PERFORMANCE VERIFICATION STATEMENT
For Seabird Scientific HydroCycle Phosphate Analyzer

TECHNOLOGY TYPE: Nutrient Sensors

APPLICATION: In situ estimates of PO_4 for coastal moored deployments

PARAMETERS EVALUATED: Accuracy, precision, range response and reliability

TYPE OF EVALUATION: Laboratory and Field Performance Verification

DATE OF EVALUATION: Testing conducted from January 2015 to November 2016

EVALUATION PERSONNEL: T. Johengen, H. Purcell, M. Tamburri, D. Loewensteiner, G.J. Smith, D. Schar, Margaret McManus, GordonWalker.

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. Tom Johengen
ACT Chief Scientist
CIGLER- University of Michigan
4840 S. State Street
Ann Arbor, MI 48108 USA
Email: Johengen@umich.edu

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The Alliance for Coastal Technology (ACT) conducted a sensor verification study of in situ nutrient analyzers during 2016 to characterize performance measures of accuracy, precision and reliability. The verification including a week of laboratory testing along with three moored field deployments in freshwater, estuarine, and oceanic coastal environments. Laboratory tests of accuracy, precision, and range were conducted at the University of Maryland's Chesapeake Biological Laboratory (CBL) in Solomons, MD. A series of five tests were conducted to evaluate performance under controlled challenge conditions including: concentration range, temperature, salinity, turbidity, and dissolved organic carbon. All laboratory tests were conducted in 250 L polypropylene tanks using RO water as the initial matrix, within a temperature controlled room. Instruments sampled from a common, well-mixed, test tank maintained at a documented level of known challenge condition. Instruments were set-up by the manufacturer daily prior to the start of each individual laboratory test, exposed to each test condition for a period of three hours, and programmed to sample at a minimum frequency of 30 minutes. Reference samples were collected every 30 minutes for five timepoints during corresponding instrument sampling times for each test.

For the laboratory concentration range challenge the absolute difference between the HydroCycle-PO4 and reference measurement across all timepoints for trials C0 – C5 ranged from -0.0163 to 0.0145 mgP/L, with a mean of -0.0039 ±0.0090 mgP/L. A linear regression of the measurement difference versus concentration was not significant ($p=0.36$; $r^2=0.03$), however measurement offsets were increasingly negative between C0 and C4, at which point there was a large reversal and the offset became positive. An assessment of precision was performed by computing the standard deviations and coefficients of variation of the five replicate measurements for C1 – C5 concentration trials. The standard deviation of the mean ranged from 0.0005 to 0.0020 mgP/L across the five trials, and the coefficient of variation ranged from 0.14 to 5.78 percent. For the laboratory temperature challenge at 5 ^{o}C, the absolute difference between instrument and reference measurement across all timepoints for trials C2 – C4 ranged from -0.0140 to -0.0046 mgP/L, with a mean of -0.0087 ±0.0032 mgP/L. Measurement differences were significantly different for the C2 and C4 trials at 5 versus 20 ^{o}C, but the temperature effect was in the opposite direction between those trials. The measurement difference at C3 was nearly identical for the two temperatures. Therefore no clear pattern of a temperature effect on accuracy was observed. For the laboratory salinity challenge performed at the C3 concentration level, the absolute difference between instrument and reference measurement across all timepoints for the three added salinity levels ranged from -0.0124 to 0.0086 mgP/L, with a mean of -0.0018 ±0.0053 mgP/L. There was a statistically significant response to increased salinity with the offsets increasing in a positive direction as salinity increased. A linear regression of the measurement differences versus salinity ($p<0.0001$; $r^2=0.70$) had a slope of 0.0005 and intercept of- 0.013. The average offset at a salinity of 30 was 0.015 mgP/L higher than at zero salinity. For the laboratory turbidity challenge, performed at the C3 concentration level, the absolute difference between instrument and reference measurement across all timepoints for the two added turbidity levels ranged from -0.0226 to -0.0113 mgP/L, with a mean of -0.0170 ±0.0047 mgP/L. A linear regression of the measurement differences versus turbidity was significant ($p=0.0001$; $r^2=0.68$), with a slope of -0.0005 and intercept of -0.010, however the trend line was clearly forced by the large decrease at 100 NTU where the offset was 0.009 mgP/L more negative (under-predicted) than results observed at the 0 and 10 NTU trials. For the laboratory DOC challenge, performed at the C3 concentration level, the absolute difference between instrument and reference measurement across all timepoints for the two added DOC levels ranged from -0.0112 to -0.0074 mgP/L, with a mean of -0.0096 ±0.0014

mgP/L. A linear regression of the measurement differences versus DOC concentration was highly significant ($p=0.0005$; $r^2=0.62$) with a slope of 0.0002 and intercept of -0.013. The measurement difference generally became less negative as DOC concentration increased.

A 32 day field deployment occurred from May 26 through June 27 in the Maumee River, at the facilities of the Bowling Green, Ohio Water Treatment Plant. The HydroCycle-PO4 operated continuously during the entire 32 day deployment sampling at hourly intervals, but of the 763 possible data points 666 were flagged by the instrument as bad. Accepting the flagged data, and omitting 131 outliers (<-0.004 or >0.250), the HydroCycle-PO4 generated 632 observations out of a possible 763 for a data completion result of 82.8%. The average and standard deviation of the measurement difference between instrument and reference PO_4 measurements for each matched pair (n=43 of a possible 51 observations) over the total deployment was -0.022 ± 0.029 mgP/L with a total range of -0.110 to 0.033 mgP/L. There was no significant trend in the measurement difference over time as estimated by linear regression ($p= 0.20$; $r^2=0.04$). A linear regression of instrument versus reference measurement was not significant ($p=0.93$; $r^2 = 0.0002$) and the HydroCycle-PO4 generally under-predicted concentrations.

An 84 day moored field test was conducted in Chesapeake Bay from July 18 to October 10, 2016. The HydroCycle-PO4 had issues communicating with the datalogger, and lost data for the first two days of the deployment. The instrument then stopped reporting data on 9/19 due to the power cable getting frayed by rubbing against the floating dock. The HydroCycle reported 728 of a possible 730 measurements while operational. The average and standard deviation of the measurement difference between instrument and reference PO_4 measurements for each matched pair (n=71 of a possible 103 observations) over the total deployment was -0.005 ±0.005 mgP/L, with the total range of differences between -0.018 to 0.003 mgP/L. There no significant trend in measurement difference over time as estimated by linear regression ($p=0.89$; $r^2=0.0003$) over the deployment period. A linear regression of the data was significant ($p<0.0001$; $r^2 = 0.56$), but with a slope of only 0.459 and intercept of 0.003. The HydroCycle-PO4 did not accurately measure concentrations above 0.015 mgP/L.

A one month long moored field test was conducted in Kaneohe Bay from October 3, 2016 to November 2, 2016. The HydroCycle-PO4 operated successfully for the entire 29 days of the deployment sampling at hourly intervals, and returning 720 of a possible 720 measurements for a data completion result of 100%. However one value was omitted as an outlier due to its value being more than 3x higher than the maximum reference value. The average and standard deviation of the differences between instrument and reference readings over the entire deployment (n=70 out of a possible 71) was -0.0025 ± 0.0012 mgP/L, with a total range in the differences of -0.0048 to -0.0003 mgP/L. There was a small but statistically significant trend in the measurement difference over time ($p<0.0001$; $r^2 = 0.54$) during the deployment, with a slope of -0.0001 mgP/L/d. A linear regression of instrument versus reference measurements was not significant ($p=0.065$; $r^2 = 0.063$) and HydroCycle-PO4 did not accurately differentiate concentrations within this fairly narrow 0.004 mgP/L range.

BACKGROUND AND OBJECTIVES

The Alliance for Coastal Technologies (ACT) is a NOAA and EPA-funded partnership of research institutions, state and regional resource managers, and private sector companies that are interested in developing, improving, and applying sensor technologies for studying and monitoring coastal environments. ACT was established on the premise that instrument validation of existing and emerging technologies is essential to support both coastal science and resource management. The overall goals of ACT's verification program are to provide industry with an opportunity to have a third-party test their instruments in both controlled laboratory settings and in diverse field applications within a range of coastal environments, and to provide users of this technology with an independent and credible assessment of instrument performance.

ACT partnered with the multi-agency Challenging Nutrients Coalition on the Nutrient Sensor Challenge to help address the environmental and ecological problems associated with nutrient pollution. A critical step in this process is facilitating the development and adoption of the next-generation of *in-situ* nutrient sensors and analyzers. To that end, the ACT Technology Verification model was applied to the Nutrient Sensor Challenge to test instrument performance in laboratory and field tests against reference water samples analyzed using EPA-approved standard methods.

The report within contains the test results for the Seabird Scientific HydroCycle-PO4 phosphate Analyzer during the ACT Performance Verification. A synthesis of the testing protocols and reference sample analysis are provided below. A complete copy of the verification protocols is available on the ACT website at the following link: http://www.act-us.info/nutrients-challenge/Download/Nutrient_Challenge_Test%20Protocols_PV16_01.pdf

INSTRUMENT TECHNOLOGY TESTED

The HydroCycle-PO4 is a wet chemical instrument for continuous monitoring of phosphate in both fresh and coastal waters. ACT evaluated its predecessor, the Cycle-PO4 in 2007 (ACT TD08-03). Although designed for estuarine research, the Cycle-PO4 was adopted by many sectors in water quality and thus the HydroCycle-PO4 was engineered for environmental monitoring. A prototype at the time of the Nutrient Sensor Challenge, the HydroCycle-PO4 has since been optimized and replaced the Cycle-PO4 as Sea-Bird Scientifics phosphate instrument (http://wetlabs.com/hydrocycle).

Based on the chemical method of Murphy and Riley[1] the instrument makes use of selective reagents that form a colored complex for which the absorbance is proportional to "reactive" phosphate concentration. In general, this method captures the free ortho-phosphate, but the term "reactive" refers to the potential for reaction with a small amount of extremely labile phosphates (e.g., ATP, polyphosphates) and potential surface reactive particle-bound phosphate. Mathematical theory is from the Beer-Lambert law, which utilizes absorption of light by the colored complex to generate a relationship between signal and phosphate concentration as follows:

Equation (1): $$T = e^{-ab} \rightarrow \frac{-\ln T}{b} = a = a^{*}c$$

Where a = absorption coefficient (m^{-1}), a = absorptivity ($M^{-1}m^{-1}$), b = pathlength (m), T = transmittance (light through sample/100% transmission), c = concentration (M)*

The 100% transmission is determined after the 5-cm optical cell (880nm) is filled with ambient water sample and the pumps have been turned off. Afterwards, reagents are mixed with sample and a stopped-flow method is used to observe the reaction until it is complete in. A slope-threshold algorithm is used to select the detector counts to average for the sample transmission. A NIST traceable on-board standard, factory calibration, a 2.3 μgP/L detection limit, and an accurate calibration mean you can trust data, line-up grab sample records, and swap sensors with minimal offsets. Real-time quality control flags have been developed to speed up troubleshooting and provide confidence in data quality and uptime.

Autonomous operation, low power, and 5-month reagent stability enable extended deployments in remote locations. Over 1500 samples per service, up to 4 samples/hour, and both SDI-12 & RS-232 permit capturing higher frequency phosphate data than before. The HydroCycle-PO4 has increased reliability over its predecessor, the Cycle-PO4, and many other wet-chemical sensors, by maximizing data reliability and sensor uptime. Fluidics have been advanced to provide >90% uptime free of bubbles for data quality in high oxygen saturation environments and to ensure it stabilizes rapidly after deployment. Filter life has been extended to minimize clogging and enhance data quality during high sediment events.

PERFORMANCE EVALUTION TEST PLAN

These Test Protocols are based on consensus recommendations of the ACT Technical Advisory Committee, ACT staff, and participating Manufacturers. In summary, the test:

- utilized standard, approved laboratory analytical methods to provide best possible measure of the 'true' nutrient concentration from reference samples, which served as performance standards against which instrument estimations were compared internally by the individual developer;
- conducted all reference sample analysis at the state certified Nutrient Analytical Services Laboratory (NASL) of the Chesapeake Biological Laboratory (CBL), Solomons, MD to determine true nutrient concentrations using USEPA approved methodologies (see details below);
- included a laboratory evaluation of instrument performance;
- included three moored/dock-based field trials under a wide range of environmental conditions including freshwater, estuarine and marine ecosystems with varying nutrient concentrations and water quality characteristics (e.g. turbidity).

All ACT personnel involved in the Nutrient Sensor Verification were trained on standardized water sample collection, storage and shipping methods. ACT staff was available to assist in the physical deployment and recovery of all submitted test instruments and were responsible for the data management of test instrument results. Challenge participants were responsible for initial set-up and calibration of their instrument. If requested, ACT provided the chemicals and nutrient standards needed for instrument set-up and calibration. All laboratory nutrient analyses of the independent reference samples were conducted at the CBL NASL using standardized automated wet chemistry. All numerical data were recorded to three significant decimals where appropriate and nutrient concentrations reported in elemental mass units as mgN/L or mgP/L for nitrate+nitrite (NO_{23}), nitrate (NO_3^-) or phosphate (PO_4^{3-}), respectively.

Laboratory Tests

Laboratory tests of accuracy, precision, and range were conducted at the University of Maryland's Chesapeake Biological Laboratory (CBL) in Solomons, MD. A series of five tests were conducted to evaluate performance under controlled challenge conditions including: concentration range, temperature, salinity, turbidity, and dissolved organic carbon (details below). All Laboratory tests were conducted in polypropylene tank using RO water as the initial matrix, within a temperature controlled room. All instruments sampled from a common, well-mixed, test tank of approximately 250L volume, maintained at a documented level of known challenge condition. Instruments were set-up by the manufacturer daily prior to the start of each individual laboratory tests. Instruments were exposed to each test condition for a period of three hours and programmed to sample at a minimum frequency of 30 minutes. Reference samples were collected every 30 minutes for five timepoints during instrument sampling times for each test. Laboratory tests included the following 'controlled' challenge conditions:

Test 1: Accuracy and Precision over a broad concentration range

- Tested response across a broad range of concentrations representative of natural waters.
  - Concentration levels for NO_3 (mgN/L): 0.005, 0.1, 1.0, 5, 10, and 50
  - Concentration levels for PO_4 (mgP/L): 0.002, 0.01, 0.05, 0.1, 0.5, and 2.0
- The range test was split into two separate tests with concentrations for levels 1-4 conducted on day 1 and the last two concentrations tested on day 6 due to time constraints. Note that the starting level on day 6 was mistakenly set to 5 mgN/L and the 10 mgN/L level was not actually tested.
- Three hour sampling windows were provided at each of the six concentrations during which instruments measured concentrations at a minimum frequency of every 30 minutes.
- Discrete reference samples were collected every 30 minutes, corresponding to instrument sampling times, to generate five comparative measurements to assess accuracy and precision against reference values.
- RO water was used as the test matrix to which known amounts of nutrient salts (KNO_3 and K_2HPO_4) were added. Analysis of ambient blanks indicated a small amount of inorganic nutrients in the RO water.
- Tests were conducted at 20 ^oC in a temperature controlled room with samples drawn from a common well-mixed 250L test tank.

Test 2: Temperature Response

- Instrument response was tested for three concentrations, corresponding to levels C2, C3, and C4 from the range test, at temperatures of 5 ^oC versus the temperature of 20 ^oC on the first day.
- Temperature was regulated and maintained within a temperature controlled room and independently verified in the test tank with an YSI EXO2 reading at 15 min intervals.
- Instruments were equilibrated to the new 5 ^oC test temperature overnight.
- Instruments were exposed for three hours at each of the 3 concentrations with reference samples collected every 30 minutes following an initial 30 minute equilibration period to each condition.

Test 3: Salinity Response

- Accuracy and precision was tested over three additional salinities (10-20-30) at the C3 concentration level of the range test at 20^{o}C.
- Salinity levels were developed using Instant Ocean additions to the RO water matrix, which could have contributed trace amounts of nutrients, but would have measured in the final reference samples.
- Instruments were exposed for three hours at each of the 3 concentrations with reference samples collected every 30 minutes following an initial 30 minute equilibration period to each condition.

Test 4: Turbidity Response

- Accuracy and precision were tested over two elevated turbidity levels (approximately 10 and 100 NTU) at the C3 concentration level of the range test at 20 ^{o}C.
- Test tanks were continuously mixed with submersed pumps but there was some settling of the material as noted by continuous monitoring with the EXO2 sonde and analysis of discrete turbidity samples on the Hach 2100.
- Turbidity concentrations were established using Elliot Silt Loam reference material (cat # 1B102M) available from the International Humic Substances Society (http://www.humic-substances.org) added into RO water matrix.
- Instruments were exposed for three hours at each of the 3 concentrations with reference samples collected every 30 minutes following an initial 30 minute equilibration period to each condition.

Test 5: DOC Response

- Accuracy and precision were tested against two DOC levels (1 and 10 mg/L) at the C3 concentration level of the range test at 20 ^{o}C.
- DOC concentrations were established using the Upper Mississippi River Natural Organic Matter standard (cat# 1R110N) available from the International Humic Substances Society (http://www.humic-substances.org) added to RO water matrix.
- Instruments were exposed for three hours at each of the 3 concentrations with reference samples collected every 30 minutes following an initial 30 minute equilibration period to each condition.

Field Tests

In situ field performance evaluations of the test instruments were conducted under extended mooring deployments at three ACT Partner Institution sites covering freshwater, estuarine, and marine conditions. Site specific details for each test site were as follows:

Freshwater Deployment: The freshwater deployment occurred on the Maumee River in Waterville, OH for one month duration and provided a high nutrient, high turbidity test environment. The ACT Partner at the University of Michigan established a flow-through system on the Maumee River near Waterville Ohio (83.74 ^{o}N; 41.48 ^{o}W), located within the pump house of the City of Bowling Green Municipal Water Treatment Plant. Instruments were deployed in a 180 gallon flow-through tank with a water depth of approximately 0.8m and exchange time of

approximately 10 minutes. The Maumee River main stem flows 137 km before flowing into the Maumee Bay of Lake Erie at the city of Toledo, Ohio. The Maumee watershed is the largest watershed of any Great Lakes river with 8,316 square miles. The majority of the watershed is cultivated crop land, mostly corn and soybeans, though concentrated areas of pasture are located in the northwestern and southeastern areas of the watershed.

Estuarine deployment : The estuarine deployment occurred at the research pier of the Chesapeake Biological Laboratory in Solomons, MD for three month duration and provided for variable salinity and nutrient levels within a highly productive and biofouling environment. The ACT Partner at Chesapeake Biological Laboratory (CBL), University of Maryland Center for Environmental Science, has established a Technology Verification Field Test Site on a fixed pier (38.32 ^{o}N;76.45 ^{o}W), with an average depth of 2.1 m at the mouth of the Patuxent River, a tributary of the Chesapeake Bay. The deployment frame was arranged so that all of the sample inlets for the instruments remain at a fixed depth of 1 m below the water surface using a floating dock. The Chesapeake is a nutrient rich estuary with a watershed that encompasses portions of six states and the District of Columbia. Water temperatures at the test site ranged from 20 to 31°C and salinity ranged from 12.7 to 16.9 psu during the Verification.

Marine deployment: The marine deployment occurred in Kaneohe Bay at the Hawaii Institute of Marine Biology field lab for one month duration and provided a full salinity, low nutrient test condition. The ACT Partner at the Hawaii Institute of Marine Biology (HIMB) is part of the University of Hawaii with a field site established on the Kaneohe Bay Barrier Reef flat (21.43 ^{o}N;157.79 ^{o}W) in waters ~16 m deep. The deployment frame was arranged so that all of the sample inlets for the instruments remain at a fixed depth of 1 m below the water surface using a floating dock. Kaneohe Bay sits on the northeast, or windward, side of Oahu. Water temperatures at this site varied between 24.5 and 27.9°C and salinities were between 27.3 and 34.8 psu during the Verification.

Instrument Setup - Prior to deployment, all instruments were set up and calibrated as required at the field sites by a manufacturer representative, with assistance provided by ACT staff as necessary. The manufacturer supplied or specified to ACT all specific materials and hardware (chemicals, power cords, cables, weights, etc.) needed to deploy the test instrument according to requirements defined for each field site. ACT staff worked with the manufacturer to design an appropriate sensor deployment configuration at each site and arranged instruments in a manner so that a single representative field sample could be collected without the potential of interference between instruments. No servicing of the instruments was to occur during the test deployment period unless observed physical damage had occurred from natural events and a repair or replacement was deemed necessary. Instruments were set up as self-recording, either internally or to an external data logger, and programmed to record data based on a time interval that allowed instruments to function for the specified number of days for the respective deployment. Specific sampling intervals varied among test instruments, but with a stated goal of 15 minute sampling intervals if possible and two-hour intervals at maximum. A sampling schedule was established so that all instruments being tested at the same time had a common sampling time point at a minimum frequency of 2 hours. Internal clocks were set to local time and synchronized against the time standard provided by www.time.gov.

Reference Water Sampling Schedule – The reference sampling schedule generated between 50 - 100 comparative reference samples and was structured to examine changes in nutrient concentrations over daily to monthly time scales. Specifically, once each week ACT staff

conducted an intensive sampling event that consisted of four consecutive samples spaced at two-hour intervals. For the remaining four days of the week, ACT staff sampled once or twice per day, spaced out to cover early morning and late-afternoon timepoints or anticipated flow or tidal events. 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.

Reference Water Sample Collection - A standard 2L Van Dorn bottle was used at the CBL and HI field sites to collect reference water samples for laboratory nutrient concentration analysis. For the riverine test site a 1L acid-cleaned, polypropylene bottle was filled directly from the flow-through tank. For the tank sampling, the sampling bottle was rinsed three times before filling. For the mooring sites, the Van Dorn bottle was lowered to the same depth and as close as physically possible to the sampling inlets of all instruments and less than 1 m from any individual sampling inlet and soaked at sampling depth for 1 minute prior to sampling. The water sample was then transferred to an acid washed 1L polypropylene bottle after three initial rinses of the field sample. All environmental reference samples were processed within 10 minutes of collection while wearing clean laboratory gloves to minimize potential sources of contamination. The sample was filtered through a 47mm Whatman GFF filter into an acid cleaned vacuum flask. The first 50 ml of filtrate were discarded as a rinse. The remaining filtrate was distributed into 8 individual acid-cleaned, 30 ml polypropylene bottles to provide three analytical replicates each for NO_3 and PO_4 plus two replicates to hold as back-ups. All final sample bottles were rinsed once before filling and filled no more than ¾ full to allow adequate headspace for freezing. The final reference samples were immediately frozen and remained so until shipment to CBL-NASL for analysis.

Sample Handling and Chain of Custody - All collected reference samples at each test site were dated and coded according to site and sample sequence. Each 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. Samples were shipped on dry ice to CBL-NASL for nutrient analysis within approximately two weeks of collection. Shipping containers were sent for next morning delivery, or the soonest possible delivery time possible from a given shipping location. All samples, including the condition shipped and received, were recorded onto Chain of Custody (COC) forms and a copy sent with the samples. The COC specified time, date, sample location, unique sample number, requested analyses, sampler name, and required turnaround time, time and date of transaction between field and laboratory staff, and name of receiving party at the laboratory. NASL confirmed receipt and condition of samples within 24 hours of their arrival by signing and faxing a copy of the form to the test site.

Reference Sample Analysis

Phosphate concentrations for all reference and quality control samples were determined by the NASL at CBL following their Standard Operating Procedures Manual (CEES, UMD, Publication Series No. SS-80-04-CBL). The methodology is based on U.S. EPA Method 365.1, in Methods for chemical analysis of water and wastes (United States Environmental Protection Agency, Office of Research and Development. Cincinnati, Ohio. Report No. EPA-600-4-79-020 March 1979). In brief, ammonium molybdate and antimony potassium tartrate react in an acidic medium with dilute solutions of phosphate to form an antimony-phospho-molybdate complex. The complex is reduced to an intensely blue-colored complex by ascorbic acid. The color produced is proportional to the phosphate concentration present in the sample.

Nitrate and nitrite concentrations for all reference and quality control samples were determined by the NASL at CBL following their Standard Operating Procedures Manual (CEES, UMD, Publication Series No. SS-80-04-CBL). The methodology is based on U.S. EPA Method 353.2, in Methods for chemical analysis of water and wastes (United States Environmental Protection Agency, Office of Research and Development. Cincinnati, Ohio. Report No. EPA-600-4-79-020 March 1979). In brief, nitrate is reduced to nitrite using the cadmium reduction method. The nitrite is then determined by diazotizing with sulfanilamide and coupling with N-1-naphthylethylenediamine di hydrochloride to form a color azo dye. The absorbance measured at 540 nm is linearly proportional to the concentration of nitrate + nitrite in the sample. Nitrate concentrations are obtained by subtracting nitrite values, which have been separately determined without the cadmium reduction procedure.

All laboratory nutrient analyses were conducted on an Aquakem 250 auto-analyzer. For phosphates, a statistically-determined method of detection limit for this instrument of 0.0007 mgP/L was established by prior laboratory studies for a wide range of salinities. An expected working concentration range for this Verification and SOP was between 0.002 and 1.48 mgP/L. The detection limits for nitrate and nitrite were similarly established at 0.0007 mgN/L and 0.0006 mgN/L respectively. The typical working concentration range for the nitrate method and SOP is between 0.0049 – 5.6 mgN /L. The typical working concentration range for the nitrite method and SOP is between 0.0042 – 0.28 mgN /L. The system contains an auto-dilutor to bring any higher concentrations down to the established linear calibration range. A sample reagent blank is analyzed in conjunction with every sample as part of the routine operation of the Aqaukem 250. Approximately 40 samples per hour can be analyzed. All internal standards were verified and calibrated using certified external nutrient standards (such as Spex Certi-Prep or NIST). In addition, Field Trip Blanks and Field Sample Spike Additions (defined below) were conducted once per week by ACT as part of established quality assurance/quality control (QA/QC) protocols.

RESULTS OF LABORATORY TEST

Accuracy

HydroCycle-PO4 measurements and corresponding reference measurements for the lab concentration range challenge are shown in figure 1. Results for the highest concentration are excluded from any numerical or statistical comparisons because of its extreme range, but were included in the test to help identify maximum detection potential. The absolute difference between instrument and reference measurement across all timepoints for trials C0 – C5 ranged from -0.0163 to 0.0145 mgP/L, with a mean of -0.0039 ±0.0090 mgP/L. The means for each trial are given in Table 1. A plot of the absolute difference between HydroCycle-PO4 and reference measurement is shown in the bottom panel of figure 1. A linear regression of the measurement difference versus concentration was not significant ($p=0.36$; $r^2=0.03$), largely because of the reversal in the offset (over-prediction) at the highest concentration.

Table 1. Accuracy results for laboratory testing of the HydroCycle-PO4 analyzer assessed by absolute difference (mgP/L) and percent error between instrument and reference measurements for the concentration range test.

Trial	Reference	Hydrocycle-PO4	Absolute Diff	% Error
C0	0.0059	0.0030	**-0.0029**	**48.8**
C1	0.0105	0.0089	**-0.0016**	**15.4**
C2	0.0189	0.0152	**-0.0037**	**19.5**
C3	0.0621	0.0496	**-0.0126**	**20.2**
C4	0.1159	0.1021	**-0.0138**	**11.9**
C5	0.4059	0.4177	**0.0118**	**2.9**

Precision

An assessment of precision was performed by computing the standard deviations and coefficients of variation of the five replicate measurements for each of the C1 – C5 concentration trials. The standard deviation of the mean ranged from 0.0005 to 0.0020 mgP/L across the five trials, and the coefficient of variation ranged from 0.14 to 5.78 percent (Table 2).

Table 2. Precision assessment of the HydroCycle-PO4 analyzer during the laboratory concentration range testing. Variance is reported as the standard deviation and coefficient of variation of five replicate measurements collected at 30 minute intervals in a well-mixed tank maintained at known uniform conditions.

	Mean PO_4 (mgP/L)		Standard Deviation		Coefficient of Variation	
Trial	*Reference*	*HydroCycle-PO4*	*Reference*	*HydroCycle-PO4*	*Reference*	*HydroCycle-PO4*
C1	0.0105	0.0089	0.0004	0.0005	4.21	5.78
C2	0.0189	0.0152	0.0001	0.0007	0.66	4.58
C3	0.0621	0.0496	0.0005	0.0020	0.75	4.10
C4	0.1159	0.1021	0.0003	0.0018	0.23	1.78
C5	0.4059	0.4177	0.0023	0.0006	0.56	0.14

Lab Concentration Range Challenge

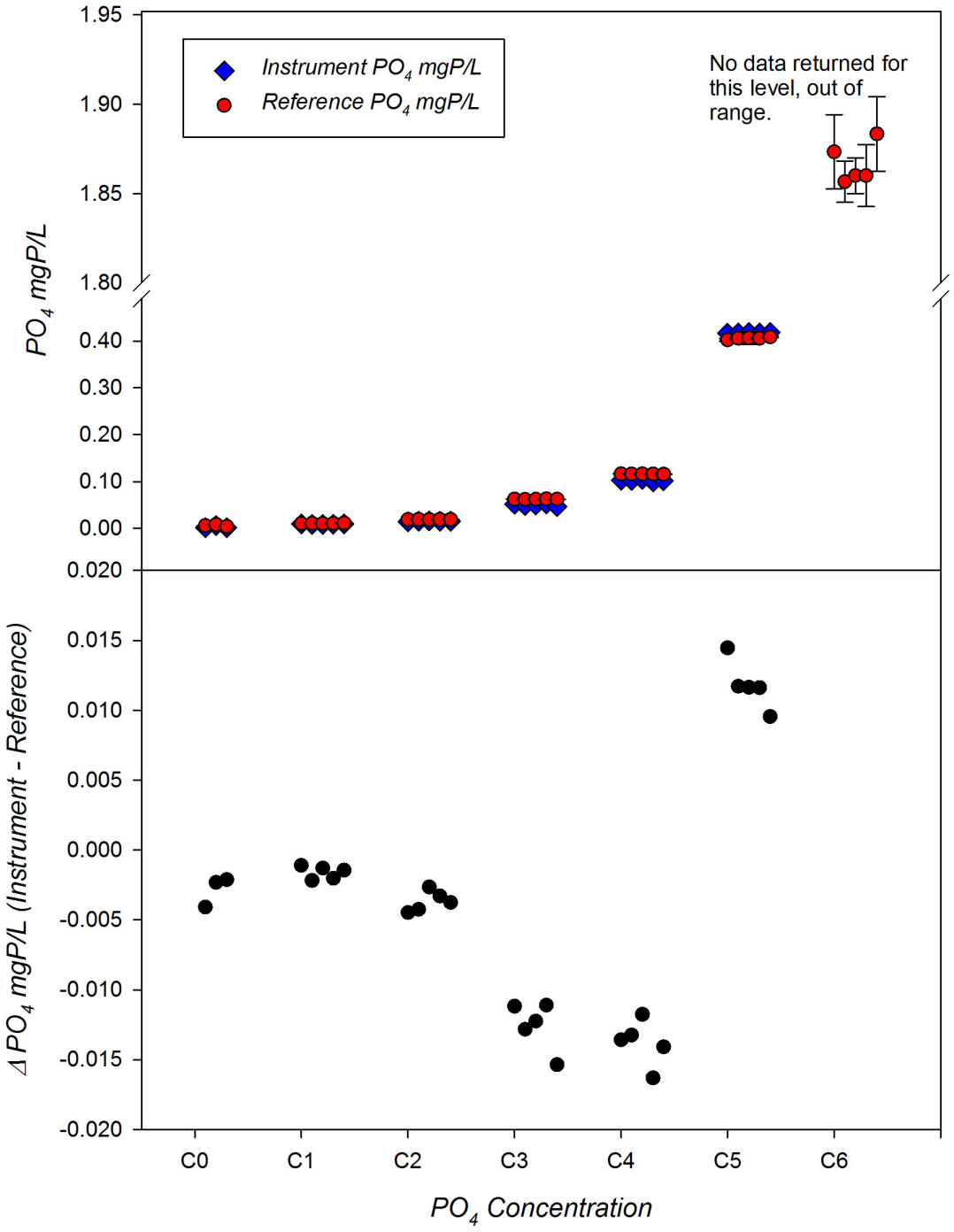


Figure 1. *Top Panel:* Plot of instrument (blue dots) and reference (red dots) measurements of PO_4 in the laboratory concentration range challenge covering ambient plus 6 concentration ranges. Five replicate measurements were made at each concentration level along with three measurements at ambient level. *Bottom Panel:* Plot of the absolute difference in mgP/L between HydroCycle-PO4 and reference measurement.

Time series results of ambient water quality conditions for the salinity, turbidity, and DOC matrix challenges are presented in figure 2. Final test concentrations of turbidity and DOC were slightly below the stated target levels, and there was noticeable settling of turbidity at the highest addition level, but confirm the overall challenge conditions being tested.

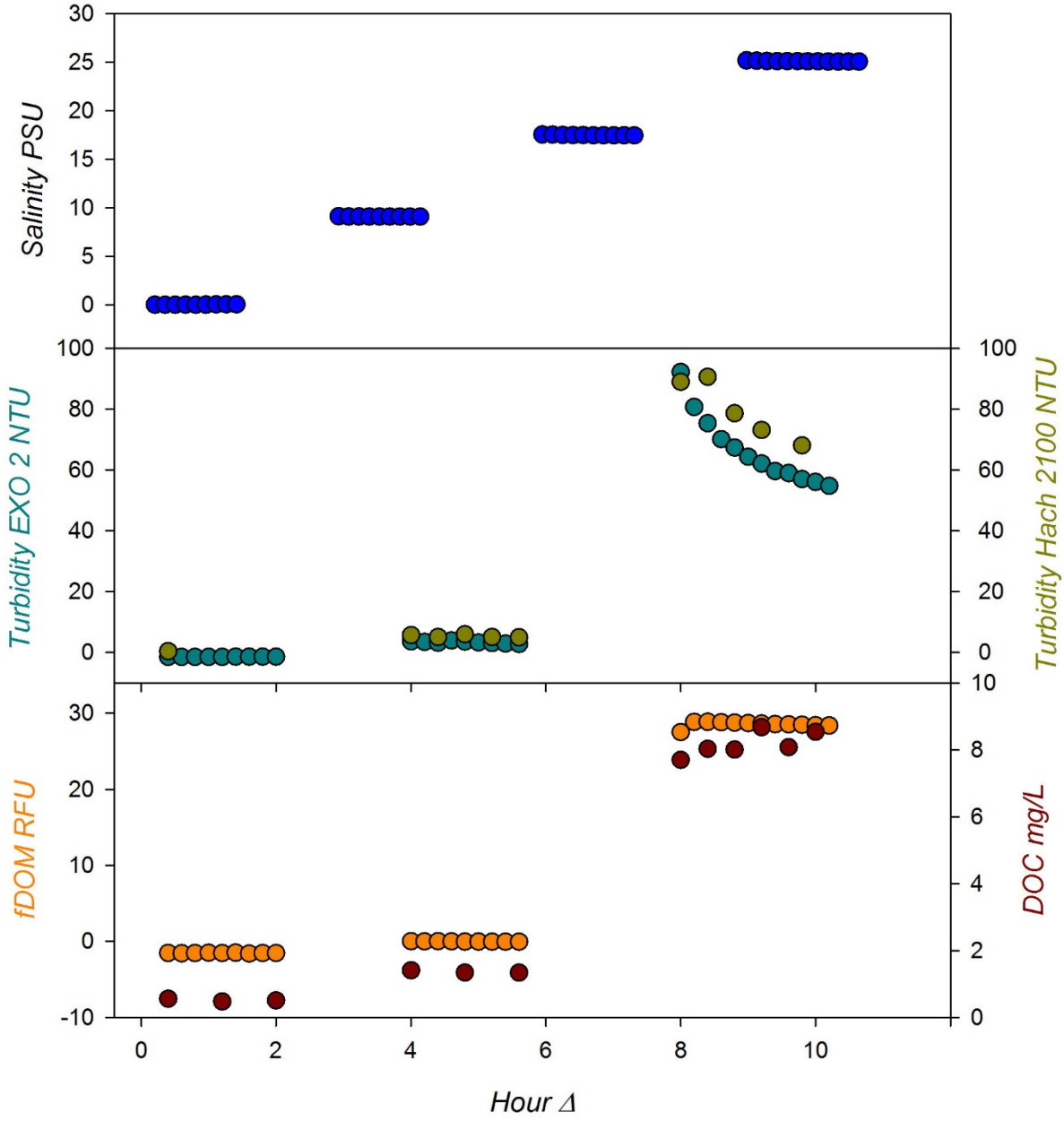


Figure 2. *Top Panel: In situ* salinity measured by EXO2 sonde in the laboratory salinity challenge covering ambient plus 3 salinity ranges. *Middle Panel: In situ* turbidity measured by EXO2 sonde (teal) and on grab samples by a Hach 2100 Turbidimeter (olive) during the laboratory turbidity challenge covering ambient plus 2 additions. *Bottom Panel: In situ* fDOM measured by EXO2 sonde (orange) and DOC of discrete samples (dark red) during the DOC challenge covering ambient plus 2 additions.

Results of the laboratory temperature challenge at 5 ^{o}C are shown in figure 3. The absolute difference between instrument and reference measurement across all timepoints for trials C2 – C4 ranged from -0.0140 to -0.0046 mgP/L, with a mean of -0.0087 ±0.0032 mgP/L. The means for each trial are given in Table 3. Measurement differences were significantly different for the C2 and C4 trials at 5 versus 20 ^{o}C, but the temperature effect was in the opposite direction between those trials. The measurement difference at C3 was nearly identical for the two temperatures. Therefore no clear pattern of a temperature effect on accuracy was observed.

Table 3. Summary of accuracy results for temperature trials assessed by absolute difference (mgP/L) and percent error between instrument and reference measurements.

Trial	Reference	HydroCycle-PO4	Absolute Diff	% Error
C2	0.0109	0.0043	**-0.0066**	**61.0**
C3	0.0547	0.0422	**-0.0125**	**22.8**
C4	0.1040	0.0969	**-0.0071**	**6.8**

Results of the laboratory salinity challenge, performed at the C3 concentration level, are shown in figure 4. The absolute difference between instrument and reference measurement across all timepoints for the three added salinity levels ranged from -0.0124 to 0.0086 mgP/L, with a mean of -0.0018 ±0.0053 mgP/L. The means for each salinity trial are given in Table 4. The zero salinity results are taken from the initial concentration challenge on day 1. The lack of a constant concentration across the three salinity is noted and indicated a potential reaction with the added salt solution. There was a statistically significant response to increased salinity with the offsets increasing in a positive direction as salinity increased. A linear regression of the measurement differences versus salinity ($p<0.0001$; $r^2=0.70$) had a slope of 0.0005 and intercept of- 0.013. The average offset at salinity 30 was around 0.015 mgP/L higher than for the zero salinity.

Table 4. Summary of accuracy results for salinity trial assessed by absolute difference (mgP/L) and percent error between instrument and reference measurements.

Trial	Reference	HydroCycle-PO4	Absolute Diff	% Error
0	0.0621	0.0496	**-0.0126**	**20.2**
10	0.0443	0.0383	**-0.0060**	**13.6**
20	0.0385	0.0375	**-0.0011**	**2.7**
30	0.0297	0.0315	**0.0018**	**6.0**

Results of the laboratory turbidity challenge, performed at the C3 concentration level, are shown in figure 5. The absolute difference between instrument and reference measurement across all timepoints for the two added turbidity levels ranged from -0.0226 to -0.0113 mgP/L, with a mean of -0.0170 ±0.0047 mgP/L. The means for each turbidity trial are given in Table 5. Results for the zero turbidity level are taken from the initial concentration challenge on day 1. The measurement difference for the 100 NTU trial was 0.009 mgP/L more negative (under-predicted) than results observed at the 0 and 10 NTU trials. This increased negative offset corresponded to a doubling in the relative error to over 40%. A linear regression of the measurement differences versus turbidity was significant ($p=0.0001$; $r^2=0.68$), with a slope of -0.0005 and intercept of -0.010, however the trend line was clearly forced by the large decrease at 100 NTU.

Table 5. Summary of accuracy results for turbidity trials assessed by absolute difference (mgP/L) and percent error between instrument and reference measurements.

Trial	Reference	HydroCycle-PO4	Absolute Diff	% Error
0	0.0621	0.0496	**-0.0126**	**20.2**
10	0.0525	0.0398	**-0.0127**	**24.1**
100	0.0520	0.0306	**-0.0214**	**41.1**

Results of the laboratory DOC challenge, performed at the C3 concentration level, are shown in figure 6. The absolute difference between instrument and reference measurement across all timepoints for the two added DOC levels ranged from -0.0112 to -0.0074 mgP/L, with a mean of -0.0096 ±0.0014 mgP/L. The means for each of the DOC trials are given in Table 6. Results for the zero DOC level are taken from the initial concentration challenge on day 1. The measurement difference generally became less negative as DOC concentration increased. The offset for the 10 mg/L trial was about 20 percent less under-predicted then for the 1 mg/L trial. A linear regression of the measurement differences versus DOC concentration was highly significant ($p=0.0005$; $r^2=0.62$) with a slope of 0.0002 and intercept of -0.013.

Table 6. Summary of accuracy results for Laboratory testing assessed by absolute difference (mgP/L) and percent error between instrument and reference measurements for each individual trial condition within each matrix challenge.

Trial	Reference	HydroCycle-PO4	Absolute Diff	% Error
0	0.0621	0.0496	**-0.0126**	**20.2**
1	0.0560	0.0454	**-0.0105**	**18.8**
10	0.0762	0.0675	**-0.0087**	**11.4**

Lab Temperature Challenge

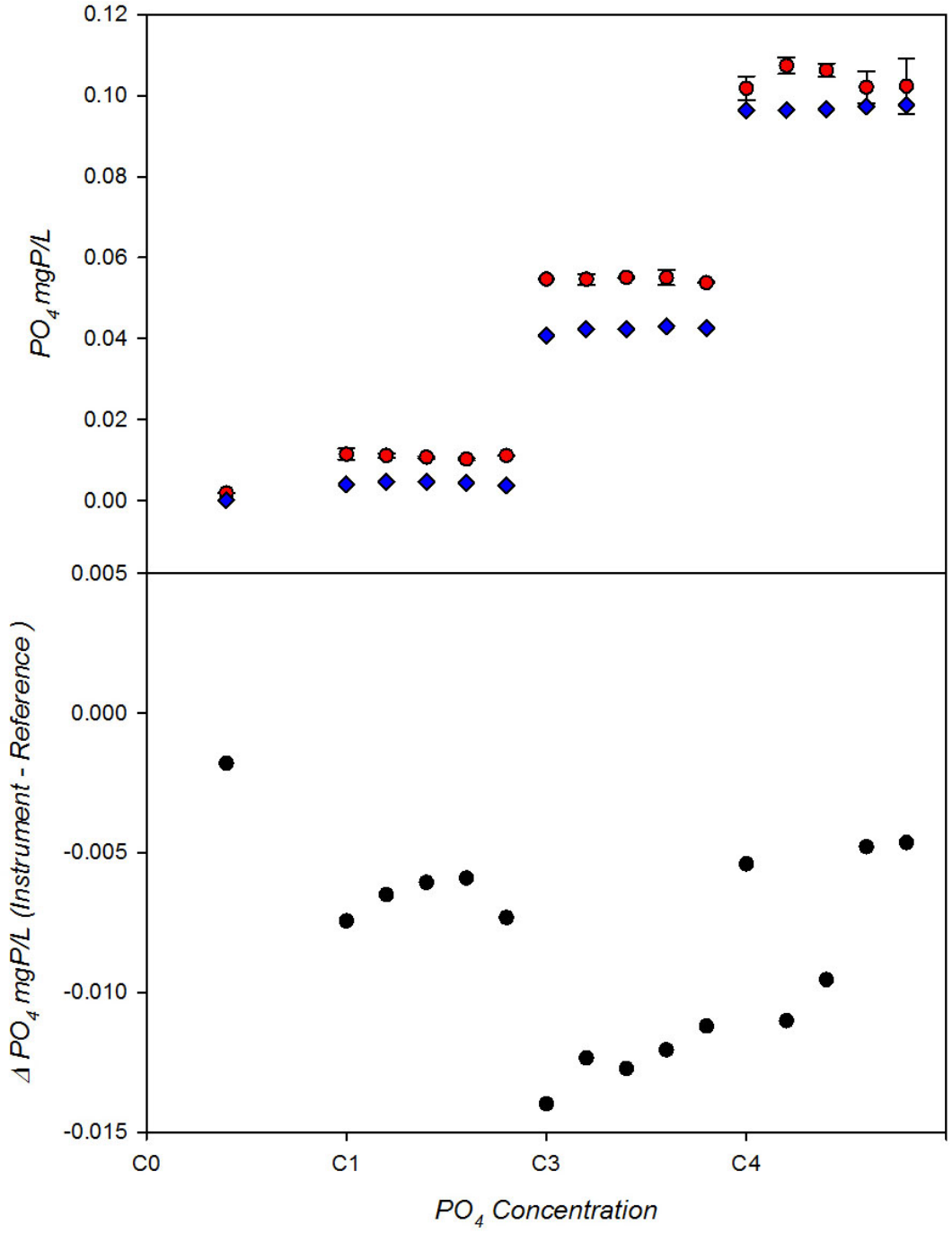


Figure 3. *Top Panel:* Plot of instrument (blue dots) and reference (red dots) measurements of PO_4 (mgP/L) in the temperature response challenge covering concentration ranges C2 – C4 measured at 5 ^{o}C test conditions. Five replicate measurements were made at each concentration level along with one measurement at ambient level. *Bottom Panel:* Plot of the absolute difference between HydroCycle-PO4 and reference measurement.

Lab Salinity Challenge

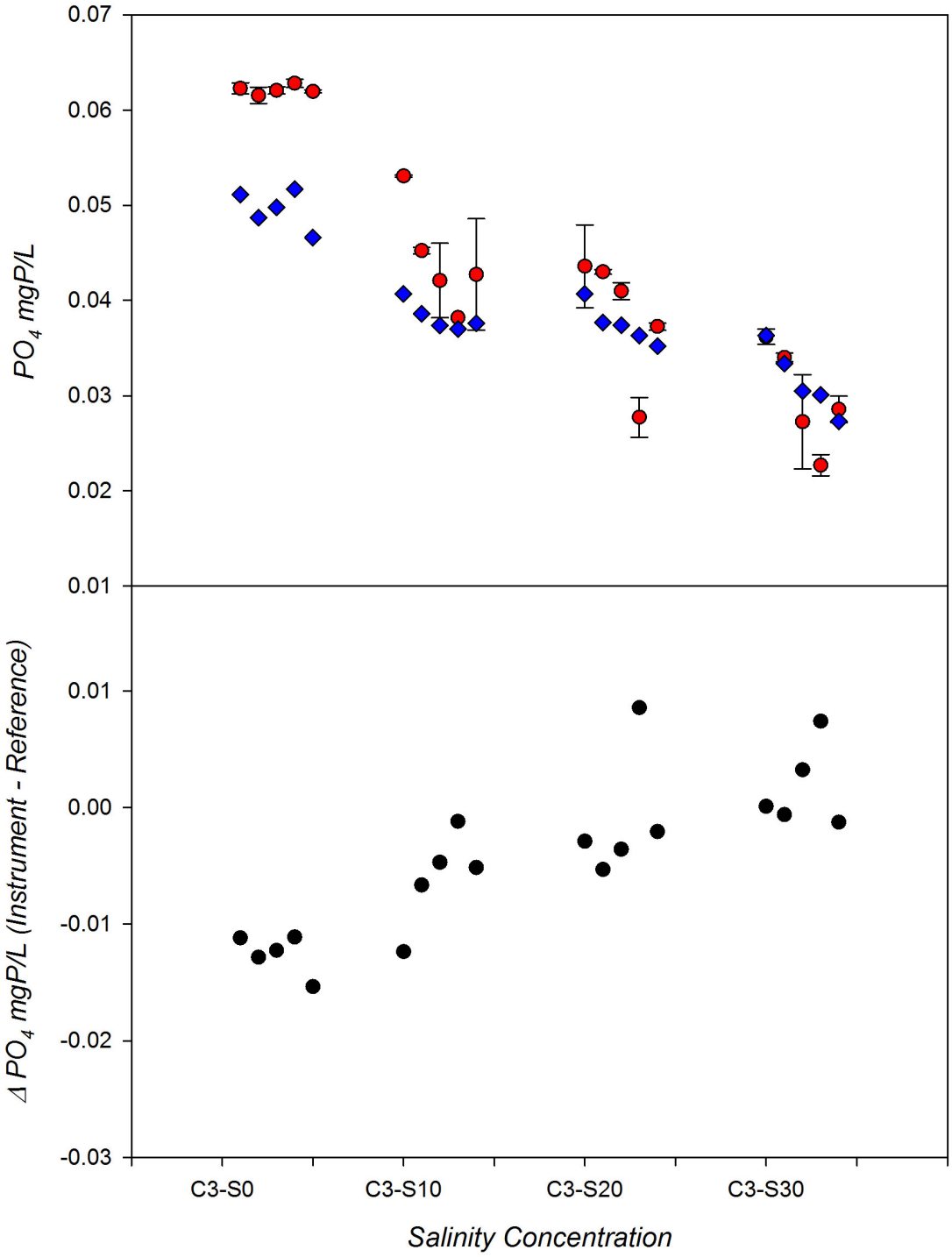


Figure 4. *Top Panel:* Plot of instrument (blue dots) and reference (red dots) measurements of PO_4 (mgP/L) at four salinity levels for the C3 concentration. Five replicate measurements were made at each concentration level. *Bottom Panel:* Plot of the absolute difference between HydroCycle-PO4 and reference measurement.

Lab Turbidity Challenge

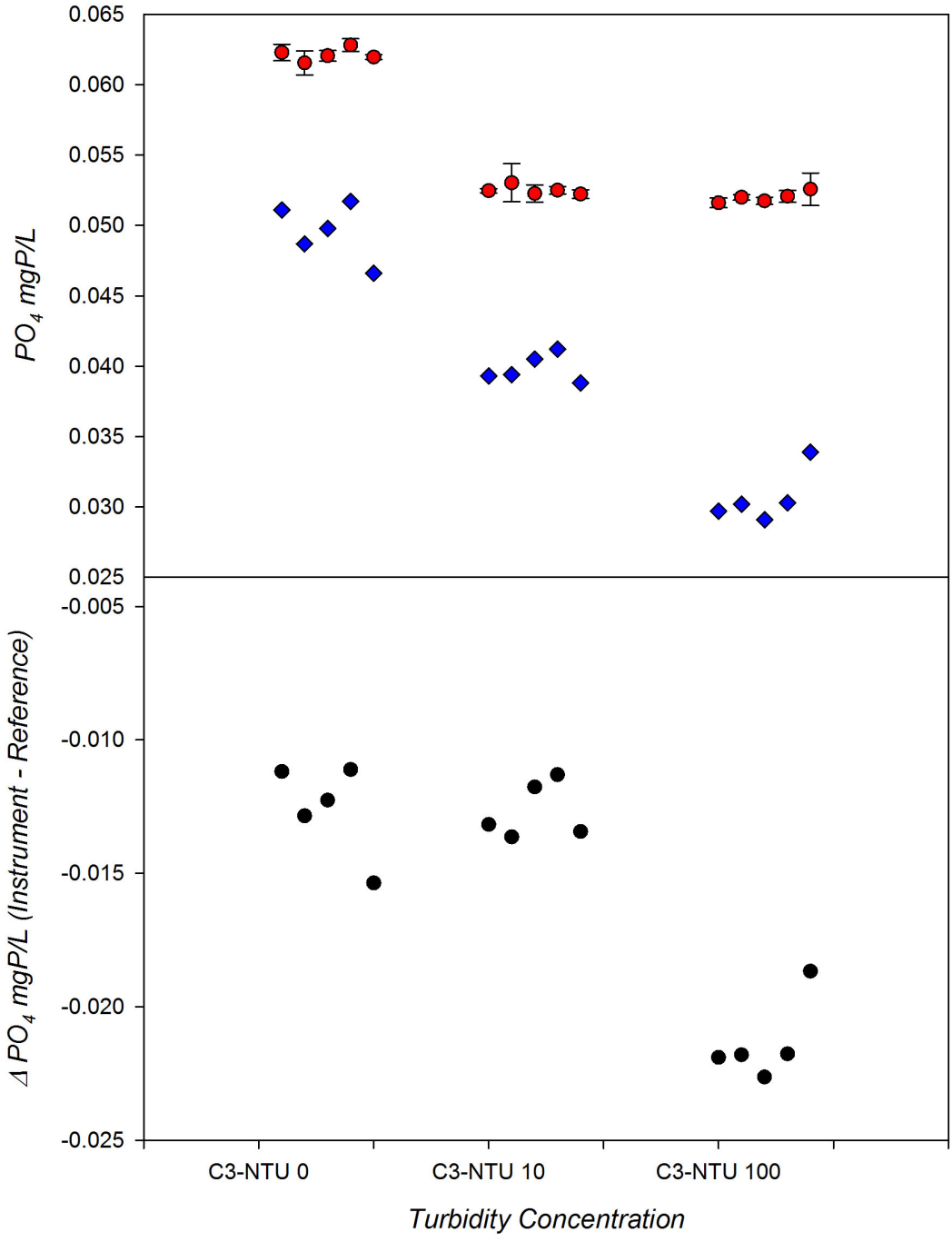


Figure 5. *Top Panel:* Plot of instrument (blue dots) and reference (red dots) measurements of PO_4 (mgP/L) at three turbidity levels for the C3 concentration. Five replicate measurements were made at each concentration level. *Bottom Panel:* Plot of the absolute difference between HydroCycle-PO4 and reference measurement.

Lab DOC Challenge

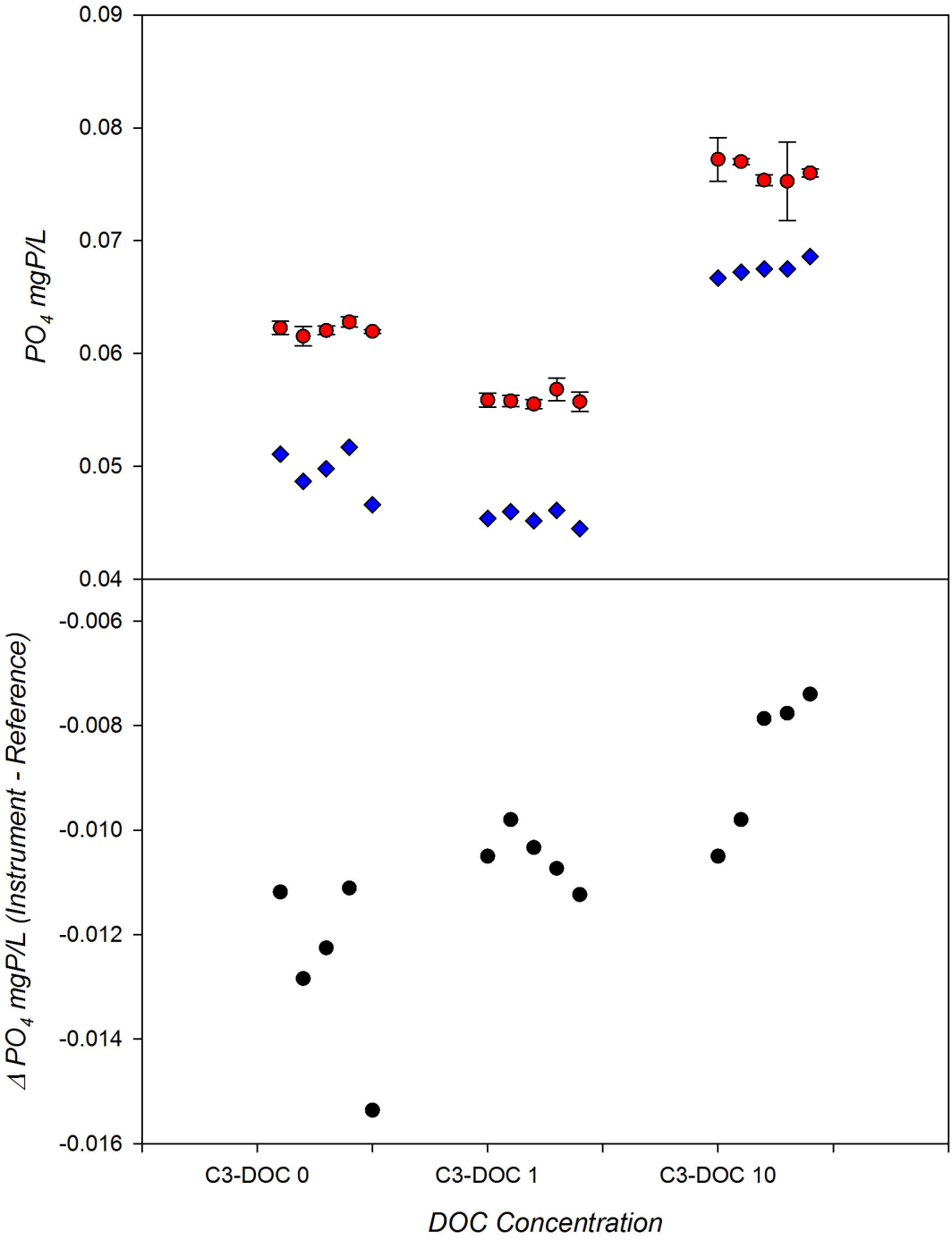


Figure 6. *Top Panel:* Plot of instrument (blue dots) and reference (red dots) measurements of PO_4 (mgP/L) at three DOC levels for the C3 concentration. Five replicate measurements were made at each concentration level. *Bottom Panel:* Plot of the absolute difference between HydroCycle-PO4 and reference measurement.

A summary of measurement differences between the HydroCycle-PO4 and reference sample for each trial of each laboratory challenge is presented together in figure 7. For most of the Lab trials the HydroCycle-PO4 under-predicted concentrations, but with a total range for all trials from -0.021 to 0.012 mgP/L. Measurement offset increased at higher concentration during the range test but unexpectedly changed in the direction of the offset at the C5 trial. There was no clear temperature or turbidity effect on measurement accuracy across the levels tested. There were significant trends in measurement offset during the salinity and DOC trials with the HydroCycle-PO4 measurement increasing at higher levels, which resulted in the measurement deltas to be less negative. Results of measurement differences averaged across all trials within each of the challenge matrices are presented in Table 7.

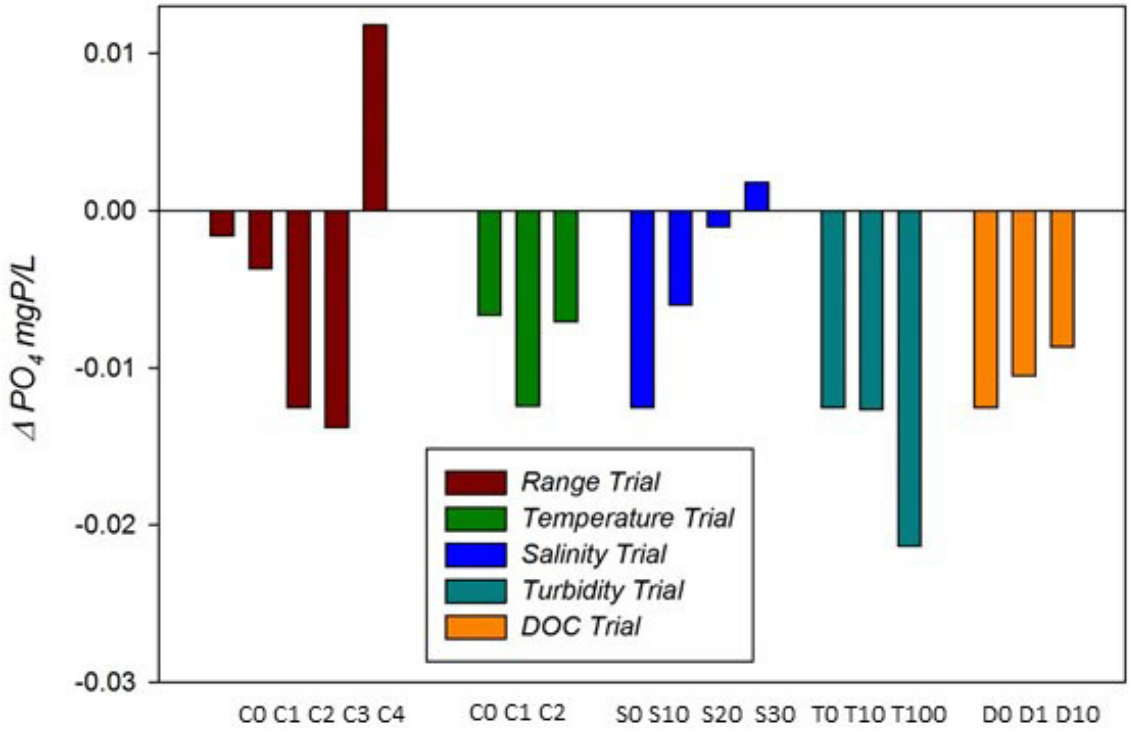


Figure 7. Global summary of difference between instrument and reference measurements for all laboratory tests at each trial conditions for the HydroCycle-PO4 analyzer.

Table 7. Measurement differences in mgP/L (min, max, mean, stdev) between instrument and reference concentrations averaged across all trials within a laboratory challenge.

HydroCycle-PO4	Range	Temp	Salinity	Turbidity	DOC
min	-0.0138	-0.0125	-0.0060	-0.0214	-0.0105
max	0.0118	-0.0066	0.0018	-0.0127	-0.0087
mean	**-0.0040**	**-0.0087**	**-0.0018**	**-0.0170**	**-0.0096**
stdev	0.0103	0.0032	0.0039	0.0061	0.0013

RESULTS of FIELD TESTS

Moored field tests were conducted to examine the performance of the HydroCycle-PO4 to consistently track natural changes in PO_4 over extended field deployments with durations of 31-84 days. In addition, field tests examined the reliability of the instrument, i.e., the ability to maintain integrity or stability of data collection over time. Reliability was determined by quantifying the percent of expected data that was recovered and useable. The performance of the HydroCycle-PO4 was examined in three separate field tests at various ACT Partner sites to include a range of biogeochemical conditions. The range and mean for temperature and salinity for each test site is presented in Table 8. The reference temperature and conductivity data was measured by RBR thermistors and a SeaBird SBE 26 or Xylem EXO2 sonde that were mounted at the same sampling depth as the test instrument.

Immediately before and after each deployment, samples of the on-board standards were taken from the instrument for comparison against a reference measurement and to assess their stability over the course of the deployment (Table 9).The HydroCycle-PO4 was calibrated and programmed for deployment by the manufacturer representative.

Table 8. Range and average for temperature, and salinity at each of the test sites during the sensor field deployments. Temperature and salinity were measured by RBR temperature loggers and a SeaBird SBE 26 or a Xylem EXO2 mounted on the instrument rack or in the tank for the duration of the deployment.

SITE (deployment period/duration)		**Temperature** ($^{\circ}$C)	**Salinity** (PSU)
Maumee River	Min.	20.1	0.0
26May – 27Jun	Max.	27.7	0.3
(n = 32 days)	Mean	24.3	0.2
Chesapeake Bay	Min.	20.0	12.7
18Jul – 10Oct	Max.	31.1	16.9
(n = 84 days)	Mean	27.2	14.7
Kaneohe Bay	Min.	24.5	27.3
3Oct – 2Nov	Max.	27.9	34.8
(n = 31 days)	Mean	26.3	34.2

Table 9. Results of the pre-deployment and post-deployment standard check for the HydroCycle-PO4 for each deployment site. (n.d. denotes no data for that observation.)

Deployment Site	**Expected PO_4 mgP/L**	**Pre PO_4 mgP/L**	**Post PO_4 mgP/L**
UM	0.3000	0.3390	0.3143
CBL	0.3000	0.2896	0.2270
HIMB	n.d.	n.d.	n.d.

Deployment at Maumee River Bowling Green, Ohio

A 32 day deployment occurred from May 26 through June 27 in the Maumee River, at the facilities of the Bowling Green, Ohio Water Treatment Plant (Figure 8). The deployment site was located at 41.48° N, 83.74° W, in a flow-through tank located in the water treatment plant pump house. The pump house is located above the Maumee, approximately 200 m up river from the water treatment intake and approximately 35 km from the Maumee outflow into Lake Erie. River water was continuously pumped into a 180 gallon test tank where it was mixed using two submerged pumps. The residence time in the tank was approximately 10 minutes. The instrumentation was suspended within the tank with the sampling inlet 0.2 m off the bottom.

Figure 8. Aerial view of the Maumee River (left) and the flow through deployment tank (right).

Time series results of ambient conditions for river discharge, temperature, specific conductivity, turbidity and chlorophyll are given in figure 9. Temperature ranged from 20.5 – 27.7°C, specific conductivity from 423 - 689 µS/cm, turbidity from 8 – 681 NTU, and chlorophyll from 4.5 – 131 µg/L over the duration of the field test.

The HydroCycle-PO4 operated continuously during the entire 32 day deployment sampling at hourly intervals, but of the 763 possible data points 666 were flagged by the instrument as bad. Accepting the flagged data, and omitting 131 outliers (<-0.004 or >0.250), the HydroCycle-PO4 generated 632 observations out of a possible 763 for a data completion result of 82.8%. Time series results of the HydroCycle-PO4 measurements and corresponding reference PO_4 results are given in figure 10 (top panel). PO_4 measured by the HydroCycle-PO4 ranged from -0.004 to 0.248 mgP/L compared to a range of 0.001 to 0.118 mgP/L within the reference samples.

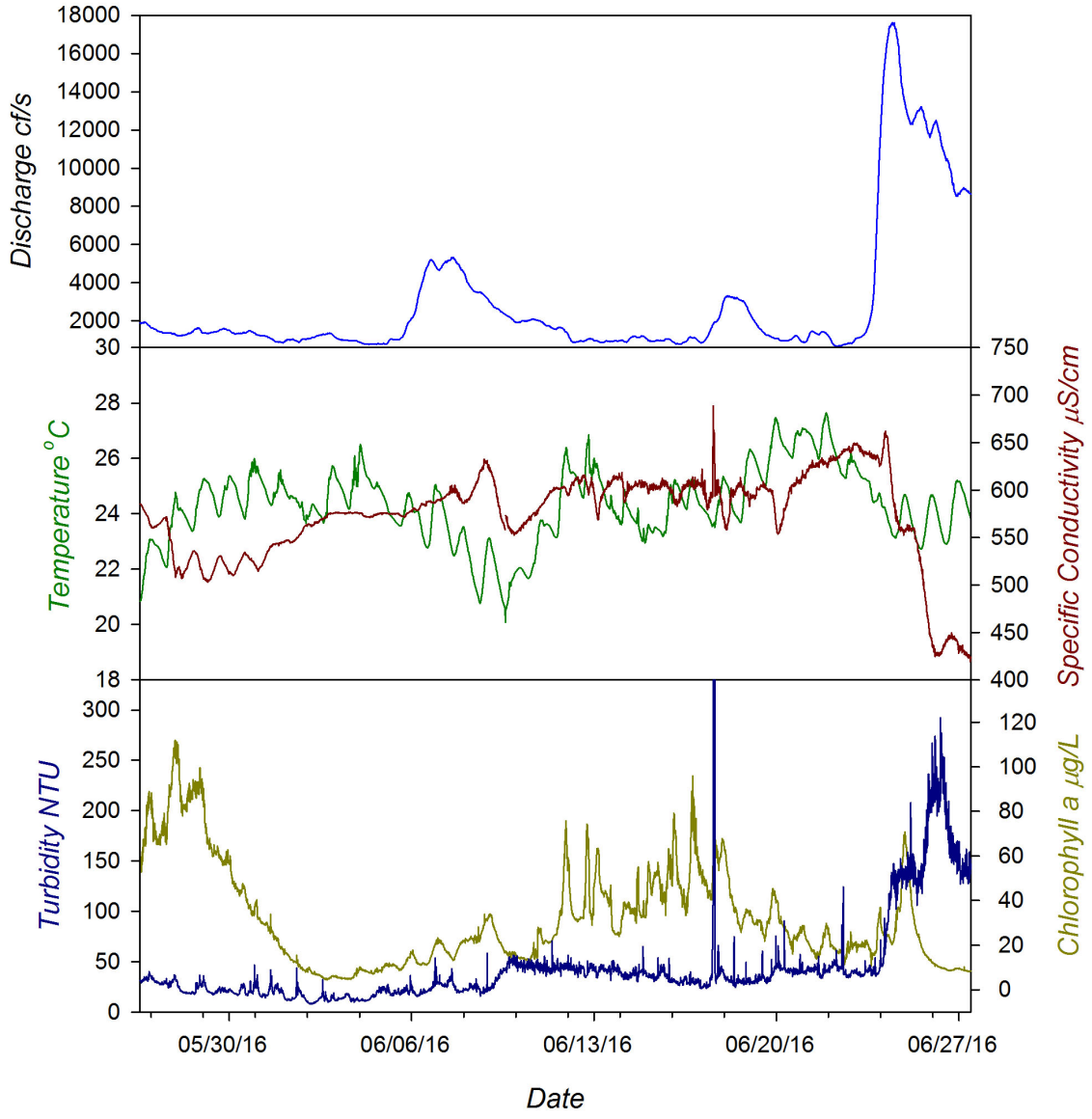


Figure 9. Environmental conditions encountered during the 32 day freshwater deployment in the Maumee River at Waterville, OH. *Top Panel:* Variation in river discharge over the term of the deployment. *Middle Panel:* Variation in temperature (green) and Conductivity (red) at the depth of the sensors, measured by an EXO 2 Sonde. *Bottom Panel:* Time series of turbidity (blue) and chlorophyll (dark yellow) as measured by the EXO 2 Sonde. The large spike in turbidity (681 NTU) was produced during a nutrient addition test when sediment accumulated on the bottom was stirred up from additional mixing of the tank.

The time series of the difference between instrument and reference PO_4 measurements for each matched pair (n=43 of a possible 51 observations) is given in the bottom panel of figure 10. Eight of the 51 possible comparisons were omitted due to their extreme range. The average and standard deviation of the measurement difference over the total deployment was -0.022 ± 0.029 mgP/L with a total range of -0.110 to 0.033 mgP/L. There was no significant trend in the measurement difference over time as estimated by linear regression (p= 0.20; r^2=0.04).

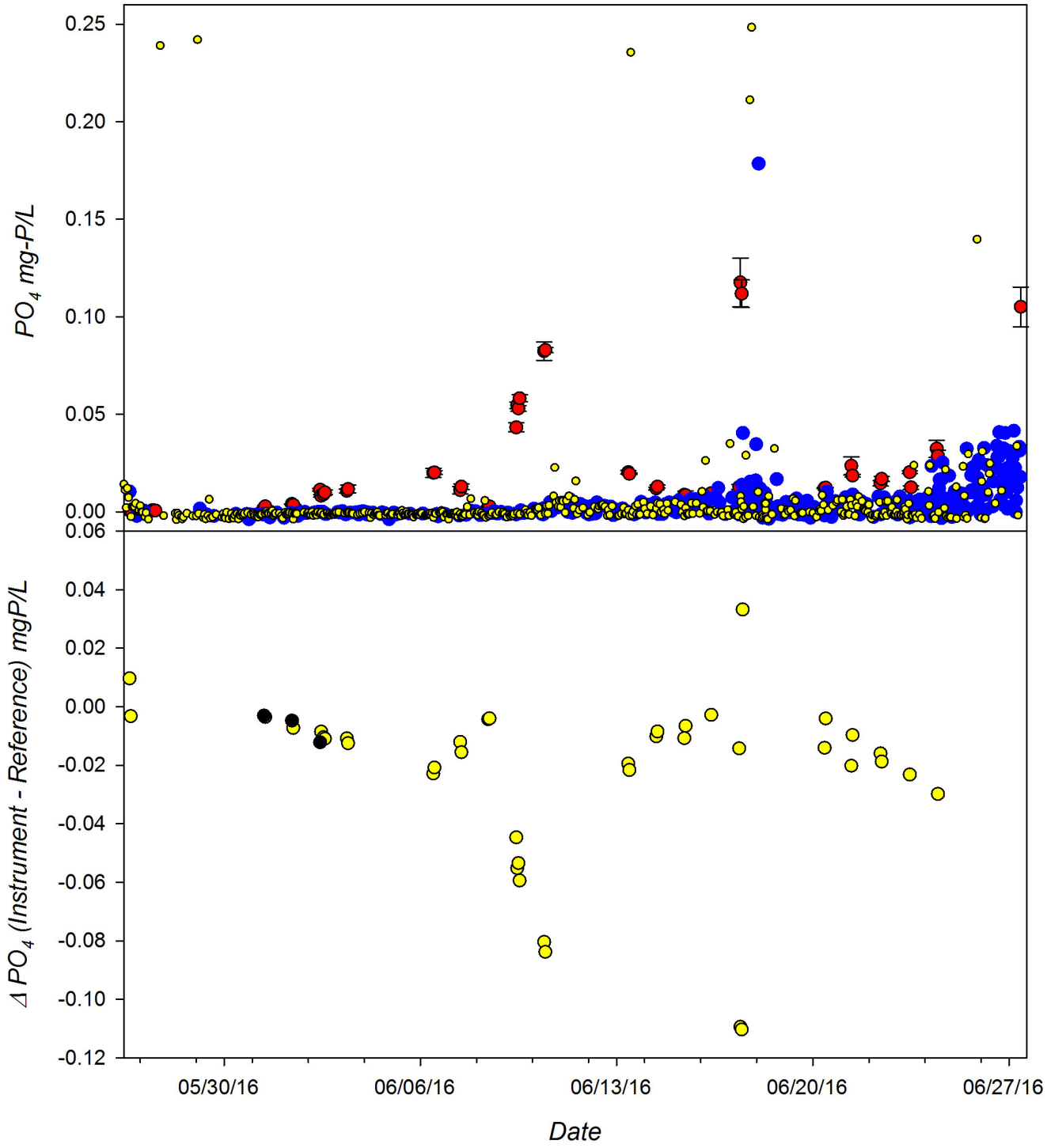


Figure 10. *Top Panel:* Time series plot of the HydroCycle-PO4 measurement (blue dots) and reference measurements (red dots) of phosphate in mgP/L. The yellow dots represent instrument flagged data, but results were accepted for comparisons. *Bottom Panel:* Time series plot of the difference between the HydroCycle-PO4 and reference measurements of phosphate in mgP/L (instrument – reference) during the freshwater deployment in the Maumee River at Waterville, OH.

A cross-plot of all matched observations for the Maumee River deployment is given in figure 11. A linear regression of instrument versus reference measurement was not significant ($p=0.93$; $r^2 = 0.0002$). In general the HydroCycle-PO4 under-predicted concentrations.

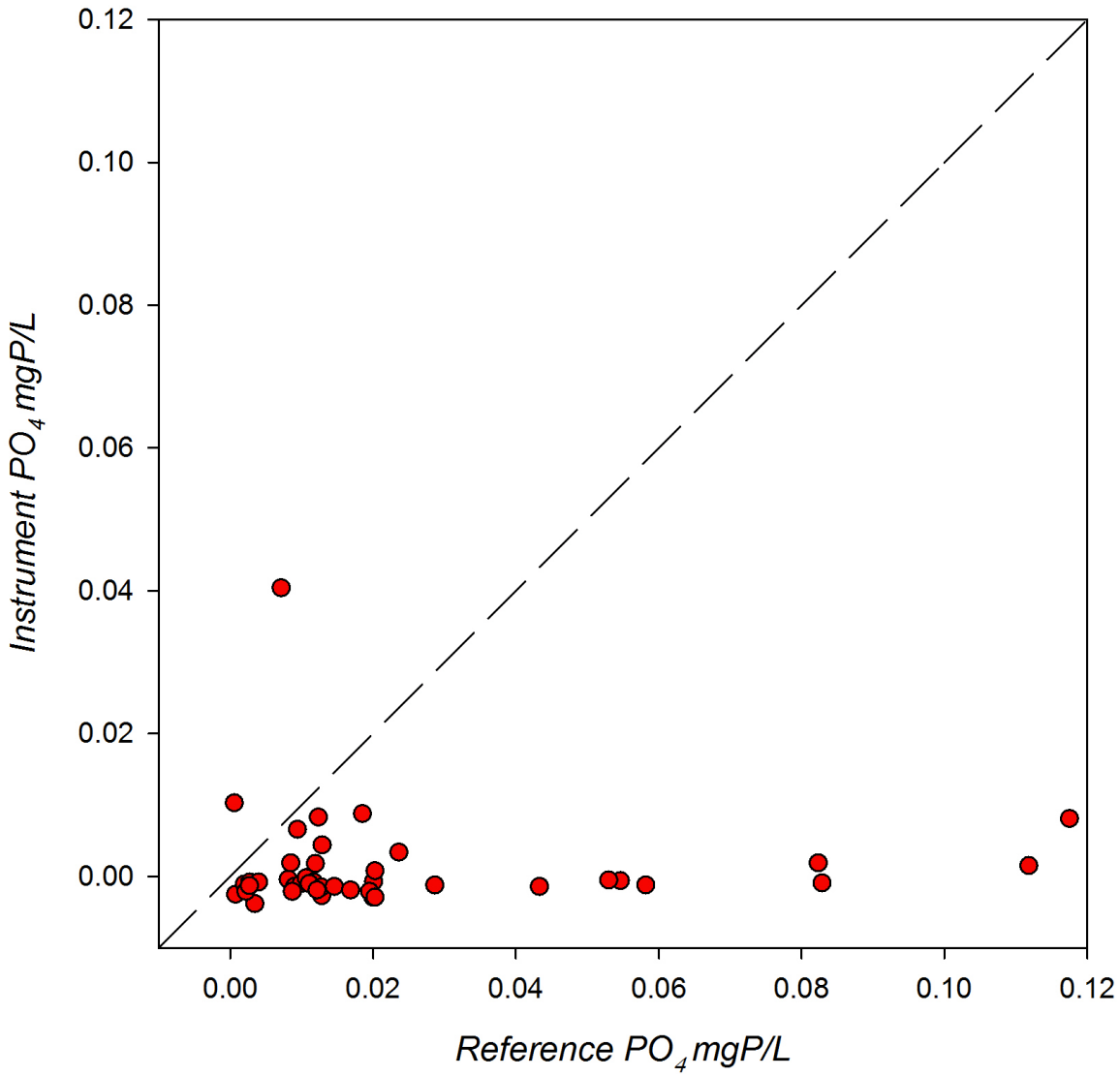


Figure 11. Maumee River field response plot for the 32 day deployment of the HydroCycle-PO4 compared to reference PO_4 samples. The plotted line represents a 1:1 correspondence.

Photographs of test instrument before and after the field deployment to indicate potential impact of biofouling (Figure 12).

Figure 12. Photographs of the HydroCycle-PO4 prior to and following a 32 day field test in the Maumee River.

Deployment at Chesapeake Biological Laboratory (CBL)

An 84 day moored field test was conducted in Chesapeake Bay from July 18 to October 10, 2016. The deployment was located at 38.32°N, 76.45°W attached to the side of a floating pier at the mouth of the Patuxent River (Figure 13.) The site was brackish with an average water depth of 2.2 m at the test site.

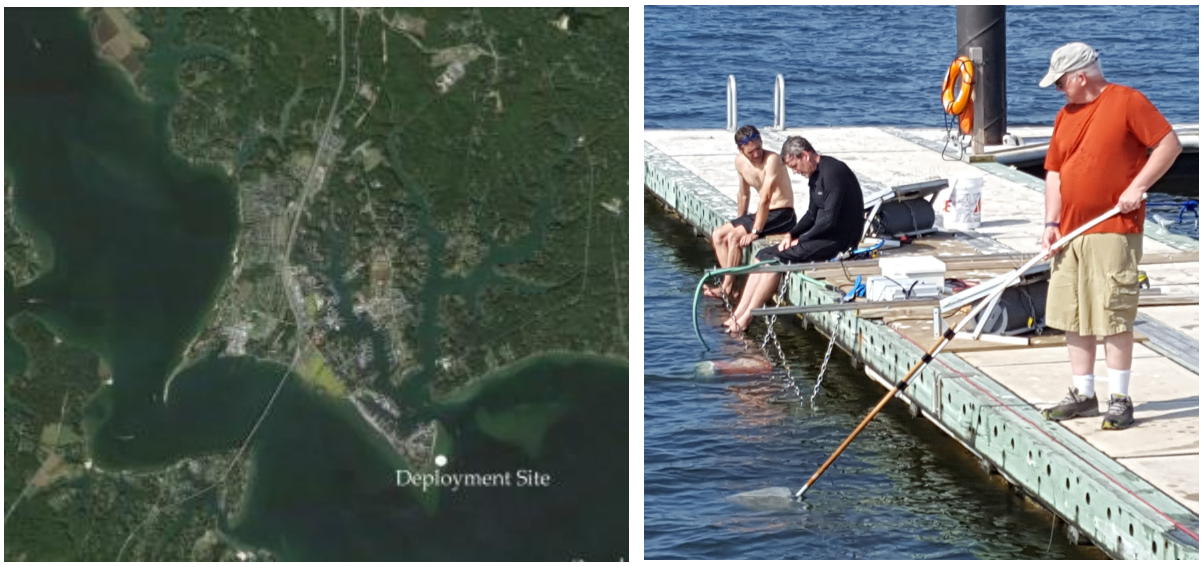


Figure 13. Aerial view of CBL deployment site (left) and instrument deployment rack off the dock during deployment.

Time series results of ambient conditions for tidal height, temperature, salinity, turbidity and chlorophyll are given in figure 14. Temperature ranged from 20.0 to 31.3°C, salinity from 12.7 to 16.9 PSU, turbidity from 0.5 to 936.3 NTU and chlorophyll from 0.2 to 97.1 µg/L over the duration of the field test.

The HydroCycle-PO4 had issues communicating with the datalogger, it was retrieved and reprogrammed but lost data for the first two days of the deployment. The instrument stopped reporting data on 9/19 due to the power cable fraying. The HydroCycle reported 728 of a possible 730 measurements while operational, but this was only 72% of the total possible 1006 values for the entire deployment. During the working deployment, two values were flagged by the instrument as bad data and there were no outliers. Time series results of the HydroCycle-PO4 and corresponding reference PO_4 results are given in figure 15 (top panel). For the working interval, the range of values reported by the HydroCycle-PO4 was 0.001 to 0.024 mgP/L, compared to 0.003 to 0.034 mgP/L within reference samples.

The bottom panel of figure 15 presents the time series of the difference between the HydroCycle-PO4 and reference PO_4 for each matched pair (n=71 comparisons out of a total of 103, (32 missing data points from retrieval/repair period and power failure). The average and standard deviation of the measurement difference for the deployment was -0.005 ±0.005 mgP/L, with the total range of differences between -0.018 to 0.003 mgP/L. There no significant trend in measurement difference over time as estimated by linear regression ($p=0.89$; $r^2=0.0003$) over the deployment period.

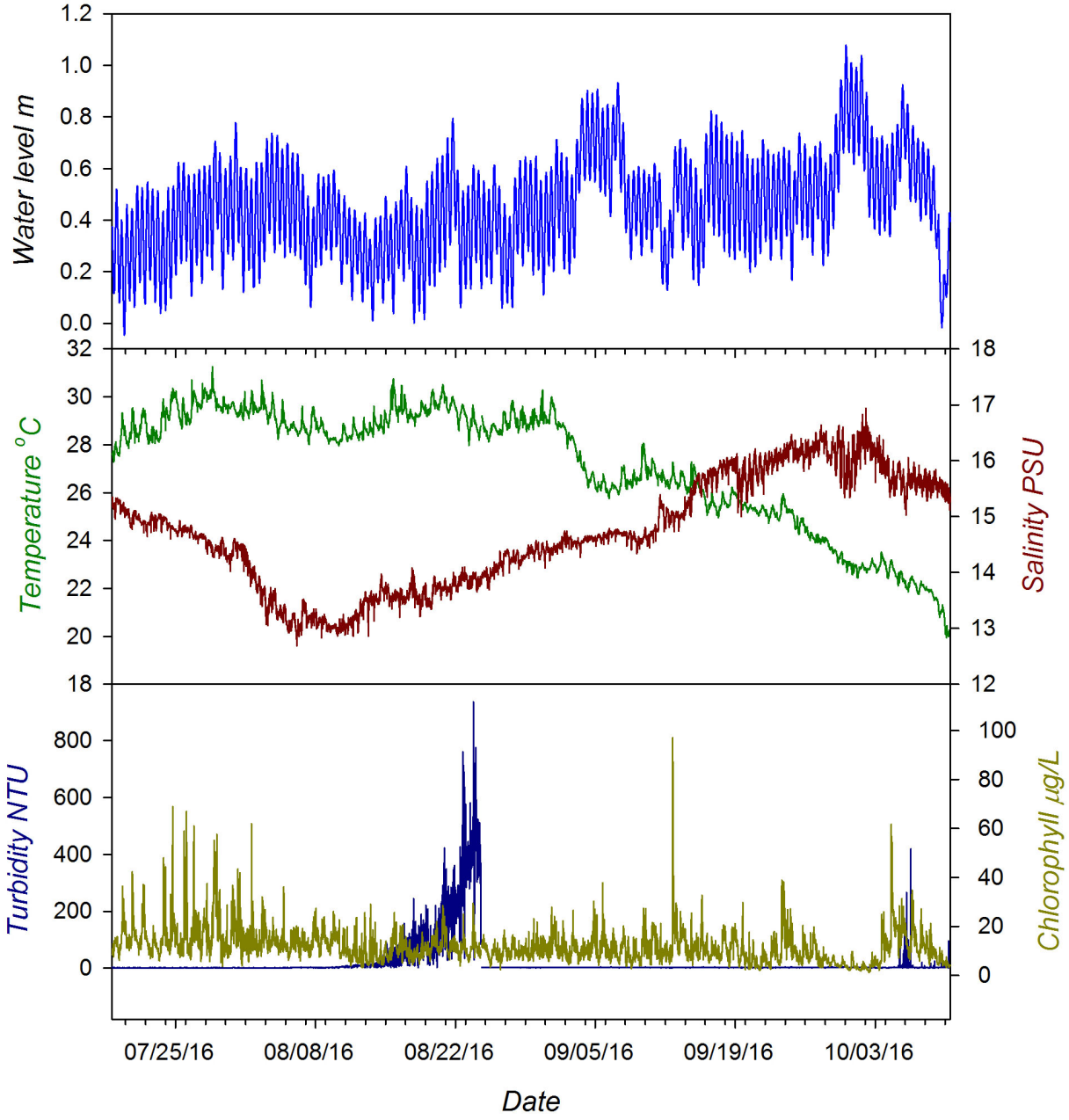


Figure 14. Environmental conditions encountered during the 84 day CBL floating dock deployment. Test sensor array deployed at 1 m fixed depth, variation in local tidal heights indicate active water flow around instrument (*Top Panel*). Variation in temperature (green) and salinity (red) at depth of instrument sensor detected by an EXO2 sonde and two RBR Solo thermistors (*Middle Panel*). Variation in turbidity (blue) and chlorophyll (dark yellow) at depth of instrument sensor detected by an EXO2 sonde (*Bottom Panel*).

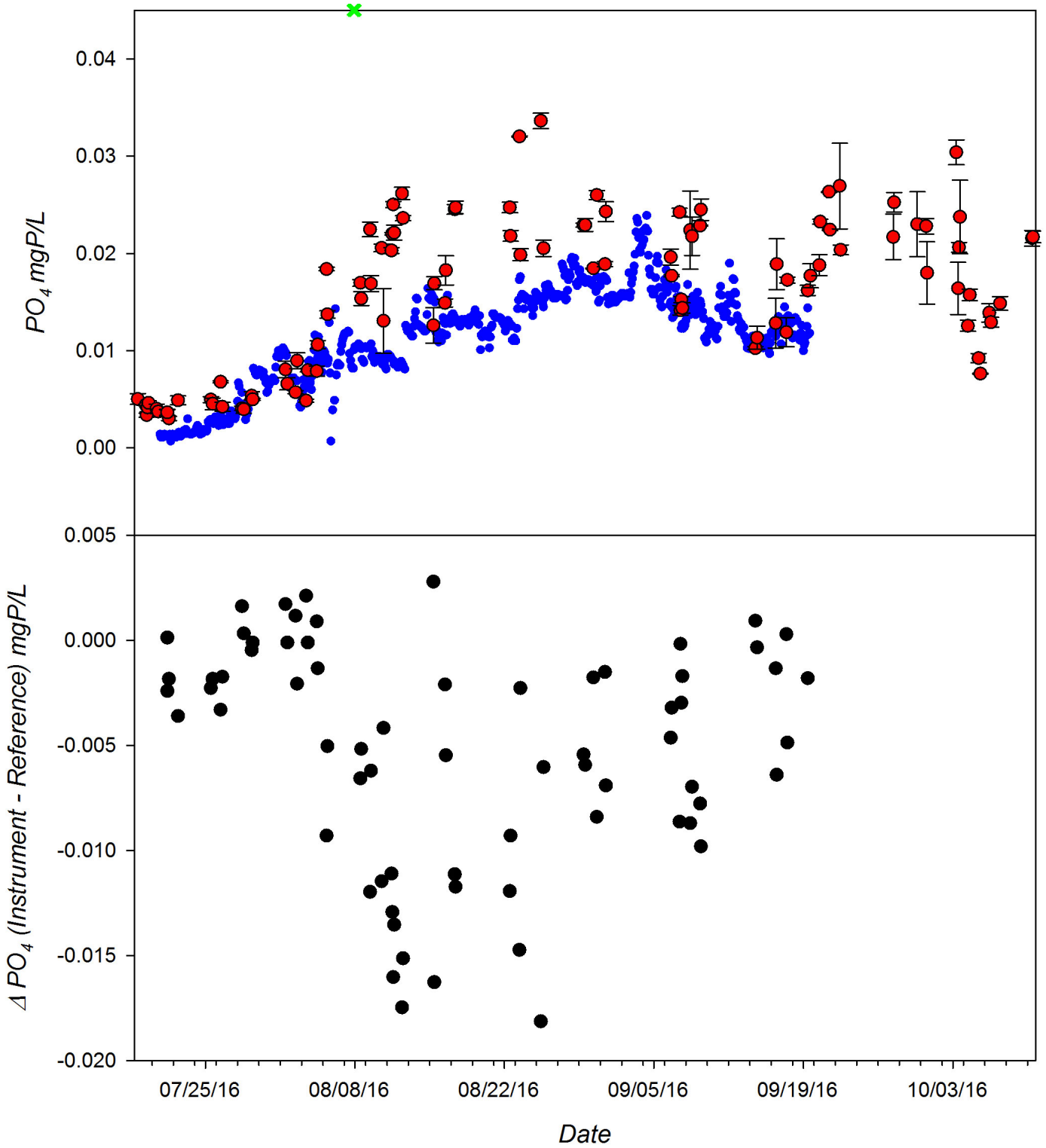


Figure 15. Time series of PO_4 measured by the HydroCycle-PO4 during the 84 day CBL field trial. *Top Panel*: Continuous PO_4 recordings from instrument (blue circles) and PO_4 of adjacent grab samples (red circles). The green crosses at the top of figure represent flagged data (not values) and are plotted on the date of occurrence. *Bottom Panel*: The difference in measured PO_4 relative to reference samples (Instrument mgP/L – Reference mgP/L) observed during deployment.

A cross-plot of the matched observations for the CBL deployment is given in figure 16. A linear regression of the data was significant ($p<0.0001$; $r^2 = 0.56$), but with a slope of only 0.459 and intercept of 0.003. The HydroCycle-PO4 did not accurately measure concentrations above 0.015 mgP/L.

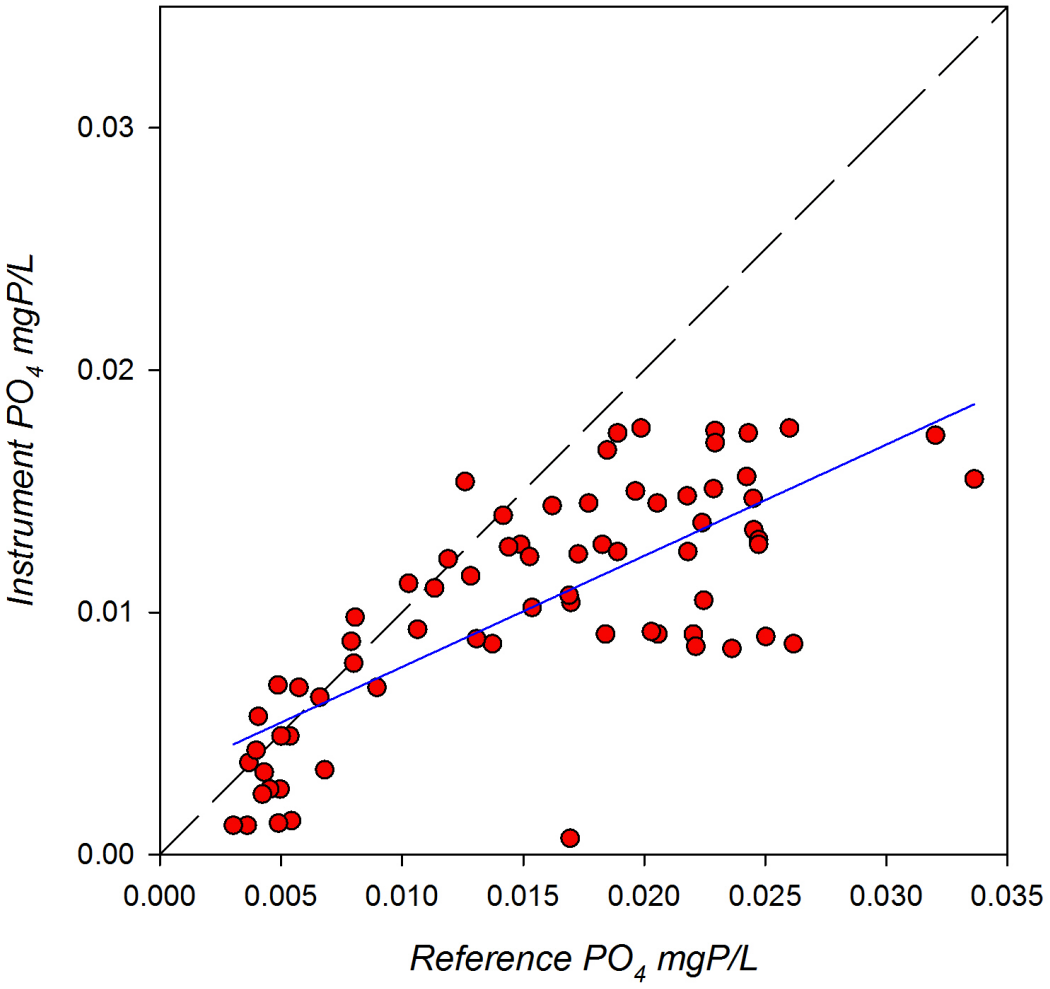


Figure 16. CBL field response plot for HydroCycle-PO4 compared to reference PO_4 samples. The plotted line represents a 1:1 correspondence, the blue line represents the linear regression.

Photographs of the HydroCycle-PO4 before and after the field deployment to indicate potential impact of biofouling (Figure 17).

Figure 17. Photographs of the HydroCycle-PO4 instrument prior to and following the CBL field trial.

Deployment off Coconut Island in Kaneohe Bay, Hawaii

A one month long moored field test was conducted in Kaneohe Bay from October 3, 2016 to November 2, 2016. The deployment site was located at 21.43° N x 157.79° W, on a floating dock anchored off Coconut Island (HIMB) in a depth of approximately 16 meters (Figure 18). 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.

Figure 18. Aerial view of HIMB deployment site (left) and instrument rack in-situ (right).

Time series results of ambient conditions for tidal height, temperature, and salinity are given in figure 19. Temperature at the sensor level ranged from 24.5 to 27.9 °C and salinity from 27.3 to 34.8 PSU over the duration of the field test

The HydroCycle-PO4 operated successfully for the entire 29 days of the deployment, sampling at hourly intervals. Time series results of the HydroCycle-PO4 and corresponding reference PO_4 results are given in figure 20 (top panel). During the deployment the HydroCycle-PO4 returned 720 instrument measurements of a possible 720 measurements for a data completion result of 100%, however one measurement was omitted due to an extreme range (>3X the reference). The range of accepted values reported by the HydroCycle-PO4 analyzer was 0.001 to 0.008 mgP/L, compared to the range within reference samples of 0.0024 to 0.0061 mgP/L. The bottom panel of figure20 presents the time series of the measurement difference between the HydroCycle-PO4 and reference PO_4 for each matched pair. The average and standard deviation of the differences between instrument and reference readings (n=70 out of a possible 71) was -0.0025 ± 0.0012 mgP/L, with a total range in the differences of -0.0048 to -0.0003 mgP/L. There was a small but statistically significant trend in the measurement difference over time ($p<0.0001$; $r^2 = 0.54$) during the deployment, with a slope of -0.0001 mgP/L/d.

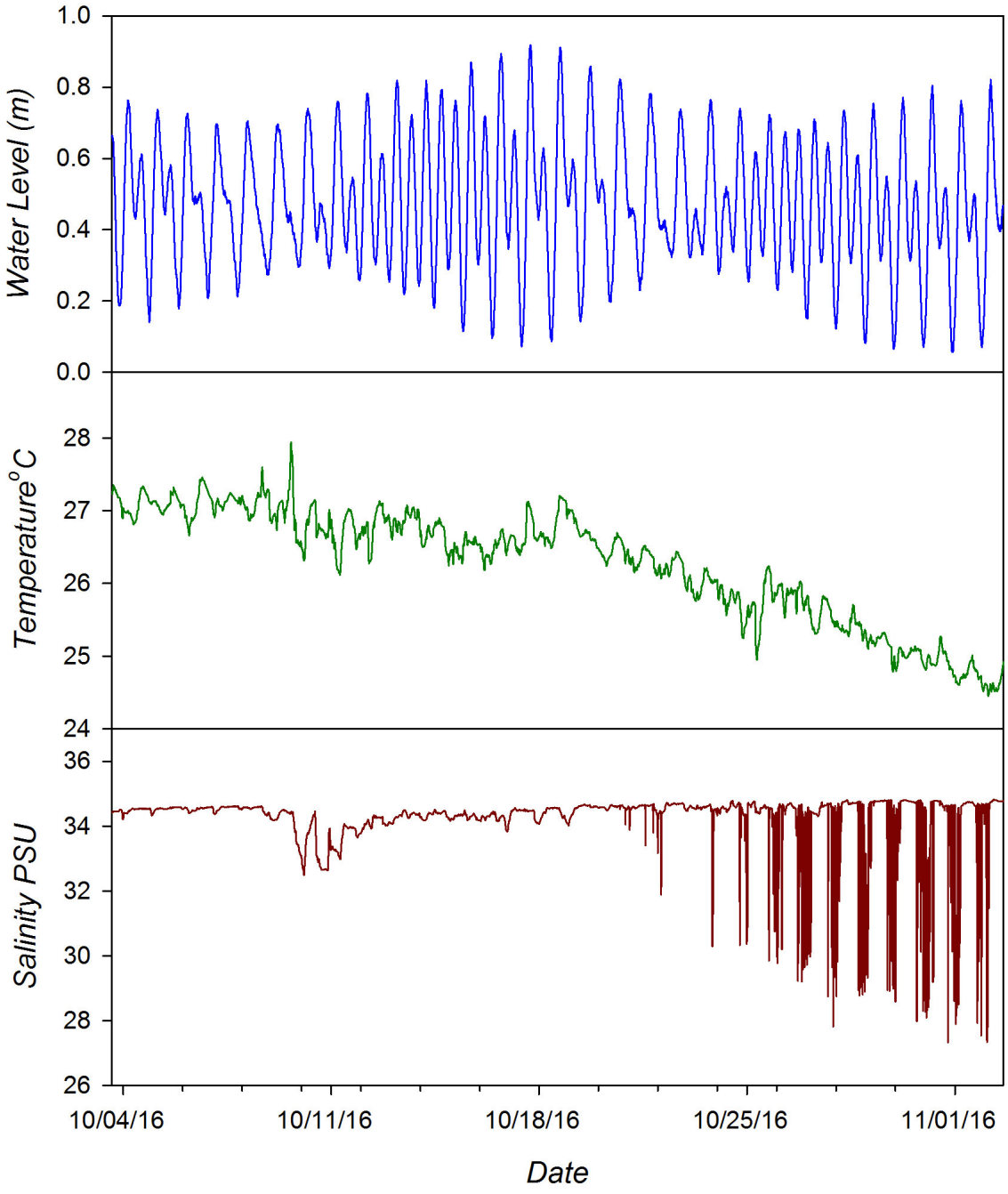


Figure 19. Environmental conditions encountered during the one month HIMB deployment on a floating dock off Coconut Island Test sensor array deployed at 1 m fixed depth, variation in local tidal heights indicate active water flow around instrument (*Top Panel*). Variation in temperature (green) and Salinity (red) at depth of instrument sensor detected by an SBE 26 and two RBR Solo thermistors (*Middle Panel and Bottom Panel*).

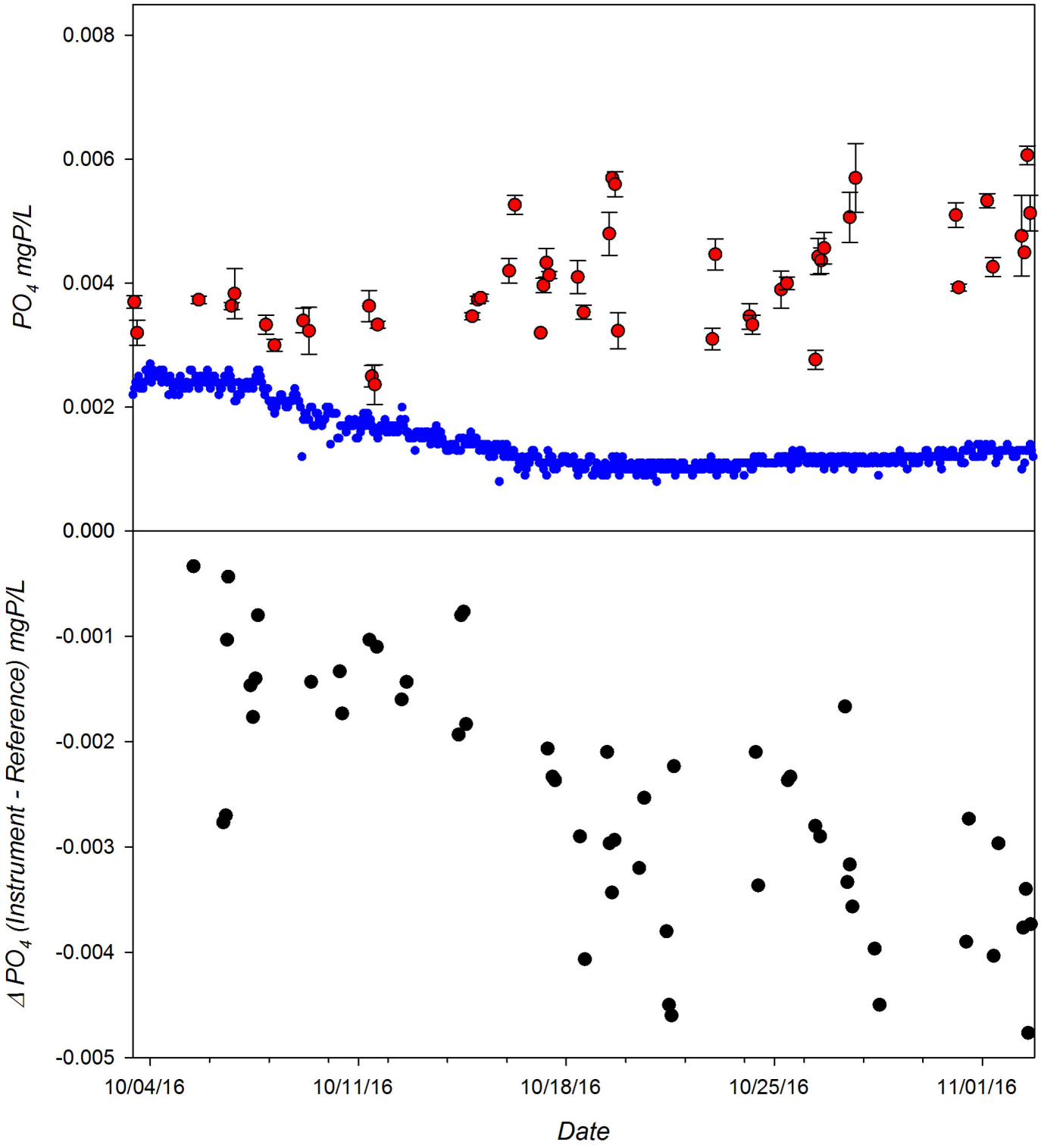


Figure 20. *Top panel:* Time series of PO_4 measured by the HydroCycle-PO4 deployed during the one month HIMB field trial. Continuous PO_4 recordings from instrument (blue dots) and PO_4 of adjacent grab samples (red circles.) *Bottom Panel:* Time series of the difference between the HydroCycle-PO4 and reference measurements for each matched pair (Instrument mgP/L – Reference mgP/L).

A cross-plot of the matched observations for the HIMB deployment is given in figure 21. A linear regression of the data was not significant ($p=0.065$; $r^2 = 0.063$). The Hydrocycle-PO4 could not accurately measure concentrations at ambient range of 0.003 to 0.006 mgP/L during this field test.

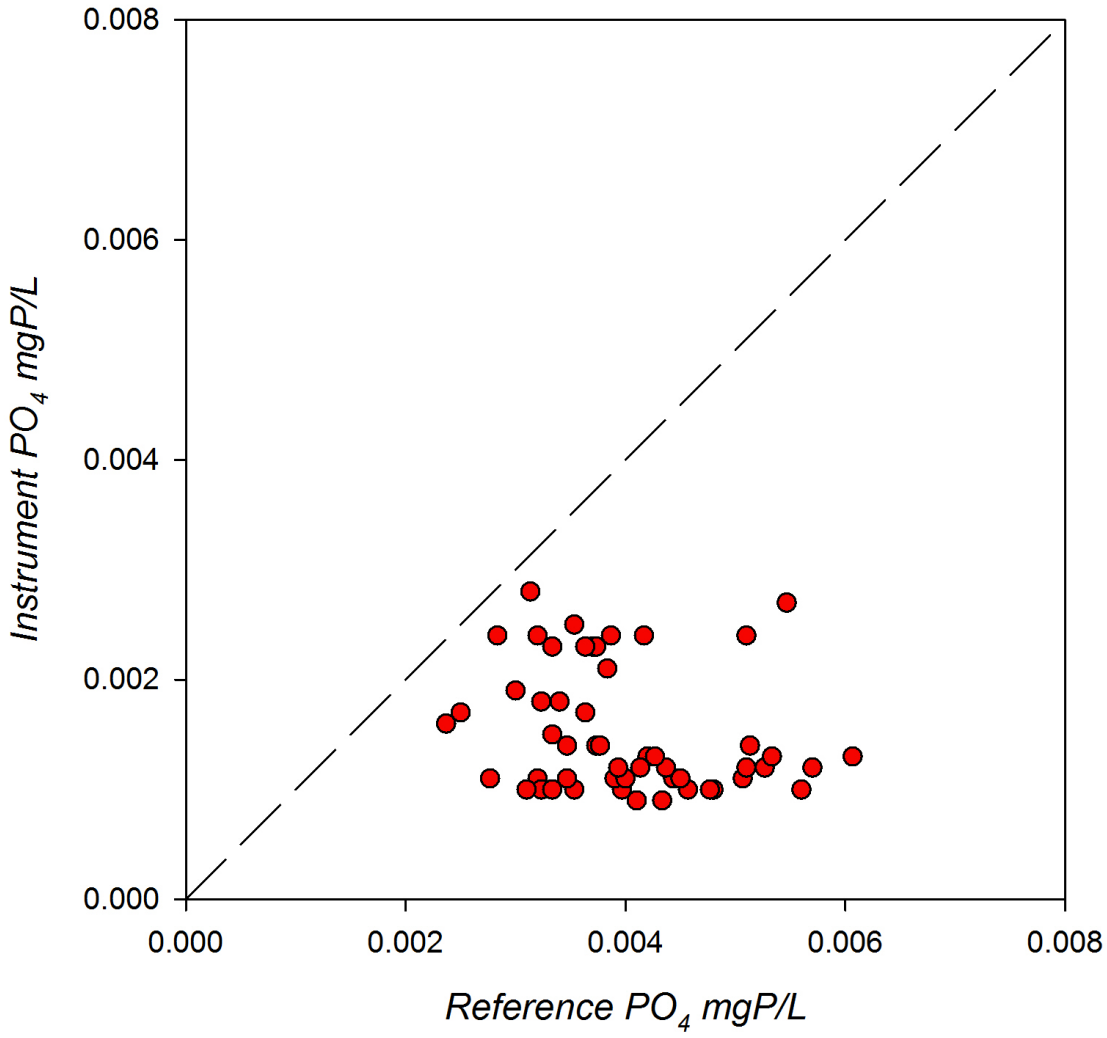


Figure 21. HIMB field response plot of HydroCycle-PO4 compared to reference PO_4 samples. The plotted line represents a 1:1 correspondence.

Photographs of and example of the test instrument prior to deployment and the test instrument after the HIMB field deployment to indicate potential impact of biofouling (Figure 22).

Figure2. Photographs of the HydroCycle-PO4 prior to and following the one month HIMB field trial.

A global summary of instrument versus reference readings for all three field deployment sites are plotted in figure 23. The HydroCycle-PO4 did not accurately measure concentrations above 0.02 mgP/L. A linear regression of all the data was significant ($p=0.029$; $r^2 = 0.03$) but with a meaningless slope of 0.064 and an intercept of 0.004. The data comparison across all three field tests covered a concentration range of 0.001 to 0.120 mgP/L.

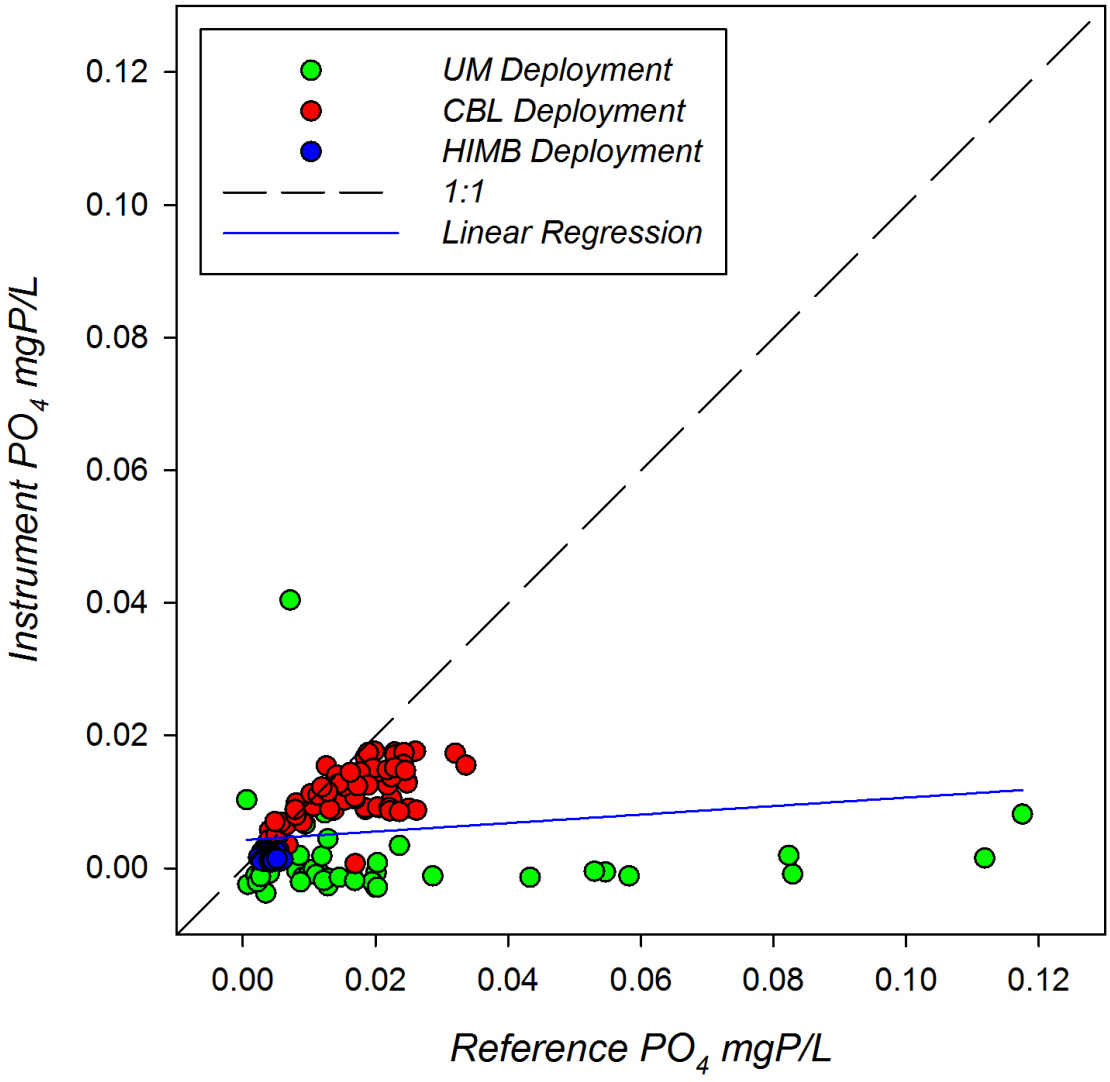


Figure 23. Global response plot for the HydroCycle-PO4 observed during the three ACT field trials. Insert shows the CBL and HIMB deployments enlarged. Black dotted line represents a 1:1 correspondence, the blue line represents the linear regression.

QUALITY ASSURANCE AND QUALITY CONTROL

All technology evaluations conducted by ACT comply with its Quality Management System (QMS), which includes the policies, objectives, procedures, authority, and accountability needed to ensure quality in work processes, products, and services. A QMS provides the framework for quality assurance (QA) functions, which cover planning, implementation, and review of data collection activities and the use of data in decision making, and quality control. The QMS also ensures that all data collection and processing activities are carried out in a consistent manner, to produce data of known and documented quality that can be used with a high degree of certainty by the intended user to support specific decisions or actions regarding technology performance. ACT's QMS meets U.S. Environmental Protection Agency quality standards for environmental data collection, production, and use, and the requirements of ISO/IEC 17025:2005(E), *General requirements for the competence of testing and calibration laboratories*.

An effective assessment program is an integral part of a quality system. The ACT Quality Assurance (QA) Manager independently conducted Technical Systems Audits (TSA) of field tests at Maumee River field trial during May 25-28, 2016, a TSA of the Laboratory test at the Chesapeake Biological Laboratory during July 10-18, 2016 and a data quality review of the reference data sets from all tests conducted during the Nutrient Challenge.

Technical System Audits

A TSA is a thorough, systematic, on-site qualitative audit of sampling and measurement processes and procedures associated with a specific technology evaluation. The objectives of the TSAs conducted during this evaluation were to assess and document the conformance of on-site testing procedures with the requirements of the Test Protocols, the ACT Quality Assurance Project Plan (QAPP), and associated Standard Operating Procedures (SOPs).

The TSA was conducted in accordance with the procedures described in n EPA's *Guidance on Technical Audits and Related Assessments for Environmental Data Operations (EPA QA/G-7)* and ISO 19011, *Guidelines for Quality and/or Environmental Management Systems Auditing*. A TSA checklist based on the Test Protocols was prepared prior to the audits and reviewed by the ACT Director and Senior Scientist. The TSA assessed ACT personnel, the test and analytical facilities, equipment maintenance and calibration procedures, sample collection, analytical activities, record keeping, and QC procedures. Reference sample handling and chain-of-custody by NASL were observed during the laboratory test at CBL.

During the audits, the QA Manager met with ACT technical staff involved in the evaluation and asked them to describe the procedures followed. All procedures were observed; and logbooks, data forms, and other records were reviewed.

Key components of the audit included:

- Assessment of Quality Assurance/Quality Control:
  - Adequacy of procedures, and
  - Adherence to procedures.
- Assessment of Sample System:
  - Sample collection,
  - Analytical procedures, and

  - Documentation.
- Assessment of Data and Document Control:
  - Chain of custody, and
  - Documentation.

The TSAs' findings were positive. The field and laboratory tests were implemented consistent with the Test Protocols, QAPP, and SOPs. Minor deviations were documented in laboratory records. There were no deviations which may have had an effect on data quality for the test. All phases of the implementation of the tests reviewed during the audits 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 evaluation and demonstrated expertise in pertinent procedures. Communication and coordination among all personnel was frequent and effective. Internal record keeping and document control was 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 were collected as described in the Test Protocols and SOPs. Examination of maintenance and calibration logs provided evidence of recent and suitable calibration of sampling and analytical equipment.

Data Quality

Data Verification, Validation, and Assessment.

Data review is conducted to ensure that only sound data that are of known and documented quality and meet technology evaluation quality objectives are used in making decisions about technology performance. Data review processes are based in part on two EPA guidance documents: *Guidance on Environmental Data Verification and Data Validation* (QA/G-8) [EPA, 2002] and *Guidance on Technical Audits and Related Assessments for Environmental Data Operations* (QA/G-7) [EPA, 2000].

The data were verified and validated to evaluate whether the data have been generated according to the Test Protocols and satisfied acceptance criteria. Data verification evaluates the completeness, correctness, and consistency of the data sets against the requirements specified in the Test Protocols, measurement quality objectives (MQOs), and any other analytical process requirements contained in SOPs.

The ACT QA Manager reviewed the reference data sets from all field and laboratory tests. The number of reference samples collected at each site and the laboratory tests are in Table 10. A total of 346 reference samples were collected for the field and laboratory tests. The overall reference data set included 3,666 distinct analyses.

Table 10. The number of reference samples collected during the laboratory test and at each field site.

Site	No. of Samples	No. of Replicates per Sample	No. of Analytes[1/] Measured in Each Replicate	No. of Measurements
Maumee River	61	3	3	549
CBL – Field	120	3	3	1080
CBL – Lab	92	5	3	1380
Hawaii	73	3	3	657
Total	346			3,666
[1/] NO_2; NO23; PO_4				

The data review verified that the sampling and analysis protocols specified in the Test Protocols were followed, and that the ACT measurement and analytical systems performed in accordance with approved methods, based on:

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

Data validation uses the outputs from data verification and included inspection of the verified field and laboratory data to determine the analytical quality of the data set. A representative set of approximately 10% of the reference data was traced in detail from 1) raw data from field and laboratory logs, 2) data transcription, 3) data reduction and calculations, to 4) final reported data. Validation of the data sets established:

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

The data validation also confirmed that the data were accumulated, transferred, summarized, and reported correctly. There is sufficient documentation of all procedures used in the data collection and analysis to validate that the data were collected in accordance with the evaluation's quality objectives.

A Data Quality Assessment (DQA) is the third and final process of the overall data assessment. It is a scientific and statistical evaluation of validated data to determine if the data are of the right type, quality, and quantity to support conclusions on the performance of the technologies. The DQA determined that the test's data quality objectives, described in Section 7.1 of the Test Protocols and Section 3.4 and Appendix B of the ACT QAPP (ACT, 2016), were achieved. This evidence supports conclusions that:

- The sampling design performed very well and is very robust with respect to changing conditions.
- Sufficient samples were taken to enable the reviewer to see an effect if it were present.

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 to collect, analyze, interpret, and report data as specified in the Test Protocols, QAPP, and SOPs. The ADQ determined 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.

Table 11. Results of Field Duplicates (FD) for the Maumee River mooring test.

Date/Time	Rep	PO_4	Mean	Std Dev	ABS Diff	CV%
6-16-16 9:00	FD1	0.0094	0.008	0.0016	0.002	19.2
	FD2	0.0072				
6-17-16 12:00	FD1	0.0071	0.008	0.0005	0.001	7.11
	FD2	0.0079				
6-20-16 10:00	FD1	0.0122	0.012	0.0003	0.0005	2.59
	FD2	0.0117				
6-23-16 11:00	FD1	0.0203	0.020	0.0001	0.0001	0.489
	FD2	0.0202				

Table 12. Results of Field Duplicates (FD) for the Chesapeake Bay, MD mooring test.

Date/Time	Rep	PO_4	Mean	Std Dev	ABS Diff	CV%
7-20-16 10:00	FD1	0.0040	0.0039	0.0002	0.0003	6.1
	FD2	0.0037				
7-26-16 14:00	FD1	0.0042	0.0039	0.0005	0.0007	12.09
	FD2	0.0036				
8-2-16 10:00	FD1	0.0057	0.0057	0.0000	0.0001	0.827
	FD2	0.0057				
8-10-16 16:00	FD1	0.0131	0.0148	0.0024	0.0034	16.14

	FD2	0.0164				
8-23-16 12:00	FD1	0.0199	0.0197	0.0003	0.0004	1.56
	FD2	0.0194				
9-8-16 10:00	FD1	0.0224	0.0249	0.0035	0.0050	14.11
	FD2	0.0274				
9-16-16 12:00	FD1	0.0189	0.0195	0.0008	0.0011	4.00
	FD2	0.0200				
10-4-16 14:00	FD1	0.0157	0.0144	0.0019	0.0027	13.27
	FD2	0.0130				
10-10-16 10:00	FD1	0.0216	0.0221	0.0007	0.0010	3.10
	FD2	0.0225				

Table 13. Results of Field Duplicates (FD) for the Kaneohe Bay, HI mooring test

Date/Time	**Rep**	**NO_3**	**Mean**	**Std Dev**	**ABS Diff**	**CV %**
10-6-16 14:00	FD1	0.0035	0.0036	0.000	0.0001	.664
	FD2	0.0036				
10-12-16 11:00	FD1	0.0034	0.0033	0.0002	0.0003	5.77
	FD2	0.0031				
10-17-16 9:00	FD1	0.0035	0.0034	0.0001	0.0001	2.07
	FD2	0.0034				
10-26-16 9:00	FD1	0.0039	0.0040	0.0002	0.0003	4.68
	FD2	0.0042				
11-1-16 9:00	FD1	0.0053	0.0053	0.0001	0.0002	2.25
	FD2	0.0052				

Table 14. Results of Field Trip Blanks all deployments.

Maumee River		**Chesapeake Bay**		**Kaneohe Bay**	
Field Blank ID	PO_4 (Std Dev)	Field Blank ID	PO_4 (Std Dev)	Field Blank ID	PO_4 (Std Dev)
GLFB1	0.0008 (0.0001)	CBLFB1	0.0027 (0.0001)	HIFB1	0.0017 (0.0000)
GLFB2	0.0012 (0.0003)	CBLFB2	0.0026 (0.0001)	HIFB2	0.0016 (0.0002)
GLFB3	0.0021 (0.0001)	CBLFB3	0.0014 (0.0001)	HIFB3	0.0013 (0.0002)
GLFB4	0.0027 (0.0004)	CBLFB4	0.0011 (0.0003)	HIFB4	0.0013 (0.0002)
--	--	--	--	HIFB5	0.0010 (0.0001)
Mean (Std Dev)	0.0017 (0.001)	Mean (Std Dev)	0.002 (0.001)	Mean (Std Dev)	0.001 (0.0003)
Grand Mean (Std Dev)					0.002 (0.0007)

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. Suzanne Bricker, National Oceanic and Atmospheric Administration, Dr. Brian Pellerin, U.S. Geological Survey, Dr. Dwane Young, U.S. Environmental Protection Agency, Dr. Matt Cohen, University of Florida, Dr. R. David Holbrook, National Institute for Standards and Technology, Mr. Chris Gross, U.S. Department of Agriculture NRCS, Dr. Joe Rudek, Environmental Defense Fund for their advice and direct participation in various aspects of this evaluation. Earle 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.

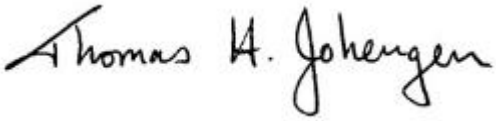
June 1, 2017

Date

Approved By: Dr. Mario Tamburri
ACT Executive Director

June 1, 2017

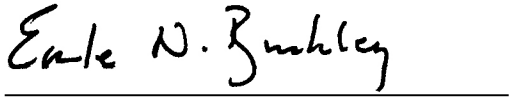
Date



Approved By: Dr. Tom Johengen
ACT Chief Scientist

June 1, 2017

Date



Approved By: Dr. Earle Buckley
Quality Assurance Supervisor

HydroCycle-PO4 ACT Evaluation Seabird Scientific Comments:

Sea-Bird Scientific would like to thank the ACT and EPA Nutrient Sensor Challenge teams that worked to make the field and lab evaluations possible. Below are comments in reference to specific elements that we feel warrant further explanation. You may contact Sea-Bird Scientific directly in you have specific questions regarding this report. Please call Sea-Bird Scientific at 1-541-929-5650 and ask for Corey Koch, HydroCycle-PO4 product manager.

It is useful to note a difference in the HydroCycle-PO4 and EPA method 365.1 arises from a filter pore size of 5-10 um and performing the reaction immediately on the sample. In some environments and sampling conditions, we expect these method differences to results in differences from grab samples. The real-time quality control flags referred to in the ACT evaluations are useful in understanding the results. In addition to the overall flags of good, bad, and suspect, flags cover out-of-range, calibration spike and mixing issues, bubbles, low signal, and high noise. The trends of data flags going from good to suspect often predict subsequent bad data, which permits users to intervene when an error condition starts to occur.

Sea-Bird recognizes the importance and value of 3rd party testing and evaluation and strongly supports the ACT program. The ACT protocols test instruments with respect to basic operation and more importantly how they work in the field. For long term monitoring programs the initial cost of an instrument is quickly dwarfed by the operational and maintenance costs, particularly in remote and/or high fouling environments. It is paramount to report good data for as long as possible, to have the ability to easily identify and remove bad data, and to know when data is suspect to minimize downtime.

Corey Koch
Product Manager
ckoch@seabird.com