

1 Intercomparison of carbonate chemistry measurements on 2 a cruise in northwestern European shelf seas

3
4 **M. Ribas-Ribas¹, V. M. C. Rérolle^{1,2}, D. C. E. Bakker³, V. Kitidis⁴, G. A. Lee³, I.
5 Brown⁴, E. P. Achterberg^{1,5} and T. Tyrrell¹**

6 [1] {Ocean and Earth Science, National Oceanography Centre Southampton, University of
7 Southampton, Southampton, United Kingdom}

8 [2] {National Oceanography Centre Southampton, Southampton, United Kingdom}

9 [3] {Centre for Ocean and Atmospheric Sciences, School of Environmental Sciences,
10 University of East Anglia, Norwich Research Park, Norwich, United Kingdom}

11 [4] {Plymouth Marine Laboratory, Prospect Place, Plymouth, United Kingdom}

12 [5] {GEOMAR, Helmholtz Centre for Ocean Research, Kiel, Germany}

13 Correspondence to: M. Ribas-Ribas (M.Ribas-Ribas@soton.ac.uk)

14 15 **Abstract**

16 Four carbonate system variables were measured in surface waters during a cruise aimed at
17 investigating ocean acidification impacts traversing northwestern European shelf seas in the
18 summer of 2011. High resolution surface water data were collected for partial pressure of
19 carbon dioxide ($p\text{CO}_2$; using two independent instruments) and pH using the total pH scale
20 (pH_T), in addition to discrete measurements of total alkalinity and dissolved inorganic carbon.
21 We thus overdetermined the carbonate system (four measured variables, two degrees of
22 freedom) which allowed us to evaluate the level of agreement between the variables on a
23 cruise whose main aim was not intercomparison and thus where conditions were more
24 representative of normal working conditions. Calculations of carbonate system variables from
25 other measurements generally compared well with direct observations of the same variables
26 (Pearson's correlation coefficient always ≥ 0.94 ; mean residuals were similar to the respective
27 accuracies of the measurements). We therefore conclude that four of the independent datasets
28 of carbonate chemistry variables were of high quality. A diurnal cycle with maximum
29 amplitude of $41 \mu\text{atm}$ was observed in the difference between the $p\text{CO}_2$ values obtained by

1 the two independent analytical $p\text{CO}_2$ systems, and this was partly attributed to irregular
2 seawater flows to the equilibrators and partly to biological activity inside the seawater supply
3 and one of the equilibrators. We discuss how these issues can be addressed to improve
4 carbonate chemistry data quality on future research cruises.

5

6 **1 Introduction**

7 Accurate determination of the inorganic carbon system is a key requirement for ocean
8 acidification studies, as it forms the basis for assessments of biological and biogeochemical
9 responses to changes in ocean carbonate chemistry as a result of rising atmospheric CO_2
10 concentrations. It is also essential for the determination of the air-sea fluxes of CO_2 ,
11 calculation of carbon budgets and estimation of anthropogenic CO_2 concentrations in different
12 water masses. When the carbonate system is overdetermined, it is possible to test if the
13 different variables are consistent with one another. This requires that more than two of the
14 measurable variables (total dissolved inorganic carbon (C_T), total alkalinity (A_T), pH_T , and
15 partial pressure or fugacity of CO_2 ($p\text{CO}_2, f\text{CO}_2$)) are determined.

16 Several at-sea intercomparison studies have taken place in recent years that compared
17 different $p\text{CO}_2$ instruments. Körtzinger et al. (1996) carried out what may have been the first
18 intercomparison study in coastal waters between two similarly designed underway $p\text{CO}_2$
19 systems. They found a remarkable agreement between the two simultaneously measured
20 $p\text{CO}_2$ datasets even though the spatial variability in surface $p\text{CO}_2$ in the North Sea was high.
21 The average difference was $0.2 \mu\text{atm}$ (standard deviation = $1.2 \mu\text{atm}$), indicating no
22 systematic difference. The difference tended to be highest during the most pronounced $p\text{CO}_2$
23 gradients. Körtzinger et al. (2000) reported on a comprehensive shipboard, international
24 intercomparison exercise which used one discrete and seven underway systems for the
25 measurement of $f\text{CO}_2$. This exercise showed that underway $f\text{CO}_2$ can be determined to a high
26 level of precision ($\pm 2 \mu\text{atm}$) with a variety of equilibrator and system designs.

27 Other workers have undertaken at-sea intercomparisons of different variables. For instance,
28 Johnson et al. (1999) compared C_T , $f\text{CO}_2$ and A_T measurements during the same
29 intercomparison exercise as reported for $f\text{CO}_2$ by Körtzinger et al. (2000). These scientists
30 found a systematic $f\text{CO}_2$ overestimation of $9 \mu\text{atm}$ when calculated from C_T and A_T
31 measurements relative to observed $f\text{CO}_2$. Lamb et al. (2001) investigated 25 cruises in the
32 Pacific Ocean where at least two of the four inorganic carbon variables were determined.

1 They examined the consistency of the dataset using Certified Reference Material (CRM)
2 analyses, precision of at-sea replicate analyses, agreement between shipboard analyses and
3 replicate shore based analyses, comparison of deep water values at locations where two or
4 more cruises overlapped or crossed, consistency with other hydrographic parameters and
5 internal consistency between multiple carbon variables measurements. Using all this evidence
6 the carbonate data was adjusted for inconsistencies and a combined dataset was constructed,
7 which showed that C_T and A_T had an estimated overall accuracy of $3 \mu\text{mol kg}^{-1}$ and
8 $5 \mu\text{mol kg}^{-1}$, respectively.

9 Other studies pointed out some inconsistencies: Millero et al. (2002) noted that the use of pH_T
10 and C_T from field measurements from the Atlantic, Indian, Southern and Pacific oceans
11 yielded standard errors (1σ) of $\pm 22.3 \mu\text{atm}$ in calculated $p\text{CO}_2$ and $\pm 4.3 \mu\text{mol kg}^{-1}$ in
12 calculated A_T . Lueker et al. (2000) noted that observed values of $p\text{CO}_2$ above $500 \mu\text{atm}$ were
13 by, on average, 3.35 % (if $f\text{CO}_2$ was $500 \mu\text{atm}$ that will be $17 \mu\text{atm}$) higher than $p\text{CO}_2$
14 calculated from C_T and A_T . This tendency towards a larger differences between measured
15 $p\text{CO}_2$ and calculated $p\text{CO}_2$ at higher $p\text{CO}_2$ levels was also observed by McElligott et al.
16 (1998), suggesting that it might result from inaccuracy in the formulation of the solubility
17 coefficient of CO_2 in seawater (K_0). However, this apparent discrepancy has not yet been
18 explained satisfactorily (Dickson, 2010). It is possible that an unidentified acid-base system
19 affects the calculation of $p\text{CO}_2$ or that one or more dissociation constants for acid-base
20 equilibria are not well parameterised at high $p\text{CO}_2$ (Dickson, 2010).

21 The aims of our study were to evaluate the quality of our observations of inorganic carbon
22 variables and to investigate differences between observed and calculated variables in order to
23 identify means of improving data quality. Our study differs from some previous work in two
24 respects: firstly, our study was undertaken in surface waters of shelf seas where spatial
25 variability is high; and, secondly, the study was not designed at the outset as an
26 intercomparison exercise, which normally involves placing all the instruments in one
27 laboratory, sampling from a single seawater supply and an intense focus on every aspect of
28 the carbonate chemistry measurements. Instead, the instruments were in three separate
29 laboratories, with samples taken from four different seawater outlets and the operators
30 conducting multiple tasks as part of the multidisciplinary research activities undertaken on the
31 cruise. Therefore our findings are more representative of a typical multidisciplinary research
32 cruise.

1

2 **2 Material and methods**

3 The data used in this study were collected in the period 06 June to 07 July 2011 during the
4 RRS *Discovery* research cruise D366 in northwestern European shelf seas. The cruise formed
5 part of the UK Ocean Acidification Research Programme. Two variables of the carbonate
6 system ($p\text{CO}_2$ and pH_T), plus salinity and sea surface temperature, were measured at a high
7 temporal resolution (every 5 min for one $p\text{CO}_2$ system ($p\text{CO}_2$ -1), every 6 min for pH_T , and
8 every 1 min for a second $p\text{CO}_2$ system ($p\text{CO}_2$ -2)). These instruments received a continuous
9 flow of water from the ship's underway continuous seawater supply (intake positioned at ca. 5
10 m depth). In addition, nutrients (nitrate plus nitrite, phosphate and silicate), C_T and A_T were
11 sampled every 2 h from the underway supply, and also collected in surface waters sampled by
12 CTD casts (samples obtained from the sampling bottle closest to 5 m depth; typically between
13 2.0 and 8.2 m). Continuous temperature and conductivity data were obtained from a Sea-Bird
14 Electronics SBE45 thermosalinograph (TSG) installed on the ship's underway supply.
15 Discrete surface water samples for salinity (S) were collected every 4 h in order to calibrate
16 the conductivity measurements. Discrete salinity samples were analysed using a salinometer
17 (Guildline Autosol 8400B). Photosynthetically Active Irradiance (PAR), radiation between
18 400 and 700 nm was measured as part of the ship's meteorological parameters with a 2-pi
19 sensor (Skye Instruments, model SKE 510) positioned at 10 m height.

20

21 **2.1 Carbonate chemistry analysis**

22 2.1.1 Partial pressure of CO_2

23 Quasi-continuous measurements of $p\text{CO}_2$ in surface water and marine air were undertaken
24 using two different instruments (hereafter $p\text{CO}_2$ -1 and $p\text{CO}_2$ -2). The $p\text{CO}_2$ -1 and $p\text{CO}_2$ -2
25 systems undertook 6,187 and 26,671 measurements of surface water $p\text{CO}_2$ during the cruise,
26 respectively.

27 System 1: System $p\text{CO}_2$ -1 was an underway $p\text{CO}_2$ instrument (PML-Dartcom *Live pCO2*) as
28 described in detail by Hardman-Mountford et al. (2008), with the modified 'vented'
29 equilibrator introduced by Kitidis et al. (2012). The instrument was located in a mid-ship
30 chemistry laboratory. The system used a vented-showerhead equilibrator, with ambient light

1 blocked out, to equilibrate seawater CO₂ with a headspace. In order to maintain atmospheric
2 pressure in the equilibrator headspace, the unit was vented to a second equilibrator, which in
3 turn was vented to the atmosphere via a 2 m coil of stainless steel tubing (1.5 mm internal
4 diameter). The equilibrator was fitted with 2 platinum resistance thermometers (Pico
5 Technology, model PT100) and a water-jacket supplied with seawater from the ship's
6 underway seawater system. A seawater flow of 1.6 L min⁻¹ was maintained across the
7 equilibrator. The average warming between the ship's underway seawater intake and the
8 equilibrator was 0.5 °C (standard deviation = 0.1 °C). Atmospheric measurements of CO₂
9 were taken from an intake located forward on the deck above the ship's bridge. Both gas
10 streams from the equilibrator headspace and the air inlet were dried in a Peltier cooler (-20
11 °C). Mixing ratios of CO₂ and water in the marine air and equilibrator headspace were
12 determined by infrared detection (LI-840, LI-COR). Measurements were referenced against
13 secondary calibration gases from BOC Gases (UK) with known CO₂ mixing ratios (0, 251.3
14 and 446.9 μmol CO₂ mol⁻¹) in synthetic air mixtures (21 % oxygen and 79 % nitrogen). All
15 calibration gases underwent pre- and post-cruise calibration against certified primary
16 standards from the National Oceanic and Atmospheric Administration (NOAA), which had
17 values of 244.9 and 444.4 μmol CO₂ mol⁻¹.

18
19 System 2: System *p*CO₂-2 was an underway *p*CO₂ system located in a container laboratory
20 positioned on the aft-deck of the ship. The instrument setup and calibration procedures are as
21 described by Bakker et al. (2007), with the exception of the vented equilibrator. The
22 percolating packed bed-type equilibrator was identical to the one described by Schuster and
23 Watson (2007). The equilibrator of transparent perspex was positioned next to the window of
24 the container without blinds. Atmospheric samples were taken from an air inlet located
25 forward on the deck above the ship's bridge. Samples from the equilibrator headspace and
26 marine air were partially dried by being passed through an electric cool box at about 2 °C,
27 prior to analysis. Mixing ratios of CO₂ and water in the marine air and equilibrator headspace
28 were determined by infrared detection with a LI-COR LI7000. The LI-COR was calibrated
29 using secondary gas standards BOC Gases (UK) with CO₂ mixing ratios of 2.4, 260.9, 364.2
30 and 473.1 μmol CO₂ mol⁻¹ in an artificial air mixture (21 % oxygen, 79 % nitrogen). All
31 calibration gases underwent pre- and post-cruise calibration against certified primary
32 standards from the NOAA, which had values of 251.6, 347.2 and 448.8 μmol CO₂ mol⁻¹. The
33 seawater flow to the container laboratory was highly variable throughout the cruise. This was

1 due to the location of the container downstream of an intermittently large water demand for
2 an experiment. The water flow was regulated to a maximum of 1.8 L min^{-1} , to avoid flooding
3 of the equilibrators and CO_2 analyser during sudden spikes in supply. The water flow tended to
4 gradually decrease to very low flow over 6 to 12 h. Two platinum resistance thermometers
5 positioned in the upper and lower part of the seawater stream determined the temperature of
6 the seawater in the equilibrators (a PT probe (Omega) with modified electronics). Average
7 warming of the seawater between the intake and the equilibrators was estimated as $0.5 \text{ }^\circ\text{C}$
8 (standard deviation = $0.4 \text{ }^\circ\text{C}$). The large temperature deviations reflected the irregular
9 seawater flow to the equilibrators. In addition, the post-cruise temperature calibration of the
10 PT100 sensors showed excessive drift of $4\text{-}5 \text{ }^\circ\text{C}$ relative to the pre-cruise calibration. The
11 absolute calibration of the temperature sensors was therefore deemed unreliable. Equilibrator
12 temperatures from 13 June (09:57) to 17 June (17:12) have been reduced by $0.7 \text{ }^\circ\text{C}$, while
13 equilibrator temperatures after 19 June (21:57) have been increased by $0.7 \text{ }^\circ\text{C}$ to remove the
14 negative temperature changes.

15
16 The precision of both LI-COR's $p\text{CO}_2$ measurements was $1 \text{ } \mu\text{atm}$, established using standard
17 gases. We estimated different accuracies for the two systems: $4 \text{ } \mu\text{atm}$ for the system $p\text{CO}_2\text{-}1$
18 and $10 \text{ } \mu\text{atm}$ for system $p\text{CO}_2\text{-}2$. The $p\text{CO}_2$ was computed from the CO_2 mixing ratios and the
19 ship's barometric pressure corrected from 18 m height to sea level, and corrected for seawater
20 vapour pressure (Weiss and Price, 1980). Sea surface $p\text{CO}_2$ data were corrected to sea surface
21 temperature to account for the warming between the seawater intake and the equilibrators
22 (Takahashi et al., 1993). The accuracies of the temperature measurements inside the
23 equilibrators were estimated to be $0.02 \text{ }^\circ\text{C}$ and $5 \text{ }^\circ\text{C}$ for $p\text{CO}_2\text{-}1$ and $p\text{CO}_2\text{-}2$ respectively.

24 $p\text{CO}_2\text{-}1$ measurements were backdated by 1 min and $p\text{CO}_2\text{-}2$ measurements by 3 min to
25 account for the travel time of the seawater between the seawater intake and the respective
26 equilibrators. The time offsets (1 min and 3 min) between seawater intake and equilibrators
27 were chosen objectively as those producing the minimum standard deviation between paired
28 equilibrator and intake temperatures. The intercomparison exercise was carried out on the
29 datasets after they had been adjusted according to the procedures just described, including
30 corrections to in-situ seawater temperature described above.

1 2.1.2 pH_T

2 Surface water pH_T was measured continuously with an automated instrument located in the
3 mid-ship chemistry laboratory and connected to the ship's underway seawater supply. The
4 pH_T system undertook 29,950 measurements. The surface distribution and the processes that
5 controls pH_T in this cruise have been described by Rérolle et al. (2014). The measurement
6 technique, described by Rérolle et al. (2013), was based on a colorimetric method using
7 Thymol Blue as pH indicator (Clayton and Byrne, 1993; Rérolle et al., 2012). pH_T was
8 determined on the total pH scale. Measurements were made every 6 min with a precision of
9 1 mpH (Rérolle et al., 2013). Three bottles of Tris pH buffer provided by Dr Andrew Dickson
10 (Scripps Institution of Oceanography, Marine Physical Laboratory, University of California
11 San Diego, USA) were analysed at the beginning, middle and end of the cruise to check the
12 accuracy of the pH_T measurements, which was 4 mpH. The Thymol Blue extinction
13 coefficients were determined in the laboratory following the cruise, applying the salinity and
14 temperature ranges observed during the cruise, with the indicator's dissociation constant taken
15 from Zhang and Byrne (1996). Measurements at sea were made at the seawater temperature
16 plus 0.2 °C due to warming between the seawater intake and the pH instrument. In order to
17 minimise absorbance interference by particulates an in-line filter (0.45 μm pore size, Millex
18 HP syringe filter MilliporeExpress® (PES) membrane 33 mm diameter, Millipore) was
19 placed at the entry of the sample tube. Chromophoric Dissolved Organic Matter (CDOM)
20 only absorbs weakly in the visible where Thymol Blue absorbance is measured (<3 % at 435
21 nm and <1% at 596 nm) and is accounted for in the seawater blank. CDOM absorbance
22 interference is thereby cancelled out. Additionally, measurements at the wavelength 750 nm
23 (not affected by Thymol Blue indicator) were used to monitor for sample turbidity and
24 instrument drift.

25

26 2.1.3 Dissolved inorganic carbon and total alkalinity

27 Samples for C_T and A_T analysis were collected from the underway seawater supply (322
28 samples in total), and shallow depths sampled (64 samples) using Ocean Test Equipment
29 bottles on the CTD frame following procedures detailed in Bakker et al. (2007). All samples
30 were fixed with 50 μL of saturated mercuric chloride (HgCl_2) solution per 250 mL seawater.
31 The samples were analysed in duplicate on replicate 250 mL samples bottles. Two VINDTAs
32 3C (Versatile Instrument for the Determination of Titration Alkalinity; Marianda) were used

1 to determine C_T and A_T , with CRMs (batch 107) analysed in duplicate for C_T and A_T at the
2 beginning, middle and end of each use of a coulometric cell. One VINDTA was used for
3 surface water samples and the other one for CTD casts. For consistency, we checked offsets
4 between those underway and CTD samples which were less than 30 min apart. Because
5 underway sampling was often stopped when we were on station, only 19 stations could be
6 checked. The average offsets between CTD and surface water samples from the continuous
7 supply were 0.3 and 0.5 $\mu\text{mol kg}^{-1}$ for C_T and A_T , respectively. This is below the accuracy of
8 the method and so we merged the data from the two types of sampling and the two VINDTA
9 3C instruments. The concentration of C_T was determined using coulometric analysis (Johnson
10 et al., 1987). Analysis for A_T was carried out by potentiometric titration with hydrochloric
11 acid to the carbonic acid end point (Dickson, 1981). The accuracies of the C_T and A_T
12 measurements were 2.0 and 1.5 $\mu\text{mol kg}^{-1}$ and the precisions 1.7 and 1.2 $\mu\text{mol kg}^{-1}$,
13 respectively (159 CRMs analysed in duplicate). The combined carbonate chemistry dataset is
14 available via the British Oceanographic Data Centre at
15 [https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/f56e35bc-635e-0ab5-](https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/f56e35bc-635e-0ab5-e044-000b5de50f38/)
16 [e044-000b5de50f38/](https://www.bodc.ac.uk/data/published_data_library/catalogue/10.5285/f56e35bc-635e-0ab5-e044-000b5de50f38/) with doi: 10.5285/f56e35bc-635e-0ab5-e044-000b5de50f38.

17

18 **2.2 Nutrients**

19 Analyses of nitrate and nitrite, phosphate and silicate were undertaken using a segmented
20 flow auto-analyser (Skalar San+) following methods described by Kirkwood (1989). Samples
21 were stored in 25 mL polycarbonate vials and kept refrigerated at approximately 4 °C until
22 analysis (conducted within 12 h after sampling). Nutrient concentrations were used for the
23 calculations of the carbonate chemistry system.

24

25 **2.3 Carbonate chemistry calculations**

26 We applied the CO2SYS programme (MATLAB version) (Lewis and Wallace, 1998; Van
27 Heuven et al., 2011) to all possible pairs of pH_T , pCO_2 , C_T and A_T measurements to calculate
28 the other variables, using the carbonate equilibria constants described by Mehrbach et al.
29 (1973) and refitted by Dickson and Millero (1987) (here after Mehrbach constants). We used
30 the Mehrbach constants because they led to the smallest inconsistencies between different
31 high-accuracy measurements in previous observational studies (Clayton et al., 1995; Lee et

1 al., 1997; McElligott et al., 1998; Wanninkhof et al., 1999). However, we also compared
2 against results calculated using constants from Roy et al. (1993), Lueker et al. (2000) and
3 Millero et al. (2006) (Sect. 3.3).

4 For the dissociation constant of boric acid we used Dickson (1990b), for bisulphate ions
5 Dickson (1990a) and for the ratio of total boron to salinity, we used Uppström (1974), but
6 also compared against results calculated using Lee et al. (2010) (Sect. 3.3).

7

8 **2.4 Intercomparison**

9 The sampling frequencies of $p\text{CO}_2$ -1 and pH_T were both around 5 min, but the measurements
10 were not synchronised and were undertaken simultaneously (within 1 min) on only 208
11 occasions. However, it was possible to interpolate pH_T determinations (with a maximum
12 interval of 5 min) and thereby obtain values at comparable times.

13 Data from the underway temperature, salinity, PAR, $p\text{CO}_2$ and pH_T measurements were
14 retrieved at the times of nutrient, C_T and A_T measurements.

15 Statistical analyses were used to determine the level of agreement between observed and
16 calculated carbonate system variables:

17 - Pearson's correlation coefficient (r): A measure of the degree of linear dependence between
18 two variables.

19 - Mean Residual (MR): Average difference between two variables, e.g. between the observed
20 values and the values calculated from measurements of a pair of other carbonate variables;
21 MR will be negative if the observed values are on average lower than the calculated values.

22 - Root Mean Square Error (RMSE): Square root of the mean of the squared differences
23 between the observed and calculated values.

24

25 Some properties of individual variables (as opposed to comparisons) are also used in Sect. 3:

26 • Accuracy is an expression of the lack of bias and relates to the degree of agreement of
27 a measured value with the true value (as determined using a CRM).

28 • Uncertainty characterizes the range of values within which the true value is asserted to
29 lie with some level of confidence. Uncertainty is derived from inaccuracy and
30 imprecision of measurements, and also from propagation of errors for calculated
31 variables.

1 In this study, uncertainties in calculated values were determined by a Monte Carlo approach
2 as follows: 1) The original carbonate chemistry variable values in the dataset were input into
3 the CO2SYS program (MATLAB version) (Van Heuven et al., 2011); 2) Artificial random
4 errors (normally distributed according to the central limit theorem, with a mean of zero and
5 standard deviation equal to the accuracy of measurement) were calculated using a random
6 number generator; 3) New carbonate chemistry variable values (the original ones plus the
7 randomly generated errors) were input into CO2SYS. Calculated $p\text{CO}_2\text{-1}$ and calculated
8 $p\text{CO}_2\text{-2}$ have the same uncertainty because they depend only on the accuracies of the
9 variables from which they are calculated. They are therefore identical for both $p\text{CO}_2$ systems.
10 The calculated uncertainty of parameters calculated from $p\text{CO}_2\text{-2}$ as one of the input variables
11 is higher than those from $p\text{CO}_2\text{-1}$ because the measurement accuracy was higher. This Monte
12 Carlo approach has previously been used by Juranek et al. (2009) to calculate uncertainties in
13 calcium carbonate saturation states.

14 In this paper we use accuracies and calculated uncertainties as benchmarks. We use them to
15 provide an assessment of what it is reasonable to expect for a match between observed and
16 calculated values. We compare values of accuracy with MRs, and calculated uncertainties
17 with RMSEs, to evaluate if the calculated variables are in “good” agreement with the
18 measured variables. All comparison resulted in MR less than or equal to accuracy (with the
19 exception of predictions based on $p\text{CO}_2$ and pH and A_T from C_T and $p\text{CO}_2\text{-1}$). In terms of
20 RMSE, all comparisons resulted in RMSE less than twice the uncertainty, except those
21 involving measured $p\text{CO}_2\text{-2}$. On this basis we conclude that there is generally good agreement
22 between measured and calculated variables, except those comparison involving measured
23 $p\text{CO}_2\text{-2}$ or calculations from pH_T and $p\text{CO}_2$. For example, if the MR and RMSE between
24 measured and calculated (from A_T and $p\text{CO}_2\text{-1}$) C_T are $-1.7 \mu\text{mol kg}^{-1}$ and $3.9 \mu\text{mol kg}^{-1}$,
25 whereas the accuracy is $2.0 \mu\text{mol kg}^{-1}$ and the calculated uncertainty is $2.5 \mu\text{mol kg}^{-1}$, then we
26 conclude that there is good agreement between the calculated and measured C_T values. As
27 another example, a RMSE of 0.008 and a MR of 0.001 demonstrates a good agreement
28 between measured and calculated (from C_T and A_T) pH_T if the calculated uncertainty
29 according to the Monte Carlo approach is 0.005 and the measurement accuracy is 0.004. A
30 smaller MR on its own does not demonstrate a better agreement; it should always be put in
31 context with RMSE, accuracy and uncertainty.

32

1 **3 Results and discussion**

2 **3.1 Comparison between two $p\text{CO}_2$ systems**

3 The $p\text{CO}_2$ datasets obtained using the $p\text{CO}_2$ -1 and $p\text{CO}_2$ -2 systems were significantly
4 correlated ($r = 0.956$, $p < 0.001$, $df = 2679$) (Fig. 1). We used a major axis model II regression
5 because both $p\text{CO}_2$ datasets included uncertainty ($p\text{CO}_2$ -1 and $p\text{CO}_2$ -2 was not one dependent
6 controlled and one independent variable). We used the R code for Model II Regression
7 (Legendre, 2014). The resulting equation of the regression is $p\text{CO}_2$ -1 = $0.9 (\pm 2.1) + 0.99 (\pm$
8 $0.01) \times p\text{CO}_2$ -2. Confidence intervals are used for testing the null hypothesis of a slope of 1
9 and a y-intercept of 0 (Quinn and Keough, 2002). The 95 % confidence interval of the slope
10 does include the value 1 and the 95 % confidence interval of the intercept does include the
11 value 0. The comparison between both sets of $p\text{CO}_2$ data revealed a mean residual of $-2 \mu\text{atm}$
12 ($p\text{CO}_2$ -1 minus $p\text{CO}_2$ -2). The RMSE was $10 \mu\text{atm}$.

13 Körtzinger et al. (2000) reported that even after correction of all differences between
14 equilibrator temperature readings and following a time synchronization procedure, the
15 remaining mean residual in their study was ca. $2 \mu\text{atm}$ for most of their cruise. The study by
16 Körtzinger et al. (2000) can be described as an ideal open ocean exercise with all instruments
17 sharing a common seawater supply, positioned in the same laboratory and sharing common
18 calibration gases. Considering that our cruise took place in coastal waters with strong
19 gradients in temperature, salinity and chemical variables (Körtzinger et al., 1996; Bozec et al.,
20 2005), with the instruments situated in different laboratories and using different calibration
21 gases, and known water flow problems (Sect. 2.1.1), we conclude that an average difference
22 of $2 \mu\text{atm}$ and a RMSE of $10 \mu\text{atm}$ were good outcomes. The non-ideal conditions of our
23 intercomparison did not negatively affect the overall consistency (we obtained comparable
24 results to the study by Körtzinger et al. (2000)). This result is also comparable with a previous
25 (the only other) coastal water intercomparison, described by Körtzinger et al. (1996). In this
26 study, where there was highly variable spatial $p\text{CO}_2$ distribution in the southern North Sea,
27 the average difference between observed values was $0.2 \mu\text{atm}$ (standard deviation =
28 $1.2 \mu\text{atm}$).

29 There have been a number of intercalibration exercises of $p\text{CO}_2$ systems in an indoor seawater
30 pool at the National Institute for Environment Studies, Japan (in 1993, 1998, 2003 and 2009)

1 (Katayama et al., 1999; IOCCP, 2004; Pierrot et al., 2009). Most of the instruments showed
2 good agreement (within 2 μatm).

3 The MR of the observed $p\text{CO}_2\text{-1}$ compared with $p\text{CO}_2$ calculated from C_T and A_T was 3 μatm
4 ($n = 43$), and of $p\text{CO}_2\text{-2}$ was 1 μatm ($n = 156$), both MR within the measurement accuracy of
5 the instrument (Table 1). The accuracies of C_T and A_T were $\pm 2.0 \mu\text{mol kg}^{-1}$ and
6 $\pm 1.5 \mu\text{mol kg}^{-1}$ respectively, and this translates into a propagated $p\text{CO}_2$ uncertainty of
7 4 μatm . This is a clear example of where the MR does not provide whole story, because
8 $p\text{CO}_2\text{-2}$ compared to calculated from C_T and A_T had a smaller MR but higher RMSE and
9 lower r than $p\text{CO}_2\text{-1}$ (Table 1). Lower MR does not necessarily mean that there is a better
10 agreement, because positive values may compensate negative values, as is the case here. The
11 ranges of the residuals were -7 to 14 μatm and -29 to 39 μatm , for $p\text{CO}_2\text{-1}$ and $p\text{CO}_2\text{-2}$
12 respectively. Both of the mean residuals were well within the expected accuracy of $p\text{CO}_2$
13 calculated from C_T and A_T measurements (Millero, 2007). The residuals of the two $p\text{CO}_2$
14 datasets are presented in Fig. 2. A diurnal cycle was observed in the residuals after 8 days of
15 the cruise, with the amplitude increasing over time and reaching a maximum difference
16 between the two $p\text{CO}_2$ systems on julian day 177 (Figs. 2 and 3). The average difference
17 ($p\text{CO}_2\text{-1}$ minus $p\text{CO}_2\text{-2}$) was -2 μatm and the maximum difference was 41 μatm .

18 Temperature forms a critical parameter for $p\text{CO}_2$ calculations, with for example temperature
19 differences of 2 $^\circ\text{C}$ translate into $p\text{CO}_2$ differences of 32 μatm . The two analytical $p\text{CO}_2$
20 systems were not in the same laboratory and therefore water spent different lengths of time
21 and took different routes between the ship's seawater intake and the two equilibrators, and
22 therefore warmed differently. This temperature effect was more important when the ship sailed
23 through strong surface water temperature gradients. Rapid changes in seawater temperature of
24 up to 2-3 $^\circ\text{C min}^{-1}$ were observed on julian day 180. However, such rapid temperature
25 changes at the seawater intake may not fully translate into $p\text{CO}_2$ changes in the equilibrator,
26 as equilibrator $p\text{CO}_2$ is effectively integrated over the equilibration time (around 8 min for
27 CO_2 in both equilibrators). Small differences in the equilibration time between the two
28 systems, i.e. how rapidly they respond to a change in seawater $p\text{CO}_2$, may account for some
29 of the observed differences between observations by $p\text{CO}_2\text{-1}$ and $p\text{CO}_2\text{-2}$. Moreover, no trend
30 was observed between the difference in measured $p\text{CO}_2\text{-1}$ and $p\text{CO}_2\text{-2}$ versus the difference
31 between the temperature in equilibrator 1 and the sea surface temperature, whereas a positive
32 relationship was observed for equilibrator 2 (Fig. 4). The discrepancy between the two $p\text{CO}_2$

1 systems was negative at low light levels (at night $p\text{CO}_2\text{-1}$ was smaller than $p\text{CO}_2\text{-2}$) and
2 positive during daylight hours ($p\text{CO}_2\text{-1}$ larger than $p\text{CO}_2\text{-2}$) (Fig. 3). This pattern is consistent
3 with respiration at night and photosynthesis during the day in the seawater supply to the
4 $p\text{CO}_2\text{-2}$ equilibrator or in the equilibrator itself. The length of the seawater pipes to
5 equilibrator 2 was about twice as long as to equilibrator 1. Furthermore, the equilibrator of
6 $p\text{CO}_2\text{-2}$ was subject to direct daylight, compared to the $p\text{CO}_2\text{-1}$ equilibrator which was
7 shielded from light. A multiple regression analysis was performed to estimate the relative
8 importance of two factors (the temperature difference between the two equilibrators (ΔT_{eq})
9 and PAR) in determining the size of the $p\text{CO}_2$ differences ($\Delta p\text{CO}_2 = p\text{CO}_2\text{-1} - p\text{CO}_2\text{-2}$). A
10 correlation coefficient of 0.47 between the predictor variables (ΔT_{eq} and PAR) indicated that
11 they are not strongly correlated. Tolerance (or the inverse of the variance inflation factor) is
12 0.78 indicating no colinearity problems as this value is well above the tolerance threshold of
13 0.1 (Quinn and Keough, 2002). The analysis indicated that the correlation of $\Delta p\text{CO}_2$ with
14 ΔT_{eq} and PAR was statistically significant ($p < 0.0001$; $F = 587.6$) and that the two parameters
15 together explained 38 % of the $\Delta p\text{CO}_2$ variance. ΔT_{eq} and PAR were found to individually
16 explain 17 % and 18 % respectively of the variance in $\Delta p\text{CO}_2$. Thus we conclude that the
17 diurnal pattern in the $p\text{CO}_2$ differences between the two instruments were caused by a
18 combination of: (a) biological activity in the seawater system or in equilibrator 2, and (b)
19 variations in water flow and temperature perturbations in $p\text{CO}_2\text{-2}$ (Sect. 2.1.1). The first issue
20 should be addressed in future studies by protection of seawater tubing and equilibrators from
21 light and by regular cleaning of the seawater intake and equilibrators. In addition, the flow
22 rate of the ship's seawater supply should be kept constant, while the water flow to
23 equilibrators should be sufficiently large and should be kept constant by using a water flow
24 controller. Furthermore, the temperature measurements inside the equilibrator need to be
25 accurate, in agreement with Körtzinger et al. (2000) and as implemented in the Surface Ocean
26 CO_2 Atlas (SOCAT) by making an accuracy of 0.05 °C for the equilibrator temperature a pre-
27 requisite for data set quality flags A and B (Pfeil et al., 2013; Bakker et al., 2014). The
28 maximum $p\text{CO}_2$ discrepancies observed in our study (up to 20 μatm) are larger than the 5
29 μatm criterion for cross-over discrepancies in order to achieve flag A status (the highest
30 quality) in the new SOCAT version 3 (Wanninkhof et al., 2013), further emphasising the
31 desirability of following the recommendation of this paper.

32

1 3.2 Intercomparison of measured and calculated variables

2 The results of the intercomparison between observed carbonate chemistry variables and those
3 calculated from different pairs of measured variables are presented in Table 1. Statistical
4 techniques were used to evaluate the agreement between the observed and the calculated
5 values. The comparison between observed pH_T and pH_T calculated from observed C_T and A_T
6 showed a mean residual of 0.001 pH_T units and a RMSE of 0.008 pH_T units (Table 1). This
7 compared to the 0.004 pH_T units accuracy of the measurements. The linear correlation
8 coefficient (r) between observed and calculated values was 0.952. pH_T calculated from C_T and
9 A_T had an uncertainty of 0.005 pH_T units. The RMSE corresponded to twice the accuracy of
10 the pH_T measurements, and slightly less than twice the uncertainty expected from the
11 calculation, so we conclude that there was good agreement between calculated and measured
12 pH_T .

13 The value of pH_T calculated from $p\text{CO}_2\text{-1}$ and A_T had a RMSE of 0.006 when compared to
14 measured pH_T , and the same calculation with $p\text{CO}_2\text{-2}$ led to a RMSE of 0.013 (Table 1). The
15 calculated values of pH_T using $p\text{CO}_2\text{-1}$ and A_T were therefore better (lower RMSE) than pH_T
16 calculated from C_T and A_T , whereas the calculations using $p\text{CO}_2\text{-2}$ and A_T had a higher
17 RMSE. Calculations of pH_T from the combination of $p\text{CO}_2$ with either C_T or A_T may be
18 expected to yield more accurate estimates than calculations of pH_T from C_T and A_T (Table 1),
19 because they do not require reliable estimates of the second dissociation constant of carbonic
20 acid and are relatively insensitive to uncertainties in C_T and A_T (Millero, 2007). However, this
21 is not always the case, as shown here.

22 $p\text{CO}_2$ calculated from C_T and A_T compared to the $p\text{CO}_2\text{-1}$ and $p\text{CO}_2\text{-2}$ observational datasets
23 showed RMSE values of 6 μatm and 12 μatm , and mean residuals of 3 and 1 μatm (Table 1).
24 This compares to an accuracy associated with direct measurement of $p\text{CO}_2$ of 4 μatm for
25 $p\text{CO}_2\text{-1}$ and 10 μatm for $p\text{CO}_2\text{-2}$. $p\text{CO}_2$ calculated from C_T and A_T is predicted to have an
26 uncertainty of 4 μatm . We therefore conclude that our calculated dataset is in good agreement
27 with the measured dataset.

28 Lueker et al. (2000) carried out a similar field-based intercomparison with comparable
29 measurement quality and overall uncertainty. At lower $f\text{CO}_2$, Lueker et al.'s differences
30 between calculated and measured $f\text{CO}_2$ were slightly lower than in our study, but at high $f\text{CO}_2$
31 they obtained slightly higher differences. Lueker et al. (2000) reported that the mean relative
32 difference between measured $f\text{CO}_2$ and $f\text{CO}_2$ calculated from C_T and A_T (for $f\text{CO}_2$ less than

1 500 μatm) was 0.07 % (standard deviation = 0.50 %). For $f\text{CO}_2$ above 500 μatm , there was a
2 mean relative difference of 3.3 % (standard deviation = 1.2 %). For example, for an $f\text{CO}_2$ of
3 500 μatm this corresponds to a difference of 16.5 μatm and a standard deviation of 6.0
4 μatm . $p\text{CO}_2$ and pH_T do not make a good pair for predicting other variables because CO_2 and
5 hydrogen ion concentration are smaller than carbonate and bicarbonate concentration.
6 Therefore, relatively small errors in CO_2 and/or hydrogen ion propagate into relatively large
7 errors in carbonate and bicarbonate concentration, when the system is computed from $p\text{CO}_2$
8 and pH_T . This is also indicated by our data, where the mean residuals for A_T and C_T (observed
9 minus calculated) are comparable to the calculated uncertainties, but are both one order of
10 magnitude greater from $p\text{CO}_2$ and pH_T than the accuracy and precision of observations (Table
11 1).

12 C_T calculated from A_T and $p\text{CO}_2$ had RMSE values of 3.9 $\mu\text{mol kg}^{-1}$ and 7.2 $\mu\text{mol kg}^{-1}$
13 compared to the measurement datasets of systems $p\text{CO}_2$ -1 and 2, and mean residuals of -1.7
14 and -0.3 $\mu\text{mol kg}^{-1}$ (Table 1). This compares to an accuracy associated with direct
15 measurement of C_T of 2 $\mu\text{mol kg}^{-1}$. A_T calculated from C_T and $p\text{CO}_2$ -1 had a MR of to
16 2.1 $\mu\text{mol kg}^{-1}$ and a RMSE value of 4.6 $\mu\text{mol kg}^{-1}$ and the calculated uncertainty was
17 3 $\mu\text{mol kg}^{-1}$ (Table 1). This compares to an accuracy associated with direct measurement of A_T
18 of 1.5 $\mu\text{mol kg}^{-1}$.

19 Several papers have raised the issue of the impact of organic acids on computations of the
20 CO_2 system in coastal waters (Kim et al., 2006; Hernández-Ayón et al., 2007; Kim and Lee,
21 2009). Dissolved organic matter (DOM) produced by phytoplankton during photosynthesis
22 potentially makes a significant contribution to seawater total alkalinity, the magnitude of
23 which has been reported to vary depending on the phytoplankton species (Kim and Lee,
24 2009). Dissolved organic matter in this context acts as an additional acid-base pair in
25 seawater. Analytically, this would interfere with our determination of A_T (by HCl titration).
26 We can therefore compare measured A_T values to the respective calculated values from e.g.
27 C_T and pH_T . As shown in Fig. 5, there is only a weak correlation ($r^2 = 0.06$) between
28 dissolved organic carbon (DOC) and A_T discrepancy, of the sign we should expect if DOM
29 was affecting A_T but not statistically significant. Similar results were as also obtained in an
30 upwelling environment (Loucaides et al., 2012). Therefore, we found no evidence of a
31 significant contribution to seawater total alkalinity from DOM. During three days in the same
32 coccolithophore bloom we collected additional samples of filtered C_T and A_T from the

1 underway seawater supply. The filtering was carried out using an in-line filter (Sartorius
2 Sartobran 300 Sterile capsule, 0.45 μm pore size). In this way we studied the effects of the
3 presence of calcite mineral particles in the surface waters on C_T and A_T measurements. The
4 average differences between unfiltered and filtered samples were 2.4 $\mu\text{mol kg}^{-1}$ and
5 3.7 $\mu\text{mol kg}^{-1}$ for C_T and A_T , respectively, with values in the unfiltered samples being higher.
6 The differences between replicates of filtered samples were on average higher than those of
7 unfiltered samples. The differences between measured pH_T and pH_T calculated from filtered
8 and unfiltered C_T and A_T were -0.003 and 0.005, respectively. The difference between
9 measured $p\text{CO}_2\text{-1}$ and $p\text{CO}_2\text{-2}$ and $p\text{CO}_2$ calculated from filtered and unfiltered C_T and A_T
10 were -4 μatm in both cases. An influence of dissolution of calcite particle on C_T and A_T
11 measurements would be expected to lead to filtered C_T and A_T being lower than unfiltered
12 with discrepancies twice as large for A_T as for C_T . This ratio was however not exactly
13 observed in our study, and hence it is unclear from our observations whether dissolution of
14 CaCO_3 particles affected C_T and A_T measurements. Filtration of samples for C_T potentially
15 introduces a further error through CO_2 loss by turbulence and ebullition, which can affect the
16 C_T measurement, although we took precautions to avoid bubbles in the filter.

17 So far in this paper we have checked the consistency of the carbonate system and assessed the
18 quality of the carbonate system measurements. Overall, the sizes of the offsets between
19 measured and calculated values agreed well with expectations based on theoretical
20 calculations and expected accuracies of measurements. Four out of the five independent
21 datasets of carbonate chemistry variables are deemed to be of high-quality and therefore
22 suitable to be used as a basis for evaluations of the impacts of ocean acidification by other
23 scientists on the same cruise, see for instance the papers by Poulton et al. (2014) and Young et
24 al. (2014).

25 Next, we examine possible reasons for discrepancies observed during the $p\text{CO}_2$ measurements
26 using $p\text{CO}_2\text{-2}$ (points 1 to 3 below) and make 'good practice' recommendations from
27 investigation of the remaining discrepancies (points 4 and 5 below). Some of these
28 recommendations are new and some are not but we think it is good to compile them all:

- 29 1. The duration of seawater transit in the ship's underway system from the seawater intake to
30 the point of sample collection or measurement (1 to 3 min) varied between the carbonate
31 chemistry measurements. It is important to recognise the period of time it takes for the
32 seawater to arrive in the equilibrator of a $p\text{CO}_2$ instrument. This will allow correction for

1 the difference between the intake temperature and the equilibrator temperature. This has
2 already been emphasized by Dickson et al. (2007). The omission of a time correction
3 results in unrealistic spikes in the difference between seawater temperature and the
4 temperature inside the equilibrator. Appropriate temperature correction will reduce
5 artificial variability in $p\text{CO}_2$, even though the average $p\text{CO}_2$ value is still likely to be
6 correct. In addition, both the underway seawater system and the equilibrators tend to
7 smooth out short-lived signals in temperature and $p\text{CO}_2$, because of the time the water
8 spent travelling along the underway water supply and inside the equilibrator. These delay-
9 times can affect the measurement reliability by smoothing out strong gradients. These
10 effects are particularly important in regions with rapid changes in carbonate chemistry and
11 sea water temperature, for example in shelf sea regions with freshwater inputs, in
12 continental shelf break regions with enhanced vertical mixing (internal tides or upwelling)
13 and in regions with sea ice melt. The effects of strong gradients on $p\text{CO}_2$ have been noted
14 by Körtzinger et al. (1996) and Körtzinger et al. (2000).

15 2. To obtain high quality $p\text{CO}_2$ datasets we recommend special care be taken with the
16 operation of the equilibrator systems, including a) careful control of the seawater supply
17 and the water flow through the equilibrator (Körtzinger et al., 1996); b) accurate
18 temperature readings (Körtzinger et al., 2000; Pierrot et al., 2009; Bakker et al., 2014) and
19 c) prevention of phototrophic growth in the equilibrator by complete shielding from light,
20 although non-phototrophic microbial growth will not be prevented. It should be noted that
21 shading of the equilibrator will not stop respiration in the pipes or equilibrator. The latter
22 can be calculated assuming a residence time of 12 minutes (2 minutes in the ship's pipe
23 network and 10 minutes in the equilibrator), community respiration of $54 \mu\text{mol L}^{-1} \text{d}^{-1}$ in
24 shelf sea waters (highest rate in Holligan et al. (1984)) and a respiratory quotient of 1.
25 This 'worst case' scenario results in the production of $0.4 \mu\text{mol L}^{-1}$ DIC which equates to
26 $0.75 \mu\text{atm } p\text{CO}_2$, using the average A_T , C_T , SST, salinity, Silicate and Phosphate
27 concentrations from D366 (calculated using CO2SYS). This value is clearly within the
28 uncertainty of our measurements. Nevertheless, regular cleaning of the equilibrator and
29 pipe network, where possible, is recommended in order to prevent the build-up of
30 microbial mats and associated respiration.

31 3. When comparing a pair of variables, it is important to examine the residuals as a function
32 of time as well as constructing a scatterplot of one variable against the other. For example,

1 from examination of Fig. 1 alone we would not have identified the diurnal variation
2 between the two $p\text{CO}_2$ systems (Fig. 2).

- 3 4. Differences between the recorded sampling time and the actual sampling time also need to
4 be taken into account. C_T and A_T are discrete measurements, while $p\text{CO}_2$ and pH_T are
5 near-continuous measurements. In addition to the transit times between the intake and the
6 instrument, $p\text{CO}_2$ is an integrated measurement over the timescale of equilibration (around
7 8 min) and pH_T is an integrated measurement over the timescale of filling the sample
8 chamber (ca. 60 seconds). When comparing carbonate chemistry datasets, corrections
9 should be made for the asynchronous times of sample collection for the different
10 variables.
- 11 5. It is recommended to characterise the extinction coefficients of each batch of pH-
12 indicator-dye on the instrument used for ship-board pH analysis, rather than use published
13 values (Clayton and Byrne, 1993; Zhang and Byrne, 1996; Hopkins et al., 2000; Mosley et
14 al., 2004; Gabriel et al., 2005; Liu et al., 2011). This is particularly important where the
15 indicator has not been purified (Yao et al., 2007) and where the detection system has a
16 wider optical bandwidth than that used in the literature to characterize the indicator (here
17 the optical bandwidth used for our pH system is 15-20 nm (Rérolle et al., 2013)). The
18 discrepancy between pH_T values calculated with our coefficients and values calculated
19 with coefficients from Zhang and Byrne (1996) was about 0.02 pH_T units (Rérolle et al.,
20 2013). Rérolle et al. (2013) estimated that about 0.005 pH_T units of the observed
21 discrepancy was due to impurities in the indicator and about 0.015 was due to the wider
22 bandpass detection window in our ship-board pH system. This recommendation has been
23 made before by Liu et al. (2011) and Yao et al. (2007).

25 **3.3 Ratio of total boron to salinity and carbonate constants**

26 The calculations of the carbonate chemistry variables were undertaken for a second time using
27 another ratio of total boron to salinity (Lee et al., 2010) for the combinations of pairs
28 involving A_T (because the ratio of total boron to salinity only influences calculations
29 involving A_T). We then compared the results obtained against those using the original ratio of
30 total boron to salinity from Uppström (1974). Statistically significant differences (ANOVA,
31 p-values < 0.001) were seen for both $p\text{CO}_2$ and pH_T calculated from C_T and A_T using the
32 different ratios of total boron to salinity. There were also significant differences between C_T

1 values calculated from A_T and pH_T and from A_T and pCO_2-2 using the different constants, and
2 A_T calculated from both C_T and pH_T and C_T and pCO_2-2 . There were, however, no statistically
3 significant differences when calculating pH_T from pCO_2-1 or pCO_2-2 and A_T , or pCO_2-1 and
4 pCO_2-2 calculated from pH_T and A_T (Fig. 6 and Table 2). This was because the calculation of
5 pH_T from pCO_2 and A_T is mainly driven by pCO_2 and not A_T (and the same when calculated
6 pCO_2 from pH_T and A_T). The residuals were on the whole smaller when using the ratio of total
7 boron to salinity from Uppström (1974) compared to using Lee et al. (2010) (Fig. 6). The
8 differences between mean residuals were $5 \mu atm$, $0.007 pH_T$ units, $2.9 \mu mol kg^{-1}$ and
9 $3.2 \mu mol kg^{-1}$ for pCO_2 , pH_T , C_T and A_T , respectively when subtracting results obtained
10 using the constants from Uppström (1974) from those obtained using the constants from Lee
11 et al. (2010). These discrepancies might not be so significant in an ocean acidification
12 context, but they are substantial in terms of air-sea flux calculations in coastal waters. Overall,
13 for our ranges of temperature and salinity, Uppström (1974) gave the best results when
14 compared to observed values for carbonate chemistry variables.

15 The calculations of the carbonate chemistry variables were undertaken for a third time using
16 other sets of carbonate constants (Roy et al., 1993; Lueker et al., 2000; Millero et al., 2006).
17 We then compared the results obtained against those using the original carbonate constants,
18 from Mehrbach and the total boron to salinity ratio of Uppström (1974). We omitted the pair
19 pH_T-pCO_2 from consideration because they are not a good pair for the calculations.

20 Statistically significant differences (ANOVA, p -values < 0.001) were noted when using the
21 Roy et al. (1993) constants (hereinafter Roy) to calculate pCO_2-1 and pCO_2-2 from C_T and A_T ,
22 C_T or pH_T or A_T and pH_T . The use of Lueker et al.'s (2000) constants (hereinafter Lueker)
23 produces results that do not yield statistically significant differences from Mehrbach when
24 calculating pCO_2-1 and pCO_2-2 from C_T and A_T , C_T or pH_T or A_T and pH_T . The use of Millero
25 et al.'s (2006) constant (hereinafter Millero) did not yield statistically significant differences
26 from Mehrbach when calculating pCO_2-2 from C_T and A_T , C_T or pH_T or A_T and pH_T . The
27 maximum mean residuals were always observed when using Roy's constants (up to $23 \mu atm$
28 when calculating pCO_2-2 from C_T and A_T) (Table 3). The calculated pCO_2 values using the
29 constants of Roy were significantly higher than those by Mehrbach in other studies
30 (Wanninkhof et al., 1999). McElligott et al. (1998) showed a good agreement between
31 measured pCO_2 and pCO_2 calculated from C_T and A_T using Mehrbach at the sea surface.

1 There were no significant differences when calculating pH_T from pCO_2-1 and either C_T or A_T .
2 However, there were statistically significant differences when using Roy for calculating pH_T
3 from pCO_2-2 and C_T or A_T . Of particular importance, the mean residual is significantly higher
4 when using Roy's constants to calculate pH_T from C_T and A_T (Table 3). McElligott et al.
5 (1998) noted that all four CO_2 parameters measured during the NOAA Equatorial Pacific CO_2
6 cruises were internally consistent when using the constants of Mehrbach et al. (1973) refitted
7 by Dickson and Millero (1987) at 20-25 °C, if spectrophotometric pH_T values were increased
8 by 0.0038 (value proposed by DelValls and Dickson (1998)).

9 For calculations of C_T and A_T , the constants by Roy led to a mean residual further from 0
10 (statistically significant) than using Mehrbach (large differences between paired values). The
11 reasons for the occasional significant differences between constants are not entirely clear and
12 investigating the discrepancies is beyond the scope of this work. Objective comparison of the
13 constants is difficult because of differences between the ranges of temperature and salinity
14 over which the constants were measured, and the curve fitting procedures differed
15 (Wanninkhof et al., 1999).

16 The constants by Mehrbach were determined in artificial seawater using the seawater pH scale
17 with a range in temperature between 2 and 35 °C and in salinity between 20 and 40. The
18 constants by Lueker were measured in natural seawater using the total pH scale with a range
19 in temperature of 2-35 °C and in salinity of 19-43. Lueker et al. (2000) converted the
20 constants by Mehrbach to the total hydrogen ion pH scale. The constants by Roy were
21 measured in artificial seawater using the total pH scale with a range in temperature of 0-45 °C
22 and in salinity of 5-45. The constants by Roy have been used for studies in the polar waters
23 (Jutterström and Anderson, 2005; Chierici et al., 2011). The constants by Millero were
24 measured in natural seawater using the seawater pH scale with a range in temperature of 0-
25 50 °C and in salinity of 1-50. Dickson et al. (2007) reported that the constants by Lueker were
26 in reasonable agreement with those reported by Roy et al. (1993). In our study this did not
27 appear to be the case.

28 Overall, the use of constants by Lueker gave similar results to those of Mehrbach (perhaps
29 unsurprisingly because they are refitted from constants by Mehrbach) but led to larger mean
30 residuals. Although using constants by Millero did not usually lead to statistically differences
31 compared to Mehrbach, the residuals were overall higher than when using either the constants
32 by Mehrbach or Lueker. Our dataset confirms that the constants by Mehrbach provided the

1 best quality results as previously demonstrated by other studies examining the internal
2 consistency in the laboratory (Lee et al., 1996; Lueker et al., 2000) and in the field
3 (Wanninkhof et al., 1999; Lee et al., 2000).

4

5 **4 Conclusions**

6 Our results show that it is possible to obtain good consistency between measurements of
7 different variables of the carbonate system, even outside the somewhat artificial conditions of
8 an intercomparison exercise. However, our retrospective intercomparison revealed several
9 sources of discrepancies, leading to the following recommendation for best practice: a)
10 undertake characterization of the pH-indicator-dye in order to obtain correct extinction
11 coefficients for the dye and analytical system in the temperature and salinity range used; b)
12 take into account the transit time of seawater from the intake to the equilibrator, when
13 comparing $p\text{CO}_2$ with other simultaneous measurements; c) examine residuals as a function of
14 time in order to detect temporal biases in measurements; d) prevent phototrophic growth in
15 $p\text{CO}_2$ equilibrators by completely shielding them from exposure to light.

16 We obtained smaller average residuals when using the ratio of total boron to salinity from
17 Uppström (1974) rather than Lee et al. (2010) and when using the carbonate constants by
18 Mehrbach et al. (1973) (refitted by Dickson and Millero (1987)) rather than Roy et al.'s
19 (1993), Lueker et al.'s (2000) or Millero et al.'s (2006). As found in other studies (Millero,
20 1995; Cullison Gray et al., 2011), the variables pH_T and $p\text{CO}_2$ are far from an ideal pair for
21 calculation of C_T or A_T , emphasizing the desirability of developing a C_T or A_T sensor capable
22 of autonomous high resolution measurements.

23

24 **Acknowledgements**

25 This work is a contribution to the UK Ocean Acidification Research Programme
26 (UKOA) which is jointly funded by the Department for Environment, Food and Rural Affairs
27 (DEFRA), the Natural Environment Research Council (NERC) and the Department for
28 Energy and Climate Change (DECC) under grant agreement no. NE/H017348/1. We thank
29 the captain and crew of RRS *Discovery* for assistance during the cruise and N. Wager for her
30 help with e-folding times. N. Greenwood, J.-P. Gattuso and two anonymous reviewers
31 provided valuable comments on earlier versions of this work.

1 **References**

- 2 Bakker, D. C. E., Nielsdóttir, M. C., Morris, P. J., Venables, H. J., and Watson, A. J.: The
3 island mass effect and biological carbon uptake for the subantarctic Crozet Archipelago,
4 *Deep-Sea Res Pt. II*, 54, 2174-2190, doi: 10.1016/j.dsr2.2007.06.009, 2007.
- 5 Bakker, D. C. E., Pfeil, B., Smith, K., Hankin, S., Olsen, A., Alin, S. R., Cosca, C., Harasawa,
6 S., Kozyr, A., Nojiri, Y., O'Brien, K. M., Schuster, U., Telszewski, M., Tilbrook, B., Wada,
7 C., Akl, J., Barbero, L., Bates, N. R., Boutin, J., Bozec, Y., Cai, W. J., Castle, R. D., Chavez,
8 F. P., Chen, L., Chierici, M., Currie, K., de Baar, H. J. W., Evans, W., Feely, R. A., Fransson,
9 A., Gao, Z., Hales, B., Hardman-Mountford, N. J., Hoppema, M., Huang, W. J., Hunt, C. W.,
10 Huss, B., Ichikawa, T., Johannessen, T., Jones, E. M., Jones, S. D., Jutterström, S., Kitidis, V.,
11 Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Manke, A. B., Mathis, J. T.,
12 Merlivat, L., Metzl, N., Murata, A., Newberger, T., Omar, A. M., Ono, T., Park, G. H.,
13 Paterson, K., Pierrot, D., Ríos, A. F., Sabine, C. L., Saito, S., Salisbury, J., Sarma, V. V. S. S.,
14 Schlitzer, R., Sieger, R., Skjelvan, I., Steinhoff, T., Sullivan, K. F., Sun, H., Sutton, A. J.,
15 Suzuki, T., Sweeney, C., Takahashi, T., Tjiputra, J., Tsurushima, N., van Heuven, S. M. A.
16 C., Vandemark, D., Vlahos, P., Wallace, D. W. R., Wanninkhof, R., and Watson, A. J.: An
17 update to the Surface Ocean CO₂ Atlas (SOCAT version 2), *Earth Syst. Sci. Data*, 6, 69-90,
18 10.5194/essd-6-69-2014, 2014.
- 19 Bozec, Y., Thomas, H., Elkalay, K., and de Baar, H. J.: The continental shelf pump for CO₂ in
20 the North Sea—evidence from summer observation, *Mar. Chem.*, 93, 131-147, 2005.
- 21 Chierici, M., Fransson, A., Lansard, B., Miller, L. A., Mucci, A., Shadwick, E., Thomas, H.,
22 Tremblay, J. E., and Papakyriakou, T. N.: Impact of biogeochemical processes and
23 environmental factors on the calcium carbonate saturation state in the Circumpolar Flaw Lead
24 in the Amundsen Gulf, Arctic Ocean, *Journal of Geophysical Research: Oceans (1978–2012)*,
25 116, 2011.
- 26 Clayton, T. D., and Byrne, R. H.: Spectrophotometric seawater pH measurements: total
27 hydrogen ion concentration scale calibration of *m*-cresol purple and at sea results, *Deep-Sea*
28 *Res. Pt. I*, 40, 2115-2129, doi: 10.1016/0967-0637(93)90048-8, 1993.
- 29 Clayton, T. D., Byrne, R. H., Breland, J. A., Feely, R. A., Millero, F. J., Campbell, D. M.,
30 Murphy, P. P., and Lamb, M. F.: The role of pH measurements in modern oceanic CO₂-

- 1 system characterizations: Precision and thermodynamic consistency, *Deep-Sea Res Pt. II*, 42,
2 411-429, 1995.
- 3 Cullison Gray, S. E., DeGrandpre, M. D., Moore, T. S., Martz, T. R., Friederich, G. E., and
4 Johnson, K. S.: Applications of *in situ* pH measurements for inorganic carbon calculations,
5 *Mar. Chem.*, 125, 82-90, 2011.
- 6 Dickson, A. G.: An exact definition of total alkalinity and a procedure for the estimation of
7 alkalinity and total inorganic carbon from titration data, *Deep-Sea Res.*, 28, 609-623, 1981.
- 8 Dickson, A. G., and Millero, F. J.: A comparison of the equilibrium constants for the
9 dissociation of carbonic acid in seawater media, *Deep-Sea Res.*, 34, 1733-1743, doi:
10 10.1016/0198-0149(87)90021-5, 1987.
- 11 Dickson, A. G.: Standard potential of the reaction $\text{AgCl(s)} + 1/2\text{H}_2(\text{g}) = \text{Ag(s)} + \text{HCl(aq)}$ and the
12 standard acidity constant of the bisulfate ion in synthetic seawater from 273.15 K to 318.15
13 K, *J. Chem. Thermodyn.*, 22, 113-127, doi: 10.1016/0021-9614(90)90074-z, 1990a.
- 14 Dickson, A. G.: Thermodynamics of the dissociation of boric acid in synthetic seawater from
15 273.15 to 318.15 K, *Deep Sea Research Part A. Oceanographic Research Papers*, 37, 755-766,
16 1990b.
- 17 Dickson, A. G., Sabine, C. L., and Christian, J. R.: Guide to best practices for ocean CO₂
18 measurements, *PICES Special Publication*, 3, 2007.
- 19 Dickson, A. G.: Interactive comment on "On CO₂ perturbation experiments: over-
20 determination of carbonate chemistry reveals inconsistencies" by Hoppe, C. J. M. Langer, G.,
21 Rokitta, S.D., Wolf-Gladrow, D.A. and Rost, B., *Biogeosciences Disc.*, 7, C1286-C1293,
22 2010.
- 23 Gabriel, M., Forja, J., Rubio, J., and Gómez-Parra, A.: Temperature and salinity dependence
24 of molar absorptivities of thymol blue: Application to the spectrophotometric determination
25 of pH in estuarine waters, *Cienc. Mar.*, 31, 309-318, 2005.
- 26 Hardman-Mountford, N. J., Moore, G., Bakker, D. C., Watson, A. J., Schuster, U., Barciela,
27 R., Hines, A., Moncoiffé, G., Brown, J., Dye, S., Blackford, J., Somerfield, P. J., Holt, J.,

- 1 Hydes, D. J., and Aiken, J.: An operational monitoring system to provide indicators of CO₂-
2 related variables in the ocean, *ICES J. Mar. Sci.*, 65, 1498-1503, 2008.
- 3 Hernández-Ayón, J. M., Zirino, A., Dickson, A. G., Camiro-Vargas, T., and Valenzuela-
4 Espinoza, E.: Estimating the contribution of organic bases from microalgae to the titration
5 alkalinity in coastal seawaters, *Limnology and Oceanography-Methods*, 5, 225-232, 2007.
- 6 Hopkins, A. E., Sell, K. S., Soli, A. L., and Byrne, R. H.: In-situ spectrophotometric pH
7 measurements: The effect of pressure on thymol blue protonation and absorbance
8 characteristics, *Mar. Chem.*, 71, 103-109, 2000.
- 9 IOCCP: Ocean surface *p*CO₂, data integration and database development, 14-17 January
10 2004, IOCCP report 2, IOC Worskshop Report 211, available at:
11 http://www.ioccp.org/images/01Underway/WR211_eo.pdf (last access 28 April 2014),
12 UNESCO Tsukuba, Japan, 2004.
- 13 Johnson, K. M., Sieburth, J. M., Williams, P. J. L., and Brandstrom, L.: Coulometric total
14 carbon-dioxide analysis for marine studies - Automation and calibration, *Mar. Chem.*, 21,
15 117-133, 1987.
- 16 Johnson, K. M., Körtzinger, A., Mintrop, L., Duinker, J. C., and Wallace, D. W. R.:
17 Coulometric total carbon dioxide analysis for marine studies: measurement and internal
18 consistency of underway TCO₂ concentrations, *Mar. Chem.*, 67, 123-144, doi:
19 10.1016/s0304-4203(99)00055-9, 1999.
- 20 Juranek, L., Feely, R., Peterson, W., Alin, S., Hales, B., Lee, K., Sabine, C., and Peterson, J.:
21 A novel method for determination of aragonite saturation state on the continental shelf of
22 central Oregon using multi-parameter relationships with hydrographic data, *Geophys. Res.*
23 *Lett.*, 36, L24601, doi:10.1029/2009GL040778, 2009.
- 24 Jutterström, S., and Anderson, L. G.: The saturation of calcite and aragonite in the Arctic
25 Ocean, *Mar. Chem.*, 94, 101-110, <http://dx.doi.org/10.1016/j.marchem.2004.08.010>, 2005.
- 26 Katayama, H., Karasudani, T., Ishii, K., Marubayashi, K., and Ueda, H.: Development of a
27 gas-liquid equilibrator for estimating CO₂ flux at the ocean surface, *Journal of Atmospheric*
28 *and Oceanic Technology*, 16, 1450-1455, 1999.

- 1 Kim, H.-C., Lee, K., and Choi, W.: Contribution of phytoplankton and bacterial cells to the
2 measured alkalinity of seawater, *Limnol. Oceanogr.*, 51, 331-338, 2006.
- 3 Kim, H. C., and Lee, K.: Significant contribution of dissolved organic matter to seawater
4 alkalinity, *Geophys. Res. Lett.*, 36, 2009.
- 5 Kirkwood, D.: Simultaneous determination of selected nutrients in sea water, International
6 Council for the Exploration of the Sea (ICES) CM, 100, 29, 1989.
- 7 Kitidis, V., Hardman-Mountford, N. J., Litt, E., Brown, I., Cummings, D., Hartman, S.,
8 Hydes, D., Fishwick, J. R., Harris, C., Martinez-Vicente, V., Woodward, E. M. S., and Smyth,
9 T. J.: Seasonal dynamics of the carbonate system in the Western English Channel, *Cont. Shelf*
10 *Res.*, 42, 30-40, doi: 10.1016/j.csr.2012.04.012, 2012.
- 11 Körtzinger, A., Thomas, H., Schneider, B., Gronau, N., Mintrop, L., and Duinker, J. C.: At-
12 sea intercomparison of two newly designed underway $p\text{CO}_2$ systems—encouraging results,
13 *Mar. Chem.*, 52, 133-145, 1996.
- 14 Körtzinger, A., Mintrop, L., Wallace, D. W. R., Johnson, K. M., Neill, C., Tilbrook, B.,
15 Towler, P., Inoue, H. Y., Ishii, M., Shaffer, G., Saavedra, R. F. T., Ohtaki, E., Yamashita, E.,
16 Poisson, A., Brunet, C., Schauer, B., Goyet, C., and Eiseheid, G.: The international at-sea
17 intercomparison of $f\text{CO}_2$ systems during the R/V *Meteor* Cruise 36/1 in the North Atlantic
18 Ocean, *Mar. Chem.*, 72, 171-192, 2000.
- 19 Lamb, M. F., Sabine, C. L., Feely, R. A., Wanninkhof, R., Key, R. M., Johnson, G. C.,
20 Millero, F. J., Lee, K., Peng, T. H., Kozyr, A., Bullister, J. L., Greeley, D., Byrne, R. H.,
21 Chipman, D. W., Dickson, A. G., Goyet, C., Guenther, P. R., Ishii, M., Johnson, K. M.,
22 Keeling, C. D., Ono, T., Shitashima, K., Tilbrook, B., Takahashi, T., Wallace, D. W. R.,
23 Watanabe, Y. W., Winn, C., and Wong, C. S.: Consistency and synthesis of Pacific Ocean
24 CO_2 survey data, *Deep-Sea Res Pt. II*, 49, 21-58, doi: 10.1016/S0967-0645(01)00093-5, 2001.
- 25 Lee, K., Millero, F. J., and Campbell, D. M.: The reliability of the thermodynamic constants
26 for the dissociation of carbonic acid in seawater, *Mar. Chem.*, 55, 233-245, 10.1016/s0304-
27 4203(96)00064-3, 1996.

- 1 Lee, K., Millero, F. J., and Wanninkhof, R.: The carbon dioxide system in the Atlantic Ocean,
2 J Geophys Res-Oceans, 102, 15693-15707, 10.1029/97jc00067, 1997.
- 3 Lee, K., Millero, F. J., Byrne, R. H., Feely, R. A., and Wanninkhof, R.: The recommended
4 dissociation constants for carbonic acid in seawater, Geophys. Res. Lett., 27, 229-232, 2000.
- 5 Lee, K., Kim, T. W., Byrne, R. H., Millero, F. J., Feely, R. A., and Liu, Y. M.: The universal
6 ratio of boron to chlorinity for the North Pacific and North Atlantic oceans, Geochim.
7 Cosmochim. Acta, 74, 1801-1811, doi: 10.1016/j.gca.2009.12.027, 2010.
- 8 Legendre, P.: lmodel2: Model II Regression, R package version 1.7-2, [http://CRAN.R-](http://CRAN.R-project.org/package=lmodel2)
9 [project.org/package=lmodel2](http://CRAN.R-project.org/package=lmodel2), 2014.
- 10 Lewis, E., and Wallace, D. W. R.: Program developed for CO₂ system calculations.
11 ORNL/CDIAC-105, Carbon Dioxide Information Analysis Center, Oak Ridge National
12 Laboratory, U.S., Department of Energy, Oak Ridge, Tennessee, 1998.
- 13 Liu, X., Patsavas, M. C., and Byrne, R. H.: Purification and characterization of meta-cresol
14 purple for spectrophotometric seawater pH measurements, Environ. Sci. Technol., 45, 4862-
15 4868, 2011.
- 16 Loucaides, S., Tyrrell, T., Achterberg, E. P., Torres, R., Nightingale, P. D., Kitidis, V., Serret,
17 P., Woodward, M., and Robinson, C.: Biological and physical forcing of carbonate chemistry
18 in an upwelling filament off northwest Africa: Results from a Lagrangian study, Global
19 Biogeochem. Cy., 26, 2012.
- 20 Lueker, T. J., Dickson, A. G., and Keeling, C. D.: Ocean *p*CO₂ calculated from dissolved
21 inorganic carbon, alkalinity, and equations for K₁ and K₂: Validation based on laboratory
22 measurements of CO₂ in gas and seawater at equilibrium, Mar. Chem., 70, 105-119, 2000.
- 23 McElligott, S., Byrne, R., Lee, K., Wanninkhof, R., Millero, F., and Feely, R.: Discrete water
24 column measurements of CO₂ fugacity and pH_T in seawater: A comparison of direct
25 measurements and thermodynamic calculations, Mar. Chem., 60, 63-73, 1998.
- 26 Mehrbach, C., Culberso, C. H., Hawley, J. E., and Pytkowic, R. M.: Measurement of apparent
27 dissociation constants of carbonic acid in seawater at atmospheric pressure, Limnol.
28 Oceanogr., 18, 897-907, 1973.

1 Millero, F. J.: Thermodynamics of the carbon dioxide system in the oceans, *Geochim.*
2 *Cosmochim. Acta*, 59, 661-677, [http://dx.doi.org/10.1016/0016-7037\(94\)00354-O](http://dx.doi.org/10.1016/0016-7037(94)00354-O), 1995.

3 Millero, F. J., Pierrot, D., Lee, K., Wanninkhof, R., Feely, R., Sabine, C. L., Key, R. M., and
4 Takahashi, T.: Dissociation constants for carbonic acid determined from field measurements,
5 *Deep-Sea Res. Pt. I*, 49, 1705-1723, 2002.

6 Millero, F. J., Graham, T. B., Huang, F., Bustos-Serrano, H., and Pierrot, D.: Dissociation
7 constants of carbonic acid in seawater as a function of salinity and temperature, *Mar. Chem.*,
8 100, 80-94, doi: 10.1016/j.marchem.2005.12.001, 2006.

9 Millero, F. J.: The marine inorganic carbon cycle, *Chem. Rev.*, 107, 308-341, doi:
10 10.1021/cr0503557, 2007.

11 Mosley, L. M., Husheer, S. L., and Hunter, K. A.: Spectrophotometric pH measurement in
12 estuaries using thymol blue and *m*-cresol purple, *Mar. Chem.*, 91, 175-186, 2004.

13 Pfeil, B., Olsen, A., Bakker, D. C. E., Hankin, S., Koyuk, H., Kozyr, A., Malczyk, J., Manke,
14 A., Metzl, N., Sabine, C. L., Akl, J., Alin, S. R., Bates, N., Bellerby, R. G. J., Borges, A.,
15 Boutin, J., Brown, P. J., Cai, W. J., Chavez, F. P., Chen, A., Cosca, C., Fassbender, A. J.,
16 Feely, R. A., González-Dávila, M., Goyet, C., Hales, B., Hardman-Mountford, N., Heinze, C.,
17 Hood, M., Hoppema, M., Hunt, C. W., Hydes, D., Ishii, M., Johannessen, T., Jones, S. D.,
18 Key, R. M., Körtzinger, A., Landschützer, P., Lauvset, S. K., Lefèvre, N., Lenton, A.,
19 Lourantou, A., Merlivat, L., Midorikawa, T., Mintrop, L., Miyazaki, C., Murata, A.,
20 Nakadate, A., Nakano, Y., Nakaoka, S., Nojiri, Y., Omar, A. M., Padin, X. A., Park, G. H.,
21 Paterson, K., Perez, F. F., Pierrot, D., Poisson, A., Ríos, A. F., Santana-Casiano, J. M.,
22 Salisbury, J., Sarma, V. V. S. S., Schlitzer, R., Schneider, B., Schuster, U., Sieger, R.,
23 Skjelvan, I., Steinhoff, T., Suzuki, T., Takahashi, T., Tedesco, K., Telszewski, M., Thomas,
24 H., Tilbrook, B., Tjiputra, J., Vandemark, D., Veness, T., Wanninkhof, R., Watson, A. J.,
25 Weiss, R., Wong, C. S., and Yoshikawa-Inoue, H.: A uniform, quality controlled Surface
26 Ocean CO₂ Atlas (SOCAT), *Earth Syst. Sci. Data*, 5, 125-143, 10.5194/essd-5-125-2013,
27 2013.

- 1 Pierrot, D., Neill, C., Sullivan, K., Castle, R., Wanninkhof, R., Lüger, H., Johannessen, T.,
2 Olsen, A., Feely, R. A., and Cosca, C. E.: Recommendations for autonomous underway $p\text{CO}_2$
3 measuring systems and data-reduction routines, *Deep-Sea Res Pt. II*, 56, 512-522, 2009.
- 4 Poulton, A. J., Stinchcombe, M. C., Achterberg, E. P., Bakker, D. C. E., Dumousseaud, C.,
5 Lawson, H. E., Lee, G. A., Richier, S., Suggett, D. J., and Young, J. R.: Coccolithophores on
6 the north-west European shelf: calcification rates and environmental controls, *Biogeosciences*
7 *Discuss.*, 11, 2685-2733, 10.5194/bgd-11-2685-2014, 2014.
- 8 Quinn, G. P., and Keough, M. J.: *Experimental design and data analysis for biologists*,
9 Cambridge University Press, 2002.
- 10 Rérolle, V., Floquet, C. F., Harris, A. J., Mowlem, M. C., Bellerby, R. R., and Achterberg, E.
11 P.: Development of a colorimetric microfluidic pH sensor for autonomous seawater
12 measurements, *Anal. Chim. Acta*, 786, 124-131, 2013.
- 13 Rérolle, V. C. M., Ribas-Ribas, M., Kitidis, V., Brown, I., Bakker, D. C. E., Lee, G. A., Shi,
14 T., Mowlem, M. C., and Achterberg, E. P.: Controls on pH in surface waters of northwestern
15 European shelf seas, *Biogeosciences Disc.*, 11, 943-974, doi: 10.5194/bgd-11-943-2014,
16 2014.
- 17 Rérolle, V. M. C., Floquet, C. F. A., Mowlem, M. C., Bellerby, R., Connelly, D. P., and
18 Achterberg, E. P.: Seawater-pH measurements for ocean-acidification observations, *Trends*
19 *Anal. Chem.*, 40, 146-157, doi: 10.1016/j.trac.2012.07.016, 2012.
- 20 Roy, R. N., Roy, L. N., Vogel, K. M., Porter-Moore, C., Pearson, T., Good, C. E., Millero, F.
21 J., and Campbell, D. M.: The dissociation constants of carbonic acid in seawater at salinities 5
22 to 45 and temperatures 0 to 45 C, *Mar. Chem.*, 44, 249-267, 1993.
- 23 Schuster, U., and Watson, A. J.: A variable and decreasing sink for atmospheric CO_2 in the
24 North Atlantic, *J Geophys Res-Oceans*, 112, C11006, doi: 10.1029/2006jc003941, 2007.
- 25 Takahashi, T., Olafsson, J., Goddard, J. G., Chipman, D. W., and Sutherland, S. C.: Seasonal
26 variation of CO_2 and nutrients in the high-latitude surface oceans- A comparative study.,
27 *Global Biogeochem. Cy.*, 7, 843-878, doi: 10.1029/93GB02263, 1993.

- 1 Uppström, L. R.: The boron/chlorinity ratio of deep-sea water from the Pacific Ocean, Deep
2 Sea Research and Oceanographic Abstracts, 21, 161-162, 1974.
- 3 Van Heuven, S., Pierrot, D., Rae, J. W. B., Lewis, E., and Wallace, D. W. R.: CO2SYS v 1.1,
4 MATLAB program developed for CO₂ system calculations, ORNL/CDIAC-105b. Carbon
5 Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. DoE, Oak Ridge,
6 TN., 2011.
- 7 Wanninkhof, R., Lewis, E., Feely, R. A., and Millero, F. J.: The optimal carbonate
8 dissociation constants for determining surface water *p*CO₂ from alkalinity and total inorganic
9 carbon, Mar. Chem., 65, 291-301, 1999.
- 10 Wanninkhof, R., Bakker, D. C. E., Bates, N., Olsen, A., Steinhoff, T., and Sutton, A. J.:
11 Incorporation of Alternative Sensors in the SOCAT Database and Adjustments to Dataset
12 Quality Control Flags. <http://cdiac.ornl.gov/oceans/Recommendationnewsensors.pdf>, Carbon
13 Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of
14 Energy, 10.3334/CDIAC/OTG.SOCAT_ADQCF, 2013.
- 15 Weiss, R. F., and Price, B. A.: Nitrous-oxide solubility in water and seawater, Mar. Chem., 8,
16 347-359, doi: 10.1016/0304-4203(80)90024-9, 1980.
- 17 Yao, W., Liu, X., and Byrne, R. H.: Impurities in indicators used for spectrophotometric
18 seawater pH measurements: Assessment and remedies, Mar. Chem., 107, 167-172, 2007.
- 19 Young, J. R., Poulton, A. J., and Tyrrell, T.: Morphology of *Emiliana huxleyi* coccoliths on
20 the North West European shelf – is there an influence of carbonate chemistry?,
21 Biogeosciences Discuss., 11, 4531-4561, 10.5194/bgd-11-4531-2014, 2014.
- 22 Zhang, H. N., and Byrne, R. H.: Spectrophotometric pH measurements of surface seawater at
23 in-situ conditions: Absorbance and protonation behavior of thymol blue, Mar. Chem., 52, 17-
24 25, 1996.
- 25
- 26

1 Table 1. Results of comparisons between direct measurements and values calculated (using
2 the software CO2SYS) from measurements of other variables. r is Pearson's correlation
3 coefficient, df is degrees of freedom, RMSE is root mean square error and MR is mean of the
4 Residuals. The measurement accuracy is from Sect. 2.1. Note that calculated $p\text{CO}_2\text{-1}$ and
5 calculated $p\text{CO}_2\text{-2}$ have the same uncertainty because they depend only on the accuracies of
6 the variables from which they are calculated. They are therefore identical for both $p\text{CO}_2$
7 systems. Calculated uncertainty of parameters calculated using $p\text{CO}_2\text{-2}$ as an input variable is
8 higher than those using $p\text{CO}_2\text{-1}$, because the measurement accuracy is higher.

9

Measured variable	Input variables for calculation	Comparison statistics				Benchmarks	
		r	df	RMSE*	MR*	Calculated Uncertainty*	Measurement accuracy*
$p\text{CO}_2\text{-1}$ (μatm)	C_T and pH_T	0.984	31	5	3	4	
	A_T and pH_T	0.990	27	6	2	4	4
	C_T and A_T	0.987	43	6	3	4	
$p\text{CO}_2\text{-2}$ (μatm)	C_T and pH_T	0.949	97	11	4	4	
	A_T and pH_T	0.947	93	11	4	4	10
	C_T and A_T	0.937	156	12	1	4	
pH_T	C_T and A_T	0.952	218	0.008	0.001	0.005	
	A_T and $p\text{CO}_2\text{-1}$	0.991	27	0.006	0.002	0.003	
	A_T and $p\text{CO}_2\text{-2}$	0.951	93	0.013	0.004	0.011	0.004
	C_T and $p\text{CO}_2\text{-1}$	0.984	31	0.006	0.004	0.004	
	C_T and $p\text{CO}_2\text{-2}$	0.943	97	0.014	0.004	0.013	
A_T ($\mu\text{mol kg}^{-1}$)	C_T and pH_T	0.996	218	4.4	-0.2	3.1	
	C_T and $p\text{CO}_2\text{-1}$	0.997	43	4.6	2.1	3.0	
	C_T and $p\text{CO}_2\text{-2}$	0.991	156	8.6	0.4	7.0	1.5
	pH_T and $p\text{CO}_2\text{-1}$	0.802	27	36.3	-12.8	35.5	
	pH_T and $p\text{CO}_2\text{-2}$	0.596	93	75.6	-23.1	72.8	
C_T ($\mu\text{mol kg}^{-1}$)	A_T and pH_T	0.994	218	4.0	0.2	2.4	
	A_T and $p\text{CO}_2\text{-1}$	0.997	43	3.9	-1.7	2.5	
	A_T and $p\text{CO}_2\text{-2}$	0.989	156	7.2	-0.3	5.8	2.0
	pH_T and $p\text{CO}_2\text{-1}$	0.680	31	32.2	-18.8	31.9	
	pH_T and $p\text{CO}_2\text{-2}$	0.528	97	69.8	-21.9	63.4	

11

1 *These columns have the same units as the first column.

2

3 Table 2. Directly measured pH_T , partial pressure of CO_2 (pCO_2-1), dissolved inorganic
 4 carbon (C_T) and total alkalinity (A_T) are compared to values calculated for the combinations
 5 of pairs involving A_T using different ratios of total boron to salinity. RMSE is root mean
 6 square error and MR is mean of the residuals. The two different ratios of total boron to
 7 salinity are from Lee et al. (2010) and Uppström (1974). The carbonate constants are from
 8 Mehrbach et al. (1973) refitted by Dickson and Millero (1987).

9

Measured variable	Ratio of total boron to salinity:	Lee		Uppström	
	Calculated variable from	RMSE*	MR*	RMSE*	MR*
pCO_2-1 (μatm)	A_T and pH_T	5	1	6	2
	C_T and A_T	10	8	6	3
pCO_2-2 (μatm)	A_T and pH_T	11	3	11	4
	C_T and A_T	14	6	12	1
pH_T	C_T and A_T	0.012	- 0.005	0.008	0.001
	A_T and pCO_2-1	0.006	0.002	0.006	0.002
	A_T and pCO_2-2	0.013	0.003	0.013	0.004
A_T ($\mu mol kg^{-1}$)	C_T and pH_T	6.6	2.9	4.4	-0.2
	C_T and pCO_2-1	7.3	5.3	4.6	2.1
	C_T and pCO_2-2	10.7	3.6	8.6	0.4
C_T ($\mu mol kg^{-1}$)	A_T and pH_T	6.1	-2.7	4.0	0.2
	A_T and pCO_2-1	6.1	-4.4	3.9	-1.7
	A_T and pCO_2-2	8.9	-3.0	7.2	-0.3

10

11 * These columns have the same units as the first column.

12

13 Table 3. Comparison of directly measured pH_T , partial pressure of CO_2 (pCO_2), dissolved
 14 inorganic carbon (C_T) and total alkalinity (A_T) to values calculated using different carbonate
 15 constants. RMSE is root mean square error and MR is mean of the residuals. The two

1 different carbonate constants are from Roy et al. (1993), (Millero et al., 2006), Mehrbach et
2 al. (1973) refitted by Dickson and Millero (1987) and Lueker et al. (2000). The ratio of total
3 boron to salinity is from Uppström (1974).

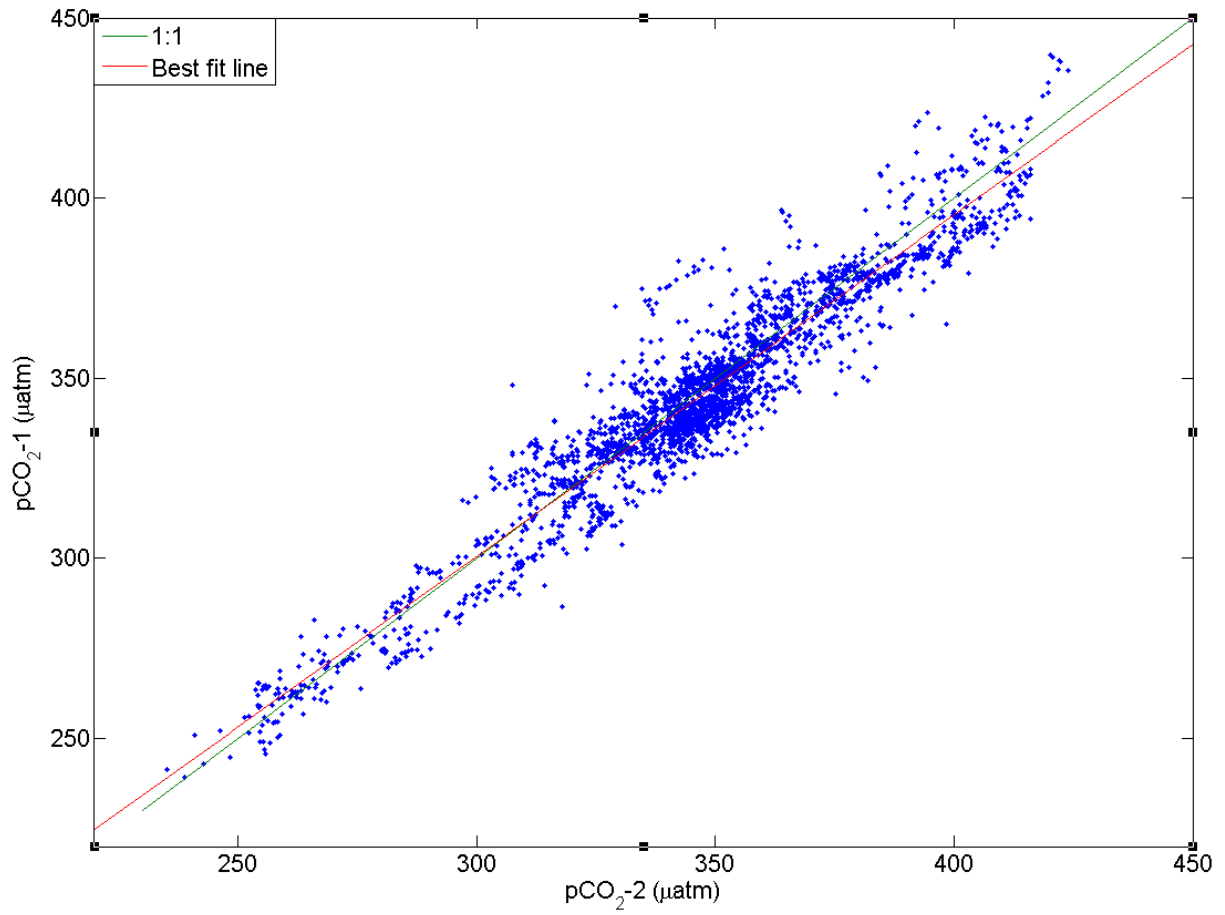
4

Carbonate constants:		Mehrbach		Roy		Millero		Lueker	
Measured variable	Calculated variable from	RMSE*	MR*	RMSE*	MR*	RMSE*	MR*	RMSE*	MR*
$p\text{CO}_2\text{-1}$ (μatm)	C_T and pH_T	5	3	10	11	8	8	6	4
	A_T and pH_T	6	2	9	8	8	7	6	3
	C_T and A_T	6	3	26	25	8	5	6	3
$p\text{CO}_2\text{-2}$ (μatm)	C_T and pH_T	11	4	14	11	13	9	12	5
	A_T and pH_T	11	4	14	10	13	9	12	5
	C_T and A_T	12	1	28	23	13	3	12	0
pH_T	C_T and A_T	0.008	0.001	0.022	-0.016	0.007	0.004	0.007	0.002
	A_T and $p\text{CO}_2\text{-1}$	0.006	0.002	0.010	0.008	0.009	0.007	0.007	0.003
	C_T and $p\text{CO}_2\text{-1}$	0.006	0.004	0.012	0.012	0.009	0.009	0.007	0.005
	A_T and $p\text{CO}_2\text{-2}$	0.013	0.004	0.015	0.010	0.015	0.009	0.013	0.005
	C_T and $p\text{CO}_2\text{-2}$	0.014	0.004	0.018	0.013	0.016	0.010	0.014	0.006
A_T ($\mu\text{mol kg}^{-1}$)	C_T and pH_T	4.4	-0.2	12.4	9.0	3.9	-1.8	4.0	-1.1
	C_T and $p\text{CO}_2\text{-1}$	4.6	2.1	17.8	16.3	5.8	3.5	4.4	1.8
	C_T and $p\text{CO}_2\text{-2}$	8.6	0.4	20.0	14.5	9.6	1.8	8.5	0.1
C_T ($\mu\text{mol kg}^{-1}$)	A_T and pH_T	4.0	0.2	11.4	-8.3	3.6	1.6	3.7	1.0
	A_T and $p\text{CO}_2\text{-1}$	3.9	-1.7	14.8	-13.6	4.9	-3.0	3.7	-1.5
	A_T and $p\text{CO}_2\text{-2}$	7.2	-0.3	16.5	-12.1	8.0	-1.5	7.1	-0.1

1

2 * These columns have the same units as the first column.

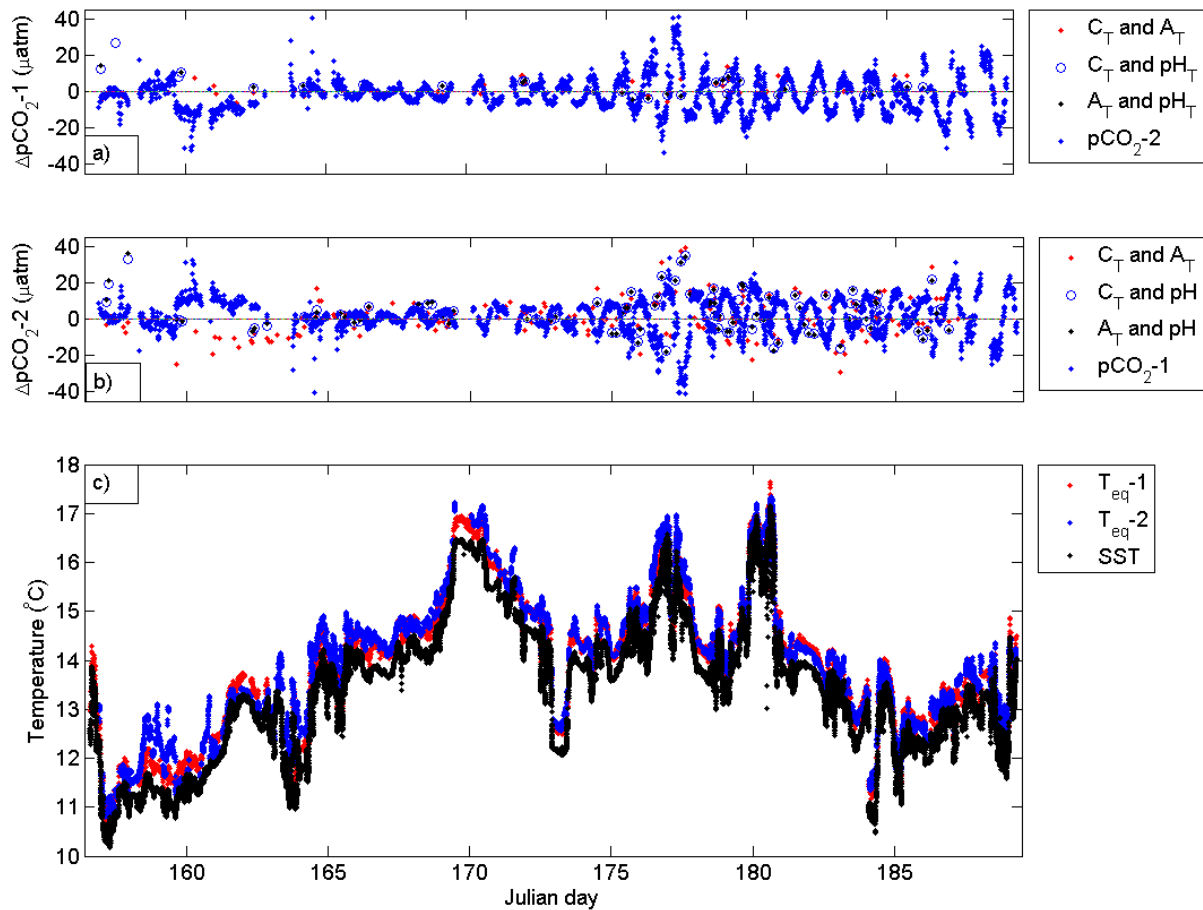
1



2

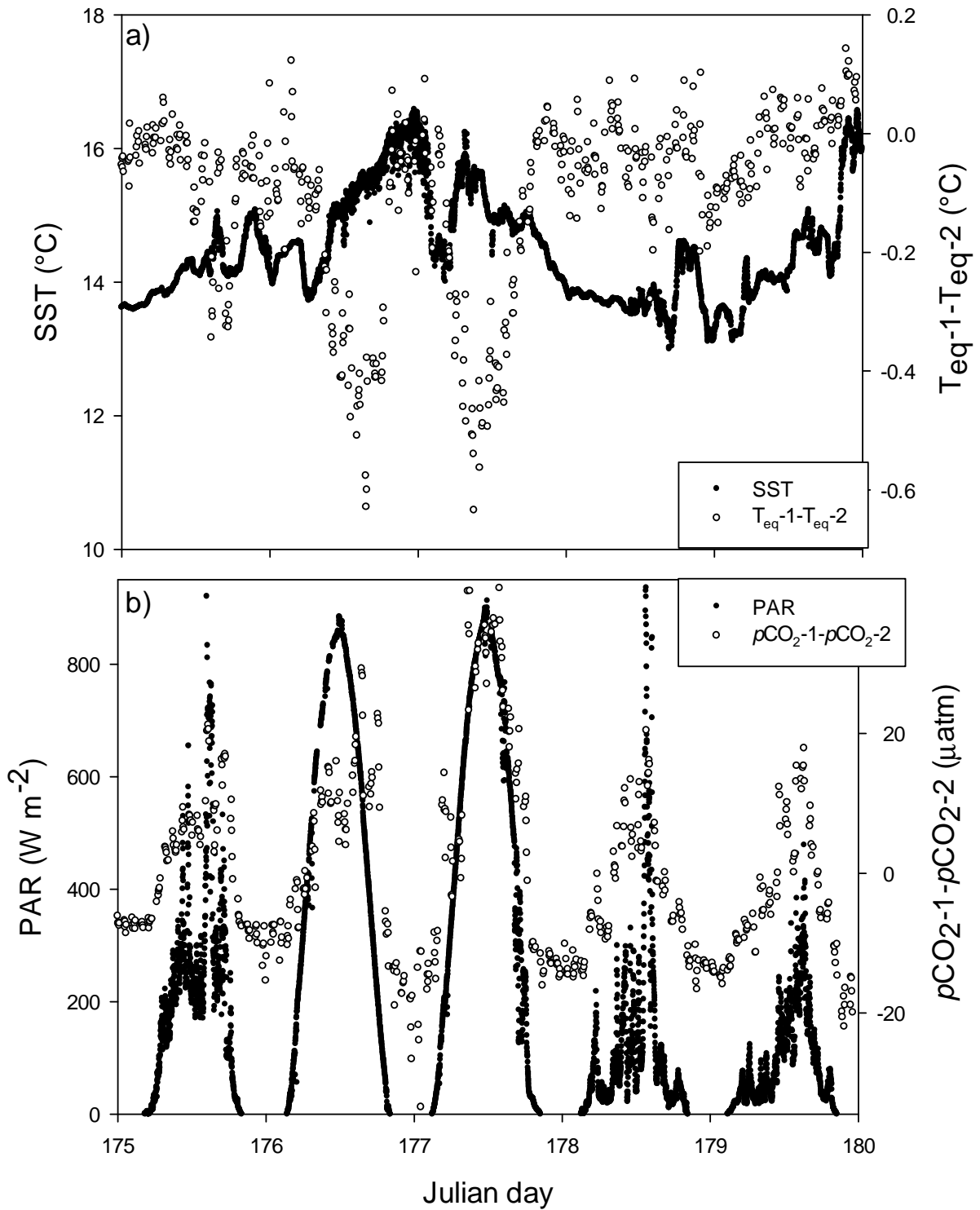
3

4 Figure 1. Relationship between two independently measured $p\text{CO}_2$ datasets. The line of
5 perfect agreement (1:1 line, in green) and the best fit line (in red) are also shown.



1
2
3
4
5
6
7
8
9

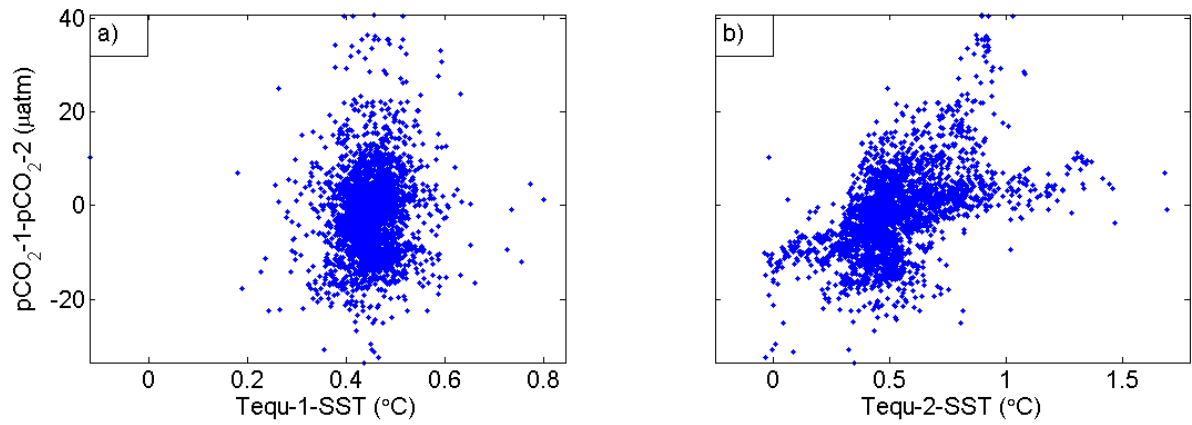
Figure 2. pCO_2 residuals of pCO_2-1 (a) and pCO_2-2 (b) compared to other measured or calculated pCO_2 values. Different symbols show different residuals: against pCO_2 from C_T and A_T (red plus symbols); against pCO_2 from C_T and pH_T (blue circles); against pCO_2 from A_T and pH_T (blue plus symbols); and against pCO_2 from the other measured pCO_2 (2 in a and 1 in b, blue dots). c) In-situ temperature, temperature of equilibrator 1, and temperature of equilibrator 2.



2

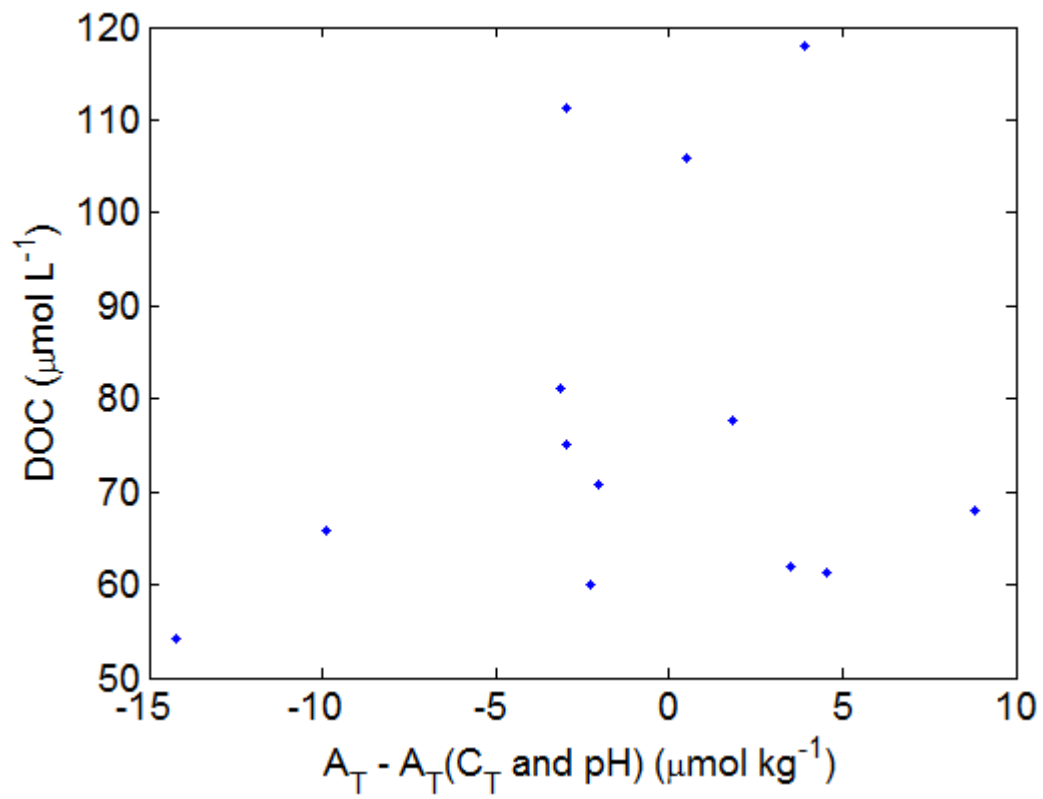
3 Figure 3. Comparison between a) the in-situ sea surface temperature (SST (°C), in black
 4 circles) and the difference in temperatures inside the equilibrators of instruments 1 and 2 (°C;
 5 white circles) and b) the photosynthetically active radiance (PAR ($W m^{-2}$), in black circles)

1 and the $p\text{CO}_2$ difference between instruments 1 and 2 (μatm ; white circles) measured over
2 five days.
3



1
2
3
4
5
6

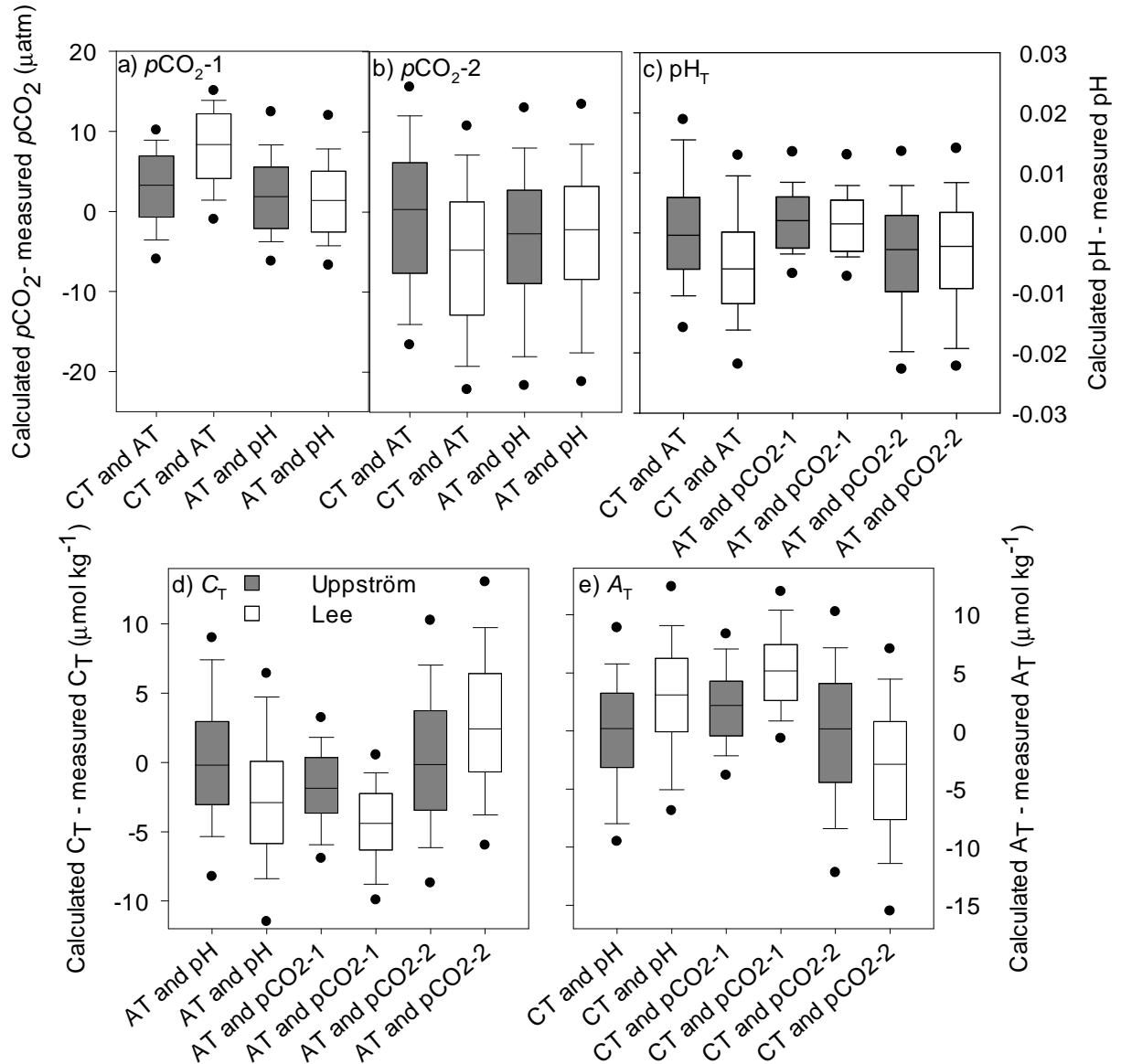
Figure 4. Difference between the two $p\text{CO}_2$ datasets and a) the difference between the temperature inside equilibrator 1 and sea surface temperature, and b) the difference between the temperature inside equilibrator 2 and sea surface temperature.



1
2
3
4

Figure 5. Dissolved organic carbon (DOC) and total alkalinity (A_T) discrepancy (measured A_T minus calculated A_T from C_T and pH_T)

1



2

3 Figure 6. Box-and-whisker plots of the residuals between measured $p\text{CO}_2\text{-1}$ (a), $p\text{CO}_2\text{-2}$ (b),
 4 pH_T (c), C_T (d), A_T (e) and the respective estimates calculated from different pairs of
 5 measured variables (denoted on the x axis) for the two sets of ratios of total boron to salinity.
 6 Grey plots depict the distributions of residuals using the constants of Upström (1974), and
 7 white plots denote those using Lee et al. (2010). The boxes show the median and the 25th and
 8 75th percentiles; dots are the 5th and 95th percentiles. Table 2 presents the RMSE and MR.