



REPORT ON PRODUCT FOR BGC-ARGO-BASED ESTIMATION OF AIR-SEA CO₂ FLUX IN COMBINATION WITH OTHER OPERATIONAL OBSERVATION NETWORKS

Ref.: D4.8_V1.0

Date: 18/11/2022

**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





Disclaimer:

This Deliverable reflects only the author's views and the European Commission is not responsible for any use that may be made of the information contained therein.

Document Reference

Project	Euro-Argo RISE - 824131
Deliverable number	D4.8
Deliverable title	Report on product for BGC-Argo-based estimation of air-sea CO ₂ flux in combination with other operational observation networks
Description	Report on product for BGC-Argo-based estimation of air-sea CO ₂ flux in combination with other operational observation networks
Work Package number	4
Work Package title	Extension to biogeochemical parameters
Lead Institute	GEOMAR
Lead authors	Arne Körtzinger, Tobias Steinhoff, Cathy Wimart-Rousseau
Contributors	Fabrizio D'Ortenzio
Submission date	18.11.2022
Due date	31.03.2022 – revised due date 30.09.2022
Comments	
Accepted by	Fabrizio D'Ortenzio

Document History

Version	Issue Date	Author	Comments
Ver. 1.0	18/11/2022	Arne Körtzinger	

EXECUTIVE SUMMARY

Based on our findings in the subpolar North Atlantic, float-based pH measurements are currently not capable of systematically reaching a quality that yields observation-based surface ocean $p/f\text{CO}_2$ data at the required accuracies of the “weather” ($\pm 10 \mu\text{atm}$) or “climate” goals ($\pm 2 \mu\text{atm}$) as defined by the ocean carbon community. The situation in the study region may be more challenging than in other regions of the world ocean and therefore does not necessarily reflect the overall status of float-based pH measurements. The results illustrate, however, the uncertainty and lack of objective criteria associated with the choice of the reference method for the delayed-mode pH correction (CANYON-B or LIR) as well the reference depth for this correction. This results in potentially significant but typically unknown uncertainty associated with the resulting accuracy of fully corrected delayed mode pH data.

This is the perhaps most important scientific obstacle that currently prevents the operational inclusion of float-based $f\text{CO}_2$ data in the “Surface Ocean CO_2 Atlas”. We note that currently no formal pathway for inclusion of the data into SOCAT exists. The current philosophy of pH data correction against climatological reference data at one single depth in the deep ocean therefore requires critical evaluation. Our results point at the need of an additional reference point, ideally at the surface where the resulting $f\text{CO}_2$ data are of highest importance and would have the potential to very significantly augment the current observation network.

We argue that dedicated crossover analyses with carbon measurements by the “ship-of-opportunity” network (SOCONET/ICOS) could be employed in assessing and perhaps systematically correcting float-based pH measurements at the surface. Our regional focus and limited dataset limits the statistical significance of our findings but may nevertheless show a way forward. We propose that the scientific community explores this in a systematic way.

Clearly, float-based pH measurements provide wonderful perspectives and a completely new approach to carbon observations in the global ocean. As is typical for novel sensor technology and was also experienced for float-based oxygen observations, it is long way, however, from having a suitable sensor to delivering routine quality-assured data. The combination of BGC-Argo and ICOS, as explored further in the H2020 project GEORGE, holds great potential to i) enhance data quality and ii) reduce the bias in derived estimates of sea-air CO_2 fluxes (Denvil-Sommer et al., 2021).

Disclaimer:

This document represents the situation described in D4.7 which – at the time of data evaluation and writing of the report – was primarily based on data from 2021. We have updated the analysis with 2022 data. 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 sensor failure. This has both reduced the amount and compromised the quality of the data from the 5 floats. The results shown in D4.7 and D4.8 are therefore based largely on 3 earlier pH/ O_2 floats with much higher data quality.

However, as some pH floats and Ship-of-Opportunity line (SOOP) 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.



TABLE OF CONTENT

1	Background.....	4
1.1	Chemistry of the Marine Carbon Dioxide System.....	4
1.2	SOCAT Quality Control	5
1.3	Quality Requirements for ICOS Stations.....	7
1.4	Uncertainties of Calculations in the Marine CO ₂ System	7
1.5	Comparison between SOOP-based and Float-based Surface Carbon Observations	10
2	Results	12
2.1	North Atlantic pH Float Pilot Study	12
2.2	SOOP-based pH Observations.....	14
2.3	SOOP-Float pH Crossover Analysis	15
3	Conclusions & Recommendations.....	17
4	References	18

1 Background

The Global Carbon Budget, a prime product of international carbon research, is produced by a large group of scientists from dozens of research institutions worldwide working under the umbrella of the Global Carbon Project (GCP). The budget is in its 15th year now (Friedlingstein et al., 2022) and provides an in-depth look at the amount and fate of fossil fuels that nations around the world release. This prime product of the international carbon research community represents the best knowledge of the mean, variations, and trends in the perturbation of CO₂ in the environment since 1750, *i.e.* the onset of the industrial era.

The components of the annually released carbon budget include separate and independent estimates for (1) the CO₂ emissions from fossil fuel combustion and oxidation from all energy and industrial processes, (2) the emissions resulting from deliberate human activities on land and their partitioning among (3) the growth rate of atmospheric CO₂ concentration, the uptake of CO₂ (4) in the ocean and (5) on land (dto.).

The reported estimate of the global ocean anthropogenic CO₂ sink is calculated as the average of two estimates. The first estimate is derived as the mean over an ensemble of eight global ocean biogeochemistry models and the second estimate is obtained as the mean over an ensemble of seven observation-based data products, all of which are based on SOCAT, the “Surface Ocean CO₂ Atlas” (dto.). SOCAT is a publicly available synthesis product for quality-controlled, surface ocean *f*CO₂ (fugacity of carbon dioxide) observations made by the international marine carbon research community (Bakker et al., 2016). It is considered the best and largest such *f*CO₂ data product and therefore the backbone of the quantification of the ocean CO₂ sink.

In order for Argo float-based pH measurements to be most useful in the context of the quantification of the ocean CO₂ sink, a conversion to *f*CO₂ and subsequent inclusion into the SOCAT database would be required. For this, the float-based surface *f*CO₂ data would have to meet the quality requirements defined by SOCAT.

1.1 Chemistry of the Marine Carbon Dioxide System

A series of chemical equilibria govern the dissolution of CO₂ in water and its reaction with water. This gives rise to a suite of different chemical species, which are connected through the corresponding equilibrium reactions; these are: CO₂ in aquatic solution, carbonic acid, bicarbonate and carbonate ions (*i.e.*, CO₂(aq), H₂CO₃, HCO₃⁻, CO₃²⁻). 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 CO₂ system in seawater (Dickson et al., 2007).

These four measurable parameters are:

- Total dissolved inorganic carbon (DIC)
- Total alkalinity (TA)
- Fugacity/partial pressure of CO₂ in gas phase in chemical equilibrium with seawater (*f*CO₂/*p*CO₂)
- Total hydrogen ion concentration (pH)

It is possible, in theory, to obtain a complete description of the marine CO₂ system in a sample of sea water at a particular temperature and pressure provided that the following information is known (Dickson et al., 2007):

- Solubility constant for CO₂ in sea water (*K*₀),
- Equilibrium constants for each of the acid-base pairs that are assumed to exist in the solution,

- Total concentrations of all the non-CO₂ acid-base pairs,
- Values of at least two of the four CO₂-related parameters: DIC, TA, f/p CO₂, pH

If f CO₂ cannot be measured directly by BGC-Argo, which is due to limitations in sensor technology, two other of the four measurement parameters of the marine CO₂ system are required for its calculation. Therefore, float-based pH – the only marine CO₂ system parameter currently measurable operationally from floats – needs to be combined with a second parameter. In the present absence of suitable sensors for p CO₂, DIC or TA, a reasonable work-around is to predict TA using algorithms that employ variables such as T, S, pressure (P), and O₂, which are 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). In a regional context, simple TA vs. salinity regressions can often be employed with good success (Millero et al., 1998). Generally, such TA predictions have a typical overall uncertainty on the order of 6 μ mol 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 algorithms based on it (or other data products) is globally not uniform. Regional and/or seasonal biases are therefore to be expected, which have the potential to compromise the accuracy of f CO₂ calculated from measured pH and predicted TA.

1.2 SOCAT Quality Control

The SOCAT quality control is defined in a cookbook that is updated regularly to reflect improvement in quality procedures and methods (Lauvset et al., 2021). In the quality control process, a data set quality control flag is assigned to each data set contributing to SOCAT. Only data sets with a flag of A, B, C, D and E (with A being the most accurate one, see Table 1 for detail) will be finally included in the SOCAT data products. To assign the quality flag, it is necessary to evaluate both the data and metadata of a given data set.

In its current state, SOCAT only contains direct surface water f CO₂ measurements recorded by f CO₂ instruments. So it does not include f CO₂ values calculated from other carbon parameters, such as pH, TA or DIC (Bakker et al., 2016). Therefore, currently no pathway exists to ingest f CO₂ data calculated from float-based pH measurements into SOCAT. A future inclusion of BGC-Argo surface f CO₂ observations into SOCAT would therefore require joint discussions between BGC-Argo and SOCAT communities. Besides a general willingness of SOCAT to open an entry path for calculated f CO₂, a solid and effective quality control on the BGC-Argo side would be an indispensably prerequisite.

In the absence of a defined input pathway in SOCAT for calculated f CO₂ data, no criteria have been defined for this yet. An orientation for the f CO₂ quality requirement of SOCAT are the currently defined quality control flags for direct f CO₂ measurements (Table 1). In terms of the minimum accuracy requirement, the following three accuracy levels have been defined: ± 2 μ atm (flags A + B), ± 5 μ atm (flags C + D), and ± 10 μ atm (flag E). It is noteworthy that the main SOCAT synthesis f CO₂ data product as well as the SOCAT gridded f CO₂ data product explicitly exclude the flag E data as the accuracy is deemed inadequate for the main purpose of the SOCAT data. The E flag f CO₂ data are reported, however, in a separate data file that is typically not included in SOCAT-based analyses such as the Global Carbon Budget (Friedlingstein et al., 2022).

In order to be most widely useful, float pH-based f CO₂ data therefore would have to reach an accuracy of ± 5 μ atm or better. It is unlikely, that calculated f CO₂ data which don't have a demonstrable accuracy of ± 5 μ atm will ever get included in the main SOCAT synthesis product. It is also noteworthy that the highest quality flag A is only assigned to data with an estimated accuracy of ± 2 μ atm which is further corroborated by the existence of a high-quality crossover with another data set (flagged A or B).

Table 1: Data set $f\text{CO}_2$ quality control flags for SOCAT version 3 and later. All criteria need to be met for assigning a flag of A to E (Lauvset et al., 2021).

Flag	Criteria ^a
A	(1) Accuracy of calculated $f\text{CO}_2\text{rec}$ (at SST ^b) is better than 2 μatm . (2) A high-quality cross-over ^{c,d} with another data set (also flagged A or B) is available. (3) Followed approved methods/SOP ^e criteria. (4) Metadata documentation complete. (5) Data set QC was deemed acceptable.
B	(1) Accuracy of calculated $f\text{CO}_2\text{rec}$ (at SST) is better than 2 μatm . (2) Followed approved methods/SOP criteria. (3) Metadata documentation complete. (4) Data set QC was deemed acceptable.
C	(1) Accuracy of calculated $f\text{CO}_2\text{rec}$ (at SST) is better than 5 μatm . (2) Did not follow approved methods/SOP criteria. (3) Metadata documentation complete. (4) Data set QC was deemed acceptable.
D	(1) Accuracy of calculated $f\text{CO}_2\text{rec}$ (at SST) is better than 5 μatm . (2) Did or did not follow approved methods/SOP criteria. (3) Metadata documentation incomplete. (4) Data set QC was deemed acceptable.
E	(Primarily for alternative sensors) (1) Accuracy of calculated $f\text{CO}_2\text{rec}$ (at SST) is better than 10 μatm . (2) Did not follow approved methods/SOP criteria. (3) Metadata documentation complete. (4) Data set QC was deemed acceptable.
S (Suspend)	(1) More information is needed for data set before flag can be assigned (2) Data set QC revealed non-acceptable data and (3) Data are being updated (part or the entire data set).
X (Exclude)	The data set duplicates another data set in SOCAT.
N (New)	Data submitted to SOCAT that has not undergone independent data set quality control.
U (Updated)	Data re-submitted to SOCAT following updates by the data provider. Will be quality controlled as if new.
Q	A data set with conflicting flags, usually different flags in different regions.

These accuracy requirements are in agreement with the essential ocean variables (EOV) as defined by the Global Ocean Observing System (GOOS) and the approach of “weather” and “climate” goals taken by the Global Ocean Acidification Observing Network (GOA-ON) (Newton et al., 2015):

The “weather goal” is defined as measurements of a quality sufficient to identify relative spatial patterns and short-term variations and support the understanding of mechanistic responses to and impact on local, immediate ocean acidification dynamics. This implies an uncertainty of: $\text{pH} \pm 0.02$, $\text{TA/DIC} \pm 10 \mu\text{mol kg}^{-1}$ and $f\text{CO}_2 \pm 2.5\%$ (= 10 μatm at 400 μatm).

The “climate goal” is defined as measurements of a quality sufficient to assess long-term trends with a defined level of confidence and support detection of the long-term anthropogenically driven changes in hydrographic conditions and carbon chemistry over multi-decadal time scales. This implies an uncertainty of: pH ± 0.003 , TA/DIC $\pm 2 \mu\text{mol kg}^{-1}$ and $f\text{CO}_2 \pm 0.5\%$ (= 2 μatm at 400 μatm).

The SOCAT flags A and B meet the “climate goal” definition and flag E fulfils the “weather goal”, while the C and D flags represent an intermediate level of accuracy. It is an expressed goal of SOCAT to reduce that amount of C/D-flagged data and increase that of A/B.

1.3 Quality Requirements for ICOS Stations

The European Infrastructure “Integrated Carbon Observation System” (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).

Marine stations of ICOS deliver high quality $f\text{CO}_2$ data to the Carbon Portal (CP) and SOCAT. To ensure that every station fulfils the quality requirements of ICOS, they undergo a two-step labeling process. In the first step, the station is evaluated on whether or not it can provide high-quality data according to ICOS standards. In the second step, the station must prove this by sending data and metadata to the Ocean Thematic Centre (OTC) showing that it measures data with the desired quality.

The labelling scheme of ICOS-Oceans (Skjelvan et al., 2021) identifies two types of station (*i.e.*, Ship of Opportunity Line (SOOP) and Fixed Ocean Station (FOS)) which each come in two different classes (*i.e.*, “Class 1” and “Class 2”) according to the set of variables measured (Table 2).

Table 2: Minimum measurement requirements for ICOS marine stations (Skjelvan et al., 2021).

Station Class	Ship of Opportunity Line (SOOP)	Fixed Ocean Station (FOS)
Class 1	$f\text{CO}_2 (\pm 2 \mu\text{atm})$	$f\text{CO}_2 (\pm 10 \mu\text{atm})$ TA or DIC Oxygen
Class 2	TA or DIC Oxygen	At surface: Nutrients (nitrate, silicate and phosphate)

The significantly reduced requirement for $f\text{CO}_2$ accuracy measured by FOS recognizes the currently existing limitations of ocean $f\text{CO}_2$ measurement capabilities. The understanding in ICOS is that fluxes based on $f\text{CO}_2$ data from FOS cannot be determined to the limited accuracy desired for large scale carbon budgets (unless a higher accuracy can be demonstrated).

1.4 Uncertainties of Calculations in the Marine CO₂ System

When marine CO₂ system variables are used to calculate others, errors both from the input parameters as well as from the thermodynamic constants used are propagated. This has been explored by Orr et al. (2018) and applied to specific ICOS cases (Steinhoff, 2020). The case relevant for BGC-Argo is the combination of pH (measured directly by floats) and TA (predicted via specific algorithms) to calculate $f\text{CO}_2$ (Figs. 1 and 2).

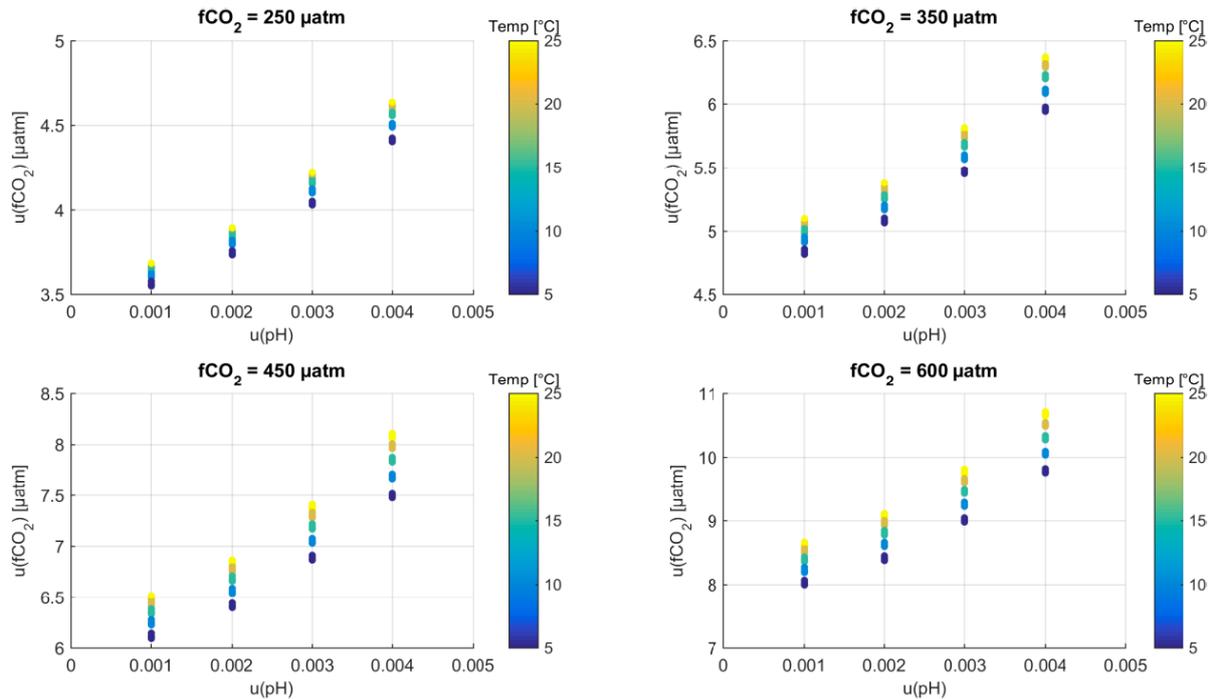


Fig. 1: Resulting uncertainty in $f\text{CO}_2$ ($u(f\text{CO}_2)$, y-axis) when using pH and TA as input parameters vs. uncertainty in pH (x-axis). The results are shown at four different $f\text{CO}_2$ levels (separate panels) and five different temperatures (color-coded).

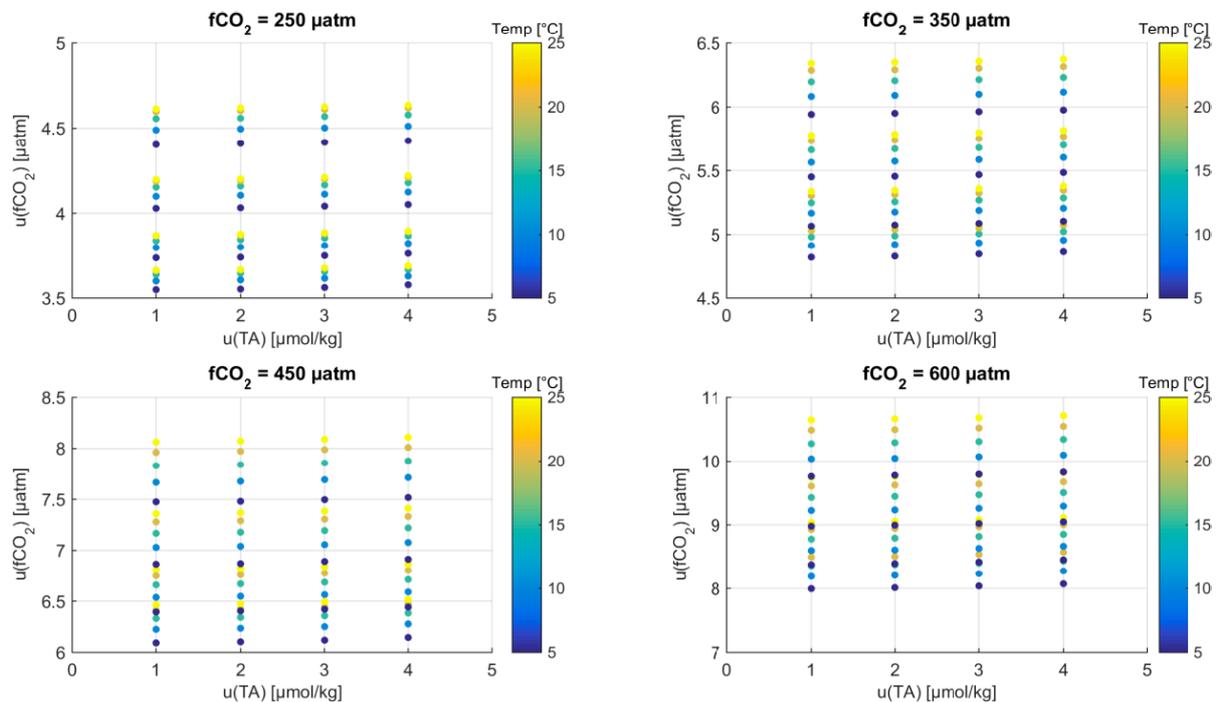


Fig. 2: Resulting uncertainty in $f\text{CO}_2$ ($u(f\text{CO}_2)$, y-axis) when using pH and TA as input parameters vs. uncertainty in TA (x-axis). The results are shown at four different $f\text{CO}_2$ levels (separate panels) and five different temperatures (color-coded).

According to the analyses of Steinhoff (2020), $f\text{CO}_2$ can be calculated with an uncertainty of $< 10 \mu\text{atm}$ for $f\text{CO}_2$ values below *ca.* $600 \mu\text{atm}$ when using TA and pH as input with uncertainties of $u(\text{pH}) < 0.003$ and $u(\text{TA}) < 4 \mu\text{mol kg}^{-1}$. This calculation based on a protocol developed by Orr et al. (2018) includes a formal error propagation of the errors of all thermodynamic constants applied in calculations. As such it can be viewed as a conservative error estimate. The overall contribution of $u(\text{TA})$ to the overall uncertainty of $f\text{CO}_2$ is comparatively small. Thus, the currently achievable uncertainty in predicted TA of about $6 \mu\text{mol kg}^{-1}$ (Bittig et al., 2018; Carter et al., 2018), which corresponds to an uncertainty in pH of about 0.001, appears to not be a major obstacle. It should be noted, however, that regional/seasonal biases in estimated TA may occur and that during times and in regions of high surface concentrations of phosphate and silicate a small additional uncertainty in TA arises from the fact that these nutrient concentrations cannot be measured by floats. Another potential complication is the existence of uncharacterizable organic alkalinity in coastal waters or during phytoplankton bloom situations which can render TA a somewhat ill-defined property (Kerr et al., 2021).

In order to achieve the minimum $f\text{CO}_2$ accuracy of $\pm 5 \mu\text{atm}$ as required for SOCAT quality flags C and D, the required $u(\text{pH})$ varies between 0.0035 and 0.0075 over the $f\text{CO}_2$ range 250–550 μatm and the temperature range 5–25°C with $u(\text{TA}) = 0$ (Fig. 3) when only the error propagation of the input variables are included (an not those of the thermodynamic constants). This estimate is therefore more optimistic than the formal estimate cited above. Including $u(\text{TA})$ would reduce the required $u(\text{pH})$ to about 0.0025 – 0.0065 in this optimistic approach. This, however, is currently not systematically met with the reported uncertainties of $u(\text{pH}) = 0.007$ (Maurer et al., 2021). So the major limiting factor for the achievable overall accuracy of $f\text{CO}_2$ is the level of uncertainty to which float-based pH can be improved.

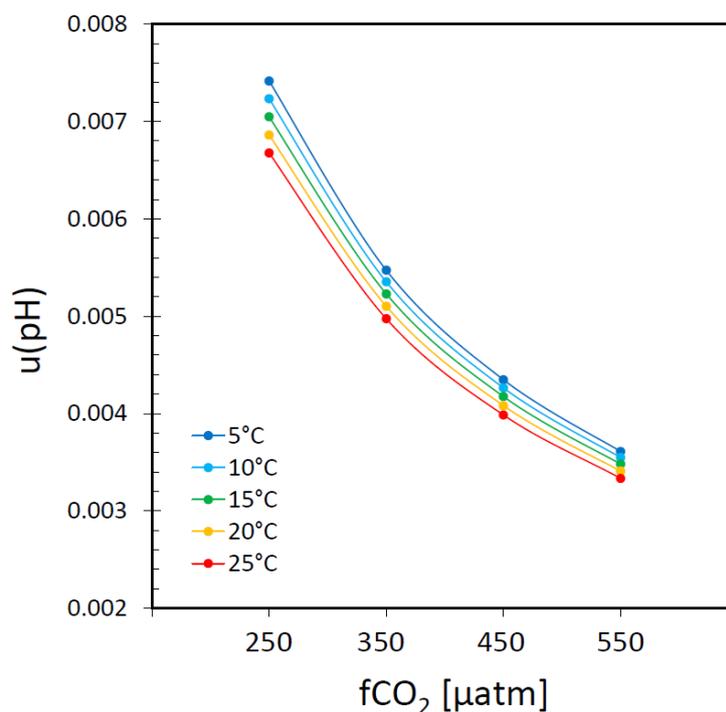


Fig. 3: Maximum allowed pH uncertainty ($u(\text{pH})$) for calculation of $f\text{CO}_2$ with an uncertainty of $< 5 \mu\text{atm}$ from TA as a function of $f\text{CO}_2$ and temperature. Note that the uncertainty of TA was not included. If $u(\text{TA}) = 6 \mu\text{mol kg}^{-1}$ is included, the allowed $u(\text{pH})$ is lower by approx. 0.001.

Williams et al. (2017) state a standard uncertainty of 2.7% (or $11 \mu\text{atm}$ at $400 \mu\text{atm}$) for $p\text{CO}_2$ calculated from float-pH and estimated TA for the Southern Ocean based on a thorough analysis of all potential

sources of uncertainty. This adds to the evidence that float-based pH is currently far from the level of uncertainty required for potential inclusion in the SOCAT database (better than $\pm 5 \mu\text{atm}$ for flags C and D). Even the achievement of the flag F introduced by SOCAT for sensor data from fixed-point observatories (uncertainty $< 10 \mu\text{atm}$) is currently not demonstrated.

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

The global Ship-of-Opportunity network (SOOP) is operated in the European Research ICOS (see above) and the 'Surface Ocean CO₂ Reference Observing Network' (SOCONET). SOCONET is a volunteer group of established operators who provide quality global surface ocean CO₂ data. Participants perform automated measurements of surface water and atmospheric CO₂ from Ships-of-Opportunity and moorings.

The data from these networks are submitted routinely and according to specific requirements to the SOCAT, where they undergo defined quality control measures. SOCAT is a synthesis activity for quality-controlled, surface ocean $f\text{CO}_2$ observations by the international marine carbon research community (>100 contributors). The data are publicly available, discoverable and citable. The SOOP network and the SOCAT data product are relevant for BGC-Argo in two ways:

- (1) BGC-Argo float-based pH observations can be converted to $f\text{CO}_2$ and potentially be submitted to the SOCAT data-base. This, however, requires better understanding of the quality of pH-based surface $f\text{CO}_2$ observations by the Argo observation network as discussed above.
- (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.

For this purpose, SOCONET/ICOS, which typically only features routine $p\text{CO}_2/f\text{CO}_2$ observations, need to be augmented and harnessed for BGC-Argo. The addition of TA as a second analytical variable of the marine CO₂ system to the suite of autonomous SOOP observations has been demonstrated to be methodologically feasible now (Seilmann et al., 2019; 2020a; 2020b). We have therefore added TA to the measurement portfolio of the 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 two purposes in the context of BGC-Argo:

- (1) From high-quality co-located and synchronous SOOP-based $f\text{CO}_2$ 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 $f\text{CO}_2$, 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.

As a further augmentation of the 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.

In order to compare SOOP-based and float-based surface carbon observations, we searched for direct crossovers between float surfacings and SOOP passages. SOCAT has defined a crossover by the following criterion x that combines distance d (in km) and time t (in d), where one day of separation in time is heuristically equivalent to 30 km of separation in space. In SOCAT, the value used for x is 80 km:

$$\sqrt{(\Delta d)^2 + (30 \cdot \Delta t)^2} \leq x \text{ km}$$

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.

2 Results

2.1 North Atlantic pH Float Pilot Study

Within the North Atlantic pilot study a total of 10 pH/O₂ floats from two manufacturers have been purchased and deployed as part of the pilot study so far (Table 3, Fig. 4). Further floats will be deployed in 2023.

Table 3: BGC-Argo floats with pH and O₂ 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	Inactive, 107 cycles, faulty GPS system
6904110	Provor/nke	July 2021	Subpolar N.A.	Active, 77 cycles, pH sensor failure after few profiles
6904111	Provor/nke	July 2021	Subpolar N.A.	Active, 95 cycles, pH sensor failure after few profiles
6904112	Provor/nke	July 2021	Subpolar N.A.	Active, 76 cycles Compromised pH sensor quality
6904114	Provor/nke	August 2021	Labrador Sea	Active 66 cycles Compromised pH sensor quality
6904115**	Provor/nke	August 2021	Labrador Sea	Inactive, 63 cycles, pH sensor failure after few profiles
6904231	Provor/nke	October 2022	Labrador Sea	Active, 20 cycles pH quality not yet assessed

*As of Nov. 5, 2022.

** Float recovery in August, 2022.

Unfortunately, the 10 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 (Communication Ken Johnson, AST-23 meeting, April 2022) 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 a little more than half of the floats are at all useful towards the goals of the pilot study.

We note that in our analyses the two long-lasting floats deployed in 2018 (3901668 + 3901669) showed the most stable behaviour of the pH sensor, and are thus assumed to represent the more optimal case

for achievable performance of this current technology. Floats deployed in 2021 (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.

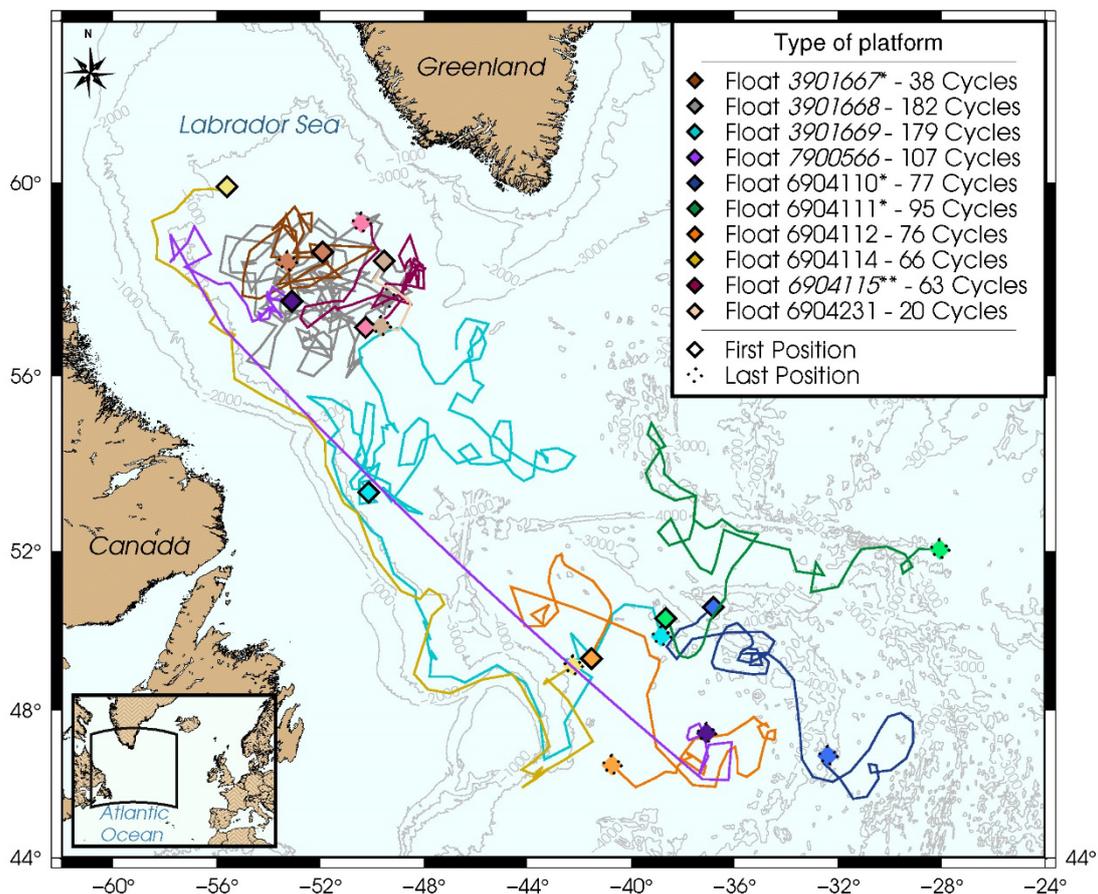


Fig. 4: Map of the Northwest Atlantic with Labrador Sea and North Atlantic Current showing the trajectories of all 10 pH/O₂ floats deployed so far in our pilot study. In the legend, floats in *italics* are inactive. * Float with a faulty pressure and/or pH sensor. ** Float recovered. Dotted points show the last locations as of November 5th, 2022.

On all floats, the standard delayed-mode cookbook procedures for pH (pump-offset correction, CANYON-B reference, segment method, 1500 dbar reference level, 0.005 pH units mCP dye offset correction, temperature correction) were applied according to Johnson et al. (2017).

We note that the two reference methods CANYON-B vs. LIR yielded corrected pH values that differed by about 0.02 pH units. The choice of the reference method, therefore, incurs a huge uncertainty much above the desired accuracy level. Also, the choice of the reference depth (*e.g.*, 1500 vs. 1900 dbar) contributed about 0.01 to the uncertainty of pH. As no objective criteria exist to guide the choice of both reference method and depth, the resulting uncertainty is on the order of a few hundred of a pH unit which is far in excess of tolerable levels. This may be a peculiarity of the subpolar North Atlantic with its complex water masses, deep convection, and deep penetration of anthropogenic CO₂. It points, however, at a critical shortcoming of the current delayed-mode quality control which requires an additional control mechanism, ideally at the surface, for independent assessment and possibly also the correction of float-pH data.

2.2 SOOP-based pH Observations

On the North Atlantic SOOP line operated by us under ICOS (DE-SOOP-Atlantic Sail), 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 (2 trans-Atlantic crossings each) the instrument was calibrated in port against fresh CRM (Certified Reference Material) provided by the Dickson Lab (batch 190, accuracy ± 0.0014 pH units).

Each system calibration typically has about 8 (5-10) repeat pH measurements on a single, freshly opened CRM bottle. These pre- and post-calibration runs are pretty stable for each meta-cresol purple (mCP) indicator bag which lasts for 4 roundtrips (Fig. 5). Some variability between mCP batches and even bags of the same batch is evident. Therefore an individual pH correction is applied to each mCP bag based on the error-weighted mean of the 5 calibration runs bracketing 4 roundtrips of the vessel. The standard error of this weighted mean is typically < 0.002 pH units. The overall accuracy of SOOP-pH is estimated to be ± 0.003 pH units. The data indicate a small temporal drift of the pH measurements between individual mCP bag which does not show within a given bag. It is currently unclear whether this is due to some deterioration in the system's optical path or aging within the stored mCP indicator bags.

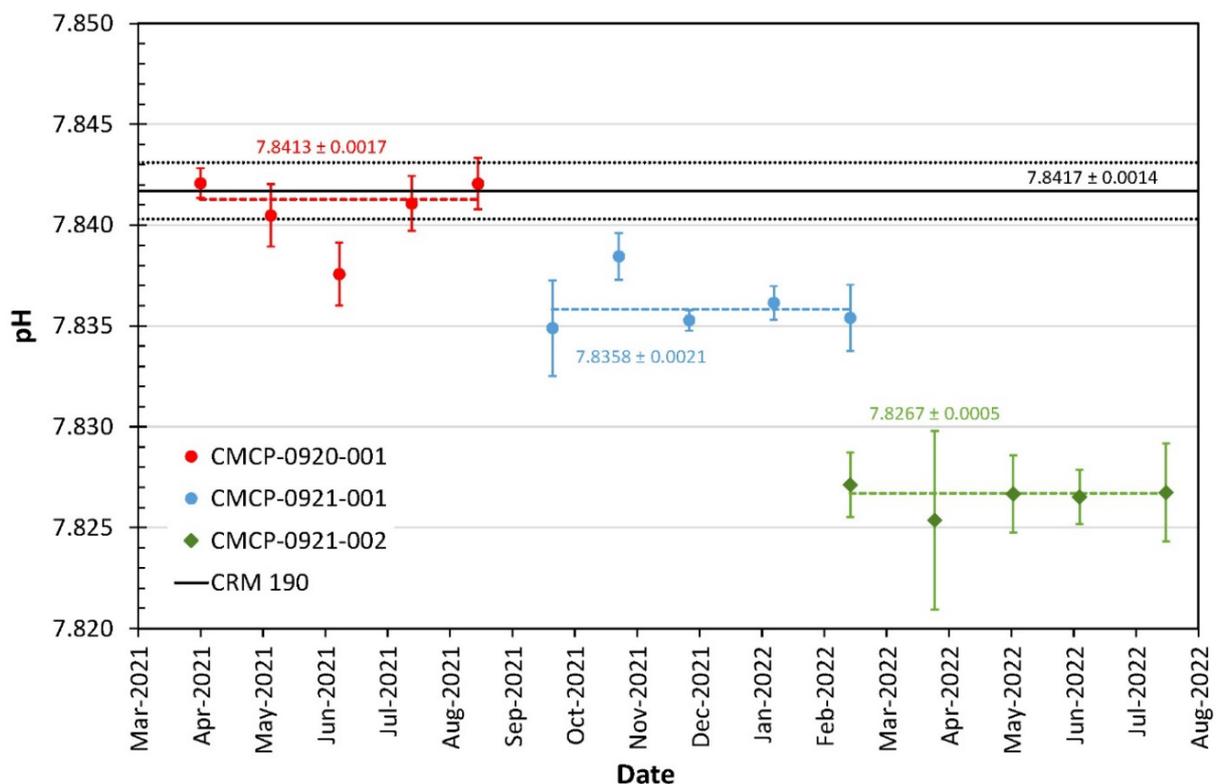


Fig. 5: pH measurements performed on CRM batch 190 with the Contros HydroFIA™ pH system before and after each 5-week roundtrip of our Ship-of-Opportunity M/V Atlantic Sail. Four consecutive roundtrips are always carried out with one reagent bag of the mCP indicator (CMCP-mmyy of batch preparation-number of bag). Adjustments of measured pH to the nominal pH value assigned to the CRM (7.8417 ± 0.0014 at 25°C) are based on the error-weighted mean of all CRM measurements carried out per individual mCP bag (typically at 5 different times with 5-10 repeat measurements each, values shown in plot).

2.3 SOOP-Float pH Crossover Analysis

We have adopted the SOCAT crossover criterion which tries to optimize between spatial and temporal mismatch to search for SOOP-float crossovers. We have explored a larger range in x (80-300 km) and we have also experimented with additional crossover criteria such as maximum offset in temperature and/or salinity. Given the limitations of our dataset (mostly due to massive manufacturing problems of the 2020/1 pH sensor series) no robust recommendations can be drawn from these experiments. To operationalize this approach, regional refinements and most likely additional crossover criteria (*e.g.*, max. temperature difference) and procedures (*e.g.*, correction of SOOP pH to the temperature of the float pH) will have to be developed by further studies.

SOOP data falling into a given SOCAT-like crossover criterion 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 statistical weight.

For 2021 and the major part of 2022, we achieved several crossovers between surfacings of three pH/O₂ floats (WMO 3901669, WMO 6904111, WMO 6904112) and our SOOP line in the subpolar North Atlantic. All of these are in the western subpolar North Atlantic where the coincidence of Atlantic waters of southern provenience and Labrador Sea waters of polar origin complicate the interpretation of crossover qualities.

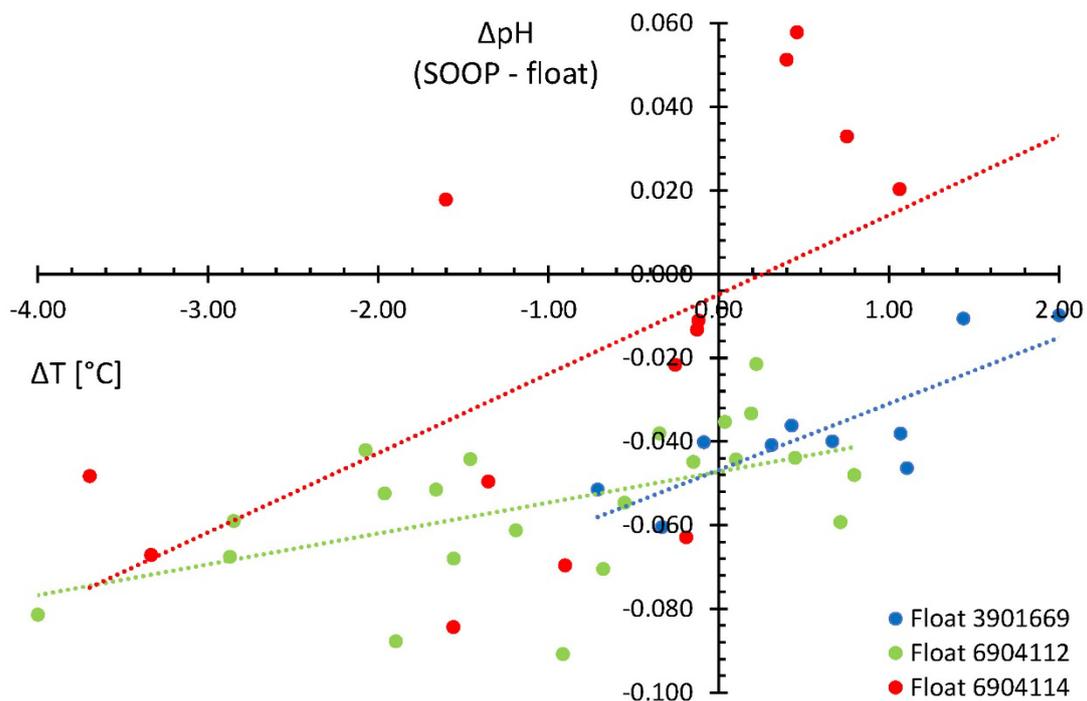


Fig. 6: Offset between SOOP pH and fully corrected float pH (y-axis) as a function of temperature difference (x-axis) for crossovers ($x = 400$ km, $\Delta T = 4^\circ\text{C}$) of three different floats.

In order to yield a larger number of crossovers we applied a rather large search window of $x = 400$ km. The resulting list of crossovers was reduced by the addition requirement $-4.0 < \Delta T < 4.0$ which excluded a few data points. For each crossover, the SOOP-pH was corrected from SOOP-T to float-T using SOOP-TA. We assume that remaining differences in pH are not driven by differences in T but represent a bias in SOOP-pH and possibly a contribution from different water mass properties represented in the SOOP and float data. We then plotted ΔpH (SOOP – float) as a function of ΔT (Fig. 6). The ΔpH at $\Delta T = 0$ should

then represent the pH offset between SOOP and float. By fitting a linear regression to the data, the pH offset can be estimated more robustly. We think this regression using crossovers achieved with a relatively wide search window yields a more robust ΔpH estimate as an average of a small number of crossovers found with a smaller search window. Another option would be to apply a temperature correction of the SOOP pH data to the float T at each crossover. Clearly, the method of choice for the detection and evaluation of SOOP-float crossovers needs to be explored more in-depth with a larger dataset of floats from different regions.

The results of our analysis (Table 4) show for two of the three floats (3901669 and 6904112) a relative good linear relationship between ΔpH and ΔT which allows the estimation of the pH offset at $\Delta T = 0$ with a reasonably small error. Interestingly, the resulting ΔpH is almost identical for the two floats. For the third float (6904114), more scatter leads to a less well-constrained pH offset, which in fact is indistinguishable from $\Delta\text{pH} = 0$. The mean ΔT of the crossovers is within $\pm 1^\circ\text{C}$ for each float. The corresponding ΔS of these crossovers is on the order ± 0.5 . Calculating the pH offset as a function of ΔS yields ΔpH values which are statistically indistinguishable from the ones based on ΔT (but have slightly larger uncertainty). This indicates that the water mass correction achieved through the regression approach is generally effective.

Table 4: Statistics of the crossover analysis for SOOP and float pH data.

Float	$\Delta\text{pH at } \Delta T=0$		ΔT		ΔS	
	mean	std dev	mean	std dev	mean	std dev
3901669	-0.047	0.004	0.87	1.00	0.41	0.50
6904112	-0.047	0.004	-0.98	1.29	-0.40	0.58
6904114	-0.005	0.010	-0.56	1.57	-0.56	1.57

3 Conclusions & Recommendations

Although the results only represent a limited number of floats, few crossovers and a relatively small regional domain the following conclusions and recommendations can already be made:

Reasonably consistent results from the subpolar North Atlantic indicate that float pH may be biased by several hundredths of a pH unit at the surface. In fact, two of the three floats suggest a virtually identical pH correction by -0.047 , *i.e.*, the float pH is systematically higher on average by this amount when compared to SOOP pH. With an uncertainty of ± 0.01 in pH corresponding to an uncertainty in $p\text{CO}_2$ of about $10 \mu\text{atm}$, anything above that 0.01 threshold appears not suitable for estimating the air-sea CO_2 flux from float-based pH observations even under conditions of the “weather goal”. These early findings therefore warrant further and more sophisticated analyses to better constrain 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 as the upper 2000 m of the column show widespread penetration of anthropogenic CO_2 and hence temporal change. Estimates of the error contribution from the choice of the deep ocean reference method and the reference depth do not explain, however, the full magnitude of the offset to SOOP pH at the surface. This may hint at the existence of a depth-variable component of the pH offset which would – similar to oxygen – make an additional surface reference necessary.

With larger amounts of SOOP-float crossover data and optimized crossover criteria, an independent quality control and perhaps even correction of float surface pH may be achieved in areas of good SOOP coverage. SOOP could thus be implemented in the DMQC routine for float pH. In the future, algorithms synchronizing float profiling routines with SOOP line schedules and projected crossings may even be employed to actively chase crossovers in a systematic way. This is something that should be explored in a systematic way.

4 References

- Bakker, D. C. E., Pfeil, B. Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., Smith, K., Cosca, C., Harasawa, S., Jones, S. D., Nakaoka, S., Nojiri, Y., Schuster, U., Steinhoff, T., Sweeney, C., Takahashi, T., Tilbrook, B., Wada, C., Wanninkhof, R., Alin, S. R., Balestrini, C. F., Barbero, L., Bates, N. R., Bianchi, A. A., Bonou, F., Boutin, J., Bozec, Y., Burger, E. F., Cai, W.-J., Castle, R. D., Chen, L., Chierici, M., Currie, K., Evans, W., Featherstone, C., Feely, R. A., Fransson, A., Goyet, C., Greenwood, N., Gregor, L., Hankin, S., Hardman-Mountford, N. J., Harlay, J., Hauck, J., Hoppema, M., Humphreys, M. P., Hunt, C. W., Huss, B., Ibáñez, J. S. P., Johannessen, T., Keeling, R., Kitidis, V., Körtzinger, A., Kozyr, A., Krasakopoulou, E., Kuwata, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lo Monaco, C., Manke, A., Mathis, J. T., Merlivat, L., Millero, F. J., Monteiro, P. M. S., Munro, D. R., Murata, A., Newberger, T., Omar, A. M., Ono, T., Paterson, K., Pearce, D., Pierrot, D., Robbins, L. L., Saito, S., Salisbury, J., Schlitzer, R., Schneider, B., Schweitzer, R., Sieger, R., Skjelvan, I., Sullivan, K. F., Sutherland, S. C., Sutton, A. J., Tadokoro, K., Telszewski, M., Tuma, M., Van Heuven, S. M. A. C., Vandemark, D., Ward, B., Watson, A. J., Xu, S. (2016). A multi-decade record of high quality fCO₂ data in version 3 of the Surface Ocean CO₂ Atlas (SOCAT). *Earth System Science Data*, 8, 383–413, doi:10.5194/essd-8-383-2016.
- 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
- Denvil-Sommer, A., M. Gehlen, and M. Vrac. 2021: Observation system simulation experiments in the Atlantic Ocean for enhanced surface ocean pCO₂ reconstructions. *Ocean Sci.*, 17, 1011–1030, doi: 10.5194/os-17-1011-2021.
- Dickson, A. G., C. L. Sabine, and J. R. Christian (Eds.), 2007: Guide to Best Practices for Ocean CO₂ Measurements. PICES Special Publication 3, 191 pp.
- 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. Lujikx, 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 System Science 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(6), 3249–3256, doi:10.1021/acs.analchem.5b04653.
- Johnson, K. S., J. N. Plant, and T. L. Maurer (2017). Processing BGC-Argo pH data at the DAC level, doi:10.13155/57195.

- Kerr, D. E., P. J. Brown, A. Grey, B. P. Kelleher (2021). The influence of organic alkalinity on the carbonate system in coastal waters. *Marine Chemistry*, 237, 104050, doi: 10.1016/j.marchem.2021.104050.
- Lauvset, S., K. Currie, N. Metzl, S.-I. Nakaoka, D. Bakker, K. Sullivan, A. Sutton, K. O'Brien, A. Olsen (2021). *SOCAT Quality Control Cookbook for SOCAT version 7*, 16 pp.
- Millero, F. J., K. Lee, and M. Roche, 1998: Distribution of alkalinity in the surface waters of the major oceans. *Mar. Chem.*, 60, 111–130.
- Newton J.A., Feely R. A., Jewett E. B., Williamson P., and Mathis J., 2015: *Global Ocean Acidification Observing Network: Requirements and Governance Plan. Second Edition, GOA-ON*, http://www.goa-on.org/docs/GOA-ON_plan_print.pdf.
- 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. Schirnack, 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.
- Orr, J. C., J. M. Epitalon, and J. P. Gattuso (2015). Comparison of Ten Packages That Compute Ocean Carbonate Chemistry, *Biogeosciences*, 12, 1483–1510.
- Orr, J. C., J. M. Epitalon, A. G. Dickson, and J. P. Gattuso (2018). Routine Uncertainty Propagation for the Marine Carbon Dioxide System. *Marine Chemistry*, 207, 84–107.
- 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.
- Skjelvan, I., T. Steinhoff, R. Battisti, E. Jeansson, S. Jones, C. Landa, S. Lauvset, A. Olsen, B. Pfeil, and A. Rutgersson (2021). *ICOS Ocean Station Labelling STEP 2, Version 6.1, ICOS Ocean Thematic Centre*, doi: 10.18160/8SDC-K4FR.
- Steinhoff, T. (2020). *Uncertainty analysis for calculations of the marine carbonate system for ICOS-Oceans stations, IOCS OTC, NORCE Norwegian Research Centre AS, Bergen, Norway*, doi: 10.18160/vb7c-z758.
- Williams, N. L., L. W. Juraneck, 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. Juraneck, 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, Y. Takeshita (2017). Calculating surface ocean pCO₂ from biogeochemical Argo floats equipped with pH: An uncertainty analysis. *Global Biogeochem. Cycles*, 31, 591–604, doi:10.1002/2016GB005541.