

In addition to the Argo cookbook method for correction of float pH (Johnson et al., 2018), it is desirable to establish further measures to independantly estimate pH data accuracy and ideally also to apply corrections, which has been our main focus within Euro-Argo RISE. We have explored the following measures:

2.2.1 CTD Cast with Discrete Water Sampling at Deployment Location

A measure that is often discussed in the context of float data quality control are CTD casts with discrete water sampling performed at the time and place of the float deployment. This would require an established reference method – in this case for discrete pH measurements – which itself can be traced back to an external reference. For marine CO_2 variables this is typically assured by measuring 'CO_2 in Seawater Reference Material' procured from the Andrew Dickson Laboratory (Scripps Institution of Oceanography, UC San Diego, USA). While the CRM is certified only for DIC and TA it is also stable for pH and the Dickson Lab is sharing their quality pH measurements that is performed on each CRM batch. A second requirement is to carry out the seawater sampling strictly in compliance with the respective standard operation procedure (SOP 1) defined in the 'Guide to Best Practices for Ocean CO_2 Measurements' (Dickson et al., 2007).

In addition, it would be advisable to modify the float's profiling routine to allow for a first full depth profiling quickly after deployment, possibly followed by 1 or 2 follow-up profiles at 24 h intervals. This would allow for improved matching of float profile(s) and deployment cast.

2.2.2 Crossovers between CTD Cast with Water Sampling and Float Profiles

Under normal circumstances it would be nearly impossible to specifically aim for crossovers between CTD casts and float profiles during a float's lifetime. However, for field work in a defined working area where floats are present and some degree of flexibility exits in the workplan of a research cruise an opportunistic crossover planning may be possible.

In Euro-Argo RISE, we have successfully achieved one such crossover in the Labrador Sea with a perfect match in space and a 1-day mismatch in time (Table 1).

Table 1: *Crossover between pH profiles from BGC Argo float and CTD rosette hydrocast in the central Labrador Sea (time refers to end of profile).*

Profile	Time	Latitude	Longitude
Float WMO 3901669 profile 122	15. August 2020, 10:26 UTC	52.955°N	48.600°W
R/V Maria S. Merian Cruise MSM94 station 78, CTD cast	16. August 2020, 05:36 UTC	52.953°N	48.600°W

2.2.3 Comparison between SOOP-based and Float-based Surface Carbon Observations

The global Ship-of-Opportunity network (SOOP) is operated in the European Research Infrastructure 'Integrated Carbon Observation System' (ICOS) and the 'Surface Ocean CO_2 Reference Observing Network' (SOCONET). It is an existing observatory that has been proposed as potential reference for quality control of float-based pH in the surface ocean. We therefore explain briefly the SOOP carbon observation network.

ICOS provides standardised and open data from more than 140 measurement stations across 14 European countries. The stations observe greenhouse gas concentrations in the atmosphere as well as carbon fluxes between the atmosphere, the land surface and the oceans. Thus, ICOS is rooted in three domains: Atmosphere, Ecosystem and Ocean. ICOS Ocean monitors greenhouse gases in the Atlantic Ocean and the Nordic, Baltic and Mediterranean Seas from instrumented Ships-of-Opportunity (SOOP) and Fixed Ocean Stations (FOS).

SOCONET is volunteer group of established operators who provide quality global surface ocean CO_2 data. Participants perform automated measurements of surface water and atmospheric CO_2 from Ships-of-Opportunity and moorings.

SOCAT: The data from these networks are submitted routinely and according to specific requirements to the 'Surface Ocean CO_2 Atlas' (SOCAT) where they undergo defined quality control measures. SOCAT hence represents a synthesis activity for quality-controlled, surface ocean fCO_2 observations by the international marine carbon research community (>100 contributors). SOCAT data are publicly available, discoverable and citable and they enable quantification of the ocean carbon sink and ocean acidification and evaluation of ocean biogeochemical models. SOCAT currently is the most important ocean data repository needed to calculate the ocean CO_2 sink in the annually released Global Carbon Budget (e.g., Friedlingstein et al., 2022). The SOOP network and the SOCAT data product are relevant for BGC-Argo in two ways (Fig. 1):

(1) BGC Argo float-based pH observations can be converted to fCO_2 and potentially be submitted to the SOCAT data-base. This, however, requires better understanding of the quality of pH-based surface fCO_2 observations by the Argo observatory. Only then float-based carbon observations can be ingested by SOCAT and receive adequate quality flags.

(2) The SOOP network and the SOCAT database represent a data repository that can actively be employed in the quality assessment and to some extent also in the routine quality control of Argo float-based pH. This is an idea we are specifically exploring in Euro-Argo RISE.

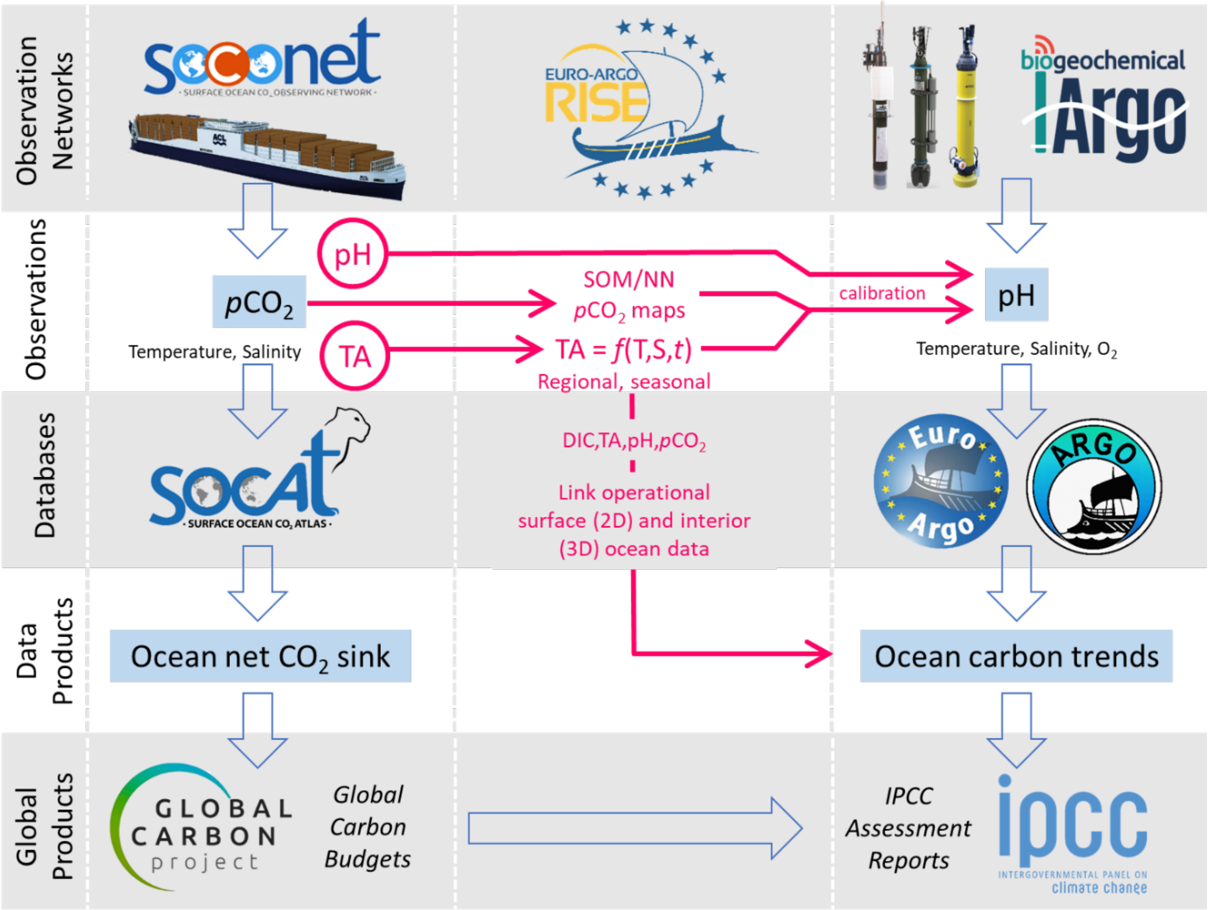

Fig. 1: *Concept for synergistic coupling of SOCONET and BGC-Argo observatories in the context of marine carbon observations.*

For this purpose, SOCONET/ICOS observations need to be augmented and harnessed for BGC-Argo. The standard-SOOP only features routine pCO_2/fCO_2 observations. The addition of total alkalinity (TA) as second analytical variable of the marine CO_2 system to the suite of autonomous SOOP observations has been demonstrated to be methodologically feasible now (Seelmann et al., 2019; 2020a; 2020b). We have therefore added TA to the measurement portfolio of our existing North Atlantic SOOP line (ICOS station DE-SOOP-Atlantic Sail) in 2019 using the Contros HydroFIA™ TA system (-4H-JENA engineering GmbH, Jena, Germany). This serves three purposes:

(1) From high-quality co-located and synchronous SOOP-based fCO_2 and TA observations pH can be calculated at relatively high quality to serve as a direct comparison for BGC-Argo floats.

(2) SOOP-based routine TA observations allow for improved parameterizations of (surface) TA which can be employed in the conversion of float pH into fCO_2 which is required for integration into the SOCAT database. Such improved TA parameterizations may help to reduce regional and/or seasonal biases present in surface TA algorithms.

(3) The SOCONET/ICOS observatory benefits tremendously from the addition of a 2nd CO_2 system variable as this allows for a complete analytical description of the marine CO_2 system and hence hugely improved interpretation potential of the acquired data.

As a further augmentation of the our existing North Atlantic SOOP we have implemented autonomous pH measurements using the Contros HydroFIA™ pH system (-4H-JENA engineering GmbH, Jena, Germany) with similar spectrophotometric technology in 2021. With these, a direct comparison with the BGC-Argo float network is possible which potentially allows for a direct and uncompromised quality assessment.

The comparison between SOOP-based and float-based surface carbon observations can in principle be carried in two ways:

(1) Search for direct crossovers between float surfacings and SOOP passages. We have explored this method in detail with variable search windows in both time and space and demonstrate the potential of the method.

(2) Self-organizing map neural network techniques (SOM/NN) can be used to produce pCO_2 maps with can be converted into pH maps using parameterisations of surface TA. If available, SOOP-based pH measurements can be used to directly produce pH maps. These pH maps can then be sampled at the time and space of the floats surfacings for pH reference data. We have not explored this approach due to the current limitations of the dataset available to us.

3 Results and Lessons Learned

Through great synergy with the operational German Argo Programme as well as several nationally funded research projects (OA-TWS-IOC, DArgo2025, C-SCOPE) we were able to start in 2018 and carry out beyond today a pilot study for float-based carbon observations in the subpolar North Atlantic and their synergistic coupling with an existing SOOP line. A total of 9 pH/O_2 floats from two manufacturers have been purchased and deployed as part of the pilot study so far (Table 2, Fig. 2). Further floats will be deployed in 2022 and 2023.

Table 2: *BGC-Argo floats with pH and O_2 sensors deployed in the Labrador Sea and subpolar North Atlantic as part of a pilot study for float-based carbon observations.*

WMO	Float Type	Depl. Time	Depl. Area	Status
3901667	Apex/Webb	June 2018	Labrador Sea	Inactive, 38 cycles, faulty pressure sensor → warranty replacement 7900566
3901668	Apex/Webb	June 2018	Labrador Sea	Inactive, 182 cycles
3901669	Apex/Webb	June 2018	Labrador Sea	Inactive, 179 cycles
7900566	Apex/Webb	August 2020	Labrador Sea	Active, 96 cycles, faulty GPS system
6904110	Provor/nke	July 2021	Subpolar N.A.	Active, 58 cycles, pH sensor failure after few profiles
6904111	Provor/nke	July 2021	Subpolar N.A.	Active, 77 cycles, pH sensor failure after few profiles
6904112	Provor/nke	July 2021	Subpolar N.A.	Active, 57 cycles Compromised pH sensor quality
6904114	Provor/nke	August 2021	Labrador Sea	Active 45 cycles Compromised pH sensor quality
6904115	Provor/nke	August 2021	Labrador Sea	Active, 49 cycles, pH sensor failure after few profiles

Unfortunaly, the nine deployed floats suffered from an unusually high number of manufacturer-related technical issues or failures either of the pressure sensor (3901167, warranty replacement by 7900566), the GPS system (7900566) or the pH sensor itself (6904110, 6904111, 6904112, 6904114, 6904115). The latter was related to a problem with the reference electrode which was reported to have occurred over the serial number range 10000 to 11117 and caused the affected pH sensors to deliver reduced data quality and in many cases show early sensor failure. This has severely compromised both the quality and the amount of pH data. In the end only little more than half of the floats are at all useful towards to goals of the pilot study.

We note the observation that in our analyses the two long-lasting floats deployed in 2018 (WMO 3901668 + 3901669) showed the best results and are thus assumed to represent the optimum case for achievable performance of this current technology. Floats deployed in 2021 (WMO 690411x) with the later reported manufacturing issues show significantly less promising results in almost all characteristics. It has to be seen how the manufacturing quality will develop in the future.

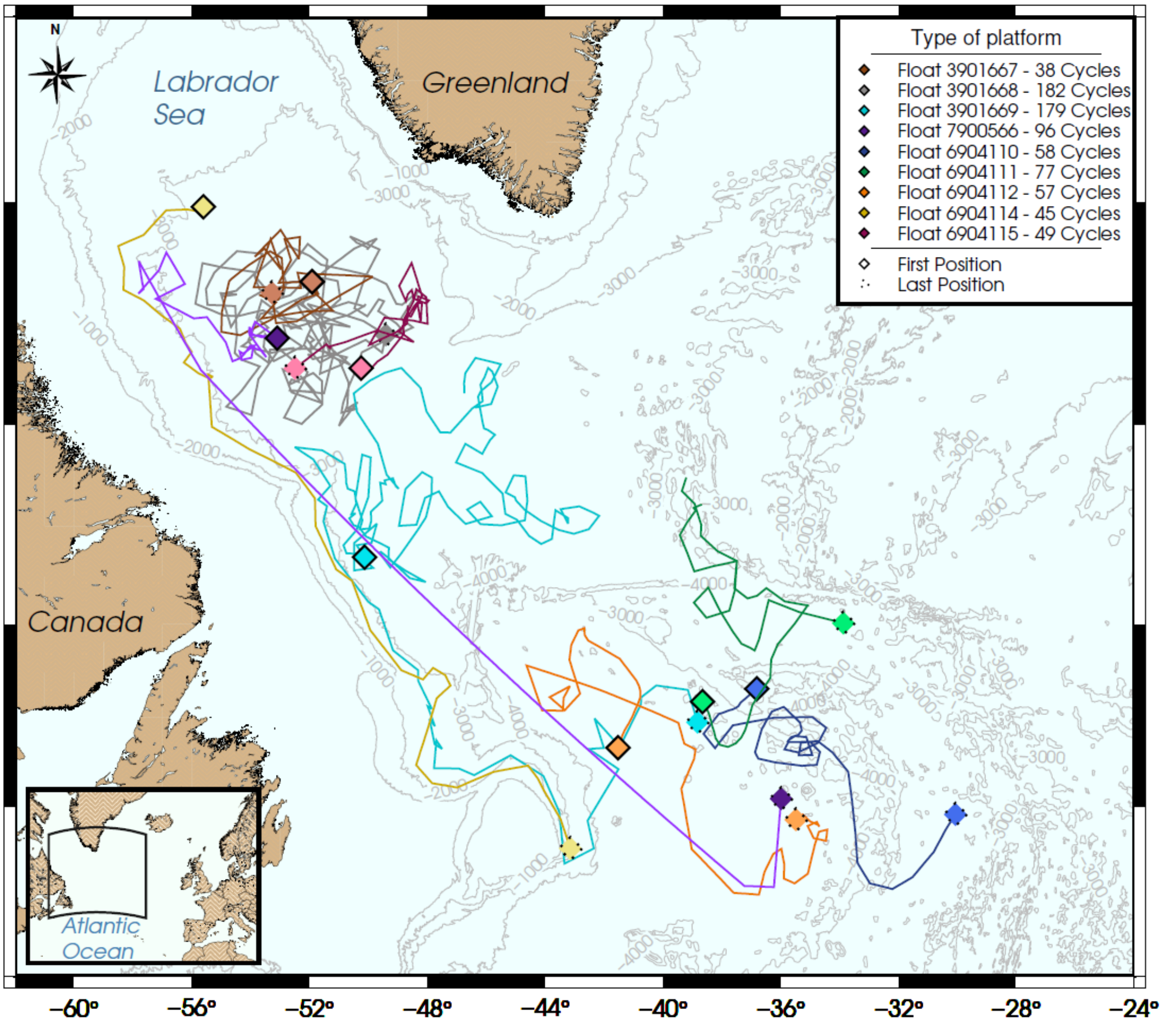

Fig. 2: *Map of the Northwest Atlantic with Labrador Sea and North Atlantic Current showing the trajectories of all 9 pH/O₂ floats deployed so far in our pilot study.*

In the following we will briefly illustrate the results and lessons learned from this pilot study and the parallel augmentation of the SOOP line.

3.1 Cookbook Procedures

3.1.1 Pump Offset

The Apex floats change from unpumped to pumped pH measurement mode during profiling at 750 dbar which can lead to a step-like offset created by the development of a flowing potential in the sensor. This offset was visible on all our Apex floats and had a magnitude of about 0.003 pH units (Fig. 3).

For the correction of the pump offset two linear regressions were fitted to data section above (700-730 dbar) and below (760-840 dbar) the depth at which the pump was switched on (750 dbar). Both linear regressions were then extrapolated to 750 dbar and the difference between the two extrapolated values taken as the pump offset (Fig. 4), applied to the lower unpumped section to make it consistent with the upper pumped section.

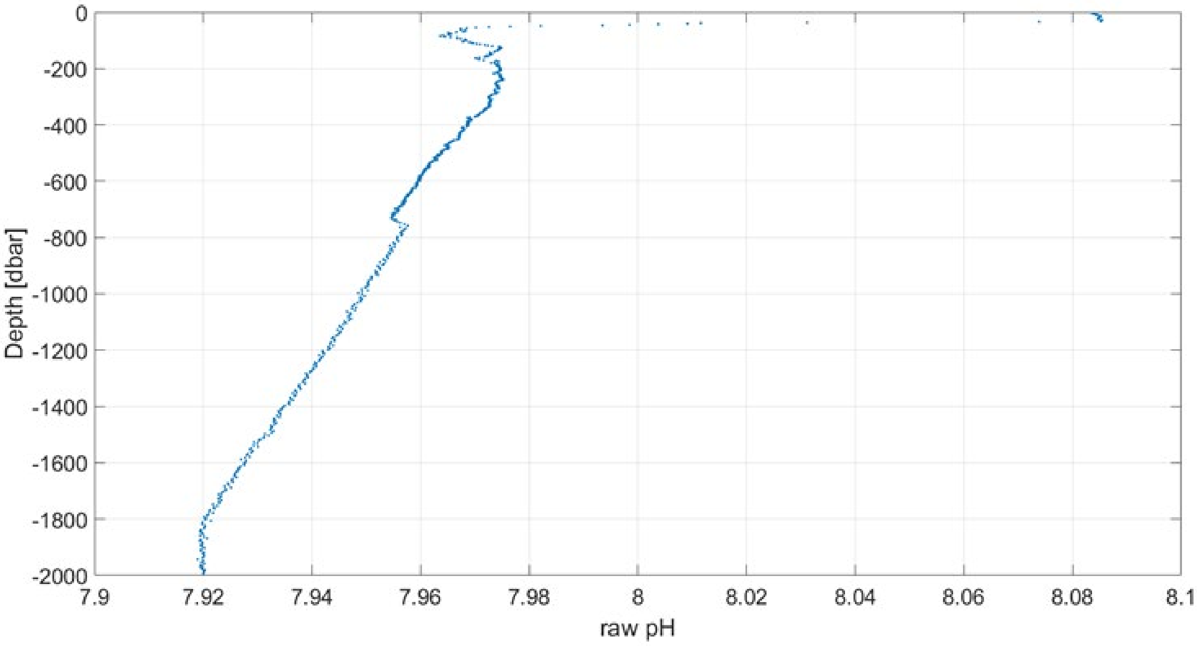

Fig. 3: *Example of the step-like jump in the pH profile that occurred when the CTD pump was switched on during at 700 dbar profiling (Apex float) .*

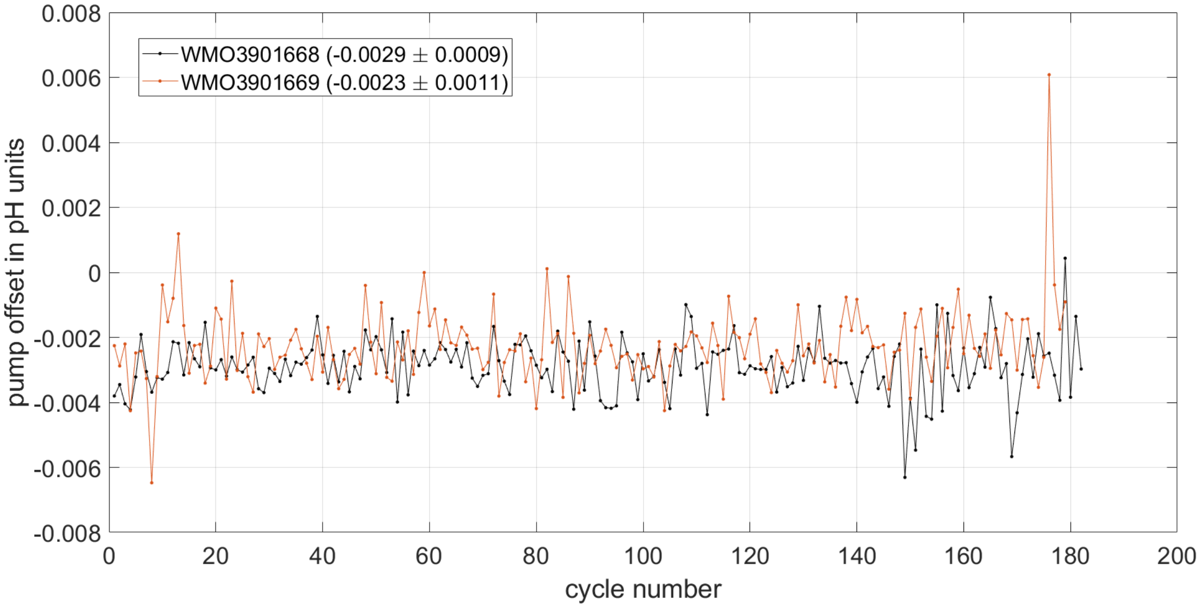

Fig. 4: *Time series of the estimated pump offset during the lifetime of floats WMO 3901668 and WMO 3901669.*

The pump offset appears to be rather stable both between floats of identical make and within a given float's lifetime. The pH uncertainty incurred by the effect of variability in the flowing potential is therefore assumed to be on the order of 0.001 pH units (Fig. 4).

3.1.2 Calculation of Major Offset Correction

Uncorrected float pH showed a significant offset to the pH from reference profiles with float-pH being consistently too low on all floats and all profiles. For the correction, we employed two reference methods: the locally interpolated pH regression (LIPHR) method (Carter et al., 2018) and the CANYON-B/CONTENT method (Sauzede et al., 2017; Bittig et al., 2018).

The calculations were carried out with two types of software code: the SAGE GUI (SOCCOM Assessment and Graphical Evaluation graphical user interface) developed at MBARI (Maurer et al., 2021) and a Matlab code developed by Tobias Steinhoff at GEOMAR. The latter code provides two calculation modes: a profile-by-profile mode, where individual corrections are calculated for every profile, and a segment mode nearly identical to the one in SAGE, where corrections are calculated for suitable segments of consecutive profiles (determined by a cost function) via linear least squares fit. Direct comparisons of these software codes revealed neglible differences (<0.001 pH units). After inspection of the data it is not yet fully clear to us whether the segment or point-by-point method is superior. The results of the calculation variants are shown for an examplary profile in Fig. 5. We already note here that the LIPHR method produced corrected pH values higher by about 0.02 pH units than those of the CANYON-B/CONTENT method.

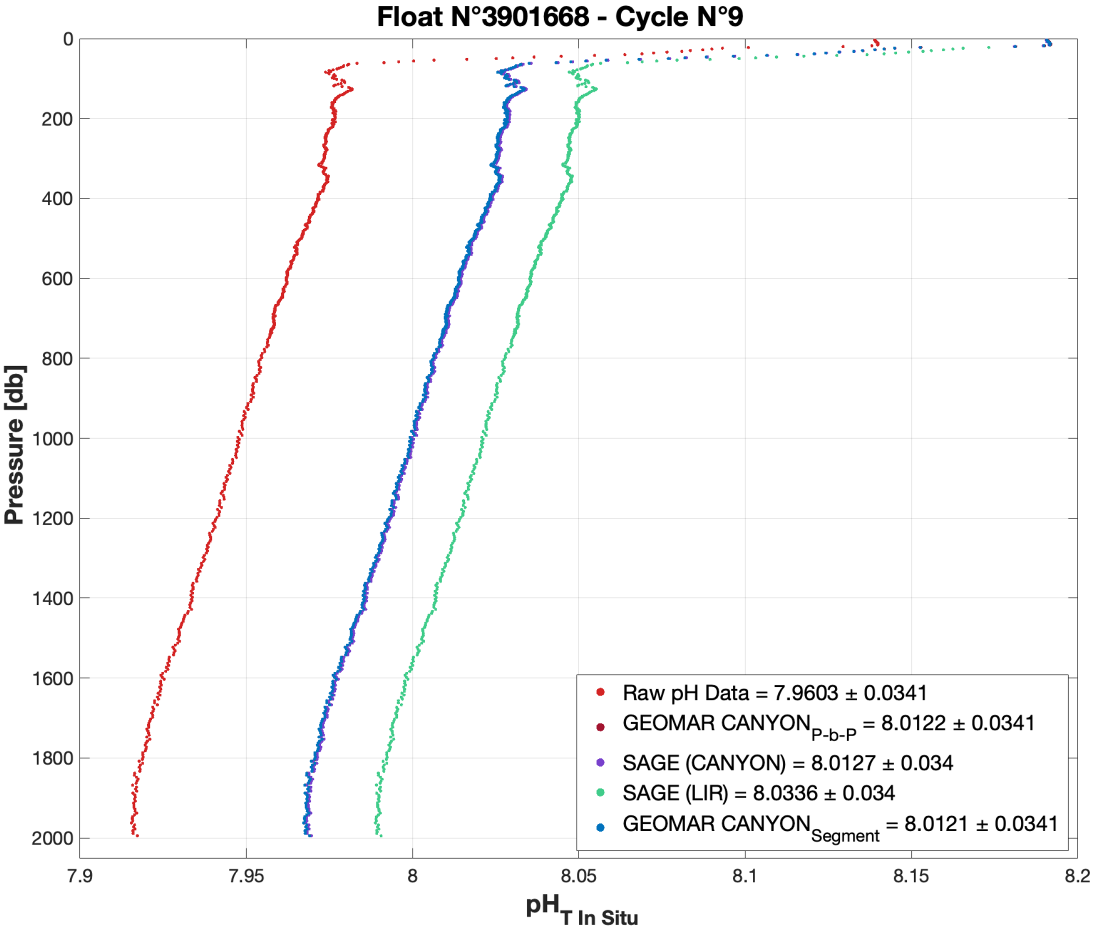

Fig. 5: *Example of float pH profile (WMO 3901668, profile 9). Shown are raw pH (light red), pH corrected by SAGE against LIPHR (green), pH corrected by SAGE against CANYON-B/CONTENT (purple), pH corrected with GEOMAR software against CANYON-B/CONTENT using the point-by-point (dark red) and segment method (blue). The adjustment suggested by Carter et al. (2018) to correct for the bias associated with the wide-spread use of non-purified m-cresol purple pH indicator dye was not applied. A temperature correction was applied to GEOMAR algorithm which mimics the way the offset to the reference potential K_0 of the pH sensor is applied in SAGE. Note that the blue and purple lines coincide in the plot as the two corrections agree to within 0.0006 pH units. Also the two GEOMAR methods (point-by-point and segment) agree to within 0.0001 pH units and are indistinguishable.*

For the purpose of the further analysis and discussion, we decided to calculate corrections by using SAGE GUI and GEOMAR software (both in segment mode) for the pressure range 1480-1520 dbar and against CANYON-B reference methods. The resulting corrections were applied uniformly across the entire pH profile (Fig. 6). Calculated offsets were lowest (<0.06 pH units) and relatively stable over time with the 2018 and 2020 Apex pH floats (green, cyan and blue dots). A small but systematic positive pH drift of about 0.002 yr^{-1} and 0.003 yr^{-1} was found for floats WMO 3901668 and WMO 3901669, respectively.

In contrast, offsets showed rapid initial drift with stabilization at an offset of about 0.15 to 0.2 pH units (except for float WMO 69041114 where a further deterioration is visible with the most recent profiles). The long-term pH drift of this sensor series is about on the order of 0.01 yr^{-1} to 0.025 yr^{-1}, i.e an order of magnitude higher than for the well-performing pH sensor discussed above. In additon the drift does not show a simple pattern.

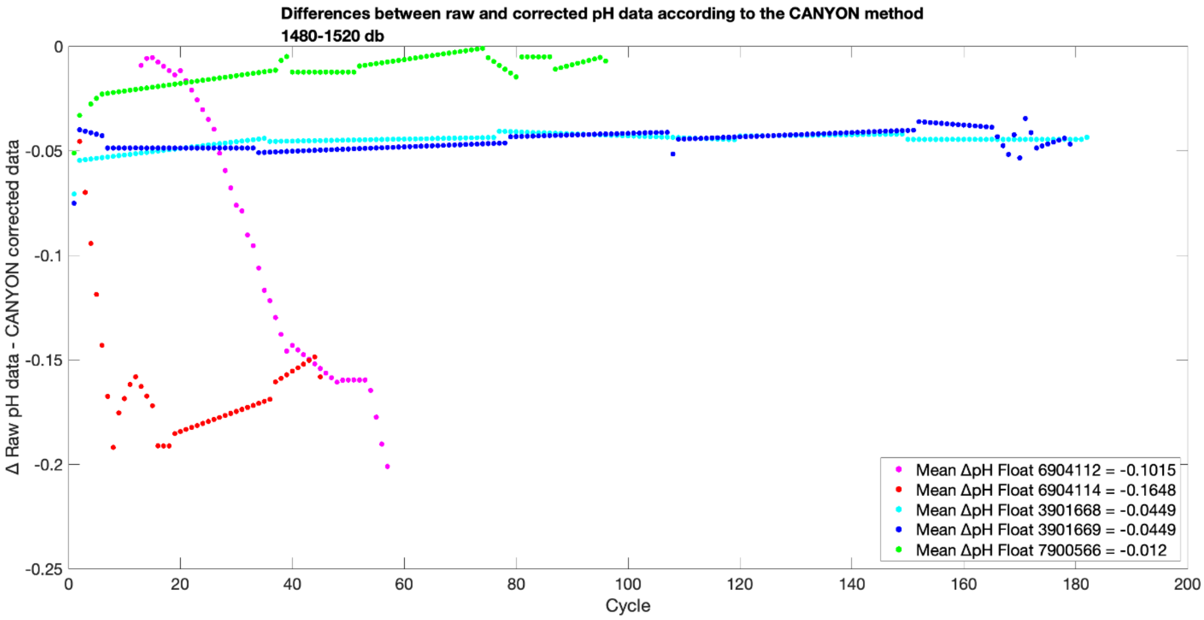

***Fig. 6:** Differences between raw pH data and pH data corrected according to the CANYON-B/CONTENT method over the pressure range 1480-1520 db.*

Our results also show that the choice of reference depth is critical in the subpolar North Atlantic Ocean where interannually varying deep convection (Fig. 7), water mass formation as well as decadal variability affect water masses at depth greater than 1000 m. For comparability purposes we followed the established procedures (1480-1520 dbar reference) and have not yet explored this in great detail. However, this situation introduces significant uncertainty into the pH correction and calls for an indepedant accuracy check at the surface.

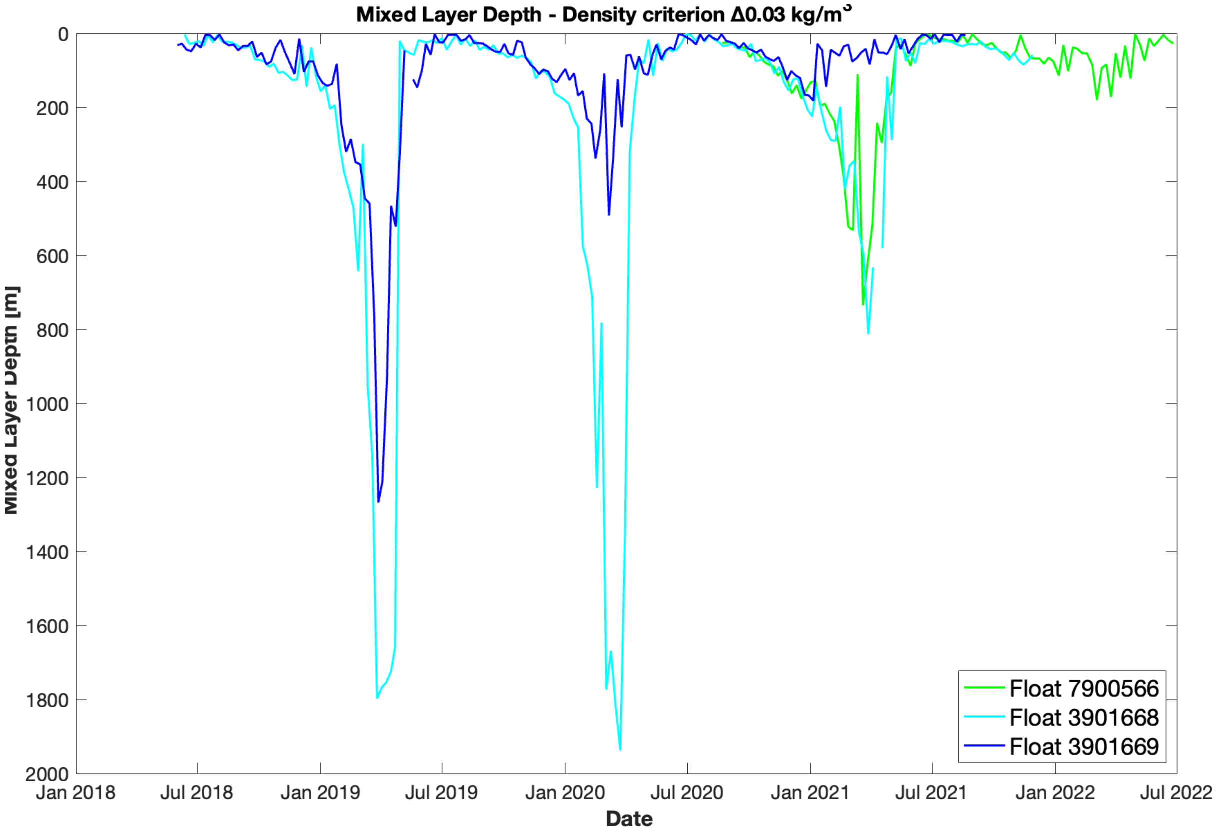

Fig. 7: *Mixed layer depth in the Labrador Sea region based on the observation from three floats of our pilot study (WMO 3901668, WMO 3901669, WMO 7900566). An MLD threshold criterion in potential density of Δσ = 0.03 kg m*$^{-3}$ *from the value at 10 m was chosen.*

3.2 Assessment of Final Accuracy of Float pH

3.2.1 CTD Cast with Discrete Water Sampling at Deployment Location

After deployment, the pH sensor showed an initial acclimation phase of 10-15 days over which the pH reading at drift depth, i.e. under presumably very stable pH conditions, shows a signficant drift towards stabilization (Fig. 8). This drift pattern appears too long to allow unquestionable matching with deployment casts with discrete sampling even when profiling routines are adapted to early and repeated profiling. A further complication of this is that during deployment cruises not always CTD rosette systems, let alone trained participants for reliable water sampling according to SOP 1 of the DOE Handbook (Dickson et al., 2007), are available.

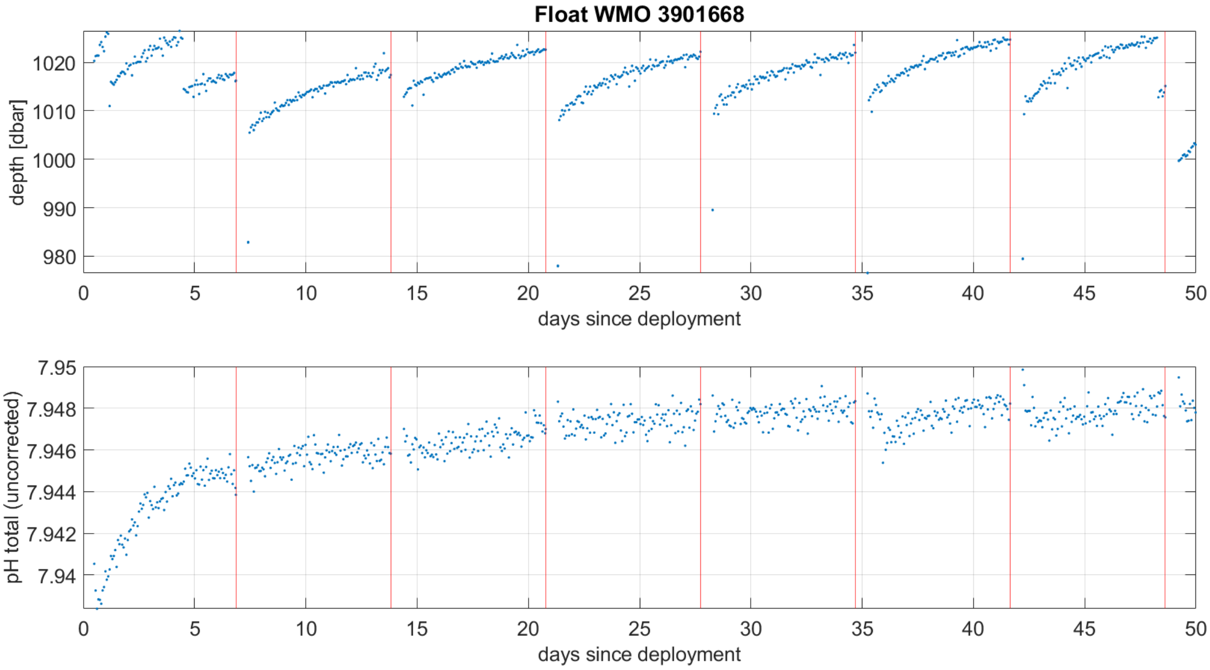

Fig. 8: *Time series of pH measured by float WMO 3901669 at drift depth (1000 dbar) since deployment. Only data between park start (PST) and end time (PET) shown. Red lines indicate time of profiling.*

The observed pH during the acclimation phase showed strongly differing magnitude even between the two well-performing sister floats WMO 3901668 (0.008) and WMO 3901669 (0.025) which prevents application of a uniform correction for the initial drift. Time series pH at drift depth do not show a clear stabilization and early sign of sensor deterioration with the problematic 2020 pH sensor series (Fig. 9, WMO 6904115).

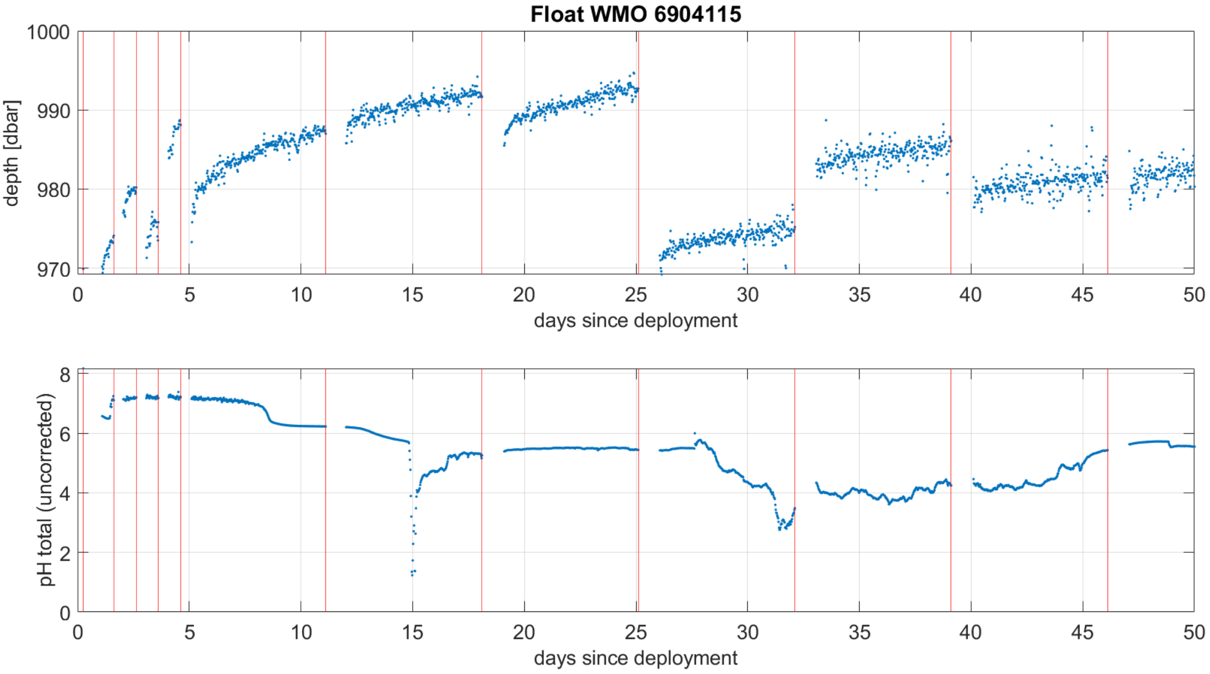

Fig. 9: *Time series of pH measured by float WMO 6904115 at drift depth (1000 dbar) since deployment. Only data between park start (PST) and end time (PET) shown. Red lines indicate time of profiling.*

We have therefore decided to not explore the deployment cast option at any greater depth. Otherwise the availability of cruises fully suitable for float deployment would have been significantly smaller and we could not have carried out the pilot study as it is now.

3.2.2 Crossover between CTD Cast with Water Sampling and Float Profile

We achieved a single crossover between a float pH profile (corrected by GEOMAR software employing CANYON-B/CONTENT and point-by-point method) and a CTD hydrocast with discrete pH sampling (Fig. 10). The pH samples were measured on the Contros HydroFIA™ pH system (estimated 1σ precision: 0.002 pH units) in discrete mode against CRM (assigned 1σ accuracy: 0.0014 pH units). The CTD cast was carried out at a distance of about 200 m and 19 hours to the location and time of the float surfacing of profile 122. Profiles 121 and 123 took place 6.5 days prior and past profile 122 and at a distance of 65 and 36 km, respectively. When assuming a float drift linear in time and space, the temporal mismatch of float and co-located CTD profile corresponds to 4-8 km. This is on the same scale as the Rossby radius of deviation (<10 km) and hence typical eddy dimensions in the Labrador Sea. This needs to be taken into account when comparing pH profiles of float and CTD.

Discrete pH data measured at 25°C and atmospheric pressure were corrected to in-situ temperature and in-situ pressure using the CO2sys software (van Heuven et al., 2011) and the TA values measured (VINDTA system, CRM correction) on the same samples. Methodologically they should therefore be consistent with float pH. The matching of discrete and float pH data was performed in density space (rather than depth space) to avoid biases from internal waves.

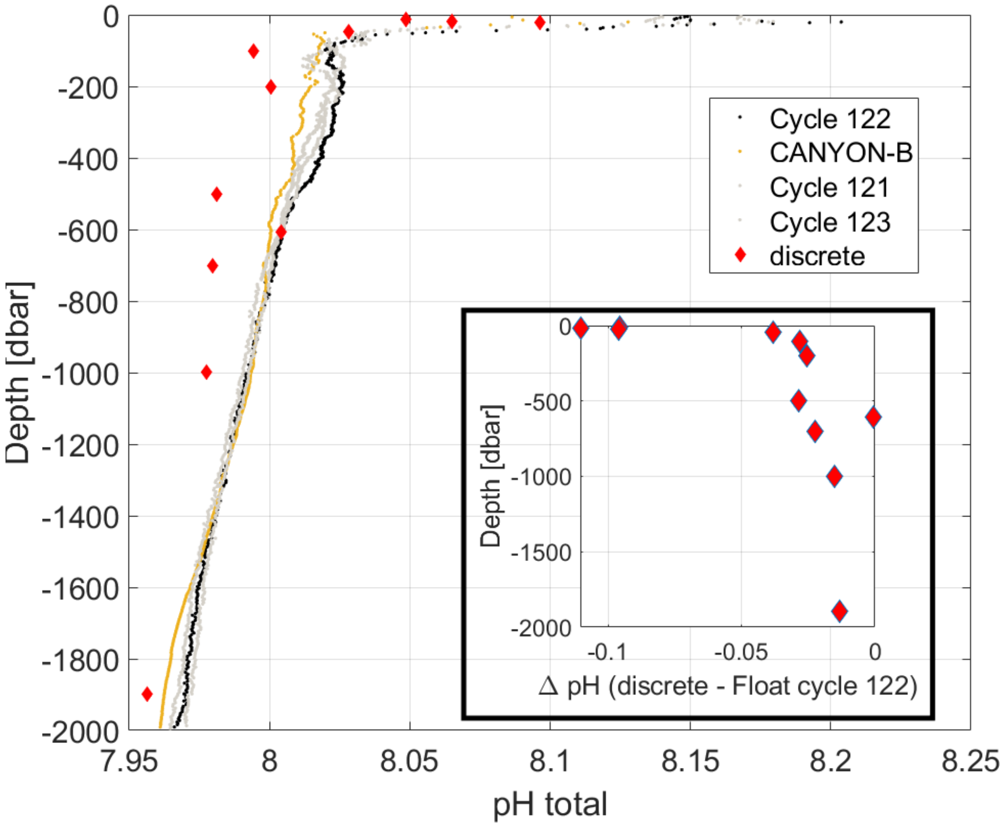

***Fig. 10:** WMO 3901669 pH profiles 121, 122 and 123 (corrected to the CANYON-B/CONTENT reference profile) and the discrete pH reference profile corrected to in-situ temperature and pressure (left panel) and absolute pH differences (discrete minus float) after matching in density space.*

The comparison shows that the CANYON-B/CONTENT reference profile to which correction was made at 1480-1520 dbar shows significant deviations from the three float pH profiles, particularly in the pressure ranges 100-700 dbar and > 1600 dbar. The latter shows that the deep ocean correction at an arbitrarily chosen depth at grater than 1000 m is not straightforward and in this case induced uncertainty of about 0.01 pH units. This may to some extent be a special characteristic of the Labrador Sea with its extreme winter convections depth of up to 1500 m. In this particular case this may speak for a deeper reference level (e.g., 1900 dbar) which would shift results by about 0.01 pH units but still leave the overall disagreement between corrected float pH and discrete pH rather disappointing. With this single example of a crossover with a discrete pH profile it is impossible to provide a final judgement about the achievable float accuracy. It clearly points however at the limitations of such crossovers as a method of assessing or improving accuracy of float pH.

3.2.3 Comparison between SOOP-based and Float-based Surface Carbon Observations

On our SOOP line, pH was measured at 15 min intervals using the Contros HydroFIA™ pH system in unattended mode. The continuous flow of uncontaminated and bubble-free seawater was provided via a cross-flow filter to avoid particle contamination of the optical path. Before and after each 5-week roundtrip (each with 2 trans-Atlantic crossings) the instrument was calibrated in port against fresh CRM provided by the Dickson Lab (batch 190, accuracy 0.0014 pH units). These pre- and post-calibration runs are rather stable for each meta-cresol purple (mCP) indicator bag which lasts for 4 roundtrips (Fig. 11). Some variability between mCP batches and even bags from the same batch is evident. Therefore an individual pH correction is applied to each mCP bag. The overall accuracy of SOOP-pH is about 0.003 pH units.

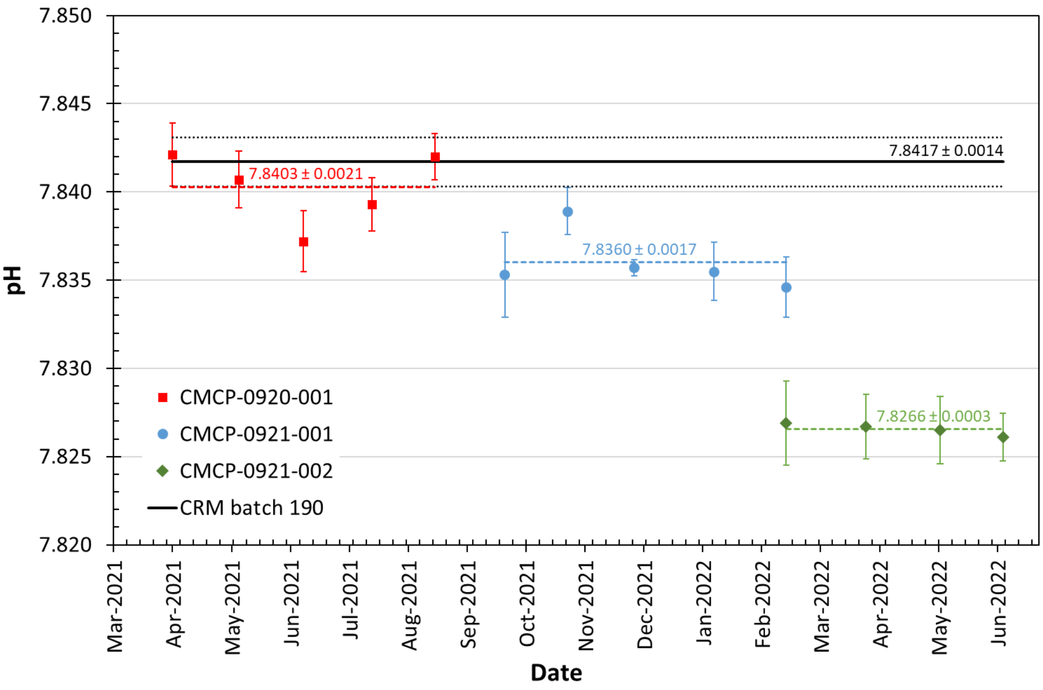

Fig. 11: *pH measurements performed on CRM batch 190 with the Contros HydroFIA™ pH system before and after each 5-week roundtrip of our Ship-of-Opportunity M/V Atlantic Sail. Four consecutive roundtrips are always carried out with one reagent bag of the mCP indicator (CMCP-mmyy of batch preparation-number of bag). Adjustments of measured pH to the nominal pH value assigned to the CRM (7.8417 ± 0.0014 at 25°C) is based on the mean of all CRM measurements carried out per individual mCP bag (typically at 5 different times, values shown in plot).*

Simple search windows in time and space were applied to search for crossovers. The approach needs to be optimized between the interest to find as many crossovers as possible and the interest to make the crossover matches as representative as possible. This balance probably needs further investigation taking into account regional decorrelation length scale of surface pH (or $p/f\mathrm{CO}_2$). Also additional crossover criteria (e.g., max. temperature difference) and procedures (e.g., correction of SOOP pH to the temperature of the float pH) probably need to be developed. In this regard, we consider our analyses somewhat preliminary.

We haven chosen a time window of ±3-7 days and a space window of 100-200 km radius for our analyses. SOOP data falling into the respective crossover criteria for a given float surfacing were extracted and averaged. Likewise the mixed layer pH data (typically upper 15 m) of a float surfacing were extracted and averaged. Standard deviations of these averages give an indication of the coherence of the extracted data portions and hence their statistic weight.

For 2021, we achieved several crossovers between surfacings of four pH/O_2 floats (WMO 3901669, WMO 6904110, WMO 6904111, WMO 6904112) and our SOOP line in the subpolar North Atlantic (Fig. 12). More crossovers are indentified in 2022 and await final processing and quality contol of the SOOP data. An updated version of this deliverable report will therefore be provided towards the end of 2022.

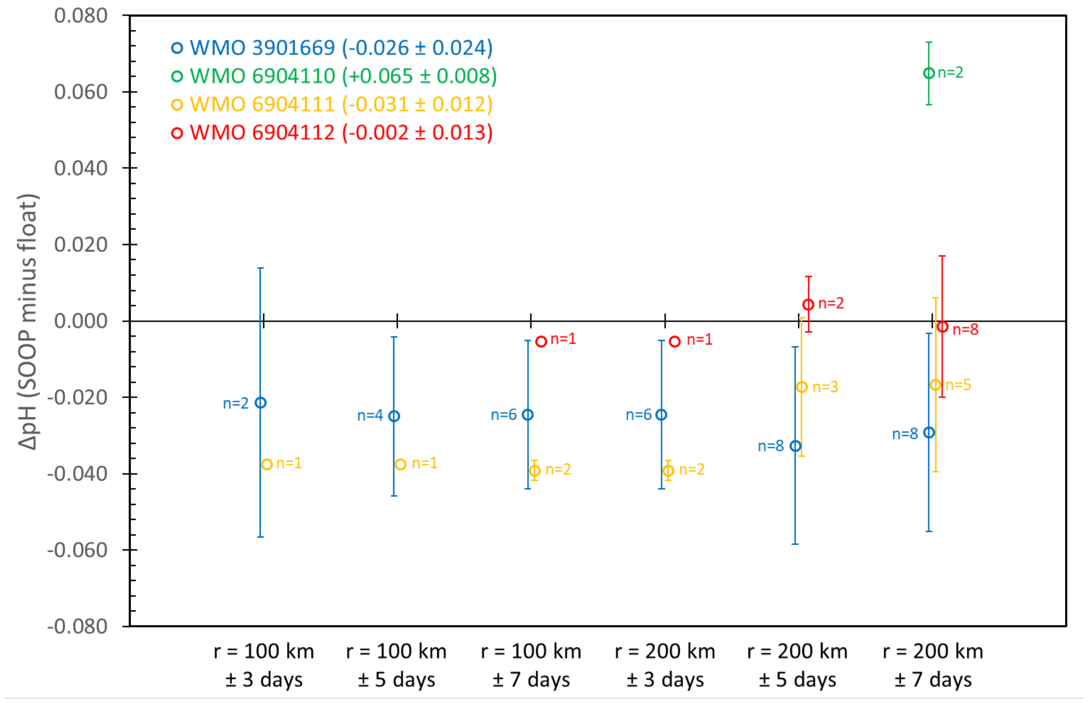

Fig. 12: *Offset between SOOP pH and fully corrected float pH at crossovers. The 6 sets of criteria applied in the crossover search are shown at the bottom of figure. For each set of criteria the average pH offset (± 1σ) is shown together with number of crossovers found. No further selection criteria were applied. The legend provides for each float the average offset (± mean 1σ) across all sets of selection criteria.*

Although the results only represent relatively short periods of time (1-3 months) of regional overlap with the SOOP line in 2021 two conclusions can already be drawn:

(1) With larger amounts of SOOP-float crossover data and optimized crossover criteria an independant quality control and perhaps correction of float surface pH can be achieved. SOOP could thus be instrumented in DMQC routine for float pH. Algorithms synchronizing float profiles

routine with SOOP line schedules and projected crossings may even be employed to actively chase crossovers in a more systematic way.

(2) More or less consistent results found for each float indicate that float pH may be biased by up to several hundredths of a pH unit at the surface. The four floats represented in this limited analysis show range from virtually no offset to +0.06 and -0.03 pH units. With an uncertainty of 0.01 in pH corresponding to an uncertainty in pCO_2 of about 10 µatm, anything about that a 0.01 thresholds appears intolerable for estimating the air-sea CO_2 flux from float-based pH observations. These early findings therefore warrant further and more sophisticated analyses to better constraint float pH at the surface. At least in the subpolar North Atlantic, the established at depth correction does not seem to yield adequate pH accuracy at the surface. This uncertainty may partly be incurred by the regional complication of finding a reliable at-depth reference.

(3) The SOOP crossover analysis yields a positive pH bias of 0.026 ± 0.024 (σ of 6 crossovers) for float WMO 3901669 at the surface. Applying a uniform correction by this amount to the pH profiles of this float will bring the pH data in much better agreement with crossover with the CTD cast (Fig. 10). This result, despite the limited statistical significance of the two independent crossovers analyses, adds credibility to the use of crossover analyses in float pH QC and points at significant issues with the accuracy of float-based pH observations.

4 Conclusions & Perspectives

The following conclusions can be drawn from the results from our pH/O_2 float pilot study – carried out in synergistically combined effort between Euro-Argo RISE and German-funded collaborative projects :

- The sub-polar North Atlantic is a challeging environment for both DMQC and independant quality assessment of float-based pH. Particulary in the western basin, the anthropogenic CO_2 imprint can be traced to depths greater than the floats' maximum profiling depth. Therefore even the deeper depth ranges are affected by ocean acidification. In addition, interannually variable deep convection and water mass formation introduce further transient signals in the deep ocean's CO_2 system. Therefore no quiescent and stable reference layer can be identified for application of the DMQC pH correction. This points at a general shortcoming of a correction scheme that relies on deep hydrography datasets which are affected to an unknown degree by the transient signals.
- The results of our pH/O_2 float pilot study point at significant issues with pH data quality, at least in the surface ocean. Accuracy assessments based on crossovers with autonomous pH measurements from a SOOP line point at offsets in excess of 0.01 pH units which correspondens to about 10 µatm uncertainty in $p/f\mathrm{CO}_2$, i.e. the limit for inclusion into the SOCAT database. While still limited by the number of crossovers achieved so far, a relatively consistent picture emerges for every tested float.
- SOOP-Argo crossovers appear to be a promising method for assessing and improving float-based pH observations in the surface ocean in regions with SOOP coverage. For this purpose, improved crossover search and correction algorithms need to be developed. Targeted float surfacings near expected SOOP passages could be considered to systematically improve the quantity and quality of crossovers. An independant surface reference for float pH seems necessary to meet required accuracy goals. In addition, a closer link between the BGC-Argo and SOCONET observatories would be mutually beneficial to both networks.
- The robustness of the conclusions to be drawn from out pH/O_2 float pilot study is limited given the reduced number of best-performing pH sensors and hence much smaller amount of highest quality data. This warrants continued attention to the quality of float-based pH in the surface ocean.
- The pH/O_2 float pilot study suffered tremendously from manufacturing-related problems. Particulary the recent serious problems with the pH sensor's reference electrode cause a high likelihood of strongly reduced sensor lifetimes plus enhanced and more variable drift characteristics. This points at an urgent need to improve high manufacturing quality.
- Even a well-performing pH sensor was found to not meet the minimum requirement of pH accuracy of <0.01 at the surface. This points at the need to further improve on the SeaFET pH sensor characteristics, independant of solving the recent manufacturing quality issues.
- Further work on alternative pH sensors – both in terms of manufacturers as well as in terms of sensor principles – is encouraged.

References:

Bittig, H. C. and A. Körtzinger, 2015: Tackling oxygen optode drift: Near-surface and in-air oxygen optode measurements on a float provide an accurate in-situ reference. J. Atmos. Ocean. Techn., 32, 1536-1543, doi: 10.1175/JTECH-D-14-00162.1.

Bittig, H. C. and A. Körtzinger, 2017: Technical note: Update on response times, in-air measurements, and in situ drift for oxygen optodes on profiling platform. Ocean Sci., 13, 1-11, doi: 10.5194/os-13-1-2017.

Bittig H. C., T. Steinhoff, H. Claustre, B. Fiedler, N. L. Williams, R. Sauzède, A. Körtzinger, and J.-P. Gattuso, 2018: 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. Front. Mar. Sci., 5, 328, doi: 10.3389/fmars.2018.00328.

Carter, B. R., R. A. Feely, N. L. Williams, A. G. Dickson, M. B. Fong, and Y. Takeshita, 2018: Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate. Limnol. Oceanogr.: Methods, 16, 119–131. doi: 10.1002/lom3.10232

Dickson, A. G., C. L. Sabine, and J. R. Christian (Eds.), 2007: Guide to Best Practices for Ocean CO_2 Measurements. PICES Special Publication 3, 191 pp.

Fiedler, B., P. Fietzek, N. Vieira, P. Silva, H. C. Bittig, and A. Körtzinger, 2013: In situ CO_2 and O_2 measurements on a profiling float. J. Atmos. Ocean. Technol., 30, 112-126, doi: 10.1175/JTECH-D-12-00043.1.

Friedlingstein, P., M. W. Jones, M. O'Sullivan, R. M. Andrew, D. C. E. Bakker, J. Hauck, C. Le Quéré, G-P. Peters, W. Peters, J. Pongratz, S. Sitch, J. G. Canadell, P. Ciais, R. B. Jackson, S. R. Alin, P. Anthoni, N. R. Bates, M. Becker, N. Bellouin, L. Bopp, T. T. T. Chau, F. Chevallier, L. P. Chini, M. Cronin, K. I. Currie, B. Decharme, L. M. Djeutchouang, X. Dou, W. Evans, R. A. Feely, L. Feng, T. Gasser, D. Gilfillan, T. Gkritzalis, G. Grassi, L. Gregor, N. Gruber, Ö. Gürses, I. Harris, R. A. Houghton, G. C. Hurtt, Y. Iida, T. Ilyina, I. T. Luijkx, A. Jain, S. D. Jones, E. Kato, D. Kennedy, K. Klein Goldewijk, J. Knauer, J. I. Korsbakken, A. Körtzinger, P. Landschützer, S. K. Lauvset, N. Lefèvre, S. Lienert, J. Liu, G. Marland, P. C. McGuire, J. R. Melton, D. R. Munro, J. E. M. S. Nabel, S.-I. Nakaoka, Y. Niwa, T. Ono, D. Pierrot, B. Poulter, G. Rehder, L. Resplandy, E. Robertson, C. Rödenbeck, T. M. Rosan, J. Schwinger, C. Schwingshackl, R. Séférian, A. J. Sutton, C. Sweeney, T. Tanhua, P. P. Tans, H. Tian, B. Tilbrook, F. Tubiello, G. R. van der Werf, N. Vuichard, C. Wada, R. Wanninkhof, A. J. Watson, D. Willis, A. J. Wiltshire, W. Yuan, C. Yue, X. Yue, S. Zaehle, and J. Zeng, 2022: Global Carbon Budget 2021. Earth Syst. Sci. Data, 14, 1917–2005, doi: 10.5194/essd-14-1917-2022.

Johnson, K. S., H. W. Jannasch, L. J. Coletti, V. A. Elrod, T. R. Martz, Y. Takeshita, R. J. Carlson, and J. G. Connery, 2016: Deep-Sea DuraFET: A pressure tolerant pH sensor designed for global sensor networks. Anal. Chem., 88, 3249-3256, doi: 10.1021/acs.analchem.5b04653.

Johnson, K. S., J. N. Plant, and T. L. Maurer, 2018: Processing BGC-Argo pH data at the DAC level. doi: 10.13155/57195.

Martz, T. R., J. G. Connery, and K. S. Johnson, 2010: Testing the Honeywell Durafet® for seawater pH applications. Limnol. Oceanogr.: Methods, 8, 172-184.

Maurer T. L., J. N. Plant, and K. S. Johnson, 2021: Delayed-Mode Quality Control of Oxygen, Nitrate, and pH Data on SOCCOM Biogeochemical Profiling Floats. Front. Mar. Sci., 8, 683207, doi: 10.3389/fmars.2021.683207.

Olsen, A., N. Lange, R. M. Key, T. Tanhua, H. C. Bittig, A. Kozyr, M. Álvarez, K. Azetsu-Scott, S. Becker, P. J. Brown, R. B. Carter, L. Cotrim da Cunha, R. A. Feely, S. van Heuven, M. Hoppema, M. Ishii, E. Jeansson, S. Jutterström, C. S. Landa, S. K. Lauvset, P. Michaelis, A. Murata, F. F. Pérez, B. Pfeil, C. Schirnick, R. Steinfeldt, T. Suzuki, B. Tilbrook, A. Velo, R. Wanninkhof, and R. J. Woosley, 2020: An updated version of the global interior ocean biogeochemical data product, GLODAPv2.2020. Earth Syst. Sci. Data, 12, 3653–3678, doi: 10.5194/essd-12-3653-2020.

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

Seelmann, K., S. Aßmann, and A. Körtzinger, 2019: Characterization of a novel autonomous analyzer for seawater total alkalinity: Results from laboratory and field tests. Limnol. Oceanogr.: Methods, doi: 10.1002/lom3.10329.

Seelmann, K., T. Steinhoff, S. Aßmann, and A. Körtzinger, 2020a: Enhance ocean carbon observations: Successful implementation of a novel autonomous total alkalinity analyzer on a Ship of Opportunity. Front. Mar. Sci., 7, doi: 10.3389/fmars.2020.571301.

Seelmann, K., M. Gledhill, S. Aßmann, and A. Körtzinger, 2020b. Impact of impurities in bromocresol green indicator dye on spectrophotometric total alkalinity measurements. Ocean Sci., 16, 535-544, doi: 10.5194/os-16-535-2020.

van Heuven, S., D. Pierrot, J. W. B. Rae, E. Lewis, and D. W. R. Wallace, 2011: MATLAB Program Developed for CO_2 System Calculations. ORNL/CDIAC-105b. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee, doi: 10.3334/CDIAC/otg.CO2SYS_MATLAB_v1.1

Williams, N. L., L. W. Juranek, K. S. Johnson, R. A. Feely, S. C. Riser, L. D. Talley, J. L. Russell, J. L. Sarmiento, and R. Wanninkhof, 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., L. W. Juranek, R. A. Feely, K. S. Johnson, J. L. Sarmiento, L. D. Talley, A. G. Dickson, A. R. Gray, R. Wanninkhof, J. L. Russell, S. C. Riser, and Y. Takeshita, 2017: Calculating surface ocean pCO_2 from biogeochemical Argo floats equipped with pH: An uncertainty analysis, Glob. Biogeochem. Cycles, 31, 591–604, doi: 10.1002/2016GB005541.