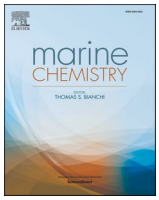
Marine Chemistry 250 (2023) 104247

Contents lists available at ScienceDirect

Marine Chemistry

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



Review article

Re-evaluation of carbonic acid dissociation constants across conditions and the implications for ocean acidification

Ryan J. Woosley *, Ji-Young Moon

Massachusetts Institute of Technology, Center for Global Change Science, Department of Earth, Atmospheric, and Planetary Sciences, Cambridge, MA 02139, USA

ARTICLE INFO

Keywords:
Ocean acidification
Internal consistency
Carbonic acid dissociation constants
Carbon system models

ABSTRACT

With the increasing threat of ocean acidification and the important role of the oceans in the global carbon cycle, highly precise, accurate, and intercomparable determination of inorganic carbon system parameters is required. Thermodynamic relationships enable the system to be fully constrained using a combination of direct measurements and calculations. However, calculations are complicated by many formulations for dissociation constants (over 120 possible combinations). To address these important issues of uncertainty and comparability, we evaluated the various combinations of constants and their (dis)agreement with direct measurements over a range of temperature (−1.9–40 °C), practical salinity (15–39) and pCO_2 (150–1200 μatm). The results demonstrate that differences between the calculations and measurements are significantly larger than measurement uncertainties, meaning the oft-stated paradigm that one only needs to measure two parameters and the others can be calculated does not apply for climate quality ocean acidification research. The uncertainties in calculated pH_t prevent climate quality pH_t from being calculated from total alkalinity (TA) and dissolved inorganic carbon (DIC) and should be directly measured instead. However, climate quality TA and DIC can often be calculated using measured pH and DIC or TA respectively. Calculations are notably biased at medium-to-high pCO_2 values (~500–800 μatm) implying models underestimate future ocean acidification. Uncertainty in the dissociation constants leads to significant uncertainty in the depth of the aragonite saturation horizon (>500 m in the Southern Ocean) and must be considered when studying calcium carbonate cycling. Significant improvements in the precision of the thermodynamic constants are required to improve pH_t calculations.

1. Introduction

The oceans play a vital role in Earth's carbon cycle. The oceans are the largest reservoir of carbon on Earth's surface (Canadell et al., 2021; Houghton, 2007) and they act to mitigate climate change through the uptake and storage of anthropogenic carbon emissions (Caldeira and Wickett, 2003). This anthropogenic carbon storage acts to change the chemistry of the oceans through the dissociation of carbonic acid, resulting in a decrease in ocean pH and carbonate ion concentration (Doney et al., 2009), a process termed ocean acidification (OA). Such alterations in the ocean chemistry are known to have many, often detrimental, effects on ocean life (Andersson et al., 2015; Busch et al., 2015; Doney et al., 2020), particularly organisms that produce calcium carbonate ($CaCO_3$) shells (Gazeau et al., 2013; Sutton et al., 2015). The inorganic carbon system is complex (Millero, 2007) and monitoring or modeling changes in ocean chemistry, along with the resulting impacts on ocean life and ecosystems, requires highly accurate and precise measurements that are comparable across locations, laboratories, and decades (Newton et al., 2015; Woosley, 2021a).

The inorganic carbon system can be directly observed through four measurable parameters, total alkalinity (TA), dissolved inorganic carbon (DIC), partial pressure or fugacity of CO_2 (pCO_2 and fCO_2 respectively), and -log($[H^+]$) (pH), or can be calculated using thermodynamic relationships (Park, 1969) along with ancillary parameters such as individual ion concentrations and $CaCO_3$ saturation states. For simplicity, pH on the total scale (pH_t) will be used here (Dickson, 1993; Dickson et al., 2016). This ability has led to the mantra that one only needs to measure two parameters and the rest can be calculated. In principle any two can be measured, but in practice some pairs are better than others (McLaughlin et al., 2015; Millero, 2007) as a result of differences in the propagation of uncertainties for the different calculations (Orr et al., 2018). Historically TA and DIC are the most often measured (Olsen et al., 2016) despite the large uncertainties that result in calculated pH_t and fCO_2 (Millero, 2007). However, development and improvements in

* Corresponding author.
E-mail address: rwoosley@mit.edu (R.J. Woosley).

https://doi.org/10.1016/j.marchem.2023.104247
Received 6 April 2023; Received in revised form 12 May 2023; Accepted 14 May 2023
Available online 16 May 2023

spectrophotometric (spec.) pH_t (Carter et al., 2013; Clayton and Byrne, 1993; Liu et al., 2011) have resulted in a large increase in pH_t measurements in the last few decades.

The rapid increase in the number of pH_t measurements has allowed a better understanding of the inorganic carbon system model and associated uncertainties through over-determinization of the system (Álvarez et al., 2020; Carter et al., 2018; Fong and Dickson, 2019; Woosley, 2021a; Woosley et al., 2017). The result has been the discovery of a pH_t (and by extension pCO_2) dependent discrepancy when comparing calculated and measured pH_t values (Carter et al., 2018). Many explanations have been explored, including uncertainty in the dissociation constants (K_i*), measurement uncertainty, or an unidentified component of alkalinity, but the issue remains unresolved (Álvarez et al., 2020; Fong and Dickson, 2019; Takeshita et al., 2021). Recent work has also renewed debate in the presence of organic or unidentified alkalinity (A_x) in open ocean seawater (Álvarez et al., 2020; Fong and Dickson, 2019; Hunt, 2021; Kerr et al., 2021; Millero et al., 2002; Sharp and Byrne, 2020, 2021). A non-negligible A_x would bias calculations involving TA as it would overestimate the carbonate alkalinity.

Although the above inconsistencies are small relative to background concentrations, they are significant enough to have important implications for studying the marine carbon cycle and OA. In the open surface ocean, pH is decreasing at a global average rate of approximately -0.0018 units yr^{-1} (Jiang et al., 2023; Waters et al., 2011). However, in coastal and coral reef systems OA is highly variable in both time and space (Shaw et al., 2012). The different rates lead to different levels of precision and accuracy requirements for studying OA in different environments. Such requirements resulted in the development of 'weather' and 'climate' quality carbon measurement criteria (Newton et al., 2015). Weather quality research requires uncertainties of approximately 0.02 in pH_t and 10 μmol kg_{sw}^{-1} in TA and DIC and is suitable for highly variable environments. Climate quality research is more precise and required in open ocean systems, with uncertainties of 0.003 in pH_t and 2 μmol kg_{sw}^{-1} in TA and DIC (Newton et al., 2015). All inorganic carbon system parameters can be measured to climate level precision when best practices are followed (Dickson et al., 2007), but that is not always the case for calculated values (Orr et al., 2018).

A significant source of uncertainty in calculated values comes from the large number (>20) of formulations for the carbonic acid dissociation constants and other parameters needed to perform the calculations. Currently there is no consensus on which constants are 'best' or exactly how comparable measured and calculated values are. The lack of consensus has significant implications for the long-term study of OA and comparability among studies, specifically for modeling of OA. Most carbon system models use TA and DIC due to their semi-conservative nature (Orr and Epitalon, 2015) and calculate pH_t. However, given the uncertainty in calculated pH_t from this pair, it is unknown if that modeled pH_t can achieve climate level precision or the implications if it cannot.

Internal consistency is a common method to evaluate the carbon system model (constants) (Byrne et al., 1999; Lee and Millero, 1997; Wanninkhof et al., 1999). A system is said to be internally consistent when the same value is determined (within known uncertainties) regardless of the method used to determine it. Here, internal consistency is determined by comparing measured and calculated carbon parameters, namely pH_t. Internal consistency is achieved when the calculated values agree with the measured values to within measurement uncertainty (~0.003 for pH_t). This work expands on the work of Woosley (2021a) by evaluating the constants and calculations over a range of practical salinity (S_p), temperature, and pCO_2 (15–40, −1.9–40 °C, and ~ 150–1200 μatm respectively). The results are then used to make recommendations for measurements and calculations as well as the potential biogeochemical implications.

2. Methods

The methods used here are very similar to those used in Woosley (2021a) with modifications to cover a range of S_P and pCO_2. Natural seawater was used for all experiments. Twenty-six batches of seawater were prepared with S_p between 15 and 39 and pCO_2 ~ 150–1200 μatm at room temperature. Batch 1 was a test batch and will not be discussed further. Batches 2 and 3 were prepared by modifying the pCO_2 or S_p of certified reference material (CRMs). The remaining batches were prepared using oligotrophic Atlantic surface seawater collected from 39° 46.406′ N and 70° 53.065′ W on October 10, 2019, with $S_p = 35.67$. For each batch, a total of 20 L of water was prepared by adjusting nominally to the desired S_p via dilution with 18.2 MΩ ultra-pure water or through slow evaporation by gentle heating in a gravity convection oven at 60 °C. Water was filtered through a 0.2 μm filter (either a Millipore Optiscale 47 capsule 0.5/0.2 μm filter or Pall Scientific Supor Acropak 800 0.8/0.2 μm filter) twice; first when preparing and mixing the batch and second when transferring the batch to the equilibration bottle. Water was poisoned with ~0.02% *v*/v saturated mercuric chloride when preparing the individual batches to prevent biological growth.

The pCO_2 (and DIC) of each batch was modified by bubbling with gas of the desired molar fraction of CO_2 (xCO_2), nominally 150, 400, 600, 800, or 1200 ppm with the balance air (Airgas Inc., Billerica, MA USA). The gas was flown through a 1 L borosilicate glass gas washing bottle containing ultra-pure water in order to moisten the gas and minimize evaporation of the seawater during gas equilibration. The gas then flowed through the batch of seawater contained in a 20 L borosilicate glass gas washing bottle. During equilibration, the water was mixed vigorously using a magnetic stir bar. The gas washing bottle was wrapped in ¼ inch neoprene insulating foam to minimize temperature variability during bottling. The temperature of the water was generally stable to ~0.2 °C during bottling. The outflow gas was passed through a Li-Cor 865 (Li-Cor Inc., Lincoln, NE USA) in order to monitor the xCO_2 of the seawater and determine when the batch was equilibrated. The Li-Cor used factory calibration and did not follow best practices for flowing pCO_2 seawater measurements (Pierrot et al., 2009); therefore, the accuracy of the measurements is unknown and cannot be used in calculations for analyzing the results. Once the xCO_2 became stable, the gas was allowed to flow for an additional one to two hours to ensure the batch was well mixed before bottling.

Once equilibrated, samples were collected for measurement of S_p, and nutrients (reactive silicate and phosphorus) into acid washed 50 mL falcon tubes. After bottling, a second sample was collected for S_p. The S_p was measured using a Mettler-Toledo LLC (Columbus, OH USA) conductivity probe (InLab 731-SM) calibrated immediately prior to analysis using a Yellow Spring Instrument (Yellow Springs, OH USA) 50 Siemen conductivity standard. It was determined that the accuracy of the conductivity probe was lower than expected. Thus, starting with batch 10, salinity was measured using both the conductivity probe and an Anton-Paar GmbH (Graz, Austria) DMA35 vibrating tube densitometer. Density was converted to practical salinity using the TEOS-10 Matlab software package (McDougall and Baker, 2011). The average of the salinity collected at the beginning and end of bottling was used for all calculations. For batches 2–10, S_p from the conductivity probe was used and for batches 11–26, S_p from the densitometer was used. The accuracy and precision of the conductivity probe S_p are estimated to be 0.5 and 0.01, and for the densitometer S_p, 0.1 and 0.01 respectively. All salinity values are practical salinity, thus salinity and S_p are used interchangeably.

Nutrients (reactive silicate and phosphorus) were measured immediately after bottling using the molybdate colorimetric method (Strickland and Parsons, 1972) modified according to NOAA National Centers for Environmental Information (NCEI) protocols (https://www.nodc.noaa.gov/archive/arc0001/9900162/2.2/data/0-data/jgofscd/Files/protocols/). An Agilent Technologies Inc. (Santa Clara, CA USA) 8454 spectrophotometer with a 1 cm pathlength cell was used. For reactive phosphorus the limit of detection is estimated to be 0.01 μmol kg_{sw}^{-1}, and

values less than or equal to that were considered zero for all calculations. The average standard deviations were 0.05 and 0.13 for phosphate and silicate respectively. The values for each batch can be found in Table S1. The uncertainty in nutrients does not make a significant contribution to the uncertainty in the calculations.

The batches were bottled into 250 mL borosilicate glass (Pyrex) reagent bottles (Corning Inc. Corning, NY USA). Prior to use, bottles and caps were washed using Micro-90 cleaning solution and baked in a muffle furnace at 560 °C for 6.5 h. To fill the bottles, water was first allowed to flow for ~2 s in order to clear the silicone sampling tube. The bottle was then filled from the bottom without rinsing, similar to how CRMs are filled (personal communications, Andrew Dickson, University of California, San Diego). Differing from how CRMs are filled, once filled to the brim the sampling tube was removed with water still flowing in order to keep the bottle completely full. A precise and reproducible headspace was then created using a pipette to remove 8.48 mL, allowing for ~1% by volume headspace once the cap was inserted. After pipetting out the headspace, a cap pre-greased with Apiezon L vacuum grease (Manchester, United Kingdom) was placed in the bottle creating a gas tight seal. The cap was then held in place using a plastic hose clamp and rubber band. A total of 44 bottles per batch were filled. Bottling took ~45 min. The xCO_2 was continuously monitored throughout bottling to check for drift. All batches were stable with a standard deviation of xCO_2 during bottling ≤1.2 ppm except for batch 25 which showed a slight trend and had a standard deviation of 1.8 ppm. The DIC and pH_t of batch 25 also indicate a trend, which is evidenced by the larger uncertainty in measurements from this batch (Table S1). Batch 26 also had a very slight trend, but the standard deviation was 0.63 ppm. A trend was not evident in the DIC and pH_t of batch 26 (Table S1).

Bottles were numbered in the order in which they were filled. The bottles were then randomly divided into 11 sets of 4 using a random number generator (random.org) with one bottle in each set coming from bottles 1–11, 12–22, 23–33, and 33–44. This method ensured that bottle treatments were randomly assigned, but also that the bottles were distributed throughout the bottling process to account for any variability in the batch during bottling. From each batch, one set of 4 bottles was used to measure TA, 1 set to measure DIC, and 9 sets to measure pH_t at 9 different temperatures.

The TA was measured by open cell HCl titration using a custom system designed and built by the laboratory of Andrew Dickson (University of California, San Diego). The instrument is the same type used to certify CRMs for TA (Dickson et al., 2003). Samples were analyzed for TA within 2 months of bottling. The DIC was measured via coulometry (Johnson, 1992) by the Carbon Program at NOAA Pacific Marine Environmental Laboratory. DIC samples were analyzed within one year of bottling. The mean standard deviation of TA and DIC measurements was 0.99 and 1.17 μmol kg_{sw}^{-1} respectively. For purposes of discussion an uncertainty of 2 μmol kg_{sw}^{-1} is used for both, although for some batches the standard deviation was larger (Table S1).

The remaining bottles were used to measure pH_t over the oceanic range of temperatures (−1.9 to 40 °C). One set of four bottles for each batch was measured at 40, 35, 30, 25, 20, 15, 10, 5 °C and approximately 0.2 °C above the freezing point of the batch, which varied with salinity. The bottles were equilibrated to the measurement temperature (± ~0.1 °C) for a minimum of one hour once the bath temperature stabilized prior to analysis in a Versacool refrigerated recirculating bath (ThemoFischer Scientific, Newington, NH USA). The pH_t at each temperature was measured using a custom designed automated spec. pH system similar to Carter et al. (2013) and described in more detail in Woosley (2021a). Two batches of purified meta-cresol purple (mCp) obtained from the laboratory of Dr. Robert H. Byrne (University of South Florida) were used. Indicator batches 1 and 3 (as described in Woosley (2021b)) were used, and the purity of the indicator was confirmed following the recommendations of (Woosley, 2021b). To ensure continuity between indicator batches, measurements of TRIS were made at the switchover, and agreed to within 0.0003 (N = 4). For every treatment, the set of four bottles was each measured four times for a total of 16 measurements per condition. The pH_t was calculated using the equations of Liu et al. (2011) and the mean temperature for each set of measurements. The temperature during analysis was measured using a Fluke Calibration (Everett, WA USA) 1523 platinum resistance thermometer located in the water-jacketed cell holder immediately adjacent to the 10 cm quartz micro-volume spectrophotometer cell (Starna Cells Inc., Atascadero, CA USA). The temperature was generally stable to ~0.01 °C during analysis. The amount of indicator was varied for each measurement, alternating starting with high or low amounts of indicator between bottles. This procedure allowed for the correction of the pH_t for the perturbation caused by the indicator itself, and to detect exchange of CO_2 between the bottle and lab atmosphere. For temperatures 35 and 40 °C, gas exchange was detected in the third and fourth measurements from each bottle (reps C and D). These values have been excluded from analysis for all batches. These two temperatures thus have only 8 measurements per batch and generally have a higher uncertainty than the lower temperatures (Table S1).

The perturbation of pH_t caused by the addition of the indicator was adjusted individually for each condition (Woosley, 2021a). The isosbestic point (488 nm) was fit to pH_t using the MatLab 'fitlm' function with RobustFit 'on.' Since RobustFit was used, primary quality control was very conservative under the assumption that questionable points would receive a lower weight during fitting and minimize their influence. The intercept of the fit is the pH_t with zero indicator and the value used for all calculations. The standard error in the intercept was taken to be the uncertainty in the measured pH_t value (Table S1). It is important to note that this is a measure of the uncertainty in the measurement and is not the same as the instrument precision often reported. This uncertainty includes instrument precision as well as other sources of uncertainty including uncertainty in the dye perturbation correction and bottle to bottle variability. The precision of the instrument is 0.0005 as determined from measurements of TRIS and CRMs. The spectrophotometer bandwidth and absorbance (A) calibrations were verified using National Institute of Standards and Technology (NIST) traceable certified references. The spectral bandwidth was checked with a holmium glass filter and A was checked using neutral density filters (both at 440, 546, 635, 434, 578, and 730 nm wavelengths; Starna Cells, Inc.) approximately monthly. The accuracy of the spectral bandwidth was within manufacturer specifications (< 0.5 nm), and the absorbances were within the expanded uncertainty of the reference material over an absorbance range of 0.3–1.2 (± 0.0027 A). The raw absorbance values for all measurements including quality flags along with pH_t, TA, DIC, S_p, and nutrients for each batch are archived in BCO-DMO (https://www.bco-dmo.org/project/813194).

The mean standard error for pH_t was 0.0032. For purposes of discussion, a measurement uncertainty of 0.003 is used. There are some conditions where the standard error was larger than 0.003. Generally, the conditions with higher uncertainties are high temperatures (≥ 35 °C), low S_p (≤ 20), or batch 25 and 26 as noted above. As discussed further in the results and discussion, most of these conditions are outside the valid range of the pH_t indicator calibration and are excluded from analysis. However, some of the scatter can be explained by conditions with higher uncertainty.

These experiments used purified indicator, but the broad conclusions are likely applicable to measurements made with impure indicator although with significantly higher uncertainties. A full evaluation of impure indicators is beyond the scope of this work.

2.1. Data analysis

The pH_t was calculated using the CO2sys version 3.1.1 for Matlab (Sharp et al., 2020) from TA and DIC along with temperature, S_p, and nutrients. All calculated pH values are on the total scale. Calculations were performed using all possible combinations of 10 different sets of carbonic acid dissociation constants: Mehrbach et al. (1973) as refit by

Dickson and Millero (1987) or by Lueker et al. (2000) as well as Cai and Wang (1998), Goyet and Poisson (1989), Millero (2010), Millero et al. (2002, 2006), Mojica Prieto and Millero (2002), Schockman et al. (Submitted), and Waters and Millero (2013). In addition, three bisulfate dissociation constants, K_{HSO4-} (Dickson, 1990; Khoo et al., 1977; Waters and Millero, 2013), two hydrogen fluoride dissociation constants, K_{HF} (Dickson and Riley, 1979; Perez and Fraga, 1987), and two values for the total boron (B_T) concentration (Lee et al., 2010; Uppström, 1974) resulting in 120 possible combinations of constants needed to calculate the inorganic carbon system parameters. These are the same constants evaluated previously (Woosley, 2021a) but with the addition of Schockman et al. (Submitted) which are newly determined, Cai and Wang (1998) which are often used in low salinity waters, and Goyet and Poisson (1989) which were determined in artificial seawater, but are still sometimes used in polar regions because some studies have found them to be the most internally consistent for high latitudes (Laika et al., 2009; Wanninkhof et al., 1999). Schockman and Byrne (2021) were determined from 15 to 35 °C, however, they have extended the constants to low temperatures, and we were provided with a prepublication version of the constants for evaluation (Katelyn Schockman, personal communication), therefore we use their constants over the full temperature range (Schockman et al. Submitted). The internal consistency of each calculation was then determined by subtracting the calculated pH_t from the measured pH_t ($\Delta pH_t = pH_{t(meas)} - pH_{t(calc)}$). A positive value means the calculation underestimates the measured value and a negative value means the calculation overestimates the measured value. The closer to zero, the more internally consistent the pH_t is. The internal consistency of DIC and TA was also examined, using the same procedure as for pH_t. However, only the pH_t and TA internal consistency will be discussed in detail.

In the definition of TA (Dickson, 1981), organic alkalinity is generally assumed to be either not present or present at negligible quantities. Most carbon system calculation software like CO2sys use this assumption and do not include a term for A_x. To assess the impact that a hypothetical A_x component would have on internal consistency, the above calculations were carried out again with a defined amount of A_x. For simplicity, and because CO2sys does not have a way to directly include additional forms of alkalinity, a defined amount of A_x was subtracted from the measured TA. Millero (2007) found that a hypothetical A_x of 8 μmol kg_{sw}^{-1} improved internal consistency of fCO_2, particularly at high fCO_2 values. Fong and Dickson (2019) found that A_x of 4 0.3 μmol kg_{sw}^{-1} was needed after making other adjustments to achieve the most internally consistent pH_t. Sharp and Byrne (2021) found values of A_x of approximately −5–10 μmol kg_{sw}^{-1} using Gulf of Mexico water and CRMs. For the calculations done here, it is assumed that the starting seawater ($S_p = 35.67$) had 4 μmol kg_{sw}^{-1} of A_x. The A_x of the batches was therefore in proportion to salinity:

$$A_x = 4/35.67^{*} S_p \quad (1)$$

assuming that the ultra-pure water used for dilution contained no alkalinity and that gentle heating for the high salinity batches did not result in a loss or alteration of A_x. No direct measurements of A_x were made (Kerr et al., 2021), making this a purely hypothetical scenario. The value of 4 μmol kg_{sw}^{-1} was chosen based on Fong and Dickson (2019) and is similar to the amount estimated by Sharp and Byrne (2021) for CRMs when using B_T of Lee et al. (2010). A value of 8 μmol kg_{sw}^{-1} was tested and found to produce a very large (and unrealistic) over correction. It should be noted that under this scenario A_x decreases with salinity, but that does not necessarily represent what happens in the natural ocean where coastal regions with low salinity can have high values of A_x (Ko et al., 2016).

3. Results and discussion

The internal consistency of the measurements was evaluated by comparing measured and calculated values. There is a large uncertainty in the accuracy of pH_t of >0.01 (Dickson, 1993) due to methodological limitations of potentiometric measurements in seawater as well as for calibrating the indicator for spec. pH_t measurements. This uncertainty in accuracy is significantly larger than the measurement uncertainty. The reason for this large discrepancy is described in detail in several other publications (DelValls and Dickson, 1998; Dickson et al., 2016). For this analysis, internal consistency is said to exist when the measured and calculated values agree within the measurement uncertainty as determined in the prior section (generally ~0.003 for pH_t and 2 μmol kg_{sw}^{-1} for TA and DIC), and agrees well with the requirements of climate quality research (Newton et al., 2015). It cannot be emphasized enough that internal consistency is not a measure of accuracy. Therefore, both the measurements and calculated values may agree, but could be off from the 'true' value. Alternately a set of constants could be regarded as not internally consistent, however the calculated values could be more accurate than a set of constants determined to be internally consistent (e.g. as a result of ignoring A_x). This is especially true for pH_t due to the large difference in uncertainties in accuracy and precision. Despite the limitations of internal consistency, it is still extremely useful because studying the marine carbon cycle and specifically OA relies on detecting changes over time. Such analysis requires highly consistent and comparable values across decades and many different laboratories and methods.

The 225 different conditions and 120 possible combinations of constants produces hundreds of figure panels. Only a few representative examples will be given here, however, additional figures and raw data needed to create them are available in the supplemental material and archived at the biological & chemical oceanography data management office (https://www.bco-dmo.org/project/813194). For ease of discussion, the various constants will be abbreviated by authors' last initial(s) and year of publication (e.g. Millero et al. (2002) will be referred to as M02). The study of Dickson and Millero (1987) provided multiple fits to different data. Since only the refit of Mehrbach et al. (1973) is discussed here it will be referred to simply as DM87. A full list of the constants, their abbreviations, ranges of applicability, estimated uncertainty, and other pertinent information can be found in Table 1.

No set of constants is internally consistent (within the uncertainty of the measurements, ± 0.003) over the full range of conditions evaluated (Fig. S1). Yet, there are conditions under which each set of constants is internally consistent. For all sets of constants, there are conditions under which the results are outside the range of the large uncertainty in the accuracy of pH_t (0.01–0.02). Excluding conditions outside the range that the indicator is calibrated, ΔpH_t ranges from −0.06 to 0.06. Excluding the constants of M02, which have a very limited salinity range only reduces the range to −0.05 to 0.05. This scatter is of similar magnitude to what is found in field measurements (Carter et al., 2018). On average the calculations tend to overestimate measured pH_t (i.e., negative ΔpH_t) although positive values are not uncommon.

Fig. 1 shows all the results limited to conditions under which the indicator is well calibrated. The results here broadly agree with those of Woosley (2021a) over similar conditions (open surface ocean in near equilibrium with the current atmosphere). Also, in agreement with that work, using the B_T concentration of L10 generally produces more internally consistent results, but there are some exceptions. Most exceptions involve the constants of GP89 or S23. Calculations using the B_T of U74, produce higher calculated pH_t values (more negative internal consistency values) than those using L10. This result is expected since the B_T of U74 is lower than L10 resulting in a higher carbonate alkalinity when using U74. The impact of B_T on the calculations will be discussed in more detail in the following sections. Except where noted, further discussion uses the B_T of L10.

Three different values of K_{HSO4-} were considered (Table 1). Woosley (2021a) found the differences to be insignificant but noted that it may become significant in cold waters at high salinities. The results here agree with Woosley (2021a); the differences between the three K_{HSO4-}

Table 1
References, abbreviations, and applicable ranges along with other pertinent information for each constant or quantity evaluated in this study.

Reference	Abbreviation	Salinity range	Temp. range (°C)	Stated uncertainty[a]	Original data/ refit[b]	Media
Carbonic Acid Dissociation Constants (K^*_1 and K^*_2)						
Mehrbach et al. (1973) refit by Dickson and Millero (1987)[c]	DM87	19–40	2–35	pK^*_1 0.011 pK^*_2 0.020	R	Real seawater
Lueker et al. (2000)[c]	L00	19–40	2–35	pK^*_1 0.011 pK^*_2 0.020	R	Real seawater
Mojica Prieto and Millero (2002)	MPM02	12–45	5–45	pK^*_1 0.011 pK^*_2 0.020	O	Real seawater
Millero et al. (2002)	M02	33–37	−1.6–38	pK^*_1 0.010 pK^*_2 0.016	O	Real seawater
Millero et al. (2006)	M06	0.1–50	1–50	pK^*_1 0.014 pK^*_2 0.022	O	Real seawater
Millero (2010)	M10	0.1–50	1–50	pK^*_1 0.016 pK^*_2 0.022	R	Real seawater
Waters and Millero (2013) (with Waters et al., 2014)	WM13	0.1–50	1–50	pK^*_1 0.014 pK^*_2 0.022	R	Real seawater
Schockman et al. (Submitted)[e]	S23	19.6–41	15–35	pK^*_1 0.014 pK^*_2 0.010	O	Real seawater
Cai and Wang (1998)	CW98	0–40	2–35	pK^*_1 0.030† pK^*_2 0.080†	R	Real and Artificial Seawater
Goyet and Poisson (1989)	GP89	10–50	−1–40	pK^*_1 0.01 pK^*_2 0.02	O	Artificial Seawater
Total Boron Concentration						
Uppström (1974)	U74	–	–	B/Cl 0.005	O	Real seawater
Lee et al. (2010)	L10	–	–	B/Cl 0.0009	O	Real seawater
Bisulfate Dissociation Constant						
Khoo et al. (1977)	K77	20–45	5–40	$\ln(K_{HSO4^-})$ 0.042	O	Artificial Seawater
Dickson (1990)	D90	5–45	0–45	$\ln(K_{HSO4^-})$ 0.042	O	Artificial Seawater
Waters and Millero (2013)	WM13	5–45	0–45	$\ln(K_{HSO4^-})$ 0.05†	R	Artificial Seawater
Hydrogen Fluoride Dissociation Constant						
Dickson and Riley (1979)[d]	DR79	26.7 and 34.6	25	unknown†	R	Real seawater
Perez and Fraga (1987)	PF87	10–40	9–33	$\ln(K_{HF})$ 0.10	O	Real seawater
meta-Cresol Purple (mCp) Indicator						
Liu et al. (2011)	L11	20–40	5–35	pH_t 0.001 pH_t 0.002–0.01†	O	TRIS Buffers
Müller and Rehder (2018)	MR18	5–35	0–40	(S_P dependent)	O/R	TRIS Buffers
Douglas and Byrne (2017)	DB17	5–35	0–40	pH_t 0.004†	R	TRIS Buffers

[a] As determined individually by each study. Dissociation constants are reported as 2 times σ. For most studies, σ is a measure of the goodness of fit, but methods of determination may vary by study and may not be directly comparable. The uncertainties for MR18 and DB17 are generally relative to L11.
† Indicates uncertainty in parameter was unclear in the original study and a best estimate is given here where possible.
[b] O = Original Data, R = Refit.
[c] Woosley (2021a) incorrectly stated the salinity range as 19–35.
[d] Original data at 2 salinities and 1 temp, extended using enthalpy and infinite dilution data, but not stated over what range equations are applicable. Real seawater is assumed, though not explicitly stated.
[e] Prepublication version of equations provided by Dr. Katelyn Schockman (University of South Florida).

are generally comparable to or smaller than spec. pH_t instrument precision (0.0005), even at low temperatures in high salinity water. D90 and WM13 agree with each other better than with K77, in agreement with earlier work (Millero, 1995; Waters and Millero, 2013; Woosley, 2021a). For consistency with earlier studies and the recommendation of recent modeling work (Humphreys et al., 2022), further discussion will use D90. Results for all K_{HSO4^-} can be found in the supplemental material.

The K_{HF} is also important for these calculations. Woosley (2021a) found small but non-negligible differences between the options included in CO2sys (DR79 and PF87), and the results here agree. The differences are small but can be several times larger than the pH_t instrument precision (0.0005). The maximum difference is $\sim|0.0019|$, making K_{HF} secondary to the B_T and carbonic acid dissociation constants. In general, the differences are largest at high temperatures (≥30 °C) for low salinities (≤ 35) and low temperatures (≤ 10 °C) for high salinities (40). For constants determined on the total pH scale (L00 and S23) the choice does not matter (when using the total scale) as it cancels out in the calculations, contributing a rounding error at most. It is unknown which K_{HF} is more accurate and there is no clear reason to choose one over the other. Although, PF87 does include more direct measurements over a broader range of temperature and salinity, neither covers the full range of temperature and salinity considered here. The applicable range of DR79 is unclear from that paper (see Table 1). Averaging over all calculations the mean, median, and standard deviation are the same. For consistency with Woosley (2021a) further discussion will use PF87 (unless otherwise stated), but under certain conditions, particularly those noted above, the differences are non-negligible. At $S_p = 35$, DR79 tends to be slightly better than PF87. Therefore, K_{HF} should be considered when estimating uncertainties. DM87 and L00 are both based on the same data (Mehrbach et al., 1973) but on different pH scales. They are nearly identical over their applicable salinity range (19–40). Differences can generally be attributed to the choice of K_{HF}, since treatment of Fluoride is the main difference between the scales. There are also

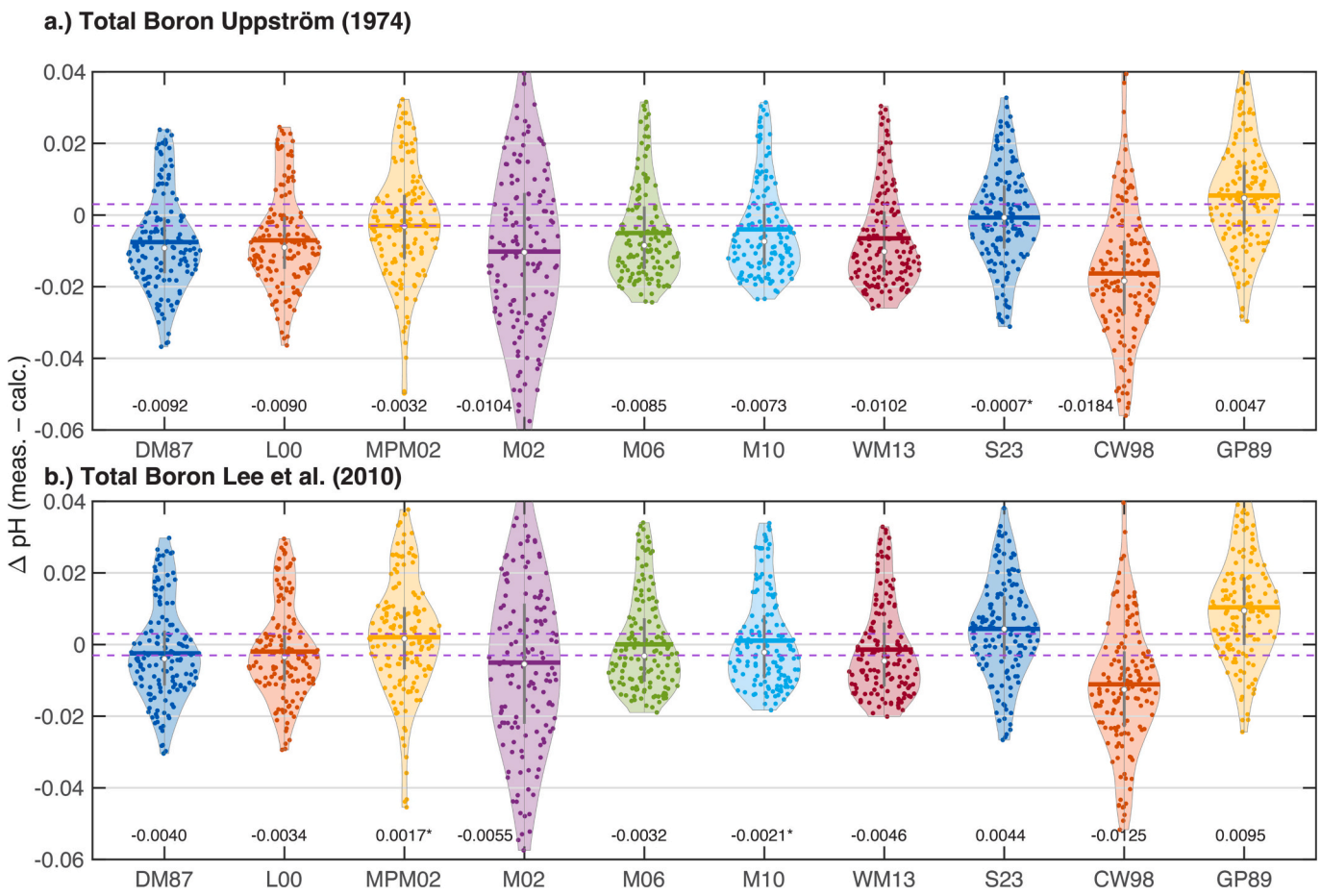


Fig. 1. Violin plots showing pH_t internal consistency for all conditions for which the indicator is calibrated (T = 5–35 °C, S_p = 20–40) for a.) Total Boron of U74 and b.) Total Boron of L10 using K_{HSO4} of D90 and K_{HF} of PF87. White dots and text are median values. * indicates median is internally consistent. Colored horizontal lines are the means. Dashed purple lines indicate internal consistency bounds. Constant abbreviation meanings can be found in Table 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

small differences at extreme temperatures. Differences are largest at lower temperatures (≤15 °C).

The constants of M02, which were determined from field data using back calculation techniques were internally consistent under a very narrow range of conditions, but the scatter is too large to be considered internally consistent for most applications. The constants of CW98, which are often used as a link between low and high salinity waters performed poorly under nearly all conditions and had very large scatter. MPM02 is also internally consistent on average, but with more scatter (fatter tails) than other constants. These sets of constants (M02, MPM02, and CW98) will not be discussed further.

Taylor Diagrams (Fig. S2) support the broad conclusions. As a whole, over the conditions in which the indicator is calibrated, M10 is technically the best. However, M06, WM13, DM87, and L00 all fall nearly on top of each other. S23 also performs very similarly to the others but with a slightly lower correlation coefficient and improved precision.

3.1. pH_t internal consistency with salinity

The various carbonic acid dissociation constants cover differing ranges of salinity (Table 1) with all covering the open ocean salinity ranges, and a few applicable to low salinity or even freshwater. The salinity range evaluated here (nominally 15–39) extends beyond the range of many of the constants, which must be considered when interpreting the results. More importantly, the meta-cresol purple indicator was only calibrated down to a salinity of 20 by Liu et al. (2011). There are calibrations for low salinities (Douglas and Byrne, 2017; Müller and Rehder, 2018). The calibration of MR18 is based on buffers prepared according to Müller et al. (2018) for low salinities and L11 for high salinities. Recent modeling work has brought the applicability of those buffers to seawater into question, suggesting a systematic bias (Clegg et al., 2022b). The calibration of DB17 relied on low salinity measurements made using impure indictor and correcting for the impurities. Compounding the issues at low salinities is that the total scale is only well constrained down to S_p = 20 (DelValls and Dickson, 1998).

For nominal salinities of 15 (Fig. 2), none of the constants produced internally consistent results. The best median internal consistency was still −0.0164 (using GP89, L10 and temperatures 5–35 °C). Given the calibration of the indicator, this result is unsurprising. When using the MR18 indicator calibration, the measured pH_t increased by ~0.01 relative to L11. Although the internal consistency improved, none of the results were internally consistent. M06 and M10 performed better than the other sets, but were not internally consistent, especially when considering the scatter. Measured pH_t using the indicator calibration of DB17 falls about halfway between the values when using L11 and MR18. A thorough evaluation of the different indicator calibrations is beyond the scope of this work, but these results show that the uncertainties at low salinities are significantly larger than at higher salinities and that issues likely exist with both the indicator calibration and the constants. Although extending the total scale to low salinities is an area of active research, the current state of knowledge prevents an evaluation of the constants for S_p < 20, even for constants determined at low salinities. Measurements at S_p = 15 are excluded from further analysis. The issues with indicator calibration highlight the need for raw absorbances to

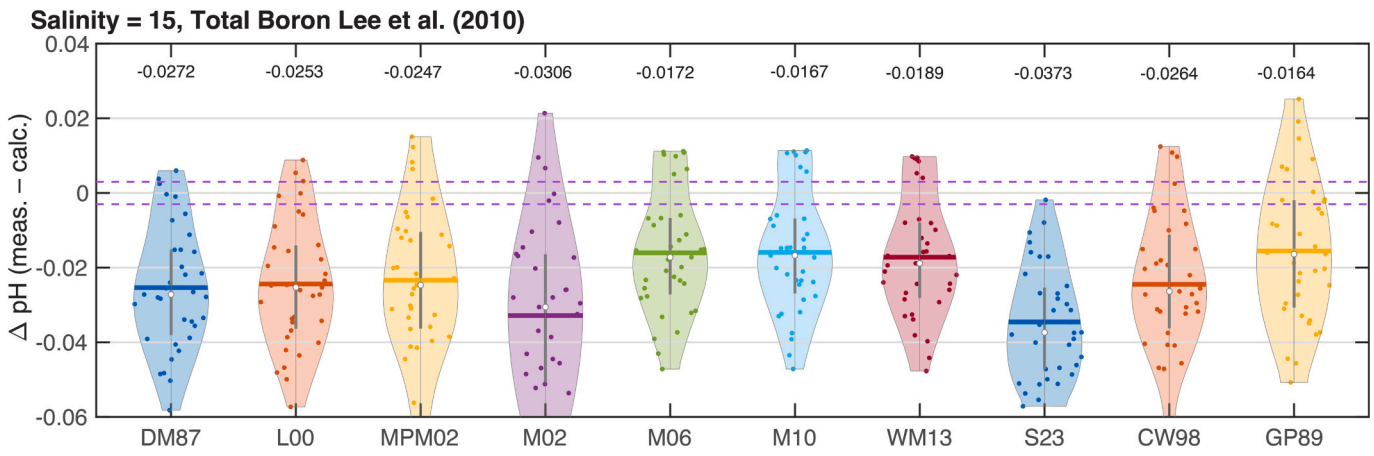


Fig. 2. Violin plots showing pH_t internal consistency for all conditions at $S_p = 15$ and $T = 5$–35 °C using the B_T of L10, K_{HSO4} of D90 and K_{HF} of PF87. White dots and text are median values. Colored horizontal lines are the means. Dashed purple lines indicate internal consistency bounds. Constant abbreviation meanings can be found in Table 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

always be made available so that values can be recalculated when improved calibrations become available. These results at $S_p = 15$ also highlight the need for improving all aspects of the inorganic carbon system for low salinity coastal systems.

At all other salinities (Figs. S3–S6) and temperature ranges from 5 to 35 °C, DM87, L00, M06, M10, WM13, and S23 perform comparably with median values that are internally consistent, or close to, but with significant scatter and ranges outside the accuracy of pH_t. MPM02 is also similar, but with larger scatter and range. GP89 is the least internally consistent, although there is improved consistency when using B_T of U74. GP89 tends to have a positive ΔpH_t, indicating an underestimation of measured pH_t, while the others tend to overestimate measured pH_t. The higher calculated pH_t is likely why GP89 has better internal consistency when using U74 for B_T. GP89 is the only set evaluated here

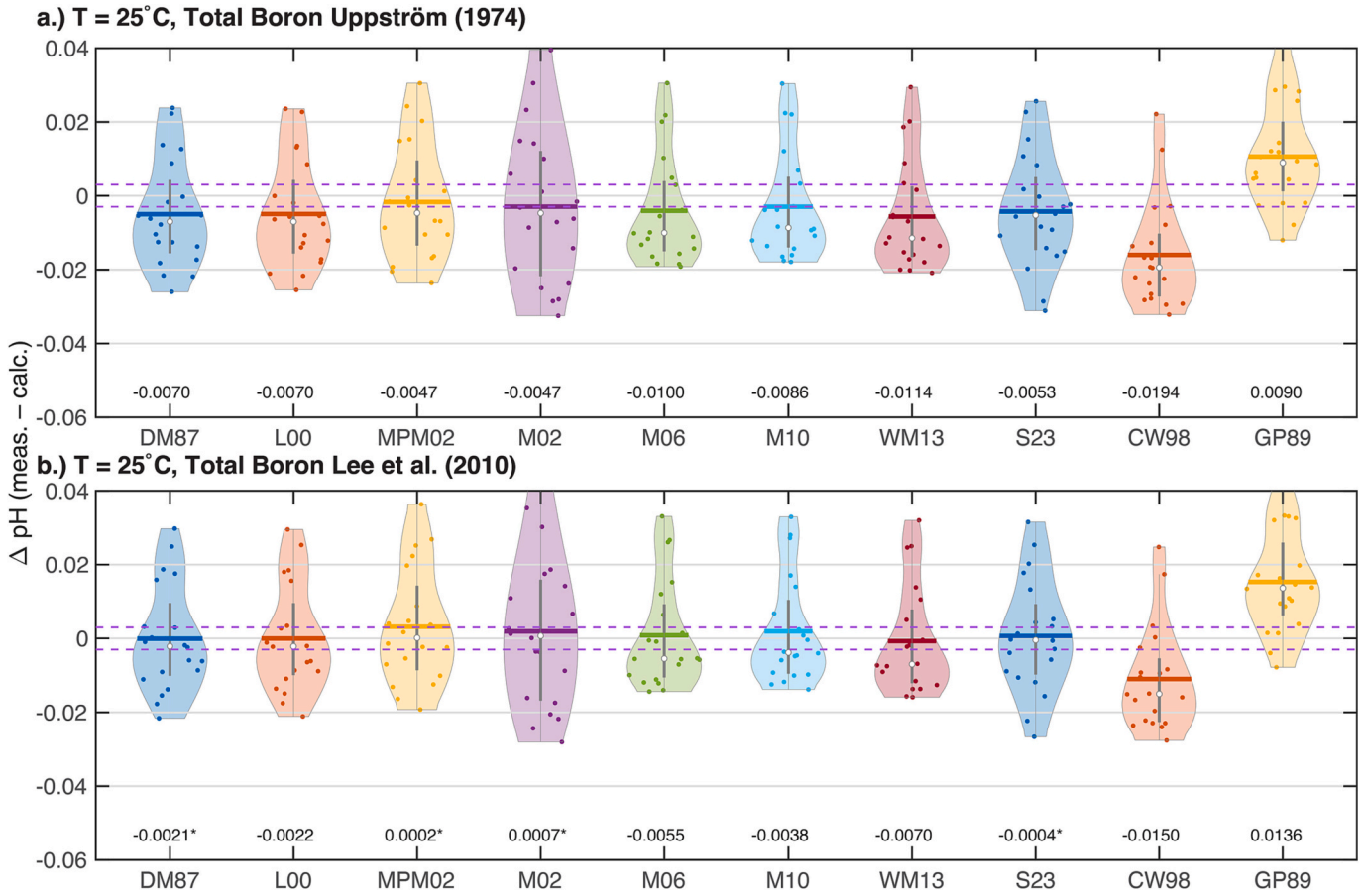


Fig. 3. Violin plots showing pH_t internal consistency for all conditions for $T = 25$ °C and $S_P = 20$–40 for a.) Total Boron of U74 and b.) Total Boron of L10 using K_{HSO4} of D90 and K_{HF} of PF87. White dots and text are median values. * indicates median is internally consistent. Colored horizontal lines are the means. Dashed purple lines indicate internal consistency bounds. Constant abbreviation meanings can be found in Table 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

determined in artificial seawater and constants in artificial seawater tend to be biased higher, relative to constants in real seawater (Dickson and Millero, 1987; Mojica Prieto and Millero, 2002). In general, salinities between 25 and 35 have more internally consistent conditions than higher or lower salinities, although still with considerable scatter. For $S_p = 39$ the B_T of U74 was usually more internally consistent.

For open ocean conditions (near $S_P = 35$ and all pCO_2 examined), the Taylor Diagrams are similar to all conditions (Fig. S2), but with S23 being the 'best' and DM87 and L00 being very similar. Although M06, M10, and WM13 are slightly worse, it is not statistically significant.

3.2. pH_t internal consistency with temperature

As with salinity, the indicator is not calibrated over the full temperature range covered here and is only valid from 5 to 35 °C. GP89 is the only set of constants calibrated below zero, but the median value is not internally consistent at negative temperatures. M06, M10, and WM13 do have median values that are internally consistent, but with significant scatter, at negative temperatures (Fig. S7). At 40 °C (Fig. S14) no set had a median value that was internally consistent. Therefore, negative temperatures and 40 °C were excluded from further analysis.

For 20–25 °C (Figs. 3 and S11), the temperatures at which most spec. pH_t measurements are made, S23 had the best median value, but DM87 and L00 had similar results. At temperatures ≤15 °C, M06, M10, and WM13 have the best median values, with DM87 and L00 having median values just outside the internal consistency cut off. S23 has decreasing consistency with decreasing temperatures. GP89 is sometimes used in polar waters due to early work showing them to be the best at cold temperatures (Laika et al., 2009; Wanninkhof et al., 1999). Indeed, their best internal consistency is at very low temperatures; however, M06, M10, and WM13 have better consistency than GP89 at all temperatures, particularly at low temperatures, suggesting these constants are better at polar temperatures than GP89. Recent modeling work showed GP89 to have more scatter than other sets of constants (Clegg et al., 2022a) which is also shown in the Taylor diagrams (Fig. S2). At high temperatures (20–35 °C, Figs. S11–14) DM87, L00, M06, M10, WM13, and S23 have reasonable median internal consistency across the range of temperatures, although their median values are not always internally consistent.

In general, M06 and M10 have comparatively good internal consistency over the full range of temperatures (although other constants are better at 25 °C where measurements are typically made). WM13 is similar to M06 and M10 but the median values tend to be slightly more negative often making them not internally consistent. DM87 and L00 have reasonable internal consistency over the broad range of temperatures, but the median value is not internally consistent at the extremes. S23 has the best internal consistency for temperatures ≥20 °C but is significantly worse at lower temperatures (Figs. S7–S10). GP89 is only internally consistent in the 5–10 °C range. The important caveat is that all constants have significant scatter at all temperatures, with the majority of points not being internally consistent. Although DM87, L00, and S23 are better at typical pH_t measurement temperatures (20–25 °C), M06 and M10 are better over the broad range of temperatures.

3.3. pH_t internal consistency with pCO_2

The batches of seawater were bubbled with gases of different CO_2 concentrations (balance air). The result is seawater at the bottling temperature (generally ~22–23 °C) with pCO_2 nominally equal to the concentration of the gas cylinder. Since pCO_2 is a strong function of temperature (~4.23% $°C^{-1}$ (Takahashi et al., 1993)) the pCO_2 of the bottles at the pH_t measurement temperature varies significantly for a given batch. For simplicity, the pCO_2 value used in this discussion will be based on the concentration of the gas cylinder and corresponds approximately to the pCO_2 of the batch at room temperature. The nominal concentrations were 150 (very low, Fig. S15), 400 (low, Fig. S16), 600 (medium, Fig. 4), 800 (high, Fig. S17), and 1200 μatm (very high, Fig. S18).

Several prior studies have found a decrease in internal consistency with increasing pCO_2 (Lueker et al., 2000; Millero, 2007). The results here are more complicated. The internal consistency is best at low (≤ 400 μatm) pCO_2 and then indeed deteriorates at medium pCO_2 (600 μatm, Fig. 4) but begins to improve at high pCO_2 (800 μatm). At very high pCO_2 (1200 μatm) conditions the median value is internally consistent (or close), but with significant scatter and a bimodal distribution. This is actually in semi-agreement with the verification experiments of Lueker et al. (2000) which showed large offsets at 600 and 800 μatm but improved internal consistency (albeit with significant scatter) at values > ~ 1300 μatm (Figs. 4, S15–S18). The bimodal distribution at 1200 μatm can be attributed to salinity, with S_p 25–35, being somewhat internally consistent and low (≤ 20) and high (39) being internally inconsistent. The pattern is similar at 600 μatm, but with no salinity being internally consistent. As with salinity, M06 and M10 are generally the best over the full range, but DM87 and L00 do better at 400 μatm pCO_2. S23 performs best at medium to high pCO_2 (> 400 μatm), but worse at lower values. Only GP89 has an internally consistent median value at 600 and 800 μatm pCO_2. It is important to note that the uncertainty in measured pH_t was significantly larger for many temperatures at 150 μatm (Table S1) making the conclusions for these treatments more uncertain. The reason for the higher measurement uncertainty is unknown but could be related to the indicator calibration at these very high pH_t values (> 8.2), or the batches not being fully equilibrated as noted for some batches in the methods.

3.4. pH_t internal consistency and excess alkalinity

As mentioned in the methods, there is increasing evidence of an unidentified minor component of TA (A_x). To assess the impact A_x would have on pH_t calculations, the internal consistency calculations were repeated with a small amount of A_x (Eq. (1); see methods). In general, the inclusion of A_x (Fig. 5) shifted the ΔpH_t in the positive direction (i.e. calculated pH_t was lower than when A_x was not considered). When using the B_T of L10, there appears to be an over correction with a positive shift of the median pH_t internal consistency of ~0.08–0.09, making none of the constants internally consistent (median value). However, when using B_T of U74 the median values are shifted into the internally consistent range. The scatter of the data points is largely unchanged with the addition of A_x. The switch from better internal consistency with L10 to better internal consistency with U74 when A_x is added makes sense because the magnitude of A_x is similar to the difference in borate alkalinity between the two B_T values (Lee et al., 2010). This would suggest that the comparatively large value (8 μmol kg_{sw}^{-1}) estimated by Millero et al. (2002) is partially explained by B_T. The overall median hides some of the trends. At 600 ppm pCO_2, there is still a bimodal distribution, but one mode is now internally consistent. For $25 \leq S_p \leq 39$, internal consistency is improved for DM87, L00, M06, M10, and WM13; but is significantly worsened for S23. The improvement for most constants supports the presence of non-negligible A_x, but the differing results for S23 do not. Prior studies have found A_x to improve internal consistency at high pCO_2/DIC most (Millero, 2007), but that is not the case here. This result could indicate that deep water A_x varies from that of the surface waters used here. From these results it is clear that A_x alone cannot explain poor internal consistency and that there must be non-negligible errors in the constants or measurements as well, in agreement with the conclusions of Fong and Dickson (2019). The results also suggest that, for the seawater used here, 4 μmol kg_{sw}^{-1} is too large for A_x, although other uncertainties could be obscuring the true amount of A_x. Uncertainty in the accuracy of B_T is a significant hinderance to determining A_x.

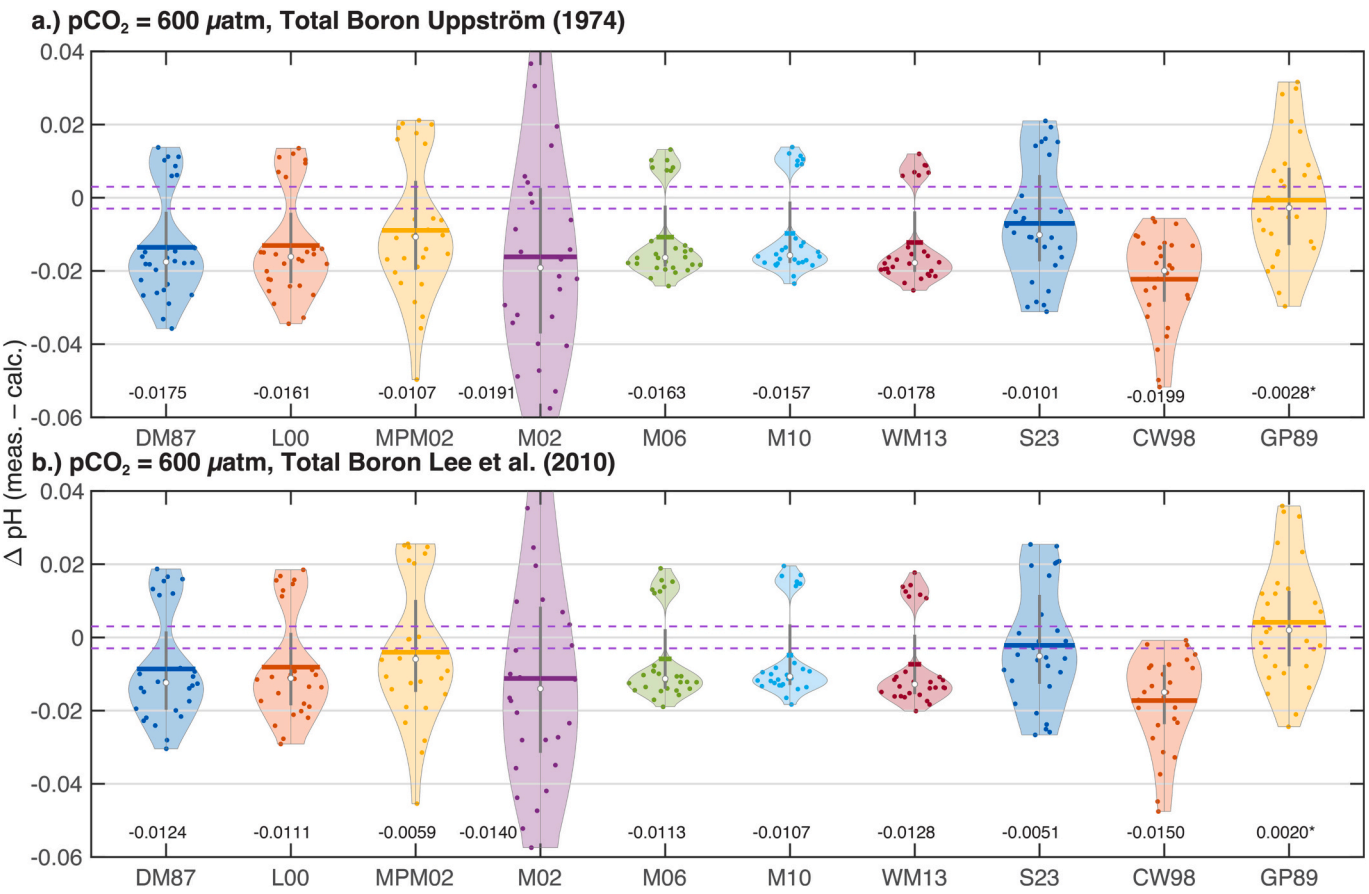


Fig. 4. Violin plots showing pH_t internal consistency for all conditions bubbled with pCO_2 of ~600 μatm (balance air at room temperature), T = 5–40 °C, and S_P = 20–40 for a.) Total Boron of U74 and b.) Total Boron of L10 using K_{HSO4} of D90 and K_{HF} of PF87. White dots and text are median values. * indicates median is internally consistent. Colored horizontal lines are the means. Dashed purple lines indicate internal consistency bounds. Constant abbreviation meanings can be found in Table 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

3.5. pH_t internal consistency discussion

Making broad conclusions from the above results is challenging. On average M06 and M10 are internally consistent and D87, L00, WM13, and S23 are nearly so. Yet, the scatter is so large that the majority of conditions are not internally consistent within the estimated uncertainty of spec. pH_t measurements (~0.0030). At most, only ~25% of conditions are internally consistent for a given set of constants. Looking at individual conditions, it becomes apparent that each of the above constants does best under a different set of conditions, such that on average they are quite comparable. Considering the median value as well as Taylor Diagrams (Fig. S2), M10 is the best over the range of conditions for which the indicator is calibrated (S = 20–40, t = 5–35 °C); however, the others mentioned above are not statistically different. The difference is under what specific conditions a given set of constants is most internally consistent.

The reason for these results is easy to explain by comparing differences in the pK_i* of the different formulations. Considering pK_1* and pK_2* at S_p = 35 across temperatures (Fig. 6a and c), it becomes clear that, with a few exceptions (M02, CW98, and GP89) all of the pK_1*s agree within the error of the fit (2σ, Table 1), even extending beyond their valid range. The differences are similar for pK_2* except that S23 is statistically different at low and high temperatures. A similar comparison for salinity at 25 °C (Fig. 6b and d) show that pK_1* tends to agree with each other over the salinity range they were determined (with the same exceptions as temperature). The results are similar for pK_2*, but there is significant divergence below S_P = 20, which is expected given many constants are not calibrated at low salinities and the uncertainties are higher. In all of the comparisons, it is worth noting that differences between constants are not linear. The non-linearity combined with the uncertainty in the constants explains why the internal consistency calculations for most constants are comparable on average, but vastly different for given conditions. It also helps to explain why S23 and GP89 have somewhat different patterns than the other constants. The non-linearity also partially explains some of the scatter in the results. All of this suggests differences in the internal consistency can largely be attributed to residuals of the fits and that the uncertainty in the constants is too large to meet the needs of climate quality research when calculating pH_t from TA and DIC. Further improvement to the marine inorganic carbon system model will require a significant improvement in the precision of each pK*, particularly pK_2*.

These results should not be interpreted as these 6 sets of constants being equal or that the choice of carbonic acid dissociation constants is unimportant. There are real, and significant, differences between the constants that must be taken into consideration. One could consider the range of conditions covered by their study and choose the one that performs the 'best.' However, that makes comparison among studies difficult if not impossible. Alternately, the 'best' constants over the broad range of oceanic conditions could become the standard. Based on these results that would be M10. At a minimum, one must take into consideration the added uncertainties of using different constants when comparing studies or when using calculated values to fill in missing measured data. Not having a universal set of constants is not only undesirable, but problematic.

Internal consistency was defined here as agreement between measured and calculated values within the pH_t measurement

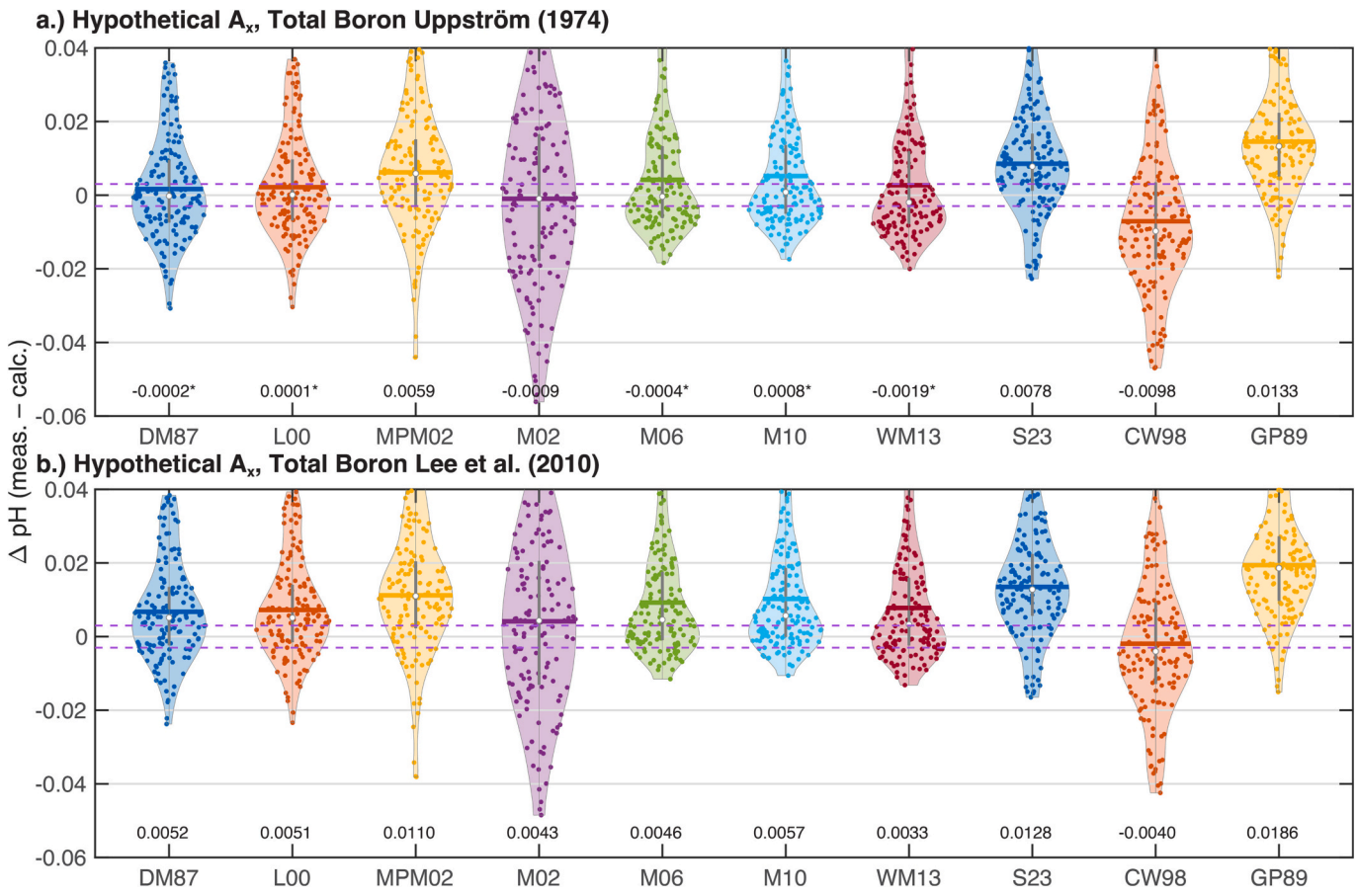


Fig. 5. Violin plots showing pH_t internal consistency with a hypothetical excess alkalinity, A_x under all conditions for which the indicator is calibrated (T = 5–35 °C, S_p = 20–40) for a.) Total Boron of U74 and b.) Total Boron of L10 using K_{HSO4} of D90 and K_{HF} of PF87. White dots and text are median values. * indicates median is internally consistent. Colored horizontal lines are the means. Dashed purple lines indicate internal consistency bounds. Constant abbreviation meanings can be found in Table 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

uncertainty (not accuracy). This definition was chosen because it agrees with the climate quality level (Newton et al., 2015), and at this level, calculated data could be considered equivalent to measured data. In other words, disparate data collection methods would result in non-disparate data, the goal of internal consistency. Although several constants produced reasonable median values, the large scatter is highly problematic. Perhaps the uncertainty in the calculation should also be taken into consideration when determining what constitutes internally consistent. The CO2sys add-on of Orr et al. (2018) provides an easy way to estimate uncertainty in the calculation by propagating the estimated uncertainties of the constants and measurements. From these calculations it becomes clear that the uncertainty in calculated pH_t is not constant and covers a very large range (~ 0.005–0.1). The large range helps to explain why there is such large scatter in the ΔpH_t values; for some conditions the calculation uncertainty is orders of magnitude larger than our defined internal consistency bounds. The large uncertainty in pH_t calculated from DIC and TA has been shown many times (McLaughlin et al., 2015; Millero, 1995, 2007; Orr et al., 2018). DIC and TA have long been recognized as a poor input pair for carbon system calculations. Combining the large calculation uncertainty with the enormous scatter demonstrated here it has become increasingly clear that calculated pH_t does not meet the requirements for most weather or climate level studies of ocean acidification, and that pH_t should be measured rather than calculated. One could argue that because OA focuses on changes rather than absolute values, having a biased pH_t is not of primary concern. The results here and elsewhere (Fassbender et al., 2021) suggest otherwise, as the bias is non-linear and highly dependent on temperature, salinity, and DIC, such that accounting for it would be more challenging than measuring pH_t directly. Somewhat comforting is that the results are generally close to internally consistent for open ocean surface water conditions (salinity near 35 and pCO_2 near 400 µatm) where ocean acidification is most acute, meaning current and historical measurements may be comparable to calculations even if not quite achieving climate quality. Nevertheless, at pCO_2 between ~500–800 µatm, conditions expected in the next several decades, calculated pH_t tends to be significantly overestimated relative to the measured values. This implies that models, which generally calculate pH from TA and DIC (Orr and Epitalon, 2015), are underestimating ocean acidification for the second half of this century; approximately 0.01–0.02 at a pCO_2 of 600 µatm, which is equivalent to almost a decade at current acidification rates. Although the exact bias will depend on the specific regional conditions and constants used.

3.6. TA and DIC internal consistency

If pH_t cannot be calculated to the required precision, it brings into question the oft cited statement "one must only measure two carbon system parameters, and the others can be calculated." At least by today's level of understanding of the marine inorganic carbon system. If it is not true, then what are the implications for parameters that cannot be directly measured (e.g., saturation state) and for models that calculate rather than model pH_t? Measuring three parameters is uncommon, and all four is extremely rare, making calculations necessary. Just as the propagation of uncertainties in calculated pH_t varies by condition, the chosen measured parameters strongly influence the uncertainties in calculated values. It has long been known that measurement of a

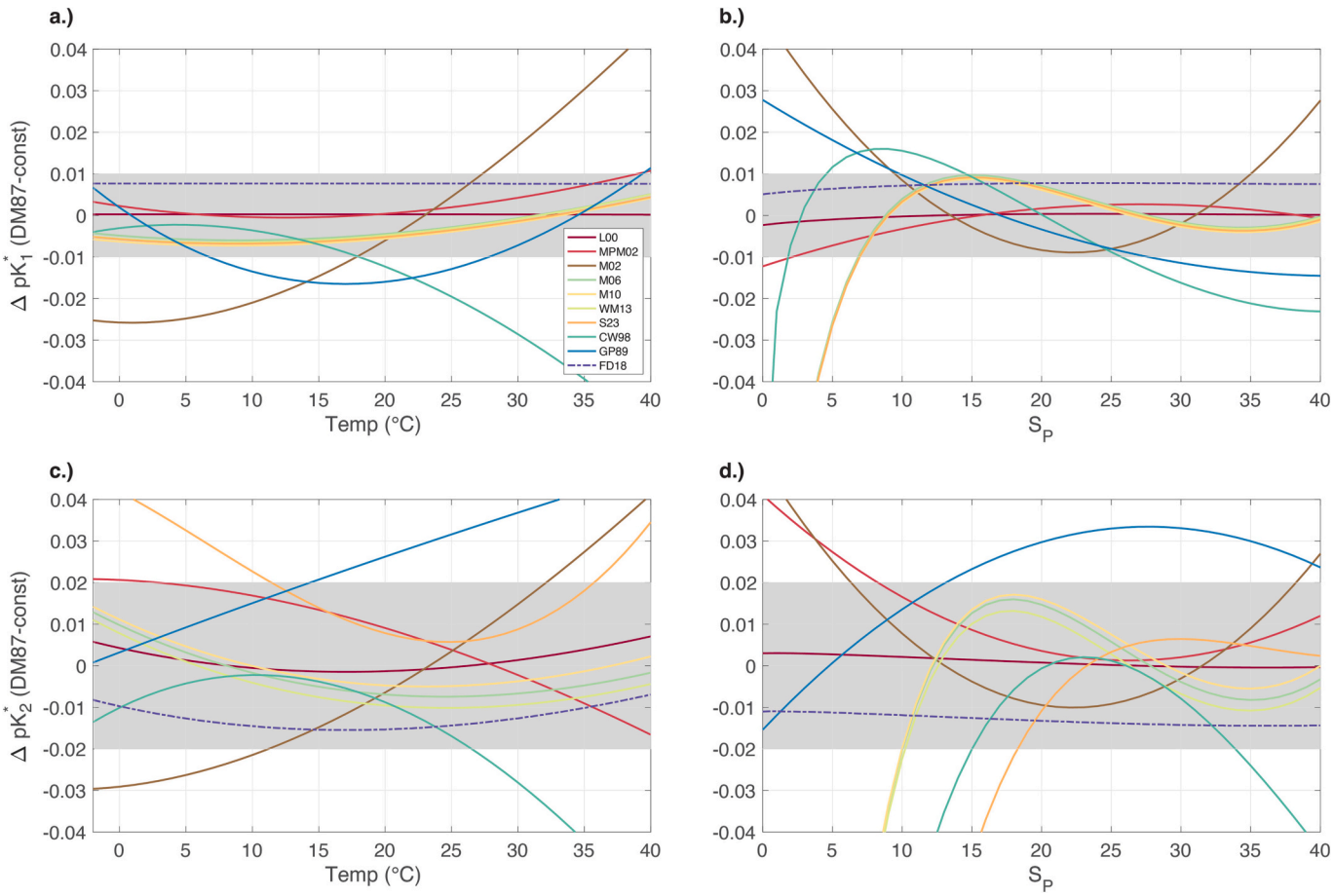


Fig. 6. Comparison of the carbonic acid dissociation constants relative to DM87 ($\Delta pK_i^* = DM87 - constant$) for pK_1^* as a function of temperature (a) at $S_P = 35$ and salinity (b) at T = 25 °C, and pK_2^* as a function of temperatures (c) at $S_P = 35$ and salinity (d) at T = 25 °C. Gray shading indicates the uncertainty in the constants based on DM87. Meanings of the abbreviations can be found in Table 1.

temperature and pressure independent parameter (TA or DIC) paired with a temperature and pressure dependent parameter (pH_t or pCO_2 (fCO_2)) produces lower uncertainties in the calculated values (McLaughlin et al., 2015; Millero, 2007).

Considering TA internal consistency, where $\Delta TA = (TA_{meas} - TA_{calc})$, over the conditions where the pH_t indicator is calibrated (Fig. 7), all constants except GP89 and CW98 have median values within the measurement uncertainty (2 μmol kg_{sw}^{-1}) of TA. As with pH_t internal consistency, the constants of M10 were the most internally consistent, but DM87, L00, M06, WM13 and S23 were not statistically different. Although there is considerable scatter outside the internal consistency bounds, the majority is within ±5 μmol kg_{sw}^{-1}, and aside from a few conditions, all are within ±10 μmol kg_{sw}^{-1}. Even for $S_P = 15$ (not shown), although not internally consistent, the median is only slightly offset, with nearly all conditions within ±10 μmol kg_{sw}^{-1}. Patterns over different conditions and between constants were very similar to those using pH_t, but with lower scatter, and therefore improved internal consistency. As with pH_t, 600 μatm pCO_2 had a bimodal distribution, with almost no conditions being internally consistent, while higher pCO_2 had median values that were internally consistent. At 1200 μatm pCO_2, nearly all conditions were ± 5 μmol kg_{sw}^{-1} for the above-mentioned constants. For GP89, the median value was not internally consistent when using B_T of L10, but was when using U74, however, there was increased scatter compared to the other constants. The total estimated uncertainty in the calculations (Orr et al., 2018), using default individual uncertainties, ranges from 2 to 10 μmol kg_{sw}^{-1}. Thus, unlike pH_t, the calculated and measured TA almost always agree within the estimated calculation uncertainties. Therefore, open ocean TA can almost always be calculated to the weather quality limit (10 μmol kg_{sw}^{-1}), and can be calculated to climate level precision of 2 μmol kg_{sw}^{-1} (Newton et al., 2015) in approximately 30–35% of conditions.

For most sets of constants, there is a small positive bias in ΔTA. The bias suggests a small amount of A_x. When including A_x, as was done with pH_t internal consistency, there is a negative shift in median ΔTA, with a small (~5%) increase in the number of conditions that are internally consistent. However, this is largely a rearrangement of which conditions are internally consistent as the added A_x is an over correction for most of the conditions that were internally consistent without A_x. A notable exception is the constants of S23. The median value of all data (where mCp is calibrated) is slightly negative, and the addition of A_x makes the constants inconsistent, on average. The addition of A_x tends to make the internal consistency when using B_T of U74 better than L10, again suggesting that the generally lower internal consistency using U74 is a result of lower accuracy in the B_T value. The disparate results of S23 compared to the other constants suggests A_x may not be necessary to achieve internal consistency, its magnitude may be smaller than other studies suggest, or A_x is not approximately constant in the ocean. Results using DIC internal consistency are nearly identical except with opposite sign (a negative value in ΔTA results in a positive value in ΔDIC).

3.7. Fong and Dickson adjustments

The results so far suggest uncertainties in multiple aspects of the calculations lead to large biases and scatter that prevent broad internal consistency. Due to the non-linearity of the system and several sources of uncertainty being indistinguishable, accounting for errors in one

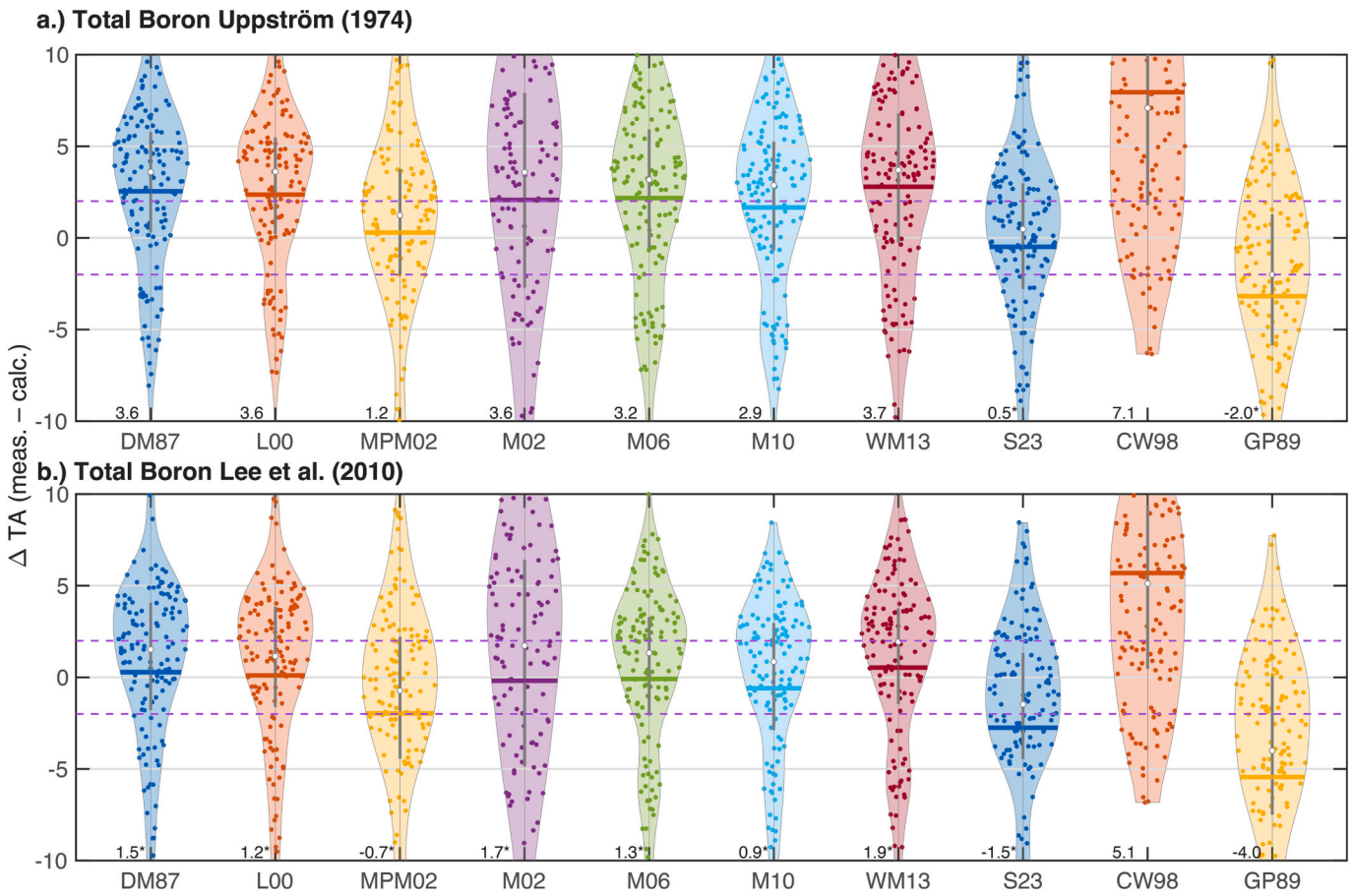


Fig. 7. Violin plots showing TA internal consistency for all conditions for which the indicator is calibrated (T = 5–35 °C, S_p = 20–40) for a.) Total Boron of U74 and b.) Total Boron of L10 using K_{HSO4} of D90 and K_{HF} of PF87. White dots and text are median values. * indicates median is internally consistent. Colored horizontal lines are the means. Dashed purple lines indicate internal consistency bounds. Constant abbreviation meanings can be found in Table 1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

parameter (e.g., including A_x or applying adjustments to the constants) often leads to unclear or mixed results making it nearly impossible to identify the specific causes of the offsets or scatter. Fong and Dickson (2019) considered how accounting for uncertainties in many different components of the calculations might improve internal consistency. They determined a suit of possible adjustments to constants, B_T, measured TA and DIC, and A_x from Pacific Ocean field data. The adjustments are not a unique combination, but they are realistic. The adjustments to the carbonic acid dissociation constants were applied to L00. The median ΔpH_t was improved relative to L00, but scatter did not improve with a long positive tail being present. Much of the large positive ΔpH_t values were at 1200 µatm pCO_2. This contrasts with their results where high CO_2 deep waters were most improved. There are a few possible reasons for the discrepancy. They used field data, while laboratory manipulated surface water was used here. A_x may vary between surface and deep waters, or between Atlantic and Pacific seawater. The analytical methods were nearly identical (automated spec. pH_t with purified indicator, open cell TA titration, and coulometric DIC), but there may be small inter-laboratory differences. A broader range of oceanic conditions was examined here, which could imply that their results are only applicable to a narrow range of S_P that exists in the Pacific Ocean. However, considering only S_p = 35, the improvement relative to L00 is negligible (median increased by 0.0008).

One of their major conclusions was that the field data could not be made internally consistent with adjustments to the constants and measured values alone. A small but non-negligible contribution of A_x was required. The slight negative bias in most of the constants here supports that conclusion. However, S23 was unique in that it had a slight positive bias suggesting A_x was not necessary to explain the results. Also, TA internal consistency suggested a nearly negligible amount of A_x. These disparate results could be due to methodological differences or real differences in the seawaters examined. Unfortunately, this study is unable to resolve the lingering question of A_x in open ocean waters.

3.8. Implications for saturation state

$CaCO_3$ cycling is strongly dependent on the saturation state (Ω), which describes the thermodynamic stability of the minerals with Ω = 1 being in equilibrium (Millero, 2007). The depth at which Ω = 1 for aragonite is the aragonite saturation horizon (ASH) and defines the thermodynamic division between dissolution and precipitation. Theoretically, aragonite should not dissolve above the ASH, but multiple lines of evidence suggest >50% of $CaCO_3$ dissolves in the upper ocean (Feely et al., 2002; Milliman et al., 1999) and models often find dissolution above the ASH is required to reproduce observations (Buitenhuis et al., 2019), creating a long standing and unresolved paradox (Woosley et al., 2012). Resolving this paradox relies in part on knowing the depth of the ASH precisely, however uncertainty in the depth is rarely considered.

Uncertainty in Ω from the different constants is approximately ±2.5% for temperatures ≤25 °C but can be >10% at higher temperatures. Using GLODAPv2.2022 (Lauvset et al., 2022), ASH was calculated for the global ocean. Using B_T of L10 generally had a deeper ASH in agreement with prior work (Orr and Epitalon, 2015). Since most studies showing shallow dissolution were conducted before 2010, it would imply that more $CaCO_3$ dissolves above the ASH than previously estimated if L10 is more accurate. Considering both the carbonic acid

dissociation constants and B_T, the depth of the ASH can vary by as much as 680 m based on the entire GLODAPv2.2022 database. Fig. 8 shows representative sections for the major basins. The uncertainty is largest in the Southern Ocean (south of 50°S) where pteropods are abundant (Bednaršek et al., 2012) and shallow $CaCO_3$ dissolution rates are highest (Feely et al., 2004). The uncertainty is also large in the North Atlantic where the depth of the ASH exceeds.

2000 m (Feely et al., 2004) and cold water corals can be found above the ASH (García-Ibáñez et al., 2021). In the North Pacific uncertainties from the constants are insignificant for ASH, but in other oceans the uncertainty range is generally 100–400 m. These calculations only consider the uncertainty contribution from the constants and not measurement uncertainty. Prior work has shown measurement uncertainty in DIC and TA can contribute ~200 m of uncertainty to ASH (Patsavas et al., 2015), meaning the estimates presented here are likely underestimations of the overall uncertainty in ASH. The uncertainty has significant implications for cycling of $CaCO_3$ and contributes to our poor understanding of its dissolution in shallow waters.

4. Conclusions

Over a broad range of temperatures, salinities, and pCO_2 conditions, the carbonic acid dissociation constants of Millero (2010) when using B_T of Lee et al., 2010 and Schockman et al. (Submitted) when using B_T of Uppström (1974) had the best performance in estimating pH_t from TA and DIC. However, other constants performed comparably (when using B_T of Lee et al. (2010)) and no set was always (or even mostly) internally consistent within the pH_t measurement uncertainty (~0.0030) across all conditions tested. Each set of constants performed well under some conditions, but poor under others. When considering the uncertainties in the fits of the constants we contend that the results imply that, with a few exceptions, all of the constants are statistically indistinguishable and much of the differences can be attributed to residuals in the fits. However, the lack of broad internal consistency means that significant and non-negligible differences exist and often prevent measurements and calculations from being equivalent, hindering inter-comparability, especially when different sets of constants are used. Such disparity prevents many calculated pH_t values from obtaining climate quality levels. The calculations are particularly biased at pCO_2 values between ~500–800 μatm (at ~22 °C) where pH_t calculated from TA and DIC overestimates measured pH_t, indicating that many models are underestimating ocean acidification for the latter half of this century. A reduction in the uncertainties in the carbonic acid dissociation constants (i.e., improved precision), particularly pK_2^*, will be required to make pH_t calculations and measurements internally consistent at the climate quality level. Uncertainties at salinities <20 are too large to draw any conclusions indicating significant improvement in both measurements and constants is required.

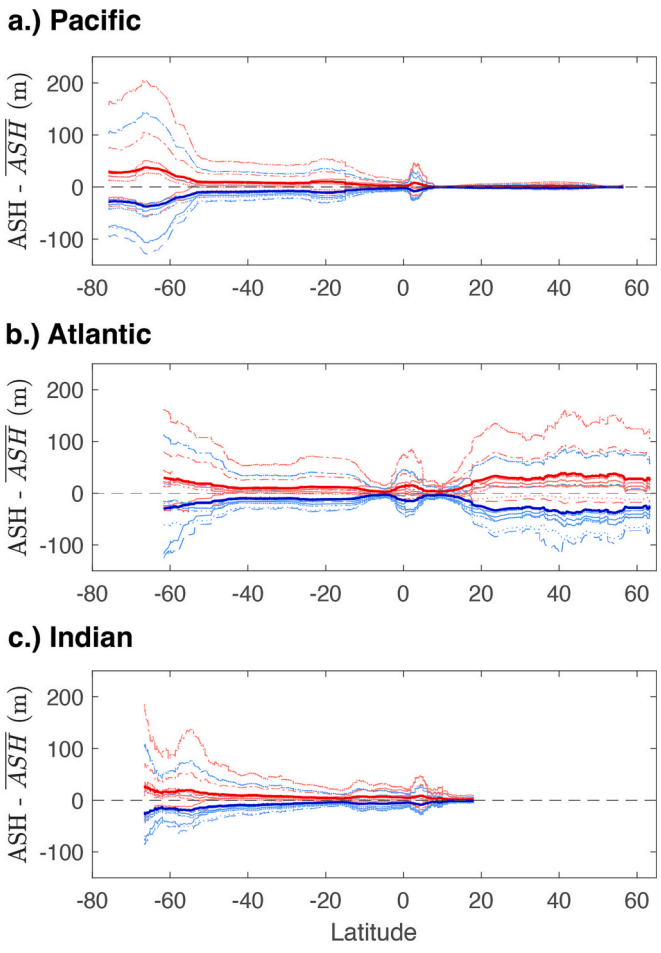


Fig. 8. Difference in the depth of the aragonite saturation horizon (ASH) from the mean of all values($\overline{ASH}$) calculated from TA and DIC using GLODAPv2.2022 and constants of DM87, L00, MPM02, M06, M10, WM13, GP89 and S23 and B_T of U74 (red) and L10 (Blue) for a.) Pacific (P16), b.) Atlantic (A16) and c.) Indian (I08/I09). Thick lines are the means for each B_T value. A moving average method was used to smooth the curves to improve figure clarity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

In agreement with prior work (Lee et al., 2010; Woosley, 2021a) the use of the B_T concentration of Lee et al. (2010) generally produces more internally consistent results and is recommended for most purposes. A notable exception is when using the newly determined constants of Schockman et al. (Submitted). Most constants tend to have a negative ΔpH_t implying the presence of a small, but non-negligible, amount of A_x, but the constants of Schockman et al. (Submitted) run counter to that, making these results inconclusive on the existence and importance of A_x. An important caveat is that these experiments only used surface waters and the presence or amount of A_x may vary by location and depth, hindering the ability of these experiments to address this important question. The disparate results for B_T further highlight the difficulty in determining A_x as uncertainties between the two are often indistinguishable (Sharp and Byrne, 2021). Current literature disagrees on the uncertainty in B_T (Fong and Dickson, 2019; Lee et al., 2010; Orr et al., 2018). Improving or increasing certainty in the accuracy of B_T is required to address questions of A_x.

While pH_t calculated from TA and DIC does not meet the analytical requirements of many studies, calculations of TA or DIC from pH_t and DIC or TA respectively often do. Therefore, it is strongly recommended that all future studies measure pH_t rather than calculate it from DIC and TA. Instead DIC or TA should be calculated if all parameters cannot be directly measured. Models that calculate pH_t from TA and DIC likely underestimate ocean acidification for conditions expected in second half of this century. The implications and solutions should be investigated. We exclude making statements about fCO_2 calculations since measurements were not made in these experiments, but other recent work has explored this in field measurements (García-Ibáñez et al., 2022) and uncertainty estimates suggest results would be similar in that uncertainties in calculations of fCO_2 from TA and DIC are too large for many purposes.

Not all carbon system parameters can be measured. Ancillary parameters, such as $CaCO_3$ saturation state, must still be calculated. Uncertainties arising from the use of different sets of constants were explored. Uncertainty in the depth of the aragonite saturation horizon can be significant (>400 m) from the constants alone. Uncertainty in the saturation horizon from measurement uncertainty can also be significant (Patsavas et al., 2015), but was not considered here. This uncertainty has important implications for cycling and dissolution of particulate inorganic carbon. Studies that use Ω, particularly when distinguishing between dissolution and preservation of $CaCO_3$, must

consider the constants in estimating the uncertainty of the saturation horizon, particularly in determining if and how much $CaCO_3$ dissolves in supersaturated waters.

Declaration of Competing Interest

None.

Data availability

All measured data including TA, DIC, pH_t, S, and nutrients as well as raw absorbance values for pH_t are archived at BCO-DMO (https://www.bco-dmo.org/project/813194). Measured values can be found in Table S1. Calculated pH_t values are available in the supplemental material.

Acknowledgements

This work was funded by NSF grant OCE- 1923312. We are deeply indebted to the Carbon Program at NOAA Pacific Marine Environmental Laboratory for analyzing DIC and Dr. Katelyn Schockman, University of South Florida, for providing their new constants before publication so that they could be included in this analysis. We also thank Andrew Babbin for his editorial comments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.marchem.2023.104247.

References

Álvarez, M., Fajar, N.M., Carter, B.R., Guallart, E.F., Pérez, F.F., Woosley, R.J., Murata, A., 2020. Global Ocean spectrophotometric pH assessment: consistent inconsistencies. Environ. Sci. Technol. https://doi.org/10.1021/acs.est.9b06932 acs.est.9b06932.

Andersson, A.J., Kline, D., Edmunds, P., Archer, S., Bednaršek, N., Carpenter, R., Chadsey, M., Goldstein, P., Grottoli, A., Hurst, T., King, A., Kübler, J., Kuffner, I., Mackey, K., Menge, B., Paytan, A., Riebesell, U., Schnetzer, A., Warner, M.J., Zimmerman, R., 2015. Understanding Ocean acidification impacts on organismal to ecological scales. Oceanography 28, 16–27. https://doi.org/10.5670/oceanog.2015.27.

Bednaršek, N., Mozina, J., Vogt, M., O'Brien, C., Tarling, G.A., 2012. The global distribution of pteropods and their contribution to carbonate and carbon biomass in the modern ocean. Earth Syst. Sci. Data 4, 167–186. https://doi.org/10.5194/essd-4-167-2012.

Buitenhuis, E.T., Le Quéré, C., Bednaršek, N., Schiebel, R., 2019. Large contribution of pteropods to shallow CaCO3 export. Glob. Biogeochem. Cycles 33, 458–468. https://doi.org/10.1029/2018gb006110.

Busch, D.S., O'Donnell, M., Hauri, C., Mach, K., Poach, M., Doney, S.C., Signorini, S.R., 2015. Understanding, characterizing, and communicating responses to ocean acidification: challenges and uncertainties. Oceanography 28, 30–39. https://doi.org/10.5670/oceanog.2015.29.

Byrne, R.H., McElligott, S., Feely, R.A., Millero, F.J., 1999. The role of pHT measurements in marine CO2-system characterizations. Deep Sea Res. Part I Oceanogr. Res. Pap. 46, 1985–1997. https://doi.org/10.1016/S0967-0637(99)00031-X.

Cai, W.-J.J., Wang, Y., 1998. The chemistry, fluxes, and sources of carbon dioxide in the estuarine waters of the Satilla and Altamaha Rivers, Georgia. Limnol. Oceanogr. 43, 657–668. https://doi.org/10.4319/lo.1998.43.4.0657.

Caldeira, K., Wickett, M.E., 2003. Anthropogenic carbon and ocean pH. Nature 425, 365. https://doi.org/10.1038/425365a.

Canadell, J.G., Monteiro, P.M.S., Costa, M.H., da Cunha, L.C., Cox, P.M., Eliseev, A.V., Henson, S., Ishii, M., Jaccard, S., Koven, C., Lohila, A., Patra, P.K., Piao, S., Rogelj, J., Syampungani, S., Zaehle, S., Zickfeld, K., 2021. Global carbon and other biogeochemical cycles and feedbacks, climate change 2021: the physical science basis. In: Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. https://doi.org/10.1017/9781009157896.007.674.

Carter, B.R., Radich, J.A., Doyle, H.L., Dickson, A.G., 2013. An automated system for spectrophotometric seawater pH measurements. Limnol. Oceanogr. Methods 11, 16–27. https://doi.org/10.4319/lom.2013.11.16.

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

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 Oceanogr. Res. Pap. 40, 2115–2129. https://doi.org/10.1016/0967-0637(93)90048-8.

Clegg, S.L., Humphreys, M.P., Waters, J.F., Turner, D.R., Dickson, A.G., 2022a. Chemical speciation models based upon the Pitzer activity coefficient equations, including the propagation of uncertainties. II. Tris buffers in artificial seawater at 25 °C, and an assessment of the seawater 'Total' pH scale. Mar. Chem. 244, 104096 https://doi.org/10.1016/j.marchem.2022.104096.

Clegg, S.L., Waters, J.F., Turner, D.R., Dickson, A.G., 2022b. Chemical speciation models based upon the Pitzer activity coefficient equations, including the propagation of uncertainties. III. Seawater from the freezing point to 45 °C, including acid-base equilibria. Mar. Chem., 104196 https://doi.org/10.1016/j.marchem.2022.104196.

DelValls, T.A., Dickson, A.G., 1998. The pH of buffers based on 2-amino-2-hydroxy-methyl-1,3-propanediol ('tris') in synthetic sea water. Deep Sea Res. Part I Oceanogr. Res. Pap. 45, 1541–1554. https://doi.org/10.1016/S0967-0637(98)00019-3.

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. Pap. 28, 609–623. https://doi.org/10.1016/0198-0149(81)90121-7.

Dickson, A.G., 1990. Standard potential of the reaction: AgCl(S) + 12H2(g) = ag(s) + HCl(aq), and the standard acidity constant of the ion HSO4- in synthetic sea water from 273.15 to 318.15 K. J. Chem. Thermodyn. 22, 113–127. https://doi.org/10.1016/0021-9614(90)90074-Z.

Dickson, A.G., 1993. The measurement of sea water pH. Mar. Chem. 44, 131–142. https://doi.org/10.1016/0304-4203(93)90198-W.

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. Pap. 34, 1733–1743. https://doi.org/10.1016/0198-0149(87)90021-5.

Dickson, A.G., Riley, J.P., 1979. The estimation of acid dissociation constants in seawater media from potentiometric titration with strong base. Mar. Chem. 7, 89–99.

Dickson, A.G., Afghan, J.D., Anderson, G.C., 2003. Reference materials for oceanic CO2 analysis: a method for the certification of total alkalinity. Mar. Chem. 80, 185–197. https://doi.org/10.1016/S0304-4203(02)00133-0.

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

Dickson, A.G., Camoes, M.F., Spitzer, P., Fisicaro, P., Stoica, D., Pawlowicz, R., Feistel, R., 2016. Metrologia metrological challenges for measurements of key climatological observables. Part 3: seawater pH related content. Metrologia 53, R26–R39. https://doi.org/10.1088/0026-1394/53/1/R26.

Doney, S.C., Fabry, V.J., Feely, R.A., Kleypas, J.A., 2009. Ocean acidification: the other CO $_2$ problem. Annu. Rev. Mar. Sci. 1, 169–192. https://doi.org/10.1146/annurev.marine.010908.163834.

Doney, S.C., Shallin Busch, D., Cooley, S.R., Kroeker, K.J., 2020. Annual Review of Environment and Resources the Impacts of Ocean Acidification on Marine Ecosystems and Reliant Human Communities. https://doi.org/10.1146/annurev-environ-012320.

Douglas, N.K., Byrne, R.H., 2017. Spectrophotometric pH measurements from river to sea: calibration of mCP for $0 \leq S \leq 40$ and $278.15 \leq T \leq 308.15$ K. Mar. Chem. 197, 64–69. https://doi.org/10.1016/J.MARCHEM.2017.10.001.

Fassbender, A.J., Orr, J.C., Dickson, A.G., 2021. Technical note: interpreting pH changes. Biogeosciences 18, 1407–1415. https://doi.org/10.5194/bg-18-1407-2021.

Feely, R.A., Sabine, C.L., Lee, K., Millero, F.J., Lamb, M.F., Greeley, D., Bullister, J.L., Key, R.M., Peng, T.-H., Kozyr, A., Ono, T., Wong, C.S., 2002. In situ calcium carbonate dissolution in the Pacific Ocean. Glob. Biogeochem. Cycles 16. https://doi.org/10.1029/2002gb001866, 91-1-91–12.

Feely, R.A., Sabine, C.L., Lee, K., Berelson, W., Kleypas, J., Fabry, V.J., Millero, F.J., 2004. Impact of anthropogenic CO2 on the CaCO3 system in the oceans. Science 305 (80), 362–366. https://doi.org/10.1126/science.1097329.

Fong, M.B., Dickson, A.G., 2019. Insights from GO-SHIP hydrography data into the thermodynamic consistency of CO2 system measurements in seawater. Mar. Chem. 211, 52–63. https://doi.org/10.1016/j.marchem.2019.03.006.

García-Ibáñez, M.I., Bates, N.R., Bakker, D.C.E., Fontela, M., Velo, A., 2021. Cold-water corals in the subpolar North Atlantic Ocean exposed to aragonite undersaturation if the 2 °C global warming target is not met. Glob. Planet. Chang. 201 https://doi.org/10.1016/j.gloplacha.2021.103480.

García-Ibáñez, M.I., Takeshita, Y., Guallart, E.F., Fajar, N.M., Pierrot, D., Pérez, F.F., Cai, W.J., Álvarez, M., 2022. Gaining insights into the seawater carbonate system using discrete fCO2 measurements. Mar. Chem. 245 https://doi.org/10.1016/j.marchem.2022.104150.

Gazeau, F., Parker, L.M., Comeau, S., Gattuso, J.P., O'Connor, W.A., Martin, S., Pörtner, H.O., Ross, P.M., 2013. Impacts of ocean acidification on marine shelled molluscs. Mar. Biol. 160, 2207–2245. https://doi.org/10.1007/s00227-013-2219-3.

Goyet, C., Poisson, A., 1989. New determination of carbonic acid dissociation constants in seawater as a function of temperature and salinity. Deep Sea Res. Part A Oceanogr. Res. Pap. 36, 1635–1654. https://doi.org/10.1016/0198-0149(89)90064-2.

Houghton, R.A., 2007. Balancing the global carbon budget. Annu. Rev. Earth Planet. Sci. 35, 313–347. https://doi.org/10.1146/annurev.earth.35.031306.140057.

Humphreys, M.P., Waters, J.F., Turner, D.R., Dickson, A.G., Clegg, S.L., 2022. Chemical speciation models based upon the Pitzer activity coefficient equations, including the propagation of uncertainties: artificial seawater from 0 to 45 °C. Mar. Chem. 244, 104095 https://doi.org/10.1016/j.marchem.2022.104095.

Hunt, C.W., 2021. Alkalinity and Buffering in Estuarine, Costal, and Shelf Waters. University of New Hampshire.

Jiang, L.-Q., Kozyr, A., Relph, J.M., Ronje, E.I., Kamb, L., Burger, E., Myer, J., Nguyen, L., Arzayus, K.M., Boyer, T., Cross, S., Garcia, H., Hogan, P., Larsen, K., Parsons, A.R.,

2023. The Ocean Carbon and Acidification Data System, pp. 1–12. https://doi.org/10.1038/s41597-023-02042-0.
Johnson, K.M., 1992. Single-Operator Multiparameter Metabolic Analyzer (SOMMA) for Total Carbon Dioxide (CT) with Coulometric Detection. Operator's Manual. Brookhaven National Laboratory, Upton, NY. https://doi.org/10.2172/10194787.
Kerr, D.E., Brown, P.J., Grey, A., Kelleher, B.P., 2021. The influence of organic alkalinity on the carbonate system in coastal waters. Mar. Chem. 237, 104050 https://doi.org/10.1016/j.marchem.2021.104050.
Khoo, K.H., Ramette, R.W., Culberson, C.H., Bates, R.G., 1977. Determination of hydrogen ion concentrations in seawater from 5 to 40.Degree.C: standard potentials at salinities from 20 to 45%. Anal. Chem. 49, 29–34. https://doi.org/10.1021/ac50009a016.
Ko, Y.H., Lee, K., Eom, K.H., Han, I.-S., 2016. Organic alkalinity produced by phytoplankton and its effect on the computation of ocean carbon parameters. Limnol. Oceanogr. 61, 1462–1471. https://doi.org/10.1002/lno.10309.
Laika, H.E., Goyet, C., Vouve, F., Poisson, A., Touratier, F., 2009. Interannual properties of the CO 2 system in the Southern Ocean south of Australia. Antarct. Sci. 21, 663–680. https://doi.org/10.1017/S0954102009990319.
Lauvset, S.K., Lange, N., Tanhua, T., Bittig, H.C., Olsen, A., Kozyr, A., Alin, S., Álvarez, M., Azetsu-Scott, K., Barbero, L., Becker, S., Brown, P.J., Carter, B.R., Da Cunha, L.C., Feely, R.A., Hoppema, M., Humphreys, M.P., Ishii, M., Jeansson, E., Jiang, L.Q., Jones, S.D., Lo Monaco, C., Murata, A., Müller, J.D., Pérez, F.F., Pfeil, B., Schirnick, C., Steinfeldt, R., Suzuki, T., Tilbrook, B., Ulfsbo, A., Velo, A., Woosley, R. J., Key, R.M., 2022. GLODAPv2.2022: the latest version of the global interior ocean biogeochemical data product. Earth Syst. Sci. Data 14, 5543–5572. https://doi.org/10.5194/essd-14-5543-2022.
Lee, K., Millero, F.J., 1997. The carbon dioxide system in the Atlantic Ocean. J. Geophys. Res. 201, 15693–15707.
Lee, K., Kim, T.-W., Byrne, R.H., Millero, F.J., Feely, R.A., Liu, Y.-M., 2010. The universal ratio of boron to chlorinity for the North Pacific and North Atlantic oceans. Geochim. Cosmochim. Acta 74, 1801–1811. https://doi.org/10.1016/j.gca.2009.12.027.
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, 4862–4868. https://doi.org/10.1021/es200665d.
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. Mar. Chem. 70, 105–119. https://doi.org/10.1016/S0304-4203(00)00022-0.
McDougall, T.J., Baker, P.M., 2011. Getting Started with TEOS-10 and the Gibbs Seawater (GSW) Oceanographic Toolbox. SCOR/IAPSO WG127.
McLaughlin, K., Weisberg, S.B., Dickson, A.G., Hofmann, G.E., Newton, J.A., Aseltine-Neilson, D., Barton, A., Cudd, S., Feely, R.A., Jefferds, I., Jewett, E., King, T., Langdon, C., McAfee, S., Pleschner-Steele, D., Steele, B., 2015. Core principles of the California current acidification network: linking chemistry, physics, and ecological effects. Oceanography 28, 160–169. https://doi.org/10.5670/oceanog.2015.39.
Mehrbach, C., Culberson, C.H., Hawley, J.E., Pytkowicx, R.M., 1973. Measurement of the apparent dissociation constants of carbonic acid in seawater at atmospheric pressure. Limnol. Oceanogr. 18, 897–907. https://doi.org/10.4319/lo.1973.18.6.0897.
Millero, F.J., 1995. Thermodynamics of the carbon dioxide system in the oceans. Geochim. Cosmochim. Acta 59, 661–677. https://doi.org/10.1016/0016-7037(94)00354-O.
Millero, F.J., 2007. The marine inorganic carbon cycle. Chem. Rev. 107, 308–341. https://doi.org/10.1021/cr0503557.
Millero, F.J., 2010. Carbonate constants for estuarine waters. Mar. Freshw. Res. 61, 139–142. https://doi.org/10.1071/MF09254.
Millero, F.J., Pierrot, D., Lee, K., Wanninkhof, R., Feely, R.A., Sabine, C.L., Key, R.M., Takahashi, T., 2002. Dissociation constants for carbonic acid determined from field measurements. Deep Sea Res. Part I Oceanogr. Res. Pap. 49, 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. Mar. Chem. 100, 80–94. https://doi.org/10.1016/j.marchem.2005.12.001.
Milliman, J.D., Troy, P.J., Balch, W.M., Adams, A.K., Li, Y.-H., Mackenzie, F.T., 1999. Biologically mediated dissolution of calcium carbonate above the chemical lysocline? Deep Sea Res. Part I Oceanogr. Res. Pap. 46, 1653–1669. https://doi.org/10.1016/S0967-0637(99)00034-5.
Mojica Prieto, F.J., Millero, F.J., 2002. The values of $pK_1 + pK_2$ for the dissociation of carbonic acid in seawater. Geochim. Cosmochim. Acta 66, 2529–2540. https://doi.org/10.1016/S0016-7037(02)00855-4.
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. Mar. Sci. 5, 177. https://doi.org/10.3389/fmars.2018.00177.
Müller, J.D., Bastkowski, F., Sander, B., Seitz, S., Turner, D.R., Dickson, A.G., Rehder, G., 2018. Metrology for pH measurements in brackish waters—part 1: extending electrochemical pHT measurements of TRIS buffers to salinities 5–20. Front. Mar. Sci. 5, 176. https://doi.org/10.3389/fmars.2018.00176.
Newton, J.A., Feely, R.A., Jewett, E.B., Williamson, P., Mathis, J., 2015. Global Ocean Acidification Observing Network: Requirements and Governance Plan Second Edition GOA-ON Global Ocean Acidification Observing Network.
Olsen, A., Key, R.M., van Heuven, S., Lauvset, S.K., Velo, A., Lin, X., Schirnick, C., Kozyr, A., Tanhua, T., Hoppema, M., Jutterström, S., Steinfeldt, R., Jeansson, E., Ishii, M., Pérez, F.F., Suzuki, T., 2016. The Global Ocean data analysis project version 2 (GLODAPv2) – an internally consistent data product for the world ocean. Earth Syst. Sci. Data 8, 297–323. https://doi.org/10.5194/essd-8-297-2016.
Orr, J.C., Epitalon, J.M., 2015. Improved routines to model the ocean carbonate system: Mocsy 2.0. Geosci. Model Dev. 8, 485–499. https://doi.org/10.5194/gmd-8-485-2015.
Orr, J.C., Epitalon, J.-M., Dickson, A.G., Gattuso, J.-P., 2018. Routine uncertainty propagation for the marine carbon dioxide system. Mar. Chem. 207, 84–107. https://doi.org/10.1016/j.marchem.2018.10.006.
Park, P.K., 1969. Oceanic CO_2 system: an evaluation of ten methods of investigation. Limnol. Oceanogr. 14, 179–186. https://doi.org/10.4319/lo.1969.14.2.0179.
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. Mar. Chem. 176, 9–20. https://doi.org/10.1016/j.marchem.2015.06.022.
Perez, F.F., Fraga, F., 1987. Association constant of fluoride and hydrogen ions in seawater. Mar. Chem. 21, 161–168. https://doi.org/10.1016/0304-4203(87)90036-3.
Pierrot, D., Neill, C., Sullivan, K., Castle, R., Wanninkhof, R., Lüger, H., Johannessen, T., Olsen, A., Feely, R.A., Cosca, C.E., 2009. Recommendations for autonomous underway pCO2 measuring systems and data-reduction routines. Deep Sea Res. Part II Top. Stud. Oceanogr. 56, 512–522. https://doi.org/10.1016/J.DSR2.2008.12.005.
Schockman, K.M., Byrne, R.H., 2021. Spectrophotometric determination of the bicarbonate dissociation constant in seawater. Geochim. Cosmochim. Acta. https://doi.org/10.1016/j.gca.2021.02.008.
Schockman K.M.; Byrne R.H.; Carter B.R.; Feely R.A., Submitted Open-Ocean Characterization of the CO2-System: Spectrophotometric Determination of the Bicarbonate Dissociation Constant ($20 \leq SP \leq 40$, $3 \leq T \leq 35$ °C) and Examination of Internal Consistency.
Sharp, J.D., Byrne, R.H., 2020. Interpreting measurements of total alkalinity in marine and estuarine waters in the presence of proton-binding organic matter. Deep Sea Res. Part I Oceanogr. Res. Pap. 165, 103338 https://doi.org/10.1016/j.dsr.2020.103338.
Sharp, J.D., Byrne, R.H., 2021. Technical note: excess alkalinity in carbonate system reference materials. Mar. Chem. 233, 103965 https://doi.org/10.1016/j.marchem.2021.103965.
Sharp, J.D., Pierrot, D., Humphreys, M.P., Epitalon, J.-M., Orr, J.C., Lewis, E.R., Wallace, D.W.R., 2020. CO2SYSv3 for MATLAB. https://doi.org/10.5281/zenodo.4023039.
Shaw, E.C., Mcneil, B.I., Tilbrook, B., 2012. Impacts of ocean acidification in naturally variable coral reef flat ecosystems. J. Geophys. Res. 117, 3038. https://doi.org/10.1029/2011jc007655.
Strickland, J.D.H., Parsons, T.R., 1972. A Practical Handbook of Seawater Analysis, 2nd ed. Fisheries Research Board of Canada, Ottawa.
Sutton, A.J., Manzello, D., Gintert, B., 2015. Coupling chemical and biological monitoring to understand the impact of ocean acidification on coral reef ecosystems. Oceanography 28, 28–29. https://doi.org/10.5670/oceanog.2015.28.
Takahashi, T., Olafssond, J., Goddard, J.G., Chipman, D.W., Sutherland, S.C., 1993. Seasonal variations of CO_2 and nutrients in the high-latitude surface oceans: a comparative study. Glob. Biogeochem. Cycles 7, 843–878.
Takeshita, Y., Warren, J.K., Liu, X., Spaulding, R.S., Byrne, R.H., Carter, B.R., DeGrandpre, M.D., Murata, A., Watanabe, S. Ichi, 2021. Consistency and stability of purified meta-cresol purple for spectrophotometric pH measurements in seawater. Mar. Chem. 236, 104018 https://doi.org/10.1016/j.marchem.2021.104018.
Uppström, L.R., 1974. The boron/chlorinity ratio of deep-sea water from the Pacific Ocean. Deep-Sea Res. Oceanogr. Abstr. 21, 161–162. https://doi.org/10.1016/0011-7471(74)90074-6.
Wanninkhof, R., Lewis, E.R., Feely, R.A., Millero, F.J., 1999. The optimal carbonate dissociation constants for determining surface water pCO_2 from alkalinity and total inorganic carbon. Mar. Chem. 65, 291–301. https://doi.org/10.1016/S0304-4203(99)00021-3.
Waters, J.F., Millero, F.J., 2013. The free proton concentration scale for seawater pH. Mar. Chem. 149, 8–22. https://doi.org/10.1016/j.marchem.2012.11.003.
Waters, J.F., Millero, F.J., Sabine, C.L., 2011. Changes in South Pacific anthropogenic carbon. Glob. Biogeochem. Cycles 25. https://doi.org/10.1029/2010GB003988.
Waters, J.F., Millero, F.J., Woosley, R.J., 2014. Corrigendum to "The free proton concentration scale for seawater pH"[MARCHE:149(2013)8-22. Mar. Chem. 165, 66–67. https://doi.org/10.1016/j.marchem.2014.07.004.
Woosley, R.J., 2021a. Evaluation of the temperature dependence of dissociation constants for the marine carbon system using pH and certified reference materials. Mar. Chem. 229, 103914 https://doi.org/10.1016/j.marchem.2020.103914.
Woosley, R.J., 2021b. Long-term stability and storage of meta-cresol purple solutions for seawater pH measurements. Limnol. Oceanogr. Methods. https://doi.org/10.1002/lom3.10462.
Woosley, R.J., Millero, F.J., Grosell, M., 2012. The solubility of fish-produced high magnesium calcite in seawater. J. Geophys. Res. Oceans 117, 1–5. https://doi.org/10.1029/2011jc007599.
Woosley, R.J., Millero, F.J., Takahashi, T., 2017. Internal consistency of the inorganic carbon system in the Arctic Ocean. Limnol. Oceanogr. Methods 15, 887–896. https://doi.org/10.1002/lom3.10208.