1 Technical Note: Stability of tris pH buffer in artificial seawater
2 stored in bags

3 Wiley H. Wolfe[1], Kenisha M. Shipley[1], Philip J. Bresnahan[2], Yuichiro Takeshita[3], Taylor Wirth[1],
4 Todd R. Martz[1]

5 [1]Scripps Institution of Oceanography, University of California San Diego, La Jolla, 92093, USA
6 [2]Department of Earth and Ocean Sciences, University of North Carolina Wilmington, Wilmington, 28403, USA
7 [3]Monterey Bay Aquarium Research Institute, Moss Landing, 95093, USA

8 *Correspondence To:* Philip J. Bresnahan Jr. (bresnahanp@uncw.edu)

Abstract

Equimolar tris (2-amino-2-hydroxymethyl-propane-1,3-diol) buffer in artificial seawater is a well characterized and commonly used standard for oceanographic pH measurements. We evaluated the stability of tris pH when stored in flexible, gas impermeable bags across a variety of experimental conditions, including bag type, tris batch, and storage in air vs. seawater over 300 days. Bench-top spectrophotometric pH analysis revealed that the pH of tris stored in bags drifted at a rate of -0.0058 ± 0.0008 yr^{-1} (mean slope $\pm$ 95% confidence interval of slope). Analyses of total dissolved inorganic carbon confirmed that a combination of CO_2 infiltration and/or microbial respiration led to the observed decrease in pH. Eliminating bagged tris pH drift remains a goal, yet the pH drift rate of 0.006 yr^{-1} is lower than many processes of interest and demonstrates the value of bagged tris to sensor calibration and validation of autonomous in situ pH measurements.

1. Introduction

Ocean pH is a key measurement used for tracking biogeochemical processes such as photosynthesis, respiration, and calcification (Takeshita et al., 2016); and represents perhaps the most recognized variable associated with ocean acidification (OA), the decrease in ocean pH due to the uptake of anthropogenic carbon dioxide (Doney et al., 2009). OA progresses with a global average pH decline of 0.002 per year in the surface open ocean (Bates et al., 2014), and the accumulated and projected near-term effects of OA have been shown to have deleterious effects on many calcifying organisms (Cooley and Doney, 2009). Beyond the narrow scope of calcifiers, organismal response is complex, exhibiting varied responses across processes such as reproduction, growth rate, and sensory perception. Organismal responses are further complicated by their impact on ecosystem level dynamics, such as altering competition and predator-prey relationships (Doney et al., 2020). Furthermore, pH effects are often exacerbated by concomitant stressors, such as decreased dissolved oxygen or increased temperature. Ultimately, OA will affect humans through impacts on fisheries, aquaculture, and shoreline protection (Branch et al., 2013; Doney et al., 2020).

The quality of pH measurement required to observe various phenomena is often broken into "climate" and "weather" levels of uncertainty (Newton et al., 2014), or 0.02 and 0.003, respectively. Discrete sampling has been shown to be capable of meeting the climate level of uncertainty when best practices are followed, yet many labs do not consistently meet this standard (Bockmon and Dickson, 2015). Furthermore, while discrete, bench-top methodologies can be the most accurate, the ocean's vast size limits the oceanographic community's ability to make ship-based discrete pH measurements, to roughly once per decade throughout most of the ocean (Sloyan et al., 2019). The sparsity of ship-board measurements hinders our ability to assess sub-decadal processes, such as seasonal cycles or bloom events, over much of the ocean (Karl, 2010), and highlights the need for autonomous, high-frequency pH measurements. Technological advancements have led to more routine autonomous pH measurements over the past decade, providing opportunities to fill some gaps in time and space in discrete sampling programs (e.g, Byrne, 2014; Martz et al., 2015; Lai et al., 2018; Wang et al., 2019; Tilbrook et al., 2019). Globally, pH sensors now operate on hundreds of autonomous platforms including moorings and profiling floats, delivering unique datasets in the form of Eulerian and depth resolved Lagrangian time series (Johnson et al., 2017; Bushinsky et al., 2019; Sutton et al., 2019).

45   While sensors increase data coverage, many sensor-based pH measurements, particularly on moored systems, continue

46   to fall short of both climate and weather levels of uncertainty, as highlighted in the intercomparison tests carried out

47   by the Alliance for Coastal Technologies (ACT, 2012) and by the Wendy Schmidt Ocean Health XPRIZE (Okazaki

48   et al., 2017).

49   Independent validation is typically required for autonomous sensors to meet both weather and climate levels

50   of uncertainty. For example, autonomous underway pCO$_2$ systems (Pierrot et al., 2009), moorings (Bushinsky et al.,

51   2019), and autonomous surface vehicles (Chavez et al., 2017; Sabine et al., 2020) are able to provide climate quality

52   observations with an uncertainty of ± 2 μatm because traceable standard gases are frequently measured in situ. For pH

53   measurements on profiling floats (Johnson et al., 2016), sensor performance is validated by comparing to a deep

54   reference pH field that is calculated using empirical algorithms (Williams et al., 2016; Bittig et al., 2018; Carter et al.,

55   2018). This approach has demonstrated the ability to obtain high quality pH measurements from a network of profiling

56   floats (Johnson et al., 2017) but requires measurements in the deep, comparatively stable ocean. It is atypical for other

57   pH sensors, including coastal moored sensors, to have an automated or remote validation. Therefore, on such

58   deployments, validation has largely relied on discrete samples taken alongside the sensor (Bresnahan et al., 2014;

59   McLaughlin et al., 2017; Takeshita et al., 2018), which presents unique challenges; primarily that spatiotemporal

60   discrepancy can lead to errors of > 0.1, especially in highly dynamic systems (Bresnahan et al., 2014).

61   Similar to the method in use by pCO$_2$ systems, one approach to validate in situ pH sensors is by measuring a

62   reference material or pH standard, one or more times during a sensor deployment. The most commonly used standard

63   for oceanographic pH measurement is equimolar tris (2-amino-2-hydroxymethyl-propane-1,3-diol) buffer in artificial

64   seawater, hereafter referred to as tris (DelValls and Dickson, 1998). The pH of tris has been characterized over a range

65   of temperature, salinity, and pressure (DelValls and Dickson, 1998; Rodriguez et al., 2015; Takeshita et al., 2017;

66   Müller et al., 2018), allowing for accurate calculation of tris pH across a wide range of marine conditions. Furthermore,

67   when stored in borosilicate bottles and under ideal conditions, these buffers have been shown to be stable to better

68   than 0.0005 over a year (Dickson, 1993; Nemzer and Dickson, 2005), making tris a good candidate for in situ

69   validation of long term deployments of autonomous pH sensors. To be utilized for in situ applications, the reference

70   solution must be stored in flexible bags (as in, Hales et al., 2005; Seidel et al., 2008; Sayles and Eck, 2009; Spaulding

71   et al., 2014; Wang et al., 2015; Lai et al., 2018). Recently, in situ sensor validation using bagged tris was demonstrated

72   by Lai et al. (2018) during a 150 day deployment of an autonomous pH sensor, where the tris standard was measured

73   in situ every 5 days. However, the stability of tris when stored in bags has not been quantified systematically using

74   spectrophotometric bench-top pH measurement techniques recommended as best practices (Dickson et al., 2007).

75   In this work we quantified the stability of tris stored in bags for 300 days. Tris from four separately prepared

76   batches was stored in two bag types either in a lab or submerged in seawater. In addition, one batch was stored in

77   borosilicate bottles in the lab as a control. Spectrophotometric pH measurements were made approximately every two

78   months on each bag of tris. Throughout the experiment, Certified Reference Materials (CRMs) for CO$_2$ (Dickson,

79   2001) were used to assess the stability of the spectrophotometric pH system.

2. Methods

80

81   Two bag types were tested for storing tris (Figure 1). Bag type 1 is based on a design used in the "Burke-o-

82 Lator" system (Hales et al., 2005; Bandstra et al., 2006), made from PAKDRY 7500 barrier film (IMPAK P75C0919).

83 The barrier film is made of layers of polyester and nylon with a sealant layer of metallocene polyethylene. Two 23 x

84 48 cm (9" x 19") sheets were heat sealed on three sides, forming a pocket, and a 1.9 cm (¾") diameter hole was cut

85 into one of the pocket walls for the bulkhead fitting and bulkhead nut (McMaster-Carr 8674T55). The bulkhead was

86 sealed into the wall with a silicone gasket (McMaster-Carr 9010K13), washer (McMaster-Carr 95649A256), and

87 coated with silicone sealant (McMaster-Carr 74955A53). A "push-to-connect" ball valve fitting (McMaster-Carr

88 4379K41), was attached to the bulkhead. The modified pocket was rinsed, dried, and heat sealed along the final edge

89 to create a ~4 L bag. Bags were left to dry for at least 24 hours before filling. Bag type 2 was a 3 L Cali-5-Bond bag

90 purchased from Calibrated Instruments and used without modification. It is a multi-layer bag made of plastic,

91 aluminium foil (to prevent liquid and gas permeation), a layer of inert HDPE (to form a non-reactive inner wall) and,

92 a polycarbonate Stopcock Luer valve.

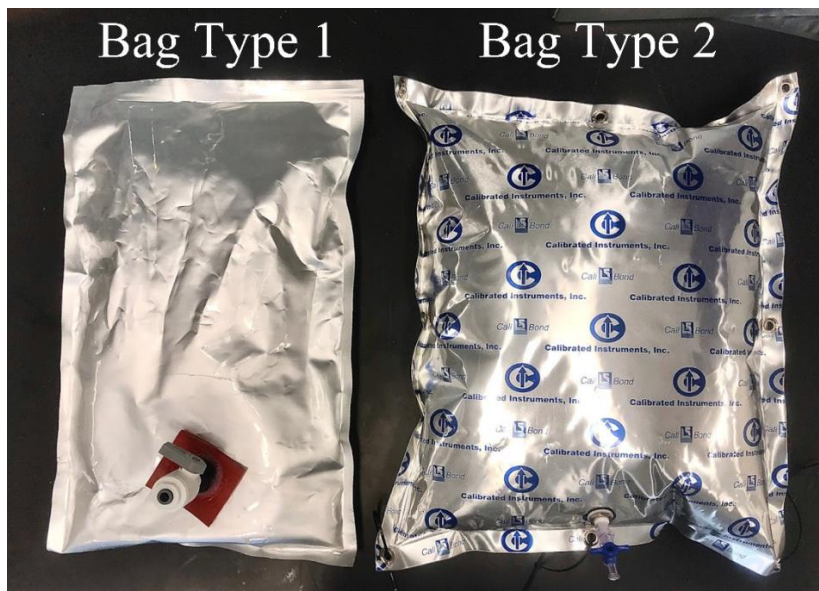



Figure 1: A picture of bag type 1 and 2 used to store tris in this study.

93   In this experiment, four batches of tris were prepared following the procedure in DelValls and Dickson

94 (1998), using off-the-shelf reagents with no additional standardization or purification (e.g. recrystallization of salts).

95 The focus of this paper is stability of bagged tris over time and does not prioritize obtaining highly accurate equimolar

96 tris (as would be necessary for characterization of thermodynamic constants, for example). The calculated pH of tris

97 in this study was 8.2652 at 20°C, based on quantity of reagents used. This is 0.0135 higher than the pH of equimolar

98 tris, 8.2517 at 20°C (DelValls and Dickson, 1998). The pH discrepancy was due to a unit error in the measurement of

99 HCl (our preparation used mol/L rather than the prescribed mol/kg-sol). Reagents used include: tris (2-amino-2-

100 hydroxymethyl-propane-1,3-diol), Fisher Scientific, PN: T395-1, lot 170360; NaCl, Fisher Scientific, PN: S641-212,

101   lot 127252; Na_2SO_4, Fisher Scientific, PN: S421-1, lot 134837; KCl, Fisher Scientific, PN: P217-500, lot 174416;

102     $MgCl_2$, Teknova, PN: M0304, lot M030427E1401; $CaCl_2$, Amresco, PN: E506-500mL, lot 0982C089; HCl, Fisher

103   PN: SA48-1, lot 175004 (batch 1, 2 and 3), lot 188768 (batch 4).

104   Three stability tests were initiated at different times over the course of 18 months. The initiation of a given

105   test is defined as the date of preparation of the tris used in that test. A summary of the differences between these tests

106   is shown in Table 1 and described here. Each bag has a unique identifier in the format of "Batch #, Bag #, Lab or

107   Tank." Test 1: Tris batches 1 and 2 in bag types 1 and 2 were used. Each bag was rinsed before filling; 3 times with

108   deionized water (DI), 5 times with ultrapure water (> 18 MΩ resistivity) and 3 times with 200 mL of tris. Tris bags

109   were stored on a lab bench or in a 5,000 L test tank filled with ozone-sterilized, filtered seawater. Test 2: Tris from

110   batch 3 was used and followed the same procedure as test 1. Bag type 2 experienced delamination of exterior layers

111   when stored in seawater and was not used in further testing.

112   Test 3: Tris from batch 4 and only bag type 1 were used. Each bag was rinsed before filling; 3 times with DI,

113   5 times with ultrapure water, and at least 6 times with 200 mL of tris. Multiple bags of type 1 containing the same

114   batch of tris were stored on a lab bench; these bags are differentiated with letters A to D. Bag 1A was rinsed with 200

115   mL of tris 6 times before filling. Additionally, tris from batch 4 was stored in borosilicate bottles following the

116   procedure in Nemzer and Dickson (2005). Bags and bottles were stored on a lab bench (no tank storage). In addition

117   to pH measurements, dissolved inorganic carbon (C_T) was measured on both bagged and bottled tris to see if changes

118   in pH were due to increased CO_2. C_T samples were measured using a custom-built system based on an infrared analyser

119   (LI-COR 7000) similar to systems used by O'Sullivan and Millero (1998) and Friederich et al. (2002). The precision

120   of the C_T measurements was ± 1.4 μmol/kg (pooled standard deviation, $n_{samples}$=15, $n_{measurements}$=44).

121   **Table 1: Tris preparation and storage.**

	Bag Type	Tris Batch	Date Made	Storage Location	Rinse Procedure	C_T Measured
Test 1	1 & 2	1 & 2	13 Dec 2017	Lab & Tank	3x DI, 5x ultrapure, 3x tris	No
Test 2	1 & 2	3	13 April 2018	Lab & Tank	3x DI, 5x ultrapure, 3x tris	No
Test 3	1 & bottle	4	26 February 2019	Lab	3x DI, 5x ultrapure, ≥ 6x tris	Yes

122

123   Tris pH was measured every 55 ± 20 days (mean ± standard deviation of measurement interval) throughout

124   the experiment. The pH of tris was measured in triplicate at each time point with spectrophotometry using m-cresol

125   purple as the indicator dye (mCP, Aldrich, lot MKBH6858V) using the system described in Carter et al. (2013).

126   Absorbance measurements were made in a 10-cm jacketed cell, and the temperature was measured directly adjacent

127   to the cell outflow using a NIST-traceable thermometer (± 0.1 °C, QTI DTU6028P-001-SC). Blank and sample were

128   held for 3 minutes in the jacketed flow cell prior to absorbance measurements. On average, temperature was stable to

129   within a 0.02 °C range over the course of the day; the mean temperature throughout the experiment was 20.09 ± 0.23

130   °C (1 σ), although temperature was 0.6 °C higher than the average on one measurement day. Spectrophotometric pH

131   measurements are reported at 20 °C by adjusting the measured pH value at the measured cell temperature T_C

132   (pH_{spec,T_C}) to 20 °C ($pH_{spec,20°C}$) using the known temperature dependence of tris (pH_{tris}) as follows:

$$pH_{spec,20°C} = pH_{spec,T_C} - (pH_{tris,T_C} - pH_{tris,20°C}) \tag{1}$$

133   pH_{tris,T_C} and $pH_{tris,20°C}$ were calculated using Eq. (18) in DelValls and Dickson (1998). Throughout the experiment,

134   CRMs (procured from A. Dickson, Scripps Institution of Oceanography) for seawater C_T and total alkalinity were

135   measured regularly to verify instrument performance (Dickson, 2001). A time-series of CRM measurements over the

136   duration of the work described here showed no systematic drift. (Fig. A1 in Appendix A).

137   To account for pH-dependent errors from impurities in unpurified mCP, a pH-dependent correction factor

138   was determined based on the protocol outlined in Liu et al. (2011). Briefly, pH of seawater was measured subsequently

139   using impure dye (pH_{impure}) and purified dye (pH_{pure}) over a range of pH between 7.4 to 8.2 at approximately 0.2

140   intervals. Triplicate measurements were made at each pH. A second order pH-dependent error was observed as

141   previously described, following the equation ($R^2 = 0.975$, RMSE = 0.000434):

$$pH_{pure} = pH_{impure} - 0.0047777 \times pH_{impure}^2 + 0.0668875 \times pH_{impure} - 0.2359740 \tag{2}$$

142   All pH_{spec} reported in this study has this correction applied.

143   Measurements of tris batches 1 and 2 made in the first 150 days have been removed from the data set due to

144   procedural changes made to the spectrophotometric pH system to correct for problems with temperature equilibration.

145   Following normal practices for spectrophotometric pH measurement, outliers were removed from the data set if the

146   absorbance at 760 nm was above 0.005 or below –0.002 (indicative of a measurement problem, such as a bubble or

147   lamp drift), resulting in the removal of 2 out of 163 measurements. Additionally, outliers were removed from the data

148   set if they were greater than three standard deviations from the mean of a measurement triplicate, where standard

149   deviation is calculated as using all sets of triplicates (1 standard deviation = 0.0004, n = 55), resulting in the removal

150   of 2 of 161 remaining measurements. The remaining 159 measurements were used for the analysis presented here. An

151   analysis of variation, or ANOVA, was used to detect the dependence of the results on tris batch, bag/bottle type and

152   storage location. Analysis was performed using MATLAB R2020a and the standard function "anovan()." pH and C_T

153   data are available via the UC San Diego Library Digital Collections (Wolfe et al., 2021).

154   **3.   Results & Discussion**

155   Figure 2 depicts $pH_{tris,20C}$, stored in either a bag or bottle, as a function of time and is subdivided for tests 1,

156   2, and 3. A negative, linear drift was observed for all bags or bottles. A linear regression was calculated for each

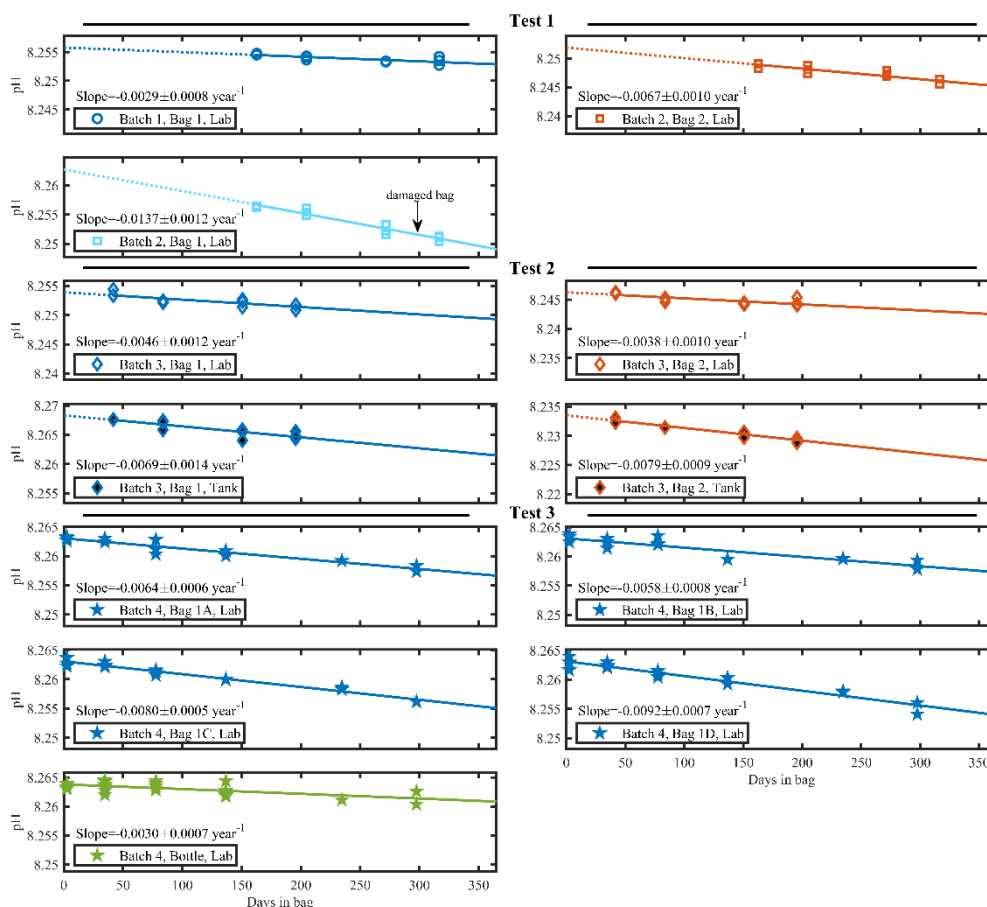
157   experimental condition and, in the cases where measurements at t = 0 were removed due to protocol changes described

158   above, the line is extrapolated back to t = 0, shown by the dotted line. The measured or extrapolated y-intercept is

159   reported as the initial pH in Table 2. In all tests, trendlines are extrapolated to t = 365 days to illustrate observed and

160   predicted drift over the course of a year as shown by the solid line. For ease of visual comparison, the y-axis of each

161   subplot has an identical pH range of 0.017.

162

163

164   **Figure 2: Rates of pH drift in tris buffer. Bag type 1 is shown in blue (light blue for the damaged bag of type 1), 2 in orange**
165   **and bottle in green. Tris batch 1 is depicted as circles, 2 as squares, 3 as diamonds and 4 as stars. Storage location in tank**
166   **has a black fill and lab symbols have no fill. The solid line is a linear regression starting at the first included pH measurement**
167   **and ending 365 days after the tris was bagged. The dotted line illustrates the extrapolation back to 0 days stored in bag**
168   **when measurements at t = 0 do not exist. The range of the y-axis scale is fixed at 0.017 pH for all subplots.**

169

Table 2: Linear regression statistics from trendlines shown in Fig. 1 and 2. The last row shows the regression statistics for tris from all batches, in either bag type, stored in the lab or test tank. Slope and intercept are shown as mean ± 95% confidence intervals. * Indicates the outlier (Batch 2, Bag 1, Lab) caused by a damaged bag. The outlier, "Batch 2, Bag 1, lab", was not used in the "All Batches, All Bags, Lab or Tank" composite. † In all batches, all bags, lab or tank, the slope was calculated with a linear fit of all (non-outlier) tris measurements. The RMSE is the mean RMSE of all (non-outlier) bag fits. ‡ The calculated tris pH was calculated at 20°C; however, this calculated pH is 0.0135 higher than equimolar tris as noted above (DelValls and Dickson, 1998).

Batch & Storage Method	Slope (mpH yr^{-1})	Intercept (Initial pH)	RMSE (mpH)	r^2	n
Batch 1, Bag 1, Lab	-2.9 ± 1.7	8.2558 ± 0.0012	0.43	0.59	12
Batch 2, Bag 1, Lab*	-13.7 ± 2.7	8.2627 ± 0.0018	0.61	0.94	11
Batch 2, Bag 2, Lab	-6.7 ± 2.2	8.2519 ± 0.0015	0.55	0.82	12
Batch 3, Bag 1, Lab	-4.6 ± 2.7	8.2539 ± 0.0010	0.62	0.62	11
Batch 3, Bag 1, Tank	-6.9 ± 3.2	8.2683 ± 0.0012	0.73	0.73	11
Batch 3, Bag 2, Lab	-3.8 ± 2.1	8.2462 ± 0.0008	0.54	0.61	12
Batch 3, Bag 2, Tank	-7.9 ± 2.1	8.2335 ± 0.0008	0.44	0.92	9
Batch 4, Bag 1A, Lab	-6.4 ± 1.3	8.2630 ± 0.0005	0.64	0.90	14
Batch 4, Bag 1B, Lab	-5.8 ± 1.8	8.2631 ± 0.0008	0.91	0.79	15
Batch 4, Bag 1C, Lab	-8.0 ± 1.0	8.2631 ± 0.0004	0.49	0.96	15
Batch 4, Bag 1D, Lab	-9.2 ± 1.6	8.2631 ± 0.0007	0.80	0.92	15
Batch 4, Bottle, Lab	-3.0 ± 1.4	8.2638 ± 0.0005	0.81	0.44	25
All Batches, All Bags, Lab or Tank†	-5.8 ± 0.8	–	0.72	0.66	126
Calculated tris pH‡	–	8.2652	–	–	–

Only bags from test 3, using tris batch 4 and bag type 1, have direct initial pH measurements and replicate bags. Initial pH measurements of these 4 bags were 8.2630 ± 0.0007 (mean ± standard deviation, n = 12). Importantly, the very low standard deviation suggests that a single initial pH measurement is representative of all replicate bags filled with a single tris batch, if the preparation procedure used in test 3 is followed. This inter-bag consistency is beneficial because it reduces the number of initial pH measurements required when filling multiple bags. There is also strong agreement in initial pH measurements between bagged and bottled tris in test 3, with the initial pH of bottled tris 0.0007 higher than bagged tris (8.26327 ± 0.0004, n = 6). The differences in filling procedure or impurities between bags and bottles in test 3 appear to have little effect on the initial pH. The mean initial pH of tris batch 4 is 0.002 (n = 5) lower than calculated $pH_{tris,20°C}$ (Fig. A2). This difference between the mean initial pH of tris batch 4 and calculated $pH_{tris,20°C}$ is similar in direction and magnitude to those reported in other studies: DeGrandpre et al. (2014) reported –0.0012 ±0.0025 and Müller and Rehder (2018) reported -0.002 to -0.008 (measured pH minus pH_{tris,T_C}). With standard laboratory equipment and off-the-shelf reagents, an uncertainty of 0.006 is expected in prepared tris (Paulsen and Dickson, 2020). Measurements were also made on Dickson standard tris (batch T35) using the same instrument and the pH was 0.0019 higher than the calculated $pH_{tris,20°C}$ (n = 2). In tests 1 and 2, the initial pH was extrapolated from a linear regression. The extrapolated initial pH values are more variable and lower (on average) than those directly

193   measured (Fig. A2). These differences may be a result of the extrapolation or different experimental variables such as

194   the increased rinsing of bags, or the single bag type and storage location used in test 3.

195

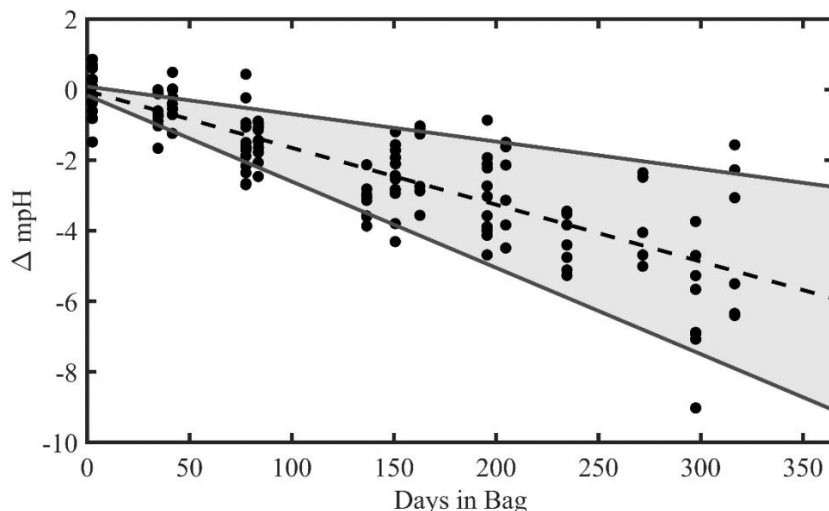



Figure 3: The dots represent every measurement made on a (non-damaged) bag of tris. The dotted line is the "All Bags, All Batches, Lab or Tank" regression. The grey shaded region is the observational 95% confidence interval (CI). The CI is intended to estimate the future pH of a tris bag (with known initial pH and an unmeasured bag specific drift rate). The upper and lower bounds are -0.0028 and -0.0091 pH per year, respectively.

200   Figure 3 depicts a composite of all test results as the change from the initial pH of tris

201    $\left(\Delta pH = pH_{spec,20°C}^{t=day} - pH_{spec,20°C}^{t=0}\right)$ as a function of time elapsed since bagging. By definition, $\Delta pH = 0$ at t = 0. We

202   illustrate these pH changes as ΔmpH ($\Delta pH \times 1000$) due to their small magnitude. A linear regression on all pH

203   measurements, excluding the outlier of "Batch 2, Bag 1, Lab", of tris stored in bag types 1 or 2, has a slope of –0.0058

204   ± 0.0008 yr^{-1} (mean ± 95% C.I.). The outlier (Batch 2, Bag 1, Lab) was excluded due to noticeable damage to the bag

205   (see Fig. A3 in Appendix A), which is believed to have caused it to drift at more than two times the average rate of

206   the other bags. This observation highlights the importance of maintaining bag integrity, particularly during use in the

207   field. For Figure 3, the only comparable example found in the literature is the work of Lai et al. (2018). In this work,

208   Lai et al. (2018) used bagged tris for sensor calibration, with in situ tris measurements made over 150 days. Lai et al.

209   (2018) did not report a drift in the pH of bagged tris over the deployment; however, the reported precision of the

210   SAMI-pH in situ instrument (± 0.003) would not resolve the expected drift shown in our Figure 3. Therefore, the

211   results of Lai et al. (2018) are not inconsistent with our study.

212   A significant increase in C_T was observed for all types of bags and bottles in Experiment 3 (Figure 4). A high

213   correlation between solution pH and C_T was observed, with a slope of –0.0029 ± 0.0006 pH per 100 µmol kg^{-1} (n =

214   14, r^2 = 0.70), suggesting that the drift in tris pH was primarily driven by an increase in CO_2. The observed slope

215   agrees closely with a theoretical prediction of a linear decrease in pH of –0.0024 per 100 µmol kg^{-1} of C_T added (over

216   the range of C_T observed). The linear decrease was calculated using an ad hoc acid–base equilibrium model of seawater

217   including tris in addition to the CO_2 and other minor acid–base systems. There are two possible sources of the

218   increasing C_T: gas exchange of CO_2 with the environment and microbial respiration within the storage vessel. Gas

219   exchange should not be a significant source of CO_2 for tris stored in a borosilicate bottle, as this is the standard

220   equipment used to store seawater CO_2 and tris buffers and has been designed to minimize gas exchange (Dickson et

221   al. 2007). Therefore, it is likely that respiration was the primary driver for the increase in C_T for tris stored in bottles.

222   On average, pH drift of tris stored in bags was larger than that in the standard bottle (Figure 2), indicating either an

223   additional source of CO_2 from gas exchange, or larger amounts of respiration. Distinguishing between these two

224   theorized sources would require measurements of additional parameters such as dissolved organic carbon.

225

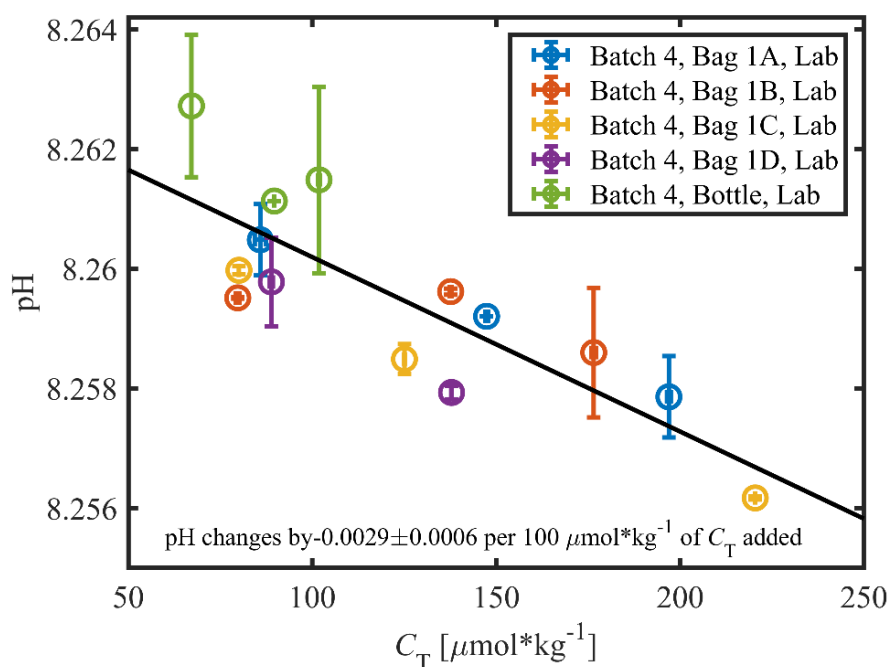



Figure 4: pH plotted against C_T shows a linear relationship between the two parameters in tris buffer with a slope of –0.0029 pH for every 100 μmol kg^{-1} of C_T added. The measurements shown are from three sampling occurrences between 130–300 days stored on bags and bottles used in Test 3. Only two measurements are shown for "Batch 4, Bag 1D, Lab" because it ran empty before C_T were made.

230   The pH stability of tris could be improved by reducing either source of C_T: gas exchange or microbial

231   respiration. For bags, CO_2 may diffuse through the fittings, gasket, or bag walls, particularly if damaged. The relatively

232   small breaks in the aluminium foil layer caused "Batch 2, Bag 1, Lab" to drift more than twice as fast as the average

233   bag. Storage bag, fitting, and gasket material, as well as careful handling, are therefore important factors in minimizing

234   gas exchange. For example, silicone is permeable to CO_2, and thus could have been a path of gas exchange into the

235   tris for this experiment. As noted above, Nemzer and Dickson (2005) found an almost negligible drift of 0.5 mpH yr^{-1}

236   in bottled tris. Our bottled tris drifted at –3.0 mpH yr^{-1} (n = 10 bottles measured over 161 days), approximately half

237   the rate of the tris stored in bags. While –3.0 mpH yr^{-1} is near the detection limit of our measurements, it suggests that

238   the bottling protocol used in this study was not as well controlled as that of Nemzer and Dickson (2005). For example,

239   the Dickson lab regularly uses an annealing oven to combust all trace organic films that may persist on glass bottles,

240   but in our study, bottles were not annealed. Although bags cannot be annealed, future steps that may be worth

241   consideration to reduce microbial respiration in bags include addition of a biocide to the tris solution, acid cleaning

242   the bags, and using ultraviolet light to remove organics from the ultrapure water used to prepare tris. There are some

243   disadvantages to these proposed steps. Addition of a biocide may not be ideal for use in sensitive environments if the

244   tris is discharged after use and would alter the composition of the solution slightly. While rinsing or prolonged soaking

245   of the bags with an acid may help to remove organics, it is unclear if it would have negative effects on the integrity of

246   the bags.

247   Both bag type 1 and 2 experienced problems with structural integrity during this experiment. A single type 2

248   bag experienced delamination of exterior bag layers when stored submerged in seawater, causing the eventual tearing

249   and failure of the bag when handling. Bag type 2 was not used in test 3 due to this failure. It should be noted that other

250   studies successfully used bag type 2 submerged the bag in seawater for less time than in this experiment (Sayles and

251   Eck, 2009; Aßmann et al., 2011; Wang et al., 2015). A single bag type 1 had the subtler problem of small breaks in

252   the aluminium foil bag layer, likely causing an increased pH drift rate. In non-damaged bags, factors such as bag

253   type/bottle, lab/tank storage, or tris batch did not have statistically significant (p-value < 0.05) correlations with the

254   pH drift of tris (p-values 0.12, 0.11 and 0.09, respectively). The results of the ANOVA support that tris can be held in

255   bag type 1 or 2 and stored in a lab or tank and the pH will drift similarly regardless of storage method for up to 300

256   days. Additional bag types could be tested, such as bags made by Pollution Measurement Corp. used by Lai et al.

257   (2018) or Scholle DuraShield used by Takeshita et al. (2015).

258   **4. Conclusions**

259   This article describes our characterization of the stability of tris buffer in artificial seawater when stored in

260   flexible and purportedly gas-impermeable bags. Several different tests, initiated over the course of a year and a half

261   and lasting up to 300 days, exhibited average drift rates of -5.8 mpH yr^{-1}. In comparison, tris stored in standard

262   borosilicate bottles was shown to have a drift rate of -3.0 mpH yr^{-1}. For yearlong deployments, an expected pH change

263   of -0.0058 is well below the weather quality threshold of 0.02 pH units. This low drift rate demonstrates the value of

264   bagged tris for in situ validation of autonomous pH sensors (regardless of sensor operating principles), particularly in

265   highly dynamic areas where repeatability of calibration based on discrete samples is challenging. Given the thorough

266   characterization of tris over wide ranges of environmental variables, this contribution can aid in the traceability and

267   intercomparability of pH sensor measurements.

268   Periodic measurement of bagged tris in situ would allow for detection of sensor drift. Most in situ pH sensors

269   are deployed in the euphotic zone which allows for biofouling and sedimentation in coastal areas, leading to sensor

270   drift (Bresnahan et al., 2014) that could be identified and potentially corrected. Such periodic calibration/validation

271   would aid in identifying sensor issues and allow for greater consistency and continuity between a timeseries and

272   planned or vicarious crossovers where an automated calibration can be used to augment or replace pre- and post-

273   deployment calibrations/validations.

Author contribution

WW performed formal analysis, visualization, and writing – original draft preparation. KS and TW contributed to investigation and writing – review & editing. PB, YT, and TM contributed to funding acquisition, conceptualization, formal analysis, and writing – review & editing.

Competing interests

The authors declare that they have no conflict of interest.

Data availability

pH and C_T data are available via the UC San Diego Library Digital Collections (Wolfe et al., 2021).

Acknowledgements

We thank May-Linn Paulsen and Andrew Dickson lab for sharing their tris expertise throughout this project. We thank the National Science Foundation Ocean Technology and Interdisciplinary Coordination (NSF-OTIC 1736905 and NSF-OTIC 1736864) and the David and Lucile Packard Foundation for supporting this work.

References

ACT: Protocols for the Performance Verification of In Situ pH Sensors, Alliance for Coastal Technologies, https://doi.org/10.25607/OBP-331, 2012.

Aßmann , S., Frank, C., and Kortzinger, A.: Spectrophotometric high-precision seawater pH determination for use in underway measuring systems, Ocean Sci., 7, 597-607, https://doi.org/10.5194/os-7-597-2011, 2011.

Bandstra, L., Hales, B., and Takahashi, T.: High-frequency measurements of total CO_2: Method development and first oceanographic observations, Mar. Chem., 100, 24-38, https://doi.org/10.1016/j.marchem.2005.10.009, 2006.

Bates, N., Astor, Y., Church, M., Currie, K., Dore, J., Gonzalez-Davila, M., Lorenzoni, L., Muller-Karger, F., Olafsson, J., and Santana-Casiano, J.: A Time-Series View of Changing Surface Ocean Chemistry Due to Ocean Uptake of Anthropogenic CO_2 and Ocean Acidification, J. Oceanogr., 27, 126-141, https://doi.org/10.5670/oceanog.2014.16, 2014.

Bittig, H. C., Steinhoff, T., Claustre, H., Fiedler, B., Williams, N. L., Sauzède, R., Körtzinger, A., and Gattuso, J.-P.: An Alternative to Static Climatologies: Robust Estimation of Open Ocean CO_2 Variables and Nutrient Concentrations From T, S, and O_2 Data Using Bayesian Neural Networks, Front. Mar. Sci., 5, https://doi.org/10.3389/fmars.2018.00328, 2018.

Bockmon, E. E., and Dickson, A. G.: An inter-laboratory comparison assessing the quality of seawater carbon dioxide measurements, Mar. Chem., 171, 36-43, https://doi.org/10.1016/j.marchem.2015.02.002, 2015.

Branch, T. A., DeJoseph, B. M., Ray, L. J., and Wagner, C. A.: Impacts of ocean acidification on marine seafood, Trends Ecol. Evol., 28, 178-186, https://doi.org/10.1016/j.tree.2012.10.001, 2013.

Bresnahan, P. J., Martz, T. R., Takeshita, Y., Johnson, K. S., and LaShomb, M.: Best practices for autonomous measurement of seawater pH with the Honeywell Durafet, Methods Oceanogr., 9, 44-60, https://doi.org/10.1016/j.mio.2014.08.003, 2014.

Bushinsky, S. M., Takeshita, Y., and Williams, N. L.: Observing changes in ocean carbonate chemistry: our autonomous future, Curr. Clim, 5, 207-220, https://doi.org/10.1007/s40641-019-00129-8, 2019.

Byrne, R. H.: Measuring Ocean Acidification: New Technology for a New Era of Ocean Chemistry, Environ. Sci. Technol., 48, 5352-5360, https://doi.org/10.1021/es405819p, 2014.

312 Carter, B., Radich, J., Doyle, H., and Dickson, A.: An automated system for spectrophotometric seawater pH
313   measurements, Limnol. Oceanogr. Methods, 11, 16-27, https://doi.org/10.4319/lom.2013.11.16, 2013.
314 Carter, B. R., Feely, R. A., Williams, N. L., Dickson, A. G., Fong, M. B., and Takeshita, Y.: Updated methods for
315   global locally interpolated estimation of alkalinity, pH, and nitrate, Limnol. Oceanogr. Methods, 16, 119-
316   131, https://doi.org/10.1002/lom3.10232, 2018.
317 Chavez, F., Pennington, J. T., Michisaki, R., Blum, M., Chavez, G., Friederich, J., Jones, B., Herlien, R., Kieft, B.,
318   Hobson, B., Ren, A., Ryan, J., Sevadjian, J., Wahl, C., Walz, K., Yamahara, K., Friederich, G., and Messié,
319   M.: Climate Variability and Change: Response of a Coastal Ocean Ecosystem, J. Oceanogr., 30, 128-145,
320   https://doi.org/10.5670/oceanog.2017.429, 2017.
321 Cooley, S. R., and Doney, S. C.: Anticipating ocean acidification's economic consequences for commercial fisheries,
322   Environ. Res. Lett., 4, 8, https://doi.org/10.1088/1748-9326/4/2/024007, 2009.
323 DeGrandpre, M. D., Spaulding, R. S., Newton, J. O., Jaqueth, E. J., Hamblock, S. E., Umansky, A. A., and Harris, K.
324   E.: Considerations for the measurement of spectrophotometric pH for ocean acidification and other studies,
325   Limnol. Oceanogr. Methods, 12, 830-839, https://doi.org/10.4319/lom.2014.12.830, 2014.
326 DelValls, T., and Dickson, A.: The pH of buffers based on 2-amino-2-hydroxymethyl-1,3-propanediol ('tris') in
327   synthetic sea water, Deep Sea Res. Part I, 45, 1541-1554, https://doi.org/10.1016/S0967-0637(98)00019-3,
328   1998.
329 Dickson, A. G.: pH buffers for sea-water media based on the total hydrogen-ion concentration scale, Deep Sea Res.
330   Part I, 40, 107-118, https://doi.org/10.1016/0967-0637(93)90055-8, 1993.
331 Dickson, A. G.: Reference materials for oceanic CO_2 measurements, J. Oceanogr., 14, 21-22, 2001.
332 Dickson, A. G., Sabine, C. L., and Christian, J. R.: Guide to Best Practices for Ocean CO_2 Measurements, PICES
333   Special Publication 3, North Pacific Marine Science Organization, Sidney, British Columbia, 191 pp., 2007.
334 Doney, S. C., Fabry, V. J., Feely, R. A., and Kleypas, J. A.: Ocean acidification: the other CO_2 problem, Annu. Rev.
335   Mar. Science, 1, 169-192, https://doi.org/10.1146/annurev.marine.010908.163834, 2009.
336 Doney, S. C., Busch, D. S., Cooley, S. R., and Kroeker, K. J.: The impacts of ocean acidification on marine ecosystems
337   and reliant human communities, Annu. Rev. Environ. Resour., 45, https://doi.org/10.1146/annurev-environ-
338   012320-083019, 2020.
339 Friederich, G., Walz, P., Burczynski, M., and Chavez, F.: Inorganic carbon in the central California upwelling system
340   during the 1997–1999 El Niño–La Niña event, Prog. Oceanogr., 54, 185-203, https://doi.org/10.1016/S0079-
341   6611(02)00049-6, 2002.
342 Hales, B., Takahashi, T., and Bandstra, L.: Atmospheric CO_2 uptake by a coastal upwelling system, Global
343   Biogeochem. Cycles, 19, https://doi.org/10.1029/2004gb002295, 2005.
344 Johnson, K. S., Jannasch, H. W., Coletti, L. J., Elrod, V. A., Martz, T. R., Takeshita, Y., Carlson, R. J., and Connery,
345   J. G.: Deep-Sea DuraFET: A Pressure Tolerant pH Sensor Designed for Global Sensor Networks, Anal.
346   Chem., 88, 3249-3256, https://doi.org/10.1021/acs.analchem.5b04653, 2016.
347 Johnson, K. S., Plant, J. N., Coletti, L. J., Jannasch, H. W., Sakamoto, C. M., Riser, S. C., Swift, D. D., Williams, N.
348   L., Boss, E., Haëntjens, N., Talley, L. D., and Sarmiento, J. L.: Biogeochemical sensor performance in the
349   SOCCOM   profiling   float   array,   J.   Geophys.   Res.:   Oceans,   122,   6416-6436,
350   https://doi.org/10.1002/2017jc012838, 2017.
351 Karl, D. M.: Oceanic ecosystem time-series programs: Ten lessons learned, J. Oceanogr., 23, 104-125,
352   https://doi.org/10.5670/oceanog.2010.27, 2010.
353 Lai, C.-Z., DeGrandpre, M. D., and Darlington, R. C.: Autonomous Optofluidic Chemical Analyzers for Marine
354   Applications: Insights from the Submersible Autonomous Moored Instruments (SAMI) for pH and pCO_2,
355   Front. Mar. Sci., 4, https://doi.org/10.3389/fmars.2017.00438, 2018.
356 Liu, X. W., Patsavas, M. C., and Byrne, R. H.: Purification and Characterization of meta-Cresol Purple for
357   Spectrophotometric   Seawater   pH   Measurements,   Environ.   Sci.   Technol.,   45,   4862-4868,
358   https://doi.org/10.1021/es200665d, 2011.
359 Martz, T. R., Daly, K. L., Byrne, R. H., Stillman, J. H., and Turk, D.: Technology for ocean acidification research
360   needs and availability, J. Oceanogr., 28, 40-47, https://doi.org/10.5670/oceanog.2015.30, 2015.
361 McLaughlin, K., Dickson, A., Weisberg, S. B., Coale, K., Elrod, V., Hunter, C., Johnson, K. S., Kram, S., Kudela, R.,
362   Martz, T., Negrey, K., Passow, U., Shaughnessy, F., Smith, J. E., Tadesse, D., Washburn, L., and Weis, K.
363   R.: An evaluation of ISFET sensors for coastal pH monitoring applications, Reg. Stud. Mar. Sci., 12, 11-18,
364   https://doi.org/10.1016/j.rsma.2017.02.008, 2017.
365 Müller, J., Bastkowski, F., Sander, B., Seitz, S., Turner, D., Dickson, A., and Rehder, G.: Metrology for pH
366   Measurements in Brackish Waters-Part 1: Extending Electrochemical pH(T) Measurements of TRIS Buffers
367   to Salinities 5-20, Front. Mar. Sci., 5, https://doi.org/10.3389/fmars.2018.00176, 2018.

368 Müller, J. D., and Rehder, G.: Metrology of pH Measurements in Brackish Waters—Part 2: Experimental
369   Characterization of Purified meta-Cresol Purple for Spectrophotometric pHT Measurements, Front. Mar.
370   Sci., 5, 177, https://doi.org/10.3389/fmars.2018.00177, 2018.
371 Nemzer, B., and Dickson, A.: The stability and reproducibility of Tris buffers in synthetic seawater, Mar. Chem., 96,
372   237-242, https://doi.org/10.1016/j.marchem.2005.01.004, 2005.
373 Newton, J., Feely, R., Jewett, E., Williamson, P., and Mathis, J.: Global ocean acidification observing network:
374   requirements and governance plan, GOA-ON, Washington, USA, 57, 2014.
375 O'Sullivan, D. W., and Millero, F. J.: Continual measurement of the total inorganic carbon in surface seawater, Mar.
376   Chem., 60, 75-83, https://doi.org/10.1016/s0304-4203(97)00079-0, 1998.
377 Okazaki, R. R., Sutton, A. J., Feely, R. A., Dickson, A. G., Alin, S. R., Sabine, C. L., Bunje, P. M. E., and Virmani,
378   J. I.: Evaluation of marine pH sensors under controlled and natural conditions for the Wendy Schmidt Ocean
379   Health XPRIZE, Limnol. Oceanogr. Methods, 15, 586-600, https://doi.org/10.1002/lom3.10189, 2017.
380 Paulsen, M. L., and Dickson, A. G.: Preparation of 2-amino-2-hydroxymethyl-1, 3-propanediol (TRIS) pHT buffers
381   in synthetic seawater, Limnol. Oceanogr. Methods, 18, 504-515, https://doi.org/10.1002/lom3.10383, 2020.
382 Pierrot, D., Neill, C., Sullivan, K., Castle, R., Wanninkhof, R., Lüger, H., Johannessen, T., Olsen, A., Feely, R. A.,
383   and Cosca, C. E.: Recommendations for autonomous underway pCO$_2$ measuring systems and data-reduction
384   routines, Deep Sea Res. Part II, 56, 512-522, https://doi.org/10.1016/j.dsr2.2008.12.005, 2009.
385 Rodriguez, C., Huang, F., and Millero, F. J.: The partial molal volume and compressibility of Tris and Tris–HCl in
386   water and 0.725 m NaCl as a function of temperature, Deep Sea Res. Part I, 104, 41-51,
387   https://doi.org/10.1016/j.dsr.2015.06.008, 2015.
388 Sabine, C., Sutton, A., McCabe, K., Lawrence-Slavas, N., Alin, S., Feely, R., Jenkins, R., Maenner, S., Meinig, C.,
389   and Thomas, J.: Evaluation of a new carbon dioxide system for autonomous surface vehicles, J. Atmos.
390   Oceanic Technol., 37, 1305-1317, https://doi.org/10.1175/JTECH-D-20-0010.1, 2020.
391 Sayles, F. L., and Eck, C.: An autonomous instrument for time series analysis of TCO$_2$ from oceanographic moorings,
392   Deep Sea Res. Part I, 56, 1590-1603, https://doi.org/10.1016/j.dsr.2009.04.006, 2009.
393 Seidel, M. P., DeGrandpre, M. D., and Dickson, A. G.: A sensor for in situ indicator-based measurements of seawater
394   pH, Mar. Chem., 109, 18-28, https://doi.org/10.1016/j.marchem.2007.11.013, 2008.
395 Sloyan, B. M., Wanninkhof, R., Kramp, M., Johnson, G. C., Talley, L. D., Tanhua, T., McDonagh, E., Cusack, C.,
396   O'Rourke, E., McGovern, E., Katsumata, K., Diggs, S., Hummon, J., Ishii, M., Azetsu-Scott, K., Boss, E.,
397   Ansorge, I., Perez, F. F., Mercier, H., Williams, M. J. M., Anderson, L., Lee, J. H., Murata, A., Kouketsu, S.,
398   Jeansson, E., Hoppema, M., and Campos, E.: The Global Ocean Ship-Based Hydrographic Investigations
399   Program (GO-SHIP): A Platform for Integrated Multidisciplinary Ocean Sci., Front. Mar. Sci., 6,
400   https://doi.org/10.3389/fmars.2019.00445, 2019.
401 Spaulding, R. S., DeGrandpre, M. D., Beck, J. C., Hart, R. D., Peterson, B., De Carlo, E. H., Drupp, P. S., and Hammar,
402   T. R.: Autonomous in Situ Measurements of Seawater Alkalinity, Environ. Sci. Technol., 48, 9573-9581,
403   https://doi.org/10.1021/es501615x, 2014.
404 Sutton, A. J., Feely, R. A., Maenner-Jones, S., Musielwicz, S., Osborne, J., Dietrich, C., Monacci, N., Cross, J., Bott,
405   R., and Kozyr, A.: Autonomous seawater pCO$_2$ and pH time series from 40 surface buoys and the emergence
406   of anthropogenic trends, Earth Syst. Sci. Data, 421, https://doi.org/10.5194/essd-11-421-2019, 2019.
407 Takeshita, Y., Frieder, C. A., Martz, T. R., Ballard, J. R., Feely, R. A., Kram, S., Nam, S., Navarro, M. O., Price, N.
408   N., and Smith, J. E.: Including high-frequency variability in coastal ocean acidification projections,
409   Biogeosciences, 12, 5853-5870, https://doi.org/10.5194/bg-12-5853-2015, 2015.
410 Takeshita, Y., McGillis, W., Briggs, E. M., Carter, A. L., Donham, E. M., Martz, T. R., Price, N. N., and Smith, J. E.:
411   Assessment of net community production and calcification of a coral reef using a boundary layer approach,
412   J. Geophys. Res.: Oceans, 121, 5655-5671, https://doi.org/10.1002/2016JC011886, 2016.
413 Takeshita, Y., Martz, T. R., Coletti, L. J., Dickson, A. G., Jannasch, H. W., and Johnson, K. S.: The effects of pressure
414   on pH of Tris buffer in synthetic seawater, Mar. Chem., 188, 1-5, 10.1016/j.marchem.2016.11.002, 2017.
415 Takeshita, Y., Johnson, K. S., Martz, T. R., Plant, J. N., and Sarmiento, J. L.: Assessment of Autonomous pH
416   Measurements for Determining Surface Seawater Partial Pressure of CO$_2$, J. Geophys. Res.: Oceans, 123,
417   4003-4013, https://doi.org/10.1029/2017jc013387, 2018.
418 Tilbrook, B., Jewett, E. B., DeGrandpre, M. D., Hernandez-Ayon, J. M., Feely, R. A., Gledhill, D. K., Hansson, L.,
419   Isensee, K., Kurz, M. L., Newton, J. A., Siedlecki, S. A., Chai, F., Dupont, S., Graco, M., Calvo, E., Greeley,
420   D., Kapsenberg, L., Lebrec, M., Pelejero, C., Schoo, K. L., and Telszewski, M.: An Enhanced Ocean
421   Acidification Observing Network: From People to Technology to Data Synthesis and Information Exchange,
422   Front. Mar. Sci., 6, 21, https://doi.org/10.3389/fmars.2019.00337, 2019.

423 Wang, Z. A., Sonnichsen, F. N., Bradley, A. M., Hoering, K. A., Lanagan, T. M., Chu, S. N., Hammar, T. R., and
424   Camilli, R.: In Situ Sensor Technology for Simultaneous Spectrophotometric Measurements of Seawater
425   Total Dissolved Inorganic Carbon and pH, Environ. Sci. Technol., 49, 4441-4449,
426   https://doi.org/10.1021/es504893n, 2015.
427 Wang, Z. A., Moustahfid, H., Mueller, A. V., Michel, A. P. M., Mowlem, M., Glazer, B. T., Mooney, T. A., Michaels,
428   W., McQuillan, J. S., Robidart, J. C., Churchill, J., Sourisseau, M., Daniel, A., Schaap, A., Monk, S.,
429   Friedman, K., and Brehmer, P.: Advancing Observation of Ocean Biogeochemistry, Biology, and
430   Ecosystems With Cost-Effective in situ Sensing Technologies, Front. Mar. Sci., 6, 22,
431   https://doi.org/10.3389/fmars.2019.00519, 2019.
432 Williams, N. L., Juranek, L. W., Johnson, K. S., Feely, R. A., Riser, S. C., Talley, L. D., Russell, J. L., Sarmiento, J.
433   L., and Wanninkhof, R.: Empirical algorithms to estimate water column pH in the Southern Ocean, Geophys.
434   Res. Lett., 43, 3415-3422, https://doi.org/10.1002/2016gl068539, 2016.
435 Wolfe, W. H., Shipley, K. M., Bresnahan, P. J., Takeshita, Y., Wirth, T., Martz, T. R.: Data from: Technical note:
436   stability of tris pH buffer in artificial seawater stored in bags. UC San Diego Library Digital
437   Collections. https://doi.org/10.6075/J0QC022G, 2021.