



PERFORMANCE VERIFICATION STATEMENT For the Eureka Manta2 pH Sensor

TECHNOLOGY TYPE:	pH sensors
APPLICATION:	In situ estimates of pH for coastal moored deployments
PARAMETERS EVALUATED:	Response linearity, accuracy, precision and reliability
TYPE OF EVALUATION:	Laboratory and Field Performance Verification
DATE OF EVALUATION:	Testing conducted from December 2012 through July 2014
EVALUATION PERSONNEL:	T. Johengen, G.J. Smith, D. Schar, M. Atkinson, H. Purcell, D. Loewensteiner, Z. Epperson and M. Tamburri

NOTICE:

ACT verifications are based on an evaluation of technology performance under specific, agreed-upon protocols, criteria, and quality assurance procedures. ACT and its Partner Institutions do not certify that a technology will always operate as verified and make no expressed or implied guarantee as to the performance of the technology or that a technology will always, or under circumstances other than those used in testing, operate at the levels verified. ACT does not seek to determine regulatory compliance; does not rank technologies nor compare their performance; does not label or list technologies as acceptable or unacceptable; and does not seek to determine “best available technology” in any form. The end user is solely responsible for complying with any and all applicable federal, state, and local requirements.

This document has been peer reviewed by ACT Partner Institutions and a technology-specific advisory committee and was recommended for public release. Mention of trade names or commercial products does not constitute endorsement or recommendation by ACT for use.

Questions and comments should be directed to: Dr. Mario Tamburri
c/o Alliance for Coastal Technologies
Chesapeake Biological Laboratory
PO Box 38 / 146 Williams Street
Solomons, Maryland 20688, USA
Email: Tamburri@umces.edu

TABLE OF CONTENTS

EXECUTIVE SUMMARY	3
BACKGROUND AND OBJECTIVES	5
INSTRUMENT TECHNOLOGY TESTED	6
PH SCALES.....	7
SUMMARY OF VERIFICATION PROTOCOLS.....	8
ANALYSIS OF REFERENCE SAMPLES	9
FIELD TEST	9
RESULTS OF MOORED FIELD TESTS.....	14
QUALITY ASSURANCE/QUALITY CONTROL	39
QUALITY CONTROL SAMPLES	39
EXTERNAL QUALITY ASSURANCE AND QUALITY CONTROL ASSESSMENTS	42
REFERENCES.....	44
ACKNOWLEDGEMENTS	46
APPENDIX I.....	47

EXECUTIVE SUMMARY

The Alliance for Coastal Technology (ACT) conducted a sensor verification study of in situ pH sensors during 2013 and 2014 to characterize performance measures of accuracy and reliability in a series of controlled laboratory studies and field mooring tests in diverse coastal environments. A ten week long laboratory study was conducted at the Hawaii Institute of Marine Biology and involved week long exposures at a full range of temperature and salinity conditions. Tests were conducted at three fixed salinity levels (0.03, 22, 35) at each of three fixed temperatures (10, 20, 30 °C). Ambient pH in the test tank was allowed to vary naturally over the first five days. On the sixth day the pH was rapidly modified using acid/base additions to compare accuracy over an extended range and during rapid changes. On the seventh day the temperature was rapidly shifted to the next test condition. On the tenth week a repeated seawater trial was conducted for two days while the temperature was varied slowly over the 10 – 30 °C range. Four field-mooring tests were conducted to examine the ability of test instruments to consistently track natural changes in pH over extended deployments of 4-8 weeks. Deployments were conducted at: Moss Landing Harbor, CA; Kaneohe Bay, HI; Chesapeake Bay, MD; and Lake Michigan, MI. Instrument performance was evaluated against reference samples collected and analyzed on site by ACT staff using the spectrophotometric dye technique following the methods of Yao and Byrne (2001) and Liu et al. (2011). A total of 263 reference samples were collected during the laboratory tests and between 84 – 107 reference samples were collected for each mooring test. This document presents the results of the Eureka Manta2 pH sensor which is a glass electrode with a KCl reference electrode. The Manta2 was submitted for testing after the Laboratory study was completed under a new Request for Technologies, therefore only results for the Field testing component are presented.

At Moss Landing Harbor the field deployment test was conducted over 28 days with a mean temperature and salinity of 16.6 °C and 33. The measured ambient pH range from our 84 discrete reference samples was 7.933 – 8.077. At this site two instrument sondes were deployed with identical pH sensors, however, one was enclosed in a copper mesh screen to serve as an anti-fouling measure and one in a plain sensor guard. The Manta2 with the copper screen operated for 23 consecutive days before experiencing a power failure. Ambient pH measured by this unit ranged from 7.95 to 8.63. The average and standard deviation of the measurement difference between the Manta2 and reference pH was 0.258 ± 0.181 with a total range of -0.014 to 0.551 (N=64). The Manta2 with no anti-fouling operated continuously for the entire 28 days of the deployment, resulting in 2579 observations at 15 minute intervals. Ambient pH measured by this Manta2 sonde ranged from 8.28 to 8.60. The average and standard deviation of the measurement difference between this Manta2 and reference pH was 0.512 ± 0.059 with a total range of 0.333 to 0.584 (N=84).

At Kaneohe Bay the field deployment test was conducted over 88 days with a mean temperature and salinity of 24.5 °C and 34.4. The measured ambient pH range from our 101 discrete reference samples was 7.814 – 8.084. The Manta2 sonde operated continuously throughout the first 67 days of deployment measuring at 30 minute intervals, however, a tunicate colonized directly on the pH sensor bulb causing it to fracture and the data beginning on January 30th were not useable for comparisons (pH instantly dropped from 8.1 to 6.6 indicating failure). Ambient pH measured by the Manta2 ranged from 7.85 to 8.39. The average and standard deviation of the measurement difference between instrument and reference was 0.17 ± 0.02 with a total range in the differences of 0.12 to 0.20 (N=75).

At Chesapeake Bay the field deployment test was conducted over 30 days with a mean temperature and salinity of 5.9 °C and 12.8. The measured ambient pH range from our 107 discrete reference samples was 8.024 – 8.403. The Manta2 operated successfully over the entire deployment and generated 2756 observations based on its 15 minute sampling interval. Ambient pH measured by the Manta2 ranged from 8.14 to 8.54. The average and standard deviation of the measurement difference between instrument and reference pH was 0.12 ± 0.02 , with the total range of differences from 0.09 to 0.16 (N=107).

At Lake Michigan the field deployment test was conducted over 29 days with a mean temperature and salinity of 21.2 °C and 0.03. The measured pH range from our 98 discrete reference samples was 8.013 - 8.526. The Manta2 operated continuously over the entire deployment and generated 2680 observations based on its 15 minute sampling interval. Ambient pH measured by the Manta2 ranged from 7.86 to 8.51. The average and standard deviation of the difference between instrument and reference pH was -0.07 ± 0.04 with a total range of -0.20 to -0.004 (N=98).

A summary plot of all four field tests indicates that the Manta2 responded consistently with good agreement to reference pH measurements at three of the four test sites. No explanation or observed failure was discovered to explain the performance at the Moss Landing test site. Results for the brackish and saline test sites show an expected offset due to calibration and scale differences between NBS buffers and the pH_{total} scale reported by the spectrophotometric dye measurement.

Lastly, it is worth emphasizing that the continuous 15 – 30 minute time-series provided by the test instrument was able to resolve a significantly greater dynamic range and temporal resolution than could be obtained from discrete reference samples. There were no obvious changes in the differences between instrument and reference measurements during the duration of the mooring test, indicating that biofouling and instrument drift had not significantly affected measurement performance over these deployment durations. Continuous in situ monitoring technologies, such as the Eureka Manta 2, provide critical research and monitoring capabilities for helping to understand and manage important environmental processes such as carbonate chemistry and ocean acidification, as well as numerous other environmental or industrial applications.

BACKGROUND AND OBJECTIVES

Instrument performance verification is necessary so that effective existing technologies can be recognized and so that promising new technologies can be made available to support coastal science, resource management and ocean observing systems. To this end, the NOAA-funded Alliance for Coastal Technologies (ACT) serves as an unbiased, third party testbed for evaluating sensors and sensor platforms for use in coastal environments. ACT also serves as a comprehensive data and information clearinghouse on coastal technologies and a forum for capacity building through workshops on specific technology topics (visit www.act-us.info).

As part of our service to the coastal community, ACT conducted a performance verification of commercially available, in situ pH sensors through the evaluation of objective and quality assured data. The goal of ACT's evaluation program is to provide technology users with an independent and credible assessment of instrument performance in a variety of environments and applications. Therefore, the data and information on performance characteristics was focused on the types of information that users most need.

The fundamental objectives of this Performance Verification were to: (1) highlight the potential capabilities of particular in situ pH sensors by demonstrating their utility in a range of coastal environments; (2) verify the claims of manufacturers on the performance characteristics of commercially available pH sensors when tested in a controlled laboratory setting, and (3) verify performance characteristics of commercially available pH sensors when applied in real world applications in a diverse range of coastal environments.

In 2012 ACT completed a customer needs and use assessment on pH sensors. Scientists, resource managers, and other users of these technologies were asked about their current use or application of these instruments, their perceptions of limitations or problems with the technology, and the most important criteria they use when selecting a sensor or instrument package. The results of these assessments, conclusions from the 2005 Measurement of Dissolved Inorganic Carbon Speciation in Natural Waters workshop, and discussions with the Technical Advisory Committee were used to identify the main applications and key parameters that ACT evaluates in this Technology Verification.

Over 170 coastal resource managers, environmental health agency representatives, manufacturers, and scientific researchers were surveyed regarding their needs, uses and current practices of making pH measurements in the environment. Overall, 42 individuals responded to the survey, all of which are using, or manufacturing, in situ pH sensors. Respondents were allowed to select more than one answer when appropriate so percentages do not add up to 100%. The types of pH sensors being used varied between potentiometric (45%), ion-selective field-effect transistor (32%), and spectrophotometric (using indicator dye) 20%, but with a quarter of respondents using more than one type. The greatest area of use among respondents was academic research (76%) followed by state and federal resource management (40%). The environmental applications were broad with coastal ocean greatest (74%), followed by estuarine (64%), followed by open-ocean (41%), and followed by freshwater (38%). Correspondingly, these environments included temperature ranges from -5 to 50 °C, with medians of 5°C and 28°C when responses were binned into low and high ranges. Similarly, salinities ranged from 0 – 100, with low and high bin medians of 15 and 35, respectively. The range of pH measured by the respondents in these applications was between 4.0 and 11.0, with low and high bin median values of 7.0 and 8.3. Remote deployment was the most common method of use (74%),

followed by depth profiling (50%), then hand-held portable use (48%), then flow-through systems (26%). Respondents used a variety of calibration procedures including commercial buffers (68%), CO₂ chemistry (35%), seawater CRMs (23%), pH indicator dyes (18%), and supplied by manufacturer (13%). The four areas where respondents expressed the greatest concern over the use of in situ pH sensors were ruggedness (49%), calibration life (46%), level of measurement uncertainty (43%), and reliability (41%). The complete needs and use assessment reports can be found at:

http://www.act-us.info/Download/Customer_Needs_and_Use/pH/index.html

INSTRUMENT TECHNOLOGY TESTED

ACT released an additional request-for-technologies for companies to participate in the pH verification after the initial Laboratory component had taken place. The Eureka Manta2 instrument was submitted during this call and therefore this statement only includes results for the Moored Field Deployment component of this pH sensor Verification.

The Eureka pH sensing system uses a robust glass sensor optimized for stability and reliability in the pH 2 - pH 12 range. The pH sensor is short, minimizing the amount of buffer solution needed during calibration. The system is equipped with a separate, refillable reference electrode with a replaceable junction cap, designed to reduce operational costs and increase accuracy. Most combination pH electrodes fail due to reference electrode problems, requiring the entire probe to be replaced.

The Eureka calibration software allows the user to calibrate with either 2 or 3 reference buffers, in any order. While calibration is often done with pH 4, 7, & 10 buffers, the user has the option to use any pH buffer, in an attempt to increase the accuracy of the system in a particular sample water. The calibration software also alerts the user if the values of the sensor are outside of the expected range, alerting the user to potential buffer contamination or data entry errors. The software creates and maintains a Calibration Log that records all calibration results, including the temperature, millivolt reading, and pre- and post- readings in each buffer. The logging software will also display and record probe millivolt readings, as well as pH values.

Eureka uses a high impedance, high resolution circuit to measure the output of the pH probe. The circuit board is sealed with a conformal coating and designed to maintain isolation of the probe connections to avoid faulty readings from electrical leakage or moisture condensation. The probe is also electrically disconnected from the instrument during sleep intervals, to avoid any possible leakage currents that might compromise the probe's lifetime or performance.

In addition, the unique design of the reference electrode improves the performance of the system in several ways:

- The electrolyte chamber is easily refillable, and designed to purge the probe of air bubbles and clean the Teflon junction during refilling, to optimize accuracy.
- The reference electrode uses a porous Teflon junction to reduce fouling; its large surface area ensures constant electrolyte diffusion, minimizing the junction potential errors often encountered when ceramic junctions get dirty.
- The large volume of reference electrolyte extends the field deployment time up to 6 weeks, and allows the addition of KCl pellets to even further extend deployment time.

- The large volume of electrolyte coupled with the relatively fast-diffusing Teflon junction also decreases junction potential errors (dirty junctions can cause errors exceeding 0.5 pH units). This also improves the accuracy of the system in low-conductivity waters.
- The Eureka reference electrolyte is not gelled, eliminating the common junction potential errors of a slightly depleted or contaminated gel.
- The Eureka reference electrode is filled with 3.3 M KCl instead of the typical saturated KCl solution, eliminating KCl precipitation in cold waters that can cause errors of up to ± 0.2 pH units. When extended deployment times (>6 weeks) are required, Eureka can also provide fast-dissolving KCl pellets (results may be temperature corrected to reduce KCl precipitation errors).

The Manta2 was calibrated by ACT staff using commercial NBS buffers prior to each moored field deployment. A two point calibration was done using pH buffers 7 and 10 from Fisher Scientific following the standard operating procedures provided by the company at a training workshop.

pH SCALES

Four pH scales are commonly used to describe the acidity of an aqueous solution: (1) the free hydrogen ion concentration scale, (2) the total hydrogen ion concentration scale, (3) an H^+ concentration scale termed the ‘seawater scale’ that is numerically quite similar to the total hydrogen ion concentration scale, and a fourth pH scale that is expressed in terms of an ‘apparent activity’. The relationship between the apparent activity scale and the other three scales is not thermodynamically well defined.

The free hydrogen ion concentration scale is expressed as $pH = -\log [H^+]$, where $[H^+]$ indicates the concentration of a free hydrated hydrogen ion. The total scale (pH_T) is written as $pH_T = -\log [H^+]_T$ where $[H^+]_T \approx [H^+] + [HSO_4^-]$, and the seawater scale pH (pH_{sws}) is written as $pH_{sws} = -\log [H^+]_{sws}$ where $[H^+]_{sws} \approx [H^+] + [HSO_4^-] + [HF^0]$.

The various pH scales are inter-related by the following equations:

$$a_H = 10^{pH(NBS)} = f_H * [H^+]/(1 \text{ mol/kg}), \text{ and} \quad (1)$$

$$[H^+] = [H^+]_{tot} / (1 + {}_TS/K_{SO4}) = [H^+]_{sws} / (1 + {}_TS/K_{HSO4} + {}_TF/K_{HF}) \quad (2)$$

where a_H is an apparent activity, f_H is an apparent activity coefficient for the H^+ ion, ${}_TS$ and ${}_TF$ are total concentrations of SO_4^{2-} and fluoride ions in seawater, and K_{HSO4} and K_{HF} are HSO_4^- and HF^0 dissociation constants in seawater.

Because electrodes transferred between dilute buffers (e.g., NBS buffers) and concentrated electrolytes (such as seawater) are subject to variations in the liquid junction potentials of reference electrodes (an effect whose magnitude is generally specific to a particular electrode), the f_H proportionality factor that relates the apparent activity scale to concentration scales is, in part, influenced by factors that are not strictly thermodynamic. Consequently, use of the apparent activity scale is generally discouraged for measurements in seawater.

Relationships between the different concentration scales depend on temperature, salinity, and pressure. At 25°C, salinity 35, and 1 atm, relationships between seawater scale, total scale and free scale pH (pH_{sws} , pH_{T} and pH) are given approximately as:

$$pH \text{ minus } pH_{\text{T}} = 0.108 \quad (4)$$

$$pH_{\text{T}} \text{ minus } pH_{\text{sws}} = 0.008 \quad (5)$$

$$pH_{\text{NBS}} \text{ minus } pH_{\text{sws}} \approx 0.13 \quad (6)$$

Equation (4) was derived using the K_{SO_4} characteristics given in DOE (1994). Alternative descriptions of the HSO_4^- dissociation equilibrium can produce significant variations in the offset between total scale and free scale pH. Determinations of K_{F} in seawater are sufficiently coherent that deviations from the relationship given in eqn. (5) are quite small. The relationship between pH_{NBS} and pH_{sws} (eqn. 6) is based on the discussion of Culberson (1981). It is important to note that f_{H} , and therefore eqn. (6), is electrode-dependent. Measurements of a_{H} on the NBS scale are expressed in terms of mol/kg- H_2O , while the other scales express concentrations in terms of mol/kg_(sw).

Spectrophotometric measurements of pH in seawater using purified meta-cresol purple are generally expressed on the total H^+ scale. Using Tris seawater buffers, glass electrodes are also commonly calibrated on the total scale (pH_{T}).

There are significant changes in junction potentials when electrodes are calibrated in dilute media and then utilized for measurements in high ionic strength media such as seawater. In such cases there is no direct way to relate pH_{NBS} values to pH_{T} or pH_{sws} values unless electrodes are calibrated in both dilute buffers (NIST buffers) and also artificial seawater buffers such as Tris.

SUMMARY of VERIFICATION PROTOCOLS

The protocols used for this performance verification were developed in conference with ACT personnel, the participating instrument manufacturers and a technical advisory committee. The protocols were refined through direct discussions between all parties during two pH Sensor Performance Verification Protocol Workshops held on 18-20 June, 2012 and 26-28 June, 2013 in Ann Arbor, MI. All ACT personnel involved in this verification were trained on use of instruments by manufacturer representatives and on standardized water sampling, storage, analysis and shipping methods during a training workshop held on 3-6 Dec, 2012 in Kaneohe, HI. The manufacturer representatives and the ACT Chief Scientist verified that all staff were trained in both instrument and sample collection protocols.

This performance verification report presents instrument output in derived pH values reported over time as directly downloaded from the test instruments or captured through independent dataloggers. A summary of the testing protocols is provided below. A complete description of the testing protocols is available in the report, *Protocols for the Performance Verification of In Situ pH Sensors* (ACT PV12-01) and can be downloaded from the ACT website (<http://www.act-us.info/Download/Evaluations/pH/Protocols/index.html>).

Analysis of Reference Samples

The reference pH measurements were the pH of discrete water samples as determined by pH indicator dyes, either, meta-cresol purple or phenol red (freshwater) (Liu et al. 2011, Yao and Byrne 2001 respectively). All dyes used throughout the Verification were prepared and characterized by Bob Byrne at the University of South Florida. An initial reading for background correction was taken on each cell and then dye (from Byrne's lab) was introduced into each cell (10 μ l of purified 10 mM meta-cresol purple for samples with salinity > 5 and 10 μ l of 10 mM phenol red for freshwater samples), mixed thoroughly, and the dye R ratio measured on an Agilent 8453 spectrophotometer (the R ratio is the ratio of absorbance measured at wavelengths corresponding to the peak absorbance of the acid and base forms of the dye).

The R value of each cell was measured a minimum of 5 times in succession in the field test and 10 times in succession for the lab test. pH was calculated from the R ratio, temperature and salinity according to the equations published in Liu et al. 2011 (meta-cresol purple) or Yao and Byrne 2001 (phenol red); both pH total scale and pH free scale are reported (Liu et al. 2011, Eq. 12). The actual sample pH at ambient temperature was calculated using CO₂-Sys with dissociation constants from Millero et al. (2006) and an estimated value of the alkalinity that is based on the average of the samples analyzed by SIO. The anticipated level of accuracy of these measurements is better than 0.01 pH for seawater ($S \geq 20$), but actual laboratory based accuracy and precision levels were quantified and reported for the brackish and freshwater environments.

Moored Field Tests

Moored field tests were conducted at four ACT Partner Institution sites covering freshwater, estuarine, and open-ocean conditions. The test sites include Moss Landing Marine Laboratories (MLML), CA; Hawaii Institute of Marine Biology (HIMB), HI; Chesapeake Biological Laboratory (CBL), MD; and Lake Michigan, MI. The field tests occurred in sequence, starting with Moss Landing on August 22, 2013. The duration of the moored deployment tests were 4 weeks, except at the Hawaii test site where instruments were deployed for a period of 12 weeks. Instruments were removed from the water only after the test period was complete.

Instrument Setup – Prior to deployment, all instruments were set up at the field site by a trained ACT staff member following established protocols that had been developed in collaboration with the manufacturers at the training session that occurred in December, 2012 (at the beginning of the ACT laboratory test). Manufacturers that did not conduct an in-person training session sent written protocols and video demonstration of desired handling procedures. As appropriate, and where requested by the manufacturer, submitted test instruments were initially calibrated by ACT staff using a common batch of commercial, non-dyed, NIST traceable buffers at a pH of 7.02 and 10.06. (Each sensor was calibrated with buffer from its own bottle, but they were all from the same LOT and shipment). All calibrations occurred in a constant temperature bath at 20.0°C (except at HI which was done at 25°C). Any deviation from this temperature was noted and used to adjust the stated pH values of the buffers as defined on the bottles.

After calibration and prior to deployment (again where appropriate for each instrument) instruments were exposed to additional certified reference solutions. At Moss Landing, HIMB

and CBL, instruments were exposed to batches of Seawater CRMs produced by the CO2-QC facility at Scripps Institution of Oceanography. (CRMs were batched into a single container before distributing into individual calibration cups for each sensor. Three cuvettes were filled immediately upon mixing the batch and again at the end of pouring out the solution to measure pH and ensure consistency of each aliquot.) For the freshwater test in Lake Michigan instruments were exposed to the NIST buffers or lake water characterized by spectrophotometric analysis by ACT personnel as appropriate for the instrument. All pre-deployment instrument measurements of the reference buffers were in a water bath at a constant temperature (20 °C, except at HIMB which were at 25 °C) and all solutions and instruments were pre-equilibrated at this temperature.

For the deployment, instruments were programmed to record data based on a time interval that allowed for a 30 day (or 90 day for HI) deployment. Intervals were selected such that there was a common 30 (or 60) minute interval achieved by all instruments. This schedule allowed us to coordinate our reference sampling for all instruments. Internal clocks were set to local time and synchronized against the time standard provided by www.time.gov. In high flow coastal environments, clock drift could lead to significant bias.

Instrument Deployment – A photograph of each individual instrument and the entire instrument rack was taken just prior to deployment and just after recovery to provide a qualitative estimate of biofouling during the field tests. Instruments were set-up as self-recording on a deployment rack and arranged so that a single representative field sample could be collected within 1 meter of any individual sampling inlet. The deployment frames were arranged so that all of the instruments remain at a fixed depth of 1 m below the water surface (using a float system or fixed dock in environments not affected by tidal changes or strong wave action). Two calibrated SeaBird CTD packages and four RBR Solo thermistors were attached to the mooring at each test site in order to accurately characterize the temperature and salinity heterogeneity surrounding the mooring. In these dynamic coastal regions it was critical to understand the spatial and temporal dynamics of the water column in order to interpret pH measurements appropriately.

Instrument Retrieval – After the instruments were retrieved and cleaned of all removable fouling according to written procedures provided by the manufacturer, a final exposure test was conducted in the CRM buffers as defined above. For instruments that pumped samples through an inlet, the inlet was connected to the manufacturer's supplied storage bag to ensure that no air was introduced prior to exposure in the buffers.

Reference Water Sampling Schedule – The sampling frequency was structured to examine changes in pH over daily and weekly time scales. Specifically, an intensive sampling event was conducted once a week that consisted of 5 sample collections within a day. During four additional days of each week, there was a minimum of twice per day sampling, scheduled in a manner to capture as much diurnal variation as possible. The initial intensive sampling event occurred within the first two days of the deployment after all instruments had been deployed, and the final intensive sampling event occurred during the last two days of the deployment. The schedule provided a higher density of comparative data at the beginning when instruments should have been functioning at optimum performance and again after the challenge of a four or twelve week deployment. The sampling schedule resulted in at least 60 reference samples paired with instrument measurements. For the 12 week deployment test at HIMB, the sampling scheme was modified to spread out a similar number of samples over the extended time period. All

sampling times were recorded on logsheets and entered into a database for final data comparisons.

Reference Water Sample Collection – A standard 2.2 or 4.2L Van Dorn water sampler was used at each field test site to collect water samples for reference pH measurements. Water sample collections were timed to correspond directly with the instrument readings or sample intake. The water sampler was lowered to the same depth of the instrument sampling inlets, and as close as physically possible to the inlets (distant by no more than 0.5 m horizontal distance). The water sampler was soaked at sampling depth for 1 minute prior to sampling. If water was not flowing, the sampler was moved to ensure that it was flushed with the ambient water. The water sampler was triggered to match the programmed sampling times of each instrument. Three replicate pH samples were collected in clean 10 cm glass spectrometer cuvettes from each individual field sample. The cuvettes were gravity filled from PTFE tubing connected to the sampling bottle spigot such that the volume of the cuvette was exchanged three times (about 20-30 secs) before a final sample is collected. Care was taken to ensure that no bubbles were left within the cuvette before sealing the ports with their PTFE stoppers. Cuvettes were stored at ambient or slightly cooled conditions during transport to the laboratory (travel times at various field sites ranged from 5 to 20 minutes) until they were placed into the temperature equilibration chamber at the specified temperature for analysis. In addition, at the freshwater site an additional 300 ml BOD bottle was filled to allow for a lab-based electrode pH measurement on each field reference sample. All samples were equilibrated to 25°C and analyzed on a lab electrode calibrated daily before use.

Twice a week (day 1 and day 5) duplicate water samples were collected to characterize all of the CO₂ parameters. (This sample collection occurred in conjunction with our field duplicate sampling and one set of external partner samples came from each Van Dorn. This protocol was mostly to facilitate limitations of volume but also helped to evaluate heterogeneity at the mooring.) Water samples were collected and preserved for pH, pCO₂, TCO₂ and TA measurements following standard oceanic protocols (Dickson et al., 2007). In brief, a 500 mL glass sample bottle (supplied by Andrew Dickson, SIO) was filled from the Van Dorn sampling bottle using PTFE tubing and preserved with 110 µL of saturated HgCl₂ solution. Reference samples shipped to Oregon State University (Burke Hales, OSU) were analyzed for pCO₂ and TCO₂. Reference samples shipped to Scripps Institution of Oceanography (Andrew Dickson) were analyzed for TA and pH. Reference samples shipped to the University of South Florida (Robert Byrne) were analyzed for pH. These external sample analyses provided further estimation of the uncertainty of ACT's reference sample measurements through both direct pH measurement and calculated pH from carbonate parameters using CO₂-Sys (Pierrot et al. 2006). However, because these samples were not handled in a manner identical to the unpreserved reference samples that were analyzed within a few hours of collection, they will not be used in any published direct assessment against the test instruments without agreement of the TAC.

All sampling procedures were practiced by ACT staff prior to the beginning of the evaluation, and the Chief Scientist verified that all staff were trained in both instrument and sample collection protocols. A preservation test was done at each site to ensure samples that were shipped for external analysis were stable.

Sample Handling and Custody Requirements - All collected reference water samples for all four field tests were handled in the same manner. All reference samples were processed according to methods defined within these protocols and cited Standard Operating Procedure

documents (Dickson et al. 2007). Each reference sample was dated and coded according to site, test condition and sample sequence. The actual sample container was labeled with a number for identification. The reference sample number was used in all laboratory records and Chain-of-Custody (COC) forms to identify the sample. The COC specified time, date, sample location, unique sample number, requested analyses, sampler name, required completion time, date and time of sample transaction, and name of receiving party for the samples. Proper labeling of sample bottles was critical. The COC was a mechanism by which samples could be tracked through the various phases of the process: collection, shipping, receiving, logging, sample prep/extraction, analysis and final data QA/QC review. Transfer of reference samples from field personnel to laboratory personnel was also recorded on the COC and records were maintained in the laboratory with the names and signatures of persons leaving and receiving the custody. All logs were duplicated weekly. The original log was retained at the ACT Partner site and a copy was sent to the ACT Chief Scientist. Accumulated samples to be analyzed by outside laboratories were shipped for analysis at the end of the month long deployment tests, and monthly for the extended HI deployment. Samples stored on site were routinely inspected by ACT personnel to assure proper preservation and label integrity. All reference samples not immediately analyzed on site by ACT staff were accompanied by the sample collection sheet and COC forms.

Analytical Methods for Reference Samples - Three spectrometer cells (10 cm cylindrical cell) were filled at the deployment site for each field reference sample, and transported directly to the lab. All analysis was done at a fixed temperature for a given test site. The measurement temperature was 25°C for HIMB, 20°C for MLML, 15°C for CBL and 25°C for Lake Michigan. The temperatures for CBL and Lake Michigan were estimated to be near the mid-point of the expected 30 day ambient range. Filled cells were incubated in a cell warmer (manufactured at USF according to Bob Byrne's specifications) to reach the specified analytical temperature (typically between 30 min to 1 hour). The Agilent had a thermal jacket surrounding the spectrophotometer cells that was continuously flushed with the same water bath that also supplied the cell warmer. Past experience at Byrne's lab has shown that this system can maintain cells at a constant temperature within $\pm 0.1^\circ\text{C}$ or better. After the initial dye reading, the sample was re-blanked and then a second 10 μl aliquot of dye was added and the R-ratio re-measured to enable a correction for the effect of the dye addition on the sample pH. By performing the perturbation measurement on each sample we could directly calculate the appropriate adjustment for each sample individually. The final temperature of the solution in the cuvette was measured with a bead thermistor upon completion of the second dye reading and recorded on the datasheets to define any deviation from the specified reading temperature. In addition, at the freshwater site an unpreserved sample collected in a 300 ml BOD bottle was incubated at the same temperature as the equilibration chamber and cell jacket until it reached a constant temperature of 25°C, and the pH was subsequently read on a recent two-point calibrated Metrohm electrode. The electrode was calibrated daily at 25.0°C.

After each use the Van Dorn sampler, fill tubing and cuvettes were thoroughly rinsed with deionized water to prevent any build-up of salts or dye.

Ancillary Environmental Data - At each of the mooring test sites, two calibrated CTD packages were attached to the test rack and positioned to best characterize the salinity surrounding the mooring. The CTDs provided an independent record of conductivity and temperature measured at 15 minute intervals. In addition, four RBR Solo temperature sensors

were deployed at each site to establish an accurate temperature history for the site. The RBR sensor has a stated accuracy and resolution of 0.002°C and 0.00005°C respectively. In conjunction with each water sample collection, technicians recorded basic site-specific conditions on standardized log sheets including: date and time, weather conditions (e.g., haze, % cloud cover, rain, wind speed/direction), air temperature, recent large weather events or other potential natural or anthropogenic disturbances, tidal state and distance from bottom of sensor rack, and any obvious problems or failures with the instruments. Datasheets were transmitted on a weekly basis to the ACT Chief Scientist for data archiving and ACT personnel performance QA/QC.

Each test site either established or identified the closest meteorological station (and river discharge gauge where appropriate) that could record air temperature, humidity, directional wind speed, precipitation on a continuous basis to help identify the timing and intensity of any event based changes at the field test locations.

Ancillary data was used in a qualitative sense to understand the history of weather patterns and changes in ambient water quality conditions. These data were not used for any direct calibration, correction, or statistical comparison to the reported salinity test data.

RESULTS of MOORED FIELD TESTS

Moored field tests were conducted to examine the performance of the Eureka Manta to consistently track natural changes in pH over extended deployment durations of 4-8 weeks. In addition, the field tests examined the reliability of the instrument, i.e., the ability to maintain integrity or stability of the instrument and data collections over time. Reliability was determined by quantifying the percent of expected data that was recovered and useable. The performance of the Manta was examined in four separate field deployment tests at various ACT Partner test sites to include a range of biogeochemical conditions. The range and mean for temperature and salinity for each test site is presented in Table 2. The final reference temperature data was computed from the mean of two RBR thermistors and the SeaBird SBE that were mounted at the same sampling depth as the test instrument. The reference temperature data and salinity were used to compute ambient pH results from the laboratory results using a CO₂-Sys program (Pierrot et.al., 2006). Immediately before each deployment the test instrument was exposed to a certified or internal reference sample for 3-4 measurements to confirm good working status and provide an initial offset against its latest calibration (Table 3).

Table 1. Range and average for temperature, and salinity at each of the test sites during the sensor field deployment. Temperature and salinity were measured by two RBR temperature loggers and a SeaBird SBE 26 (or SBE26plus) mounted on the instrument rack and the duration of the deployment.

SITE (deployment period/duration)		Temperature (°C)	Salinity
Moss Landing Harbor	Min.	14.9	30.7
26Aug – 22Sept	Max.	18.6	33.7
(n = 28 days)	Mean	16.6	33.0
Kaneohe Bay	Min.	22.3	32.8
15Nov – 10Feb	Max.	26.5	35.2
(n = 88 days)	Mean	24.5	34.4
Chesapeake Bay	Min.	3.4	9.3
12Mar – 10Apr	Max.	10.1	14.0
(n = 30 days)	Mean	5.9	12.8
Lake Michigan	Min.	11.6	0.024
21Jun – 19Jul	Max.	24.9	0.035
(n = 29 days)	Mean	21.2	0.031

Table 2. Results of the pre-deployment check for the Manta at each deployment site. The stated pH of the reference solution is based on ACT's spectrophotometric analysis completed at the time of exposure. Seawater CRM provided by Andrew Dixon's Lab (Scripps Institution of Oceanography). NIST buffer prepared by ACT personnel from salts supplied by Kenneth Pratt (NIST).

Deployment Site	Reference Solution	Measurement Temperature	Reference Solution pH	Instrument pH
MLML	Seawater CRM	20 °C	7.97	8.04
HIMB	Seawater CRM	25 °C	7.86	7.93
CBL	Seawater CRM	20 °C	8.00	8.05
GLERL	NIST Buffer	20 °C	9.23	9.14

Moss Landing Marine Laboratories Field Deployment Site

A month-long moored field test was conducted at the Moss Landing Marine Laboratories from August 26 to September 22, 2013. The deployment site was located at 36.80°N, 121.79°W, in the Small Boat Facilities area of Moss Landing Harbor (Photo 1). This secure deployment site was located in the Harbor on the junction of the northern tributary of the Salinas River and Elkhorn Slough National Estuarine Reserve on the central coast of California. Instrumentation was deployed off a floating dock in waters with a tidal range of 2 meters and a maximum depth below the dock of 4 meters.

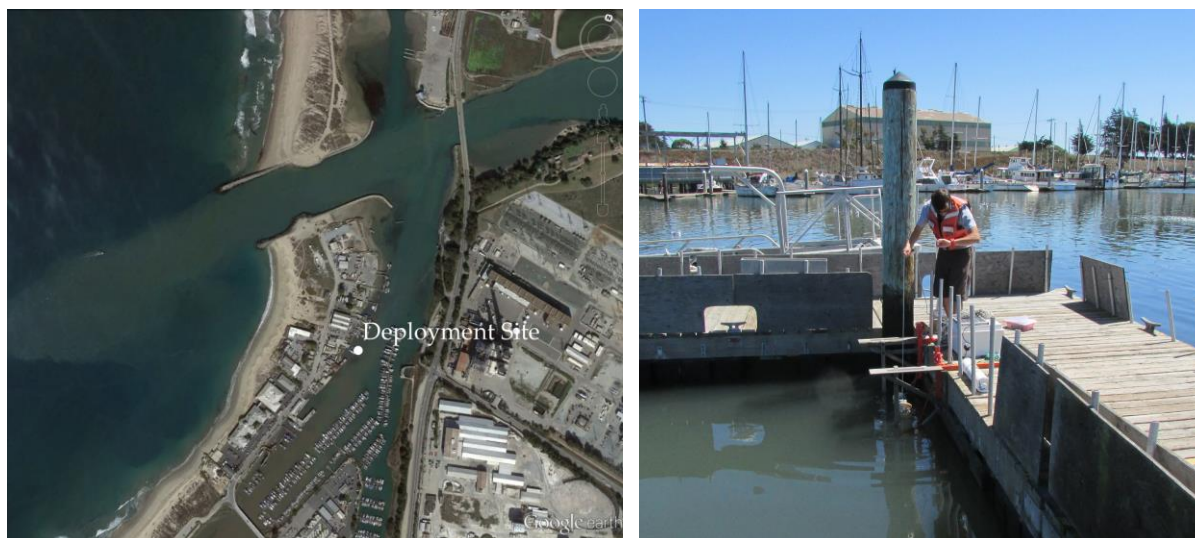


Photo 1. Aerial view of MLML Harbor (left) and dockside mooring deployment (right).

Time series results of ambient conditions for tidal height, temperature, and salinity are given in figure 1. Temperature ranged from 14.9 to 18.6 and salinity from 30.7 to 33.7 over the duration of the field test. The bottom panel displays the maximum difference recorded between all reference thermistors mounted at the same depth and different locations across the mooring rack. The average temperature difference observed across the space of the mooring rack was 0.11 °C with a maximum of 0.8 °C. The difference between the test instruments instantaneous temperature reading and the final reference temperature will contribute to the variability and offset to the calculated ambient reference pH. It is not possible to quantify the exact difference between the averaged reference temperature and that measured by the test instrument, but a difference of 1 °C would amount to an approximate offset of 0.015 pH units at ambient conditions. Additional differences resulting from variation in the actual chemistry of the water across this space may exist but will be minimized because the sampling bottle integrates across the mooring space.

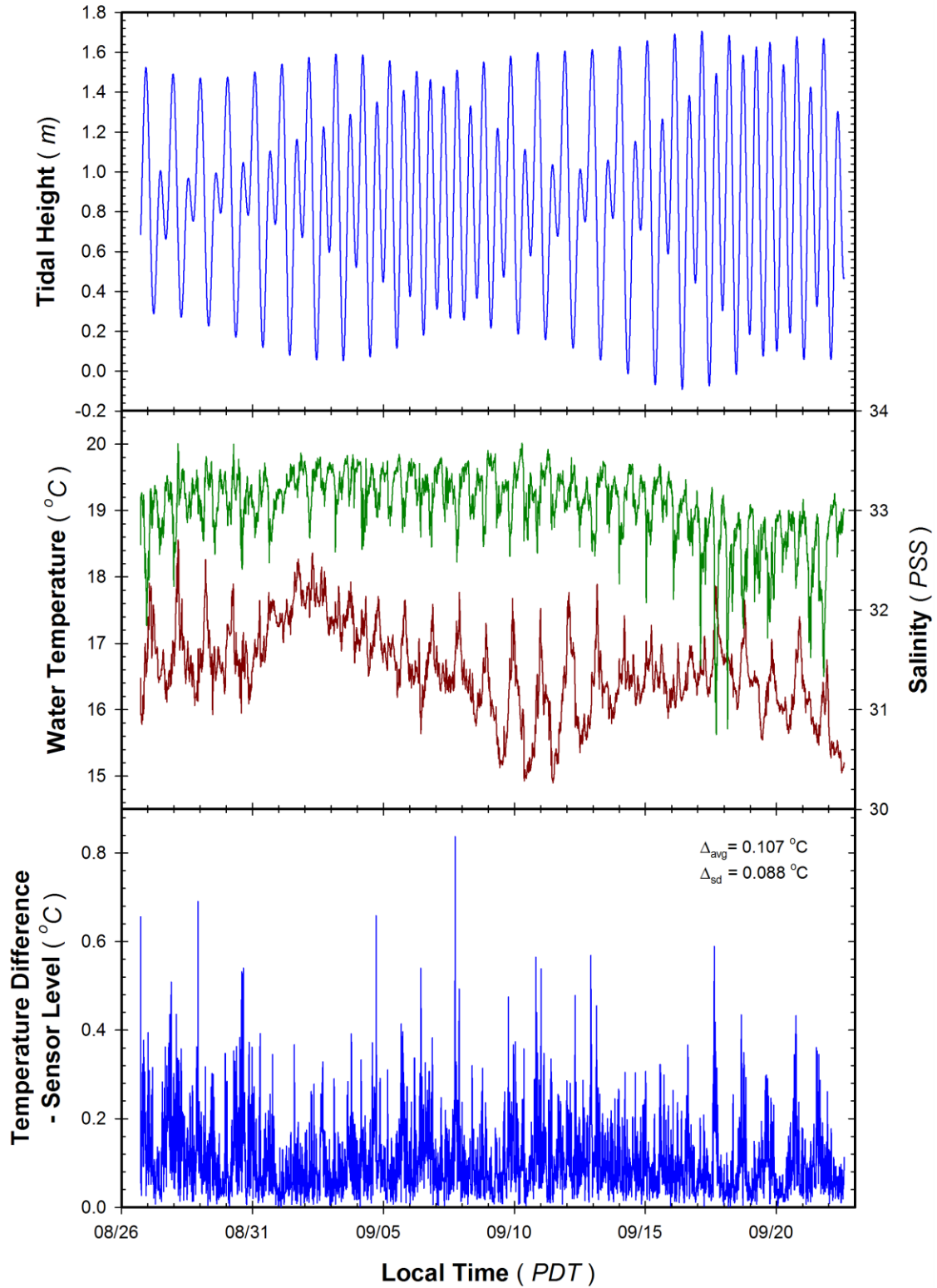


Figure 1. Environmental conditions encountered during deployment at the MLML Small Boats Dock. Test sensor array deployed at 1 m fixed depth, variation in local tidal heights indicate active water flow around instrument (*Top Panel*). Variation in salinity (green) and temperature (red) at depth of instrument sensor detected by an SBE 26 and two RBR Solo thermistors (*Middle Panel*). Temperature range determined from max-min temperatures detected by RBR and SBE thermistors spanning instrument sensor array (*Bottom Panel*).

Two separate sondes were used at this deployment, one containing antifouling copper mesh within the sensor guard and one with just a plain guard. The Manta2 with antifouling operated for 23 consecutive days before experiencing a power failure, corresponding to 2128 observations and a data return rate of 83% for the total deployment. Time series results of the Manta2 with antifouling and corresponding reference pH results are given in figure 2. Ambient pH measured by the Manta2 ranged from 7.95 to 8.63, compared to the range measured from reference samples of 7.933 to 8.077. The bottom panel presents the time series of the difference between instrument and reference pH measurements for each matched pair (N=64 observations). The average and standard deviation of the measurement difference (Manta2 – Reference) over the total deployment was 0.258 ± 0.181 with a total range of -0.014 to 0.551. As noted in the text on pH scales, a difference of approximately 0.13 pH units could be expected because of the difference in calibrating with an NBS buffer versus our reference measurement based on the total pH scale.

A cross-plot of the matched observations for the Manta2 with the Copper Screen is given in figure 3. The linear regression of instrument versus reference pH values was statistically significant but with a very low goodness of fit and high slope ($p=0.001$, $r^2 = 0.16$, slope = 2.89). In general the instrument measured variation did not meaningfully correspond to the variation in ambient pH as measured from the reference samples.

The Manta2 with no antifouling operated continuously for the entire 28 days of the deployment, resulting in 2579 observations at 15 minute intervals. Time series results of the Manta2 without copper mesh and corresponding reference pH results are given in figure 4. Ambient pH measured by this Manta2 sonde ranged from 8.28 to 8.60, compared to the range captured by the reference measurements of 7.933 to 8.077. The bottom panel presents the time series of the difference between instrument and reference pH measurements for each matched pair (N=84 observations). The average and standard deviation of the measurement difference (Manta2 – Reference) over the total deployment was 0.512 ± 0.059 with a total range of 0.333 to 0.584. As noted in text on pH scales, a difference of approximately 0.13 pH units could be expected because of the difference in calibrating with an NBS buffer versus our reference measurement based on the total pH scale. The range in pH values measured by the Manta2 was nearly three times greater than that measured by the reference samples.

A cross-plot of the matched observations is given in figure 5. The linear regression between the instrument measurement and reference values was not statistically significant ($p=.024$, $r^2 = 0.06$) and again the instrument measured variation in pH did not closely match the variation in ambient pH measured from the reference samples.

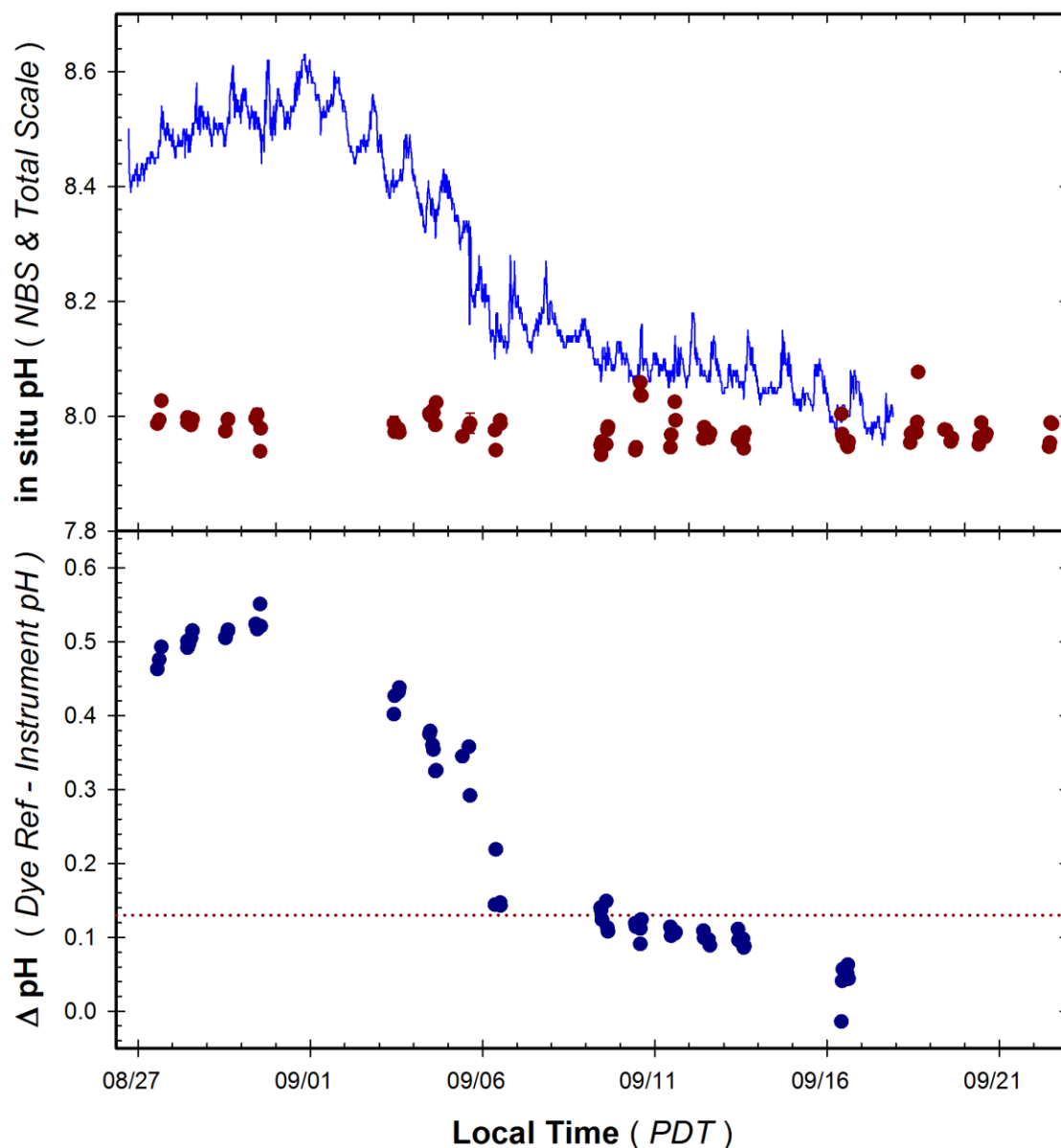


Figure 2. Time series of pH measured detected by Manta2 sonde deployed with an antifouling copper screen during the 1 month MLML field trial. *Top Panel:* Continuous pH_{NBS} recordings from instrument (blue line) and pH_{Total} of adjacent grab samples determined by mCP dye (red circles). *Bottom Panel:* Difference in measured pH relative to reference samples (Instrument – Dye Ref) observed during deployment. Dashed line represents offset (ca 0.13) predicted due to difference in pH_{NBS} and pH_{Total} scales used by the Manta2 sensor and the dye reference method

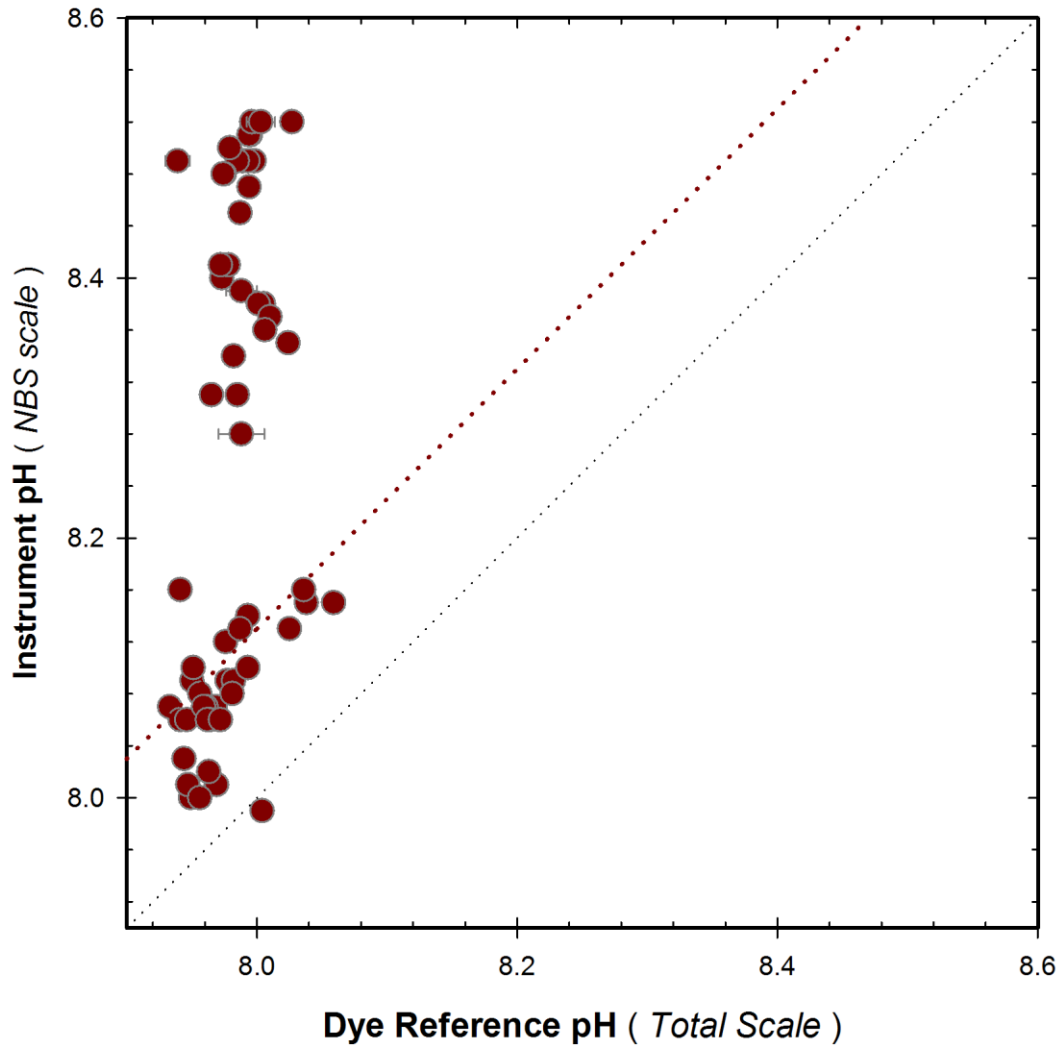


Figure 3. MLML field response plot of Manta2 pH sensor deployed with antifouling copper screen compared to reference pH_{Total} samples. Fine dotted line represents 1:1 correspondence, coarse dotted line represents correspondence predicted based on difference between pH_{NBS} and pH_{Total} in seawater (pH_{NBS} = pH_{Total} + 0.13).

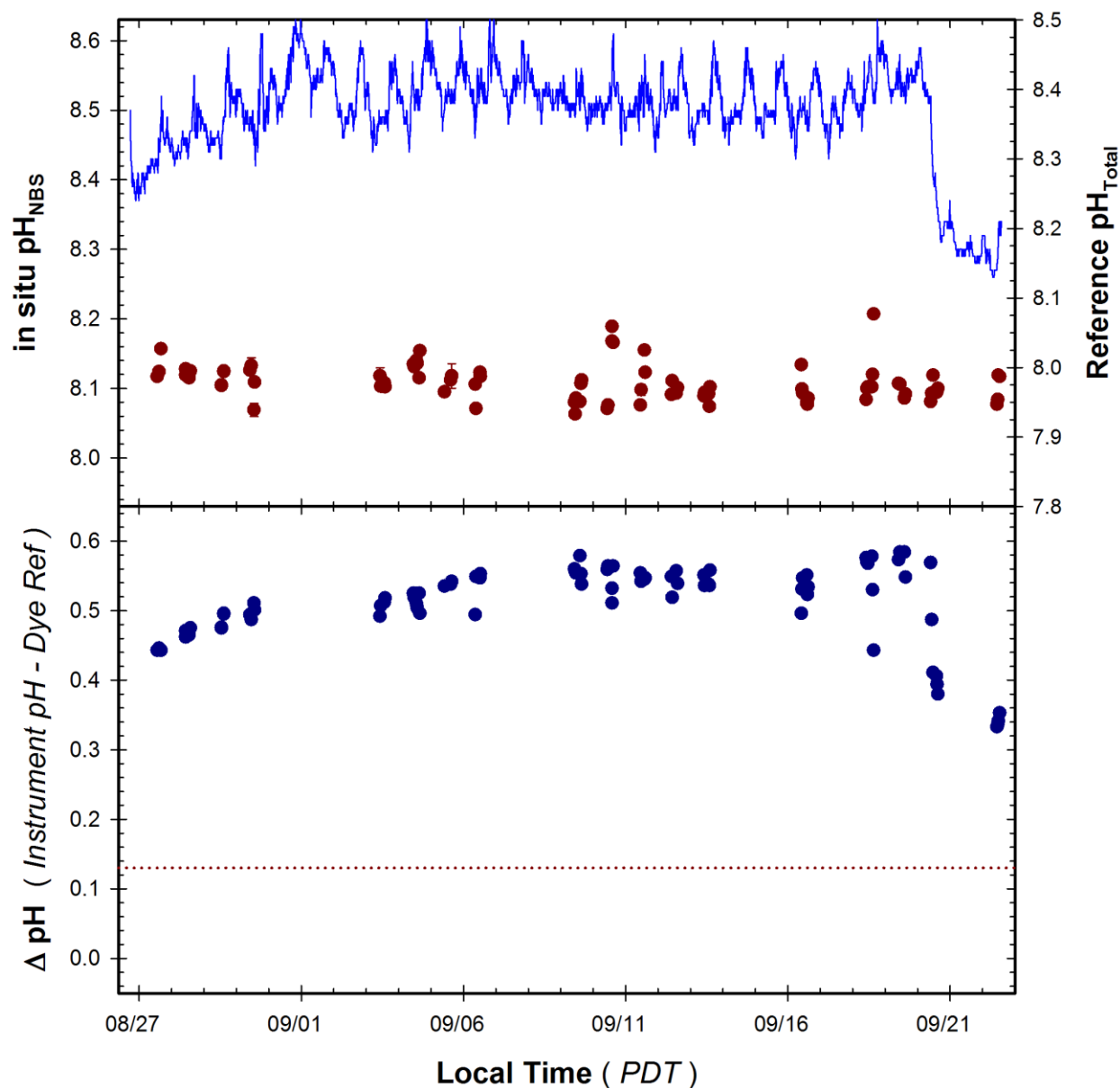


Figure 4. Time series of pH measured detected by Manta2 sonde deployed without antifouling copper screen during the 1 month MLML field trial. *Top Panel:* Continuous pH_{NBS} recordings from instrument (blue line) and pH_{Total} of adjacent grab samples determined by mCP dye (red circles). *Bottom Panel:* Difference in measured pH relative to reference samples (Instrument – Dye Ref) observed during deployment. Dashed line represents offset (ca 0.13) predicted due to difference in pH_{NBS} and pH_{Total} scales used by the Manta2 pH sensor and the dye reference method.

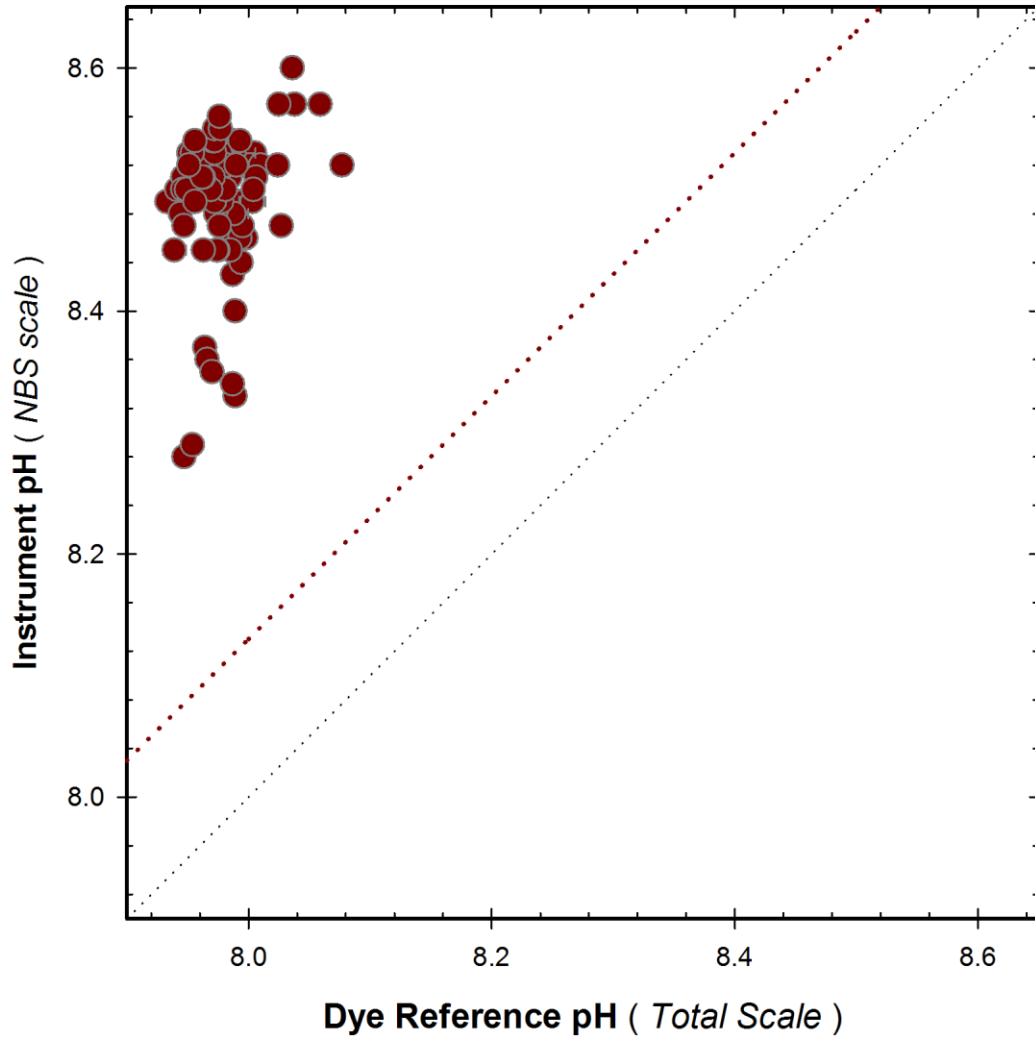


Figure 5. MLML field response plot of Manta2 pH sensor deployed without antifouling copper screen compared to reference pH_{Total} samples. Fine dotted line represents 1:1 correspondence, coarse dotted line represents correspondence predicted based on difference between pH_{NBS} and pH_{Total} in seawater (pH_{NBS} = pH_{Total} + 0.13).

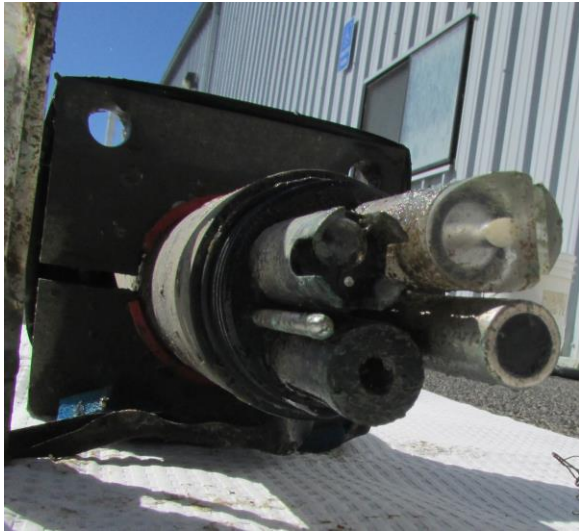


Photo 2. Eureka Manta2 prior to and following 1 month deployment for MLML field test.

Moored Deployment off Coconut Island in Kaneohe Bay, Hawaii

A three month-long moored field test was conducted in Kaneohe Bay from November 15, 2013 to February 10, 2014. The deployment site was located at 21.46° N, 157.80° W in the back-reef region of the barrier reef offshore of Coconut Island (HIMB) in a depth of 3 meters (Photo 3). Kaneohe Bay, located on the eastern side of Oahu, Hawaii is a complex estuarine system with a large barrier coral reef, numerous patch reefs, fringing reefs, and several riverine inputs. Tides in Kaneohe Bay are semi-diurnal with mean tidal amplitude of approximately 68 cm day. The pH instruments were mounted on the Crimp II Buoy (right), about one meter below the surface.



Photo 3. HIMB deployment site (left) and sampling at deployment buoy (right).

Time series results of ambient conditions for tidal height, temperature, and salinity are given in figure 6. Temperature ranged from 22.3 to 26.5 and salinity from 32.8 to 35.2 over the duration of the field test. The bottom panel displays the maximum difference recorded between reference thermistors mounted at the same depth across the mooring rack. The average temperature difference observed across the space of the mooring rack was 0.04 ± 0.10 °C, with a maximum of 0.65 °C. As noted above, it is not possible to quantify the exact difference between the averaged reference temperature and that measured by the test instrument, but differences resulting from this variation and variation in the chemistry of the water across will be minimized because the sampling bottle integrates across the mooring space.

The Manta2 sonde operated throughout the entire 88 days of the deployment measuring at 30 minute intervals, however, the sonde showed a significant shift in measurement response (pH dropped from a mean of 8.1 to 6.6) starting on January 30th, after 67 days of deployment. Upon retrieval, it was discovered that a tunicate had colonized directly on the pH sensor bulb causing it to fracture. These last eleven days of data were not used in the direct comparison with reference samples so the useable number of instrument observations was reduced to 3642, with 75 matched reference timepoints. The time series results of the Manta2 and corresponding reference pH results are given in figure 7. Ambient pH measured by the Manta2 ranged from 7.85 to 8.39, compared to a range captured by the reference measurements of 7.814 to 8.084. The bottom panel presents the time series of the difference between Manta2 and reference measurements for each matched pair (N=75). The average and standard deviation of the differences were 0.17 ± 0.02 with a total range in the differences of 0.12 to 0.20.

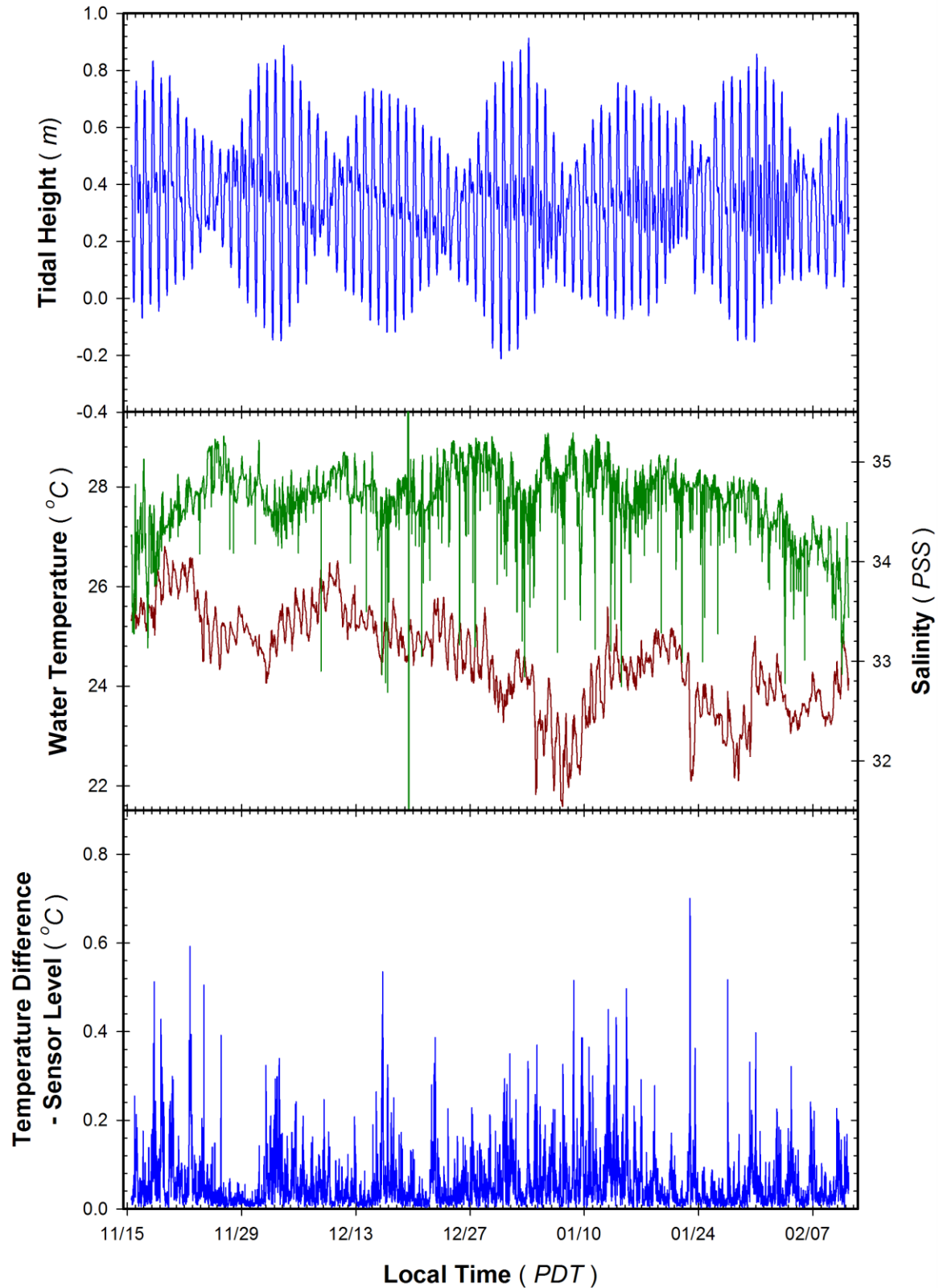


Figure 6. Environmental conditions encountered during the 3 month HIMB offshore deployment on the NOAA CRIMP Buoy. Test sensor array deployed at 1 m fixed depth, variation in local tidal heights indicate active water flow around instrument (*Top Panel*). Variation in salinity (green) and temperature (red) at depth of instrument sensor detected by an SBE 26 and two RBR Solo thermistors (*Middle Panel*). Temperature range determined from max-min temperatures detected by RBR and SBE thermistors spanning instrument sensor array (*Bottom Panel*).

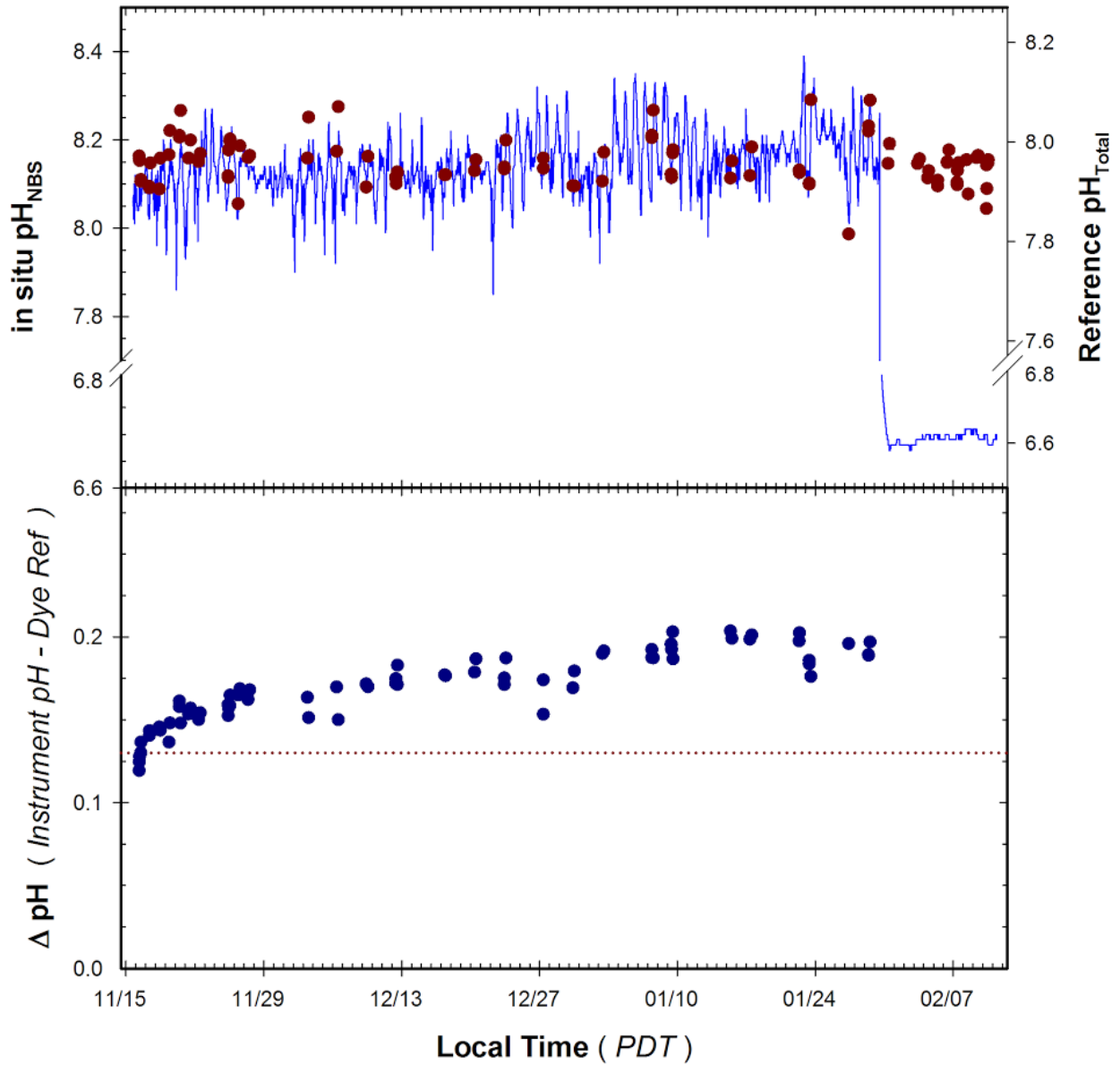


Figure 7. Time series of pH measured by Manta2 sonde deployed with antifouling copper screen during the 3 month HIMB trial. *Top Panel:* Continuous pH_{NBS} recordings from instrument (blue line) and pH_{Total} of adjacent grab samples determined by mCP dye (red circles). *Bottom Panel:* Difference in measured pH relative to reference samples (Instrument – Dye Ref) observed during deployment. Dashed line represents offset (ca 0.13) predicted due to difference in pH_{NBS} and pH_{Total} scales used by the Manta2 pH sensor and the dye reference method.

A cross-plot of the matched observations is given in figure 8. A linear regression of the data was highly significant ($p < 0.001$, $R^2 = 0.84$) with a slope of 0.98 and intercept of 0.33.

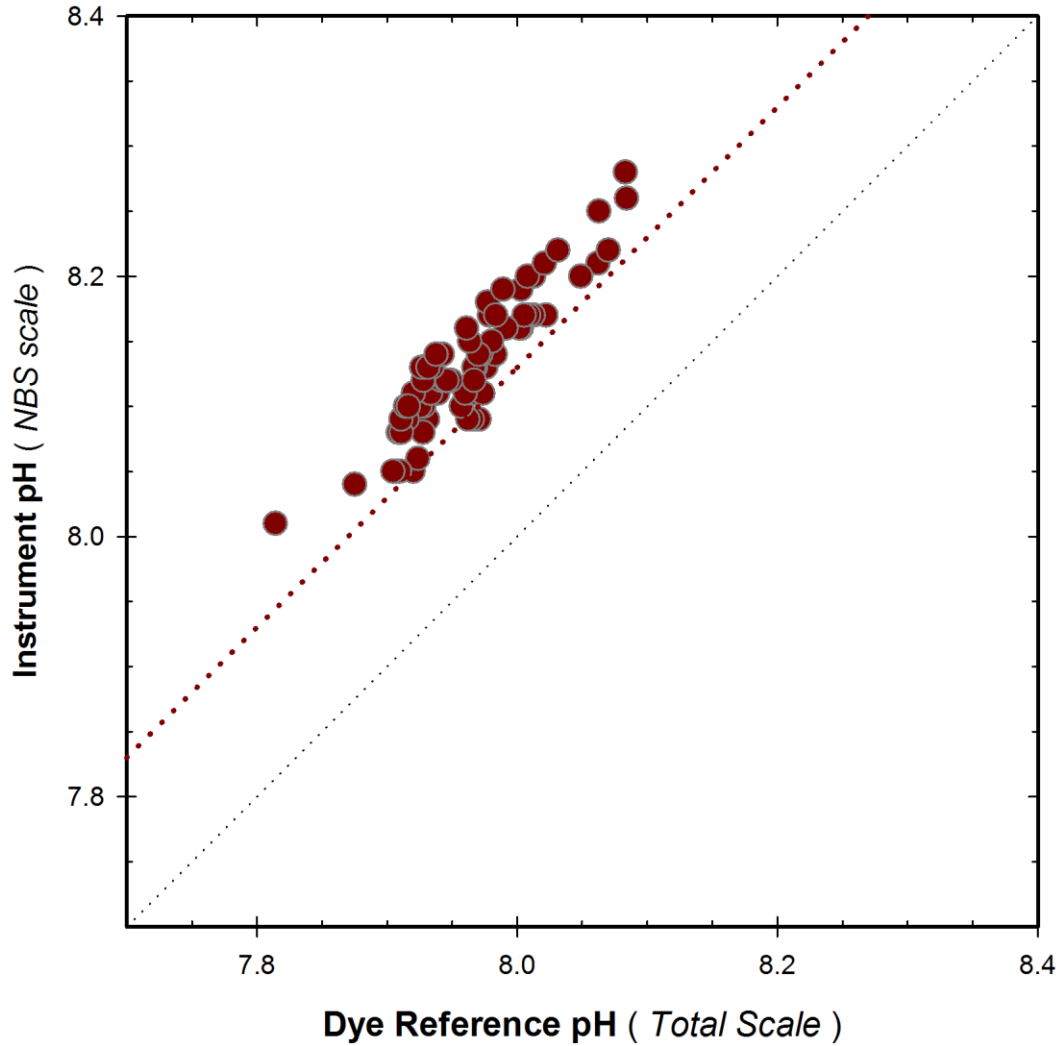


Figure 8. HIMB field response plot of Manta2 pH sensor deployed with antifouling copper screen compared to reference pH_{Total} samples. Fine dotted line represents 1:1 correspondence, coarse dotted line represents correspondence predicted based on difference between pH_{NBS} and pH_{Total} in seawater ($pH_{NBS} = pH_{Total} + 0.13$). Period of sensor failure (02/01 – 02/11) depicted in Fig 7 excluded for clarity.

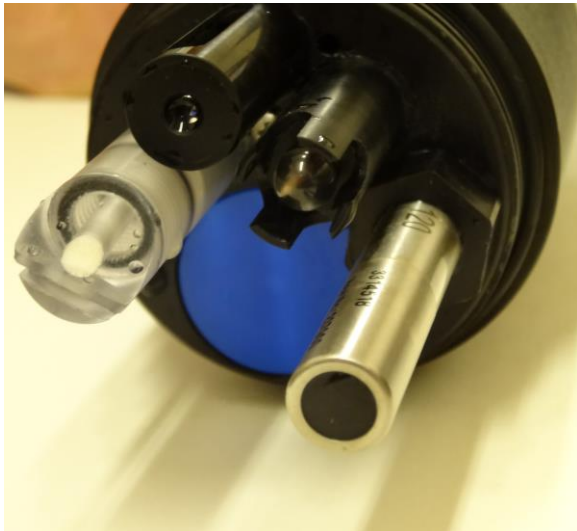


Photo 4. The Eureka Manta prior to and following the 3 month deployment in Kaneohe Bay. Note the broken bulb in the close up post deployment photo.

Moored Deployment at Chesapeake Biological Laboratory (CBL)

A month-long moored field test was conducted in Chesapeake Bay from March 3 to April 10, 2014. The deployment was located at 38.32°N, 76.45°W attached to the end of a pier at the mouth of the Patuxent River (Photo 5.) The average water depth of the test site was 2.2 m. The site was brackish with salinity ranging from 9.1 PSU to 14.3 PSU during the deployment and temperature ranging from 3.3 °C to 10.2 °C.



Photo 5. Aerial view of CBL deployment site (left) and duplicate sampling at mooring rack.

Time series results of ambient conditions for tidal height, temperature, and salinity are given in figure 9. Temperature ranged from 3.4 to 10.1 and salinity from 9.3 to 14.0 over the duration of the field test. The bottom panel displays the maximum difference recorded between all reference thermistors (RBR Solo and SBE26) mounted at the same depth and different locations across the mooring rack. The average temperature difference observed across the space of the mooring rack was 0.04 ± 0.06 °C, with a maximum of 1.38 °C. As noted above, it is not possible to quantify the exact difference between the averaged reference temperature and that measured by the test instrument, but differences resulting from this variation and variation in the chemistry of the water across will be minimized by the sampler integrating across the mooring space.

The Manta2 operated successfully over the entire deployment from March 12 through April 10 and generated 2756 observations based on its 15 minute sampling interval. Time series results of the Manta2 and corresponding reference pH results are given in figure 10. Ambient pH measured by the Manta ranged from 8.14 to 8.54, compared to the range captured by the reference measurements of 8.024 to 8.403. The bottom panel presents the time series of the difference between the Manta2 and reference pH for each matched pair (n=107 observations). The average and standard deviation of the measurement difference was 0.12 ± 0.02 , with the total range of differences from 0.09 to 0.16.

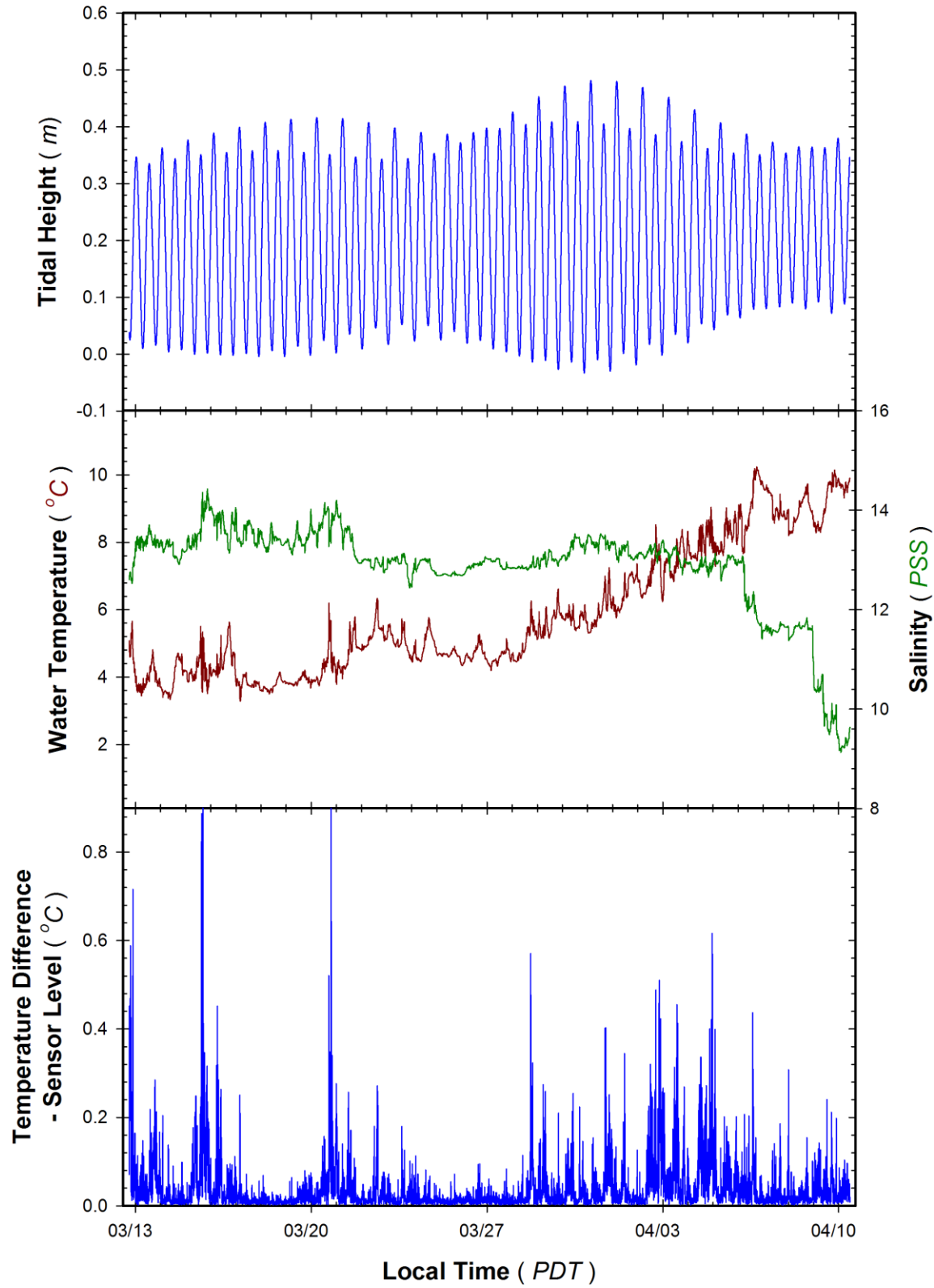


Figure 9. Environmental conditions encountered during the 1 month CBL floating dock deployment. Test sensor array deployed at 1 m fixed depth, variation in local tidal heights indicate active water flow around instrument (*Top Panel*). Variation in salinity (green) and temperature (red) at depth of instrument sensor detected by an SBE 26 and two RBR Solo thermistors (*Middle Panel*). Temperature range determined from max-min temperatures detected by RBR and SBE thermistors spanning instrument sensor array (*Bottom Panel*).

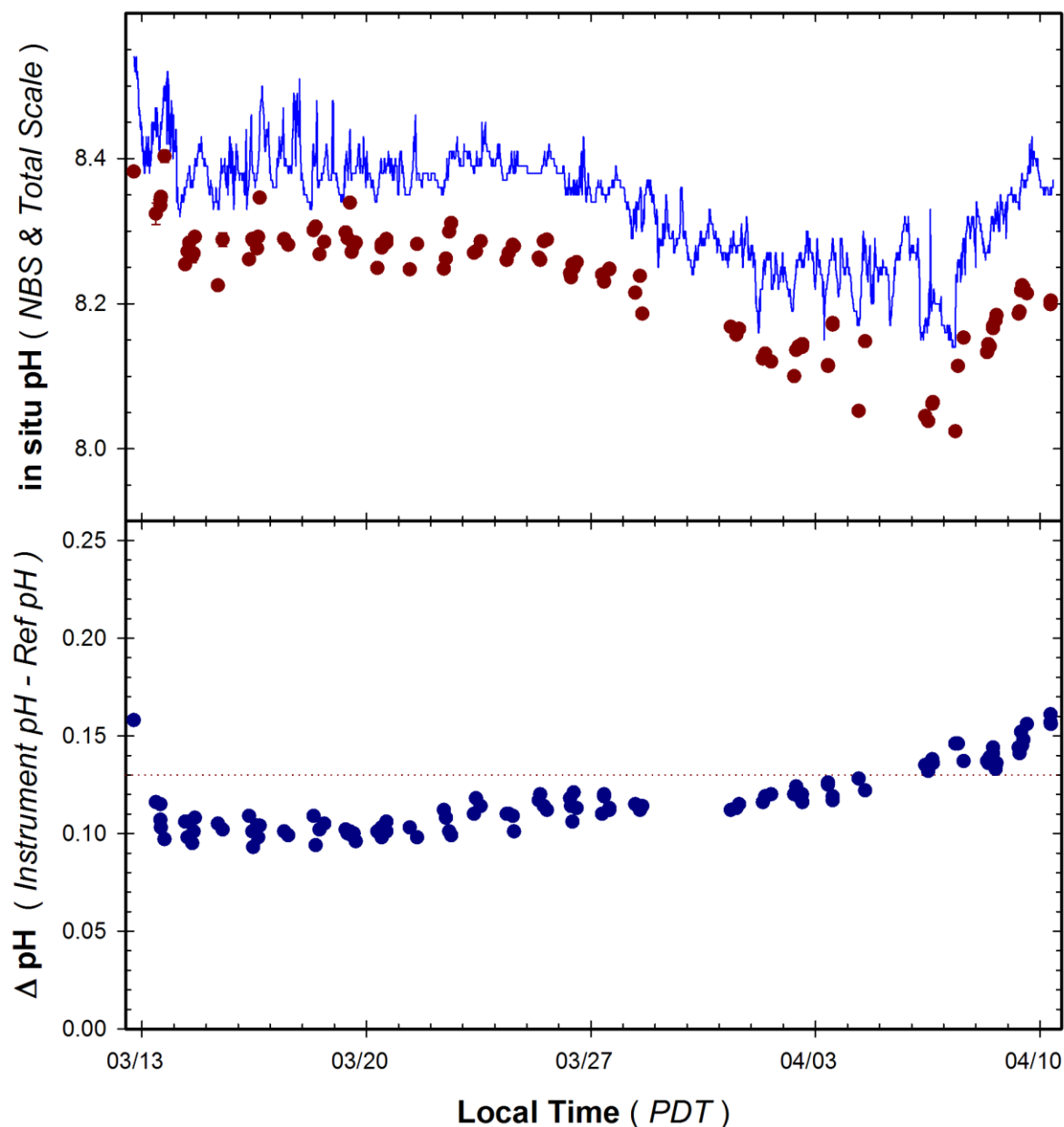


Figure 10. Time series of pH measured by Manta2 sonde deployed with antifouling copper screen during the 1 month CBL trial. *Top Panel:* Continuous pH_{NBS} recordings from instrument (blue line) and pH_{Total} of adjacent grab samples determined by phenol red dye (red circles). *Bottom Panel:* Difference in measured pH relative to reference samples (Instrument – Dye Ref) observed during deployment. Dashed line represent offset (ca 0.13) predicted due to difference in pH_{NBS} and pH_{Total} scales used by the Manta2 sensor and the dye reference method.

A cross-plot of the matched observations is given in figure 11. A linear regression of the data was highly significant ($p < 0.001$, $R^2 = 0.96$) with a slope of 0.88 and intercept of 1.14.

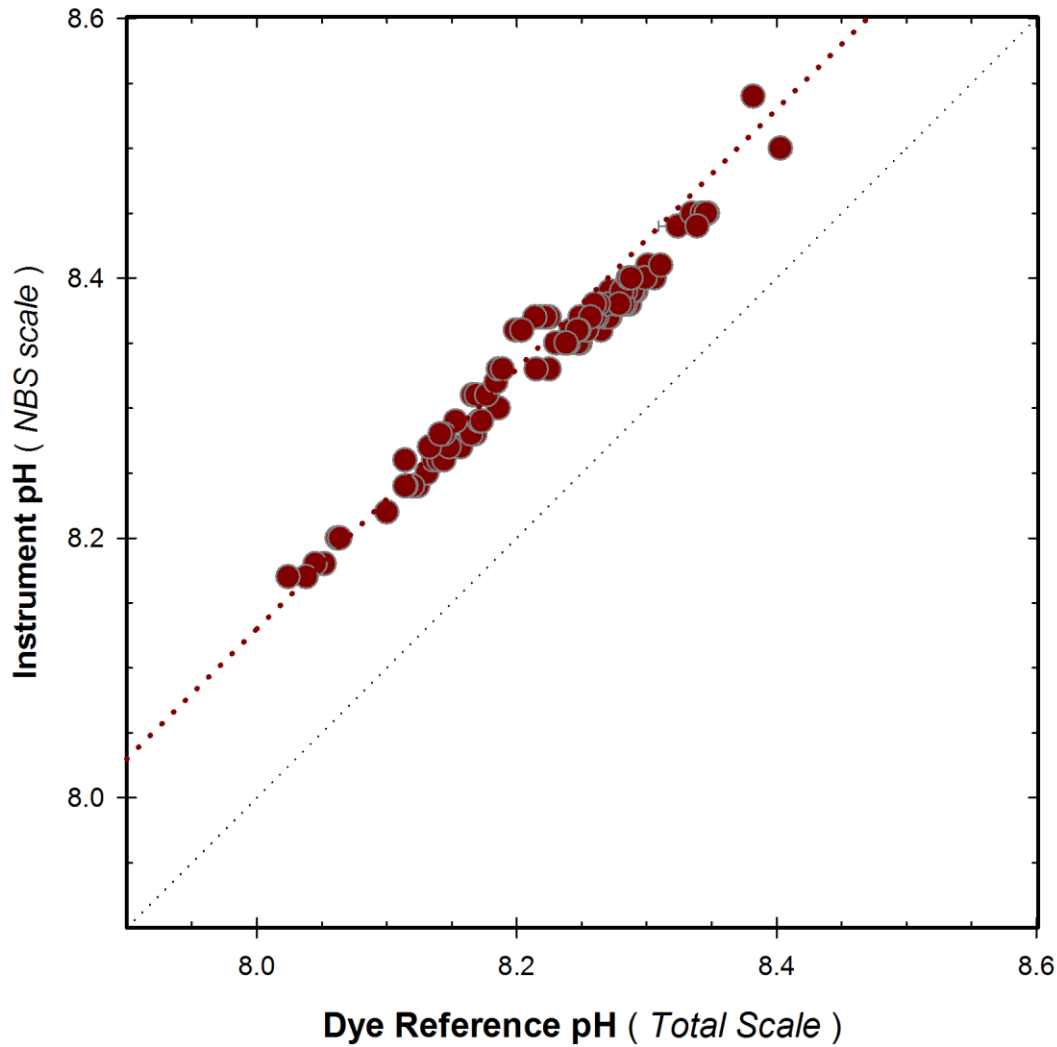


Figure 11. CBL field response plot of Manta2 pH sensor deployed with antifouling copper screen compared to reference pH_{Total} samples. Fine dotted line represents 1:1 correspondence, coarse dotted line represents correspondence predicted based on difference between pH_{NBS} and pH_{Total} in seawater ($\text{pH}_{\text{NBS}} = \text{pH}_{\text{Total}} + 0.13$).



Photo 6. The Eureka Manta prior to and following the one month deployment for the CBLfield tests.

Great Lakes Field Test Site

A month-long moored field test was conducted in Lake Michigan from June 21 to July 19, 2014. The Great Lakes deployment site was located at 43.23°N, 86.34°W on a fixed pier at the Lake Michigan Field Station of the NOAA Great Lakes Environmental Research Laboratory, in Muskegon, Michigan. The site is located at the outfall of Lake Muskegon into Lake Michigan with a depth of approximately 3 meters. The temperature range during deployment was 11.6°C through 24.9°C and salinity averaged 0.031.

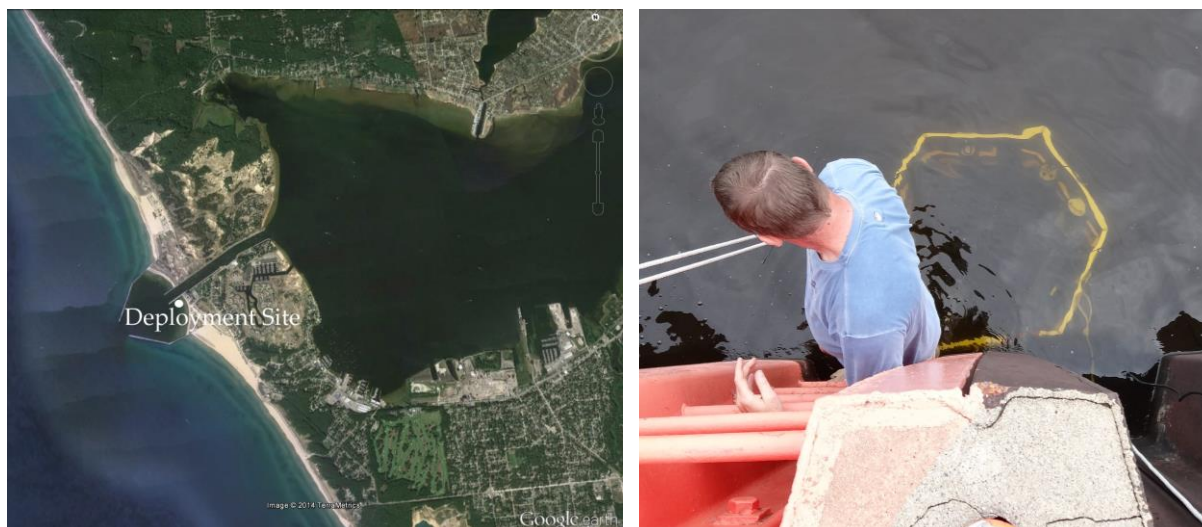


Photo 7. Aerial view of Lake Michigan and Muskegon Lake (left) and rack deployment at GL Site

Time series results of ambient conditions for water depth, temperature, and salinity are given in figure 12. Temperature ranged from 11.6 to 24.9 and salinity from 0.024 to .035 over the duration of the field test. While there are no tidal cycles at this site there was significant exchange between the open lake and the protected basin within the break wall as evidenced by changes in salinity and water depth. The bottom panel displays the maximum difference recorded between all reference thermistors mounted at the same depth across the mooring rack. The average temperature difference observed across the space of the mooring rack was 0.08 ± 0.18 °C, with a maximum of 2.53 °C. As noted above, it is not possible to quantify the exact difference between the averaged reference temperature and that measured by the test instrument, but differences resulting from this variation and variation in the chemistry of the water across will be minimized because the sampling bottle integrates across the mooring space.

The Manta2 operated throughout the entire deployment and generated 2680 observations based on its 15 minute sampling interval. The ambient pH measured by the Manta2 ranged from 7.86 to 8.51 compared to the range captured by the reference measurements of 8.013 to 8.526. (Fig.13). The bottom panel presents the time series of the difference between the Manta2 and reference pH for each matched pair (n=98 observations). The average and standard deviation of the difference for the entire deployment was -0.07 ± 0.04 with a total range of -0.20 to -0.004.

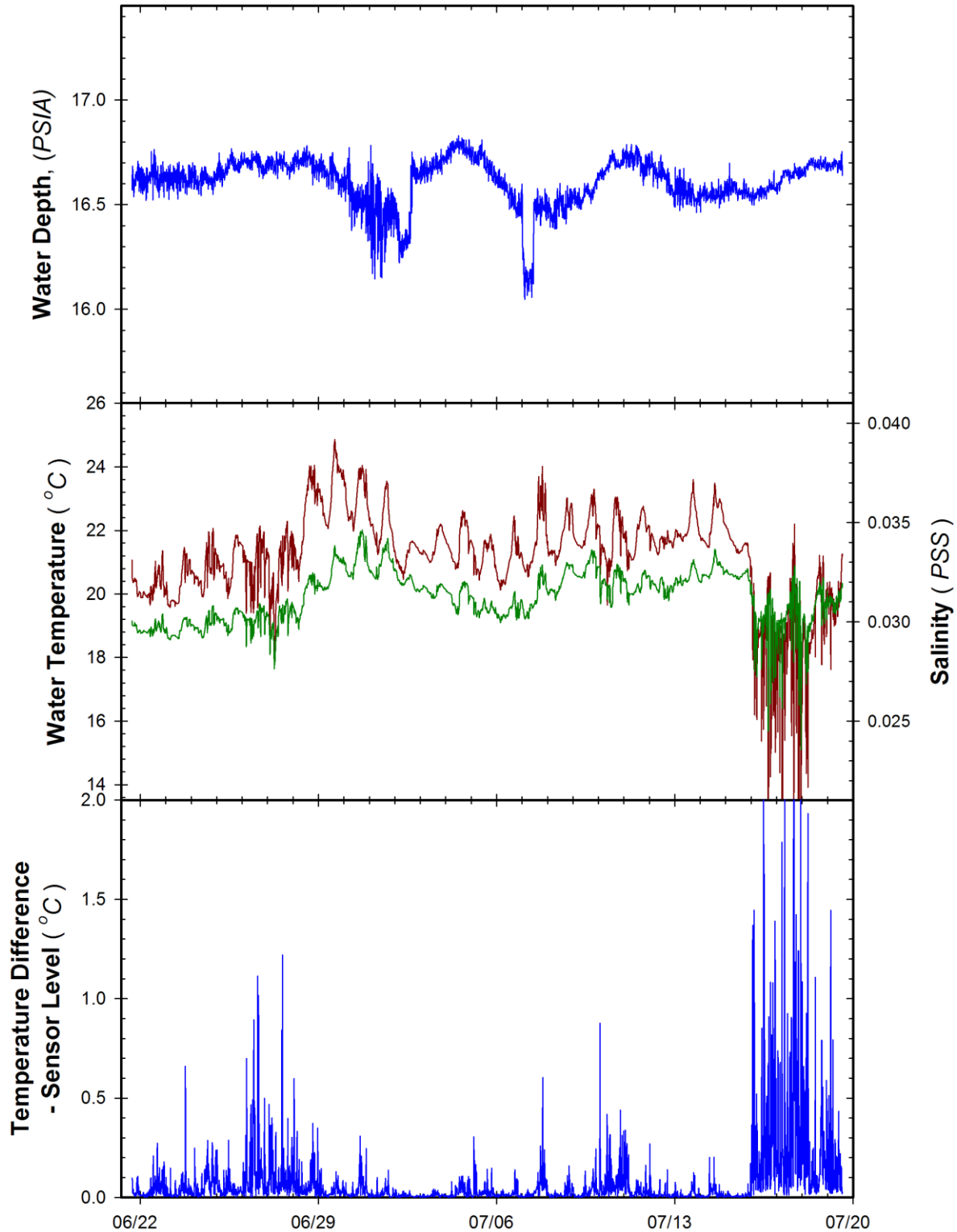


Figure 12. Environmental conditions encountered during the 1 month freshwater deployment adjacent to the NOAA GLERL Pier. Test sensor array deployed on bottom, variation in local water depth measured as PSIA by an SBE 26PLUS Wave gauge indicating active water flow around instrument (*Top Panel*). Variation in salinity (green) and temperature (red) at depth of instrument sensor detected by SBE 26 and two RBR Solo thermistors (*Middle Panel*). Temperature range determined from max-min temperatures detected by RBR and SBE thermistors spanning instrument sensor array (*Bottom Panel*). Storm event contributed to higher variation in environmental conditions near end of deployment period.

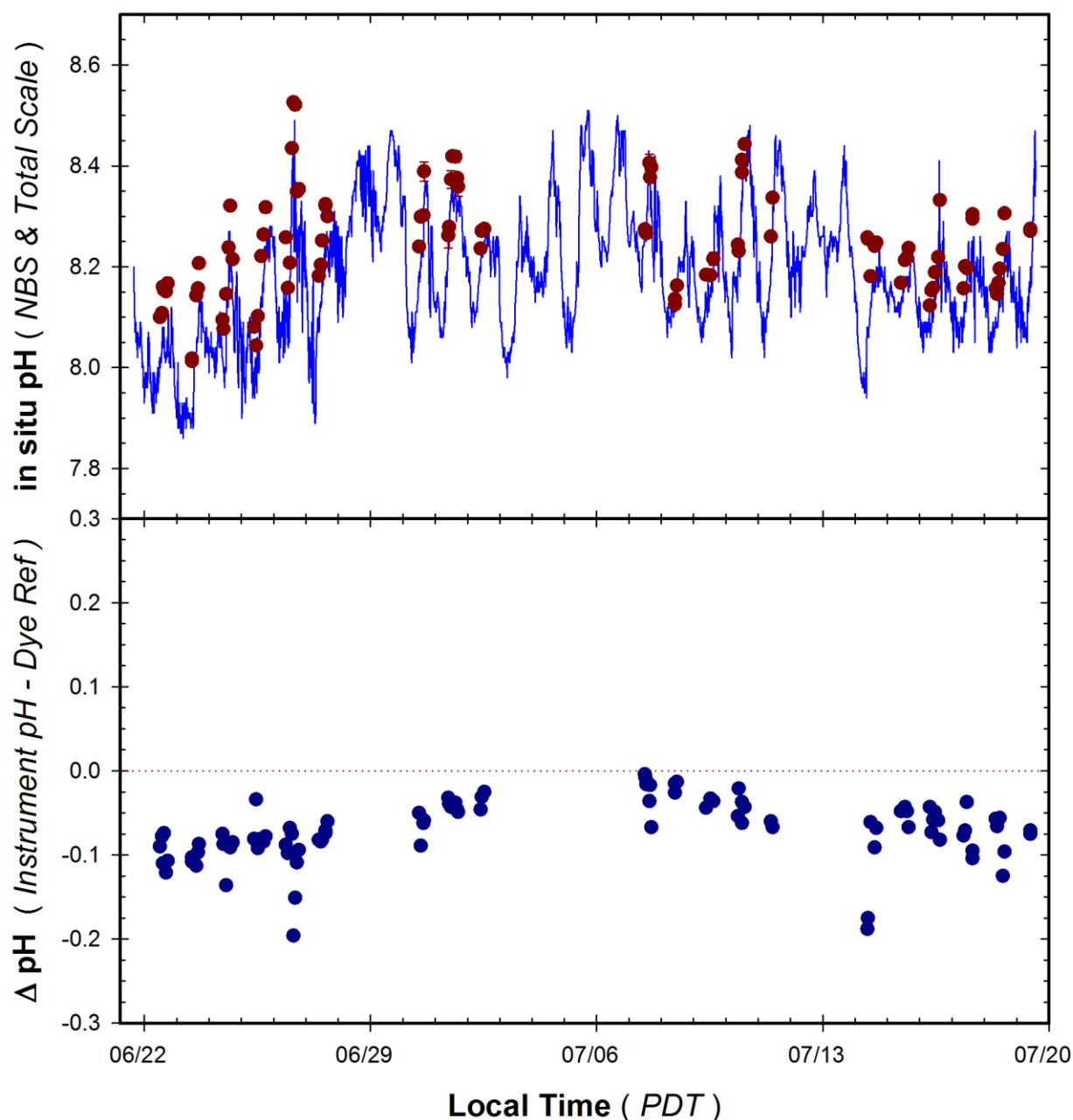


Figure 13. Time series of pH measured by Manta2 pH sensor deployed with antifouling copper screen during the 1 month GLERL trial. *Top Panel:* Continuous pH_{NBS} recordings from instrument (blue line) and pH_{Total} of adjacent grab samples determined by phenol red dye (red circles). *Bottom Panel:* Difference in measured pH relative to reference samples (Instrument – Dye Ref) observed during deployment. Dashed line represent offset (ca 0.13) predicted due to difference in pH_{NBS} and pH_{Total} scales used by the Manta2 sensor and the dye reference method.

A cross-plot of the matched observations is given in figure 14. The linear regression was highly significant ($p < 0.001$, $R^2 = 0.89$) with a slope of 0.99 and intercept of -0.03.

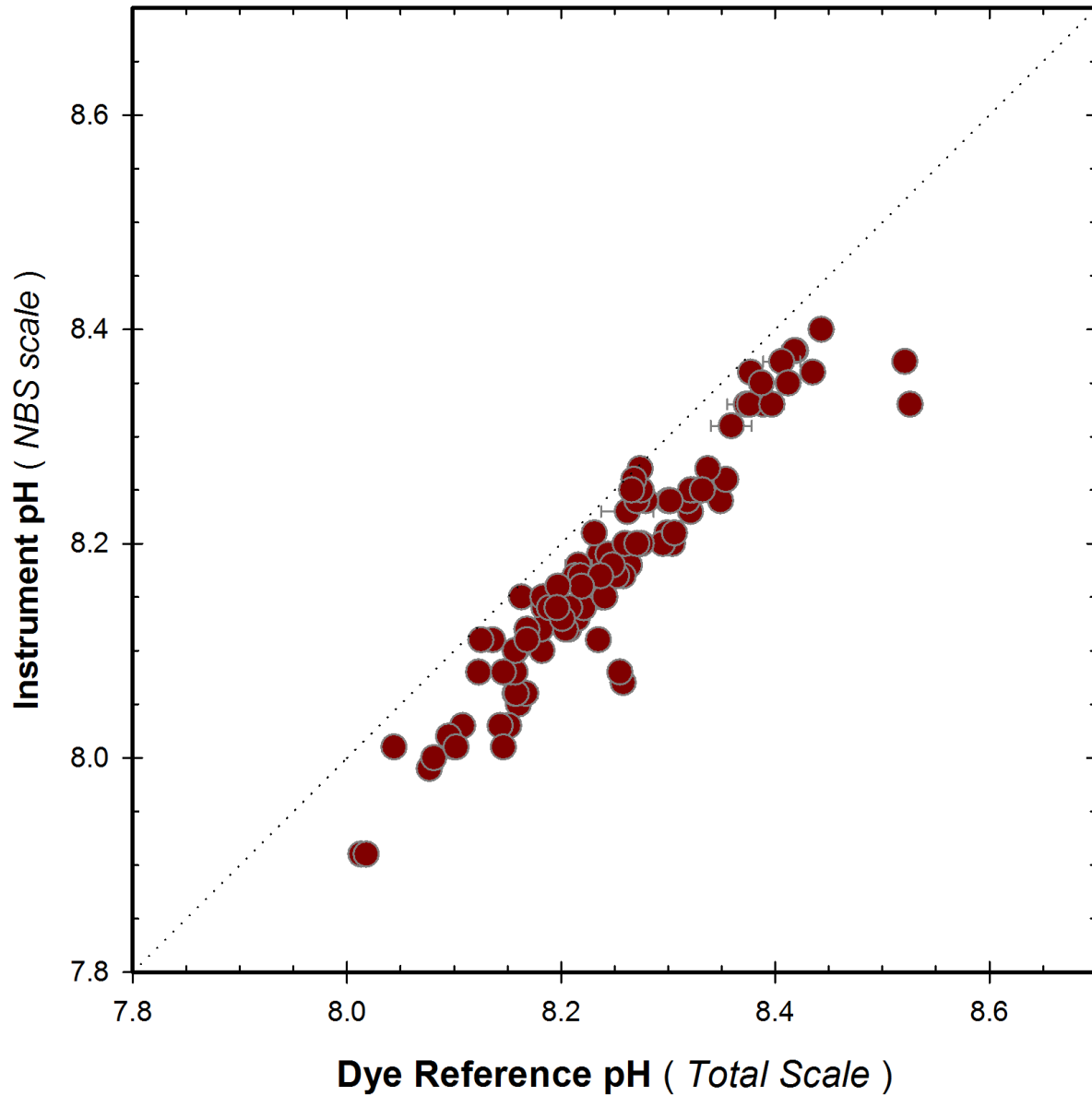


Figure 14. GLERL field response plot of Manta2 pH sensor deployed with antifouling copper screen compared to reference pH_{Total} samples. Fine dotted line represents 1:1 correspondence, expected agreement between pH_{NBS} and pH_{Total} in freshwater.



Photo 8. The Eureka Manta prior to and following the one month deployment in Lake Michigan.

A summary plot of the Manta2 versus dye reference pH for all sites is shown in figure 15. The Manta2 responded consistently with good agreement to reference pH measurements at three of the four test sites. No explanation or observed failure was discovered to explain the performance at the Moss Landing test site. Results for the brackish and saline test sites show the expected offset due to calibration and scale differences. The low ionic freshwater test and phenol red based dye measurement shows a closer agreement to a direct 1:1 relationship.

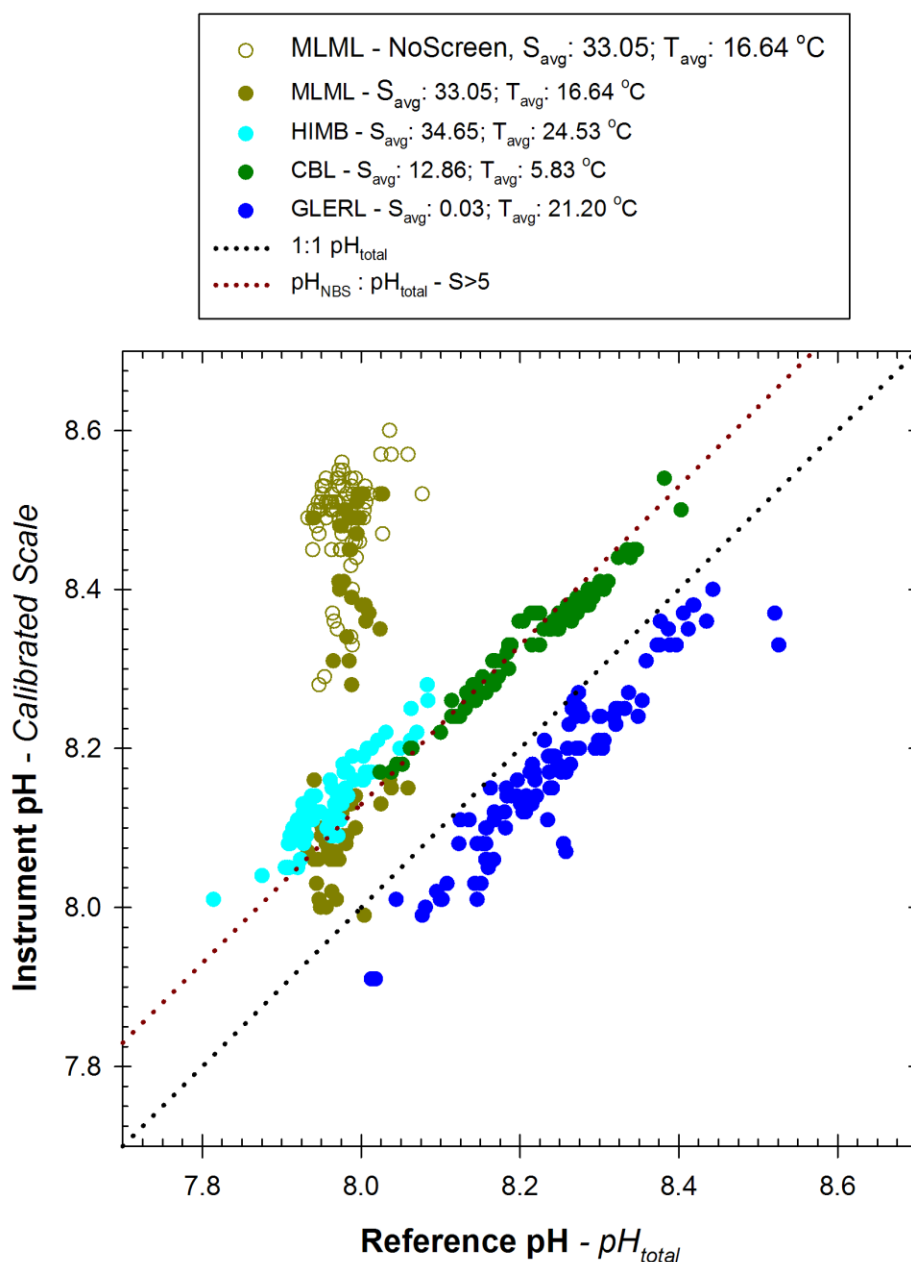


Figure 15. Global response plot for Manta2 pH sensors observed during ACT field trials. Black dotted line represents 1:1 correspondence, red dotted line represents correspondence predicted based on difference between pH_{NBS} and pH_{Total} in seawater ($pH_{NBS} = pH_{Total} + 0.13$).

QUALITY ASSURANCE AND QUALITY CONTROL

A variety of procedures were established for this verification to ensure the highest quality of data possible. Foremost was the development of detailed, scientifically reviewed testing protocols and state of the art analytical methodologies. Each reference sample generated was analyzed in triplicate and any result with a standard deviation of greater than 0.01 pH units was initially flagged. If there was an obvious reason to exclude any of the triplicates due to handling errors, that triplicate was eliminated and a new value computed from the remaining two replicates. Reference values with a standard deviation of greater than 0.02 were discarded. External standards were run weekly to verify consistency in performance of the spectrophotometric dye analysis.

Quality Control Samples

Each site also conducted weekly field duplicates and one field trip blank, with the exception that no field trip blank was conducted in Hawaii and only one field duplicate was conducted during the Moss Landing moored deployment. Results of these QA/QC samples are presented below in Tables 3 – 6. The global average of the standard deviation among field duplicates for all field test sites was 0.005 ± 0.008 (n=21), with two values exceeding our expected quality threshold of better than 0.01 pH units. These results attest to the representativeness of our sampling to water mass being analyzed by the test instruments and to consistent sample handling. The three field trip blanks that were performed had a standard deviation among the paired samples of 0.003, 0.002, and 0.002. The absolute difference between the paired samples was 0.004, 0.003, and 0.003. The field trip blanks attest to proper cleaning and handling of sampling equipment and sample processing.

Table 3. Results of Field Duplicates (FD) and Field Trip Blank (FTB) for the Moss Landing Harbor, CA mooring test. The field trip blank was determined with a Seawater CRM from the Dickson Lab at Scripps Institute of Oceanography.

Date/Time	Rep	Temp	Salinity	pH	Std Dev	Mean	Std Dev	ABS Diff
9-10-13 14:00	FD1	15.0	33.0	8.059	.003	8.049	.015	.021
	FD2			8.038	.001			
9-5-13 11:20	FTB1	22.0	35.0	7.974	.001	7.972	.003	.004
	FTB2			7.970	.001			

Table 4. Results of Field Duplicates (FD) for the Kaneohe Bay, HI mooring test. No field trip blank was determined at this test site.

Date/Time	Rep	Temp	Salinity	pH	Std Dev	Mean	Std Dev	ABS Diff
11-17-13 10:00	FD1	25.2	33.64	7.907	.001	7.908	.002	.003
	FD2			7.910	.001			
11-20-14 11:00	FD1	26.1	34.36	8.012	.001	8.011	.002	.003
	FD2			8.009	.002			
11-27-14 10:00	FD1	24.7	35.03	7.968	.001	7.968	0	0
	FD2			7.968	.001			
12-9-13 10:00	FD1	25.7	34.72	7.909	.001	7.909	0	0
	FD2			7.909	.001			
12-23-14 10:30	FD1	25.0	24.7	7.949	.007	7.947	.003	.004
	FD2			7.945	.001			
1-7-14 10:30	FD1	22.6	34.89	8.013	.001	8.010	.003	.005
	FD2			8.008	.001			
1-23-14 10:00	FD1	22.5	34.76	7.914	.001	7.915	.002	.002
	FD2			7.916	0			
2-6-14 10:00	FD1	23.4	33.99	7.958	.002	7.959	.001	.002
	FD2			7.959	0			
2-9-14 9:30	FD1	23.4	33.43	7.968	.001	7.968	0	0
	FD2			7.968	.001			

Table 5. Results of Field Duplicates (FD) and Field Trip Blank (FTB) for the Chesapeake Bay, MD mooring test. The field trip blank was determined with a mixture of Seawater CRMs from the Dickson Lab at Scripps Institution of Oceanography.

Date/Time	Rep	Temp	Salinity	pH	Std Dev	Mean	Std Dev	ABS Diff
3-20-14 11:30	FD1	4.4	13.87	8.111	.003	8.113	.034	.0048
	FD2			8.115	.004			
3-20-14 15:00	FD1	4.6	13.98	8.124	.004	8.121	.004	.005
	FD2			8.119	.003			
3-27-14 10:00	FD1	4.3	13.02	8.065	.004	8.065	.000	.000
	FD2			8.065	.001			
3-27-14 14:00	FD1	4.7	12.87	8.090	.002	8.090	.001	.0013
	FD2			8.089	.002			
4-3-14 9:30	FD1	6.9	13.19	7.992	.001	7.991	.001	.0015
	FD2			7.991	.009			
4-3-14 13:00	FD1	7.5	13.27	8.054	.001	8.055	.001	.001
	FD2			8.056	.007			
4-10-14 8:00	FTB1	9.5	9.28	8.122	.003	8.120	.002	.003
	FTB2			8.118	.003			

Table 6. Results of Field Duplicates (FD) and Field Trip Blank (FTB) for the Great Lakes, MI mooring test. The field trip blank was determined with a mixture of commercial NBS buffers.

Date/Time	Rep	Temp	Salinity	pH	Std Dev	Mean	Std Dev	ABS Diff
6-23-14 11:15	FD1	19.7	0.029	8.013	.003	8.016	.004	.005
	FD2			8.018	.003			
6-27-14 14:30	FD1	21.0	0.030	8.324	.004	8.322	.002	.003
	FD2			8.321	.008			
7-8-14 10:15	FD1	21.1	0.031	8.136	.001	8.131	.008	.011
	FD2			8.125	.004			
7-10-14 12:00	FD1	21.0	0.031	8.387	.010	8.400	.018	.025
	FD2			8.412	.006			
7-17-14 15:15	FD1	20.4	0.032	8.304	.008	8.300	.006	.009
	FD2			8.295	.007			
7-19-14 10:15	FD1	20.4	0.032	8.275	.002	8.273	.003	.004
	FD2			8.271	.008			
7-14-14 9:15	FTB1	22.0	0.032	8.258	.005	8.257	.002	.003
	FTB2			8.255	.003			

External Quality Assurance and Quality Control Assessments

An effective assessment program is an integral part of ACT's quality system for performance evaluations of marine sensor technologies. The ACT Quality Assurance (QA) Manager independently conducted three technical systems audits (TSA), an audit of data quality (ADQ), and a Data Quality Assessment (DQA) for the pH verification. The audits were conducted in accordance with the procedures described in EPA's *Guidance on Technical Audits and Related Assessments for Environmental Data Operations (EPA QA/G-7)*.

Technical System Audit - A Technical Systems Audit (TSA) is a thorough, systematic, on-site qualitative audit of ACT's sampling and measurement processes and procedures associated with a specific technology verification. The objective of a TSA is to assess and document the conformance of on-site testing procedures with the requirements of the Test Protocols, and associated SOPs. The ACT QA Manager conducted the following three TSAs over the course of the verification:

- The laboratory tests at the Hawaii Institute of Marine Biology (HIMB), Coconut Island, Kaneohe Bay, Oahu, HI during December 11-15, 2012.
- The field tests at HIMB during November 19-21, 2013.
- The field tests at the Chesapeake Biological Laboratory, Solomons, MD during March 11-14, 2014

All of the audits included a review of staff, test procedures (sample collection, sample analysis, data processing, etc.), facilities, and documentation to assure compliance with the ACT Test Protocols and Standard Operating Procedures (SOPs) for the pH sensor verification.

The TSA of the laboratory tests at HIMB determined that there were several deviations in the conduct of the tests from the Test Protocols. Minor deviations, such as schedule changes and the collection of triplicate versus duplicate samples, were documented in laboratory records. Other deviations from the Test Protocols included deviations in the sequence of test conditions and dropping one type of burst sampling. These deviations also were documented and explanations of the rationale for the changes were provided. It was determined that none of the deviations in the Test Protocols had any effect on data quality for the verification and no corrective action was required.

The TSA's findings for the two field tests were positive in nature and indicated that these components of the pH sensor verification were being implemented in a manner consistent with the Test Protocols and SOPs. All phases of the implementation of the test reviewed during the TSA were acceptable and performed in a manner consistent with ACT data quality goals. The overall quality assurance objectives of the test were met.

ACT personnel are well-qualified to implement the verification, and demonstrated expertise in pertinent procedures. Communication and coordination among all personnel was frequent and effective. ACT's internal record keeping and document control is well organized. The ACT staff understands the need for QC, as shown in the conscientious development and implementation of a variety of QC procedures.

All samples and instrument measurements were collected, analyzed and cataloged as described in the Test Protocols and SOPs. For samples transported and analyzed by other affiliated laboratories, chain-of-custody (COC) protocols were practiced, specifying time, date, sample location, unique sample number, requested analyses, sampler name, time and date of transaction between field and laboratory staff, and name of receiving party at the laboratory. Examination of maintenance and calibration logs provided evidence of recent and suitable calibration of sampling and analytical equipment.

Audit of Data Quality - The ACT QA Manager conducted an Audit of Data Quality (ADQ) on verified data to document the capability of ACT's data management system (hardcopy and electronic) to collect, analyze, interpret, and report data as specified in the Test Protocols and SOPs. The ADQ implemented an examination of data after they had been collected and verified by ACT personnel. A representative set of no less than 10% of the data on core parameters was traced in detail from raw data from field and laboratory logs and instrument readouts, data transcription or transference, data manipulation, data reduction to summary data, data calculations, and final reported data.

The ADQ determined that ACT's sample measurement system performed in accordance with the performance goals specified in the Test Protocols and that the data were accumulated, transferred, reduced, calculated, summarized, and reported correctly. There is sufficient documentation of all procedures used in the data collection and analysis to verify that the data have been collected in accordance with ACT quality objectives.

Data Quality Assessment - A Data Quality Assessment (DQA) is a scientific and statistical evaluation of validated data to determine if the data are of the right type, quality, and quantity to support their intended use. Data review is conducted to ensure that only sound data that are of known and documented quality and meet ACT technology verification quality objectives are used in making decisions about technology performance. Two steps were followed to determine the validity of the analytical data – data verification and validation.

Data verification evaluates the completeness, correctness, and consistency of the test data sets against the requirements specified in the Test Protocols. For the pH sensor verification, this step verified that:

- The raw data records were complete, understandable, well-labeled, and traceable;
- All data identified in the Test Protocols were collected;
- Instrument calibration and QC criteria were achieved;
- Data calculations were accurate.

Data validation established:

- Required sampling methods were used;
- Sampling procedures and field measurements met performance criteria;
- Required analytical methods were used;
- QC measures were obtained and criteria were achieved.

The DQA determined that ACT's data quality objectives were achieved. The DQA supports conclusions that:

- The sampling design performed very well and was very robust with respect to changing conditions.
- Sufficient samples were taken to enable the reviewer to see an effect if it were present.
- Data on the performance of the pH sensors are unambiguous, and the vendors and buyers can make informed choices about the performance of a sensor with a high level of certainty.

REFERENCES

American National Standards Institute and American Society for Quality Control (ANSI/ASQC). 2004. Quality Systems for Environmental Data Collection and Environmental Technology Programs: Requirements with Guidance for Use. E4-2004. Milwaukee, WI.

Culberson, C.H. 1981. Direct potentiometry, in: Whitfield, M., and Jagner, D. (Eds.), Marine Electrochemistry: a Practical Introduction. Wiley, pp. 187–261.

Dickson, A.G., Sabine, C.L. and Christian, J.R. (Eds.) 2007. Guide to Best Practices for Ocean CO₂ Measurements. PICES Special Publication 3:191.

Easley R.A. and Byrne, R.H. 2012. Spectrophotometric calibration of pH electrodes in seawater using purified m-cresol purple. Environmental Science and Technology 46:5018-5024.

Liu, X., Patsavas, M., and Byrne, R.H. 2011. Purification and characterization of meta-cresol purple for spectrophotometric seawater pH measurements. *Environmental Science and Technology* 45:4862-4868.

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

National Environmental Laboratory Accreditation Conference (NELAC). 2003. NELAC Standard. EPA/600/R-04/003. Washington, DC.

Pierrot, D., Lewis, E., and Wallace, D.W.R. 2006. MS Excel Program Developed for CO₂ System Calculations. ORLN/CDIAC-105, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy. Oak Ridge, TN.

United States Department of Energy (U.S. DOE). 1994. Handbook of Methods for the Analysis of the Various Parameters of the Carbon Dioxide System in Sea Water; version 2, A.G. Dickson and C. Goyet, (eds.) ORNL/CDIAC-74.

United States Environmental Protection Agency (U.S. EPA). 2003. Guidance on Assessing Quality Systems. EPA/240/R-03/002: EPA QA/G-3. Washington, DC.

United States Environmental Protection Agency (U.S. EPA). 2006a. Guidance on Technical Audits and Related Assessments for Environmental Data Operations. EPA QA/G-7: EPA/600/R-99/080. Washington, D.C.

United States Environmental Protection Agency (U.S. EPA). 2006b. Data Quality Assessment: A Reviewer's Guide. EPA/240/B-06/002: EPAQA/G-9R. Washington, DC.

United States Environmental Protection Agency (U.S. EPA). 2006c. Data Quality Assessment: Statistical Tools for Practitioners EPA/240/B-06/003: EPA QA/G-9S. Washington, DC

United States Environmental Protection Agency (U.S. EPA). 2008. Guidance on Environmental Data Verification and Data Validation EPA/240/R-02/004: EPA QA/G-8. Washington, DC

Yao, W., and Byrne, R. H. 2001. Spectrophotometric determination of freshwater pH using bromocresol purple and phenol red. *Environmental Science and Technology* 35:1197-1201.

ACKNOWLEDGEMENTS:

We wish to acknowledge the support of all those who helped plan and conduct the verification test, analyze the data, and prepare this report. In particular we would like to thank our Technical Advisory Committee, Dr. Robert Byrne, University of South Florida, Dr. Andrew Dickson (Chair), Scripps Institution of Oceanography, University of California, San Diego, Dr. Burke Hales, Oregon State University, Mr. Scott McLean, Oceans Networks Canada, University of Victoria, Dr. Kenneth Pratt, National Institute of Standards and Technology, Dr. Chris Sabine, NOAA Pacific Marine Environmental Laboratory, Dr. Simone Alin, NOAA Pacific Marine Environmental Laboratory, Dr. Rik Wanninkhof, NOAA Atlantic Oceanographic and Meteorological Laboratory for their advice and direct participation in various aspects of this evaluation. E. Buckley also provided critical input on all aspects of this work and served as the independent Quality Assurance Manager. This work has been coordinated with, and funded by, the National Oceanic and Atmospheric Administration, Integrated Ocean Observing System program.

February 26, 2015

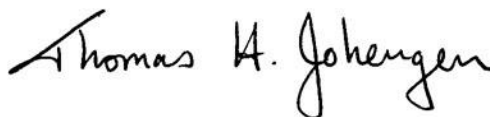
Date



Approved By: Dr. Mario Tamburri
ACT Executive Director

February 26, 2015

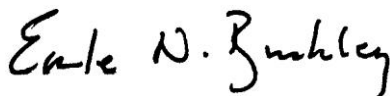
Date



Approved By: Dr. Tom Johengen
ACT Chief Scientist

February 26, 2015

Date



Approved By: Dr. Earle Buckley
Quality Assurance Supervisor

APPENDIX I

Company Response Letter to be Submitted