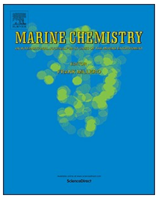
Marine Chemistry 211 (2019) 52–63

Contents lists available at ScienceDirect

Marine Chemistry

journal homepage: www.elsevier.com/locate/marchem



Insights from GO-SHIP hydrography data into the thermodynamic consistency of CO_2 system measurements in seawater

Michael B. Fong*, Andrew G. Dickson

Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

ARTICLE INFO

Keywords:
CO_2 system
Internal consistency
Organic alkalinity
Spectrophotometric pH

ABSTRACT

Due to advances in technology, routine seawater pH measurements of excellent repeatability are becoming increasingly common for studying the ocean CO_2 system. However, the accuracy of pH measurements has come into question due to a widespread observation, from a large number of carefully calibrated state-of-the-art CO_2 measurements on various cruises, of there being a significant pH-dependent discrepancy between pH that was measured spectrophotometrically and pH calculated from concurrent measurements of total dissolved inorganic carbon (C_T) and total alkalinity (A_T), using a thermodynamic model of seawater acid-base systems. From an analysis of four recent GO-SHIP repeat hydrography datasets, we show that a combination of small systematic errors in the dissociation constants of carbonic acid (K_1 and K_2), the total boron-salinity ratio, and in C_T and A_T measurements are likely responsible for some, but not all of the observed pH-dependent discrepancy. The residual discrepancy can only be fully accounted for if there exists a small, but meaningful amount ($\sim 4\ \mu\text{mol kg}^{-1}$) of an unidentified and typically neglected contribution to measured A_T, likely from organic bases, that is widespread in the open ocean. A combination of these errors could achieve consistency between measured and calculated pH, without requiring that any of the shipboard measurements be significantly in error. Future research should focus on establishing the existence of organic alkalinity in the open ocean and constraining the uncertainty in both CO_2 measurements and in the constants used in CO_2 calculations.

1. Introduction

Quantifying long-term changes in the carbon cycling of the ocean due to the uptake of anthropogenic CO_2 from the atmosphere requires accurate characterization of the CO_2 system in seawater. Questions that are fundamental to ocean carbon cycle research, such as calculation of the air-sea flux of CO_2, the calcium carbonate ($CaCO_3$) saturation horizon (the depth below which $CaCO_3$ dissolution is thermodynamically favorable), and the anthropogenic CO_2 inventory of the ocean, all depend on reliable measurements of seawater CO_2 parameters and an accurate characterization of CO_2 system equilibria in seawater (as well as of the equilibria of all other acid-base species in seawater – particularly if total alkalinity is one of the measured parameters).

Conventionally (Park, 1969; Skirrow, 1975; Takahashi et al., 1970), it is stated that only two measured CO_2 parameters – usually from the set: pH, partial pressure of CO_2 ($p(CO_2)$), total alkalinity (A_T), and dissolved inorganic carbon (C_T) – are required to characterize the CO_2 system in seawater. That is, if two of the four parameters from the above set are measured, the other two parameters can be calculated. Of course, this requires a knowledge of the various equilibrium constants for all the acid dissociation reactions considered and other information such as B_T/S, the total boron/salinity ratio, as well as the total concentrations of other acid-base systems present, such as phosphate or silicate. If more than two CO_2 parameters are measured on a suite of samples (*e.g.*, Clayton et al., 1995; Takahashi et al., 1970), then systematic discrepancies observed between the measured values of particular CO_2 parameters and the values calculated from other measured CO_2 parameters have often been attributed to systematic errors in the available sets of equilibrium constants for the CO_2 system (i.e., K_1 and K_2), enabling a preferred set of such constants to be identified, although different studies have disagreed on the preferred set of constants (see *e.g.* Clayton et al., 1995 vs. Lee et al., 2000 and Wanninkhof et al., 1999).

Although a significant number of studies have demonstrated reasonable consistency between seawater $p(CO_2)$, A_T, and C_T measurements using constants based on those originally published by Mehrbach et al. (1973), and provided that $p(CO_2) < 500$ µatm (*e.g.,* Chen et al., 2015; Lee et al., 2000; Lueker et al., 2000; Patsavas et al., 2015; Wanninkhof et al., 1999), no published studies have shown such agreement between measured spectrophotometric pH and the pH

---

* Corresponding author.

E-mail address: mbfong@ucsd.edu (M.B. Fong).

https://doi.org/10.1016/j.marchem.2019.03.006
Received 19 September 2018; Received in revised form 24 January 2019; Accepted 11 March 2019

calculated from measurements of A_T and C_T. On the contrary, measurements from a variety of cruises (see *e.g.*, Carter et al., 2018; Carter et al., 2013; McElligott et al., 1998; Williams et al., 2017) show clearly that there is a seemingly systematic discrepancy between the measured and calculated pH values that is itself a function of pH. Such a discrepancy can, in principle, be attributed either to systematic errors in the spectrophotometric pH measurements themselves and/or to errors in the measurements of A_T and C_T combined with errors in the thermodynamic model of acid-base reactions in seawater that is used to infer pH from such measurements. Furthermore, as the speciation of seawater acid-base systems is itself a function of pH, systematic errors in any of the input parameters used to calculate pH will necessarily lead to a pH-dependent error in the calculated pH.

The consistency of spectrophotometric pH with the other CO_2 parameters has important implications for the utility of pH as a measured biogeochemical parameter. Although pH measurements in seawater have historically been considered unreliable (Keeling, 1968), recent advances in measurement technology, in particular the advent of spectrophotometric pH, have positioned pH to be suitable for routine measurement and as a potentially useful parameter for studying the ocean CO_2 system (Dickson, 1993). With its excellent short-term precision (repeatability ~0.0004 in pH, Clayton and Byrne, 1993), spectrophotometric pH may be particularly desirable for use in CO_2 system calculations, due to the potentially small contribution from pH to the overall imprecision of the calculated parameter (Clayton et al., 1995; McElligott et al., 1998; Patsavas et al., 2015). pH is also a popular choice for autonomous sensors (*e.g.*, Martz et al., 2010; Seidel et al., 2008), which are either based on, or calibrated using, spectrophotometric pH measurements. pH sensors have been developed for use on profiling floats (Johnson et al., 2016), and a network of floats has been deployed in the Southern Ocean, with one of its goals being to calculate $p(CO_2)$ from float-measured pH (Williams et al., 2017) and A_T estimated from a locally interpolated regression (Carter et al., 2018). If, however, there is a systematic error in spectrophotometric pH of the magnitude suggested by the discrepancy between measured and calculated pH (potentially greater than 0.01 pH units), this would severely limit the reliability of pH data and the use of pH to calculate other CO_2 parameters.

The large number of carefully calibrated state-of-the-art CO_2 measurements made on repeat hydrography cruises makes it possible to evaluate the likely quality of CO_2 measurements as well as our understanding of CO_2 system thermodynamics. This paper examines measurements from four GO-SHIP cruises (Global Ocean Ship-Based Hydrographic Investigations Program), in which our laboratory participated (measuring pH and A_T), and will consider how systematic errors in the dissociation constants of carbonic acid (i.e., K_1, K_2,) the boric acid dissociation constant (K_B), B_T/S, A_T, and C_T measurements, as well as how potentially unaccounted for acid-base species, might be responsible for the observed pH-dependent pH discrepancy.

2. Methods

2.1. Choice of data sets

Data from four GO-SHIP repeat hydrography cruises (Fig. 1) were examined in this study: 2014 P16S, 2015 P16N, 2016 I08S, and 2016 I09N (Expocodes: 320620140320; 33RO20150410, 33RO20150525; 33RR20160208; 33RR20160321). These were chosen as they are recent cruises on which our laboratory made measurements of pH (by spectrophotometry using purified *m*-cresol purple indicator – using the approach developed by Carter et al., 2013) and A_T (following SOP 3b in Dickson et al., 2007, a variant of the open-cell method of Dickson et al., 2003); C_T was measured by the usual extraction / coulometric technique (SOP 2 in Dickson et al., 2007) by scientists from NOAA. A_T and C_T measurements were standardized to CO_2-in-seawater Reference Materials produced by our laboratory. These four cruises represent the South Pacific, the North Pacific, the Southern Ocean (Indian Ocean sector), and the Indian Ocean, respectively. Full depth profiles for pH, A_T, C_T, temperature, salinity, and nutrients (phosphate, silicate, nitrate, and nitrite) were measured on all four cruises. Other than excluding data with bad quality flags for any of the measurements for a particular bottle and adjusting the measured pH value to 25 °C as needed (the spectrophotometric pH was measured at 20 °C on 2014 P16S), the data sets were not adjusted further before analysis. There were 10,018 samples altogether for which all of these parameters were measured successfully. The data and cruise reports are available at NOAA National Centers for Environmental Information (https://www.nodc.noaa.gov/ocads/oceans/RepeatSections/).

2.2. Approach

Values of ΔpH were calculated at 25 °C and a gauge pressure of zero dbar (ambient atmospheric pressure) – the conditions at which pH_{spec} was usually measured – for water samples from each of these cruises using the equation below:

$$\Delta pH = pH_{spec} - pH_{calc}(C_{T,}\, A_T, \ldots). \tag{1}$$

The ellipsis stands for the full thermodynamic model used to calculate pH from A_T and $C_{T.}$ We computed values for ΔpH using CO2SYS for MATLAB (van Heuven et al., 2011) and the data for equilibrium constants, etc. detailed in Table 1. Each value of ΔpH calculated requires, in addition to the measured values of the carbonate parameters, a salinity value (used to estimate equilibrium constants and the total boron concentration) and the measured total concentrations of phosphate and of silicate.

These ΔpH values are plotted for each of these cruises as a function of pH in Fig. 2 (left-hand panels). For each cruise, these discrepancies have a clear pH-dependence, and they appear to be well represented by a straight line that can be fit using a simple unweighted least-squares to the ΔpH values shown. In each case, this least-squares line has a significant non-zero slope, and there is a significant non-zero mean for ΔpH (Table 2).

A uniform adjustment to any one of the parameters used to estimate pH_{calc} will necessarily change each value of ΔpH. However, it does not change them identically, as the sensitivity of pH_{calc} to each of these parameters is a function of pH. To a reasonable approximation, the modified values for ΔpH resulting from such an adjustment (or a combination of such adjustments) also lie almost on a straight line, but with a changed slope and mean ΔpH. Any individual point for ΔpH is the result of a combination of both systematic and random errors. However, as the adjustments do not significantly change the distribution of ΔpH residuals, their effect on the slope of the least-squares line and the mean ΔpH primarily reflects adjustments for systematic errors. (See Fig. 3 for examples of how changing both K_1 and K_2 affect the distribution of ΔpH.) Our goal therefore is to choose a set of adjustments that, when applied, results in a distribution of ΔpH that has no significant slope, and for which the mean ΔpH is essentially zero. Ideally, of course, any proposed adjustments will not seem implausible, nor will they significantly worsen the apparent consistency previously found between seawater $p(CO_2)$, A_T, and C_T measurements. Furthermore, the final distribution of ΔpH should ideally reflect its likely precision (i.e., reflecting the contributions of only random errors, which can be inferred from the known precisions of the measurements of pH, A_T, and C_T).

2.3. Identifying possible systematic errors

The 2015 P16N cruise covered the widest range of pH (see Fig. 2) and was therefore considered initially. The first question to ask is: how would the distribution of ΔpH change if alternate formulations for the CO_2 acid dissociation constants (K_1 and K_2) were used? We therefore

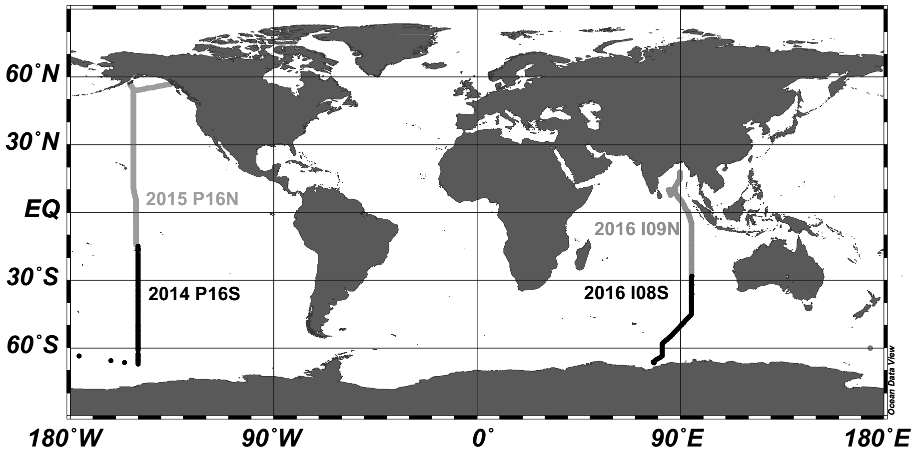


Fig. 1. Map showing cruise tracks for the four cruise data sets discussed here: 2014 P16S, 2015 P16N, 2016 I08S, and 2016 I09N.

Table 1
Parameters used to calculate ΔpH (see Eq. (1)). The various constants listed here are explicitly chosen in CO2SYS-MATLAB. Other constants not listed are implicitly chosen in the program. Values given for the initial estimated systematic error are used in determining the vector lengths in Fig. 4a and, for pH, defining the tolerance in the mean value of ΔpH of the adjusted dataset (as in Fig. 4b).

Parameter	Source	Initial estimated systematic error
pH	Shipboard measurement	0.004[a]
A_T	Shipboard measurement	1 µmol kg^{-1} [b]
C_T	Shipboard measurement	1 µmol kg^{-1} [b]
pK_1	Lueker et al., 2000	0.0075[c]
pK_2	Lueker et al., 2000	0.015[c]
pK_B	Dickson, 1990	0.004
B_T/S	Lee et al., 2010	0.03 (*relative error*) [d]
A_X	Assumed = 0 µmol kg^{-1}	4 µmol kg^{-1}

[a] Estimate of the systematic error in assigning pH values to Tris buffers in synthetic seawater due to the assumption that the activity coefficient of HCl in the buffers is the same as that in pure synthetic seawater (DelValls and Dickson, 1998; Müller and Rehder, 2018).

[b] Estimated bias of a dataset where A_T and C_T were measured with state-of-the-art methods and standardized to CO$_2$-in-seawater Reference Materials.

[c] Estimated from the degree of agreement between different formulations of the same constants over their entire salinity range, as described in Orr et al. (2018).

[d] The difference between the ratios of Lee et al. (2010) and Uppström (1974) is 3.9%. Thus, we allow an adjustment in this ratio up to 3%.

repeated the calculation for this cruise using alternate published values for the dissociation constants (measured in natural seawater) that are available within CO2SYS-MATLAB (Fig. 3). Clearly, there is no set of CO$_2$ constants that removes the observed slope of ΔpH vs. pH, although the mean value of ΔpH does change noticeably depending on the particular set of CO$_2$ constants used.

We then considered, one by one, the effect of the following potentially important systematic errors in the model for the calculation of ΔpH: constant relative errors in K_1 and K_2 (i.e., constant offsets in the pK values from those of Lueker et al., 2000); a constant relative error in K_B (relative to Dickson, 1990); a constant relative error in the boron/salinity ratio (relative to Lee et al., 2010); a constant absolute error in measured values of C_T and A_T; and, finally, the effect of there being additional, unidentified, acid-base systems present in seawater (Dickson, 1992) that we treat (simplistically) as if they contribute a constant amount (A_X) to the measured alkalinity at all places in a particular ocean region. Thus, this can be considered as a systematic error in our interpretation of the measured total alkalinity in terms of the contributing species. We decided to ignore potential systematic errors in the parameters associated with the phosphate and silicate systems, as these systems are present at low concentrations and thus their likely uncertainties are less important.

One way to consider the implications of such model adjustments is to ask how each individual adjustment would affect the mean ΔpH and the slope of a least-squares line through the entire distribution of ΔpH for a particular cruise, as a consequence of the pH-dependent sensitivity of pH$_{calc}$ to an adjustment. In Fig. 4a, we display this in the form of a vector diagram showing, for each of the potential adjustments, the effect (averaged over the whole 2015 P16N data set) both on the mean value of ΔpH (y-coordinate) and on the slope of a least-squares line of ΔpH as a function of pH (x-coordinate). Each vector indicates the direction and magnitude of change in these coordinates and is displayed with a length corresponding to the effect of an initial estimate of the likely magnitude of systematic error in each of the parameters (Table 1) – with the exception of A_X – and starting at the point indicating the original (unadjusted) mean ΔpH and the slope of its dependence on pH for the 2015 P16N dataset. As the concentration of A_X is unknown, we display in Fig. 4a a vector representing the effect of a small amount of A_X (4 µmol kg^{-1}). Thus, the vector i representing an adjustment of a single parameter (i.e., ΔpK_1, ΔpK_2, $\Delta B_T/S$, ΔK_B, ΔC_T, ΔA_T, or A_X) can be written as (Δx_i, Δy_i). Ideally, a combination of plausible adjustments can be found such that ΔpH ends up distributed with a mean close to zero and without any significant dependence on pH. It is important to note, however, that our approach which assumes a uniformly distributed, constant amount of A_X would not adjust appropriately for the effects of a non-uniform distribution of A_X within a particular cruise (see later discussion).

2.4. Computational approach

Using the approach described in the previous section, any plausible set of adjustments (correcting for potential systematic errors) will be a sum of the vectors shown in Fig. 4a (scaled as needed) that starts at the original coordinates and reaches a target area for which the mean ΔpH is not significantly different from zero ($+0.004 \geq$ mean $\Delta\text{pH} \geq -0.004$: an estimate of the systematic error in assigning pH values to the buffers used to calibrate spectrophotometric pH) and for which the slope of a least-squares regression of ΔpH against pH is also not significantly different from zero ($+0.001 \geq$ slope ≥ -0.001: approximately the mean 95% confidence interval for the slopes of the lines shown in Figure 2 and listed in Table 2). Ideally, the magnitude of individual proposed adjustments would not substantially exceed the estimated systematic error for the proposed parameters (i.e., the scale factors $-1 \leq a_i \leq +1$). (Of course, for A_X the magnitude is unknown, and there is no constraint on the corresponding value of a.) Hence, the

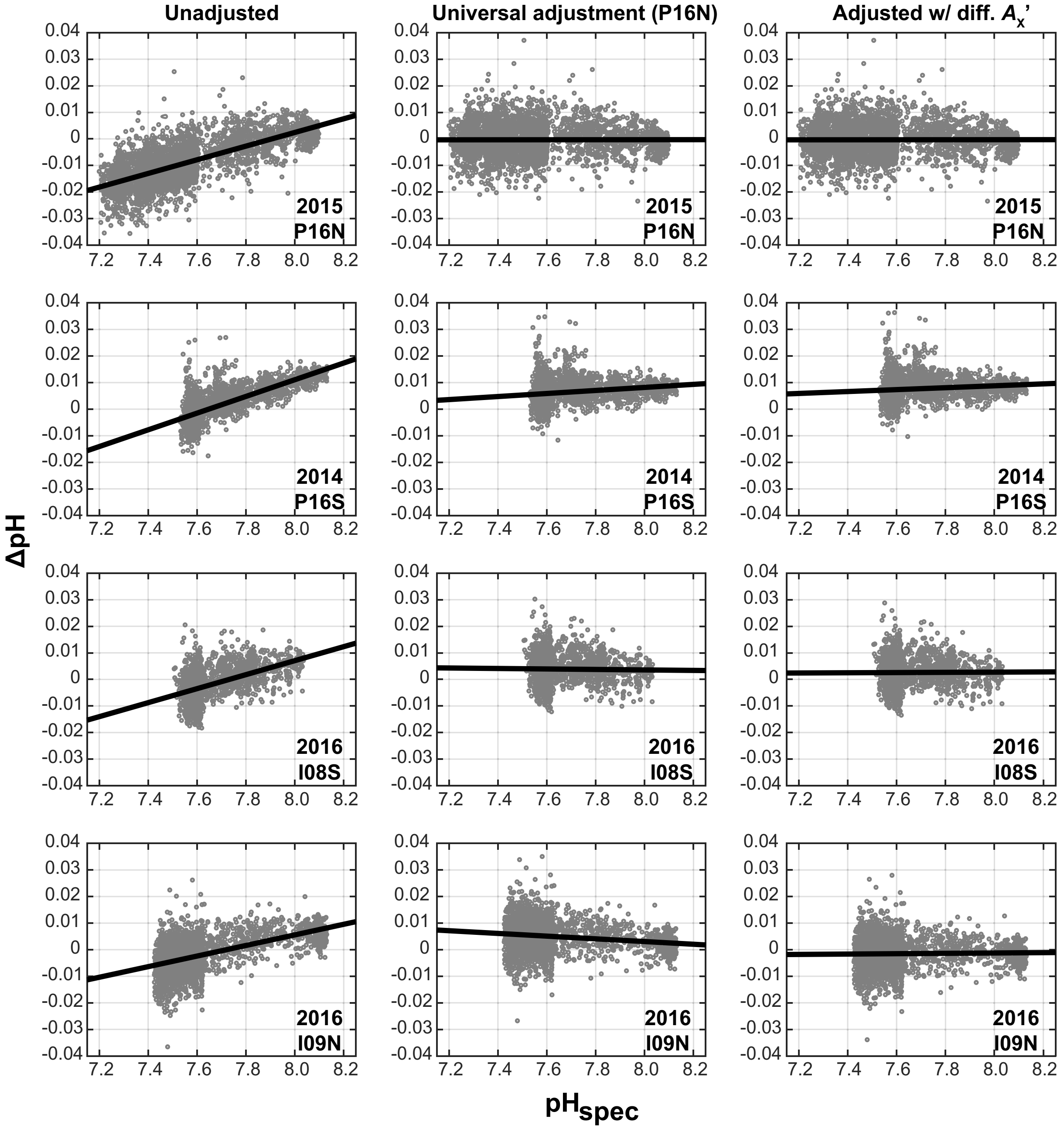


Fig. 2. Values of ΔpH calculated at 25°C and a gauge pressure of zero dbar plotted against measured pH for each cruise. Eq. (1) and the model in Table 1 were used to calculate ΔpH in the panels on the left. ΔpH in the center panels were calculated using adjustments to the model estimated from the 2015 P16N dataset (see also Table 3). ΔpH in the right panels were calculated with the same adjustments to pK_1, pK_2, and the total boron-salinity ratio as in the center panels but with a different amount of apparent excess alkalinity (A_X'), optimized individually for each cruise (see also Table 4).

Table 2
Regression statistics for the data shown in Figure 2 (left panels) for ΔpH *versus* pH (at 25°C and a gauge pressure of zero dbar).

	Slope ± std. error	Intercept	*x*-intercept	R^2	mean ΔpH ± std. dev.	*n*
2015 P16N	0.0257 ± 0.00044	–0.203	7.91	0.52	–0.0092 ± 0.0083	3166
2014 P16S	0.0315 ± 0.00057	–0.241	7.65	0.53	0.0009 ± 0.0062	2673
2016 I08S	0.0264 ± 0.0011	–0.204	7.73	0.26	–0.0016 ± 0.0064	1661
2016 I09N	0.0199 ± 0.00059	–0.154	7.72	0.31	–0.0018 ± 0.0069	2518

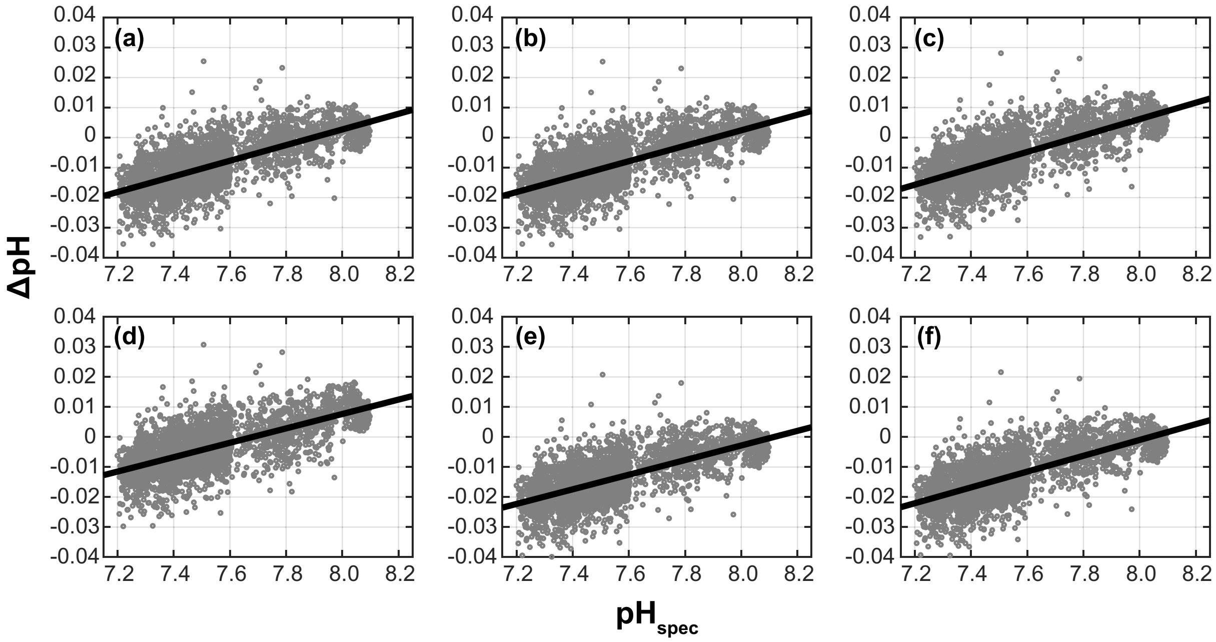


Fig. 3. Values of ΔpH for 2015 P16N calculated using Eq. (1) and the model in Table 1, but with alternate formulations for K_1 and K_2: the modified Mehrbach constants of (a) Dickson and Millero (1987) and (b) Lueker et al. (2000), as well as the constants of (c) Mojica Prieto and Millero (2002), (d) Millero et al. (2002), (e) Millero et al. (2006), and (f) Millero (2010).

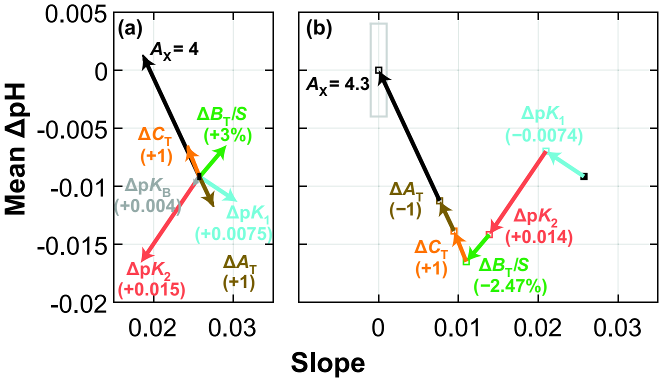


Fig. 4. Vector diagrams showing (a) how individual systematic errors in the various parameters (Table 1) used to calculate ΔpH would affect the value of the slope (of ΔpH *versus* pH) and the mean value of ΔpH for the 2015 P16N dataset and (b) a combination of adjustments (scaled versions of the vectors in Fig. 4a) obtained using fgoalattain (see also Table 3) that would bring the value of the slope and the mean ΔpH to near zero for the same dataset. The gray box defines a near-zero goal for both the slope and mean ΔpH of –0.001 to +0.001 and –0.004 to +0.004, respectively.

net effect of such a set of adjustments can be described by the equation:

$$\mathbf{0} \approx \begin{cases} x = x_0 + a_1\,\Delta x_1 + a_2\,\Delta x_2 ... + a_n\,\Delta x_n \\ y = y_0 + a_1\,\Delta y_1 + a_2\,\Delta y_2 ... + a_n\,\Delta y_n \end{cases}. \qquad (2)$$

As in Section 2.3, x is the slope of a regression of ΔpH against measured pH, and y is the mean value of ΔpH for the same data set; the subscript 0 refers to the initial starting value (calculated with no adjustments); the subscripts $i = 1, 2, \ldots n$ refer to six of the adjustment vectors in Fig. 4a: ΔpK_1, ΔpK_2, $\Delta B_T/S$, ΔC_T, ΔA_T, and A_X. The vector representing the effect of systematic error in the acid dissociation constant for boric acid is small and approximately co-linear to that for the dissociation of bicarbonate (K_2). We therefore chose not to adjust it independently, as we believe it has the smaller uncertainty. A vector diagram such as in Fig. 4b can be used to visualize the implications of such a set of adjustments. Our aim is to determine if the discrepancies in ΔpH can be eliminated while keeping assumed systematic errors small enough that they can be considered plausible (or at least cannot be ruled out). Hence, the vectors representing our proposed set of adjustments for systematic errors (Fig. 4b) are scaled versions of the vectors in Fig. 4a, with absolute magnitudes no larger than the values specified in Table 1. For ΔA_T and ΔC_T, we chose limits of ± 1 μmol kg^{-1}; for ΔpK_1, ± 0.0075 (~1.74% in K_1); for ΔpK_2, ± 0.015 (~3.5% in K_2); we omitted consideration of any error in K_B as it would not be practical to distinguish it from any systematic error in K_2 (see above); for $\Delta B_T/S$, we allowed a relative error of up to 0.03 (3%); for the omitted alkalinity component, A_X, there was no *a priori* limit chosen (though any systematic error in A_T measurement, and to a large extent any systematic error in C_T, is necessarily inseparable from this particular systematic error using our approach – see later discussion). Although our adjustment limits are not rigorous estimates of the systematic error in the various parameters, they represent plausible magnitudes (see footnotes in Table 1) that are unlikely to be exceeded by a significant amount. We discuss later how our assumptions about the likely systematic error in the various parameters in Table 1 may affect our solution.

We used the MATLAB function fgoalattain with these constraints on the magnitudes of the various vectors and the goal constraints described in this section (+0.001 ≥ slope ≥ –0.001; +0.004 ≥ mean ΔpH ≥ –0.004) to choose a set of vectors that achieved our aim. This multi-objective goal attainment function uses a sequential quadratic programming algorithm and finds a solution that minimizes the relative difference between the values of the two objective functions (x and y) in Eq. (2) and the goals while also satisfying various constraints for the solution variables (the scale factors a_i). We first applied fgoalattain to the 2015 P16N dataset and then in turn to the other datasets. To define the objective functions for fgoalattain, we numerically evaluated the adjustment vectors (Δx_i, Δy_i) separately for each dataset (as in Fig. 4a), as these vectors may be slightly different depending on the range in composition of the seawaters analyzed. For a single dataset, however, these vectors are effectively constant over the range of adjustments we examined, and hence, the net effect of a set of adjustments can be calculated by summing the scaled vectors as in Eq. (2).

Table 3
Adjustments estimated for each of the four cruises by using fgoalattain for each cruise independently. All datasets, with the exception of 2014 P16S, met the goal of a near-zero slope (of ΔpH *versus* pH) and near-zero mean ΔpH after these adjustments.

	pK_1	pK_2	B_T/S in %	C_T µmol kg^{-1}	A_T µmol kg^{-1}	A_X µmol kg^{-1}
2015 P16N	−0.0074	+0.014	−2.47	1.0	−1.0	4.3
2014 P16S	−0.0075	+0.015	−3.0	−1.0	−1.0	7.3
2016 I08S	−0.0075	+0.015	−3.0	−1.0	−1.0	5.3
2016 I09N	−0.0075	+0.014	−2.25	0.7	−0.9	2.4

3. Results

3.1. Initial examination of cruises

The discrepancy ΔpH between the spectroscopic pH measured on the ship and the pH calculated from the shipboard measurements of A_T and C_T (using the various constants shown in Table 1) has a similar pattern (pH-dependent slope) for each of the cruises we studied (left-hand panels of Fig. 2; Table 2). However, the least squares line fit to each dataset is not the same for all cruises. The slope is not identical for all cruises, nor is the *x*-intercept (the pH where the regressed value of ΔpH = 0). The mean value of ΔpH varies as each cruise encompasses a different range of pH values. A data set with a greater number of points at low pH, where ΔpH is most negative, would be expected to have more a more negative mean ΔpH (see Table 2).

3.2. Seeking plausible adjustments that eliminate the systematic discrepancies in ΔpH

The results from applying fgoalattain are shown in Fig. 4b for the 2015 P16N data set. Similar calculations were carried out separately for each of the individual data sets considered (Table 3). As the estimated adjustments for the various constants (pK_1, pK_2, B_T/S) might reasonably be expected to be of global applicability, we also carried out calculations holding these adjustments at the values suggested for 2015 P16N and simply optimizing for the apparent A_X for each separate cruise (Table 4).

It should be recognized that our approach does not obtain a unique solution (see also the discussion in Section 4.3). Rather, it can only demonstrate that a plausible solution exists (a set of potential systematic errors and their implied magnitudes given our *a priori* constraints). In addition, an examination of Fig. 4a indicates that the vector representing the effects of a systematic error in the measurement of A_T, ΔA_T, is exactly opposite in direction to that for an omitted alkalinity component, A_X. Also, the vector representing a systematic error in the measurement of C_T is approximately co-linear with that for a systematic error in the measurement of A_T, although of opposite sign. Consequently, our approach cannot be expected to distinguish reliably between these, and thus the combination: $A_X{}' = A_X - \Delta A_T + \Delta C_T$, may well be better defined than any of its individual terms. This is discussed further in Section 4.

3.3. Implication of proposed adjustments for the calculation of $A_X{}'$ from pH and C_T

A common expression for seawater total alkalinity (Dickson et al., 2003) is

$$\begin{aligned} A_T \;=\; & [HCO_3^-] + 2[CO_3^{2-}] + [B(OH)_4^-] + [OH^-] \\ & + [HPO_4^{2-}] + 2[PO_4^{3-}] + [SiO(OH)_3^-] + [NH_3] + [HS^-] + \ldots \\ & - [H^+] - [HSO_4^-] - [HF] - [H_3PO_4] + \ldots \end{aligned} \tag{3}$$

where the ellipses indicate additional minor base or acid species that are either unidentified or present in such small amounts that they need not be considered. The concentrations of NH_3 and HS^- are typically so low that they are unimportant in oxygenated open ocean water. Furthermore, other conceivable inorganic acid-base species are also be expected to be present only at very low concentrations. It is thus usual to imagine that any discrepancy between the measured total alkalinity and that calculated from Eq. (3),

$$A_{T\,meas} - A_{T\,calc}(C_T,\, \mathrm{pH}, \ldots), \tag{4}$$

is indicative of the presence of measurable amounts of organic bases in the seawater sample (Cai et al., 1998), and it has been suggested that organic particles such as phytoplankton or bacterial cells can also contribute (Kim et al., 2006), though these are scarcer in open ocean samples than in the coastal region studied by Kim et al. In addition, this difference will also include the effects of any systematic errors in the measured values of pH, A_T and C_T, as well as in the various constants etc. that go into the calculation of A_T from pH and C_T. If the adjustments proposed for pK_1, pK_2, and B_T/S (based on 2015 P15N) are correct, then this difference will essentially be the previously defined term $A_X{}'$.

In Fig. 5 (left-hand panels), we plot this difference as a function of depth for each of the four data sets, both using the usual recommended estimates for the various constants, etc. (Table 1), as well as showing the effect of repeating these calculations (Fig. 5, right-hand panels) using the adjusted values for pK_1, pK_2, and the ratio B_T/S that were estimated from the 2015 P16N data set.

3.4. Implication of proposed adjustments for the calculation of $p(CO_2)$ from A_T and C_T

As noted in our introduction, modified versions (changed to a concentration-based pH scale) of the Mehrbach et al. (1973) acid dissociation constants for carbon dioxide (Dickson and Millero, 1987; Lueker et al., 2000) have been widely recommended over the past 20 years or so, largely because the estimates of $p(CO_2)$ calculated from measurements of A_T and C_T were found to be in reasonable agreement with measured values (Wanninkhof et al., 1999), at least for underway $p(CO_2)$ where $p(CO_2)$ is rarely above 500 µatm. Lueker et al. reported similar observations for a laboratory study where seawater $p(CO_2)$ was modified (and measured) and A_T and C_T were also measured.

We felt it would be appropriate to evaluate quite how much our proposed adjustments might affect these earlier observations. We therefore recalculated $p(CO_2) = f(A_T, C_T, B_T, K_1, K_2, \ldots)$ for the Lueker et al. set of measurements, and compared it (as Lueker et al. did) with

Table 4
Regression statistics for the data (ΔpH *versus* pH) from each of the four cruises, using a common set of adjustments for pK_1, pK_2, and B_T/S (−0.0074, +0.014, and −2.47%, respectively, as listed in Table 3 for 2015 P16N and also shown in Fig. 4b) and optimizing the individual cruise adjustment for *apparent excess alkalinity* ($A_X{}'$).

	$A_X{}'$ µmol kg^{-1}	Slope ± std. error	Intercept	R^2	mean ΔpH ± std. dev.	*n*
2015 P16N	6.3	−0.00050 ± 0.00045	0.004	0.00039	0.0005 ± 0.0059	3166
2014 P16S	6.6	0.00353 ± 0.00059	−0.019	0.013	0.0077 ± 0.0044	2673
2016 I08S	5.5	0.00046 ± 0.0011	−0.001	0.0001	0.0025 ± 0.0057	1661
2016 I09N	3.5	0.00071 ± 0.00060	−0.007	0.00055	−0.0015 ± 0.0058	2518

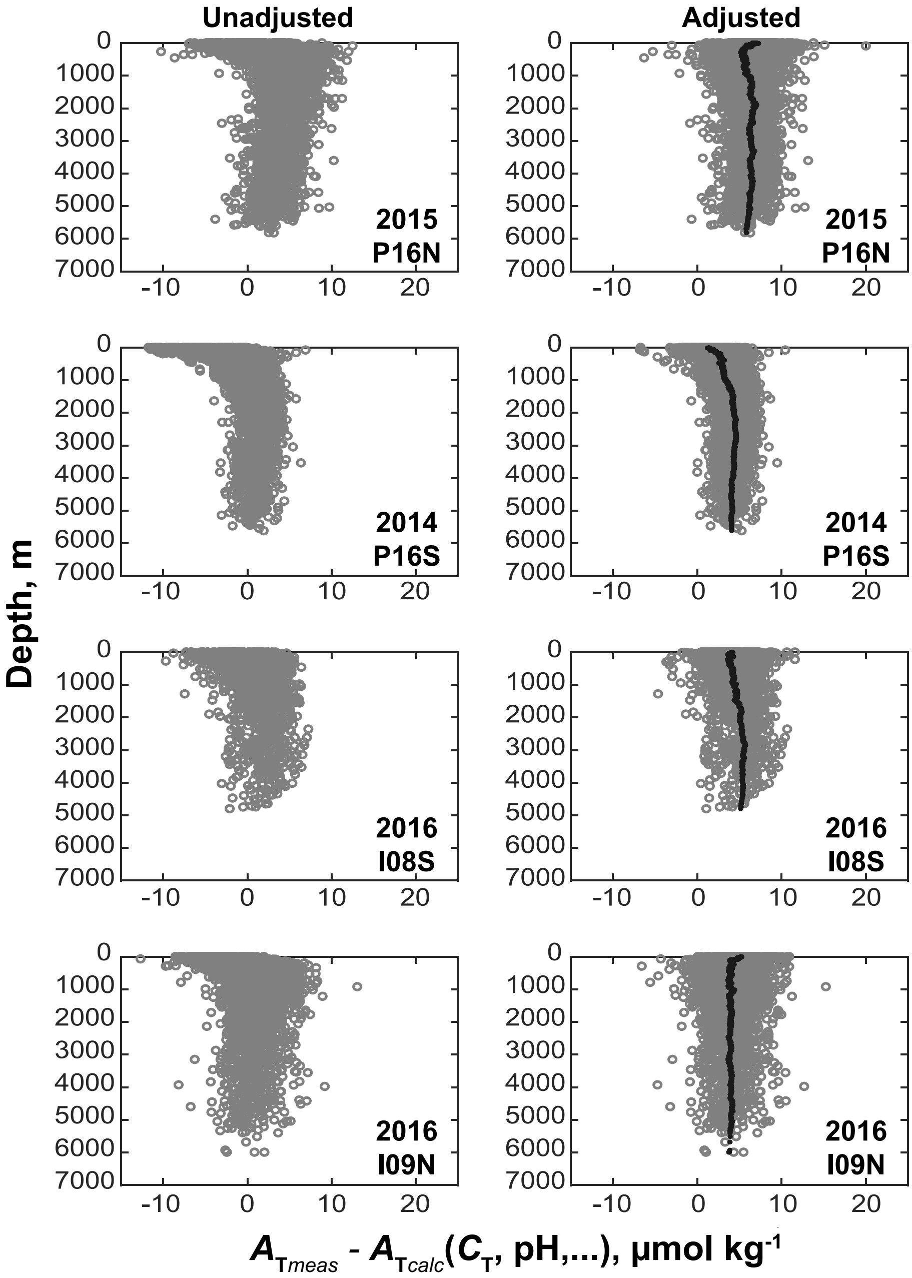


Fig. 5. Estimation of *excess alkalinity* as a function of depth for each of the four cruises. The left panels show the *excess alkalinity* as calculated from Eq. (4) and the model in Table 1, whereas the right panels show the *excess alkalinity* after adjusting pK_1, pK_2, and the total boron-salinity ratio as proposed for the 2015 P16N dataset (Fig. 4b). The solid black line is a 100-point running mean of the data.

the measured $p(CO_2)$ values. As our adjustments are appropriate only to 25°C, we only considered the measurements of Lueker et al. at ~25°C. We also did not adjust the A_T and C_T data of Lueker et al., as these measurements were expected to be of higher quality than shipboard measurements. The effect of our set of proposed adjustments (as estimated from the 2015 P16N cruise data set and neglecting systematic errors in A_T and C_T) on these differences is shown in Fig. 6.

4. Discussion

4.1. Proposed causes of the pH-dependent discrepancy

As noted in the introduction, there is a seemingly systematic discrepancy between the spectrophotometrically measured pH values and pH values inferred from measured values of A_T and C_T, using a model for the acid-base processes occurring in seawater. Although, in theory, this could be caused by systematic problems with the spectrophotometric measurement of pH, we discount such a possibility for a couple of reasons. First, our laboratory has some unpublished measurements where the pH of a series of buffers based on Tris in synthetic seawater over a range of pH (7.6 – 8.3) was measured both using Harned cells and spectrophotometrically using purified *m*-cresol purple. The discrepancies were small (< 0.004 in pH) and did not show the same systematic variation. Second, a poster at the recent 2018 Ocean Sciences Meeting (Walz et al., 2018) demonstrated reasonable agreement (< 0.005 in pH) between spectrophotometric pH and pH measured using an IS-FET sensor for a seawater whose pH was changed between 8.5 and 7. Even a discrepancy in the calibration of the acid-dissociation constant for *m*-cresol purple would not introduce a significant pH-dependent discrepancy, though it would affect the mean ΔpH value. One remaining possibility is

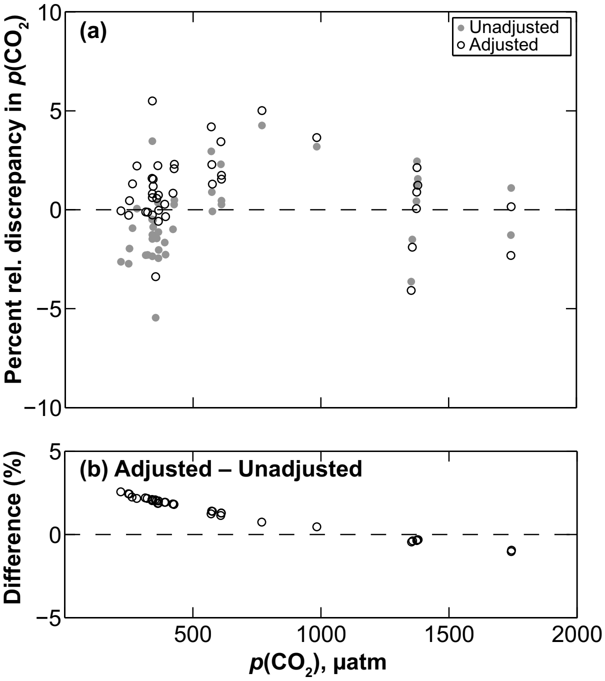


Fig. 6. (a) Recalculation of the data of Lueker et al. (2000), showing how the adjustments proposed for 2015 P16N (neglecting systematic errors in measured C_T and A_T) affect the observed percent relative discrepancy in $p(CO_2)$. Only the data at ~25°C were considered. The relative discrepancy was calculated as the difference — measured $p(CO_2)$ *minus* $p(CO_2)$ calculated from A_T, C_T, and the constants in Table 1, etc. — relative to measured $p(CO_2)$. (b) The difference between the adjusted and unadjusted data shown in Fig. 6a.

that errors in the calibration of the dye's optical properties might introduce a pH dependence of ΔpH; however, the disagreement between the two such published calibrations (DeGrandpre et al., 2014; Liu et al., 2011) could only be responsible for ~10% of the slopes shown in the left-hand panels of Fig. 2 and ~0.003 of the observed mean ΔpH.

The hypothesis that underlies our procedure is that, as a consequence of the pH-dependence of acid-base reactions in aqueous systems, the error in the pH computed for a seawater sample resulting from systematic error in any of the input parameters will itself be a function of pH (i.e., the composition of the system). This is true not only for the acid-dissociation constants, K_1, K_2, K_B etc., but also for the total concentrations such as A_T, C_T, or B_T. We therefore seek to propose a set of postulated systematic errors that could be "corrected," thus eliminating the observed discrepancy.

An examination of the results for this process on our archetype (the 2015 P16N dataset) shows that adjustments estimated using our approach can indeed remove the observed problem. The adjusted data no longer exhibit an apparent dependency of ΔpH on pH, and there is no significant offset of the adjusted mean ΔpH from the hoped for zero value. This is apparent in the central top panel of Fig. 2, where the remaining variance in ΔpH (s^2 = 0.0059^2) can be considered to be due to the variance resulting from measurement imprecision (for pH, A_T, and C_T) as well as any inhomogeneity in the distribution of our imagined A_X parameter. The shipboard measurements of pH, A_T, and C_T were determined (from measurements on CO_2 in seawater reference materials) to have precisions (expressed as standard deviations) of ~0.0010 in pH, and ~1.0 µmol kg^{-1} for A_T and C_T. A simple variance analysis for the 2015 P16N cruise data set would then imply that the inhomogeneity in A_X (again expressed as a standard deviation) was about 1.7 µmol kg^{-1}.

4.2. Plausibility of proposed adjustments to the 2015 P16N dataset

As noted above, a key touchstone for our proposed adjustments for systematic errors was that they should not be implausible. Figure 4b (and Table 3) details the adjustments found to be necessary to "fix" the 2015 P16N data. As was noted, no adjustment was made to K_B; the suggested adjustment for K_1 was +1.71% (–0.0074 in pK_1) and to K_2 –3.33% (0.014 in pK_2). Neither of these exceeds published estimates for the likely standard uncertainty of these parameters (Dickson, 2010a; Orr et al., 2018) and, indeed, they are only about 1.5 times the measurement precision (estimated from the fit to an interpolation equation – Lueker et al., 2000). The proposed adjustment to B_T/S is –2.47%, putting the adjusted value between the published value of Lee et al. (2010) and that of Uppström (1974), which differ by about 3.9%. The most striking adjustment proposed is an unidentified contribution to measured total alkalinity that we refer to as A_X. The adjustments for A_T and C_T (–1.0 and +1.0 µmol kg^{-1} respectively) are similar to the estimated standard uncertainty of the reference materials used to quality control the shipboard measurements of these parameters (i.e., the maximum magnitude permitted) and appear to have been chosen by fgoalattain so as to minimize the magnitude of A_X. The amount content of A_X needed to ensure that the measured and calculated pH values are consistent with one another was estimated as 4.3 µmol kg^{-1}. This is a significant quantity, as the usual shipboard precision of total alkalinity measurements is between 1.0 and 1.5 µmol kg^{-1}.

If there is indeed a contribution to measured alkalinity that is not normally considered in the acid-base model of open-ocean seawater, it may be relevant to examine its vertical distribution directly. One way to achieve this is to estimate $A_X{}'$ for each seawater sample on a cruise by subtracting the estimated inorganic contributions to alkalinity from the measured value (i.e., $A_X{}' = A_{Tmeas} - A_{Tcalc}(C_T,\text{pH},\ldots)$, as in Eq. (4)) and to recognize that the calculated value for $A_X{}'$ represents an apparent *excess alkalinity* (i.e., also including the effects of potential systematic errors in A_T and C_T). Of course, it would likely be appropriate to use our adjusted values for K_1, K_2, and B_T/S when calculating $A_X{}'$. The results (both with and without the proposed adjustments to K_1, K_2, and B_T/S) are plotted in Fig. 5. As can be seen, the corrections proposed to correct the apparent relationship between ΔpH and pH also change the depth distribution of the *excess alkalinity* ($A_X{}' = A_X - \Delta A_T + \Delta C_T$). Without any adjustments, there is a clear problem. At shallower depths, there are lots of negative values of this term, implying some combination of measurement errors in A_T and/or C_T that becomes unreasonably large the closer one gets to the surface. Once the proposed adjustments are incorporated, nearly all values of *excess alkalinity* are positive. Furthermore, there is little variation in the value with depth. As it is likely that there is also little variability in the systematic errors of A_T and C_T within a particular cruise data set, this observation implies that our *excess alkalinity* is itself fairly uniformly distributed with depth along P16N. The mean value of $A_X{}'$ for 2015 P16N is ~6.3 µmol kg^{-1}, as would be expected from the values of the adjustments proposed in Table 3. Also, the standard deviation of this value (2.2 µmol kg^{-1}) matches that inferred from the scatter of ΔpH for the same dataset ($2.2^2 \approx 1.7^2 + 1.0^2 + 1.0^2$).

It is interesting to note that Patsavas et al. (2015) reported values for $A_X{}'$ based on data from relatively shallow waters off the coast of California, in the Gulf of Mexico, and off the US East Coast that look in many ways similar to the distributions shown in the left-hand panels of Fig. 5, with an increasing proportion of negative $A_X{}'$ values towards the surface. Similar results were reported by Yang et al. (2015) for offshore waters in the Northern Gulf of Mexico.

A second line of evidence that suggests that the proposed adjustments cannot be dismissed out of hand is that they do not significantly affect the $p(CO_2)$ computed for the dataset reported by Lueker et al. (2000), and may even act to improve the agreement with measured values slightly (Fig. 6). With our proposed adjustments, the average percent relative discrepancy in $p(CO_2)$ improved from -1.30% to 0.76%

for $p(\mathrm{CO_2}) < 500$ μatm, while it increased slightly (< 0.3%) for $p(\mathrm{CO_2}) > 500$ μatm. Thus, although our adjustments do not entirely eliminate the $p(\mathrm{CO_2})$ discrepancy, particularly at high $p(\mathrm{CO_2})$, they do not contradict previous observations regarding the internal consistency of the $\mathrm{CO_2}$ system in seawater nor add meaningful additional uncertainty to the calculation of $p(\mathrm{CO_2})$ from A_T and C_T.

4.3. Uniqueness of solution

Of course, as was noted above – there is nothing unique about the proposed adjustments shown in Fig. 4b, nor can we be confident that they are accurate. Nevertheless, given the deviation of the original ΔpH dataset for 2015 P16N from "ideal," it is clear from an examination of Fig. 4b that a significant adjustment to the alkalinity as well as adjustments to K_1, K_2 and the B_T/S ratio at once reduces the observed slope in ΔpH against pH significantly, centers the ΔpH values around zero (see Fig. 2), and provides a fairly uniform depth distribution of $A_\mathrm{X}{'}$ (Fig. 5).

Some sense of the robustness of our conclusions can be drawn from a careful scrutiny of Fig. 4. An examination of Fig. 4a shows both the magnitude of the problem and the potential for a solution. The goal (for 2015 P16N) is to propose adjustments that simultaneously reduce the magnitude of the slope of the calculated ΔpH points shown in the top left panel of Fig. 2, and also reduce the average deviation of ΔpH from zero. As initially the data from this cruise exhibit a slope of –0.0257 and a mean ΔpH of –0.0092, this requires a net vector (Fig. 4b) that changes these values to close to zero (within our proposed tolerances, discussed elsewhere). It is apparent from an examination of Fig. 4a, which shows the effects of individual adjustments, that no single vector is even aimed in the right direction (even when one recalls that changing the sign of an adjustment will simply alter the direction of the vector by 180°). Nevertheless, it is apparent from Fig 4a, and even more clearly from Fig. 4b that one could (simplistically) consider the problem as a sum of three clearly independent vectors: one showing the implications of adjusting for systematic errors in pK_1; a second which can be thought of as a sum of three approximately co-linear effects resulting from adjusting for systematic errors in pK_2, pK_B, and B_T/S; and the third which is the sum of the effects resulting from systematic errors in the measurements of A_T and C_T as well as any effect resulting from omitting a component of A_T (which we have designated as A_X). Additionally, as discussed in section 4.1, the disagreement between two published sets of optical coefficients for *m*-cresol purple (DeGrandpre et al., 2014; Liu et al., 2011) has a small effect on the slope and mean ΔpH, and this effect is equivalent to a vector pointing in approximately the same direction as the ΔpK_2 in Fig. 4a, but with a third of the length. An examination of Fig. 4b, and of the angles involved, will indicate that these three vector combinations could – in principle – be combined in many ways to end up at the desired goal, by simply changing their magnitudes (lengths) to achieve this.

However, not all such combinations are equally realistic. For example, the likely systematic error on pK_1 is probably not large, as a number of independent studies (e.g., Mehrbach et al., 1973; Hansson, 1973; Roy et al., 1993; Mojica-Prieto & Millero, 2002; Millero et al., 2006) come up with very similar values of pK_1 (Millero et al., 2006; Millero, 2007) – exhibiting a total range of < 0.03 in pK_1 for the original measurements themselves at around 25 °C, and significantly less for the discrepancies between the various fitting functions. If we were to allow our estimate for the systematic error on pK_1 (at 25 °C) to double to 0.015, an examination of Fig. 4b would suggest that the other two vector combinations would each be shorter than shown, and that the magnitude of the third combination (the quantity we have called $A_\mathrm{X}{'}$) would be ~5 $\mu\mathrm{mol\ kg^{-1}}$, as opposed to 6.3 $\mu\mathrm{mol\ kg^{-1}}$ (Fig. 4b; Table 4). Insofar as the shipboard measurements of A_T and C_T are checked carefully against measurements on $\mathrm{CO_2}$-in-seawater reference materials (Dickson, 2010b) whose likely uncertainty is believed to be small (about 1 $\mu\mathrm{mol\ kg^{-1}}$ for measurements of either A_T or C_T), it seems very unlikely that this apparent discrepancy could be attributed entirely to systematic errors in A_T and C_T, especially when one considers that not only the magnitude, but also the sign of such errors would need to be such that they reinforce each other to this extent. A more rigorous assessment of the uncertainties for each of the various parameters in Table 1 will be needed before further progress might be made to inferring an unambiguous result.

4.4. How might the suggested adjustments vary from cruise to cruise?

The center panels in Fig. 2 show the effect of simply applying the adjustments found for 2015 P16N to all four of the data sets. Clearly, these improve the situation substantially, but are not an equally perfect solution for each of the data sets. There are residual slopes for two of the data sets (2014 P16S and 2016 I09N) and also a meaningful offset in the mean value of ΔpH, particularly for 2014 P16S.

If our proposed adjustments for 2015 P16N are indeed true for that region, then wherever one is in the oceans, it would be reasonable to expect that adjustments suggested to the constants K_1, K_2, and B_T/S should remain essentially the same. However, we are less confident that the distribution of either A_X or the systematic errors in measurements of total alkalinity or total dissolved inorganic carbon are necessarily uniform around the world, or from cruise to cruise. We therefore decided to look at the three other cruises (2014 P16S, 2016 I08S, and 2016 I09N), but now keeping the proposed adjustments to K_1, K_2, and B_T/S identical to those chosen for 2015 P16N, and only varying $A_\mathrm{X}{'}$.

Table 4 shows the resulting estimates of $A_\mathrm{X}{'}$ and the regression statistics that were obtained in this way for each of four cruises considered. Also, the right-hand panels in Fig. 2 show the effect of this approach to adjustment on each data set. It is apparent that the results are still significantly improved over the unadjusted data, and every cruise implies the existence of a significant amount of A_X. However, the 2014 P16S data set still does not attain our desired goals for slope and ΔpH (see next section).

4.5. The exception to test our rule?

As noted above, for the 2014 P16S dataset, we were unable to identify adjustments that achieved our goals ($+0.001 \geq \text{slope} \geq -0.001$; $+0.004 \geq \text{mean ΔpH} \geq -0.004$), whether with the adjustments of Table 3 (independently estimated for each cruise) or with those of Table 4 (using a single set of adjustments for the constants K_1, K_2, and B_T/S). The best we could achieve with the adjustments in Table 3 was a slope of 0.0013 and a mean ΔpH of ~0.008. The reasons for this discrepancy are not clear. One hypothesis is that there are significant cruise-to-cruise discrepancies in the analytical measurements, particularly of pH. However, an examination of deeper data (> 1,500 m) from a station common to both cruises (P16N Station 1 / P16S Station 88 – 16.4° S, 150° W) confirmed that probable cruise-to-cruise discrepancies could not be responsible for the residual slope and mean ΔpH of 2014 P16S. Additionally, a more detailed analysis of the deep water discrepancies from multiple cruise crossovers has been done for the Global Data Analysis Project version 3, and no adjustments were recommended for A_T and C_T for any of the cruises we examined (Are Olsen, personal communication).

We therefore decided to see if the term ΔpH estimated for the various cruises (after applying the adjustments in Table 4) showed any strong depth dependence (Fig. 7). With the exception of 2014 P16S, they do not. But, for these South Pacific data there is a clear difference at depths shallower than ~2,000 m (pH > ~7.65), where ΔpH increases significantly as one goes shallower. This suggests an apparent decrease in *alkalinity excess* for the same depth range (see Fig. 5) which has the effect of increasing the apparent slope of the relationship between ΔpH and pH. Thus, for this data set in particular, our assumption that A_X (and hence $A_\mathrm{X}{'}$) can be considered constant throughout a cruise (both in its depth variation, as well as along the cruise track) is

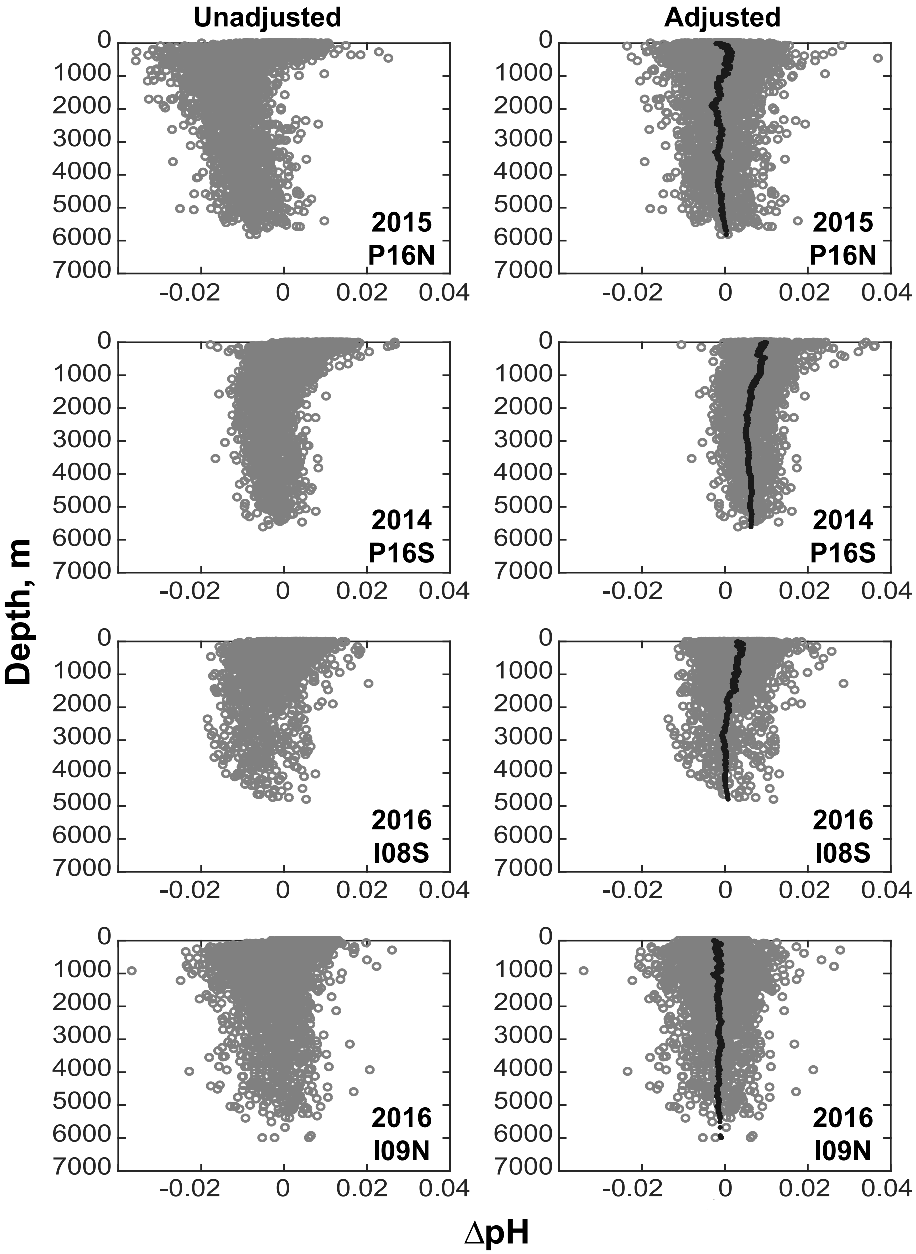


Fig. 7. Values of ΔpH calculated at 25°C and a gauge pressure of zero dbar plotted against depth for each cruise. The unadjusted model in Table 1 was used to calculate ΔpH in the left panels, while ΔpH in the right panels were calculated with a common set of adjustments for pK_1, pK_2, and B_T/S (–0.0074, +0.014, and –2.47%, respectively, as in Fig. 4b), but different *apparent excess alkalinity* for each cruise (as in Table 4). The solid black line is a 100-point running mean of the data.

problematic. As a result, it is not possible to achieve a simple distribution of ΔpH around a zero mean (as was found for 2015 P16N) without explicitly accounting for the changes in A_X.

4.6. A potential organic source for A_X

Organic acids and bases have been proposed as an unaccounted component in seawater total alkalinity (A_X) responsible for an observed apparent *excess alkalinity* (*e.g.,* Patsavas et al., 2015; Yang et al., 2015). One plausible candidate for an organic source for A_X in the open ocean may be carboxylic acids, which have been identified, through nuclear magnetic resonance (NMR) spectroscopy, as a ubiquitous and significant component of marine DOC, present in carboxylic-rich aliphatic matter (CRAM), which comprises a major fraction of refractory DOC, and in heteropolysaccharides (Hertkorn et al., 2006; Hertkorn et al., 2013). If the total concentration of organic bases (i.e., compounds with a p$K \geq$ 4.5, as in Dickson, 1981) is on the order of several μmol kg^{-1}, as suggested by our proposed A_X values, then this implies that functional groups with p*K*s ranging from 4.5–6 (within the range for carboxylic acids) would contribute significantly to A_X, as their basic forms comprise > 95% of their total concentration at seawater pH.

Hertkorn et al. (2013) characterized the composition of DOC isolated by solid-phase extraction (SPE) in the Atlantic, and their data suggests an upper limit for the concentration of carboxylic acids in the extracted DOC of ~3-5 μmol kg^{-1}, based on the total of all the ^{13}C NMR resonances that might be attributable to carboxyl groups. These values are therefore consistent with our proposed amounts of A_X in Table 3. It may also be possible that the concentration of carboxylic acids in marine DOC is higher than these values, as the extraction techniques used to isolate carboxylic acids (i.e., ultrafiltration and SPE) may not recover all carboxyl compounds in seawater. The DOC extraction efficiency of SPE varies with the type of resin used and can range from 8-

79% (Mopper et al., 2007), and neither SPE nor ultrafiltration techniques recover the low molecular weight, hydrophilic fraction of marine DOC, which may comprise up to 33% of the total DOC and is largely uncharacterized (Zigah et al., 2017).

The results of Hertkorn et al. (2013) also suggest that the distribution of A_X, if due to carboxyl groups, may not necessarily be uniform. At their sampling site in the Atlantic, the abundance of carboxyl groups, as inferred from their contribution to the total proton NMR integral, was observed to increase with depth, similar to the increase in A_X' with depth observed in the P16S dataset (Fig. 5). A detailed analysis of the distribution of A_X in the ocean may provide additional insights, but is beyond the scope of this paper.

5. Conclusions

It seems, from a careful evaluation of the data from a group of four GO-SHIP cruises, where measurements of pH, A_T, and C_T were made using state of the art techniques that, despite this, there are apparent inconsistencies between the measured values of both pH and total alkalinity, and the values of each of these two parameters as calculated from the other two measured parameters. Our evaluation of the discrepancies in pH for the GO-SHIP cruise 2015 P16N (in the North Pacific) leads us to the conclusion that these inconsistencies do not arise simply from the choice of CO_2 constants, but that they also indicate likely systematic uncertainties in the values of K_1, K_2, and B_T/S, as well as the widespread presence of an unidentified contribution to the measured total alkalinity (A_X) – likely from organic bases – that is usually ignored when using total alkalinity as one of the measured "CO_2 parameters" for open ocean seawater. These conclusions also apply to the other cruises we examined, although for 2014 P16S, it appears that there was a systematic distribution of A_X with depth that caused our simple approach (that assumed an essentially constant amount of A_X throughout the cruise region) not to work as well. Furthermore, when calculating $p(CO_2)$ from high-quality measurements of A_T and C_T, the dominant source of uncertainty for surface water conditions are in the equilibrium constants K_1 and K_2, rather than in the measurements (Orr et al., 2018). For changes in K_1 and K_2 of the magnitude given here, calculated $p(CO_2)$ would be expected to have an uncertainty of under 3%, as is seen in Fig. 6.

If our proposed adjustments to pK_1 (–0.0074), pK_2 (+0.014), and the B_T/S ratio (–2.47%), estimated from an examination of the data from 2015 P16N are used when calculating A_T from measurements of pH and C_T, the resulting estimates for apparent *alkalinity excess* ($A_{T\mathit{meas}}$ – $A_{T\mathit{calc}}$) seem plausible (both in amount and in vertical distribution), thus lending weight to the likely significance of such adjustments (bearing in mind the caution that the proposed adjustment to pK_2 is more correctly thought of as the sum of adjustments to pK_2 and to pK_B – see section 2.4). It should, however, be noted that the proposed revised values of pK_1 and pK_2 do not correspond to any of the various sets of constants examined in Fig. 3. Furthermore, adjustments to these constants alone cannot adequately correct the computed values of ΔpH for the seeming dependence on pH; however, incorporating an additional explicit correction for an *alkalinity excess* (assumed to be constant along a particular cruise track, though possibly different in different oceanic regions) does indeed make the data for three of our four cruises (2015 P16N; 2016 I08S; 2016 I09N) reasonably internally consistent (see Fig. 2 right-hand panels and Table 4).

The fourth cruise (2014 P16S) is more problematic. Our approach results in a noticeable residual slope and a clearly non-zero value for the mean value of ΔpH. A closer examination suggests that the distribution of the so-called *alkalinity excess* with depth computed from this dataset is not uniform, exhibiting a relatively constant value at depths below ~2,000 m, and lower values in the upper ocean. It seems that this non-uniform distribution is likely responsible for the residual slope and large offset in the mean ΔpH.

So what measurements are needed to better understand, and even resolve these observations? This is, perhaps, not straightforward, as the measurements discussed here (both the shipboard data and the various constants) are considered to be state-of-the-art, and it will not be simple to improve upon them. We suggest that a key first step may well be to verify that *thermodynamic consistency* in the sense described here can be achieved in a system where it is known that there is no additional, unidentified, acid-base system present (and hence no evidence of measurable *excess alkalinity*). Two alternatives exist: either measurements in synthetic seawater made up from carefully purified salts or measurements in a (once) natural seawater where an effort has been made to oxidize any residual organic material without leaving residues that affect subsequent measurements of pH, A_T and C_T. Another potential line of investigation is to examine additional GO-SHIP cruises (and other cruises where state-of-the-art measurements of pH, A_T, and C_T were made) and better assess how this putative A_X´ might be distributed around the world's oceans. In the past, it has usually been identified as a feature of coastal environments, for example Patsavas et al. (2015) suggest that $A_X´ \approx 4\ \mu\text{mol kg}^{-1}$ for waters with $S \leq 35$ (but without explicitly suggesting the possibility of other inconsistencies). It may be that a geographic (and depth) distribution would give clues as to the likely sources/sinks of this material. Finally, it may be practical to use methods such as that described by Cai et al. (1998) and Yang et al. (2015) who back-titrated seawater samples that had been stripped of CO_2 and interpreted their data imagining the organic alkalinity as due to a mixture of bases with differing p*K*s.

Another avenue to explore would be to better quantify the likely standard uncertainties of the measurements themselves and of the various constants used in the calculations. This may ultimately require new measurement approaches to reduce the uncertainties. If the uncertainties were well known, then it would be simpler to assess the significance of any observed inconsistencies.

Finally, we feel we should reiterate: the proposed adjustments do indeed improve the apparent thermodynamic consistency of the measurements described here. However, that – of itself – is not sufficient proof that the adjusted values are necessarily correct. Still, we feel that our insights should be considered either when using current GO-SHIP CO_2 data or when planning further work to acquire such state-of-the-art ocean CO_2 data.

Contributions

Both authors contributed to the conception of ideas for this research and to the writing of the manuscript. M.B. Fong made some of the measurements at sea and analyzed the cruise datasets.

Acknowledgements

We would like to thank the crew and scientists aboard the Ronald H. Brown, the Nathaniel B. Palmer, and the Roger Revelle, without which these cruises would not have been successful. Scientists from NOAA-AOML and NOAA-PMEL measured dissolved inorganic carbon on all of these cruises. We would also like to thank Robert H. Byrne, Marta Alvarez, and one anonymous reviewer for their critical comments that helped improve this manuscript. This work was supported by the US National Science Foundation (PLR1425989, OCE1657799, and OCE1436748, through a sub-contract from the University of Miami).

References

Cai, W.-J., Wang, Y., Hodson, R.E., 1998. Acid-base properties of dissolved organic matter in the estuarine waters of Georgia, USA. Geochimica et Cosmochimica Acta 62 (3), 473–483.

Carter, B.R., Radich, J.A., Doyle, H.L., Dickson, A.G., 2013. An automated system for spectrophotometric seawater pH measurements. Limnol. Oceanogr. 11 (1), 16–27.

Carter, B.R., et al., 2018. Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate. Limnol. Oceanogr. 16 (2), 119–131.

Chen, B., Cai, W.-J., Chen, L., 2015. The marine carbonate system of the Arctic Ocean:

Assessment of internal consistency and sampling considerations, summer 2010. Marine Chem. 176, 174–188.

Clayton, T.D., Byrne, R.H., 1993. Spectrophotometric seawater pH measurements: total hydrogen ion concentration scale calibration of m-cresol purple and at-sea results. Deep Sea Res. Part I 40 (10), 2115–2129.

Clayton, T.D., et al., 1995. The role of pH measurements in modern oceanic CO_2-system characterizations: Precision and thermodynamic consistency. Deep Sea Res. Part II 42 (2), 411–429.

DeGrandpre, M.D., et al., 2014. Considerations for the measurement of spectrophotometric pH for ocean acidification and other studies. Limnol. Oceanogr. 12 (12), 830–839.

DelValls, T.A., Dickson, A.G., 1998. The pH of buffers based on 2-amino-2-hydroxymethyl-1,3-propanediol ('tris') in synthetic sea water. Deep Sea Res. Part I 45 (9), 1541–1554.

Dickson, A.G., 1981. An exact definition of total alkalinity and a procedure for the estimation of alkalinity and total inorganic carbon from titration data. Deep Sea Res. Part A. Oceanogr. Res. Papers 28 (6), 609–623.

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

Dickson, A.G., 1992. The development of the alkalinity concept in marine chemistry. Marine Chem. 40 (1), 49–63.

Dickson, A.G., 1993. The measurement of sea water pH. Marine Chem. 44 (2), 131–142.

Dickson, A.G., 2010a. The carbon dioxide system in seawater: equilibrium chemistry and measurements. In: Riebesell, U., Fabry, V.J., Hansson, L., Gattuso, J.-P. (Eds.), Guide to Best Practices for Ocean Acidification Research and Data Reporting. Publications Office of the European Union Luxembourg, pp. 260.

Dickson, A.G., 2010b. Standards for ocean measurements. Oceanography 23 (3), 34–47.

Dickson, A.G., Millero, F.J., 1987. A comparison of the equilibrium constants for the dissociation of carbonic acid in seawater media. Deep Sea Res. Part A. Oceanogr. Res. Papers 34 (10), 1733–1743.

Dickson, A.G., Afghan, J.D., Anderson, G.C., 2003. Reference materials for oceanic CO_2 analysis: a method for the certification of total alkalinity. Marine Chem. 80 (2), 185–197.

Dickson, A.G., Sabine, C.L., Christian, J.R., 2007. Guide to Best Practices for Ocean CO_2 Measurements. PICES Special Publication 191 p.

Hansson, I., 1973. A new set of acidity constants for carbonic acid and boric acid in sea water. Deep Sea Res. Oceanogr. Abstr. 20 (5), 461–478.

Hertkorn, N., et al., 2006. Characterization of a major refractory component of marine dissolved organic matter. Geochimica et Cosmochimica Acta 70 (12), 2990–3010.

Hertkorn, N., Harir, M., Koch, B.P., Michalke, B., Schmitt-Kopplin, P., 2013. High-field NMR spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. Biogeosciences 10 (3), 1583–1624.

van Heuven, S., Pierrot, D., Rae, J.W.B., Lewis, E., Wallace, D.W.R., 2011. MATLAB program developed for CO_2 system calculations. ONRL/CDIAC-105b. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee.

Johnson, K.S., et al., 2016. Deep-Sea DuraFET: A pressure tolerant pH sensor designed for global sensor networks. Anal. Chem. 88 (6), 3249–3256.

Keeling, C.D., 1968. Carbon dioxide in surface ocean waters: 4. Global distribution. J. Geophys. Res. 73 (14), 4543–4553.

Kim, H.-C., Lee, K., Choi, W., 2006. Contribution of phytoplankton and bacterial cells to the measured alkalinity of seawater. Limnol. Oceanogr. 51 (1), 331–338.

Lee, K., Millero, F.J., Byrne, R.H., Feely, R.A., Wanninkhof, R., 2000. The recommended dissociation constants for carbonic acid in seawater. Geophys. Res. Lett. 27 (2), 229–232.

Lee, K., et al., 2010. The universal ratio of boron to chlorinity for the North Pacific and North Atlantic oceans. Geochimica et Cosmochimica Acta 74 (6), 1801–1811.

Liu, X., Patsavas, M.C., Byrne, R.H., 2011. Purification and characterization of meta-cresol purple for spectrophotometric seawater pH measurements. Environ. Sci. Technol. 45 (11), 4862–4868.

Lueker, T.J., Dickson, A.G., Keeling, C.D., 2000. Ocean pCO_2 calculated from dissolved inorganic carbon, alkalinity, and equations for K_1 and K_2: validation based on laboratory measurements of CO_2 in gas and seawater at equilibrium. Marine Chem. 70 (1), 105–119.

Martz, T.R., Connery, J.G., Johnson, K.S., 2010. Testing the Honeywell Durafet® for seawater pH applications. Limnol. Oceanogr. 8 (5), 172–184.

McElligott, S., et al., 1998. Discrete water column measurements of CO_2 fugacity and pH_T in seawater: A comparison of direct measurements and thermodynamic calculations. Marine Chem. 60 (1), 63–73.

Mehrbach, C., Culberson, C.H., Hawley, J.E., Pytkowicz, R.M., 1973. Measurement of the apparent dissociation constants of carbonic acid in seawater at atmospheric-pressure. Limnol. Oceanogr. 18 (6), 897–907.

Millero, F.J., 2007. The marine inorganic carbon cycle. Chem. Rev. 107 (2), 308–341.

Millero, F.J., 2010. Carbonate constants for estuarine waters. Marine Freshwater Res. 61 (2), 139–142.

Millero, F.J., et al., 2002. Dissociation constants for carbonic acid determined from field measurements. Deep Sea Res. Part I Oceanogr. Res. Papers 49 (10), 1705–1723.

Millero, F.J., Graham, T.B., Huang, F., Bustos-Serrano, H., Pierrot, D., 2006. Dissociation constants of carbonic acid in seawater as a function of salinity and temperature. Marine Chem. 100 (1), 80–94.

Mojica Prieto, F.J., Millero, F.J., 2002. The values of $pK_1 + pK_2$ for the dissociation of carbonic acid in seawater. Geochimica et Cosmochimica Acta 66 (14), 2529–2540.

Mopper, K., Stubbins, A., Ritchie, J.D., Bialk, H.M., Hatcher, P.G., 2007. Advanced instrumental approaches for characterization of marine dissolved organic matter: extraction techniques, mass spectrometry, and nuclear magnetic resonance spectroscopy. Chem. Rev. 107 (2), 419–442.

Müller, J.D., Rehder, G., 2018. Metrology of pH measurements in Brackish Waters—Part 2: experimental characterization of purified meta-Cresol purple for spectrophotometric pH_T measurements. Front. Marine Sci. 5 (177).

Orr, J.C., Epitalon, J.-M., Dickson, A.G., Gattuso, J.-P., 2018. Routine uncertainty propagation for the marine carbon dioxide system. Marine Chem. 207, 84–107.

Park, P.K., 1969. Oceanic CO_2 system: An evaluation of ten methods of investigation. Limnol. Oceanogr. 14 (2), 179–186.

Patsavas, M.C., Byrne, R.H., Wanninkhof, R., Feely, R.A., Cai, W.-J., 2015. Internal consistency of marine carbonate system measurements and assessments of aragonite saturation state: Insights from two U.S. coastal cruises. Marine Chem. 176, 9–20.

Roy, R.N., et al., 1993. The dissociation constants of carbonic acid in seawater at salinities 5 to 45 and temperatures 0 to 45°C. Marine Chem. 44 (2), 249–267.

Seidel, M.P., DeGrandpre, M.D., Dickson, A.G., 2008. A sensor for in situ indicator-based measurements of seawater pH. Marine Chem. 109 (1), 18–28.

Skirrow, G., 1975. The dissolved gases–carbon dioxide. In: Riley, J.P., Skirrow, G. (Eds.), Chemical Oceanography. Academic Press, New York, pp. 1–192.

Takahashi, T., et al., 1970. A carbonate chemistry profile at the 1969 Geosecs Intercalibration Station in the eastern Pacific Ocean. J. Geophys. Res. 75 (36), 7648–7666.

Uppström, L.R., 1974. The boron/chlorinity ratio of deep-sea water from the Pacific Ocean. Deep Sea Res. Oceanogr. Abstr. 21 (2), 161–162.

Walz, P.M., Takeshita, Y., Coletti, L.J., Jannasch, H.W., Johnson, K.S., 2018. Comparison of spectrophotometric and ISFET-potentiometric pH in seawater. In: Ocean Sciences Meeting, Portland, Oregon.

Wanninkhof, R., Lewis, E., Feely, R.A., Millero, F.J., 1999. The optimal carbonate dissociation constants for determining surface water pCO_2 from alkalinity and total inorganic carbon. Marine Chem. 65 (3), 291–301.

Williams, N.L., et al., 2017. Calculating surface ocean pCO_2 from biogeochemical Argo floats equipped with pH: An uncertainty analysis. Global Biogeochem. Cycles 31 (3), 591–604.

Yang, B., Byrne, R.H., Lindemuth, M., 2015. Contributions of organic alkalinity to total alkalinity in coastal waters: A spectrophotometric approach. Marine Chem. 176, 199–207.

Zigah, P.K., et al., 2017. Allochthonous sources and dynamic cycling of ocean dissolved organic carbon revealed by carbon isotopes. Geophys. Res. Lett. 44 (5), 2407–2415.