AN EVALUATION OF THE PERFORMANCE OF AN ISFET pH SENSOR IN A DYNAMIC ESTUARINE SYSTEM

by

Stephen Fisher Gonski

A thesis submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Master of Science in Marine Studies

Fall 2016

AN EVALUATION OF THE PERFORMANCE OF AN ISFET

pH SENSOR IN A DYNAMIC ESTUARINE SYSTEM

by

Stephen Fisher Gonski

Approved: _____

Wei-Jun Cai, Ph.D.
Professor in charge of thesis on behalf of the Advisory Committee

Approved: _____

Mark A. Moline, Ph.D.
Director of the School of Marine Science and Policy

Approved: _____

Mohsen Badiey, Ph.D.
Acting Dean of the College of Earth, Ocean, and Environment

Approved: _____

Ann L. Ardis, Ph.D.
Senior Vice Provost for Graduate and Professional Education

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to my advisor Dr. Wei-Jun Cai for his guidance and assistance throughout the entirety of my project. I am extremely grateful for everything he has taught me and the opportunities he has given me.

I would also like to express my sincere gratitude to Dr. Karl Booksh and Dr. William Ullman for serving on my committee. Both have supported me throughout both my undergraduate and graduate studies at the University of Delaware. Dr. Booksh's technical insights into the operational theory of chemical sensors was invaluable to my project. Dr. Ullman's extensive working knowledge of the Murderkill Estuary-Delaware Bay system was indispensable, and his infectious enthusiasm for Oceanography and scientific discovery has passed on to me over the entire time I have known him. His mentorship has truly shaped my career.

I would to like to thank Dr. Wei-Jen Huang, Dr. Chris Main, Dr. Tye Pettay, Dr. Najid Hussain, Dr. Baoshan Chen, and Dr. Janet Reimer for all their assistance during the various phases of my project.

I would like to thank Dr. Todd Martz, Dr. Philip Bresnahan Jr., Dr. Yui Takeshita, and Taylor Wirth of the Scripps Institute of Oceanography for all their help with the various technical aspects of my project.

I would to thank my Aunt Martha & Uncle Mark O'Neill, for all the encouragement, rent-free housing, and the help they gave me when it came to building things for the duration of my Master's work at the University of Delaware.

I am extremely grateful to my lab mates for all their help throughout the course of my project, and I would like to thank my friends and family for all their support.

Finally, I would like to thank Dr. Chris Winn and Dr. Sam Khang who first introduced me to seawater pH and set me on this course in 2013 when I worked with them at Hawai'i Pacific University prior to transferring back to UD.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The feasibility of the Honeywell Durafet to the measurement of pH in a dynamic, productive, high-fouling, highly-turbid estuarine environment was investigated. Three different configurations of the SeapHOx sensor equipped with a Honeywell Durafet and its integrated internal (Ag/AgCl reference electrode containing a 4.5M KCl gel liquid junction) and external (solid-state chloride ion selective electrode, Cl-ISE) reference electrodes were deployed between April 2015 to August 2015 and September 2015 to August 2016 in the Murderkill Estuary-Delaware Bay System (Delaware, USA). Comprehensive sensor maintenance was performed every 1-2 weeks to ensure proper continuous sensor operation. Discrete bottle samples for dissolved inorganic carbon (DIC), total alkalinity (TA), and pH_{NBS} were collected every 30 minutes to calibrate raw sensor output on more than 20 tidal cycle sampling trips. During these sampling trips, the full range of biogeochemical exchange between the fresher Murderkill Estuary outflow and the more saline Delaware Bay water endmembers was captured. The sensor pH collected during the summer of 2016 using the furthest refined SeapHOx configuration exhibited good agreement with the independent reference pH used. Accordingly, the sensor pH had a root-mean squared error (RMSE) ranging between 0.011 and 0.036 pH units across the full salinity range of the deployment environment relative to both pH_T calculated from measured DIC and TA and pH_{NBS} measured with a glass electrode corrected to pH_T at *in-situ* conditions.

In this environment, the Honeywell Durafet proved capable of making high-resolution, high-frequency pH measurements ranging between 6.8 and 8.4. Natural pH fluctuations ranging from <0.5 pH units to >1 pH unit were routinely captured. A number of deficiencies in existing deployment guidelines and calibration protocol for Durafet-based biogeochemical sensors specific to estuarine and coastal ocean systems were identified and highlighted. Further, aspects of electrode response requiring further investigation were highlighted. A set of recommendations for the future utilization of these sensors in estuarine and coastal systems also resulted from the present work. The present work effectively demonstrated the viability of the Honeywell Durafet to the accurate and dependable measurement of pH as a part of future estuarine and coastal ocean CO_2 chemistry studies. The present work was a vital step in linking the parallel emerging trends in seawater pH metrology associated with the spectrophotometric measurement of pH using purified colorimetric indicator dyes and the electrochemical measurement of pH using the Honeywell Durafet over the full temperature and salinity range of natural waters. When these trends finally converge, the excellent accuracy of the Honeywell Durafet characterized under open-ocean conditions (<0.01 pH units) should be consistently achievable in estuarine and coastal ocean systems as well.

Chapter 1

INTRODUCTION

The rapid rise in atmospheric carbon dioxide released by human activities since the onset of the industrial revolution and its potential effects on the oceans has been of great concern in recent years (Doney *et al.*, 2009a; Doney *et al.*, 2009b; Kleypas *et al.*, 2006). As atmospheric carbon dioxide concentrations increase, up to one-third of the released CO_2 diffuses into the oceans, thereby altering the chemistry of the seawater through a process called ocean acidification (Caldeira and Wickett, 2003). The effects of this acidification are manifested in the subsequent decline in pH of not only surface seawater in the open ocean, estimated to be about 0.1 pH units to date (Raven *et al.*, 2005), but in natural waters of systems that link the terrestrial and marine biospheres such as estuaries and the coastal ocean as well (Duarte *et al.*, 2013). The importance of pH as a universal environmental indicator is beyond doubt. The parameter not only provides information on the equilibrium processes within the marine CO_2 system and other acid-base systems, but can also govern mineral saturation states, trace metal speciation, nutrient fluxes, nutrient bioavailability, and intra-cellular functions such as calcification, photosynthesis, and respiration in marine organisms. Because pH exerts such great control over a plethora of biological, biogeochemical, and physical properties, pH is often referred to as the "master" variable for these processes in aquatic systems (Bates, 1982; Byrne *et al.*, 1988; Clayton *et al.*, 1995; Orr *et al.*, 2005). Accordingly, pH and its effects on marine,

estuarine, and limnological systems continue to be a highly contentious topic throughout the oceanographic community.

Sørensen first described pH on a concentration scale in 1909 using the following equation:

$$\mathrm{pH} = -\log_{10}([\mathrm{H}^+]) \quad (1)$$

where $[\mathrm{H}^+]$ is the concentration of hydrogen ions in solution. This simplistic definition of pH ignores the underlying complexity attributed to this parameter (Zeebe and Wolf-Gladrow, 2001) Departing from the initial convention, pH can be better expressed in terms of the activity (or "effective" concentration) of hydrogen ions contained in solution, thus producing the Effective Hydrogen Ion Activity Scale $(\mathrm{pa_H})$ scale as described below:

$$\mathrm{pa_H} = -\log_{10}([\mathrm{a_H}]) = -\log_{10}(\gamma_{\mathrm{H}}\mathrm{m_H}) \quad (2)$$

where $\mathrm{a_H}$ is the hydrogen ion activity and γ_{H} is the ion activity coefficient of the molal concentration of hydrogen ions $(\mathrm{m_H})$ in solution (Sørensen and Linderstrøm-Lang, 1924). Following the early work of Sørensen (1909; 1924), the $\mathrm{pa_H}$ scale was eventually refined into the NBS Scale in 1965. This scale is based on a series of standard freshwater buffer solutions with assigned pH values closely resembling the pH values measured on the pa_H scale. This convention was employed as a guide for initial calibration of the NBS buffers (Zeebe and Wolf-Gladrow, 2001).

The reference state of the NBS Scale is pure water with an ionic strength of about 0.1 mol kg^{-1} and seawater has a higher ionic strength of about 0.7 mol kg^{-1}. Consequently, many researchers observed problems with the use of the NBS Scale for

various applications, including Bates (1975), Hansson (1973a; 1973b), Johnson *et al.* (1977), and Sillen (1968). Moreover, the fundamental problem with the NBS Scale is its relationship to the pa_H scale (Hawley and Pytkowicz, 1973; Pytkowicz *et al.*, 1974). Though the NBS Scale is not normally used for measuring seawater pH, it is still employed in freshwater (Zeebe and Wolf-Gladrow, 2001) and brackish water systems (Joesoef *et al.*, 2015).

Currently, seawater pH is defined in terms of a hydrogen ion concentration and is measured on a concentration scale. The three concentration scales employed at present are the free hydrogen ion concentration scale ($\mathrm{pH_F}$), total hydrogen ion concentration scale ($\mathrm{pH_T}$), and seawater scale ($\mathrm{pH_{SWS}}$). These scales can be defined by the following conventions:

$$\mathrm{pH_F} = -\log_{10}([\mathrm{H^+}]_\mathrm{F}) \tag{3}$$

$$\mathrm{pH_T} = -\log_{10}([\mathrm{H^+}]_\mathrm{T}) = -\log_{10}([\mathrm{H^+}]_\mathrm{F} + [\mathrm{HSO_4^-}]) \tag{4}$$

$$\mathrm{pH_{SWS}} = -\log_{10}([\mathrm{H^+}]_\mathrm{SWS}) = -\log_{10}([\mathrm{H^+}]_\mathrm{F} + [\mathrm{HSO_4^-}] + [\mathrm{HF}]) \tag{5}$$

where $[\mathrm{H^+}]_\mathrm{i}$ is the concentration of hydrogen ions in solution on the scale of interest; $[\mathrm{HSO_4^-}]$ is the concentration of sulfate ions in solution; and $[\mathrm{HF}]$ is the concentration of hydrogen fluoride ions in solution (Zeebe and Wolf-Gladrow, 2001). Each concentration scale is progressively more inclusive of ($\mathrm{H^+}$) contributions from various species present in seawater.

In accordance with the operational definition of the NBS Scale, all three concentration scales remain traceable back to a series of standard buffers, for which artificial seawater (ASW) that resembles the composition and ionic strength of natural

seawater has been adopted as their reference state (Dickson, 1993a; Millero *et al.*, 1993). The ASW buffers normally utilized are 0.04 equimolal buffers with a high buffer capacity and good storage stability (Dickson, 1993a). The four main bases used to prepare ASW buffers are 2-aminopyridine (AMP) (Bates and Erickson, 1986), tetrahydro-1,4-isohexane (morpholinium) (Czerminski *et al.*, 1982), 2-amino-2-methyl-1,3-propanediol (BIS) (Bates and Calais, 1981), and 2-amino-2-hydroxymethyl-1,3-propanediol (TRIS) (DelValls and Dickson, 1998; Nemzer and Dickson, 2005; Ramette *et al.*, 1977). The pH values of these four buffers on the total scale (pH_T) at 25°C and 35 PSU are 6.770, 8.573, 8.810, and 8.094, respectively (DelValls and Dickson, 1998; Dickson, 1993a). Since the nominal pH for most natural seawater is between 7.4 – 8.2, the only calibrated seawater buffer with a pH that falls within that pH range is TRIS, which makes it the primary seawater buffer recommended when developing pH measurement methodologies and performing pH measurements of natural waters (Dickson *et al.*, 2007).

The first methodology developed for the determination of pH in natural waters was potentiometry. Potentiometry is a branch of electroanalytical chemistry, in which a potentiometric cell is used to measure the electrical potential of a solution between two electrodes. A potentiometric cell consists of a reference electrode that is designed to maintain a known constant electrical potential regardless of changes in pH or ionic activity of the solution, and a pH measurement electrode that develops a voltage response proportional to the $[H^+]$ in the solution of interest (Skoog *et al.*, 2006). When using a potentiometric cell, the sample pH, pH(X), is defined to be relative to the

standard pH of the buffer, $\mathrm{pH(S)}$, calibrated with the standard hydrogen electrode using the following equation:

$$\mathrm{pH(X)} = \mathrm{pH(S)} + \frac{\mathrm{E_S} - \mathrm{E_X}}{\mathrm{k}} \quad (6)$$

where $\mathrm{E_S}$ is the electrical potential of the standard buffer solution measured with the glass electrode and $\mathrm{E_X}$ is the electrical potential of the unknown sample (Dickson *et al.*, 2007). k is the Nernst slope and can be calculated via:

$$\mathrm{k} = \{\mathrm{RTln(10)}\}/\mathrm{F} \quad (7)$$

where R is the Gas Constant (8.3145 J mol^{-1} K^{-1}), T is the temperature in Kelvin, and F is the Faraday Constant (96,487 C mol^{-1}) (Bates and Guggenheim, 1960).

The second methodology developed for the performance of pH measurements of natural waters is spectrophotometry. A spectrophotometric analysis is defined as any method that uses light absorption, emission, reflection, or scattering to measure chemical concentrations. In spectrophotometric pH determinations, absorbance is related to chemical concentration through Beer's Law via the following equation:

$$\mathrm{A} = \mathrm{\varepsilon bc} \quad (8)$$

where ε is the molar absorptivity in units of $M^{-1}cm^{-1}$; b is the path length; and c is the concentration. Since ε and b are both constants, the absorbance of a compound is directly proportional to c (Harris, 2007). Likewise, this principle can be applied to the spectrophotometric pH determinations of natural waters through the addition of a pH-sensitive colorimetric sulfonepthalein indicator dye (Rérolle *et al.*, 2012). pH-sensitive indicator dyes are diprotic compounds comprised of two weak acid-base pairs,

H_2I/HI^- and HI^-/I^{-2}, that absorb light at different wavelengths per their specific color and molar absorptivity. For spectrophotometric pH determinations, the second deprotonation reaction is the one of interest (Clayton and Byrne, 1993). When using colorimetric indicator dyes, pH is calculated from the second indicator acid dissociation constant (K_2) and the measured ratio of the peak absorbances of the HI^- and I^{-2} species (Clayton and Byrne, 1993; King and Kester, 1989; Robert-Baldo *et al.*, 1985; Yao and Byrne, 1998; Zhang and Byrne, 1996). The peak absorbances of the acidic and basic forms of the indicator dye are measured at their wavelengths of maximum absorption. These specific wavelengths are inherent of the indicator dye and vary based on the indicator dye used (Rérolle *et al.*, 2012).

The utilization of these pH-sensitive indicator dyes as the vehicles for the determination of the pH of natural waters dates to the first quarter of the 20th Century (McClendon, 1917). The development and refinement of these methodologies has continued (Byrne, 1987; Byrne and Breland, 1989; Clayton and Byrne, 1993; King and Kester, 1989; Liu *et al.*, 2011). Of the numerous pH-sensitive indicator dyes available, the most preferred is meta-Cresol Purple (mCP) as it better characterizes the full pH range of natural waters (Seidel *et al.*, 2008). This colorimetric indicator dye is ideal for pH measurements over the pH range of 7.2 – 8.2 (Easley and Byrne, 2012) in both the surface and deeper depths of the seawater column (Clayton and Byrne, 1993). The HI^- species of mCP has an absorption maximum at 434 nm, while the I^{-2} species of mCP has its absorption maximum at 578 nm (Clayton and Byrne, 1993).

The work of Clayton and Byrne (1993) produced the first equation for the spectrophotometric determination of pH on the total scale (pH_T) using mCP valid over a temperature range of $293 \leq T/K \leq 303$ and a salinity range of $30 \leq S \leq 37$, which takes the form expressed below:

$$pH_T = \log K_2 + \log\left(\frac{R - e_1}{e_2 - R \cdot e_3}\right) \tag{9}$$

where K_2 is the second stoichiometric acid dissociation constant of mCP; e_i correspond to the molar absorptivity ratios; and R corresponds to $A_{578}^{\lambda}/A_{434}^{\lambda}$. In accordance with Beer's Law, the second half of the above equation is equal to $[I^{-2}]/[HI^{-}]$ (Clayton and Byrne, 1993). Recently, it has been shown that vendor-specific impurities have been present in the mCP used in the spectrophotometric pH determinations of natural waters over the years, which can yield errors up to 0.018 pH units (Yao *et al.*, 2007). To combat this, additional purification procedures for removing impurities from mCP via HPLC (Liu *et al.*, 2011) and flash chromatography (Patsavas *et al.*, 2013a) have been developed. The introduction of purified mCP was accompanied by the development a new, slightly modified equation for the spectrophotometric determination of natural waters on the total scale (pH_T) valid over $278.15 \leq T/K \leq 308.15$ and only $20 \leq S \leq 40$, as given below:

$$pH_T = \log(K_2^T e_2) + \log\left(\frac{R - e_1}{1 - R \cdot \frac{e_3}{e_2}}\right) \tag{10}$$

where each term continues to correspond to the previously defined parameter or quantity (Liu *et al.*, 2011).

A side-by-side comparison of the two prevailing pH measurement methodologies permits an assessment of their accuracy and precision. The accuracy of each measurement methodology can be traced back to the series of TRIS Buffers in ASW used in their development. Accordingly, the accuracy of the pH measurements would be limited by the accuracy of the buffers, which is about ±0.004 pH units (Dickson *et al.*, 2007). The spectrophotometric pH determinations using unpurified or purified indicator dyes consistently yield better precision (±0.0004 to 0.001 pH units) (Carter *et al.*, 2013; Clayton and Byrne, 1993; Hammer *et al.*, 2014; Liu *et al.*, 2011) than potentiometric pH measurement methodologies using various electrode couples (±0.001 to 0.003 pH units) (Millero *et al.*, 1993). In addition, a variety of operational considerations for both measurement methodologies generally favor the spectrophotometric method over the potentiometric method. However, the preference of one pH measurement technique over the other often differs between individual laboratories, research groups, and scientists, as well as the concentration scale used to report their pH measurements (Marion *et al.*, 2011).

A continuing problem with these two methodologies is that they are presently limited to $S > 20$ where the standards are well calibrated. In addition, the convention of defining pH in terms of an activity or a concentration below this salinity threshold continues to be disputed. The effects of this discord propagate through all the major aspects associated with the measurement and use of pH in natural waters of $S < 20$. The standardization needed to govern and inform research incorporating the measurement and treatment of pH over this salinity range continues to elude the

scientific community (Dickson *et al.*, 2016). Ultimately, a well-defined and traceable set of reference materials is needed to develop, calibrate, and test suitable pH measurement methodologies for natural waters of S < 20 (Dickson *et al.*, 2016; Millero, 1986).

It is only recently that a growing movement within the seawater pH metrology community seeking to develop a high-precision spectrophotometric pH measurement methodology using purified indicator dyes applicable over the full temperature and salinity range of natural waters has emerged. Spurred on by the findings of Yao *et al.* (2007), steps have been taken to incorporate purified indicator dyes in subsequently developed spectrophotometric pH measurement methodologies (Lai *et al.*, 2016; Liu *et al.*, 2011; Patsavas *et al.*, 2013a; 2013b; Soli *et al.*, 2013). Concurrently, parallel work contributing to the development of such measurement methodologies that would facilitate pH measurements under sub-zero temperatures (Papadimitrou *et al.*, 2016), at higher pressures (800-900 bar) (Hopkins *et al.*, 2000; Soli *et al.*, 2013), and below open-ocean salinities (French *et al.*, 2002; Gabriel *et al.*, 2005; Hammer *et al.*, 2014; Lai *et al.*, 2016; Mosley *et al.*, 2004; Yao and Byrne, 2001) has also been completed. The calibration of a new series of TRIS buffers prepared by varying the ratios of TRIS:TRIS-HCl added to them, using a Harned Cell, produced buffers of similar composition to natural waters of S < 20 that possess the required buffering capacity to perform the much needed method development for this task (Pratt, 2014).

Chapter 2

AN EVALUATION OF THE PERFORMANCE OF AN ISFET pH SENSOR IN A DYNAMIC ESTUARINE SYSTEM

2.1 Introduction

In recent years, greater consideration has been given to the characterization of pH in contemporary studies of the chemistry of natural waters. This type of work is central to investigations conducted over a full spectrum of temporal and spatial scales (Martz *et al.,* 2010). An integral part of this work has encompassed the development of *in-situ* autonomous biogeochemical sensor packages capable of operating at high frequencies in remote areas with minimal user interaction over extended periods of time (weeks to months), over which continuous high-quality data of all types may be collected (Bresnahan *et al.,* 2014; DeGrandpre *et al.*, 1999; Martz *et al.*, 2010). Measurements of this type help resolve the temporal and spatial patterns and trends that traditional discrete water sampling in many different environments is unable to capture (Ben-Yakov *et al.*, 1974). Presently, there are two major classes of chemical sensors that are being utilized to accomplish this goal when it comes to pH: Reagent-based Optical Chemical Sensors (ROCS) and Electrochemical Sensors (ECS) (Hulanicki *et al.*, 1991).

ROCS employ an analyte-selective reagent (usually a pH-sensitive colorimetric indicator dye) to convert the analyte signal into an optical signal proportional to analyte concentration, whereas ECS transform the effect of the electrochemical interaction between the analyte-electrode couple into a useful signal

(Hulanicki *et al.*, 1991). Each type of sensor has the potential to be adapted to autonomous measurements of pH. However, with the refinement of the spectrophotometric pH measurement methodologies using pH-sensitive colorimetric indicator dyes in the last 30 years (Byrne, 1987; Clayton and Byrne, 1993; Liu *et al.*, 2011; Robert-Baldo *et al.*, 1985; Zhang and Byrne, 1996;) and its wide-spread use over electrochemical pH measurement methodologies, the development of these sensing systems has been steered in favor of ROCS over ECS (Martz *et al.*, 2010). This trend has been amplified by the introduction of cost-effective and operationally stable optical components (such as waveguides and LED light sources) (DeGrandpre *et al.*, 1999; Rérolle *et al.*, 2013) and smaller, better-quality spectrophotometers capable of being integrated into various field-portable ROCS packages (Friis *et al.*, 2004). The extensive literature of pH sensors based on ROCS designs developed in the last 10-15 years utilizing various optical properties like absorbance (Martz *et al.*, 2003; Rérolle *et al.*, 2013; Seidel *et al.*, 2008) and fluorescence using immobilized dye spots (Clarke *et al.* 2015; Larsen *et al.*, 2011) supports this view.

Despite the prevailing trends exhibited in recent autonomous pH sensor development, ROCS incorporating pH-sensitive colorimetric indicator dyes are not without their limitations. Such sensors are composite systems that incorporate intricate arrangements of sensor components consisting of both bulkier hardware like pumps and valves and more delicate optical media like LEDs and wave guides that must properly work in sequence over a prolonged measurement cycle to achieve the desired result. As such, some types of ROCS are known for their complicated operation and

slow instrumental response times (Martz *et al.*, 2010; Rérolle *et al.* 2016). Moreover, their elevated power requirements (Rérolle *et al.*, 2016) and sometimes component-specific operational temperature dependence (Bagshaw *et al.*, 2016) makes them undesirable for long-term deployments (e.g. > 6 months) in some systems (Rérolle *et al.*, 2016). Even more, the sensitivity of some ROCS to freeze-thaw cycles would limit their use in polar environments and at near-zero to sub-zero temperatures experienced *in-situ* in range of different deployment environments (Bagshaw *et al.*, 2016). As an added caveat to ROCS operation, they are routinely subject to optical interferences if the waters they sample are not filtered from a variety of sources that can instrumental slow response time and compromise optical measurements used to calculate pH (Liebsch *et al.*, 2001). Ultimately, this could severely limit their use in highly turbid waters commonly found in estuarine and coastal marine ecosystems.

Autonomous pH sensors based on ROCS designs are reputed to produce calibration-free measurements of pH given that they are only dependent on the optical properties inherent to the indicator dye used, thus resulting in repeatable, drift-free continuous measurements (DeGrandpre *et al.*, 1999; Martz *et al.*, 2003). However, since measurements consume indicator, this places a simple constraint on ROCS operation (Rérolle *et al.*, 2016). Indicator dyes are also known to be less stable in solution form and subject to temporal degradation, as demonstrated in measurements of pH using their benchtop counterparts (Dickson *et al.*, 2007). Indeed, this caveat of ROCS operation would result in erroneous, unrepresentative measurements, which only further adds to their complexity and uncertain reliability. Additionally, ROCS

operation is limited in the range of environmental temperature and salinity manifested in the constraints of the thermodynamic constants of the indicator dye (e.g. constants of purified mCP outlined in Liu *et al.* (2011)) used to calculate pH (Bagshaw *et al.*, 2016; Rérolle *et al.*, 2016). Likewise, indicator dyes used in ROCS pH measurements may have to be purified as well (Bagshaw *et al.*, 2016; Rérolle *et al.*, 2016) due to well-known pH errors attributed to dye impurities (e.g. mCP - Yao *et al.*, 2007). A viable, robust, simple, and straightforward alternative to pH sensors based on ROCS designs was found in the form of a pH-sensitive Ion-Selective Field Effect Transistor (ISFET) (Martz *et al.*, 2010). For a more detailed discussion of ISFET operational theory, see Section 2.2.7.1.

The Honeywell Durafet is the current preeminent ISFET pH sensor used in open-ocean seawater CO_2 chemistry and ocean acidification studies (Bresnahan *et al.*, 2014; Martz *et al.*, 2010). The Durafet is well adept to this purpose given its quick response time and low noise effects due to its low impedance (Martz *et al.*, 2010), its stable and consistent linear response with temperature (Takeshita *et al.*, 2014), and its well-characterized signal stability at 1 atm (Sandifer and Voycheck, 1999). The Durafet is now paired with two quality reference electrodes, an Ag/AgCl reference electrode containing a 4.5M KCl gel liquid junction (internal) and a solid-state chloride ion selective electrode (Cl-ISE) (external) (Bresnahan *et al.*, 2014; Martz *et al.*, 2010). These characteristics justify the operation of the Honeywell Durafet on single calibration point over weeks to months without periodic recalibration. Operating under these circumstances, the Honeywell Durafet has repeatedly

demonstrated a short-term precision (hours) and long-term precision (weeks to months) under controlled laboratory conditions of ±0.0005 pH units and ±0.005 pH units, respectively (Martz *et al.*, 2010). Also, it routinely exhibited the desirable long-term precision of better than ±0.01 pH units under *in-situ* open-ocean conditions (Bresnahan *et al.*, 2014). The published precisions of pH measurements made with the Honeywell Durafet at open-ocean salinities (S=30-36) brackets the precision needed to perform general water quality work (±0.1 pH units) (Millero, 1986; Yao and Byrne, 2001), general estuarine chemistry work (±0.02 pH units) (Head, 1985), long-term ocean acidification monitoring (±0.001-0.003 pH units) (Bates, 2007; Byrne *et al.*, 2010; Dore *et al.*, 2009; González-Davila *et al.*, 2007; Midorikawa *et al.*, 2010; Newton *et al.*, 2015) and to constrain the marine CO_2 system in seawater (±0.001 pH units) (Yao and Byrne, 2001). When combined with the ease of use associated with its sensor packages and straightforward calibration approach (Bresnahan *et al.*, 2014), the worldwide use of the Honeywell Durafet has grown very quickly (Frieder *et al.*, 2012; Martz *et al.*, 2014; Matson *et al.*, 2011; Price *et al.*, 2012).

The Honeywell Durafet and its dual-reference electrode configuration were originally integrated into two autonomous biogeochemical sensors known as the SeaFET and SeapHOx. The SeaFET is the passively flushed sensor that lacks an active pumping mechanism and is limited to the collection of pH data. The SeapHOx is the actively flushed sensor equipped with a Sea-Bird Electronics 5M submersible pump capable of making measurements of temperature, salinity (Sea-Bird Electronics Conductivity-Temperature Sensor – SBE37), pH, and oxygen (Aanderaa Data

Instruments 4835 Optode). All of the SeapHOx components deploy sequentially in one flow stream (Bresnahan *et al.*, 2014). Since the introduction of these sensor packages, they have been customized and adopted to the measurement of pH over the full environmental range of temperature and salinity characteristic of natural waters using a variety of different platforms.

Besides the use of the Honeywell Durafet under open-ocean conditions, it's use for the measurement of pH at high pressures (Johnson *et al.*, 2016), at near-zero temperatures (Bagshaw *et al.*, 2016; Matson *et al.*, 2011; Rérolle *et al.*, 2016), and at low ionic strengths (Bagshaw in prep.; cited in Bagshaw *et al.*, 2016) has been demonstrated or discussed. The Honeywell Durafet can also be used underway on hydrographic cruises (Rérolle *et al.*, 2016), for profiling down to 3000 meters using the Deep Sea Durafet (Johnson *et al.*, 2016), and on mobile oceanographic monitoring platforms such as the WavepHOx (Bresnahan *et al.*, 2016) and Argo Floats (Johnson *et al.*, 2016). The versatility of the Honeywell Durafet suggests that it will become the sensor of choice in the future to complement the application of a spectrophotometric pH measurement methodology for dilute estuarine waters. Hence, there is a clear need to ascertain the feasibility of Durafet use in natural waters of $S < 20$ characteristic of estuarine and coastal ocean systems.

2.2 Materials & Methods

2.2.1 Study Site

The Murderkill Estuary is a well-mixed, turbid, and nutrient-replete tributary (Voynova *et al.*, 2015) of the significantly larger Delaware Bay. The Delaware Bay is 215 km long, 18 km wide at its mouth up to about 44 km at widest point further upstream (Wong *et al.*, 2009). It drains a 36,570 km^2 watershed that includes large parts of Delaware, New Jersey, Pennsylvania, and the state of New York (Sharp *et al.*, 2009). The Murderkill Estuary discharges into the Delaware Bay at Bowers, DE approximately 39 km upstream of the Bay mouth (Voynova *et al.*, 2015). This small tributary estuary covers about 20 km between its discharge point at Bowers, DE and its upstream point at Frederica, DE. It has an average channel width of 50 meters and average depth of 4.5 meters (Wong *et al.*, 2009). The two dominant system endmembers in the Murderkill Estuary-Delaware Bay System are the fresher Murderkill Estuary outflow and the more saline Delaware Bay water. The Murderkill River and Estuary and the Delaware Bay are both turbid due to elevated suspended sediment concentrations attributed to the numerous sources of suspended sediments in their watersheds, their tidal currents, and the erosion of ancient, muddy sedimentary deposits (deWitt & Daiber, 1974; Ullman *et al.*, 2013; Voynova *et al.*, 2015).

The Murderkill Watershed is composed of 275 km^2, of which 55% is agricultural/farmland, 11% is covered by forests, 14% is urban, and 17% is covered by wetlands (comprised of 36% tidal wetlands and 60% freshwater wetlands) (DNREC, 2006). Moreover, the Kent County Regional Wastewater Treatment Facility

(KCRWTF) which serves about 130,000 total customers drains into the Murderkill Estuary around 10.5 km upstream from its mouth. This subjects the estuary to greater population pressures than its predominantly agricultural watershed would suggest (KCBPW, 2012). It is important to note that the inputs from the watershed's wetlands and marshes and inputs from the wastewater treatment plant comprise minor third and fourth endmembers in the Murderkill Estuary-Delaware System (Dr. William Ullman, personal communication). Summed up, these sources contribute to high nutrient concentrations in the Murderkill Estuary that can fuel episodic phytoplankton blooms that can persist over periods of hours, days, or longer (Voynova *et al.*, 2015), which provide ideal food sources for zooplankton and other estuarine heterotrophs (Cloern, 1996).

2.2.2 Deployment Design

Over the course of the ten-month deployment period with SeapHOx SP053 from September 2015 to August 2016 and the preceding 4.5-month deployment with SeapHOx SP033, three different sensor configurations were used (Figure 1). Each new subsequent configuration incorporated modifications to sensor design meant to improve its applicability to the measurement of pH in an estuarine system. SeapHOx Configuration v1.0 (SeapHOx SP033) (Figure 1, Left Panel) was simply the sensor as designed for open-ocean deployments following its assembly and testing at the Scripps Institute of Oceanography. This unit was deployed from 08 April 2015 to 26 August 2015 in the Murderkill Estuary-Delaware Bay System, but was removed in late August

2015 after sensor failure was detected. SeapHOx Configuration v2.0 (SeapHOx SP053) (Figure 1, Middle Panel) was equipped with a new flow path consisting of rounded tubing connections of a standardized cross-sectional area or inner diameter (R-3400 Tubing, 1/2” I.D. x 3/4” O.D., McMaster-Carr Part No. 5255K26) between the SB37, SBE 5M submersible pump, and flow housing that lacked all but one of the right-angle elbow connectors found on the original sensor flow path. The design of the SeapHOx SP053 flow path was further refined into SeapHOx Configuration v3.0 (Figure 1, Right Panel) by abandoning the remaining right-angle elbow connector at the outflow of the SBE37 conductivity cell and integrating a rounded tubing connection (R-3400 Tubing, 3/8” I.D. x 5/8” O.D, McMaster-Carr Part No. 5255K24) via a tight-seal barbed tube reduction fitting (1/2” I.D. x 3/4" O.D., McMaster-Carr Part No. 5463K639) to the 1/2” I.D. tubing on the remainder of the flow path. It must be noted that thicker tubing better retained its shape when rounded tubing connections were fashioned from it, which minimized constrictions in the new flow path. Because of the modifications, the flow path was considerably longer. Accordingly, a much longer flushing time was used for SeapHOx Configurations v2.0 and v3.0 (e.g. 60 seconds for 09 May 2016 to 09 June 2016 deployment and 70 seconds for 20 July 2016 and 24 August 2016 deployment) than what is recommended for open-ocean deployments (25 seconds). The increased flushing time substantially reduced the build-up of sediment and its associated fouling of the flow path of the instruments. Finally, a sampling interval of 30 minutes was chosen for the present work.

2.2.3 Sensor Preparation

Prior to each sensor redeployment in the present work, sufficient steps were taken to verify proper sensor operation and ensure the maintenance of such operation over periods of time between scheduled sensor maintenance trips. The necessary tasks conform with the Best Practices outlined in Bresnahan *et al.* (2014) and those detailed in a similar section of Rivest *et al.* (2016). Briefly, following sensor removal, the sensor was completely disassembled to facilitate comprehensive interior and exterior sensor cleaning procedures and the removal of biofouling organisms of all types. Trends in power consumption over the course of the previous deployment manifested in recorded battery voltages were also investigated and the old battery pack was replaced if warranted. Upon sensor reassembly, all O-rings were replaced (if needed) and regreased, all cable connectors were regreased, and any corroded or fouled screws found anywhere on the exterior of sensor were replaced.

Following sensor reassembly, the SBE 5M submersible pump was removed, submerged, and tested to ensure proper operation. Optode operation was also tested in air pending the measurement of ~270 µM at ~100% oxygen saturation to verify performance against the original optode calibration as detailed in Bresnahan *et al.* (2014). Using filtered seawater (salinity ~ 30-31), tests were conducted on the response of both electrodes to ensure measurements returned to within nominal voltage ranges for each reference electrode described in Bresnahan *et al.* (2014). Once all the tests of all sensor components were completed, the sensor body was wrapped with 2” wide white EZ Tear Construction Tape (Micronova Mfg., Inc., Part No. EZT-

2WH) followed by overlapping layers of copper tape. Copper tape was employed as the primary antifouling measure for the present work consistent with the recommendations of Bresnahan *et al.* (2014), while the construction tape greatly simplified the process of removing it after subjection to multiple months of *in-situ* estuarine conditions. The SeapHOx was also fitted with a replacement U-shaped copper pipe inlet at that time as well. Then, using the "Test Deployment" mode available on SeapHOx units of a build similar to SP053, electrodes were conditioned in filtered seawater taken from the lower Delaware Bay while the sensor was continuously powered and the pump turned off for the periods of time between sensor deployment periods. Electrodes remained stored in seawater during transport to the deployment site as well. For a more detailed evaluation of the applicability of the recommended open-ocean electrode conditioning protocol for a sensor deployments in an estuarine system, see Section 2.4.1.

2.2.4 Field Deployment

For the duration of the sensor deployments, the SeapHOx units were deployed alongside the suite of co-located sensors comprising the Kent County Land-Ocean Biogeochemical Observatory (LOBO) (http://kentcounty.loboviz.com/) (Table 1) except during December 2015-March 2016 when those sensors were undergoing annual maintenance and conditioning. The LOBO deployment platform consisted of an aluminum cage with removable struts mounted on a trolley made of 1" thick galvanized steel raised and lowered by means of a marine winch onto an I-Beam

situated on a pier adjacent to the mouth of the estuary that terminated with a concrete slab set 2.5 feet above the estuary floor. SeapHOx SP033 (SeapHOx Configuration v1.0) (Figure 1, Left Panel) was deployed upright. Thus, sensor performance suffered greatly from the effects of sediment accumulated over the electrode surfaces (Figure 2, Top Left/Top Right Panels) as indicated in the sustained pH decrease captured by both reference electrodes (Figure 3). Over the course of this deployment, various measures to combat sediment accumulation inside the flow housing were employed (Figure 2, Bottom Left/Bottom Right Panels) with only limited success. Ultimately, sediment was still able to accumulate over periods of variable turbidity conditions between sensor maintenance trips as it settled out of suspension from the waters left inside the flow housing after each sampling cycle once pumping ceased.

When deploying subsequent SeapHOx configurations with SeapHOx SP053, the sensor was attached to the LOBO cage using a pair of aluminum brackets joined to the white SeapHOx brackets (Figure 4). A stainless-steel hose clamp (9-3/8” x 12-1/4” clamp ID, McMaster-Carr Part No. 5011T44) was also wrapped around the struts and attached sensor with pieces of gasket rubber separating the aluminum from the stainless steel as a fail-safe stabilization measure. Prior to deployment, the aluminum brackets and all aluminum LOBO components were coated with black Interlux Pacifica Plus antifouling paint (Interlux Paint, LLC., Union, NJ, USA). To prevent a repeat of the SeapHOx SP033 failure, SeapHOx SP053 was deployed upside down so any sediment would accumulate away from the electrodes inside the flow housing. For future sensor deployments in turbid environments, an upside down or horizontal

orientation that minimizes sediment effects is recommended. Likewise, either orientation is also more neutral with respect to the strain placed on the pump associated with pumping against gravity. The flow path was plumbed into the original outflow point of the flow housing to ensure it filled bottom-to-top in its new orientation keeping the electrodes and optode always immersed. In addition, the copper pipe inserts integrated as a secondary antifouling measure into the flow path of SeapHOx Configuration v1.0 were replaced by a U-shaped copper pipe inlet (Mueller Industries, MCTP-W Type ACR Refrigeration/AC Copper Pipe, 1/2" O.D.) bent to a 180^{o} angle. The U-shaped inlet was placed at the sensor inflow point to prevent the passive settling of sediment and/or active invasion of fouling organisms into the sensor flow path. Because multiple metals were used for various parts of the LOBO assembly, sufficient anti-corrosion measures were employed.

2.2.5 Sampling Approach

For a sensor deployment in a dynamic, productive, high-fouling, highly-turbid estuarine system characterized by high turbidity and large salinity changes over relatively short-time scales (Ullman *et al.*, 2013; Voynova *et al.*, 2015), the discrete sampling regiment adopted to calibrate the sensor should not be based on the static-point approach practiced by operators of these sensors in open-ocean environments. Instead, the prevailing biogeochemical controls should be integrated into the discrete sampling regiment and discrete bottle samples should be collected over a period long enough to capture the full fluctuations in the parameters of interest in the deployment

environment. If this dynamic-point approach is not practiced, sensor operators risk introducing unintentional bias into any subsequent sensor calibrations.

The Murderkill Estuary-Delaware Bay System is essentially a river mouth and flood-dominant system (Dzwonkowski *et al.*, 2013; Wong *et al.*, 2009) characterized by significant discharge asymmetry resulting in much slower ebb tides and more rapid flood tides (Dzwonkowski *et al.*, 2013). It is known that both high-frequency tidal fluctuations and low-frequency subtidal controls of spring-neap tides, winds, and large storm events drive local biogeochemistry (Dzwonkowski *et al.*, 2013; Voynova *et al.*, 2015; Wong *et al.*, 2009). Thus, these biogeochemical controls were considered when planning sampling trips, which allowed for tidal cycle sampling to be adopted where discrete bottle samples were collected over the course of 10-12 hour periods. For this reason, the full range of salinity and pH associated with the periods of time when one or another system endmember dominates and periods of mixing between the two endmembers interspersed between were captured. Hence, this reduced the chances of introducing bias into reference pH datasets used to the perform the sensor calibration in terms of salinity, pH, or system endmember. Accordingly, sampling trips were also synchronized with favorable conditions related to subtidal controls of local biogeochemistry such as winds and the temporal responses of both system endmembers to large storm events that represent significant freshwater inputs to the watershed (Voynova *et al.*, 2015). This allowed the sensor operator to maximize the biogeochemical range the dynamic-point calibration covers. To this end, the incorporation of local biogeochemical controls into dynamic-point sensor calibrations

performed on Durafet-based biogeochemical sensors must continue in the future in similar deployment environments when practical.

2.2.6 Sensor Maintenance

Sensor maintenance was continuously performed every 7-10 days from May 2016 to August 2016, following previously compiled SOPs designed to minimize the effects of sediment and fouling organisms on sensor performance between maintenance trips. The degree of fouling observed on the SeapHOx over a 1-2 week period in an estuarine system (Figure 5) was comparable to that of much longer open-ocean SeapHOx deployments. Fouling appeared to be considerably worse from March to October coinciding with the most productive times of year in the Murderkill Estuary-Delaware Bay System (Ullman *et al.*, 2013; Voynova *et al.*, 2015), whereas it was only minimal over the winter months from November to February. SeapHOx maintenance was performed concurrently with that of the co-located Kent County LOBO equipment. This reduced the chances of degraded sensor performance attributed to fouling organisms or sediment not directly in contact with the SeapHOx, but could still indirectly affect data quality through altering the chemistry of the waters directly surrounding the instruments sampled by the SeapHOx.

2.2.7 pH Calculation

The SeapHOx reports a pair of pH values based on the dual-reference electrode design of this sensor package. These pH values and their associated calibration

constants must be calculated from sensor voltages, *in-situ* temperature, and *in-situ* salinity measured over the course of each individual sampling cycle. In the subsequent section, a review of the ISFET operational principle and the derivations of the essential equations comprising the operation theory of each reference electrode is discussed.

2.2.7.1 ISFET Operational Theory

Conventional ISFET sensors are a subdivision of Chemical Sensitized Field Effect Transistors (CHEMFETs) similar to a Metal Oxide Semiconductor Field Effect Transistor (MOSFET) technology (Hulanicki *et al.*, 1991). However, an ISFET lacks the metal gate electrode over the conduction channel (Martz *et al.*, 2010). ISFETs generally operate by applying a constant drain source voltage and employing a feedback circuit to maintain it (Martz *et al.*, 2010) (Figure 6). The metal gate electrode was replaced with a thin insulating layer of amphoteric material capable of facilitating both surficial protonation and deprotonation reactions that covers the surface of the conduction channel located between the source and the drain (Martz *et al.*, 2010). The function of the metal gate electrode is satisfied by direct exposure of the insulating material to solution (Rérolle *et al.*, 2012). The pH of the solution is then determined by voltage measurements between a reference electrode and the sensing layer (Martz *et al.*, 2010) manifested in the change in the drain-source current usually held constant (Hulanicki *et al.*, 1991) as controlled by the site-binding protonation-deprotonation process occurring the solution-insulator interface (Martz *et al.*, 2010).

In mathematical terms, ISFET response can be calculated via the following equation (Janata, 2009):

$$\Delta V_G = \frac{2.303RT}{z_i F} \log\left(\frac{a_i}{\sum_j K_{i,j} a_j^{z_i/z_j}} + 1\right) \quad (11)$$

where ΔV_G is the ISFET response, R is the Gas Constant, T is the temperature in Kelvin, F is the Faraday Constant, z_i is the charge of the ion of interest, z_j is the charge of any interfering ions present, a_i is the activity of the ion of interest, a_j is the activity of any interfering ions present, and $K_{i,j}$ are the intrinsic dissociation constants of the ion of interest and any interfering ions present, respectively. Like other ion-selective sensors, the measured change in the drain-source current or ISFET response (ΔV_G) is proportional to the activity and concentration of the ion of interest in solution (Mikhelson, 2013) and hence, the pH of the solution (Janata, 2009). For a more comprehensive explanation on the possible effects of other interfering ions present in equation 11, see Section 2.2.9.

For the application of the proton-sensitive Honeywell Durafet to the measurement of seawater pH on a concentration scale using a conventional reference electrode that exhibits a Nernstian response to the free chloride ion, equation 11 can be rewritten into the general forms given below (Martz *et al.*, 2010):

$$E_{sensor} = E^*_{sensor} - \left(\frac{RT\ln(10)}{F} \times \log(a_H a_{Cl})\right) \quad (12)$$

$$E_{sensor} = E^*_{sensor} - \left(\frac{RT\ln(10)}{F} \times \log(\gamma_H \gamma_{Cl})\right) - \left(\frac{RT\ln(10)}{F} \times \log(m_H m_{Cl})\right) \quad (13)$$

where E_{sensor} is the measured voltage, E^{*}_{sensor} is the calibration constant, and a_i is the activity of either H^+ or Cl^-, γ_i is the ion activity coefficient of either H^+ or Cl^-, and m_i is the molal concentration of either H^+ or Cl^-. In addition, a nominal +1 gain added to the corresponding quantity is programmed into the operation of all sensors built around the Honeywell Durafet effectively bringing equations 12 and 13 in line with equation 11. Using the logarithm laws, equation 13 can be rearranged to isolate pH or the $-\log(m_H)$ as seen below:

$$E_{sensor} = E^{*}_{sensor} - \left(\frac{RT\ln(10)}{F} \times [\log(\gamma_H \gamma_{Cl} m_{Cl}) - pH]\right) \quad (14)$$

Temperature-dependent standard potentials for the reference electrode configuration incorporating the Honeywell Durafet are instead referred to as calibration constants (Bresnahan *et al.*, 2014) designated by an asterisk (*) (e.g. E^{*}_{sensor}) over a nought symbol (^{o}) (e.g. E^{o}_{sensor}) because of the inter-sensor variability observed in these quantities characterized from previous work (Martz *et al.*, 2010).

2.2.7.2 FET|INT and FET|EXT

As designed, the SeapHOx can make two separate pH measurements using its two integrated reference electrodes – FET|INT (Ag/AgCl) and FET|EXT (Cl-ISE) – designated by the notations of pH^{INT} and pH^{EXT}, respectively (Martz *et al.*, 2010). In the case of FET|INT, the necessary parameters denoted by 'INT' can be calculated via the following equations (Martz *et al.*, 2010):

$$E_{INT} = E_j + E^{*}_{INT} - S \times \log(m_H) \quad (15)$$

$$E^*_{INT} = \{E^*(FET|INT) - S\times \log[\gamma_H(sw)] - S\times \log[a_{Cl}(\text{ref gel})]\} \quad (16)$$

$$pH^{INT} = \frac{(E_{INT} - E^*_{INT})}{S} \quad (17)$$

where S is the Nernst slope ($S = \{RTln(10)\}/F$); E_j is the liquid junction potential; a_{Cl} is the chloride ion activity of the KCl reference gel in FET|INT; and γ_H and m_H represent the ion activity coefficient and molality of the hydrogen ion in solution. Since FET|INT can measure dissolved H^+ in solution and dissolved Cl^- in the reference gel, it always has a repeatable chloride ion activity (a_{Cl}). Because of this, $\log(\gamma_H\gamma_{Cl}m_{Cl})$ is incorporated into the operational definition of the voltage converted to pH via the temperature-dependent Nernst slope in equation 17 (Martz *et al.*, 2010).

In the case of FET|EXT, it does not demonstrate a repeatable a_{Cl} (Martz *et al.*, 2010). To a similar effect, a set of equations detailed in Martz *et al.*, (2010) can be used to calculate the key parameters as denoted by 'EXT' as well:

$$E_{EXT} = E^*(FET|INT) - S\times \log(\gamma_H\gamma_{Cl}) - S\times \log(m_{Cl}) - S\times \log(m_H) \quad (18)$$

$$E^*_{EXT} = E^*(FET|EXT) \quad (19)$$

$$pH^{EXT} = \frac{(E_{EXT} - E^*_{EXT}) + S\times \log(\gamma_H\gamma_{Cl}m_{Cl})}{S} \quad (20)$$

where S is the Nernst slope ($S = \{RTln(10)\}/F$) and γ_i and m_i represent activity coefficients and molality of either H^+ or Cl^-, respectively. Since FET|EXT measures dissolved hydrogen chloride (HCl) in solution, the mean activity coefficient of hydrogen chloride (HCl) ($\gamma_\pm(HCl)^2 = \gamma_H\gamma_{Cl}$) calculated from temperature and salinity (Khoo *et al.*, 1977) and the concentration of the free chloride ion in seawater

(m_{Cl}) calculated from salinity (Dickson *et al.*, 2007) are used to calculate pH^{EXT} (Bresnahan *et al.*, 2014; Martz *et al.*, 2010). Using the measured voltage, *in-situ* temperature, and *in-situ* salinity, pH^{INT} and pH^{EXT} are calculated assuming a 100% Nernst slope (e.g. 59.16 mV/pH at 25ºC) and a constant dE^*_{sensor}/dT (e.g. $dE^*_{INT}/dT = -0.001101$ V/℃ *and* $dE^*_{EXT}/dT = -0.001048$ V/℃) (Bresnahan *et al.*, 2014; Martz *et al.*, 2010).

2.2.8 Calibration

Conventional calibration protocol for electrochemical pH measurement methods consist of employing a series of traceable reference materials (e.g. TRIS Buffers in ASW) representative of the composition of natural waters at those salinities to calibrate and periodically recalibrate the electrodes or sensors to be used. When performing an electrochemical pH calibration using one of the three concentration scales for work in low and intermediate salinities, it has been recommended that a single standard reference buffer at an intermediate salinity between freshwater and seawater (e.g. $15 \leq S \leq 20$) or a series of standard buffers spanning the full salinity range encountered be used (Whitfield *et al.*, 1985). Yet, the problem with the former is that the errors associated with any measurements made greatly increase when deviating from the single calibration salinity (Whitfield *et al.*, 1985), while the problem with the latter rests in the difficulties in the preparation and preservation of each individual buffer (Easley and Byrne, 2012). In the case of the latter, it has been demonstrated that a 10-unit change in the salinity between the seawater buffer solution

and natural seawater sample yields errors of up to 0.028 pH units and even a difference of only 2 salinity units results in errors of 0.005 pH units (Easley and Byrne, 2012). Certainly, this highlights the need for finer salinity incrementation between successive buffers and only adds to the complexity of this method (Easley and Byrne, 2012; Whitfield *et al.* 1985).

The difficulties of standard buffer calibration methods for electrochemical pH determinations below S=20 were circumvented by employing the calibration method for Durafet-based biogeochemical sensors recommended by Bresnahan *et al.* (2014). This type of sensor calibration is dependent on the natural variability inherent to the deployment environment (Hofmann *et al.*, 2011; Kline *et al.*, 2012) rather than a comparison against standard reference buffers. Such an *in-situ* or field calibration is performed through the collection of *in-situ* pH data via a series of discrete samples coinciding with the time of sensor measurements measured using well-established benchtop methods and then calibrating a working Durafet to those field measurements (Bresnahan *et al.*, 2014).

Using this approach, the primary control on the quality of the sensor calibration and subsequent sensor time series is directly related to the amount and quality of the discrete samples collected over the course of a sensor deployment. An optimal discrete sampling regiment is characterized by >10 usable discrete bottle samples (Rivest *et al.*, 2016). On the other hand, calibrating a sensor time series to only one discrete bottle sample can produce sensor inaccuracies of ~0.1 pH units (Bresnahan *et al.*, 2016) and is strictly recommended against (Bresnahan *et al.*, 2014)

due to the limitations it places on the evaluation of sensor failure or effects of biofouling (Rivest *et al.*, 2016). A second control is related to the benchtop pH measurement method chosen to measure the pH of the discrete bottle samples. Past experiences dictate that the pH measurements of discrete bottle samples from open-ocean deployments be performed using spectrophotometric pH measurement methods using pH-sensitive colorimetric indicators known for their high precision and minimal operator interaction (Carter *et al.*, 2013). On the other hand, the application of spectrophotometric pH measurement methodologies is not straightforward in lower salinity waters as previously discussed in Section 2.1. Therefore, this limits the applicability of spectrophotometric pH measurement methods to sensor deployments in natural waters below S=20. Yet, other pH measurement methodologies exist and pH can be calculated from two marine CO_2 system parameters – dissolved inorganic carbon (DIC), partial pressure of CO_2 (pCO_2), and total alkalinity (TA) – so a sensor calibration in an estuarine environment is still achievable.

2.2.8.1 Field Measurements

Dissolved inorganic carbon (DIC), total alkalinity (TA), and pH_{NBS} were measured at the sensor deployment site located at the confluence of the Murderkill Estuary and Delaware Bay at Bowers Beach, DE on 01 June 2016 from 0900-1900 and 02 August 2016 from 0800-1930. A 1200-Watt gasoline-powered generator (Champion Power Equipment) was used to supply power to all equipment over the course of these sampling trips. Water column samples were continuously collected

every 30 minutes coinciding with sensor measurements (00:00 and 00:30) made at the same frequency using a peristaltic pump by way of a 12.5' length of tubing lowered into the water and sufficiently weighted to maintain a fixed position. The tubing was placed at the approximate location of the sensor relative to the pier its deployment structure was mounted on. Measurements of *in-situ* temperature, salinity, and pressure were taken from the SeapHOx.

2.2.8.2 Analytical Methods

Samples for DIC and TA were collected usually in duplicate following filtration through Whatman 0.45 µm Polyethersulfone (PES) filters (GE Healthcare Bio-Sciences, Pittsburgh, PA, USA) by bottom-filling into triple-rinsed 250-mL borosilicate bottles. The samples were then fixed with 100 µL of saturated mercuric chloride ($HgCl_2$) solution. Approximately every 30 minutes, filters were regenerated by rinsing the filter backwards then forwards three times using a reservoir of deionized (DI) water and a second peristaltic pump following individual sampling cycles. Upon returning from the field, the samples were preserved in 4^{o}C for future analysis (Cai and Wang, 1998; Jiang *et al.*, 2008). DIC was determined through acid extraction by quantifying the released CO_2 using an infrared gas analyzer (AS-C3 Apollo Scitech). TA was measured by Gran Titration (Gran, 1950; 1952) using a semi-automated open-cell titration system (Cai *et al.*, 2010; Huang *et al.*, 2012). All measurements were calibrated against certified reference materials (CRM, provided by A.G. Dickson from Scripps Institute of Oceanography).

Samples for pH_{NBS} were collected unfiltered always in duplicate via bottom-filling into triple-rinsed 125-mL clear glass bottles. pH_{NBS} was measured using an Orion Dual Star pH/ISE Benchtop Meter equipped with a new Orion 8302BNUMD Ross Ultra Glass Triode pH/ATC Combination Electrode (Thermo Fisher Scientific Inc., Beverly, MA, USA) within 3-5 minutes of sample collection at *in-situ* temperature. The pH electrode was calibrated every 2-3 hours using three National Bureau of Standards (NBS) traceable freshwater pH buffers of 4.01, 7.00, and 10.01 stored in 20 mL scintillation vials thermostatted in water taken from the Murderkill Estuary for at least 30 minutes prior to the calibration. During all pH_{NBS} measurements, parafilm was used to minimize CO_2-exchange between the sample and the surrounding atmosphere. Also, all pH measurements were performed in the shade under an umbrella to minimize electrode drift caused by exposure to direct sunlight. Fresh 20 mL aliquots of the pH buffers were also used for each sampling trip to minimize the errors in pH_{NBS} measurements potentially imparted by the pH buffers themselves.

2.2.8.3 Independent Reference pH

From the measurements of DIC, TA, and pH_{NBS}, two sets of independent reference pH emerged that were used to calibrate the raw sensor time series. The first independent reference pH was pH_T calculated from measured DIC and TA at *in-situ* temperature, salinity, and pressure (pH^{disc}_{DIC-TA}) using the inorganic carbon dissociation

constants from Millero *et al.* (2006), the bisulfate ion acidity constant of Dickson (1990), and the boron-to-chlorinity ratio of Lee *et al.*, (2010) in the Excel macro CO2SYS (Pierrot *et al.*, 2006). These constants were used for all other CO2SYS calculations performed. Calculations of $\mathrm{pH_T}$ using the inorganic carbon dissociation constants from Mehrbach *et al.*, 1973 refit by Dickson and Millero (1987) were also performed. However, preliminary results of the present work found that the constants from Millero *et al.* (2006) were more suited to calculations of $\mathrm{pH_T}$ in estuarine waters ($S < 30$).

The second independent reference pH was $\mathrm{pH_T}$ calculated from the field measurements of $\mathrm{pH_{NBS}}$ made on discrete bottle samples with a pH electrode. To use these pH measurements, they were first corrected to the SeapHOx temperature by applying the difference between values of $\mathrm{pH_{NBS}}$ calculated from measured DIC and TA at the measurement temperature and at the SeapHOx temperature to them. This set of $\mathrm{pH_{NBS}}$ values was designated $\mathrm{pH_{NBS}^{field}}$. Following this step, a second correction based on the difference between $\mathrm{pH_{NBS}}$ and $\mathrm{pH_T}$ calculated from measured DIC and TA at *in-situ* temperature (T), salinity (S), and pressure (P) solely attributed to the difference in the pH scales ($\Delta\mathrm{pH_{scales}}$) was then applied to $\mathrm{pH_{NBS}^{field}}$ to generate the second independent reference pH ($\mathrm{pH_{elec}^{disc}}$) per the following conventions:

$$\Delta\mathrm{pH_{scales}}(\mathrm{T,S,P}) = \mathrm{pH_{NBS}^{DIC-TA}}(\mathrm{T,S,P}) - \mathrm{pH_T^{DIC-TA}}(\mathrm{T,S,P}) \quad (21)$$

$$\mathrm{pH_{elec}^{disc}}(\mathrm{T,S,P}) = \mathrm{pH_{NBS}^{field}}(\mathrm{T,S,P}) - \Delta\mathrm{pH_{scales}}(\mathrm{T,S,P}) \quad (22)$$

By applying this pair of corrections, the pH_{elec}^{disc} was converted to the total scale (pH_T) used in the sensor calibration and *in-situ* environmental conditions (temperature, salinity, and pressure) were conserved.

2.2.8.4 Departure from Open-Ocean Calibration Approach

The remoteness of open-ocean sensor deployment sites can limit access to the sensor over the course of long-term deployments and the ability to collect quality discrete bottle samples. This often results in the need for complicated, time-intensive collection methods for those samples (e.g. via Niskin Bottle by Scuba Diver) (Rivest *et al.*, 2016). It would also prevent the performance of regularly scheduled maintenance needed for the continuous collection of high-quality data in dynamic, productive, high-fouling, highly-turbid environments. Thus, tradition has dictated that a sensor time series generated from such deployments be treated as one continuous time series from beginning to end with pre-deployment calibration and/or post-deployment calibrations sometimes attached to validate sensor performance (Bresnahan *et al.*, 2014). In the case of an estuarine sensor deployment in an accessible location characterized by an adequate number of quality independent reference pH measurements from quality discrete bottle samples and regularly scheduled sensor maintenance, a departure from this convention is justified.

The results of such sensor deployments leave the sensor operator with many shorter-duration sensor time series that may be calibrated independently since the sensor operation is essentially reset following the conclusion of each maintenance

session. These independent sensor time series can be combined after the sensor deployment to build an aggregated sensor time series. This more frequent and rigorous calibration approach provides a better way of accounting for the natural variability inherent to the estuarine deployment environment. Ultimately, this approach would facilitate an evaluation of the performance of the Honeywell Durafet in an estuarine environment over timescales bracketing those performed in controlled laboratory environments (hours to days) (Bresnahan *et al.*, 2014; Martz *et al.*, 2010) and in the open ocean (multiple months) (Bresnahan *et al.*, 2014; Rivest *et al.*, 2016).

2.2.8.5 Types of Calibrations

At present, there are two contrasting methods that have been used to calibrate Durafet-based biogeochemical sensors: pH Domain and Raw Signal Domain. The first calibration method is comparative and only requires access to the sensor pH (pH^{sensor}) and the independent reference pH. The second method is based on the raw signal domain of the data which is much more intricate and requires intimate knowledge of the sensor operational theory. The raw signal domain requires access to the measured voltages and the independent reference pH as well. In the following sections, we provide a brief overview of each calibration type.

2.2.8.5.1 pH Domain

Essentially, the calibration method encompassing the pH domain is a simple comparison of the raw, uncalibrated sensor output (pH^{sensor}) against an independent

reference pH (e.g. pH measured from discrete bottle samples - pH^{disc}). Property-property plots of pH^{sensor} (dependent variable) vs. pH^{disc} (independent variable) provide valuable insight into sensor performance reflected in the degree of deviation of the sensor output from the truest pH of the water seen by the sensor at that time (pH^{disc}) (Bresnahan *et al.*, 2014). These deviations can be characterized from the intercept (sensor offset - c_0) and slope (sensor gain - c_1) (Bresnahan *et al.*, 2016) generated from a Model II least squares fit of the data (Peltzer, 2007). If significant deviations in the intercept and slope from 0 and 1, respectively, are observed, this suggests bias in the sensor and/or reference pH used in the comparison. Calibrating a raw sensor time series only using this approach is not recommended as it limits the identification of and the differentiation between all the potential sources of error, thereby introducing significant sampling error to the final sensor time series (Bresnahan *et al.*, 2014).

2.2.8.5.2 Raw Signal Domain

The raw signal domain provides a more robust and dependable means of calibrating a raw sensor output. It is done through the calculation and application of a new set of calibration constants (E^{*}_{sensor}) to displace the raw sensor measurements to bring them in-line with the chosen reference pH (Bresnahan *et al.*, 2014). To report a reasonable range of pH values for a given deployment environment, all SeapHOx units are programmed with a pair of initial calibration constants corrected to the calibration

temperature of 25°C specific to each reference electrode when they are first assembled, $E^*_{INT,0}$ (T = 25℃) (Ag/AgCl) and $E^*_{EXT,0}$ (T = 25℃) (Cl-ISE). These values serve as nominal placeholders used to calculate pH^{INT} and pH^{EXT} using equations 15-17 and 18-20, respectively (Martz, 2012). The values of $E^*_{INT,0}$ (T = 25℃) and $E^*_{EXT,0}$ (T = 25℃) programmed into SeapHOx SP053 were -0.4347 V and -1.4070 V, respectively. Many researchers elect to perform pre-deployment calibrations, thus yielding a more dependable and realistic set of calibration constants that can be utilized from the onset of sensor deployment (Bresnahan *et al.*, 2014). However, since this was a pilot deployment of a Durafet-based biogeochemical sensor in an estuarine environment with no available comparisons, a system of rigorous *in-situ* sensor deployment calibrations was pursued.

The calculation of new sets of calibration constants is simplified via the use of a MATLAB script that uses inputs of the measured sensor voltages ($E_{INT}(T)$/ $E_{EXT}(T)$), *in-situ* temperature (T), *in-situ* salinity, and the independent reference pH at *in-situ* temperature ($pH_T(T)/pH_F(T)$) corresponding to the same time stamp of an individual sensor measurement. For the FET|INT, the new calibration constants at *in-situ* temperature, $E^*_{INT}(T)$, and the final calibration constants corrected to 25°C, $E^*_{INT,f}$ (T = 25℃), are calculated via the following equations (Bresnahan *et al.*, 2014):

$$E^*_{INT}(T) = E_{INT}(T) - S(T)\times pH_T(T) \quad (23)$$

$$E^*_{INT,f}\,(T = 25℃) = E^*_{INT}(T) + \frac{dE^*_{INT}}{dT}(25℃ - T) \quad (24)$$

where S(T) corresponds to the Nernst slope at *in-situ* temperature and $\mathrm{d}E^*_{INT}/\mathrm{d}T$ corresponds to its previously defined quantity. For the FET|EXT, $\mathrm{E}^*_{\mathrm{EXT}}(\mathrm{T})$ and $\mathrm{E}^*_{\mathrm{EXT,f}}\,(\mathrm{T}=25℃)$ are calculated as follows (Bresnahan *et al.*, 2014):

$$\mathrm{E}^*_{\mathrm{EXT}}(\mathrm{T}) = \mathrm{E}_{\mathrm{EXT}}(\mathrm{T}) + \mathrm{S}\times\log(\gamma_{\mathrm{H}}\gamma_{\mathrm{Cl}}\mathrm{m}_{\mathrm{Cl}}) - \mathrm{S}\times\mathrm{pH}_{\mathrm{F}}\,(\mathrm{T}) \tag{25}$$

$$\mathrm{E}^*_{\mathrm{EXT,f}}\,(\mathrm{T}=25℃) = \mathrm{E}^*_{\mathrm{EXT}}(\mathrm{T}) + \frac{\mathrm{d}E^*_{EXT}}{\mathrm{d}T}(25℃ - \mathrm{T}) \tag{26}$$

where γ_{i} represents the ion activity coefficients of either H^+ or Cl^-, $\mathrm{m_i}$ represents the molality of either H^+ or Cl^-, and $\mathrm{d}E^*_{EXT}/\mathrm{d}T$ corresponds to its previously defined quantity. Throughout this process, values of $\mathrm{E}^*_{\mathrm{EXT}}(\mathrm{T})$ and $\mathrm{E}^*_{\mathrm{EXT,f}}\,(\mathrm{T}=25℃)$ are calculated from independent reference values of $\mathrm{pH_F}$ corresponding to the time stamp of an individual sensor measurement (Bresnahan *et al.*, 2014).

Once the values of $\mathrm{E}^*_{\mathrm{INT,f}}\,(\mathrm{T}=25℃)$ and $\mathrm{E}^*_{\mathrm{EXT,f}}\,(\mathrm{T}=25℃)$ are set, the application of these values and subsequent recalculation of all pH values is expedited by a second MATLAB script designed to batch process all sensor measurements in a single sensor time series (Bresnahan *et al.*, 2014). The pH values are recalculated using modified forms of equations 17 and 20 for $\mathrm{pH}^{\mathrm{INT}}$ and $\mathrm{pH}^{\mathrm{EXT}}$, respectively as follows:

$$\mathrm{pH}^{\mathrm{INT}}_{\mathrm{final}} = \frac{\left(\mathrm{E}_{\mathrm{INT}} - \mathrm{E}^*_{\mathrm{INT,f}}\,(\mathrm{T}=25℃)\right)}{\mathrm{S}} \tag{27}$$

$$\mathrm{pH}^{\mathrm{EXT}}_{\mathrm{final}} = \frac{\left(\mathrm{E}_{\mathrm{EXT}} - \mathrm{E}^*_{\mathrm{EXT,f}}\,(\mathrm{T}=25℃)\right) + \mathrm{S}\times\log(\gamma_{\mathrm{H}}\gamma_{\mathrm{Cl}}\mathrm{m}_{\mathrm{Cl}})}{\mathrm{S}} \tag{28}$$

where pH_{final}^{INT} and pH_{final}^{EXT} correspond to the final calibrated pH values measured with the internal and external reference electrodes, respectively. In this final step, standard practice dictates that the pH_F taken from E_{EXT} be converted to pH_T, thereby making the two pH values measured by the different reference electrodes directly comparable (Bresnahan *et al.*, 2014).

When pH values populating a final sensor time series are calibrated, all raw sensor measurements (voltages, *in-situ* temperature, and *in-situ* salinity) are constants and used to recalculate the calibration constants based on the reference pH chosen. This means the calibration constants and the pH values are dependent one another such that the difference between the raw sensor pH and the independent reference pH is the principle control on the magnitude of the new calibration constant. This caveat propagates through the calibration process since the only thing that changes when calibrating a sensor time series is the calibration constant. So, the difference between the final calibrated sensor pH (pH_{final}^{sensor}) and raw uncalibrated sensor pH (pH_{raw}^{sensor}) is dependent on and inversely related to the difference between the initial ($E_{sensor,0}^{*}\ (T = 25℃)$) and final ($E_{sensor,f}^{*}\ (T = 25℃)$) calibration constants according to the quantity $\left(E_{sensor} - E_{sensor,f}^{*}\ (T = 25℃)\right)$ in equations 27 and 28. When redefining the calibration constant and calibrating a sensor time series, the only change is the contribution of the calibration constant to the total voltage, $\left(E_{sensor} - E_{sensor,f}^{*}\ (T = 25℃)\right)$, converted to pH via the temperature-dependent Nernst slope. Since both calibration constants are negative quantities, if

$E^*_{sensor,f}\,(T = 25℃) < E^*_{sensor,0}\,(T = 25℃)$, then $\left(E_{sensor} - E^*_{sensor,f}\,(T = 25℃)\right) > \left(E_{sensor} - E^*_{sensor,0}\,(T = 25℃)\right)$ and $pH^{sensor}_{final} > pH^{sensor}_{raw}$, and vice versa.

2.2.9 Assumptions & Limitations

The deployment of Durafet-based biogeochemical sensors in an estuarine environment experiencing lower salinities than those in open-ocean environments ($S \geq 30$) is not without its assumptions. The main assumption associated with the present work was assuming 100% Nernstian response in pH calculations over the full range of salinity encountered. Previous work has demonstrated that the dual-reference electrode configuration is capable of repeatedly exhibiting 100% Nernstian response over a wide range of pH (2-12) and a narrow salinity range (20-35) in seawater media (Takeshita *et al.*, 2014). It was also assumed the effects of the major interfering cations found in seawater referred to in equation 11 are negligible (leading to pH error of <0.00005 over $20 \leq S \leq 35$) as demonstrated by Takeshita *et al.* (2014). This allowed those quantities to be excluded from equations 12 and 13. Related to the Cl-ISE, its cross-sensitivity to bromide was found not to be a substantial source of error over $20 \leq S \leq 35$ (pH error of <0.003) (Takeshita *et al.*, 2014). As such, it was assumed that the Cl-ISE should continue to operate properly if the Cl^- to Br^- ratio of any new deployment environment is similar to that of open-ocean systems (Rérolle *et al.*, 2016). Finally, it was assumed the Debye-Hückel relationship used to calculate the mean activity

coefficient of HCl ($\gamma_H\gamma_{Cl}$) constrained by Khoo *et al.* (1977) was valid outside of its published salinity range (S=20-45).

The deployment of Durafet-based biogeochemical sensors in an estuarine environment experiencing lower salinities than those in open-ocean environments ($S \geq 30$) is also not without its potential limitations. The internal reference electrode (Ag/AgCl) contains a liquid junction, which could possibly impart errors of 0.01-0.14 pH units (Dickson, 1993b; Easley and Byrne, 2012) to the pH^{INT} time-series only because of liquid junction potential effects (Bresnahan *et al.*, 2014; Martz *et al.*, 2010). The effects of liquid junction potential can be especially problematic in dynamic environments that routinely experience large freshwater inputs (Bresnahan *et al.*, 2014) since these effects cannot be effectively quantified or predicted (Bates, 1973). Other possible limitations on electrode performance include – (1) hysteresis effects, (2) sensor drift (Bresnahan *et al.*, 2014; Martz *et al.*, 2010), and (3) streaming effects at the electrode surface under differential flow conditions during individual sampling cycles (Bagshaw *et al.*, 2016). Alternatively, under very low ionic strength (low salinity) conditions, if too few electrons are present, it could compromise voltage measurements made with any electrode couple (Bates, 1973).

2.3 Results

2.3.1 Discrete Sample pH Comparisons

The four sets of relevant pH values ($pH_{raw}^{INT}, pH_{raw}^{EXT}, pH_{DIC-TA}^{disc}, pH_{elec}^{disc}$) coinciding with the sampling trips on 01 June 2016 (Figure 7) and 02 August 2016

(Figure 8) were compared against one another as a preliminary check of the quality of the discrete sampling strategy. On both occasions, discrete bottle samples were collected on sampling trips that bracketed the time of at least one low tide and one high tide. On both occasions, the relative temporal trends between all four measurements of pH were conserved across the full range of pH and salinity encountered during both periods of slower decline on the ebb tide and more rapid rise on the flood tide. Yet, pH^{disc}_{DIC-TA} was always slightly higher than pH^{disc}_{elec}. This difference may be attributed to a number of factors, either singly or in combination – e.g. small errors in the analysis of discrete bottle samples (Bresnahan *et al.*, 2014; Patsavas *et al.*, 2015) and/or thermodynamic constants used to calculate pH^{disc}_{DIC-TA} and pH^{disc}_{elec} (Patsavas *et al.*, 2015; Rérolle *et al.*, 2016), existence of environmental pH gradients of varying magnitudes (Bresnahan *et al.*, 2014), and other sources of sampling/handling error (ex. CO_2 outgassing out of or invasion into samples during sample collection) (Rérolle *et al.*, 2016). A final, albeit important possibility may be contributions of excess alkalinity (Patsavas *et al.*, 2015; Ribas-Ribas *et al.*, 2014) which have been known to increase in magnitude when working in coastal waters (Cai *et al.*, 1998; Yang *et al.*, 2015).

On 01 June 2016 Figure 7), a total of 38 discrete bottle samples for DIC and TA were collected from 0900-1900. Duplicate samples for DIC and TA were collected during periods of mixing between endmembers and single samples for DIC and TA were collected when the deployment environment was inundated by the relatively

uniform, higher salinity Delaware Bay water. These efforts yielded 21 individual calibration points for pH_{DIC-TA}^{disc} coinciding with the times of sensor measurements. However, due to the time needed for the thermal equilibration of the pH buffers in the field, duplicate sampling for pH_{NBS}^{field} did not start until 0930 which resulted in only 20 calibration points for pH_{elec}^{disc}. Over the course of the day, the pH_T values of the discrete bottle samples varied from about pH ~7.1-8.3 with localized morning maxima of pH ~8.11-8.13 at about S=21.8, localized early afternoon minima of pH ~7.18 at about S=9.0, and localized evening maxima of pH ~8.30 at about S=22. The large range of pH captured on this sampling trip can attributed to the large, but episodic, phytoplankton blooms that characterize this system during the early summer under favorable conditions of warm, sunny weather (Voynova *et al.*, 2015) and strong winds (Wong *et al.*, 2009). The salinity range varied from S=8.99-22.31 with the lowest salinities being observed around low tide in the early afternoon and the highest salinities observed around high tide in the evening.

On 02 August 2016 (Figure 8), a total of 43 discrete bottle samples for DIC and TA were collected from 0800-1930. Duplicate samples for DIC and TA were collected during periods of mixing between endmembers and single samples for DIC and TA were collected when the deployment environment was inundated by the relatively uniform, higher salinity Delaware Bay water. Duplicate samples for pH_{NBS}^{field} were always collected. These efforts produced 24 individual calibration points coinciding with the times of sensor measurements for both pH_{DIC-TA}^{disc} and pH_{elec}^{disc}.

Throughout the day, the pH_T values of the discrete bottle samples varied from about pH ~7.0-7.9 with localized early afternoon maxima of pH ~7.71-7.74 at about S=26.6, localized late afternoon minima of pH ~7.00-7.10 at about S=21.1, and localized evening maxima of pH ~7.75-7.85 at about S=26.7. The dampened range of pH captured on this sampling trip may be reflective of well-characterized decrease in primary production routinely observed in the system during the mid-late summer relative to the early summer (Voynova *et al.*, 2015). In addition, the narrower pH range may also be attributed to a series of short, violent storms characterized by strong, prolonged winds experienced in the area on the day before the sampling trip. These storms may have potentially churned up sediment from the river bottom thereby increasing turbidity and contributed to periods of cloudiness on the day of sampling. Also, strong NE winds may have facilitated the propagation of the more saline Delaware Bay water upstream into the Murderkill Estuary thereby limiting exchange between system endmembers (Dzwonkowski *et al*., 2013; Wong *et al.*, 2009) and instantaneous primary production in the system as well (Voynova *et al.*, 2015). Furthermore, the salinity fluctuation over a longer period was much smaller (S=21.06-27.06) and the absolute salinity values were much higher. These trends both aligned with the dryer mid-late summer trends observed in previous years in the Murderkill Estuary-Delaware Bay System (Kent County Land-Ocean Biogeochemical Observatory, http://kentcounty.loboviz.com/).

2.3.2 Raw Sensor pH vs. Independent Reference pH

The raw uncalibrated sensor pH calculated from measured voltages using both the internal and external reference electrodes were compared against the two sets of independent reference pH generated from discrete sampling trips on 01 June 2016 (Figures 9 and 10) and 02 August 2016 (Figures 11 and 12). The relatively good agreement between the raw sensor pH and the independent reference pH is manifested in their near 1:1 linear relationship ($\mathrm{pH}_{\mathrm{sensor}}^{\mathrm{raw}} = \mathrm{pH}^{\mathrm{disc}}$). This demonstrates the viability of the Honeywell Durafet to the collection of high-resolution pH data in a dynamic, productive, high-fouling, highly-turbid estuarine system. Moreover, the good agreement between the raw uncalibrated sensor pH and the independent reference pH considered to be the "true" pH of the natural waters sampled by the sensor validated the scope and design of the discrete sampling regiment.

In contrast, the apparent deviation of some of the data points from the near 1:1 linear relationships suggest the presence of additional controls on data quality not accounted for in Section 2.3.l. A collection of data in the upper salinity ranges from both sampling days exhibit significant deviations. This may be indicative of small spatiotemporal mismatches between the discrete bottle samples and sensor measurements attributed to: (1) environmental pH gradients (Bresnahan *et al.*, 2014), (2) influences of the pier the sensor deployment platform is mounted to on system mixing at different stages of flow, and/or (3) rapid biogeochemical fluctuation during periods of rapid, intense mixing (e.g. flood tide). Conversely, the variability among the lowest salinity data point from 01 June 2016 may be indicative of water column

stratification during periods of slack tide when the horizontal flow drops to near-zero values. In the absence of horizontal flow, there is almost no vertical mixing. This could lead to more subtle pH differences between the water the sensor measured and the water captured in the discrete bottle samples. In other words, the physics of the system becomes part of the natural variability inherent to the deployment environment integrated into the sensor calibration. Frequent rigorous calibration of the sensors is needed to tease apart the multitude of potential controls of sensor performance and data quality.

2.3.3 Calculating New Sets of $\mathbf{E^{*}_{sensor,f}}$ ($\mathbf{T = 25}$℃)

The raw sensor time series of each reference electrode were recalibrated using both sets of independent reference pH values: (1) $\mathrm{pH}^{\mathrm{disc}}_{\mathrm{DIC-TA}}$ and (2) $\mathrm{pH}^{\mathrm{disc}}_{\mathrm{elec}}$. As an additional feature to the evaluation of the viability of the external reference electrode to the measurement of pH in estuarine and coastal ocean systems, $\mathrm{pH}^{\mathrm{EXT}}_{\mathrm{raw}}$ from 01 June 2016 was recalibrated to two different values of $\mathrm{E}^{*}_{\mathrm{EXT,f}}$ ($\mathrm{T} = 25$℃): (A) using all good calibration points over full range of salinity and (B) using only the good calibration points above S=20. Henceforth, the calibration constants and subsequently recalibrated sensor time series will be referred to by the alphanumeric designation of their respective calibration method – (1)/(2) according to the independent reference pH used and (A)/(B) according to any salinity constraints of the calibration. The alphanumeric designations can be found in the subscript of all parameter names (e.g.

$pH_{final,1}^{INT}$ corresponds to the final sensor time series calculated from voltage measured using the internal reference electrode calibrated against pH_T calculated from measured DIC and TA (pH_{DIC-TA}^{disc})). All alphanumeric designations were summarized in Table 2.

2.3.3.1 Preliminary Comparisons

The new sets of calibration constants for the internal reference electrode on 01 June 2016 (Figure 13) and 02 August 2016 (Figure 14) and the external reference electrode on 01 June 2016 (Figure 15) and 02 August 2016 (Figure 16) were compared. In accordance with the inverse relationship between $E_{sensor,f}^{*}$ $(T = 25℃)$ and pH_{final}^{sensor}, sets of calibration constants calculated from pH_{DIC-TA}^{disc} were always less than those calculated from pH_{elec}^{disc}. Moreover, periods of time corresponding to the dominance of the higher salinity Delaware Bay water exhibited the least variance, while periods of mixing with the Murderkill Estuary outflow exhibited the greatest variance. More importantly, the $E_{EXT,f}^{*}$ $(T = 25℃)$ profile of 01 June 2016 exhibited a steadily decreasing trend at S < 20. However, provided the sensor still obeyed the Nernst equation at S < 20, the differences in $E_{EXT,f}^{*}$ $(T = 25℃)$ between successive measurements should be minimized. This seems to have been the case for some of the calibration points over the S < 20 range on 01 June 2016.

Every set of calibration constants calculated from both sets of independent reference pH for both reference electrodes on both sampling days all exhibited the

feature of a notable decrease in $E^{*}_{sensor,f}$ $(T = 25℃)$ at the beginning of flood tide. A spatiotemporal mismatch between the discrete bottle sample and sensor measurement is the most likely source of this feature consistent with the findings of Bresnahan *et al.* (2014). The critical examination of the profiles associated with new calculated calibration constants was used to identify outliers to be excluded from subsequent analyses.

2.3.3.2 Setting Final Values of $E^{*}_{sensor,f}$ $(T = 25℃)$

After eliminating all outliers, each $E^{*}_{INT,f}$ $(T = 25℃)$ and $E^{*}_{EXT,f}$ $(T = 25℃)$ was set to an average value based on many discrete samples (Bresnahan *et al.*, 2014) collected over a range of environmental conditions. This was done to minimize the anomaly between the sensor pH and reference pH (Bresnahan *et al.*, 2014; Bresnahan *et al.*, 2016). Essentially, the anomaly between the sensor pH and reference pH was set to 0. The root mean squared error (RMSE) calculated from a Model II least squares fit of the comparison becomes the degree to which the constraint of the difference between the sensor pH and reference pH can be characterized, or the accuracy of the sensor pH relative to the chosen reference pH (Dr. Todd Martz, personal communication). The average calibration constants used to recalibrate all raw sensor time series were summarized in Table 3.

2.3.4 Calibrated Sensor pH vs. Independent Reference pH

2.3.4.1 01 June 2016

The calibrated sensor pH of $\mathrm{pH}^{\mathrm{INT}}_{\mathrm{final,1}}$ (Figure 17), $\mathrm{pH}^{\mathrm{EXT}}_{\mathrm{final,1A}}$ (Figure 18), $\mathrm{pH}^{\mathrm{EXT}}_{\mathrm{final,1B}}$ (Figure 19), $\mathrm{pH}^{\mathrm{INT}}_{\mathrm{final,2}}$ (Figure 20), $\mathrm{pH}^{\mathrm{EXT}}_{\mathrm{final,2A}}$ (Figure 21), and $\mathrm{pH}^{\mathrm{EXT}}_{\mathrm{final,2B}}$ (Figure 22) from 01 June 2016 generally exhibited good agreement when compared against their respective independent reference pH with an intercept or sensor offset (c_0) close to 0 and a slope or sensor gain (c_1) close to 1 (Table 4). However, more calibration points were incorporated into the calibration of the raw sensor pH calculated from measured voltages using internal reference electrode than with the external reference electrode. Ultimately, if the salinity range of the discrete samples used in the sensor calibration was representative of the salinity range of the raw sensor measurements, this should reduce the unintentional bias introduced to the final sensor time series, specifically to the sensor measurements falling outside of the salinity range over which discrete bottle sample were collected.

When compared, $\mathrm{pH}^{\mathrm{INT}}_{\mathrm{final,1}}$ had a sensor offset and sensor gain further away from 0 and 1, respectively, than $\mathrm{pH}^{\mathrm{INT}}_{\mathrm{final,2}}$. This suggested that calibrating raw sensor output to $\mathrm{pH}^{\mathrm{disc}}_{\mathrm{DIC-TA}}$ results in the overestimation of the sensor gain. This introduces bias to independent reference pH and the subsequently calibrated sensor time series due to contributions of excess alkalinity. Conversely, calibrating raw sensor output against a direct pH measurement ($\mathrm{pH}^{\mathrm{disc}}_{\mathrm{elec}}$) eliminated the effects of excess alkalinity and resulted in the inclusion of two additional low salinity calibration points (S~9).

Yet, there was a tradeoff associated with this action quantified by an increase in the RMSE of $\mathrm{pH}_{\mathrm{final},2}^{\mathrm{INT}}$ of 0.008 pH units.

When $\mathrm{pH}_{\mathrm{final},1\mathrm{A}}^{\mathrm{EXT}}$ and $\mathrm{pH}_{\mathrm{final},2\mathrm{A}}^{\mathrm{EXT}}$ were calibrated, a range of S=14.84-22.31 was used. When compared, $\mathrm{pH}_{\mathrm{final},2\mathrm{A}}^{\mathrm{EXT}}$ was characterized by a sensor offset and sensor gain closer to 0 and 1, respectively. This indicated that the calibration traced back to the direct measurements of $\mathrm{pH}_{\mathrm{NBS}}^{\mathrm{field}}$ yielded better agreement between the calibrated sensor pH and the independent reference pH. In contrast, the sensor gain associated with $\mathrm{pH}_{\mathrm{final},1\mathrm{A}}^{\mathrm{EXT}}$ was found to be the highest for 01 June 2016. Like $\mathrm{pH}_{\mathrm{final},1}^{\mathrm{INT}}$, bias attributed to the effects of excess alkalinity was also introduced to the calibrated sensor time series of $\mathrm{pH}_{\mathrm{final},1\mathrm{A}}^{\mathrm{EXT}}$. Despite this, since the bias present in $\mathrm{pH}_{\mathrm{final},1\mathrm{A}}^{\mathrm{EXT}}$ was not manifested in $\mathrm{pH}_{\mathrm{final},2\mathrm{A}}^{\mathrm{EXT}}$, any error associated with the operation of the external reference electrode below S < 20 down to S~15 was shown to be minimal.

The sensor time series, $\mathrm{pH}_{\mathrm{final},1\mathrm{B}}^{\mathrm{EXT}}$ and $\mathrm{pH}_{\mathrm{final},2\mathrm{B}}^{\mathrm{EXT}}$, were calibrated using only those measurements above S=20. In both cases, sensor offsets and sensor gains were close to 0 and 1, respectively. This confirmed the satisfactory performance of external reference electrode at salinities at S > 20 as characterized by Martz *et al.*, (2010), Bresnahan *et al.* (2014), and Takeshita *et al.* (2014). The errors associated with c_0 and c_1 for the comparisons of $\mathrm{pH}_{\mathrm{final},1\mathrm{B}}^{\mathrm{EXT}}$ and $\mathrm{pH}_{\mathrm{final},2\mathrm{B}}^{\mathrm{EXT}}$ versus their respective independent reference pH were significantly greater than the others probably due the limited number of calibration points and the higher variability present within a limited pH range. Quality discrete samples are needed to develop a suitable calibration for the raw

sensor output (Bresnahan *et al.*, 2014; Rivest *et al.*, 2016). A sensor gain slightly less than 1 for $\mathrm{pH}_{\mathrm{final,1B}}^{\mathrm{EXT}}$ also indicated that excess alkalinity had a decreased effect on the S > 20 data.

2.3.4.2 02 August 2016

The calibrated sensor pH of $\mathrm{pH}_{\mathrm{final,1}}^{\mathrm{INT}}$ (Figure 23), $\mathrm{pH}_{\mathrm{final,1}}^{\mathrm{EXT}}$ (Figure 24), $\mathrm{pH}_{\mathrm{final,2}}^{\mathrm{INT}}$ (Figure 25), and $\mathrm{pH}_{\mathrm{final,2}}^{\mathrm{EXT}}$ (Figure 26) from 02 August 2016 also generally exhibited good agreement when compared against their respective set of independent reference pH with sensor offsets (c_0) close to 0 and sensor gain (c_1) close to 1 (Table 5). Unlike the previously discussed sampling day, 02 August 2016 was characterized higher salinities (S > 20), a narrower salinity range, and slower rates of salinity change. Raw sensor time series recalibrated using $\mathrm{pH}_{\mathrm{DIC-TA}}^{\mathrm{disc}}$ exhibited greater sensor gains than those recalibrated with $\mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}$. This may be indicative of the effects of excess alkalinity like what was observed on 01 June 2016. However, when the sensor offsets and gains of $\mathrm{pH}_{\mathrm{final}}^{\mathrm{INT}}$ were compared with $\mathrm{pH}_{\mathrm{final}}^{\mathrm{EXT}}$, they were much closer to 0 and 1, respectively. This may indicate bias present in the sensor time series of the internal reference electrode attributed to the previously discussed sources.

2.4 Discussion

2.4.1 Electrode Performance in an Estuarine System

A single pair of calibrated sensor time series from each sampling day calibrated using the same independent reference pH - $\mathrm{pH}_{\mathrm{final,2}}^{\mathrm{INT}} + \mathrm{pH}_{\mathrm{final,2A}}^{\mathrm{EXT}}$ for 01 June 2016 and $\mathrm{pH}_{\mathrm{final,2}}^{\mathrm{INT}} + \mathrm{pH}_{\mathrm{final,2}}^{\mathrm{EXT}}$ for 02 August 2016 – was used in lieu of the six total pairs. The sensor time series calibrated using $\mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}$ were used because the time series were always characterized by better Model II fit parameters and were theoretically free of the effects of excess alkalinity.

2.4.1.1 General Electrode Performance

For the duration of the sensor deployments from 9 May 2016 to 09 June 2016 calibrated using the discrete bottle samples collected on 01 June 2016 (Figure 27A) and 20 July 2016 to 24 August 2016 calibrated using the discrete bottle samples on 02 August 2016 (Figure 28A), voltages measured with both the internal and external reference electrodes align very well across the full spectrum of salinity characteristic of a flood-dominant, estuarine system. Likewise, the pH values calculated from the measured voltages exhibit good agreement across the lower and more intermediate salinities encountered in the 09 May 2016 to 09 June 2016 deployment (Figure 27B) of S=3.25-25.00 and the consistently higher salinities characterizing the 20 July 2016 to 24 August 2016 deployment (Figure 28B) of S=15.33-29.83. During these two sensor deployments, pH fluctuations ranging from <0.5 pH units to >1 pH unit were

routinely captured over the course of single tidal cycles over that full salinity range. Notable departures in the agreement between both the voltages and the subsequently calculated pH values were observed during periods of pH change during the ebb and flood tides. These departures were most noticeable at the minima of the profiles where both pH and salinity were generally at their lowest values, and at the maxima of the profiles when both pH and salinity returned to their higher values on the flood tide.

The nominal seawater voltage ranges for E_{INT} and E_{EXT} characteristic of open-ocean conditions were reported to be 0.03 to 0.1 V and -0.95 to -0.80 V, respectively (Bresnahan *et al.*, 2014). When compared, the voltage ranges for E_{INT} characterizing the two estuarine deployments were much lower at around -0.025 to 0.055 V. This was to be expected in an estuarine system characterized by consistently lower salinities and sub-pH 7 values at low tide which resulted in lower measured voltages for E_{INT}. On the other hand, when the voltage ranges of E_{EXT} were compared, the minima fell slightly outside of the nominal range for seawater at -0.98 to -0.86 V. Similarly, a different voltage range for E_{EXT} would be expected for an estuarine system that experiences a lower salinity range and more substantial simultaneous temperature, salinity, and pH variability.

2.4.1.2 Temporal Evolution of $\Delta pH^{INT-EXT}$ Anomaly

The dual-reference electrode configuration incorporated into Durafet-based biogeochemical sensors like the SeapHOx is not an absolute necessity for their use in different environments (Bresnahan *et al.*, 2014). In open-ocean deployment

environments, this configuration provides a simple, yet powerful method of detecting the effects of fouling or sensor failure through the temporal evolution of $\Delta pH^{INT-EXT}$ anomalies (Bresnahan *et al.*, 2014, Rivest *et al.*, 2016). In an estuarine system, both the instantaneous salinity experienced and the relative change in salinity between successive sensor measurements can negatively impact electrode. The $\Delta pH^{INT-EXT}$ anomaly is the only quasi-suitable metric used to gauge sensor performance over time to look for the effects of drift, fouling, and sensor failure in the present work as well. Average non-zero $\Delta pH^{INT-EXT}$ anomalies were observed during both the 09 May 2016 to 09 June 2016 (0.006±0.063 pH units) and 20 July 2016 to 24 August 2016 (-0.008±0.020 pH units) deployments with conditioning periods removed. The $\Delta pH^{INT-EXT}$ anomalies observed in the present work were much greater than those characterized by Bresnahan *et al*. (2014) (<0.005 pH units) at $30 \leq S \leq 36$. However, the current $\Delta pH^{INT-EXT}$ anomalies are insignificant when natural pH fluctuations at least one order of magnitude greater than the anomaly are routinely captured.

During both sensor deployments, 09 May 2016 to 09 June 2016 (Figure 29) and 20 July 2016 to 24 August 2016 (Figure 30), the $\Delta pH^{INT-EXT}$ anomalies followed the same patterns. On the ebb tide, a positive $\Delta pH^{INT-EXT}$ anomaly ($pH_{final}^{INT} > pH_{final}^{EXT}$) was consistently observed under prolonged periods of salinity decrease on the ebb tide which bottomed out at the lowest salinities coinciding with the greatest influences of the fresher Murderkill Estuary outflow. This was followed by rapid, sharp decrease to a negative $\Delta pH^{INT-EXT}$ anomaly ($pH_{final}^{INT} < pH_{final}^{EXT}$) coinciding with the tide change

and the first flood tide measurements. Afterwards, the $\Delta pH^{INT-EXT}$ anomaly approaches a near-zero value as the deployment environment becomes inundated with the higher salinity Delaware Bay water for increased periods of time prior to the next ebb tide. The result is an asymptotic profile for each tidal excursion across the full range of salinity experienced during both sensor deployments.

The magnitude of the $\Delta pH^{INT-EXT}$ anomalies was directly related to the magnitude of the salinity fluctuation, but these quantities were inversely related with respect to the exhibited trends. As characterized by Bresnahan *et al.* (2014) under open-ocean salinities, the source of this anomaly is most like a salinity lag arising from the inadequate flushing of the instrument flow path and flow housing. In other words, the salinity of the water sampled by the rapidly flushed conductivity-temperature sensor (SBE37) was different than the salinity of the water seen by the electrodes inside the more slowly flushed flow housing (Bresnahan *et al.*, 2014). Also, this anomaly may also be attributed to a slower response time of one or both reference electrodes at $S < 20$ and/or under large, rapid *in-situ* salinity changes. This would be manifested in the inadequate reconditioning or re-equilibration of the reference electrodes to the characteristic salinities experienced under dynamic flow conditions in an estuarine system. This demonstrates the need to characterize the performance of the dual-reference electrode configuration at the extremes of environmental conditions typical of estuarine and coastal ocean systems in a controlled laboratory setting.

2.4.1.3 Quality Control Considerations

Operators of Durafet-based biogeochemical sensors routinely report their Quality Control (QC) procedures used to flag, and if necessary, exclude data from final datasets incorporated into publications (Bresnahan *et al.*, 2014; Matson *et al.*, 2011; Rivest *et al.*, 2016) and major data repositories (Rivest *et al.*, 2016). These procedures use the measured pH ranges (Rivest *et al.*, 2016) and variability of pH with time (Bresnahan *et al.*, 2014; Rivest *et al.*, 2016) as data quality tools. Measurements found to fall outside of accepted pH ranges of the deployment environment may be flagged and excluded. If the sensor pH is changing significantly faster than *in-situ* pH determined from other methods, sensor measurements may also be flagged and excluded as well (Rivest *et al.*, 2016). For sensor deployments carried out in well-studied open-ocean systems, these QC procedures are reasonable (Bresnahan *et al.*, 2014; Rivest *et al.*, 2016). In rapidly changing estuarine systems, these QC procedures are less justifiable. Due to the physical and chemical variability in our estuarine system, we have chosen to report all sensor measurements to inform future work done in similar environments. However, the performance of future sensor deployments in similar systems under similar conditions would help develop a set of recommended QC procedures moving forward.

2.4.1.4 Sensor Drift

In a dynamic, productive, high-fouling, highly-turbid estuarine system characterized by a broad range of *in-situ* environmental conditions, there is a greater

possibility for sustained sensor drift that could be attributed to any number of sources (Bresnahan *et al.*, 2014; Martz *et al.*, 2010). However, due to the limitations of the experimental design, it would be difficult to attribute any anomalies observed in the data solely to sensor drift based on the complexity of sensor operation in this type of setting. To definitively attribute anomalies witnessed in the data to sensor drift, measurements or calculations of additional sets of independent reference pH made over a comparable frequency and range of environmental conditions would be needed. The current sensor deployment suffered from the fact that the only means of pH comparison available for the sensor output was by means of comparison against pH measured or calculated from discrete bottle samples collected over a significantly lower frequency. Even so, the accessibility of the deployment site allowed for the performance of regularly scheduled comprehensive sensor maintenance that should limit the occurrence, longevity, and magnitude of sensor drift in the present work. This was evident from the large pH fluctuations that were consistently captured throughout the present sensor deployment.

2.4.2 Effects of Excess Alkalinity

2.4.2.1 Excess Alkalinity Calculations

In estuarine and coastal ocean systems, the effects of excess alkalinity, which is alkalinity present in an aquatic system not accounted for by the marine CO_2 system has been well-characterized (Cai *et al.*, 1998; Yang *et al.*, 2015). Contributions to excess alkalinity can come in the form of organic material derived from humic

particles (Cai *et al.*, 1998; Yang *et al.*, 2015) as well as planktonic and bacterial cells (Kim & Lee, 2009; Ko *et al.*, 2016). Excess alkalinity is often referred to as organic alkalinity if it proves to strongly correlate with total dissolved organic matter (DOM), total dissolved inorganic carbon (DOC), or a similar metric (Cai *et al.*, 1998; Kim & Lee, 2009). Excess alkalinity is manifested as the difference in alkalinity (ΔTA) between measured alkalinity (TA_{meas}) and alkalinity calculated from two other measured marine CO_2 system parameters, which for the present work are pH_{NBS}^{field} and DIC, (TA_{calc}^{pH-DIC}):

$$\Delta TA = TA_{meas} - TA_{calc}^{pH-DIC} \quad (29)$$

When excess alkalinity is present, any inorganic parameter calculated using the measured alkalinity, such as pH_T calculated from measured DIC and TA will yield overestimations of them (Cai *et al.*, 1998; Patsavas *et al.*, 2015). This can be especially problematic when performing an intercomparison of marine carbonate chemistry measurements for a specific aquatic system (Patsavas *et al.*, 2015). However, even when excess alkalinity is present, it may not always have a statistically significant effect on any subsequent marine CO_2 system analyses (Loucaides *et al.*, 2012; Ribas-Ribas *et al.*, 2014).

2.4.2.2 Excess Alkalinity in the Murderkill Estuary-Delaware Bay System

The apparent effect of excess alkalinity on the two sets of independent reference pH was clearly distinguishable with values of pH_{DIC-TA}^{disc} consistently greater

than pH_{elec}^{disc} for all usable calibration points on both 01 June 2016 and 02 August 2016. When calculated, average values of ΔTA on 01 June 2016 (Figure 31A) and 02 August 2016 (Figure 32A) were 12.6±1.8 µmol kg^{-1} and 14.7±2.8 µmol kg^{-1}, respectively. Contributions of excess alkalinity generally increased with decreasing salinity which led to greater overestimations of pH_{DIC-TA}^{disc} relative to pH_{elec}^{disc} (Figures 31B and 32B) when measured total alkalinities were lower. This aligned well with trends described from comparisons of sensor pH and independent reference pH in preceding sections. However, excess alkalinity was not consistently strongly correlated with salinity over the course of sampling on 01 June 2016 (r^2 = 0.6564) or 02 August 2016 (r^2 = 0.2390), nor with DOC calculated from fluorescence using the summer 2013 relationship from Voynova *et al.* (2015) for the same deployment site on 01 June 2016 (r^2 = 0.0787) or 02 August 2016 (r^2 = 0.0832). Nevertheless, these results aligned well with those of other marine CO_2 chemistry studies performed in other parts of the Murderkill Estuary watershed near its confluence with the Delaware Bay (Ullman & Aufdenkampe, unpublished data).

Excess alkalinity was also calculated using a constant mole-ratio of excess alkalinity to total dissolved organic carbon (DOC) (Patsavas *et al.*, 2015). To do this, a constant organic-base-to-DOC ratio of 0.122 mol-alk/mol-C (Cai *et al.*, 1998) and DOC calculated from fluorescence (Voynova *et al.*, 2015) were used. Excess alkalinity in this system did not correlate well with the excess alkalinity calculated from DOC using this ratio on 01 June 2016 (r^2 = 0.0787) or 02 August 2016 (r^2 = 0.0832). No such mole ratio of excess alkalinity to DOC would suffice to describe the

contributions of excess alkalinity for this system. When summed up, it is possible that excess alkalinity is not related to total DOC but to a specific, uncharacterized fraction of the total and/or there is more than one source of excess alkalinity present in this system. Excess alkalinity never exceeded more than 1-2% of total measured alkalinity. The effects of excess alkalinity on the relative trends in pH observed in this system are insignificant when pH fluctuations as large as 0.5 pH units to a full pH unit are routinely experienced over tidal excursions alone.

2.4.2.3 Quantifying Effects of Excess Alkalinity on Sensor Output

When setting $E^{*}_{sensor,f}\,(T = 25℃)$ to an average value to minimize the anomaly between the sensor pH and pH^{disc}_{DIC-TA}, the effects of excess alkalinity captured in discrete bottle samples are averaged out. However, this approach does produce multiple calibrated sensor time series with those calibrated using pH^{disc}_{DIC-TA} exhibiting elevated pH signals relative to those calibrated using pH^{disc}_{elec}. Given the inverse, dependent relationship between $E^{*}_{sensor,f}\,(T = 25℃)$ and pH^{sensor}_{final}, the magnitude of the difference between the corresponding pH values in each sensor time series calibrated using pH^{disc}_{DIC-TA} and pH^{disc}_{elec} will be controlled by the difference in the values of $E^{*}_{sensor,f}\,(T = 25℃)$ used for calibration.

Moreover, the temperature response of the Honeywell Durafet is partly controlled by the dependence of the calibration constant with temperature (directly related to $E^{*}_{sensor,f}\,(T = 25℃)$) (Martz *et al.*, 2010). Because of this, the small degree

and nature of the variability characterizing the differences between values of $\mathrm{pH}_{\mathrm{final}}^{\mathrm{sensor}}$ calibrated using $\mathrm{pH}_{\mathrm{DIC-TA}}^{\mathrm{disc}}$ and $\mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}$ over time will depend on the relationship between $\mathrm{E}_{\mathrm{sensor,f}}^{*}\ (\mathrm{T} = 25℃)$ and temperature. These relationships are governed by $\mathrm{d}E_{INT}^{*}/\mathrm{d}T$ and $\mathrm{d}E_{EXT}^{*}/\mathrm{d}T$ in equations 24 and 26 for $\mathrm{pH}_{\mathrm{final}}^{\mathrm{INT}}$ and $\mathrm{pH}_{\mathrm{final}}^{\mathrm{EXT}}$, respectively. For the 01 June 2016 calibration (Figure 33), the difference between the values of $\mathrm{E}_{\mathrm{INT,f}}^{*}\ (\mathrm{T} = 25℃)$ was 0.0016 V, which translated into a difference in $\mathrm{pH}_{\mathrm{final}}^{\mathrm{INT}}$ of 0.0268-0.0281 pH units. For the 02 August 2016 calibration (Figure 34), the difference between the values of $\mathrm{E}_{\mathrm{INT,f}}^{*}\ (\mathrm{T} = 25℃)$ was 0.0026 V, which translated into a difference in $\mathrm{pH}_{\mathrm{final}}^{\mathrm{INT}}$ of 0.0429-0.0439 pH units. In both cases, the difference between the values of $\mathrm{pH}_{\mathrm{final}}^{\mathrm{INT}}$ matched up against a profile indicative of an inverse relationship with temperature. This adequately describes relationship between $\mathrm{E}_{\mathrm{INT}}^{*}$ and temperature characterized by Martz *et al.* (2010). Similar comparisons of $\mathrm{pH}_{\mathrm{final}}^{\mathrm{EXT}}$ calibrated using $\mathrm{pH}_{\mathrm{DIC-TA}}^{\mathrm{disc}}$ and $\mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}$ produced similar results since the equations used for $\mathrm{E}_{\mathrm{EXT,f}}^{*}\ (\mathrm{T} = 25℃)$ and $\mathrm{pH}_{\mathrm{final}}^{\mathrm{EXT}}$ are of an equivalent type and the same calibration protocols were employed.

2.4.2.4 Recommendations for the Treatment of Excess Alkalinity in Future Work

Our results show that existing calibration protocols for Durafet-based biogeochemical sensors cannot be used to detect, quantify, and correct for contributions of excess alkalinity to sensor measured pH values in estuarine and coastal ocean systems. Sampling for DOC and/or DOM and measuring additional

marine CO_2 system parameters (e.g. DIC or pCO_2) may be needed to gauge the magnitude of the impacts excess alkalinity has on pH measurements in these settings. If pH calculated from measured alkalinity and a second marine CO_2 system parameter serves as the primary reference pH, it is recommended that all values be corrected for the effects of excess alkalinity prior to performing the sensor calibration. If a strong correlation exists between excess alkalinity and another measured parameter such as DOC, DOM, or salinity measured with comparable frequency, such a correction should be feasible.

2.4.3 Electrode Conditioning in an Estuarine System

2.4.3.1 Electrode Conditioning at Beginning of Sensor Deployment

Conditioning periods at the beginning of sensor deployments can endure for a few hours to several days depending on the pre-deployment procedures undertaken to address electrode conditioning. If left unconditioned prior to deployment, the external reference electrode takes significantly longer to condition than the internal reference electrode due to the time needed for Br^- ions to replace Cl^- ions in the AgCl solid solution found in the Cl-ISE (Bresnahan *et al.*, 2014). To minimize these conditioning periods, electrodes should be stored in seawater, be continuously powered for 5-10 days prior to deployment, and should remain stored in seawater while the instrument is transported to the deployment site as well (Bresnahan *et al.*, 2014).

These recommended procedures were used leading up to the redeployment of the sensor on 09 May 2016. Filtered seawater taken from the lower Delaware Bay with

a salinity range of S=29-31 throughout the year was used as the conditioning medium. The seawater was stored inside the flow housing of the SeapHOx from 13 April 2016 to 09 May 2016 prior to its redeployment for a total of 27 days. The conditioning medium was replaced every 5 days prior to the sensor deployment on 09 May 2016. This was followed by a conditioning period of 74 hours at the beginning of the sensor deployment (Figure 35). At that time, the calculated $\Delta pH^{INT-EXT}$ anomaly returned to a consistently stable value of < ±0.01 pH units while the deployment site was inundated with the more saline Delaware Bay water for extended periods of time. Without an independent reference pH measured or calculated with a comparable frequency, the $\Delta pH^{INT-EXT}$ anomaly was found to be an appropriate metric with which to gauge the success of the electrode conditioning procedure.

The causes of this prolonged conditioning period were mostly likely attributed to the time needed to accomplish a stable flow of ions across the liquid junction of the internal reference electrode and the previous discussed replacement of Cl^- ions with Br^- ions in the solid AgCl solution of the external reference electrode as outlined in Bresnahan *et al.* (2014). With salinities ranging from 5.9 < S < 23.1, the implications of large changes in $[Br^-]$ and $[Cl^-]$ accompanying such large *in-situ* salinity fluctuations could result in a conditioning period of this length in an estuarine system. If the initial ion replacement process inside the Cl-ISE was slower than the tidal action, it is certainly possible that the conditioning period required for pH^{INT} was much shorter since the conditioning signal was dominated by pH^{EXT}. For this reason,

the development of a suitable pre-deployment electrode conditioning protocol for sensor deployments in estuarine and coastal ocean systems is needed.

2.4.3.2 Intra-Deployment Electrode Conditioning

At any time when the electrodes are not immersed and allowed to dry out, a conditioning period of varying length will always be required for the electrodes to recondition and recover their performance before dependable sensor pH measurements are again made (Bresnahan *et al.*, 2014). The implications of these shorter, more numerous conditioning periods on data loss can be severe when regularly scheduled sensor maintenance was performed when the sensor was out of the water for 1-3 hours. Throughout the course of present work, the sensor was redeployed with – (1) only air in the flow housing or (2) filtered seawater stored in the flow housing – to determine the time required for the electrodes to recondition after each sensor maintenance trip. This was done by comparing sensor pH and pH_T calculated from measured DIC and TA (pH^{Cond}_{DIC-TA}) of discrete bottle samples collected following sensor maintenance trips.

The results from 20 July 2016 (Figure 36) following the sensor redeployment with filtered seawater stored in the flow housing and from 12 August 2016 (Figure 37) with only air stored in the flow housing were both promising. On 20 July 2016, discrete bottle samples were collected up to 6.5 hours after the first sensor measurement at 1200 at the end of a flood tide until the start of the following flood tide. In contrast, on 12 August 2016, discrete bottle samples were collected up to 6

hours after the first sensor measurement at 1130 (end of the ebb tide) until the point in time where the deployment site was inundated with the higher salinity Delaware Bay water. In both cases, a near-zero $\Delta pH^{Sensor-Cond}$ anomaly was achieved for both reference electrodes. Because of this, it can reasonably be assumed that under either condition, the reference electrodes adequately recondition within 6 hours of the first sampling cycle after sensor redeployment following sensor maintenance trips. Accordingly, a nominal 6-hour reconditioning period was assumed for all May 2016 to August 2016 data used in subsequent analyses for the present work.

Given that near-zero $\Delta pH^{Sensor-Cond}$ anomalies were achieved within 3.5 hours of the first sensor measurement after redeployment under both conditions, the shorter conditioning periods most likely arise from the fact the sensors simply needed to be reimmersed and made "wet". Incidentally, the departure from zero $\Delta pH^{Sensor-Cond}$ anomalies demonstrated the difficulties of constraining intra-deployment conditioning periods with respect to each reference electrode in a dynamic estuarine environment. These difficulties relate to the challenge of collecting quality discrete bottle samples over changing tidal conditions and the presence of the salinity lag addressed in Section 2.4.1.2. These intra-deployment conditioning periods are most likely heavily dependent on local temperature and salinity conditions at the deployment site at different times of the year. Hence, this would make their length site-dependent and deployment-specific. It is recommended that filtered seawater or natural waters taken from the deployment site be stored in the flow housing of a SeapHOx before its redeployment following sensor maintenance. This is

recommended because a narrower range for the $\Delta pH^{Sensor-Cond}$ anomalies were yielded under this condition and it conforms with existing recommended electrode conditioning protocol (Bresnahan *et al.*, 2014).

2.4.4 Recommendations

2.4.4.1 Choice of Independent Reference pH

2.4.4.1.1 Murderkill Estuary-Delaware Bay Results

The property-property plots of pH_{final}^{sensor} vs. pH^{disc} emphasize strong agreement between the sensor pH and reference pH. Our results show that pH_{elec}^{disc} provided a suitable means of calibrating raw sensor output. This is supported by the sensor gains and offsets describing their comparisons yielding deviations less than or equal to their counterparts calibrated using pH_{DIC-TA}^{disc} from 1 and 0, respectively. This indicates that comparable or reduced bias was present in comparisons made with pH_{elec}^{disc} versus pH_{DIC-TA}^{disc}. Calibrating sensor output against pH_{elec}^{disc} worked to a disproportionately better effect on pH_{final}^{EXT} than pH_{final}^{INT}. This is attributed to the fact that the ΔpH_{scales} decreases with decreasing salinity, so in effect, both pH^{EXT} and ΔpH_{scales} possessed some degree of salinity sensitivity.

Elevated values of the calculated sensor offsets yielded from Model II least squares fits of pH_{final}^{sensor} vs. pH^{disc} often make them more abstract and difficult to directly compare. Instead, property-property plots between the perturbations in pH^{disc} and pH_{final}^{sensor} from the minimum value of pH^{disc} characterized by perturbation

variables of $\mathrm{pH}^{\mathrm{disc}\prime} = \mathrm{pH}^{\mathrm{disc}} - \min(\mathrm{pH}^{\mathrm{disc}})$ and $\mathrm{pH}_{\mathrm{final}}^{\mathrm{sensor}\prime} = \mathrm{pH}_{\mathrm{final}}^{\mathrm{sensor}} - \min(\mathrm{pH}^{\mathrm{disc}})$ put Model II fit parameters in a more realistic context in which the yielded sensor gains (c_1 – slopes) and offsets (c_0 – intercepts) can be better investigated (Bresnahan *et al.*, 2016). Such property-property plots (Figures 38 and 39) produce Model II fit sensor offsets much closer to 0 (Tables 6 and 7). Consistent with Bresnahan *et al.* (2016), comparisons of the results of the present work were made against the results of evaluations of seven widely-used pH sensors by the Alliance for Coastal Technologies (ACT, 2012: http://www.act-us.info/evaluations.php) completed using pH perturbation property-property plots. The ACT results produced standard deviations in $\mathrm{pH}^{\mathrm{sensor}} - \mathrm{pH}^{\mathrm{disc}}$ anomalies of 0.01-0.1 pH units under *in-situ* conditions. In the case of the present work, good agreement between all pairs of $\mathrm{pH}_{\mathrm{final}}^{\mathrm{sensor}}$ and $\mathrm{pH}^{\mathrm{disc}}$ values were demonstrated and similar sensor validation protocol yielded standard deviations or root mean squared errors (RMSE) of 0.01-0.04 pH units across both reference electrodes over a range of environmental conditions. This demonstrates the viability of the Honeywell Durafet in the collection of dependable high-frequency, high-resolution, high-precision pH data over periods of weeks to months in dynamic, productive, high-fouling, highly-turbid estuarine and coastal ocean systems regardless of the independent reference pH used.

2.4.4.1.2 Practical Considerations

Our results indicate that both pH^{disc}_{DIC-TA} and pH^{disc}_{elec} provided an adequate means of calibrating raw sensor output. Both were used to help confirm aspects of electrode response at higher salinities (S > 20) characterized from previous work (Bresnahan *et al.*, 2014; Bresnahan *et al.*, 2016; Martz *et al.*, 2010; Takeshita *et al.*, 2014) and to describe electrode performance in an estuarine system. Moreover, our results of pH^{INT}_{final} also demonstrate the Honeywell Durafet continues to exhibit near-ideal Nernstian response to at least S=8.99 well below the S=20 threshold characterized by Takeshita *et al.* (2014). In addition, the elevated particle loads characterizing the Murderkill Estuary-Delaware Bay System (Ullman *et al.*, 2013) do not appear to exert a significant and measurable influence on instantaneous sensor measurements given the similar agreement of pH^{sensor}_{final} with pH^{disc}_{DIC-TA} (filtered) and pH^{disc}_{elec} (unfiltered). This may also need to be studied further. It is difficult to determine long-term effects of turbidity on electrode response without comparisons against an independent reference pH measured over a similar frequency. In this deployment environment, it is known that the total suspended solids concentrations tend to be higher under stronger tidal currents, over prolonged periods of strong winds, and following strong localized storm events (Voynova *et al.*, 2015). So, any long-term effects of turbidity on electrode performance may also be subject to similar controls. However, turbidity appears to not significantly affect the relative trends in pH captured over the course of the present work.

The calibration of raw sensor output to pH calculated from measured DIC and TA at *in-situ* conditions has proven to work in the past (Bresnahan *et al.*, 2016; Rérolle *et al.*, 2016). However, the potential overestimation of pH_{DIC-TA}^{disc} due to contributions of excess alkalinity in estuarine and coastal ocean deployment environments cannot be ignored. Sets of independent reference pH generated from pH calculated from two other measured marine CO_2 system parameters can be used to calibrate a Honeywell Durafet in the future. This approach also offers the advantage of providing the sensor operators with a better understanding of the underlying controls on the marine CO_2 system in a specific environment (Macleod *et al.*, 2015) via its constraint and its variability over a range of different environmental conditions (Bresnahan *et al.*, 2016). Employing the calculation of pH_T from measured marine CO_2 system parameters as a means of calibration also allows the sensor operators to maximize the number of parameter combinations used to calculate an independent reference pH, to better constrain electrode response, and to improve the quality of the calibrations. While this is beneficial, it is not always as practical given the labor-intensive requirement that such a calibration presents (Bresnahan *et al.*, 2016). In the case of the present sensor deployment, a total of 486 discrete bottle samples for DIC and TA were collected on more than 20 sampling days throughout the year. These efforts produced 255 calibration points over the ten-month sensor deployment from September 2015 to August 2016. This translated into over 125 person hours attributed to sampling alone and does include time dedicated to other concomitant tasks. This sort of comprehensive sensor calibration method is not always possible.

In contrast, since marine CO_2 system parameters can be measured more accurately than they can be calculated, calibrating a working Honeywell Durafet to direct measurements of pH is a better approach. For the present work, an independent reference pH traceable to direct measurements of pH_{NBS} made with a glass electrode calibrated with freshwater pH buffers was satisfactorily used to calibrate a Honeywell Durafet. While there will be uncertainty in any pH measurement associated with using a standard with a fixed salt concentration as a reference for all pH measurements, under the pretext that the pH_{NBS} measurements are free of electrode drift and biofouling, their correction to pH_T at *in-situ* conditions substantially reduces this uncertainty. Given the well-documented problems associated with using freshwater pH buffers in seawater pH applications, this calibration method has been largely overlooked. More importantly, the low cost and wide availability of freshwater pH buffers provide a cost-effective and straightforward, yet robust calibration method for future sensor deployments in estuarine and coastal ocean systems.

The calibration method chosen for a sensor deployment in an estuarine or coastal ocean system is subject an array of practical considerations. These considerations can be heavily deployment-specific and governed by the deployment site chosen (e.g. accessibility), resources available to the sensor operators (e.g. funding or instrumentation), available manpower, and time that can be spent on the different phases of the sensor deployment. To assist with the planning and execution of the present sensor deployments, the set of published Best Practices for the use of Durafet-based biogeochemical sensors outlined in Bresnahan *et al.* (2014) was indispensable,

and the workflow plan for pH sensor deployments detailed in Rivest *et al.* (2016) should further streamline them in the future.

Since the independent reference pH used in the present work was collected over a much lower frequency, it is imperative that future sensor deployments in estuarine and coastal ocean systems include the deployment of co-located, independent sensors used to measure an additional marine CO_2 system parameter (e.g. pCO_2) and/or utilize regional empirical marine CO_2 system relationships (e.g. Alin *et al.* (2012)) for the southern California Current System) as recommended by Bresnahan *et al.* (2014). This would generate the set of high-frequency independent reference pH that is needed to examine the long-term trends in electrode response in this new type of environment. The present work clearly indicates that the employment of multiple sets of independent reference pH are a necessary tool for distinguishing between bias in sensor pH and reference pH in-line with the recommendations of Bresnahan *et al.* (2014). The choice of calibration method does not have not have a universal answer, but all aspects of future sensor deployments must carefully be considered before deciding on the type and breadth of calibration method to be used and the design of the discrete sampling regiment with which to carry it out.

2.4.4.2 Modifications to SeapHOx Design

As discussed previously in Section 2.2.2, the open-ocean design of the SeapHOx was not practical in a dynamic, productive, high-fouling, highly-turbid estuarine system. As designed, the SeapHOx was not capable of dependable high-

resolution, high-frequency pH measurements over periods of days to weeks between sensor maintenance trips to validate its application to such an environment that was needed for the present work. A second SeapHOx unit modified for use in this type of deployment environment was successfully deployed for a period of about 10 months between September 2015 to August 2016. The results of the present work yielded a comprehensive set of recommendations to be considered for a subsequent sensor design specific to estuarine and coastal ocean deployment environments.

The effects of high turbidity have the potential to introduce bias not attributed to degraded electrode response into raw sensor output and subsequently calibrated sensor time series. This may lead to the flagging and exclusion of large numbers of sensor measurements after undergoing routine QC procedures. The magnitude of data loss attributed to high turbidity increases significantly in the absence of regularly scheduled sensor maintenance. In the present work, the degradation of pump performance over time was experienced due to the formation of choke points in the instrument flow path especially around areas where the cross-sectional area or the direction of the flow path changed. These features lead to progressively shorter periods of slower, weaker water flow through the instrument.

By employing the flow path modifications discussed in Section 2.2.4, the periods of time characterized by full pump operation and strong water flow through the instrument is significantly extended. So, employing similar modifications on SeapHOx flow paths is recommended for their future use in turbid waters. Even with a modified flow path, particles were still routinely observed inside the impeller chamber

of SBE 5M submersible pump. The presence of these particles around the plane of impeller rotation can be especially problematic. Given the variable pump performance experienced in this sensor deployment, it is recommended that SeapHOx units designed for use in estuaries and the coastal ocean be equipped with a stronger pump more adept to pumping waters with high particle loads. Ultimately, the Sea-Bird Electronics (SBE) 5P submersible pump (Sea-Bird Scientific, Bellevue, WA) should prove well suited to this purpose if the stronger, faster flow through the flow housing of the SeapHOx does not compromise of the integrity of the sensing surfaces.

Currently, the communication with any SeapHOx unit is limited by way of a direct RS-232 connection with a computer through a suitable terminal program such as HyperTerminal or Tera Term. On the other hand, the recently developed and tested mobile oceanographic sensing platform incorporating the dual-reference electrode design integrated into the SeapHOx known as the WavepHOx described by Bresnahan *et al.* (2016) achieved methods of wireless communication between the instrument and the sensor operator using an iPhone/iPad. Moving forward, it is recommended that subsequently designed SeapHOx units for use in deployment environments in close enough proximity to land (typical most potential estuarine and coastal ocean environments) within range of telecommunication networks (e.g., 4G/3G networks) and/or WiFi networks adept at facilitating remote communication between sensor and sensor operator incorporate wireless communication capabilities as well. With respect to the present work, wireless communications to help stop and start deployment loops, to cycle the sensor in and out of sleep mode, and to download sensor data remotely

would have been preferred over removing the sensor from the water each time and manually performing all these tasks. The ability to download data remotely would also minimize data loss attributed to intra-deployment conditioning that occurred each time after sensor removal. Wireless communications would allow the sensor operator increased capability to monitor aggregate sensor performance throughout a deployment, thus greatly improving the chances of success in challenging deployment environments like the one used in the present work.

The current open-ocean design of the SeapHOx is characterized by the absence of technology that allows the sending of data to and the viewing of data by the sensor operators, PIs, and stakeholders remotely in real time. There exist data loggers and communications systems to improve the capability to monitor underwater instruments. The STOR-X Submersible Data Logger (Satlantic L.P., Halifax, Nova Scotia, Canada), integrated into the co-located suite of biogeochemical sensors of the Kent County Land-Ocean Biogeochemical Observatory (http://kentcounty.loboviz.com/), is capable of and has proven extremely useful in monitoring and maintaining underwater instruments during long-term deployments in the same deployment environment (Dr. William Ullman, personal communication). The benefits of action with respect to these tasks greatly outweighs the costs of inaction as the focuses of ocean acidification and marine CO_2 chemistry research further expand into the less well-characterized dynamic estuarine and coastal ocean systems (Andersson *et al.*, 2015; Cai *et al.*, 2011; Duarte *et al.*, 2013; Weisberg *et al.*, 2016).

2.4.5 Sensor Redundancy

Related work with the Honeywell Durafet over comparable timescales carried out under rigorously-controlled laboratory conditions established small degrees of characteristic inter-sensor variability in key parameters associated with sensor operation (e.g. $\mathrm{d}E^*_{INT}/\mathrm{d}T$, $\mathrm{d}E^*_{EXT}/\mathrm{d}T$) (Martz *et al.*, 2010). Because of this, sensor redundancy (the use of multiple units) was a key feature in subsequent work (Bresnahan *et al.*, 2014; Takeshita *et al.*, 2014). For the present work, one SeapHOx unit with one Honeywell Durafet was used in a highly variable estuarine system. Further work is needed to properly constrain and interpret electrode response under the conditions experienced over the present sensor deployment to sufficiently substantiate its results.

Chapter 3

SUMMARY & CONCLUSIONS

Pilot deployments of the SeapHOx sensor equipped with the Honeywell Durafet were carried out between April 2015 to August 2015 and September 2015 to August 2016 in the Murderkill Estuary-Delaware Bay System (Delaware, USA). The present work encompassed the operation of a Honeywell Durafet in a dynamic, productive, high-fouling, highly-turbid estuarine system. In this setting, salinities lower than the characteristic open-ocean ranges (S=30-36) under which sensor performance had previously been investigated (Bresnahan *et al.*, 2014) ranging between S=3.25 and S=29.33 were consistently experienced. Sensor performance was also evaluated over the dynamic salinity conditions on tidal excursions that characterize this system. The sensor pH collected during May 2016 to August 2016 using the furthest refined SeapHOx configuration and discrete sampling strategy exhibited good agreement with the independent reference pH used. The sensor pH was characterized by a root-mean squared error (RMSE) ranging between 0.011 and 0.036 pH units across the full environmental salinity range relative to both pH_T calculated from measured DIC and TA and pH_{NBS} corrected to pH_T at *in-situ* conditions. Despite the challenges of operating a SeapHOx in an estuarine system, the results were very promising.

In this environment, the Honeywell Durafet was capable of consistently capturing natural pH fluctuations nominally ranging from <0.5 pH units to >1 pH unit

occurring in the surrounding waters across the more saline Delaware Bay water and the fresher Murderkill Estuary outflow endmembers mixed along the salinity gradient over a range of timescales. Sensor performance demonstrated the versatility of the Honeywell Durafet and reinforced its viability to the measurement of pH as a part of future estuarine and coastal ocean CO_2 chemistry studies. A number of deficiencies in existing deployment guidelines and calibration protocol for Durafet-based biogeochemical sensors were identified and highlighted in this work. Aspects of electrode response requiring further investigation were also highlighted. Moreover, a comprehensive set of recommendations for the future utilization of these sensors in similar systems resulted from the present work. As the emerging parallel trends in seawater pH metrology for natural waters of $S < 20$ converge, the highly desirable accuracy of the Honeywell Durafet achieved under open-ocean conditions of better than 0.01 pH units should be attainable in estuarine and coastal ocean systems in the near-future as well.

TABLES

Table 1. Deployment Details.

Site	Regime	Lat.	Lon.	Deployment Dates	Sensor Packages	SeapHOx Build	Validation Approaches
Murderkill Estuary-Delaware Bay System (Bowers Beach, DE)	Flood-Dominant Estuarine	39.05 ^{o}N	75.39 ^{o}W	08 April 2015 to 26 Aug 2015	SeapHOx SP033, LOBO Sensors	v1.0	Not Applicable due to Sensor Failure
Murderkill Estuary-Delaware Bay System (Bowers Beach, DE)	Flood-Dominant Estuarine	39.05 ^{o}N	75.39 ^{o}W	17 Sep 2015 to 09 Dec 2015	SeapHOx SP053, LOBO Sensors	v2.0	DIC-TA (226 samples) pH_{NBS} (214 samples)
Murderkill Estuary-Delaware Bay System (Bowers Beach, DE)	Flood-Dominant Estuarine	39.05 ^{o}N	75.39 ^{o}W	11 Dec 2015 to 04 April 2016	SeapHOx SP053	v2.0	DIC-TA (116 samples) pH_{NBS} (108 samples)
Murderkill Estuary-Delaware Bay System (Bowers Beach, DE)	Flood-Dominant Estuarine	39.05 ^{o}N	75.39 ^{o}W	09 May 2016 to 24 Aug 2016	SeapHOx SP053, LOBO Sensors	v3.0	DIC-TA (125 samples) pH_{NBS} (138 samples)

Table 2. Independent reference pH, salinity constraints, and alphanumeric designations corresponding to each sensor calibration method for each reference electrode on each sampling day.

Sampling Day	Reference Electrode	Independent Reference pH	Salinity Constraints	Alphanumeric Designation
01 June 2016	Internal	pH^{disc}_{DIC-TA}	None	1
01 June 2016	External	pH^{disc}_{DIC-TA}	None	1A
01 June 2016	External	pH^{disc}_{DIC-TA}	Only S>20	1B
01 June 2016	Internal	pH^{disc}_{elec}	None	2
01 June 2016	External	pH^{disc}_{elec}	None	2A
01 June 2016	External	pH^{disc}_{elec}	Only S>20	2B
02 August 2016	Internal	pH^{disc}_{DIC-TA}	None	1
02 August 2016	External	pH^{disc}_{DIC-TA}	None	1
02 August 2016	Internal	pH^{disc}_{elec}	None	2
02 August 2016	External	pH^{disc}_{elec}	None	2

Table 3. Average values of final calibration constants ($E^{*}_{sensor,f}$ ($T = 25℃$)) applied to all raw sensor time series.

Sampling Day	Reference Electrode	Calibration Method	Salinity Constraints	Final Calibration Constant
01 June 2016	Internal	1	None	-0.4369 ± 0.0017
01 June 2016	External	1A	None	-1.4080 ± 0.0013
01 June 2016	External	1B	Only S>20	-1.4075 ± 0.0010
01 June 2016	Internal	2	None	-0.4353 ± 0.0022
01 June 2016	External	2A	None	-1.4064 ± 0.0007
01 June 2016	External	2B	Only S>20	-1.4064 ± 0.0008
02 August 2016	Internal	1	None	-0.4400 ± 0.0011
02 August 2016	External	1	None	-1.4109 ± 0.0008
02 August 2016	Internal	2	None	-0.4374 ± 0.0012
02 August 2016	External	2	None	-1.4083 ± 0.0007

Table 4. Root-mean squared error (RMSE), sensor offset (c_0), and sensor gain (c_1) calculated from Model II least squares fits for all final sensor pH and reference pH comparisons from the 01 June 2016 sampling day.

Sampling Day	Reference Electrode	Calibration Method	Salinity Constraints	RMSE	c_0 (intercept)	c_1 (slope)
01 June 2016	Internal	1	None	0.0275	-0.1795±0.1683	1.0225±0.0210
01 June 2016	External	1A	None	0.0174	-0.5848±0.1741	1.0724±0.0215
01 June 2016	External	1B	Only S>20	0.0158	0.0882±0.3860	0.9891±0.0473
01 June 2016	Internal	2	None	0.0358	-0.0784±0.1592	1.0101±0.0202
01 June 2016	External	2A	None	0.0114	-0.0308±0.1063	1.0039±0.0132
01 June 2016	External	2B	Only S>20	0.0123	-0.1557±0.3258	1.0191±0.0399

Table 5. Root-mean squared error (RMSE), sensor offset (c_0), and sensor gain (c_1) calculated from Model II least squares fits for all final sensor pH and reference pH comparisons from the 02 August 2016 sampling day.

Sampling Day	Reference Electrode	Calibration Method	Salinity Constraints	RMSE	c_0 (intercept)	c_1 (slope)
02 August 2016	Internal	1	None	0.0159	0.2281±0.1072	0.9698±0.0142
02 August 2016	External	1	None	0.0121	-0.0892±0.0814	1.0119±0.0108
02 August 2016	Internal	2	None	0.0149	0.3655±0.0982	0.9514±0.0131
02 August 2016	External	2	None	0.0105	0.0560±0.0689	0.9927±0.0092

Table 6. Root-mean squared error (RMSE), sensor offset (c_0), and sensor gain (c_1) calculated from Model II least squares fits for $pH^{sensor'}$ vs. $pH^{disc'}$ property-property comparisons from the 01 June 2016 sampling day.

Sampling Day	Reference Electrode	Calibration Method	Salinity Constraints	RMSE	c_0 (intercept)	c_1 (slope)
01 June 2016	Internal	1	None	0.0275	-0.0160±0.0171	1.0225±0.0210
01 June 2016	External	1A	None	0.0174	-0.0405±0.0129	1.0724±0.0215
01 June 2016	External	1B	Only S>20	0.0158	0.0006±0.0076	0.9891±0.0473
01 June 2016	Internal	2	None	0.0358	-0.0068±0.0175	1.0101±0.0202
01 June 2016	External	2A	None	0.0114	-0.0018±0.0085	1.0039±0.0132
01 June 2016	External	2B	Only S>20	0.0123	-0.0025±0.0071	1.0191±0.0399

Table 7. Root-mean squared error (RMSE), sensor offset (c_0), and sensor gain (c_1) calculated from Model II least squares fits for $pH^{sensor'}$ vs. $pH^{disc'}$ property-property comparisons from the 02 August 2016 sampling day.

Sampling Day	Reference Electrode	Calibration Method	Salinity Constraints	RMSE	c_0 (intercept)	c_1 (slope)
02 August 2016	Internal	1	None	0.0159	0.0148±0.0076	0.9698±0.0142
02 August 2016	External	1	None	0.0121	-0.0051±0.0058	1.0019±0.0108
02 August 2016	Internal	2	None	0.0149	0.0244±0.0072	0.9514±0.0131
02 August 2016	External	2	None	0.0105	0.0045±0.0050	0.9927±0.0092

FIGURES

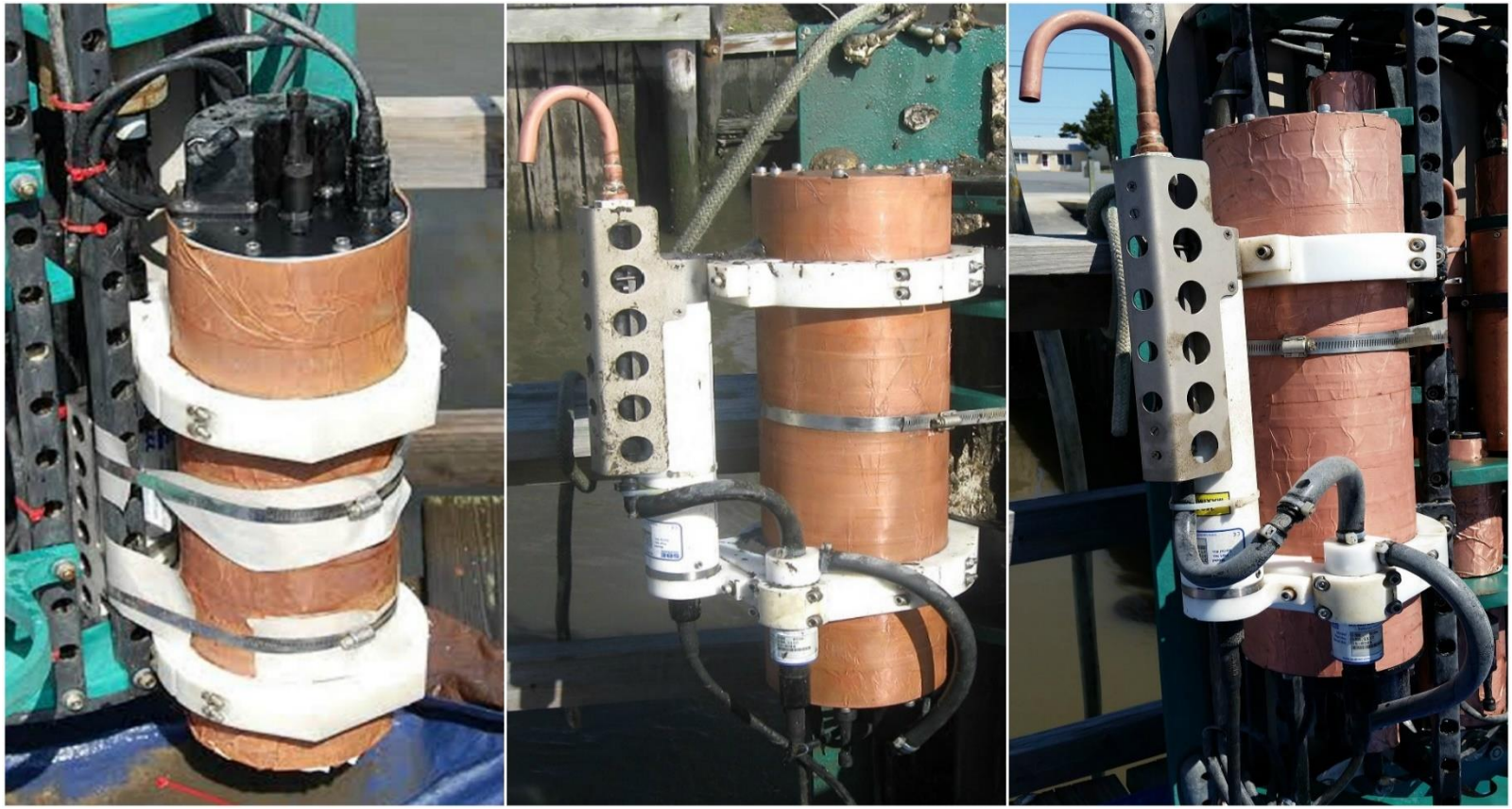

Figure 1. Different SeapHOx configurations used over the course of the project: Left Panel – v1.0, Middle Panel – v2.0, and Right Panel – v3.0.

Figure 2. Example (Top Left Panel) and extent (Top Right Panel) of sediment accumulation inside SeapHOx flow housing experienced over 2 weeks during the 02 April 2016 to 26 August 2016 SeapHOx deployment. Unsuccessful preventive measures of outfitting flow housing with multiple outflow points (Bottom Left Panel) and employing a hydrodynamic filter stuffed with fiberglass mesh (Bottom Right Panel) to combat sediment accumulation inside SeapHOx flow housing.

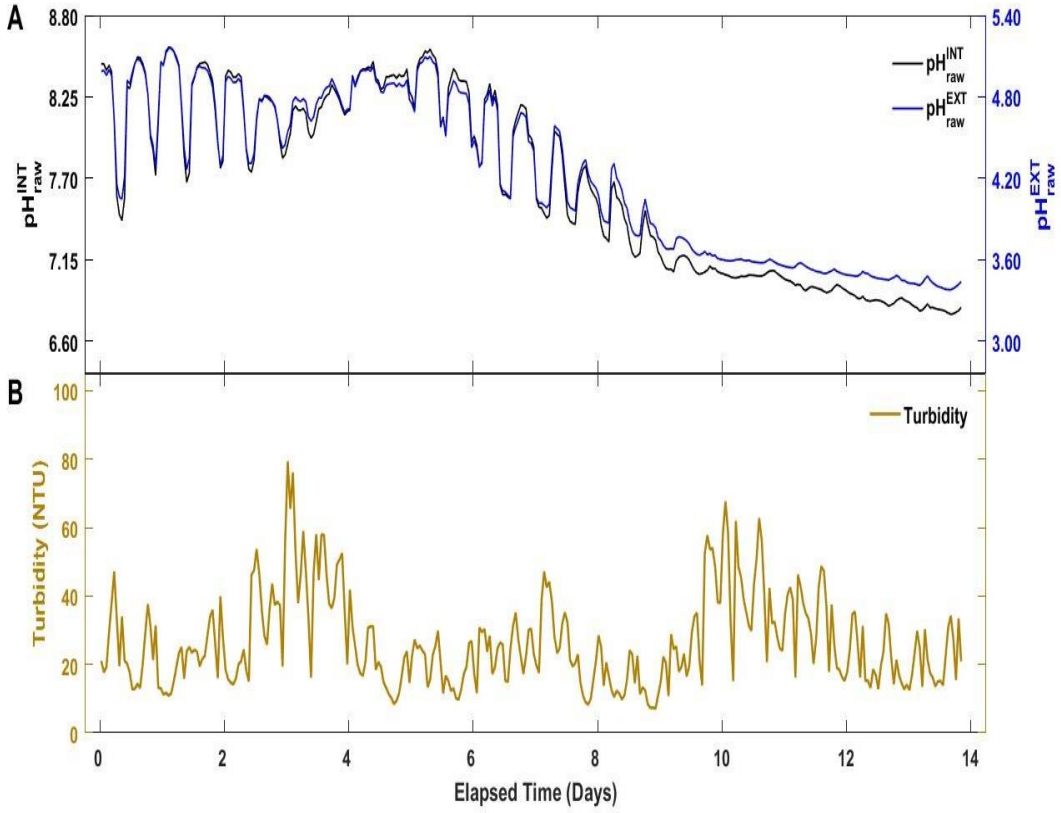

Figure 3. Measured (A) pH_{raw}^{INT} (solid black line) and pH_{raw}^{EXT} (solid blue line) during periods of sediment accumulation inside SeapHOx flow housing and (B) turbidity (solid gold line) from 21 April 2015 to 05 May 2015.

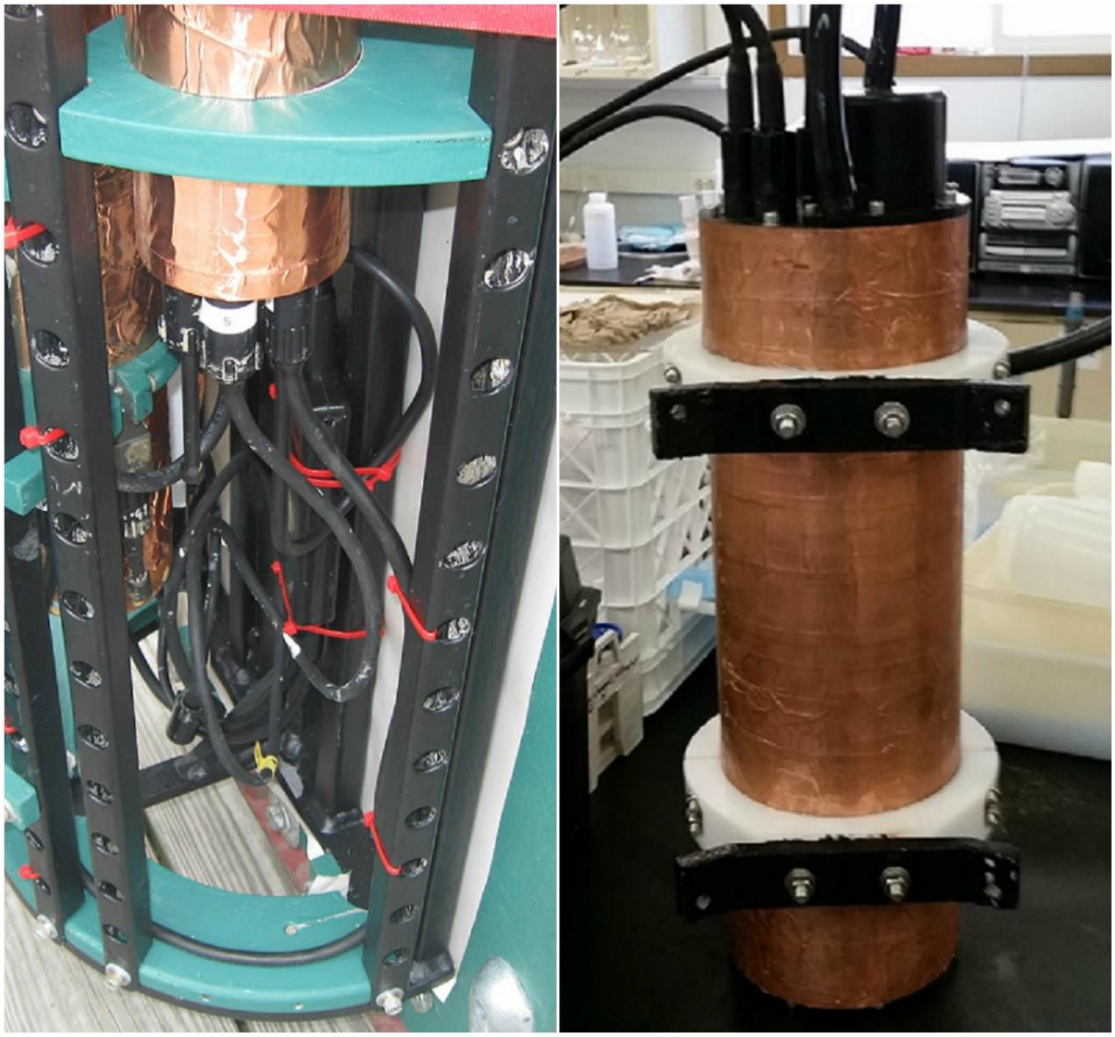

Figure 4. Overview of SeapHOx field deployment method for September 2015 to August 2016 SeapHOx deployments.

Figure 5. Examples of biofouling observed within 1-2 weeks experienced during SeapHOx deployments from summer 2016 (Left Panel) and summer 2015 (Right Panel).

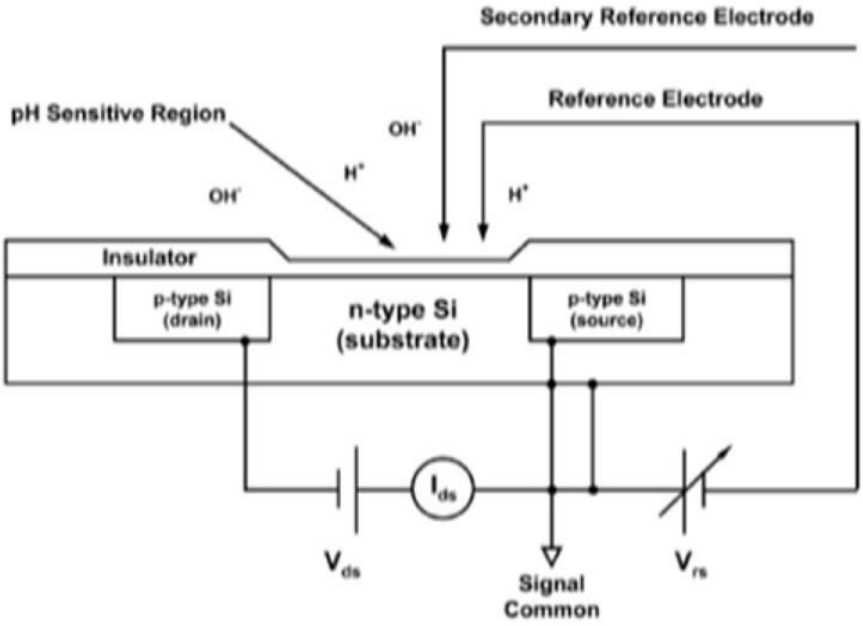

Figure 6. Functional implementation of the ISFET operating principle (from Martz *et al.*, 2010).

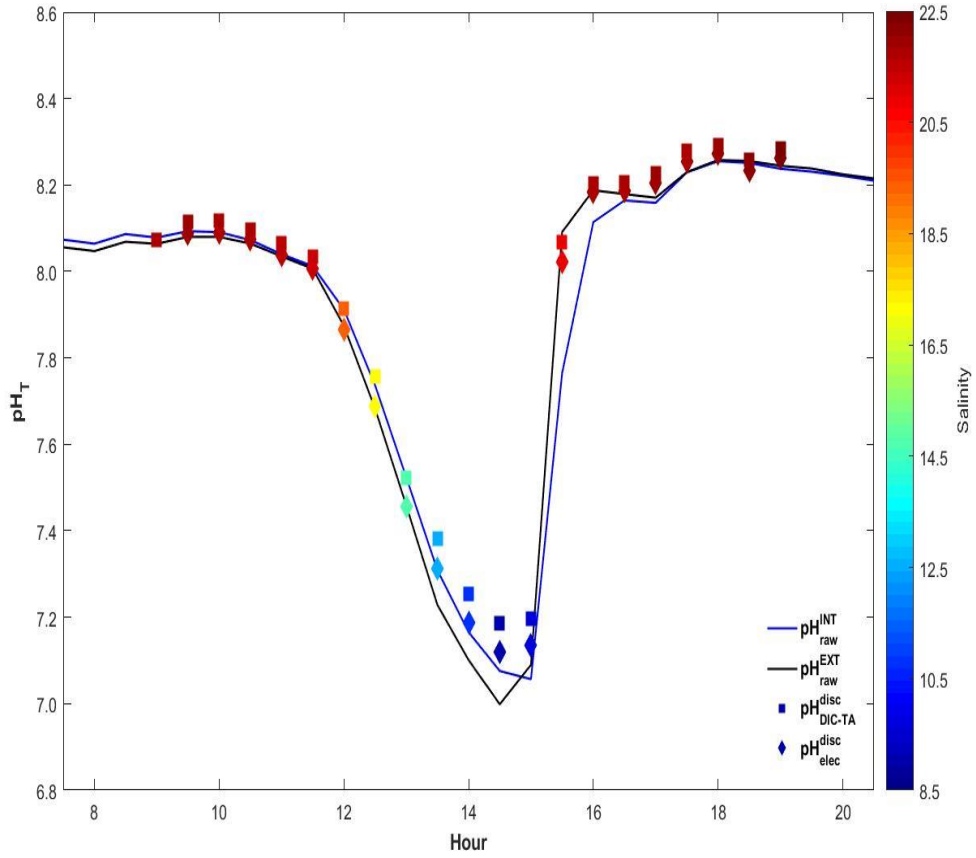

Figure 7. pH_{raw}^{INT} (solid blue line), pH_{raw}^{EXT} (solid black line), calculated pH_{DIC-TA}^{disc} (squares), and calculated pH_{elec}^{disc} (diamonds) on the total hydrogen ion concentration scale (pH_T) as a function of salinity from 0900-1900 on the 01 June 2016 sampling day.

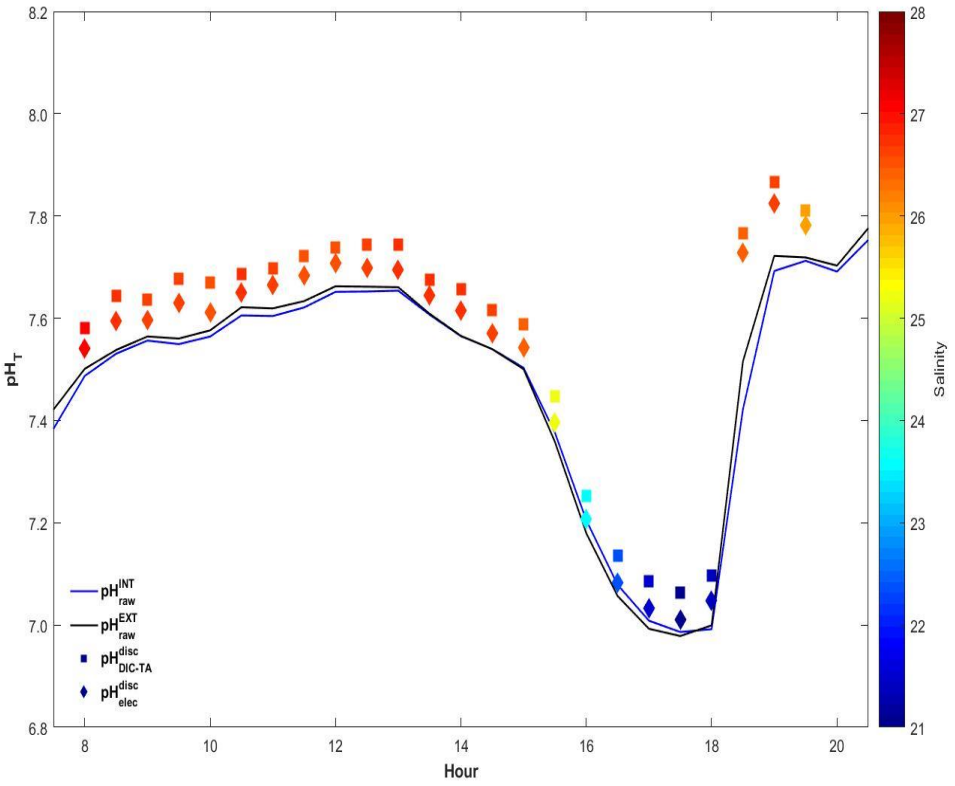

Figure 8. pH_{raw}^{INT} (solid blue line), pH_{raw}^{EXT} (solid black line), calculated pH_{DIC-TA}^{disc} (squares), and calculated pH_{elec}^{disc} (diamonds) on the total hydrogen ion concentration scale (pH_T) as a function of salinity from 0800-1930 on the 02 August 2016 sampling day.

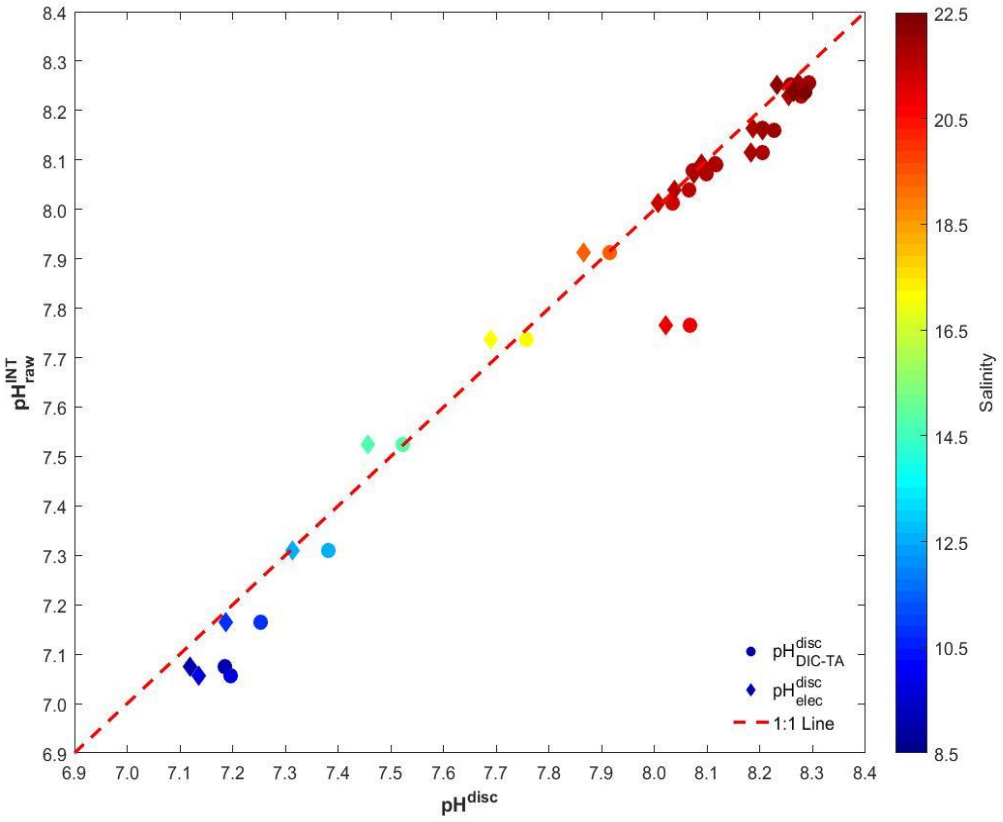

Figure 9. Property-property plot of $\mathrm{pH}_{\mathrm{raw}}^{\mathrm{INT}}$ vs. $\mathrm{pH}_{\mathrm{DIC-TA}}^{\mathrm{disc}}$ (circles) and $\mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}$ (diamonds) as a function of salinity from the 01 June 2016 sampling day shown relative to a 1:1 $\left(\mathrm{pH}^{\mathrm{INT}} = \mathrm{pH}^{\mathrm{disc}}\right)$ relationship (dashed red line).

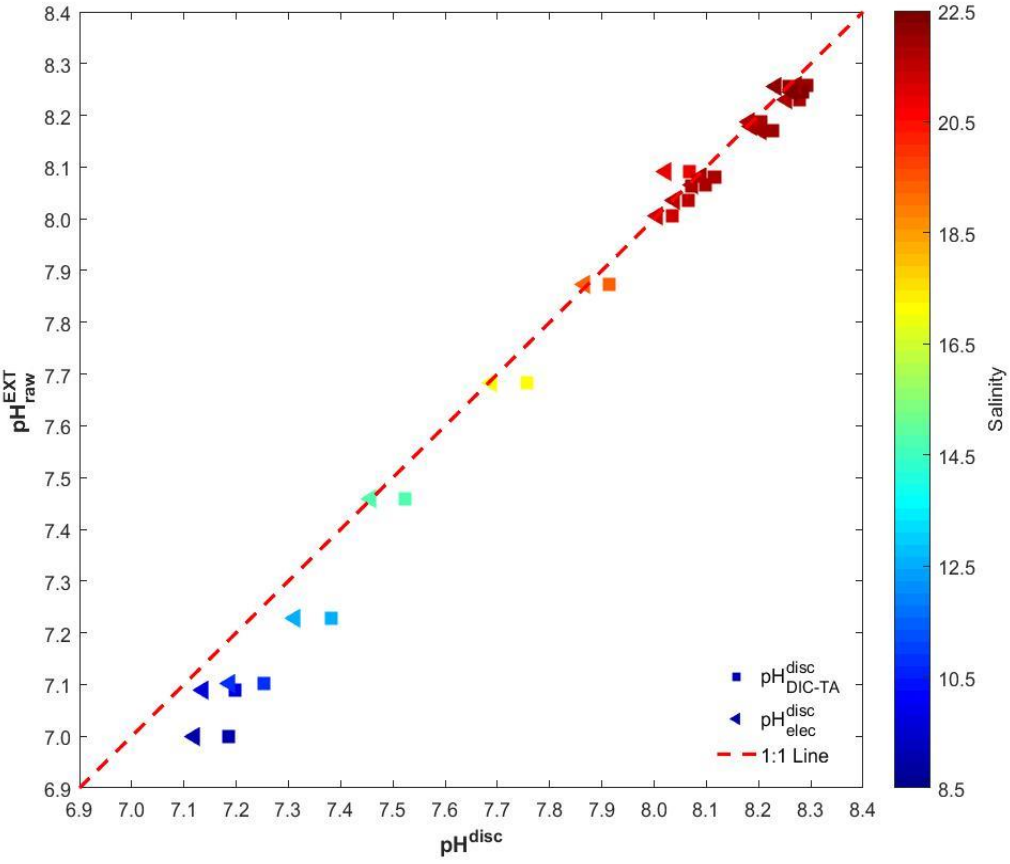

Figure 10. Property-property plot of $\mathrm{pH}_{\mathrm{raw}}^{\mathrm{EXT}}$ vs. $\mathrm{pH}_{\mathrm{DIC-TA}}^{\mathrm{disc}}$ (squares) and $\mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}$ (triangles) as a function of salinity from the 01 June 2016 sampling day shown relative to a 1:1 $\left(\mathrm{pH}^{\mathrm{EXT}} = \mathrm{pH}^{\mathrm{disc}}\right)$ relationship (dashed red line).

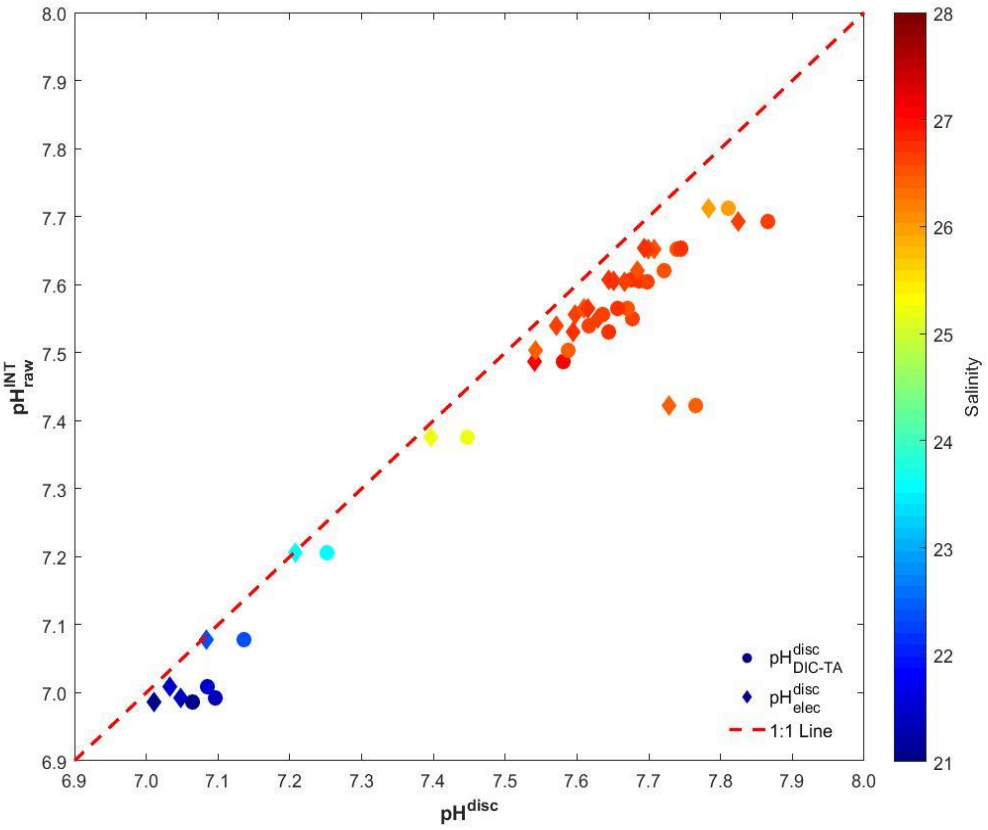

Figure 11. Property-property plot of $\mathrm{pH}_{\mathrm{raw}}^{\mathrm{INT}}$ vs. $\mathrm{pH}_{\mathrm{DIC-TA}}^{\mathrm{disc}}$ (circles) and $\mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}$ (diamonds) as a function of salinity from the 02 August 2016 sampling day shown relative to a 1:1 $\left(\mathrm{pH}^{\mathrm{INT}} = \mathrm{pH}^{\mathrm{disc}}\right)$ relationship (dashed red line).

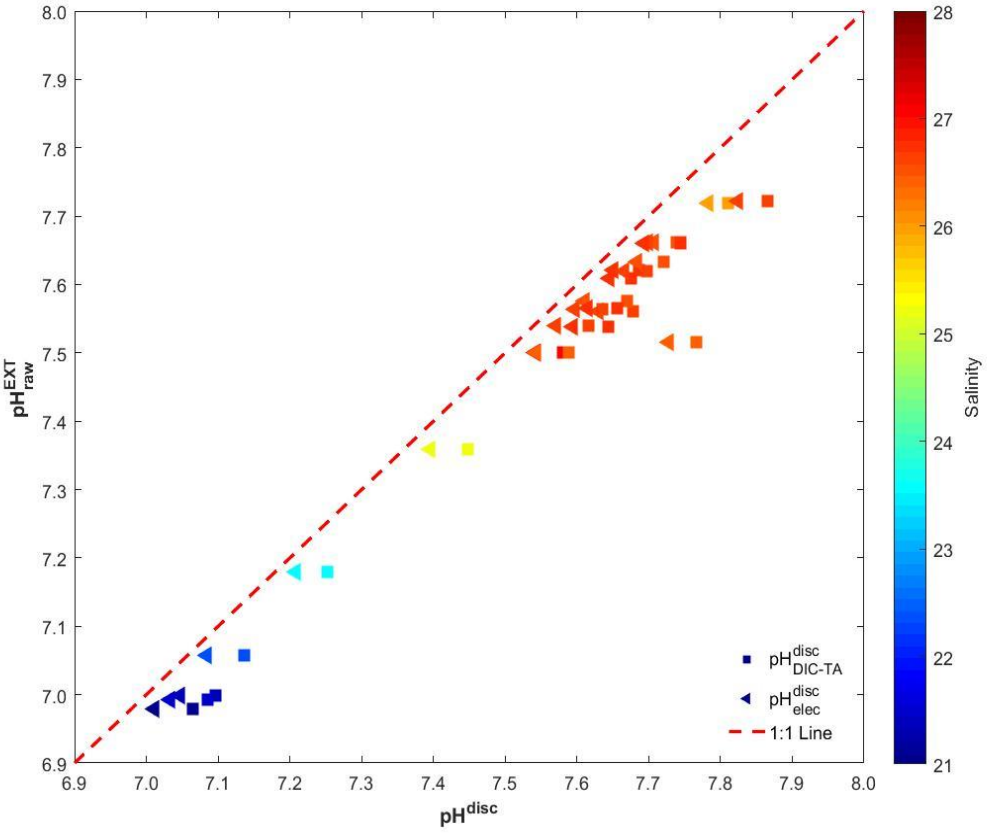

Figure 12. Property-property plot of $\mathrm{pH}_{\mathrm{raw}}^{\mathrm{EXT}}$ vs. $\mathrm{pH}_{\mathrm{DIC-TA}}^{\mathrm{disc}}$ (squares) and $\mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}$ (triangles) as a function of salinity from the 02 August 2016 sampling day. shown relative to a 1:1 $\left(\mathrm{pH}^{\mathrm{EXT}} = \mathrm{pH}^{\mathrm{disc}}\right)$ relationship (dashed red line).

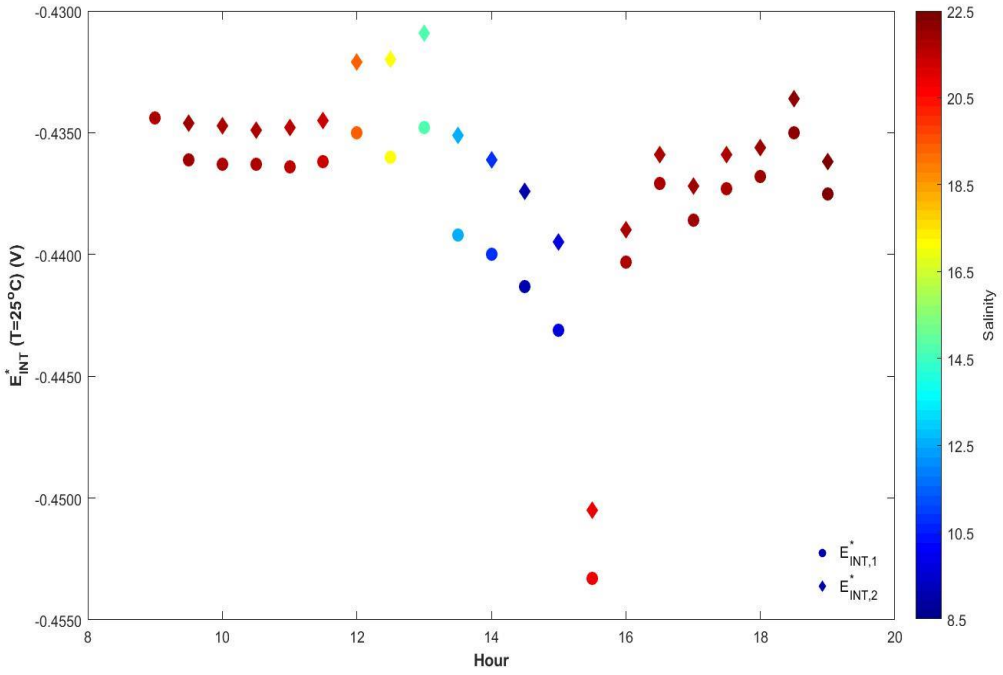

Figure 13. Calculated values of $E^{*}_{INT,1}$ (T = 25°C) from pH^{disc}_{DIC-TA} (circles) and $E^{*}_{INT,2}$ (T = 25°C) from pH^{disc}_{elec} (diamonds) for all measurements as a function of salinity from the 01 June 2016 sampling day.

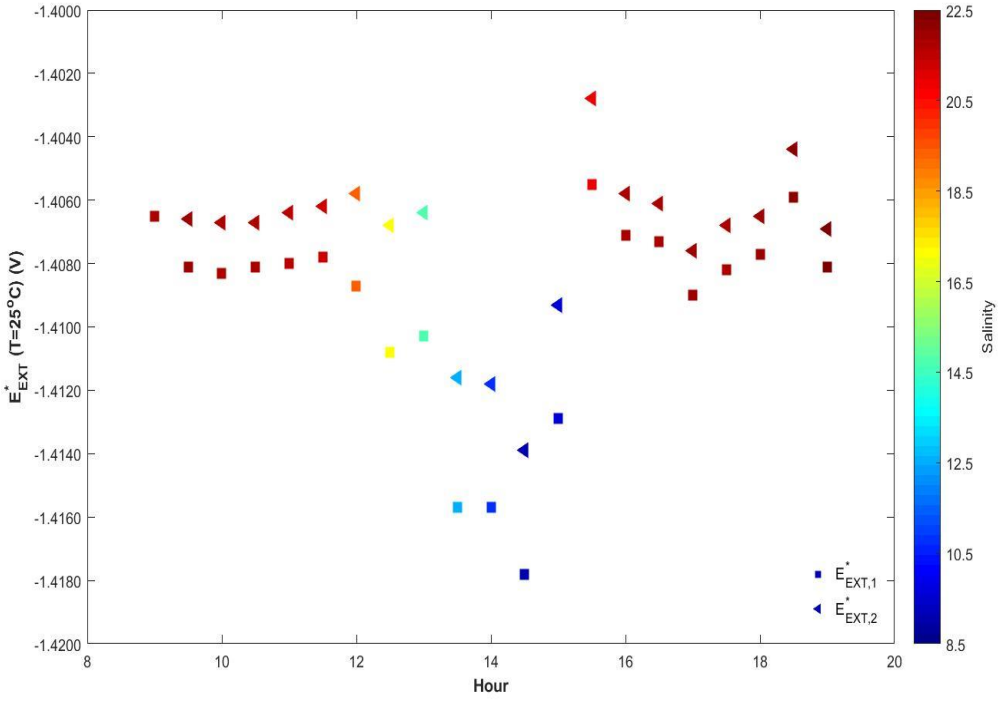

Figure 14. Calculated values of $E^{*}_{EXT,1}$ (T = 25℃) from pH^{disc}_{DIC-TA} (squares) and $E^{*}_{EXT,2}$ (T = 25℃) from pH^{disc}_{elec} (triangles) for all measurements as a function of salinity from the 01 June 2016 sampling day.

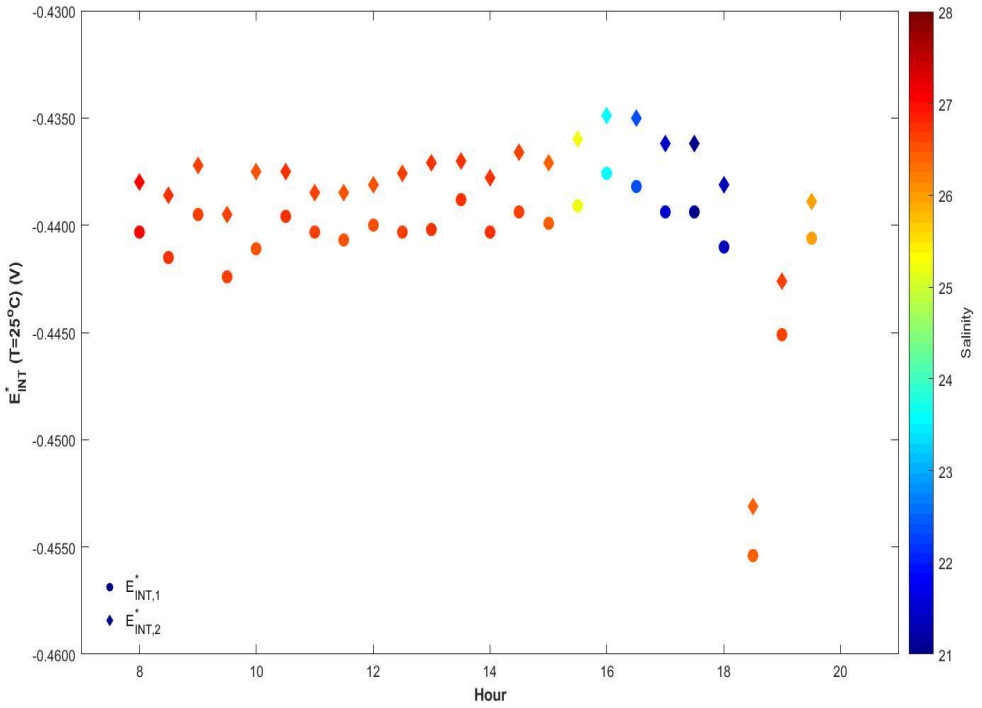

Figure 15. Calculated values of $E^{*}_{INT,1}$ (T = 25°C) from pH^{disc}_{DIC-TA} (circles) and $E^{*}_{INT,2}$ (T = 25°C) from pH^{disc}_{elec} (diamonds) for all measurements as a function of salinity from the 02 Aug 2016 sampling day.

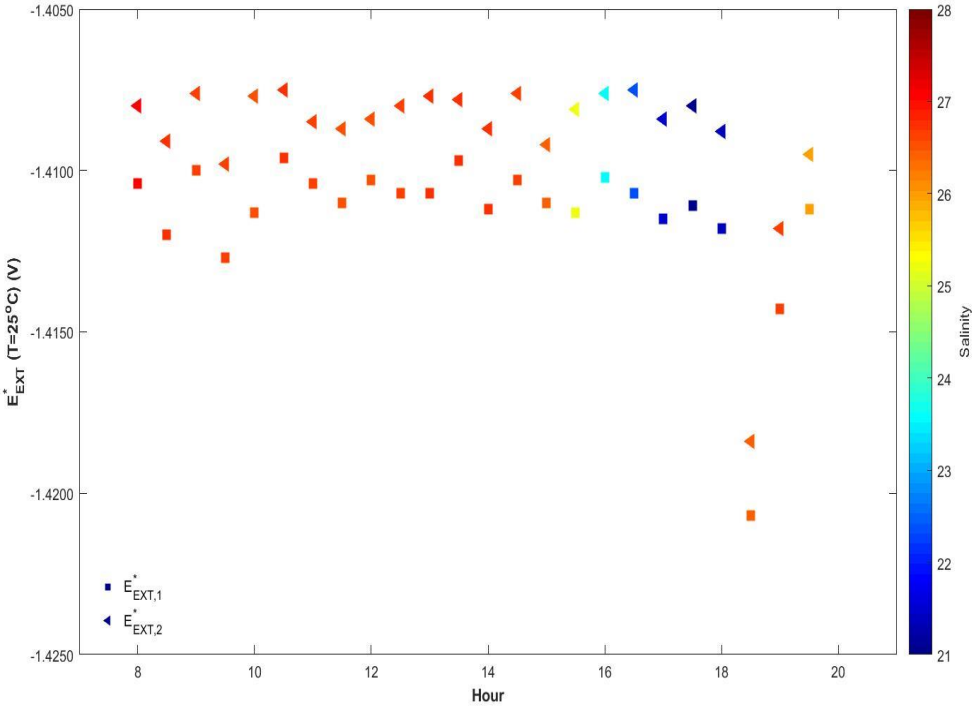

Figure 16. Calculated values of $E^*_{EXT,1}$ (T = 25°C) from pH^{disc}_{DIC-TA} (squares) and $E^*_{EXT,2}$ (T = 25°C) from pH^{disc}_{elec} (triangles) for all measurements as a function of salinity from the 02 August 2016 sampling day.

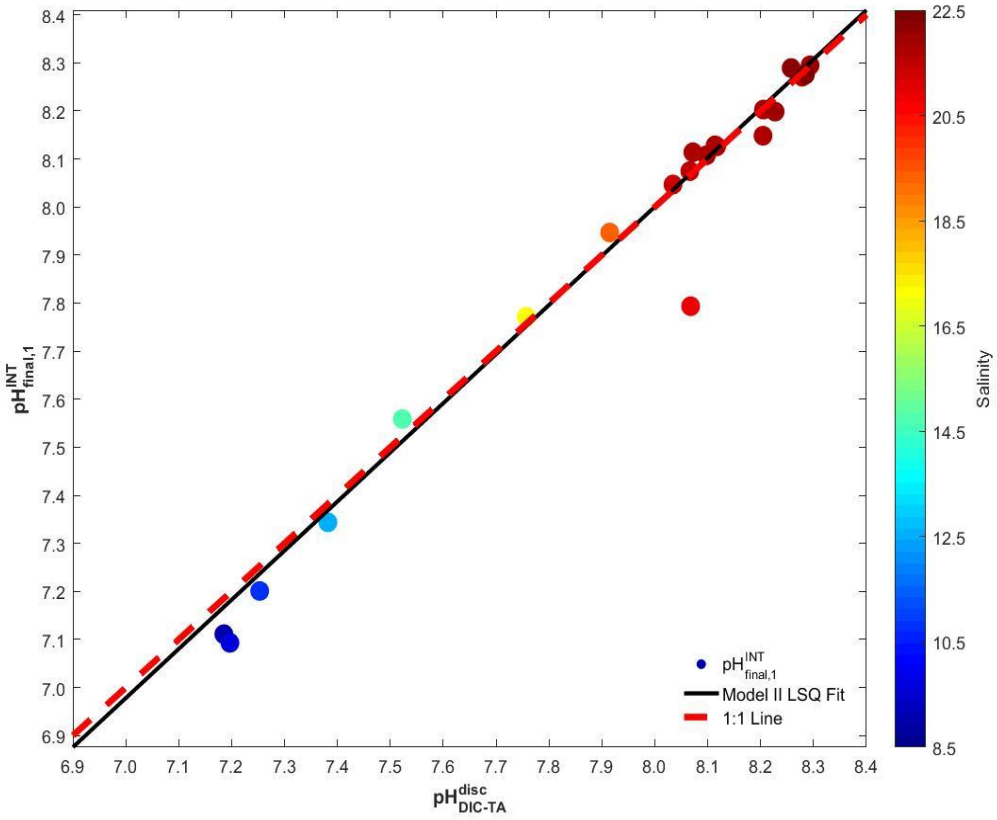

Figure 17. Property-property plot of $\mathrm{pH}^{\mathrm{INT}}_{\mathrm{final,1}}$ vs. $\mathrm{pH}^{\mathrm{disc}}_{\mathrm{DIC-TA}}$ as a function of salinity from the 01 June 2016 sampling day. The solid black line represents a Model II least squares fit and the dashed red line represents 1:1 $\left(\mathrm{pH}^{\mathrm{INT}}_{\mathrm{final,1}} = \mathrm{pH}^{\mathrm{disc}}_{\mathrm{DIC-TA}}\right)$ relationship.

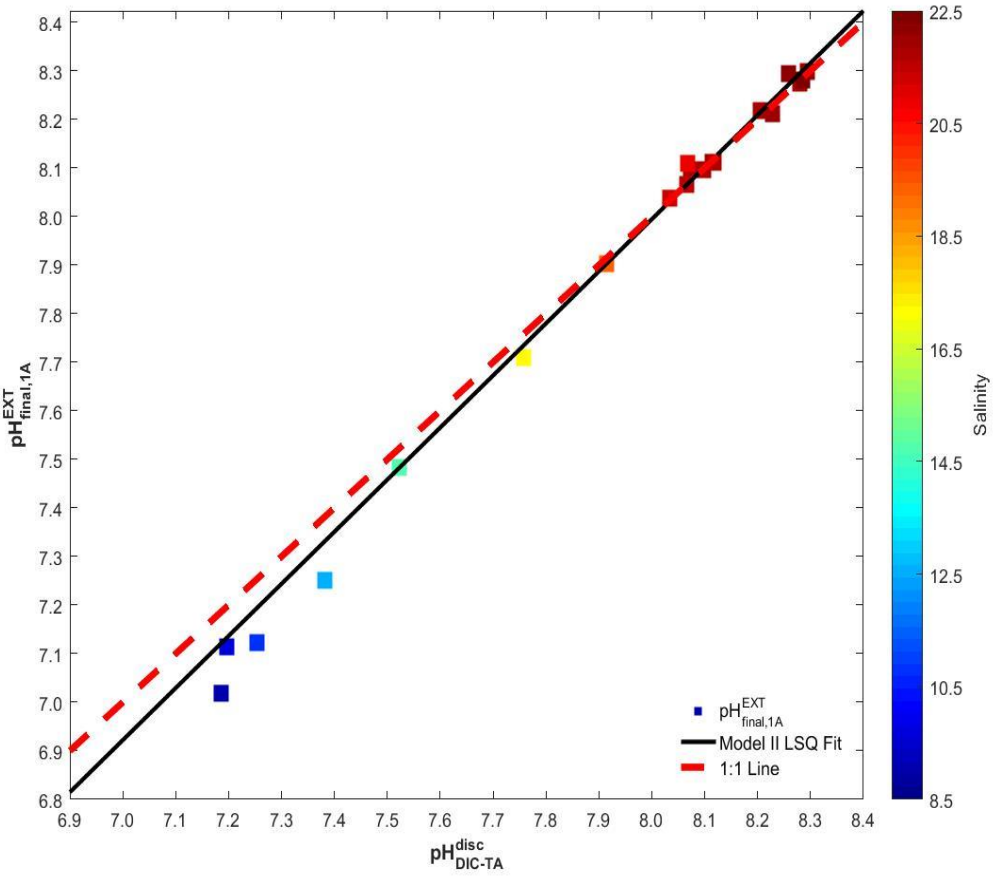

Figure 18. Property-property plot of $\mathrm{pH}^{\mathrm{EXT}}_{\mathrm{final,1A}}$ vs. $\mathrm{pH}^{\mathrm{disc}}_{\mathrm{DIC-TA}}$ as a function of salinity from the 01 June 2016 sampling day. The solid black line represents a Model II least squares fit and the dashed red line represents 1:1 $\left(\mathrm{pH}^{\mathrm{EXT}}_{\mathrm{final,1A}} = \mathrm{pH}^{\mathrm{disc}}_{\mathrm{DIC-TA}}\right)$ relationship.

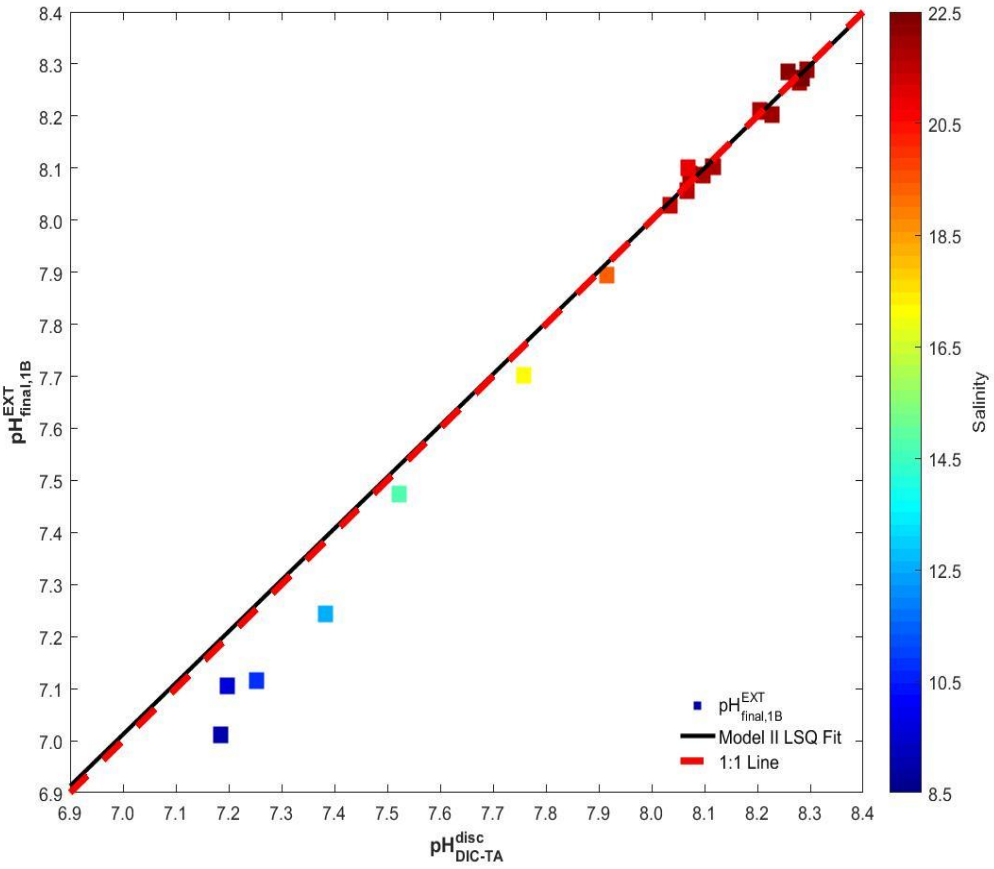

Figure 19. Property-property plot of $\mathrm{pH}^{\mathrm{EXT}}_{\mathrm{final,1B}}$ vs. $\mathrm{pH}^{\mathrm{disc}}_{\mathrm{DIC-TA}}$ as a function of salinity from the 01 June 2016 sampling day. The solid black line represents a Model II least squares fit and the dashed red line represents 1:1 $\left(\mathrm{pH}^{\mathrm{EXT}}_{\mathrm{final,1B}} = \mathrm{pH}^{\mathrm{disc}}_{\mathrm{DIC-TA}}\right)$ relationship.

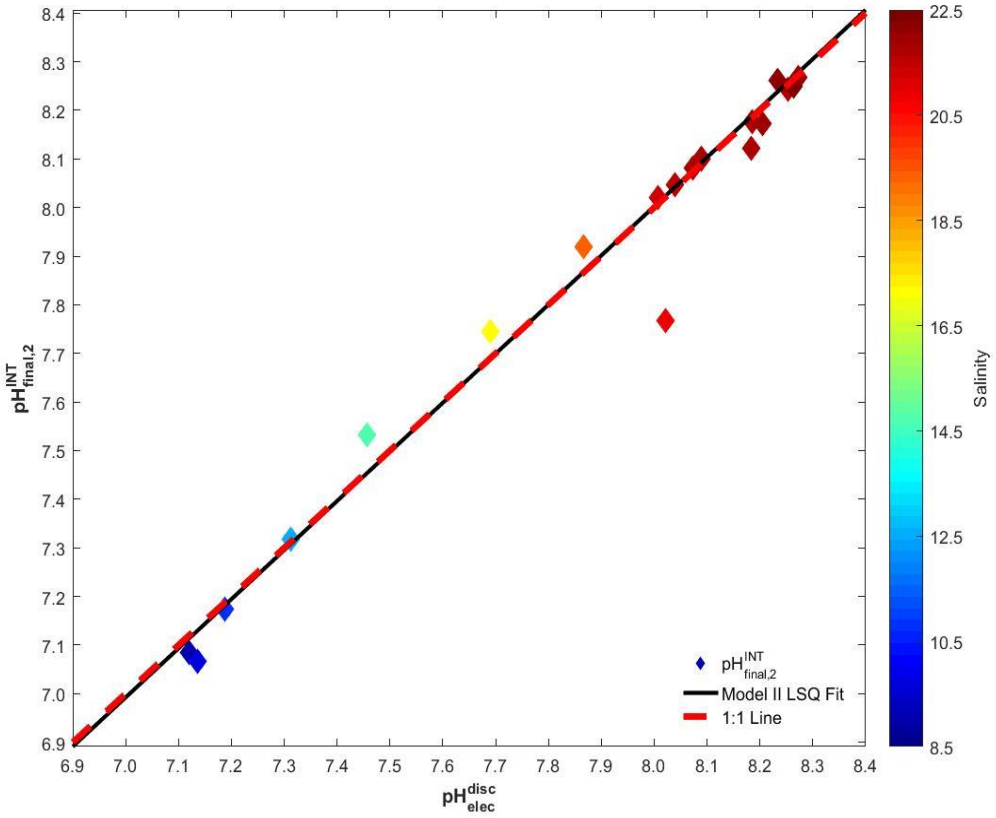

Figure 20. Property-property plot of $\mathrm{pH}_{\mathrm{final,2}}^{\mathrm{INT}}$ vs. $\mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}$ as a function of salinity from the 01 June 2016 sampling day. The solid black line represents a Model II least squares fit and the dashed red line represents 1:1 $\left(\mathrm{pH}_{\mathrm{final,2}}^{\mathrm{INT}} = \mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}\right)$ relationship.

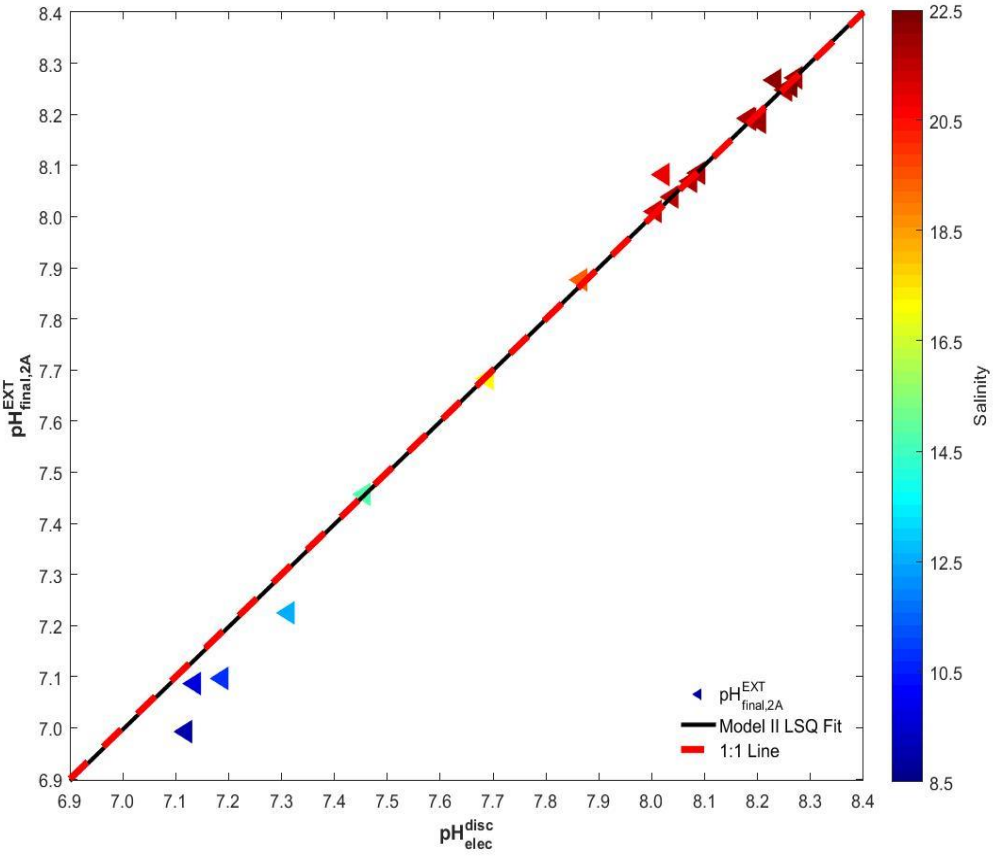

Figure 21. Property-property plot of $\mathrm{pH}_{\mathrm{final,2A}}^{\mathrm{EXT}}$ vs. $\mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}$ as a function of salinity from the 01 June 2016 sampling day. The solid black line represents a Model II least squares fit and the dashed red line represents 1:1 $\left(\mathrm{pH}_{\mathrm{final,2A}}^{\mathrm{EXT}} = \mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}\right)$ relationship.

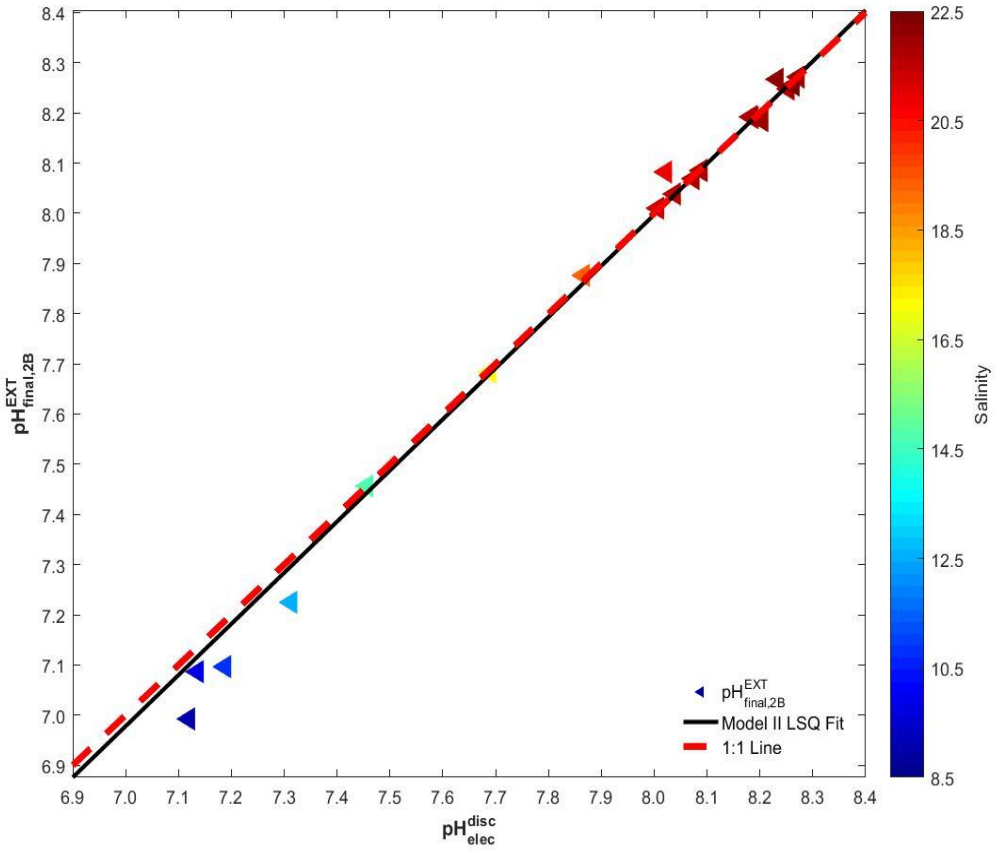

Figure 22. Property-property plot of $\mathrm{pH}^{\mathrm{EXT}}_{\mathrm{final,2B}}$ vs. $\mathrm{pH}^{\mathrm{disc}}_{\mathrm{elec}}$ as a function of salinity from the 01 June 2016 sampling day. The solid black line represents a Model II least squares fit and the dashed red line represents 1:1 $\left(\mathrm{pH}^{\mathrm{EXT}}_{\mathrm{final,2B}} = \mathrm{pH}^{\mathrm{disc}}_{\mathrm{elec}}\right)$ relationship.

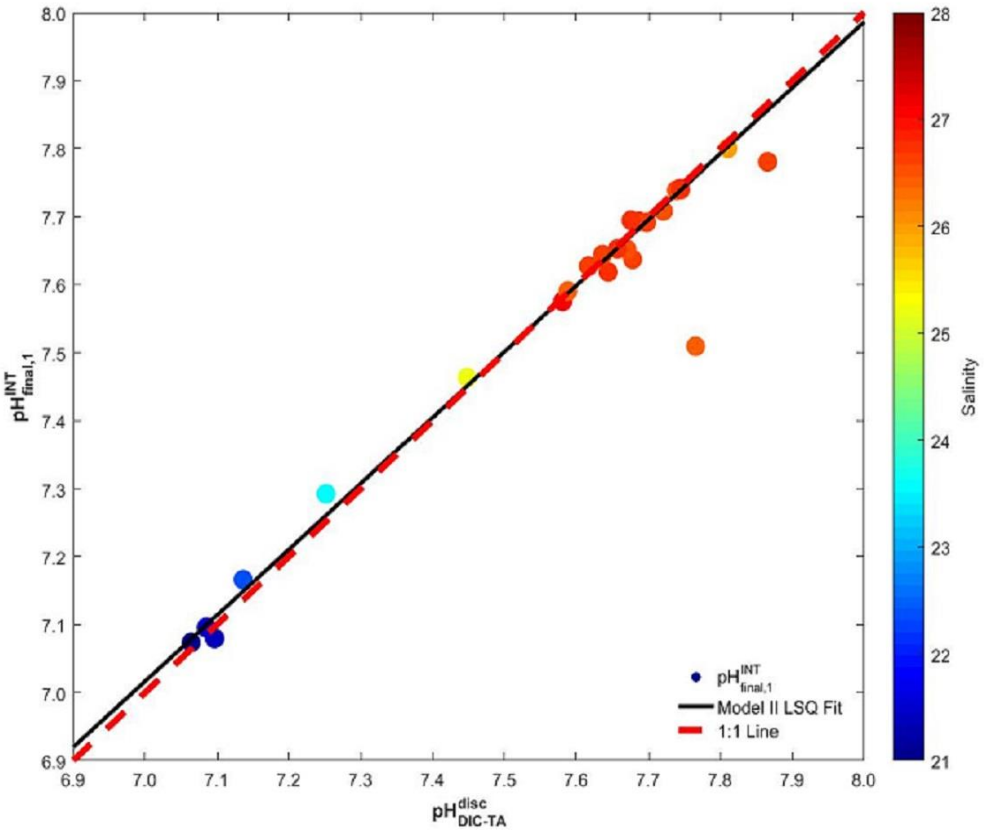

Figure 23. Property-property plot of $\mathrm{pH}^{\mathrm{INT}}_{\mathrm{final,1}}$ vs. $\mathrm{pH}^{\mathrm{disc}}_{\mathrm{DIC-TA}}$ as a function of salinity from the 02 August 2016 sampling day. The solid black line represents a Model II least squares fit and the dashed red line represents 1:1 $\left(\mathrm{pH}^{\mathrm{INT}}_{\mathrm{final,1}} = \mathrm{pH}^{\mathrm{disc}}_{\mathrm{DIC-TA}}\right)$ relationship.

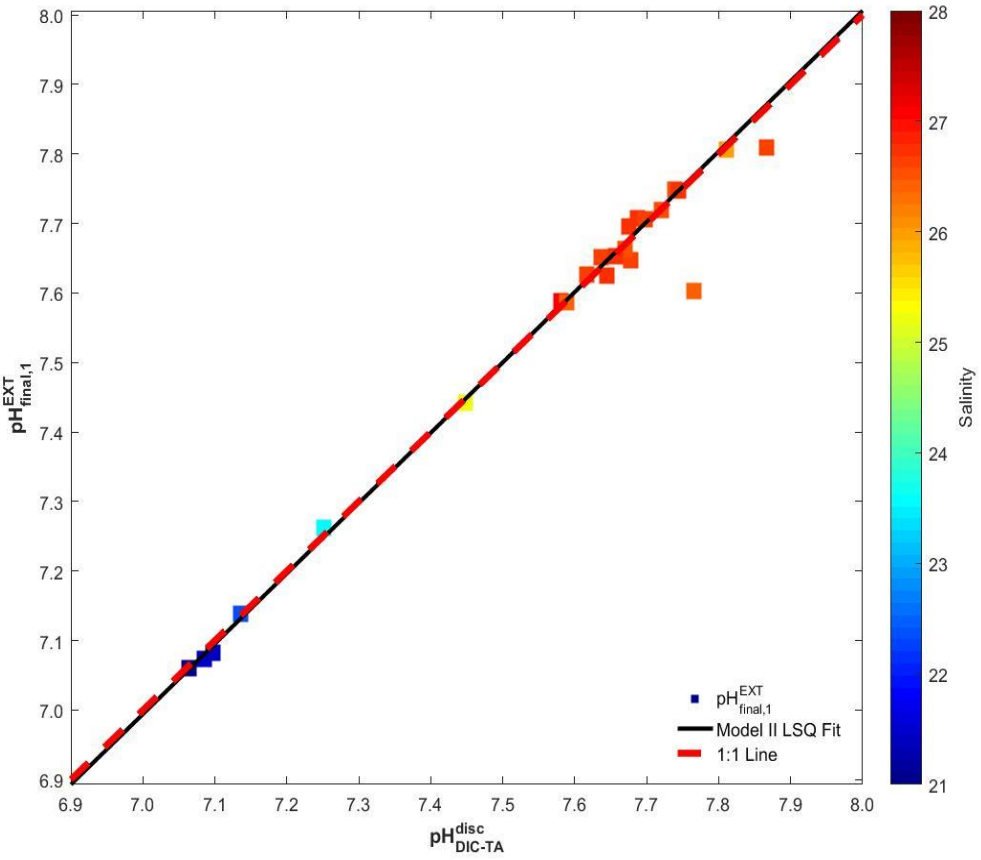

Figure 24. Property-property plot of $\mathrm{pH}^{\mathrm{EXT}}_{\mathrm{final,1}}$ vs. $\mathrm{pH}^{\mathrm{disc}}_{\mathrm{DIC-TA}}$ as a function of salinity from the 02 August 2016 sampling day. The solid black line represents a Model II least squares fit and the dashed red line represents 1:1 $\left(\mathrm{pH}^{\mathrm{EXT}}_{\mathrm{final,1}} = \mathrm{pH}^{\mathrm{disc}}_{\mathrm{DIC-TA}}\right)$ relationship.

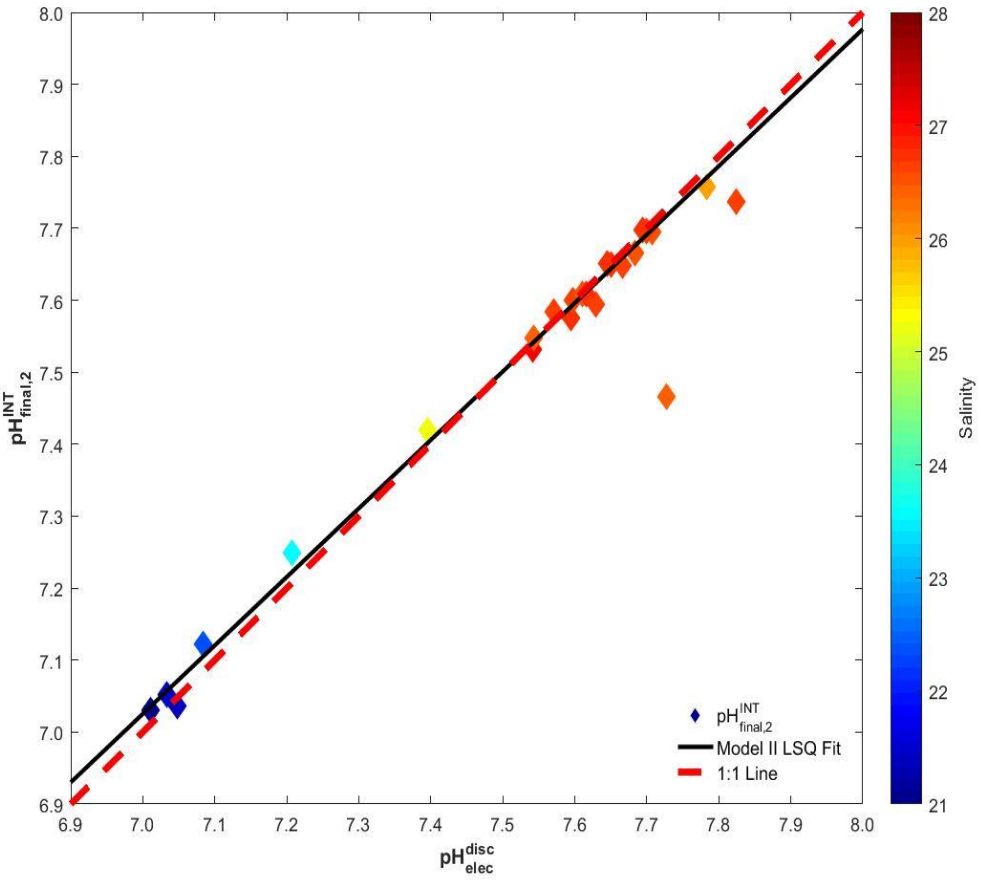

Figure 25. Property-property plot of $\mathrm{pH}_{\mathrm{final,2}}^{\mathrm{INT}}$ vs. $\mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}$ as a function of salinity from the 02 August 2016 sampling day. The solid black line represents a Model II least squares fit and the dashed red line represents 1:1 $\left(\mathrm{pH}_{\mathrm{final,2}}^{\mathrm{INT}} = \mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}\right)$ relationship.

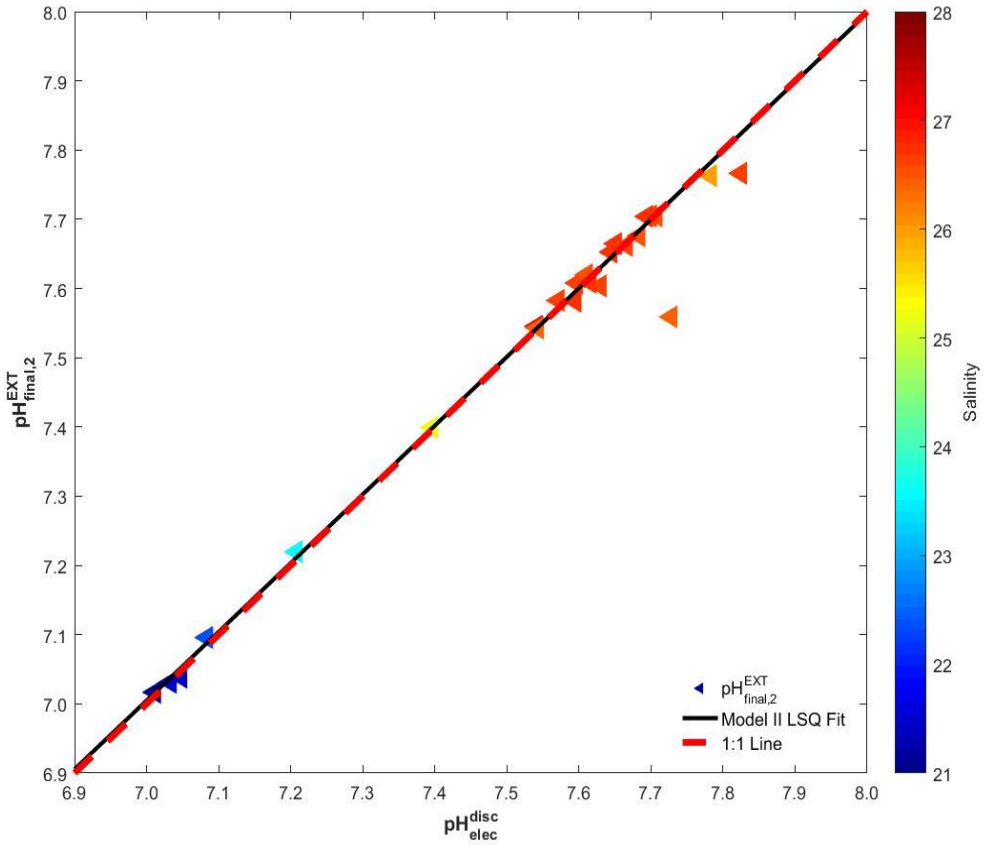

Figure 26. Property-property plot of $\mathrm{pH}_{\mathrm{final,2}}^{\mathrm{EXT}}$ vs. $\mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}$ as a function of salinity from the 02 August 2016 sampling day. The solid black line represents a Model II least squares fit and the dashed red line represents 1:1 $\left(\mathrm{pH}_{\mathrm{final,2}}^{\mathrm{EXT}} = \mathrm{pH}_{\mathrm{elec}}^{\mathrm{disc}}\right)$ relationship.

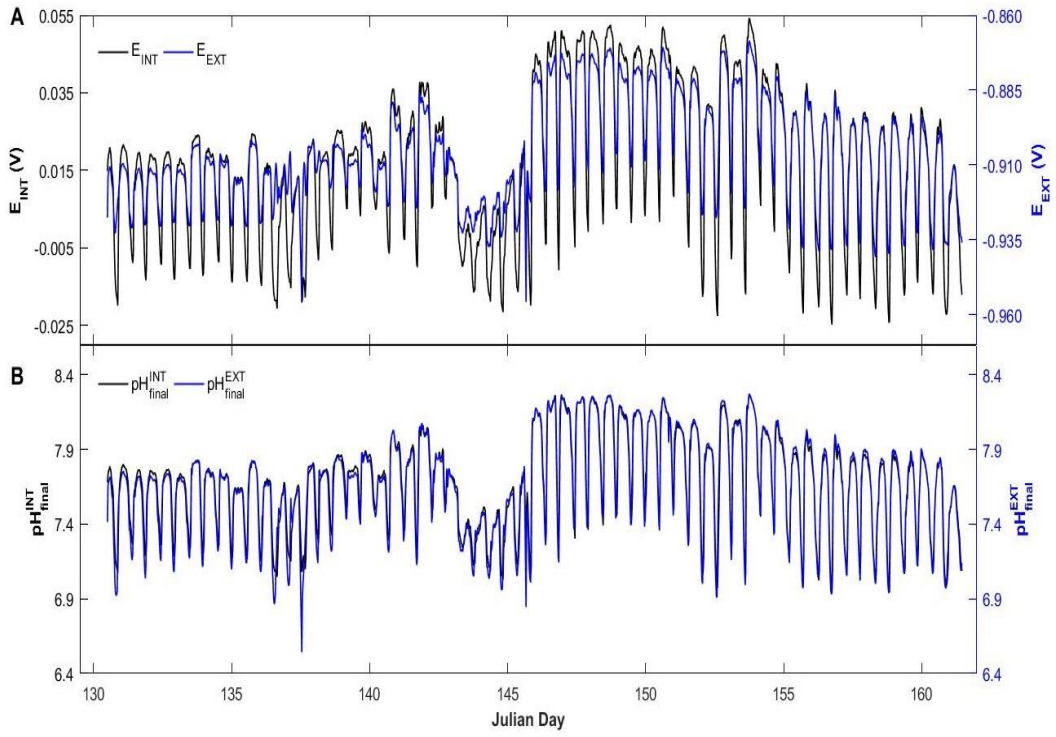

Figure 27. Murderkill Estuary-Delaware Bay System pH time-series from 09 May 2016 to 09 June 2016. (A) Raw sensor voltages for the internal (black) and external (blue) reference electrodes show four shorter-term deployment combined into one longer-term deployment. (B) pH calculated using the internal (black) and external (blue) reference electrodes.

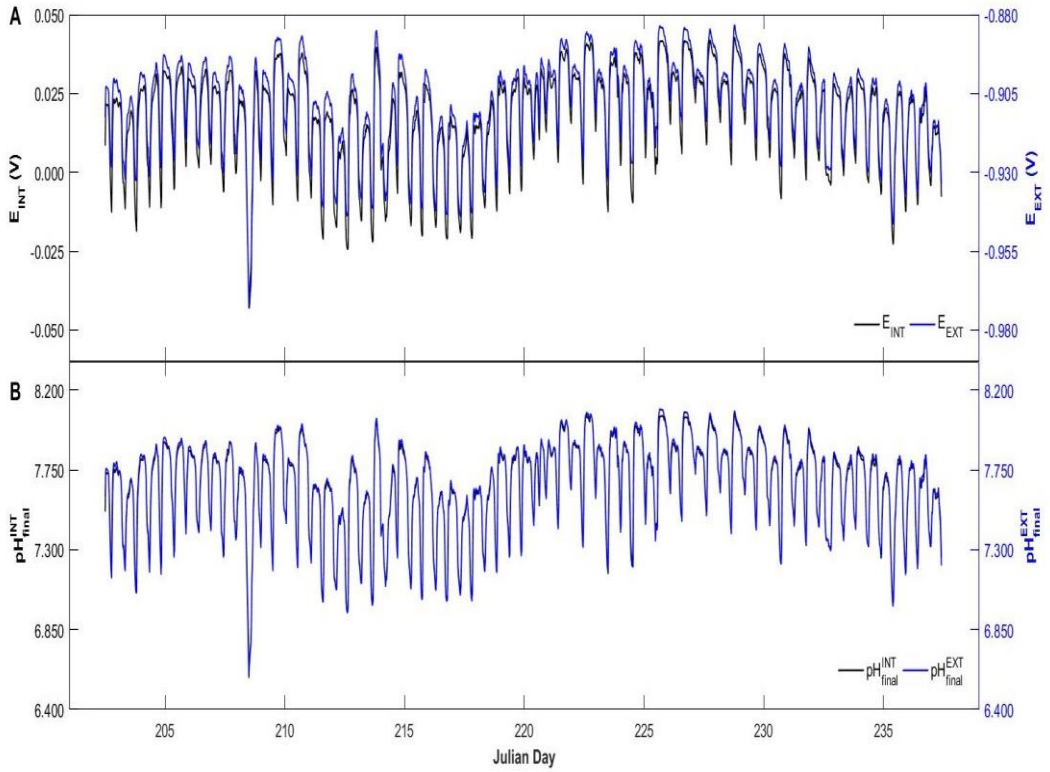

Figure 28. Murderkill Estuary-Delaware Bay System pH time-series from 20 July 2016 to 24 August 2016. (A) Raw sensor voltages for the internal (black) and external (blue) reference electrodes show four shorter-term deployment combined into one longer-term deployment. (B) pH calculated using the internal (black) and external (blue) reference electrodes.

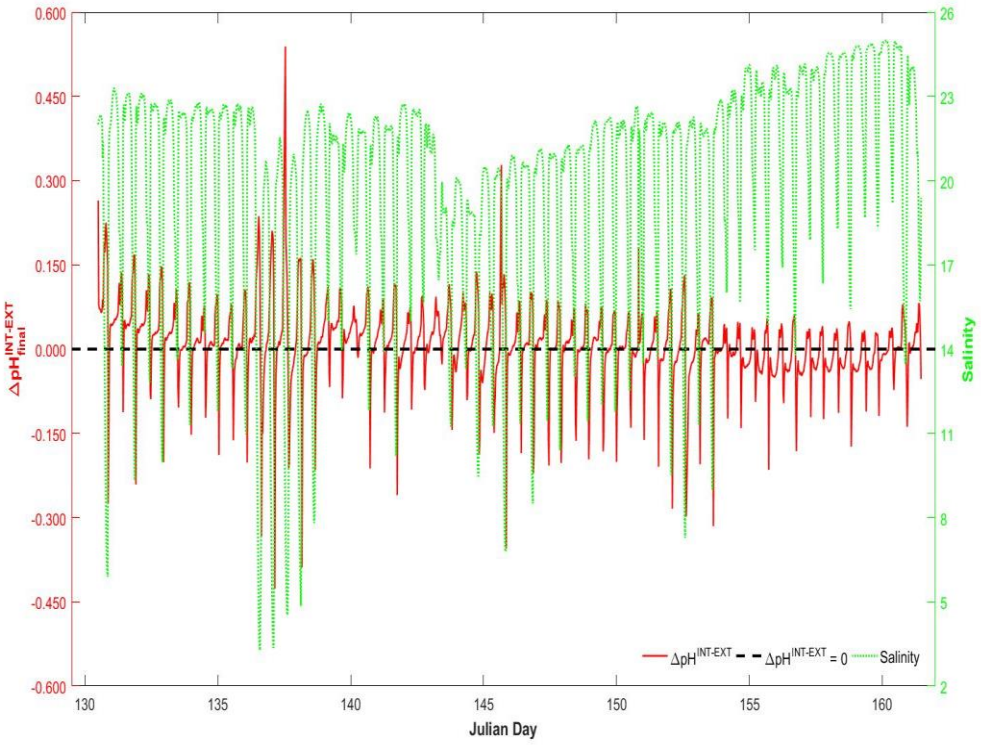

Figure 29. Calculated $\Delta\mathrm{pH}^{\mathrm{INT-EXT}}$ anomaly (solid red line) shown relative to a zero $\Delta\mathrm{pH}^{\mathrm{INT-EXT}}$ anomaly (dashed black line) and *in-situ* salinity (dotted green line) from 09 May 2016 to 09 June 2016. Uncharacteristically large positive $\Delta\mathrm{pH}^{\mathrm{INT-EXT}}$ anomalies coincide with the first measurements of conditioning periods that follow sensor maintenance.

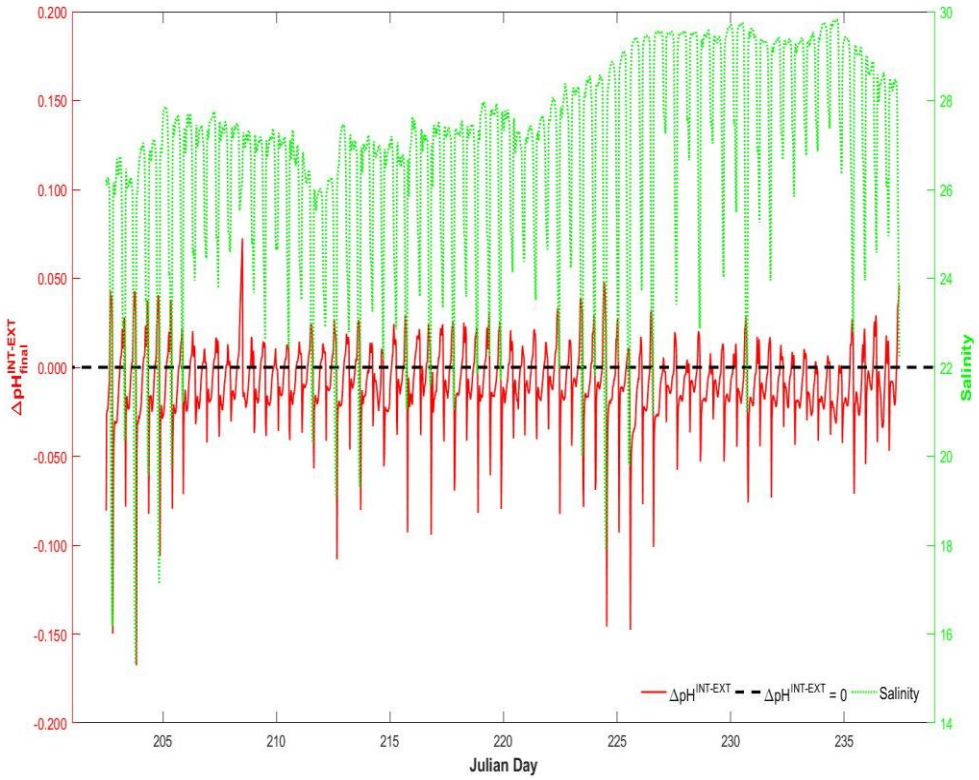

Figure 30. Calculated $\Delta\mathrm{pH}^{\mathrm{INT-EXT}}$ anomaly (solid red line) shown relative to a zero $\Delta\mathrm{pH}^{\mathrm{INT-EXT}}$ anomaly (dashed black line) and *in-situ* salinity (dotted green line) from 20 July 2016 to 24 August 2016.

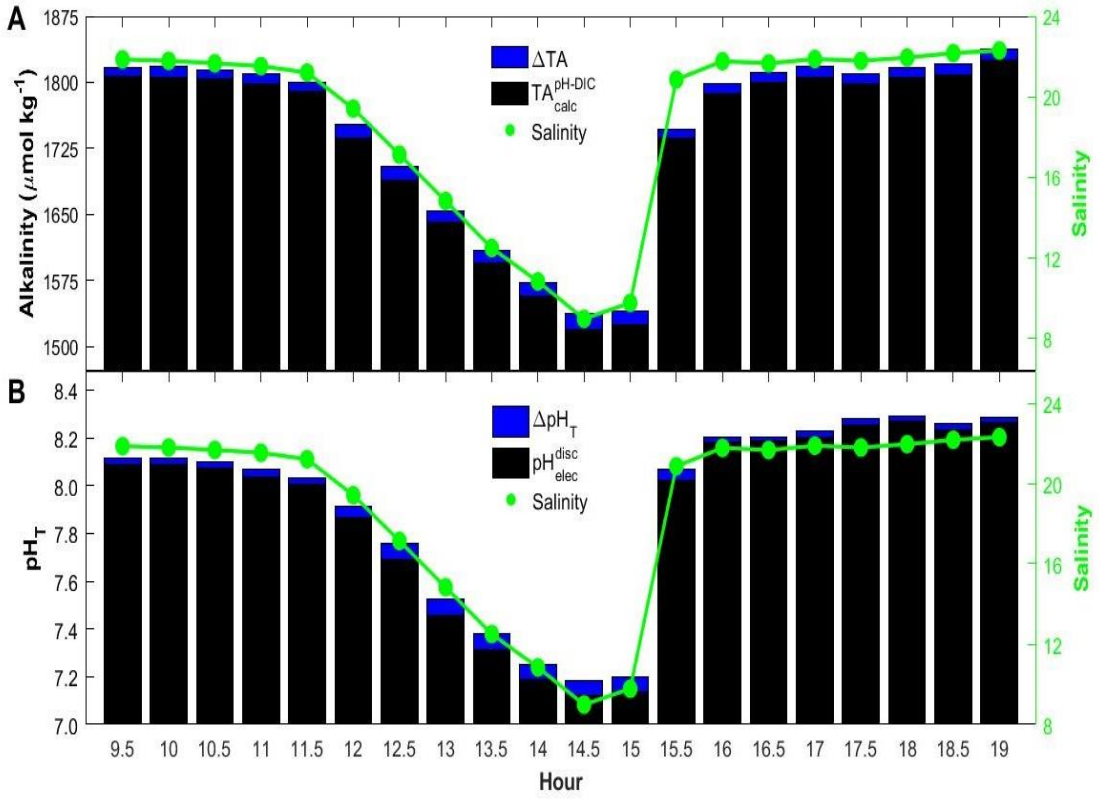

Figure 31. Contributions of excess alkalinity to total measured alkalinity on the 01 June 2016 sampling day. (A) Excess alkalinity (ΔTA) (blue) and alkalinity calculated from pH^{field}_{NBS} and DIC (black) shown relative to *in-situ* salinity (green). (B) ΔpH due to contributions of excess alkalinity (blue) and pH^{disc}_{elec} assumed to be free of the contributions of excess alkalinity (black) shown relative to *in-situ* salinity (green).

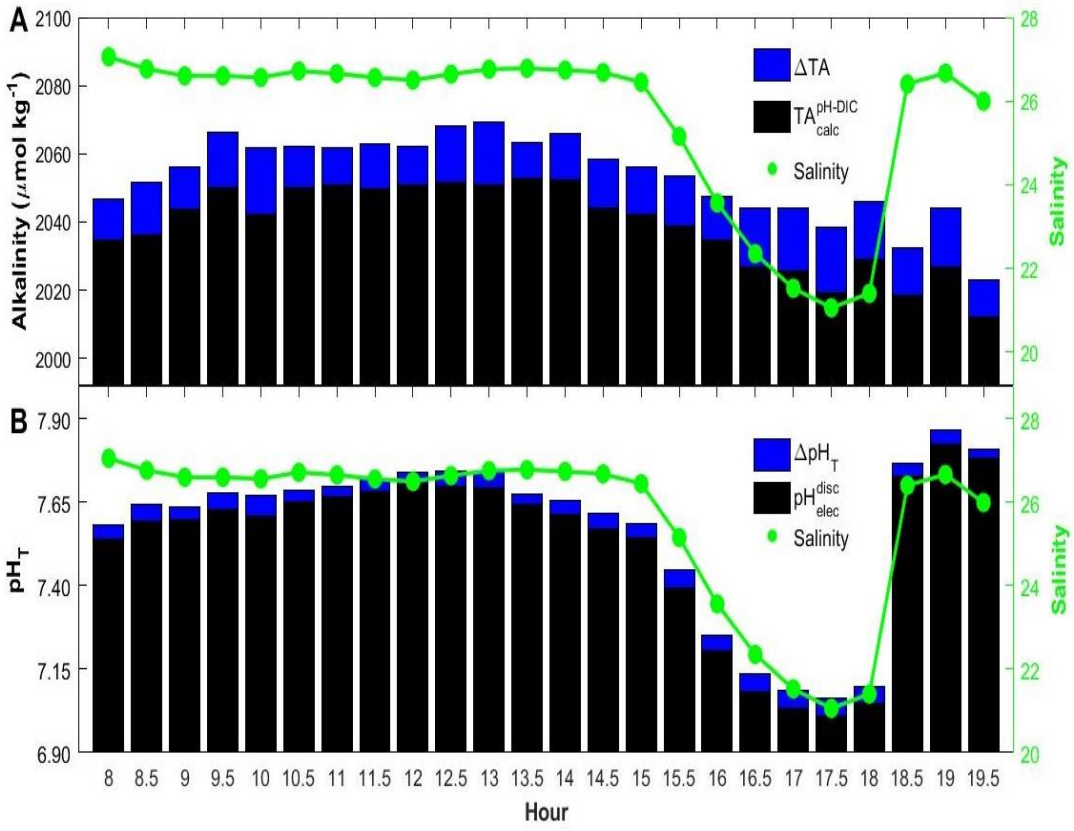

Figure 32. Contributions of excess alkalinity to total measured alkalinity on the 02 August 2016 sampling day. (A) Excess alkalinity (ΔTA) (blue) and alkalinity calculated from pH_{NBS}^{field} and DIC (black) shown relative to *in-situ* salinity (green). (B) ΔpH due to contributions of excess alkalinity (blue) and pH_{elec}^{disc} assumed to be free of the contributions of excess alkalinity (black) shown relative to *in-situ* salinity (green).

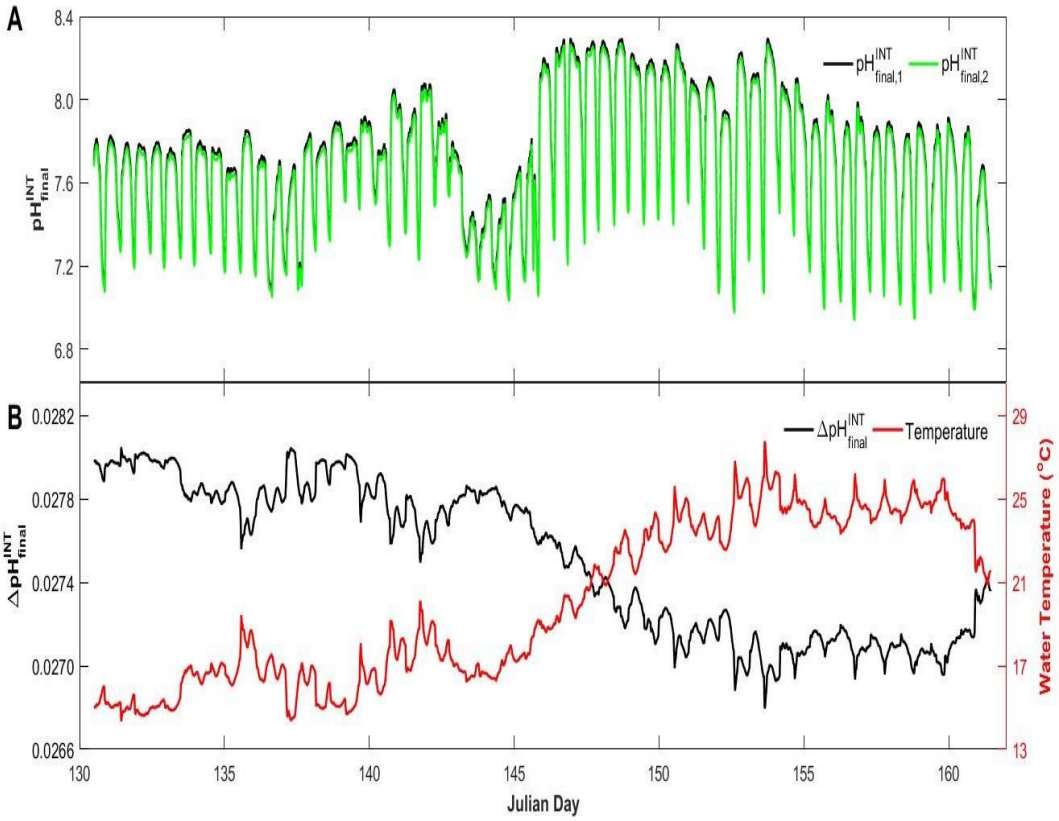

Figure 33. Composite of excess alkalinity effects time-series for 09 May 2016 to 09 June 2016 SeapHOx deployment. (A) $\mathrm{pH}^{\mathrm{INT}}_{\mathrm{final},1}$ (black) calibrated using $\mathrm{pH}^{\mathrm{disc}}_{\mathrm{DIC-TA}}$ assumed to incorporate errors due to excess alkalinity and $\mathrm{pH}^{\mathrm{INT}}_{\mathrm{final},2}$ (green) calibrated using $\mathrm{pH}^{\mathrm{disc}}_{\mathrm{elec}}$ assumed to be free of errors due to excess alkalinity. (B) $\Delta\mathrm{pH}^{\mathrm{INT}}_{\mathrm{final}}$ anomaly (black) calculated from difference between $\mathrm{pH}^{\mathrm{INT}}_{\mathrm{final},1}$ and $\mathrm{pH}^{\mathrm{INT}}_{\mathrm{final},2}$ shown relative to *in-situ* water temperature (red).

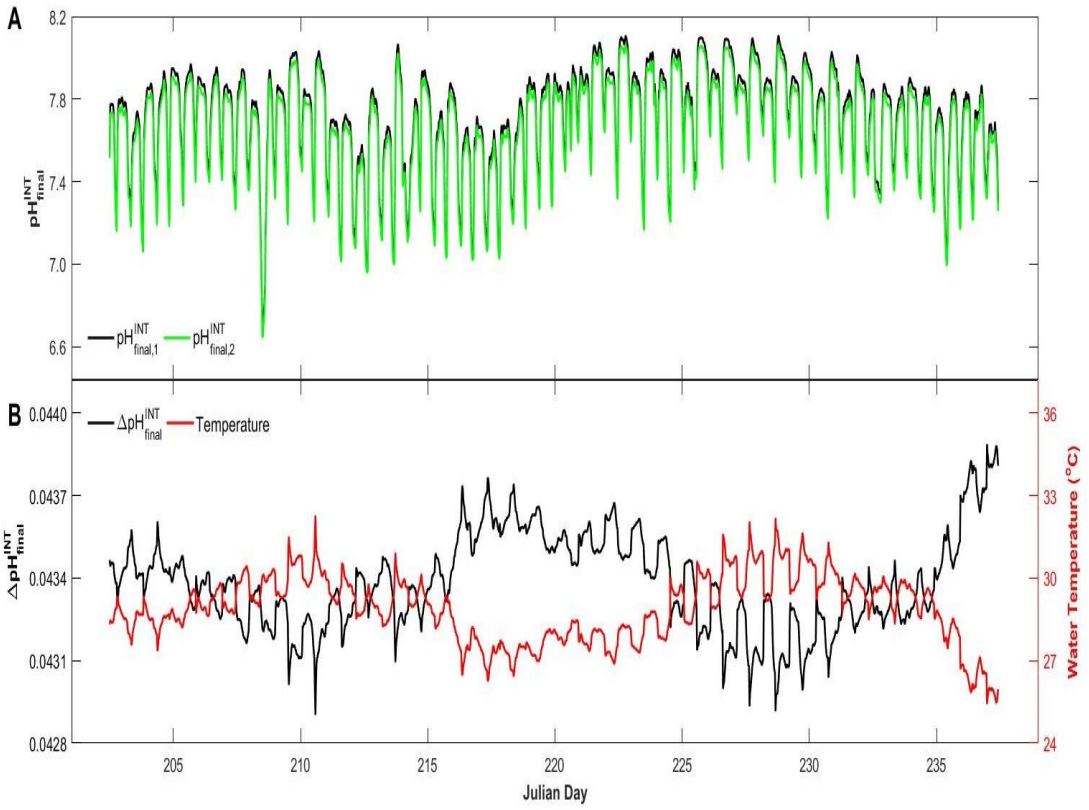

Figure 34. Composite of excess alkalinity effects time-series for 20 July 2016 to 24 August 2016 SeapHOx deployment. (A) $pH^{INT}_{final,1}$ (black) calibrated using pH^{disc}_{DIC-TA} assumed to incorporate errors due to excess alkalinity and $pH^{INT}_{final,2}$ (green) calibrated using pH^{disc}_{elec} assumed to be free of errors due to excess alkalinity. (B) ΔpH^{INT}_{final} anomaly (black) calculated from difference between $pH^{INT}_{final,1}$ and $pH^{INT}_{final,2}$ shown relative to *in-situ* water temperature (red).

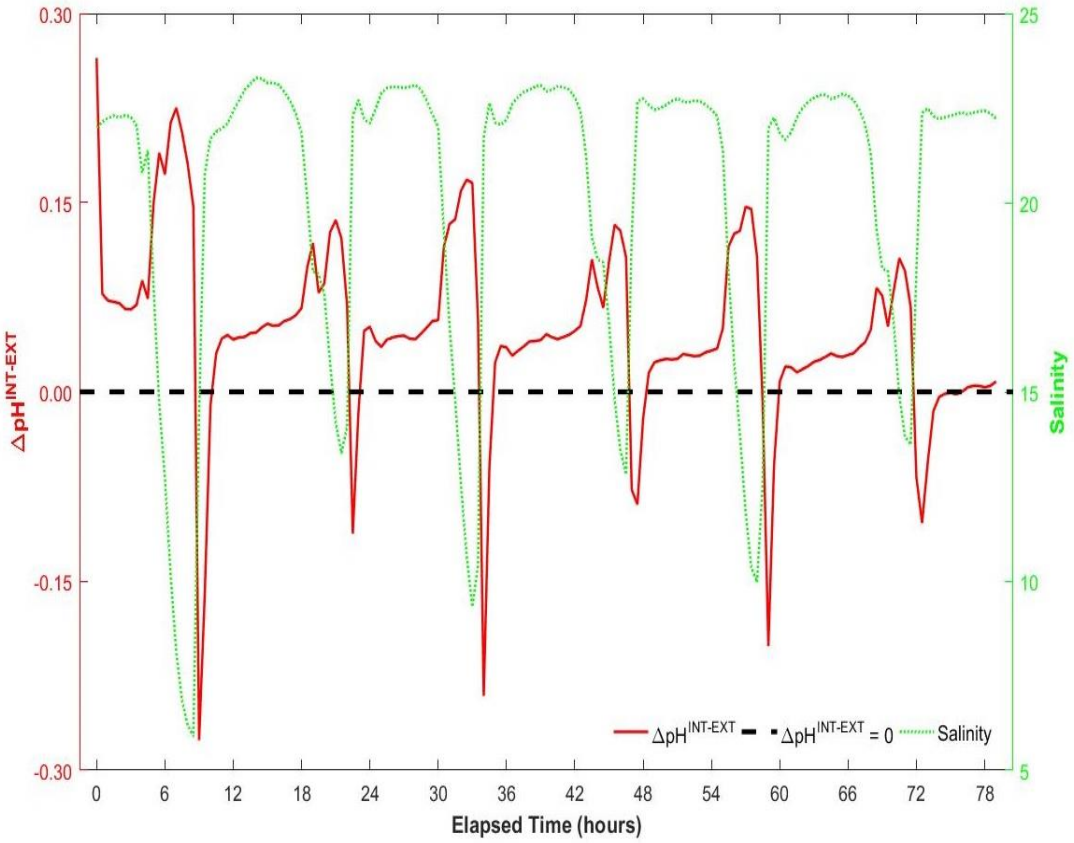

Figure 35. Conditioning period at the start of the 09 May 2016 to 24 August 2016 SeapHOx deployment. Calculated $\Delta pH^{INT-EXT}$ anomaly (solid red line) shown relative to a zero $\Delta pH^{INT-EXT}$ anomaly (dashed black line) and *in-situ* salinity (dotted green line). Conditioning period starts on 9 May 2016 at 1200 (Hour 0) and goes until 12 May 2016 at 1900 (Hour 79).

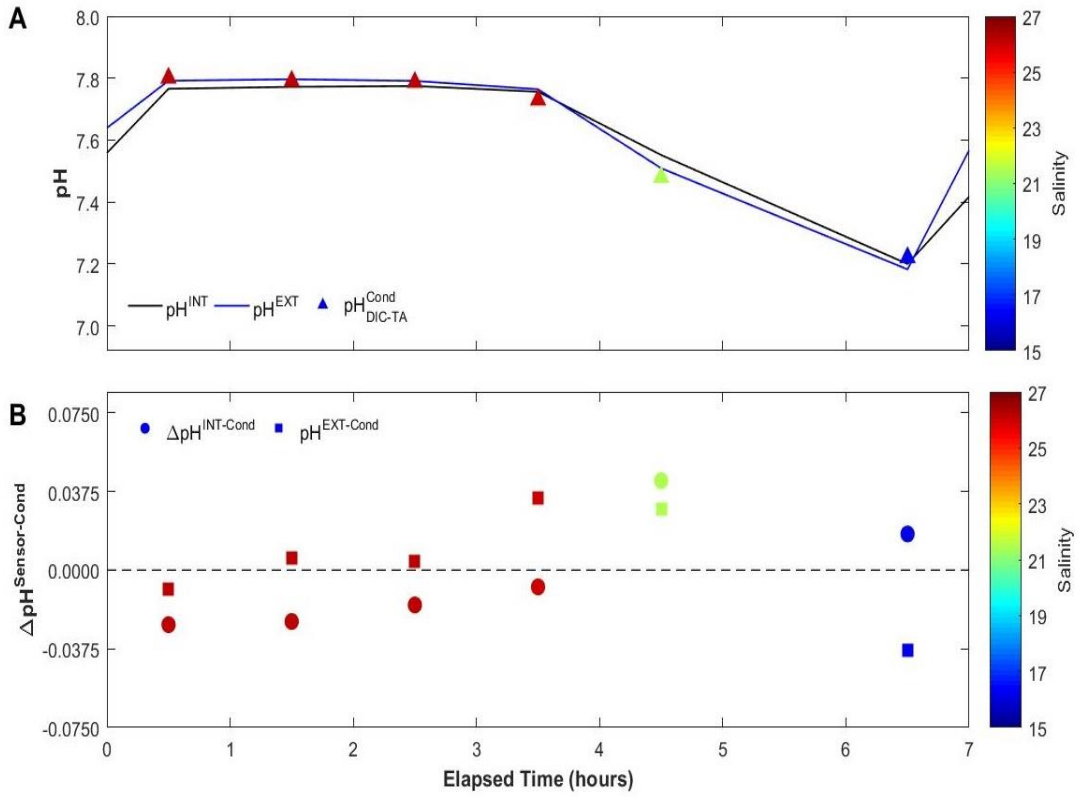

Figure 36. Composite of intra-deployment conditioning time-series following sensor maintenance on 20 July 2016. (A) sensor measured pH^{INT} (solid black line) and pH^{EXT} (solid blue line) shown relative to pH_T of conditioning samples calculated from DIC-TA (pH^{Cond}_{DIC-TA} – triangles) as a function of salinity. (B) $\Delta pH^{INT-Cond}$ (circles) and $\Delta pH^{EXT-Cond}$ (squares) as a function of salinity. Conditioning period starts on 20 July 2016 at 1200 (Hour 0) and goes until 20 July 2016 at 1900 (Hour 7). Conditioning samples collected at Hour 5.5 (1730) partially froze in storage and were lost.

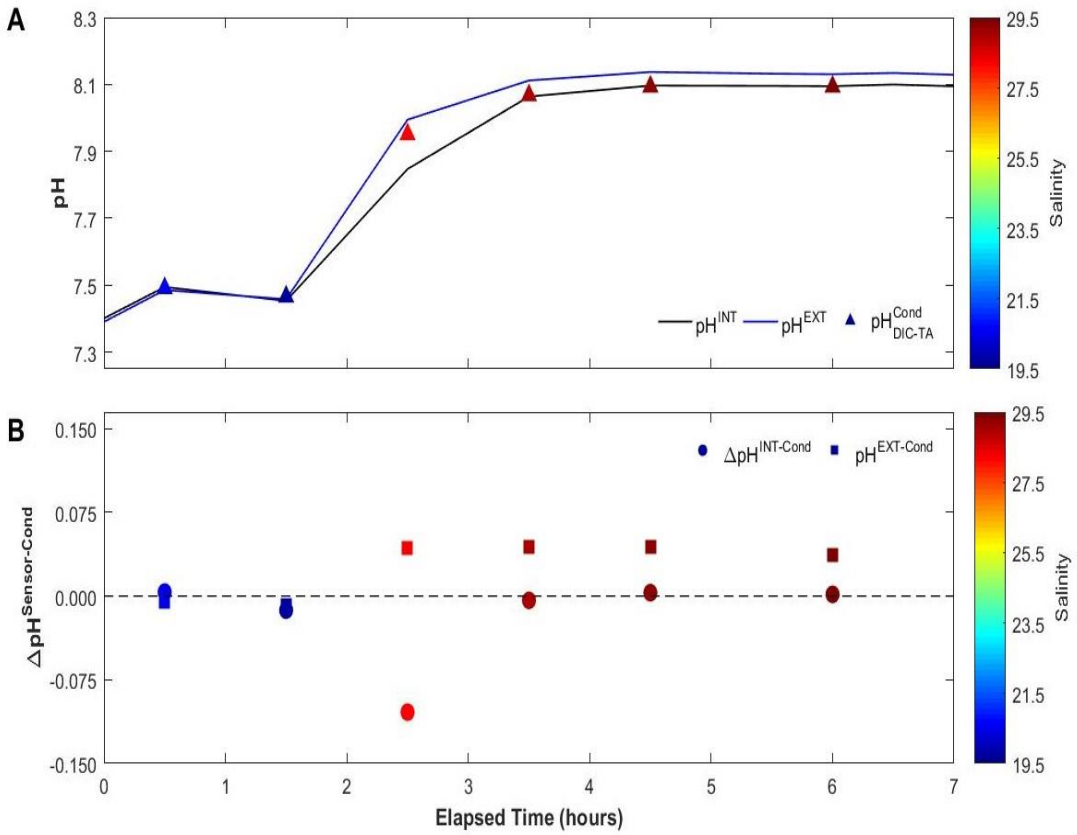

Figure 37. Composite of intra-deployment conditioning time-series following sensor maintenance on 12 August 2016. (A) sensor measured pH^{INT} (solid black line) and pH^{EXT} (solid blue line) shown relative to pH_T of conditioning samples calculated from DIC-TA (pH^{Cond}_{DIC-TA} – triangles) as a function of salinity. (B) $\Delta pH^{INT-Cond}$ (circles) and $\Delta pH^{EXT-Cond}$ (squares) as a function of salinity. Conditioning period starts on 12 August 2016 at 1130 (Hour 0) and goes until 12 August 2016 at 1830 (Hour 7). Conditioning samples collected at Hour 5.5 (1700) were collected 7 minutes after the hour effectively missing the sensor measurement and were excluded.

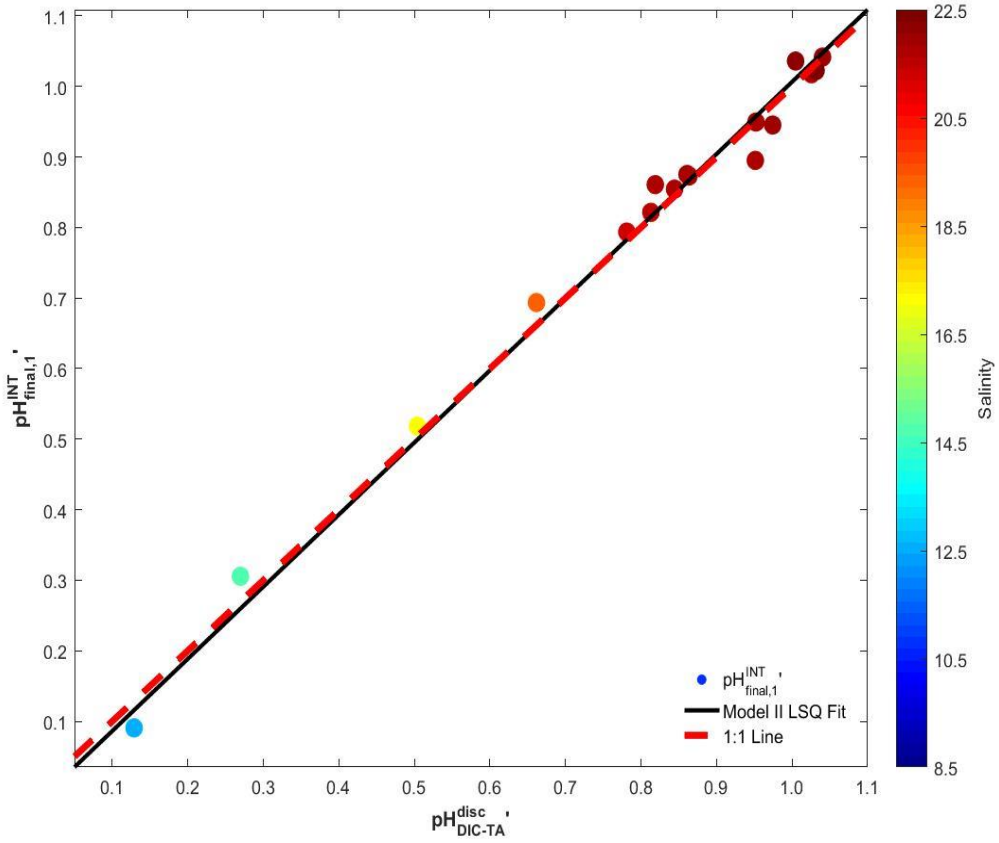

Figure 38. Property-property plot of $\mathrm{pH}_{\mathrm{final,1}}^{\mathrm{INT}}{}'$ vs. $\mathrm{pH}_{\mathrm{DIC-TA}}^{\mathrm{disc}}{}'$ as a function of salinity from the 01 June 2016 sampling day. The solid black line represents a Model II least squares fit and the dashed red line represents 1:1 $\left(\mathrm{pH}_{\mathrm{final,1}}^{\mathrm{INT}}{}' = \mathrm{pH}_{\mathrm{DIC-TA}}^{\mathrm{disc}}{}'\right)$ relationship.

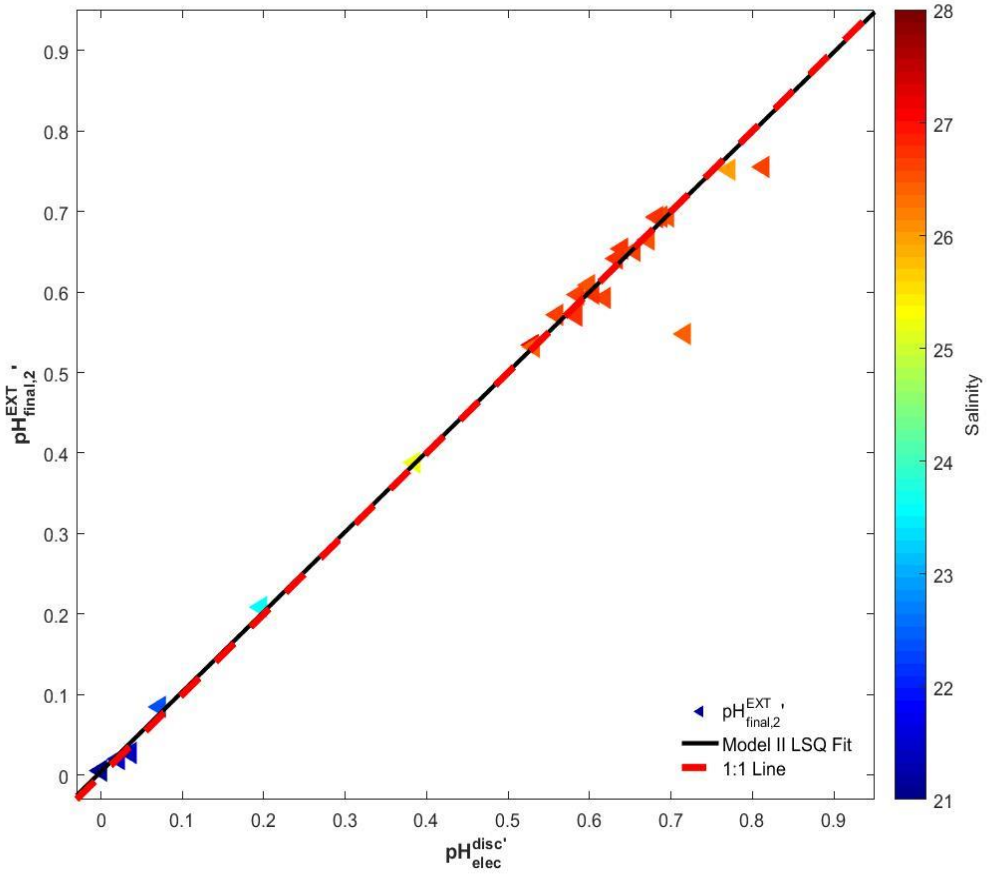

Figure 39. Property-property plot of $pH_{final,2}^{EXT\ \prime}$ vs. $pH_{elec}^{disc\ \prime}$ as a function of salinity from the 02 August 2016 sampling day. The solid black line represents a Model II least squares fit and the dashed red line represents 1:1 $\left(pH_{final,2}^{EXT\ \prime} = pH_{elec}^{disc\ \prime}\right)$ relationship.

REFERENCES

ACT. (2012). Alliance for Coastal Technologies: Protocol for the Verification of *In Situ* pH Sensors.

Alin, S. R., Feely, R. A., Dickson, A. G., Hernández-Ayón, J. M., Juranek, L. W., Ohman, M. D., and Goericke, R. (2012). Robust empirical relationships for estimating the carbonate system of the southern California Current System and application to CalCOFI hydrographic cruise data (2005-2011). *Journal of Geophysical Research, 117*(C5).

Andersson, A. J., Kline, D. I., Edmunds, P. J., Archer, S. D., Bednarsek, N., Carpenter, R. C., Chadsey, M., Goldstein, P., Grottoli, A. G., Hurst, T. P., King, A. L., Kübler, J. E., Kuffner, I. B., Mackey, K. R. M., Menge, B. A., Paytan, A., Riebesell, U., Schnetzer, A., Warner, M. E., and Zimmerman, R. C. (2015). Understanding ocean acidification impacts on organismal to ecological scales. *Oceanography*, *28*(2).

Bagshaw, E. A., Beaton, A., Wadham, J. L., Mowlem, M., Hawkings, J. R., and Tranter, M. (2016). Chemical sensors for in situ data collection in the cryosphere. *Trends in Analytical Chemistry, 82*, 348-357.

Bates, N. R. (2007). Interannual variability of the oceanic CO_2 sink in the subtropical gyre of the North Atlantic Ocean over the last 2 decades. *Journal of Geophysical Research: Oceans, 112*(C9).

Bates, R. G. (1973). *Determination of pH: Theory and Practice* (2nd ed.). Wiley-Interscience.

Bates, R. G. (1975). pH scales for seawater In E. D. Goldberg (Ed), The nature of seawater, *Dahlem Konferenzen* (pp. 313-338).

Bates, R. G. (1982). pH measurements in the marine environment. *Pure and Applied Chemistry, 54*(1), 229–232.

Bates, R. G., and Calais, J. G. (1981). Thermodynamics of the dissociation of bis h+ in seawater from 5 to 40c. *Journal of Solution Chemistry, 10*(4), 269–279.

Bates, R. G., and Erickson, W. P. (1986). Thermodynamics of the dissociation of 2-aminopyridinium ion in synthetic seawater and a standard for pH in marine systems. *Journal of Solution Chemistry, 15*(11), 891-901.

Bates, R. G., and Guggenheim, E. A. (1960). Report on the standardization of pH and related terminology. *Pure and Applied Chemistry, 1*(1), 163-168.

Ben-Yaakov, S., Ruth, E., and Kaplan, I. R. (1974). Calcium carbonate saturation in northeastern Pacific: in situ determinations and geochemical implications. *Deep Sea Research and Oceanographic Abstracts, 21*(3), 229-243.

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

Bresnahan Jr., P. J., Wirth, T. R., Martz, T. R., Anderssen, A. J., Cyronak, T., D'Angelo, S., Pennise, J., Melville, W. K., Lenain, L., and Statom, N. (2016). A sensor package for mapping pH and oxygen from mobile platforms. *Methods in Oceanography, 17*, 1-13.

Byrne, R., Kump, L., and Cantrell, K. (1988). The influence of temperature and pH on trace metal speciation in seawater. *Marine Chemistry, 25*(2), 163-181.

Byrne, R. H. (1987). Standardization of standard buffers by visible spectrometry. *Analytical Chemistry, 59*(10), 1479-1481.

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. (2011). Estuarine and coastal ocean carbon paradox: CO_2 sinks or sites of terrestrial carbon incineration? *Annual Review of Marine Science, 3*, 123-145.

Cai, W. J., Hu., X., Huang, W. J., Jiang, L. Q., Wang, Y., Peng, T. H., and Zhang, X. (2010). Alkalinity Distributions in the western North Atlantic Ocean margins. *Journal of Geophysical Research: Oceans, 115*, 1-15.

Cai, W. J., and Wang, Y. (1998). The chemistry, fluxes, and sources of carbon dioxide in the estuarine waters of the Satilla and Altamaha Rivers, Georgia. *Limnology and Oceanography, 43*(4), 657-668.

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. (2003). Anthropogenic carbon and ocean pH. *Nature, 425*(6956), 365.

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

Clarke, J. S., Achterberg, E. P., Rérolle, V. M. C., Bey, S. A. K., Floquet, C. F. A., and Mowlem, M. C. (2015). Characterisation and deployment of an immobilised pH sensor spot towards surface ocean pH measurements. *Analytical Chimica Acta, 897*, 69-80.

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., and Feely, R. A. (1995). The role of pH measurements in modern oceanic CO_2-system characterization: Precision and thermodynamic consistency. *Deep Sea Research Part II: Topical Studies in Oceanography*, *42*(2-3), 411-429.

Cloern, J. E. (1996). Phytoplankton bloom dynamics in coastal ecosystems: a review with some general lessons from sustained investigation of San Francisco Bay, California. *Reviews of Geophysics, 34*(2), 127-168.

Czerminski, J. B., Dickson, A. G., and Bates, R. G. (1982). Thermodynamics of the dissociation of morpholinium ion in seawater from 5 to 40^{o}C. *Journal of Solution Chemistry, 11*(2), 79-89.

DeGrandpre, M. D., Baehr, M. M., and Hammar, T. R. (1999). Calibration-free optical chemical sensors. *Analytical Chemistry, 71*(6), 1152-1159.

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.

deWitt, P., and Daiber, F. C. (1974). The hydrography of the Murderkill Estuary, Delaware. *Chesapeake Science, 15*, 84-95.

Dickson, A. G. (1990). Standard potential of the reaction: AgCl (s) + 1/2 H_2 (g) = Ag (s) + HCl (aq), and the standard acidity constant of the ion HSO_4^- in synthetic sea water from 273.15 to 318.15 K. *The Journal of Chemical Thermodynamics, 22*(2), 113-127.

Dickson, A. G. (1993a). pH buffers for sea water media based on the total hydrogen ion concentration scale. *Deep-Sea Research, 40*(1), 107-118.

Dickson, A. G. (1993b). The measurement of sea water pH. *Marine Chemistry, 44*(2-4), 131-142.

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.

Dickson, A. G., and Millero, F. J. (1987). A comparison of the equilibrium constants for the dissociation of carbonic acid in seawater media. *Deep Sea Research Part A: Oceanographic Research Papers, 34*(10), 1733-1743.

Dickson, A. G., Sabine, C. L., and Christian J. R. (2007). Guide to best practices for ocean CO_2 measurements.

DNREC. (2006). TMDLs for Bacteria for the Murderkill River Watershed, Delaware.

Doney, S., Fabry, V., Feely, R., and Kleypas, J. (2009a). Ocean acidification: the other CO_2 problem. *Annual Review of Marine Science, 1,* 169-192.

Doney, S. C., Balch, W. M., Fabry, V. J., and Feely, R. A. (2009b). Ocean acidification: a critical emerging problem for the ocean sciences. *Oceanography, 22*(4), 16-25.

Dore, J. E., Lukas, R., Sadler, D. W., Church, M. J., and Karl, D. M. (2009). Physical and biogeochemical modulation of ocean acidification in the central North Pacific. *Proceedings of the National Academy of Sciences, 106*(30), 12235-12240.

Duarte, C. M., Hendriks, I. E., Moore, T. S., Olsen, Y. S., Steckbauer, A., Ramajo, L., Carstensen, J., Trotter, J. A., and McCulloch, M. (2013). Is Ocean Acidifcation an Open-Ocean Syndrome? Understanding Anthropogenic Impacts on Seawater pH. *Estuaries and Coasts, 36,* 221-236.

Dzwonkowski, B., Wong, K. C., and Ullman, W. J. (2013). Water level and velocity characteristics of a salt marsh channel in the Murderkill Estuary, Delaware. *Journal of Coastal Research, 30*(1), 63-74.

Easley, R. A., and Byrne, and R. H. (2012). Spectrophotometric calibration of pH electrodes in seawater using purified m-cresol purple. *Environ. Sci. Technol., 46*(9), 5018-5024.

French, C. R., Carr, J. J., Dougherty, E. M., Eidson, L. A. K., Reynolds, J. C., and DeGrandpre, M. D. (2002). Spectrophotometric pH measurements of freshwater. *Analytica Chimica Acta, 453*, 13-20.

Frieder, C. A., Nam, S. H., Martz, T. R., and Levin, L. A. (2012). High temporal and spatial variability of dissolved oxygen and pH in a nearshore California kelp forest. *Biogeosciences, 9*(10), 3917-3930.

Friis, K., Kortzinger, A., and Wallce, D. W. R. (2004). Spectrophotometric pH measurement in the ocean: Requirements, design, and testing of autonomous charge-coupled device detector system. *Limnol. Oceanogr. Methods, 2*, 126-136.

Gabriel, M. D., Forja, J. M., Rubio, J. A., and Gómez-Parra, A. (2005). Temperature and salinity dependence of molar absorptivities of thymol blue: Application to the spectrophotometric determination of pH in estuarine waters. *Ciencias Marinas, 31*(1B), 309-318.

González-Davila, M., Santana-Casiano, J. M., and González-Davila, E. F. (2007). Interannual variability of the upper ocean carbon cycle in the northeast Atlantic Ocean. *Geophysical Research Letters, 34*(7).

Gran, G. (1950). Determination of the equivalence point in potentiometric titrations. *Acta Chemica Scandinavica, 4*, 559-577.

Gran, G. (1952). Determination of the equivalence point in potentiometric titrations–Part II. *The Analyst, 77*(920), 661-671.

Hammer, K., Schneider, B., Kulinski, K., and Schulz-Bull, D. E. (2014). Precision and accuracy of spectrophotometric pH measurements at environmental conditions in the Baltic Sea. *Estuarine, Coastal, and Shelf Science, 146*, 24-32.

Hansson, I. (1973a). A new set of acidity constants for carbonic acid and boric acid in sea water. *Deep Sea Research and Oceanographic Abstracts, 20*(5), 461-478.

Hansson, I. (1973b). A new set of pH-scales and standard buffers for sea water. *Deep Sea Research and Oceanographic Abstracts, 20*(5), 479-491.

Harris, D. C. (2007). *Quantitative Chemical Analysis* (7th ed.). W.H. Freeman and Company.

Hawley, J. E., and Pytkowicz, R. M. (1973). Interpretation of pH measurements in concentrated electrolyte solutions. *Marine Chemistry, 1*(3), 245-250.

Head, P. C. (1985). *Practical estuarine chemistry: a handbook*. Columbia University Press.

Hofmann, G. E., Smith, J. E., Johnson, K. S., Send, U., Levin, L. A., Micheli, F., Paytan, A., Price, N. N., Peterson, B., Takeshita, Y., Matson, P. G., Derse-Crook, E., Kroeker, K. J., Gambi, M., Rivest, E. B., Frieder, C. A., Yu, P. C., and Martz, T. R. (2011). High-frequency dynamics of ocean pH: a mutli-ecosystem comparison. *PLoS One, 6*.

Hopkins, A. E., Sell, K. S., Soli, A. L., and Byrne, R. H. (2000). In-situ spectrophotometric pH measurements: the effects of pressure on thymol blue protonation and absorbance characteristics. *Marine Chemistry, 71*(1-2), 103-109.

Huang, W. J., Wang, Y., and Cai, W. J. (2012). Assessment of sample storage techniques for total alkalinity and dissolved inorganic carbon in seawater. *Limnology and Oceanography: Methods, 10*(9), 711-717.

Hulanicki, A., Glab, S., and Ingman, F. (1991). Chemical Sensors: Definitions and Classification. *Journal of Pure & Applied Chemistry, 63*(9), 1247-1250.

Janata, J. (2009). *Chemical Sensors* (2nd ed.). Springer-Verlag Press.

Jiang, L. Q., Cai, W. J., Wannikhoff, R., Wang, Y., and Luger, H. (2008). Air-sea CO_2 fluxes on the US South Atlantic Bight: spatial and seasonal variability. *Journal of Geophysical Research: Oceans (1978-2012), 113*(C7).

Joesoef, A., Huang, W. J., Gao, Y., and Cai, W. J. (2015). Air-water fluxes and sources of carbon dioxide in the Delaware Estuary: spatial and seasonal variability. *Biogeosciences, 12*, 6085-6101.

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.

Johnson, K. S., Voll, R., Curtis, C. A., and Pytkowicz, R. M. (1977). A critical examination of the NBS pH Scale and the determination of titration alkalinity. *Deep-Sea Research*, *24*(10), 915-926.

Kent County Board of Public Works (KCBPW). (2012). Kent County Regional Wastewater Treatment Facilities. http://www.co.kent.de.us/media/804272/WTF-FACT-SHEET-2013_1.pdf (accessed 25.02.2016).

Khoo, K. H., Ramette, R. W., Culberson, C. H., and Bates, R. G. (1977). Determination of hydrogen ion concentrations in seawater from 5 to 40^{o}C: Standard potentials at salinities from 20 to 45. *Analytical Chemistry*, *49*(1), 29-34.

Kim H. C., and Lee, K. (2009). Significant contribution of dissolved organic matter to seawater alkalinity. *Geophysical Research Letter*s, *36*(20).

King, D., and Kester, D. R. (1989). Determination of seawater pH from 1.5 to 8.5 using coulometric indicators. *Marine Chemistry*, *26*(1), 5-20.

Kleypas, J., Feely, R., Fabry, V., Langdon, C., Sabine, C., and Robbins, L. (2006). Impacts of ocean acidification on coral reefs and other marine calcifiers: A guide for future research, report of a workshop held by NSF. NOAA, and the USGS.

Kline, D. I., Teneva, L., Schneider, K., Miard, T., Chai, A., Marker, M., Headley, K., Opdyke, B., Nash, M., Valetich, M., Caves, J. K., Russell, B. D., Connell, S. D., Kirkwood, B. J., Brewer, P., Peltzer, E., Silverman, J., Caldeira, K., Dunbar, R. B., Koseff, J. R., Monismith, S. G., Mitchell, B. G., Dove, S., and Hoegh-Guldberg, O. (2012). A short-term in situ CO2 enrichment experiment on Heron Island (GBR). *Scientific Reports*, *2*, 413.

Ko, Y. H., Lee, K., Eom, K. H., and Han, I. S. (2016). Organic alkalinity produced by phytoplankton and its effect on the computation of the ocean carbon parameters. *Limnology and Oceanography*.

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*.

Larsen, M., Borisov, S. M., Grunwald, B., Klimant, I., and Glud, R. N. (2011). A simple and inexpensive high resolution color ratiometric planar optode imaging approach: applications to oxygen and pH sensing. *Limnology and Oceanography: Methods, 9*(9), 348-360.

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.

Liebsch, G., Kilmant, I., Krause, C., and Wolfbeis, O. S. (2001). Fluorescent imaging of pH with optical sensors using time domain dual lifetime referencing. *Analytical Chemistry, 73*(17), 4354-4363.

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

Loucaides, S., Tyrrell, T., Achterberg, E. P., Torres, R., Nightingale, P. D., Kitidis, V., Serret, P., Woodward, M., and Robinson, C. (2012). Biological and physical forcing of carbonate chemistry in an upwelling filament off northwest Africa: Results from a Lagrangian study. *Global Biogeochemical Cycles, 26*(3).

MacLeod, C. D., Doyle, H. L., and Currie, K. I. (2015). Technical Note: Maximising accuracy and minimizing cost of a potentially regulated ocean acidification simulation system. *Biogeosciences, 12*(3), 713-721.

Marion, G. M., Millero, F. J., Camoes, M. F., Spitzer, P., Feistel, R., and Chen, C. T. (2011). pH of seawater. *Marine Chemistry, 126*(1-4), 89-96.

Martz, T. (2012). SeaFET & SeapHOx Users Guide. University of California, San Diego. Scripps Institute of Oceanography.

Martz, T., Send, U., Ohman, M. D., Takeshita, Y., Bresnahan, P. J., Kim, H. J., and Nam, S. (2014). Dynamic variability of biogeochemical ratios in the Southern California Current System. *Geophysical Research Letters, 41*(7), 2496-2501.

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. *Analytica Chimica Acta, 75*, 1844-1850.

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

Matson, P. G., Martz, T. R., and Hofmann, G. E. (2011). High-frequency observations under Antarctic sea ice in the southern Ross Sea. *Antarctic Science, 23*(06), 607-613.

McClendon, J. F. (1917). The standardization of a new colorimetric method for the determination of the hydrogen ion concentration, CO_2 tension, and CO_2 and O_2 content of seawater, of animal heat, and of CO_2 of the air, with a summary of similar data on bicarbonate solutions in general. *Journal of Biological Chemistry*, *30*(2), 265-288.

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

Midorikawa, T., Ishii, M., Saito, S., Sasano, D., Kosugi, N., Motoi, T., Kamiya, H., Nakadate, A., Nemoto, K., and Inoue, H. Y. (2010). Decreasing pH trend estimated from 25-yr time series of carbonate parameters in the western North Pacific. *Tellus B*, *62*(5), 649-659.

Mikhelson, K. N. (2013). *Ion-Selective Electrodes.* Springer-Verlag Press.

Millero, F. J. (1986). The pH of estuarine waters. *Limnology and Oceanography, 31*(4), 839–847.

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

Millero, F. J., Zhang, J. Z., Fiol, S., Sotolongo, S., Roy, R. N., Lee, K., and Mane, S. (1993). The use of buffers to measure pH for seawater. *Marine Chemistry, 22*(2-4), 143-152.

Mosley, L. M., Husheer, S. L. G., and Hunter, K. A. (2004). Spectrophotometric pH measurement in estuaries using thymol blue and m-cresol purple. *Marine Chemistry*, *91*(1-4), 175-186.

Nemzer B. V., and Dickson, A. G. (2005). The stability and reproducibility of tris buffers in synthetic seawater. *Marine Chemistry*, *96*(3-4), 237-242.

Newton, J. A., Feely, R. A., Jewett, E. B., Williamson, P., and Mathis, J. (2015). Global Ocean Acidification Observing Network: Requirements and Governance Plan. 2nd ed., *GOA-ON*.

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 impacts on calcifying organisms. *Nature*, *437*(7059), 681-686.

Papadimitrou, S., Loucaides, S., Réroll,e V., Achterberg, E. P., Dickson, A. G., Mowlem, M., and Kennedy, H. (2016). The measurement of pH in saline and hypersaline media at sub-zero temperatures: Characterization of Tris buffers. *Marine Chemistry*.

Patsavas, M. C., Byrne, R. H., and Liu, X. (2013a). 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., and Liu, X. (2013b). Physical-chemical characterization of purified cresol red for spectrophotometric pH measurements in seawater. *Marine Chemistry, 155*, 158-164.

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 US coastal cruises. *Marine Chemistry, 176*, 9-20.

Peltzer, E. (2007). Model II least squares fit: lsqfitma.m. http://www.mbari.org/staff/etp3/regress/lsqfitma.m.

Pierrot, D., Lewis, E., and Wallace, D. W. R. (2006). CO2SYS DOS Program Developed for CO_2 System Calculations. *ORNL/CDIAC-105, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, US Department of Energy, Oak Ridge, TN*.

Pratt, K. W. (2014). Measurement of pH_T values of TRIS Buffers in artificial seawater at varying mole rations of TRIS:TRIS-HCl. *Marine Chemistry, 162*, 89-95.

Price, N. N., Martz, T. R., Brainard, R. E., and Smith, J. E. (2012). Diel variability in sewatet pH relates to calcification and benthic community structure on coral reefs. *PloS one, 7*(8), e43843.

Pytkowicz, R. M., Ingle, S. E., and Mehrbach, C. (1974). Invariance of apparent of equilibrium constants with pH. *Limnology and Oceanography, 19*(4), 665-669.

Ramette, R. W., Culberson, C. H., and Bates, R. G. (1977). Acid-base properties of tris(hydroxymethyl) aminomethance (tris) buffers in sea water from 5 to 40^{o}C. *Analytical Chemistry, 49*(6), 867-870.

Raven, J., Caldeira, K., Eldersfield, H., Hoegh-Guldberg, O., Liss, P., Riebesell, U., Shepherd, J., Turley, C., and Watson, A. (2005). Ocean acidification due to increasing atmospheric carbon dioxide. *The Royal Society*.

Rérolle, V., Ruiz-Pino, D., Rafizadeh, M., Loucaides, S., Papadimitrou, S., Mowlem, M., and Chen, J. (2016). Measuring pH in the Arctic Ocean: Colorimetric method or SeaFET? *Methods in Oceanography, 17*, 32-49.

Rérolle, V. M. C., Floquet, C. F. A., Harris, A. J., Mowlem, M. C., Bellerby, R. R., and Achterberg, E. P. (2013). Development of a colorimetric microfluidic pH sensor for autonomous seawater measurements. *Analytica Chimica Acta, 786*, 124-131.

Rérolle, V. M. C., Floquet, C. F. A., Mowlem, M. C., Bellerby, R. R., Connelly, D. P., and Achterberg, E. P. (2012). Seawater-pH measurements for ocean-acidification observations. *Trends in Analytical Chemistry, 40*, 146-157.

Ribas-Ribas, M., Rérolle, V., Bakker, D. C. E., Kitidis, V., Lee, G., Brown, I., Achterberg, E. P., Hardman-Mountford, N., and Tyrrell, T. (2014). Intercomparison of carbonate chemistry measurements on a cruise in northwestern European shelf seas. *Biogeosciences, 11*, 4339-4355.

Rivest, E. B., O'Brien, M., Kapsenberg, L., Gotschalk, C. C., Blanchette, C. A., Hoshijima, U., and Hofmann, G. E. (2016). Beyond the benchtop and the benthos: Dataset management planning and design for time series of ocean carbonate chemistry associated with Durafet-based pH sensors. *Ecological Informatics*.

Robert-Baldo, G. L., Morris, M. J., and Byrne, R. H. (1985). Spectrophotometric determination of seawater pH using phenol red. *Analytical Chemistry, 57*(13), 2564-2567.

Sandifer, J., and Voycheck, J. (1999). A review of biosensor and industrial applications of pH-ISFETs and an evaluation of Honeywell's "DuraFET." *Microchimica Acta, 131*, 91-98.

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*(1-2), 18-28.

Sharp, J. H., Yoshiyama, K., Parker, A. E., Schwartz, M. C., Curless, S., Beauregard, A. Y., Ossolinski, J. E., and Davis, A. R. (2009). A biogeochemical view of estuarine eutrophication: lessons from the Delaware Estuary. *Estuaries and Coasts*, *32*(6), 1023-1043.

Sillen, G. (1968). Master variables and activity scales.

Skoog, D., Holler, F. J., and Crouch, S. R. (2006). *Principles of Instrumental Analysis* (6th ed.). Cengage Learning, Inc.

Soli, A. L., Brody, J. P., 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.

Sørensen, S. P. L. (1909). Enzymstudien. II. Mitteilung. Uber die Messung and die Bedeutung de Wasserstoffionenkonzentration bei enzymatische Prozessen. *Biochem. Z., 21*, 131-304.

Sørensen S. P. L., and Linderstrøm-Lang, K. (1924). On the determination and values of $\pi 0$ in electromotoric measurements of hydrogen ion concentrations. *Compt. rend. du Lab. Carlsberg, 15*, 1-40.

Takeshita, Y., Martz, T. R., Johnson, K. S., and Dickson, A. G. (2014). Characterization of an Ion-Selective Field Effect Transistor and Chloride Ion-Selective Electrodes for pH Measurements in Seawater. *Analytical Chemistry, 86*(22), 1189-1195.

Ullman, W. J., Aufdenkampe, A., Hays, R. L., and Dix, S. (2013). Nutrient Exchange between a Salt Marsh and the Murderkill Estuary, Kent County, Delaware.

Voynova, Y. G., Lebaron, K. C., Barnes, R. T., and Ullman, W. J. (2015). *In situ* response of bay productivity to nutrient loading from a small tributary: The Delaware Bay-Murderkill Estuary tidally-coupled biogeochemical reactor. *Estuarine, Coastal, and Shelf Science, 160*, 33-48.

Weisberg, S. B., Bednarsek, N., Feely, R. A., Chan, F., Boehm, A. B., Sutula, M., Ruesink, J. L., Hales, B., Largier, J. L., and Newton, J. A. (2016). Water quality criteria for an acidifying ocean: Challenges and opportunities for improvement. *Ocean & Coastal Management, 126*, 31-41.

Whitfield, M., Butler, R. A., and Covington A. K. (1985). The determination of pH in estuarine waters I. Definition of pH scales and selection of buffers. *Oceanological Acta, 8*(4), 423-432.

Wong, K. C., Dzwonkowski, B., and Ullman, W. J. (2009). Temporal and spatial variability of sea level and volume flux in the Murderkill Estuary. *Estuarine, Coastal, and Shelf Science, 84*, 440-446.

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.

Yao, W., and Byrne, R. H. (1998). Simplified seawater alkalinity analysis: Use of linear array spectrometers. *Deep Sea Research Part I: Oceanographic Research Papers, 45*(8), 1383-1392.

Yao, W., and Byrne R. H. (2001). Spectrophotometric determination of freshwater pH using bromocresol purple and phenol red. *Environ. Sci. Technol., 35*(6), 1197-1201.

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.

Zeebe, R. E., and Wolf-Gladrow, D. A. (2001). *CO_2 in seawater: equilibrium, kinetics, isotopes* (No. 65). Gulf Professional Publishing.

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.