Measuring pH in the Arctic Ocean: colorimetric method or SeaFET?

Victoire Rérolle[a], Diana Ruiz-Pino[a], Mehrad Rafizadeh[a], Socratis Loucaides[b], Stathys Papadimitriou[c], Matthew Mowlem[b], Jianfang Chen[d]

a. Sorbonne Universités (UPMC, Univ Paris 06)-CNRS-IRD-MNHN, LOCEAN Laboratory, 75005, Paris, France
b. National Oceanography Centre, Southampton, SO143ZH, United Kingdom
c. Ocean Sciences, College of Natural Sciences, Bangor University, Menai Bridge, LL59 5AB, United Kingdom
d. Second institute of Oceanography, Hangzhou City, 310012, China

Corresponding author: Victoire.rerolle@locean-ipsl.upmc.fr
Université Pierre & Marie Curie - boîte 100 - 4, Place Jussieu - 75252 PARIS CEDEX 05

Co-authors email addresses:
ruizlod@locean-ipsl.upmc.fr
Mehrad.RAFIZADEH@locean-ipsl.upmc.fr
s.loucaides@noc.ac.uk
s.papadimitriou@bangor.ac.uk
matm@noc.ac.uk
biogeo_chen@hotmail.com

Abstract

The suitability of the colorimetric method in a custom-made instrumental set-up and the commercial potentiometric SeaFET® electrode sensor to measure pH in surface oceanic water in the Arctic was investigated during the Chinese Arctic Research Expedition (CHINARE) in summer 2014. The instruments were set up in parallel on the on-board underway seawater supply for 65 days, enabling comparison in various conditions in the Arctic Ocean from the Chukchi Sea to the ice-covered high latitudes (81°N) and the open-water North-western Pacific Ocean. Overall, the instruments yielded pH datasets of similar high quality (method uncertainty ≤ 0.010). Detailed comparison with the parallel colorimetric pH measurements indicated that the measurements with the SeaFET external electrode in the low salinity ice-covered area were problematical and that the internal reference electrode failed after almost 2 months of cruise. Reasons for discrepancies between the data from the two instruments and recommendations for the use of either instrument for pH measurements in the Arctic are discussed. Finally, the investigation yielded a reliable high-resolution pH dataset in surface waters along a transect from the Pacific to the Arctic Ocean. Large pH variations were observed in the ice-free Arctic surface waters, with pH ranging between 7.98 and 8.49. The highest pH values were observed at the ice edge, whereas a relatively invariable pH (8.02 ± 0.02) was measured in under-ice seawater in the ice-covered Arctic Ocean. The high resolution surface seawater pH data set obtained here could be used as reference to detect the on-going acidification rate in the Pacific Arctic.

Key words

pH, Arctic Ocean, sensors, colorimetric, SeaFET, acidification

1. Introduction

The increasing concentration of CO_2 in the atmosphere has significant impacts on the global climate, which lead to dramatic environmental changes in the Arctic [IPCC, 2013]. Reduction of the Arctic sea ice cover as a result of global warming opens up new prospects in the Arctic, such as the opening of new shipping routes and resource exploitation. However, ongoing environmental changes are also likely to have a significant cost. Questions arise about how the marine carbon cycle will respond to these changes and, particularly, how the uptake and sequestration of atmospheric CO_2 in the Arctic Ocean will impact oceanic pH (acidification: pH decrease) via the CO_2-seawater reaction system (oceanic carbonate system). The Arctic Ocean is particularly vulnerable to acidification because the large freshwater inputs from continental glacier melts transported by rivers reduce alkalinity, making it less effective at neutralizing the acidifying effect of CO_2 [*Bates and Mathis*, 2009]. Additionally, the cold temperature of the Arctic Ocean facilitates the absorption of anthropogenic CO_2 from the atmosphere through increased solubility. Due to the intense ventilation of its subsurface layers, the Arctic Ocean is estimated to contain nearly two times the global mean concentration of anthropogenic CO_2 [*Tanhua et al.*, 2009]. Important increases of anthropogenic carbon concentrations have recently been estimated in the intermediate layers of the Arctic Ocean [*Ericson et al.*, 2014], leading to a significant acidification of these water masses. In order to determine ongoing changes in the carbonate chemistry of the Arctic Ocean, it is essential to expand observational capacity.

The spatial and temporal distribution of pH and its variability remain unknown in most of the Arctic regions. Most of the existing pH data from the Arctic Ocean have been calculated from measurements of other parameters of the oceanic carbonate system, typically total dissolved inorganic carbon (DIC) and total alkalinity (TA), with the inherent errors of these calculations (e.g. [*Millero et al.*, 2002]). Generally, the Arctic Ocean remains one of the less documented oceans in terms of the regional dynamics of the carbonate system (e.g. GLODAP and SOCAT databases [*Sabine et al., 2005; Bakker et al., 2014*]). Available oceanographic data have been collected mostly in summer and across the shallow and large shelf [*Bates and Mathis*, 2009; *Miller et al.*, 2014]. Currently, the carbonate system parameter with the most abundant dataset is the partial pressure of CO_2 (pCO_2) as a result of the development of automated shipboard pCO_2 analysers running continuously on the underway water supply of scientific R/Vs, which are relatively easy to maintain [*Bakker et al.*, 2014]. The first surface oceanic pH values measured with an automated instrument along a cruise transect in the Arctic has recently been reported by Tynan et al. [2016] for the Fram strait region. The high resolution of the dataset allowed the investigators to demonstrate that local advection pre-determines the horizontal distribution of the carbonate system whereas biological activity enhances its high variability within water masses. Automated shipboard and in situ sensors are therefore essential to obtain continuous data with high spatial and temporal resolution in order to resolve natural variations of the carbonate system and underlying changes due to anthropogenic activities [*Takeshita et al.*, 2015].

Quality pH measurements are critical for long-term ocean acidification research (precision ≤ 0.003 pH unit; GOA-ON report; Newton *et al.*, 2015) but challenging to obtain with an automated instrument, particularly in the Arctic Ocean where salinity and temperature can be low and surface ice (sea ice) is extensive. The challenge is not only to find pH sensors with the longevity, stability, reliability, and robustness required for deployments in the Arctic, but also the measurement method

must be capable of accurate pH measurement in Arctic conditions. Several automated underway pH systems have recently been developed based on the colorimetric method [*Assmann et al.*, 2011; *Rérolle et al.*, 2012]. An advantage of the colorimetric method is that the temperature and salinity dependence of the (purified) pH indicator dye *meta*-Cresol Purple (mCP) is already well-characterised in the temperature (T) and practical salinity (S) ranges $5 \leq T \leq 35$ °C and $20 \leq S \leq 40$ [*Liu et al.*, 2011]. The colorimetric method can provide high quality data (accuracy ~0.005 pH unit and precision ≤ 0.0005 pH unit, [*Carter et al.*, 2013]) in wide ranges of salinity and temperature with a negligible drift over long periods of time [*Rérolle et al., 2012*]. However, this method has the drawback that it consumes indicator, which, as it should be purified to achieve high accuracy measurements [*Yao et al.*, 2007], is currently not easy to obtain. Additionally, colorimetric analysers are usually relatively complex, with pumps, valves, mixer, and fragile optical components which make them expensive, power hungry and potentially unreliable over long (> 6 month) deployments. An alternative method is potentiometric pH determination using either glass electrodes [*Seiter and DeGrandpre*, 2001] or ion-sensitive field-effect transistors (ISFET) [*Shitashima et al.*, 2002]. Potentiometric sensors have the key advantages of being relatively small, not consuming reagents, and having fast response times allowing for continuous high frequency measurements. On the other hand, these sensors can drift over time and require co-localised discrete samples to correct possible calibration offset [*Bresnahan et al.*, 2014]. Potentiometric sensors can also be sensitive to salinity and temperature changes. A recently developed ISFET-based pH sensor (SeaFET®) has been demonstrated to have great stability over 9 months of deployment at surface seawater [*Bresnahan et al.*, 2014; *Martz et al.*, 2010]. The use of the SeaFET pH sensors is increasing rapidly worldwide because of its ease of use and good performance in seawater [*Bresnaha et al.*, 2014; *Hofmann et al.*, 2011].

In this study, we compare the performance of a custom-made colorimetric analyser and a SeaFET sensor during a transect in the Arctic Ocean. We assess the quality of the pH data obtainable from these pH instruments and discuss practical considerations regarding instrumentation challenges. To this end, both instruments were connected to the underway seawater supply of the *MV XueLong* during the 6th Chinese Arctic Research Expedition (CHINARE) in summer 2014. Finally, a unique high-resolution pH dataset is presented for the Pacific sector of the Arctic Ocean.

2. Materials and methods

The 6th Chinese Arctic Research Expedition (CHINARE) took place between the 11th of July and the 23rd of September 2014 aboard the icebreaker *MV XueLong* (“Snow Dragon”). The ship visited various oceanic areas in the North Pacific, the Bering Sea, the Chukchi Sea, the Canada Basin, and the Arctic Ocean where it was stationed for an 8-day ice camp from the 18th (157.6°W;80.9°N) to the 25th (156.2°W;81.1°N) of August (Figure 1).

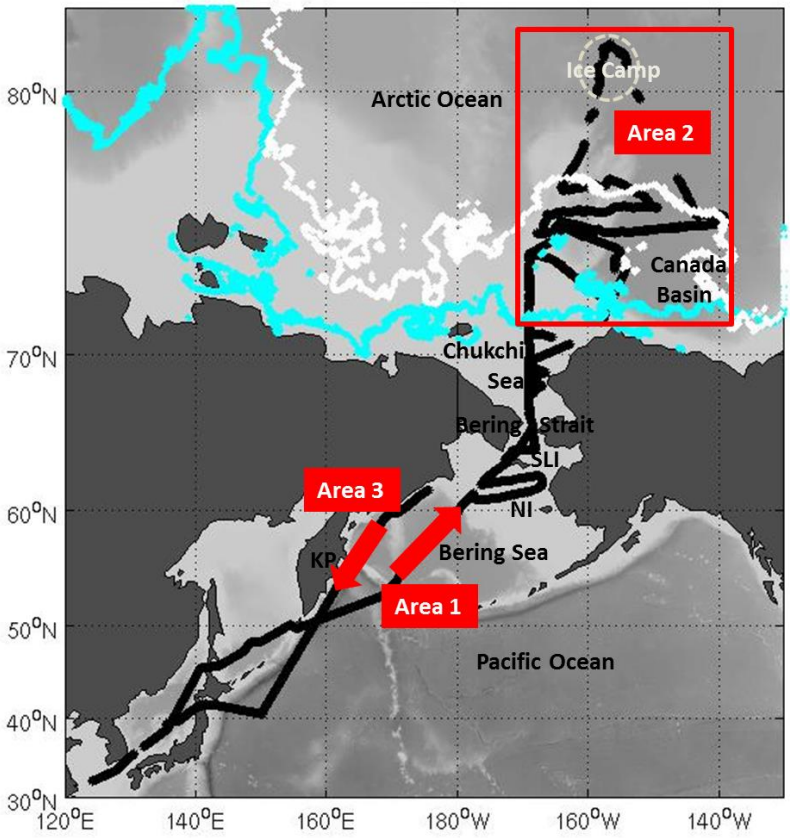


Figure 1: Map of the *MV Xue Long* transect (black dots) during the 6th Chinese Arctic Research Expedition in summer 2014. The areas of the three datasets are indicated in red, while the location of the ice camp is indicated in grey. Area1 on the outbound leg was ice-free, Area2 was ice-covered and includes the ice camp, and Area3 on the return leg was ice-free. Red arrows indicate the outbound and return legs of the transect. The blue and white dots show the sea-ice edge contours on the weeks of the 25th of July and the 3rd of September, respectively. Ice concentration data were obtained from the satellite AMSR2 database [*Beitsch et al.*, 2013] and sea-ice edge contours are defined for ice concentration between 25 and 40%. Saint Lawrence Island (SLI), Nunivak Island (NI), and Kamchatka Peninsula (KP) are also indicated for reference.

2.1. Cruise Set Up

Two pH instruments were running in parallel: a commercially available SeaFET Ocean pH sensor from Satlantic [*Martz et al.*, 2010] and a custom-made colorimetric bench-top analyser using *meta*-cresol purple indicator [*Rérolle et al.*, 2013].The two pH instruments were set up to run continuously on the underway seawater supply (intake at ~6 m depth) in the on-board wet lab from the 14th of July in the Japan Sea until the 16th of September in the Pacific Ocean east of the Kamchatka Peninsula. The underway water flow was irregular and when it was too low to run all the instruments, only the SeaFET sensor was left on and the colorimetric analyser was switched off. As a result, the colorimetric sensors had to be stopped for 30% of the length of the cruise (see Figure 3.d for gaps in dataset). Through the entire transect, 68,313 data point were initially collected with the SeaFET instrument and 2,451 with the colorimetric sensor. After data quality control (QC), a number of data points were removed from the data set due to the low quality of either surface seawater salinity (SSS), surface seawater temperature (SST), sensors measurement temperature or high gradient salinity data. The first three days (14-16/07/2014) of SeaFET measurements were also omitted from

the dataset because of the discrepancy between internal and external reference data (Figure 5.a). After QC, 11 % of the SeaFET pH data and 3 % of the colorimetric pH data were removed. Finally, 2,115 data points represented simultaneous (<1 min difference) measurements with the two instruments and are used here for sensor performance comparison. Cruise pH and hydrological data are shown in Figure 3. The pH is reported in the total proton scale throughout the document.

The SeaFET instrument is a potentiometric system based on ion-sensitive field-effect transistors (ISFET) [*Martz et al.*, 2010]. The instrument combines one Durafet pH sensor and two different reference electrodes referred to as internal reference and external reference by the manufacturer. The principle of the two reference electrodes will not be detailed here (see *Martz et al.* [2010] for further details) but the key aspect is that the external reference is sensitive to salinity and is more stable than the internal one, which is not sensitive to salinity within the typical seawater conditions (30-36) [Bresnahan et al., 2014]. The two SeaFET sensing pairs (Durafet+Internal reference and Durafet+External reference electrodes) will be referred as SeaFET_int and SeaFET_ext. The instrument was connected to the underway seawater supply using the flow-cell provided by the manufacturer. The accuracy of the SeaFET sensor is 0.02 according to manufacturer's specifications. The instrument was set up in a continuous data recording mode using the SeaFETCom® software to log the data directly on the computer. The data were recorded using the sensor's measurement temperature (T_SeaFET) probe and a value of 35 for salinity. The instrument was new (serial number: 106) and the calibration coefficients from the manufacturer were used (Appendix A). Measurements were made every 4 seconds.

The colorimetric pH measurement method is based on the use of an indicator dye (here, *m*-Cresol Purple, mCP) with an acidic (HI^-) and a basic (I^{2-}) form having different colour. The indicator is added to the seawater sample and the relative amount of the two forms of the dissociated indicator, deduced from absorption measurements, indicates the pH of the solution: $pH = pK_{ind} + \log_{10}((R-e_1)/(e_2-Re_3))$, with $R = A_1/A_2$,
$e_1 = \varepsilon_1(HI^-)/\varepsilon_2(HI^-)$, $e_2 = \varepsilon_1(I^{2-})/\varepsilon_2(HI^-)$, $e_3 = \varepsilon_2(I^{2-})/\varepsilon_2(HI^-)$, A_1, A_2= absorbance at wavelengths 1 and 2, the maxima of the indicator acid and basic forms, and $\varepsilon_i(I^{2-})$, $\varepsilon_i(HI^-)$ = extinction coefficients at wavelength *i* (*i* = 1, 2) of the basic (I^{2-}) and acid (HI^-) forms of the dye (see *Rérolle et al.* [2012] for detailed description of the method). The colorimetric analyser used during the cruise was adapted from the microfluidic system detailed in *Rérolle et al.* [2013]. The colorimetric measurement cell was immersed in a sampling chamber to reduce warming of the sample and possible gas exchange with the atmosphere. A tungsten light source was coupled to a blue LED in order to obtain suitable light spectra for the use of purified mCP indicator. The purified mCP indicator solution was prepared in 0.7 M NaCl, and the mCP calibration coefficients from *Liu et al.* (2011), valid for T = 5 – 35 °C and S = 20 – 40, were used to determine pH from absorbance measurements. The validity of extending the use of these coefficients to temperatures lower than 5 °C was evaluated by measuring the pH of a S = 35 Tris buffer in synthetic seawater and non-equimolal Tris/Tris-H^+ composition (0.02 mol kg^{-1} Tris + 0.040 mol kg^{-1} Tris-H^+) in the laboratory. This validation was performed with a pH set-up identical to the shipboard system at 0 and 25 °C. The pH of this buffer was characterized electrochemically with the Harned Cell to near its freezing point (Papadimitriou et al., manuscript submitted to *Marine Chemistry*). Measurements were made every 6-7 minutes.

Discrete seawater samples for DIC and TA were collected from the underway seawater supply (n = 25). The samples were collected using standard protocols [*A G Dickson et al.*, 2007] in 500 mL

borosilicate glass bottles and poisoned with a saturated mercuric chloride solution for storage. The samples were shipped back to the laboratory (LOCEAN, UPMC Paris) by cargo and were analysed for DIC and TA five months after the cruise using a custom-made potentiometric method based on *Edmond* [1970]. The accuracy of the analysis was estimated to be within 4 µmol kg^{-1} for DIC and TA based on analyses of certified reference material (Batch 139, Scripps Institution of Oceanography, University of California San Diego). The pH of the discrete underway seawater samples was calculated from the DIC and TA data using version 2.1 of CO2Sys.xls [*Pierrot et al.*, 2006] and the carbonic acid dissociation constants of (*i*) *Roy et al.* [1993], defined for low temperatures, and (*ii*) the re-fit of Mehrbach et al. (1973) data by *A Dickson and Millero* [1987] as recommended by *A G Dickson et al.* [2007]. The boric acid dissociation constant of *Dickson* [1990a], the bisulphate ion acidity constant of *Dickson* [1990b], and the boron-to-chlorinity ratio of *Lee et al.* [2010] were used in the computations.

Continuous SST and SSS data were obtained from the Sea-Bird Electronics SBE21 ThermoSalinoGraph (TSG) installed on the ship's underway supply. Because the two pH instruments were set up in a different laboratory from the TSG set-up, the temperature of the seawater increased between the TSG and the pH instruments. The measurement temperatures of the colorimetric and SeaFET® instruments were on average 1.2 ± 0.5 °C and 1.9 ± 0.6 °C higher than SST, respectively, with larger discrepancy at lower SST. All the pH data have therefore been corrected to in situ temperature using the equation from Millero et al. (2007), dpH/dT = –0.001582. Low salinity values were observed during the transect through the ice. During the ice camp phase of the study, discrete salinity measurements (S_{WTW}) were made with a hand-held conductivity probe (WTW ProfiLine Cond 197i, Germany) in under-ice water sampled from 5 m depth through a hole drilled in the ice.

2.2. Data analysis

2.2.1. Geographic grouping of pH measurements

The data were split in three sub-sets corresponding to measurements taken in the ice-free Area1 on the outbound leg of the cruise, the ice-covered Area2, and the ice-free Area3 on the return cruise leg (Figure 2). Area2 also comprises data collected in the low salinity (SSS < 30) and cold temperature (SST < 6 °C) open surface water off the edge of the ice-covered area. Definition of the areas covered by the three datasets was done using temperature and salinity criteria as detailed in Table 1 and shown in Figures 1 and 2. The SST was about 20 °C at the start and end of the transect in the Pacific Ocean and decreased to –0.9°C in the ice-covered Arctic Ocean (Figure 2.a). Following a similar pattern, the salinity ranged between 30 and 35 in the Pacific Ocean and an average 26 in the Arctic Ocean (Figure 2.b).

Table 1: Geographic and hydrological characteristics of the Areas defined to study the pH datasets from CHINARE 2014.

Dataset	Start date	End date	Start GPS location	End GPS location	SSS	SST
Area1 (outbound leg)	14/07/14 02h00	30/07/14 17h33	40.37N 136.81E	71.00N –167.69E	29.36-34.31	4.17 to 24.30
Area2 (ice-covered leg)	30/07 17h34	07/09 19h47	71.01N -167.71E	72.53N –168.98E	19.89-30.46	–0.87 to 6.12
Area3 (return leg)	07/09/14 19h48	16/09/14 05h38	72.53N -168.98E	42.36N 151.42E	30.10-33.32	2.80 to 19.48

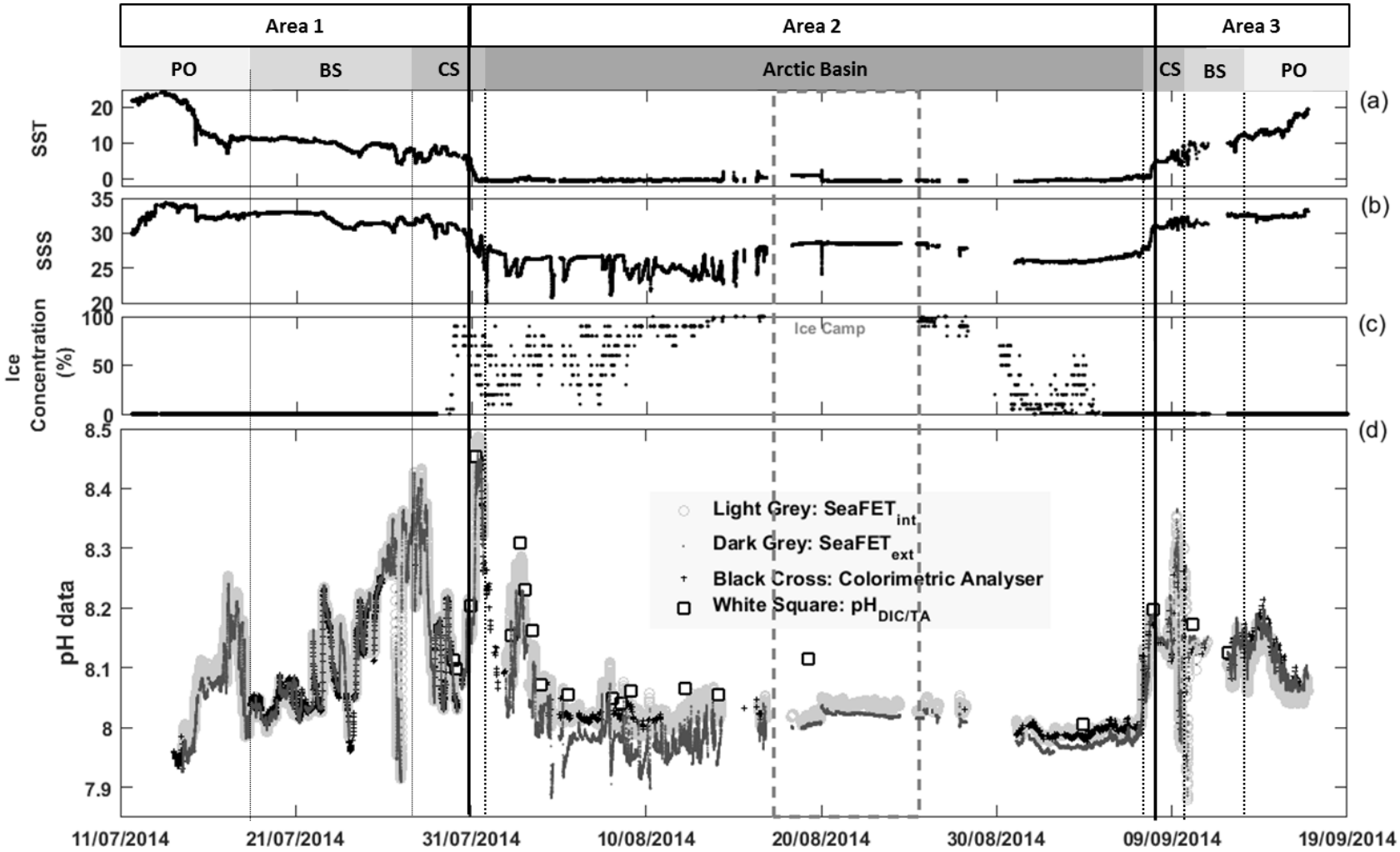


Figure 2: In situ seawater surface temperature (a), salinity (b), ice concentration (c), and pH (total proton scale) (d) during the CHINARE2014 research cruise. Light grey circles represent pH data from SeaFET_int, dark grey dots pH data from SeaFET_ext, black crosses indicate the colorimetric pH data, and open squares the pH calculated from discrete DIC and TA analyses using the carbonic acid constants from *Roy et al.* [1993]. Data corresponding to Area1 (ice-free outbound), Area2 (Arctic ice-covered), and Area3 (ice-free return leg) are delimited with the vertical solid lines and the ice camp with the vertical dotted lines. The Pacific Ocean (PO), Bering Strait (BS), Chukchi Sea (CS), and Arctic Basin areas are also indicated.

2.2.2. Measurement precision and uncertainty

The short term precision of the instruments was estimated at nine CTD stations where the ship was stationary for at least 1 hour. Given the various frequencies of the instruments, precision was estimated by both the standard deviation of the data obtained in 1 hour of measurement and the standard deviation of 10 consecutive measurements. Results were similar for both methods and only results from consecutive measurements are discussed.

To estimate the uncertainty in the salinity data, a comparison was made between underway salinity (S_{UW}) and surface CTD data collected with a Sea-Bird Electronics SBE-19 (S_{CTD}). The impact of this uncertainty on colorimetric and SeaFET_ext pH datasets was estimated by comparing pH colorimetric and SeaFET_ext obtained with S_{UW} and then pH colorimetric and SeaFET_ext obtained with $S_{UW}+S_{uncertainty}$.

The uncertainty of the colorimetric pH dataset is evaluated as the square root of the sum of squares of the estimated uncertainty of the method and uncertainties due to errors in the measurement temperature and salinity. The uncertainty of the method was estimated as the square root of the sum of squares of the discrepancy between the certified and measured pH of the Tris buffer and the uncertainty in the Tris buffer characterisation (0.004 pH unit; Papadimitriou et al., submitted to Marine Chemistry). The colorimetric measurement temperature uncertainty (0.5 °C) was estimated from the standard deviation amongst the three thermistors immersed in the colorimetric chamber.

The uncertainty in the comparison between calculated pH from DIC and TA and colorimetric pH was calculated as the square root of the sum of squares of the estimated uncertainties due to lack of synchronicity between the discrete and continuous pH sampling, discrete sample handling and storage, and computational errors due to uncertainty in the thermodynamic constants of the marine carbonate system and uncertainty in DIC, TA and colorimetric pH measurements. These calculations are detailed is Appendix C.

2.2.3. Quantitative comparison of pH datasets

The agreement between the colorimetric, SeaFET_int and SeaFET_ext data sets (Figure 3) was evaluated using Pearson's correlation coefficient (*r*), mean and standard deviation of residuals (MR and SR), and the root mean square error (RMSE) of the linear fit. Data sets are considered to compare well when $r > 0.9$, with MR and SR ≤ 0.01 pH unit. The correlation between each pair of pH dataset was computed using a multiple linear regression in Matlab (function *regress*) [*Chatterjee and Hadi, 1986*].

Factors that may give rise to data discrepancies between the sensors were investigated by Pearson's correlation between pH residuals and other measured environmental parameters (pH datasets, SSS, SST, T_SeaFET, T_colorimetric and date), as well as inter-correlation between these parameters. A stepwise regression (*stepwisefit*) resulted in keeping the parameters *pH datasets*, *SSS*, and *SST* to determine the factors influencing distribution variability of pH residuals along the cruise transect [*Draper and Smith, 1998*]: $pH\_residuals = x_{pH\_SeaFET} \cdot pH\_SeaFET + x_{SSS} \cdot SSS + x_{SST} \cdot SST$, with $x_{parameter}$ = regression coefficient for each parameter. The percent relative contribution from each parameter to the pH dataset discrepancy was calculated as follows: $RMSE \cdot x_{parameter}/sum(x_{parameters})$. All parameters were standardized (centred to 0 and variance scaled to 1) before performing the multi-linear regressions: $p_{standardized} = (p_i - p_{mean})/\sigma_p$, where p = parameter and σ_p = parameter standard deviation.

2.2.4. Final pH dataset

A final data set (pH_final) was obtained by combining the data from SeaFET_int sensor in Area1 and Area2 with the data from SeaFET_ext sensor in Area3. Data from the colorimetric sensor were used to fill gaps in the dataset, consisting mostly of 259 data points over ten hours on the 14th of July and over seven hours on the 17th of July.

3. Results and Discussion

Surface seawater temperature, salinity, pH and ice concentration datasets collected during CHINARE 2014 are presented Figure 2.

3.1. Measurement precision and uncertainty: results

Short term precision of the two SeaFET electrodes and the colorimetric analyser was 0.0005 and 0.0010, respectively (n = 10, see 2.2.2.Measurements precision and uncertainty).

3.1.1. Salinity observations: uncertainty in the iced-covered region and its contribution to pH uncertainty

During transit in the ice-covered Area2, S_{UW} was between 0.02 to 4.00 units lower than S_{CTD} ($S_{uncertainty}=S_{UW} - S_{CTD} = -1.8 \pm 1.3$, n = 24, Figure 3). No such discrepancy was observed with the temperature sensor (Appendix B). The conductivity sensor does not appear to have drifted during the cruise because S_{UW} compared well in areas visited both on the outbound and on the return legs of the cruise. Additionally, there was good agreement between $S_{UW} = 28.49 \pm 0.02$ (n = 5735) during the 8-day long stationary ice-camp phase and $S_{WTW} = 28.2 \pm 0.2$ (n = 3) measured with the hand-held salinometer in under-ice seawater. A small part of the salinity discrepancy during transit in Area2 may be due to sampling mismatch between the underway inlet and the CTD cast as observed in the outbound ice-free Area1, where $S_{UW} - S_{CTD} = -0.04 \pm 0.41$ (n = 59). However, the comparatively larger discrepancy between the two salinity datasets in Area2 indicates additional contributing factors, such as potentially problematical measurements with the underway salinometer due to ice fragments or air bubbles, or both, pumped in the underway supply during transit through the ice-covered region. Due to the unaccounted-for discrepancy between S_{UW} and S_{CTD} in the ice-covered Area2 stated above, the corresponding uncertainty in pH_SeaFET_ext and pH_colorimetric was estimated by comparing the pH data processed using S_{UW} and $S_{UW}+1.8$. In this comparison, pH_colorimetric($S_{UW}+1.8$) was found to be 0.002 ± 0.001 pH unit lower than pH_colorimetric(S_{UW}). In contrast, pH_SeaFET_ext($S_{UW}+1.8$) was found to be 0.027 ± 0.005 pH unit higher than pH_SeaFET_ext(S_{UW}). Because the pH_colorimetric data from Area2 are less susceptible to uncertainty in salinity, they are considered to be more reliable than the pH_SeaFET_ext data. The pH_SeaFET_int is not considered to be sensitive to salinity within the salinity range encountered in this study and, so, will not be affected by the salinity uncertainty during transit in this region of the Arctic Ocean.

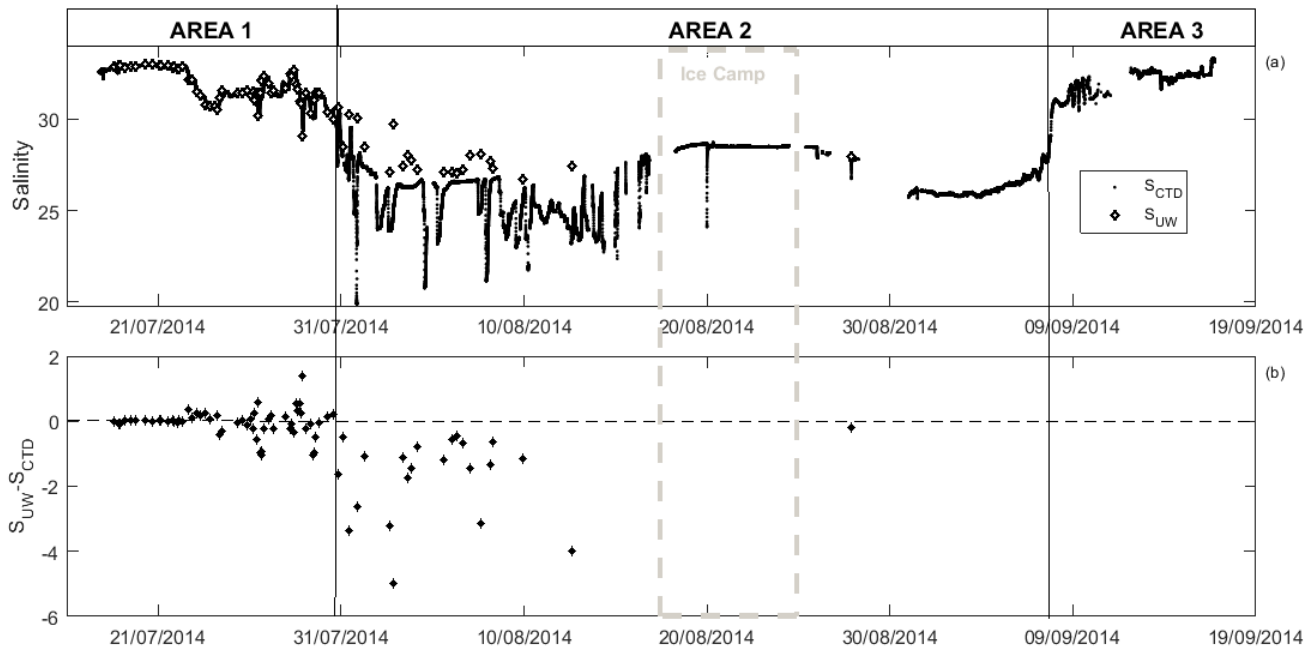


Figure 3: Comparison of underway salinity (S_{UW}) data with salinity data from surface CTD casts (S_{CTD}). a) S_{UW} is shown in black dots and S_{CTD} in white diamonds. b) Salinity residuals (S_{UW}-S_{CTD}). Data corresponding to Area1 (ice-free outbound), Area2 (Arctic, ice-covered), and Area3 (ice-free return leg) are delimited by the vertical solid lines and the ice camp by the vertical dotted lines.

3.1.2. Comparison between measured and calculated pH data

Similar trends were observed between the measured pH and the pH calculated from discrete DIC and TA data (Figure 2d). However, a discrepancy of 0.034 ± 0.036 pH unit was obtained between measured and calculated pH using carbonic acid constants from Roy et al. [1993] (0.045 ± 0.035 with constants from Mehrbach et al. [1973] refitted by A Dickson and Millero [1987]). This discrepancy is likely due to lack of synchronicity between the discrete and continuous pH sampling, discrete sample handling and storage, and computational errors due to uncertainty in the thermodynamic constants of the marine carbonate system. The overall uncertainty in the comparison between calculated pH from DIC and TA and colorimetric pH was estimated to be 0.041 pH unit (see Materials and methods section and Appendix C for calculation details). This is the same order of magnitude as the observed discrepancy between measured and calculated pH.

3.1.3. Colorimetric measurement uncertainty

The discrepancy of our measured buffer pH from the electrochemical value was between 0.001 pH unit at 25 °C and 0.005 pH unit at 0°C (see Materials and methods section). An error of 0.006 pH unit is obtained from the maximum discrepancy between the certified and measured Tris buffer pH (0.005 pH unit) and the uncertainty in the Tris buffer characterisation (0.004 pH unit; *Papadimitriou et al.,* submitted to *Marine Chemistry*). The error due to the salinity uncertainty is estimated to be within 0.002 pH unit (see section 3.1.1.), while the error due to the colorimetric measurement temperature uncertainty is 0.002 pH unit. The combined measurement uncertainty is thus computed to be 0.007 pH unit, and, based on this, the colorimetric dataset (pH_colorimetric) from the entire cruise is considered of good quality and will be used as reference data to study pH_SeaFET data quality.

3.2. Comparison of pH sensors

The results from the statistical comparison of the sensor data are presented in Table 2 and Figure 4. The coefficients from the multi-linear regression are detailed in Appendix D. The correlation between the data from both sensors was high ($r > 0.9$) but significant discrepancies with mean residuals (MR) higher than 0.010 pH unit were observed in some areas. The highest MR was found between pH_SeaFET_int and pH_SeaFET_ext ($MR_{All\_SeaFET} = 0.016 \pm 0.022$ pH unit, Table 2). The two sensors compared well in open ocean water (Area1) with residuals pH_SeaFET_int - pH_SeaFET_ext = 0.002 ± 0.005 but the comparison worsened with residuals about an order of magnitude higher in Area2 (positive residuals) and Area3 (negative residuals). Data from SeaFET_int compared well with the colorimetric data from open ocean water in Area1 and the ice-covered Area2, with residuals between 0.000 ± 0.007 and 0.002 ± 0.011 pH unit. In Area3, the SeaFET_int data differed from the colorimetric data, with similar discrepancies to those observed between SeaFET_int and SeaFET_ext. Finally, the data from SeaFET_ext compared well with the colorimetric data in the open water regions in Area1 and Area3 (absolute residuals ≤ 0.002 ± 0.016) whereas large discrepancies were observed in the ice-covered Area2 (residuals = –0.026 ± 0.016). Reasons for these discrepancies will be discussed in 3.4. SeaFET sensor performance.

Figure 4 shows that a large part of the residuals between the two SeaFET datasets and between SeaFET_ext and the colorimetric analyser is due to the salinity distribution in the entire transect and in Area2. In contrast, in Area3, temperature seems to be the predominant factor in explaining the residuals between the two SeaFET datasets and between SeaFFET_int and the colorimetric dataset, which in fact was probably to be due to a fault in the internal electrode (see 3.4.2 drift of the internal electrode).

Table 2: Comparison between the underway data from the colorimetric and SeaFET sensors, divided into three datasets (Area1, Area2 and Area3: for definitions, see section *Materials and methods*): *r* = Pearson's correlation coefficient, MR = mean standard deviation of residuals, and SR = standard deviation of residuals. Datasets where comparison is good (mean residuals ≤ 0.010) are highlighted in bold.

Sensors Y vs X	Dataset	Correlation	Residuals (X-Y)		NB Points
		r	MR	SR	
SeaFET_ext vs SeaFET_int	All	0.994	0.016	0.022	60787
	Area1	**0.997**	**0.002**	**0.005**	16877
	Area2	0.998	0.029	0.018	36333
	Area3	0.928	-0.017	0.004	7577
Colorimetricvs SeaFET_int	All	0.989	-0.001	0.010	2110
	Area1	**0.997**	**0.000**	**0.007**	1495
	Area2	**0.991**	**0.002**	**0.011**	459
	Area3	0.952	-0.017	0.020	156
Colorimetricvs SeaFET_ext	All	0.977	-0.007	0.014	2110
	Area1	**0.999**	**-0.002**	**0.006**	1495
	Area2	0.972	-0.026	0.016	459
	Area3	**0.997**	**-0.002**	**0.016**	156

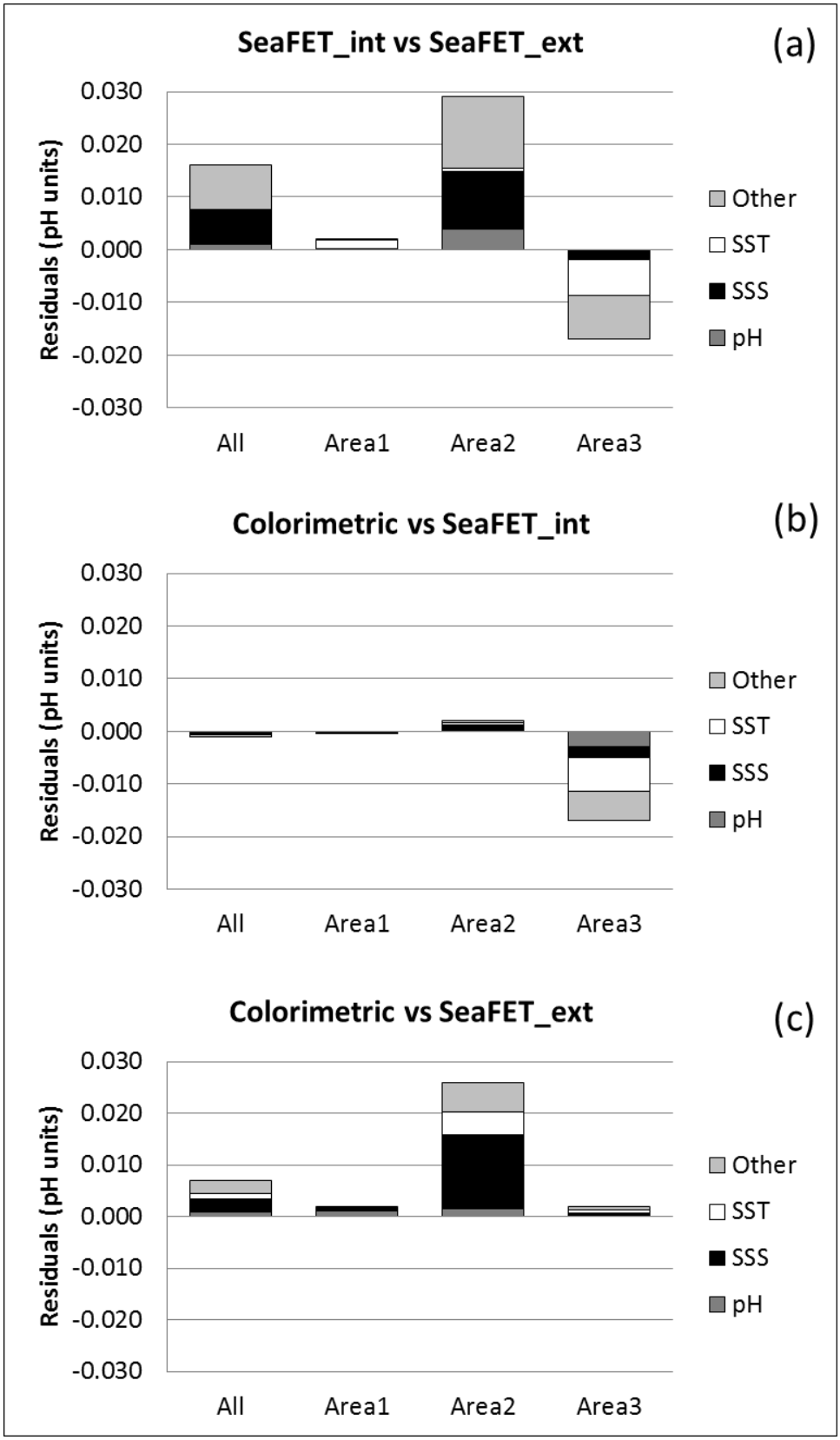


Figure 4: Results of the multi-linear regression analysis using pH_SeaFET, SSS, and SST to explain residuals between pH datasets. Underway data are divided into three datasets (Area1, Area2, and Area3; see section *Materials and methods*). (a) Residuals = pH_SeaFET_int – pH_SeaFET_ext. (b) Residuals = pH_colorimetric – pH_SeaFET_int. (c) Residuals = pH_colorimetric – pH_SeaFET_ext.

3.3. Colorimetric analyser: Performance at near-zero temperatures

The colorimetric method can be used at temperatures down to –1.8 °C without any further adaptation as long as a good mixing between the indicator and the seawater is ensured (solution viscosity will increase at low measurement temperature). However, the mCP has only been characterised above 5 °C by *Liu et al.* [2011] and care should be taken when using equations outside their definition range. In this case, the extinction coefficients and dissociation constants (pKa') have been shown to be close to linear with temperature [*DeGrandpre et al.*, 2014; *Liu et al.*, 2011] and extending equations to lower temperatures may lead to negligible errors. This assumption is supported by comparison with equations from *DeGrandpre et al.* [2014] defined for a temperature down to 0 °C showing a discrepancy lower than 0.010 in pKa' at 0 °C. During this investigation, the SSS remained within the range of the *Liu et al.* [2011] mCP characterisation. As an additional check, the performance of the custom-made colorimetric set-up used in this investigation was evaluated with a Tris buffer solution prior to the cruise with good agreement (see section 3.1.3.), which indicates that the Liu et al. (2011) indicator calibration coefficients are suitable for the minimum measurement temperature in this study (0.36 °C). Until relevant buffer data become available, extrapolation to below-zero temperatures of the existing oceanographic equations is not recommended because the indicator parameters behaviour close to the freezing point of seawater is unknown.

Although the colorimetric method can be easily extended *a priori* to operate at near-zero temperatures, deploying an in situ colorimetric pH instrument is not straightforward. First, in situ colorimetric systems typically use LEDs as light source with wavelengths known to shift at low temperature [Chhajed et al., 2005]. So, even if an LED were centred on the desired wavelength at 20 °C, it may be several nanometres away at 0 °C. It is then required to check the calibration not only of the pH indicator dye, but also the entire sensor at the temperature range of deployment. Second, a colorimetric analyser requires a pumping system and valves to control the indicator and sample flows. Power settings to activate the different components of an autonomous instrument are usually optimised in order to minimise the power requirement and may not be the same for temperatures typical of the temperate and tropical oceanic waters and for temperatures near or below zero as in Arctic waters. For example, the power requirements of each component, particularly pumps and valves, may vary with temperature. Battery output will also decrease at low temperature. Instrument settings must then be optimised depending on the deployment conditions. It is therefore highly recommended to check with the manufacturer that the sensor has been calibrated and set up for the intended deployment conditions.

3.4. SeaFET sensor performance

3.4.1. Conditioning of the electrode

The SeaFET instrument requires time for warming up and conditioning of the electrodes [*Bresnahan et al.*, 2014], which can vary with instrument and, particularly, deployment conditions. Our SeaFET instrument required three days to stabilize based on the difference between pH_SeaFET_int and pH_SeaFET_ext converging to zero after three days of deployment (Figure 6a). To reduce this instrument conditioning time requirement, it is recommended to power the sensor and to store it in natural seawater for about a week prior to deployment [*Bresnahan et al.*, 2014].

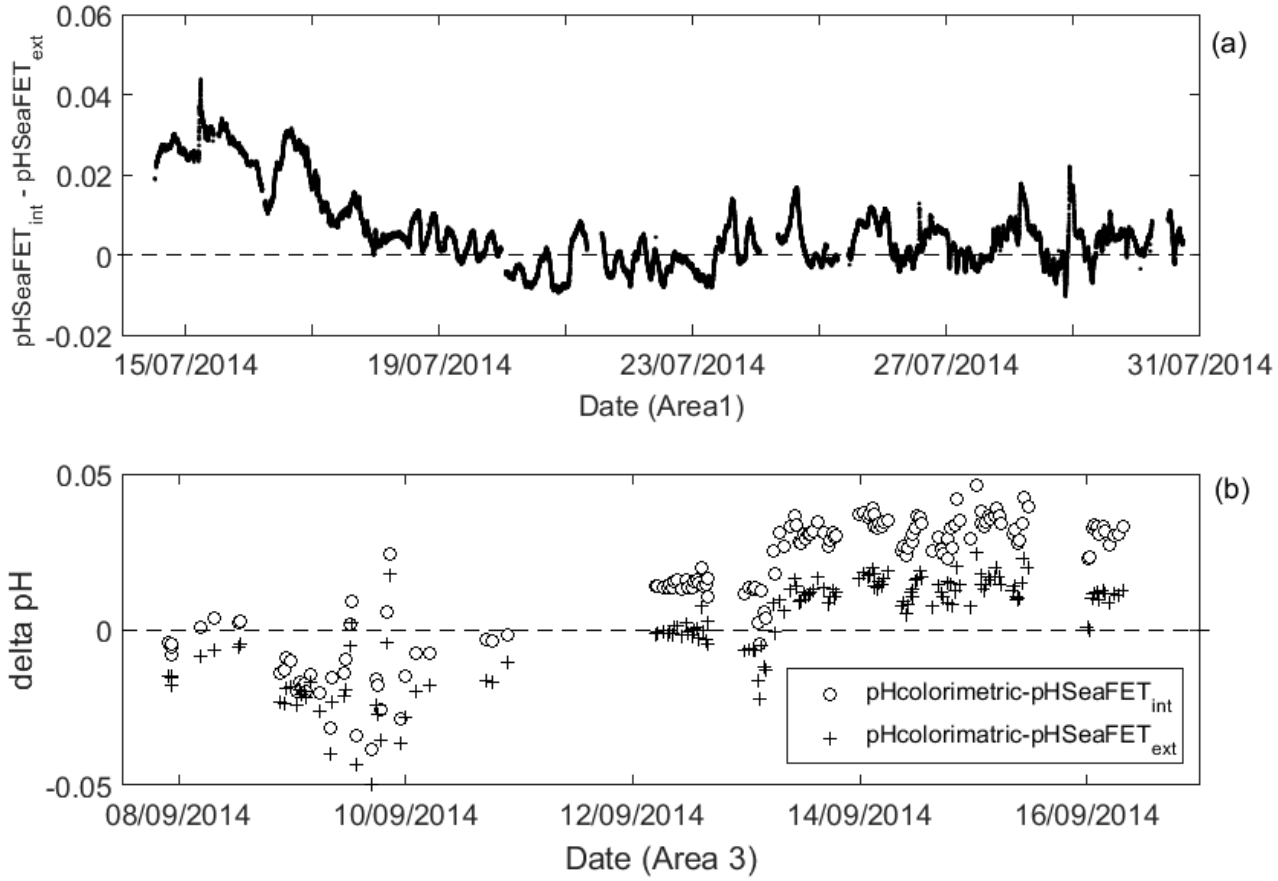


Figure 5: (a) Evidence for electrode conditioning time requirement: the difference between pH_SeaFET_int and pH_SeaFET_ext overtime in Area1; (b) Evidence for internal reference drift: the difference between pH_colorimetric and pH_SeaFET_int/ext overtime in Area3.

3.4.2. Drift of the internal electrode

Based on the residuals relative to both SeaFET_ext and colorimetric pH measurements of about 0.017 pH unit, the SeaFET_int did not perform well in Area3 (Figures 4b and 5b) despite similar measurement temperature and salinity conditions as in Area1 (Figure 2 and Table 1). The SeaFET_int data quality appears to worsen over time with the discrepancy between SeaFET_int and SeaFET_ext increasing from 0.010 to more than 0.020 pH unit in a week (from the 8th to the 16th of August, Figure 5b). A similar sensor drift has been observed by *Bresnahan et al.* [2014] and was attributed to deterioration of the internal reference electrode. The SeaFET used in our study was an early version of SeaFET instruments fitted with an internal electrode probably prone to early deterioration and drift (Van Dommelen, Satlantic, pers. comm.).

3.4.3. Performance in open ocean conditions and at low measurement temperatures

Good comparison with residuals lower than 0.010 pH unit (Table 2 and Figure 4) between either pH_SeaFET_int or pH_SeaFET_ext and pH_colorimetric in Area1 shows that SeaFET_int and SeaFET_ext sensors performed well in open ocean conditions (29 < S < 34 and 4 < T < 25 °C). Indeed, the pH discrepancy (within 0.007 pH unit) was within the colorimetric measurement uncertainty. The reliability of the SeaFET measurements was therefore better than 0.010 pH unit without the need to apply a correction using a discrete in situ sample after deployment as recommended in *Bresnahan et al.* [2014]. The SeaFET instrument was operated at measurement temperatures down to 1.0 °C with a mean T_SeaFET = 1.9 ± 0.8 °C in Area2. Contrary to salinity, low measurement

temperatures do not seem to affect the quality of the SeaFET sensor data (Area2 in Figure 4b and 4c). According to SeaFET User Manual, the instrument operating measurement temperature range is down to 0 °C, which would prevent deploying the sensor in situ in several areas of the Arctic where below-zero temperatures occur, although several users have reported such use without issue (IOCCP Sensors summer course participants) but this still requires further investigations.

3.4.4. Low salinity data quality and the external reference sensor

The estimated uncertainty in pH_SeaFET_ext in Area2 due to salinity uncertainty (0.027 pH unit, see section 3.1.1.) is of the same order of magnitude as the mean discrepancy observed between pH_SeaFET_ext and the two other pH datasets (Table 2). It is therefore likely that the poor quality of the salinity data in Area2 mostly explains the poor comparison between pH_SeaFET_ext and the other two pH datasets of this study. Takeshita et al. (2014) showed that the external reference electrode (Cl-ISE) had a near-Nernstian behaviour to the chloride ion activity between salinity 20 and 35. It is therefore expected that the SeaFET_ext sensor should perform well with the range of salinity met in the ice-covered, low salinity regions of the Arctic ocean provided that the Cl^- to Br^- ratio is similar to that in the open ocean areas. However, during the ice-camp phase of this study, when S_{UW} was reliable and stable around 28.5 for 8 days, a mean discrepancy of 0.015 ±0.004 pH unit (n = 5677) was still observed between the pH_SeaFET_int and pH_SeaFET_ext data sets. We do not anticipate sea ice formation and decay to alter the chloride ion composition in Arctic oceanic waters but, until further tests are conducted along these lines, the reason for this discrepancy remains unidentified.

3.5. Choice of sensor

The comparison with the bench-top colorimetric analyser showed that high quality pH data could be obtained by the ISFET with the internal reference electrode whereas the external reference electrode did not perform well in the ice-covered area. The quality of the data from both internal and external reference electrodes did not appear to be affected by the low measurement temperature to 1 °C in the Arctic region. The SeaFET instrument can therefore be used to measure pH in the Arctic but it is advisable to use the internal electrode data in ice-melt low salinity conditions. Long term stability of the instrument in such environment also requires further investigation.

The choice of instrument will ultimately depend on the user constraints and deployment conditions. The colorimetric method presents the key advantage that the instrument calibration can be tested in the laboratory without need for an in situ re-calibration, but regular checks with discrete samples or Tris buffer measurement are always advisable during deployment. Also, the colorimetric measurements are reproducible in the long term [Byrne et al., 2010]. Although difficult and expensive to obtain, the purified mCP indicator dye has been carefully characterized in the open ocean temperature and salinity range and can be reasonably extended to 0 °C. However, colorimetric instruments require valves and pumps which can be problematic and make these instruments often less reliable at low temperature. On the other hand, the SeaFET instrument has been shown to provide high quality measurements even at low temperature and the simpler instrumentation make it currently more amenable to in situ deployments at low temperature. However, when using the SeaFET instrument, it is strongly recommended to use regularly an independent procedure, such as discrete colorimetric pH data, or calculated pH from DIC/TA data, to re-calibrate SeaFET at in situ conditions and monitor sensor drift (see [*Bresnahan et al.*, 2014] for

detailed in situ calibration procedure). This is particularly critical for long term deployments. Finally, independent of the choice of sensor, it is recommended to check with the manufacturer that the instrument is calibrated and set up for the expected deployment conditions.

3.6. A new Arctic pH data set

Figure 6 presents the geographic distribution of surface pH_final along the outbound and return legs of our transect in the investigated regions of the North Pacific and Arctic oceans. The overall uncertainty of this dataset is estimated to be within 0.01 pH unit. This dataset can be used for marine carbonate chemistry studies and long-term ocean acidification monitoring.

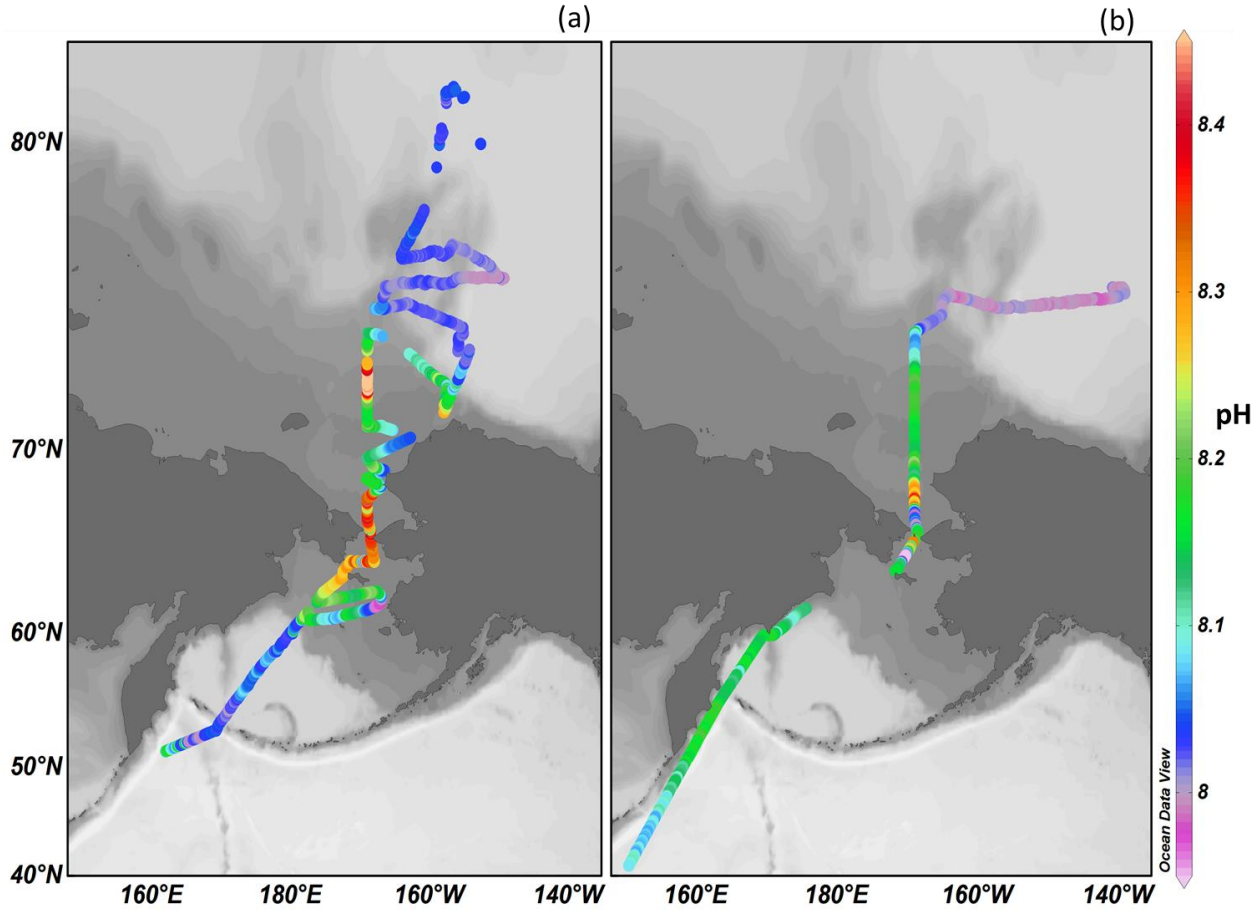


Figure 6: Surface pH data collected during CHINARE 2014. (a) Surface pH data collected from the 17th of July until the 28th of August. (b) Surface pH data collected from the 31st of August until the 16th of September.

Our finalized pH data set shows large pH variations in the explored Arctic regions across the shelves in ice-free and ice-covered surface seawater. Hot spots with a pH increase of up to 0.40 pH unit relative to the other regions of the transect were seen in the Bering Strait and the Chukchi Sea (Figure 6). In contrast, the pH remained relatively stable in under-ice seawater well into the sea ice zone, within 0.02 pH unit from a mean value of 8.02, which is comparable to the reported pH = 8.0 in the ice-covered Fram Strait by *Tynan et al.,* [*2016*]. Also, in a similar pattern to that reported in the Fram Strait (pH maximum = 8.5; *Tynan et al.,* [*2016*]), the highest pH values observed here (maximum = 8.49) were measured at the ice edge on the 31st of July at 73.00 °N and 168.97 °W, where, as evidenced by the brown aggregates observed in the water, we probably encountered a large diatom bloom. On the return leg of the transect, when we transited the same area, the pH was

8.18, which is 0.3 pH unit lower than a month before during the putative diatom bloom. This pH change is in the same order of magnitude as the estimated 0.25 pH change in the course of an ice-edge bloom in the Fram Strait by Tynan et al. [2016]. High pH values (up to 8.41) were also measured in the Bering Strait on the way north. The lowest pH values (7.93) was measured in the Bering Strait north of Saint Lawrence Island (SLI, Figure 1) at 64.36 °N and 169.45 °W where colder (-2 °C) and more saline waters (+1 practical salinity unit compared to surrounding waters) were locally encountered probably due to upwelled water as evidenced by the well-mixed water column. Low pH values (down to 7.96 pH units) were also measured close to the Alaskan coast at 61.70 °N and 167.65 °W north of Nunivak Island (NI, Figure 1).This low pH area maybe be due to the plume of the Yukon river water discharge as evidenced by fresher water (S = 30.4) compared to surrounding water (S > 30.8). Lower aragonite saturation state has also been observed in the area in 2012 [*Mathis et al.*, 2015], which would imply low pH values for these water masses.

4. Conclusion

A SeaFET pH sensor and a bench-top colorimetric pH analyser comparison was performed in surface Arctic and North Pacific waters in boreal summer 2014. The comparison allowed testing of the increasingly common SeaFET pH sensor for use in Arctic conditions, where the highest and fastest anthropogenic acidification is expected. The two instruments were used successfully to measure low temperature seawater pH in the Arctic Ocean, provided that the SeaFET internal reference dataset was used in the low salinity ice-covered area instead of the SeaFET external reference dataset. The SeaFET instrument presents the advantage of being simple to use, while the lack of mechanical components make it more reliable than current colorimetric instruments in low temperature conditions. However, for long term deployment without access to the instrument for recalibration, a colorimetric analyser would be favoured. Further development is still required for pH measurements in below-zero temperature hypo-saline and hyper-saline polar waters. Finally, a unique high resolution (15 measurement $minute^{-1}$) and high quality surface water pH dataset (uncertainty ≤ 0.010 and precision ≤ 0.001 pH unit) was obtained using the two instruments in different oceanic water masses in open water in the North Pacific, Bering Sea, and Chukchi Sea, as well as in the ice-covered region of the Arctic Ocean. Large pH variations were observed on the Arctic shelves, with pH (total proton scale) ranging between 7.98 and 8.49 whereas the under-ice seawater pH was relatively stable at 8.02 ± 0.02. The large spatial and seasonal pH variability observed here is not yet described by climatological models in the Arctic Oceans. As a result of this and the rigorous data quality assessment, we propose this novel and reliably measured pH data for use as reference to detect the on-going acidification rate in the Arctic. This highlights the pressing need for more extensive measurements in space and time in order to understand the dynamics of the carbonate system in the Arctic Ocean in the face of ongoing atmospheric CO_2 changes.

Acknowledgements

This work is supported by funding from the ICE-ARC programme from the European Union 7th Framework Programme, grant number 603887 (2014-2017) and Chinese Polar Project No. Chinare 03-04 (2012-2016). All the CHINARE 2014 meta-data were issued by the Data-sharing Platform of Polar Science (http://www.chinare.org.cn) maintained by Polar Research Institute of China (PRIC) and Chinese National Arctic & Antarctic Data Center (CN-NADC). We would like to extend a special thanks to the captain and crew of the *MV XueLong* for their support during the cruise. We would also like to thank Lei Ruibo for supplying ice concentration data. The cruise was financially supported

and organized by the Polar Research Institute of China (PRIC) and the Second Institute of Oceanography, SOA, Hangzhou City, China. Participation to the cruise was supported by the exchange program Cai Yuanpei (project Campus France, ICAR 30481PM). The research leading to the development of the colorimetric pH sensors has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreements n° 237868 (SENSEnet) and n° 614141 (SenseOCEAN), and from NERC UK grant NE/J011096/1.

References

Assmann, S., C. Frank, and A. Koertzinger (2011), Spectrophotometric high-precision seawater pH determination for use in underway measuring systems, *Ocean Sci.*, *7*(5), 597-607.

Bakker, D., S. Hankin, A. Olsen, B. Pfeil, K. Smith, S. Alin, C. Cosca, B. Hales, S. Harasawa, and A. Kozyr (2014), An update to the Surface Ocean CO2 Atlas (SOCAT version 2), *Earth*.

Bates, N., and J. Mathis (2009), The Arctic Ocean marine carbon cycle: evaluation of air-sea CO 2 exchanges, ocean acidification impacts and potential feedbacks, *Biogeosciences*, *6*(11), 2433-2459.

Beitsch, A., L. Kaleschke, and S. Kern (2013), AMSR2 ASI 3.125 km Sea Ice Concentration Data, V0.1, edited by U. o. H. Institute of Oceanography, Germany, digital media (ftp-projects.zmaw.de/seaice/).

Bresnahan Jr, P. J., T. R. Martz, Y. Takeshita, K. S. Johnson, and M. LaShomb (2014), Best practices for autonomous measurement of seawater pH with the Honeywell Durafet, *Methods in Oceanography*, *9*(0), 44-60.

Byrne, R. H., S. Mecking, R.A. Feely, and X. Liu (2010), Direct observations of basin-wide acidification of the North Pacific Ocean, Geophys.l Res. Let., 37.

Carter, B., J. Radich, H. Doyle, and A. Dickson (2013), An automated system for spectrophotometric seawater pH measurements, *Limnol. Oceanogr. Methods*, *11*(1), 16-27.

Chatterjee, S., and A. S. Hadi (1986), Influential Observations, High Leverage Points, and Outliers in Linear Regression, *Statistical Science*, 1, 379–416.

Chhajed, S., Y. Xi, Y.-L. Li, Th. Gessmann and E. F. Schubert (2005), Influence of junction temperature on chromaticity and color-rendering properties of trichromatic white-light sources based on light-emitting diodes, *Journal of Applied Physics*, 97, 054506 .

DeGrandpre, M. D., R. S. Spaulding, J. O. Newton, E. J. Jaqueth, S. E. Hamblock, A. A. Umansky, and K. E. Harris (2014), Considerations for the measurement of spectrophotometric pH for ocean acidification and other studies, *Limnol. Oceanogr. Methods*, *12*(12), 830-839.

Dickson, A., and F. Millero (1987), A comparison of the equilibrium constants for the dissociation of carbonic acid in seawater media, *Deep Sea Res. (A Oceanogr. Res. Pap)*, *34*(10), 1733-1743.

Dickson, A. G. (1990a), Thermodynamics of the dissociation of boric acid in synthetic seawater from 273.15 to 318.15 K, *Deep Sea Res. (A Oceanogr. Res. Pap)*, *37*(5), 755-766.

Dickson, A. G. (1990b), Standard potential of the reaction: AgCl(s) + $12H_2$(g) = Ag(s) + HCl(aq), and and the standard acidity constant of the ion HSO_4^- in synthetic sea water from 273.15 to 318.15 K, *The Journal of Chemical Thermodynamics*, *22*(2), 113-127.

Dickson, A. G., C. L. Sabine, and J. R. Christian (2007), Guide to best practices for ocean CO_2 measurements, *PICES special publication*, *3*.

Draper, N. R., and H. Smith (1998), Applied Regression Analysis, Hoboken, NJ: Wiley-Interscience, 307–312.

Edmond, J. M. (1970), High precision determination of titration alkalinity and total carbon dioxide content of sea water by potentiometric titration, paper presented at Deep Sea Research and Oceanographic Abstracts, Elsevier.

Ericson, Y., A. Ulfsbo, S. van Heuven, G. Kattner, and L. G. Anderson (2014), Increasing carbon inventory of the intermediate layers of the Arctic Ocean, *Journal of Geophysical Research: Oceans*, *119*(4), 2312-2326.

Hofmann, G. E., et al. (2011), High-Frequency Dynamics of Ocean pH: A Multi-Ecosystem Comparison, *Plos One*, *6*(12).

IPCC(2013), Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the IntergovernmentalPanel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley(eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp.

Lee, K., T.-W. Kim, R. H. Byrne, F. J. Millero, R. A. Feely, and Y.-M. Liu (2010), The universal ratio of boron to chlorinity for the North Pacific and North Atlantic oceans, *Geochim. Cosmochim. Acta*, *74*(6), 1801-1811.

Liu, X., M. C. Patsavas, and R. H. Byrne (2011), Purification and Characterization of meta-Cresol Purple for Spectrophotometric Seawater pH Measurements, *Environ. Sci. Technol.*, 4862–4868.

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.

Mehrbach, C., C. H. Culberson, J. E. Hawley, and R. M. Pytkowicz (1973), Measurement of the apparent dissociation constants of carbonic acid in seawater at atmospheric pressure, *Limnol. Oceanogr.*, 897-907.

Miller, L. A., Macdonald, R. W., McLaughlin F., Mucci, A., Yamamoto-Kawai, M., Giesbrecht, K. E., Williams, W. L. (2014). Changes in the marine carbonate system of the western Arctic: patterns in a rescued dataset. Polar Research 33, 20577, doi:10.3402/polar.v343.20577.

Millero, F. J., D. Pierrot, K. Lee, R. Wanninkhof, R. Feely, C. L. Sabine, R. M. Key, and T. Takahashi (2002), Dissociation constants for carbonic acid determined from field measurements, *Deep Sea Res. (I Oceanogr. Res. Pap.)*, *49*(10), 1705-1723.

Newton, J.A., R. A.Feely,E. B. Jewett, P.Williamson and J.Mathis (2015), Global Ocean Acidification Observing Network: Requirements and Governance Plan.2nd Edition, GOA-ON.

Papadimitriou, S., S. Loucaides, V. Rérolle, E. Achterberg, A. G. Dickson, M. Mowlem, H. Kennedy (submitted), The measurement of pH in saline and hypersaline media at sub-zero temperatures: Characterization of Tris buffers, *Mar. Chem.*

Pierrot, D., E. Lewis, and D. W. R. Wallace (2006), MS Excel Program Developed for CO_2 System Calculations. ORNL/CDIAC-105a. Carbon Dioxide Information Analysis Center, edited by U. S. D. o. E. Oak Ridge National Laboratory, Oak Ridge, Tennessee.

Rérolle, V. M. C., C. F. A. Floquet, M. C. Mowlem, D. P. Connelly, E. P. Achterberg, and R. R. G. J. Bellerby (2012), Seawater-pH measurements for ocean-acidification observations, *TrAC, Trends Anal. Chem.*, *40*(0), 146-157.

Rérolle, V. M. C., C. F. A. Floquet, A. J. K. Harris, M. C. Mowlem, R. R. G. J. Bellerby, and E. P. Achterberg (2013), Development of a colorimetric microfluidic pH sensor for autonomous seawater measurements, *Anal. Chim. Acta*, *786*(0), 124-131.

Roy, R. N., L. N. Roy, K. M. Vogel, C. Porter-Moore, T. Pearson, C. E. Good, F. J. Millero, and D. M. Campbell (1993), The dissociation constants of carbonic acid in seawater at salinities 5 to 45 and temperatures 0 to 45 C, *Mar. Chem.*, *44*(2), 249-267.

Sabine, C., R. M. Key, A. Kozyr , R. A. Feely, R. Wanninkhof, F. J. Millero, TH. Peng, J. L. Bullister, and K. Lee (2005), Global ocean data analysis project (GLODAP): results and data, ORNL/CDIAC-145, NDP-083. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee, 110 pp.

Seiter, J., and M. DeGrandpre (2001), Redundant chemical sensors for calibration-impossible applications, *Talanta*, *54*(1), 99-106.

Shitashima, K., M. Kyo, Y. Koike, and H. Henmi (2002), Development of in situ pH sensor using ISFET, IEEE.

Takeshita, Y., T. Martz, K. S. Johnson and A. G. Dickson (2014), Characteriszation of an Ion sensitive Field Effect Transistor and Choride Ion Selective Electrodes for pH Measurements in Seawater, *Analytical Chemistry*, 86, 11189-11195.

Takeshita, Y., C. Frieder, T. Martz, J. Ballard, R. Feely, S. Kram, S. Nam, M. Navarro, N. Price, and J. Smith (2015), Including high frequency variability in coastal ocean acidification projections, *Biogeosci. Disc.*, *12*(9).

637 Tanhua, T., E. P. Jones, E. Jeansson, S. Jutterström, W. M. Smethie, D. W. R. Wallace, and L. G.
638 Anderson (2009), Ventilation of the Arctic Ocean: Mean ages and inventories of anthropogenic CO2
639 and CFC-11, *Journal of Geophysical Research: Oceans*, *114*(C1), C01002.
640 Yao, W., X. Liu, and R. H. Byrne (2007), Impurities in indicators used for spectrophotometric
641 seawater pH measurements: Assessment and remedies, *Mar. Chem.*, *107*(2), 167-172.
642 Tynan, E., J. S. Clarke, M. P. Humphreys, M. Ribas-Ribas, M. Esposito, V. M. C. **Rerolle, S.** Thorpe, T.
643 Tyrrell, E. P. Achterberg (2016), Physical and biogeochemical controls on the variability in surface pH
644 and calcium carbonate saturation states in the Atlantic sectors of the Arctic and Southern Oceans,
645 *Deep-Sea Res. Part II*.

646

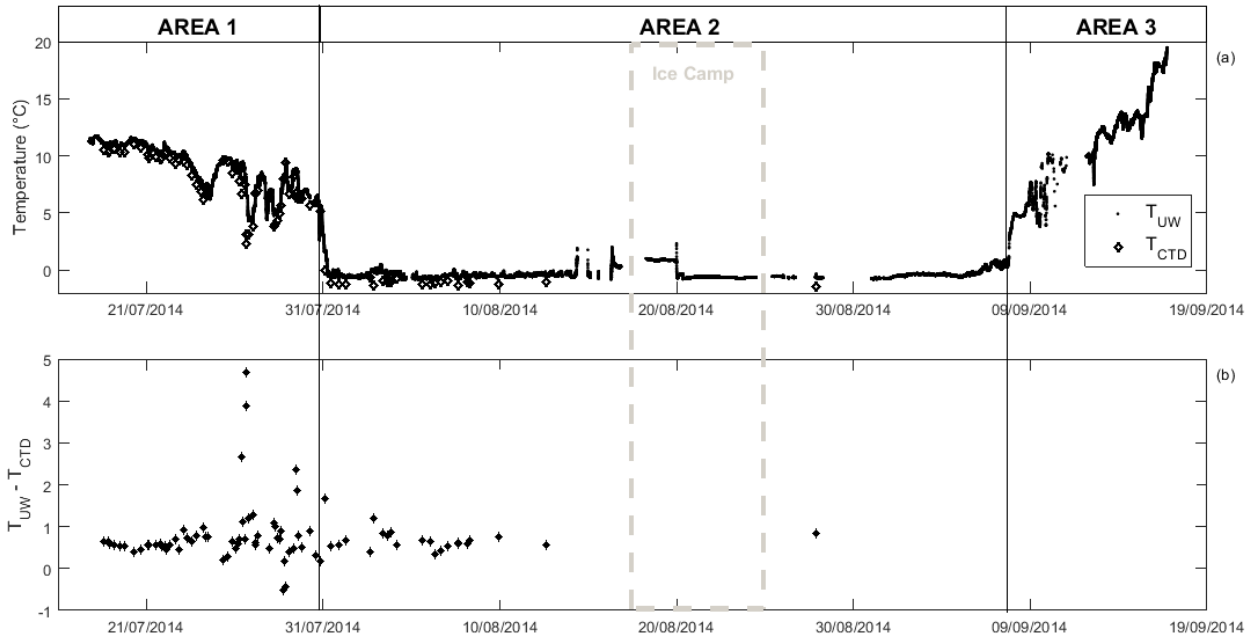
Appendix

Appendix A: SeaFET instrument calibration coefficients from the manufacturer

FILE-VERSION	3.0.0
APP-VERSION	SeaFETCom 1.1.2_149
INSTR-PACKAGE	SFET_0106_001.xml|SeaFET0106|1.0|2014-01-08-0400|22489
TIME-STAMP	2014-07-14T21:33:03+0200
FW-VERSION	SeaFET v3.4.1
SERIAL-NUMBER	106
CAL_PHINT_OFFSET_COEFF	-6.13E-02
CAL_PHINT_SLOPE_COEFF	-1.10E-03
CAL_PHEXT_OFFSET_COEFF	-1.08E+00
CAL_PHEXT_SLOPE_COEFF	-1.05E-03

Appendix B: Comparison of underway temperature (T_{UW}) data with temperature data from surface CTD casts (T_{CTD}).

a) T_{UW} is shown in closed circles and T_{CTD} in open diamonds. b) Temperature residuals (T_{UW}-T_{CTD}). Data corresponding to Area1 (ice-free outbound), Area2 (Arctic ice-covered), and Area3 (ice-free return) are delimited with the vertical solid lines and the ice camp with the vertical dotted lines.



Appendix C: Estimation of pH uncertainty in comparison between measured and calculated pH

1. DIC/TA uncertainty:
    - DIC/TA measurement accuracy: 4 µmol kg^{-1}
    - Error due to sample handling and storage (average of standard deviation between replicates values): 3 µmol kg^{-1}
        - Uncertainty in DIC/TA: $\sqrt{4^2 + 3^2}$ = 5 µmol kg^{-1}

    pH uncertainty due to uncertainty in DIC/TA = <u>0.021 pH unit</u>. pH uncertainty was estimated by computing pH with DIC ± 5 µmol kg^{-1} and/or TA ± 5 µmol kg^{-1} and calculating the standard deviation of the pH values obtained.
2. **pH uncertainty due to sample and sensor data mismatch = <u>0.001 pH unit</u>**. Estimated from standard deviation of five minutes of pH measurement.
3. **pH colorimetric method uncertainty = <u>0.007 pH unit</u>**.(see paragraph 3.4.2 for estimation details).
4. Temperature uncertainty:
    - Uncertainty in temperature= 0.04 °C.
    - Seawater pH sensitivity to temperature = 0.015 pH $degC^{-1}$.

    pH uncertainty due to temperature uncertainty = <u>0.001 pH unit</u>.
5. **pH uncertainty due to uncertainty in thermodynamics constants = <u>0.035 pH unit</u>.** Uncertainty estimated from standard deviation of pH values obtained using various combinations of thermodynamics constants (7 combinations).

➔ **Overall pH uncertainty in calculated and colorimetric pH datasets comparison =** $\sqrt{0.021^2 + 2\times0.001^2 + 0.007^2 + 0.035^2}$ **= <u>0.041 pH unit</u>.**

Appendix D: Results of the multi-linear regressions between pH dataset residuals and pH_SeaFET, SSS and SST (all the parameters are standardised).

RMSE is the Root Mean Squared Error.x_pH, x_SSS and x_SST are the coefficients resulting from the correlation:pH_residuals=x_pH·pH_SeaFET+x_SSS·SSS+x_SST·SST. The percentage columns indicate how much of the pH dataset residuals are explained by each parameter.

Sensors	Dataset	RMSE	x_pH	x_SSS	x_SST	fstat	% pH	% SSS	% SST
SeaFEText vs SeaFETint	All	0,47	0,14	-0,97	0,04	70801	6	40	2
	Area1	0,94	0,00	-0,05	-0,30	1060	0	13	78
	Area2	0,53	0,30	-0,84	-0,05	30517	14	38	2
	Area3	0,51	-0,02	-0,20	-0,74	7215	1	11	39
Colorimetric vs SeaFETint	All	0,66	0,30	1,21	-0,47	923	10	40	16
	Area1	0,90	0,59	-0,24	0,21	121	51	21	18
	Area2	0,78	0,00	0,89	0,18	146	0	64	13
	Area3	0,67	0,22	-0,17	-0,49	63	17	13	38
Colorimetric vs SeaFEText	All	0,64	0,37	1,15	-0,47	1000	12	37	15
	Area1	0,88	0,65	-0,23	0,21	141	53	19	17
	Area2	0,78	0,08	0,75	0,23	101	6	55	17
	Area3	0,68	0,20	-0,18	-0,51	62	16	14	39