LIMNOLOGY
and
OCEANOGRAPHY: METHODS

ASLO

Limnol. Oceanogr.: Methods 2021

doi: 10.1002/lom3.10462

Long-term stability and storage of meta-cresol purple solutions for seawater pH measurements

Ryan J. Woosley *

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

Abstract

Changes in seawater pH resulting from anthropogenic influence, termed ocean acidification, have significant implications. Monitoring and evaluating ocean acidification requires highly precise measurements comparable among many laboratories over decades. Short-term repeatability can be achieved with spectrophotometric methods, but comparability among studies is much less certain. The indicator used plays an important role in measurement quality. Meta-cresol purple (mCp), which is used for most seawater pH measurements, is known to break down under exposure to ultraviolet light, which can impact the long-term stability of the indicator, and thus the quality of the measurements. Here, certified reference material and 2-hydroxymethyl-1,3-propanediol (TRIS) buffers were used to assess the long-term stability of mCp solutions. Purified indicator solutions were found to be stable (within 0.0025) at room temperature in a dark container for at least 5.3 years. Uncertainties in unpurified indicators made the stability assessment inconclusive. Such long-term stability minimizes one potential source of uncertainty when comparing measurements and can reduce costs and waste by not prematurely disposing of indicator solutions that remain useable. Changes in the pH of the indicator solution indicate the indicator perturbation correction should be regularly determined. The $A_{434\text{imp}}$ correction method for determining impurities could be used as a quality assurance measure by making measurements over the life of the indicator solution to monitor for changes in the solution. However, the choice of molar absorptivities required for the calculation impacts the magnitude of the correction and thus warrants further study to improve best practices for making corrections to impure indicators.

Ocean pH has been decreasing rapidly as a result of the uptake of anthropogenically produced CO_2 from the atmosphere (Sabine et al. 2004; Waters et al. 2011; Woosley et al. 2016; Carter et al. 2017; Woosley and Millero 2020). This process, termed ocean acidification, has significant implications for many biological (Kroeker et al. 2010; Andersson et al. 2015; Bednaršek et al. 2019) and physical/chemical processes (Millero et al. 2009; Millero and DiTrolio 2010), particularly those involving calcium carbonate (Feely et al. 2004; Woosley et al. 2012). As pH is defined by a logarithmic scale, even small changes in pH reflect significant changes in the acid–base balance of seawater. Therefore, detection of ocean acidification requires high-quality, accurate, and precise measurements (Newton et al. 2015; Dickson et al. 2016). Given that ocean acidification occurs over decades and centuries, detection and monitoring require measurements made by multiple different laboratories to be directly comparable over years and decades. Such comparability requires consistency in methods, and clearly defined protocols with standardization whenever possible.

Highly precise pH measurements (with short term [a few days] standard deviations of ~ 0.0005) are readily achieved using simple spectrophotometric methods (Clayton and Byrne 1993; Liu et al. 2011). These methods rely on a pH-sensitive indicator dye with characterized physical and chemical properties. Concerning the typical pH range of seawater, meta-cresol purple (mCp) is most often used (Clayton and Byrne 1993) although other indicators are available (Zhang and Byrne 1996; French et al. 2002; Patsavas et al. 2013). A full and accurate characterization of an indicator dye (e.g., indicator pK and extinction coefficients) is paramount to achieving accurate and comparable pH measurements (Clayton and Byrne 1993; Liu et al. 2011; DeGrandpre et al. 2014). The indicator itself is a proton donor/acceptor, causing a small but detectable perturbation to the pH of the sample, and requiring a corresponding correction (Clayton and Byrne 1993; Carter et al. 2013). Impurities in commercially

*Correspondence: rwoosley@mit.edu

available indicators have been shown to cause pH-dependent errors (Yao et al. 2007; Liu et al. 2011; Patsavas et al. 2013). The mCp indicator is also known to be sensitive to ultraviolet (UV) light, which can break down the indicator, potentially causing errors and instability in the indicator solution and reducing the quality of measurements (Carter et al. 2013). Aside from UV radiation, oxidation can also alter the chemical properties of the solution, meaning the storage vessel may also be important for long-term stability. Therefore, solutions are often stored in dark containers with small headspaces to minimize interactions with the atmosphere.

Indicator solution stability has been assessed at the short-term scale of individual measurements (Carter et al. 2013), but long-term stability has not. This knowledge gap leads to uncertainty and questions of indicator solution preparation and storage. Indicator purification is expensive, thus revealing the long-term stability of the dye would reduce cost and time expenditure. Each batch of indicator requires the pH of the solution to be carefully adjusted, making it desirable to prepare indicator solutions in large batches. Yet, if the indicator breaks down rapidly, smaller batches prepared as needed would be required to maintain data quality. Regarding practicality for a research cruise, there is debate whether the better practice is preparing a large batch in the laboratory (possibly months ahead of time) or making several smaller batches throughout the cruise, requiring dedicated hours of precious sea time and resulting in potential variability between batches. Additional questions relate to whether a single batch can be used for multiple cruises or for autonomous platforms over the span of months or years (Seidel et al. 2008; Rérolle et al. 2013; Reggiani et al. 2016; Pinto et al. 2019).

To address the feasibility for making larger batches of indicator solutions that can be used across cruises or integrated for deployment on autonomous platforms, an experiment was undertaken to evaluate the long-term stability and storage of multiple mCp indicator solution batches over 5.3 years. The results of this work will be useful in maximizing efficiency and reducing time and cost expenditure while also ensuring that the highest-quality measurements are maintained.

Methods

A custom-designed automated spectrophotometric pH system similar to that of Carter et al. (2013) was used. Briefly, the system consists of an Agilent 8454 UV–Vis spectrophotometer (Agilent Technologies) connected to a 10-mL Kloehn v6 syringe pump (Norgren, Inc.) that draws the sample, mixes in the indicator, and injects the sample into a 10-cm quartz micro-volume spectrophotometric cell (Starna, Inc.). The sample and cell holder were thermostated to 25 ± 0.1°C with a Versacool recirculating water bath (Fisher Scientific). Temperature was measured with a platinum resistance thermometer (Fluke 1523, Fluke Calibration). Each analysis took approximately 5 min. Further details of the instrument and procedures can be found in Woosley (2021). The spectrophotometer spectral 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.) before the start of the experiments. 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 ($< \pm 0.0027$ A).

Eight separate batches of indicator solutions made between 2007 and 2019 were compared to the reference indicator solution (Batch 3) prepared immediately prior to the experiment in September 2020. Details of each indicator solution, including date of production, purity, and storage container, are provided in Table 1. Two batches were stored in different types of containers (P18_2016 and ARC01), allowing for the influence of storage container on stability to be assessed. All indicator solutions were stored in dark containers, or the container was covered (with either black electrical tape or aluminum foil) to minimize UV exposure and were stored at room temperature (generally ~ 20°C, but not monitored or strictly controlled). Both pure and impure indicator solutions were analyzed as the older batches were prepared before purified indicator became available. All pure indicators were obtained from Robert H. Byrne (University of South Florida). For pure indicators, the pH on the total scale (pH_t) was calculated using the equations of Liu et al. (2011). All impure indicators are believed to have been prepared from the same lot of Sigma-Aldrich mCp (87H3629), but this cannot be verified as the lot number was only reported for A16. The impure indicators were prepared before purification was developed and purified indicator became widely available. As a result, a correction to pure indicator was not determined at the time that they were prepared. Determining individual corrections now is not possible because it would, by definition, make the solutions agree with the reference indicator. This is problematic for this evaluation study. Attempts were made to use different correction methods to account for impurities. The uncertainty in the corrections was found to be too great to provide a quantitative evaluation of the impure indicator stability. Thus, the impure indicator stability was inconclusive and will only be discussed qualitatively.

Measurements were made to determine the absorbance of impurities at the 434 nm wavelength following Douglas and Byrne (2017). A 2-L solution of 0.7 M NaCl adjusted to a pH of ~ 12 with NaOH (final concentration ~ 0.01 M) was prepared. For each batch of indicator solution, the absorbances of the NaCl solution were measured three times following the same protocol as the certified reference material (CRM) and 2-hydroxymethyl-1,3-propanediol (TRIS) samples (Woosley 2021). From these measurements, the correction factor for impurities which absorb at 434 nm was determined according to:

Table 1. Details of the production and storage of the different batches of indicator used in these experiments.

Batch	Date produced	Age (yr)	Purity*	Solution	Container	Initial pH	Final pH†	A_{434imp} (L11)‡	A_{434imp} (D14)‡
3 (ref)§	14 Sept 2020	0.00	Pure	NaCl	Labtainer (HDPE)‖	7.8	7.8	−0.0018	−0.0002
1	7 Jun 2019	1.27	Pure	NaCl	Labtainer (HDPE)‖	7.8	7.6	−0.0011	0.0005
Falkor180624	6 Jun 2018	2.28	Pure	NaCl	Glass serum bottle	7.8	7.5	−0.0012	−0.0002
I07N	7 Nov 2017	2.85	Pure	NaCl	Glass vial	7.8	7.7	−0.0008	0.0001
P18_2016	9 Aug 2016	4.10	Pure	NaCl	PE vial	7.8	7.5	−0.0011	0.0001
P18_2016Bag	9 Aug 2016	4.10	Pure	NaCl	Tedlar PVF bag	7.8	7.4	−0.0003	0.0008
ARC01	27 May 2015	5.31	Pure	Milli-Q	Brown HDPE	7.8	7.5	−0.0001	0.0010
ARC01_Bag	27 May 2015	5.31	Pure	Milli-Q	Tedlar PVF bag	7.8	7.2	−0.0008	0.0001
A16	19 Jun 2013	7.24	Impure	Milli-Q	Glass vial	7.8	7.2	0.0024	0.0036
P06	22 Jul 2009	11.16	Impure	Milli-Q	PE vial	7.8	6.4	0.0017	0.0024
P06_Adj	22 Jul 2009	11.16	Impure	Milli-Q	PE vial	7.8	7.8	0.0075	0.0082
P18_2007	1 Sept 2007	13.05	Impure	Milli-Q	Glass vial	7.8	7.4	0.0013	0.0026

*All pure indicators were obtained from Robert H. Byrne (University of South Florida). All impure indicators are thought to be from Sigma-Aldrich (Lot #87H3629).
†Measured using a nanodrop spectrophotometer on 21 September 2020.
‡A_{343imp} determined using the methods of Douglas and Byrne (2017), L11 uses the $\frac{e_3}{e_2}$ of Liu et al. (2011) and D14 uses the $\frac{e_3}{e_2}$ of DeGrandpre et al. (2014).
§Reference indicator to which other indicators are compared.
‖Labtainer (ThermoFischer, Inc.) is a layered bioprocessing container of PE, ethylene vinyl-alcohol, polyester plastic (inner to outer layers).

$$A_{434imp} = \left(1 - \frac{e_3}{e_2} \times R_{obs}\right) \times A_{434obs,} \quad (1)$$

where A_{434imp} is the batch specific correction factor for the absorbance at 434 nm, $\frac{e_3}{e_2}$ is the molar absorptivity. Two different values of $\frac{e_3}{e_2}$ were considered. One was determined an artificial seawater media (Liu et al. 2011), the other was determined in NaCl (DeGrandpre et al. 2014). R_{obs} is the observed absorbance ratio, and A_{434obs} is the observed absorbance at 434 nm. Both observed variables are corrected for background absorbance at 730 nm.

The sample R is then adjusted using the batch specific A_{434imp} according to:

$$R_{pure} = R_{obs}\left(1 + \frac{A_{434imp}}{A_{434obs} - A_{434imp}}\right), \quad (2)$$

where R_{pure} is the adjusted absorbance ratio of the sample and is used to calculate pH_t with the equations of Liu et al. (2011). Note that the A_{434imp} adjustment was determined approximately 5 months after the stability experiments, which adds a small, but likely negligible, amount of uncertainty.

Between analysis of each indicator solution batch the indicator line was detached from the system and flushed with canned air, and connections were wiped with a low-lint Kimwipe to ensure all of the previous indicator was removed. The line was then flushed with at least two times the volume of the line with the new indicator solution. The system was then flushed with seawater twice and two "junk" seawater measurements were made to ensure no carry over from the prior indicator batch. All stability measurements were made within 6 d of the preparation of the reference "Batch 3" indicator. As there is no certified pH_t standard, the stability was assessed as the difference between the reference indicator (Batch 3) and each indicator batch ($\Delta pH_t = pH_{ref} - pH_{ind}$).

CRMs from batch 189 were obtained from Andrew Dickson (Scripps Institution of Oceanography). A 3-L batch of TRIS buffer was prepared according to Paulsen and Dickson (2020). The TRIS was prepared as two separate batches (2 and 1 L), then combined, and mixed overnight before bottling into 125-mL borosilicate glass serum bottles and sealed with butyl rubber stoppers and aluminum seals. For each batch of indicator, four bottles of the CRM were measured four times each with varying amounts of indicator for a total of 16 pH_t CRM measurements per indicator. One bottle of TRIS was also measured four times per indicator batch.

The pH_t perturbation caused by the addition of the indicator was corrected for by extrapolation to zero indicator for CRM measurements. A linear fit of the isosbestic point (488 nm absorbance) vs. pH_t was determined using Matlab® (version 2019b) robust fit function with default tuning parameters. The pH_t without any indicator added is the intercept of the fit, and the standard error of the intercept is taken as the standard error of the measurement. The standard errors of the pure indicators ranged from 0.0005 to 0.0033 (Table 2) and comprises both the precision of the analysis and bottle-to-bottle variability of the CRM. No dye perturbation correction was made for the TRIS measurements due to the buffering capacity of TRIS. The mean and standard error of the TRIS measurements were used in the stability analysis. The standard error of the TRIS measurements ranged from 0.0001 to 0.0023 (Table 2).

Table 2. Measured pH_t, standard error, and ΔpH_t (reference – indicator) for CRM and TRIS for each indicator batch.

	CRM			TRIS		
Batch	**pH_t**	**SE**	**ΔpH_t**	**pHt**	**SE**	**ΔpH_t**
3*	7.8704	0.0010	–	8.0840	0.0004	–
1	7.8699	0.0033	0.0006	8.0816	0.0001	0.0023
Falkor180624	7.8681	0.0007	0.0024	8.0814	0.0023	0.0025
I07N	7.8689	0.0009	0.0015	8.0863	0.0002	−0.0023
P18_2016	7.8695	0.0005	0.0009	8.0851	0.0005	−0.0012
P18_2016Bag	7.8685	0.0007	0.0019	8.0830	0.0003	0.0009
ARC01	7.8704	0.0006	0.0000	8.0836	0.0001	0.0004
ARC01_Bag	7.8680	0.0011	0.0024	8.0840	0.0003	−0.0001

*Reference indicator to which other indicators are compared.

Results and discussion

Evaluation of $A_{434\text{imp}}$

The method of Douglas and Byrne (2017) for determination of absorbances due to impurities is simple. The method can act as a quality assurance measure by confirming that solutions of pure indicator are still pure or as a quality control measure by correcting for impurities in impure solutions. This method assumes that most of the impurities absorb at the 434 nm wavelength and that they can be quantified as $A_{434\text{imp}}$ (*see* "Methods" section). The values of $A_{434\text{imp}}$ were determined for each batch of indicator including the purified batches (Table 1; Fig. 1). According to the theory of the method, a pure indicator would have $A_{434\text{imp}} = 0$, and an impure indicator would have a positive value.

When the $\frac{e_3}{e_2}$ value of Liu et al. (2011) was used to calculate $A_{434\text{imp}}$, the pure indicators had slightly negative values (Fig. 1a), with the impure indicators having positive values. Liu et al. (2011) determined the value in artificial seawater. When the $\frac{e_3}{e_2}$ value of DeGrandpre et al. (2014), which was determined in NaCl, was used, the pure indicators were essentially zero and impure indicators had positive values (Fig. 1b). The difference in $A_{434\text{imp}}$ when using different values of $\frac{e_3}{e_2}$ is likely due to media effects that are known to exist for these parameters (Liu et al. 2011). $A_{434\text{imp}}$ is determined in NaCl rather than natural seawater due to problems of precipitation at the high pH of the measurements. There are other possible explanations for the differences between the $\frac{e_3}{e_2}$ values, which include improvements to the indicator purification process since Liu et al. (2011), impurities at 578 nm which have been ignored, or inaccuracies due to limits of the spectrophotometer used. One of the challenges with this adjustment method is obtaining measurements within the linear range of the spectrophotometer as the absorbances at 578 nm are high and those at 434 nm are near zero (< 0.14). The linearity of the spectrophotometer was validated using NIST traceable certified references (*see* "Methods" section). Any absorbance value higher than 1.2 was excluded from analysis. Subtraction of the background absorbance determined at 730 nm helps to account for noise at low absorbances, but the overall signal to noise ratio may still be too large.

Qualitatively, the $A_{434\text{imp}}$ measurements can be used as a purity verification step. The $A_{434\text{imp}}$ value using the $\frac{e_3}{e_2}$ of DeGrandpre et al. (2014) of the pure indicators was near zero, confirming their purity at least at 434 nm. Thus, regular measurements of $A_{434\text{imp}}$ on a given batch of indicator over time could potentially be used as a quality assurance measure where a change to positive $A_{434\text{imp}}$ values indicates the solution is no longer stable. The threshold at which $A_{434\text{imp}}$ indicates an unstable batch would need to be determined.

Quantitively, interpreting $A_{434\text{imp}}$ becomes challenging. Applying the correction to the impure indicators here produced values in worse agreement with the reference indicator than no correction at all. The youngest impure indicator used here was 7.2 years making it difficult to interpret the results and make conclusions applicable to other studies because measurements were not made over time. It could be that the solutions were unstable and $A_{434\text{imp}}$ is not able to account for the changes. It could also be that there are unaccounted for impurities at 578 nm, or that the signal to noise ratio of $A_{434\text{imp}}$ using the Agilent 8454 spectrophotometer is too small.

The choice of $\frac{e_3}{e_2}$ is also uncertain for seawater samples. Given the known media effect (Liu et al. 2011), the values determined in artificial seawater should theoretically be better than values determined in NaCl. However, the negative value for pure indicators suggests a bias in those values. The difference in pH values between the two was ~0.001–0.003, which introduces a small but significant bias. Further work is warranted to determine the best practices for correcting for impurities. It is therefore suggested that any study using $A_{434\text{imp}}$ to adjust measurements made with impure indicators verify the adjustment with paired pure–impure measurements in the same media as the samples and provide raw absorbance data, so results can be

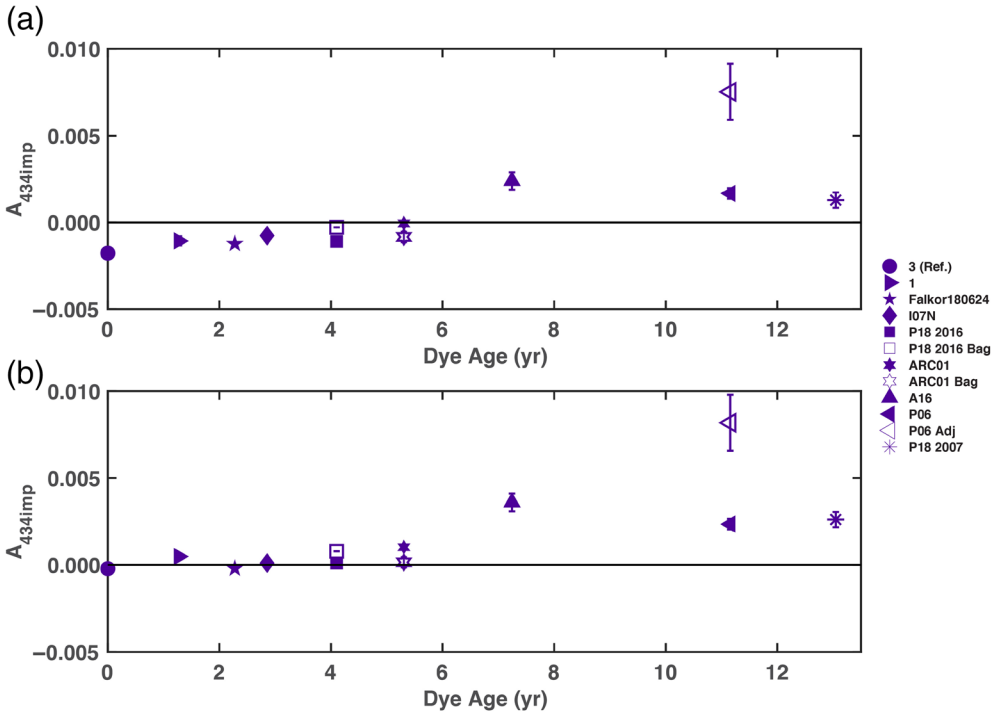

Fig. 1. The mean and standard deviation of A_{434imp} determined for each batch of indicator solution according to methods of Douglas and Byrne (2017). (**a**) The values using molar absorbtitivies of Liu et al. (2011), which were determined in an artificial seawater media, and (**b**) the values using molar absorbtitivies of DeGrandpre et al. (2014), which were determined in NaCl.

updated in the future if needed. Further evaluation of A_{434imp} is not possible with the experiments presented here.

mCp stability analysis

The ΔpH_t for the CRMs and TRIS of the pure indicator solutions is shown in Fig. 2. The pH_t and standard error values are given in Table 2. The indicator was stable (within 0.0025) for at least 5.3 years for both CRMs and TRIS buffers. The short-term precision (standard error) of these measurements is often smaller than 0.0025, but the long-term precision (including bottle-to-bottle variability) for CRMs based on cruise measurements is typically ~ 0.003 (Millero et al. 2014, 2016), which is taken as the overall precision of the measurements here.

For the pure indicators, the A_{434imp} adjustment can qualitatively act as further verification of the stability of the indicator. The A_{434imp} adjustment using $\frac{e_3}{e_2}$ of DeGrandpre et al. (2014) was essentially zero ($\text{mean} \pm \text{SD} = -0.0009 \pm 0.0005$). That A_{434imp} is stable over all of the pure indicators suggests it could be used to qualitatively verify the purity of the indicator by making measurements of A_{434imp} over time. No correction was applied to the pure indicators, the A_{434imp} measurements were only used as a qualitative assessment of the solutions.

For the impure indicators, the uncertainties in accounting for impurities were too large to assess the stability. Nevertheless, some important observations resulted from the measurements. The P06 indicator batch showed obvious signs of degradation. The color, which is typically a dark purple, had become a pale orange. The pH of the indicator solution had also dropped significantly (Table 1). In addition, the amount of indicator required to obtain a reliable measurement (peak absorbance 0.3–1.2) was four times that of other indicators. An aliquot of the solution was taken, and the pH was adjusted with NaOH to 7.8 (P06_adj). Measurements of CRMs and TRIS had worse agreement with the reference indicator and the A_{434imp} was significantly larger. The P06 indicator was stored in a polyethylene (PE) plastic bottle and sealed with parafilm. These qualitative results may indicate that because plastic containers have some oxygen permeability, the indicator oxidizes slowly over time, reducing the effective concentration. It is also possible that the parafilm wrapped around the cap did not create a perfect seal allowing oxygen to diffuse into the bottle over time. If true, the process is very slow as other (younger) indicator solutions were stored in plastic containers and did not show signs of degradation (Table 1). This result also suggests that indicator volume changes could be used as a quality assurance measure to assess the stability of individual solution batches.

The influence of storage container on stability can be further evaluated because two batches of indicator were aliquoted and stored in two separate types of containers, P18_2016 and ARC01. For P18_2016, the indicator was stored in a tedlar gas sampling bag or a PE bottle. The difference in CRM pH_t

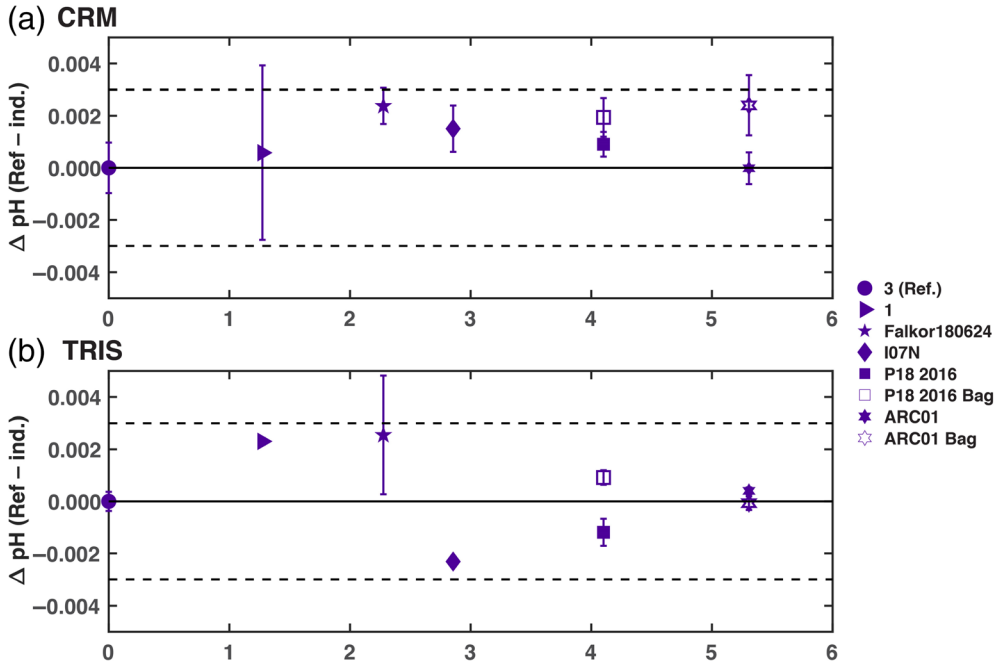

Fig. 2. The difference in pH between the reference indicator (3) and the individual pure indicator batches for (**a**) CRM batch 189 and (**b**) TRIS buffer. Error bars represent standard error of the measurements. If error bars are not visible, they are smaller than the symbol size. Dashed lines represent overall estimated precision in CRM and TRIS values (0.003) based on typical long line cruise measurements (Millero et al. 2014, 2016). Impure indicators are not shown as there was too much uncertainty in applying a correction to pure indicator.

between the two containers was 0.0010, and for TRIS was 0.0021. For ARC01, the indicator was stored in a tedlar gas sampling bag or in a high-density polyethylene (HDPE) brown plastic bottle. The differences between the two containers for CRMs and TRIS were 0.0001 and 0.0004, respectively. The differences between storage containers of the same batch are within the overall uncertainty of the measurements and are indistinguishable, but the values for P18_2016 are larger than the standard errors (Table 2). The $A_{434\text{imp}}$ values for these batches are also indistinguishable suggesting none of them have degraded significantly. From these measurements, the storage container does not appear to be a primary influence on indicator stability over at least 5.3 years but may become important over longer time periods. Glass, HDPE, and tedlar bags appear to be stable over at least 5.3 years. Additional experiments would be needed to confirm that the PE plastic caused the instabilities in P06, but such containers should probably be avoided for long-term storage.

The pH of all indicator solutions did decrease over time (Table 1). This may be a result of chemical modification of the indicator itself or due to continued dissolution of CO_2 during storage. CO_2 uptake could occur slowly through the walls of the bottle when stored in plastic, or when air is exchanged in the headspace as the container is opened periodically with use. Aside from P06, the decreases are small and not likely to have an impact on the stability of the indicator but could affect the indicator perturbation on a seawater sample (Chierici et al. 1999). Thus, the indicator perturbation should be determined regularly throughout the life of the indicator. For Batch 1, the indicator correction was similar in March 2020 (Woosley 2021) and September 2020. Although not systematic, these results would suggest that the indicator perturbation is stable over several months.

These experiments only cover a sample pH range of ~ 7.8–8.1. It is possible that impurities created by the time-dependent breakdown of the indicator may affect lower pH values differently than the higher values examined here. Unfortunately, a large volume of stable and characterized low pH seawater to test this is not available; however, the impurities found by Liu et al. (2011) had a larger effect at higher pH than lower pH values. Recent work (Takeshita et al. 2020) has also shown spectrophotometric pH-dependent errors using pure indicators are not indicator related. It seems likely that the results presented here would be similar for lower pH waters, though it cannot be verified with these experiments. The pH range of the indicator solutions used (Table 1) was also very small, and it is unknown if indicator solutions prepared at higher pH values will exhibit the same stability.

Conclusions

The quality of mCp indicator dye solutions commonly used for seawater pH measurements is paramount for high-quality pH measurements required to monitor and evaluate ocean acidification over decadal timescales. The indicator requires

purification of the solid obtained from chemical manufacturers, which is expensive, and only a few laboratories have the required equipment and skill to perform the purification. Such limitations compound the undesirability of waste, and as such, knowing the stability and proper long-term storage conditions of the indicator solutions will enable the efficient characterization of seawater pH by many groups at minimal costs in time and money. The experiments here show that the indicator solution is stable (± 0.0025) for at least 5.3 years when stored in a well-sealed opaque container at room temperature. For storage beyond 5 years PE plastic may cause solution degradation, but more experiments are needed for confirmation. The pH of the indicator solution does decrease over time which would impact the indicator perturbation correction. The perturbation correction should therefore be determined regularly. Given the importance of the indicator to pH measurements, full details of the indicator solution and its preparation and storage should always be clearly described.

The $A_{434\text{imp}}$ method for determining indicator impurities can be used as a quality assurance measure to qualitatively verify indicator stability. Regular measurements of $A_{434\text{imp}}$ on a batch of indicator could potentially be used to detect changes in the indicator over time. When using the method to make corrections to measurements that used impure indicators the choice of value used for $\frac{e_3}{e_2}$ makes a small but non-negligible difference, thought to be the result of a media effects. Further research is warranted. It is suggested that studies using the $A_{434\text{imp}}$ method to adjust sample pH validate the adjustment using paired pure-impure measurements in media similar to the samples.

The results presented here indicate that mCp is generally very stable when protected from UV light and oxidation, and that larger batches can be prepared and used over several cruises or experiments. Preparation of small batches as needed is not required. The small batches require significant time to prepare. Therefore, knowledge of mCp stability and storage determined here can be used to save time and prevent waste arising from the disposal of indicator solutions that are still useable.

References

Andersson, A. J., and others. 2015. Understanding ocean acidification impacts on organismal to ecological scales. Oceanography **28**: 16–27. doi:10.5670/oceanog.2015.27

Bednaršek, N., and others. 2019. Systematic review and meta-analysis toward synthesis of thresholds of ocean acidification impacts on calcifying pteropods and interactions with warming. Front. Mar. Sci. **6**: 227. doi:10.3389/fmars.2019.00227

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

Carter, B. R., and others. 2017. Two decades of Pacific anthropogenic carbon storage and ocean acidification along GO-SHIP sections P16 and P02. Global Biogeochem. Cycl. **31**: 1–22. doi:10.1002/2016GB005485

Chierici, M., A. Fransson, and L. G. Anderson. 1999. Influence of m-cresol purple indicator additions on the pH of seawater samples: Correction factors evaluated from a chemical speciation model. Mar. Chem. **65**: 281–290. doi:10.1016/S0304-4203(99)00020-1

Clayton, T. D., and R. H. Byrne. 1993. Spectrophotometric seawater pH measurements: Total hydrogen ion concentration scale calibration of m-cresol purple and at-sea results. Deep. Res. I **40**: 2115–2129. doi:10.1016/0967-0637(93)90048-8

DeGrandpre, M. D., R. S. Spaulding, J. O. Newton, E. J. Jaqueth, S. E. Hamblock, A. A. Umansky, and K. E. Harris. 2014. Considerations for the measurement of spectrophotometric pH for ocean acidification and other studies. Limnol. Oceanogr. Methods **12**: 830–839. doi:10.4319/lom.2014.12.830

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

Douglas, N. K., and R. H. Byrne. 2017. Achieving accurate spectrophotometric pH measurements using unpurified meta-cresol purple. Mar. Chem. **190**: 66–72. doi:10.1016/J.MARCHEM.2017.02.004

Feely, R. A., C. L. Sabine, K. Lee, W. Berelson, J. Kleypas, V. J. Fabry, and F. J. Millero. 2004. Impact of anthropogenic CO_2 on the $CaCO_3$ system in the oceans. Science **305**: 362–366. doi:10.1126/science.1097329

French, C. R., J. J. Carr, E. M. Dougherty, L. A. Eidson, J. C. Reynolds, and M. D. DeGrandpre. 2002. Spectrophotometric pH measurements of freshwater. Anal. Chim. Acta **453**: 13–20. doi:10.1016/S0003-2670(01)01509-4

Kroeker, K. J., R. L. Kordas, R. N. Crim, and G. G. Singh. 2010. Meta-analysis reveals negative yet variable effects of ocean acidification on marine organisms. Ecol. Lett. **13**: 1419–1434. doi:10.1111/j.1461-0248.2010.01518.x

Liu, X., M. C. Patsavas, and R. H. Byrne. 2011. Purification and characterization of meta-cresol purple for spectrophotometric seawater pH measurements. Environ. Sci. Technol. **45**: 4862–4868. doi:10.1021/es200665d

Millero, F. J., R. Woosley, B. DiTrolio, and J. Waters. 2009. Effect of ocean acidification on the speciation of metals in seawater. Oceanography **22**: 72–85. doi:10.5670/oceanog.2009.98

Millero, F. J., and B. DiTrolio. 2010. Use of thermodynamics in examining the effects of ocean acidification. Elements **6**: 293–298. doi:10.2113/gselements.6.5.293

Millero, F. J., J. Sharp, R. J. Woosley, C. Rodriguez, J. Levy, J. Williamson, J. Burns, and K. Mastropole. 2014. Global ocean repeat hydrography study: pH and total alkalinity measurements in the Atlantic Ocean. A16 North and South Aug 2013–Feb 2014. CDIAC Oak Ridge National Laboratory. https://www.ncei.noaa.gov/data/oceans/ncei/ocads/data/0157363/A16N_A16S_2013_2014_UM_ALKALI_PH_Technical_Report.pdf

Millero, F. J., R. J. Woosley, A. R. Margolin, and F. Huang. 2016. Global ocean repeat hydrography study: pH, total alkalinity, and total CO_2 measurements in the Arctic Ocean. CDIAC, Oak Ridge National Laboratory.

Newton, J. A., R. A. Feely, E. B. Jewett, P. Williamson, and J. Mathis. 2015. Global ocean acidification observing network: requirements and governance plan second edition GOA-ON global ocean acidification observing network. GOA-ON.

Patsavas, M. C., R. H. Byrne, and X. Liu. 2013. Purification of meta-cresol purple and cresol red by flash chromatography: Procedures for ensuring accurate spectrophotometric seawater pH measurements. Mar. Chem. **150**: 19–24. doi:10.1016/J.MARCHEM.2013.01.004

Paulsen, M., and A. G. Dickson. 2020. Preparation of 2-amino-2-hydroxymethyl-1,3-propanediol (TRIS) pH_T buffers in synthetic seawater. Limnol. Oceanogr. Methods **18**: 504–515. doi:10.1002/lom3.10383

Pinto, V. C., C. F. Araújo, P. J. Sousa, L. M. Gonçalves, and G. Minas. 2019. A low-cost lab-on-a-chip device for marine pH quantification by colorimetry. Sens. Actuators B **290**: 285–292. doi:10.1016/J.SNB.2019.03.098

Reggiani, E. R., A. L. King, M. Norli, P. Jaccard, K. Sørensen, and R. G. J. Bellerby. 2016. FerryBox-assisted monitoring of mixed layer pH in the Norwegian Coastal Current. J. Mar. Syst. **162**: 29–36. doi:10.1016/J.JMARSYS.2016.03.017

Rérolle, V. M. C., C. F. A. Floquet, A. J. K. Harris, M. C. Mowlem, R. R. G. J. Bellerby, and E. P. Achterberg. 2013. Development of a colorimetric microfluidic pH sensor for autonomous seawater measurements. Anal. Chim. Acta **786**: 124–131. doi:10.1016/J.ACA.2013.05.008

Sabine, C. L., and others. 2004. The oceanic sink for anthropogenic CO_2. Science **305**: 367–371. doi:10.1126/science.1097403

Seidel, M. P., M. D. DeGrandpre, and A. G. Dickson. 2008. A sensor for in situ indicator-based measurements of seawater pH. Mar. Chem. **109**: 18–28. doi:10.1016/J.MARCHEM.2007.11.013

Takeshita, Y., K. S. Johnson, L. J. Coletti, H. W. Jannasch, P. M. Walz, and J. K. Warren. 2020. Assessment of pH dependent errors in spectrophotometric pH measurements of seawater. Mar. Chem. **223**: 103801. doi:10.1016/J.MARCHEM.2020.103801

Waters, J. F., F. J. Millero, and C. L. Sabine. 2011. Changes in South Pacific anthropogenic carbon. Global Biogeochem. Cycl. **25**: GB4011. doi:10.1029/2010GB003988

Woosley, R. J. 2021. Evaluation of the temperature dependence of dissociation constants for the marine carbon system using pH and certified reference materials. Mar. Chem. **229**: 103914. doi:10.1016/j.marchem.2020.103914

Woosley, R. J., F. J. Millero, and M. Grosell. 2012. The solubility of fish-produced high magnesium calcite in seawater. J. Geophys. Res. Ocean. **117**: 1–5. doi:10.1029/2011JC007599

Woosley, R. J., F. J. Millero, and R. Wanninkhof. 2016. Rapid anthropogenic changes in CO_2 and pH in the Atlantic Ocean: 2003–2014. Global Biogeochem. Cycl. **30**: 70–90. doi:10.1002/2015GB005248

Woosley, R. J., and F. J. Millero. 2020. Freshening of the western Arctic negates anthropogenic carbon uptake potential. Limnol. Oceanogr. **65**: 1834–1846. doi:10.1002/lno.11421

Yao, W., X. Liu, and R. H. Byrne. 2007. Impurities in indicators used for spectrophotometric seawater pH measurements: Assessment and remedies. Mar. Chem. **107**: 167–172. doi:10.1016/J.MARCHEM.2007.06.012

Zhang, H., and R. H. Byrne. 1996. Spectrophotometric pH measurements of surface seawater at in-situ conditions: Absorbance and protonation behavior of thymol blue. Mar. Chem. **52**: 17–25. doi:10.1016/0304-4203(95)00076-3

Acknowledgments

The author would like to thank Laney Flanagan for preliminary tests in preparation of the experiments presented here and Frank J. Millero for providing the opportunity to collect most of these different indicator batches. This work was supported by the National Science Foundation Division of Ocean Sciences (Award No. 1923312) and the MIT mTerra Catalyst fund.

Conflict of Interest

None declared.

Submitted 08 March 2021
Revised 16 July 2021
Accepted 27 September 2021

Associate editor: Gregory Cutter