UNIVERSITY OF CALIFORNIA SAN DIEGO

Uncertainty of spectrophotometric pH measurements in seawater and implications for ocean carbon chemistry

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy

in

Oceanography

by

Michael Brian Fong

Committee in charge:

Professor Andrew G. Dickson, Chair
Professor Andreas J. Andersson
Professor Clifford P. Kubiak
Professor Todd R. Martz
Professor Lynne Talley

2021

The dissertation of Michael Brian Fong is approved, and it is acceptable in quality and form for publication on microfilm and electronically.

University of California San Diego

2021

DEDICATION

This dissertation is dedicated to my parents, my elders, and my ancestors for their sacrifices that made possible the opportunities I have today and to my high school science teacher Ms. Katherine Melvin, who sparked my interest in oceanography.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

ALS	–	Alternating Least Squares
CO2SYS	–	Computer program for CO_2 system calculations
CRAM	–	Carboxylic-rich aliphatic matter
CRM	–	Certified Reference Material
DOC	–	Dissolved Organic Carbon
EPPS	–	3-[4-(2-Hydroxyethyl)-1-piperazinyl]propanesulfonic acid
FB4	–	ID for a lot of purified m-cresol purple
GLODAP	–	Global Ocean Data Analysis Project
GOA-ON	–	Global Ocean Acidification Observing Network
GO-BGC	–	Global Ocean Biogeochemistry Array
GO-SHIP	–	Global Ocean Ship-Based Hydrographic Investigations Program
GUI	–	Graphical User Interface
HPLC	–	High performance liquid chromatography
I08S	–	The name of a section in the southern Indian Ocean and Southern Ocean
I09N	–	The name of a section in the northern Indian Ocean
ISFET	–	Ion-sensitive Field Effect Transistor
mCP	–	*m*-cresol purple
MCR	–	Multivariate Curve Resolution
NIST	–	National Institute of Standards and Technology
NMR	–	Nuclear magnetic resonance spectroscopy
NOAA-AOML	–	National Oceanic and Atmospheric Administration Atlantic Oceanographic and Meteorological Laboratory
NOAA-PMEL	–	National Oceanic and Atmospheric Administration Pacific Marine Environmental Laboratory
P16N	–	The name of a section in the northern Pacific Ocean
P16S	–	The name of a section in the southern Pacific Ocean
PCA	–	Principal Components Analysis
SAMI-pH	–	Submersible Autonomous Moored Instrument for pH
SIMPLISMA	–	Simple-to-Use Interactive Self-Modeling Mixture Analysis
SOCCOM	–	Southern Ocean Carbon and Climate Observations and Modeling Project
SPE	–	Solid-phase extraction for DOC
SRM 930D	–	NIST Standard Reference Material for transmittance, consisting of a set of neutral density glass filters
TRIS	–	2-Amino-2-(hydroxymethyl)propane-1,3-diol

LIST OF SYMBOLS

General

A_{434}	–	Absorbance at 434 nm
A_{578}	–	Absorbance at 578 nm
A_{iso}	–	Absorbance at the isosbestic wavelength (488.1 nm at 25°C for *m*-cresol purple)
A_{λ}	–	Absorbance at the wavelength λ
A_T	–	Total alkalinity
B_T/S	–	Total boron-salinity ratio
$CaCl_2$	–	Calcium chloride
CO_2	–	Carbon dioxide
CO_3^{2-}	–	Carbonate ion
C_T	–	Dissolved inorganic carbon
ε_{λ}	–	Molar absorption coefficient at the wavelength λ for *m*-cresol purple
e_1, e_2, e_3	–	Absorption coefficient ratios for *m*-cresol purple
H^+	–	Hydrogen ion
H_2CO_3	–	Carbonic acid
HCl	–	Hydrochloric acid
HCO_3^-	–	Bicarbonate ion
HI^-	–	Singly protonated form of *m*-cresol purple
I^{2-}	–	Fully deprotonated form of *m*-cresol purple
K_1, K_2	–	First and second dissociation constant of carbonic acid
$K_a(H_2I^-)$	–	First acid dissociation constant of *m*-cresol purple
$K_a(HI^-)$	–	Second acid dissociation constant of *m*-cresol purple
ℓ	–	Cell pathlength
$MgCl_2$	–	Magnesium chloride
Na_2HPO_4	–	Sodium hydrogen phosphate
Na_2SO_4	–	Sodium sulfate
NaCl	–	Sodium chloride
NaH_2PO_4	–	Sodium dihydrogen phosphate
NaOH	–	Sodium hydroxide
$p(CO_2)$	–	Partial pressure of CO_2
$p(K_a(HI^-)\cdot e_2)$	–	Negative logarithm of the product of $K_a(HI^-)$ and e_2
R	–	Ratio of the absorbance at 578 nm to the absorbance at 434 nm
S	–	Salinity

Chapter 2

$CaCO_3$	–	Calcium carbonate
pH_{spec}	–	Spectrophotometrically measured pH
$pH_{calc}(C_T, A_T,\ldots)$	–	pH calculated from dissolved inorganic carbon and total alkalinity, etc.
ΔpH	–	Difference between spectrophotometrically measured pH and calculated pH (i.e., $pH_{spec} - pH_{calc}(C_T, A_T,\ldots)$)
A_X'	–	Apparent excess alkalinity
K_B	–	Boric acid dissociation constant
$B(OH)_4^-$	–	Borate ion
OH^-	–	Hydroxide ion
PO_4^{3-}	–	Phosphate ion
$SiO(OH)_3^-$	–	Silicate ion
NH_3	–	Ammonia
HS^-	–	Bisulfide ion
HSO_4^-	–	Bisulfate ion
HF	–	Hydrogen fluoride
HPO_4^{2-}	–	Hydrogen phosphate
H_3PO_4	–	Phosphoric acid
A_{Tmeas}	–	Measured total alkalinity
$A_{Tcalc}(C_T,pH,\ldots)$	–	Total alkalinity calculated from dissolved inorganic carbon and pH

Chapter 3

$\mathbf{A}$	–	Matrix of spectra
$\mathbf{A}^\dagger$	–	Mean-centered matrix of spectra
$\mathbf{A}^*$	–	Matrix containing the normalized spectra of *m*-cresol purple measured in HCl, acetate, and NaOH solutions
$\mathbf{a_{proj}}$	–	Vector of projected spectrum
$\mathbf{C}$	–	Matrix containing the relative concentrations of the three species of *m*-cresol purple in the HCl, acetate, and NaOH solutions
$\mathbf{c}_{HI^-}$, $\mathbf{c}_{I^{2-}}$	–	Vectors of relative contributions of the HI^- and I^{2-} absorption to the sample spectra
D	–	Deuterium
E^0	–	Standard potential for an electrochemical cell
W	–	Tungsten
$\mathbf{S}_{HI^-}$, $\mathbf{S}_{I^{2-}}$	–	Absorption spectra for the HI^- and I^{2-} forms of *m*-cresol purple
$\mathbf{t}_i$	–	Vector of scores for the *i*th principal component
$\mathbf{v}_i^T$	–	Vector of the loading spectrum for the *i*th principal component

Chapter 4

$A_{T,dye}$	–	Total alkalinity of the dye stock solution
$A_{T,samp}$	–	Total alkalinity of the sample
$A_{T,samp+dye}$	–	Total alkalinity of the sample with dye added
$C_{T,dye}$	–	Dissolved inorganic carbon concentration in the dye stock solution
$C_{T,samp}$	–	Dissolved inorganic carbon concentration in the sample
$C_{T,samp+dye}$	–	Dissolved inorganic carbon concentration in the sample with dye added
f_{dye}, f_{samp}	–	Fraction of dye solution and sample, respectively, in the sample-dye mixture
pH_{dye}	–	pH of the dye stock solution
pH_{ref}	–	Reference pH value used in simulations
pH_{sim}	–	Simulated pH value
$\Delta pH/\Delta[mCP]_T$	–	Perturbation to the sample pH per unit change in total dye concentration
R'	–	Sample absorbance ratio after the first dye addition
$\Delta R/\Delta[mCP]_T$	–	Pertubation to the absorbance ratio R per unit change in total dye concentration
$\Delta R/\Delta V$	–	Perturbation to the absorbance ratio R per unit volume of dye addition
s_{meas}	–	Standard error of the slope of pH versus total dye concentration
t	–	Student's t value
u_c	–	Combined standard uncertainty
U_{model}	–	Expanded uncertainty (95% confidence level) of the modeled slope of pH versus total dye concentration
$U_{model–meas}$	–	Expanded uncertainty (95% confidence level) of the difference between the modeled and measured slopes of pH versus total dye concentration
u_{rand}	–	Random uncertainty, expressed as a standard deviation
u_{sys}	–	Systematic uncertainty, expressed as a standard deviation

Chapter 5

${}_{434}A_{imp}$	–	Impurity absorption at 434 nm
${}_{434}A_{NaOH}$	–	Absorbance at 434 nm measured in a NaOH solution
${}_{434}A_{samp}$	–	Absorbance at 434 nm measured in a sample solution
$\mathbf{c_{Imp}}$	–	Vector of relative contribution of the impurity absorption to the sample spectra
$\mathbf{E}$	–	Matrix of residuals
pH_{imp}	–	pH measured with an impure dye
pH_{pure}	–	pH measured with a purified dye
R_{NaOH}	–	Absorbance ratio measured in a NaOH solution

R_{samp}	–	Absorbance ratio of a sample solution, measured with an impure dye
R_{pure}	–	Absorbance ratio of a sample solution, measured with an impure dye and adjusted for the impurity absorption at 434 nm
$\mathbf{S_{Imp}}$	–	Impurity absorption spectrum for the Acros *m*-cresol purple

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

My advisor Andrew Dickson was critical to this research and to my development as a scientist. I am thankful for his guidance in shaping these research projects, the valuable insights he provided, and his support over the years. I also thank my committee for their support and guidance.

The work in **Chapter 3** and **Chapter 5** could not be possible without the help of many people. Charles Branham, who I first met at a Gordon Research Conference, introduced me to the field of chemometrics and provided a unique perspective to these chapters. Foraying into this unfamiliar field was an exciting learning experience for me. Many chemometric techniques, I learned, are applicable to problems that the spectrophotometric pH community has been struggling with, and I hope to apply more chemometrics to my future research.

Kate Boyle, Jennifer Cragan, and Huge Doyle were my predecessors in the Dickson Lab who I never met, but I stood on their shoulders. Kate Boyle's meticulous records in her laboratory notebook provided both insights and validation that the struggles I experienced were not uniquely my own. Trisha Nguyen was an undergraduate summer intern in the Dickson lab who performed preliminary experiments on the characterization of *m*-cresol purple. Manuel Belmonte performed the phosphate buffer measurements in **Chapter 3** and helped to uncover a problem in spectrophotometric pH measurements that we did not know we had.

I thank Emily Bockmon for giving me advice in my early years as a graduate student, for encouraging me to write **Chapter 2**, and for producing the modified batches of seawater reference materials which I used in **Chapter 4** and helping me run dissolved inorganic carbon measurements on *m*-cresol purple solutions.

The work in **Chapter 2** involved personally spending 102 days at sea, in both fair and rough conditions. This would not have been possible without the support of the crew and scientists on the Ronald H. Brown and the Roger Revelle. I did not participate in the 2014 P16S cruise, but I thank the crew and scientists on the Nathaniel B. Palmer for their hard work. David Cervantes deserves special thanks for taking extra efforts on the cruises to ensure that operations ran smoothly and that our experience was as comfortable as possible. August Pereira, Britain Richardson, and Heather Page helped me collect water samples and kept me from becoming overwhelmed. Through this work, I became involved with the Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) community and the Ocean Carbonate System Intercomparison Forum (OCSIF). I am honored to have the opportunity to contribute to these efforts, and I am excited about the direction in which our field is heading.

During many of the nights that I toiled in the lab far too late, UC San Diego's Triton Rides drove me home when I finally called it a day. I also want to acknowledge UC San Diego for the incredible job it has done in keeping the campus community safe during the COVID-19 pandemic. Admittedly, the research in this dissertation would have benefited from some additional measurements after learning why things went wrong, but this was made challenging by the conditions surrounding the pandemic.

Finally, my time at SIO would certainly have been quite miserable if it were not for the community I had—members of the Dickson Lab, my office and lab-mate, May-Linn Paulsen, the Marine Chemistry and Geochemistry curricular group, and my friends, Jessica Ng, Daniel Yee, Garfield Kwan, Cecilia Chow, Renny Ng, Emily Wei, Jasmeet Dhaliwal, Ned Richards, Matthew Pendergraft, Rob Lampe, Brendan Stephens, Julia Dohner, Maitreyi Nagarkar, Matthew Cook, Eric Lybrand, Travis Courtney, and Brian Chow.

Chapter 2, in full, is a reprint of the material as it appears in Marine Chemistry. Fong, M.B. and Dickson, A.G., 2019. Insights from GO-SHIP hydrography data into the thermodynamic consistency of CO_2 system measurements in seawater. Marine Chemistry, 211: 52-63. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in part, is currently being prepared for submission for publication of the material. Fong, M.B., Branham, C.W., and Dickson, A.G. The dissertation author was the primary investigator and author of this paper.

Chapter 4, in part, is currently being prepared for submission for publication of the material. Fong, M.B. and Dickson, A.G. The dissertation author was the primary investigator and author of this paper.

Chapter 5, in part, is currently being prepared for submission for publication of the material. Fong, M.B., Branham, C.W., and Dickson, A.G. The dissertation author was the primary investigator and author of this paper.

VITA

2012 Bachelor of Arts, University of California Berkeley (Marine Science)

2021 Doctor of Philosophy, Scripps Institution of Oceanography, University of California San Diego (Oceanography)

PUBLICATIONS

Fong, M.B. and Dickson, A.G., 2019. Insights from GO-SHIP hydrography data into the thermodynamic consistency of CO_2 system measurements in seawater. Marine Chemistry, 211: 52-63.

Carter, B.R., Feely, R.A., Williams, N.L., Dickson A.G., Fong, M.B., and Takeshita, Y., 2018. Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate. Limnology and Oceanography: Methods, 16(2): 119-131.

Bishop, J.K.B., Fong, M.B. and Wood, T.J., 2016. Robotic observations of high wintertime carbon export in California coastal waters. Biogeosciences, 13(10): 3109-3129.

Bauman, S.J., Costa, M.T., Fong, M.B., House, B.M., Perez, E.M., Tan, M.H., Thornton, A.E., and Franks, P.J.S., 2014. Augmenting the Biological Pump: The Shortcomings of Geoengineered Upwelling. Oceanography, 27(3): 17-23.

ABSTRACT OF THE DISSERTATION

Uncertainty of spectrophotometric pH measurements in seawater and implications for ocean carbon chemistry

by

Michael Brian Fong

Doctor of Philosophy in Oceanography

University of California San Diego, 2021

Professor Andrew G. Dickson, Chair

As pH measurement technologies evolve and are increasingly used to study the ocean's CO_2 system, there is a need to assess the uncertainty of seawater pH measurements, particularly those based on the spectrophotometric method using an indicator dye, which is used to calibrate alternative pH measurement methods, including autonomous pH sensors. In this dissertation, I investigated methodological and instrumental contributions to uncertainty in spectrophotometric pH measurements, developed tools for their quality control, and evaluated

their likely quality based on their consistency with other seawater CO_2 parameters. The accuracy of seawater pH measurements has been questioned due to observations from open ocean cruises of a significant pH-dependent discrepancy between measured pH and pH calculated from dissolved inorganic carbon (C_T) and total alkalinity (A_T), using a thermodynamic model of seawater acid-base systems. Based on an analysis of high quality CO_2 measurements on four open ocean cruises, I showed that a combination of plausible biases in the constants and assumptions used to calculate pH can explain the observed inconsistencies and that there is likely an unaccounted, possibly organic, contribution to the A_T measured in the open ocean. Next, I developed methods for analyzing spectra to identify and possibly correct instrumental contributions to spectrophotometric pH measurement error, such as those that may arise from degradation of the spectrophotometer lamp. In another chapter, laboratory and chemical modeling experiments were conducted to evaluate how adjustments for effect of the indicator addition on the sample pH contribute to uncertainty in the final pH result. Assumptions about the indicator's behavior can result in error in the pH correction, but is minor, except when using high dye concentrations in short pathlength cells. Finally, I evaluated the performance of pH measurements using impure *m*-cresol purple, which contains absorbing impurities that can bias pH measurements. Calibrated impure dye measurements can still be inconsistent with purified dye measurements depending on the assumptions about the impurity absorption behavior or the purity of the reference dye. This dissertation contributes to understanding of the uncertainty of spectrophotometric pH measurements and our ability to use ocean pH data to characterize the CO_2 system.

Chapter 1

Introduction

1.1 Motivation and scope of dissertation:

pH is an important property of aquatic systems, affecting a wide range of chemical and biological processes. Because pH is often straightforward to measure, there is a desire by marine chemists to use ocean pH data to study biogeochemical processes, particularly the carbon dioxide system. When carbon dioxide dissolves in water, it reacts with water to form a weak acid, which further dissociates and establishes equilibrium with three different forms of inorganic carbon: carbonic acid (H_2CO_3), bicarbonate ion (HCO_3^-) and carbonate ion (CO_3^{2-}). These chemical reactions also release hydrogen ions (H^+), lowering the pH. Because the rates of these reactions are fast, pH is a good indicator of the state of the CO_2 system at equilibrium, providing information about the relative proportions of the inorganic carbon species. However, pH by itself is not useful for characterizing the seawater CO_2 system unless another CO_2 parameter is also measured. Marine chemists commonly measure at least two of the following four parameters to fully describe the ocean's CO_2 system: pH, dissolved inorganic carbon (C_T), total alkalinity (A_T), and partial pressure of CO_2 ($p(CO_2)$). When two CO_2 parameters are measured, the other parameters can be calculated from a chemical equilibrium model, provided that the equilibrium constants for the various acid dissociation reactions in seawater and the total concentrations of all the acid-base systems present are also known (Park, 1969; Skirrow, 1975; Takahashi et al., 1970).

As the oceans continue to absorb the anthropogenic CO_2 emitted to the atmosphere from the burning of fossil fuels, there is an increasing need to understand the resulting long-term impacts on the carbon cycling of the oceans and on marine life. The oceans are the largest

sink for carbon over decadal to centennial timescales and have absorbed approximately 27% of total anthropogenic CO_2 emissions (Khatiwala et al., 2013; Sabine and Tanhua, 2010). This accumulation of CO_2 in the oceans has resulted in a decrease in seawater pH, a process called ocean acidification. Because changes in pH reflect changes in the equilibrium state of the CO_2 system, it has implications for a variety of processes that are dependent on the inorganic carbon composition of the solution. For instance, the continued ability of the oceans to mitigate climate change by absorbing atmospheric CO_2 depends on the CO_2 buffering capacity of seawater, which will decline as the pH of the oceans decreases (Fassbender et al., 2017; Riebesell et al., 2009). Calcifying marine organisms may also be adversely impacted by lower seawater pH, which decreases the carbonate ion concentration and increases the thermodynamic favorability of the dissolution of calcium carbonate (Caldeira and Wickett, 2003; Doney et al., 2009; Fabry et al., 2008; Orr et al., 2005).

While pH is relevant to characterizing the inorganic carbon composition of seawater, it provides only partial information. It follows, therefore, that understanding the overall uncertainty of CO_2 measurements and in the thermodynamic models used for CO_2 calculations is critical to the accurate interpretation of the CO_2 system. Of the four commonly measured CO_2 parameters, pH is the least well-constrained in terms of its uncertainty, yet is increasingly used to quantify anthropogenically-driven long-term changes in the ocean's CO_2 system (Byrne et al., 2010). In this dissertation, I investigated methodological and instrumental contributions to uncertainty in spectrophotometric pH measurements and evaluated the implications for using pH to make inferences about the CO_2 system. This research will advance the CO_2 chemistry community by improving understanding of the uncertainty of a widely used method for

measuring seawater pH and highlighting remaining gaps in our knowledge of pH measurement uncertainty and of the thermodynamics of the CO_2 and other acid-base systems in seawater.

1.2 Advances in pH measurement technology

Historically, marine chemists were interested in using measurements of seawater pH to infer the global distribution of CO_2 in the ocean but were limited by the poor accuracy and reproducibility of the pH measurements (Keeling, 1968). The development of the spectrophotometric pH method for use in seawater in the 1980s (Byrne and Breland, 1989; Clayton and Byrne, 1993; Zhang and Byrne, 1996) was a significant technological advance that made possible routine seawater pH measurements with remarkable short-term repeatability (~0.0004; Clayton and Byrne, 1993), thus improving the ability of marine chemists to use pH to meaningfully interpret the ocean CO_2 system (Dickson, 1993).

As a colorimetric method using a pH-indicating dye, spectrophotometric pH determination provides relatively stable measurements without the need for frequent calibration by the user. These advantages have undoubtedly led to the popularity of spectrophotometric pH and its implementation on a wide variety of systems, including semi-automated benchtop systems for discrete sampling (Carter et al., 2013), continuous flow and underway systems (Bellerby et al., 2002; Dickson, 1998; Tapp et al., 2000), and autonomous sensors (Seidel et al., 2008). Spectrophotometric pH is also considered a benchmark method for pH determination in natural waters (Dickson, 2010; Dickson et al., 2007) and has often been used to calibrate alternative pH measurement methods, such as glass electrodes (Easley and Byrne, 2012), pH sensors based on Ion Sensitive Field Effect Transistors (Bresnahan Jr. et al., 2014; Johnson et al., 2016; Martz et al., 2010), and custom-designed "do-it-yourself" instruments (Yang et al., 2014).

The excellent repeatability of spectrophotometric pH measurements is advantageous for reducing the overall imprecision when calculating other CO_2 parameters (Clayton et al., 1995; McElligott et al., 1998; Patsavas et al., 2015) and offers hope that the measurements are sufficiently sensitive to study changes in surface ocean pH, which is decreasing at a rate of ~0.002 units per year due to rising atmospheric $p(CO_2)$ (Pörtner et al., 2014). However, reliably quantifying long-term changes in pH requires consideration of the reproducibility of the measurements over longer timescales and between laboratories, which may use different instruments and operating procedures. Systematic errors may be less important when estimating differences in a value but will necessarily affect the accuracy in determining the absolute pH values and when using pH to estimate other CO_2 parameters. The quality of spectrophotometric pH measurements depends, therefore, on the overall uncertainty, which includes both systematic and random contributions (over different timescales and across different measurement conditions) and will be larger than the repeatability (i.e., the short-term precision of measurements obtained under the same set of conditions).

The standard uncertainty (i.e., uncertainty expressed as a standard deviation) of spectrophotometric pH measurements has been estimated to be ~0.005 to 0.01 (Carter et al., 2013), but some contributions, such as from the uncertainty in the indicator properties and instrumental contributions, are not well-understood. Additionally, an inter-laboratory study of seawater CO_2 measurements found that the comparability of spectrophotometric pH measurements was within 0.02 units across 27 laboratories and within 0.004 across 7 laboratories that used indicators which were carefully purified to remove impurities which may interfere with the pH measurement (Bockmon and Dickson, 2015). Thus, the available evidence suggests that the uncertainty of spectrophotometric pH measurements still fall short of

the goal of a standard uncertainty of 0.003 proposed by the Global Ocean Acidification Observing Network (GOA-ON) for the purposes of detecting long-term climate-related changes in the CO_2 system (Newton et al., 2014).

1.3 pH measurements and ocean carbon observational strategies

Large scale CO_2 observational efforts, such as time series stations (Bates et al., 2014), ship-based repeat hydrography surveys (Talley et al., 2016), collection of global sea surface CO_2 partial pressure observations (Bakker et al., 2016), and international collaborative networks for ocean acidification data exchange and research (e.g., GOA-ON), are essential for documenting long-term changes in ocean carbon chemistry.

pH measurements have played a major role in current CO_2 observational efforts and are expected to become increasingly important to emerging observational strategies. As a measure of acidity, pH provides direct quantification of ocean acidification (Byrne et al., 2010) and is therefore among the essential parameters measured on the Global Ocean Ship-Based Hydrographic Investigations Program (GO-SHIP). The number of seawater pH measurements is also rapidly increasing with the development of new pH measurement technologies, promoted by programs such as the Wendy Schmidt Ocean Health XPRIZE Challenge for the development of affordable and accurate pH sensors. Currently the only instruments capable of measuring a CO_2 parameter over the full range of temperatures and pressures in the oceans, autonomous pH sensors provide high resolution observations of the CO_2 system in regions where data collection from ships is limited due to challenging conditions. These measurements are typically used along with estimates of A_T from locally interpolated regression algorithms to calculate other CO_2 parameters (Bushinsky et al., 2019b; Carter et al., 2018; Williams et al., 2017). Measurements from pressure-tolerant pH sensors on profiling floats (Johnson et al.,

2016) have been used to estimate air-sea CO_2 fluxes in the Southern Ocean, where ship-based $p(CO_2)$ observations are scarce in the winter months, a critical period when significant CO_2 outgassing occurs (Bushinsky et al., 2019a; Gray et al., 2018). Due to the success of the Southern Ocean Carbon and Climate Observations and Modeling (SOCCOM) project, a $53 million grant has recently been funded to implement the Global Ocean Biogeochemistry (GO-BGC) Array, which plans to eventually deploy a total of 1,000 profiling floats with various chemical and biological sensors (including pH sensors) around the globe.

1.4 Implications of bias in spectrophotometric pH measurements

Pressure-tolerant autonomous pH sensors (such as those based on Ion Sensitive Field Effect Transistors) are calibrated directly with spectrophotometric pH measurements and indirectly through drift adjustments that aim to match the sensor pH with the pH expected at depth in a particular location of the ocean, estimated from algorithms trained with shipboard spectrophotometric pH measurements of discrete seawater samples (Johnson et al., 2016; Williams et al., 2016). Thus, any significant biases in spectrophotometric pH measurements would seriously limit the utility of the sensors for observing the CO_2 system.

The accuracy of spectrophotometric pH measurements has recently come into question due to a widespread observation, from many open ocean cruises where state-of-the-art CO_2 measurements were made, of there being a significant discrepancy between spectrophotometric pH measurements and pH calculated from C_T and A_T, which varies as a function of pH (Álvarez et al., 2020; Carter et al., 2018; Carter et al., 2013; McElligott et al., 1998; Williams et al., 2017). These discrepancies can be larger than the decadal climate-related changes in pH and have been problematic for the creation of an internally-consistent CO_2 data product (Olsen et al., 2016), for the use of pH measurements from profiling floats to quantify air-sea CO_2 fluxes

(Bushinsky et al., 2019a; Gray et al., 2018), and for determining calcium carbonate saturation states (Naviaux et al., 2019). Historical spectrophotometric pH measurements are known to have significant biases due to colored impurities in the indicator dye, which affect the determination of pH (Yao et al., 2007). However, the pH discrepancies observed on cruises persisted even after the development and use of purified *m*-cresol purple (Liu et al., 2011; Patsavas et al., 2013). Furthermore, a recent study comparing spectrophotometric pH against ISFET sensor-measured pH (calibrated against spectrophotometric pH at a single pH) across the full seawater pH range showed good agreement (Takeshita et al., 2020), suggesting that pH-dependent errors in spectrophotometric pH are unlikely to contribute significantly to the observed discrepancies, but instead, the problem may lie in the calculated pH values, potentially from errors in the various equilibrium constants needed for the calculations, in the input measurements C_T and A_T, and in estimating the total concentrations of boron and other acid-base systems in seawater. However, despite the internally consistent results between spectrophotometric pH and ISFET sensor-measured pH, the true accuracy of spectrophotometric pH measurements (relative to the primary method of pH measurement) over the full seawater pH range is not well-known. Constraining the overall uncertainty of spectrophotometric pH measurements is therefore urgently needed to both enhance the community's confidence in the method and to better understand the internal consistency of seawater CO_2 measurements.

1.5 Uncertainty in spectrophotometric pH measurements

The sources of uncertainty in spectrophotometric pH measurements can be thought of as belonging to several major categories, as outlined in **Figure 1.1** and discussed in more detail in this section. The upper half of **Figure 1.1** shows contributions to the uncertainty in the

determination of the dye properties, which include the purity of the dye, the performance of the spectrophotometer used to characterize the dye, and the metrological traceability of the indicator calibration to primary pH standards and the internationally-recognized definition of pH. The uncertainty in the dye properties can lead to bias in spectrophotometric pH measurements, as the typical user and the community as a whole rely on a single set of published values for the dye properties. The bottom half of **Figure 1.1** pertains to uncertainty contributions that affect the variance of individual pH measurements. These include the uncertainty in measuring absorbance, which is related to the spectrophotometer performance, and adjustments made to the pH measurement results to correct for the effect of the indicator on the sample pH, which relies on assumptions about the behavior of the dye over different measurement conditions. Additionally, if an impure dye is used, the calibrations or adjustments needed to achieve consistency with purified dye measurements (dotted arrow in **Figure 1.1**) is yet another source of uncertainty in spectrophotometric pH measurements.

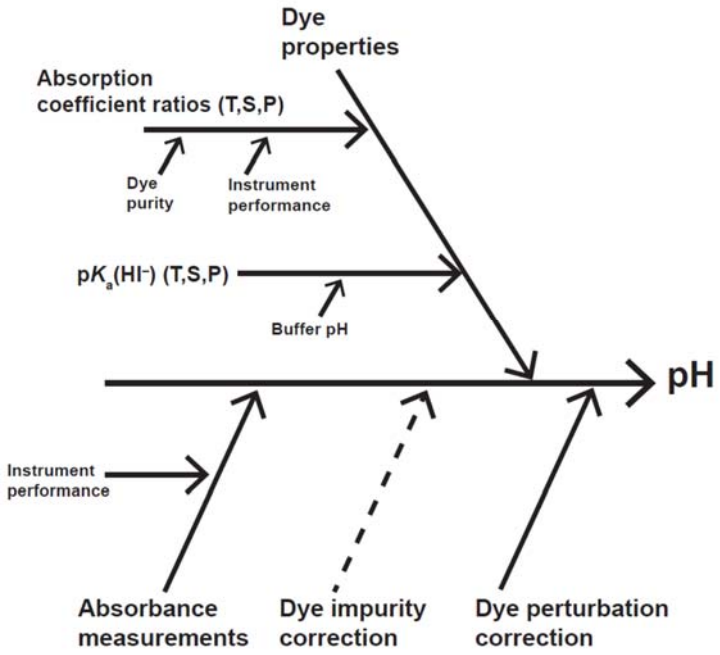

Figure 1.1 Diagram of major uncertainty contributions in spectrophotometric pH measurements, based on **Eq. 2**.

Sulfonephthalein indicators are diprotic acids which exist in predominantly two forms in the pH range of most natural waters (i.e., $HI^- \overset{K_a}{\rightleftarrows} H^+ + I^{2-}$, where I represents the indicator). The absorption spectra of the two forms of dye are substantially different, and thus the composite spectrum of the sample solution with dye will reflect the relative proportions of the dye species, which is a function of pH (**Eq. 1**).

$$\mathrm{pH} = \mathrm{p}K_a(\mathrm{HI}^-) + \log\left(\frac{[\mathrm{I}^{2-}]}{[\mathrm{HI}^-]}\right) \tag{1}$$

The concentration ratio $\frac{[\mathrm{I}^{2-}]}{[\mathrm{HI}^-]}$ is typically inferred from the ratio of the absorbances at the two wavelengths corresponding to the absorbance maxima of the I^{2-} and HI^- species ($R = A_{578}/A_{434}$, **Figure 1.2**) and from the molar absorption coefficient ratios of the species (e_1, e_2, e_3) as in **Eq. 2.**

$$\mathrm{pH} = -\log(K_a(\mathrm{HI}^-)) + \log\left(\frac{R - e_1}{e_2 - \mathrm{R}\, e_3}\right) \quad (2)$$

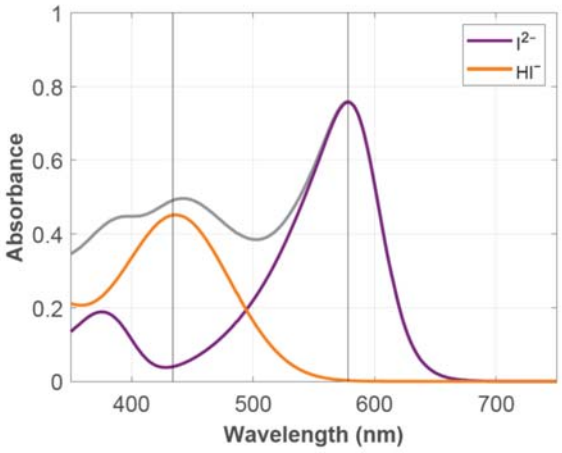

Figure 1.2. Example spectrum of purified mCP in seawater. The composite spectrum (gray) is the sum of contributions from the I^{2-} and HI^- species, which have absorbance maxima at 578 and 434 nm respectively.

The molar absorption coefficient ratios are determined from measurements of the spectra of a dye solution at high pH, in which nearly all the dye is present in the I^{2-} form, and at low pH, in which nearly all the dye is in the HI^- form (Liu et al., 2011). Once the molar absorption coefficient ratios are known, the dye dissociation constant $K_a(HI^-)$ can be determined from **Eq. 2** by measuring the R in buffers which have been assigned pH values based on Harned cell measurements, thus establishing traceability to the primary method of pH measurement (Buck et al., 2002; Milton and Quinn, 2001). These indicator properties have been characterized over a range of temperatures, pressures, and salinities (including brackish and freshwater) for *m*-cresol purple (mCP), the most widely used indicator for seawater pH measurements (Lai et al., 2016; Liu et al., 2011; Loucaides et al., 2017; Müller and Rehder, 2018; Soli et al., 2013), and a metrologically traceable characterization of purified mCP was

recently performed using Harned cell-characterized pH buffer standards (Müller and Rehder, 2018).

The current calibration of the spectrophotometric pH indicator has a number of deficiencies that remains to be fully addressed. Of the uncertainties described in **Figure 1.1**, the contribution from the dye properties is the most important. The measured molar absorption coefficient ratios are sensitive to impurities in the dye (Liu et al., 2011; Yao et al., 2007) and to the bandpass and performance of the spectrophotometer used to characterize the dye (DeGrandpre et al., 2014). Hence, the published values of the dye properties may not necessarily be compatible with measurements made on spectrophotometers with a significantly different bandpass or using a different batch of dye with a substantially different level of impurities. Although purified indicators are now produced by several laboratories in limited quantities, there is evidence that not all purified dyes are of identical purity and pH performance (Takeshita et al., submitted). A project is ongoing at the National Institute of Standards and Technology (NIST) to produce and characterize a standard reference material for pure mCP, which will ultimately reduce the uncertainty in our knowledge of the dye properties.

A more fundamental uncertainty in the calibration of spectrophotometric pH measurements is related to the assignment of pH values to the buffer standards used to determine $K_a(HI^-)$, which has an unknown uncertainty from assumptions about the activity coefficients of HCl in the buffers and from the lack of traceability of conventional definitions of seawater pH to the International System of Units (Dickson et al., 2016). Although these issues will not be addressed in this dissertation, resolving these problems remain urgent priorities for the pH community.

Of secondary importance compared to the uncertainty in the dye properties are the uncertainty contributions from differences in spectrophotometer performance and operational procedures (i.e., bottom half of **Figure 1.1**), which have not been thoroughly evaluated. Given that wavelength and absorbance errors from poorly functioning spectrophotometers can contribute to error in spectrophotometric pH measurements (DeGrandpre et al., 2014), there is a need to develop quality control procedures and to investigate how differences in performance between spectrophotometers of the same design and changes in an individual spectrophotometer's performance affect the long-term stability of pH measurements. Analyzing the full spectra collected on some systems may provide additional information about the contribution of spectrophotometer performance to pH measurement uncertainty. The use of full spectra for spectrophotometric pH determination has been proposed as an alternative to the conventional method based on the ratio of absorbances at two wavelengths (Ohline et al., 2007). Although this method purportedly offers superior precision, its overall uncertainty compared to the conventional Ratio Method is unknown.

Another source of uncertainty related to operational procedures is from the dye perturbation correction. Spectrophotometric pH measurements require an adjustment for the effect of the dye addition, which will necessarily change the sample pH as the indicator itself is an acid. There are several different approaches to achieving such a correction, each with its own assumptions and uncertainties which have yet to be rigorously evaluated (Aβmann et al., 2011; Chierici et al., 1999; Clayton and Byrne, 1993; Li et al., 2020; Martz et al., 2003; Seidel et al., 2008).

Finally, as purified dyes are expensive and limited in availability, many users have a desire and need for obtaining accurate pH measurements with impure dyes. This could

potentially be achieved by several different ways: 1) characterizing the properties of the impure dye (Clayton and Byrne, 1993), 2) determining the offsets in the pH measurements relative to a reference (ideally pure) dye (Liu et al., 2011; Yao et al., 2007), or 3) adjusting for the estimated impurity absorption at 434 nm (Douglas and Byrne, 2017). However, it may be difficult to calibrate impure dyes against a purified dye and evaluate the accuracy if the user's sample of reference purified dye was not absolutely pure or if the published properties for purified dye were not truly representative of a pure dye. Estimating the impurity absorption adjustment (Douglas and Byrne, 2017) typically relies on assumptions about the properties of a pure dye and the impurity absorption behavior, which can vary between different batches of impure dyes. There is, therefore, additional uncertainty when using impure dyes, which can show up as a bias and in the variance in spectrophotometric pH measurements. This dissertation aims to better constrain some of the uncertainties discussed in this section and understand their implications for using pH measurements to study the ocean CO_2 system.

1.6 Dissertation Outline

Chapter 2 of this dissertation (published as Fong and Dickson, 2019) investigated potential causes for the pH-dependent discrepancies between measured and calculated pH observed on open ocean cruises. The study analyzed datasets from four GO-SHIP repeat hydrography cruises to evaluate how systematic errors in the carbonic acid dissociation constants (K_1 and K_2), the total boron-salinity ratio (B_T/S), and in the C_T and A_T measurements individually affect the pH discrepancies and then determined how a combination of plausible adjustments for these errors might reduce the mean discrepancy in each dataset to zero and eliminate the pH-dependence. Achieving these two goals required, in addition to adjustments in K_1, K_2 and B_T/S, acknowledging an additional contribution to the measured A_T, potentially from

organic bases. Although organic alkalinity has been proposed as being important in coastal and estuarine waters (Cai et al., 1998; Patsavas et al., 2015; Yang et al., 2015), these results suggest, for the first time, that organic alkalinity may be ubiquitous in the open ocean. If true, then CO_2 calculations using A_T as one of the input parameters will always be in error without accounting for organic alkalinity contributions. Additionally, the presence of organic alkalinity may result in systematic misinterpretation of A_T titrations (Sharp and Byrne, 2020). The proposed adjustments, which have been verified in an independent analysis of the complete Global Ocean Data Analysis Project (GLODAP) v2 dataset after 2010 (Álvarez et al., 2020), suggest that the lack of thermodynamic consistency in the CO_2 system can be rationalized with plausible systematic errors in the various constants used in CO_2 calculations, without requiring that any of the CO_2 measurements (including pH) be significantly in error. There is thus a need to better constrain the uncertainty in the various constants, a concern that has been echoed by others particularly in regards to the values of the carbonic acid dissociation constants at low temperature and high pressure (Raimondi et al., 2019; Sulpis et al., 2020; Woosley et al., 2017)

Chapter 3 developed methods for using full spectral data for the quality control of spectrophotometric pH measurements, based on the Full Spectrum Method of Ohline et al. (2007) and Principal Components Analysis (PCA). These methods, which are easily implemented on diode array spectrophotometers, are useful for identifying and potentially correcting instrumental contributions to pH measurement error. This study showed that using the Full Spectrum Method together with the Ratio Method can provide useful quality control information, but the Full Spectrum Method requires reliable data on the dye species absorption spectra, ideally measured on the user's spectrophotometer. PCA-based methods, on the other hand, do not require information on the dye species absorption spectra. Systematic errors in

absorbance, which appeared to have developed from the degradation of the deuterium lamp on the spectrophotometer, were identified from PCA on datasets of spectra. These absorbance errors can result in errors in the dye perturbation correction and in the measured dye properties but are unlikely to contribute significantly to the pH-dependent pH discrepancies observed on open ocean cruises and examined in **Chapter 2**. It is recommended that laboratories using diode array spectrophotometers analyze the spectra from regular measurements of a stable batch of seawater or buffered solution to monitor for changes in spectrophotometer performance affecting pH measurement quality.

Chapter 4 evaluated the uncertainty in the adjustments made to the measured pH of seawater samples for the effect of the indicator dye and provided recommendations for minimizing this source of uncertainty. The dye perturbation correction can be achieved by extrapolating measurements to zero dye concentration for each individual sample (Aβmann et al., 2011; Martz et al., 2003; Seidel et al., 2008), inferring the necessary adjustment from a correction curve based on dye additions to samples over a range of pH (Clayton and Byrne, 1993), or estimating the adjustment from a chemical equilibrium model (Chierici et al., 1999). This study used a combination of dye addition experiments, chemical modeling, and numerical simulations to evaluate the uncertainty in the various dye perturbation correction approaches. It was found that modeled and measured pH perturbations may not be perfectly consistent with each other potentially due to the systematic errors in measuring absorbance identified in **Chapter 3**. Additionally, the use of the correction curve approach may result in a systematic error in the pH correction due to the assumption of linearity, a finding corroborated by another study (Li et al., 2020). However, for open ocean samples, this error only becomes large (>0.01) at the high dye concentrations used in a 1 cm cell and can be minimized by using a low pH dye

solution, which linearizes the correction curve. These findings show that the dye perturbation corrections are a small source of uncertainty in spectrophotometric pH measurements and provide insights into how users of short pathlength cells can reduce the uncertainty of their dye perturbation adjustments.

Chapter 5 examined the likely accuracy and consistency of spectrophotometric pH measurements using impure mCP. The properties of a batch of purified and impure mCP were characterized, and these data were used together with other laboratory measurements and numerical simulations to evaluate the performance of three different published methods for calibrating impure dyes. For the most accurate calibration of impure mCP, it is recommended to either characterize the properties of the dye or measure the dye batch-specific pH offsets relative to a purified dye. A simpler approach based on adjustments for the estimated impurity absorption at 434 nm (Douglas and Byrne, 2017) may have greater uncertainty due to assumptions about the impurity absorption behavior. Additionally, the properties for the batch of purified mCP characterized in this study differed from published values (Liu et al., 2011) in a way that suggested a small impurity contribution in the published properties and potentially an error in the $pK_a(HI^-)$. The uncertainty in the properties of purified mCP can result in inconsistencies between impure and purified mCP pH measurements larger than 0.005 at pH > 8. These results highlight the need for a re-evaluation of the properties of purified mCP and development of quality control procedures for the purification of mCP. Users of impure dyes should consider their measurement quality needs in the context of the findings in this study.

1.7 Broader implications and future research

As pH measurements become increasingly important for studies of the ocean's CO_2 system, there is an urgent need to better quantify their uncertainty. This dissertation contributes

to the community's understanding of the uncertainty of spectrophotometric pH measurements and provides tools for their quality control. Although the uncertainty of seawater pH measurements is still not fully understood, the quality of seawater pH measurements have undoubtedly improved over time with the development of spectrophotometric pH measurements using purified indicators, such that it is possible to identify systematic trends in the offsets between measured and calculated CO_2 parameters which point to gaps in our understanding of the CO_2 and other seawater acid-base systems. Accurate CO_2 parameter measurements and the characterization of CO_2 and other seawater acid-base equilibria are critical to quantifying long-term changes in the carbon cycling of the ocean. The potential causes for the thermodynamic inconsistencies in the seawater CO_2 system and remaining challenges with pH measurements highlighted by this work are issues that the community will need to investigate and resolve. This work is part of a larger and continuing community effort to better constrain the uncertainty in seawater CO_2 measurements and in the thermodynamic models of the CO_2 system. At the time of writing, the community has already made progress in investigating some of these outstanding questions with recent studies addressing the effects of organic acids and bases on the interpretation of A_T measurements (Sharp and Byrne, 2020), re-determining K_2 using a spectrophotometric method (Schockman and Byrne, 2021), evaluating systematic errors in K_1 and K_2 at low temperatures (Sulpis et al., 2020; Woosley, 2020), and verifying the total boron-salinity ratio (Olafsson et al., 2020). These problems will undoubtedly continue to be revisited and scrutinized in the decades to come.

1.8 References

Álvarez, M., Fajar, N.M., Carter, B.R., Guallart, E.F., Pérez, F.F., Woosley, R.J. and Murata, A., 2020. Global Ocean Spectrophotometric pH Assessment: Consistent Inconsistencies. Environmental Science & Technology, 54(18): 10977-10988.

Aβmann, S., Frank, C. and Kortzinger, A., 2011. Spectrophotometric high-precision seawater pH determination for use in underway measuring systems. Ocean Science, 7(5): 597-607.

Bakker, D.C.E., Pfeil, B., Landa, C.S., Metzl, N., O'Brien, K.M., Olsen, A., Smith, K., Cosca, C., Harasawa, S., Jones, S.D., Nakaoka, S., Nojiri, Y., Schuster, U., Steinhoff, T., Sweeney, C., Takahashi, T., Tilbrook, B., Wada, C., Wanninkhof, R., Alin, S.R., Balestrini, C.F., Barbero, L., Bates, N.R., Bianchi, A.A., Bonou, F., Boutin, J., Bozec, Y., Burger, E.F., Cai, W.J., Castle, R.D., Chen, L., Chierici, M., Currie, K., Evans, W., Featherstone, C., Feely, R.A., Fransson, A., Goyet, C., Greenwood, N., Gregor, L., Hankin, S., Hardman-Mountford, N.J., Harlay, J., Hauck, J., Hoppema, M., Humphreys, M.P., Hunt, C.W., Huss, B., Ibánhez, J.S.P., Johannessen, T., Keeling, R., Kitidis, V., Körtzinger, A., Kozyr, A., Krasakopoulou, E., Kuwata, A., Landschützer, P., Lauvset, S.K., Lefèvre, N., Lo Monaco, C., Manke, A., Mathis, J.T., Merlivat, L., Millero, F.J., Monteiro, P.M.S., Munro, D.R., Murata, A., Newberger, T., Omar, A.M., Ono, T., Paterson, K., Pearce, D., Pierrot, D., Robbins, L.L., Saito, S., Salisbury, J., Schlitzer, R., Schneider, B., Schweitzer, R., Sieger, R., Skjelvan, I., Sullivan, K.F., Sutherland, S.C., Sutton, A.J., Tadokoro, K., Telszewski, M., Tuma, M., van Heuven, S.M.A.C., Vandemark, D., Ward, B., Watson, A.J. and Xu, S., 2016. A multi-decade record of high-quality fCO2 data in version 3 of the Surface Ocean CO_2 Atlas (SOCAT). Earth Syst. Sci. Data, 8(2): 383-413.

Bates, N.R., Astor, Y.M., Church, M.J., Currie, K., Dore, J.E., Gonzalez-Davila, M., Lorenzoni, L., Muller-Karger, F., Olafsson, J. and Santana-Casiano, J.M., 2014. A Time-Series View of Changing Surface Ocean Chemistry Due to Ocean Uptake of Anthropogenic CO_2 and Ocean Acidification. Oceanography, 27(1): 126-141.

Bellerby, R.G.J., Olsen, A., Johannessen, T. and Croot, P., 2002. A high precision spectrophotometric method for on-line shipboard seawater pH measurements: the automated marine pH sensor (AMpS). Talanta, 56(1): 61-69.

Bockmon, E.E. and Dickson, A.G., 2015. An inter-laboratory comparison assessing the quality of seawater carbon dioxide measurements. Marine Chemistry, 171: 36-43.

Bresnahan Jr., P.J., Martz, T.R., Takeshita, Y., Johnson, K.S. and LaShomb, M., 2014. Best practices for autonomous measurement of seawater pH with the Honeywell Durafet. Methods in Oceanography, 9: 44-60.

Buck, R.P., Rondinini, S., Covington, A.K., Baucke, F.G.K., Brett, C.M.A., Camoes, M.F., Milton, M.J.T., Mussini, T., Naumann, R., Pratt, K.W., Spitzer, P. and Wilson, G.S., 2002. Measurement of pH. Definition, standards, and procedures (IUPAC Recommendations 2002). Pure and Applied Chemistry, 74(11): 2169-2200.

Bushinsky, S.M., Landschützer, P., Rödenbeck, C., Gray, A.R., Baker, D., Mazloff, M.R., Resplandy, L., Johnson, K.S. and Sarmiento, J.L., 2019a. Reassessing Southern Ocean Air-Sea CO_2 Flux Estimates With the Addition of Biogeochemical Float Observations. Global Biogeochemical Cycles, 33(11): 1370-1388.

Bushinsky, S.M., Takeshita, Y. and Williams, N.L., 2019b. Observing Changes in Ocean Carbonate Chemistry: Our Autonomous Future. Current Climate Change Reports, 5(3): 207-220.

Byrne, R.H. and Breland, J.A., 1989. High precision multiwavelength pH determinations in seawater using cresol red. Deep Sea Research Part A. Oceanographic Research Papers, 36(5): 803-810.

Byrne, R.H., Mecking, S., Feely, R.A. and Liu, X., 2010. Direct observations of basin-wide acidification of the North Pacific Ocean. Geophysical Research Letters, 37(2).

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

Caldeira, K. and Wickett, M.E., 2003. Anthropogenic carbon and ocean pH. Nature, 425(6956): 365-365.

Carter, B.R., Feely, R.A., Williams, N.L., Dickson, A.G., Fong, M.B. and Takeshita, Y., 2018. Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate. Limnology and Oceanography: Methods, 16(2): 119-131.

Carter, B.R., Radich, J.A., Doyle, H.L. and Dickson, A.G., 2013. An automated system for spectrophotometric seawater pH measurements. Limnology and Oceanography: Methods, 11(1): 16-27.

Chierici, M., Fransson, A. and Anderson, L.G., 1999. Influence of *m*-cresol purple indicator additions on the pH of seawater samples: correction factors evaluated from a chemical speciation model. Marine Chemistry, 65(3): 281-290.

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

Clayton, T.D., Byrne, R.H., Breland, J.A., Feely, R.A., Millero, F.J., Campbell, D.M., Murphy, P.P. and Lamb, M.F., 1995. The role of pH measurements in modern oceanic CO_2-system characterizations: Precision and thermodynamic consistency. Deep Sea Research Part II: Topical Studies in Oceanography, 42(2): 411-429.

DeGrandpre, M.D., Spaulding, R.S., Newton, J.O., Jaqueth, E.J., Hamblock, S.E., Umansky, A.A. and Harris, K.E., 2014. Considerations for the measurement of spectrophotometric pH for ocean acidification and other studies. Limnology and Oceanography: Methods, 12(12): 830-839.

Dickson, A.G., 1993. The measurement of sea water pH. Marine Chemistry, 44(2): 131-142.

Dickson, A.G., 1998. Underway measurement of pH: method and meaning. In: S. Blain (Editor), Marc'hMor Workshop, IUEM, Brest, pp. 104-107.

Dickson, A.G., 2010. The carbon dioxide system in seawater: equilibrium chemistry and measurements. In: U. Riebesell, V.J. Fabry, L. Hansson and J.-P. Gattuso (Editors), Guide to best practices for ocean acidification research and data reporting. Publications Office of the European Union, Luxembourg, pp. 17-38.

Dickson, A.G., Camoes, M.F., Spitzer, P., Fisicaro, P., Stoica, D., Pawlowicz, R. and Feistel, R., 2016. Metrological challenges for measurements of key climatological observables. Part 3: Seawater pH. Metrologia, 53(1): R26-R39.

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

Doney, S.C., Fabry, V.J., Feely, R.A. and Kleypas, J.A., 2009. Ocean Acidification: The Other CO_2 Problem. Annual Review of Marine Science, 1: 169-192.

Douglas, N.K. and Byrne, R.H., 2017. Achieving accurate spectrophotometric pH measurements using unpurified meta-cresol purple. Marine Chemistry, 190: 66-72.

Easley, R.A. and Byrne, R.H., 2012. Spectrophotometric Calibration of pH Electrodes in Seawater Using Purified *m*-Cresol Purple. Environmental Science & Technology, 46(9): 5018-5024.

Fabry, V.J., Seibel, B.A., Feely, R.A. and Orr, J.C., 2008. Impacts of ocean acidification on marine fauna and ecosystem processes. ICES Journal of Marine Science, 65(3): 414-432.

Fassbender, A.J., Sabine, C.L. and Palevsky, H.I., 2017. Nonuniform ocean acidification and attenuation of the ocean carbon sink. Geophysical Research Letters, 44(16): 8404-8413.

Fong, M.B. and Dickson, A.G., 2019. Insights from GO-SHIP hydrography data into the thermodynamic consistency of CO_2 system measurements in seawater. Marine Chemistry, 211: 52-63.

Gray, A.R., Johnson, K.S., Bushinsky, S.M., Riser, S.C., Russell, J.L., Talley, L.D., Wanninkhof, R., Williams, N.L. and Sarmiento, J.L., 2018. Autonomous Biogeochemical Floats Detect Significant Carbon Dioxide Outgassing in the High-Latitude Southern Ocean. Geophysical Research Letters, 45(17): 9049-9057.

Johnson, K.S., Jannasch, H.W., Coletti, L.J., Elrod, V.A., Martz, T.R., Takeshita, Y., Carlson, R.J. and Connery, J.G., 2016. Deep-Sea DuraFET: A Pressure Tolerant pH Sensor Designed for Global Sensor Networks. Analytical Chemistry, 88(6): 3249-3256.

Keeling, C.D., 1968. Carbon dioxide in surface ocean waters: 4. Global distribution. Journal of Geophysical Research, 73(14): 4543-4553.

Khatiwala, S., Tanhua, T., Fletcher, S.M., Gerber, M., Doney, S.C., Graven, H.D., Gruber, N., McKinley, G.A., Murata, A., Rios, A.F. and Sabine, C.L., 2013. Global ocean storage of anthropogenic carbon. Biogeosciences, 10(4): 2169-2191.

Lai, C.-Z., DeGrandpre, M.D., Wasser, B.D., Brandon, T.A., Clucas, D.S., Jaqueth, E.J., Benson, Z.D., Beatty, C.M. and Spaulding, R.S., 2016. Spectrophotometric measurement of freshwater pH with purified meta-cresol purple and phenol red. Limnology and Oceanography: Methods, 14(12): 864-873.

Li, X., García-Ibáñez, M.I., Carter, B.R., Chen, B., Li, Q., Easley, R.A. and Cai, W.-J., 2020. Purified meta-Cresol Purple dye perturbation: How it influences spectrophotometric pH measurements. Marine Chemistry, 225: 103849.

Liu, X., Patsavas, M.C. and Byrne, R.H., 2011. Purification and Characterization of meta-Cresol Purple for Spectrophotometric Seawater pH Measurements. Environmental Science & Technology, 45(11): 4862-4868.

Loucaides, S., Rèrolle, V.M.C., Papadimitriou, S., Kennedy, H., Mowlem, M.C., Dickson, A.G., Gledhill, M. and Achterberg, E.P., 2017. Characterization of meta-Cresol Purple for spectrophotometric pH measurements in saline and hypersaline media at sub-zero temperatures. Scientific Reports, 7(1): 2481.

Martz, T.R., Carr, J.J., French, C.R. and DeGrandpre, M.D., 2003. A Submersible Autonomous Sensor for Spectrophotometric pH Measurements of Natural Waters. Analytical Chemistry, 75(8): 1844-1850.

Martz, T.R., Connery, J.G. and Johnson, K.S., 2010. Testing the Honeywell Durafet® for seawater pH applications. Limnology and Oceanography: Methods, 8(5): 172-184.

McElligott, S., Byrne, R.H., Lee, K., Wanninkhof, R., Millero, F.J. and Feely, R.A., 1998. Discrete water column measurements of CO_2 fugacity and pH_T in seawater: A comparison of direct measurements and thermodynamic calculations. Marine Chemistry, 60(1): 63-73.

Milton, M.J.T. and Quinn, T.J., 2001. Primary methods for the measurement of amount of substance. Metrologia, 38(4): 289-296.

Müller, J.D. and Rehder, G., 2018. Metrology of pH Measurements in Brackish Waters—Part 2: Experimental Characterization of Purified meta-Cresol Purple for Spectrophotometric pHT Measurements. Frontiers in Marine Science, 5(177).

Naviaux, J.D., Subhas, A.V., Dong, S., Rollins, N.E., Liu, X., Byrne, R.H., Berelson, W.M. and Adkins, J.F., 2019. Calcite dissolution rates in seawater: Lab vs. in-situ measurements and inhibition by organic matter. Marine Chemistry, 215: 103684.

Newton, J.A., Feely, R.A., Jewett, E.B., Williamson, P. and Mathis, J., 2014. Global Ocean Acidification Observing Network: Requirements and Governance Plan.

Ohline, S.M., Reid, M.R., Husheer, S.L.G., Currie, K.I. and Hunter, K.A., 2007. Spectrophotometric determination of pH in seawater off Taiaroa Head, Otago, New Zealand: Full-spectrum modelling and prediction of pCO_2 levels. Marine Chemistry, 107(2): 143-155.

Olafsson, J., Lee, K., Olafsdottir, S.R., Benoit-Cattin, A., Lee, C.-H. and Kim, M., 2020. Boron to salinity ratios for Atlantic, Arctic and Polar Waters: A view from downstream. Marine Chemistry, 224: 103809.

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. and 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(2): 297-323.

Orr, J.C., Fabry, V.J., Aumont, O., Bopp, L., Doney, S.C., Feely, R.A., Gnanadesikan, A., Gruber, N., Ishida, A., Joos, F., Key, R.M., Lindsay, K., Maier-Reimer, E., Matear, R., Monfray, P., Mouchet, A., Najjar, R.G., Plattner, G.K., Rodgers, K.B., Sabine, C.L., Sarmiento, J.L., Schlitzer, R., Slater, R.D., Totterdell, I.J., Weirig, M.F., Yamanaka, Y. and Yool, A., 2005. Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. Nature, 437(7059): 681-686.

Park, P.K., 1969. Oceanic CO_2 system: An evaluation of ten methods of investigation. Limnology and Oceanography, 14(2): 179-186.

Patsavas, M.C., Byrne, R.H. and Liu, X., 2013. Purification of meta-cresol purple and cresol red by flash chromatography: Procedures for ensuring accurate spectrophotometric seawater pH measurements. Marine Chemistry, 150: 19-24.

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

Pörtner, H.O., Karl, D.M., Boyd, P.W., Cheung, W.W.L., Lluch-Cota, S.E., Nojiri, Y., Schmidt, D.N. and Zavialov, P.O., 2014. Ocean systems. In: C.B. Field et al. (Editors), Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 411-484.

Raimondi, L., Matthews, J.B.R., Atamanchuk, D., Azetsu-Scott, K. and Wallace, D.W.R., 2019. The internal consistency of the marine carbon dioxide system for high latitude shipboard and in situ monitoring. Marine Chemistry, 213: 49-70.

Riebesell, U., Körtzinger, A. and Oschlies, A., 2009. Sensitivities of marine carbon fluxes to ocean change. Proceedings of the National Academy of Sciences, 106(49): 20602-20609.

Sabine, C.L. and Tanhua, T., 2010. Estimation of Anthropogenic CO_2 Inventories in the Ocean. Annual Review of Marine Science, 2: 175-198.

Schockman, K.M. and Byrne, R.H., 2021. Spectrophotometric Determination of the Bicarbonate Dissociation Constant in Seawater. Geochimica et Cosmochimica Acta.

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

Sharp, J.D. and Byrne, R.H., 2020. Interpreting measurements of total alkalinity in marine and estuarine waters in the presence of proton-binding organic matter. Deep Sea Research Part I: Oceanographic Research Papers, 165: 103338.

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

Soli, A.L., Pav, B.J. and Byrne, R.H., 2013. The effect of pressure on meta-Cresol Purple protonation and absorbance characteristics for spectrophotometric pH measurements in seawater. Marine Chemistry, 157: 162-169.

Sulpis, O., Lauvset, S.K. and Hagens, M., 2020. Current estimates of K1* and K2* appear inconsistent with measured CO_2 system parameters in cold oceanic regions. Ocean Sci., 16(4): 847-862.

Takahashi, T., Weiss, R.F., Culberson, C.H., Edmond, J.M., Hammond, D.E., Wong, C.S., Li, Y.-H. and Bainbridge, A.E., 1970. A carbonate chemistry profile at the 1969 GEOSECS Intercalibration Station in the eastern Pacific Ocean. Journal of Geophysical Research, 75(36): 7648-7666.

Takeshita, Y., Johnson, K.S., Coletti, L.J., Jannasch, H.W., Walz, P.M. and Warren, J.K., 2020. Assessment of pH dependent errors in spectrophotometric pH measurements of seawater. Marine Chemistry, 223: 103801.

Takeshita, Y., Warren, J.K., Liu, X., Spaulding, R.S., Byrne, R.H., Carter, B.R., DeGrandpre, M.D., Murata, A. and Watanabe, S.-i., submitted. Consistency and stability of purified meta-cresol purple for spectrophotometric pH measurements in seawater. Marine Chemistry.

Talley, L.D., Feely, R.A., Sloyan, B.M., Wanninkhof, R., Baringer, M.O., Bullister, J.L., Carlson, C.A., Doney, S.C., Fine, R.A., Firing, E., Gruber, N., Hansell, D.A., Ishii, M., Johnson, G.C., Katsumata, K., Key, R.M., Kramp, M., Langdon, C., Macdonald, A.M., Mathis, J.T., McDonagh, E.L., Mecking, S., Millero, F.J., Mordy, C.W., Nakano, T., Sabine, C.L., Smethie, W.M., Swift, J.H., Tanhua, T., Thurnherr, A.M., Warner, M.J. and Zhang, J.-Z., 2016. Changes in Ocean Heat, Carbon Content, and Ventilation: A Review of the First Decade of GO-SHIP Global Repeat Hydrography. Annual Review of Marine Science, 8(1): 185-215.

Tapp, M., Hunter, K., Currie, K. and Mackaskill, B., 2000. Apparatus for continuous-flow underway spectrophotometric measurement of surface water pH. Marine Chemistry, 72(2): 193-202.

Williams, N.L., Juranek, L.W., Feely, R.A., Johnson, K.S., Sarmiento, J.L., Talley, L.D., Dickson, A.G., Gray, A.R., Wanninkhof, R., Russell, J.L., Riser, S.C. and Takeshita, Y., 2017. Calculating surface ocean pCO2 from biogeochemical Argo floats equipped with pH: An uncertainty analysis. Global Biogeochemical Cycles, 31(3): 591-604.

Williams, N.L., Juranek, L.W., Johnson, K.S., Feely, R.A., Riser, S.C., Talley, L.D., Russell, J.L., Sarmiento, J.L. and Wanninkhof, R., 2016. Empirical algorithms to estimate water column pH in the Southern Ocean. Geophysical Research Letters, 43(7): 3415-3422.

Woosley, R.J., 2020. Evaluation of the temperature dependence of dissociation constants for the marine carbon system using pH and certified reference materials. Marine Chemistry: 103914.

Woosley, R.J., Millero, F.J. and Takahashi, T., 2017. Internal consistency of the inorganic carbon system in the Arctic Ocean. Limnology and Oceanography: Methods, 15(10): 887-896.

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

Yang, B., Patsavas, M.C., Byrne, R.H. and Ma, J., 2014. Seawater pH measurements in the field: A DIY photometer with 0.01 unit pH accuracy. Marine Chemistry, 160: 75-81.

Yao, W., Liu, X. and Byrne, R.H., 2007. Impurities in indicators used for spectrophotometric seawater pH measurements: Assessment and remedies. Marine Chemistry, 107(2): 167-172.

Zhang, H. and Byrne, R.H., 1996. Spectrophotometric pH measurements of surface seawater at in-situ conditions: absorbance and protonation behavior of thymol blue. Marine Chemistry, 52(1): 17-25.

Chapter 2

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

Marine Chemistry 211 (2019) 52–63

Contents lists available at ScienceDirect

Marine Chemistry

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

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

Michael B. Fong*, Andrew G. Dickson

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

ARTICLE INFO

Keywords:
CO_2 system
Internal consistency
Organic alkalinity
Spectrophotometric pH

ABSTRACT

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

1. Introduction

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

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

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

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

https://doi.org/10.1016/j.marchem.2019.03.006
Received 19 September 2018; Received in revised form 24 January 2019; Accepted 11 March 2019
Available online 13 March 2019

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

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

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

2. Methods

2.1. Choice of data sets

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

2.2. Approach

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

$$\Delta pH = pH_{spec} - pH_{calc}(C_T, A_T, \ldots). \quad (1)$$

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

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

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

2.3. Identifying possible systematic errors

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

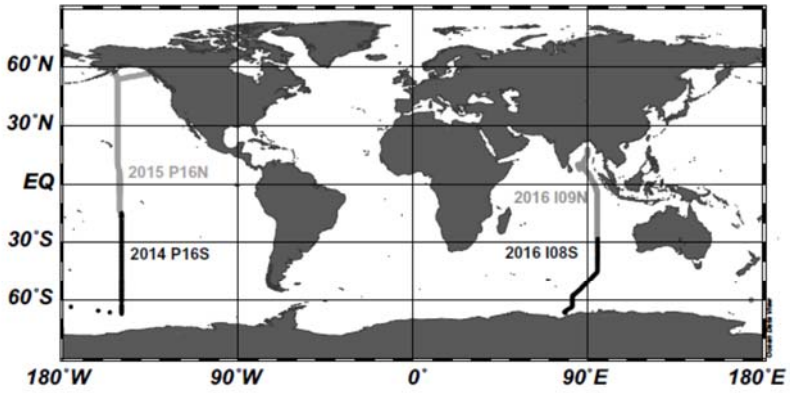

Fig. 1. Map showing cruise tracks for the four cruise data sets discussed here: 2014 P16S, 2015 P16N, 2016 I08S, and 2016 I09N.

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

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

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

[b] Estimated bias of a dataset where A_T and C_T were measured with state-of-the-art methods and standardized to CO_2-in-seawater Reference Materials.

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

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

repeated the calculation for this cruise using alternate published values for the dissociation constants (measured in natural seawater) that are available within CO2SYS-MATLAB (Fig. 3). Clearly, there is no set of CO_2 constants that removes the observed slope of ΔpH vs. pH, although the mean value of ΔpH does change noticeably depending on the particular set of CO_2 constants used.

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

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

2.4. Computational approach

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

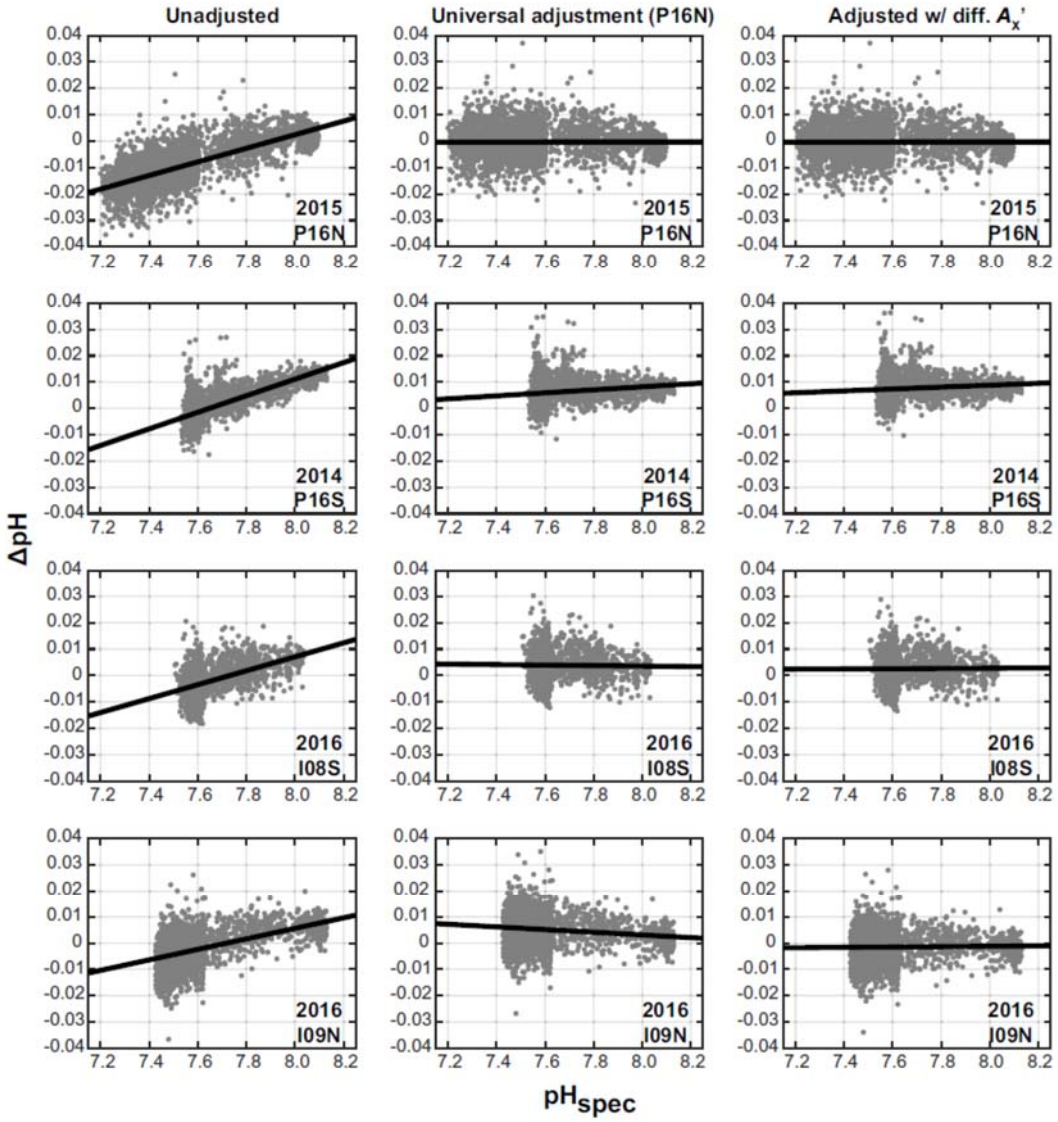

Fig. 2. Values of ΔpH calculated at 25°C and a gauge pressure of zero dbar plotted against measured pH for each cruise. Eq. (1) and the model in Table 1 were used to calculate ΔpH in the panels on the left. ΔpH in the center panels were calculated using adjustments to the model estimated from the 2015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

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

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

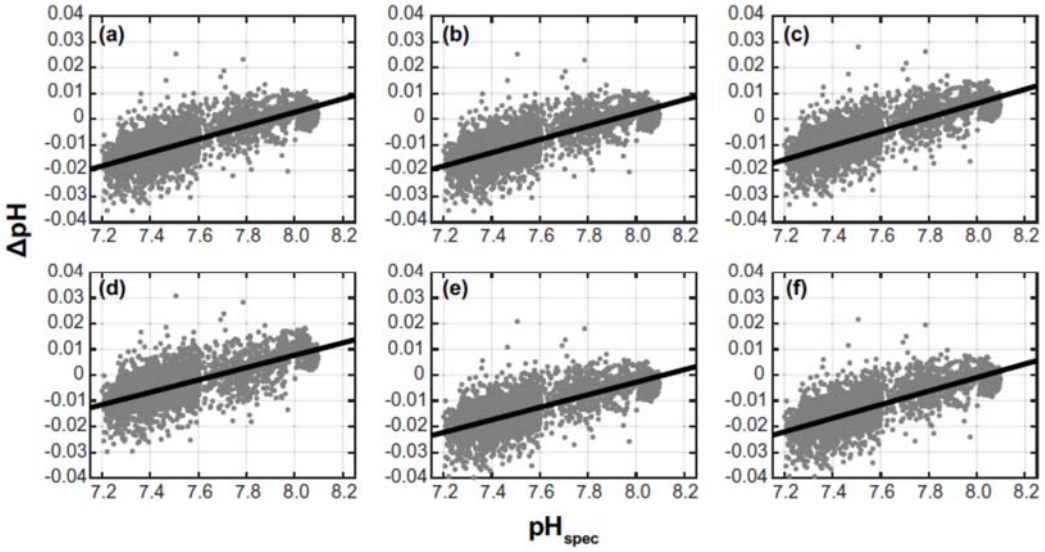

Fig. 3. Values of ΔpH for 2015 P16N calculated using Eq. (1) and the model in Table 1, but with alternate formulations for K_1 and K_2: the modified Mehrbach constants of (a) Dickson and Millero (1987) and (b) Lueker et al. (2000), as well as the constants of (c) Mojica Prieto and Millero (2002), (d)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

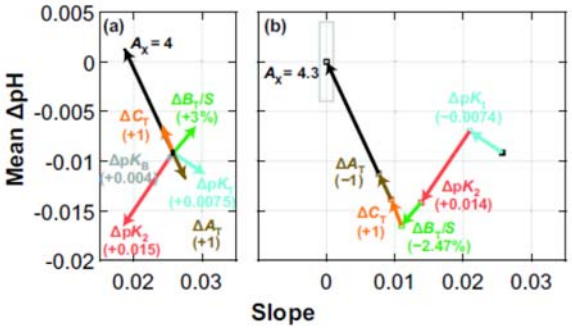

Fig. 4. Vector diagrams showing (a) how individual systematic errors in the various parameters (Table 1) used to calculate ΔpH would affect the value of the slope (of ΔpH *versus* pH) and the mean value of ΔpH for the 2015 P16N dataset and (b) a combination of adjustments (scaled versions of the vectors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

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

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

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

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

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

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

3. Results

3.1. Initial examination of cruises

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

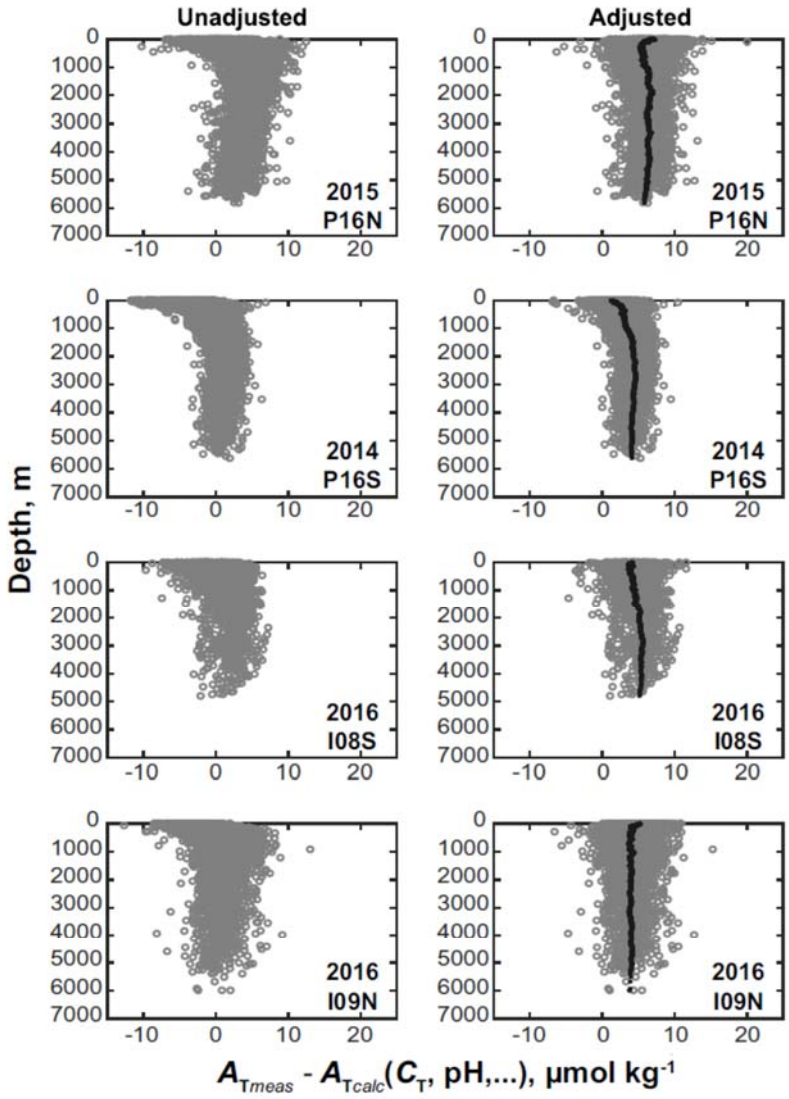

Fig. 5. Estimation of *excess alkalinity* as a function of depth for each of the four cruises. The left panels show the *excess alkalinity* as calculated from Eq. (4) and the model in Table 1, whereas the right panels show the *excess alkalinity* after adjusting pK_1, pK_2, and the total boron-salinity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

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

4. Discussion

4.1. Proposed causes of the pH-dependent discrepancy

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

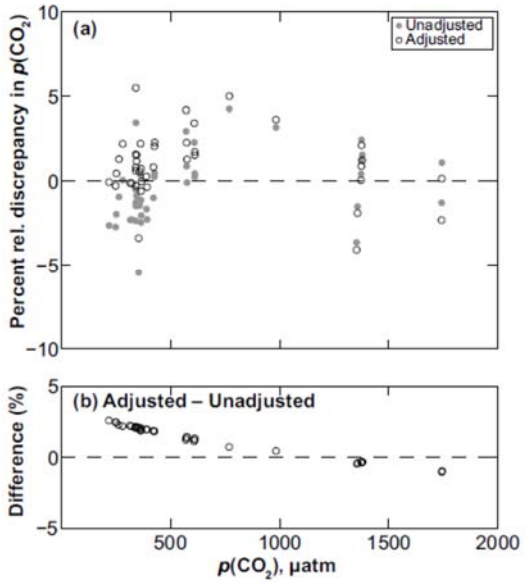

Fig. 6. (a) Recalculation of the data of Lueker et al. (2000), showing how the adjustments proposed for 2015 P16N (neglecting systematic errors in measured C_T and A_T) affect the observed percent relative discrepancy in $p(CO_2)$. Only the data at ~25°C were considered. The relative discrepancy was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

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

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

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

4.2. Plausibility of proposed adjustments to the 2015 P16N dataset

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

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

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

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

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

4.3. Uniqueness of solution

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

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

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

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

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

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

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

4.5. The exception to test our rule?

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

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

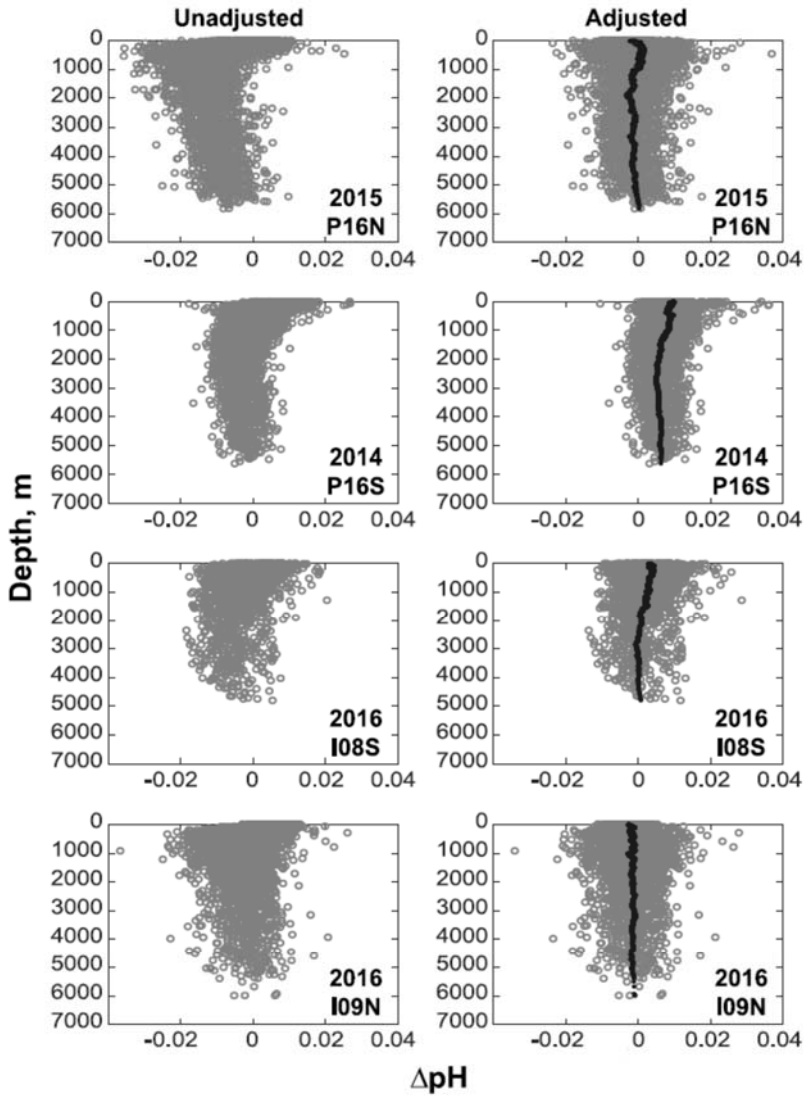

Fig. 7. Values of ΔpH calculated at 25°C and a gauge pressure of zero dbar plotted against depth for each cruise. The unadjusted model in Table 1 was used to calculate ΔpH in the left panels, while ΔpH in the right panels were calculated with a common set of adjustments for pK_1, pK_2, and B_T/S (–0.0074,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

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

4.6. A potential organic source for A_X

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

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

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

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

5. Conclusions

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

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

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

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

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

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

Contributions

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

Acknowledgements

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

References

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

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

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

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

Assessment of internal consistency and sampling considerations, summer 2010. Marine Chem. 176, 174–188.
Clayton, T.D., Byrne, R.H., 1993. Spectrophotometric seawater pH measurements: total hydrogen ion concentration scale calibration of m-cresol purple and at-sea results. Deep Sea Res. Part I 40 (10), 2115–2129.
Clayton, T.D., et al., 1995. The role of pH measurements in modern oceanic CO_2-system characterizations: Precision and thermodynamic consistency. Deep Sea Res. Part II 42 (2), 411–429.
DeGrandpre, M.D., et al., 2014. Considerations for the measurement of spectrophotometric pH for ocean acidification and other studies. Limnol. Oceanogr. 12 (12), 830–839.
DelValls, T.A., Dickson, A.G., 1998. The pH of buffers based on 2-amino-2-hydroxymethyl-1,3-propanediol ('tris') in synthetic sea water. Deep Sea Res. Part I 45 (9), 1541–1554.
Dickson, A.G., 1981. An exact definition of total alkalinity and a procedure for the estimation of alkalinity and total inorganic carbon from titration data. Deep Sea Res. Part A. Oceanogr. Res. Papers 28 (6), 609–623.
Dickson, A.G., 1990. Thermodynamics of the dissociation of boric acid in synthetic seawater from 273.15 to 318.15 K. Deep Sea Res. Part A. Oceanogr. Res. Papers 37 (5), 755–766.
Dickson, A.G., 1992. The development of the alkalinity concept in marine chemistry. Marine Chem. 40 (1), 49–63.
Dickson, A.G., 1993. The measurement of sea water pH. Marine Chem. 44 (2), 131–142.
Dickson, A.G., 2010a. The carbon dioxide system in seawater: equilibrium chemistry and measurements. In: Riebesell, U., Fabry, V.J., Hansson, L., Gattuso, J.-P. (Eds.), Guide to Best Practices for Ocean Acidification Research and Data Reporting. Publications Office of the European Union Luxembourg, pp. 260.
Dickson, A.G., 2010b. Standards for ocean measurements. Oceanography 23 (3), 34–47.
Dickson, A.G., Millero, F.J., 1987. A comparison of the equilibrium constants for the dissociation of carbonic acid in seawater media. Deep Sea Res. Part A. Oceanogr. Res. Papers 34 (10), 1733–1743.
Dickson, A.G., Afghan, J.D., Anderson, G.C., 2003. Reference materials for oceanic CO_2 analysis: a method for the certification of total alkalinity. Marine Chem. 80 (2), 185–197.
Dickson, A.G., Sabine, C.L., Christian, J.R., 2007. Guide to Best Practices for Ocean CO_2 Measurements. PICES Special Publication 191 p.
Hansson, I., 1973. A new set of acidity constants for carbonic acid and boric acid in sea water. Deep Sea Res. Oceanogr. Abstr. 20 (5), 461–478.
Hertkorn, N., et al., 2006. Characterization of a major refractory component of marine dissolved organic matter. Geochimica et Cosmochimica Acta 70 (12), 2990–3010.
Hertkorn, N., Harir, M., Koch, B.P., Michalke, B., Schmitt-Kopplin, P., 2013. High-field NMR spectroscopy and FTICR mass spectrometry: powerful discovery tools for the molecular level characterization of marine dissolved organic matter. Biogeosciences 10 (3), 1583–1624.
van Heuven, S., Pierrot, D., Rae, J.W.B., Lewis, E., Wallace, D.W.R., 2011. MATLAB program developed for CO_2 system calculations. ONRL/CDIAC-105b. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tennessee.
Johnson, K.S., et al., 2016. Deep-Sea DuraFET: A pressure tolerant pH sensor designed for global sensor networks. Anal. Chem. 88 (6), 3249–3256.
Keeling, C.D., 1968. Carbon dioxide in surface ocean waters: 4. Global distribution. J. Geophys. Res. 73 (14), 4543–4553.
Kim, H.-C., Lee, K., Choi, W., 2006. Contribution of phytoplankton and bacterial cells to the measured alkalinity of seawater. Limnol. Oceanogr. 51 (1), 331–338.
Lee, K., Millero, F.J., Byrne, R.H., Feely, R.A., Wanninkhof, R., 2000. The recommended dissociation constants for carbonic acid in seawater. Geophys. Res. Lett. 27 (2), 229–232.
Lee, K., et al., 2010. The universal ratio of boron to chlorinity for the North Pacific and North Atlantic oceans. Geochimica et Cosmochimica Acta 74 (6), 1801–1811.
Liu, X., Patsavas, M.C., Byrne, R.H., 2011. Purification and characterization of meta-cresol purple for spectrophotometric seawater pH measurements. Environ. Sci. Technol. 45 (11), 4862–4868.
Lueker, T.J., Dickson, A.G., Keeling, C.D., 2000. Ocean pCO_2 calculated from dissolved inorganic carbon, alkalinity, and equations for K_1 and K_2: validation based on laboratory measurements of CO_2 in gas and seawater at equilibrium. Marine Chem. 70 (1), 105–119.
Martz, T.R., Connery, J.G., Johnson, K.S., 2010. Testing the Honeywell Durafet® for seawater pH applications. Limnol. Oceanogr. 8 (5), 172–184.
McElligott, S., et al., 1998. Discrete water column measurements of CO_2 fugacity and pH_T in seawater: A comparison of direct measurements and thermodynamic calculations. Marine Chem. 60 (1), 63–73.
Mehrbach, C., Culberson, C.H., Hawley, J.E., Pytkowicz, R.M., 1973. Measurement of the apparent dissociation constants of carbonic acid in seawater at atmospheric-pressure. Limnol. Oceanogr. 18 (6), 897–907.
Millero, F.J., 2007. The marine inorganic carbon cycle. Chem. Rev. 107 (2), 308–341.
Millero, F.J., 2010. Carbonate constants for estuarine waters. Marine Freshwater Res. 61 (2), 139–142.
Millero, F.J., et al., 2002. Dissociation constants for carbonic acid determined from field measurements. Deep Sea Res. Part I Oceanogr. Res. Papers 49 (10), 1705–1723.
Millero, F.J., Graham, T.B., Huang, F., Bustos-Serrano, H., Pierrot, D., 2006. Dissociation constants of carbonic acid in seawater as a function of salinity and temperature. Marine Chem. 100 (1), 80–94.
Mojica Prieto, F.J., Millero, F.J., 2002. The values of $pK_1 + pK_2$ for the dissociation of carbonic acid in seawater. Geochimica et Cosmochimica Acta 66 (14), 2529–2540.
Mopper, K., Stubbins, A., Ritchie, J.D., Bialk, H.M., Hatcher, P.G., 2007. Advanced instrumental approaches for characterization of marine dissolved organic matter: extraction techniques, mass spectrometry, and nuclear magnetic resonance spectroscopy. Chem. Rev. 107 (2), 419–442.
Müller, J.D., Rehder, G., 2018. Metrology of pH measurements in Brackish Waters—Part 2: experimental characterization of purified meta-Cresol purple for spectrophotometric pH_T measurements. Front. Marine Sci. 5 (177).
Orr, J.C., Epitalon, J.-M., Dickson, A.G., Gattuso, J.-P., 2018. Routine uncertainty propagation for the marine carbon dioxide system. Marine Chem. 207, 84–107.
Park, P.K., 1969. Oceanic CO_2 system: An evaluation of ten methods of investigation. Limnol. Oceanogr. 14 (2), 179–186.
Patsavas, M.C., Byrne, R.H., Wanninkhof, R., Feely, R.A., Cai, W.-J., 2015. Internal consistency of marine carbonate system measurements and assessments of aragonite saturation state: Insights from two U.S. coastal cruises. Marine Chem. 176, 9–20.
Roy, R.N., et al., 1993. The dissociation constants of carbonic acid in seawater at salinities 5 to 45 and temperatures 0 to 45°C. Marine Chem. 44 (2), 249–267.
Seidel, M.P., DeGrandpre, M.D., Dickson, A.G., 2008. A sensor for in situ indicator-based measurements of seawater pH. Marine Chem. 109 (1), 18–28.
Skirrow, G., 1975. The dissolved gases–carbon dioxide. In: Riley, J.P., Skirrow, G. (Eds.), Chemical Oceanography. Academic Press, New York, pp. 1–192.
Takahashi, T., et al., 1970. A carbonate chemistry profile at the 1969 Geosecs Intercalibration Station in the eastern Pacific Ocean. J. Geophys. Res. 75 (36), 7648–7666.
Uppström, L.R., 1974. The boron/chlorinity ratio of deep-sea water from the Pacific Ocean. Deep Sea Res. Oceanogr. Abstr. 21 (2), 161–162.
Walz, P.M., Takeshita, Y., Coletti, L.J., Jannasch, H.W., Johnson, K.S., 2018. Comparison of spectrophotometric and ISFET-potentiometric pH in seawater. In: Ocean Sciences Meeting, Portland, Oregon.
Wanninkhof, R., Lewis, E., Feely, R.A., Millero, F.J., 1999. The optimal carbonate dissociation constants for determining surface water pCO_2 from alkalinity and total inorganic carbon. Marine Chem. 65 (3), 291–301.
Williams, N.L., et al., 2017. Calculating surface ocean pCO_2 from biogeochemical Argo floats equipped with pH: An uncertainty analysis. Global Biogeochem. Cycles 31 (3), 591–604.
Yang, B., Byrne, R.H., Lindemuth, M., 2015. Contributions of organic alkalinity to total alkalinity in coastal waters: A spectrophotometric approach. Marine Chem. 176, 199–207.
Zigah, P.K., et al., 2017. Allochthonous sources and dynamic cycling of ocean dissolved organic carbon revealed by carbon isotopes. Geophys. Res. Lett. 44 (5), 2407–2415.

2.7 Acknowledgments

Chapter 2, in full, is a reprint of the material as it appears in Marine Chemistry. Fong, M.B. and Dickson, A.G., 2019. Insights from GO-SHIP hydrography data into the thermodynamic consistency of CO_2 system measurements in seawater. Marine Chemistry, 211: 52-63. The dissertation author was the primary investigator and author of this paper.

Chapter 3

Full spectrum methods for quality control and calibration of spectrophotometric pH measurements

3.1 Abstract

We developed methods for using full spectral data for the quality control of spectrophotometric pH measurements. These methods, which are most easily implemented on diode array spectrophotometers, are useful for identifying and potentially correcting instrumental contributions to pH measurement error. Spectrophotometric pH determination typically uses the ratio of the absorbances measured at two wavelengths to calculate pH. However, the use of full spectra for pH determination has been proposed as an alternative method that can potentially offer better precision. We show that the use of the Full Spectrum Method together with the conventional Ratio Method may provide useful quality control information, but broader use of the Full Spectrum Method is limited by the need for reliable data on the indicator species absorption spectra. Furthermore, the Full Spectrum Method is more sensitive to absorbance errors, which outweighs the benefit of improved precision. We developed methods based on Principal Components Analysis of datasets of spectra to identify systematic errors in absorbance that can affect pH measurement quality. We found that an absorbance-dependent absorbance error can develop from the degradation of the deuterium lamp on the spectrophotometer and contribute to the uncertainty of dye perturbation corrections. These absorbance errors, along with other contributions related to the spectrophotometer performance, can also lead to errors in determining the indicator properties. However, the uncertainty from these instrumental contributions could not entirely explain the discrepancies we found between the current published properties of purified m-cresol purple

(mCP) and our characterization of a batch of this indicator. These results suggest a need to re-evaluate the characterization of purified mCP and quantify its uncertainty.

3.2 Introduction

Spectrophotometric pH measurement with sulfonephthalein indicator dyes is a widely used method of pH determination for seawater that has the advantage of being a relatively straightforward measurement capable of high repeatability (~0.0004 in pH, Clayton and Byrne, 1993). As large numbers of discrete seawater samples are collected for pH measurements in ocean monitoring efforts aimed at quantifying long-term anthropogenically driven changes in the seawater CO_2 system (e.g., Olsen et al., 2016), a widespread observation has emerged of there being a significant pH-dependent discrepancy between spectrophotometrically measured pH, using purified *m*-cresol purple (mCP), and pH calculated from dissolved inorganic carbon (C_T) and total alkalinity (A_T), thus calling into question the accuracy of pH measurements (Álvarez et al., 2020; Carter et al., 2018; Carter et al., 2013; Fong and Dickson, 2019; Williams et al., 2017). Although current evidence suggests that the apparent errors in spectrophotometric pH measurements are unlikely to be responsible for the observed discrepancies (Takeshita et al., 2020), there is still limited understanding of the true accuracy of pH measurements and how to best achieve quality control of spectrophotometric pH measurements.

A semi-automated implementation of the spectrophotometric pH method using a single beam diode array spectrophotometer has been developed to enable straightforward, reliable, and efficient pH measurements on discrete seawater samples (Carter et al., 2013). Diode array spectrophotometers also offer the advantage of quick acquisition of spectra over the full UV-visible range. However, the full spectral data collected on diode array instruments is underutilized in the standard spectrophotometric pH procedure, which calculates the sample pH

from the ratio of the absorbances at the two wavelengths corresponding to the absorbance maxima of the acid and base forms of the indicator dye.

The use of full spectra for spectrophotometric pH determination has been proposed as an alternative to the conventional absorbance ratio-based method due to the potential for better precision resulting from the improvement of the signal-to-noise ratio of measurements when using absorbances over a range of wavelengths (Ohline et al., 2007). However, this approach has not been widely used. One of the largest obstacles to its widespread use is, undoubtedly, the lack of information for the full absorption spectra for the acid-base forms of pure mCP (the most widely used indicator dye for seawater pH measurements), which is needed for the Full Spectrum calculation of pH. Furthermore, the uncertainty of the Full Spectrum Method has not been thoroughly evaluated.

Analyzing the full spectra collected during spectrophotometric pH measurements on diode array spectrophotometers may provide information relevant to quality control. In the Full Spectrum Method of calculating pH, the measured sample spectrum is modeled as a mixture of the spectra of the acid-base forms of the indicator dye, and thus, the residuals (difference between modeled and measured spectrum) can provide indication of unexpected contributions to the sample spectrum (Husheer, 2001), such as additional absorbing chemical species, effects from the composition of the sample medium, and instrumental contributions. Principal Components Analysis on datasets of spectra can also be used to identify changes in measurement quality.

In this study, we used Ohline et al.'s Full Spectrum Method in combination with PCA-based methods for the quality control of spectrophotometric pH measurements. These quality control approaches, which can be replicated by users of diode array spectrophotometers, were

applied to datasets of spectra collected over varying timescales to demonstrate their utility in identifying changes in spectrophotometer performance. In addition, we characterized the indicator species absorption spectra of a batch of purified mCP and used these data to evaluate how the spectrophotometer performance contributes to uncertainty in the Full Spectrum and Ratio Methods of spectrophotometric pH determination and in the characterization of the dye properties.

3.3 Materials and Procedures

3.3.1 Theory

Spectrophotometric pH determination is based on the measurement of the absorption spectrum of solutions containing a pH-sensitive indicator dye (Byrne and Breland, 1989; Clayton and Byrne, 1993; Zhang and Byrne, 1996). Sulfonephthalein indicators, such as cresol red, *m*-cresol purple, and thymol blue, are diprotic acids that exist predominantly in two forms in the pH range of most natural waters (i.e., $HI^- \rightleftarrows H^+ + I^{2-}$), and thus the pH of a solution containing dye can be determined from

$$\mathrm{pH} = \mathrm{p}K_a(\mathrm{HI}^-) + \log\left(\frac{[\mathrm{I}^{2-}]}{[\mathrm{HI}^-]}\right) \quad (1).$$

The concentration ratio ($[I^{2-}]/[HI^-]$) can be estimated from the dye solution spectrum, which is a composite of the absorbance contributions from the acidic and basic forms of the indicator. The absorbance (A) spectrum can be expressed in the form of the Beer-Lambert Law, where ℓ is the cell pathlength, λ is the wavelength, and ε is a molar absorption coefficient.

$$A_\lambda = \varepsilon_\lambda(\mathrm{HI}^-)\ell[\mathrm{HI}^-] + \varepsilon_\lambda(\mathrm{I}^{2-})\ell[\mathrm{I}^{2-}] \quad (2)$$

For *m*-cresol purple (mCP), the most widely used pH indicator in seawater, the fully deprotonated species (I^{2-}) and the singly protonated species (HI^-) have maximum absorbance at

578 and 434 nm respectively. The concentration ratio ($[I^{2-}]/[HI^{-}]$) is therefore commonly determined from the ratio of the absorbances at these two wavelengths (i.e., $R = A_{578}/A_{434}$) by rearrangement of **Eq. 2**.

$$\frac{[I^{2-}]}{[HI^{-}]} = \frac{R-\varepsilon_{578}(HI^{-})/\varepsilon_{434}(HI^{-})}{\varepsilon_{578}(I^{2-})/\varepsilon_{434}(HI^{-})-R\varepsilon_{434}(I^{2-})/\varepsilon_{434}(HI^{-})} = \frac{R-e_1}{e_2-Re_3} \quad (3)$$

Combining **Eq. 3** and **Eq. 1** gives

$$\mathrm{pH} = \mathrm{p}K_a(HI^{-}) + \log\left(\frac{R-e_1}{e_2-Re_3}\right) \quad (4)$$

The molar absorption coefficient ratios (e_1, e_2, and e_3) are determined from measurements in high and low pH solutions where nearly all the dye is present in a single form (i.e., I^{2-} or HI^{-}). As there is potentially greater uncertainty in determining the ratios e_2 and e_3, which are obtained from a combination of measurements in separate solutions with significantly different pH, Liu et al. (2011) use an equation of the following form:

$$\mathrm{pH} = -\log(K_a(HI^{-})\cdot e_2) + \log\left(\frac{R-e_1}{1-R\frac{e_3}{e_2}}\right) \quad (5)$$

where the ratios e_1 and e_3/e_2 can be determined from measurements in single solutions at low and high pH, respectively. Once e_1 and e_3/e_2 have been determined, the term $-\log(K_a(HI^{-})\cdot e_2)$ (i.e., $\mathrm{p}(K_a(HI^{-})\cdot e_2)$) can be determined from **Eq. 5** by measuring R in TRIS buffers which have been assigned a pH value. Similarly, $\mathrm{p}K_a(HI^{-})$ can be determined using **Eq. 4**. The terms e_1, e_3/e_2, and $\mathrm{p}(K_a(HI^{-})\cdot e_2)$ in **Eq. 5** have been determined in synthetic seawater solutions over a range of temperatures and salinities for purified mCP (Liu et al., 2011).

3.3.2 Determining pH from full spectra

Ohline et al. (2007) proposed an alternative method of determining pH by fitting the full spectrum of a sample solution containing dye to a linear combination of the pure HI^- and I^{2-} spectra of the indicator. An estimate of the ratio $[I^{2-}]/[HI^-]$ can be obtained from the least squares solution of the following Beer-Lambert-like equation:

$$\mathbf{A} = \mathbf{c}_{HI^-}\mathbf{S}_{HI^-} + \mathbf{c}_{I^{2-}}\mathbf{S}_{I^{2-}} \quad (6),$$

where $\mathbf{A}$ is a $n \times m$ matrix of n sample solution spectra measured over m wavelengths; $\mathbf{S}_{HI^-}$ and $\mathbf{S}_{I^{2-}}$ (both $1 \times m$ vectors) are the HI^- and I^{2-} species spectra of the indicator determined from measurements in low and high pH solutions as described earlier; and $\mathbf{c}_{HI^-}$ and $\mathbf{c}_{I^{2-}}$ (both $n \times 1$ vectors) are the relative contributions of the HI^- and I^{2-} species spectra to the overall sample spectra. The form of **Eq. 6** used in Ohline et al. (2007) also included an intercept term that corrects for uniform baseline changes in absorbance between the time the background spectrum is measured (sample solution without dye) and when the sample spectrum (sample solution with dye) is measured. However, we corrected for baseline changes by subtracting the average of the absorbances from 725 nm to 735 nm from $\mathbf{A}$ for consistency with the implementation of the baseline correction in the Ratio Method (Carter et al., 2013).

Strictly speaking, $\mathbf{c}_{HI^-}$ and $\mathbf{c}_{I^{2-}}$ are not concentrations, but as the ratios $\mathbf{c}_{I^{2-}} / \mathbf{c}_{HI^-}$ and $[I^{2-}]/[HI^-]$ are the same value, the pH can be calculated by using $\mathbf{c}_{I^{2-}} / \mathbf{c}_{HI^-}$ in **Eq. 1** provided that $pK_a(HI^-)$ is known. For the purposes of evaluating the performance of the Ratio and Full Spectrum Methods, any arbitrary value for $pK_a(HI^-)$ can be used. Unless otherwise stated, the pH values reported for the various measurements in this study were calculated using a value of 7.9913 for the $pK_a(HI^-)$ of mCP at a salinity of 35 and a temperature of 25°C based on

unpublished measurements in 2012 from our laboratory on a batch of purified mCP. We also determined the the pK_a(HI^-) and $p(K_a(HI^-)\cdot e_2)$ for the batch of purified mCP used in this study for the purposes of evaluating their uncertainty.

In addition to estimates of the sample pH, the residuals of the sample spectra can be calculated from the difference between the predicted spectra (calculated using the values of $\mathbf{c}_{HI^-}$, and $\mathbf{c}_{I^{2-}}$ obtained from the least squares solution of **Eq. 6**) and the measured spectra.

3.3.3 Principal Components Analysis of spectra

In this study, we explore the application of Principal Components Analysis of spectra for the quality control of spectrophotometric pH measurements. A brief overview of PCA and our approach is provided here. Principal Components Analysis (Pearson, 1901) can be used to decompose a data matrix into a set of orthogonal basis vectors (i.e., the eigenvectors or principal components) and corresponding "scores" (position of the data along each eigenvector) which model the statistically significant components of variance in a dataset as well as random measurement error. A potential benefit of PCA is that the effects of random errors on a dataset can be minimized by discarding the minor principal components associated with noise (Gemperline, 2006; Kramer, 1998). PCA has been applied to studies of the speciation of sulfonephthalein indicators to achieve noise reduction in spectra (Shimada and Hasegawa, 2017; Shimada et al., 2019).

The principal components and scores for a set of spectra can be computed from singular value decomposition (SVD) of the matrix **A** in **Eq. 6** (Gemperline, 2006) and used to construct the following principal components model of the data:

$$\mathbf{A} = \mathbf{t}_1\mathbf{v}_1^{\mathrm{T}} + \mathbf{t}_2\mathbf{v}_2^{\mathrm{T}} + \ldots + \mathbf{t}_m\mathbf{v}_m^{\mathrm{T}} \tag{7}$$

where the matrix **A** is modeled as a sum of the outer products of the eigenvectors (or the "loading spectra") $\mathbf{v}_i^{\mathrm{T}}$ ($1\times m$) and the scores vector $\mathbf{t}_i$ ($n\times 1$). For spectra measured over m wavelengths, m principal components are required to fully represent both the significant components of variation and random errors. The principal components are ordered in **Eq. 7** such that each successive component explains increasingly less variance in the dataset.

The m principal components in **Eq. 7** can be separated into true basis factors (which explain significant components of variation) and noise factors either by examining the eigenvalues or singular values obtained from SVD (i.e., determining the mathematical rank of the matrix **A**), which are related to the variance explained by a principal component, or from *a priori* knowledge about the data. An example PCA is shown in **Figure 3.1.** for a set of simulated mCP spectra from pH 7-8 with a normally distributed error in the absorbance (*std. dev.* = 0.00031) added at each wavelength of each spectrum. The set of spectra in **Figure 3.1a** was mean-centered prior to SVD (i.e., $\mathbf{A}^{\dagger} = \mathbf{A} - \overline{\mathbf{A}}$ is used in **Eq. 7**.). Thus, the first principal component (99.99991% of the total variance) shown in **Figure 3.1b** captures the variation of the spectra relative to the mean spectrum in the dataset. As the variations are due to changes in pH, they are largest at 434 and 578 nm. By reconstructing the spectra using only the first principal component, the noise (contained in the other principal components) is discarded. To transform $\mathbf{A}^{\dagger}$ back into the original data matrix **A**, the mean is simply added to **Eq. 7** (i.e., $\mathbf{A} = \mathbf{t}_1\mathbf{v}_1^{\mathrm{T}} + \overline{\mathbf{A}}$). As we will show later, PCA can be a powerful tool for identifying systematic errors in the measurement of the spectra, which will appear as principal components unrelated to the dye's response to pH or noise.

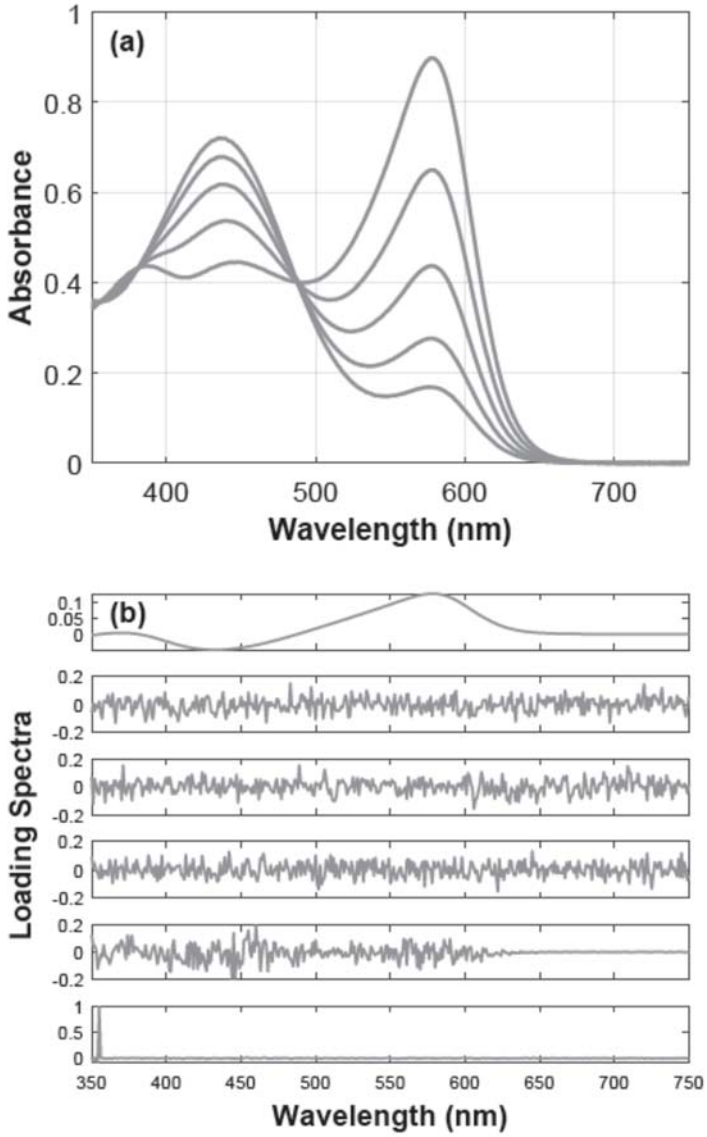

Figure 3.1. (a) Set of simulated spectra at a constant total concentration of mCP, representing pH values from 7 to 8. A normally distributed error with a standard deviation of 0.00031 was added to the absorbances at each wavelength. (b) Loading spectra for the first six principal components (arranged from top to bottom) of the set of spectra in (a) after mean-centering.

Another way in which we use PCA is for the correction of orthogonal errors in spectra by projecting them onto the orthonormal basis vectors that span the space of a reference dataset (see **Figure 3.2, Eq. 8, Eq. 9**).

$$\mathbf{t}_{proj} = \mathbf{aV} \tag{8}$$

$$\mathbf{a}_{proj} = \mathbf{t}_{proj}\mathbf{V}^{T} \tag{9}$$

As we will discuss later, this approach was used to correct for the effect of the background medium on the I^{2-} spectrum of mCP and systematic errors in the I^{2-} spectrum due to changes in the spectrophotometer lamp performance.

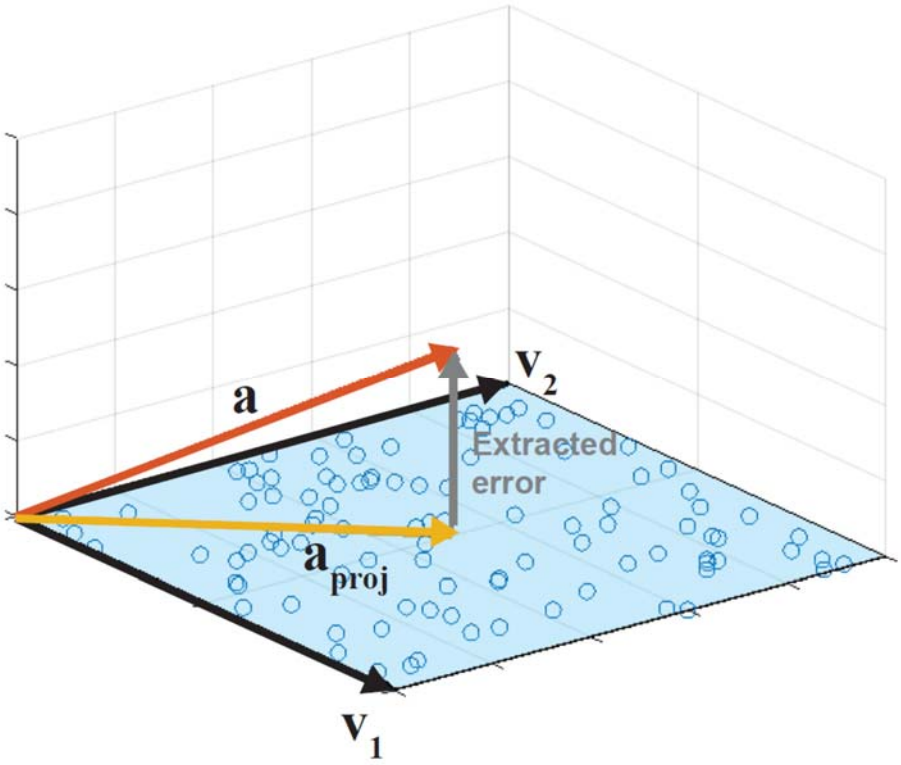

Figure 3.2. Schematic illustrating the correction of orthogonal errors in a spectrum (**a**) by projecting it onto the space spanned by the orthogonal basis vectors of the reference dataset. The points on the figure represent the position (or "scores") of the spectra in the dataset along each basis vector.

Our use of orthogonal projection is similar to the idea of Orthogonal Signal Correction (Wold et al., 1998), which removes variation in the spectral matrix **A** that is orthogonal to (and therefore not correlated with) the response variable (i.e., pH in this discussion). However, the use of OSC requires that the response variable be determined independently of the spectral measurements (Workman, 2018) Our orthogonal projection approach removes errors

orthogonal to the basis vectors that describe the pH variation of the reference dataset without having to independently determine pH. Similar methods have been applied to the calibration transfer of near-infrared spectra (Poerio and Brown, 2018).

3.3.4 Characterization of *m*-cresol purple

We characterized one lot of purified mCP (Lot FB4, obtained from Robert H. Byrne, U. of Southern Florida) at a temperature of 25°C and ionic strength of ~0.69 mol kg^{-1} to obtain the absorption spectra of each of the indicator acid-base forms required to evaluate the Full Spectrum Method.

The mCP species spectra were obtained by measuring the spectra in solutions at pH values selected so that a single form of dye dominates (pH ≈ 0, 4.5, and 12 for the H_2I, HI^-, and I^{2-} species spectra, respectively). These solutions were prepared from a ~2.5 mmol kg^{-1} mCP stock solution. The mCP stock solutions were made by sonicating the appropriate amount of mCP powder in a 0.67 mol kg^{-1} NaCl solution with ~0.025 mol kg^{-1} NaOH (added from a standardized NaOH solution) for 1 hour and then adjusting the pH of the solution, measured with a glass electrode calibrated on the NBS scale, to between 7 and 8 with ~1 mol kg^{-1} HCl. Measurements of the spectra were made with a 10 cm cell on an Agilent 8453 single beam spectrophotometer, using the semi-automated system of Carter et al. (2013). The spectrum of the I^{2-} species was measured in a ~0.67 mol kg^{-1} NaCl solution with ~0.02 mol kg^{-1} NaOH. The HI^- species spectrum was measured in a buffered solution (pH~4.5) consisting of ~0.02 mol kg^{-1} sodium acetate, ~0.01 mol kg^{-1} HCl, and ~0.67 mol kg^{-1} NaCl. The pH of the acetate buffer was measured with a glass electrode calibrated on the free hydrogen ion scale by titrating a ~0.7 mol kg^{-1} NaCl solution with HCl. The H_2I species spectrum was measured in a ~1 mol kg^{-1} HCl solution standardized by titration against borax (Vogel, 1961).

To obtain the pure spectra of the individual mCP species, the spectra measured in each of three solutions must be corrected for minor contributions from the other species. We used an approach similar to that outlined in Ohline et al. (2007) and Husheer (2001), based on the application of matrix algebra to the Beer-Lambert Law, to correct the set of spectra for the presence of minor species (**Eq. 10-12**).

$$\mathbf{A}^* = [\mathbf{a}^*_{\mathrm{HCl}} \quad \mathbf{a}^*_{\mathrm{Ac}} \quad \mathbf{a}^*_{\mathrm{NaOH}}] \tag{10}$$

$$\mathbf{C} = \begin{bmatrix} f_{\mathrm{H_2I}}(\mathrm{HCl}) & f_{\mathrm{H_2I}}(\mathrm{Ac}) & f_{\mathrm{H_2I}}(\mathrm{NaOH}) \\ f_{\mathrm{HI^-}}(\mathrm{HCl}) & f_{\mathrm{HI^-}}(\mathrm{Ac}) & f_{\mathrm{HI^-}}(\mathrm{NaOH}) \\ f_{\mathrm{I^{2-}}}(\mathrm{HCl}) & f_{\mathrm{I^{2-}}}(\mathrm{Ac}) & f_{\mathrm{I^{2-}}}(\mathrm{NaOH}) \end{bmatrix} \tag{11}$$

$$\mathbf{S} = \mathbf{A}^* \cdot \mathbf{C}^{-1} \tag{12}$$

In **Eq. 10**, $\mathbf{A}^*$ is a m ×3 matrix containing the spectra measured in the HCl, acetate buffer, and NaOH solutions, normalized by the absorbances at the appropriate isosbestic wavelengths (488.1 nm for the NaOH and acetate buffer spectra and 480.1 nm for the HCl spectrum—see Liu et al., 2011) to minimize concentration differences. We normalized the spectra by the isosbestic absorbances rather than dye concentration as we experienced difficulty fully dissolving the dye powder. The pure species spectra (**S**) are obtained by multiplying $\mathbf{A}^*$ by the inverse of **C**, a 3×3 matrix containing the relative concentrations of the three mCP species in each solution (**Eq. 11-12**). The relative concentrations of the mCP species sum to 1 in each solution, and thus, they can be calculated from the two acid dissociation constants for mCP and the pH of each solution (**Eq. 13**). We use a value of 1.487 for pK_a(H_2I) (Clayton and Byrne, 1993) and 7.9913 for pK_a(HI^-), based on our unpublished measurements on a different batch of purified mCP. The pK_a values used in **Eq. 13** should be considered estimated values and have a

small impact on the uncertainty of the estimated dye properties and on pH, as we will show later.

$$
\begin{aligned}
f_{\mathrm{H_2I}} &= \frac{1}{1+\frac{K_a(\mathrm{H_2I})}{[\mathrm{H^+}]}+\frac{K_a(\mathrm{H_2I})K_a(\mathrm{HI^-})}{[\mathrm{H^+}]}} \\
f_{\mathrm{HI^-}} &= \frac{1}{1+\frac{[\mathrm{H^+}]}{K_a(\mathrm{H_2I})}+\frac{K_a(\mathrm{HI^-})}{[\mathrm{H^+}]}} \\
f_{\mathrm{I^{2-}}} &= \frac{1}{1+\frac{[\mathrm{H^+}]^2}{K_a(\mathrm{H_2I})K_a(\mathrm{HI^-})}+\frac{[\mathrm{H^+}]}{K_a(\mathrm{HI^-})}}
\end{aligned}
\tag{13}
$$

The pure species spectra for the FB4 dye is shown in **Figure 3.3a.** The ratios e_1, e_2, and e_3 for each dye were calculated from their respective dye species spectra. We also determined the pK_a(HI^-) and $p(K_a(HI^-)\cdot e_2)$ for the FB4 dye from measurements in a single bottle of equimolar TRIS-TRIS · H^+ buffer standard in synthetic seawater (Batch T32) prepared according to DelValls and Dickson (1998), so that the full set of dye properties (i.e., e_1, e_2, e_3, pK_a(HI^-) and $p(K_a(HI^-)\cdot e_2)$) can be compared to previously published values for purified mCP (**Table 3.1**,**Table 3.2**). We obtained the values of e_2, e_3, and pK_a(HI^-) for the dye characterized by Liu et al. (2011) from personal communication with X. Liu as they did not explicitly report these values in their study.

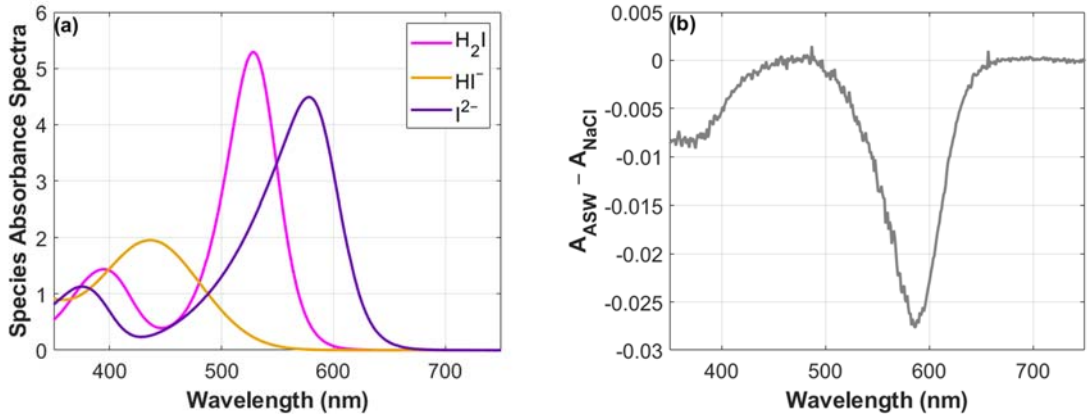

Figure 3.3. (a) Normalized absorption spectra for the three species of *m*-cresol purple, obtained from measurements in NaCl solutions with purified mCP and corrected for the contribution of minor species. The HI^- and I^{2-} spectra are normalized to an absorbance of 1at 488.1 nm, and the H_2I spectrum is normalized to an absorbance of 1 at 480.1 nm of the HI^- spectrum. (b) Difference between the I^{2-} spectrum (normalized at 488.1 nm) for purified mCP measured in a modified synthetic seawater medium and in NaCl.

Table 3.1. Absorption coefficient ratios for the FB4 batch of purified mCP and the batch of purified mCP characterized by Liu et al. (2011). The ratios for the FB4 dye was determined in NaCl solutions. The adjusted values refer to corrections to the spectra to account for differences in the absorption behavior of the dye in NaCl and seawater-like media and other potential systematic errors, while the unadjusted values have only been corrected for the contribution of minor species (see text).

	FB4		
	Unadjusted	**Adjusted**	**Liu**
e_1	0.005965	0.00667	0.005707
e_2	2.2972	2.2646	2.2267
e_3	0.12722	0.12436	0.12646
e_3/e_2	0.05538	0.05492	0.05678

Table 3.2. Values of pK_a(HI$^-$) and $p(K_a(HI^-)\cdot e_2)$ determined from measurements of a TRIS buffer standard. pK_a(HI$^-$) was calculated using **Eq. 1** for the Full Spectrum Method and **Eq. 4** for the Ratio Method. $p(K_a(HI^-)\cdot e_2)$ was calculated using **Eq. 5** for the Ratio Method and the value of pK_a(HI$^-$) calculated using the Full Spectrum Method combined with e_2. The calculations were done using either the absorption coefficient ratios and species absorption spectra for the FB4 dye or the ratios of Liu et al. (2011). The adjusted FB4 value refers to corrections to the spectra to account for differences in the absorption behavior of the dye in NaCl and seawater-like media and other potential systematic errors (see text). The values for the batch of purified mCP characterized by Liu et al. (2011) are also included for comparison.

	pK_a(HI$^-$)		$p(K_a(HI^-)\cdot e_2)$	
Dye properties used	**Ratio**	**Full Spectrum**	**Ratio**	**Full Spectrum**
FB4	8.0154	8.0148	7.6542	7.6536
Adjusted FB4	8.0099	8.0099	7.6549	7.6550
Liu	8.0001		7.6524	
Reference value (Liu et al., 2011)	7.9963		7.6486	

Liu et al. (2011) reported that the ratio e_3/e_2, determined from measurements in NaOH solutions, has a slight dependence on the ionic composition of the background medium. Therefore, they used a modified synthetic seawater medium—replacing Na_2SO_4 with NaCl and $MgCl_2$ with $CaCl_2$ to avoid precipitation of the magnesium and sulfur salts at high pH—as the background in their NaOH solutions. We evaluated the effect of the background medium on the I^{2-} spectrum and the implications for the Full Spectrum and Ratio Methods with additional measurements of the FB4 I^{2-} spectrum in NaOH and modified synthetic seawater solutions of the same composition as those used by Liu et al. (2011). These measurements were compared to those made in NaOH/NaCl solutions on the same day (**Figure 3.3b**).

3.3.5 Evaluation of uncertainty in indicator properties

The performance of the spectrophotometer used to characterize the indicator can contribute to uncertainty in the dye properties and subsequently in determining pH. We performed Monte Carlo simulations to evaluate the uncertainty in our values for the absorption coefficient ratios, the pK_a(HI$^-$), and the $p(K_a(HI^-)\cdot e_2)$ of the FB4 dye. The spectrophotometer-

related uncertainty contributions listed in **Table 3.3** and **Table 3.4** include a systematic error we identified in our measurements of the I^{2-} spectrum and the normalization of the spectra during the data processing performed to correct for minor species. As the minor species correction involves calculating the relative proportions of the dye species in each solution (**Eq. 13**), information about the pH of the three solutions used to measure the species spectra and estimates of pK_a(HI^-) and pK_a(H_2I) are required. Errors in these parameters also contribute to the uncertainty in determining the dye properties but are relatively minor contributions (**Table 3.3**, **Table 3.4**).

Table 3.3. Estimates of the contributions to uncertainty in e_1, e_2, e_3, e_3/e_2, and $\log(e_2)$ from the data processing to obtain the mCP species absorption spectra and from systematic error in the I^{2-} spectrum. The uncertainties, expressed as standard deviations, were evaluated for the FB4 batch of purified mCP.

			Resulting uncertainty in				
Source	***u***	**Distribution**	e_1	e_2	e_3	e_3/e_2	$\log(e_2)$
Isosbestic wavelength (480.1 nm)	0.26 nm	Rectangular	5.1E-06	1.4E-05	7.5E-07	2.4E-10	2.6E-06
Isosbestic wavelength (488.1 nm)	0.26 nm	Rectangular	7.9E-06	0.031	1.7E-03	5.8E-07	5.8E-03
pH of HCl solution	0.0019	Normal	4.3E-08	2.9E-07	1.6E-08	5.5E-12	5.5E-08
pH of acetate solution	0.039	Combined	2.6E-05	1.4E-04	7.8E-06	2.5E-09	2.7E-05
pH of NaOH solution	0.0013	Rectangular	1.7E-10	6.5E-07	2.5E-07	1.2E-07	1.2E-07
pK_a (H_2I)	0.2[a]	Normal	0.00016	1.1E-03	5.9E-05	2.0E-08	2.0E-04
pK_a (HI^-)	0.0067[b]	Normal	8.9E-06	1.1E-05	8.5E-07	6.4E-07	2.1E-06
Systematic error in I^{2-} spectrum	*see text*	Normal	3.5E-06	1.4E-02	0.0012	0.00067	0.0026
Combined std. uncertainty (k=1)			0.00016	0.034	0.0021	0.00067	0.0064

[a] Assumed value
[b] Standard deviation of values from Degrandpre et al. (2014), Liu (pers. comm.), and our unpublished data.

Table 3.4. Estimates of the contributions to uncertainty in pK_a(HI^-) and $p(K_a(HI^-)\cdot e_2)$ from the data processing to obtain the mCP species absorption spectra and from systematic error in the I^{2-} spectrum. The uncertainties, expressed as standard deviations, were evaluated for the FB4 dye using both the Ratio and Full Spectrum Methods.

	pK_a(HI^-)		$p(K_a(HI^-)\cdot e_2)$	
Source	**Ratio**	**Full Spectrum**	**Ratio**	**Full Spectrum**
Isosbestic wavelength (480.1 nm)	1.6E-06	7.5E-07	9.2E-07	1.8E-06
Isosbestic wavelength (488.1 nm)	0.0058	0.0058	2.1E-06	2.1E-06
pH of HCl solution	5.9E-08	6.6E-08	7.3E-09	1.4E-08
pH of acetate solution	2.2E-05	2.6E-05	4.8E-06	7.9E-07
pH of NaOH solution	2.7E-07	2.7E-07	1.5E-07	1.5E-07
pK_a (H_2I)	0.00024	0.00027	3.0E-05	5.9E-05
pK_a (HI^-)	4.7E-06	4.7E-06	2.5E-06	2.5E-06
Systematic error in I^{2-} spectrum	0.0032	0.0036	0.0008	0.0013
Combined std. uncertainty (k=1)	0.0066	0.0069	0.0008	0.0013

Our approach of normalizing the spectra by the absorbances at the isosbestic points is sensitive to small errors in the spectrophotometer wavelength calibration. Specifically, the estimate of the absorbance at the HI^-/I^{2-} isosbestic point is especially sensitive to wavelength errors, as this point lies in the strongly sloping regions of the HI^- and I^{2-} spectra (**Figure 3.3**). We calculated the isosbestic absorbance from a weighted average of the absorbances at two adjacent diodes. Assuming that the wavelength errors for our spectrophotometer are evenly distributed between ±0.5 nm (based on the manufacturer specifications), the resulting uncertainty in estimating the isosbestic wavelength from a weighted average of two diodes is 0.26 nm.

Uncertainties for the pH of three solutions were estimated from the precision of the borax titrations (for the HCl solution), the manufacturer certification for the concentration of the standardized NaOH solution, and the uncertainty of our glass electrode pH measurements (for the acetate solution). The uncertainty of the glass electrode pH measurements included

contributions from the precision of the E^0 values of the electrode (from titrating NaCl solutions) and the Nernstian behavior of the electrode (>97%, according to the specifications for the Metrohm Ecotrode Plus, which we assume implies Nernst slopes evenly distributed between 0.97 to 1).

For the uncertainty in the estimates of pK_a(HI^-), we use the standard deviation of the values from DeGrandpre et al. (2014), Liu (pers. comm.), and our own unpublished measurements on a different batch of purified mCP. The uncertainty in pK_a(H_2I) is unknown, so we assume a value of 0.2.

The uncertainties in **Table 3.3** were used to generate 1,000 simulated values for each contribution (i.e., the isosbestic wavelengths, the pH of the three solutions, and the pK_a's) by adding errors drawn from the appropriate probability distributions to a reference value. Each simulated value was used, while holding the other parameters constant at their reference values, to recalculate the corrected dye species spectra, the absorption coefficient ratios, the pK_a(HI^-), and $p(K_a(HI^-) \cdot e_2)$. The standard uncertainty (u) in these dye properties resulting from each contribution was calculated from the standard deviation of the 1,000 simulated values for contributions with normal error distributions or from the appropriate formula for contributions with rectangular (even) error distributions (**Eq. 14**, where x refers to the simulated values).

$$u(x) = \frac{0.5 \cdot (\max(x) - \min(x))}{\sqrt{3}} \tag{14}$$

The contribution from the pH of the acetate solution involved a combination of a normal (from E_0) and rectangular error distribution (from the Nernst slope). These two contributions were assessed separately and then added in quadrature to estimate the uncertainty contribution from the pH of the acetate solution.

We found systematic drifts from repeated measurements (over a period of eight months) of the spectra in NaOH solutions (see later discussion). The uncertainty contribution from systematic errors in measuring the I^{2-} spectrum was estimated from the standard deviation of the various dye properties recalculated using the ten different sets of I^{2-} spectrum measurements that we made over eight months.

Finally, the combined standard uncertainty in the dye properties (**Table 3.3**, **Table 3.4**) was calculated by adding the various uncertainty contributions in quadrature.

3.3.6 Validation measurements

A set of validation samples measured with the FB4 lot of mCP was used to evaluate our ability to model spectra in the seawater pH range as mixtures of the HI^- and I^{2-} spectra (**Eq. 6**) we determined for the dye. The validation samples consisted of an unbuffered solution ~0.02 mol kg^{-1} sodium acetate in NaCl background (~0.7 mol kg^{-1} ionic strength, adjusted to pH~7.11 with HCl), a buffered solution of ~0.08 mol kg^{-1} 3-[4-(2-Hydroxyethyl)-1-piperazinyl]propanesulfonic acid (EPPS) and ~0.042 mol kg^{-1} NaOH in NaCl background (~0.7 mol kg^{-1} ionic strength, pH~7.79), a CO_2-in-seawater Reference Material (CRM 186, pH~7.86) prepared in the Dickson Laboratory at the Scripps Institution of Oceanography (Dickson 2010), and a 0.04 mol kg^{-1} approximately equimolal TRIS/TRIS-HCl buffer in synthetic seawater (S=35, pH~8.07). Two subsamples were drawn from each individual bottle for the pH measurements. For each bottle, the pH was measured with two separate additions of the dye solution. The validation samples were measured between November 2019 and January 2020, within the same timeframe as when the FB4 dye was characterized.

The pH values were calculated with both the Full Spectrum (**Eq. 6**) and Ratio Methods (**Eq. 4**) using a $pK_a(HI^-)$ of 7.9913 and the optical absorption properties appropriate to the dye.

The residuals of the spectra obtained from processing the data with the Full Spectrum Method were calculated (**Figure 3.4a**) to examine how well the sample spectra could be modeled from the dye species spectra and to evaluate the effect of the background medium (i.e., NaCl vs. seawater) on the Full Spectrum Method. The dye perturbation to the pH of the unbuffered solutions was corrected by extrapolating the pH values from two separate dye additions (at different dye amounts) in each individual bottle to zero isosbestic absorbance ($A_{iso} = 0$, representing zero dye concentration). For the buffered solutions, the replicate pH measurements were averaged.

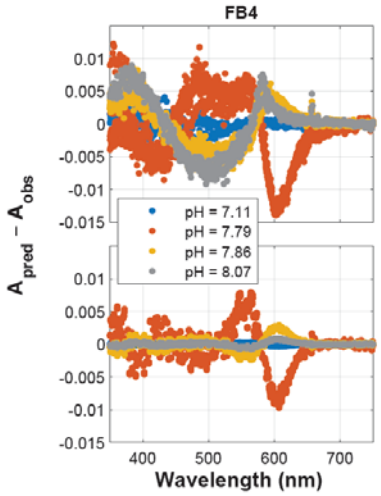

Figure 3.4. Residuals from fitting the validation sample spectra (see **Figure 3.5a**) to **Eq. 6** (top panel). The residuals were calculated using spectra normalized to $A_{iso} = 1$. The bottom panel shows the residuals calculated using the sample spectra reconstructed by retaining only the first two principal components in **Figure 3.5a** and the HI^- and I^{2-} spectra corrected for orthogonal errors relative to the validation sample datasets (**Figure 3.10**).

3.3.7 Adjustments to the dye species spectra

The residuals in **Figure 3.4** suggested that our full spectrum model (**Eq. 6**) was inadequate in fully explaining the variation of the validation sample spectra. Misfits with the model may be due to systematic errors in the dye species spectra (see later discussion). We found that adjustments to the dye species spectra were necessary to achieve agreement between

the measured and modeled spectra. These adjustments involved the orthogonal projection approach introduced earlier and are described in more detail below.

1. **Select spectra for the reference dataset**: Each set of validation sample spectra was chosen as the reference for correcting their corresponding dye species spectra. The EPPS buffer spectra were excluded as they appeared to contain an interference unique to those samples (**Figure 3.4**, pH 7.79). Including these spectra in the dataset would attract the primary principal components towards the anomalous observations and thus distort the results of PCA in **Step 2.**
2. **Reconstruct reference spectra using PCA**: The validation sample spectra were reconstructed from the first two principal components of the unit length-normalized (i.e., the original spectra were divided by their vector lengths.) and mean-centered data (**Eq. 7**). These components (**Figure 3.5**) represent the variation in the spectra due to differences in pH (PC 1) and the variability in the scaling of the spectra (PC 2), and thus discarding the other components removes the noise and error components from the spectra.
3. **Compute basis vectors:** The principal components were computed again using the set of reconstructed spectra, re-scaled to unit length and mean-centered. Rescaling the spectra introduces a third principal component, and thus the first three principal components were selected as the basis vectors.
4. **Orthogonal projection:** The HI^{-} and I^{2-} spectra for each dye were corrected for minor species (**Eq. 10-12**) and then normalized to unit length and centered relative to the means in **Step 3**. These normalized and mean-centered spectra were then projected onto the basis vectors (**Eq. 8-9**). The EPPS buffer spectra were also corrected in this way.

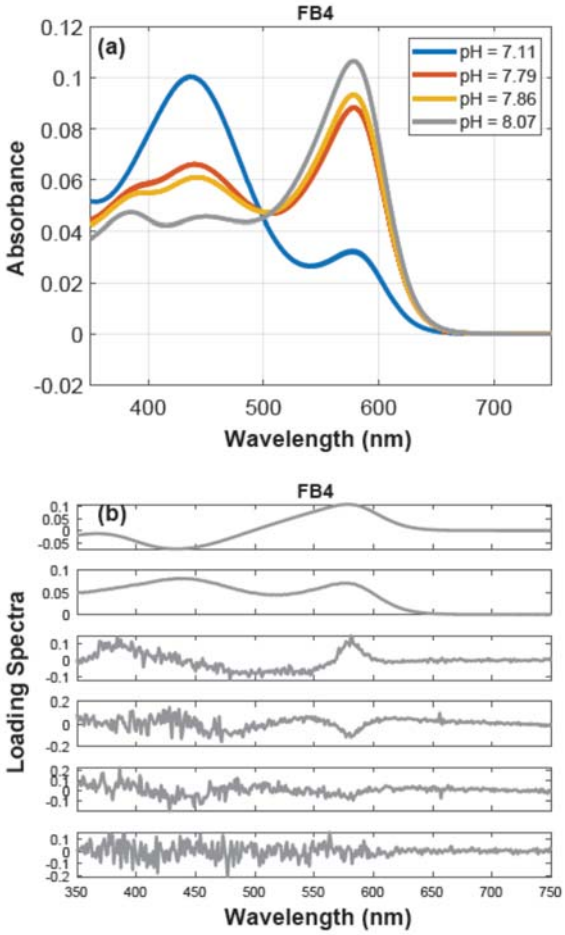

Figure 3.5. (**a**) Unit length-normalized spectra measured in NaCl and seawater solutions (pH 7.11 to 8.07) and (**b**) the loading spectra for the first six principal components of each dataset. Each dataset consists of 17 spectra, but the individual spectra may not be distinguishable due to the scaling. The spectra were mean-centered before computing the loading spectra.

The pH (calculated with the Full Spectrum and Ratio Methods) and the residuals from the Full Spectrum fitting were re-calculated using the adjusted dye species spectra and the adjusted validation sample spectra, normalized to $A_{iso} = 1$ (**Figure 3.4b**). In addition to the effect on pH, the adjustments also change the values of A_{iso} used for the dye perturbation corrections. Adjusted values of A_{iso} were calculated from the corrected spectra transformed to the same scale as the original spectra (i.e., reversing the mean-centering and multiplying the spectra by the vector lengths originally used to scale the spectra to unit length).

3.3.8 Measurements of the I^{2-} spectrum

Repeated measurements of the spectra in freshly prepared NaOH solutions were made over a period of eight months using 3 separate FB4 dye stock solutions. Significant drifts in the spectra were observed over this period, coincident with a steady decrease in the deuterium lamp intensities (**Figure 3.6**). The change in the spectra plotted in **Figure 3.6** were reported relative to measurements made on the date closest to when the set of lamps on the spectrophotometer were last replaced (May 2019). The reference date in **Figure 3.6** was Jun. 13, 2019. The differences in **Figure 3.6** were calculated using the average of multiple spectra (n=2-8) measured on each date, and the spectra were normalized to unit-length to minimize changes due to differences in the total dye concentration of the solutions.

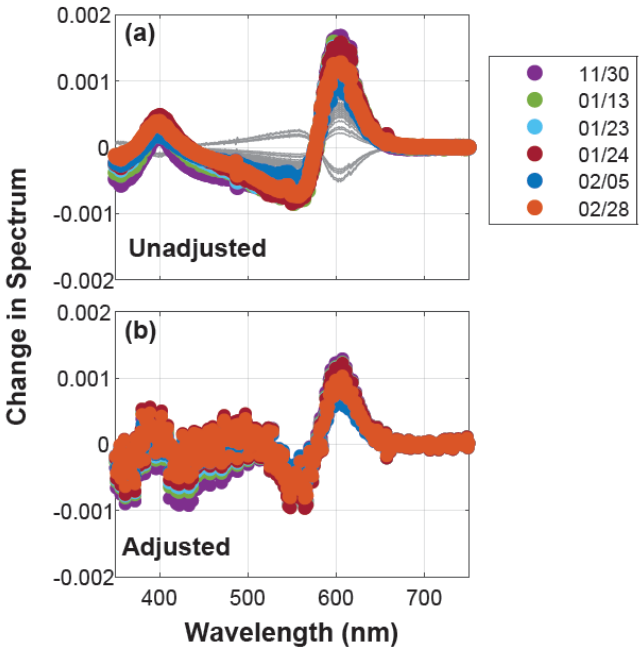

Figure 3.6. Change in the mCP spectrum measured in NaOH/NaCl solutions (pH~12) relative to the measurements made on a reference date (Jun. 13, 2019). Multiple measurements on each day were averaged to obtain a mean spectrum, and the differences were calculated using spectra normalized to unit length. Measurement dates for each dye are indicated in the legend. The gray curves in (a) show the simulated change in the spectrum resulting from the change in the wavelength of the deuterium emission peak around 486.0 nm between May 2019 to March 2020. The gray curves have been scaled upwards by a factor of 10 for visibility. The lower panels show the differences after the spectra (including the reference spectrum) have been projected onto a set of spectra representing measurements in solutions over a pH range of 7.1-8.1 (see text and also **Figure 3.4**).

To further evaluate the effect of lamp performance on the I^{2-} spectra measurements, measurements of the FB4 dye spectrum in NaOH solutions were made using two different spectrophotometers and different sets of deuterium and tungsten lamps over a period of three weeks between February to March 2020.

The first set of measurements on the main spectrophotometer (Feb. 28, 2020) used the same lamps that had been in use previously. Subsequent experiments on the main spectrophotometer tested different sets of lamps: (1) an unused set of lamps purchased in 2017 (Mar. 4, 2020), (2) a set of lamps taken from a second spectrophotometer (Mar. 5, 2020), and (3) a set of new lamps (Mar.19, 2020). Each set of measurements consisted of 6-17 spectra, which were averaged, normalized to unit length, and compared to the mean normalized spectrum obtained on a second spectrophotometer (on Mar. 3, 2020, see also **Figure 3.7**). The measurements are summarized in **Table 3.5**.

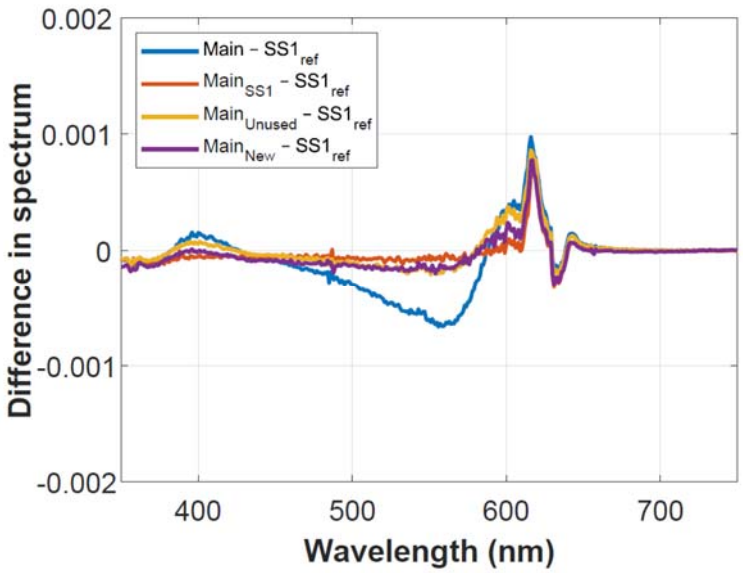

Figure 3.7. Difference between the FB4 mCP spectra measured in NaOH/NaCl solutions (pH~12) on the main spectrophotometer using different sets of lamps (indicated in the subscripts in the legend) relative to the measurements made on a second spectrophotometer (SS1). The measurements were made over the course of three weeks (between February to March 2020). Multiple measurements on each day were averaged to obtain a mean spectrum, and the differences were calculated using spectra normalized to unit length.

Table 3.5. Summary of measurements of the FB4 mCP spectra in NaOH/NaCl solutions (pH~12) on the main spectrophotometer using different sets of lamps and on a second spectrophotometer (SS1). The lamps were assigned scores on a linear scale from 0 (fail) to 5 (peak performance), based on the intensities in the 190-350 nm region for the deuterium lamp and the 950-1100 nm region for the tungsten lamp relative to the maximum intensities observed for all the lamps that have been used throughout the entire history of the main spectrophotometer.

Measurement date	Instrument	Lamp	Number of spectra	Deuterium Score	Tungsten Score
2/28/2020	Main	Main	17	2[a]	4[a]
3/3/2020	SS1	SS1	17	3[b]	5[b]
3/4/2020	Main	Unused	8	4	4
3/5/2020	Main	SS1	6	3	5
3/19/2020	Main	New	12	4	4

(a) Based on the intensities measured on 11/6/2019.
(b) Based on the intensities measured on 3/5/2019.

3.3.9 Measurements of seawater reference materials

We analyzed a dataset of measurements on a single batch of CO_2-in-seawater RM (CRM 186) over a period of five months to evaluate the use of PCA for quality control. A total

of 26 bottles of CRM 186 were measured between October 2019 to February 2020 using four separate FB4 dye stock solutions. Each bottle was measured with two to three separate dye additions using different dye amounts, and the pH values were extrapolated to $A_{iso} = 0$ to obtain the pH at zero dye concentration.

The spectra were normalized to unit length and mean-centered to compute the loading spectra and scores of the principal components of this dataset (**Figure 3.8**). In this dataset, the first principal component represents the variation in pH, while the other components represent various systematic and random errors. The components unrelated to the variation in pH were removed from the spectra to evaluate their effect on the pH values calculated with the Ratio and Full Spectrum Methods. For these adjustments, the principal components were re-computed excluding anomalous observations identified from the preliminary PCA on the complete dataset (**Figure 3.8**). The excluded measurements were later corrected by orthogonal projection onto the basis vectors constructed from the rest of the dataset.

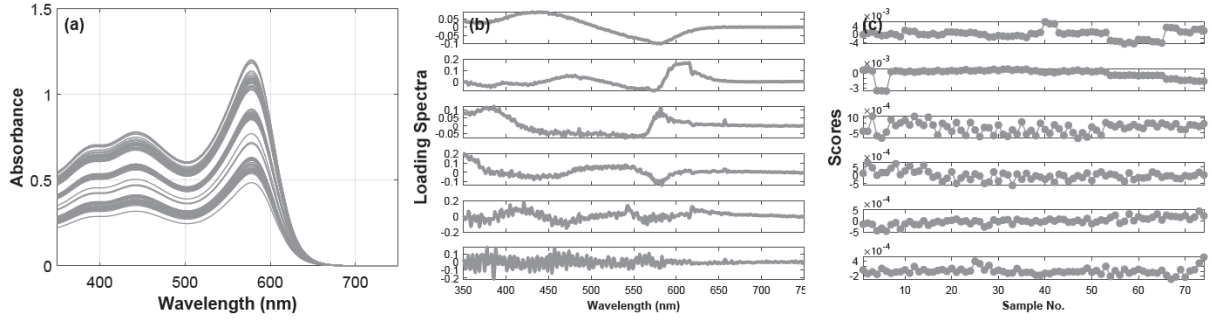

Figure 3.8. (a) Raw absorbance spectra of measurements of CRM Batch 186 between October 2019 to February 2020, using the FB4 purified mCP. The spectra were normalized to unit length and mean-centered prior to computing the (b) loading spectra and (c) scores for the first six principal components.

3.3.10 Phosphate buffer measurements

Phosphate buffers (pH~7.5), chosen for their stable pH and low temperature sensitivity, were measured on three different spectrophotometers to evaluate instrument differences. The spectra from these measurements (normalized to unit length and mean-centered) were analyzed

by PCA to further diagnose differences in instrument behavior and assess their implications on pH analysis.

A total of three separate experiments were conducted over the course of a year. In the first experiment in October 2019, a single large bottle containing 7 L of phosphate buffer in NaCl background was prepared with a total phosphate concentration of ~0.1 mol kg^{-1} and an ionic strength of ~0.7 mol kg^{-1}. The buffer was prepared from a mixture of NaH_2PO_4 and Na_2HPO_4, and the pH was adjusted to approximately 7.5 with a 1N NaOH solution. Subsamples were drawn from this single bottle of phosphate buffer for measurements on the different spectrophotometers. The same FB4 dye stock solution, 10 cm cell, and other components of the measurement system (besides the spectrophotometers) were used for all the measurements. A total of 24-46 measurements were made on each spectrophotometer using different dye amounts, and all measurements were made on a single day.

The phosphate buffer pH measured on the main spectrophotometer in the first set of experiments showed a strong dependence on the dye amount, which was suspected to be caused by an absorbance-dependent error in absorbance resulting from the degrading deuterium lamp (See later discussion.). In October 2020, the experiment was repeated using a new batch of phosphate buffer solution and a new FB4 dye stock solution. A total of 21 measurements were made on each spectrophotometer using different dye amounts. However, the set of lamps used on the main spectrophotometer in the October 2019 experiment had failed the instrument self-diagnostics criteria, so these lamps were replaced. Lamp intensities and the spectra of NIST SRM 930D glass filters were measured on each spectrophotometer prior to the phosphate buffer measurements. Finally, a third experiment, consisting of the same measurements, was

conducted on the following day using a separate bottle of phosphate buffer and a single set of lamps on all three spectrophotometers.

3.3.11 Comparison of the Ratio and Full Spectrum Methods

The contribution of random and systematic errors to the uncertainty in estimating pH with the Ratio and Full Spectrum Methods was evaluated using Monte Carlo simulations with a synthetic dataset. Five synthetic spectra representing pH values from 7 to 8 were created from linear combinations of the FB4 HI^- and I^{2-} spectra. To evaluate the effects of random errors, each spectrum was simulated 1,000 times with a normally-distributed error with a standard deviation of 0.00031 (an estimate of the repeatability of the absorbance measurements of the Agilent 8453; Carter et al., 2013) added to the absorbance at each wavelength. The pH was calculated from each simulated spectrum using both the Ratio and Full Spectrum Methods. All the uncertainties in this study were expressed as standard deviations, and thus they were calculated from the standard deviation of the simulated pH values.

The wavelength and absorbance (or photometric) accuracy of the spectrophotometer constitute systematic sources of uncertainty to spectrophotometric pH determination. We evaluated the effect of wavelength calibration errors by simulating spectra (over the pH range of 7-8) with small wavelength errors (evenly distributed between ±0.5 nm). These spectra were simulated from the reference spectra using weighted averages (scaled according to the wavelength error) of adjacent diode pairs. For each of the five pH values, 1,000 simulated spectra were generated and used to estimate the uncertainty contribution to pH calculated with the Ratio and Full Spectrum Methods.

The uncertainty contribution from absorbance errors was evaluated using calculations similar to those used to evaluate the repeatability contribution. We assumed that the

manufacturer specifications of ±0.005 photometric accuracy represents a 99% confidence interval and that the absorbance errors are constant across all wavelengths.

In addition to systematic pH uncertainty contributions from the spectrophotometer, we also evaluated the contribution from uncertainty in the indicator properties. These calculations used the perturbed indicator properties from the Monte Carlo simulations described earlier to calculate the pH of the five synthetic spectra. The Ratio Method estimates of pH were calculated using the pK_a(HI$^-$)-based **Eq. 4** as well as the $\mathrm{p}(K_a(\mathrm{HI}^-)\cdot e_2)$ -based **Eq. 5**.

3.4 Assessment

3.4.1 Indicator properties and comparison with previous studies

We found discrepancies between our values for the properties of the FB4 purified mCP and Liu et al.'s values (**Table 3.1**,**Table 3.2**) that could be partly attributed to the uncertainty in our measurements and potential systematic errors in the current published dye properties. Of the various molar absorption coefficient ratios, the discrepancies between our values of e_3/e_2 and e_2 and Liu et al.'s values have the most impact on pH. The uncertainty in our value of e_3/e_2 was dominated by the systematic error in measuring the I^{2-} spectrum (**Table 3.3**), related to the drift in our measurements as the deuterium lamp degraded in intensity (**Figure 3.6)**. The error in the I^{2-} spectrum also contributed significantly to the uncertainty in e_2, although the dominant source of uncertainty in e_2 was from the error in normalizing the NaOH and acetate buffer solution spectra at the 488.1 nm isosbestic point during the data processing for the minor species correction. As the uncertainty in our values of e_3/e_2 and e_2 were large, our values may not be significantly different from Liu et al.'s values within two times the estimated standard uncertainties in **Table 3.3** even if their uncertainty was 10% of ours.

The choice of the background medium of the solutions used to measure e_3/e_2 could contribute to differences in the observed values, but this was unlikely to be responsible for the discrepancies we found with Liu et al.'s values. We observed a difference in the I^{2-} spectrum of the FB4 dye measured in NaCl and in a modified synthetic seawater medium (**Figure 3.3b**)—both solutions measured on the same day to ensure that the differences did not represent a drift in the measurements. However, as the difference at 434 nm was small, the resulting difference in the values of e_3/e_2 was not significant. The value of e_3/e_2 measured in a NaCl medium was 0.05543±0.00044 (*mean* ± *std. dev.*, n = 6) and 0.05544±0.00061 (n = 4) in a modified synthetic seawater medium. Furthermore, the adjustments we made to the spectra with orthogonal projection, which partially corrected for background medium effects (see later discussion), did not reduce the discrepancy between our value and Liu et al.'s value for e_3/e_2 (**Table 3.1**).

Our values of e_3/e_2 were consistent with the values (from other studies and our unpublished data) determined on independent batches of purified mCP, all of which were lower than Liu et al.'s value (**Figure 3.9**). As differences in e_3/e_2 for different batches of mCP may reflect differences in the levels of absorbing impurities (Douglas and Byrne, 2017) and as impurities have been found in some batches of imperfectly purified indicators (Takeshita, submitted), a plausible explanation may be that the FB4 dye and Liu et al.'s dye were not identically pure. More research is needed to investigate the source of these discrepancies. Since many dye impurities have a wide range of absorption in the UV-visible region (Liu et al., 2011; Yao et al., 2007), analysis of the full spectra of mCP in NaOH solutions may be a useful approach for the detection of impurities.

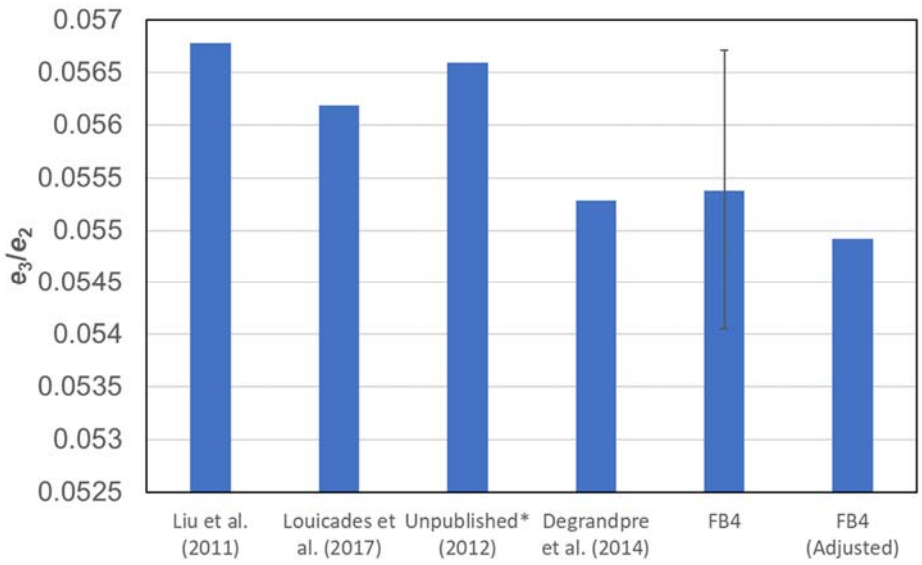

Figure 3.9. Comparison of e_3/e_2 values for independently purified batches of mCP, including various published values, the values we obtained for the FB4 batch of purified mCP, and our unpublished data from 2012 on a different batch of purified mCP. The error bars on the unadjusted FB4 value represent the estimated expanded uncertainty (2 coverage factors) for our measurements (see **Table 3.3**). The adjusted FB4 value refers to corrections to the spectra to account for differences in the absorption behavior of the dye in NaCl and seawater-like media and other potential systematic errors (see text).

Another indication of potential systematic errors in the published mCP properties was that the pH we measured in a TRIS buffer standard at 25°C and S = 35 was 0.0045 units lower than the published value of DelValls and Dickson (1998) when using the coefficients and $\mathrm{p}(K_a(\mathrm{HI}^-)\cdot e_2)$ of Liu et al. (2011), implying that the $\mathrm{p}(K_a(\mathrm{HI}^-)\cdot e_2)$ should be higher. Similar offsets were reported by Müller and Rehder (2018) and Carter et al. (2013). Accordingly, we determined higher values of $\mathrm{p}(K_a(\mathrm{HI}^-)\cdot e_2)$ (0.0056-0.0063 higher, **Table 3.2**) from the TRIS buffer measurements and the FB4 optical absorption properties, also comparable to the offsets reported by DeGrandpre et al. (2014) and Müller and Rehder (2018).

The estimated combined standard uncertainty in $\mathrm{p}(K_a(\mathrm{HI}^-)\cdot e_2)$ from the sources listed in **Table 3.4** (the same as those listed in **Table 3.3**) was 0.0008-0.0013. Thus, the uncertainty in determining the dye optical absorption properties cannot explain the discrepancy between our

value of $p(K_a(HI^-) \cdot e_2)$ for the FB4 dye and Liu et al.'s value. Sources of uncertainty we did not consider in **Table 3.4** include: (1) the uncertainty in the assigned pH value of the TRIS buffer standard (~0.004; Buck et al., 2002), (2) the impact of TRIS on the dissociation behavior of mCP (Müller and Rehder, 2018), and (3) differences in the absorption coefficient ratios between the two dyes. The correction in $p(K_a(HI^-) \cdot e_2)$ for the effect of TRIS was estimated to be approximately -0.001 for a 0.04 mol kg^{-1} equimolar TRIS-TRIS $\cdot$ H^+ buffer at 25°C and S = 35 (Müller and Rehder, 2018), and there is still a residual discrepancy of 0.0038 if the $p(K_a(HI^-) \cdot e_2)$ of the FB4 dye was calculated using Liu et al.'s absorption coefficient ratios. Thus, the latter two sources of uncertainty are unlikely to be the dominant contribution to the discrepancy. Instead, the discrepancy seems to be consistent with the uncertainty in assigning a pH value to the TRIS buffer.

3.4.2 Comparison of modeled and measured spectra

The residuals from a least squares fit of the validation sample spectra to **Eq. 6** had strong patterns (**Figure 3.4**) indicating that there may be systematic errors in our full spectrum model. We adjusted the dye species and sample spectra with the PCA and orthogonal projection approaches described earlier to investigate likely sources of error in the model.

The signal extracted from orthogonal projection of the I^{2-} spectrum was similar to the difference in the I^{2-}spectrum for NaCl versus modified synthetic seawater media, while the error extracted from projection of the HI^- spectrum appeared to be noise (**Figure 3.10**). These results suggested that, in addition to other orthogonal errors, there was likely a systematic error in using the I^{2-} spectrum determined in NaCl media in **Eq. 6** to model the spectra of seawater. From calculations with a synthetic dataset, we determined that the orthogonal projection can

correct for ~77% of the error due to differences in the background medium, as the angle between the vector representing the difference in the I^{2-} spectrum in NaCl versus seawater-like media and the basis vectors of a set of seawater spectra representing a pH range of 7-8 is approximately 130° (i.e., sin 130° ≈ 0.77).

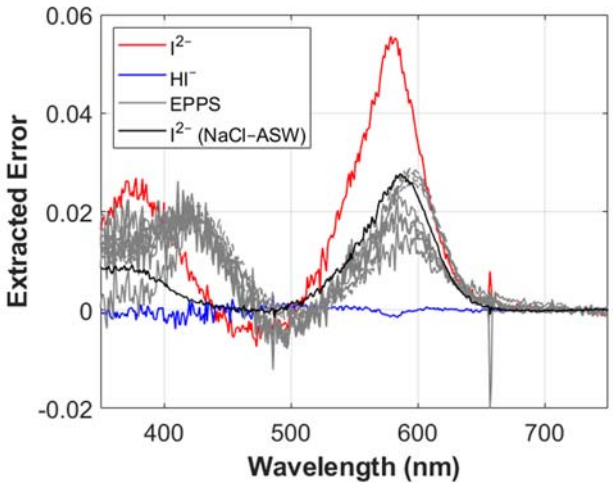

Figure 3.10. Orthogonal errors extracted from the I^{2-}, HI^{-}, and EPPS buffer spectra (pH~7.79) by projecting each spectrum onto the orthogonal basis vectors of a reference dataset (see **Figure 3.2**). The basis vectors were computed from singular value decomposition of the set of spectra (**Figure 3.5a)** reconstructed by retaining only the first two principal components in **Figure 3.5b**. The difference in the I^{2-} spectrum obtained in NaCl and a modified synthetic seawater medium is also plotted (black curve, see also **Figure 3.3b**). All the errors and differences were calculated using spectra normalized to $A_{iso} = 1$.

The residuals were greatly reduced when re-evaluated using the adjusted spectra (**Figure 3.4**). The EPPS buffer samples (pH~7.86) had a unique interference in the spectra (**Figure 3.4**). The cause of this interference was unclear, but could be an instrumental effect or from an interaction between the dye and the EPPS buffer. Although the orthogonal projection adjustment (**Figure 3.10**) improved the agreement of the EPPS buffer spectra with the full spectrum model, the residuals still had a pronounced, but different pattern.

These results demonstrate that, although the residuals from the Full Spectrum Method can provide diagnostic and quality control information, the effect of the background medium on

the dye species spectra can complicate the interpretation. Ideally, the dye species spectra should be determined in the medium appropriate to the samples or adjusted for the medium effect.

3.4.3 Quality control with the Full Spectrum Method

In addition to the information from the residuals, the agreement between the Full Spectrum and Ratio Method estimates of pH can also indicate potentially problematic measurements. For the unadjusted data, the mean discrepancy between the Full Spectrum and Ratio Method estimates of pH of individual measurements on the validation samples was 0.00095±0.00062. One measurement in particular—the third sample in the FB4 dataset (an EPPS buffer, pH~7.79)—appeared to be problematic, lying beyond two standard deviations from the mean discrepancy (**Figure 3.11**). This sample had slightly different residuals compared to the other EPPS buffer samples, with a sharp peak (extending beyond the range of the y-axis in **Figure 3.4**) at 657 nm (approximately the location of one of the Balmer emission lines), suggesting that the measurement was affected by a transient aberration in the behavior of the deuterium lamp.

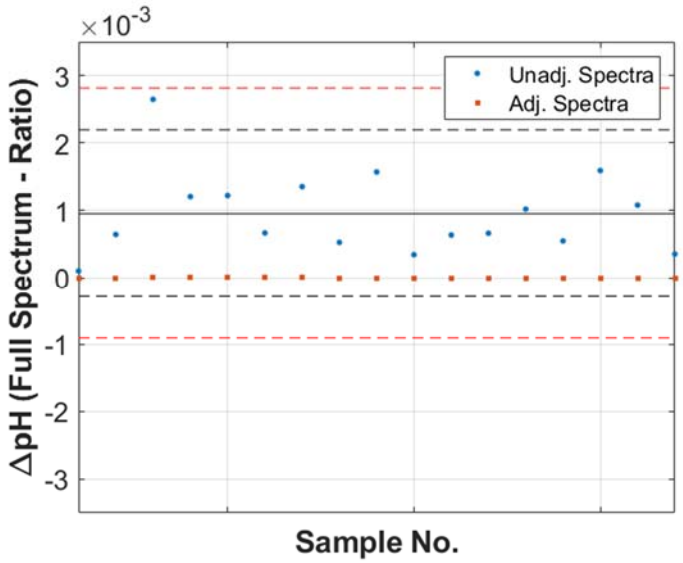

Figure 3.11. Differences between the Full Spectrum and Ratio Method estimates of pH for the individual measurements on the validation samples. The solid black line denotes the mean ΔpH of the unadjusted data, and the dashed black and red lines represent the two and three standard deviations from the mean ΔpH, respectively.

3.4.4 Effect of deuterium lamp performance

Our measurements of the I^{2-} spectrum appeared to be affected by changes in the performance of the deuterium lamp. Repeated measurements of the FB4 dye spectra in NaOH solutions over the course of eight months showed drifts (**Figure 3.6**) that became larger with increasing time elapsed since new lamps were installed on the spectrophotometer (in May 2019). During this time period (May to November 2019), the deuterium lamp intensities, measured in air, decreased significantly (64% decrease in the 190-220 nm region and 50% decrease in the 350-500 nm region), while the tungsten lamp intensities were relatively stable (<0.1% change in the 950-1100 nm region).

Data from additional measurements made over the course of several weeks in February and March 2020 supported our hypothesis of the degrading deuterium lamp being a source of systematic error in measurements of the I^{2-} spectrum. The difference in the I^{2-} spectrum measured on the main spectrophotometer, using the same set of lamps as before, and a second

instrument (**Figure 3.7**), which used a different set of lamps, showed a similar pattern to those in **Figure 3.6**. This pattern was minimized when the measurements on the main spectrophotometer were repeated using different sets of lamps which included deuterium lamps with higher intensities (**Figure 3.7**, **Table 3.5**). Although all of the lamps passed the Agilent 8453 self-diagnostic criteria for intensity (**Table 3.5**), these results suggest that moderate levels of degradation of the deuterium lamp may affect measurements of the I^{2-} spectrum. Additionally, the degradation of the deuterium lamp can result in absorbance errors that affect measurements at lower pH (see later discussion). Thus, periodic monitoring of drifts in I^{2-} spectra measurements may be useful for early detection of declining lamp performance.

The exact mechanism as to how degrading intensities on the deuterium lamp may affect the I^{2-} spectra measurements is unclear. The patterns in **Figure 3.6** resembled the first derivative of the I^{2-} spectrum, suggesting that they may be related to wavelength errors. Shifts in the deuterium emission peak around 486.0 nm, measured during the spectrophotometer self-tests between May 2019 and May 2020, can reproduce the same patterns of errors in the I^{2-} spectrum but an order of magnitude smaller than the measured changes (**Figure 3.6a, b**). Wavelength-dependent absorbance errors may explain our observations, but we were unable to test this hypothesis.

Orthogonal projection of the I^{2-} spectra should, in principle, remove first derivative-type errors, which are nearly orthogonal to the basis vectors for a set of spectra representing a pH range of 7-8. However, there was still some degree of residual error in the adjusted dataset (**Figure 3.6a, 3.4b**), suggesting that there may have been non-orthogonal errors that were not extracted by the projection.

3.4.5 Quality control of pH measurements with PCA

Our analysis of a dataset of measurements on CRM 186 over five months demonstrated the utility of PCA for identifying changes in spectrophotometer performance and other quality control issues.

A number of anomalous measurements in this dataset could be identified from examination of the loading spectra and scores for the first six principal components (**Figure 3.8**). As the data represented measurements on a stable batch of seawater with fairly homogenous pH, the scores for the first principal component, which corresponds to the variation in pH, were relatively constant, with the exception of three groups of measurements (Samples 40-42, 54-65, and 66-74). These measurements corresponded to the three groups of data from the right in **Figure 3.12**, indicated with filled symbols. Two of these groups had pH values that were on average at least 0.004 units higher than the mean value of the entire dataset.

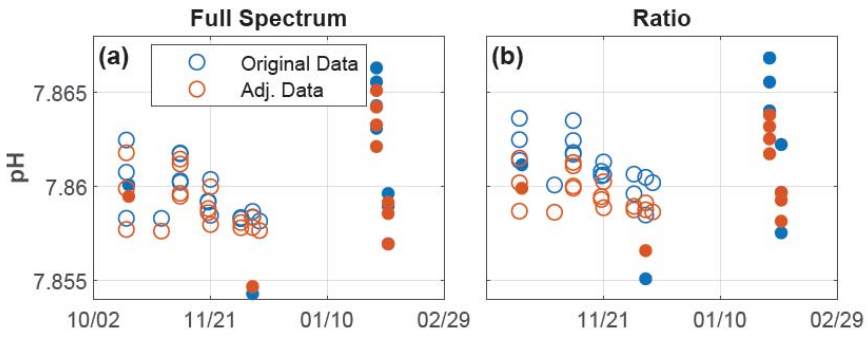

Figure 3.12. pH measured on 26 bottles of CRM Batch 186 between October 2019 to February 2020, using the FB4 purified mCP. The pH values were calculated using the (a) Full Spectrum and (b) Ratio Method (**Eq. 4**) with the FB4 dye optical absorption properties and a value of 7.9913 for pK_a(HI^-). Each point represents multiple measurements at different dye amounts on an individual bottle, extrapolated to A_{iso} = 0 to correct for the effect of the indicator on the sample pH. The pH values of the adjusted dataset (orange symbols) were calculated using the sample spectra reconstructed using only the first principal component of the set of spectra represented by the unfilled symbols and the adjustments to the FB4 HI^- and I^{2-} spectra in **Figure 3.10**. Spectra represented by the filled symbols were excluded from the PCA adjustments and were corrected by projection onto the dataset represented by the unfilled symbols (see text).

The second principal component in **Figure 3.8** was associated with two groups of measurements (Samples 4-6 and 66-74). Although the source of this component is unclear, the scores indicated that the contribution from the second principal component became persistent and increasing in magnitude after January 2020 (**Figure 3.8**), suggesting that it may be related to the declining performance of the deuterium lamp. The anomalous measurements associated with the second principal component could also be identified from the residuals obtained from the Full Spectrum Method, appearing as large deviations from the main pattern (**Figure 3.13**).

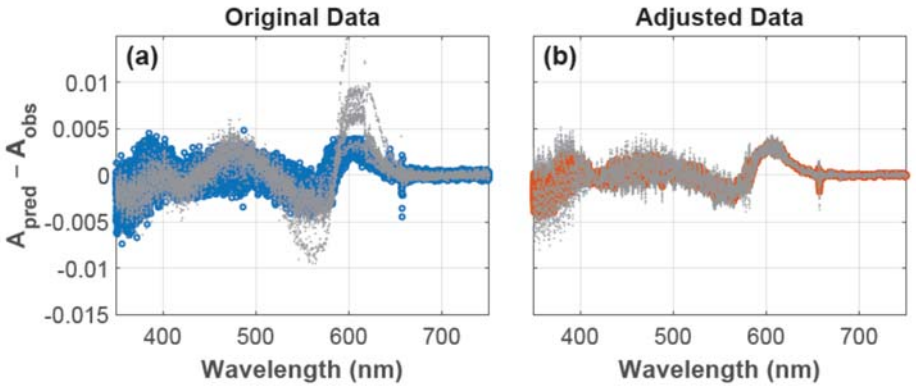

Figure 3.13. Residuals from fitting the CRM 186 spectra (see **Figure 3.8a**), normalized to $A_{iso} = 1$, to the HI^- and I^{2-} spectra for the FB4 dye (**Eq. 6**). The residuals were evaluated with and without the adjustments described in **Figure 3.12**. The filled gray symbols are the residuals for the spectra excluded from the PCA adjustments in **Figure 3.12**.

The adjustments to the spectra from discarding the error components identified from PCA and from orthogonal projection eliminated the large deviations in the residuals, improved the agreement between the Full Spectrum and Ratio Method estimates of pH, and reduced the variance in the pH values (**Figure 3.12**, **Figure 3.13**). The adjustments reduced the standard deviation of the pH values from 0.0027 and 0.0026 for the Full Spectrum and Ratio Methods, respectively, to 0.0024 and 0.0016. Thus, the variance removed from these adjustments was 0.0012 and 0.0020 for the Full Spectrum and Ratio Methods, respectively.

3.4.6 Evaluation of instrument differences

We identified an absorbance-dependent error in absorbance which contributed to the uncertainty of the pH measurements on one of our spectrophotometers. Our first set of phosphate buffer measurements (Oct. 2019) on three spectrophotometers had an instrument pair offset in pH as large as 0.0045. Additionally, on the main spectrophotometer, there was a strong dependence of the pH on the amount of dye added, as judged by the slope of pH vs. A_{iso} (**Figure 3.14**), which was at least twice as large as the slopes from the other sets of measurements (**Table 3.6**).

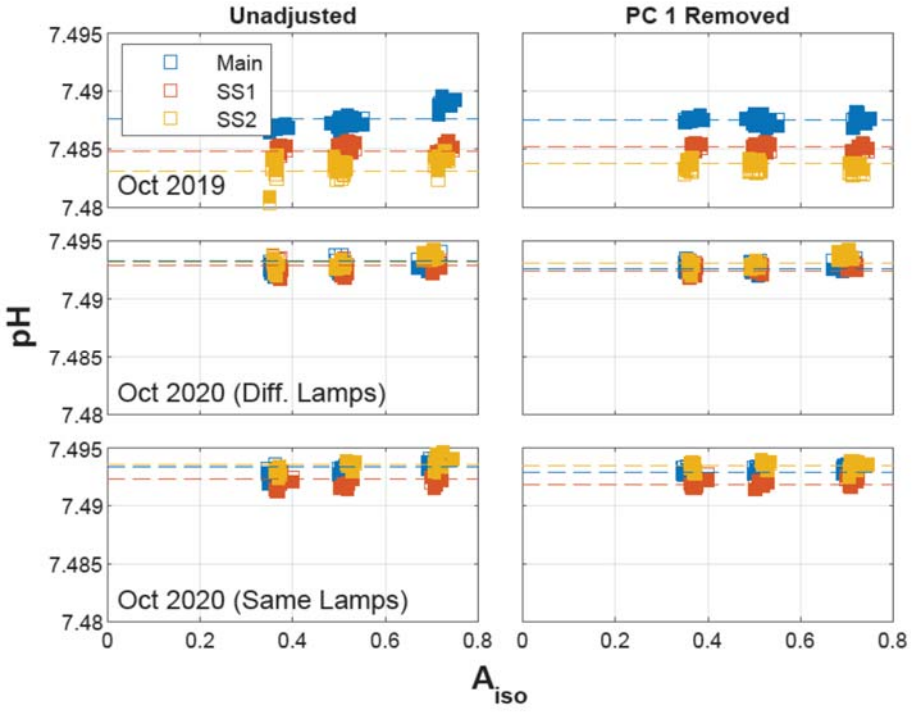

Figure 3.14. pH of three batches of phosphate buffer measured on different spectrophotometers in three separate experiments, plotted against A_{iso}. pH values were calculated using the Ratio Method (filled symbols) and the Full Spectrum Method (unfilled symbols). The dashed lines indicate the mean pH for each set of measurements. The left panels show the pH calculated from the unadjusted spectra, while the pH values in the right panels were calculated using spectra that have been adjusted by removing the first principal component of their corresponding dataset of mean-centered, unit length-normalized spectra.

Table 3.6. Summary of measurements of phosphate buffers using the FB4 purified dye on different spectrophotometers in three separate experiments. A large volume of a single batch of buffer was used for each experiment. Values of the mean pH, its standard deviation, the slope of pH *vs*. A_{iso}, the 95% confidence interval of the slope, and the number of measurements (n) are reported for each set of measurements in **Figure 3.14**. The pH and the slopes were calculated using both the Full Spectrum and Ratio Methods. Slopes that are significantly greater than zero are bolded. For each set of measurements, the second set of italicized values for the mean pH and slope were calculated using spectra that have been adjusted by removing the first principal component of their corresponding dataset of mean-centered, unit length-normalized spectra. The deuterium (D) and tungsten (W) lamps were assigned scores on a linear scale from 0 (fail) to 5 (peak performance) as in **Table 3.5**.

		Full Spectrum Method		Ratio Method		
Exp	**Instrument**	**Mean pH ± *std. dev.***	**Slope ± 95% C.I.**	**Mean pH ± *std. dev.***	**Slope ± 95% C.I.**	**n**
Oct '19	Main	7.4876±0.00070	**0.00617±0.00048**	7.4875±0.00076	**0.00606±0.00097**	46
	D = 2, W = 4	*7.4876±0.00005*	*-0.00005±0.00013*	*7.4875±0.00031*	*-0.00007±0.00085*	
	SS1	7.4848±0.00018	0.00039±0.00050	7.4852±0.00026	0.00007±0.00075	45
	D, W = n.d.	*7.4848±0.00010*	***-0.00042±0.00026***	*7.4852±0.00021*	***-0.00079±0.00055***	
	SS2	7.4830±0.00076	**0.00284±0.00219**	7.4837±0.00076	0.00208±0.00236	24
	D, W = n.d.	*7.4830±0.00012*	***-0.00063±0.00029***	*7.4837±0.00029*	***-0.00116±0.00084***	
Oct' 20 (a)	Main	7.4932±0.00048	**0.00235±0.00115**	7.4926±0.00047	**0.00199±0.00127**	21
	D = 4, W= 3	*7.4932±0.00017*	***0.00054±0.00052***	*7.4926±0.00037*	*0.00080±0.00118*	
	SS1	7.4929±0.00053	0.00035±0.00176	7.4924±0.00036	0.00066±0.00117	21
	D >5, W=5	*7.4929±0.00016*	***0.00060±0.00043***	*7.4924±0.00022*	***0.00080±0.00061***	
	SS2	7.4932±0.00038	0.00109±0.00117	7.4931±0.00053	**0.00238±0.00140**	21
	D = 4, W = 4	*7.4932±0.00034*	***0.00170±0.00083***	*7.4931±0.00053*	***0.00262±0.00128***	
Oct' 20 (b)	Main	7.4934±0.00052	**0.00310±0.00086**	7.4929±0.00052	**0.00300±0.00092**	21
	D=3, W= 3	*7.4934±0.00005*	*0.00015±0.00016*	*7.4929±0.00026*	*0.00046±0.00084*	
	SS1	7.4923±0.00048	0.00126±0.00149	7.4919±0.00041	**0.00164±0.00115**	21
	D > 5, W = 5	*7.4923±0.00012*	*0.00023±0.00041*	*7.4919±0.00030*	***0.00101±0.00090***	
	SS2	7.4935±0.00046	**0.00259±0.00082**	7.4935±0.00061	**0.00288±0.00144**	21
	D = 4, W = 4	*7.4935±0.00025*	*0.00025±0.00079*	*7.4935±0.00041*	*0.00012±0.00135*	

n.d. = No data
(a) = Different sets of lamps were used on the three spectrophotometers.
(b) = The same set of lamps were used on the three spectrophotometers.

Significant changes in the pH due to addition of indicator are unlikely in well-buffered solutions. Changes in the pH of the solution would result in the absorbances at 434 nm and 578 nm both changing in opposite directions. Instead, a closer examination of the spectra collected from measurements on the main spectrophotometer in the first experiment suggested an absorbance-dependent error in absorbance disproportionately affecting 434 nm. The differences between the individual spectra measured on the main spectrophotometer and the mean spectrum measured on a second spectrophotometer showed strong A_{iso}-dependent negative

anomalies below 450 nm (**Figure 3.15a).** The lower absorbances measured on the main spectrophotometer at 434 nm therefore led to a higher mean pH compared to the other two spectrophotometers (**Figure 3.14**).

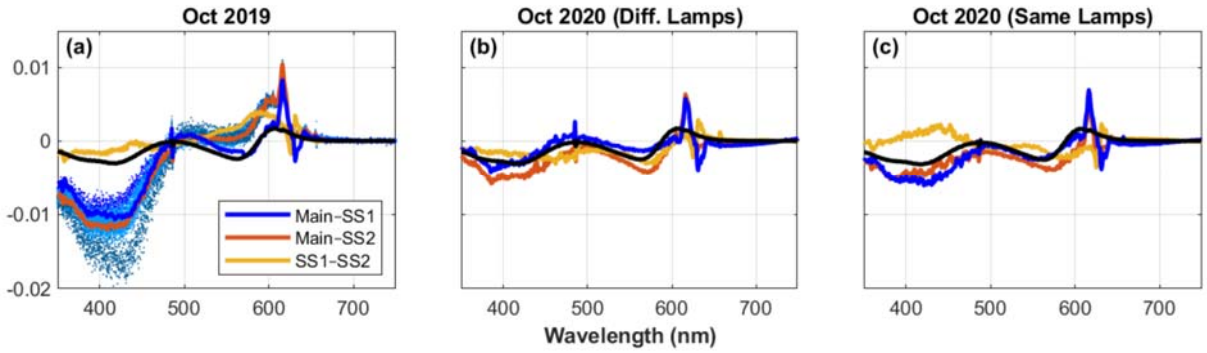

Figure 3.15. Instrument pair differences in the mean phosphate buffer spectra (normalized to $A_{iso} = 1$) for the three experiments. The black curve is the simulated difference spectrum resulting from a 0.1 nm wavelength error in the main spectrophotometer. The blue points in (a) are the differences between the individual spectra measured on the main spectrophotometer and the mean spectrum measured on the SS1 spectrophotometer. The three different shades of blue indicate three different dye amounts used in the measurements (see **Figure 3.14**).

PCA was used to further diagnose problematic components in the spectra. As the pH variation in the buffer is negligible, the principal components computed for each set of phosphate buffer measurements represent systematic and random errors. The scores for the first principal component were strongly correlated with A_{iso} for the first set of measurements on the main spectrophotometer (**Figure 3.16**), indicating the absorbance-dependent contribution of the principal component to the spectra. Subtracting the contribution of the first principal component from the spectra in this set of measurements thus eliminated the slope of pH vs. A_{iso} (**Table 3.6**). The plots in **Figure 3.16** indicated varying levels of absorbance-dependent contributions from the first principal component in the other sets of measurements, which have an effect on the slope of pH vs. A_{iso} (**Table 3.6**). However, in some cases, there were also significant and opposing contributions from the other principal components (not shown), and

thus removing the first principal component in those datasets either increased the magnitude of the slope (e.g., SS1 and SS2 in the second experiment with different sets of lamps) or reversed its sign (e.g., SS1 and SS2 in the first experiment).

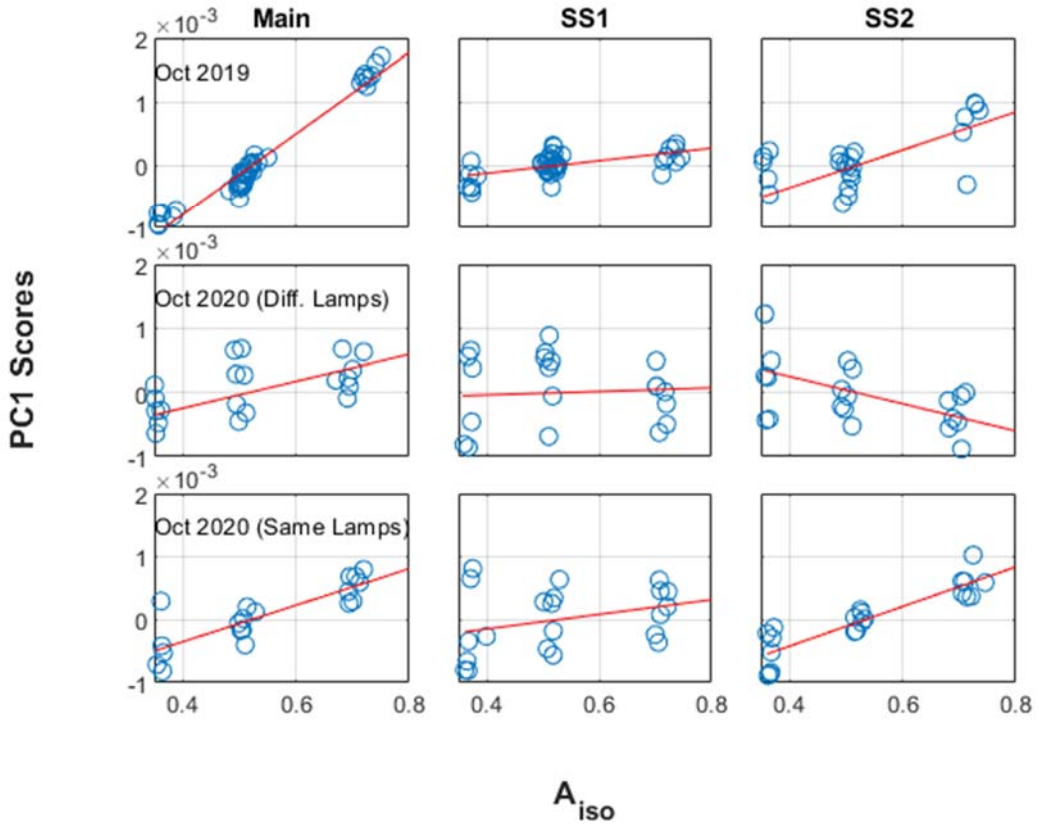

Figure 3.16. Scores of the first principal component (calculated using unit length, mean-centered spectra) plotted against A_{iso} for each set of phosphate buffer measurements on the three spectrophotometers in three different experiments.

The slopes of pH vs. A_{iso} and their 95% confidence intervals were slightly different when using pH values estimated with the Full Spectrum and Ratio Methods (**Table 3.6**). Generally, the Full Spectrum Method resulted in better fit of the points in **Figure 3.14** to the least squares line (i.e, lower standard error in y) and therefore a lower uncertainty in the slope.

Consequently, some of the slopes estimated using the Ratio Method were not significantly different from zero, while the slopes estimated with the Full Spectrum Method were significant due to the lower uncertainty (**Table 3.6**).

We were not able to reproduce the problems observed in the first set of experiments with the main spectrophotometer as the deuterium lamp used originally had failed entirely by October 2020. The deuterium lamps used in the October 2020 experiments had 1.2 to 4.5 times the intensity (in the 190-350 nm region) of the deuterium lamp used on the main spectrophotometer in October 2019. Using lamps with higher intensities, the instrument pair pH offsets were reduced to no larger than 0.0005 in the second experiment and no larger than 0.0016 in the third experiment. The instrument pair differences in the phosphate buffer spectrum also did not show the same strong negative absorbance anomalies below 450 nm as observed initially on the main spectrophotometer (**Figure 3.15**). Instead, the differences in **Figure 3.15** could largely be explained by small offsets in the wavelength calibration of the spectrophotometers (<0.1 nm as determined from wavelengths of the deuterium emission peaks reported by the spectrophotometer self-tests). Hence, none of the data in these later experiments showed a strong dependence of pH with A_{iso}.

Differences in lamp and instrument performance contributed to a standard deviation in pH of 0.0019 for the Full Spectrum Method and 0.0016 for the Ratio Method in the October 2019 experiment (calculated from the standard deviation of all the pH measurements), comparable to the variance extracted from the PCA adjustments on the CRM 186 dataset. If lamp performance issues were better controlled, the uncertainty in pH could be reduced to 0.00061 for the Full Spectrum Method and 0.00071 for the Ratio Method (estimated from a

pooled standard deviation of the pH measured on the three spectrophotometers from the Oct. 2020 experiments).

3.4.7 Performance of the Full Spectrum and Ratio Methods

One of the purported advantages of the Full Spectrum Method is the significantly improved precision achieved through the use of information over a range of wavelengths. Ohline et al. (2007) reported a repeatability of 0.0001 in pH when using the Full Spectrum Method with an Agilent 8453 spectrophotometer. Our estimates of the repeatability of pH measurements from Monte Carlo simulations varied with pH and were no larger than 0.00012 and 0.00085 at pH 7 for the Full Spectrum and Ratio Methods, respectively (**Figure 3.17**). This level of repeatability with the Full Spectrum Method was not achieved in our measurements of the phosphate buffers. The pooled standard deviations in pH from all the experiments in **Figure 3.14** were 0.00053 and 0.00055 for the Full Spectrum and Ratio Methods, respectively. As discussed in the previous section, there were likely systematic contributions to the variance of the pH measurements in these datasets. If the first principal component from each set of spectra (representing the absorbance-dependent absorbance errors) was removed, the pooled standard deviations in pH improve to 0.00016 and 0.00032.

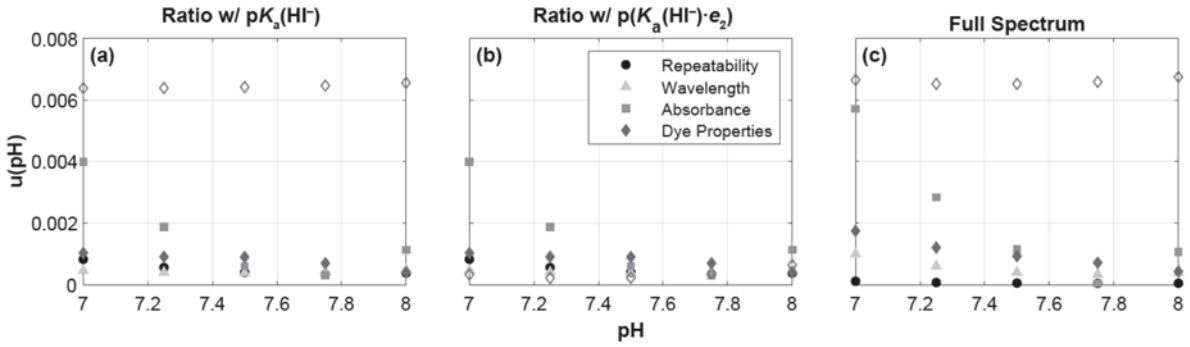

Figure 3.17. Estimates of the uncertainty in pH, expressed as standard deviations and plotted against pH, due to contributions from the repeatability of absorbance measurements, wavelength errors, absorbance errors, and errors in the dye properties. The uncertainties were evaluated for the Ratio Method using (a) **Eq. 4** and (b) **Eq. 5** and for the (c) Full Spectrum Method. The filled diamonds in the three plots represent the pH uncertainty contribution from the dye properties when the errors in pK_a(HI$^-$) and $\mathrm{p}(K_a(\mathrm{HI}^-)\cdot e_2)$ are correlated with the errors in the dye optical absorption properties. The unfilled diamonds are the contributions only from errors in the dye optical absorption properties.

Sub-nanometer wavelength errors on the Agilent 8453 should not significantly affect spectrophotometric pH accuracy as the relevant wavelengths for the Ratio Method are at fairly broad absorbance peaks rather than a strongly sloping region of the spectrum (Sharp et al., 2017). Consistent with these expectations, we estimated the uncertainty in pH due to wavelength errors to be <0.001 for both the Ratio and Full Spectrum Methods.

The pH uncertainty contribution from systematic errors in absorbance, on the other hand, is more significant. Our estimates of the contribution of absorbance errors to uncertainty in pH were as large as 0.0059 at pH 7 for the Full Spectrum Method and 0.0041 for the Ratio Method (**Figure 3.17**). Thus, although the Full Spectrum Method can potentially offer greater repeatability, it is more sensitive to systematic errors in absorbance.

Uncertainties in the indicator properties (**Table 3.3, Table 3.4**) will contribute to systematic uncertainty in the spectrophotometric pH estimate and will affect the Full Spectrum and Ratio Methods differently. Generally, the determination of the indicator properties

constitutes a small source of uncertainty to the pH estimate in both methods. Although there are potentially large uncertainties in $\log(e_2)$ and $pK_a(HI^-)$, these two sources of uncertainty are correlated. Hence, when the absorption coefficient ratios and the $pK_a(HI^-)$ are both determined, their errors largely cancel when using these properties to estimate pH. Similarly, errors in the $pK_a(HI^-)$ and in the HI^- and I^{2-} spectra cancel to a large extent, such that the resulting uncertainty in the Full Spectrum estimate of pH is also small. Our Monte Carlo simulation estimates of the contribution of the uncertainty in the indicator properties to the uncertainty in pH varied as a function of pH (**Figure 3.17**). For the Ratio Method, the uncertainty in pH varied from 0.0010 at pH 7 to 0.0004 at pH 8 and was identical whether using **Eq. 4** or **Eq. 5**. For the Full Spectrum Method, the uncertainty at low pH is larger than with the Ratio Method due to a larger contribution from the uncertainty in $pK_a(H_2I)$, which is used to correct the HI^- spectrum for the presence of small amounts of H_2I (**Figure 3.18**). As a result, the uncertainty in the Full Spectrum estimate of pH varied from 0.0018 at pH 7 to 0.0004 at pH 8.

In some cases, it may be desirable for spectrophotometric pH users to characterize the optical properties of their dye—for instance, when using an impure dye or when using an instrument with a significantly different bandpass than the reference instrument on which the published indicator properties were determined (DeGrandpre et al., 2014). If the apparent $pK_a(HI^-)$ or $p(K_a(HI^-)\cdot e_2)$ are not also determined, the indicator calibration will not have the advantage of error cancellation, and the differences between the Full Spectrum and various forms of the Ratio Method will become more important. pH calculated from the Ratio Method using **Eq. 4** will have an uncertainty of at least 0.0064 (**Figure 3.17, Figure 3.18)** due to the contributions from systematic errors in the I^{2-} spectrum and from $\log(e_2)$, which is strongly affected by errors in normalizing the HI^- and I^{2-} spectra at the isosbestic point (**Table 3.3**). The

uncertainty in the Full Spectrum estimate of pH will also have a similar uncertainty for similar reasons. The overall uncertainty in pH for both of these methods will also include the contribution from the value of pK_a(HI^-) selected for use in **Eq. 4** or **Eq. 1**. Assuming an uncertainty of 0.0067 in pK_a(HI^-), based on independent characterizations of mCP (**Table 3.3**), the overall uncertainty in pH will be at least 0.0093.

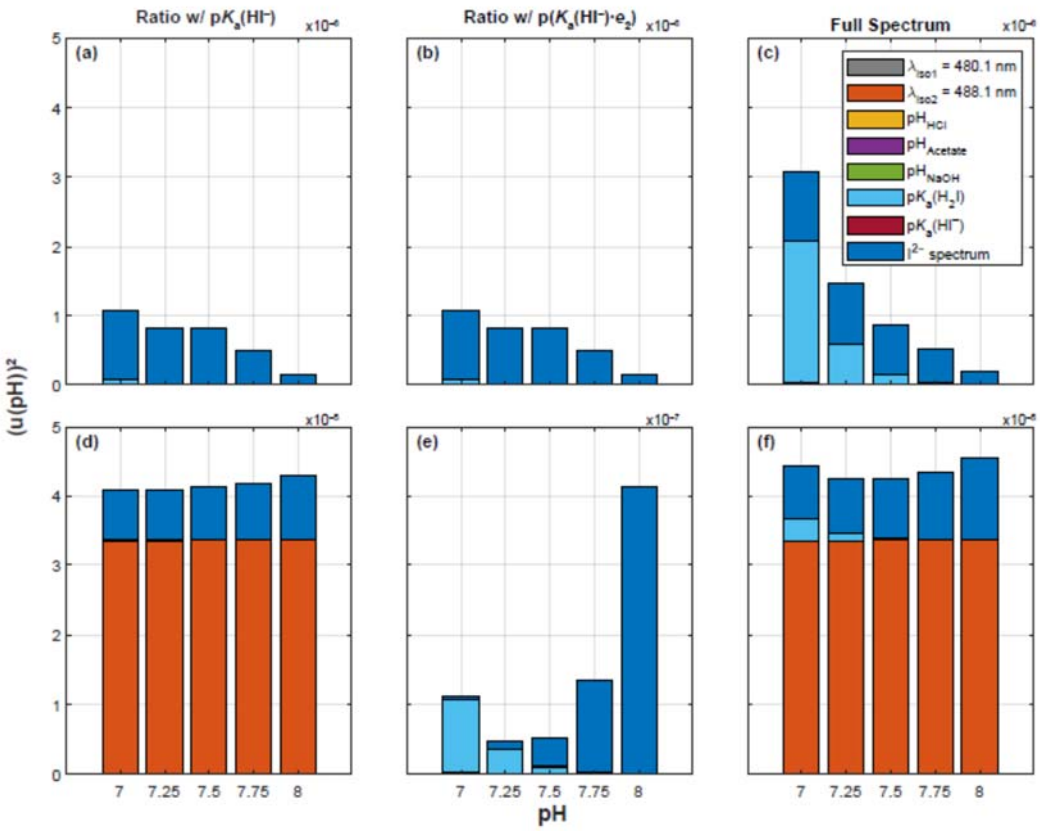

Figure 3.18. Stacked bar chart showing contributions to $(u(\mathrm{pH}))^2$, plotted against pH, from the data processing to obtain the mCP species absorption spectra. The listed contributions are the same as those in **Table 3.3**. The uncertainties were evaluated for the Ratio Method using **Eq. 4** (left panels) and **Eq. 5** (center panels) and for the Full Spectrum Method (right panels). The top panels represent the scenario in which errors in pK_a(HI^-) and $\mathrm{p}(K_a(\mathrm{HI}^-)\cdot e_2)$ are correlated with the errors in the dye optical absorption properties (filled diamonds in **Figure 3.17**). The bottom panels consider only errors in the dye optical absorption properties (unfilled diamonds in **Figure 3.17**).

For users who need to characterize the optical properties of their dye, several options are possible for reducing the uncertainty in the indicator calibration. Concurrently determining both the dye optical properties as well as the pK_a(HI^-) or $p(K_a(HI^-)\cdot e_2)$ with measurements in TRIS buffer standards will give the lowest uncertainty in the calibration. If this option is not practical, an alternative approach to normalizing the dye species spectra, such as normalizing by the total dye concentration rather than at the isosbestic point, may be reduce this source of uncertainty. We avoided the approach of normalizing the spectra by total dye concentration, as we had difficulties ensuring the complete dissolution of the indicator. The use of **Eq. 5**, which is based on $p(K_a(HI^-)\cdot e_2)$, is advantageous in principle, as this form of the Ratio Method does not require determining e_2 and is thus insensitive to the errors in normalizing the spectra. However, the use of $p(K_a(HI^-)\cdot e_2)$ is incompatible with the Full Spectrum Method, instruments with a different bandpass, and impure dyes. There is therefore a need to better constrain the pK_a(HI^-) of mCP, ideally with a method independent of the spectrophotometric method.

3.5 Discussion

We demonstrated that the analysis of full spectral information can reveal systematic errors in measuring absorbance, which may be related to subtle changes in spectrophotometer performance. Although these changes are not immediately obvious when using the conventional Ratio Method for pH determination, they can nevertheless contribute to error in spectrophotometric pH measurements. We found absorbance-dependent errors in absorbance on one of our spectrophotometers which developed when using deuterium lamps that have degraded significantly in intensity. These absorbance errors, which lead to error in the dye

perturbation correction and contribute to instrument offsets, can be identified from PCA on a dataset of spectra collected from repeated measurements on a stable batch of seawater or buffered solution.

The Full Spectrum Method of Ohline et al. (2007) can also offer useful quality control information. Systematic patterns in the residuals or large discrepancies between the pH values estimated with the Full Spectrum and Ratio Methods can alert the user to potentially problematic measurements. However, a difficulty with using the Full Spectrum Method is that the I^{2-} spectrum is sensitive to differences in the background medium, and thus, interpreting the residuals will be more complicated if the I^{2-} spectrum was measured in a different background medium than the samples. Although some researchers, including ourselves, have used a simple NaCl background when measuring the I^{2-} spectrum (DeGrandpre et al., 2014; Douglas and Byrne, 2017), these measurements require adjustments for the medium effect when using the Full Spectrum Method. Ideally, the I^{2-} spectrum should be measured in a modified synthetic seawater medium as in Liu et al. (2011).

Undoubtedly, the Full Spectrum Method of Ohline et al. (2007) introduces significantly more complexity to spectrophotometric pH measurements and requires tedious measurements to characterize the indicator species absorption spectra. Using this approach for quality control thus has a significant disadvantage compared to PCA-based methods, which do not require information about the dye species spectra. However, Ohline et al.'s Full Spectrum Method provides quality control information for individual samples at the time of measurement, while analyzing datasets of spectra with PCA can provide information about changes in measurement quality over time.

The main advantage of the Full Spectrum Method for pH determination is purportedly significantly improved precision, but its overall uncertainty compared to the Ratio Method is unknown. Previous studies have evaluated various contributions to the uncertainty of spectrophotometric pH measurements when using the Ratio Method (Carter et al., 2013; DeGrandpre et al., 2014). Using Monte Carlo simulations, we demonstrated that a seven to eight-fold improvement in the repeatability of pH measurements (relative to the Ratio Method) is possible with the Full Spectrum Method, consistent with the results of Ohline et al. (2007). However, this benefit is outweighed by the greater sensitivity of the Full Spectrum Method to the spectrophotometer absorbance accuracy, which can lead to pH-dependent errors in pH.

Systematic errors in the spectrophotometer wavelength calibration and absorbance measurements have been proposed as potentially contributing to the pH-dependent discrepancies between spectrophotometric pH measurements and pH calculated from A_T and C_T, which have been observed on open ocean cruises (Álvarez et al., 2020). We estimated that the small wavelength calibration errors on the Agilent 8453 contribute <0.001 to the uncertainty of pH measurements. Absorbance errors, on the other hand, contribute to the uncertainty of pH measurements in a pH-dependent manner. However, even with an absorbance error as large as 0.005 (the manufacturer specification for the Agilent 8453) at the two wavelengths 434 and 578 nm, the resulting pH-dependent discrepancies in pH are only a fraction of what was observed on open ocean cruises (**Figure 3.19)**. Hence, other systematic errors in the dissociation constants of carbonic acid (K_1 and K_2), the total boron-salinity ratio (B_T/S), the A_T and C_T measurements, and unidentified (potentially organic) contributions to A_T need to be considered to explain the observed inconsistencies (Fong and Dickson, 2019). Spectrophotometers should be routinely checked with absorbance and wavelength standards

and the lamp intensities carefully monitored on future repeat hydrography cruises to verify that the instrument performance is within specifications.

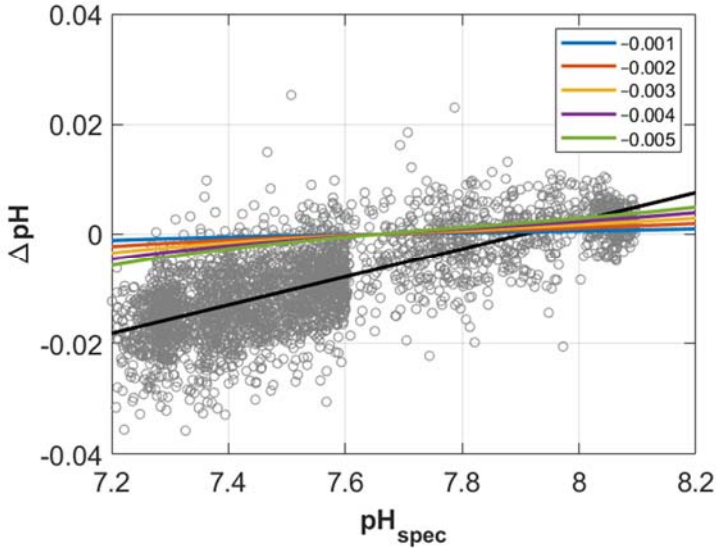

Figure 3.19. Values of ΔpH (difference between spectrophotometric pH and pH calculated from A_T and C_T) at 25°C and a gauge pressure of zero dbar plotted against measured pH for the 2015 P16N cruise (see **Figure 2** in Fong and Dickson, 2019). The various colored curves represent the error in pH resulting from a constant error in measuring the absorbance (as indicated in the legend) at 434 nm and 578 nm.

We also evaluated how the characterization of the dye properties contribute to uncertainty in determining pH with the Ratio and Full Spectrum Methods. Errors in characterizing the dye species absorption spectra are correlated with the error in determining $pK_a(HI^-)$ or $p(K_a(HI^-)\cdot e_2)$. Therefore, the resulting uncertainty in pH is small (<0.001 for the Ratio Method over the pH range of 7-8). The Full Spectrum Method is more sensitive to errors in the HI^- spectrum, resulting in a larger uncertainty at low pH (~0.0018 at pH 7). These estimates do not include the more fundamental uncertainty from the pH value assigned to the TRIS buffer used to determine $pK_a(HI^-)$ or $p(K_a(HI^-)\cdot e_2)$, which may be as large as ~0.004 (Buck et al., 2002; Pratt, 2014). Additionally, if the dye used for the sample measurements is

not of the same purity as the batch of dye that was characterized or if the instrument used for pH measurements has a significantly different bandpass than the instrument used to characterize the dye, the uncertainty in determining pH will be larger (DeGrandpre et al., 2014; Liu et al., 2011; Yao et al., 2007).

Published sets of dye properties likely contain an uncertainty contribution related to the performance of the spectrophotometer used to characterize the dye, and this can contribute to some of the discrepancies in various sets of published values for purified mCP. We identified large drifts in our measurements of the I^{2-} spectrum related to the degradation of the deuterium lamp on our spectrophotometer. It was therefore difficult to distinguish our value of e_3/e_2 for the FB4 dye from Liu et al.'s value within two times our estimated standard uncertainty. However, our value for the e_3/e_2 was within the spread of e_3/e_2 values from independent determinations on different batches of purified mCP (DeGrandpre et al., 2014; Loucaides et al., 2017; our unpublished data), which were all systematically lower than Liu et al.'s values. Insofar as differences in e_3/e_2 can be considered an indicator of differences in dye purity (Douglas and Byrne, 2017), lower e_3/e_2 values imply a purer dye. Another indication that there may be systematic errors in the current determination of purified mCP properties was that our (unadjusted) value of $\mathrm{p}(K_a(\mathrm{HI}^-)\cdot e_2)$ was 0.0056 higher than Liu et al.'s value, which is similar to the magnitude and direction of offsets reported in independent studies (DeGrandpre et al., 2014; Müller and Rehder, 2018). The most likely source of this discrepancy is an error in the pH value assigned to the TRIS buffer (based on its composition as inferred from the solution preparation process), as the other uncertainties considered in **Table 3.4** were estimated to contribute ~0.0008 to the uncertainty in $\mathrm{p}(K_a(\mathrm{HI}^-)\cdot e_2)$.

3.6 Comments and recommendations

We encourage laboratories that perform routine spectrophotometric pH measurements using diode array spectrophotometers to utilize the full spectral information for quality control. Because pH determination with the Full Spectrum Method has greater uncertainty, we recommend using it in conjunction with the Ratio Method as a quality control check only. In addition (or as an alternative), we recommend conducting PCA on datasets of spectra to identify potential systematic errors in the measurements of the spectra or other changes in measurement quality. Spectra from regular measurements of CO_2-in-seawater RMs or a buffered solution can be analyzed by PCA to evaluate contributions to the variance of pH measurements under intermediate precision conditions. Measurements on a single batch of a buffered solution are also useful for identifying absorbance-dependent absorbance errors that may lead to errors in the dye perturbation correction during routine measurements. For this evaluation, we recommend preparing a large volume of phosphate buffer like we did, as the pH of the buffer does not need to be accurately known and the phosphate buffer pH has a low temperature sensitivity compared to other buffers like TRIS.

Many of the problems we experienced, such as drifts in the measured I^{2-} spectra and absorbance-dependent absorbance errors, seem to be related to degrading intensities on the deuterium lamp. We therefore recommend carefully monitoring the lamp intensities and conducting further checks (such as the phosphate buffer measurements) once the deuterium lamp intensities have decreased by more than half of their original values when the lamp was first installed. As the I^{2-} spectra measurements seem to be particularly sensitive to changes in the deuterium lamp quality, periodic measurements in NaOH solutions can also be used as a quality control check.

Implementing the Full Spectrum Method requires the absorption spectra for the HI^- and I^{2-} species of the dye and a value for $pK_a(HI^-)$. In order to use the Full Spectrum Method for quality control purposes, we recommend users characterize their own dye, as our measurements of the HI^- and I^{2-} species spectra for purified mCP may contain systematic errors specific to our spectrophotometer and there may also be slight differences in the purity of different batches of purified mCP. However, for the sake of consistency in reporting, we recommend using the properties of Liu et al. (2011) for pH determination with purified mCP until better information on the properties of mCP is available.

When characterizing the indicator, the precautions we have outlined in this manuscript should be followed.

- The I^{2-} spectrum should be measured in a modified synthetic seawater medium, and the lamp intensities should be checked prior to measurements.
- Repeated measurements in NaOH solutions over time should be made to estimate the intermediate precision and verify that no significant drifts have occurred.
- Validation measurements should be made on seawater samples over a range of pH to check for systematic errors in the determination of the HI^- and I^{2-} spectra or other instrumental problems, based on examination of the residuals from the Full Spectrum Method.

Our study also identified potential systematic errors in the current published properties of purified mCP. A project is currently underway at the National Institutes of Standards and Technology to produce a standard reference material for mCP, which will be certified in its purity, optical absorption properties, and dissociation behavior. This effort will help constrain

the uncertainty in the properties of mCP. An indicator reference material would also facilitate the use of full spectrum methods and the development of pH quality control procedures, particularly if the data for the full species spectra is available. For instance, the HI^- and I^{2-} spectra of the reference dye can be used to apply various calibration transfer approaches (Workman, 2018) to correct for instrumental differences. A reference value for the $pK_a(HI^-)$ of pure mCP with a well-defined uncertainty will also be useful for implementing the Full Spectrum Method and for users of different bandpass instruments for which the use of $p(K_a(HI^-) \cdot e_2)$ is inappropriate. We therefore encourage further development of the methods presented in this paper.

3.7 Acknowledgments

Chapter 3, in part, is currently being prepared for submission for publication of the material. Fong, M.B., Branham, C.W., and Dickson, A.G. The dissertation author was the primary investigator and author of this paper.

The authors would like to thank Trisha Nguyen for performing the preliminary experiments on characterizing *m*-cresol purple in NaOH solutions, Manuel Belmonte for performing the phosphate buffer measurements, and Hugh Doyle for his early work on evaluating the Liu et al. characterization of purified mCP.

3.8 References

Álvarez, M., Fajar, N.M., Carter, B.R., Guallart, E.F., Pérez, F.F., Woosley, R.J. and Murata, A., 2020. Global Ocean Spectrophotometric pH Assessment: Consistent Inconsistencies. Environmental Science & Technology, 54(18): 10977-10988.

Buck, R.P., Rondinini, S., Covington, A.K., Baucke, F.G.K., Brett, C.M.A., Camoes, M.F., Milton, M.J.T., Mussini, T., Naumann, R., Pratt, K.W., Spitzer, P. and Wilson, G.S., 2002. Measurement of pH. Definition, standards, and procedures (IUPAC Recommendations 2002). Pure and Applied Chemistry, 74(11): 2169-2200.

Byrne, R.H. and Breland, J.A., 1989. High precision multiwavelength pH determinations in seawater using cresol red. Deep Sea Research Part A. Oceanographic Research Papers, 36(5): 803-810.

Carter, B.R., Feely, R.A., Williams, N.L., Dickson, A.G., Fong, M.B. and Takeshita, Y., 2018. Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate. Limnology and Oceanography: Methods, 16(2): 119-131.

Carter, B.R., Radich, J.A., Doyle, H.L. and Dickson, A.G., 2013. An automated system for spectrophotometric seawater pH measurements. Limnology and Oceanography: Methods, 11(1): 16-27.

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

DeGrandpre, M.D., Spaulding, R.S., Newton, J.O., Jaqueth, E.J., Hamblock, S.E., Umansky, A.A. and Harris, K.E., 2014. Considerations for the measurement of spectrophotometric pH for ocean acidification and other studies. Limnology and Oceanography: Methods, 12(12): 830-839.

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

Douglas, N.K. and Byrne, R.H., 2017. Achieving accurate spectrophotometric pH measurements using unpurified meta-cresol purple. Marine Chemistry, 190: 66-72.

Fong, M.B. and Dickson, A.G., 2019. Insights from GO-SHIP hydrography data into the thermodynamic consistency of CO_2 system measurements in seawater. Marine Chemistry, 211: 52-63.

Gemperline, P., 2006. Principal Component Analysis. In: P. Gemperline (Editor), Practical Guide to Chemometrics, Second Edition. CRC Press, Boca Raton, FL, pp. 69-104.

Husheer, S.L.G., 2001. On spectrophotometric pH measurement in seawater media, University of Otago, Dunedin, New Zealand, 145 pp.

Kramer, R., 1998. Chemometric Techniques for Quantitative Analysis. Marcel Dekker, Inc., New York, NY, 203 pp.

Liu, X., Patsavas, M.C. and Byrne, R.H., 2011. Purification and Characterization of meta-Cresol Purple for Spectrophotometric Seawater pH Measurements. Environmental Science & Technology, 45(11): 4862-4868.

Loucaides, S., Rèrolle, V.M.C., Papadimitriou, S., Kennedy, H., Mowlem, M.C., Dickson, A.G., Gledhill, M. and Achterberg, E.P., 2017. Characterization of meta-Cresol Purple for spectrophotometric pH measurements in saline and hypersaline media at sub-zero temperatures. Scientific Reports, 7(1): 2481.

Müller, J.D. and Rehder, G., 2018. Metrology of pH Measurements in Brackish Waters—Part 2: Experimental Characterization of Purified meta-Cresol Purple for Spectrophotometric pHT Measurements. Frontiers in Marine Science, 5(177).

Ohline, S.M., Reid, M.R., Husheer, S.L.G., Currie, K.I. and Hunter, K.A., 2007. Spectrophotometric determination of pH in seawater off Taiaroa Head, Otago, New Zealand: Full-spectrum modelling and prediction of pCO_2 levels. Marine Chemistry, 107(2): 143-155.

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. and 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(2): 297-323.

Pearson, K., 1901. LIII. On lines and planes of closest fit to systems of points in space. The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, 2(11): 559-572.

Poerio, D.V. and Brown, S.D., 2018. Dual-Domain Calibration Transfer Using Orthogonal Projection. Appl Spectrosc, 72(3): 378-391.

Shimada, T. and Hasegawa, T., 2017. Determination of equilibrium structures of bromothymol blue revealed by using quantum chemistry with an aid of multivariate analysis of electronic absorption spectra. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 185: 104-110.

Shimada, T., Tochinai, K. and Hasegawa, T., 2019. Determination of pH Dependent Structures of Thymol Blue Revealed by Cooperative Analytical Method of Quantum Chemistry and Multivariate Analysis of Electronic Absorption Spectra. Bulletin of the Chemical Society of Japan, 92(10): 1759-1766.

Takeshita, Y., Johnson, K.S., Coletti, L.J., Jannasch, H.W., Walz, P.M. and Warren, J.K., 2020. Assessment of pH dependent errors in spectrophotometric pH measurements of seawater. Marine Chemistry, 223: 103801.

Takeshita, Y.W., J.K.; Liu, X.; Spaulding, R.S.; Byrne, R.H.; Carter, B.R.; Degrandpre, M.D.; Murata, A.; Watanabe, S., submitted. Consistency and stability of purified meta-cresol purple for spectrophotometric pH measurements in seawater. Marine Chemistry.

Vogel, A.I.A., 1961. A Textbook of Quantitative Inorganic Analysis including Elementary Instrumental Analysis. Lowe & Brydone (Printers) Ltd, London.

Williams, N.L., Juranek, L.W., Feely, R.A., Johnson, K.S., Sarmiento, J.L., Talley, L.D., Dickson, A.G., Gray, A.R., Wanninkhof, R., Russell, J.L., Riser, S.C. and Takeshita, Y., 2017. Calculating surface ocean pCO_2 from biogeochemical Argo floats equipped with pH: An uncertainty analysis. Global Biogeochemical Cycles, 31(3): 591-604.

Wold, S., Antti, H., Lindgren, F. and Öhman, J., 1998. Orthogonal signal correction of near-infrared spectra. Chemometrics and Intelligent Laboratory Systems, 44(1): 175-185.

Workman, J.J., Jr., 2018. A Review of Calibration Transfer Practices and Instrument Differences in Spectroscopy. Appl Spectrosc, 72(3): 340-365.

Yao, W., Liu, X. and Byrne, R.H., 2007. Impurities in indicators used for spectrophotometric seawater pH measurements: Assessment and remedies. Marine Chemistry, 107(2): 167-172.

Zhang, H. and Byrne, R.H., 1996. Spectrophotometric pH measurements of surface seawater at in-situ conditions: absorbance and protonation behavior of thymol blue. Marine Chemistry, 52(1): 17-25..

Chapter 4

Evaluation of indicator perturbation corrections for spectrophotometric pH measurements in seawater

4.1 Abstract

Spectrophotometric pH measurements require correction for the effect of the indicator dye on the sample pH. Corrections can be determined empirically by extrapolating measurements to zero dye concentration for each individual sample or inferring the necessary adjustments from a correction curve based on dye additions (using a single batch of dye solution) to samples over a range of pH. Alternatively, the pH perturbations can be estimated from a chemical equilibrium model. This study evaluated the uncertainty of dye correction approaches with a combination of laboratory experiments, chemical modeling, and numerical simulations. We showed from numerical simulations that random errors in absorbance measurements typically dominate the total uncertainty of empirical corrections. Absorbance-dependent absorbance errors, which may have affected our measurements, can also contribute to error in the correction. We found small, but meaningful inconsistencies between measured and modeled pH perturbations that seemed to be consistent with this hypothesis. Additionally, the correction curve approach has a systematic error from the assumption of linearity, but this error only becomes large at the high dye concentrations used in short pathlength cells (>0.01 pH units in a 1 cm. cell) and can be minimized by using a low pH dye solution, which results in a more linear correction curve. When optimized, we estimate that the empirical corrections contribute <0.002 to the uncertainty in spectrophotometric pH for 5 and 10 cm cells.

4.2 Introduction

Spectrophotometric measurements of seawater pH, using an indicator dye, have become increasingly popular in the oceanographic community due to its distinct advantages of being a simple, highly precise, and calibration-free method for measuring pH. The spectrophotometric approach has been applied to a wide variety of systems for seawater pH measurements, including semi-automated benchtop systems for discrete sampling (Carter et al., 2013), continuous flow and underway systems (Bellerby et al., 2002; Tapp et al., 2000), and autonomous sensors used in moorings (Seidel et al., 2008). Its apparent robustness and excellent short-term precision (repeatability of ~0.0004 in pH, Clayton and Byrne, 1993) also makes spectrophotometric pH potentially suitable for studying decadal changes in open ocean pH related to the uptake of anthropogenic CO_2 from the atmosphere (Byrne et al., 2010) and for calculating other parameters of the CO_2 system with good precision (Clayton et al., 1995; McElligott et al., 1998; Patsavas et al., 2015).

Because the indicator dye (typically a sulfonephthalein compound) itself is an acid, the addition of dye to a seawater sample will necessarily change the sample pH. The magnitude of perturbation to the sample pH will depend on the amount of dye added to the sample as well as the difference between the composition of the dye stock solution and the sample (Chierici et al., 1999). If the pH of the dye stock solution is not adjusted to a value close to pH of the samples, the perturbation to sample pH can be as large as 0.005 pH units in a 10 cm cell (Clayton and Byrne, 1993). In shorter pathlength cells, the perturbation would be proportionately larger (e.g., twice as large in a 5 cm cell), as more dye is required to attain an adequate absorbance. Accurate determination of the sample pH therefore requires a correction for the effect of the indicator dye solution.

Several methods for dye perturbation corrections have been proposed. The correction for every individual sample can be determined by extrapolating data from measurements made at two or more dye concentrations to zero dye concentration. This approach is most convenient in continuous flow systems (e.g., Aßmann et al., 2011; Martz et al., 2003; Seidel et al., 2008), which take numerous absorbance measurements along a gradient of indicator passing through the flow cell. Another approach is to measure the perturbation in a subset of samples spanning a range of pH to construct a linear correction curve (specific to a single batch of dye stock solution) from which the perturbation for an individual sample with a particular pH can be inferred (Clayton and Byrne, 1993). The sample pH perturbation can also be estimated from a chemical equilibrium model of seawater acid-base systems with the inclusion of the indicator dye as an additional acid-base system (Chierici et al., 1999). The model-based approach requires that the composition of the dye stock solution as well as of the sample are known.

As these approaches involve different assumptions about the behavior of the dye over a range of measurement conditions, there could be differences in the final pH result and in its uncertainty depending on the correction method implemented. In this paper, we evaluate the contribution of dye perturbation corrections to the overall uncertainty of spectrophotometric pH measurements using a combination of laboratory experiments, chemical modeling, and numerical simulations. First, we developed a chemical equilibrium model, similar to that of Chierici et al. (1999), for estimating the dye perturbation to sample pH and evaluated the overall uncertainty of model-estimated pH perturbations based on uncertainty contributions from the individual model input parameters. We then performed a series of dye addition experiments with multiple batches of carefully characterized dye stock solutions and seawater samples of well-known composition and compared the measured dye perturbations to the

modeled perturbations. Finally, we used the equilibrium model to evaluate the implications of the assumptions inherent in the extrapolation and correction curve approaches and to estimate their uncertainty, considering both systematic and random contributions. Our results provide insights into how the uncertainty can be minimized when implementing dye perturbation corrections.

4.3 Materials and Procedures

4.3.1 Theory

Developed for use in seawater in the late 1980s, spectrophotometric pH determination is based on the measurement of the absorption spectrum of a solution to which an indicator dye has been added (Byrne and Breland, 1989; Clayton and Byrne, 1993; Zhang and Byrne, 1996) . Sulfonephthalein indicators, such as cresol red, *m*-cresol purple, and thymol blue, are diprotic acids, but at seawater pH, the second dissociation is the predominant reaction.

$$HI^- \rightarrow H^+ + I^{2-} \quad (1)$$

Therefore, if the acid dissociation constant (pK_a) and the speciation of the dye in the sample are known, the pH of the sample can be determined from the Henderson-Hasselbalch equation.

$$pH = pK_a(HI^-) + \log\left(\frac{[I^{2-}]}{[HI^-]}\right) \quad (2)$$

Because the absorption spectra of the acidic and basic forms of the indicator are substantially different, the composite spectrum of a dye solution can be used to estimate the concentration ratio ($[I^{2-}]/[HI^-]$). The absorbance (A) of a sample with dye in a cell of pathlength ℓ at each wavelength (λ) is the sum of absorbance contributions from the acidic and basic dye species as well as the sample background (B_λ) and instrumental error (e_λ), as given by the Beer-Lambert Law, where ε is a molar absorption coefficient.

$$\frac{A_\lambda}{\ell} = \varepsilon_\lambda(\mathrm{HI}^-)[\mathrm{HI}^-] + \varepsilon_\lambda(\mathrm{I}^{2-})[\mathrm{I}^{2-}] + B_\lambda + e_\lambda \quad (3)$$

Most commonly, the concentration ratio ($[I^{2-}]/[HI^-]$) is determined from the ratio of the absorbances at the two wavelengths corresponding to the maximum absorbance of the basic and acidic dye species. For *m*-cresol purple, the most widely used indicator in seawater, the fully deprotonated species (I^{2-}) has its maximum absorbance at 578 nm, and the singly protonated species (HI^-) has its maximum absorbance at 434 nm. Provided that the absorption coefficients of each species are known at these wavelengths and that the sample background and instrumental errors have been corrected for, the ratio ($[I^{2-}]/[HI^-]$) is related to the ratio of the absorbances at these two wavelengths (i.e., $R = A_{578}/A_{434}$) by rearrangement of **Eq. 3**.

$$\frac{[\mathrm{I}^{2-}]}{[\mathrm{HI}^-]} = \frac{R - \varepsilon_{578}(\mathrm{HI}^-)/\varepsilon_{434}(\mathrm{HI}^-)}{\varepsilon_{578}(\mathrm{I}^{2-})/\varepsilon_{434}(\mathrm{HI}^-) - R\varepsilon_{434}(\mathrm{I}^{2-})/\varepsilon_{434}(\mathrm{HI}^-)} = \frac{R - e_1}{e_2 - Re_3} \quad (4)$$

In **Eq. 4**, e_1, e_2, and e_3 are ratios of the molar absorption coefficients of the I^{2-} and HI^- species. Liu et al. (2011) calculate the pH of a sample solution with indicator from this equation rearranged as,

$$\mathrm{pH} = -\log(K_a(\mathrm{HI}^-) \cdot e_2) + \log\left(\frac{R - e_1}{1 - R\frac{e_3}{e_2}}\right) \quad (5).$$

and provide the temperature and salinity dependence for each of terms in **Eq. 5** for purified *m*-cresol purple.

4.3.2 Dye perturbation corrections

The original approach of Clayton and Byrne (1993) quantified the dye perturbation behavior for an individual batch of dye stock solution and used this information to adjust the measured absorbance ratio of each sample solution. In this approach, the absorbance ratios from two separate and approximately equal additions of dye solution to a sample were measured to calculate $\Delta R/\Delta V$, the change in R per unit volume of dye solution added. From a

number of such measurements in samples over a range of pH, the data for $\Delta R/\Delta V$ were fitted to a linear function of R after the first dye addition (i.e., $\Delta R/\Delta V = aR'+b$). The $\Delta R/\Delta V$ can be estimated from the fitted line for an individual sample and used to calculate the R that would have been observed in the absence of dye perturbation via **Eq. 6**, where V is the volume of dye solution added.

$$R = R' - V(aR' + b) \quad (6)$$

Carter et al. (2013) took a similar approach, but used the absorbance at the isosbestic wavelength (488.1 nm for purified m-cresol purple at 25°C, Liu et al., 2011) instead of volume of dye solution added as a proxy for the dye concentration in the sample. A key assumption with these correction curve approaches is that $\Delta R/\Delta[\text{mCP}]_T$ (the change in R per unit change in total dye concentration) is a linear function of R. We assess the validity of this assumption and discuss later the implications for systematic errors in dye perturbation corrections.

A second approach to dye perturbation corrections, as mentioned earlier, is to correct individual samples by extrapolating data from measurements at two or more dye concentrations to zero dye concentration. Such an approach can be implemented in several different ways—by extrapolating R or pH to zero dye concentration or to zero absorbance at the isosbestic wavelength (i.e., $A_{iso} = 0$). Like the first approach, the extrapolation approach assumes that R or pH change linearly with dye concentration and, additionally, that extrapolating R or pH is equivalent.

4.3.3 Estimating dye perturbation from an equilibrium model

The magnitude of dye perturbation can be estimated from an equilibrium model of acid-base systems in seawater by calculating the difference between the pH of the sample with and without indicator. We used CO2SYS for MATLAB (van Heuven et al., 2011), with all

subroutines modified to include the indicator equilibria of *m*-cresol purple, and the data for the equilibrium constants, etc. in **Table 4.1** to perform these calculations. For a seawater sample without any added indicator, its pH can be calculated from its dissolved inorganic carbon (C_T) and total alkalinity (A_T), along with other information such as salinity (used to estimate the equilibrium constants and total boron), total phosphate, and total silicate. For a seawater sample containing a small amount of indicator, its pH can be calculated similarly by including the indicator as an additional acid-base system in the definition of A_T. Thus, the expression for A_T for a seawater sample with indicator becomes

$$\begin{aligned} A_T = {} & [\mathrm{HCO_3^-}] + 2[\mathrm{CO_3^{2-}}] + [\mathrm{B(OH)_4^-}] + [\mathrm{OH^-}] \\ & + 2[\mathrm{PO_4^{3-}}] + [\mathrm{SiO(OH)_3^-}] + [\mathrm{NH_3}] + [\mathrm{HS^-}] + [\mathrm{I^{2-}}] \ldots \\ & - [\mathrm{H^+}] - [\mathrm{HSO_4^-}] - [\mathrm{HF}] - [\mathrm{H_3PO_4}] + \ldots \end{aligned} \quad (7).$$

For oxygenated seawater, the concentrations of NH_3 and HS^- are negligible. The contribution of the indicator to A_T (i.e., $[I^{2-}]$) is a function of the total dye concentration ($[mCP]_T$) in the sample, $K_a(HI^-)$, and $[H^+]$.

$$[\mathrm{I^{2-}}] = \frac{[\mathrm{mCP}]_T K_a(\mathrm{HI^-})}{[\mathrm{H^+}] + K_a(\mathrm{HI^-})} \quad (8)$$

Table 4.1. Parameters used in an equilibrium model (modified from CO2SYS-MATLAB) to estimate the magnitude of indicator perturbation on the sample pH. The various constants listed here are explicitly chosen in CO2SYS-MATLAB. Other constants not listed are implicitly chosen in the program. Values for the estimated standard uncertainties are from Orr et al. (2018), unless otherwise stated. Values in parentheses are the estimated uncertainties for constants in a 0.7 mol/kg NaCl medium.

Parameter	Source	Estimated Standard Uncertainty
pH_{dye}	Spectrophotometric pH	0.01
$A_{T,dye}$	Titration of seawater-dye mixture	38 μmol kg^{-1} [a]
$[mCP]_T$	Total dye concentration in stock solution	20 μmol kg^{-1} [b]
A_T	CO_2-in-seawater Reference Material	0.5 – 0.75 μmol kg^{-1} [c]
C_T	CO_2-in-seawater Reference Material	0.36 – 0.92 μmol kg^{-1} [c]
$pK_a(HI^-)$	Unpublished data; Millero et al. (2009)	0.0067[d] (0.0070[e])
pK_1	Lueker et al. (2000); Dyrssen and Hansson (1973)	0.0075
pK_2	Lueker et al. (2000); Dyrssen and Hansson (1973)	0.015
pK_B	Dickson (1990)	0.01
pK_W	Millero (1979); Dyrssen and Hansson (1973)	0.01
B_T / S	Lee et al. (2010)	0.02 (*relative uncertainty)*

[a] Pooled standard deviation of titrations from three batches of dye solution.

[b] A systematic discrepancy between the gravimetrically known concentrations and the concentrations expected from the isosbestic absorbances in the dye addition experiments suggested a concentration error of 40 μmol kg^{-1} for one stock solution, which we use as a 2*u* estimate for the standard uncertainty.

[c] Range of reported standard deviations from the Certificates of Analysis for CRMs 160, 162, 164, and 172.

[d] Standard deviation of three independent determinations of the pK_a of purified m-cresol purple from Liu (*pers. communication*), DeGrandpre et al. (2014), and our unpublished data.

[e] Estimated from a sum of quadrature of the uncertainty in (d) and a contribution from systematic error in the absorption coefficients of m-cresol purple due to the use of an impure dye in Millero et al. (2009).

Additionally, when calculating the pH of a seawater sample containing indicator, it is necessary to use values of C_T and A_T resulting from the mixture of the sample and the dye stock solution, as in **Eq. 9** and **Eq. 10,** where f_{samp} and f_{dye} are the fraction of sample and dye stock solution in the mixture, respectively.

$$C_{\mathrm{T,samp+dye}} = f_{\mathrm{samp}} C_{\mathrm{T,samp}} + f_{\mathrm{dye}} C_{\mathrm{T,dye}} \quad (9)$$

$$A_{\mathrm{T,samp+dye}} = f_{\mathrm{samp}} A_{\mathrm{T,samp}} + f_{\mathrm{dye}} A_{\mathrm{T},dye} \quad (10)$$

Therefore, the magnitude of the dye perturbation (ΔpH) can be calculated from

$$\Delta\mathrm{pH} = \mathrm{pH}\left(C_{\mathrm{T,samp+dye}}, A_{\mathrm{T,samp+dye}}, [\mathrm{mCP}]_{\mathrm{T}}, \ldots\right) - \mathrm{pH}(C_{\mathrm{T,samp}}, A_{\mathrm{T,samp}}, \ldots) \quad (11),$$

where the ellipses stand for the other data required (as described earlier) to calculate pH from C_T and A_T.

4.3.4 Preparation of dye stock solutions and sample materials

Four stock solutions of purified *m*-cresol purple (supplied by Robert H. Byrne, U. of Southern Florida) were prepared for use in dye addition experiments aimed at evaluating the extrapolation and model-based dye perturbation corrections. Three batches of dye solution were prepared in an artificial seawater background, the medium in which the pK_a of purified *m*-cresol purple was determined (Liu et al., 2011), so that the stock solution composition (required as model inputs) can be calculated. We also prepared one batch of dye stock solution in a 0.7 mol kg^{-1} NaCl background for comparison with these other experiments.

The dye stock solutions (2.5 to 4 mol kg^{-1} *m*-cresol purple) were prepared by dissolving the required weight of dye powder (in the molecular form, H_2I) in approximately 50 to 100 g of the background solution with at least 0.01 mol kg^{-1} NaOH. The solutions were stirred for an hour, and then the pH was adjusted with weighed amounts of 1 mol kg^{-1} HCl using a glass electrode calibrated on the NBS scale with buffers traceable to NIST standard reference materials (pH 7 and 10). Three batches of dye solutions were prepared with a pH between 7.1-7.3, and one batch was prepared with a pH of ~8.

The artificial seawater background (at a salinity of 35) were prepared according to the recipe of Dickson et al. (2007). To prevent precipitation of Mg and Ca salts at high pH during

the initial dissolution with NaOH (pH~12), $MgCl_2$ and $CaCl_2$ (in the form of ~1 mol kg^{-1} solutions) were added to the dye solution after the pH adjustment with HCl.

To characterize the composition of the dye solutions, we measured the pH and A_T of the solutions and calculated C_T from pH, A_T, and total dye concentration. The pH of the dye solutions was determined by measuring the R of the solutions in a 0.2 mm pathlength thermostatted flow-through cell at 25°C. An artificial seawater or NaCl solution was used for the measurement of the background spectrum. The A_T of the dye stock solutions was determined by first titrating seawater and then a mixture of dye solution and seawater, following SOP 3b in Dickson et al. (2007), a variant of the open-cell method of Dickson et al. (2003). Three ~130 g subsamples drawn from a large bottle of seawater (~10 L) were titrated, followed by titration of another three subsamples of the same seawater with a small amount of dye solution (~3 g dye solution in 130 g total). The A_T of the dye stock solution can then be calculated from **Eq. 10**.

We verified that our assumption of there being C_T (and carbonate alkalinity—see later discussion) in the dye solutions was correct by directly measuring the C_T (in addition to pH and A_T) in two separate stock solutions of impure *m*-cresol purple with a non-dispersive infrared C_T analyzer system. These solutions were prepared in an artificial seawater background at a pH of ~7.10 and ~7.95. When measuring the C_T of the dye stock solutions, care was taken to minimize contact with the atmosphere by storing the prepared dye solution in a 100 mL glass Tomopal syringe without headspace.

As it was also necessary to have seawater samples with well-defined C_T and A_T for estimating the dye perturbations from the model, we used CO_2-in-seawater Certified Reference Materials (CRMs) prepared in the Dickson Laboratory at the Scripps Institution of

Oceanography at UC San Diego (Dickson, 2010) in the dye addition experiments. Four batches of seawater CRMs were used in the experiments: CRM 160, 162, 164, and 172. Two of these batches, CRM 162 and 164, were modified to have higher C_T and A_T than a typical reference material batch, through bubbling with CO_2 gas or addition of sodium carbonate, similar to the process of Bockmon and Dickson (2015). These four batches provided seawater samples with a pH range of 7.54 to 7.91 at 25°C (**Table 4.2**). Complete information regarding the composition of these batches can be found in the Certificates of Analysis available online (https://www.nodc.noaa.gov/ocads/oceans/Dickson_CRM/batches.html).

Table 4.2. Composition of the CO_2-in-seawater Reference Materials used in the dye addition experiments. The pH values (total scale) calculated from A_T and C_T at 25°C are also given.

Batch	Salinity	C_T µmol kg^{-1}	A_T µmol kg^{-1}	pH(C_T, A_T, …)
160	33.414	2030.39	2212.44	7.84
162	33.312	2177.27	2403.72	7.91
164	33.247	2238.89	2309.32	7.55
172	33.345	2038.99	2217.40	7.83

4.3.5 Dye addition experiments

Dye addition experiments in seawater were performed with each of the four batches of dye solutions (**Table 4.3**). The first two experiments (with Dye 20170210 and Dye 20170317) were aimed at evaluating the performance of the extrapolation approach to dye perturbation correction. Weighed amounts of dye solution were added to multiple bottles (n = 4 to 6) of each batch of seawater (i.e., CRM 160, 162, and 164), and the pH of the seawater-dye mixtures was measured spectrophotometrically. Separate bottles of the same batch of seawater (without dye) were used to measure the background spectrum. The bottles were weighed beforehand and then emptied, dried, and reweighed after measurement to determine the total weight of solution and

thus the dye concentration in the sample. The total dye concentration in the samples ranged from ~2 to ~15 μmol kg^{-1}. Dye 20170210 (pH~7.2) was more acidic than the samples, so the dye additions decreased the sample pH. For Dye 20170317 (pH~8), which had a higher pH than that of the samples, the dye additions had the opposite effect. For each batch of seawater, the pH values from each set of dye addition measurements (one with a low pH dye solution and the other with a high pH dye solution) were extrapolated to zero dye concentration to evaluate the consistency of the perturbation-corrected pH values determined using two separate batches of dye solution.

Table 4.3. Summary of dye addition to seawater experiments. Information is given on the composition of the dye stock solutions, the batches and number of bottles (specified in the parentheses) of CO_2-in-seawater Reference Materials that were analyzed, the cell pathlength (ℓ) used to analyze the seawater samples, and the dye addition method (W = weighed additions to multiple bottles, K = addition with a Kloehn pump to individual bottles). Three of the dye solutions were prepared in an artificial seawater background, and one was prepared in a NaCl background. Measured values for the pH and A_T of the dye solutions are reported as *mean* ± *standard deviation* (number of measurements). Values for $A_{T,dye}$ calculated from **Eq. 12** are also given. The pH of the dye solutions and the seawater samples were both measured at T_{meas}.

Dye solution	$[mCP]_T$ mmol kg^{-1}	pH_{dye}	$A_{T,dye}$ μmol kg^{-1}	$A_{T,dye}$ (calc.) μmol kg^{-1}	T_{meas} (°C)	CRMs analyzed	ℓ (cm.)	Dye add.
20170210 in ASW	3.80	7.1836 ± 0.00060 (3)	n/a	512	25.1	B160 (7)	10	W
						B162 (7)	10	W
						B164 (6)	5	W
20170317 in ASW	2.63	8.0097 ± 0.00009 (3)	2739 ± 38 (3)	1352	25	B160 (7)	5	W
						B162 (5)	5	W
						B164 (6)	5	W
20180405 in ASW	2.95	7.1392 ± 0.00017 (3)	709 ± 39 (4)	364	25	B164 (3)	5	K
						B172 (4)	5	W
						B172 (8)	5	K
20180518 in NaCl	2.94	7.2833[a] ± 0.00077 (4)	777 ± 25 (3)	360	25	B160 (1)	5	K
						B164 (2)	5	K

[a] This pH value is on the free hydrogen ion scale and was calculated using the $pK_a(HI^-)$ of Millero et al. (2009) and values of e_1, e_2, and e_3 for purified m-cresol purple provided by Liu (*pers. communication*).

The spectrophotometric pH measurements were made with an Agilent 8453 spectrophotometer using the automated method of Carter et al. (2013). Measurements were

made at 25°C and a gauge pressure of 0 dbar (ambient atmospheric pressure), except for one set of measurements which was made at 25.1°C. For comparison with the other data, this set of data was converted to 25°C using the pH values measured at 25.1°C, the C_T values of the samples, and the data in **Table 4.1**. The amount of dye solution added to the samples was chosen with the goal of keeping the absorbances at 434 and 578 nm within the range of 0.2–1.2. Initially, we used a 10 cm cell for two sets of measurements, but later switched to a 5 cm cell to allow measurements at higher dye concentrations while keeping within the desired absorbance range.

The slopes of the lines fitted through each set of dye addition measurements (pH vs. total dye concentration) were compared to the slopes predicted from the dye perturbation model. However, the data from the experiment using Dye 20170210 was excluded as the composition for that dye solution (required as model inputs) was not fully characterized. In the last two experiments (with Dye 20180405 and 20180518), we also modified our dye addition method, using the Kloehn syringe pump on the automated pH system to deliver the dye to subsamples drawn from individual bottles (four subsamples per bottle). The total dye concentration in these samples was inferred from the A_{iso} using data from the earlier experiments in which A_{iso} was measured in samples with weighed amounts of dye solution. This modification to the dye addition method reduced the uncertainty in measuring the slope compared to the initial method of weighed additions to multiple bottles of a seawater batch, which had an added source of variance from the bottle-to-bottle variability in pH (see later discussion).

4.3.6 Calculating the slope of pH versus total dye concentration from a dye perturbation model

For each sample in **Table 4.3** (excluding those measured with Dye 20170210), values of ΔpH were calculated from the dye perturbation model (**Eq. 11**) at 16 discrete dye concentrations (ranging from 0 – 18 μmol/kg), and the slope $\Delta pH/\Delta[mCP]_T$ was estimated from a least squares line fit through the data. These model-estimated slopes were then compared with the experimentally measured slopes, as described in the previous section.

Calculating the slope $\Delta pH/\Delta[mCP]_T$ from the dye perturbation model requires estimates of the dye stock solution composition ($[mCP]_T$, C_T, and A_T) as input parameters. As our dye solutions were prepared both in an artificial seawater background and in a NaCl background, these were treated differently in the model calculations. For a dye solutions made in an artificial seawater background according to the recipe of Dickson et al. (2007), which does not include borate, fluoride, phosphate, and silicate, and assuming $C_T = 0$, the expression for the A_T of such a dye solution simplifies to

$$A_T = [I^{2-}] + [OH^-] - [H^+] - [HSO_4^-] \quad (12).$$

For dye solutions in NaCl background, the expression for A_T does not include HSO_4^-. The concentration of I^{2-} was calculated from **Eq. 8**, using the total dye concentration of the stock solution and a value for $K_a(HI^-)$ appropriate to the background medium (see **Table 4.1**). Additionally, the model calculations included the dilution of the total boron, fluoride, phosphate, and silicate of the sample by the dye stock solution, and for the addition of a dye solution prepared in a NaCl background, the calculations included the dilution of the total sulfate of the sample. It should be noted that our approach in estimating the C_T of the dye solution from A_T and pH assumes that the A_T of the dye solution has a contribution from

carbonate alkalinity. Thus, we use the following expression for the A_T of our dye solutions in an artificial seawater background:

$$A_T = [HCO_3^-] + 2[CO_3^{2-}] + [I^{2-}] + [OH^-] - [H^+] - [HSO_4^-] \quad (13).$$

A possible source of carbonate alkalinity in our dye solutions may have been from sodium carbonate contamination in the NaOH pellets we used to dissolve the dye. We discuss later the plausibility of the carbonate alkalinity assumption and its implications for our uncertainty in estimating the dye stock composition.

To assess whether discrepancies between the modeled and measured slopes were statistically significant, the uncertainty of the difference was calculated by adding in quadrature the estimate of the expanded uncertainty (95% confidence level) of the model slope and the 95% confidence interval of the measured slope (i.e., the standard error, s_{meas}, multiplied by the appropriate t value), as in **Eq. 14**.

$$U_{model-meas} = \sqrt{U_{model}^2 + (t \cdot s_{meas})^2} \quad (14)$$

Discrepancies with magnitudes smaller than the estimated uncertainty of the difference are not significantly different from zero at a 95% confidence level.

The uncertainty in the model-predicted slope was estimated by considering how uncertainties in the various model input parameters, diagrammed in **Figure 4.1**, combine into an overall uncertainty in the calculated slope. The sources of uncertainty we considered include the carbonic acid dissociation constants (K_1 and K_2), the dissociation constant of boric acid (K_B), the total boron to salinity ratio (B_T/S), the acid dissociation constant of the m-cresol purple ($K_a(HI^-)$), the total dye concentration in the dye stock solutions, the pH and A_T of the dye stock solution, and the C_T and A_T of the seawater sample. We neglected uncertainties from the equilibrium constants K_W, K_{Si}, K_{1P}, K_{2P}, and K_{3P}, as these contribute negligibly to the

uncertainty of pH calculated from A_T and C_T (Orr et al., 2018). Estimates of the standard uncertainties in each of these parameters are given in **Table 4.1**. The contributions of each of these sources of uncertainty (x_i) were summed in quadrature to calculate the combined standard uncertainty (u_c) in the calculated slope (y), as in **Eq. 15**.

$$u_c\big(y(x_1, x_2, \ldots, x_n)\big) = \sqrt{\sum_{i=1,n}\left(\frac{\partial y}{\partial x_i}\right)^2 u(x_i)^2} \quad (15)$$

In **Eq**. **15**, each of the sensitivity terms ($\partial y/\partial x_i$) were evaluated numerically. The expanded uncertainty of the model slope (at a 95% confidence level) is $2u_c$.

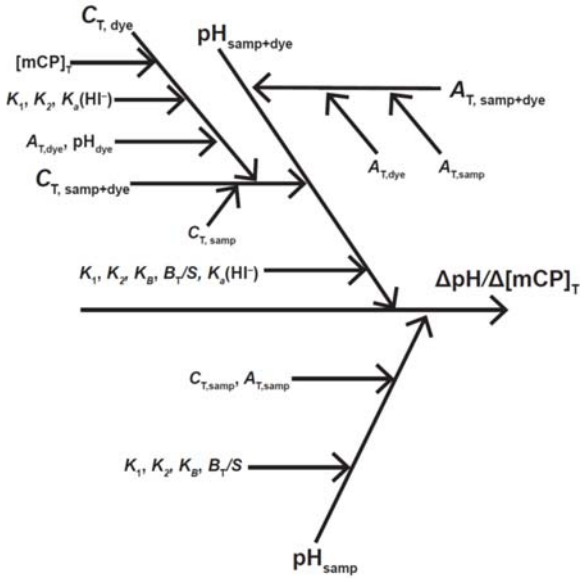

Figure 4.1. Diagram of contributions to uncertainty in the slope of pH versus total dye concentration as estimated from an equilibrium model of acid-base systems in seawater with the inclusion of m-cresol purple.

4.3.7 Simulations of systematic and random errors in empirical dye corrections

Empirical dye corrections infer the pH of the sample without dye based on measurements of the change in sample pH due to dye addition. Sources of systematic and random errors in the dye perturbation correction include the absorbance measurements used to infer pH, the way dye concentration in the samples is estimated, and assumptions about the

linearity of the dye perturbation (**Figure 4.2)**. Additionally, in the correction curve approach, it is assumed that a single correction curve is representative of the required adjustments for samples over a range of compositions. We used the dye perturbation model to simulate how systematic and random errors affect the extrapolation and correction curve approaches when implemented with double dye additions to seawater. In these simulations, the dye perturbation model was used to calculate the pH and *R* values at two different dye concentrations for seawater samples (S = 35, A_T = 2400 μmol kg^{-1}) at 11 different pH values ranging from 7.2 – 8.2. The model calculations were done for a 2 mmol kg^{-1} mCP stock solution with C_T = 0 and a pH of 7.2 and 7.7. For each seawater sample, the simulated pH and *R* values at the two dye concentrations were extrapolated to zero dye concentration to estimate the pH of the sample without dye. Values of $\Delta R/\Delta[\text{mCP}]_T$ were also calculated from the same data and regressed against the *R* values from the first dye addition to derive a dye correction curve for estimating the required adjustment in *R* for a sample measured at a third dye concentration halfway between the first two dye concentrations (see **Eq. 6**). The difference between the pH values estimated from these three approaches (i.e., extrapolating pH, extrapolating *R*, or from a dye correction curve) and the true pH value is the systematic error of the correction method (see **Figure 4.3**). The systematic errors in these three approaches were evaluated for 1, 5, and 10 cm. cells by scaling the dye concentrations used in the calculations. For a 5 cm. cell, the double dye additions were evaluated at 6 and 12 μmol kg^{-1} mCP, and the dye correction curve was used to adjust the *R* values of samples measured at 9 μmol kg^{-1} mCP.

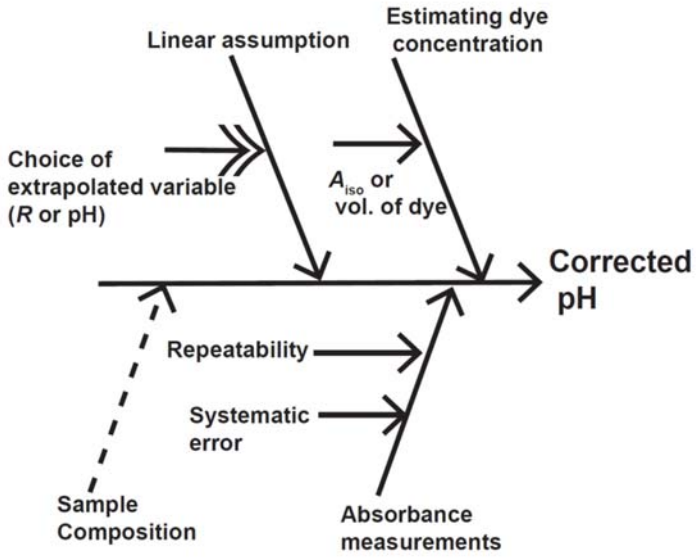

Figure 4.2. Diagram of contributions to uncertainty in empirical dye perturbation corrections. The dashed arrow indicates an uncertainty contribution unique to the correction curve approach. The double-tipped arrow indicates an uncertainty contribution unique to the extrapolation approach.

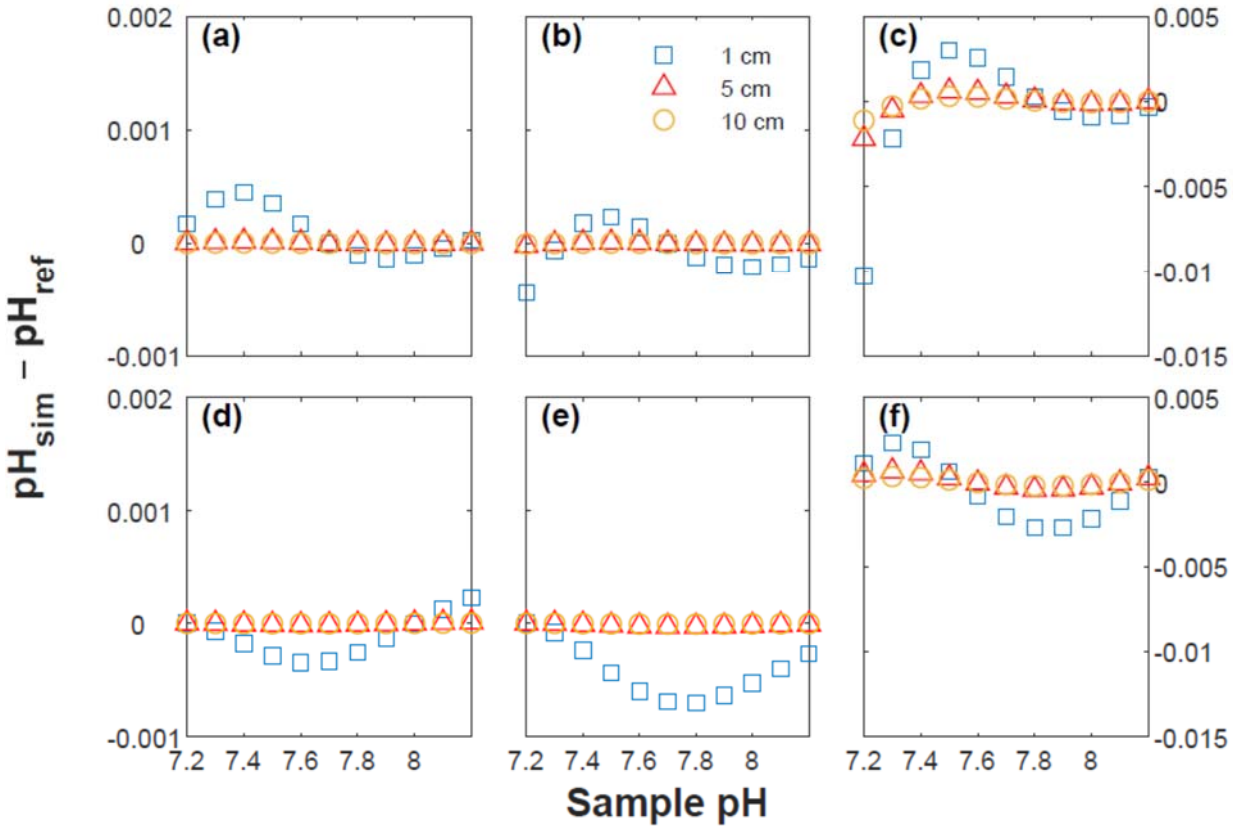

Figure 4.3. Estimated systematic errors for three empirical dye correction approaches : extrapolating pH (left panels) or *R* values (center panels) from a double dye addition to zero dye concentration and the use of a dye correction curve constructed from double dye additions to samples across a range in pH (right panels). The errors shown are the difference between the sample pH, estimated by applying the correction approaches to the model-simulated pH values at various dye concentrations, and the true reference value of the sample pH. Errors are calculated for the addition of a 2 mmol kg^{-1} dye solution with $C_T = 0$ and a pH of 7.7 (top panels) or 7.2 (bottom panels) to seawater samples with S=35, A_T = 2400 µmol kg^{-1}, and pH ranging from 7.2 – 8.2. The different symbols show the estimated errors for different pathlength cells. Note that the scale of the y-axis is different for the right panels.

To evaluate the error in using the correction curve approach to correct samples with a range of compositions, we simulated a dye correction curve for a pH 7.2 dye solution (2 mmol kg^{-1} mCP, $C_T = 0$) using values of C_T and A_T of seawater samples collected on the GO-SHIP repeat hydrography cruise 2015 P16N (Expocode: 320620140320) to calculate $\Delta R/\Delta[\text{mCP}]_T$ from the dye perturbation model, as in the previous simulation (**Figure 4.4**). The error in the correction, shown in **Figure 4.5**, was calculated from the difference between the pH calculated from the correction-curve adjusted *R* values (based on the line fit through the data in **Figure 4.4**

and **Eq. 6**) and the "true" pH values of the samples without dye, calculated from their A_T and C_T values.

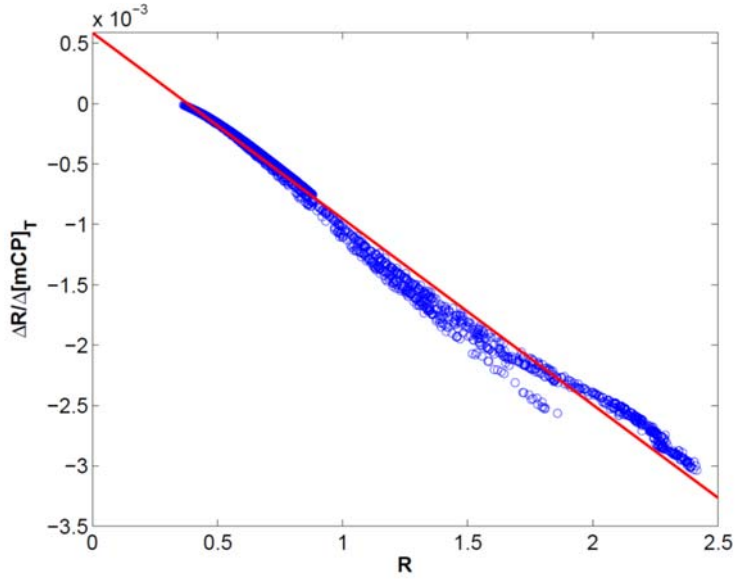

Figure 4.4. Model-simulated dye correction curve for a dye solution with a concentration of 2 mmol kg^{-1} mCP, C_T = 0, and pH of 7.2. Simulated values of $\Delta R/\Delta[\text{mCP}]_T$ are calculated for a range of natural seawater samples on the 2015 P16N cruise and plotted as a function of the R of the samples. The red line is a least squares line fit through the data.

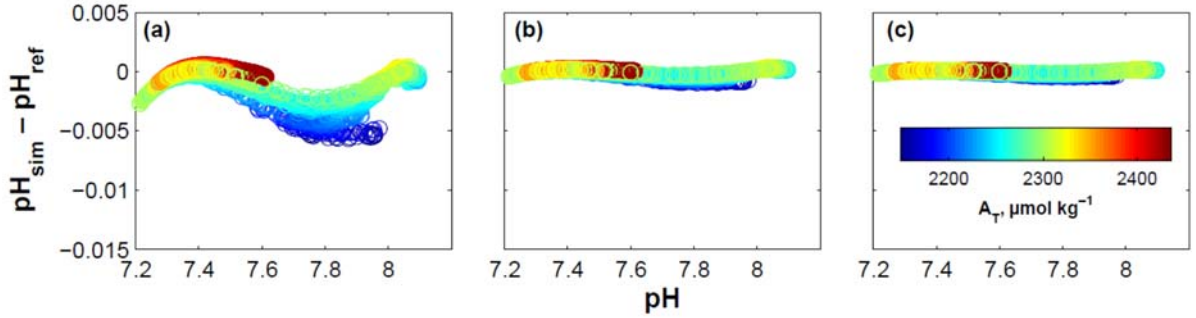

Figure 4.5. Estimated systematic errors in pH (difference from true value) when using the line fit to the data in **Figure 4.4** to estimate the required adjustment to the R of natural seawater samples with a range of A_T values. The errors are estimated for (a) 1 cm., (b) 5 cm., and (c) 10 cm. cells.

Next, we performed a Monte Carlo simulation to assess the effects of random errors in absorbance measurements and in determining the amount of dye added to the sample. The

details of this simulation are similar to the earlier simulation but with the addition of a normally-distributed random error to the absorbances at 434 nm, 578 nm, and the isosbestic wavelength (in the Carter et al., 2013 approach) or to the volume of dye solution pipetted into the cell in the Clayton and Byrne (1993) approach. For each of the 11 seawater samples (as in the previous simulations), the double dye additions were simulated 10,000 times using a random uncertainty, expressed as a standard deviation, of 0.00031 for the absorbances at 434 nm and 578 nm (an estimate of the repeatability of the absorbance measurements in the Agilent 8453; Carter et al., 2013) and 0.00044 for A_{iso} (a weighted average of the absorbances at two diodes). The uncertainty when using pipetted volumes of dye solution was estimated assuming that an Eppendorf pipette was used to pipette 5 µL (with a random uncertainty of 0.04 µL, according to the manufacturer specifications) of dye solution into the cell twice for a double dye addition and a total of 7.5 µL (with an uncertainty of 0.044 µL, estimated from the manufacturer specifications) of dye solution when using a dye correction curve to adjust a sample measured at a third dye concentration. The random uncertainty of each correction approach, shown in **Figure 4.6**, was estimated from the standard deviation of the simulated pH values (n = 10,000).

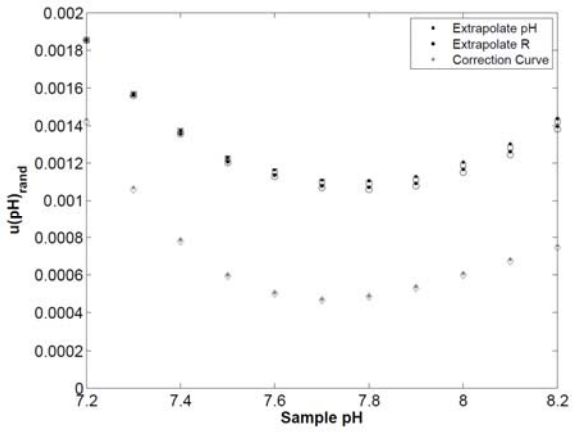

Figure 4.6. Estimates from a Monte Carlo simulation of the random contribution to the combined standard uncertainty in pH when implementing one of three empirical dye correction approaches with a double addition of a pH 7.2 dye solution (with 2 mmol kg^{-1} mCP and $C_T = 0$) to seawater samples with S = 35, A_T = 2400 µmol kg^{-1}, and pH ranging from 7.2 – 8.2. Uncertainties are estimated both for the approach of using volume of added dye solution (filled symbols) and for the approach of using the isosbestic absorbance (unfilled symbols) to estimate the dye concentration in the sample.

4.4 Assessment

4.4.1 Evaluation of the extrapolation approach to dye corrections

For each of the three batches of seawater (CRM 160, 162, and 164) in the weighed dye addition experiments, the pH values obtained by extrapolating the dye addition measurements made with two different dye solutions (at pH ~8 and pH ~7.2) to zero dye concentration were in excellent agreement (within 0.001 pH units) with each other (**Figure 4.7**). The change in the pH of the seawater samples with dye addition was also generally consistent with results from the dye perturbation model (e.g., **Figure 4.8**), which predicted that the sample should change linearly with dye concentration, increasing if the pH of the dye solution is higher than that of the samples and vice versa.

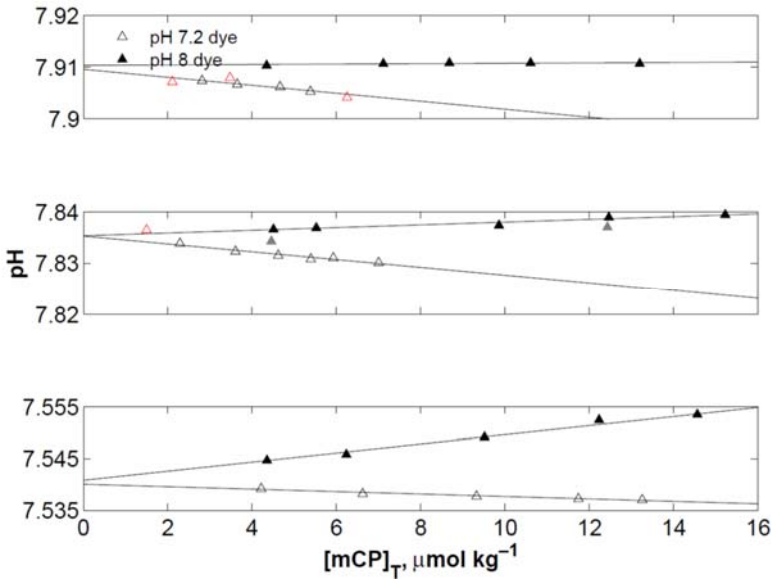

Figure 4.7. pH of three batches of CO_2-in-seawater Reference Materials, measured with two different batches of dye solution (20170210 and 20170317, see **Table 4.3** and **Table 4.4**) plotted against dye concentration. Each point represents an individual bottle of a seawater batch in which a weighed amount of dye solution was added. Each set of data was extrapolated to zero dye concentration from a least squares line fit to the data. The filled gray symbols and the unfilled red symbols were potentially bad measurements and were excluded from the regression (see text). The data measured with the pH 8 dye were converted to 25°C.

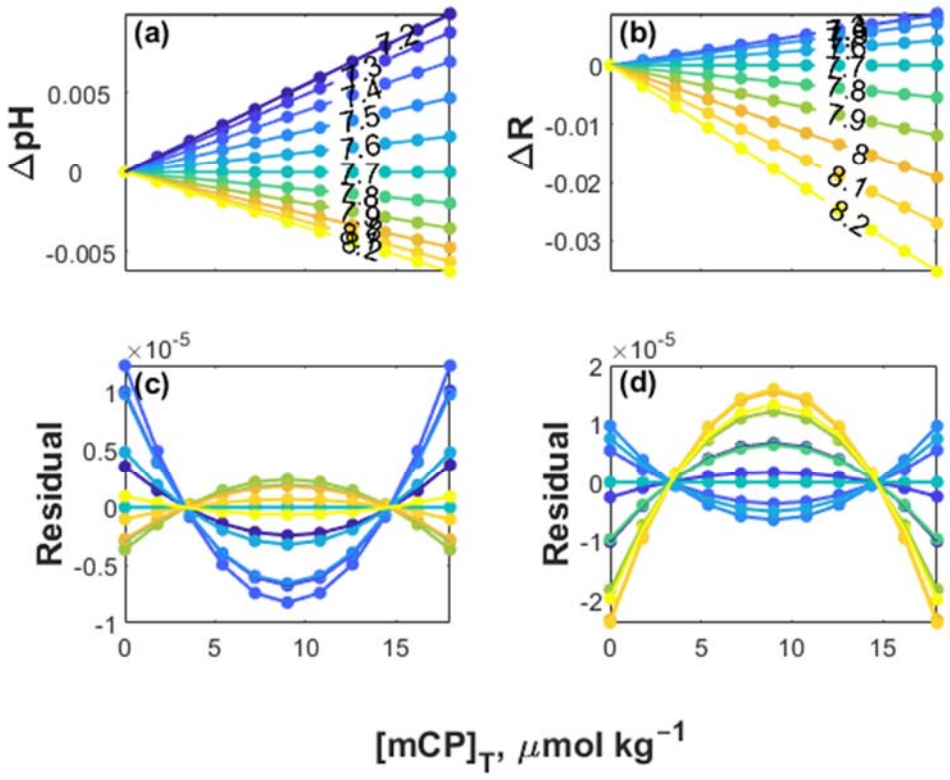

Figure 4.8. Model-simulated values of ΔpH and Δ*R* as a function of dye concentration for the addition of a pH 7.7 dye solution (2 mmol kg^{-1}) to seawater samples with an A_T of 2400 μmol kg^{-1} and pH ranging from 7.2 – 8.2 (shown in the contours). The bottom panels show the residuals of the linear fit to each set of data in the top panels and are colored the same way as the data in the top panels.

The dye perturbation model predicted that a least squares line fitted to a plot of pH versus total dye concentration, such as in **Figure 4.7**, should be highly linear with R^2 values >0.9999. However, the observed R^2 values for the fitted lines in **Figure 4.7** were <0.99 (**Table 4.4**). A few of the data points in **Figure 4.7** were identified as potentially bad measurements from a measurement of another bottle and from the nonlinearity of the data. Retaining these data points resulted in R^2 values of the fitted least squares lines as low as ~0.71. When these data points were excluded from the regression, the R^2 values of the least squares lines fitted to

each set of data was >0.93, except for one set of data, in which the pH of the samples was close to the pH of the dye solution, resulting in a near zero slope (3.66×10^{-5}) and therefore a low R^2 value (~0.42).

Table 4.4. Regression statistics for the data shown in **Figure 4.7** (outliers excluded) for pH (at 25°C) versus total dye concentration (in μmol kg^{-1}).

Dye Solution	CRM	Slope ± std. error	Intercept ± std. error	R^2
20170210	B160	–0.00076 ± 0.000084	7.8353 ± 0.00042	0.95320
pH~7.1	B162	–0.00077 ± 0.000093	7.9096 ± 0.00040	0.97163
	B164	–0.00024 ± 0.000018	7.5400 ± 0.00019	0.97868
20170317	B160	0.00026 ± 0.000040	7.8353 ± 0.00042	0.93348
pH~8	B162	0.000037 ± 0.000025	7.9104 ± 0.00023	0.42131
	B164	0.00088 ± 0.000048	7.5407 ± 0.00056	0.98814

A likely reason for the lower than expected linearity of the data in **Figure 4.7** was the contribution from bottle-to-bottle variability in pH to the variance in the data, as each set of data in **Figure 4.7** consisted of measurements of multiple bottles of an individual batch of seawater. We confirmed this hypothesis with additional experiments which compared the slopes determined from dye additions to individual bottles versus dye additions to multiple bottles of a seawater batch (**Table 4.5**). The slopes of the least squares lines determined from multiple dye additions to seawater from three individual bottles of CRM 172 agreed to within 2×10^{-5} and had R^2 values > 0.99. On the other hand, the data from dye additions to multiple bottles were noticeably less linear, with an R^2 of 0.91, and the slope of the fitted line was ~0.0001 higher than the mean value of the slope from the other three sets of data. The standard deviation of the dye perturbation-corrected pH values determined from dye additions to the individual bottles was 0.0015 (**Table 4.5**), similar to the bottle-to-bottle variability reported for the seawater reference materials (Bockmon and Dickson, 2014). In later experiments, the dye additions were done to individual bottles to reduce this source of uncertainty. Despite the shortcomings of these initial experiments, the data in **Figure 4.7** suggest that the extrapolation

approach gives consistent pH values across different batches of dye solutions and does not appear to contribute to a significant error in determining the pH of a sample without dye.

Table 4.5. Comparison of regression statistics for pH versus total dye concentration (in μmol kg^{-1}) from additions of a dye solution (20180405) to three individual bottles of CRM 172 and from weighed dye additions to four separate bottles of the same seawater batch.

Dye addition method	Slope ± std. error	Intercept ± std. error	R^2
Kloehn, single bottle	–0.00058 ± 0.000040	7.8294 ± 0.00040	0.99081
	–0.00056 ± 0.000026	7.8295 ± 0.00026	0.99590
	–0.00058 ± 0.000053	7.8268 ± 0.000053	0.99984
Weighed, multiple bottles	–0.00046 ± 0.00010	7.8285 ± 0.0010	0.91000

We also evaluated whether using A_{iso} as a proxy for dye concentration, as in Carter et al. (2013), might contribute to error in implementing the extrapolation approach. Using the same data as in **Figure 4.7**, the pH values of each set of measurements, whether extrapolated to A_{iso} = 0 or zero dye concentration, agreed to better than $2x10^{-5}$ pH units. Furthermore, when the total dye concentrations (known from weighed additions of dye) were plotted against their corresponding A_{iso} values, the intercept of a least squares line fit to the data suggested that the dye concentration at A_{iso} = 0 was not significantly different from zero (**Figure 4.9a**). However, a systematic pattern was observed in the residuals of the fit in one set of measurements (**Figure 4.9b**), which is consistent with an error of ~40 μmol kg^{-1} in the total dye concentration of that particular batch of dye solution (Dye 20180405). We use this value as an estimate of the systematic error in the dye stock solution concentration when evaluating the model uncertainties.

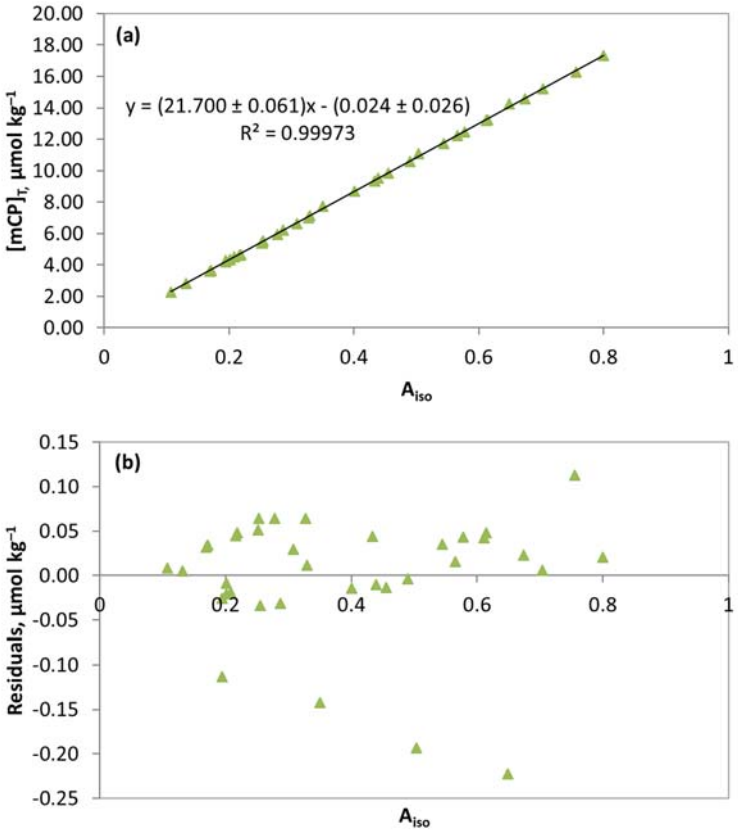

Figure 4.9. (a) Regression of total dye concentration versus the isosbestic absorbance from weighed additions of dye to seawater and (b) the residuals of the regression. These data show the relationship between total dye concentration and isosbestic absorbance appropriate for a 5 cm cell.

4.4.2 Comparison of measured and modeled dye perturbation

A small discrepancy was found between the measured and model-estimated slopes $\Delta pH/\Delta[mCP]_T$. The measured slope for each batch of seawater across three dye solutions was typically more positive than the model-estimated slope, except for one set of data (CRM 162 in the first experiment) which had a near-zero measured slope (**Table 4.6**).

Table 4.6. Comparison of the model-estimated slope of pH versus total dye concentration (in μmol kg^{-1}) and the measured slope from dye addition experiments. Values of the slopes are given, along with their estimated 95% confidence intervals and the number of degrees of freedom (for the measured slope). Differences are significant at the 95% confidence level (indicated by asterisks) when their absolute magnitudes are larger than the estimated uncertainty of the difference.

Dye solution	CRM	$Slope_{model}$ ± 95% C.I.	$Slope_{meas}$ ± 95% C.I.	df	$Slope_{meas} - Slope_{model}$	$U_{meas-model}$
20170317	B160	0.00023 ± 0.000038	0.00026 ± 0.00013	3	0.000032	0.00013
pH~8	B162	0.00011 ± 0.000035	0.000037 ± 0.000079	3	–0.000076	0.000086
	B164	0.00078 ± 0.000041	0.00088 ± 0.00013	4	0.000097	0.00014
20180405	B164	–0.00042 ± 0.000034	–0.00034 ± 0.000066[a]	6	0.000085*	0.000074
pH~7.1	B172	–0.00064 ± 0.000034	–0.00058 ± 0.000057[a]	16	0.000060	0.000066
20180518	B160	–0.00066 ± 0.000037	–0.00053 ± 0.000093	2	0.00013*	0.00010
pH~7.3	B164	–0.00043 ± 0.000037	–0.00031 ± 0.00010[a]	4	0.00012*	0.00011

[a] Pooled value

As shown in **Table 4.6**, the discrepancies between modeled and measured slopes were not significantly different in the first set of experiments (with Dye 20170317). It is likely that the larger standard error in the measured slopes (due to the contribution from bottle-to-bottle variability in pH as discussed earlier) and the limited number of dye additions resulted in insufficient statistical power to detect a significant difference between the modeled and measured slopes. Note that the uncertainty of the difference, $U_{model–meas}$, was relatively large (>0.0008). In the later experiments, the slopes were determined from dye additions to individual bottles, reducing the standard error of the slope. For three of these slopes, as indicated by the superscript in **Table 4.6**, the reported standard errors were pooled values from dye additions to individual bottles repeated multiple times with different bottles of the same batch of seawater. The pooled standard errors of the slopes measured in CRM 164 and 172 with Dye 20180405 included measurements from a six-month dye storage study (n = 8 bottles of CRM 172 and 3 bottles of CRM 164 measured with Dye 20180405—see **Table 4.3**). As a result of reducing the standard error of the measured slopes, we were able to detect significant

differences between the modeled and measured slopes in three out of four sets of measurements in the later experiments (**Table 4.6**)

It is unlikely that uncertainties in the dye perturbation model could be responsible for the discrepancy. The standard uncertainties of the model slope inferred from **Table 4.6** were quite small and suggest an uncertainty of ~0.001 in the pH correction at 45 μmol kg^{-1} mCP, a typical dye concentration for a 1 cm cell. Uncertainties in the various equilibrium constants will contribute to the uncertainty in calculating pH from A_T and C_T. However, a large degree of error cancellation occurs when calculating the difference ΔpH, and thus, the overall uncertainty in the calculated slope is small. The largest contribution to the overall uncertainty of the model slope was typically from K_2, followed by pH_{dye}, which is used to estimate the dye solution composition (**Figure 4.10**). When using dye solutions prepared in a NaCl background, the model slope will have a greater uncertainty contribution from the dissociation constant of the dye, $K_a(HI^-)$, as this value is not as well-known in a NaCl medium.

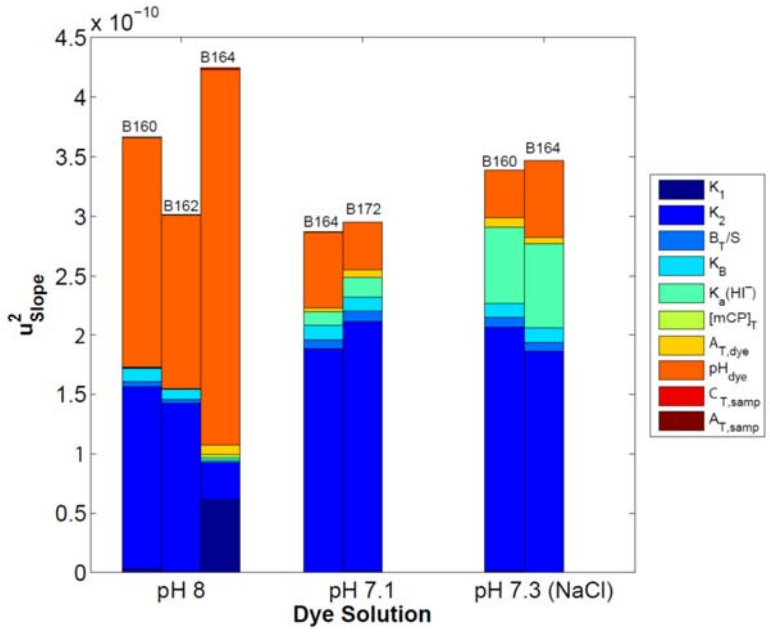

Figure 4.10. Stacked bar chart showing the contributions to u_c^2 of the model-estimated slopes ($\Delta pH/\Delta[mCP]_T$) for the addition of three separate dye solutions to seawater samples (see also **Table 4.6**). Standard uncertainties are from **Table 4.1**.

As discussed earlier, our assumption of the presence of carbonate alkalinity in the dye solutions could be a source of uncertainty in the dye perturbation model. Two lines of evidence, however, suggest that this assumption was valid. First, the measured A_T of the dye solutions was always higher than the A_T calculated from pH and total dye concentration using **Eq**. **12**. The excess alkalinity in the dye solutions ranged from 345 μmol kg^{-1} to as high as 1389 μmol kg^{-1} in one dye solution (**Table 4.3**). Secondly, we prepared two impure *m*-cresol purple solutions of similar composition to the purified dye solutions used in the dye addition experiments and measured their C_T, A_T, and pH. The C_T measured in these two solutions ranged from 288-509 μmol kg^{-1}, and the carbonate alkalinity estimated from the A_T, C_T, and total dye concentration ranged from 273-544 μmol kg^{-1} and agreed with the values of excess alkalinity estimated for these solutions to better than 90 μmol kg^{-1}. The discrepancy between excess alkalinity and carbonate alkalinity for these two solutions was consistent with a systematic

error in the pH and total dye concentration due to the use of an impure dye. As our dye solutions contained significant amounts of carbonate alkalinity, the use of pH and A_T to calculate C_T was an appropriate choice. The combination of pH and A_T also has the advantage that an error in estimating the A_T of the dye solution is correlated with and approximately equal to the error in the calculated C_T (i.e., $\Delta A_T/\Delta C_T \sim 1$), and hence, the calculated perturbation slope will nearly be the same as the true value. For dye solutions with low carbonate alkalinity, calculating the A_T of the dye solution with **Eq**. **12** (i.e., assuming zero carbonate alkalinity and zero C_T) will result in small errors in estimating the actual A_T and C_T of the dye solution and consequently a small error in the calculated slope. For a dye solution with high carbonate alkalinity, the error in the calculated slope becomes meaningful. The difference between the slope calculated using the measured pH and A_T of the dye solution and the slope calculated using the measured pH and assuming zero C_T was $< 3\times10^{-5}$ for all of the seawater samples except for CRM 164 measured with Dye 20170317, where this difference was ~0.0001. This particular dye solution had the highest A_T and excess alkalinity of all the dye solutions (**Table 4.3**) as it was prepared with more than twice the usual amount of NaOH pellets due to difficulties we had in dissolving the dye.

Since the uncertainties in the dye perturbation model could not explain the observed discrepancy between the modeled and measured slopes, another possibility to consider is a systematic error in the measured slope. We later identified an absorbance-dependent absorbance error on our spectrophotometer which had the effect of producing an apparent dye perturbation in phosphate buffer measurements (see **Chapter 3**). These absorbance errors developed on our spectrophotometer as the deuterium lamp degraded in intensity over its lifetime. When the deuterium lamp was replaced, these errors, along with apparent dye

perturbations in buffered solutions, were minimized. The slope of $\Delta pH/\Delta[mCP]_T$ in the phosphate buffer measurements using a significantly degraded deuterium lamp was ~0.00028. With newer deuterium lamps, the slopes in the phosphate buffers were smaller, but could still be as large as ~0.00014, which is on the order of the magnitude of the model vs. measured slope discrepancies in **Table 4.6**. Although the phosphate buffer measurements were not made at the same time as the seawater dye addition experiments in this study, it seems possible that our measurements were affected to some degree by similar errors.

4.4.3 Implications for the use of model-based corrections

What are the implications of the discrepancy between modeled and measured slopes on the error of the dye perturbation correction? The discrepancies we observed imply an inconsistency between the pH adjustments estimated from the modeled versus measured slopes, which would grow larger with increasing dye concentration (i.e., $\Delta pH = \Delta pH/\Delta[mCP]_T \times [mCP]_T$). The difference between the ΔpH estimated from the modeled slope and the ΔpH estimated from the measured slope, both calculated at a dye concentration appropriate for a 1 cm cell (45 μmol kg^{-1}), is shown in **Figure 4.11** for the seawater samples measured in the three experiments in **Table 4.6**. The largest of these discrepancies in ΔpH was ~0.006. In 5 and 10 cm cells, the error would be 0.0012 and 0.0006, respectively, and therefore of minor importance. Thus, if there are additional contributions to the measured slope, such as from absorbance-dependent absorbance errors, this may result in significant inconsistencies with model-based corrections for a 1 cm cell.

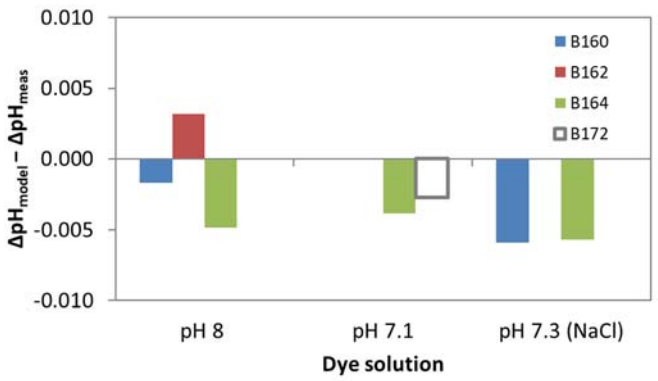

Figure 4.11. Discrepancy between the dye perturbation to sample pH estimated from the model slope of pH versus total dye concentration and from the measured slope (see **Table 4.6**) at a dye concentration of 45 μmol kg^{-1}, appropriate for a 1 cm. cell.

In the absence of problematic instrumental contributions to the slope $\Delta pH/\Delta[mCP]_T$, the model can be used estimate the dye perturbation to sample pH with low uncertainty. Assuming that the dominant sources of uncertainty to the modeled slope and the other minor sources of uncertainty from the various constants, etc. in **Table 4.1** cannot be reduced, the tolerance in the uncertainty of the sample composition, required as model inputs, is fairly lenient. The most critical condition occurs at the point of minimum buffer capacity of the sample, where the sensitivity of the slope $\Delta pH/\Delta[mCP]_T$ to the sample C_T and A_T is highest. This point occurs at approximately the midpoint between pK_1 and pK_2 (pH = 7.4066 at S = 35 and T = 25°C). **Figure 4.12** shows the tolerance in the uncertainty in sample C_T and A_T or, alternatively, the sample pH and A_T for a target uncertainty of 0.002 in pH when using model-based corrections. The tolerance is lowest for a high pH dye, but even so, an uncertainty of 6 μmol kg^{-1} in A_T and C_T or an uncertainty of 6 μmol kg^{-1} in A_T and 0.02 in pH would be adequate for estimating the sample composition for input into the model. Given the large tolerance in the uncertainty of the sample A_T, the use of a locally interpolated regression (e.g., Carter et al., 2018) would be adequate in estimating A_T to within 6 μmol kg^{-1} uncertainty. Additionally, it may be adequate

to use uncorrected pH values as a model input, as the magnitude of dye perturbation is typically smaller than the required uncertainty in the initial estimate of pH for most samples, especially if a low pH dye solution is used (**Figure 4.12b**).

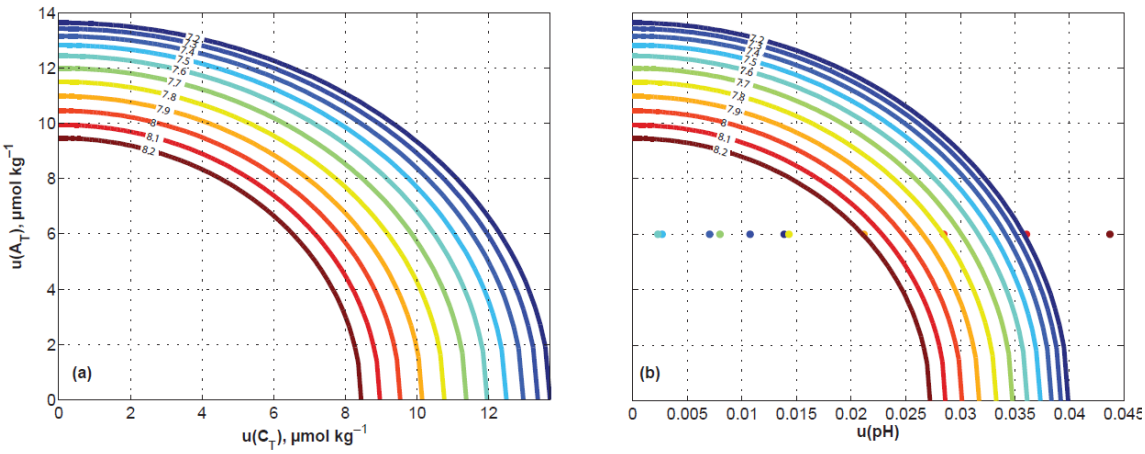

Figure 4.12. Contour plots showing the required uncertainty in the sample (a) C_T and A_T or (b) pH and A_T for using an equilibrium model to estimate the dye perturbation to sample pH with an uncertainty of 0.002 in a 1 cm cell (45 µmol kg^{-1} mCP). These estimates were calculated for a seawater sample of S=35, pH=7.4066, and A_T = 2400 µmol kg^{-1} (see text) and included the other contributions to the uncertainty of a model-estimated correction listed in **Table 4.1**. In these calculations, the dye solutions added to the samples had a concentration of 2 mmol kg^{-1} mCP, $\underline{C}_T$ = 0, and pH ranging from 7.2 – 8.2 (shown in the contours). Each point in (b) corresponds to one of the dye solutions represented by the contours and are colored the same way. The points are positioned at $u(A_T)$ = 6 µmol kg^{-1} and along the x-axis corresponding to the magnitude of the pH perturbation to the sample. Points that fall to the left of their corresponding contours have pH perturbations smaller than the required uncertainty in the initial estimate of the sample pH.

4.4.4 Systematic errors in the extrapolation approach

The dye perturbation model predicts that the pH and R of a seawater sample both change linearly with dye concentration. For each set of simulated values of pH and R in **Figure 4.8**, the least squares lines fitted to the data had R^2 values >0.999. However, a closer examination of the residuals of the fit (**Figure 4.8c, Figure 4.8d**) reveals that a straight line, although a good approximation, does not perfectly describe how pH and R change with dye concentration. Therefore, extrapolating pH or R to zero dye concentration will have a small error due to the slight non-linearity. The error in extrapolating pH values from a double dye

addition to zero dye concentration is shown in **Figure 4.3a** (plotted against the sample pH) for additions of a pH 7.7 dye solution and in **Figure 4.3d** for a pH 7.2 dye solution. The distribution of these errors with respect to the sample pH reflects how the degree of misfit between the simulated pH values and a straight line changes with sample composition. **Figure 4.8c** shows the residuals for a pH 7.7 dye solution and can be compared to distribution of the errors shown in **Figure 4.3a**. Over the sample pH interval of 7.2 to 7.4, the pattern of the residuals becomes more concave in shape and then flattens from pH 7.4 to 7.7, becoming essentially horizontal at zero at pH 7.7, where the sample and dye solution pH are the same. Above pH 7.7, the residuals develop only a slight convex shape, as the slopes of the lines above pH 7.7 do not grow as large as the slopes below pH 7.7. The errors in **Figure 4.3a** follow this pattern, increasing from pH 7.2 to 7.4 and then decreasing above pH 7.4 until becoming zero at pH 7.7. Above pH 7.7, the errors become slightly negative. The error in extrapolating to zero dye concentration grows in shorter pathlength cells (**Figure 4.3a, Figure 4.3d**), as the extrapolation is over a larger pH range at the high dye concentrations in short pathlength cells. However, the error in extrapolating pH to zero dye concentration is small and no larger than ~0.0005 in a 1 cm cell.

The error in calculating the sample pH from a value of *R* extrapolated to zero dye concentration is similar to the error in extrapolating pH values directly. However, there is an additional contribution from the way an error in the extrapolated *R* value propagates into an error in pH. The difference between the error in extrapolating pH directly (**Figure 4.3a, Figure 4.3d)** and the error in calculating pH from an extrapolated *R* value (**Figure 4.3b, Figure 4.3e**) grows at lower pH due to the higher sensitivity of pH to an error in *R* and diminishes as the

sample pH approaches the pH of the dye solution. The error from extrapolating R is also small and no larger than –0.0007 in a 1 cm cell.

The choice of extrapolating to zero dye using the volume of dye added or A_{iso} as the proxy for dye concentration will contribute negligibly to the uncertainty of the pH correction. We evaluated this contribution to uncertainty with a Monte Carlo simulation using a value of 0.38 μL (half the value of the manufacturer's claimed systematic error) for the systematic component of the standard uncertainty of pipetting 5 μL with an Eppendorf pipette and a value of 0.026 μmol kg^{-1} mCP, the standard error of the intercept in **Figure 4.9**, as an estimate of the uncertainty of extrapolating to zero dye using A_{iso}. The resulting uncertainties in the pH correction were on the order of 10^{-4} or less, and thus we neglect this contribution when calculating a combined standard uncertainty.

4.4.5 Systematic errors in the dye correction curve approach

The main source of systematic error from the use of the correction curve approach to dye corrections is from the assumption that the $\Delta R/\Delta[\text{mCP}]_T$ for a particular sample can be estimated from a linear function of R. The simulated correction curve in **Figure 4.13** shows clearly that $\Delta R/\Delta[\text{mCP}]_T$ does not change linearly over the full range of R, consistent with the results of Li et al. (2020). The resulting error in estimating $\Delta R/\Delta[\text{mCP}]_T$ for a sample propagates into an error in estimating the required adjustment in R (i.e., ΔR) and consequently the sample pH and is most substantial where the relationship between $\Delta R/\Delta[\text{mCP}]_T$ and R deviates significantly from linearity. The error in the pH correction grows with higher dye concentration as the error in estimating ΔR from a slightly wrong value of $\Delta R/\Delta[\text{mCP}]_T$ scales with dye concentration. This error can be minimized by using a low pH dye solution for which the relationship between $\Delta R/\Delta[\text{mCP}]_T$ and R is more linear, as can be seen in **Figure 4.13**. In a

1 cm cell, the error in the pH correction can be as large as –0.01 for a sample at pH 7.2 when using a dye solution with a pH of 7.7 (**Figure 4.3c**). However, if using a pH 7.2 dye solution, the error for a sample at pH 7.2 is reduced to 0.001 and the largest error is –0.0026 for a sample at pH 7.8 (**Figure 4.3f**).

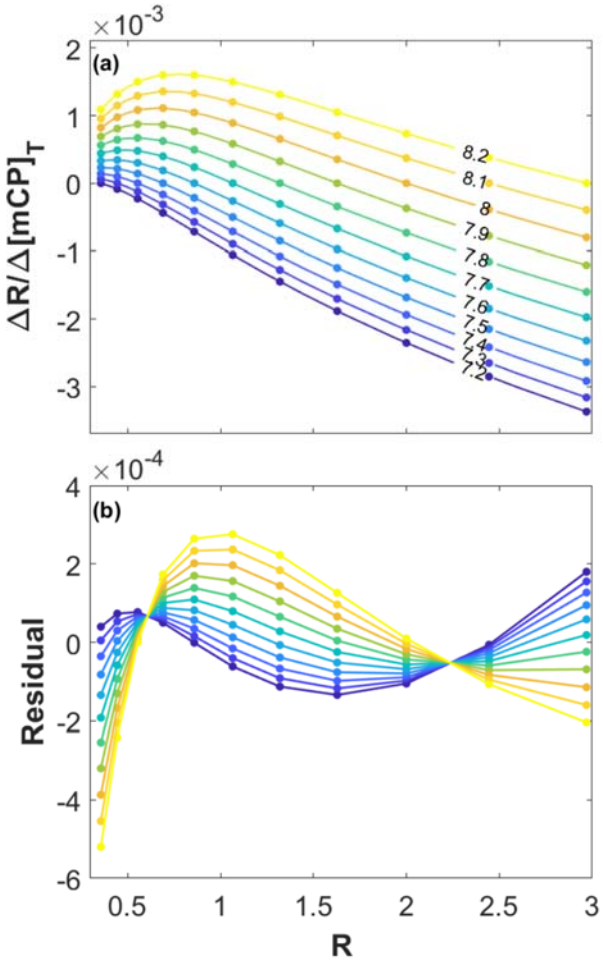

Figure 4.13. (a) Model-simulated dye correction curves for dye solutions with a concentration of 2 mmol kg^{-1} mCP, C_T = 0, and pH ranging from 7.2 – 8.2 (shown in the contours). Values of $\Delta R/\Delta[\text{mCP}]_T$ are plotted as a function of the R of the samples. In each curve, the seawater samples had S = 35, A_T = 2400 µmol kg^{-1}, and pH ranging from 7.2 – 8.2. (b) Residuals of the linear fit to each set of data in (a), plotted in the same colors.

The curves of $\Delta R/\Delta[\text{mCP}]_T$ versus R in **Figure 4.13** represent the dye perturbation behavior for samples at constant alkalinity. The range in A_T of samples collected in a surface-to-deep open ocean profile can lead to a systematic error in estimating $\Delta R/\Delta[\text{mCP}]_T$ from a

correction curve and consequently a dye concentration-dependent error in estimating the required pH correction. The systematic error resulting from variability in sample composition was evaluated using data from open ocean profiles collected along a cruise transect in the North Pacific (**Figure 4.4**, **Figure 4.5**). As the variability in A_T was greatest near the surface (<50 m), the error in correcting the pH of surface samples can be as large as ~0.006 in a 1 cm cell, but minor in 5 and 10 cm cells (**Figure 4.5**).

4.4.6 Effect of systematic errors in absorbance

As discussed earlier, during the time the data were collected, our spectrophotometer may have had an absorbance-dependent error in absorbance which contributed to the apparent measured slope $\Delta pH/\Delta[mCP]_T$. Such errors can be minimized by monitoring the lamp intensities and replacing lamps that show signs of such anomalous behavior. However, in our experience, even with new lamps, these absorbance errors may not be entirely eliminated. From simulations using the dye addition model and synthetic spectra with the addition of the absorbance-dependent absorbance errors (see **Section 4.8.1**), we estimated a systematic contribution to the uncertainty of the dye perturbation corrections to be no larger than 0.0012 at pH 8.2 (**Figure 4.15**).

4.4.7 Random errors in empirical dye corrections

Random errors in the absorbance measurements at 434 nm, 578 nm, and the isosbestic wavelength will affect the precision of the pH measurements and the dye perturbation correction. The resulting uncertainties in the adjusted pH values for the extrapolation and correction curve approaches, plotted as standard deviations in **Figure 4.6**, show a pH-dependent distribution that is a consequence of the way a constant variance in absorbance propagates into a variance in pH. Because the dye concentrations in the simulation were chosen

to obtain the same absorbances in any pathlength cell, the uncertainty estimates in **Figure 4.6** are independent of cell pathlength. For the extrapolation approach with a double dye addition, the contribution of random errors in absorbance measurements to the uncertainty in pH is <0.002. In continuous flow systems that take numerous measurements along an indicator gradient, lower uncertainties can be achieved. For a 25 point extrapolation, such as might be done in the Submersible Autonomous Moored Instrument for pH (SAMI-pH) system of Seidel et al. (2008), the random uncertainty in pH is <0.0006. For the system of Aßmann et al. (2011), which takes more than 200 measurements along an indicator gradient, the random uncertainty in pH is 0.0002.

Similarly, the greater number of points used to define a dye correction curve results in a reduction of the random uncertainty in the correction curve approach relative to a two-point extrapolation. For a 16 point correction curve, the uncertainties are <0.0015 (**Figure 4.6**). As the number of points used to define the curve increases, the uncertainties will decrease until reaching a limit of ~0.0004.

In the Clayton and Byrne (1993) approach, the volume of dye solution pipetted into the cell is used as a proxy for dye concentration. The random uncertainties in pH shown in **Figure 4.6** were also evaluated with a random error in estimating the volume of dye pipetted with an Eppendorf pipette, and the resulting uncertainties, shown by the filled symbols, are nearly identical to those estimated for the approach using A_{iso}.

4.4.8 Combined standard uncertainty estimates for empirical dye corrections

We derived estimates of the combined standard uncertainty for each empirical correction approach, following the recommendations of the International Organization for Standardization's Guide to the Expression of Uncertainty in Measurement (GUM, 1993) to

express both random and systematic sources of uncertainty as standard deviations and to propagate them together when estimating the combined standard uncertainty. Thus, the combined standard uncertainty (u_c) is calculated by summing the random and systematic contributions in quadrature (**Eq. 16**).

$$u_c = \sqrt{u_{rand}^2 + u_{sys}^2} \tag{16}$$

Typically, when their signs and magnitudes are known, systematic errors are removed before propagating uncertainties, so that u_{sys} in **Eq. 16** represents the uncertainty of the correction. Although we explicitly evaluated the systematic errors in the various empirical correction approaches, we treat them as though they were unknown (as is typically the case in practice) when calculating the combined standard uncertainty. We assumed that the values of the estimated errors (difference between simulated and true pH) for each correction approach given in **Figure 4.3** represent $2u$ estimates of those contributions to the systematic uncertainty. Additionally, we included an uncertainty contribution from absorbance-dependent absorbance errors ($u(\text{pH})_{abs}$ in **Figure 4.15**). The total systematic uncertainty u_{sys} was therefore calculated by summing in quadrature $u(\text{pH})_{abs}$ and half the absolute value of the errors in **Figure 4.3**. Together with the values of u_{rand}, estimated from the Monte Carlo simulation (**Figure 4.6**), we calculate u_c for a pH 7.2 dye solution and plot these values against the sample pH in **Figure 4.14**. These estimates of u_c represent the total uncertainty of the dye correction method at the 68% confidence level.

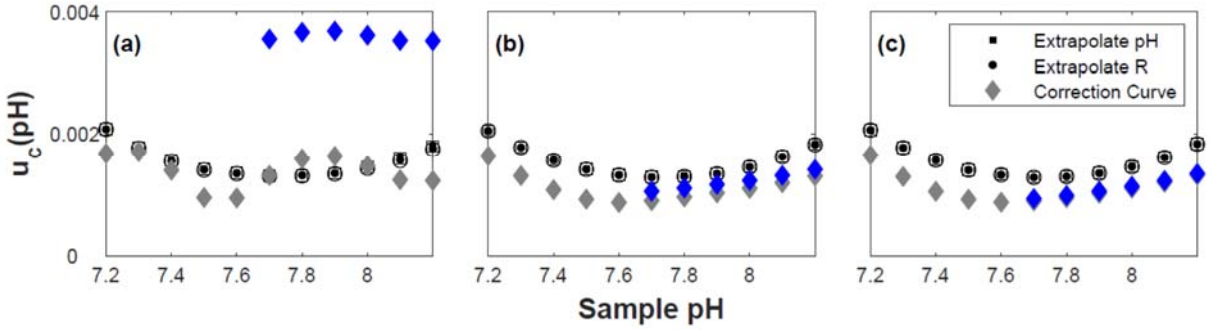

Figure 4.14. Estimates of the combined standard uncertainty in pH (including random and systematic contributions) when implementing one of three empirical dye correction approaches with a double addition of a pH 7.2 dye solution (with 2 mmol kg^{-1} mCP and C_T = 0) to seawater samples with S = 35, A_T = 2400 µmol kg^{-1}, and pH ranging from 7.2 – 8.2. Uncertainties are estimated for (a) 1 cm., (b) 5 cm., and (c) 10 cm. cells, both for the approach of using volume of added dye solution (filled symbols) and for the approach of using the isosbestic absorbance (unfilled symbols) to estimate the dye concentration in the sample. The blue diamonds include an additional uncertainty contribution when using a dye correction curve to correct samples with variable A_T.

From comparison of **Figure 4.3**, **Figure 4.6**, and **Figure 4.14**, it can be seen that the total uncertainties of the two extrapolation approaches (extrapolating pH or extrapolating *R*) are nearly identical and dominated by the random contribution. In 5 and 10 cm cells, the total uncertainty of the correction curve approach is also dominated by the random contribution, and as the random uncertainty is lower than that of the extrapolation approach, the total uncertainty is also smaller than for the extrapolation approach (**Figure 4.14b, Figure 4.14c**). In a 1 cm cell, the total uncertainty of the correction curve approach is dominated by the systematic contribution from the assumption of the linearity of the correction curve (**Figure 4.14a**). However, this systematic uncertainty can be minimized by using a low pH dye solution (~7.2), which linearizes the correction curve (**Figure 4.3c, Figure 4.3f**). When the pH of the dye solution is optimized, the total uncertainty of the correction curve approach in a 1 cm cell is comparable to that of the extrapolation approach. However, for surface samples (pH ≥ 7.7), there may be an additional systematic contribution from the error in estimating adjustments from a correction curve for samples with a wider range of A_T. When the RMS errors of the data

≤ 50 m in **Figure 4.5** was included as a contribution to the u_{sys} of samples with pH ≥ 7.7, the total uncertainty of the correction curve approach in a 1 cm cell is ~0.0035 for surface samples (**Figure 4.14a**).

4.5 Discussion

This study evaluated the systematic and random contributions to the total uncertainty of dye perturbation corrections for spectrophotometric pH measurements. The total uncertainty of the various empirical dye perturbation correction approaches (i.e., extrapolation and correction curve methods) using double dye additions is generally dominated by contributions from random errors in absorbance measurements. An advantage of the correction curve approach is that the effect of random errors is reduced from the number of measurements used to characterize the curve, and hence, the total uncertainty of the correction curve method is typically lower than that of the extrapolation approach with double dye additions. However, at the high dye concentrations used in short pathlength cells (e.g., 1 cm cells), the uncertainty of the correction curve approach becomes dominated by the systematic contribution from the assumption that the required adjustment in R can be estimated from a linear function.

The natural range of A_T in a surface-to-deep open ocean profile can lead to systematic error in estimating a dye perturbation adjustment from a correction curve, which is strictly representative of samples with constant A_T. This error is negligible for 5 and 10 cm cells, but is sizeable in 1 cm cells, particularly for surface samples where A_T is most variable. For low salinity samples, such as from estuarine environments, the range in A_T is larger than in the open ocean, and thus there may be an even larger error in using a dye correction curve to estimate the pH perturbation adjustment for these samples (Li et al., 2020).

The dominant sources of uncertainty in dye perturbation corrections estimated from an equilibrium model are from the equilibrium constants K_1 and K_2, the acid dissociation constant of *m*-cresol purple, and the pH of the dye solution. However, as the model is fairly insensitive to many of these sources of uncertainty and those listed in **Figure 4.1**, the uncertainty of model-estimated dye perturbations is small (~0.001) even in 1 cm cells. Yet, despite the model's seeming insensitivity to various sources of uncertainty, we observed a discrepancy between the model and measured pH perturbation that suggested an error in the pH correction as large as 0.006 in a 1 cm cell. A likely cause of this inconsistency may be absorbance-dependent absorbance errors on our spectrophotometer, which constitute an additional contribution to the apparent pH perturbation that is unaccounted by the model.

The estimates of the combined standard uncertainty of the empirical correction approaches we provide in this study apply to single beam spectrophotometers similar to the Agilent 8453. Carter et al. (2013) estimated that applying a dye correction curve based on acidified seawater samples to natural seawater samples (which would have a significantly different A_T) will contribute an uncertainty of 0.0012 to a spectrophotometric pH measurement with a 10 cm cell. Using our estimates of the maximum uncertainty of the correction curve method in 5 and 10 cm cells (~0.0015) along with Carter et al.'s estimates of other uncertainty contributions, the overall uncertainty of a spectrophotometric pH measurement with purified m-cresol purple would be 0.0048. Our analysis suggests that the various empirical dye correction approaches contribute a small source of uncertainty to spectrophotometric pH measurements, particularly in 5 and 10 cm cells, and are unlikely to contribute to the pH-dependent discrepancy between measured spectrophotometric pH and pH calculated from A_T and C_T that

has been observed on a number of open ocean cruises (e.g., Carter et al., 2018; Carter et al., 2013; Fong and Dickson, 2019; McElligott et al., 1998; Williams et al., 2017).

4.6 Comments and recommendations

As there are various approaches to achieving dye perturbation corrections, users should consider the results of this study to select an approach that offers the best convenience and lowest uncertainty. When using the double dye addition method with 5 and 10 cm cells, the correction curve approach typically has the lower uncertainty. As only a subset of samples (covering the full range of sample pH) need be measured with double dye additions to characterize the correction curve, the correction curve approach can also offer savings in overall measurement time and reagent expenditure and is thus preferable when a large number of samples is available. A significant systematic error can arise if a dye correction curve is used to adjust samples of significantly different composition than those represented by the curve. The approach of Carter et al. (2013) of characterizing a dye correction curve by acidifying some seawater samples to obtain points at lower pH is not recommended, as this modification to the natural range of seawater alkalinity would result in a dye perturbation that is not representative of the actual perturbation in the samples.

The extrapolation approach with double dye additions is recommended in 1 cm cells, for low salinity samples, and when the available range in sample pH is so limited that it would not be possible to obtain a well-defined correction curve. In continuous flow systems that take numerous measurements along an indicator gradient, the extrapolation approach offers the least uncertainty. When applying the extrapolation approach, extrapolating R has an additional contribution to systematic error from the propagation of errors in R to errors in pH. However,

the total uncertainty from extrapolating R is not substantially different from extrapolating pH. Thus, either approach may be used to reliably correct for dye perturbation.

The use of model-based dye perturbation corrections can potentially reduce overall sample measurement time and reagent expenditure, but its implementation requires care. We outline some recommended practices when using model-based corrections. As the use of the model to estimate dye perturbation requires estimates of the sample composition as input parameters, it would be most convenient to use model-based corrections when measuring a large number of samples that are also concurrently measured for A_T and C_T. Alternatively, the sample composition can be characterized using the uncorrected pH values together with measured A_T or A_T estimated from a locally interpolated regression (e.g., Carter et al., 2018), which will still meet the tolerance in the uncertainty in estimating the sample composition. Although the uncertainty of estimating dye perturbations from a model is, in principle, low, absorbance-dependent absorbance errors can result in inconsistencies with the model-estimated corrections, which grow large at the high dye concentrations used in short pathlength cells. The model-based corrections are therefore not recommended for use with pH measurements in 1 cm cells. Additionally, users should ideally use high quality spectrophotometers, check their instruments with absorbance standards, and monitor the lamp intensities.

When using model-based dye corrections, particular care needs to be taken in the preparation, characterization, and storage of the dye solution, so that its composition is well-known. The dye solution should be prepared by carefully weighing the required reagents. When preparing the dye solution from the molecular form of *m*-cresol purple, which requires the addition of NaOH to facilitate dissolution, it may be preferable to use standardized NaOH solutions rather than NaOH pellets, to avoid introducing sodium carbonate contamination and

carbonate alkalinity to the solution. The composition of the dye solution can then be characterized by measuring the pH of the dye solution directly in a short pathlength cell and assuming that $C_T = 0$. Provided that the carbonate alkalinity of the dye solution is low, this assumption will not result in significant error in estimating the dye perturbation. When storing dye solutions, care should be taken to protect against exposure to light and atmospheric exchange, to minimize changes in the composition of the dye solution that may affect the estimated dye perturbations. Our experience has been that ~100 mL of dye solution can be stored in a Tomopal glass syringe for six months at room temperature in a dark cabinet with no significant change in the measured slope $\Delta pH/\Delta[mCP]_T$ in two batches of seawater CRMs. Additionally, another study found that a dye solution carefully stored in a dark, climate-controlled environment provided consistent pH measurements to within ±0.001 pH units over two years (Takeshita et al., submitted).

Traditionally, it has been recommended that the pH of the dye solution be adjusted to minimize the magnitude of the dye perturbation (Chierici et al., 1999; Clayton and Byrne, 1993; Dickson et al., 2007). However, the results of this study suggest that the minimum *uncertainty* in the dye correction is not necessarily realized when the pH of the samples is close to the pH of the dye solution. The use of the correction curve approach in 1 cm cells has a large systematic error from the assumption that the correction curve is linear. However, this systematic error can be minimized by using a dye solution that is more acidic than most of the samples (pH~7.2), as this effectively linearizes the correction curve. The use of an acidic dye solution is also advantageous when using model-based dye corrections, as there is a greater tolerance in the uncertainty of the model inputs for the sample A_T and C_T or A_T and pH. For 5

and 10 cm cells, there is negligible improvement in the uncertainty (whether using empirical or model-based corrections) from the use of a more acidic dye.

When the recommendations and precautions outlined in this study are followed, the dye perturbation correction will contribute a small source of uncertainty (<0.002) to spectrophotometric pH measurements.

4.7 Acknowledgements

Chapter 4, in part, is currently being prepared for submission for publication of the material. Fong, M.B. and Dickson, A.G. The dissertation author was the primary investigator and author of this paper.

The authors would like to thank May-Linn Paulsen for preparing some of the solutions for the synthetic seawater background and Emily Bockmon for preparing the modified batches of seawater reference materials used in this study and running the C_T measurements on the *m*-cresol purple stock solutions.

4.8 Supplementary Information

4.8.1. Evaluating the contribution of absorbance-dependent absorbance errors to the uncertainty of dye perturbation corrections

We evaluated the contribution of absorbance-dependent absorbance errors to the uncertainty of dye perturbation corrections based on our measurements of the spectra of phosphate buffers at different dye amounts (see **Chapter 3**) and numerical simulations. We used the dye perturbation model to calculate the pH of seawater samples (at a temperature of 25°C, salinity of 35, total alkalinity of 2400 μmol kg^{-1}, and pH range of 7.2–8.2) at three different dye concentrations corresponding to typical amounts of dye used in 1, 5, and 10 cm cells (as in **Section 4.3.7**). Synthetic spectra representing the seawater-dye mixtures at different

dye concentrations were generated using the HI^- and I^{2-} species spectra for purified mCP (**Figure 3.3**), and an absorbance-dependent absorbance error was added to each spectrum based on the first principal component of the spectra obtained for each set of phosphate buffer measurements in **Figure 3.16** and the A_{iso} of the simulated seawater spectrum. The calculations were repeated using the principal component data in **Figure 3.16** for each phosphate buffer dataset to include the variability in the levels of absorbance-dependent absorbance errors that we observed when using different deuterium lamps. However, the initial 2019 set of measurements made with the degraded deuterium lamp was excluded, so that the simulation is more representative of typical measurement conditions. The pH was then calculated from the simulated spectra using **Eq. 5**, and the dye perturbation-corrected pH was obtained by the extrapolation and correction curve approaches. The uncertainty in pH for each simulated seawater sample was estimated from the standard deviation of the dye perturbation-corrected pH values obtained from the 8 sets of calculations (**Figure 4.15**).

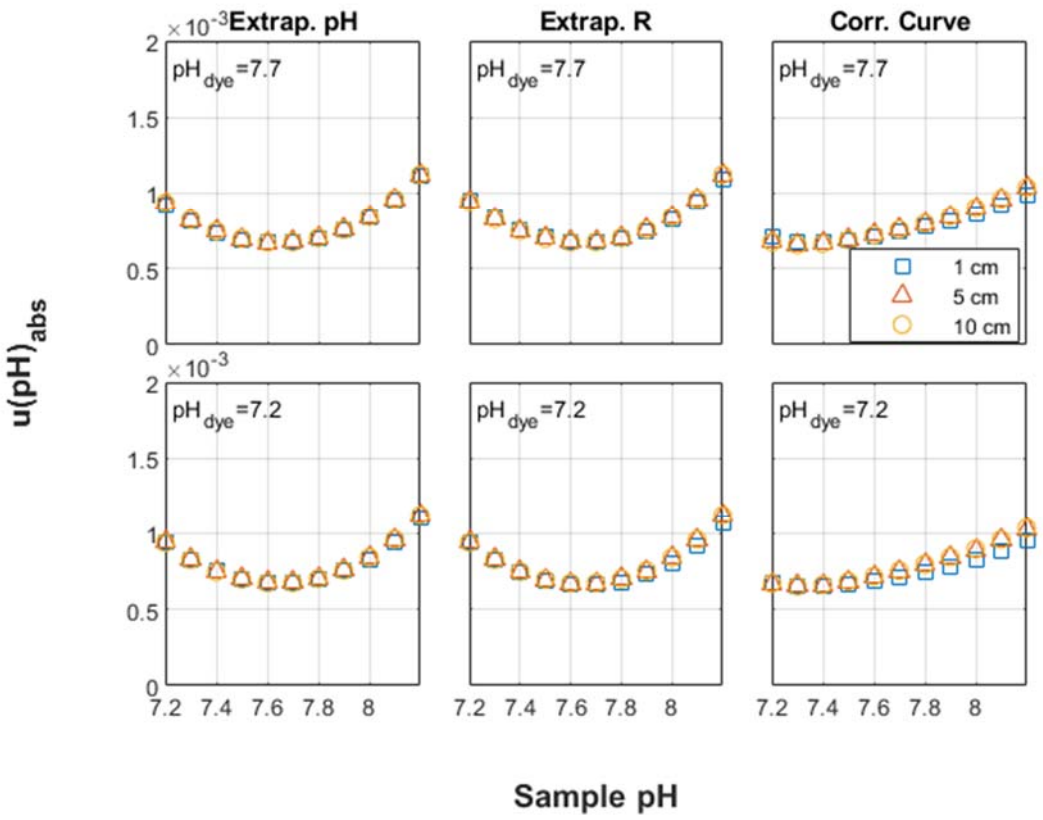

Figure 4.15. Estimated standard uncertainty in pH due to the effect of absorbance-dependent absorbance errors on the dye perturbation correction (see text). The uncertainty for three dye correction approaches were evaluated: extrapolating pH (left panels) or *R* values (center panels) from a double dye addition to zero dye concentration and the use of a dye correction curve constructed from double dye additions to samples across a range in pH (right panels). The uncertainties were estimated for the addition of a 2 mmol kg^{-1} mCP solution with $C_T = 0$ and a pH of 7.7 (top panels) or 7.2 (bottom panels) to seawater samples with S=35, A_T = 2400 µmol kg^{-1}, and pH ranging from 7.2 – 8.2 and in 1, 5, and 10 cm cells.

4.9 References

Aßmann, S., Frank, C. and Körtzinger, A., 2011. Spectrophotometric high-precision seawater pH determination for use in underway measuring systems. Ocean Sci., 7(5): 597-607.

Bellerby, R.G.J., Olsen, A., Johannessen, T. and Croot, P., 2002. A high precision spectrophotometric method for on-line shipboard seawater pH measurements: the automated marine pH sensor (AMpS). Talanta, 56(1): 61-69.

Bockmon, E.E. and Dickson, A.G., 2014. A seawater filtration method suitable for total dissolved inorganic carbon and pH analyses. Limnology and Oceanography: Methods, 12(4): 191-195.

Bockmon, E.E. and Dickson, A.G., 2015. An inter-laboratory comparison assessing the quality of seawater carbon dioxide measurements. Marine Chemistry, 171: 36-43.

Byrne, R.H. and Breland, J.A., 1989. High precision multiwavelength pH determinations in seawater using cresol red. Deep Sea Research Part A. Oceanographic Research Papers, 36(5): 803-810.

Byrne, R.H., Mecking, S., Feely, R.A. and Liu, X., 2010. Direct observations of basin-wide acidification of the North Pacific Ocean. Geophysical Research Letters, 37(2).

Carter, B.R., Feely, R.A., Williams, N.L., Dickson, A.G., Fong, M.B. and Takeshita, Y., 2018. Updated methods for global locally interpolated estimation of alkalinity, pH, and nitrate. Limnology and Oceanography: Methods, 16(2): 119-131.

Carter, B.R., Radich, J.A., Doyle, H.L. and Dickson, A.G., 2013. An automated system for spectrophotometric seawater pH measurements. Limnology and Oceanography: Methods, 11(1): 16-27.

Chierici, M., Fransson, A. and Anderson, L.G., 1999. Influence of *m*-cresol purple indicator additions on the pH of seawater samples: correction factors evaluated from a chemical speciation model. Marine Chemistry, 65(3): 281-290.

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

Clayton, T.D., Byrne, R.H., Breland, J.A., Feely, R.A., Millero, F.J., Campbell, D.M., Murphy, P.P. and Lamb, M.F., 1995. The role of pH measurements in modern oceanic CO_2-system characterizations: Precision and thermodynamic consistency. Deep Sea Research Part II: Topical Studies in Oceanography, 42(2): 411-429.

DeGrandpre, M.D., Spaulding, R.S., Newton, J.O., Jaqueth, E.J., Hamblock, S.E., Umansky, A.A. and Harris, K.E., 2014. Considerations for the measurement of spectrophotometric pH for ocean acidification and other studies. Limnology and Oceanography: Methods, 12(12): 830-839.

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

Dickson, A.G., 2010. Standards for Ocean Measurements. Oceanography, 23(3): 34-47.

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

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

Dyrssen, D. and Hansson, I., 1973. Ionic medium effects in sea water — A comparison of acidity constants of carbonic acid and boric acid in sodium chloride and synthetic sea water. Marine Chemistry, 1(2): 137-149.

Fong, M.B. and Dickson, A.G., 2019. Insights from GO-SHIP hydrography data into the thermodynamic consistency of CO_2 system measurements in seawater. Marine Chemistry, 211: 52-63.

GUM, 1993. Guide to the Expression of Uncertainty in Measurement. ISO, Geneva.

Lee, K., Kim, T.-W., Byrne, R.H., Millero, F.J., Feely, R.A. and Liu, Y.-M., 2010. The universal ratio of boron to chlorinity for the North Pacific and North Atlantic oceans. Geochimica et Cosmochimica Acta, 74(6): 1801-1811.

Li, X., García-Ibáñez, M.I., Carter, B.R., Chen, B., Li, Q., Easley, R.A. and Cai, W.-J., 2020. Purified meta-Cresol Purple dye perturbation: How it influences spectrophotometric pH measurements. Marine Chemistry, 225: 103849.

Liu, X., Patsavas, M.C. and Byrne, R.H., 2011. Purification and Characterization of meta-Cresol Purple for Spectrophotometric Seawater pH Measurements. Environmental Science & Technology, 45(11): 4862-4868.

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

Martz, T.R., Carr, J.J., French, C.R. and DeGrandpre, M.D., 2003. A Submersible Autonomous Sensor for Spectrophotometric pH Measurements of Natural Waters. Analytical Chemistry, 75(8): 1844-1850.

McElligott, S., Byrne, R.H., Lee, K., Wanninkhof, R., Millero, F.J. and Feely, R.A., 1998. Discrete water column measurements of CO_2 fugacity and pHT in seawater: A comparison of direct measurements and thermodynamic calculations. Marine Chemistry, 60(1): 63-73.

Millero, F.J., 1979. The thermodynamics of the carbonate system in seawater. Geochimica et Cosmochimica Acta, 43(10): 1651-1661.

Millero, F.J., DiTrolio, B., Suarez, A.F. and Lando, G., 2009. Spectroscopic measurements of the pH in NaCl brines. Geochimica et Cosmochimica Acta, 73(11): 3109-3114.

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

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

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

Takeshita, Y., Warren, J.K., Liu, X., Spaulding, R.S., Byrne, R.H., Carter, B.R., DeGrandpre, M.D., Murata, A. and Watanabe, S.-i., submitted. Consistency and stability of purified meta-cresol purple for spectrophotometric pH measurements in seawater. Marine Chemistry.

Tapp, M., Hunter, K., Currie, K. and Mackaskill, B., 2000. Apparatus for continuous-flow underway spectrophotometric measurement of surface water pH. Marine Chemistry, 72(2): 193-202.

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

Williams, N.L., Juranek, L.W., Feely, R.A., Johnson, K.S., Sarmiento, J.L., Talley, L.D., Dickson, A.G., Gray, A.R., Wanninkhof, R., Russell, J.L., Riser, S.C. and Takeshita, Y., 2017. Calculating surface ocean pCO_2 from biogeochemical Argo floats equipped with pH: An uncertainty analysis. Global Biogeochemical Cycles, 31(3): 591-604.

Zhang, H. and Byrne, R.H., 1996. Spectrophotometric pH measurements of surface seawater at in-situ conditions: absorbance and protonation behavior of thymol blue. Marine Chemistry, 52(1): 17-25.

Chapter 5

Accuracy and consistency of spectrophotometric pH measurements with impure *m*-cresol purple

5.1 Abstract

Although spectrophotometric pH determination with a purified indicator dye (e.g., *m-cresol purple*) provides the most accurate seawater pH measurements, there is often a desire and need to obtain pH measurements with impure dyes due to the high cost and limited availability of purified dyes. We characterized the properties of a lot of purified and impure mCP and evaluated the performance of three different published methods for calibrating impure dyes using a combination of laboratory measurements and numerical simulations. When possible, we recommend characterizing the apparent absorption coefficient ratios and acid dissociation constant of the impure dye or measuring the dye lot-specific pH offsets relative to a purified dye. Another possible approach, which requires no purified dye and less overall effort, is to estimate the impurity absorption at 434 nm from measurements in a high pH solution and use this information to correct impure dye measurements. However, this method involves assumptions about the impurity absorption behavior and can result in errors in the corrected pH values. Despite the availability of methods for calibrating impure dyes, there can be inconsistencies between impure and purified mCP pH measurements larger than 0.005 due to the uncertainty in the properties of purified mCP and inadequately purified dyes. Our results highlight the need for further study of indicator impurity absorption behavior, a re-evaluation of the properties of purified mCP, and improvement of quality control procedures for the purification of mCP.

5.2 Introduction

Spectrophotometric pH measurement using purified indicator meta-cresol purple (mCP) is considered a benchmark method for seawater pH measurements (Dickson, 2010a; Dickson et al., 2007) and is often used to calibrate alternative methods of pH measurements, such as glass electrodes (Easley and Byrne, 2012), pH sensors based on Ion Sensitive Field Effect Transistors (Bresnahan Jr. et al., 2014; Johnson et al., 2016; Martz et al., 2010), and custom-designed "do-it-yourself" instruments (Yang et al., 2014). A recent study has demonstrated that spectrophotometric pH can provide reproducible measurements (within 0.003 pH units) between multiple laboratories provided that high quality spectrophotometers and adequately purified dyes are used (Takeshita et al., submitted). The spectrophotometric pH method therefore has promising potential to provide the high quality measurements needed (Newton et al., 2014) to study the long-term changes in the ocean's carbon chemistry resulting from the uptake of anthropogenic CO_2.

Since the development of spectrophotometric pH for use in seawater (Byrne and Breland, 1989; Clayton and Byrne, 1993; Zhang and Byrne, 1996), the method has been routinely used on ship-based repeat hydrography surveys to measure decadal changes in ocean pH (Byrne et al., 2010; Talley et al., 2016). However, it was later realized that spectrophotometric pH measurements had a significant bias due to colored impurities in the indicator dye, which affected the determination of pH (Yao et al., 2007). The subsequent development of purified indicators (Liu et al., 2011; Patsavas et al., 2013) has greatly improved the consistency and accuracy of spectrophotometric pH measurements. However, due to the high cost and limited availability of purified indicators, many laboratories may still use unpurified indicators. A variety of approaches for calibrating impure dyes are possible, each

with their own advantages and disadvantages. One approach is to characterize the optical properties and apparent acid dissociation constant of the dye (Clayton and Byrne, 1993; Liu et al., 2011), but this process is tedious. Another approach is to measure the offsets in the pH values measured with the impure dye relative to a reference dye over a range of pH (Yao et al., 2007). If purified dye is available, the impure dye measurements could be adjusted to be consistent with purified dye measurements. As it would be advantageous to not be reliant on the availability of purified dye for the calibration, a simple impurity correction method was developed that capitalized on the property that many impurities absorb strongly at 434 nm (one of the wavelengths used in spectrophotometric pH determination with *m*-cresol purple), and therefore, the impurity absorption at 434 nm can be estimated from measurements of impure dye solutions at high pH, where the dye's absorption at 434 nm is minimal (Douglas and Byrne, 2017). This method was shown to achieve adequate impurity correction for some, but not all impure mCP. For any of these methods, evaluating the consistency between impure and purified dye measurements ultimately requires that the properties of pure mCP are well-known and that adequately pure samples of mCP are available for the comparison. Some lots of purified mCP have been shown to have some impurity contamination (Takeshita et al., submitted), and independent characterizations of the properties of purified mCP have shown discrepancies (DeGrandpre et al., 2014; Loucaides et al., 2017; Müller and Rehder, 2018) which will contribute to the uncertainty in implementing the various impure dye calibration methods. There is a need, therefore, to investigate the reliability of the various dye impurity correction approaches and to understand the quality of pH measurements that might be expected when using impure dyes.

In this study, we characterized a single lot of purified mCP and a single lot of unpurified mCP and used a combination of laboratory measurements with these dyes and numerical simulations to evaluate the performance of the various dye impurity correction approaches and assess their errors and consistency with purified mCP measurements. Finally, we suggest recommended practices when calibrating impure indicators.

5.3 Methods

5.3.1 Spectrophotometric pH measurements

The principles of spectrophotometric pH determination have been described in **Chapter 3** and in previous studies (Byrne and Breland, 1989; Clayton and Byrne, 1993; Zhang and Byrne, 1996). Briefly, the pH of a solution containing a pH-sensitive indicator dye, typically a sulfonephthalein compound, is determined from its absorption spectrum, which reflects the relative proportions of the indicator's acid and base forms. The pH can therefore be calculated from

$$\mathrm{pH} = \mathrm{p}K_a(\mathrm{HI}^-) + \log\left(\frac{[\mathrm{I}^{2-}]}{[\mathrm{HI}^-]}\right) \quad (1).$$

The concentration ratio ($[I^{2-}]/[HI^-]$) is typically estimated from the ratio of the absorbances at the wavelengths corresponding to the absorbance maxima of the I^{2-} and HI^- species (i.e., $R = A_{578}/A_{434}$ for *m*-cresol purple, the most widely used spectrophotometric indicator), as in **Eq. 2**.

$$\mathrm{pH} = \mathrm{p}K_a(\mathrm{HI}^-) + \log\left(\frac{R - e_1}{e_2 - Re_3}\right) \quad (2)$$

e_1, e_2, and e_3 are molar absorption coefficient ratios ($e_1 = \varepsilon_{578}(HI^-)/\varepsilon_{434}(HI^-)$, $e_2 = \varepsilon_{578}(I^{2-})/\varepsilon_{434}(HI^-)$, and $e_3 = \varepsilon_{434}(I^{2-})/\varepsilon_{434}(HI^-)$), which are determined from measurements in high and

low pH solutions where nearly all of the dye is present in a single form (i.e., I^{2-} or HI^{-}). Because the ratios e_2 and e_3, being derived from a combination of measurements in separate solutions with significantly different pH, may have a larger uncertainty, an alternative form of **Eq. 2** is often preferred. This form of the equation (**Eq. 3**) combines $pK_a(HI^{-})$ and e_2 into a single term and only requires determining two ratios, e_1 and e_3/e_2, which can be determined from measurements in single solutions at low and high pH, respectively (Liu et al., 2011).

$$\mathrm{pH} = -\log(K_a(\mathrm{HI}^{-}) \cdot \mathrm{e}_2) + \log\left(\frac{R-e_1}{1-R\frac{e_3}{e_2}}\right) \quad (3)$$

Once the absorption coefficient ratios have been determined, $pK_a(HI^{-})$ and $-\log(K_a(HI^{-}) \cdot e_2)$ (i.e., $p(K_a(HI^{-}) \cdot e_2)$) can be determined from **Eq. 2** or **Eq. 3** by measuring R in TRIS buffers which have been assigned a pH value. The terms e_1, e_3/e_2, and $p(K_a(HI^{-}) \cdot e_2)$ have been characterized over a range of temperatures and salinities for purified mCP (Liu et al., 2011).

Impure dyes can be characterized in the same way. Because the apparent values of e_1, e_2, e_3, $pK_a(HI^{-})$ and $-\log(K_a(HI^{-}) \cdot e_2)$ are affected by the absorption of the impurity which will likely vary between different lots of dye and different manufacturers, the apparent values determined for an impure dye are specific to a particular lot of dye. As the impurity absorption behavior is built into the apparent dye properties, they can be used to determine the pH of samples measured with that lot of dye.

5.3.2 Indicator characterization

We characterized one lot of purified mCP (Lot FB4, obtained from Robert H. Byrne, U. of Southern Florida) and one lot of impure mCP (Acros Lot A0320498) at a temperature of 25°C and ionic strength of ~0.69 mol kg^{-1} in a NaCl background. All measurements were made using a 10 cm cell on an Agilent 8453 single beam spectrophotometer, using the semi-

automated system of Carter et al. (2013). The set of absorption coefficient ratios e_1, e_2, and e_3 for each dye were determined from measurements of the spectra in three different solutions at pH ≈ 0, 4.5, 12, where the H_2I, HI^-, and I^{2-} species dominate, respectively. The spectra were corrected for contributions from minor species by processing the spectra using an approach similar to that used by Ohline et al. (2007) and Husheer (2001). These corrected spectra represent the pure absorption spectra for the H_2I, HI^-, and I^{2-} species. The absorption coefficient ratios reported in **Table 5.1** were determined from the HI^- and I^{2-} species spectra by calculating the ratios described in the previous section. The full species absorption spectra (350-750 nm) were used in the various simulations discussed later. Additionally, we determined $pK_a(HI^-)$ and $p(K_a(HI^-)\cdot e_2)$ for the FB4 dye from measurements in a single bottle of equimolar TRIS-TRIS·H^+ buffer standard in synthetic seawater (Batch T32) prepared according to DelValls and Dickson (1998). Further details on these measurements were reported in **Chapter 3**.

Table 5.1. Values for the dye properties of the FB4 lot of purified mCP, a lot of impure mCP (Acros), and the lot of purified mCP characterized by Liu et al. (2011). The absorption coefficient ratios for the FB4 and Acros dyes were determined in NaCl solutions. $pK_a(HI^-)$ and $p(K_a(HI^-)\cdot e_2)$ were also determined for the FB4 dye. Values of ${}_{434}A_{imp}$ (for spectra scaled to $A_{iso} = 1$) were separately calculated using the FB4 and Liu et al.'s e_3/e_2 values as the references (**Eq. 5**).

	FB4	**Acros**	**Liu et al.**
e_1	0.005965	0.007278	0.005707
e_2	2.2972	2.2555	2.2267
e_3	0.12722	0.14225	0.12646
e_3/e_2	0.05538	0.06307	0.05678
$pK_a(HI^-)$	8.0154	n/a	7.9963
$p(K_a(HI^-)\cdot e_2)$	7.6542	n/a	7.6486
${}_{434}A_{imp}$ (FB4)	0	0.0340	0.0063
${}_{434}A_{imp}$ (Liu)	–0.0063	0.0278	0

5.3.3 Correcting measurements made with impure indicators

Commercial indicators contain impurities that absorb at the wavelengths relevant to spectrophotometric pH determination, which can lead to pH-dependent biases in pH larger than 0.01 units (Liu et al., 2011). Yao et al. (2007) first recommended that samples of impure indicators be archived, so that historical pH measurements may be corrected at a later date by calibrating the impure dye against a purified dye when it becomes available. A dye lot-specific correction curve can be constructed from paired pH measurements with the impure dye and a sample of purified dye in solutions over a range of pH values (e.g., 7.2–8.2 for seawater). The pH offsets relative to the user's sample of purified dye are estimated from the difference of the *apparent* pH values measured with the impure and purified dye ($\Delta pH = pH_{imp}' - pH_{pure}'$), where both pH_{imp}' and pH_{pure}' are calculated using the properties of the reference purified dye (e.g., the absorption coefficient ratios and $p(K_a(HI^-) \cdot e_2)$ of Liu et al. (2011)). The ΔpH values are plotted against pH_{imp}' and a second-order polynomial fit to the data is used to estimate the required offset correction (ΔpH) for a sample measured with the impure dye. The corrected pH value is obtained by adding ΔpH to pH_{imp}' (**Eq. 4**).

$$pH_{corr} = pH_{imp}' - \Delta pH \quad (4)$$

Ideally, pH_{corr}, the final result obtained from **Eq. 4**, represents the true pH that would be obtained with a pure dye. This approach thus requires that the user's lot of purified dye has identical properties to the reference purified dye. The prime notation in pH_{pure}' denotes an apparent value, recognizing the possibility of differences between different lots of purified dye.

5.3.4 Determining corrections for the impurity absorption at 434 nm ($_{434}A_{imp}$)

Douglas and Byrne (2017) showed that pH measurements made with impure indicators can be meaningfully corrected by estimating the impurity absorption at 434 nm ($_{434}A_{imp}$) from measurements at high pH (pH~12). As this approach involves a straightforward measurement and requires no purified dye, it has advantages over the more tedious approach of measuring dye lot-specific pH offsets or characterizing an impure dye. However, $_{434}A_{imp}$-based corrections assume that (1) there is no impurity absorption at 578 nm and (2) the impurity absorption is independent of pH.

$_{434}A_{imp}$ is estimated by comparing the R of the impure dye measured in a NaOH solution (pH~12) to the known value for the reference purified mCP ($R = e_2/e_3$ at pH~12).

$$_{434}A_{imp} = \left(1 - \frac{R^{imp}_{NaOH}}{R^{pure}_{NaOH}}\right) {}_{434}A^{imp}_{NaOH} \quad (5)$$

In **Eq. 5**, R^{pure}_{NaOH} is the value for the reference purified mCP (Liu et al., 2011); R^{imp}_{NaOH} and $_{434}A^{imp}_{NaOH}$ are the absorbance ratio and the absorbance at 434 nm, respectively, measured with the impure dye in the NaOH solution. As $_{434}A_{imp}$ is proportional to the dye concentration, it will need to be scaled to the dye concentration in the sample. In this study, we used the isosbestic absorbance estimated at 488.1 nm (A_{iso}) from the weighted average of the absorbances at the diodes at 488 and 489 nm as a proxy for dye concentration (Carter et al., 2013). We therefore scaled $_{434}A_{imp}$ by the ratio of the A_{iso} of the sample spectrum to the A_{iso} of the impure dye NaOH spectrum (i.e., $_{434}A^{*}_{imp} = {}_{434}A_{imp} \times \left(\frac{A^{samp}_{iso}}{A^{NaOH}_{iso}}\right)$). The R values of samples measured with the impure dye are adjusted for $_{434}A_{imp}$ with **Eq. 6**.

$$R_{434corr} = R_{samp}\left(1 + \frac{{}_{434}A^{*}_{imp}}{{}_{434}A_{samp} - {}_{434}A^{*}_{imp}}\right) \quad (6)$$

Finally, the ${}_{434}A_{\mathrm{imp}}$-corrected pH is calculated from **Eq. 3** using the adjusted R values ($R_{434corr}$) and the reference dye properties.

5.3.5 Comparative measurements of pH with the Acros and FB4 dyes

A set of solutions over the pH range of approximately pH 7.1 to 8.1 was measured with the FB4 and Acros dyes to compare the pH performance of the two independently characterized dyes against each other. These measurements were also used to determine the pH offsets for the Acros dye and to evaluate the ${}_{434}A_{\mathrm{imp}}$-based corrections.

The comparison samples consisted of an unbuffered solution of ~0.02 mol kg^{-1} sodium acetate in NaCl background (~0.7 mol kg^{-1} ionic strength, adjusted to pH~7.11 with HCl), a buffered solution of ~0.08 mol kg^{-1} 3-[4-(2-Hydroxyethyl)-1-piperazinyl]propanesulfonic acid (EPPS) and ~0.042 mol kg^{-1} NaOH in NaCl background (~0.7 mol kg^{-1} ionic strength, pH~7.79), a CO_2-in-seawater Reference Material (CRM 186, pH~7.86) prepared in the Dickson Laboratory at the Scripps Institution of Oceanography (Dickson, 2010b), and a 0.04 mol kg^{-1} approximately equimolal TRIS/TRIS-HCl buffer in synthetic seawater (S=35, pH~8.07). Four subsamples were drawn from each individual bottle for the pH measurements. For each bottle, the pH was measured with two separate additions of the FB4 dye followed by two additions of the Acros dye. For the EPPS buffer measurements, a single batch of buffer was filled into six 125 mL bottles. One set of three bottles was measured with the FB4 dye, and the other set with the Acros dye.

As some of the solutions were unbuffered, the pH values required an adjustment for the effect of the indicator addition. The pH values from the two dye additions at different dye

amounts were extrapolated to zero isosbestic absorbance ($A_{iso} = 0$) for the unbuffered solutions. For the buffered solutions, the pH values were simply averaged. The difference between the pH values obtained with the Acros and FB4 dyes were plotted against pH (**Figure 5.1**). The pH values were calculated with **Eq. 2** assuming a value of 7.9913 for pK_a(HI^-) and the appropriate absorption coefficient ratios for each dye. We used a value of 7.9913 for pK_a(HI^-) to calculate pH as we did not separately determine the apparent pK_a(HI^-) for the Acros dye.

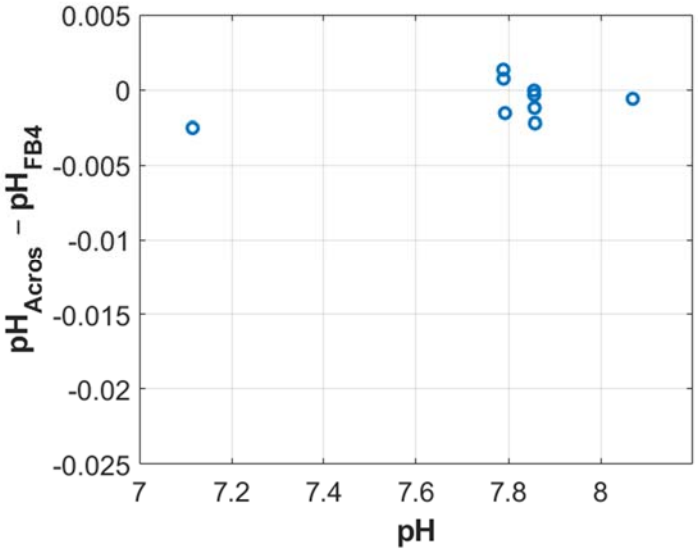

Figure 5.1. Difference in the pH measured with the Acros and FB4 dyes in NaCl and seawater solutions (pH 7.11 to 8.07) plotted against pH. The pH values were calculated with **Eq. 2** using the absorption coefficient ratios appropriate to the dye (**Table 5.1**) and a value of 7.9913 for pK_a(HI^-). Each point represents measurements made in a single bottle, except for the second group of points from the left where each of the three points represents a pair of bottles filled from a single batch of buffer (see text).

5.3.6 Assessment of errors in determining dye lot pH offsets

The set of dye properties we determined for our FB4 dye differed from those reported by Liu et al. (2011) for a different lot of purified mCP. Hence, the choice of dye properties can result in inconsistencies when correcting for dye lot pH offsets. We constructed two pH offset correction curves for the Acros dye from the comparison sample measurements, one using the

set of FB4 dye properties to calculate ΔpH and pH_{imp} and the other using the properties of Liu et al. (2011) (**Figure 5.2**).

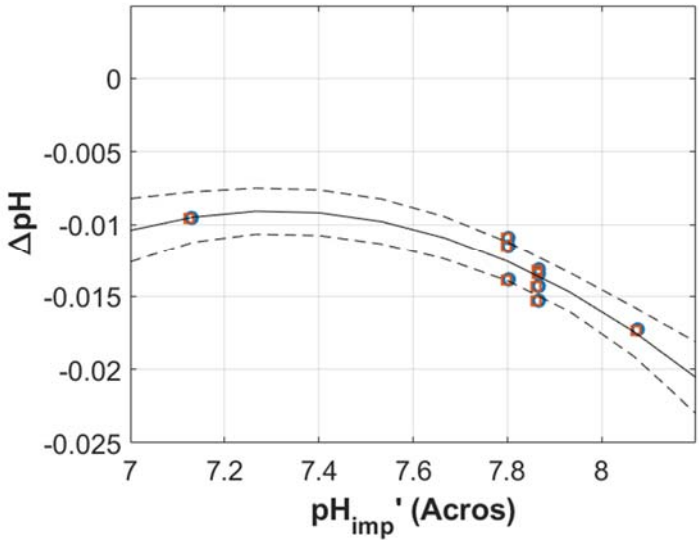

Figure 5.2. Measured pH offsets ($\Delta pH = pH_{imp}' - pH_{pure}'$) for the Acros dye relative to the FB4 dye. The measurements were made in the same solutions as in **Figure 5.1**, and the two sets of data show ΔpH values calculated using the FB4 dye properties (circles) and the properties of Liu et al. (2011) (squares). A second order polynomial curve was fit to the set of data (solid curve) calculated with the FB4 properties, and the dashed curves show the standard error of prediction.

We assessed the precision in estimating ΔpH from the correction curve with a synthetic dataset, as our measurements were not evenly distributed across the pH range. When fitting a second-order polynomial to the data, a minimum of four points is needed to have at least one degree of freedom ($df = n - 3$). A set of four spectra representing pH 7.2 to 8.2 was simulated for the FB4 and Acros dyes from linear combinations of their apparent HI^- and I^{2-} spectra. To simulate the effect of random absorbance errors on the correction curve, a normally-distributed error with a standard deviation of 0.00031 (an estimate of the repeatability of the absorbance measurements of the Agilent 8453; Carter et al., 2013) was added to the absorbance at each wavelength and the resulting perturbed pH values were used to construct the correction curve.

The correction curves were simulated 1,000 times, and the standard errors of prediction from the polynomial fit were pooled to obtain an estimate of the precision in estimating ΔpH.

Because the dye properties are a function of temperature and salinity, we simulated ΔpH correction curves at different temperatures (10°C, 20°C, and 25°C) and salinities (S = 20, 30, and 35) (**Figure 5.3**). As we did not determine the temperature and salinity-dependence for our dye properties, we used the properties of Liu et al. (2011) and Clayton and Byrne (1993) to calculate *R* values representing different pH values for purified and impure mCP and the corresponding apparent pH values (i.e., calculated with Liu et al.'s dye properties). The standard errors shown in **Figure 5.3** for each curve are pooled values from 1,000 simulations, assuming similar absorbances to the spectra in the previous simulation and the same repeatability in the absorbance measurements.

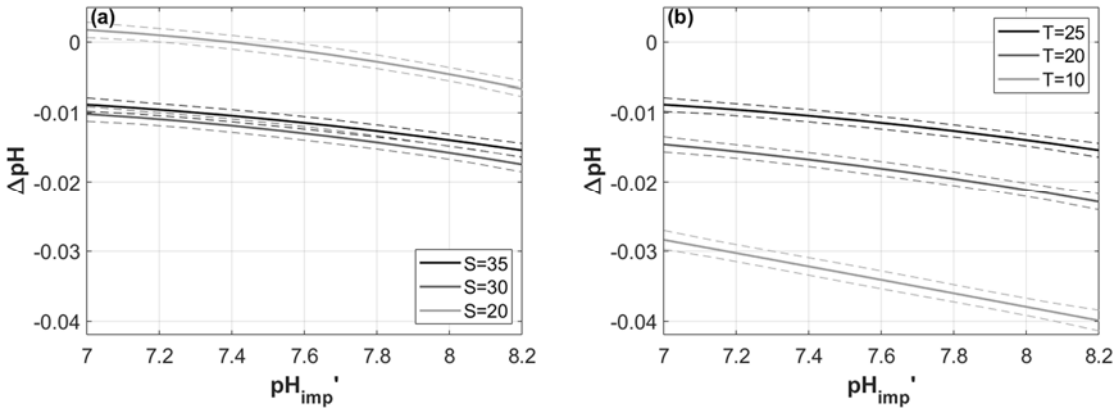

Figure 5.3. Simulated pH offset curves and their standard errors at different (a) salinities and (b) temperatures for the impure dye from Clayton and Byrne (1993). The ΔpH values were calculated using the properties of Liu et al. (2011).

5.3.7 Multivariate Curve Resolution analysis

We explicitly resolved the impurity absorption spectrum of the Acros dye with Multivariate Curve Resolution analysis and subsequently used this information in numerical

simulations to obtain insights about the impacts of assumptions about the impurity absorption behavior on ${}_{434}A_{\text{imp}}$ corrections (see next section). Multivariate Curve Resolution (MCR) is a method for deconvolving a set of mixture spectra into their pure component spectra, when given various constraints and often with limited information (de Juan et al., 2014). MCR has been applied to studies of sulfonephthalein indicators to obtain the absorption spectra of their acid-base forms (Shimada and Hasegawa, 2017; Shimada et al., 2019).

MCR analysis decomposes a data matrix of mixture spectra (**A**) into the pure component spectra (**S**) and their corresponding relative contributions (**C**) to the overall sample spectra. Impure dye spectra can be modeled as being a mixture of the HI^-, I^{2-}, and the impurity absorption spectra (**Eq. 7**).

$$\mathbf{A} = \mathbf{c}_{\mathbf{HI^-}}\mathbf{S}_{\mathbf{HI^-}} + \mathbf{c}_{\mathbf{I^{2-}}}\mathbf{S}_{\mathbf{I^{2-}}} + \mathbf{c}_{\mathbf{Imp}}\mathbf{S}_{\mathbf{Imp}} + \mathbf{E} \tag{7}$$

In **Eq. 7**, **A** is an $n \times m$ matrix of n spectra measured over m wavelengths; $\mathbf{S}_{\mathbf{HI^-}}$, $\mathbf{S}_{\mathbf{I^{2-}}}$, and $\mathbf{S}_{\mathbf{Imp}}$ are the HI^-, I^{2-}, and impurity absorption spectra; $\mathbf{c}_{\mathbf{HI^-}}$, $\mathbf{c}_{\mathbf{I^{2-}}}$, and $\mathbf{c}_{\mathbf{Imp}}$ ($n \times 1$ vectors) are the relative contributions of the components (and not concentrations, strictly speaking); and **E** is a residuals matrix ($n \times m$).

We used the MCR-ALS GUI 2.0 program in Matlab (Jaumot et al., 2015) for the MCR analysis. Given a set of input spectra (**A**), the MCR analysis first requires knowledge of the number of components in the mixture (obtained through a rank analysis or from other information) and then calculating initial estimates of the solutions **C** and **S**. In the MCR-ALS GUI 2.0 program, the initial estimate step is performed using the Simple-to-Use Interactive Self-Modeling Mixture Analysis (SIMPLISMA) algorithm (Windig and Guilment, 1991), which identifies wavelengths where the absorbances respond mainly to one of the components

in the mixture. Next, constraints are defined (see later discussion), and finally, the solutions are optimized by alternating least squares (ALS), first solving for **S** given an initial estimate of **C** and the various constraints, then solving for **C** given the estimate of **S** and the various constraints, and iteratively refining the solutions until convergence criteria are achieved. In each iterative cycle, **C** or **S** are solved by least squares with the goal of minimizing the norm of the residuals between the data reproduced from Principal Components Analysis ($\mathbf{A_{PCA}}$) using the first *n* principal components (where *n* is the number of components in the mixture) and the ALS-reproduced data ($\mathbf{A_{ALS}}$), calculated from **C** and **S**.

The uneven distribution of our comparison sample measurements would overly weight the MCR model towards the high pH data, and thus we used synthetic Acros dye spectra in **A** for the MCR analysis. Synthetic spectra representing a pH range of 7 to 8 were simulated from linear combinations of the apparent HI^- and I^{2-} spectra that we measured for the Acros dye in NaCl background. In addition to the Acros spectra, pure dye spectra also need to be included in **A**, because the detection and resolution of the impurity component by MCR, which involves Principal Components Analysis in the optimization step, requires that the set of spectra contain variation due to the impurity. We therefore included the HI^- and I^{2-} spectra of the FB4 dye in **A** and also used these as constraints in the MCR analysis. That is, given a set of synthetic Acros dye spectra and the HI^- and I^{2-} spectra of pure mCP, the MCR analysis is performed to solve for $\mathbf{S_{Imp}}$ and **C**.

Another key constraint needed to reduce the uncertainty in the MCR solutions is closure in **C**:

$$\mathbf{c}_{HI^-} + \mathbf{c}_{I^{2-}} + \mathbf{c}_{Imp} = 1 \qquad (8).$$

Achieving the closure constraint requires scaling the spectra in **A** appropriately. We normalized the spectra at the isosbestic point at 488.1 nm. The pure mCP spectra were normalized to A_{iso} = 1. The Acros spectra were scaled to a slightly lower A_{iso}, as these spectra contain an impurity contribution. The appropriate scaling factor was estimated from the apparent I^{2-} spectrum of the Acros dye by decomposing its vector length into contributions from the dye and the impurity under the same closure constraint (**Figure 5.4, Eq. 9**).

$$f \,||\, \mathbf{S}_{\mathbf{I}^{2-}_{app}} \,|| = \sqrt{(fc_{I^{2-}} \,||\, \mathbf{S}_{\mathbf{I}^{2-}} \,|| + fc_{\mathrm{Imp}} \,||\, \mathbf{S}_{\mathbf{Imp}} \,|| \cos\theta_1)^2 + (fc_{\mathrm{Imp}} \,||\, \mathbf{S}_{\mathbf{Imp}} \,|| \sin\theta_1)^2} \qquad (9)$$

In **Eq. 9**, $\mathbf{S}_{I^{2-}_{app}}$ is the apparent I^{2-} spectrum of the Acros dye; $\mathbf{S}_{I^{2-}}$ is the pure I^{2-} spectrum (obtained from the FB4 I^{2-} spectrum); $\mathbf{S_{imp}}$ is estimated from the difference between the apparent I^{2-} spectrum of the Acros dye and the FB4 I^{2-} spectrum, both normalized at 578 nm. The closure constraint requires that $c_{I^{2-}} + c_{\mathrm{Imp}} = 1$. The equation is solved with the MATLAB function fmincon to estimate *f*, which is the inverse of the scaling factor. That is, if the pure dye spectra in **A** are scaled to A_{iso} = 1, the Acros dye spectra should be scaled to $A_{iso} = 1/f$. The result of the MCR analysis is shown in **Figure 5.5**.

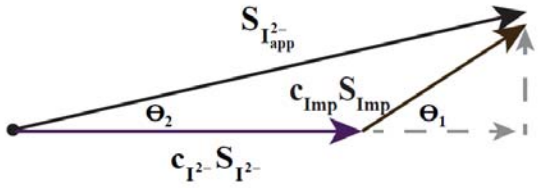

Figure 5.4. Schematic showing the decomposition of the apparent I^{2-} spectrum of an impure dye into contributions from the dye and impurity

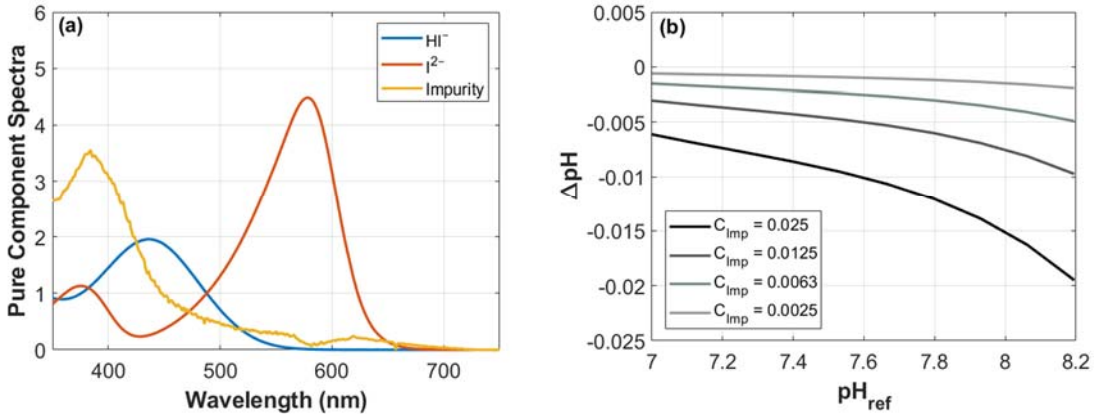

Figure 5.5. (a) Absorption spectra for the HI^- and I^{2-} species of m-cresol purple, obtained from measurements in NaCl solutions with the FB4 lot of purified mCP and the absorption spectrum for the Acros mCP estimated from Multivariate Curve Resolution. The HI^- and I^{2-} spectra are normalized to an absorbance of 1 at the isosbestic wavelength (488.1 nm). The impurity absorption spectrum is scaled so that $\mathbf{c}_{HI^-} + \mathbf{c}_{I^{2-}} + \mathbf{c}_{Imp} = 1$ in **Eq. 7** for a set of Acros dye spectra. (b) Simulated pH offsets ($\Delta pH = pH_{imp}' - pH_{pure}'$) for the impure dyes in **Table 5.2**.

5.3.8 Simulations of ${}_{434}A_{imp}$ corrections

We performed numerical simulations to evaluate how assumptions about the impurity absorption behavior and the properties of the reference purified dye affect the performance of the ${}_{434}A_{imp}$-based correction method for impure dyes with varying levels of impurities. Impure dye spectra representing pH values over the range of 7 to 8.2 were simulated from linear combinations of the HI^-, I^{2-}, and impurity absorption spectra shown in **Figure 5.5.** The proportion of impurity absorption was the same in each individual spectrum, and four sets of such spectra were simulated with the proportion of impurity absorption (c_{Imp}) varying from 0.0025 to 0.025 to represent dyes with different levels of impurities. The pH offsets relative to pure dye are plotted in **Figure 5.5** for these four simulated impure dyes, and their absorption coefficient ratios are given in **Table 5.2**.

Table 5.2. Absorption coefficient ratios and $_{434}A_{imp}$ values for the dyes used in the numerical simulations of $_{434}A_{imp}$ corrections. The pure dye ($c_{Imp} = 0$) is the FB4 dye (see **Table 1**), and the other dyes were simulated by adding different proportions of the impurity absorption spectrum shown in **Figure 5** to the pure dye HI^- and I^{2-} spectra, with the requirement that $c_{HI^-} + c_{Imp}$ and $c_{I^{2-}} + c_{Imp}$ equal 1. $_{434}A_{imp}$ values (for spectra scaled to $A_{iso} = 1$) were separately calculated using the values of e_3/e_2 for the pure dye and a dye with minor impurity contamination ($c_{Imp} = 0.005$) as the references (**Eq. 5**).

c_{Imp}	**0.025**	**0.0125**	**0.0063**	**0.005**	**0.0025**	**0**
e_1	0.007128	0.006544	0.006256	0.006196	0.006080	0.005965
e_2	2.2558	2.2756	2.2854	2.2875	2.2914	2.2972
e_3	0.14259	0.13482	0.13099	0.13019	0.12865	0.12722
e_3/e_2	0.06321	0.05925	0.05732	0.05691	0.05614	0.05538
$_{434}A_{imp}$ (Ref. $c_{Imp} = 0$)	0.0347	0.0172	0.0087	0.0069	0.0034	0
$_{434}A_{imp}$ (Ref. $c_{Imp} = 0.005$)	0.0279	0.0104	0.0018	0	–0.0035	–0.0069

To evaluate the $_{434}A_{imp}$ correction, values of $_{434}A_{imp}$ were calculated for the four dyes from **Eq. 5** using the value of e_3/e_2 of the FB4 dye ($R^{pure}_{NaOH} = e_2 / e_3$) and obtaining R^{imp}_{NaOH} and $_{434}A^{imp}_{NaOH}$ from the apparent I^{2-} spectrum of the dye (assuming the same proportion of impurity absorption as in the pH 7 to 8.2 spectra). The R values corresponding to the pH 7 to 8.2 spectra were adjusted with the appropriate $_{434}A_{imp}$ values (**Eq. 6**), and the pH was calculated using the adjusted R values, the absorption coefficient ratios of the FB4 dye, and a value of 7.9913 for $pK_a(HI^-)$ (**Eq. 2**). The difference between the $_{434}A_{imp}$-corrected pH values and the true pH values is plotted in **Figure 5.6a.**

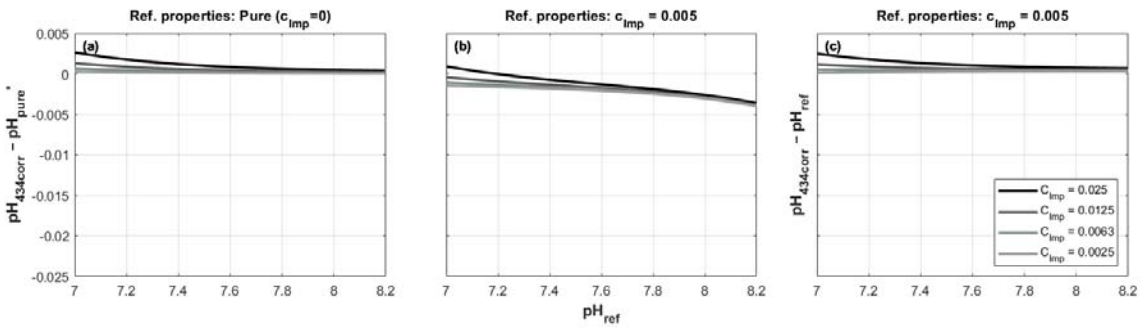

Figure 5.6. Simulated pH offsets (relative to pure dye measurements) for four impure dyes (see **Table 5.2**), calculated using $_{434}A_{imp}$-corrected pH values for the impure dye and different sets of reference dye properties. The differences in **(a)** represent the offsets relative to a pure dye. **(b)** shows the apparent offsets relative to a pure dye if the pH values were calculated using a set of reference properties for a dye with a small amount of impurity contamination (c_{Imp} = 0.005) **(c)** shows the same data as in **(b)**, but the offsets were calculated relative to the true pH values.

We also evaluated how an error in the reference dye properties will affect the performance of the $_{434}A_{imp}$ correction and the apparent consistency of the corrected pH values with pure dye measurements. In these simulations, the same calculations as described above were performed, but the reference absorption coefficient ratios represented a dye that had a small amount of impurity (c_{Imp} = 0.005). Thus, using these absorption coefficient ratios as the reference values will result in error in estimating the $_{434}A_{imp}$-corrected pH values and in assessing the agreement of the corrected pH values with pure dye measurements, as the apparent pure dye pH values will be wrong if calculated with the wrong set of coefficients. The offsets relative to the apparent pure dye pH values and to the true pH values are shown in **Figure 5.6**c.

5.4 Results

5.4.1 pH performance of the FB4 and Acros dyes

The pH values measured with the FB4 and Acros dyes in the comparison samples (calculated with the absorption coefficient ratios we determined for the two dyes and pK_a(HI^-)

=7.9913) agreed well with each other, with a mean difference of –0.00071±0.0023 (**Figure 5.1**). Although this agreement does not necessarily mean that our characterization of the dyes was without error, the similar pH performance of the two dyes indicates that the errors in the two sets of absorption coefficient ratios (**Table 5.1**) were at least similar and therefore cancelled when calculating the difference in pH.

However, the discrepancies between the absorption coefficient ratios we determined for the FB4 dye and those reported by Liu et al. (2011) for a different lot of purified mCP suggested potentially large inconsistencies in pH. For instance, the value of $\log(e_2)$ we determined for the FB4 dye was ~0.014 units higher than the value we obtained from personal communication with X. Liu, which would imply a difference of 0.014 in pH when using **Eq**. **2**. As discussed in **Chapter 3**, this discrepancy can be explained by the large uncertainty in our determination of e_2 due to the error in normalizing the HI^- and I^{2-} spectra at the 488.1 nm isosbestic point, which is particularly sensitive to any small wavelength errors on our spectrophotometer.

The difference in pH corresponding to the differences between the full set of FB4 dye properties (including $pK_a(HI^-)$ or $p(K_a(HI^-)\cdot e_2)$) and those of Liu et al. (2011) is smaller, because the errors in the absorption coefficient ratios are correlated with the error in determining $pK_a(HI^-)$ and $p(K_a(HI^-)\cdot e_2)$ and therefore cancel when calculating pH with **Eq. 2** or **Eq. 3**. Using the $pK_a(HI^-)$ and absorption coefficients that we determined for the FB4 dye, the pH values of the samples in the FB4 dataset were 0.0042 to 0.0054 higher than the values calculated using the published dye properties from Liu et al. (2011). This discrepancy is similar to the pH offset that we measured in the TRIS buffer standard, which was 0.0045 units lower

than the published value of DelValls and Dickson (1998) when using the Liu et al.'s dye properties. As the TRIS buffer was used to determine the pK_a(HI$^-$) of the FB4 dye, the pH values calculated with the FB4 dye properties will be ~0.0045 units higher than those calculated with the properties of Liu et al. (2011).

These results imply that a full characterization (i.e., determining the absorption coefficient ratios and pK_a(HI$^-$) or $\mathrm{p}(K_a(\mathrm{HI}^-)\cdot e_2)$) is needed to minimize the effects of errors in determining the absorption coefficient ratios. The resulting uncertainty in pH obtained with the pK_a(HI$^-$)-based **Eq**. **2** versus the $\mathrm{p}(K_a(\mathrm{HI}^-)\cdot e_2)$-based **Eq. 3** are similar (see **Chapter 3**), so there is no clear advantage to using one form of the pH equation over the other.

5.4.2 Errors in determining dye lot pH offsets

The pH offsets we measured for the Acros dye relative to the FB4 dye in the comparison samples (**Figure 5.2**) were essentially the same whether the apparent pH values, $\mathrm{pH_{imp}}'$ and $\mathrm{pH_{pure}}'$, were calculated with the set of FB4 dye properties or with the properties of Liu et al. (2011). Because a single set of values for the dye properties are used when calculating ΔpH, the ΔpH values reflect the difference in the pH due to the effect of impurities, as inferred from the difference in the *R* values measured with the two dyes in the same solutions. However, the final offset-adjusted pH values ($\mathrm{pH_{corr}}$) will necessarily be dependent on the set of dye properties used in the calculation of pH. As discussed in the previous section, the pH values calculated with the set of properties we determined for the FB4 dye were on average ~0.0047 higher than the values calculated with the coefficients of Liu et al. (2011), and thus the adjusted Acros pH measurements will also have this offset.

In addition to the error from assumptions about the properties of pure mCP, the dye pH offset adjustments will also have an error from estimating ΔpH from the second polynomial curve fit to the calibration ΔpH data. The standard error of estimating the required adjustment for an individual sample from the curve fit to the data in **Figure 5.2** was as large as 0.0024, due to the uneven distribution of our data across the pH range and bottle-to-bottle variance in pH for some of our samples. The pooled standard error of prediction from the Monte Carlo simulations was ~0.001 for four points evenly distributed over the range of pH 7.2-8.2 and assuming a repeatability of 0.00031 in the absorbance measurements. This level of precision should be achievable if care is taken to obtain measurements appropriately distributed across the working pH range.

The simulated ΔpH curves at different temperatures and salinities (**Figure 5.3**) showed that the ΔpH values were more sensitive to temperature than salinity. For a salinity range of 30-35, within the range of open ocean waters, the ΔpH curves will be indistinguishable from each other within their standard error of prediction. The ΔpH curve at a salinity of 20, however, was offset from the curves at salinity 30 and 35 by as much as ~0.01 pH units. Thus, care should be taken to construct the correction curve using measurements at the same temperature and salinity as the samples, particularly when working with low salinity samples.

5.4.3 Errors in ${}_{434}A_{imp}$ corrections

The choice of the reference dye properties will affect the calculation of ${}_{434}A_{imp}$ and the resulting impurity-corrected pH values (**Table 5.1**). Although, theoretically, pure dyes should have an ${}_{434}A_{imp}$ of 0, the FB4 dye had a negative ${}_{434}A_{imp}$ value (–0.0063 at $A_{iso} = 1$) when using the properties of Liu et al. (2011) as the reference. Other studies have also reported e_3/e_2 values for different lots of purified mCP that were lower than Liu et al.'s value, implying negative

$_{434}A_{imp}$ values for those dyes as well (DeGrandpre et al., 2014; Loucaides et al., 2017; Takeshita et al., submitted). The negative $_{434}A_{imp}$ values in recent lots of purified mCP may suggest trace levels of impurity contamination in the lot of mCP characterized by Liu et al., but this hypothesis requires further investigation. This uncertainty in the properties of pure mCP is reflected in our assessment of the performance of the $_{434}A_{imp}$ corrections. The apparent discrepancy in pH between the FB4 and Acros dyes in the comparison samples was –0.0006±0.0011 (*mean±std dev.*) and –0.0029±0.0012 (**Figure 5.7**) when calculated using the FB4 dye properties and Liu et al.'s properties, respectively, as the reference.

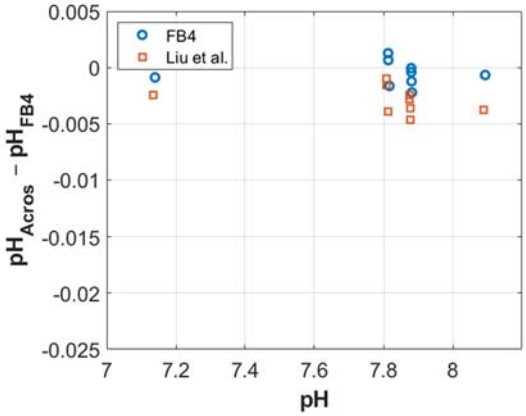

Figure 5.7. Difference between the $_{434}A_{imp}$-corrected Acros dye pH values and the FB4 dye pH values, measured in NaCl and seawater solutions (as in **Figure 5.1**). The $_{434}A_{imp}$ and pH values were calculated using the set of absorption coefficient ratios and $\mathrm{p}(K_a(\mathrm{HI}^-)\cdot e_2)$ as indicated in the legend (see also **Table 5.1**).

In addition to the uncertainty in the pure mCP properties, some of the assumptions about the impurity absorption behavior in the $_{434}A_{imp}$ correction method of Douglas and Byrne (2017) may also be problematic. The Acros dye impurity absorption spectrum, estimated from MCR analysis, showed a curious trough near 578 nm (**Figure 5.5**). Although the cause of this feature is unclear, it may be an artifact of an instrumental contribution in the set of spectra used to estimate the impurity absorption spectrum from MCR. Despite this, the estimated impurity

absorption spectrum had a small non-zero absorption at 578 nm. The effect of a small impurity absorption at 578 nm will have a greater impact on the error of the correction at low pH, where A_{578} is small, and the effect is greater at larger levels of impurity (**Figure 5.6a**). This behavior seemed to be present in the pH residuals reported by Douglas and Byrne (2017) for some lots of dye.

Douglas and Byrne (2017) reported a significant pH-dependence in the pH residuals after ${}_{434}A_{\text{imp}}$ corrections for some lots of impure dye, which they hypothesized could indicate a pH-dependence in the impurity absorption at 434 nm. We were not able to directly evaluate the possibility of a pH-dependence in the impurity absorption from the MCR analysis as we modeled the impurity absorption as a single component. Furthermore, a pH-dependence in the impurity absorption may not necessarily be resolvable as a separate component with MCR if the response is correlated with the principal component related to the pH variation in the set of spectra. The impurity absorption behavior will also vary with different lots of dye and thus requires studying multiple lots of dye from different commercial manufacturers.

The ${}_{434}A_{\text{imp}}$ corrections largely removed the pH-dependent pH offsets between the FB4 and Acros dye in the comparison samples (**Figure 5.7**), suggesting that there was not a strong pH-dependence in the impurity absorption for the Acros dye. Our simulations with synthetic dye spectra suggested that some of the residual pH-dependent offsets reported by Douglass and Byrne (2017) could be explained by the inconsistency resulting from the choice of values for the reference dye properties. For instance, if the reference dye had a small amount of impurity contamination (i.e., $c_{\text{Imp}} = 0.005$ in the simulations, giving an ${}_{434}A_{\text{imp}}$ value that approximately reflects the difference between the FB4 and Liu et al. dyes—see **Table 5.1** and **Table 5.2**), the use of its absorption coefficient ratios in the calculation of ${}_{434}A_{\text{imp}}$ and pH would result in an

apparent residual pH-dependent pH offset when comparing measurements between an impure dye and a truly pure dye (**Figure 5.6b**).

The ${}_{434}A_{\text{imp}}$-corrected pH values are essentially identical whether they are calculated using the pure dye properties or the contaminated "purified" dye properties, because the ${}_{434}A_{\text{imp}}$ correction adjusts the R values measured with the impure dye so that they are consistent with the selected set of absorption coefficient ratios of the reference purified dye. Thus, the offsets relative to the true pH values **Figure 5.6c** will be the same as those in **Figure 5.6a**, but if the ${}_{434}A_{\text{imp}}$-corrected values are compared to pure dye measurements, there will be an apparent offset that reflects the error in calculating the pure dye pH values with the wrong set of coefficients (i.e., $\text{pH}_{\text{pure}}' \neq \text{pH}_{\text{pure}}$) . These results imply that comparative measurements against a "purified" dye may not necessarily be a reliable indicator of the accuracy of the ${}_{434}A_{\text{imp}}$ correction if the user's dye and the reference dye were not identically pure.

5.5 Discussion

This study evaluated the performance of three different approaches which have been used to calibrate impure mCP for spectrophotometric measurements. The consistency of impure mCP pH measurements with purified mCP pH measurements depends critically on the accuracy of previously published properties of pure mCP. This study and others (DeGrandpre et al., 2014; Müller and Rehder, 2018) that have characterized independent lots of purified mCP found discrepancies with the most widely used published properties of Liu et al. (2011), which imply an inconsistency in pH on the order of ~0.005. The uncertainty in assigning pH values to the buffers used to determine $\text{p}K_{\text{a}}(\text{HI}^-)$ or $\text{p}(K_a(\text{HI}^-)\cdot e_2)$, about 0.004 in pH (Buck et al., 2002; Pratt, 2014), and potentially inadequate dye purification may explain the discrepancy,

but more investigation is needed to identify the source of the error. Therefore, although we demonstrated that we were able to obtain consistent pH measurements with the Acros and FB4 dyes when both were characterized on the same spectrophotometer, there may be an inconsistency with pure dye pH measurements as large as ~0.005 even when the impure dye was carefully characterized.

This uncertainty in the pure mCP properties, as well as the possibility of inadequately purified dyes (Takeshita et al., submitted), makes tenuous the goal of achieving consistency with pure dye measurements. The accuracy of corrections based on measuring the lot-specific pH offsets between an impure and "purified" dye will depend on the purity of the user's sample of purified dye as well as the choice of reference pure dye properties used to calculate pH. Some lots of inadequately purified dye have been found to have pH offsets relative to pure dye as large as ~0.008 for pH > 8 (Takeshita et al., submitted).

Similarly, ${}_{434}A_{\text{imp}}$ corrections (Douglas and Byrne, 2017) are also dependent on the choice of reference pure dye properties used in the calculations. Errors in the reference pure dye properties can affect the apparent consistency of the ${}_{434}A_{\text{imp}}$-corrected pH values with pure dye pH measurements, leading to a pH-dependent offset in the apparent pH values. However, the true error in the ${}_{434}A_{\text{imp}}$-corrected pH values is smaller (<0.003 in our simulations). Douglas and Byrne (2017) reported residual discrepancies between ${}_{434}A_{\text{imp}}$-corrected pH values and purified dye pH values as large as ~0.01 at pH > 8 for some lots of impure mCP. Although these offsets may be smaller due to the uncertainty in the purified mCP properties, they are likely still large. Our results therefore do not rule out the possibility of a pH-dependent impurity absorption behavior as suggested by Douglas and Byrne (2017), but this hypothesis is

better tested with more direct approaches such as HPLC characterization of impure dyes under different pH conditions.

Assumptions about the impurity absorption behavior are likely the largest source of uncertainty in the ${}_{434}A_{\text{imp}}$ corrections. Using the Acros dye impurity absorption spectrum resolved from MCR analysis, we showed that a small impurity absorption at 578 nm can result in meaningful error in ${}_{434}A_{\text{imp}}$ corrections at low pH (<0.003 in our simulations).

This study is the first to demonstrate the utility of using MCR to study sulfonephthalein indicator impurities. MCR can potentially be used with full spectrum methods (e.g., Ohline et al. 2007 and see **Chapter 3**) to calibrate impure dyes, but more research is needed to demonstrate the reliability of the approach. However, like the ratio-based ${}_{434}A_{\text{imp}}$ correction method, the approach we used to model the impurity absorption spectrum with MCR also assumes no pH-dependence in the impurity absorption behavior.

5.6 Conclusions and recommendations

As purified dyes are limited in availability and may be cost-prohibitive for some purposes, there is a need and desire to obtain high quality pH measurements with impure indicators. Achieving this requires the characterization of the behavior of absorbing impurities which affect the determination of pH. Users should consider the complexities and inherent uncertainties in the different dye calibration approaches when selecting the most appropriate method for their needs. We outline recommended practices for calibrating impure dyes with each approach.

For the most accurate calibration of impure mCP, we recommend either characterizing the properties of the dye or measuring the lot-specific pH offsets relative to a purified dye.

${}_{434}A_{\text{imp}}$ corrections may have a larger uncertainty than the other two impure dye calibration approaches, as it involves assumptions about the impurity absorption behavior that are not well-understood and likely vary with different lots of impure dyes. However, if the user's accuracy requirements are not stringent, the simplicity of the ${}_{434}A_{\text{imp}}$ correction method may justify its use.

Characterizing the absorption coefficient ratios and the pK_a(HI$^-$) or $\text{p}(K_a(\text{HI}^-)\cdot e_2)$ has the advantage of not requiring purified dye or relying on assumptions about its properties. It is critical that the absorption coefficient ratios and the apparent values of pK_a(HI$^-$) or $\text{p}(K_a(\text{HI}^-)\cdot e_2)$ are both determined, so that their errors, which are correlated, will cancel when calculating pH with **Eq. 2** or **Eq. 3**. A major disadvantage of this approach is that characterizing a dye involves a complex solution preparation process and may also require measurements in solutions over a range of temperatures and salinities if the dye is intended to be used over a wide range of those conditions. Furthermore, TRIS buffer pH standards are needed to determine pK_a(HI$^-$) or $\text{p}(K_a(\text{HI}^-)\cdot e_2)$.

Care should be taken to check the spectrophotometer lamp intensities prior to characterizing the dye, as moderate levels of degradation in the deuterium lamp intensity can result in absorbance errors that significantly affect the measurement of the I^{2-} spectrum (see **Chapter 3**). The lamps should be replaced if their intensities have degraded by more than 50% of their original values. These precautions also apply to the ${}_{434}A_{\text{imp}}$-based corrections, which involve measuring the I^{2-} spectrum of the impure dye.

pH offset correction curves for impure dyes should ideally be constructed using at least four samples evenly distributed across the pH range that will be encountered in application and under the same temperature and salinity conditions as in application.

Ultimately, more research is needed to evaluate the uncertainty of spectrophotometric pH measurements using impure indicators (as well as using purified indicators). There is currently insufficient understanding of the impurity absorption behavior of impure dyes to have high confidence in ${}_{434}A_{\text{imp}}$ corrections. The properties of purified mCP also need to be better constrained. The findings in this study and in others (Takeshita et al., submitted) highlight the need for better quality control in the production of purified mCP and for the re-evaluation of the properties of purified mCP.

5.7 Acknowledgements

Chapter 5, in part, is currently being prepared for submission for publication of the material. Fong, M.B., Branham, C.W., and Dickson, A.G. The dissertation author was the primary investigator and author of this paper.

Trisha Nguyen performed the preliminary experiments on characterizing *m*-cresol purple in NaOH solutions The authors would also like to thank Yui Takeshita and Regina Easley for helpful discussions and for sharing their findings from their studies on mCP impurities.

5.8 References

Bresnahan Jr., P.J., Martz, T.R., Takeshita, Y., Johnson, K.S. and LaShomb, M., 2014. Best practices for autonomous measurement of seawater pH with the Honeywell Durafet. Methods in Oceanography, 9: 44-60.

Buck, R.P., Rondinini, S., Covington, A.K., Baucke, F.G.K., Brett, C.M.A., Camoes, M.F., Milton, M.J.T., Mussini, T., Naumann, R., Pratt, K.W., Spitzer, P. and Wilson, G.S., 2002. Measurement of pH. Definition, standards, and procedures (IUPAC Recommendations 2002). Pure and Applied Chemistry, 74(11): 2169-2200.

Byrne, R.H. and Breland, J.A., 1989. High precision multiwavelength pH determinations in seawater using cresol red. Deep Sea Research Part A. Oceanographic Research Papers, 36(5): 803-810.

Byrne, R.H., Mecking, S., Feely, R.A. and Liu, X., 2010. Direct observations of basin-wide acidification of the North Pacific Ocean. Geophysical Research Letters, 37(2).

Carter, B.R., Radich, J.A., Doyle, H.L. and Dickson, A.G., 2013. An automated system for spectrophotometric seawater pH measurements. Limnology and Oceanography: Methods, 11(1): 16-27.

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

de Juan, A., Jaumot, J. and Tauler, R., 2014. Multivariate Curve Resolution (MCR). Solving the mixture analysis problem. Analytical Methods, 6(14): 4964-4976.

DeGrandpre, M.D., Spaulding, R.S., Newton, J.O., Jaqueth, E.J., Hamblock, S.E., Umansky, A.A. and Harris, K.E., 2014. Considerations for the measurement of spectrophotometric pH for ocean acidification and other studies. Limnology and Oceanography: Methods, 12(12): 830-839.

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

Dickson, A.G., 2010a. The carbon dioxide system in seawater: equilibrium chemistry and measurements. In: U. Riebesell, V.J. Fabry, L. Hansson and J.-P. Gattuso (Editors), Guide to best practices for ocean acidification research and data reporting. Publications Office of the European Union, Luxembourg, pp. 17-38.

Dickson, A.G., 2010b. Standards for Ocean Measurements. Oceanography, 23(3): 34-47.

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

Douglas, N.K. and Byrne, R.H., 2017. Achieving accurate spectrophotometric pH measurements using unpurified meta-cresol purple. Marine Chemistry, 190: 66-72.

Easley, R.A. and Byrne, R.H., 2012. Spectrophotometric Calibration of pH Electrodes in Seawater Using Purified *m*-Cresol Purple. Environmental Science & Technology, 46(9): 5018-5024.

Husheer, S.L.G., 2001. On spectrophotometric pH measurement in seawater media, University of Otago, Dunedin, New Zealand, 145 pp.

Jaumot, J., de Juan, A. and Tauler, R., 2015. MCR-ALS GUI 2.0: New features and applications. Chemometrics and Intelligent Laboratory Systems, 140: 1-12.

Johnson, K.S., Jannasch, H.W., Coletti, L.J., Elrod, V.A., Martz, T.R., Takeshita, Y., Carlson, R.J. and Connery, J.G., 2016. Deep-Sea DuraFET: A pressure tolerant pH sensor designed for global sensor networks. Analytical Chemistry, 88(6): 3249-3256.

Liu, X., Patsavas, M.C. and Byrne, R.H., 2011. Purification and Characterization of meta-Cresol Purple for Spectrophotometric Seawater pH Measurements. Environmental Science & Technology, 45(11): 4862-4868.

Loucaides, S., Rèrolle, V.M.C., Papadimitriou, S., Kennedy, H., Mowlem, M.C., Dickson, A.G., Gledhill, M. and Achterberg, E.P., 2017. Characterization of meta-Cresol Purple for spectrophotometric pH measurements in saline and hypersaline media at sub-zero temperatures. Scientific Reports, 7(1): 2481.

Martz, T.R., Connery, J.G. and Johnson, K.S., 2010. Testing the Honeywell Durafet® for seawater pH applications. Limnology and Oceanography: Methods, 8(5): 172-184.

Müller, J.D. and Rehder, G., 2018. Metrology of pH Measurements in Brackish Waters—Part 2: Experimental Characterization of Purified meta-Cresol Purple for Spectrophotometric pHT Measurements. Frontiers in Marine Science, 5(177).

Newton, J.A., Feely, R.A., Jewett, E.B., Williamson, P. and Mathis, J., 2014. Global Ocean Acidification Observing Network: Requirements and Governance Plan.

Ohline, S.M., Reid, M.R., Husheer, S.L.G., Currie, K.I. and Hunter, K.A., 2007. Spectrophotometric determination of pH in seawater off Taiaroa Head, Otago, New Zealand: Full-spectrum modelling and prediction of pCO_2 levels. Marine Chemistry, 107(2): 143-155.

Patsavas, M.C., Byrne, R.H. and Liu, X., 2013. Purification of meta-cresol purple and cresol red by flash chromatography: Procedures for ensuring accurate spectrophotometric seawater pH measurements. Marine Chemistry, 150: 19-24.

Pratt, K.W., 2014. Measurement of pHT values of Tris buffers in artificial seawater at varying mole ratios of Tris:Tris·HCl. Marine Chemistry, 162: 89-95.

Shimada, T. and Hasegawa, T., 2017. Determination of equilibrium structures of bromothymol blue revealed by using quantum chemistry with an aid of multivariate analysis of electronic absorption spectra. Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 185: 104-110.

Shimada, T., Tochinai, K. and Hasegawa, T., 2019. Determination of pH Dependent Structures of Thymol Blue Revealed by Cooperative Analytical Method of Quantum Chemistry and Multivariate Analysis of Electronic Absorption Spectra. Bulletin of the Chemical Society of Japan, 92(10): 1759-1766.

Takeshita, Y., Warren, J.K., Liu, X., Spaulding, R.S., Byrne, R.H., Carter, B.R., DeGrandpre, M.D., Murata, A. and Watanabe, S.-i., submitted. Consistency and stability of purified meta-cresol purple for spectrophotometric pH measurements in seawater. Marine Chemistry.

Talley, L.D., Feely, R.A., Sloyan, B.M., Wanninkhof, R., Baringer, M.O., Bullister, J.L., Carlson, C.A., Doney, S.C., Fine, R.A., Firing, E., Gruber, N., Hansell, D.A., Ishii, M., Johnson, G.C., Katsumata, K., Key, R.M., Kramp, M., Langdon, C., Macdonald, A.M., Mathis, J.T., McDonagh, E.L., Mecking, S., Millero, F.J., Mordy, C.W., Nakano, T., Sabine, C.L., Smethie, W.M., Swift, J.H., Tanhua, T., Thurnherr, A.M., Warner, M.J. and Zhang, J.-Z., 2016. Changes in Ocean Heat, Carbon Content, and Ventilation: A Review of the First Decade of GO-SHIP Global Repeat Hydrography. Annual Review of Marine Science, 8(1): 185-215.

Windig, W. and Guilment, J., 1991. Interactive self-modeling mixture analysis. Analytical Chemistry, 63(14): 1425-1432.

Yang, B., Patsavas, M.C., Byrne, R.H. and Ma, J., 2014. Seawater pH measurements in the field: A DIY photometer with 0.01 unit pH accuracy. Marine Chemistry, 160: 75-81.

Yao, W., Liu, X. and Byrne, R.H., 2007. Impurities in indicators used for spectrophotometric seawater pH measurements: Assessment and remedies. Marine Chemistry, 107(2): 167-172.

Zhang, H. and Byrne, R.H., 1996. Spectrophotometric pH measurements of surface seawater at in-situ conditions: absorbance and protonation behavior of thymol blue. Marine Chemistry, 52(1): 17-25.