See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/320803425

Assessment of the suitability of Durafet-based sensors for pH measurement in dynamic estuarine environments

Article in Estuarine Coastal and Shelf Science · November 2017

DOI: 10.1016/j.ecss.2017.10.020	
---------------------------------	--



Some of the authors of this publication are also working on these related projects:

Project GulfCarbon View project

Georgia Coastal Ecosystems - Long-Term Ecological Research Project View project

Estuarine, Coastal and Shelf Science 200 (2018) 152-168

Contents lists available at ScienceDirect

Estuarine, Coastal and Shelf Science

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

Assessment of the suitability of Durafet-based sensors for pH measurement in dynamic estuarine environments

Stephen F. Gonski ^{a, 1}, Wei-Jun Cai ^{a, *}, William J. Ullman ^a, Andrew Joesoef ^a, Christopher R. Main ^b, D. Tye Pettay ^a, Todd R. Martz ^c

^a School of Marine Science and Policy, University of Delaware, Newark, DE 19716, USA

^b Delaware Department of Natural Resources and Environmental Control, 89 Kings Highway, Dover, DE 19901, USA

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

ARTICLE INFO

Article history: Received 17 May 2017 Received in revised form 17 October 2017 Accepted 30 October 2017 Available online 2 November 2017

Keywords: pH Sensor Evaluation ISFET Acidification Estuaries

ABSTRACT

The suitability of the Honeywell Durafet to the measurement of pH in productive, high-fouling, and highly-turbid estuarine environments was investigated at the confluence of the Murderkill Estuary and Delaware Bay (Delaware, USA). Three different flow configurations of the SeapHOx sensor equipped with a Honeywell Durafet and its integrated internal (Ag/AgCl reference electrode containing a 4.5 M KCl gel liquid junction) and external (solid-state chloride ion selective electrode, CI-ISE) reference electrodes were deployed for four periods between April 2015 and September 2016. In this environment, the Honeywell Durafet proved capable of making high-resolution and high-frequency pH measurements on the total scale between pH 6.8 and 8.4. Natural pH fluctuations of >1 pH unit were routinely captured over a range of timescales. The sensor pH collected between May and August 2016 using the most refined SeapHOx configuration exhibited good agreement with multiple sets of independently measured reference pH values. When deployed in conjunction with rigorous discrete sampling and calibration schemes, the sensor pH had a root-mean squared error ranging between 0.011 and 0.036 pH units across a wide range of salinity relative to both pH_T calculated from measured dissolved inorganic carbon and total alkalinity and pH_{NBS} measured with a glass electrode corrected to pH_T at in situ conditions. The present work demonstrates the viability of the Honeywell Durafet to the measurement of pH to within the weather-level precision defined by the Global Ocean Acidification Observing Network (GOA-ON, \leq 0.02 pH units) as a part of future estuarine CO_2 chemistry studies undertaken in dynamic environments. © 2017 Elsevier Ltd. All rights reserved.

1. Introduction

The rapid rise in atmospheric carbon dioxide (CO_2) 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, 2009b; Kleypas et al., 2006). As atmospheric CO₂ concentrations increase, up to one-third of the released CO₂ 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 also in natural waters that link the terrestrial and marine biospheres such as estuaries and the coastal ocean where natural pH variations are also great (Duarte et al., 2013). In these latter systems, eutrophication often enhances this acidification due to oxidation of excess organic matter supported by terrestrially derived nutrients (Cai et al., 2011; Feely et al., 2010; Wallace et al., 2014). pH exerts control over so many biogeochemical processes that it is often referred to as the "master" variable in aquatic systems (Bates, 1982; Byrne et al., 1988; Clayton et al., 1995; Morel and Hering, 1993; Orr et al., 2005; Stumm and Morgan, 1996). Accordingly, pH, its measurement, and its effects on marine, estuarine, and freshwater systems is of great interest to scientists and

* Corresponding author.







E-mail address: wcai@udel.edu (W.-J. Cai).

¹ Present address: Department of Geography, University of Calgary, AB, Canada.

environmental managers.

The pH of natural waters can be defined as either an activity or a concentration of hydrogen ions in solution (Zeebe and Wolf-Gladrow, 2001). The merits and limitations of each convention as well as their associated pH scales have been widely discussed over the past 45 years (Bates, 1973, 1975; Dickson, 1984, 1993; Dickson et al., 2016; Hansson, 1973a; 1973b; Marion et al., 2011; Waters and Millero, 2013), but still lead to confusion among most of the scientific community (Dickson et al., 2016; Millero, 1986). At present, the concentration convention is mostly used in seawater studies and the activity convention is used primarily in freshwater studies (Zeebe and Wolf-Gladrow, 2001). The various pH scales and interconversions are discussed in detail by Marion et al. (2011).

There exist two principal benchtop methodologies for the determination of the pH of natural waters: (1) electrochemical methods using potentiometric cells and (2) spectrophotometric methods using diprotic sulfonepthalein pH-sensitive colorimetric indicator dyes (Dickson et al., 2007). These measurement methodologies are reviewed in detail by Rérolle et al. (2012).

A continuing problem with these two methodologies is that they are largely limited to salinity (S) > 20 in seawater where TRIS buffer standards in artificial seawater (DelValls and Dickson, 1998) are well-calibrated (Dickson et al., 2016). In addition, the indicator dyes used in the spectrophotometric measurements require purification to avoid pH errors of up to 0.02 pH units (Yao et al., 2007). When using unpurified *meta*-Cresol Purple, however, these pH errors can be minimized by employing the corrective procedures summarized in Douglas and Byrne (2017a). Further, the thermodynamic constants of available purified indicator dyes are poorly known below S = 20 (Liu et al., 2011; Patsavas et al., 2013b). These methodological shortcomings propagate through all aspects of the measurement and use of pH in brackish waters of S < 20 (Dickson et al., 2016; Millero, 1986).

Spurred on by the findings of Yao et al. (2007), the seawater pH community is collectively seeking to develop a high-precision spectrophotometric methodology using purified indicator dyes for pH measurement over the full temperature and salinity range of natural waters (DeGrandpre et al., 2014; Lai et al., 2016; Liu et al., 2011; Patsavas et al., 2013a, 2013b; Soli et al., 2013). Parallel work for pH measurements under near-zero temperatures (DeGrandpre et al., 2014; Loucaides et al., 2017; Papadimitriou et al., 2016; Rérolle et al., 2016), at high pressures (Hopkins et al., 2000; Rodriguez et al., 2015; Soli et al., 2013; Takeshita et al., 2016a), and below open-ocean salinities (French et al., 2002; Gabriel et al., 2005; Gallego-Urrea and Turner, 2017; Hammer et al., 2014; Lai et al., 2016; Mosley et al., 2004; Yao and Byrne, 2001) has also been completed. The availability of a new series of variable mole-ratio TRIS:TRIS-HCl buffers calibrated using a Harned Cell have produced buffers of similar composition to natural waters of S < 20that possess the required buffering capacity needed for method development (Pratt, 2014). However, please see a recent paper by Douglas and Byrne (2017b) for further discussion of the new spectrophotometric pH measurement methodology using purified meta-Cresol Purple for estuarine waters.

This continual refinement of benchtop pH measurement methodologies and the subsequent development of autonomous sensor packages based on those methodologies has stimulated research centered on pH and its controlling processes in natural waters (Byrne, 2014). High-frequency measurements can help resolve the temporal and spatial patterns and trends that traditional discrete water sampling is unable to capture (Johnson et al., 2007) especially in estuaries that experience substantial simultaneous temperature, salinity, and biogeochemical variability. Presently, two classes of *in situ* chemical sensors are available for pH measurement for high-frequency field studies: Reagent-based

Optical Chemical Sensors (ROCS) and Electrochemical Sensors (Hulanicki et al., 1991).

Recently, pH sensor development has been steered in favor of ROCS over Electrochemical Sensors (Martz et al., 2010). The extensive literature of pH sensors based on ROCS designs developed in the last 10-15 years utilizing various optical properties like absorbance (de Vargas Sansalvador et al., 2016; Martz et al., 2003; Rérolle et al., 2013: Seidel et al., 2008) and fluorescence (Clarke et al., 2015) or luminescence (Larsen et al., 2011) in conjunction with immobilized indicator dyes supports this view. However, attention has returned toward Electrochemical Sensors like the Honeywell Durafet due to the limitations of pH sensors based on ROCS designs associated with their complex operation, slow response times, elevated power requirements, reagent consumption and storage, optical interferences due to turbidity, and the limitations of available indicator dyes (Bagshaw et al., 2016; Martz et al., 2010; Rérolle et al., 2016). For a more thorough comparison of the two classes of pH sensors, the reader is referred to Martz et al. (2010) and Rérolle et al. (2016).

The Honeywell Durafet is a pH-sensitive Ion-Selective Field Effect Transistor (ISFET) that has been successfully used for pH measurement in seawater CO₂ chemistry and ocean acidification studies (Bresnahan et al., 2014; Martz et al., 2010). Modified Honeywell ISFETs have also been integrated into a novel solid-state sensor capable of rapid, simultaneous measurements of pH and total alkalinity in seawater (Briggs et al., 2017). The Honeywell Durafet is suitable for pH measurement in marine environments due to its quick response time (Martz et al., 2010), consistent linear response with temperature (Takeshita et al., 2014), and good signal stability (Martz et al., 2010; Sandifer and Voycheck, 1999). In controlled laboratory settings, the Honeywell Durafet is capable of precisions of better than ±0.005 pH units over weeks to months (Martz et al., 2010) and ±0.01 pH units under in situ open ocean conditions (Bresnahan et al., 2014). In coastal waters, experienced and well-trained sensor operators can routinely achieve agreement between sensor and discrete bottle sample measurements better than ± 0.02 pH units (McLaughlin et al., 2017).

Besides the use of the Honeywell Durafet under open-ocean conditions, its use for pH measurement 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 (Hawkings et al., 2016) has been demonstrated. The Honeywell Durafet has also been used underway on hydrographic cruises (Rérolle et al., 2016), for profiling down to 3000 m using the Deep Sea Durafet (Johnson et al., 2016), and on mobile oceano-graphic monitoring platforms such as the WavepHOx (Bresnahan et al., 2016), Wave Gliders (Chavez et al., 2017), and Argo Floats (Johnson et al., 2016). The versatility of the Honeywell Durafet suggests it will become the sensor of choice in the future to complement the application of a spectrophotometric pH measurement technique for dilute estuarine waters.

There is, however, a clear need to characterize the performance of the Honeywell Durafet over a broad range of time varying salinity to determine its reliability in real estuarine settings. In this work, we report our evaluation of the performance of the Honeywell Durafet integrated into a SeapHOx sensor package (originally designed and assembled by Todd R. Martz of Scripps Institution of Oceanography and recently commercialized by Sea-Bird Scientific, Bellevue, Washington, USA; described in section 2.3) in a tidallyforced, productive, highly-turbid, and high-fouling estuarine environment characterized by an extensive salinity range. We outline modifications to sensor flow-path design specific to our estuarine environment, considerations for the design of future long-term deployments of the SeapHOx sensor package in similar environments, and results from two month-long deployments carried out between 09 May 2016 to 09 June 2016 (spring) and 20 July 2016 to 24 August 2016 (summer) at the confluence of the Murderkill Estuary and Delaware Bay (Delaware, USA).

2. Materials & methods

2.1. Sensor operation

As designed, the Honeywell Durafet pH sensor in the SeapHOx calculates and reports two pH values using its two integrated reference electrodes – FET|INT (Ag/AgCl reference containing a 4.5 M KCl gel liquid junction) and FET|EXT (Cl-ISE) – designated by the notations of pH^{INT} and pH^{EXT}, respectively. pH^{INT} and pH^{EXT} are calculated via:

$$pH^{INT} = \frac{\left(E_{INT} - E_{INT}^*\right)}{S}$$
(1)

$$pH^{EXT} = \frac{\left(E_{EXT} - E_{EXT}^*\right) + S \times log(\gamma_H \gamma_{CI} m_{CI})}{S}$$
(2)

where E_{INT} and E_{EXT} are the measured sensor voltages, E_{INT}^* and E_{EXT}^* are the calibration constants, γ_i is the ion activity coefficient of either H⁺ or Cl⁻, m_{Cl} is the molal concentration of Cl⁻, and S is the temperature-dependent Nernst slope. The temperature-dependent Nernst slope is calculated via:

$$S = \frac{RT}{F} \times \ln(10) \tag{3}$$

where R is the gas constant (8.3145 J mol⁻¹K⁻¹), T is temperature in Kelvin, and F is the Faraday constant (96485 C mol⁻¹) (Martz et al., 2010). Temperature-dependent standard potentials for this electrode configuration are instead referred to as calibration constants (Bresnahan et al., 2014) designated by an asterisk (*) (e.g. E_{sensor}^{o}) versus a nought symbol (°) (e.g. E_{sensor}^{o}) (Martz et al., 2010). Using the measured voltages, *in situ* temperature, and salinity, pH^{INT} and pH^{EXT} are calculated assuming a 100% Nernst slope (e.g. 59.16 mV/ pH at 25 °C) and a constant dE_{sensor}^*/dT (e.g. $dE_{INT}^*/dT = -0.001101 \text{ V/°C}$ and $dE_{EXT}^*/dT = -0.001048 \text{ V/°C}$) (Bresnahan et al., 2014; Martz et al., 2010). For a more detailed discussion of the operating principle of the Honeywell Durafet and its two reference electrodes, see Martz et al. (2010) and Bresnahan et al. (2014).

2.2. Study site

The SeapHOx was deployed at Bowers. Delaware (Lat. 39.05°N. Long, 75.39°W) at the confluence of the larger Delaware Estuary and its smaller, well-mixed, and turbid tributary, the Murderkill Estuary, approximately 39 km upstream of the Atlantic Ocean (Voynova et al., 2015). The Murderkill watershed covers 275 km² and is largely agricultural (55%) with smaller amounts of forest (11%), urban areas (14%), and wetland (17% of which 60% are freshwater and 36% are tidal) (Ullman et al., 2013). The estuary is nutrient-rich due to the agricultural land use of its watershed and discharge from a large wastewater treatment facility approximately 10.5 km upstream of Delaware Bay that serves a more urban population of 130,000 people both north and south of the Murderkill watershed (Voynova et al., 2015). The estuary has an average channel width of 50 m and average depth of 4.5 m (Wong et al., 2009). Delaware Bay is one of the largest estuaries in the USA, 215 km long, 18–44 km wide, and drains a 36,570 km² watershed that includes parts of Delaware, New Jersey, Pennsylvania, and the

state of New York (Sharp et al., 2009; Wong et al., 2009).

2.3. Sensor packages

The SeapHOx sensor package uses an actively flushed flow driven by a Sea-Bird Electronics 5M submersible pump and includes sensors for temperature, salinity (Sea-Bird Electronics Conductivity-Temperature Sensor – SBE37), pH (Honeywell Durafet), and oxygen (Aanderaa Data Instruments 4835 Optode) (Bresnahan et al., 2014). Two SeapHOx units were used in this study: SP033 from April 2015 to August 2015 and SP053 from September 2015 to August 2016. When deployed, the configuration of SP033 (Fig. 1) was simply the SeapHOx as designed and assembled for open-ocean deployments. This sensor configuration proved to be less-suited for our estuarine application. Subsequently, SP053 (Fig. 1) was reconfigured with a modified flow path consisting of tubing with a uniform cross-sectional area and rounded connections between all sensor components. The flow path of this configuration was considerably longer than the original and the pump time of the instrument was increased from 25 s (recommended for open-ocean deployments by Martz (2012)) to 60-70 s to accommodate this modification. The increased flushing time substantially reduced the build-up of sediment in the flow housing. Finally, a sampling interval of 30 min was chosen.

Prior to each sensor redeployment, steps were taken to ensure proper operation of all sensor components. Briefly, the SBE 5M submersible pump was removed, submerged, and tested. Optode operation was also tested in air at ~100% oxygen saturation as recommended by Aanderaa Data Instruments AS (Nesttun, Norway). Using filtered seawater (salinity ~ 29-31), tests were conducted on the response of both reference electrodes to ensure measurements returned to within nominal voltage ranges for each electrode as detailed in Bresnahan et al. (2014). To minimize biofouling, 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-foil tape. Using filtered seawater taken from the lower Delaware Bay (salinity $\sim 29-31$), the recommended pre-deployment electrode conditioning protocols were followed before each sensor deployment. Detailed sensor preparation procedures are given in Bresnahan et al. (2014) and Rivest et al. (2016).

2.4. Field deployment

The SeapHOx units were deployed alongside the suite of colocated sensors comprising the Kent County Land-Ocean Biogeochemical Observatory (LOBO) (Sea-Bird Scientific, http://kentcou nty.loboviz.com/) except during December 2015—March 2016 when the LOBO sensors were undergoing annual maintenance and calibration. The LOBO deployment platform consists of an aluminum cage with removable struts mounted on a trolley made of 1" thick galvanized steel that can be winched up and down along an I-Beam track attached to a pier to a fixed vertical position, approximately 0.8 m above the estuary floor. Prior to deployment, all the aluminum components of the deployment platform were coated with antifouling paint. (Interlux Paint, LLC., Union, NJ, USA).

SeapHOx SP053 was deployed upside down so sediment would accumulate away from the electrodes inside the flow housing. The copper pipe inserts integrated as a secondary antifouling measure into the flow path of SeapHOx SP033 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° angle. The Ushaped inlet was placed at the sensor inflow point to prevent the passive settling of sediment and/or active invasion by fouling organisms into the sensor flow path. Comprehensive sensor



Fig. 1. Schematic plots of sensor configurations used in the Murderkill Estuary-Delaware Bay SeapHOx deployments. Left schematic plot shows SeapHOx SP033 in its original configuration (Configuration v1.0) deployed between 08 April 2015 and 26 August 2015. Right panel shows SeapHOx SP053 in Configuration v3.0 deployed between 09 May 2016 and 24 August 2016. Blue arrows represent the sensor flow path and indicate the direction of water flow through the sensor. Water flow in Configuration v1.0 is bottom-up through: (A) the conductivity channel of the (B) Sea-Bird Electronics Conductivity-Temperature Sensor (SBE37), (C) the Sea-Bird Electronics 5M Submersible Pump, and (D) the sensor flow housing where the electrodes are located. Water flow in Configuration v3.0 is top-down entering through its U-shaped copper pipe inlet and follows the same flow sequence as Configuration v1.0. In both configurations, the sensor flow housing fills bottom-to-top. For further context, see the Supplemental Materials for examples of the fully-assembled SeapHOx packages and a list of the materials used in the modified flow path of Configuration v3.0. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

maintenance was performed on the SeapHOx and the LOBO every 7–10 days between May 2016 and August 2016.

2.5. Sampling approach

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). High-frequency tidal fluctuations and low-frequency subtidal controls associated with spring-neap tides, winds, and large storm events drive local biogeochemistry (Dzwonkowski et al., 2013; Voynova et al., 2015; Wong et al., 2009). A tidal cycle sampling scheme was adopted for discrete sampling: samples were collected over the course of 10–12 h periods bracketing at least one low tide and one high tide. Using this sampling scheme, the full range of salinity and pH observed at this site were captured. This reduced the chances of introducing bias due to salinity, pH, or system endmember into reference pH datasets used for sensor calibration. To this end, sampling trips were synchronized to periods with large expected salinity and pH variation, such as periods with high freshwater flows following large storms. Large storm events represent significant freshwater inputs to the watershed and result in lower salinities at the sampling site during the end of the ebb tide (Voynova et al., 2015).

2.6. Sensor calibration

Conventional calibration protocols for electrochemical pH measurement methods consist of employing a series of traceable reference materials (e.g. TRIS Buffers in artificial seawater) representative of the composition of natural waters at those salinities to calibrate and periodically recalibrate the electrodes or sensors. When performing an electrochemical pH calibration in low and intermediate salinities, for example, it has been recommended that a single standard reference buffer at an intermediate salinity between freshwater and seawater (e.g. $15 \le S \le 20$) or a series of standard buffers spanning the full salinity range encountered be used (Whitfield et al., 1985). The former method leads to measurement uncertainties associated with the deviation from the salinity of the single calibration standard (Whitfield et al., 1985), while the latter method requires the preparation and preservation of many standard buffers (Easley and Byrne, 2012). Easley and Byrne (2012) demonstrated that a 10-unit difference in the salinity between the standard buffer solution and a natural seawater sample can yield uncertainties of up to 0.028 pH units and even a difference of only 2 salinity units can result in errors of 0.005 pH units. This highlights the need for finer salinity incrementation between successive buffers and the deficiency of the use of the single calibration standard (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 discrete bottle samples coinciding with the time of sensor measurements, and the determination of pH of those samples using established benchtop methods. Using the measured sensor voltages, in situ temperature, salinity, and the discrete sample pH corrected to in situ temperature, calibration constants specific to each reference electrode $(E_{INT}^*$ and $E_{EXT}^*)$ are calculated based on an average of all discrete samples. A single calibration constant specific to each reference electrode is then applied retroactively to the raw sensor pH. This redefines the contribution of the calibration constant to the voltage converted to pH via the temperature-dependent Nernst slope in equations (1) and (2) thereby displacing and aligning the sensor pH with the discrete sample pH data (Bresnahan et al., 2014). A comparison of the sensor pH and the reference pH is evaluated based on its Model II least squares fit (Peltzer, 2007) which generate fit parameters comprised of a sensor offset or intercept (c_0) and a sensor gain or slope (c_1) . Under optimal conditions, these quantities would equal 0 and 1, respectively (Bresnahan et al., 2014). The root-mean squared error (RMSE) calculated from this fit 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. For a more thorough discussion of the calibration procedure, the reader is referred to 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 number and quality of the discrete samples collected over the course of a sensor deployment. An optimal discrete sampling scheme is characterized by > 10 usable discrete bottle samples for sensor deployments of several months or more (Rivest et al., 2016). In contrast, calibrating a sensor time series of a similar length to only one discrete bottle sample can produce sensor inaccuracies of ~0.1 pH units and is not recommended (Bresnahan et al., 2014, 2016) 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 (McLaughlin et al., 2017). 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 indicator dyes (Rivest et al., 2016) known for their high precision, simple analytical procedures, and minimal operator interaction (Carter et al., 2013; DeGrandpre et al., 2014). On the other hand, the application of spectrophotometric pH measurement methodologies is not straightforward in S < 20 waters as previously discussed. Alternatively, pH can be measured via glass electrode potentiometry or calculated from two marine CO₂ 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.

2.6.1. Field measurements

Dissolved inorganic carbon (DIC), total alkalinity (TA), and pH_{NBS} were measured at the sensor deployment site on 01 June 2016 and 02 August 2016. Discrete water column samples were collected every 30 min coinciding with sensor measurements on the hour and half-hour. Samples were collected using a peristaltic pump by

way of a 3.75 m length of tubing lowered into the water to the approximate depth of the SeapHOx and sufficiently weighted to maintain a fixed position. Measurements of *in situ* temperature, salinity, and pressure were taken from the SeapHOx.

2.6.2. Analytical methods

Samples for DIC and TA were typically collected in duplicate following filtration through Whatman 0.45 um Polvethersulfone (PES) filters (GE Healthcare Bio-Sciences, Pittsburgh, PA, USA) by bottom-filling and overflowing into triple-rinsed 250-mL borosilicate glass bottles. The samples were then fixed with 100 µL of saturated mercuric chloride (HgCl₂) solution, and stored for later analysis. Approximately every 30 min, filters were regenerated by rinsing the filter backwards then forwards three times using a reservoir of deionized (DI) water and a second peristaltic pump between sampling cycles. Upon returning from the field, the samples were stored at 4 °C (Cai and Wang, 1998; Jiang et al., 2008). DIC was determined through acid extraction by quantifying the released CO₂ using an infrared gas analyzer (AS-C3 Apollo Scitech). TA was measured by Gran Titration (Gran, 1950, 1952) using a semiautomated open-cell titration system (AS-ALK2 Apollo SciTech) (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) with a precision of $\pm 2.2 \ \mu mol \ kg^{-1}$.

Unfiltered samples for pH_{NBS} were collected in duplicate via bottom-filling and overflowing into triple-rinsed narrow-neck 125mL clear soda glass bottles. pH_{NBS} was measured using an Orion Dual Star pH/ISE Benchtop Meter equipped with an Orion 8302BNUMD Ross Ultra Glass Triode pH/ATC Combination Electrode (Thermo Fisher Scientific Inc., Beverly, MA, USA) within 3–5 min of sample collection at a temperature within ±0.3 °C of *in situ* temperature recorded by the SeapHOx. The glass electrode was calibrated every 2–3 h using three National Bureau of Standards (NBS) traceable low ionic strength 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 min prior to the calibration. During all pH_{NBS} measurements, parafilm was used to seal the bottle neck and minimize CO₂-exchange between the sample and the surrounding atmosphere.

A system of rigorous *in situ* deployment calibrations was pursued using two sets of independent reference pH. 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. The inorganic carbon dissociation constants from Millero et al. (2006) were found most suited to calculations of pH_T in estuarine waters (S < 30).

The second independent reference pH was pH_T calculated from the field measurements of pH_{NBS} made on the discrete bottle samples with a glass electrode. In order to compare these bottle samples to the DIC-TA bottles samples and the *in situ* sensor data, we used the Excel macro CO2SYS (Pierrot et al., 2006) to correct the discrete pH_{NBS} measurements to the total scale (pH_T) at *in situ* temperature, salinity, and pressure using the following steps:

1) Correct the discrete pH_{NBS} measurements from the measurement temperature (pH_{NBS}^{meas}) to *in situ* temperature recorded by the SeapHOx by applying the difference in pH_{NBS} calculated from DIC and TA at the measurement and *in situ* temperatures (ΔpH_{NBS}^{temp}) to generate the temperature-corrected discrete pH_{NBS} data (pH_{NBS}^{field}).

2) Apply a correction to pH_{NBS}^{field} based on the difference between pH_{NBS} and pH_T calculated from DIC and TA at *in situ* conditions solely attributed to the difference in the pH scales (ΔpH_{scales}) to generate the second independent reference pH (pH_{elec}^{disc}) on the total scale.

Over the course of our work, the values of ΔpH_{scales} varied roughly between 0.145 and 0.155 pH units. The equations used to calculate pH_{elec}^{disc} together with the accompanying study-specific terms are discussed in greater detail in the Glossary included with the Supplementary Materials.

2.7. Assumptions & limitations

The remoteness of many open-ocean sensor deployment sites often limits the opportunities to maintain, calibrate, and verify the performance of the sensors (Rivest et al., 2016). Thus, each deployment is traditionally treated as one continuous time series with benchmark calibration samples typically collected directly following sensor deployments along with pre-deployment calibrations and/or post-deployment calibrations sometimes attached to further gauge sensor drift and validate sensor performance (Bresnahan et al., 2014). This approach, however, is not preferable when the sensors are serviced, maintained, and redeployed on a weekly basis. With an estuarine sensor deployment in an accessible location, more frequent maintenance, calibration, and verification is possible and, due to higher fouling potential, probably necessary. However, frequent and targeted discrete sampling results in subsets of data with different suites of independent calibrations using calibration constants set to average values based on independent sets of discrete bottle samples that must be aggregated over the total deployment period. In our opinion, this alternative approach is more practical for sensor calibrations in dynamic estuarine environments when frequent sensor maintenance is necessary.

The principal analytical assumption of this study is that the Honeywell Durafet exhibits the full Nernstian response over the full range of salinity encountered (S = 3–30). Previous work has demonstrated that the dual-reference electrode configuration of the SeapHOx repeatedly exhibited this response over a salinity range of 20–35 in seawater and that the Br⁻ sensitivity of the CI-ISE was only <0.003 pH units (Takeshita et al., 2014). The CI-ISE was assumed to operate properly if the Cl⁻ : Br⁻ ratio remained constant, as is the case for seawater and all but the most severe dilutions of seawater (Rérolle et al., 2016). Finally, it was assumed the empirical function used to calculate the mean activity coefficient of HCI ($\gamma_{\pm HCI}$) reported by Khoo et al. (1977) was valid outside of its published salinity range (S = 20–45).

The "internal" (Ag/AgCl) reference electrode contains a liquid junction, which could impart measurable pH errors of varying magnitude (Bates, 1973; Dickson, 1993) to the pH^{INT} time-series due to liquid junction potential effects (Bresnahan et al., 2014; Martz et al., 2010). Hysteresis due to the effects of liquid potential can be especially problematic in a river mouth environment that routinely experiences large salinity fluctuations due to freshwater inputs (Bresnahan et al., 2014).

3. Results & discussion

3.1. Discrete sample pH comparisons

The four measures of pH (pH_{raw}^{INT} , pH_{raw}^{EXT} , pH_{DIC-TA}^{disc} , pH_{elec}^{disc}) coinciding with the sampling trips on 01 June 2016 (Fig. 2a) and 02 August 2016 (Fig. 2b) were compared as a preliminary check of the quality of the discrete sampling scheme. All four measures of pH similarly showed a period of slow decline on the ebb tide and a

more rapid rise on the flood tide. The pH minima and maxima coincided with the observed salinity minima and maxima. There is a noticeable disparity between the independent reference pH and raw sensor pH during periods of rapid salinity changes most likely due to small spatiotemporal mismatches between the waters captured in the discrete bottle samples and measured by the electrodes in a highly dynamic physical and biogeochemical environment. It is also possible that this may reflect a relatively slow sensor response to a rapid salinity change.

Across all measurements, pH_{DIC-TA}^{disc} was higher than pH_{elec}^{disc} by 0.0356 ± 0.0190 and 0.0438 ± 0.0086 pH units on 01 June 2016 and 02 August 2016, respectively. This difference may be attributed to a number of factors, either singly or in combination: small errors in the analysis of discrete bottle samples (Bresnahan et al., 2014; McLaughlin et al., 2017; Patsavas et al., 2015); uncertainty in the thermodynamic constants used to calculate pH_{DIC-TA}^{disc} and pH_{elec}^{disc} (Patsavas et al., 2015; Rérolle et al., 2016); environmental pH gradients of varying magnitudes (Bresnahan et al., 2014); and other sources of sampling/handling error (McLaughlin et al., 2017; Rérolle et al., 2016). Finally, these differences may reflect contributions of excess alkalinity in coastal waters that would lead to higher calculated pH_T values (Cai et al., 2014).

On 01 June 2016 (Figure 2a), 21 calibration points for PH_{DIC-TA}^{disc} and 20 calibration points for PH_{elec}^{disc} coinciding with the times of sensor measurements were collected over a wide range of salinity. The PH_T varied from about $pH \sim 7.1-8.3$. The high pH_T values in the large range of pH captured on this sampling trip is attributable to phytoplankton blooms that characterize the mouth of the Murderkill Estuary during the early summer under strong southerly winds and warm, sunny weather (Voynova et al., 2015; Wong et al., 2009). On the other hand, a flux of fresher waters from the watershed of naturally low pH (commonly as low as pH 6.5) (Ullman et al., 2013, unpublished data) is responsible for the low pH_T values seen at the deployment site.

On 02 August 2016 (Figure 2b), 24 individual calibration points coinciding with the times of sensor measurements were collected for both pH_{DIC-TA}^{disc} and pH_{elec}^{disc} over a narrower range of higher salinities. The pH_T varied from about pH ~7.0-7.9. The dampened range of pH captured on this sampling trip may reflect the decrease in primary production previously observed during the mid-late summer relative to the late spring and early summer at the mouth of the Murderkill Estuary (Voynova et al., 2015). Alternatively or in addition, the lower pH_T maximum may also be attributed to a series of short storms on the day before sampling that churned up sediment from the estuary bottom thereby increasing water column turbidity, limiting light penetration, and reducing the impact of photosynthesis on pH. Moreover, the storms may have also contributed to periods of cloudiness on the day of sampling. It should be noted that sediment particles could exchange protons with the water thus leading to both a narrower and lower pH range (Stumm and Morgan, 1996).

3.2. Electrode performance in an estuarine environment

The raw spring and summer sensor time series of each reference electrode were recalibrated using both sets of independent reference pH: (1) pH_{DIC-TA}^{disc} and (2) pH_{elec}^{disc} determined from discrete bottle samples collected on 01 June 2016 and 02 August 2016, respectively. pH_{raw}^{EXT} from 01 June 2016 was recalibrated using two calibration constants: (A) derived from all good calibration points over the range of salinity where pH^{disc} and pH^{EXT} are linear (only S \geq 14.84) and (B) using a subset of this data collected above S = 20. Table 1 summarizes all of the possible procedures that can be used to calibrate the raw senor pH at this site. The calibrated and



Fig. 2. Values of pH_T calculated from measured DIC and TA (squares) and corrected from measurements of pH_{NBS}^{field} at *in situ* conditions (diamonds) as a function of salinity from sampling trips on (a) 01 June 2016 and (b) 02 August 2016 shown relative to raw sensor pH calculated from measured voltages using the internal (blue) and external (black) reference electrodes. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1

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}	None	1
01 June 2016	External	pH ^{disc} _{DIC-TA}	None	1A
01 June 2016	External	pH ^{disc}	Only $S > 20$	1B
01 June 2016	Internal	pH ^{disc}	None	2
01 June 2016	External	pH ^{disc}	None	2A
01 June 2016	External	pH ^{disc}	Only $S > 20$	2B
02 August 2016	Internal	pH ^{disc}	None	1
02 August 2016	External	pH ^{disc} _{DIC-TA}	None	1
02 August 2016	Internal	pH ^{disc}	None	2
02 August 2016	External	pH ^{disc} _{elec}	None	2

adjusted sensor time series are denoted by the subscript 'final' and represent the best estimate of pH at the deployment site.

Since all the various choices for the single point calibration of sensor pH result in nearly identical time series, we display and use 'final' data from only one combination of calibration constants for each deployment (2/2A for spring and 2 for summer). For the duration of the spring (Fig. 3a) and summer (Fig. 3d) deployments, the pH values calculated from the measured voltages exhibit good agreement across the wide natural temperature (Fig. 3b and e) and salinity (Fig. 3c and f) ranges that characterize the tidally-forced Murderkill Estuary-Delaware Bay System. 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. Similar to that shown in Fig. 2, larger diel pH cycles are observed with a few hours of nearly stable pH readings at high tides reflecting the more saline Delaware Bay, and sharp changes around the pH minima at the low salinities, reflecting contributions from the Murderkill Estuary outflow. Notable departures in the agreement between the pH values were observed during periods of simultaneous pH and salinity changes on the ebb and flood tides near extreme values.

All the calibrated sensor pH generally exhibited good agreement with their respective independent reference pH with sensor offsets (c_0) close to 0 and sensor gains (c_1) close to 1 across the respective salinity ranges on 01 June 2016 (Table 2/Fig. 4a–d) and 02 August 2016 (Table 2/Fig. 4e–f).

There was a good agreement between pH^{EXT}_{final,1A} and pH^{EXT}_{final,2A} with their respective independent reference pH down to S~15 in our estuarine environment consistent with previous work (Bresnahan et al., 2014; Martz et al., 2010; Takeshita et al., 2014). Below this salinity threshold, notable departures from linearity were evident (see panels (c) and (d) in Fig. 4). Under the current flow design, pH^{EXT} is sensitive to rapidly changing low salinity in our estuarine environment, as discussed below.

3.3. Evolution of $\Delta \mathbf{p} \mathbf{H}^{\mathbf{INT}-\mathbf{EXT}}$ anomalies

The dual-reference electrode configuration incorporated into Durafet-based biogeochemical sensors is not an absolute necessity for their use (Bresnahan et al., 2014). This configuration, however, provides a simple and powerful method for detecting the effects of fouling or sensor failure through direct comparisons of pH^{INT} and pH^{EXT} with time (Bresnahan et al., 2014; Rivest et al., 2016). Average anomalies (expressed as Δ pH^{INT-EXT}) were near zero during the spring (0.006 ± 0.063 pH units) and summer (-0.008 ± 0.020 pH units) deployments with conditioning periods excluded. However, salinity-driven spikes in the anomaly of ≥0.15 pH units were consistently observed with large salinity changes (Fig. 5). During

both sensor deployments, a positive $\varDelta p H^{INT-EXT}$ anomaly $(p H_{final}^{INT} > p H_{final}^{EXT})$ was consistently observed under prolonged periods of salinity decrease on the ebb tide, which bottomed out at the lowest salinities coinciding with the fresher Murderkill Estuary outflow. A rapid decrease to a negative $\varDelta p H^{INT-EXT}$ anomaly $(p H_{final}^{INT} coincided with the tide change and was followed by a near-zero anomaly.$

Bresnahan et al. (2014) also observed similar, though much smaller anomalies (< \pm 0.005 pH units) at 30 \leq S \leq 36 directly following rapid (albeit smaller than observed in this study) changes in salinity in a test tank. These anomalies suggest that a salinity lag may arise due to inadequate flushing of the instrument flow path and flow housing perhaps resulting in differences in the water sampled by the rapidly flushed conductivity-temperature sensor (SBE37) and the water seen by the electrodes inside the more slowly flushed flow housing (Bresnahan et al., 2014). Secondarily, the anomaly may also occur due to electrode reconditioning to new salinities when salinity changes rapidly, but this effect has yet to be definitively characterized. It is also possible the m_{Cl} calculated from salinity used to calculate pHEXT is not accurate at the lowest salinities observed in this work which would produce systematic errors in those values (Takeshita et al., 2014). The salinity-driven $\Delta p H^{INT-EXT}$ anomalies in estuaries will be discussed in greater detail in a future publication (Gonski et al., In Prep.). Further studies to characterize the performance of the dual-reference configuration at the extremes of and during rapid changes in environmental conditions are needed to better interpret electrode response to dynamic environments.

3.4. 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; McLaughlin et al., 2017; Rivest et al., 2016) as the principal data quality assessment tools. Typically, 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 (McLaughlin et al., 2017; Rivest et al., 2016).

These simple techniques serve primarily to identify a failed sensor rather than evaluate a working one, and in a dynamic environment may also exclude valid data. Due to the physical and



Fig. 3. Murderkill Estuary-Delaware Bay time-series from (a, b, c) spring and (d, e, f) summer deployments. Panels (a) and (d) show the final pH calculated using the internal (black) and external (blue) reference electrodes. Panels (b) and (e) show *in situ* temperature. Panels (c) and (f) show salinity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

biogeochemical variability in our estuarine environment, we have chosen to report all sensor measurements except for those associated with intra-deployment conditioning periods to inform future work done in similar environments. However, there is a need to develop a set of recommended QC procedures, specific for highly variable estuarine conditions for the future.

3.5. Electrode conditioning in an estuarine environment

3.5.1. 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 predeployment electrode conditioning procedures. In the absence of pre-conditioning, 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 artifacts, electrodes should always be stored in seawater, be continuously powered for 5–10 days prior to deployment, and transported in seawater to the deployment site (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 (S = 29–31) was used as the pre-conditioning medium. The seawater was stored inside the flow housing of the SeapHOx from 13 April 2016 to 09 May 2016 and was replaced every 5 days. The $\Delta p H^{INT-EXT}$ anomaly was used to gauge the success of the electrode conditioning procedure. A final conditioning period of 74 h in the field followed (Fig. 6), during which, the calculated $\Delta p H^{INT-EXT}$ anomaly repeatedly returned to a consistently stable value of < ±0.01 pH units during proceeding high tide periods when the more saline Delaware Bay water inundated the deployment site.

The prolonged conditioning period may be linked to the time needed to achieve a stable flow of ions across the liquid junction of the internal reference electrode and the previous discussed exchange of Cl^- with Br^- in the solid AgCl solution of the external reference electrode (Bresnahan et al., 2014). With salinities ranging between 5.9 and 23.1, large changes in $[Br^-]$ and $[Cl^-]$ associated with natural salinity fluctuations could result in an especially long conditioning period in a dynamic estuarine environment. If the steady-state ion replacement process of the Cl-ISE was slower than the rate of tidal salinity change, it is also possible that full conditioning cannot be achieved in a dynamic estuarine environment with a wide salinity range. Further development and testing of predeployment electrode conditioning protocols for this sensor in estuaries are certainly needed.

3.5.2. Intra-deployment electrode conditioning

When the electrodes are allowed to dry out, a conditioning period is always required to 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 requires the sensor to be out of the water for 1–3 h. Throughout the course of this 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 reconditioning was determined by comparing sensor pH and pH_T calculated from measured DIC and TA from discrete bottle samples collected after sensor maintenance.

Between May and August 2016, both reference electrodes adequately reconditioned within 6 h of the first sampling cycle after sensor redeployment under both conditions. A nominal 6-h intra-deployment conditioning period was assumed for all May 2016 to August 2016 data used in subsequent analyses. The relatively quick servicing time translates into these shorter conditioning periods where the sensors simply needed to be reimmersed and made "wet". Intra-deployment conditioning periods may be heavily dependent on local temperature and salinity conditions at the deployment site, and their length may change at throughout the year. This would make them site-dependent and deploymentspecific. Following sensor maintenance, it is recommended to fill the SeapHOx flow housing with filtered natural waters from the deployment site prior to redeployment which conforms with existing electrode conditioning protocol (Bresnahan et al., 2014).

3.6. Recommendations

3.6.1. Choice of independent reference pH

3.6.1.1. Murderkill Estuary-Delaware Bay results. Consistent with Bresnahan et al. (2016), comparisons 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. Property-property plots between the perturbations in pH^{disc} and pH^{sensor} from the minimum value of pH^{disc} characterized by perturbation variables of pH^{disc} = pH^{disc} – min(pH^{disc}) and pH^{sensor} = pH^{sensor} – min(pH^{disc}) scale Model II fit parameters to the pH range of a dataset making them more representative (Bresnahan et al., 2016). Such comparisons (Fig. 7a–b) produce Model II fit sensor offsets much closer to 0 (Table 3).

The ACT results for at-sea conditions produced standard deviations in pH^{sensor} – pH^{disc} anomalies ranging between 0.01 and 0.1 pH units. In the present study, similar sensor validation protocol (selected comparisons shown above in Fig. 7a-b) yielded standard deviations or root-mean squared errors ranging between 0.011 and 0.036 pH units over extensive pH ranges. This demonstrates the viability of the Honeywell Durafet to the collection of highfrequency, high-resolution, and high-precision pH data over weeks to months in dynamic, productive, high-fouling, and highlyturbid estuarine environments with extensive salinity ranges regardless of the independent reference pH used. These results align well with the long-term precisions of +0.02 pH units achieved by well-trained sensor operators of Durafet-based biogeochemical sensors in coastal waters (McLaughlin et al., 2017). Taken together, the Honeywell Durafet is capable of achieving the weather-level pH precision for ocean acidification research recommended by GOA-ON in dynamic estuarine environments with extensive salinity and pH ranges (precision \leq 0.02 pH units; GOA-ON report; Newton et al., 2015).

3.6.1.2. Practical considerations. For sensor deployments in dynamic estuarine environments, sensor calibrations should be based on a discrete "dynamic point" sampling scheme with samples collected and analyzed to capture the full range of expected pH and salinity fluctuations in the deployment environment. If this is not done, sensor operators risk introducing unintended bias into their measurements and interpretations. This approach also has the advantage of identifying other potential controls on electrode performance, most notable, turbidity. Based on the agreement of $PH_{final}^{\rm ensor}$ with $PH_{DIC-TA}^{\rm DiC}$ (filtered) and $PH_{elec}^{\rm disc}$ (unfiltered), elevated particle loads in the Murderkill Estuary-Delaware Bay System (Ullman et al., 2013) do not appear to interfere with the instantaneous sensor measurements during short-term periods at this site. Long-term effects of turbidity on electrode response, however, were not investigated and may need to be studied further.

Using pH_T calculated from two measured marine CO₂ system

Table 2

Root-mean squared error (RMSE), sensor offset (c_0), and sensor gain (c_1) calculated from Model II least squares fits of pH^{sensor}vs. pH^{disc} from the 01 June 2016 and 02 August 2016 sampling days.

Sampling Day	Reference Electrode	Calibration Method	RMSE	c ₀ (intercept)	c ₁ (slope)
01 June 2016	Internal	1	0.0275	-0.1795 ± 0.1683	1.0225 ± 0.0210
01 June 2016	External	1A	0.0174	-0.5848 ± 0.1741	1.0724 ± 0.0215
01 June 2016	External	1B	0.0158	0.0882 ± 0.3860	0.9891 ± 0.0473
01 June 2016	Internal	2	0.0358	-0.0784 ± 0.1592	1.0101 ± 0.0202
01 June 2016	External	2A	0.0114	-0.0308 ± 0.1063	1.0039 ± 0.0132
01 June 2016	External	2B	0.0123	-0.1557 ± 0.3258	1.0191 ± 0.0399
02 August 2016	Internal	1	0.0159	0.2281 ± 0.1072	0.9698 ± 0.0142
02 August 2016	External	1	0.0121	-0.0892 ± 0.0814	1.0119 ± 0.0108
02 August 2016	Internal	2	0.0149	0.3655 ± 0.0982	0.9514 ± 0.0131
02 August 2016	External	2	0.0105	0.0560 ± 0.0689	0.9927 ± 0.0092

parameters (e.g. DIC and TA) to calibrate a working Honeywell Durafet has been done previously (Bresnahan et al., 2016; McLaughlin et al., 2017; Rérolle et al., 2016; Takeshita et al., 2015). This offers the advantage of providing the sensor operators with a better understanding of the underlying controls on the marine CO₂ system in a specific environment (Bresnahan et al., 2016; MacLeod et al., 2015). Yet, this approach is very labor intensive (Bresnahan et al., 2016). In the present study, a total of 486 discrete bottle samples for DIC and TA were collected and analyzed, involving more than 125 person-hours on more than 20 sampling days to produce 255 potential calibration points over the tenmonth sensor deployment from September 2015 to August 2016. This intensity of sensor calibration is desirable, but not always possible.

A less labor-intensive calibration approach using field measurements of pH_{NBS} made with a glass electrode calibrated with low ionic strength pH buffers may be appropriate for some estuarine environments. Culberson (1981) advocates the continued employment of the NBS pH scale for pH measurement in estuaries. Even more, high quality pH measurements using a glass electrode are still achievable (Easley and Byrne, 2012), and Martz et al. (2015) agrees that similar methods of pH measurement may still be appropriate for marine CO₂ chemistry and acidification studies depending on the needs and limitations of individual researchers.

While there will be some uncertainty in any pH calibration based on standards with a fixed salt concentration as a reference for all pH measurements, under the pretext that the $\ensuremath{\mathsf{pH}_{\text{NBS}}}$ measurements are free of electrode drift and biofouling, their correction to a concentration scale at in situ conditions using simultaneously collected DIC and TA data substantially reduces this uncertainty (note if only for the purpose of determining $\varDelta pH_{scales},$ the data density need not be high). A simultaneous secondary calibration of a glass electrode in TRIS Buffer in artificial seawater on a concentration scale, however, would provide an additional check of the accuracy of the pH scale conversion and help approximate any liquid junction potential errors present in the original pH_{NBS} measurements (Martz et al., 2015; Whitfield et al., 1985). In the future, we encourage other researchers using a sensor calibration traceable to pH_{NBS} measurements to report their values of ΔpH_{scales} as we have for either method consistent with the recommendations of Butler et al. (1985). The use of low ionic strength pH buffers in estuarine pH applications for sensor calibration has been overlooked, but might provide a cost-effective alternative.

In productive coastal environments characterized by high fouling potential and high sediment loads, sensor maintenance has proven to be essential to the continuous collection of high quality pH data over periods of weeks to months between servicing trips (McLaughlin et al., 2017; Takeshita et al., 2015). As the use of Durafet-based biogeochemical sensors in estuaries expands, routine sensor maintenance that restores proper sensor operation should be integrated into future deployment designs. Independent co-located sensors to measure a redundant pH and/or an additional marine CO₂ system parameter (e.g. *p*CO₂) coupled with the utilization of regional empirical marine CO₂ system relationships (e.g. Alin et al. (2012)) should also form the basis of an ideal sensor deployment and calibration (Bresnahan et al., 2014) in dynamic estuarine environments.

Subsequent high-frequency measurements of an independent reference pH may be used to examine the short-term trends in electrode response between routine sensor maintenance. Together with a sampling approach characterized by the frequent collection of discrete samples over a much longer duration than the present work (e.g. weekly samples for at least 6 months), long-term sensor drift in estuaries may also be definitively identified and a suitable drift-correction developed. Yet, the present study primarily served to evaluate Durafet performance in an estuarine environment over an extensive salinity range and a wide range of time varying salinity. Accordingly, the identification of long-term sensor drift in dynamic estuarine environments was beyond the scope of the present experimental design. Other studies, however, have intimately examined electrode response in similar highly dynamic coastal environments and demonstrated that working sensors do not drift to a detectable amount or to within the error of the discrete bottle samples over periods of more than several months (Bresnahan et al., 2014; Kapsenberg and Hofmann, 2016).

Training and experience with Durafet-based biogeochemical sensors will also improve the quality of sensor deployments and their calibrations, compared with those of novice sensor operators (McLaughlin et al., 2017). Thus, in addition to established sensor best practices, instructional field and deployment materials, referred to as "sensor good practices" by Martz et al. (2015), encompassing the more practical aspects of Durafet operation need to be produced and compiled by, and ultimately disseminated to the user community to keep pace with the rapidly growing use of the Honeywell Durafet worldwide. While sensor best practices will be similar, sensor good practices will differ with deployment environment.

3.6.2. Treatment of excess alkalinity

Excess alkalinity is a potential source of the elevated pH_{DlC-TA}^{disc} relative to pH_{elec}^{disc} discussed in section 3.1. Excess alkalinity (Δ TA), can be calculated as the difference between measured alkalinity (TA_{meas}) and alkalinity calculated from pH_{NBS}^{field} and DIC (TA_{calc}^{pH-DIC}):

$$\Delta TA = TA_{meas} - TA_{calc}^{pH-DIC}$$
(4)

In this study, the average values of Δ TA were 12.6 \pm 1.8 μ mol kg⁻¹ out of a measured alkalinity range of 1535–1840 μ mol kg⁻¹ on 01 June 2016 and 14.7 \pm 2.8 μ mol kg⁻¹ out of a measured alkalinity



Fig. 4. Property-property plots of sensor pH vs. independent reference pH as a function of salinity. Data taken from 01 June 2016 shown in (a)–(d) and 02 August 2016 shown in (e) and (f). See Table 1 for descriptions of the comparisons. Dashed black lines represent a 1:1 relationship ($pH^{sensor} = pH^{disc}$). Note that the bottom two panels use a different salinity color scale.



Fig. 5. Calculated $\Delta p H^{INT-EXT}$ anomalies from (a) spring and (b) summer deployments. Anomaly (solid red line) shown relative to a zero $\Delta p H^{INT-EXT}$ anomaly (dashed black line) and salinity (dotted green line). Gaps in the $\Delta p H^{INT-EXT}$ anomalies represent intra-deployment conditioning periods which were filtered out. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)



Fig. 6. Conditioning period at the start of the spring 2016 SeapHOx deployment between 09 May 2016 at 1200 (Hour 0) and 16 May 2016 at 1030 (Hour 166.5). The calculated $\Delta p H^{INT-EXT}$ anomaly (solid red line) is shown relative to a zero $\Delta p H^{INT-EXT}$ anomaly (dashed black line) and salinity (dotted green line). The conditioning period lasted until 12 May 2016 at 1400 (Hour 74). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Fig. 7. Property-property plots of $pH_{final}^{sensor'}$ vs. $pH_{final}^{disc'}$ as a function of salinity. Data taken from 01 June 2016 showing (a) $pH_{final,2}^{NT'}$ vs. $pH_{elec}^{disc'}$ and 02 August 2016 showing (b) $pH_{final,2}^{EXT'}$ vs. $pH_{elec}^{disc'}$. Dashed black lines represent a 1:1 relationship ($pH^{sensor} = pH^{disc'}$). Note the different salinity color scales used in the two panels. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 3 Root-mean squared error (RMSE), sensor offset (c_0), and sensor gain (c_1) calculated from Model II least squares fits of pH^{sensor/}_{final} vs. pH^{disc/} from the 01 June 2016 and 02 August 2016 sampling days.

Sampling Day	Reference Electrode	Calibration Method	RMSE	c ₀ (intercept)	c_1 (slope)
01 June 2016	Internal	1	0.0275	-0.0160 ± 0.0171	1.0225 ± 0.0210
01 June 2016	External	1A	0.0174	-0.0405 ± 0.0129	1.0724 ± 0.0215
01 June 2016	External	1B	0.0158	0.0006 ± 0.0076	0.9891 ± 0.0473
01 June 2016	Internal	2	0.0358	-0.0068 ± 0.0175	1.0101 ± 0.0202
01 June 2016	External	2A	0.0114	-0.0018 ± 0.0085	1.0039 ± 0.0132
01 June 2016	External	2B	0.0123	-0.0025 ± 0.0071	1.0191 ± 0.0399
02 August 2016	Internal	1	0.0159	0.0148 ± 0.0076	0.9698 ± 0.0142
02 August 2016	External	1	0.0121	-0.0051 ± 0.0058	1.0019 ± 0.0108
02 August 2016	Internal	2	0.0149	0.0244 ± 0.0072	0.9514 ± 0.0131
02 August 2016	External	2	0.0105	0.0045 ± 0.0050	0.9927 ± 0.0092

range of 2020–2070 µmol kg⁻¹ on 02 August 2016. Δ TA was not consistently strongly correlated with dissolved organic carbon (DOC) calculated from fluorescence using the summer 2013 relationship from Voynova et al. (2015) ($r^2 = 0.86$) for the same deployment site on 01 June 2016 ($r^2 = 0.0787$; p = 0.43) or 02 August 2016 ($r^2 = 0.0832$; p = 0.36). These results align well with those of other marine CO₂ chemistry studies performed in other parts of the Murderkill Estuary watershed near its confluence with the Delaware Bay (Ullman et al., 2013, unpublished data). Excess alkalinity also never exceeded more than 1–2% of total measured alkalinity. Accordingly, the effects of excess alkalinity on the relative trends in pH observed in this environment are insignificant when pH fluctuations as large as >0.5 pH units are experienced over tidal excursions alone.

When setting the calibration constant to an average value to minimize the anomaly between the sensor pH and pH_T calculated from DIC and TA, the effects of excess alkalinity captured in discrete bottle samples can be easily incorporated into a sensor calibration (McLaughlin et al., 2017). In other words, this introduces a systematic pH offset into the adjusted sensor pH dataset. Therefore, existing calibration protocols for Durafet-based biogeochemical sensors cannot be used to detect, quantify, and correct for contributions of excess alkalinity to sensor pH measurements in dynamic estuarine environments on their own. Sampling for DOC and/or DOM, measuring additional marine CO_2 system parameters (e.g. DIC or pCO_2), and/or dedicated sampling for direct measurements of excess alkalinity using methods like those described in Cai et al. (1998) and Yang et al. (2015) are 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.

3.6.3. Modifications to SeapHOx design

The open-ocean design of the SeapHOx was not optimal for our estuarine application. We encountered a number of problems at our field site and made modifications to better avoid these problems. Based on our experiences, we propose a set of recommendations for subsequent sensor design specific to a dynamic estuarine deployment environment. These include:

- Flow Path: Our instrument experienced erratic and degraded pump performance over time due to choke points in the instrument flow path especially in areas where the cross-sectional area or the direction of the flow path changed. This may lead to the flagging and exclusion of a large number of sensor measurements after undergoing routine QC procedures. By standardizing the cross-sectional area of the flow path tubing and employing rounded tubing connections between sensor components, this data loss can be substantially minimized.
- 2) <u>Submersible Pump</u>: Degraded pump performance was observed between sensor maintenance trips due to high particle loads. An

incomplete flushing of the SeapHOx flow housing leads to the collection of unrepresentative instantaneous pH data. A stronger pump designed for high particle loads should be integrated as long as the stronger, faster flow does not compromise the integrity of the sensing surfaces similar to Takeshita et al. (2016b). Similarly, an upside down or horizontal sensor orientation is recommended for future deployments in similar estuarine environments. This orientation minimizes sedimentation effects and the strain placed upon the pump due to pumping turbid waters against gravity.

3) <u>Wireless Communications & Monitoring</u>: Currently, the communication with any SeapHOx unit is mediated by a direct RS-232 connection with a computer through a suitable terminal program (Martz, 2012). It would be useful to have a more sophisticated interface that would permit wireless communications between the sensor and sensor operator and the wireless transmission of data from the sensor to a central location by way of cellular and WiFi connections. As most estuarine and coastal ocean sites should be within range of land-based networks, this could permit sensor operators to monitor sensor operations remotely and to respond to sensor failures in a timely fashion. Wireless communications and real-time data capabilities would also make data available to the user community more rapidly and allow sensor operators to collect calibration samples over a wider range of observations more efficiently.

Improvements in instrument design and communication capabilities would expand the areas suitable for SeapHOx deployment and the community of potential users with interests in ocean acidification and marine CO_2 chemistry in a wider array of estuarine and coastal ocean settings.

3.7. 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. dE_{INT}^*/dT and dE_{EXT}^*/dT) (Martz et al., 2010). Because of this variability, sensor redundancy was a key feature in subsequent work (Bresnahan et al., 2014; McLaughlin et al., 2017; Takeshita et al., 2014). In this evaluation of the Honeywell Durafet, one Sea-pHOx unit with one Honeywell Durafet was used successfully in a highly variable estuarine environment and the success of this sensor was verified using a pair of validation procedures. It may be possible to reduce the need for the sensor redundancy in the future by adhering to established sensor best practices and compiling sets of accepted sensor good practices for different deployment environments.

4. Conclusions

Deployments of the SeapHOx sensor equipped with the Honeywell Durafet were carried out between April 2015 and August 2016 at the confluence of the Murderkill Estuary and Delaware Bay (Delaware, USA). This work yielded useful high quality and high frequency data concerning the role of environmental forcing on pH in a dynamic, productive, high-fouling, and highly-turbid estuarine environment characterized by a wide salinity range (S = 3.25 to 29.33). The sensor pH collected during May 2016 to August 2016 using the most refined SeapHOx configuration exhibited good agreement with the independent determinations of pH_T. When rigorously calibrated using 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 natural waters mixed along the salinity gradient, the sensor pH had a root-mean squared error ranging between 0.011 and 0.036 pH units across natural pH fluctuations of up to >1 pH unit and range of 26 salinity.

The performance of the Honeywell Durafet reinforced the versatility of this sensor and demonstrated its viability to the collection of dependable high-frequency and high-resolution pH data with GOA-ON weather-level precision over periods of weeks to months with regularly-scheduled sensor maintenance in dynamic estuarine environments. The Honeywell Durafet can now be used to elucidate small-scale pH variations and trends to help discern the impact of localized acidification on dynamic estuarine environments and characterize any subsequent ecosystem responses with certainty (Newton et al., 2015). We identified and resolved a number of deficiencies in existing deployment guidelines and calibration protocol caused by variable environmental conditions in estuaries. We also highlighted aspects of electrode response requiring further investigation and provided a set of recommendations for the future utilization of these sensors in similar environments. In the future, it is likely that the accuracy of the Honeywell Durafet (± 0.01 pH units) seen in the open-ocean should be attainable in dynamic estuarine environments with extensive salinity ranges and broad ranges of time varying salinity.

Acknowledgements

We sincerely thank Philip J. Bresnahan Jr., Yuichiro Takeshita, and Taylor Wirth from the Scripps Institute of Oceanography for their insights and support. We thank Andrew Collins, Najid Hussain, and Michael Scaboo from the University of Delaware for their assistance with sample collection in the field. Finally, we thank the staff of the University of Delaware's Marine Operations for their assistance with the sensor modifications. Financial support for this work was provided by internal funds from the University of Delaware Provost's office and the College of Earth, Ocean and Environment Dean's office (SMSP175136), a Delaware Sea Grant award (HCE-23), and a NASA grant (NNX14AM37G) awarded to the W.-J. Cai. Data used are uploaded with the manuscript.

Appendix A. Supplementary data

Supplementary data related to this article can be found at https://doi.org/10.1016/j.ecss.2017.10.020.

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., 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). J. Geophys. Res. 117 (C5).
- Bagshaw, E.A., Beaton, A., Wadham, J.L., Mowlem, M., Hawkings, J.R., Tranter, M., 2016. Chemical sensors for in situ data collection in the cryosphere. Trends Anal. Chem. 82, 348–357.
- Bates, R.G., 1973. Determination of PH: Theory and Practice, second ed. Wiley-Interscience.
- Bates, R.G., 1975. pH scales for seawater. In: Goldberg, E.D. (Ed.), The Nature of Seawater, Dahlem Konferenzen, pp. 313–338.
- Bates, R.G., 1982. pH measurements in the marine environment. Pure Appl. Chem. 54 (1), 229–232.
- Bresnahan Jr., P.J., Martz, T.R., Takeshita, Y., Johnson, K.S., LaShomb, M., 2014. Best practices for autonomous measurements of seawater pH with the Honeywell Durafet. Methods Oceanogr. 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., Statom, N., 2016. A sensor package for mapping pH and oxygen from mobile platforms. Methods Oceanogr. 17, 1–13.
- Briggs, E.M., Sandoval, S., Erten, A., Takeshita, Y., Kummel, A.C., Martz, T.R., 2017. Solid state sensor for simultaneous measurement of total alkalinity and pH in seawater. ACS Sensors 2 (9), 1302–1309.
- Butler, R.A., Covington, A.K., Whitfield, M., 1985. The determination of pH in estuarine waters. II: practical considerations. Oceanol. Acta 8 (4), 433–439.

Byrne, R., Kump, L., Cantrell, K., 1988. The influence of temperature and pH on trace metal speciation in seawater. Mar. Chem. 25 (2), 163–181.

- Byrne, R.H., 2014. Measuring ocean acidification: new technology for a new era of ocean chemistry. Environ. Sci. Technol. 48 (10), 5352–5360.
- Cai, W.-J., Hu, X., Huang, W.-J., Jiang, L.-Q., Wang, Y., Peng, T.-H., Zhang, X., 2010. Alkalinity distributions in the western north Atlantic Ocean margins. J. Geophys. Res. Oceans 115, 1–15.
- Cai, W.-J., Hu, X., Huang, W.-J., Murrell, M.C., Lehrter, J.C., Lohrenz, S.E., Chou, W.-C., Zhai, W., Hollibaugh, J.T., Wang, Y., Zhao, P., Guo, X., Gunderson, K., Dai, M., Gong, G.C., 2011. Acidification of subsurface coastal waters enhanced by eutrophication. Nat. Geosci. 4 (11), 766–770.
- Cai, W.-J., Wang, Y., 1998. The chemistry, fluxes, and sources of carbon dioxide in the estuarine waters of the Satilla and Altamaha Rivers, Georgia. Limnol. Oceanogr. 43 (4), 657–668.
- Cai, W.-J., Wang, Y., Hodson, R.E., 1998. Acid-base properties of dissolved organic matter in the estuarine waters of Georgia, USA. Geochimica Cosmochimica Acta 62 (3), 473–483.
- Caldeira, K., Wickett, M., 2003. Anthropogenic carbon and ocean pH. Nature 425 (6956), 365.
- Carter, B.R., Radich, J.A., Doyle, H.L., Dickson, A.G., 2013. An automated system for spectrophotometric seawater pH measurements. Limnol. Oceanogr. Methods 11, 16–27.
- Chavez, F.P., Sevadjian, J., Wahl, C., Friederich, J., Friederich, G.E., 2017. Measurement of pCO₂ and pH from an autonomous surface vehicle in a coastal upwelling system. Deep Sea Res. Part II Top. Stud. Oceanogr. https://doi.org/10.1016/ j.dsr2.2017.01.001 (in press).
- Clarke, J.S., Achterberg, E.P., Rérolle, V.M.C., Bey, S.A.K., Floquet, C.F.A., Mowlem, M.C., 2015. Characterisation and deployment of an immobilised pH sensor spot towards surface ocean pH measurements. Anal. Chim. Acta 897, 69–80.
- Clayton, T.D., Byrne, R.H., Breland, J.A., Feely, R.A., 1995. The role of pH measurements in modern oceanic CO₂-system characterization: precision and thermodynamic consistency. Deep Sea Res. Part II Top. Stud. Oceanogr. 42 (2–3), 411–429.
- Culberson, C.H., 1981. Direct potentiometry. In: Whitfield, M., Jagner, D. (Eds.), Marine Electrochemistry. Wiley, pp. 187–261.
- DeGrandpre, M.D., Spaulding, R.S., Newton, J.O., Jaqueth, E.J., Hamblock, S.E., Umansky, A.A., Harris, K.E., 2014. Considerations for the measurement of spectrophotometric pH for ocean acidification and other studies. Limnol. Oceanogr. Methods 12 (12), 830–839.
- de Vargas Sansalvador, I.M.P., Fay, C.D., Cleary, J., Nightingale, A.M., Mowlem, M.C., Diamond, D., 2016. Autonomous reagent-based microfluidic pH sensor platform. Sensors Actuators B Chem. 225, 369–376.
- DelValls, T.A., Dickson, A.G., 1998. The pH of buffers based on 2-amino-2hydroxymethyl-1,3-propanediol ('tris') in synthetic sea water. Deep Sea Res. Part I Oceanogr. Res. Pap. 45 (9), 1541–1554.
- Dickson, A.G., 1984. pH scales and proton-transfer reactions in saline media such as seawater. Geochimica Cosmochimica Acta 48 (11), 2299–2308.
- Dickson, A.G., 1990. Standard potential of the reaction: AgCl (s) + 1/2 H₂ (g) = Ag (s) + HCl (aq), and the standard acidity constant of the ion HSO₄ in synthetic sea water from 273.15 to 318.15 K. J. Chem. Thermodyn. 22 (2), 113–127.
- Dickson, A.G., 1993. The measurement of sea water pH. Mar. Chem. 44 (2-4), 131-142.
- Dickson, A.G., Camoes, M.F., Spitzer, P., Fisicaro, P., Stoica, D., Pawlowicz, R., Feistel, R., 2016. Metrological challenges for measurements of key climatological observables. Part 3: seawater pH. Metrologia 53 (1), R26.
- Dickson, A.G., Sabine, C.L., Christian, J.R., 2007. Guide to Best Practices for Ocean $\rm CO_2$ Measurements.
- Doney, S., Fabry, V., Feely, R., Kleypas, J., 2009a. Ocean acidification: the other CO₂ problem. Annu. Rev. Mar. Sci. 1, 169–192.
- Doney, S.C., Balch, W.M., Fabry, V.J., Feely, R.A., 2009b. Ocean acidification: a critical emerging problem for the ocean sciences. Oceanography 22 (4), 16–25.
- Douglas, N.K., Byrne, R.H., 2017a. Achieving accurate spectrophotometric pH measurements using unpurified *meta*-Cresol Purple. Mar. Chem. 190, 66–72.
- Douglas, N.K., Byrne, R.H., 2017b. Spectrophotometric pH measurements from river to sea: calibration of mCP for $0 \le S \le 40$ and $278.15 \le T \le 308.15$ K. Mar. Chem. https://doi.org/10.1016/j.marchem.2017.10.001 (in press).
- Duarte, C.M., Hendriks, I.E., Moore, T.S., Olsen, Y.S., Steckbauer, A., Ramajo, L., Carstensen, J., Trotter, J.A., McCulloch, M., 2013. Is ocean acidifcation an Open-Ocean Syndrome? Understanding anthropogenic impacts on seawater pH. Estuaries Coasts 36, 221–236.
- Dzwonkowski, B., Wong, K.-C., Ullman, W.J., 2013. Water level and velocity characteristics of a salt marsh channel in the Murderkill Estuary, Delaware. J. Coast. Res. 30 (1), 63–74.
- Easley, R.A., Byrne, R.H., 2012. Spectrophotometric calibration of pH electrodes in seawater using purified m-Cresol Purple. Environ. Sci. Technol. 46 (9), 5018–5024.
- Feely, R.A., Alin, S.R., Newton, J., Sabine, C.L., Warner, M., Devol, A., Krembs, C., Maloy, C., 2010. The combined effects of ocean acidification, mixing, and respiration on pH and carbonate saturation in an urbanized estuary. Estuar. Coast. Shelf Sci. 88 (4), 442–449.
- French, C.R., Carr, J.J., Dougherty, E.M., Eidson, L.A.K., Reynolds, J.C., DeGrandpre, M.D., 2002. Spectrophotometric pH measurements of freshwater. Anal. Chim. Acta 453, 13–20.

Gabriel, M.D., Forja, J.M., Rubio, J.A., 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 Mar. 31 (1B), 309–318.

- Gallego-Urrea, J.A., Turner, D.R., 2017. Determination of pH in estuarine and brackish waters: pitzer parameters for Tris buffers and dissociation constants for *m*-Cresol Purple at 298.15 K. Mar. Chem. 195, 84–89.
- Gran, G., 1950. Determination of the equivalence point in potentiometric titrations. Acta Chem. Scand. 4, 559–577.
- Gran, G., 1952. Determination of the equivalence point in potentiometric titrations-Part II. Analyst 77 (920), 661-671.
- Hammer, K., Schneider, B., Kulinski, K., Schulz-Bull, D.E., 2014. Precision and accuracy of spectrophotometric pH measurements at environmental conditions in the Baltic Sea. Estuar. Coast. Shelf Sci. 146, 24–32.
- Hammer, K., Schneider, B., Kulinski, K., Schulz-Bull, D.E., 2017. Acid-base properties of Baltic Sea dissolved organic matter. J. Mar. Syst. 173, 114–121.
- Hansson, I., 1973a. A new set of acidity constants for carbonic acid and boric acid in sea water. Deep Sea Res. Oceanogr. Abstr. 20 (5), 461–478.
- Hansson, I., 1973b. A new set of pH-scales and standard buffers for sea water. Deep Sea Res. Oceanogr. Abstr. 20 (5), 479–491.
- Hawkings, J., Wadham, J., Tranter, M., Telling, J., Bagshaw, E., Beaton, A., Simmons, S.-L., Chandler, D., Tedstone, A., Nienow, P., 2016. The Greenland Ice Sheet as a hot spot of phosphorous weathering and export in the Arctic. Glob. Biogeochem. Cycles 30 (2), 191–210.
- 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., Martz, T.R., 2011. High-frequency dynamics of ocean pH: a multi-ecosystem comparison. PLoS One 6 (12), e28983.
- Hopkins, A.E., Sell, K.S., Soli, A.L., Byrne, R.H., 2000. In-situ spectrophotometric pH measurements: the effects of pressure on Thymol Blue protonation and absorbance characteristics. Mar. Chem. 71 (1–2), 103–109.
- Huang, W.-J., Wang, Y., Cai, W.-J., 2012. Assessment of sample storage techniques for total alkalinity and dissolved inorganic carbon in seawater. Limnol. Oceanogr. Methods 10 (9), 711–717.
- Hulanicki, A., Glab, S., Ingman, F., 1991. Chemical sensors: definitions and classification. J. Pure Appl. Chem. 63 (9), 1247–1250.
- Jiang, L.-Q., Cai, W.-J., Wannikhoff, R., Wang, Y., Luger, H., 2008. Air-sea CO₂ fluxes on the US South Atlantic Bight: spatial and seasonal variability, 1978-2012 J. Geophys. Res. Oceans 113 (C7).
- Johnson, K.S., Jannasch, H.W., Coletti, L.J., Elrod, V.A., Martz, T.R., Takeshita, Y., Carlson, R.J., Connery, J.G., 2016. Deep-Sea DuraFET: a pressure tolerant pH sensor designed for global sensor networks. Anal. Chem. 88 (6), 3249–3256.
- Johnson, K.S., Needoba, J.A., Riser, S.C., Showers, W.J., 2007. Chemical sensor networks for the aquatic environment. Chem. Rev. 107 (2), 623–640.
- Kapsenberg, L., Hofmann, G.E., 2016. Ocean pH time-series and drivers of variability along the northern Channel Islands, California, USA. Limnol. Oceanogr. 61 (3), 953–968.
- Khoo, K.H., Ramette, R.W., Culberson, C.H., Bates, R.G., 1977. Determination of hydrogen ion concentrations in seawater from 5 to 40°C: standard potentials at salinities from 20 to 45. Anal. Chem. 49 (1), 29–34.
- Kleypas, J., Feely, R., Fabry, V., Langdon, C., Sabine, C., 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., Hoegh-Guldberg, O., 2012. A short-term in situ CO₂ enrichment experiment on Heron Island (GBR). Sci. Rep. 2, 413.
- Lai, C.-Z., DeGrandpre, M.D., Wasser, B.D., Brandon, T.A., Clucas, D.S., Jaqueth, E.J., Benson, Z.D., Beatty, C.M., Spaulding, R.S., 2016. Spectrophotometric measurement of freshwater pH with purified *meta*-Cresol Purple and Phenol Red. Limnol. Oceanogr. Methods 14 (12), 864–873.
- Larsen, M., Borisov, S.M., Grunwald, B., Klimant, I., Glud, R.N., 2011. A simple and inexpensive high resolution color ratiometric planar optode imaging approach: applications to oxygen and pH sensing. Limnol. Oceanogr. Methods 9 (9), 348–360.
- Lee, K., Kim, T.-W., Byrne, R.H., Millero, F.J., Feely, R.A., Liu, Y.-M., 2010. The universal ratio of boron to chlorinity for the North Pacific and North Atlantic oceans. Geochimica Cosmochimica Acta 74 (6), 1801–1811.
- Liu, X., Patsavas, M.C., Byrne, R.H., 2011. Purification and characterization of *meta*-Cresol Purple for spectrophotometric seawater pH measurements. Environ. Sci. Technol. 45 (11), 4862–4868.
- Loucaides, S., Rérolle, V.M.C., Papadimitriou, S., Kennedy, H., Mowlem, M.C., Dickson, A.G., Gledhill, M., Achterberg, E.P., 2017. Characterization of *meta*-Cresol Purple for spectrophotometric pH measurements in saline and hypersaline media at sub-zero temperatures. Sci. Rep. 7.
- MacLeod, C.D., Doyle, H.L., Currie, K.I., 2015. Technical Note: maximising accuracy and minimising 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., Chen, C.-T., 2011. pH of seawater. Mar. Chem. 126 (1–4), 89–96.
- Martz, T., 2012. SeaFET & SeapHOx Users Guide. University of California, San Diego (Scripps Institute of Oceanography).
- Martz, T.R., Carr, J.J., French, C.R., DeGrandpre, M.D., 2003. A submersible autonomous sensor for spectrophotometric pH measurements of natural waters. Anal.

Chim. Acta 75, 1844–1850.

- Martz, T.R., Connery, J.G., Johnson, K.S., 2010. Testing the Honeywell Durafet for seawater pH applications. Limnol. Oceangr. Methods 8, 172–184.
- Martz, T.R., Daly, K.L., Byrne, R.H., Stillman, J.H., Turk, D., 2015. Technology for ocean acidification research: needs and availability. Oceanography 28 (2), 40–47.
- Matson, P.G., Martz, T.R., Hofmann, G.E., 2011. High-frequency observations under Antarctic sea ice in the southern Ross Sea. Antarct. Sci. 23 (06), 607–613.
- McLaughlin, K., Dickson, A., Weisberg, S.B., Coale, K., Elrod, V., Hunter, C., Johnson, K.S., Kram, S., Kudela, R., Martz, T., Negrey, K., Passow, U., Shaughnessy, F., Smith, J.E., Tadesse, D., Washburn, L., Weis, K.R., 2017. An evaluation of ISFET sensors for coastal pH monitoring applications. Regional Stud. Mar. Sci. 12, 11–18.
- Millero, F.J., 1986. The pH of estuarine waters. Limnol. Oceanogr. 31 (4), 839–847. Millero, F.J., Graham, T.B., Huang, F., Bustos-Serrano, H., Pierrot, D., 2006. Dissoci-
- ation constants of carbonic acid in seawater as a function of salinity and temperature. Mar. Chem. 100 (1), 80–94.
- Morel, F.M.M., Hering, J.G., 1993. Principles and Applications of Aquatic Chemistry, first ed. Wiley-Interscience.
- Mosley, L.M., Husheer, S.L.G., Hunter, K.A., 2004. Spectrophotometric pH measurement in estuaries using Thymol Blue and m-Cresol Purple. Mar. Chem. 91 (1–4), 175–186.
- Newton, J.A., Feely, R.A., Jewett, E.B., Williamson, P., Mathis, J., 2015. Global Ocean Acidification Observing Network: Requirement and Governance Plan, second 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., Yool, A., 2005. Anthropogenic ocean acidification over the twenty-first century and its impacts on calcifying organisms. Nature 437 (7059), 681–686.
- Papadimitriou, S., Loucaides, S., Rérolle, V., Achterberg, E.P., Dickson, A.G., Mowlem, M., Kennedy, H., 2016. The measurement of pH in saline and hypersaline media at sub-zero temperatures: characterization of Tris buffers. Mar. Chem. 184, 11–20.
- Patsavas, M.C., Byrne, R.H., Liu, X., 2013a. Purification of *meta*-Cresol Purple and Cresol Red by flash chromatography: procedures for ensuring accurate spectrophotometric seawater pH measurements. Mar. Chem. 150, 19–24.
- Patsavas, M.C., Byrne, R.H., Liu, X., 2013b. Physical-chemical characterization of purified Cresol Red for spectrophotometric pH measurements in seawater. Mar. Chem. 155, 158–164.
- Patsavas, M.C., Byrne, R.H., Wanninkhof, R., Feely, R.A., Cai, W.-J., 2015. Internal consistency of marine carbonate system measurements and assessments of aragonite saturation state: insights from two US coastal cruises. Mar. Chem. 176, 9–20.
- Peltzer, E., 2007. Model II least squares fit: lsqfitma.m. http://www.mbari.org/staff/e tp3/regress/lsqfitma.m.
- Pierrot, D., Lewis, E., Wallace, D.W.R., 2006. CO2SYS DOS Program Developed for CO₂ 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-HCI. Mar. Chem. 162, 89–95.
- Raven, J., Caldeira, K., Eldersfield, H., Hoegh-Guldberg, O., Liss, P., Riebesell, U., Shepherd, J., Turley, C., Watson, A., 2005. Ocean Acidification Due to Increasing Atmospheric Carbon Dioxide. Royal Society.
- Rérolle, V., Ruiz-Pino, D., Rafizadeh, M., Loucaides, S., Papadimitriou, S., Mowlem, M., Chen, J., 2016. Measuring pH in the arctic ocean: colorimetric method or SeaFET? Methods Oceanogr. 17, 32–49.
- Rérolle, V.M.C., Floquet, C.F.A., Harris, A.J., Mowlem, M.C., Bellerby, R.R., Achterberg, E.P., 2013. Development of a colorimetric microfluidic pH sensor for autonomous seawater measurements. Anal. Chim. Acta 786, 124–131.
- Rérolle, V.M.C., Floquet, C.F.A., Mowlem, M.C., Bellerby, R.R., Connelly, D.P., Achterberg, E.P., 2012. Seawater-pH measurements for ocean-acidification observations. Trends Anal. Chem. 40, 146–157.
- Ribas-Ribas, M., Rérolle, V., Bakker, D.C.E., Kitidis, V., Lee, G., Brown, I., Achterberg, E.P., Hardman-Mountford, N., 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., 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. Ecol. Inf. 36, 209–220.
- Rodriguez, C., Huang, F., Millero, F.J., 2015. The partial molal volume and compressibility of Tris and Tris-HCl in water and 0.725 m NaCl as a function of temperature. Deep Sea Res. Pap. Part I Oceanogr. Res. Pap. 104, 41–51.
- Sandifer, J., Voycheck, J., 1999. A review of biosensor and industrial applications of pH-ISFETs and an evaluation of Honeywell's "DuraFET." Microchim. Acta 131, 91–98.
- Seidel, M.P., DeGrandpre, M.D., Dickson, A.G., 2008. A sensor for in situ indicatorbased measurements of seawater pH. Mar. Chem. 109 (1–2), 18–28.
- Sharp, J.H., Yoshiyama, K., Parker, A.E., Schwartz, M.C., Curless, S., Beauregard, A.Y., Ossolinski, J.E., Davis, A.R., 2009. A biogeochemical view of estuarine eutrophication: lessons from the Delaware Estuary. Estuaries Coasts 32 (6), 1023–1043.
- Soli, A.L., Brody, J.P., Byrne, R.H., 2013. The effect of pressure on *meta*-Cresol Purple protonation and absorbance characteristics for spectrophotometric pH measurements in seawater. Mar. Chem. 157, 162–169.
- Stumm, W., Morgan, J.J., 1996. Aquatic Chemistry: Chemical Equilibria and Rates in Natural Waters, third ed. Wiley-Interscience. Takeshita, Y., Frieder, C.A., Martz, T.R., Ballard, J.R., Feely, R.A., Kram, S., Nam, S.,
- Fakeshita, Y., Frieder, C.A., Martz, T.R., Ballard, J.R., Feely, R.A., Kram, S., Nam, S., Navarro, M.O., Price, N.N., Smith, J.E., 2015. Including high-frequency variability in coastal ocean acidification projections. Biogeosciences 12 (19), 5853–5870.
- Takeshita, Y., Martz, T.R., Coletti, L.J., Dickson, A.G., Jannasch, H.W., Johnson, K.S., 2016a. The effects of pressure on pH of Tris buffer in synthetic seawater. Mar. Chem. 188, 1–5.
- Takeshita, Y., Martz, T.R., Johnson, K.S., Dickson, A.G., 2014. Characterization of an ion-selective field effect transistor and chloride ion-selective electrodes for pH measurements in seawater. Anal. Chem. 86 (22), 1189–1195.
- Takeshita, Y., McGillis, W., Briggs, E.M., Carter, A., Donham, E., Martz, T.R., Price, N.N., Smith, J.E., 2016b. Assessment of net community production and calcification of a coral reef using a boundary layer approach. J. Geophys. Res. Oceans 121 (8), 5655–5671.
- Ullman, W.J., Aufdenkampe, A., Hays, R.L., Dix, S., 2013. Nutrient Exchange between a Salt Marsh and the Murderkill Estuary, Kent County, Delaware. Part C. Technical Report. Delaware Department of Natural Resources and Environmental Control, Dover, DE (unpublished).
- Voynova, Y.G., Lebaron, K.C., Barnes, R.T., 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. Estuar. Coast. Shelf Sci. 160, 33–48.
- Wallace, R.B., Baumann, H., Grear, J.S., Aller, R.C., Gobler, C.J., 2014. Coastal ocean acidification: the other eutrophication problem. Estuar. Coast. Shelf Sci. 148, 1–13.
- Waters, J.F., Millero, F.J., 2013. The free proton concentration scale for seawater pH. Mar. Chem. 149, 8–22.
- Whitfield, M., Butler, R.A., Covington, A.K., 1985. The determination of pH in estuarine waters. I: definition of pH scales and selection of buffers. Oceanol. Acta 8 (4), 423–432.
- Wong, K.-C., Dzwonkowski, B., Ullman, W.J., 2009. Temporal and spatial variability of sea level and volume flux in the Murderkill Estuary. Estuar. Coast. Shelf Sci. 84, 440–446.
- Yang, B., Byrne, R.H., Lindemuth, M., 2015. Contributions of organic alkalinity to total alkalinity in coastal waters: a spectrophotometric approach. Mar. Chem. 176, 199–207.
- Yao, W., 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., Byrne, R.H., 2007. Impurities in indicators used for spectrophotometric seawater pH measurements: assessment and remedies. Mar. Chem. 107 (2), 167–172.
- Zeebe, R.E., Wolf-Gladrow, D.A., 2001. CO2 in Seawater: Equilibrium, Kinetics, Isotopes (No. 65). Gulf Professional Publishing.