Ref. No. [UMCES] CBL 2016-010
ACT VS16-01

PERFORMANCE VERIFICATION STATEMENT FOR JFE AROUSB AND AROW-USB Dissolved Oxygen Sensors

TECHNOLOGY TYPE: Dissolved Oxygen sensors

APPLICATION: In situ estimates of DO 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 January 2015 to January 2016

EVALUATION PERSONNEL: T. Johengen, G.J. Smith, D. Schar, H. Purcell, D. Loewensteiner, Z. Epperson, and M. Tamburri, G. Meadows, S. Green, F. Yousef, J. Anderson.

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
CILER- 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 dissolved oxygen sensors during 2015-2016 to characterize performance measures of accuracy and reliability in a series of controlled laboratory studies and field mooring tests in diverse coastal environments. The verification included several months of Laboratory testing along with three field deployments covering freshwater, estuarine, and oceanic environments. Laboratory tests of accuracy, precision, response time, and stability were conducted at Moss Landing Marine Lab. A series of nine accuracy and precision tests were conducted at three fixed salinity levels (0, 10, 35) at each of three fixed temperatures (4, 15, 30 ^{o}C). A laboratory based stability test was conducted over 56 days using deionized water to examine performance consistency without active biofouling. A response test was conducted to examine equilibration times across an oxygen gradient of 8mg/L at a constant temperature of 15 ^{o}C. Three field-mooring tests were conducted to examine the ability of test instruments to consistently track natural changes in dissolved oxygen over extended deployments of 12-16 weeks. Deployments were conducted at: (1) Lake Superior, Houghton, MI from 9Jan – 22Apr, (2) Chesapeake Bay, Solomons, MD from 20May – 5Aug, and (3) Kaneohe Bay, Kaneohe, HI from 24Sep – 21Jan. Instrument performance was evaluated against reference samples collected and analyzed on site by ACT staff using Winkler titrations following the methods of Carignan et al. 1998. A total of 725 reference samples were collected during the laboratory tests and between 118 – 142 reference samples were collected for each mooring test. This document presents the results of two different models of the JFE Advantech RINKO optical dissolved oxygen sensors (AroUSB and AroW-USB). Both models were tested in all Laboratory trials and the fast-response AroUSB was used in the field profiling application, while the wiper based AroW-USB was used in the extended field mooring applications.

Instrument accuracy and precision for the AroUSB and AroW-USB sensors were tested under nine combinations of temperature and salinity over a range of DO concentrations from 10% to 120% of saturation. The laboratory testing set-up did result in bubbles from the sparging gases used to change DO levels occasionally becoming trapped on the sensor foil and those data where noted were excluded from any comparisons to reference samples. The means of the difference between the AroUSB and reference measurement for the nine trials ranged from -0.277 to 0.265 mg/L. A linear regression of the accepted data (n=377; $r^2 = 0.965$; $p<0.0001$) produced a slope of 1.015 and intercept of 0.098. For the AroUSB, the absolute precision, estimated as the standard deviation (s.d.) around the mean, ranged from 0.002 – 0.014 mg/L across trials with an overall average of 0.004 mg/L. Relative precision, estimated as the coefficient of variation (CV% = (s.d./mean)x100), ranged from 0.013 – 0.278 percent across trials with an overall average of 0.058%. The means of the difference between the AroW-USB and reference measurements ranged from -0.277 to 0.134 mg/L across all trials. A linear regression of the accepted data (n=257; $r^2 = 0.976$; $p<0.0001$) produced a slope of 0.969 and intercept of 0.114. The absolute precision for the AroW-USB were ranged from 0.001 – 0.012 across trials, with an overall average of 0.004 and the relative precision ranged from 0.017 – 0.247 percent across trials with an overall average of 0.051%.

For the 56 day lab stability test, the overall mean of the differences between AroUSB and reference measurements was 0.001 (± 0.326) mg/L. There was no significant trend in accuracy over time (slope = -0.0007 mg/L/d) that would indicate any type of performance drift over the duration. The overall mean of the differences between AroW-USB and reference measurements was -0.154 (± 0.319) mg/L. There was a minor drift in instrument accuracy over the deployment (slope = -0.006 mg/L/d; r^2=0.17) but the goodness of fit was low due to several outliers.

For the lab-based functional response time assessment, the calculated τ_{90} for the AroUSB was 11.8 s during high to low transitions and 7.1 s for low to high transitions covering a DO range of approximately 8 mg/L at a constant 15 °C. However, as noted in the report we incorrectly programmed the sampling rate to 10 seconds which would have a direct impact on the calculated response rate. For the AroW-USB the calculated τ_{90} was 209 s during high to low transitions and 284 s for low to high transitions for the same conditions.

At Houghton, MI the field test was conducted under the ice over 104 days with a mean temperature and salinity of 0.7 °C and 0.01. The measured DO range from our 118 discrete reference samples was 10.25 – 14.01 mg/L compared to a range of 8.669 – 15.076 mg/L reported by the AroW-USB over its 9859 observations conducted continuously at 15 minute intervals. The useable data return for the deployment was 100%. The average and standard deviation of the measurement difference between the AroW-USB and reference samples over the total deployment was 0.170 ±0.057 mg/L with a total range of 0.055 to 0.309 mg/L. A drift rate in instrument response, estimated by linear regression ($r^2=0.325$, $p<0.001$) of the difference across time, was -0.001 mg/L/d but directionally getting closer to the Winkler reference values.

At the Chesapeake Biological Lab, the field test was conducted over 78 days with a mean temperature and salinity of 25.6 °C and 10.9. The measured DO range from our 142 discrete reference samples was 4.370 – 10.858mg/L compared to a range of 2.610 – 14.510 mg/L reported by the AroW-USB over its 7270 continuous observations conducted at 15 minute intervals. The data completion rate for this deployment was 100%. The average and standard deviation of the measurement difference between the AroW-USB and reference samples over the total deployment was -0.056 ±0.131 mg/L with a total range of -0.375 to 0.392 mg/L. There was minor trend in response accuracy over the deployment (slope = -0.002 mg/L/d; $r^2 = 0.16$) but with a low predictive fit.

At Kaneohe Bay, HI the field test was conducted over 121 days with a mean temperature and salinity of 25.8 °C and 33.4. The measured DO range from our 129 discrete reference samples was 3.63 – 9.85 mg/L compared to a range of 2.329 – 10.996 mg/L reported by the AroW-USB. Fourteen percent (785 of 5653) of the continuous 30 minute observations fell more than 2 mg/L outside of a natural ambient range as determined by the pattern of Winkler reference samples and were excluded from statistical comparisons. For the accepted data (n=75 of a potential 129 comparisons), the average and standard deviation of the measurement difference between the AroW-USB and reference samples over the total deployment was 0.367 ±0.637 mg/L with a total range of -0.720 to 1.991 mg/L. The drift rate in the instrument offset based on linear regression ($r^2 = 0.74$) was 0.165 mg/L/d throughout the deployment period.

Overall, the response of the AroW-USB during field testing showed good linearity across all three salinity ranges including freshwater, brackish water, and oceanic water. The accuracy of the response curve was quite consistent across the concentration ranges observed within each test site and relatively consistent over the wide range of DO conditions (4 - 14 mg/L) across sites.

The Aro-USB was evaluated in a profiling field test in the Great Lakes at two separate locations in order to experience transitions from surface waters into both normoxic and hypoxic hypolimnion. In Muskegon Lake, the temperature ranged from 21.0 °C at the surface to 13.5 °C in the hypolimnion, with corresponding DO concentrations of 7.8 and 2.8 mg/L, respectively. In Lake Michigan, the temperature ranged from 21.0 °C at the surface to 4.1 °C in the hypolimnion, with corresponding DO concentrations of 8.6 and 12.6 mg/L, respectively. Two profiling trials were conducted at each location. The first trial involved equilibrating test instruments at the surface (3m) for ten minutes and then collecting three Niskin bottle samples at one minute intervals. Following the third sample, the rosette was quickly profiled into the hypolimnion were samples were collected immediately upon arrival and then each minute for the next 6 minutes. The second trial was performed in the reverse direction. Note for Muskegon Lake cast 1 was aborted due to bottle misfires and repeated as cast 3. In Muskegon Lake, the Aro-USB exhibited a negative bias in the colder, low DO hypolimnion and a positive bias in the warm, high DO surface. Sensor equilibration time was slightly greater going from surface to hypolimnetic conditions. The range in measurement differences between instrument and reference was -0.42 to 0.34 mg/L for cast 2 and -0.75 to 0.27 mg/L for cast 3.

In Lake Michigan, the Aro-USB exhibited a positive bias in both portions of the water column but the magnitude was higher in the cold high DO hypolimnion. Sensor equilibration time was similar between both trials, whether equilibrated at surface or depth. The range in measurement differences between instrument and reference was -0.16 to 0.53 mg/L for cast 1 and 0.18 to 0.50 mg/L for cast 2.

BACKGROUND AND OBJECTIVES

Instrument performance verification is necessary so that effective existing technologies can be recognized and 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 dissolved oxygen (DO) 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. To this end, the data and information on performance characteristics were focused on the types of information users most need.

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

INSTRUMENT TECHNOLOGY TESTED

For this performance evaluation both the JFE Advantech RINKO AroUSB and AroW-USB were tested during all laboratory trials. The AroUSB was additionally evaluated in the profiling field test in Michigan, and the AroW-USB was evaluated in the three field deployment tests at Michigan, Maryland, and Hawaii. JFE Advantech RINKO AroUSB and AroW-USB are high-accuracy and high-resolution optical dissolved oxygen (DO) sensors based on the phosphorescence principle. The sensors are coated with photo-stimulable phosphor (PSP) on the outside of the pressure-resistant acrylic optical window, measuring a phosphorescence quenching phase shift. The excitation blue LED pulse generates a red phosphorescence pulse, which in turn has an inverse correlation with the oxygen partial pressure in water (DO concentration). Since the method does not consume oxygen, there is no need for stirring. The instruments offer a user friendly 2-point calibration that compensates the sensing foil time-drift and helps ensure reliable and accurate data. The method consists of simply recalibrating the DO sensor at two saturation points (0% and 100%) using a Na2SO3 aqueous solution (0% saturation) and air saturated water (100% saturation). Depending on the application, RINKO sensors are offered in different models and among them there are:

RINKO AroUSB is an autonomously deployable time data logger with temperature and fast response optical DO sensor. The instrument has various operating modes, offering flexibility when carrying out observations. The compact size containing the data logger allows for being easily integrated on different platforms (e.g. CTDs). Compared to galvanic, clark-cell and optical DO sensors, RINKO I has the fastest response time (< 1 s with 90% response), allowing for measurements of dissolved oxygen concentration at high resolution (at sub-meter scale) and enabling continuous profiling at 0.5 m s^{-1}.

RINKO AroW-USB is an autonomously deployable/real time data logger with temperature and optical DO sensor. The instrument has a mechanical wiper in order to protect the sensing foil against accumulating bio-fouling, and therefore, allowing for long-term observations without affecting data quality.

PERFORMANCE EVALUATION TEST PLAN

Laboratory Tests

Laboratory tests of accuracy, precision, response time, and stability were conducted at Moss Landing Marine Lab. All tests were run under ambient pressure (logged hourly from a barometer at the laboratory) and involved the comparison of dissolved oxygen concentration reported by the instrument versus Winkler titration values of water samples taken from the test baths. All tests were run in thermally controlled tanks at specific temperature, salinity, and DO concentrations. Tanks were well mixed with four submersible Aquatic Ecosystem Model 5 pumps with flow rates of 25 L/min. Temperatures were controlled to within approximately 0.2°C of set point using Thermo Digital One Neslab RTE 17 circulating thermostats flowing through closed coils distributed within the tank. Four RBR temperature loggers were deployed within the tank to verify actual temperature to better than 0.02°C. Salinity was varied by addition of commercial salts (Instant Ocean) to Type 1 deionized water. Salinity was verified at the beginning and end of each test condition by analysis on a calibrated CTD. Dissolved oxygen concentrations were controlled by use of compressed gases of known oxygen concentration sparging through diffusers within the tank. Tanks were covered with a layer of floating closed-cell plastic insulation that continuously sealed the water surface and to minimize atmospheric exchange. If required by the manufacturer, instruments were only calibrated prior to the start of the first lab test, and then again prior to the stability test which began one month later. The following series of tests were conducted in the laboratory trials:

Accuracy at various T/S and DO conditions

A series of measurements were conducted under 36 discrete conditions to target 3 temperatures, 3 salinities, and at least 4 DO concentrations as follows:

- Temperature Conditions: 5, 15, 30 °C
- 3 Salinity Conditions: 0, 10, 34
- Dissolved oxygen,(% air saturation): 0% (hypoxic), 20 – 30%, 100% and $\geq$120%, (levels were achieved by mixing pure O_2 and N_2 sources with pure N_2 was used for the 0% O_2 concentration)

Tests were run such that all 4-6 DO concentrations were tested for a fixed temperature and salinity on the same day. The tests began at ambient, near air saturation, conditions following overnight equilibration of tank water to the test salinity and temperature. Subsequently DO were dropped to near 0 mg/L and increased stepwise to the highest concentration. Instruments were allowed to equilibrate at the fixed temperature and salinity for 1 h before the start of that day's trial. Sparging with each DO gas concentration was conducted for a minimum of 60 minutes prior to the start of data collection and reference sampling. For each test condition, the test instruments were programmed to sample at no slower than 1 minute intervals and reference samples were collected at 6 timepoints spaced 5 minutes apart for each of the fixed conditions. For three of the timepoints duplicate samples were collected from two different sampling ports mounted at opposite ends of the tank to access heterogeneity within the tank. Inlets of sampling ports were positioned at the depth of the sensor heads (ca. -0.5m). All reference samples were collected while the gas sparging was off and took approximately 1 minute to complete. Reference samples were processed and analyzed as defined below. The order of the test conditions were 15 then 5 then 30 ^{o}C, going from 0 then 10 then 34 salinity at each temperature.

Precision Test at various DO concentrations

Instrument precision was evaluated under stable conditions generally achieved at the start of each trial's day. Instruments were equilibrated to each test condition for a minimum of one hour prior to testing. The sampling frequency for test instruments was 1 minute with reference samples matching instrument sampling to monitor for drift in tank DO. At least 6 reference samples were collected over a 30 minute instrument precision evaluation trial. Reference samples were processed and analyzed as defined below.

Functional Response Time Test

A response time test was conducted by examining measurements during a rapid exchange across a large gradient in dissolved oxygen for a fixed temperature (15 ^{o}C) in deionized water, following the approach described in Bittig et al. 2014. The reservoirs of the thermostat baths were constantly bubbled with either N_2 gas or air to maintain discrete DO levels. A submersible pump was added to each bath to ensure uniform flow and oxygen conditions and instruments were mounted at a fixed position within the baths to minimize variance due to instrument manipulation. Instruments were programmed to measure every 10s continuously for minutes following the exchange (Note: instruments were mistakenly not programmed to measure at their highest frequency and sampling rate will affect the calculated response time). For instruments with the capability, real-time monitoring of instrument output was monitored to verify a steady state reading had been obtained. Instruments were moved from the high DO concentration to the low DO concentration and subsequently reversed to check for response hysteresis. During transitions, care was taken to minimize carryover by shaking off residual water. The sensor was then carefully inserted into the new bucket and mixed by hand to ensure no bubble entrapment and full exposure to the new solution. Reference samples from each reservoir were taken at the beginning and end of the exposure. The test instrument was equilibrated in the high DO reservoir for at least 30 min prior to the exchange to ensure temperature equilibration.

Lab-based Stability Test

A laboratory stability test was conducted to examine potential instrument drift in a non-biofouling environment. These results are contrasted to the stability of measurement accuracy observed in the long-term field mooring deployments. The test occurred over 56 days, with daily temperature fluctuations of approximately 10°C, achieved by alternating the set point of the recirculation chiller. Reference samples were collected at minimum and maximum temperatures at least 3 times per week. The test was conducted in deionized water at saturated air conditions. Tanks were well circulated and open to the atmosphere. Water in the test tank was exchanged as needed if there was any indication of biological growth. Instruments stayed continuously submerged and were not exposed to air during any water exchange. The goal of comparisons of accuracy over time between the field and a sensor deployed similarly in the laboratory is intended to provide insight into drift and reliability intrinsic to the instrument relative to changes that may result from biofouling.

Moored Field Tests

Field Deployment Sites and Conditions

A four month moored deployment was conducted at Michigan Technological University's Great Lakes Research Center dock in Houghton, MI. Instruments were deployed in January and kept under ice cover until April. Instruments were programmed to sample at a minimum frequency of once per hour. ACT collected reference samples twice per day for 4 days per week during the entire deployment. Instruments were moored at approximately 4m depth and surface access through the ice was maintained by gentle circulation with a propeller to allow deployment of the Van Dorn sampling bottle. The goal of this test application was to demonstrate instrument performance (reliability, accuracy, and stability) in winter-time environmental conditions and to demonstrate the ability to operate continuous observations under ice.

A three month moored deployment was conducted at the Chesapeake Biological Lab Pier, Solomons, MD. Instruments were deployed between May and August during a period of warming temperatures and high biological production. Instruments were moored at fixed depth of 1m on a floating dock. Instruments were programmed to sample at a minimum frequency of once per hour. ACT collected reference samples twice per day for 3 days per week and collected six samples on one day per week during the entire deployment. The intensive sampling was spaced to capture the maximum range of expected diurnal variation in dissolved oxygen concentrations. The goal of this test application was to demonstrate instrument performance (reliability, accuracy, and stability) under high biofouling conditions and over a range of salinity and temperature conditions in a coastal estuarine environment.

A four month moored deployment was conducted in a shore patch reef at the Hawaii Institute of Marine Biology (HIMB), Coconut Island, Kaneohe, HI. Instruments were deployed between September and January. Instruments were moored at approximately 1m depth on a bottom mounted PVC rack and were programmed to sample at a minimum frequency of once per hour. Some manufacturers chose to sample more frequently to demonstrate that capability. ACT collected reference samples twice per day for 3 days per week and collected six samples on one day per week during the entire deployment. The intensive sampling was spaced to capture the maximum range of expected diurnal variation in dissolved oxygen concentrations. The goal of this test application was to demonstrate instrument performance (reliability, accuracy, and stability) under high biofouling conditions in warm, full salinity coastal ocean conditions.

Field Testing Procedures

The moored deployments were run sequentially, and instrument packages were returned to manufacturers for reconditioning and calibration in between each successive field test.
Prior to each deployment, instruments were set-up and calibrated if required, as directed by the manufacturer and demonstrated at a prior training workshop. Sensors were programmed to record dissolved oxygen data at a minimum of once per hour at the top of the hour for the duration of the planned deployment. All instrument internal clocks were set to local time and updated before programming using www.time.gov as the time standard. A photograph of each individual sensor and the entire sensor rack was taken just prior to deployment and just after recovery to provide a qualitative estimate of biofouling during the field tests. In the final step before deployment, instruments were placed in a well aerated fresh water bath, with a known temperature, for 45 minutes and allowed to record three data points as a baseline reference. Reference samples were drawn at the corresponding sampling times and analyzed for dissolved oxygen using Winkler titration method described below.

All instrument packages were deployed on a single box shaped rack that allowed all sensor heads to be at the same depth, with instruments side by side and all sensor heads deployed at the closest proximity feasible. The rack was deployed so that all sensor heads remained at a fixed depth of 1 m below the water surface. A standard and calibrated CTD package was deployed at each test site and programmed to provide an independent record of conductivity and temperature at the sensor rack during each instrument sampling event. At least four additional RBR temperature loggers were placed on the rack to capture any spatial variation in the temperature across the rack.

A standard 4 L Van Dorn bottle was used at each test site to collect water samples for Winkler titrations. The bottles were lowered into the center of the sensor rack, at the same depth and as close as physically and safely possible to the sensor heads. The bottle was triggered to close at the same time as the instruments were measuring to ensure that the same water mass was compared for DO content. Three replicate 125 ml BOD bottles were filled from each reference sample and immediately fixed in the field for subsequent Winkler titration analysis as described below. The order of each sub-sample was recorded and tracked to examine any variation that arose from sample handling. Approximately 10 - 12 independent sampling events were conducted each week. At least once per week an intensive sampling event was conducted to capture the maximum diurnal range of dissolved oxygen concentrations. Once per week field duplicates were collected to examine fine-scale variability around the mooring site. Approximately 120 comparative reference samples were collected over the 3 - 4 month-long deployments.

In conjunction with each water sample collection, each deployment site also recorded site-specific conditions. The following information, logged on standardized datasheets were transmitted electronically on a weekly basis to the ACT Chief Scientist, for data archiving and site performance review:

- Date, time (local) of water sample collection.
- Barometric pressure from nearest weather station at time of water sample collection.
- Weather conditions (e.g., haze, % cloud cover, rain, wind speed/direction) and air temperature at time of water sample collection.

- Recent large weather event or other potential natural or anthropogenic disturbances.
- Tidal state and distance from bottom of sensor rack at time of water sample collection.
- Any obvious problems or failures with instruments.

ACT was responsible for accurately characterizing temperature and salinity surrounding the mooring with the goal of characterizing micro-stratification or heterogeneity surrounding the mooring. Four RBR Solo temperature loggers and two SeaBird CTDs were deployed at each mooring site. Sensors were mounted both at the instrument sampling depth and approximately 0.5 m above the sampling depth.

At the end of each mooring deployment a pre- and post-cleaned comparison of sensor response to a 100 % saturated water bath was conducted. Upon retrieval the sensor was wrapped in a damp towel and returned to the lab as quickly as possible. Prior to any cleaning, the sensor was submerged in a 100 % DO water bath (via bubbling with air) and DO recorded for a minimum of three readings after an initial 30 minute equilibration period. Then the sensor was removed from the bath and cleaned of any visible biofouling according to recommended manufacturer procedures. Following cleaning the sensor was submerged in a second 100% DO water bath to avoid any biofouling debris carryover and DO recorded for a minimum of three readings after an initial 30 minute equilibration period. Temperature of both water baths was monitored continuously and maintained at a constant condition within 0.5^{o}C. DO concentration was maintained at a constant saturated level with bubbling and confirmed by Winkler titration at the beginning and final instrument reading timepoints.

Water-Column Profiling Test Procedures

Instruments were tested in a profiling application on a CTD rosette aboard the R/V Laurentian in the Great Lakes. Profiling tests were conducted during strong thermal stratification (late August, thermal gradient of >15 °C) and in two different regions including a normoxic and hypoxic hypolimnion. The normoxic hypolimnion site was in Lake Michigan within a 100m deep water column approximately 15 km offshore of Muskegon, MI. The hypoxic site profiling was conducted in Muskegon Lake, a drowned river mouth lake adjacent to Lake Michigan.

Two full water-column CTD casts were conducted at each test site. The first trial involved equilibrating test instruments at the surface (3m) for ten minutes and then collecting three Niskin bottle samples at one minute intervals. Following the third sample, the rosette was quickly profiled into the hypolimnion where samples were collected immediately upon arrival and then each minute for the next 6 minutes. The second trial was performed in the reverse direction where instruments were equilibrated for 10 minutes within the hypolimnion, three samples collected, and then profiled into the surface and sampled at one minute intervals over the next 7 minutes. The CTD was then immediately returned to the ship for sample processing. Triplicate BOD bottles were filled from each Niskin and immediately fixed for Winkler titrations.

Reference Sample Analysis

The Winkler titration for quantifying dissolved oxygen was used as the standard for comparison. The specific method is described in detail below and is based on the procedures described in, *Measurement of primary production and community respiration in oligotrophic lakes using the*

Winkler method (Carignan et al. 1998). All Winkler titrations were done at the individual laboratory and field sites by trained ACT staff using standardized techniques and equipment.

Initial Preparation

The volumes of each BOD bottles (≈ 125 mL) were determined with a precision better than 0.005%. The volume of each bottle was measured gravimetrically (± 0.01 mL) near 20°C, after filling with degassed (boiled 10 min and cooled) distilled water. Since the procedure's precision approaches 1 $\mu g\ O_2 . L^{-1}$, particular care was taken to avoid contamination of the glassware and working space from any trace amounts of thiosulfate, iodate, I_2, and manganese. Reagents recommended by Carritt and Carpenter (1966) were used and whole bottles titrated to minimize the loss of volatile I_2 and the oxidation of iodide to I_2 at low pH.

Reagents

(1) Manganous chloride solution (3M Mn^{2+}): dissolve 300 g of $MnCl_2 \cdot 4H_2O$ in 300 mL of distilled water. Bring to 500 mL.
(2) Alkaline iodide solution (8M OH^-, 4M I^-): separately dissolve 160 g of NaOH and 300 g of NaI in ca 160 mL of distilled water. Mix with stirring and bring to 500 mL.
(3) 23N Sulfuric acid solution: slowly add 313 mL of concentrated H_2SO_4 to 175 mL of distilled water. Carefully mix and cool and bring to 500 mL.
(4) Thiosulfate titrant 0.03N: add 300 mL 0.1N $Na_2S_2O_3 \cdot 5H_2O$ (Fisher SS368-1) to 700 mL DI. The thiosulfate is standardized daily with KIO_3 according to the procedure described below. Note: The normality of thiosulfate will be adjusted to ensure that a complete sample can be titrated within one burette volume (less than 10 mLs), but kept as low as possible to maximize precision.
(5) Potassium iodate standard, 0.1000N ±0.005N commercially available stock (Fisher SP232-1).

Sample Fixing Procedures

(1) Samples were fixed immediately after collection into the BOD bottles. Filling order was noted on log sheets along with bottle and sample IDs. 1.0 ± 0.05 mL of $MnCl_2$ was dispensed just below the water surface, followed by 1.0 ± 0.05 mL of alkaline iodide using positive displacement pipettors. The pipettors were washed with distilled water every day to prevent valve and plunger malfunction due to salt crystallization.
(2) The bottle was immediately closed and shaken vigorously. The precipitate was allowed to settle for about two thirds of the bottle and shaken again to re-suspend the precipitate a second time. A water seal was immediately added to the neck of the bottle to prevent air suction by the contained water sample.
(3) Samples were stored in the dark and room temperature (ca. 20°C) and temperature variations were minimized. Samples were titrated within 18 - 24 hours of being fixed.
(4) Samples were acidified just prior to titration. With the precipitate settled to the lower third of the bottle, 1.0 ± 0.05 mL of 23N H_2SO_4 was added. The H_2SO_4 was allowed to flow gently along the neck of the bottle. The bottle was closed and shaken vigorously, until precipitate was dissolved
(5) If titration was delayed beyond the 24 hour window, the fixed sample remained stored in darkness and at a temperature equal to or slightly lower than the temperature of the samples, with a water seal maintained at all times. The sample was acidified only immediately before titration. Storage at temperatures above the sample temperature cause the loss of I_2 due to the thermal expansion of the solution of 0.025 $mL \cdot °C^{-1}$ for a 125 ml sample (Carignan et al. 1998).

Sample Titration Procedures

Whole bottles were titrated using a Metrohm automated model 916 Ti-Touch titrator equipped with a 10-mL burette and a Metrohm Pt ITrode. The Pt ring of the electrode was polished weekly. The titrator was used in the dynamic equivalence point titration (DET) mode, with a measuring point density of 4, a 1.0-µL minimum increment, and a 2 $mV \cdot min^{-1}$ signal drift condition. In this method, the solution's potential (controlled by the I_2/I^- and $S_2O_3^{2-}/S_4O_6^{2-}$– redox couples) was monitored after successive additions of titrant, where optimal increment volumes are calculated by the titrator's software. During titration, the size and rotation speed of the magnetic stirring bar was controlled in such a way that complete mixing of the I_2 generated during standardization occurred within 3 - 4 s, without vortex formation. To reduce the titration time (3 - 4 min) and I_2 volatilization, an initial volume of titrant equivalent to 85–90% of the expected O_2 concentration was added at the beginning of the titration. Because the molar volume of water and the normality of the titrant vary appreciably with temperature, care was taken to standardize the titrant and conduct all titrations of a given batch of samples at constant temperature (± 1°C).
(1) The stopper of the BOD bottle was removed and, using a wash bottle fitted with a 200-µL pipette tip, the I_2 present on the side and conical part of the stopper was rinsed into the BOD bottle with 1 - 2 mL of distilled water.
(2) BOD bottles (Corning No. 5400-125) had been selected to accommodate the displacement of the electrode without having to remove any volume of the fixed sample.
(3) The stirring bar was inserted into the bottle using plastic or stainless steel forceps.
(4) The delivery tip and the electrode were immersed, the stirrer turned on and the titration begun. The electrode was not allowed to touch the neck of the bottle.
(5) Once the titration was complete, the equivalence point volume (VT) was noted

Thiosulfate Standardization

The Thiosulfate was standardized at room temperature as the first and last step in daily analysis. Either triplicate assays of a fixed volume of iodate standard was run, or a range of volumes (≥ 3) bracketing the normal sample titration range (eg. 0.500, 1.000, 1.500, 2.000 mL for well oxygenated waters.) A clean BOD bottle and clean glassware were dedicated to this purpose.
(1) Insert a stirring bar into a 200 mL beaker.
(2) With mixing add 1.0 mL of the H_2SO_4 reagent followed by 1.0 mL of the alkaline iodide and then 1.0 mL Mn^{2+}reagent.
(3) Using a gravimetrically calibrated pipet add a suitable volume of the KIO_3 standard to the stirring solution
(4) Insert the electrode and delivery tube and immediately begin titration
(5) The normality of the thiosulfate is calculated from the equivalence point volume as Vol_{KIO3} / Vol_{Thio})* N KIO_3 using replicates of single KIO_3 volume additions or from the slope of a range of KIO_3 addition volumes.

Blank determination

Reagent blanks were determined as follows:
(1) A volume of 1-2 L of site water was brought to a boil in a clean glass reagent bottle.
(2) Boiled, degassed water was cooled and poured into 125 ml sample flasks and sparged with N_2 for no less than 30 minutes.
(3) The sample was then rapidly fixed as a normal sample, and on the auto titrator.
(4) A global reagent blank taken as the mean of samples fixed at each test site (0.078 ± 0.020, n=5) and used to correct all reference values.

RESULTS of LABORATORY TESTING

Instrument accuracy for the JFE Rinko AroUSB and AroW-USB were tested under nine combinations of temperature and salinity over a range of DO concentrations from 10% to 120% of saturation (Table 1). Results are plotted as a time series of instrument readings against the time series of comparative Winkler reference samples (Figures A-I for AroUSB; M-U for AroW-USB). The bottom panel of each figure shows the time series of the difference in instrument measurement versus corresponding reference sample. Plotted differences were limited to a range of ± 2.0 mg/L, as values exceeding this range are well outside any expected performance specification and likely occurred as a result of bubbles from the sparging gas used to vary dissolved oxygen concentrations getting trapped on the sensor foil. Comparisons of tank duplicates taken at opposite sides of the tank from between 9-18 of the timepoints during each trial showed a mean difference of 0.017 mg/L, with a range over the nine trials of 0.006 – 0.038 mg/L. Those values are within the analytical precision of the reference method and indicate conditions throughout the tank were very homogeneous and trapping of bubbles on the sensor foil happened as isolated events, which were clearly distinguishable. However, small changes in measured DO concentrations did occur during some of the sampling phases indicating the tank was slightly moderating after sparging was stopped. Those small drifts in DO concentrations were clearly captured by both instrument and reference sample measurements.

Table 1. Dissolved oxygen temperature and salinity challenge trial conditions. For each trial pre and post measurements of tank temperature (^oC) and salinity (S) were made with a calibrated SBE26+4M CTD, equilibrated in well mixed tank for 20 min until stable readings obtained.

Trial ID	Mean Temp ^oC	S.D. Temp ^oC	Mean Salinity PSU	S.D. Salinity PSU	Levels of DO Tested mg/L	Figure for AroUSB	Figure for AroW-USB
L_T15_S00	15.44	0.03	0.00	0.000	0, 2, 5,9,10,14	A	M
L_T15_S10	15.47	0.01	8.82	0.003	0, 2, 8, 9, 13	B	N
L_T15_S35	15.39	0.03	34.20	0.009	0,2,6,8,12	C	O
L_T04_S00	5.40	0.08	0.00	0.000	0,4, 12, 17	D	P
L_T04_S10	5.30	0.03	8.98	0.009	0, 5, 12, 16	E	Q
L_T04_S35	5.23	0.07	34.77	0.073	0, 4, 10, 14	F	R
L_T30_S00	30.22	0.03	0.00	0.000	0, 3, 5, 9	G	S
L_T30_S10	30.51	0.12	9.28	0.036	0, 3, 7, 10	H	T
L_T30_S35	30.61	0.07	34.43	0.050	0, 2, 6, 9	I	U

The mean and standard deviation of the differences between instrument and reference measurements for each trial (n= 24 - 42) are presented in Table 2 for both the AroUSB and the AroW-USB. Mean difference among trials for the AroUSB ranged from -0.277 to 0.265 mg/L DO, and between -0.277 to 0.134 mg/L DO for the AroW-USB. All instrument DO readings were post processed using the manufacturer's software to correct for salinity and there were no obvious patterns to the level of accuracy over the full range of salinity and temperature tested for either model.

Table 2. Summary of average offset mg/L DO between paired JFE AroUSB and AroW-USB DO measures and Winkler reference DO measures during laboratory trials.

Trial ID	AroUSB – Winkler DO			AroW-USB – Winkler DO		
	mean	*s.d.*	*n*	*mean*	*s.d.*	*n*
L_T15_S00	-0.096	0.466	42	-0.183	0.222	24
L_T15_S10	-0.277	0.245	39	-0.100	0.165	39
L_T15_S35	-0.063	0.139	32	-0.169	0.421	38
L_T04_S00	0.265	0.437	39	-0.096	0.191	39
L_T04_S10	0.141	0.237	36	-0.277	0.440	29
L_T04_S35	0.121	0.232	37	-0.145	0.114	37
L_T30_S00	0.071	0.183	40	-0.062	0.251	39
L_T30_S10	0.147	0.534	39	0.134	0.400	39
L_T30_S35	0.169	0.215	40	0.087	0.177	40

The precision of the JFE AroUSB and AroW-USB were also characterized for each of the nine temperatures and salinity trials (Table 3 and 4, respectively). Precision trials were conducted at the start of each new tank test when conditions were most stable. Instruments were equilibrated in test tanks at indicated temperature and salinities for 45 min then the subsequent 31 one minute measurements were used to estimate average tank DO (mg/L) and its variation over that interval. For AroUSB the absolute precision, estimated as the standard deviation (s.d.) around the mean, ranged from 0.002 – 0.014 mg/L across trials with an overall average of 0.004 mg/L. Relative precision, estimated as the coefficient of variation (CV% = (s.d./mean)x100), ranged from 0.013 – 0.278 percent across trials with an overall average of 0.058%.

Table 3. Characterization of the precision of the JFE AroUSB (sn 0091) over a range of temperatures and salinities.

Trial ID	Temperature	Salinity	Dissolved Oxygen Reading			
			mg/L mean	*mg/L s.d.*	*CV%*	*n*
L_T15_S00	15.44	0.00	9.382	0.005	0.054	31
L_T15_S10	15.47	8.82	9.516	0.002	0.023	31
L_T15_S35	15.39	34.20	8.324	0.003	0.032	31
L_T04_S00	5.40	0.00	12.485	0.002	0.013	31
L_T04_S10	5.30	8.98	12.218	0.003	0.024	31
L_T04_S35	5.23	34.77	10.747	0.004	0.035	31
L_T30_S00	30.22	0.00	4.919	0.014	0.278	31
L_T30_S10	30.51	9.28	7.063	0.002	0.033	31
L_T30_S35	30.61	34.43	5.984	0.002	0.030	31

Results for the AroW-USB were very similar (Table 4). Absolute precision ranged from 0.001 – 0.012 across trials, with an overall average of 0.004. Relative precision ranged from 0.017 – 0.247 percent across trials with an overall average of 0.051%.

Table 4. Characterization of the precision of the JFE AroW-USB (sn 0181) over a range of temperatures and salinities.

Trial ID	Temperature	Salinity	Dissolved Oxygen Reading			
			mg/L mean	*mg/L s.d.*	*CV%*	*n*
L_T15_S00	15.44	0.00	9.035	0.004	0.040	31
L_T15_S10	15.47	8.82	9.174	0.002	0.020	31
L_T15_S35	15.39	34.20	8.035	0.002	0.030	31
L_T04_S00	5.40	0.00	12.066	0.003	0.021	31
L_T04_S10	5.30	8.98	11.783	0.005	0.040	31
L_T04_S35	5.23	34.77	10.341	0.003	0.028	31
L_T30_S00	30.22	0.00	4.798	0.012	0.247	31
L_T30_S10	30.51	9.28	6.954	0.001	0.017	31
L_T30_S35	30.61	34.43	5.877	0.001	0.018	31

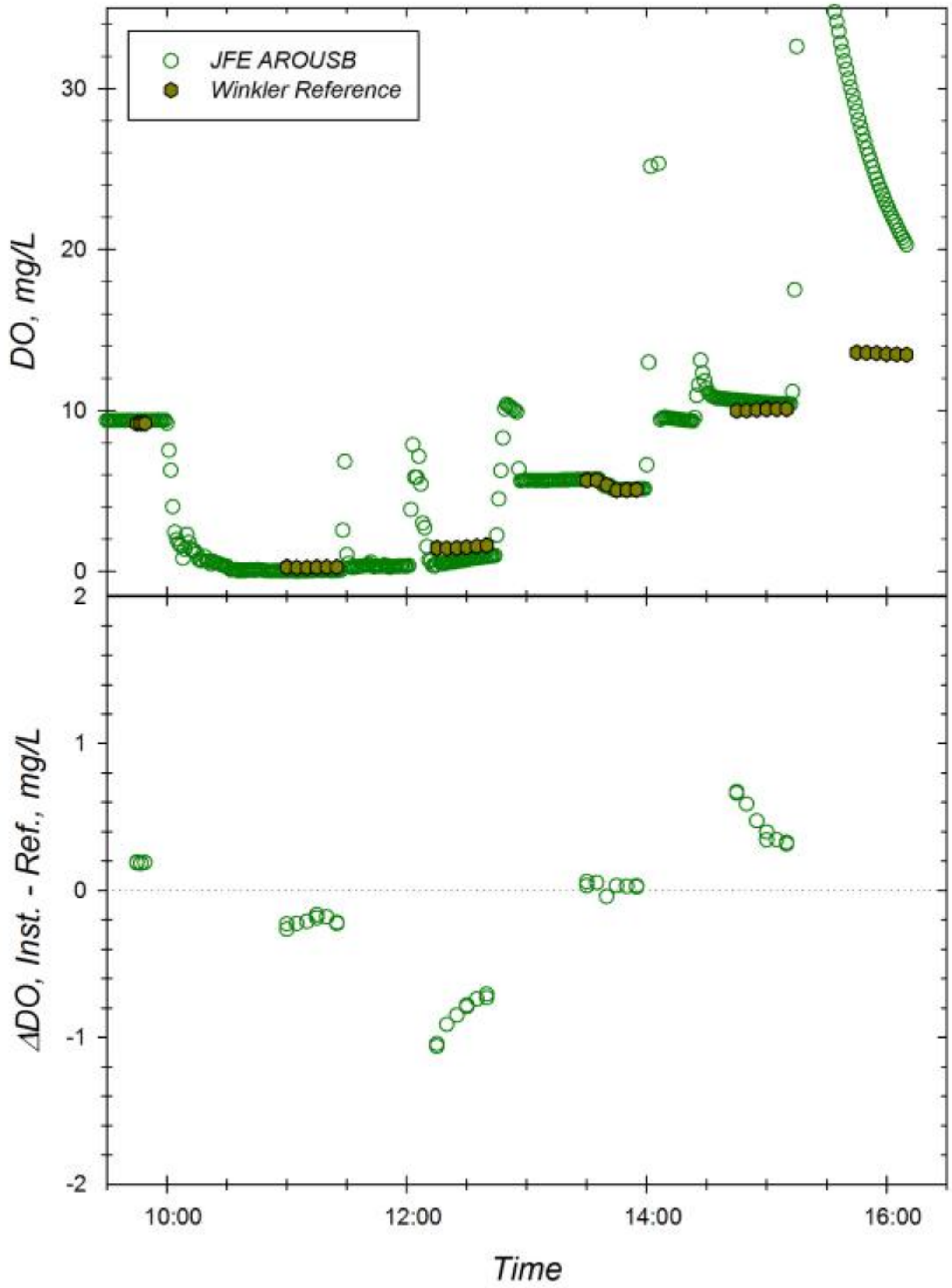

Figure A. Response of JFE AROUSB (sn 0091) to changing dissolved oxygen concentrations at 15.4 °C and S=0.00 *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. The decay in instrument signal under high DO exposure likely reflects bubble contamination of DO sensor foil. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference. Offsets beyond 2 mg/L are excluded.

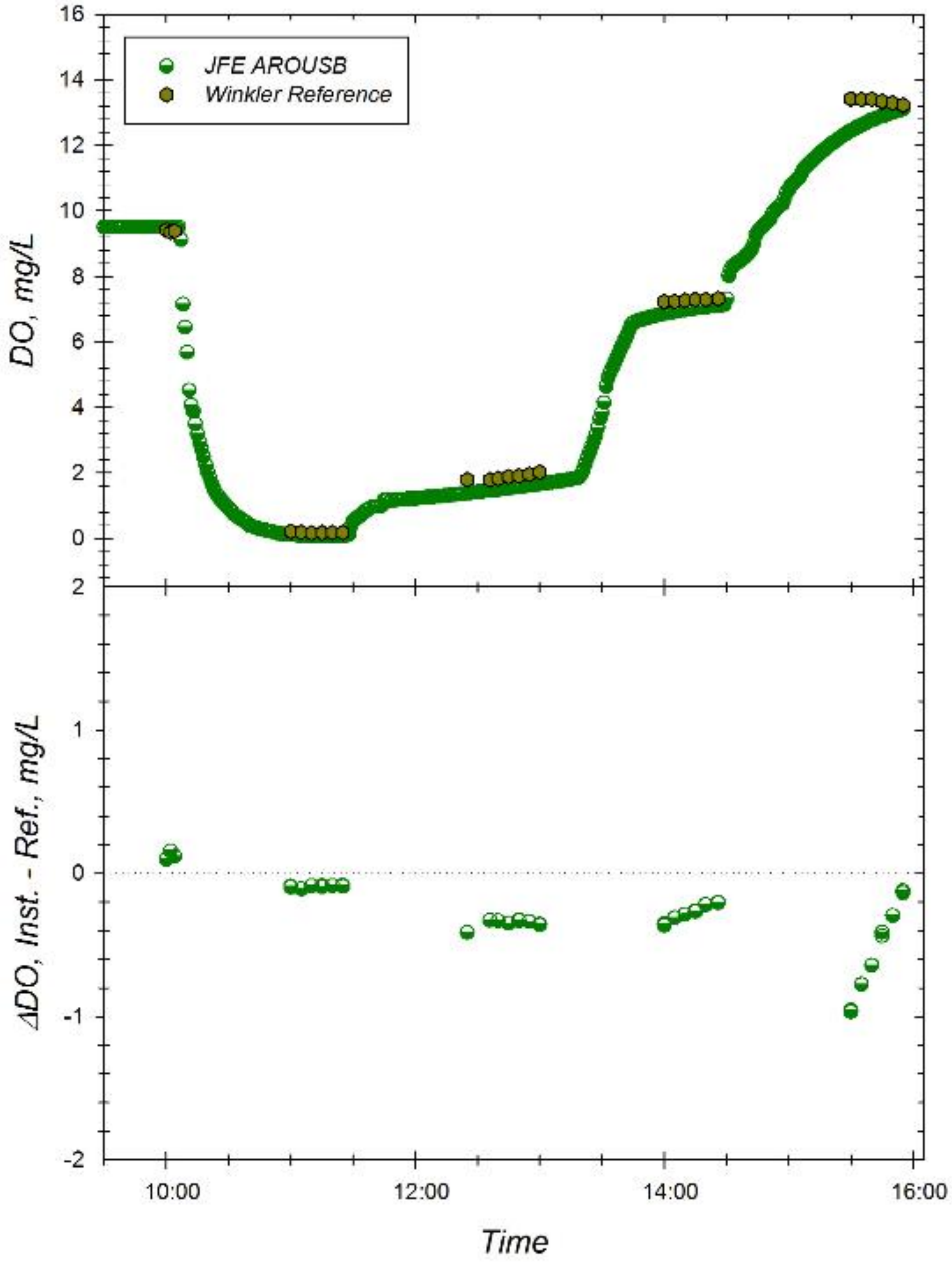

Figure B. Response of JFE AROUSB (sn 0091) to changing dissolved oxygen concentrations at 15.4 ^{o}C and S=8.82. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference.

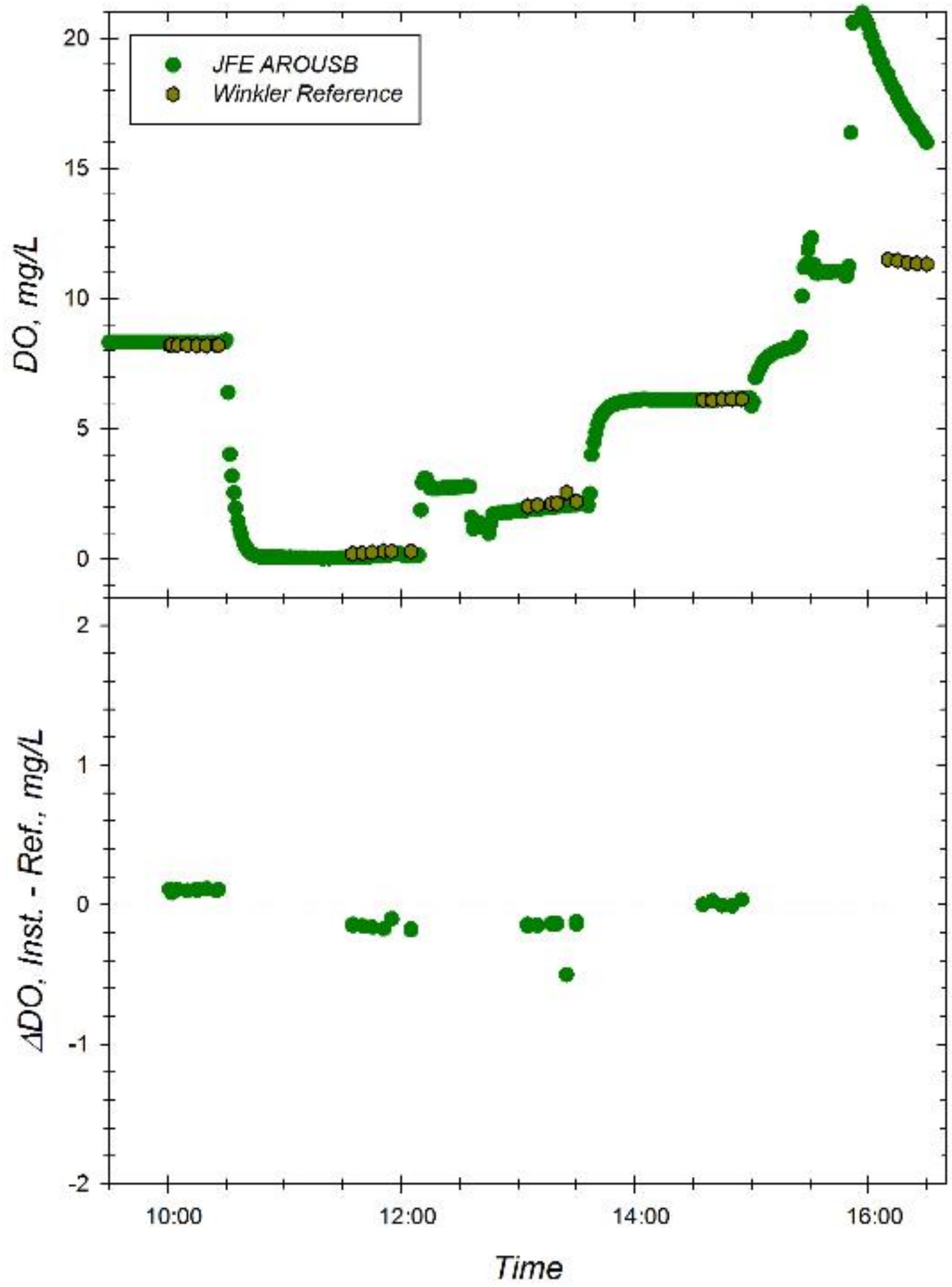

Figure C. Response of JFE AROUSB (sn 0091) to changing dissolved oxygen concentrations at 15.4 °C and S=34.20. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. The decay in instrument signal under high DO exposure likely reflects bubble contamination of DO sensor foil. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference. Offsets beyond 2 mg/L are excluded.

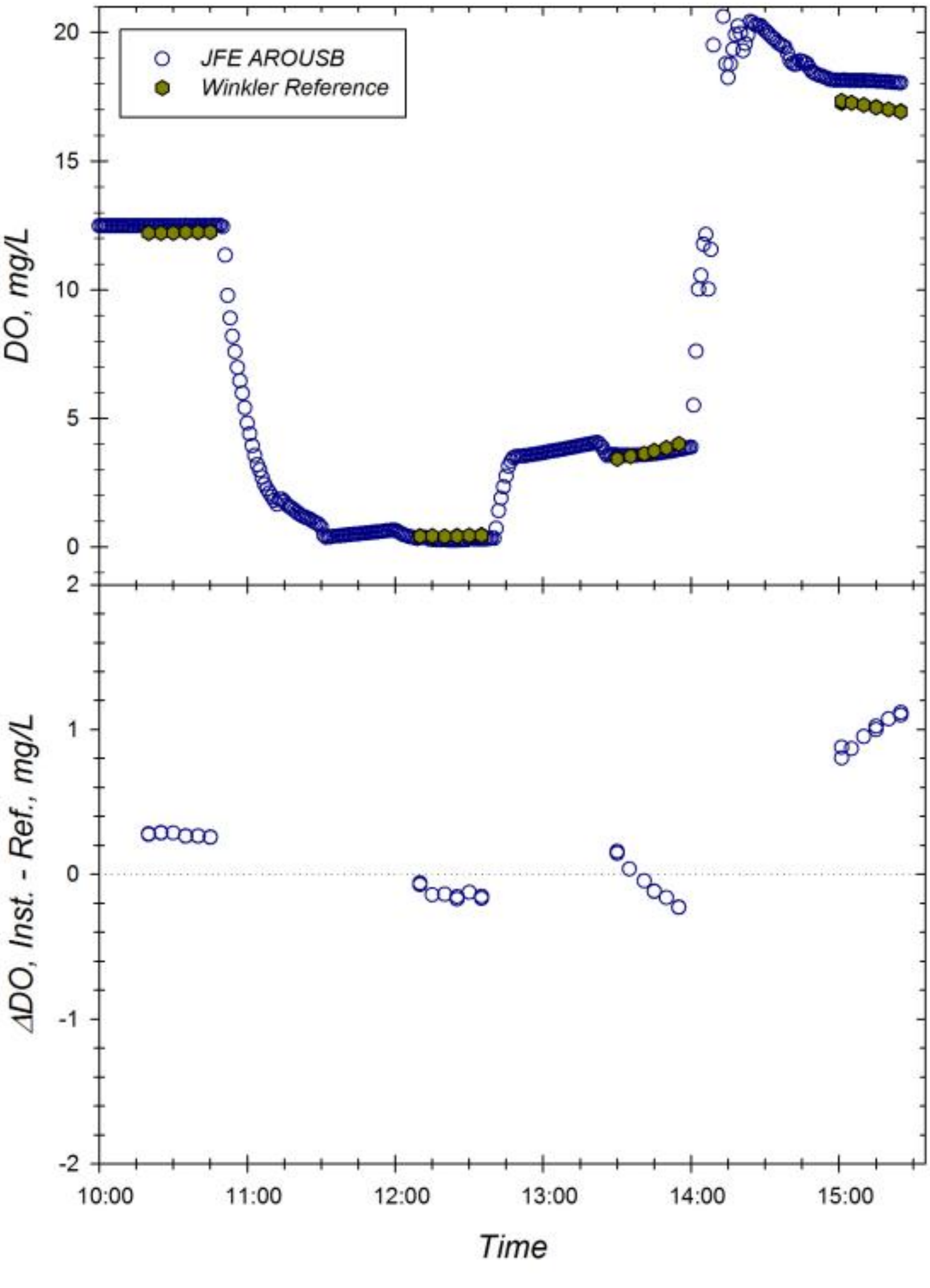

Figure D. Response of JFE AROUSB (sn 0091) to changing dissolved oxygen concentrations at 5.4 °C and S=0.00. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference.

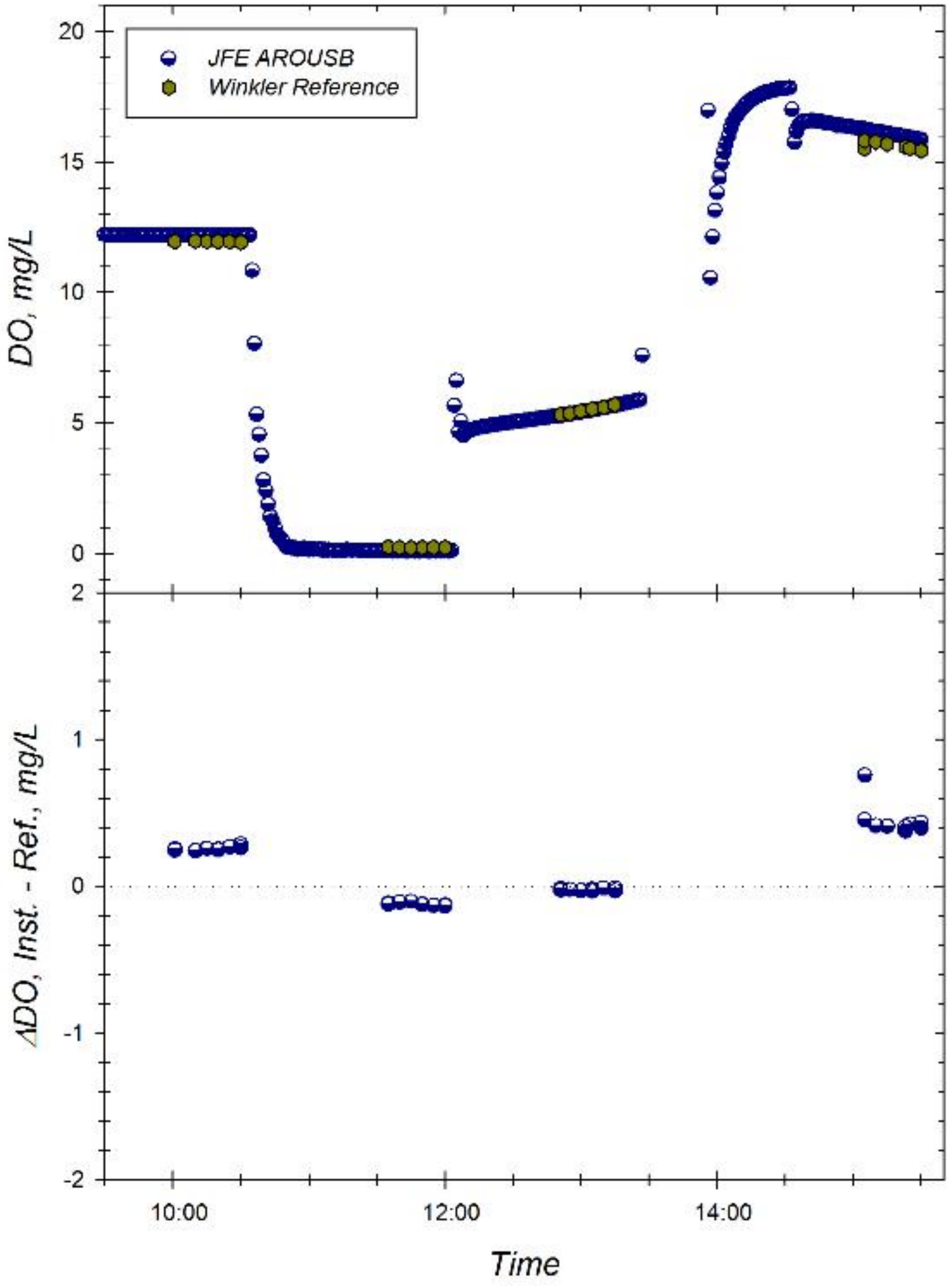

Figure E. Response of JFE AROUSB (sn 0091) to changing dissolved oxygen concentrations at 5.3 °C and S=8.98. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference.

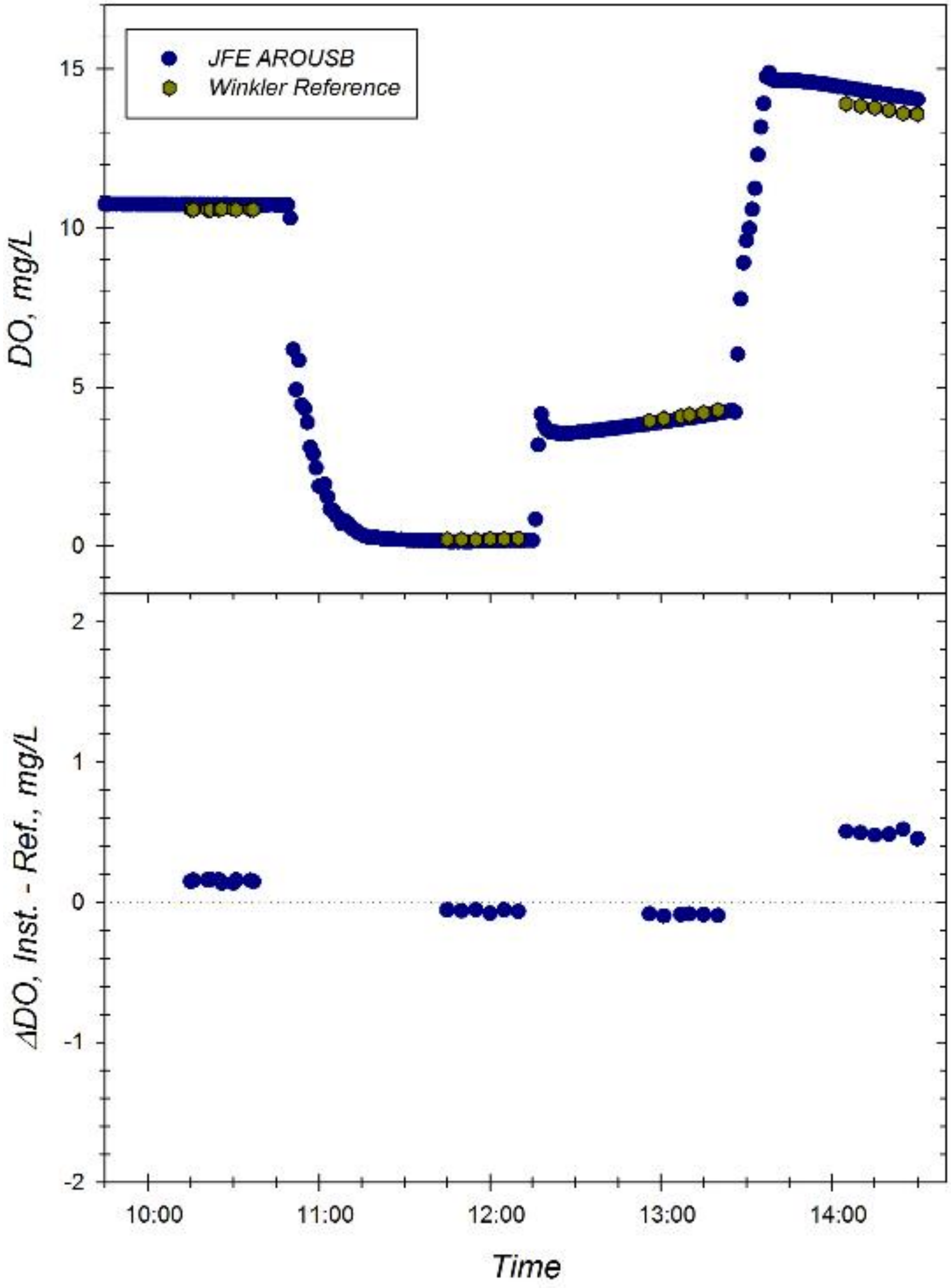

Figure F. Response of JFE AROUSB (sn 0091) to changing dissolved oxygen concentrations at 5.2 °C and S=34.77. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference. Offsets beyond 2 mg/L are excluded.

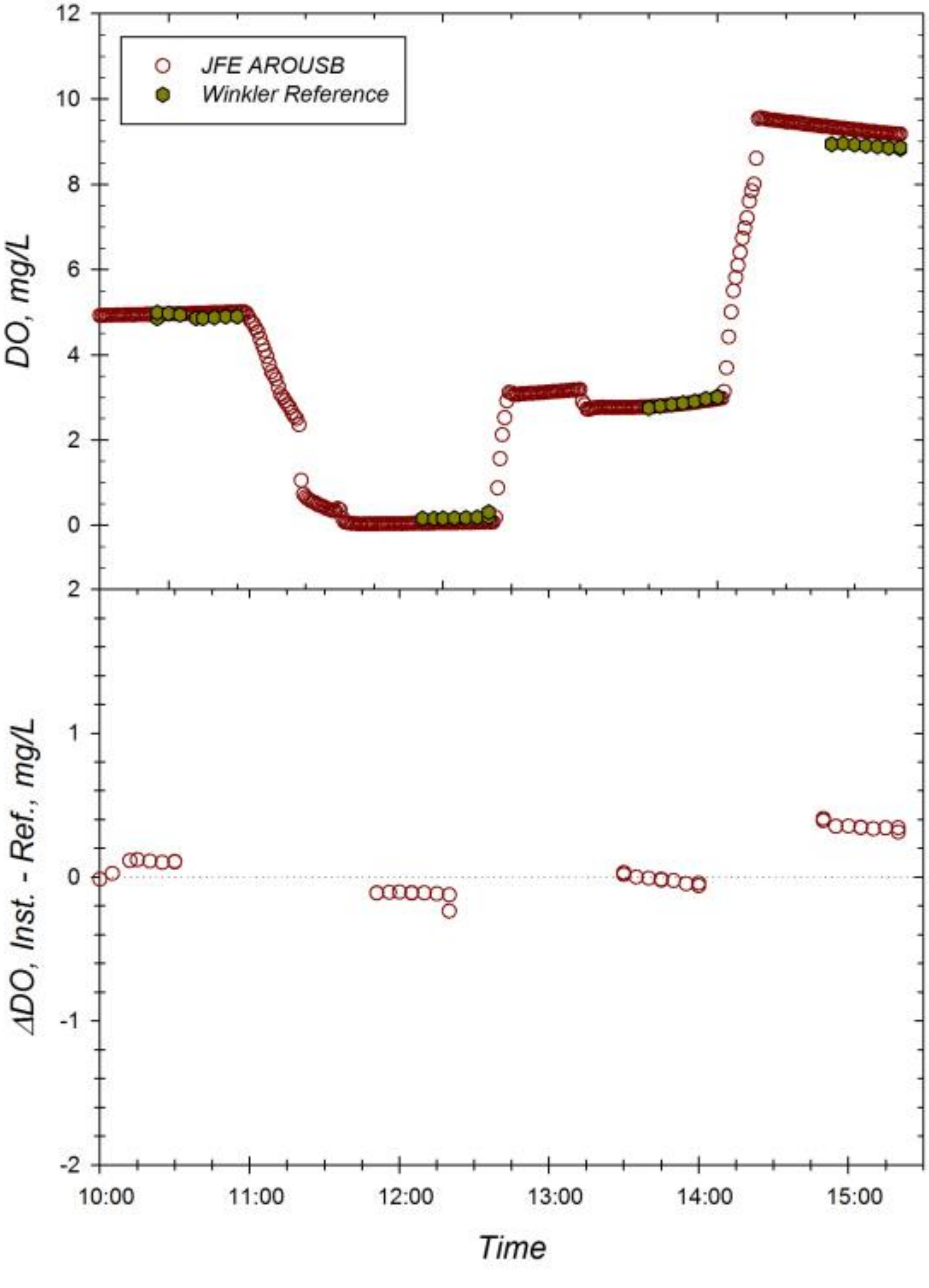

Figure G. Response of JFE AROUSB (sn 0091) to changing dissolved oxygen concentrations at 30.2 °C and S=0.00. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference.

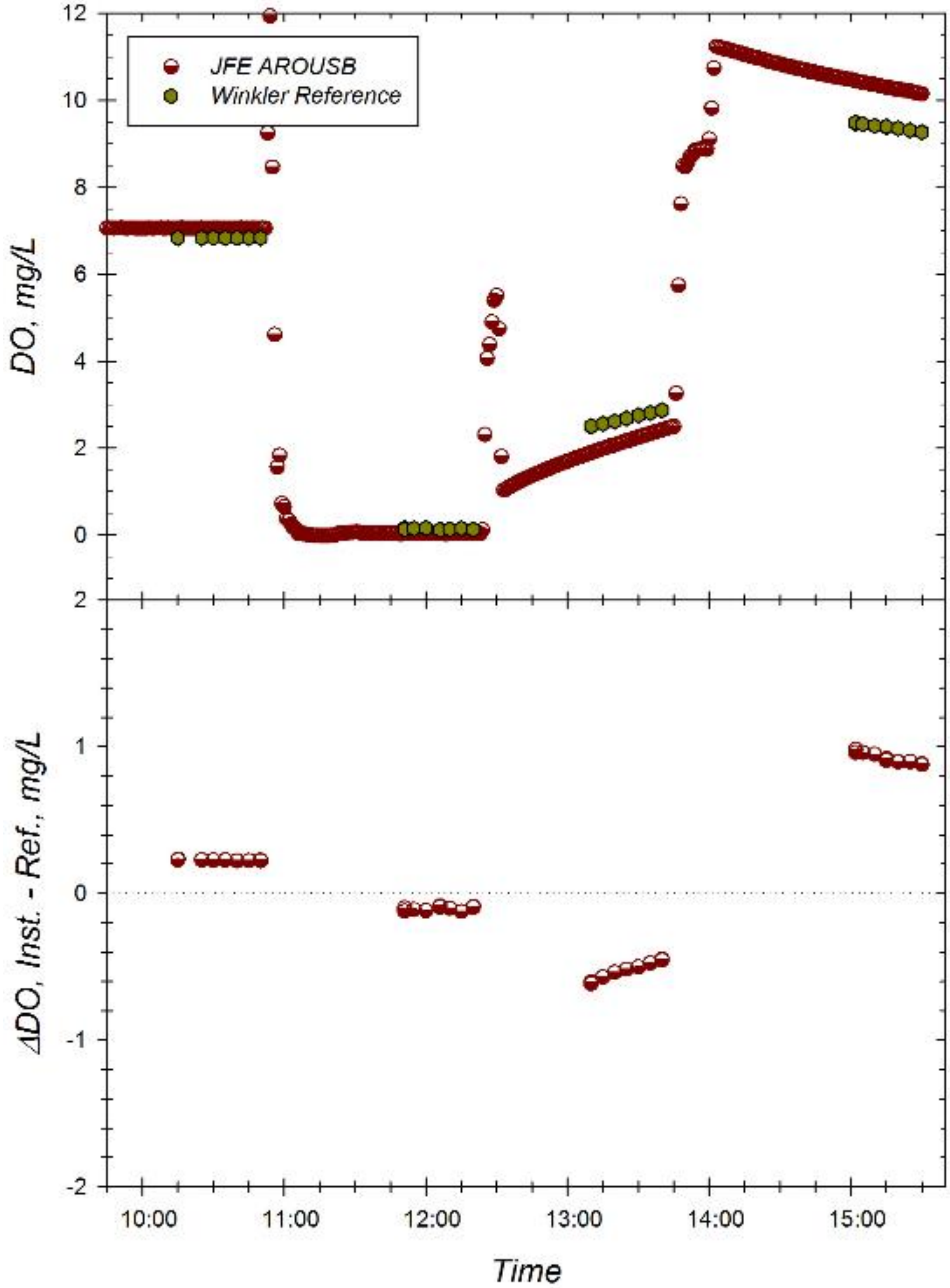

Figure H. Response of JFE AROUSB (sn 0091) to changing dissolved oxygen concentrations at 30.5 °C and S=9.28. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference.

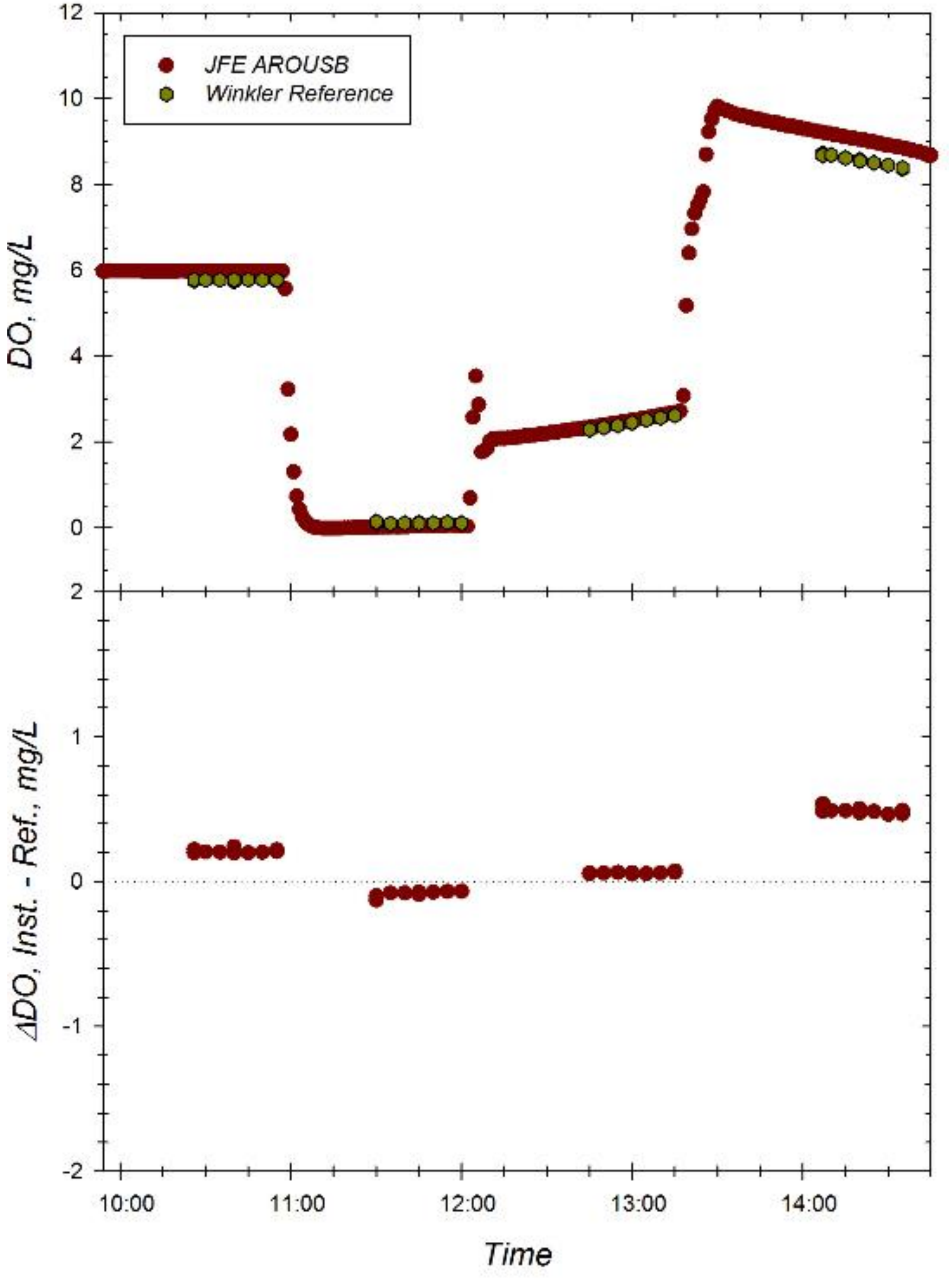

Figure I. Response of JFE AROUSB (sn 0091) to changing dissolved oxygen concentrations at 30.6 °C and S=34.43. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. The decay in instrument signal under high DO exposure likely reflects bubble contamination of DO sensor foil. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference. Offsets beyond 2 mg/L are excluded.

The global response of the AroUSB for all nine temperature-salinity trials is shown in figure J. In general, readings are well correlated across the entire range of DO concentrations (hypoxic to supraoxic) covering all temperature and salinity ranges tested. Highlighted deviation likely reflects bubble contamination of the sensor foil. A linear regression of the data (n=377; r^2 = 0.965; p<0.0001) produced a slope of 1.015 and intercept of 0.098. Highlighted data were omitted from the regression because of expected bubble contamination.

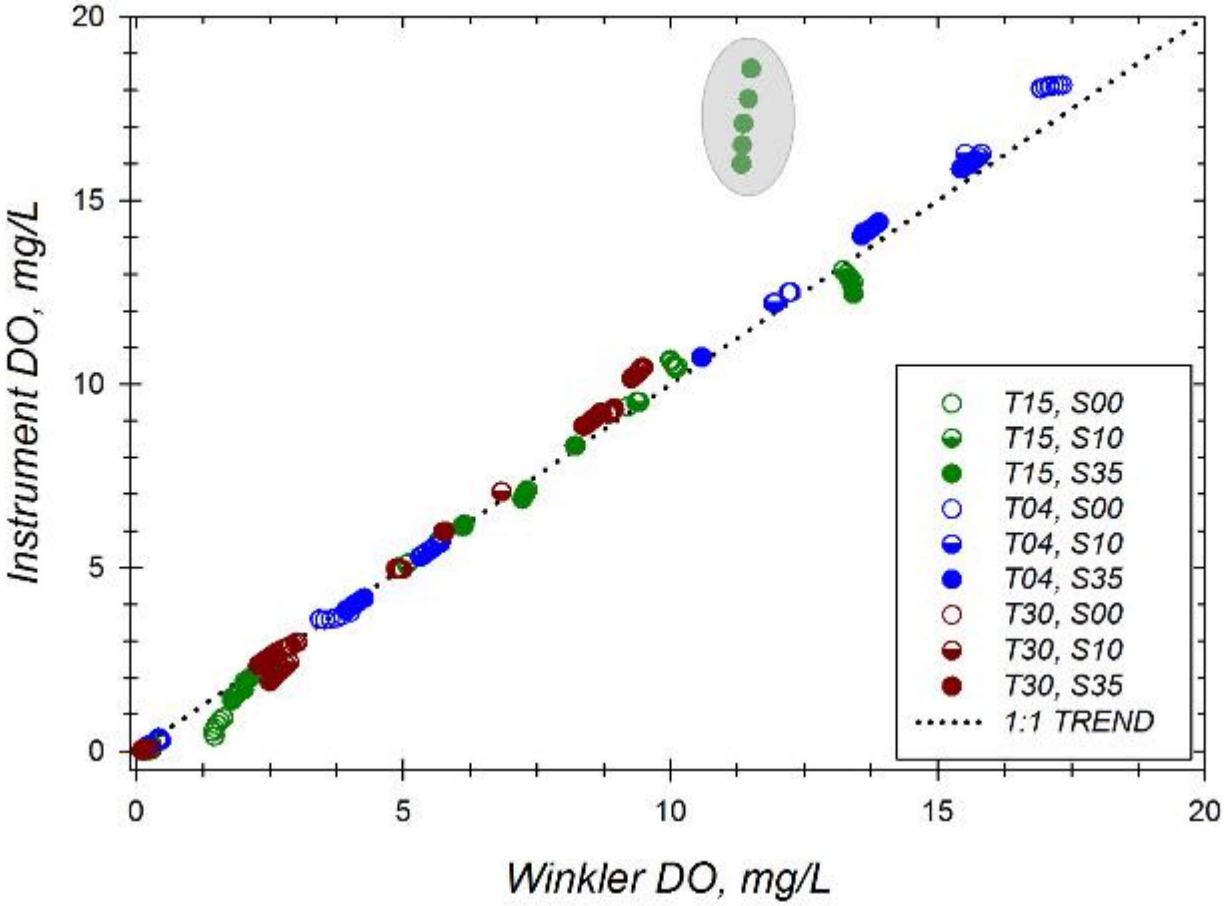

Figure J. Concordance of JFE AROUSB (sn 0091) oxygen measures paired with discrete Winkler DO reference samples. Dotted line represents perfect agreement between the two measures.

Individual response slopes and intercepts of instrument versus reference DO for each of the laboratory trials are summarized in Table 5. Slopes were similar across all three salinity trials when averaged across temperature. The response slope was slightly higher for the 30 °C trials (mean = 1.11) versus the 5 and 15 °C trials (mean = 1.04 and 1.05, respectively) when averaged across salinity.

Table 5. Summary of regression statistics for the AroUSB versus reference sample response curves for each of the nine laboratory trials.

Test ID	R2	y Intercept	Slope
LT15S00*	0.996	-0.567	1.094
LT15S10	0.998	-0.159	0.980
LT15S35*	1.000	-0.203	1.036
LT04S00	0.999	-0.257	1.062
LT04S10	1.000	-0.173	1.038
LT04S35	1.000	-0.163	1.039
LT30S00	1.000	-0.151	1.054
LT30S10	0.995	-0.437	1.121
LT30S35	1.000	-0.105	1.065

* Bubble contamination data excluded.

Results of the 56 day long-term stability and thermal stress challenge for the AroUSB are shown in figure K. The instrument was maintained in a well circulated tank and oxygen content manipulated by alternately varying water temperature set point between 15 and 25 ^{o}C several times per week of deployment. The data completion result for the stability test was 100%. The time series of instrument readings at 15 min intervals is plotted against discrete values for Winkler reference samples (*top panel*) along with the time series of the difference between instrument and reference measurements (*bottom panel*). The overall mean difference between measurements was 0.001 (s.d. = 0.326) mg/L. There was no significant trend in accuracy over time (slope = -0.0007 mg/L/d) that would indicate any type of performance drift over the duration.

Results for the functional sensor response time assessment of the AroUSB are shown in figure L. The top panel depicts the time series of 10s instrument reads during transfers between adjacent high (9.6 mg/L) and low (2.0 mg/L) DO water baths, maintained commonly at 15 ^{o}C. The bottom panel (*lower left*) depicts results fit with a 3 parameter exponential decay function: $DO_{rel} = DO_{relMin} + a\boldsymbol{e}^{-bt}$ and indicated τ calculated from fit. Data for low DO to high DO transitions (*lower right*) were treated similarly but normalized to steady state value in subsequent high DO tank and subsequently fit with an analogous 3 parameter exponential rise function: $DO_{rel} = DO_{relMin} + a(1-\boldsymbol{e}^{-bt})$ with indicated τ being directly calculated from fit. The calculated τ_{90} was 11.8 s during high to low transitions and 7.1 s for low to high transitions covering a DO range of approximately 8 mg/L at a constant 15 ^{o}C. However, the AroUSB was incorrectly programmed for a sampling rate of 10 seconds instead of its highest frequency of 1 s. We caution the reader that the sampling rate would have a direct impact on the calculated response rate as noted within the manufacturer's response.

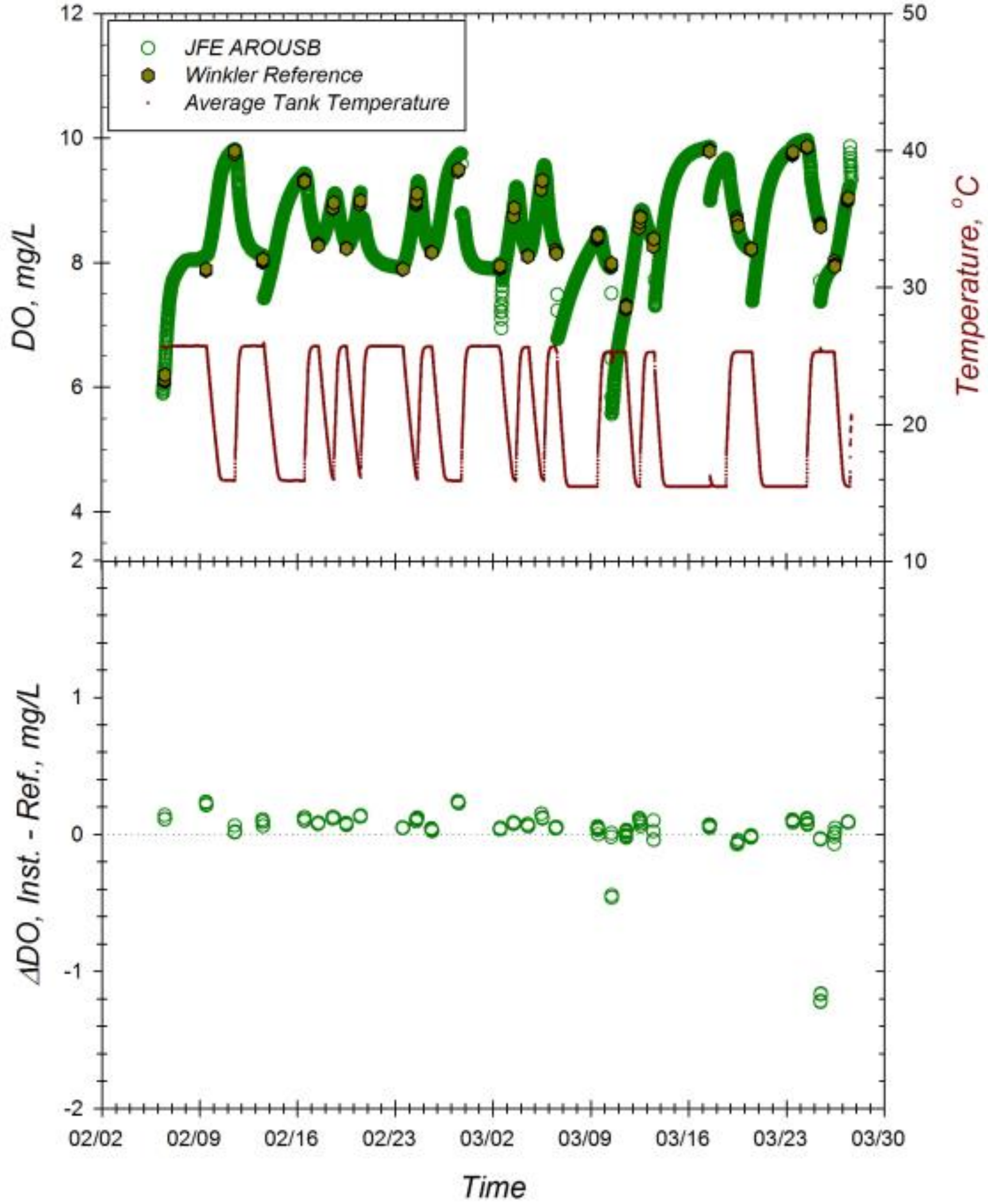

Figure K. Long-term stability and thermal stress challenge of the JFE AROUSB (sn 0091). Instrument was maintained in well circulated tank and oxygen content manipulated by alternately varying water temperature set point between 15 and 25 °C several times per week of deployment. *Top Panel*: Time series of AROUSB instrument readings representing average of ten 1s burst reads at 15min intervals with discrete values for Winkler assays of reference samples taken at sensor level. *Bottom Panel*: Offset in paired DO readings of AROUSB and Winkler reference samples. Offsets beyond 2 mg/L (1 event) are excluded for clarity.

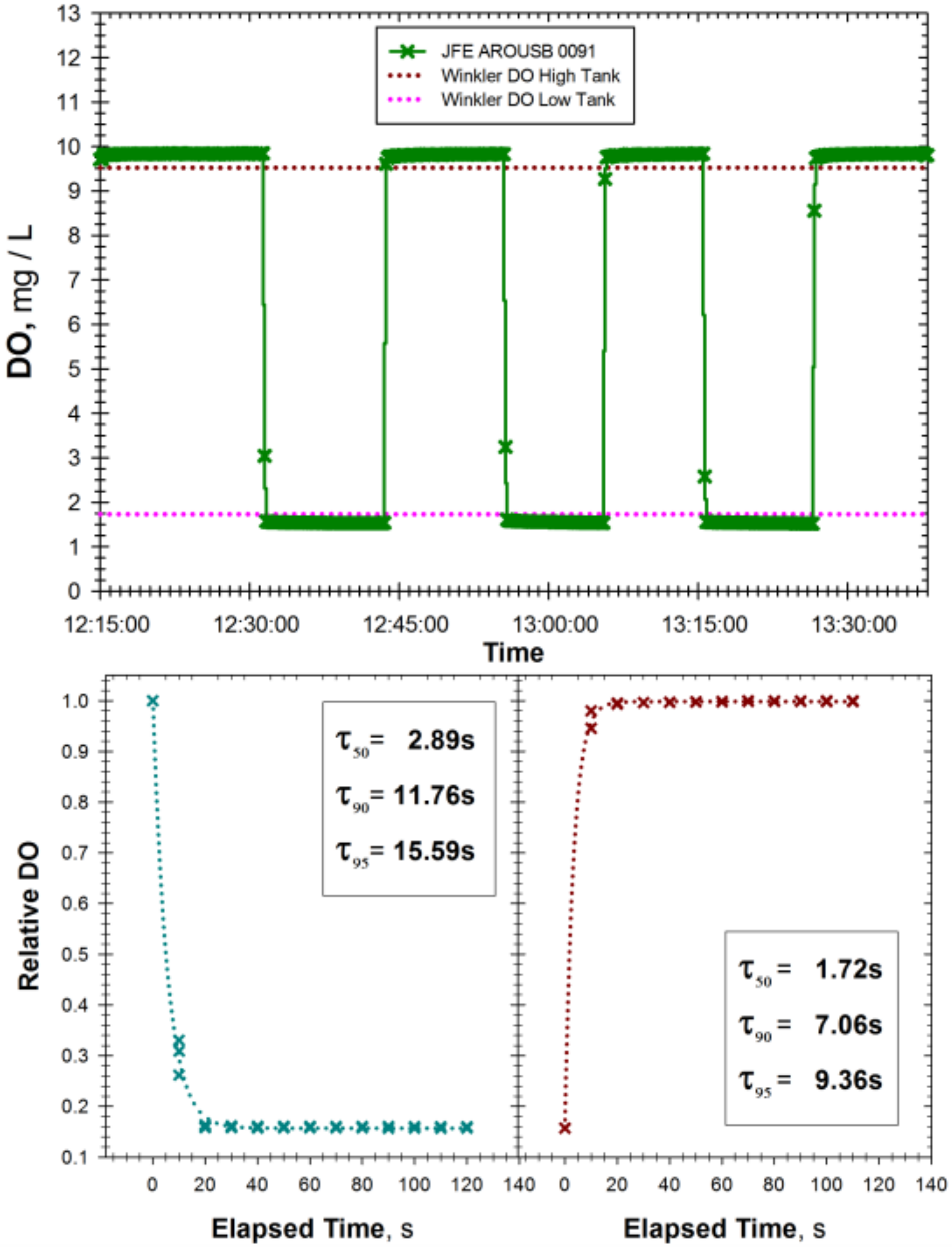

Figure L. Functional sensor response time assessment of the JFE AROUSB (sn 0091). *Top Panel*: Time series of 10s instrument reads during transfers between adjacent high and low DO water baths, maintained commonly at 15 ^{o}C. Indicated DO levels were maintained by air stone bubblers (high DO) or air+N2 (low DO) and verified by Winkler assays. *Bottom Panels*: All high DO to low DO transitions were normalized to last reading before condition switch for the indicated post transition time. Data was fit with a 3 parameter exponential decay function: $DO_{rel} = DO_{relMin} + a\boldsymbol{e}^{-bt}$ and indicated τ calculated from fit. Data for low DO to high DO transitions (*lower right*) were treated similarly but normalized to steady state value in subsequent high DO tank and subsequently fit with an analogous 3 parameter exponential rise function: $DO_{rel} = DO_{relMin} + a(1-\boldsymbol{e}^{-bt})$ with indicated τ being directly calculated from fit.

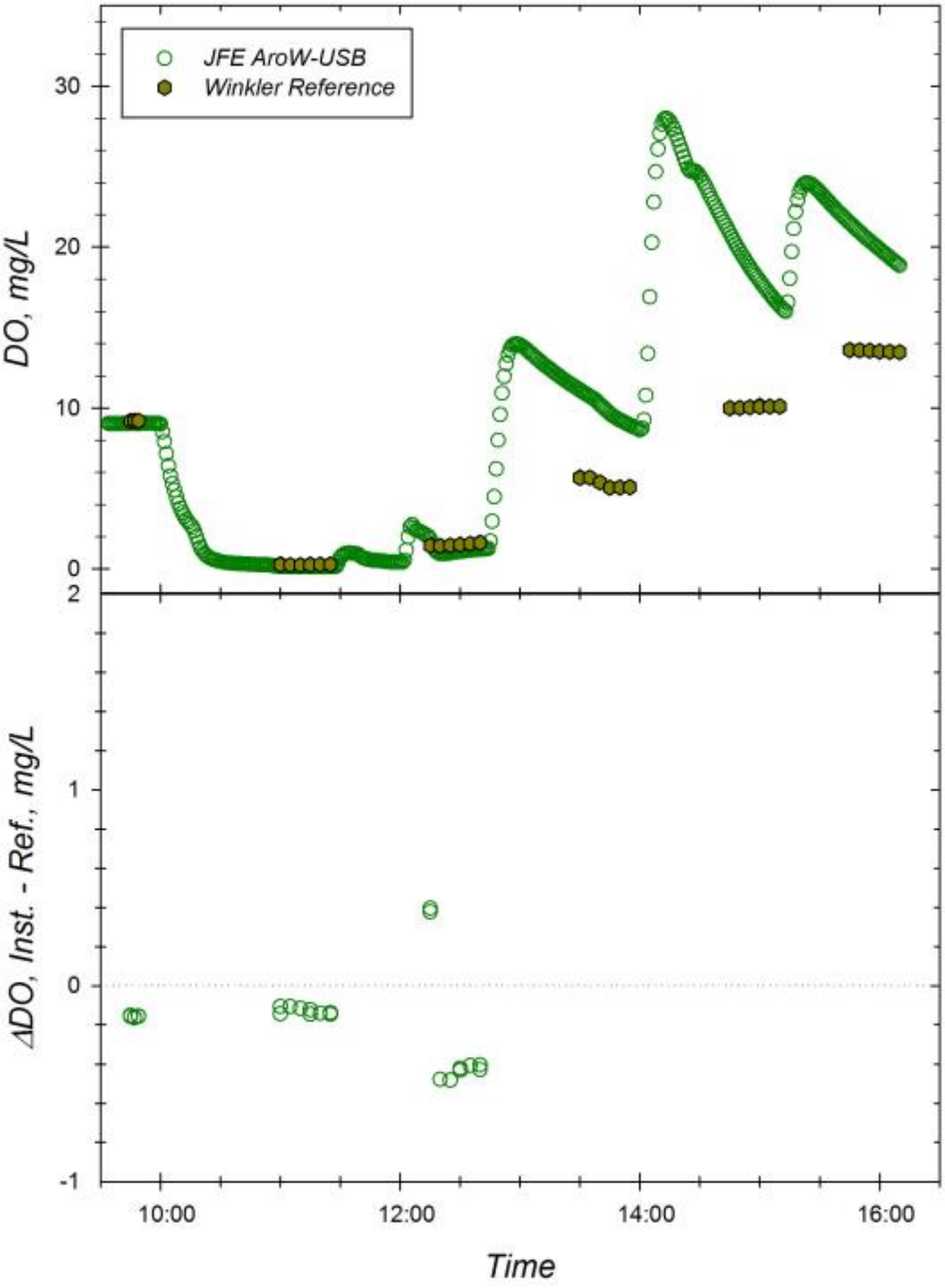

Figure M. Response of JFE AroW-USB (sn 0181) to changing dissolved oxygen concentrations at 15.4 ºC and S=0.00. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. The deviations in instrument signal under high DO exposure likely reflects bubble contamination of DO sensor foil. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference. Offsets beyond 2 mg/L are excluded as likely due to bubble interference.

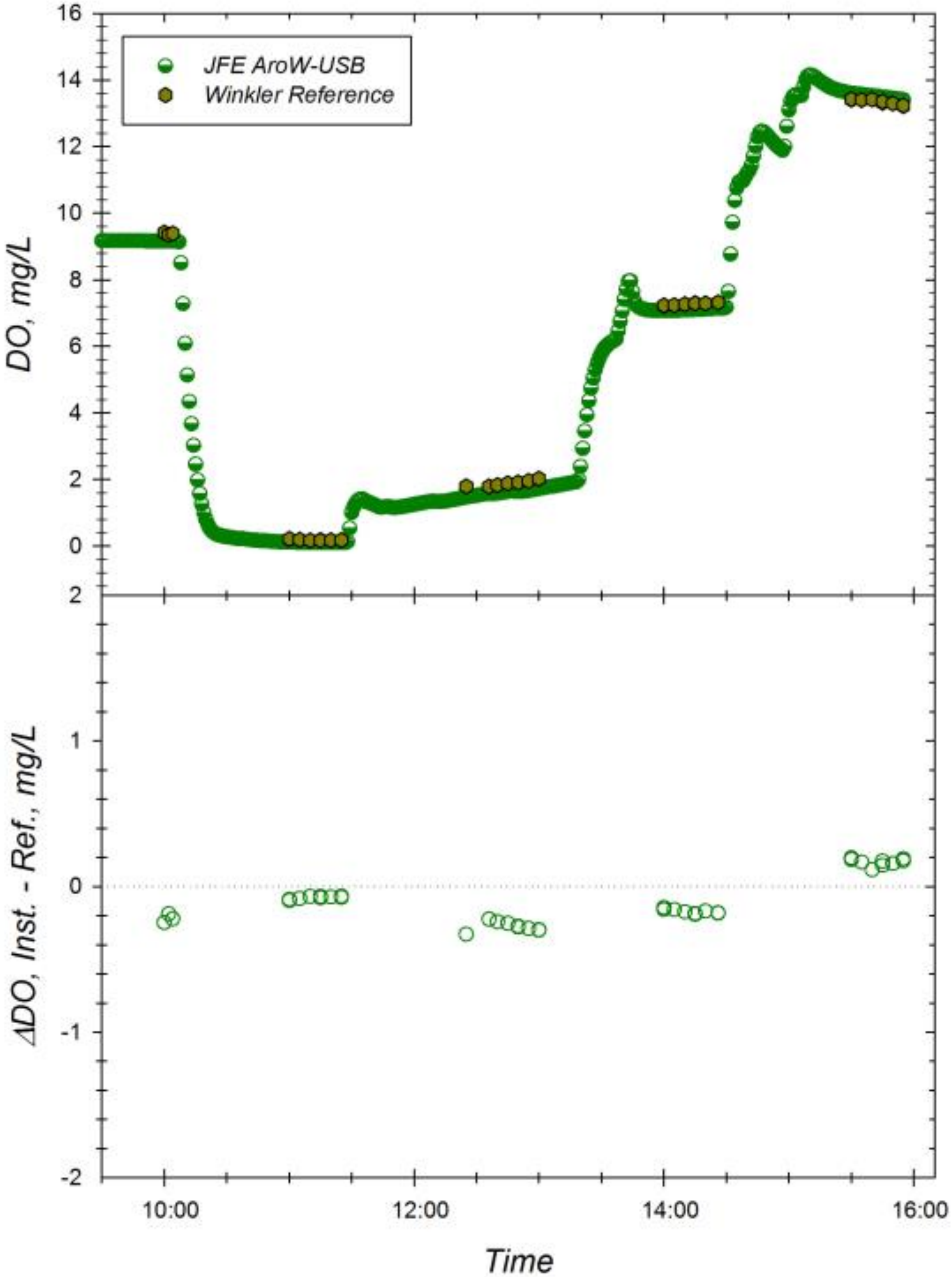

Figure N. Response of JFE AroW-USB (sn 0181) to changing dissolved oxygen concentrations at 15.4 ºC and S=8.82. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference.

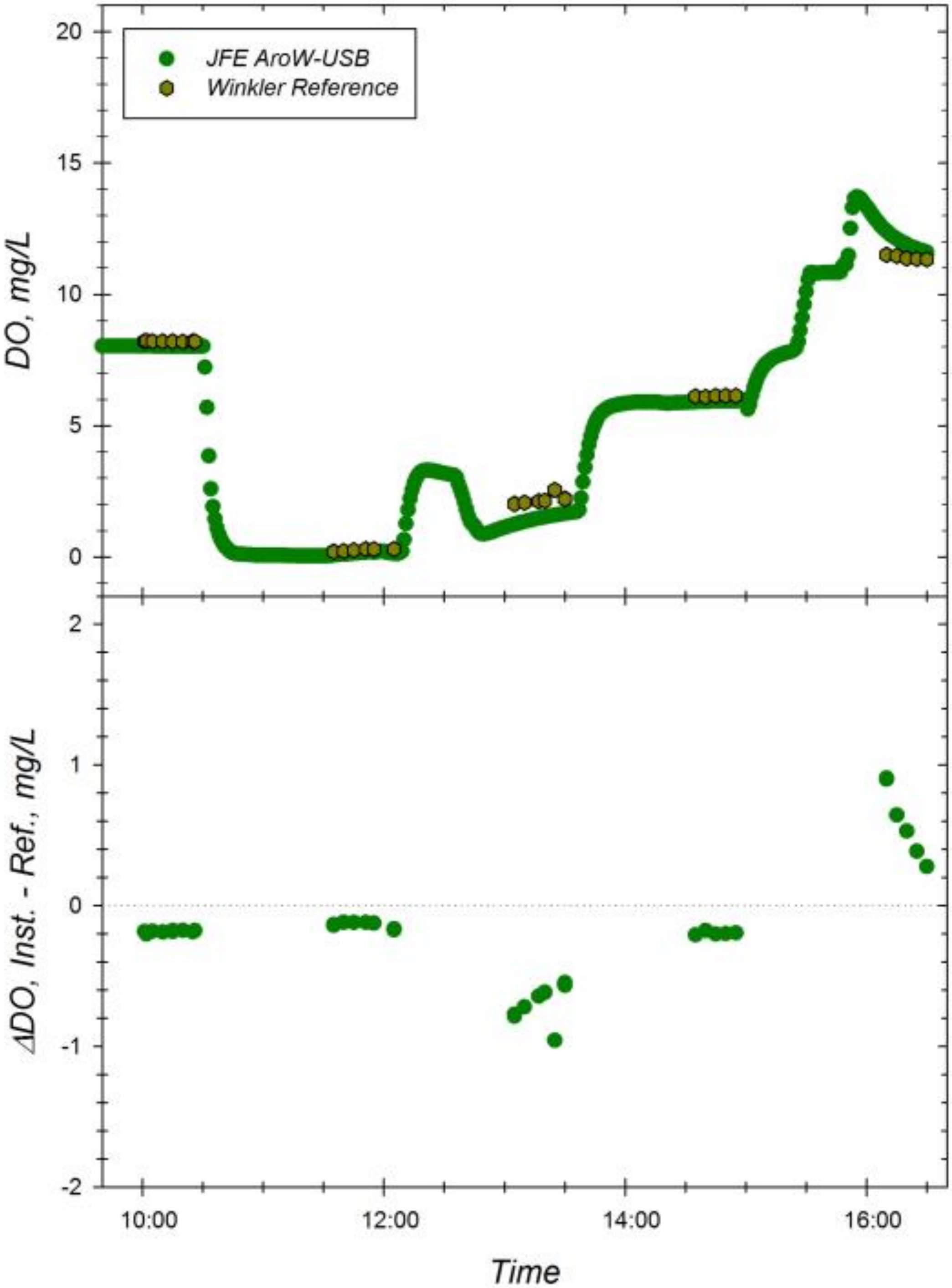

Figure O. Response of JFE AroW-USB (sn 0181) to changing dissolved oxygen concentrations at 15.4 ºC and S=34.20. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. The decay in instrument signal under high DO exposure likely reflects bubble contamination of DO sensor foil. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference. Offsets beyond 2 mg/L are excluded.

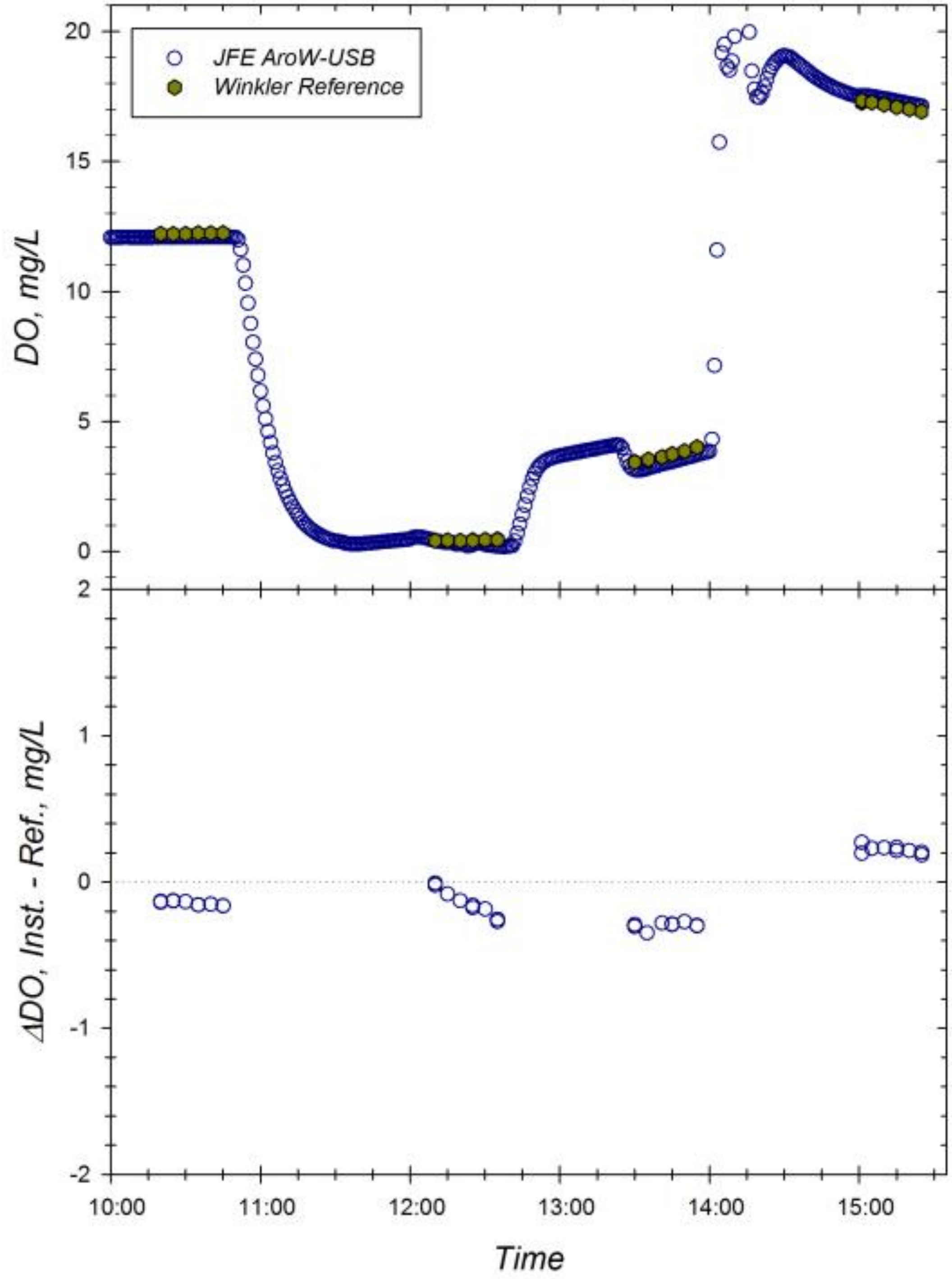

Figure P. Response of JFE AroW-USB (sn 0181) to changing dissolved oxygen concentrations at 5.4 ºC and S=0.00. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference.

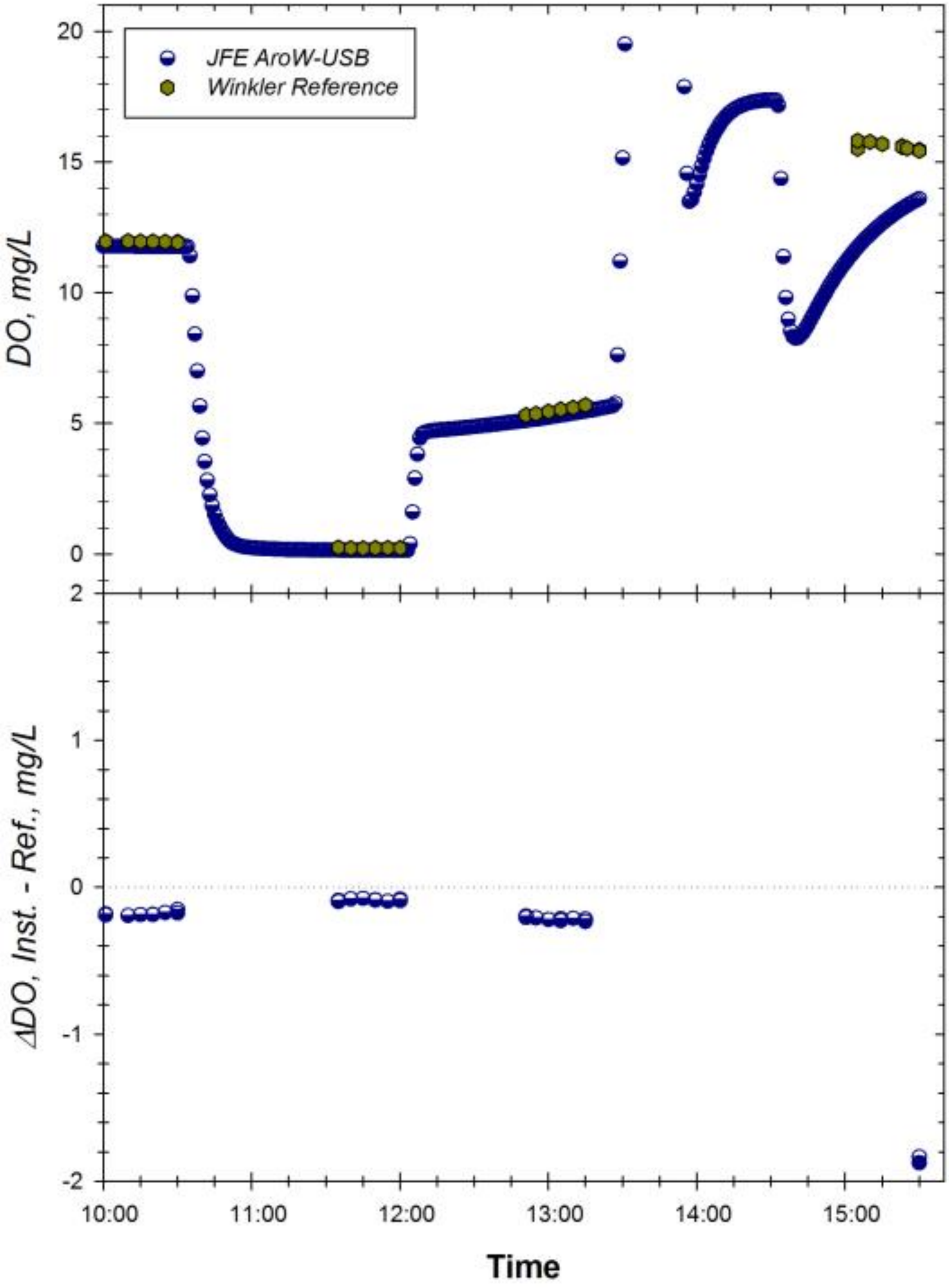

Figure Q. Response of JFE AroW-USB (sn 0181) to changing dissolved oxygen concentrations at 5.3 ºC and S=8.98. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. The decay in instrument signal under high DO exposure likely reflects bubble contamination of DO sensor foil. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference.

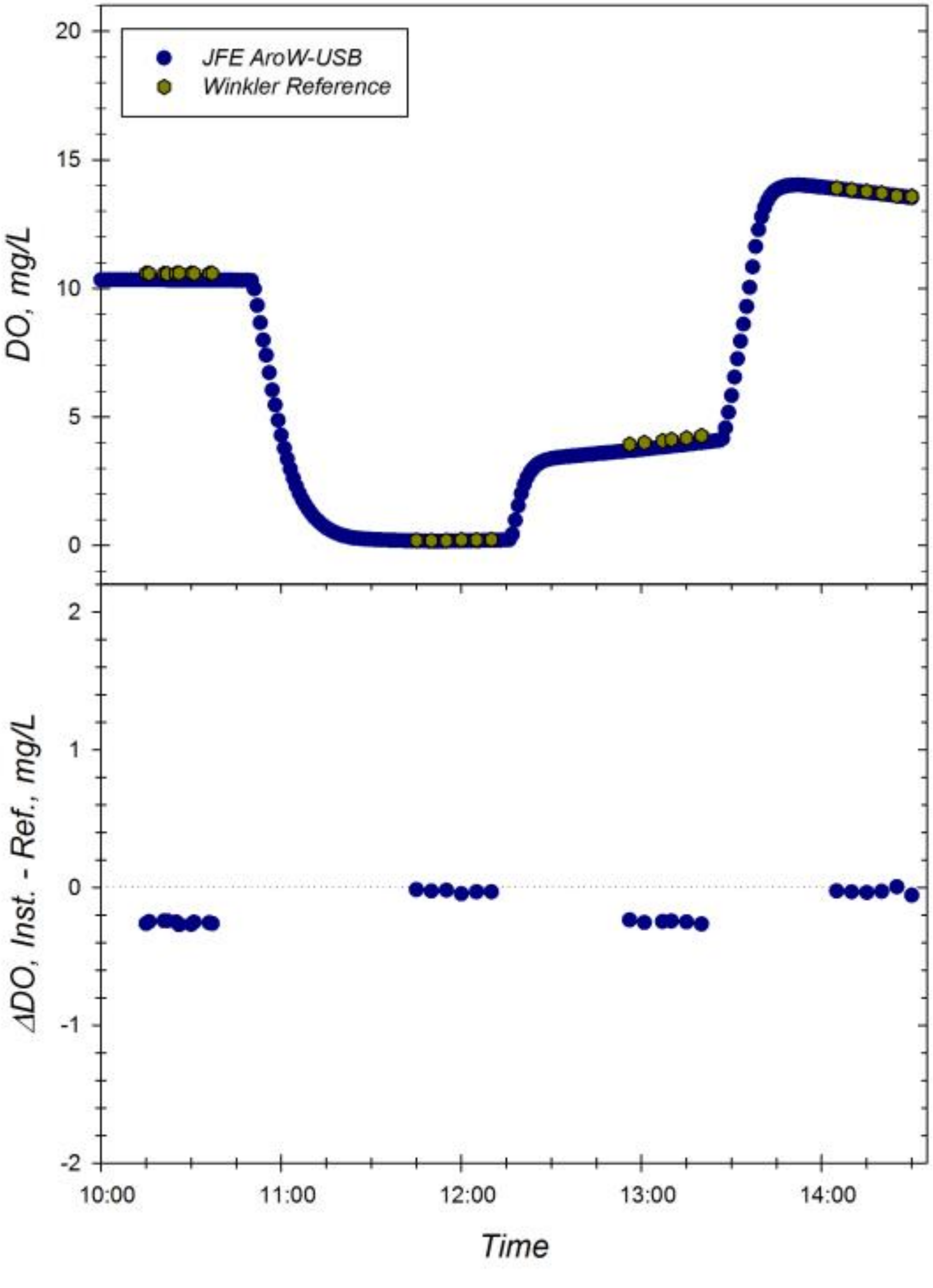

Figure R. Response of JFE AroW-USB (sn 0181) to changing dissolved oxygen concentrations at 5.2 ºC and S=34.77. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference. Offsets beyond 2 mg/L are excluded.

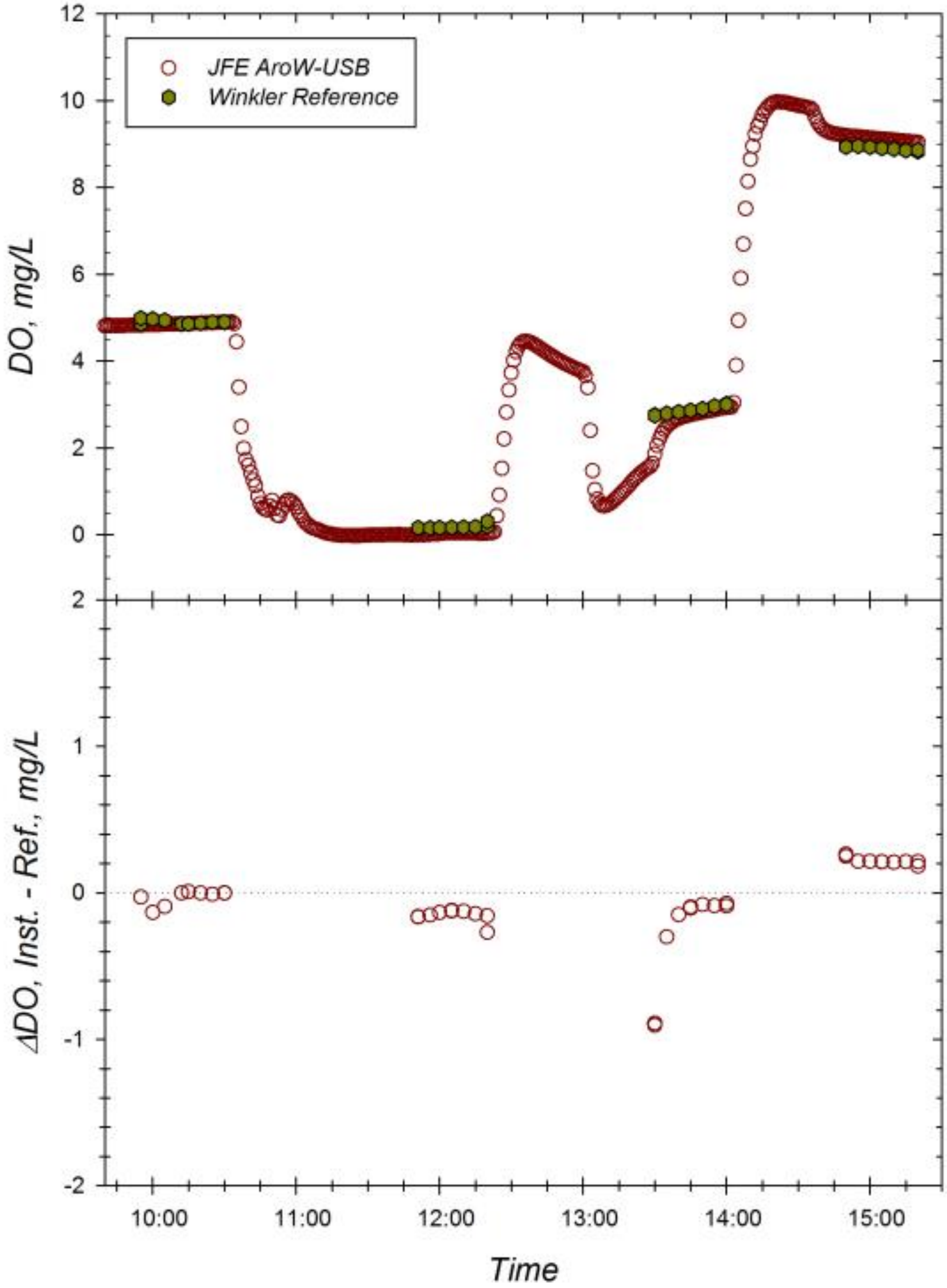

Figure S. Response of JFE AroW-USB (sn 0181) to changing dissolved oxygen concentrations at 30.2 ºC and S=0.00. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference.

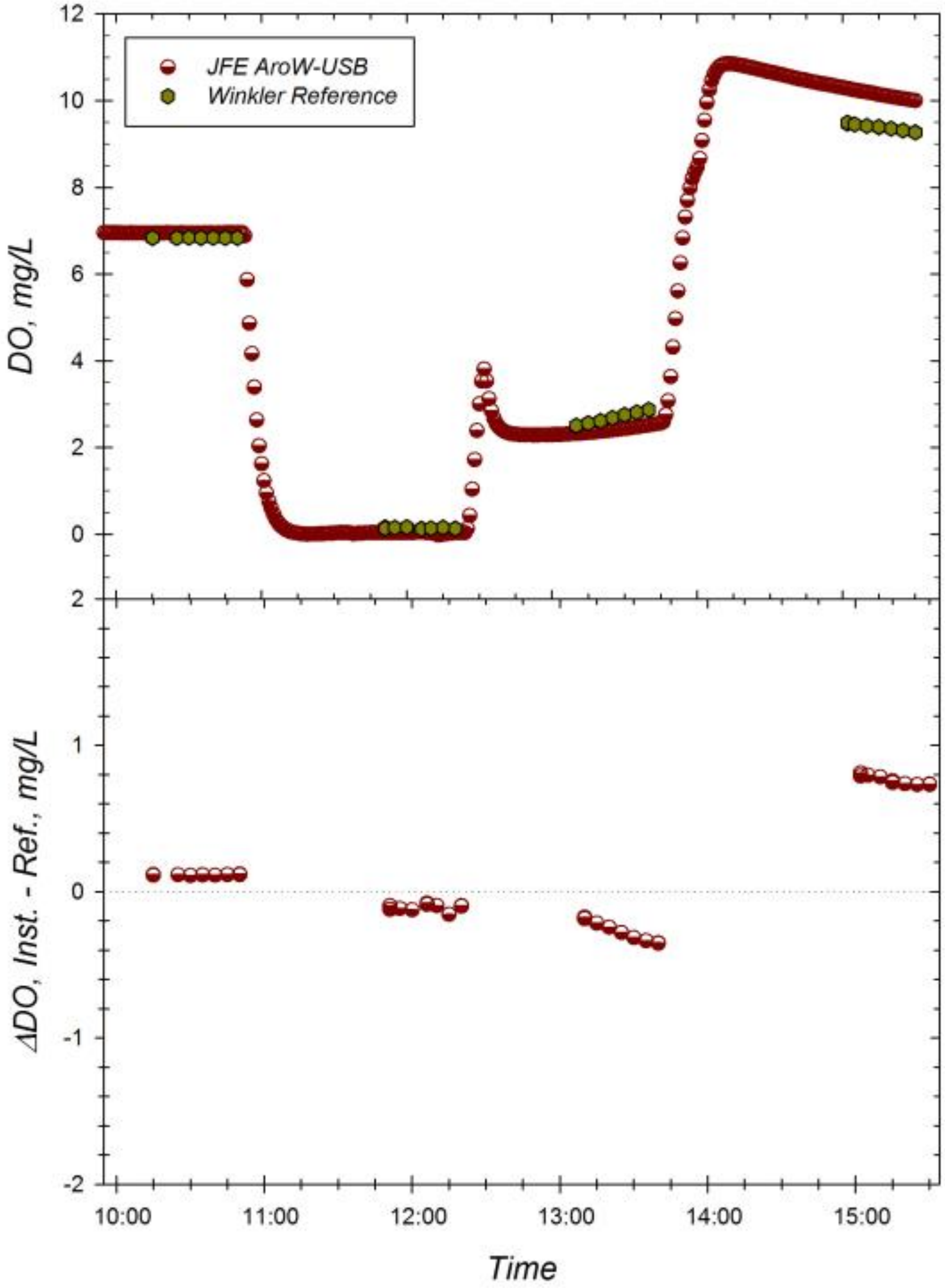

Figure T. Response of JFE AroW-USB (sn 0181) to changing dissolved oxygen concentrations at 30.5 ºC and S=9.28. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. The decay in instrument signal under high DO exposure likely reflects bubble contamination of DO sensor foil. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference.

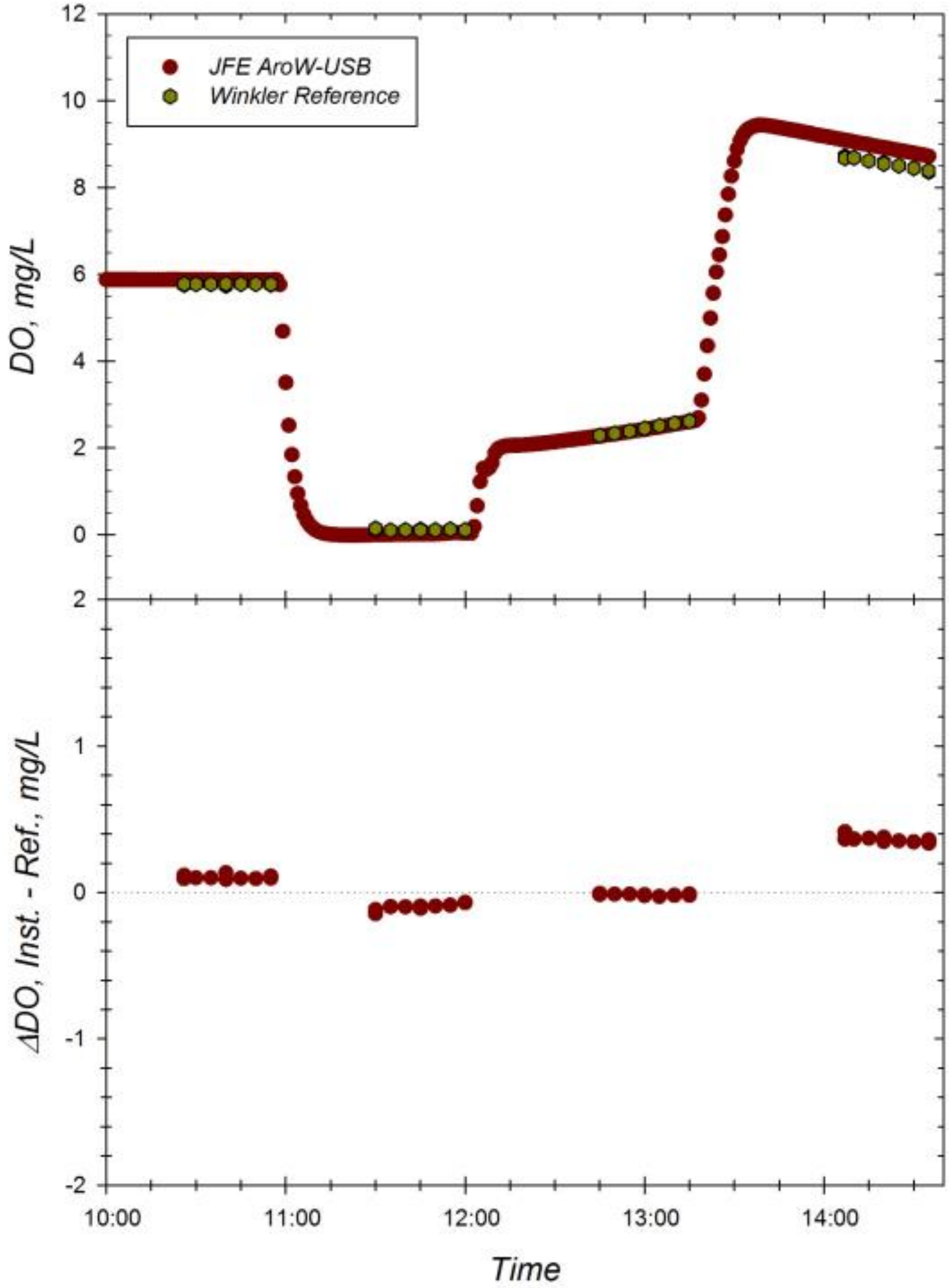

Figure U. Response of JFE AroW-USB (sn 0181) to changing dissolved oxygen concentrations at 30.6 ºC and S=34.43. *Top Panel*: Time series of instrument readings and Winkler reference samples drawn at level of sensor foil in well-mixed test tanks. *Bottom Panel*: Corresponding offset between instrument reading and Winkler DO reference sample during the trial. Dotted line represents zero offset from reference. Offsets beyond 2 mg/L are excluded.

The global response of the AroW-USB for all nine temperature-salinity trials is shown in figure V. In general readings are well correlated across the entire range of DO concentrations (hypoxic to supraoxic) covering all temperature and salinity ranges tested. Highlighted deviations are likely to reflect bubble contamination of the sensor foil. A linear regression of the data (n=257; $r^2 = 0.976$; $p<0.0001$) produced a slope of 0.969 and intercept of 0.114. Highlighted data were omitted from the regression because of expected bubble contamination.

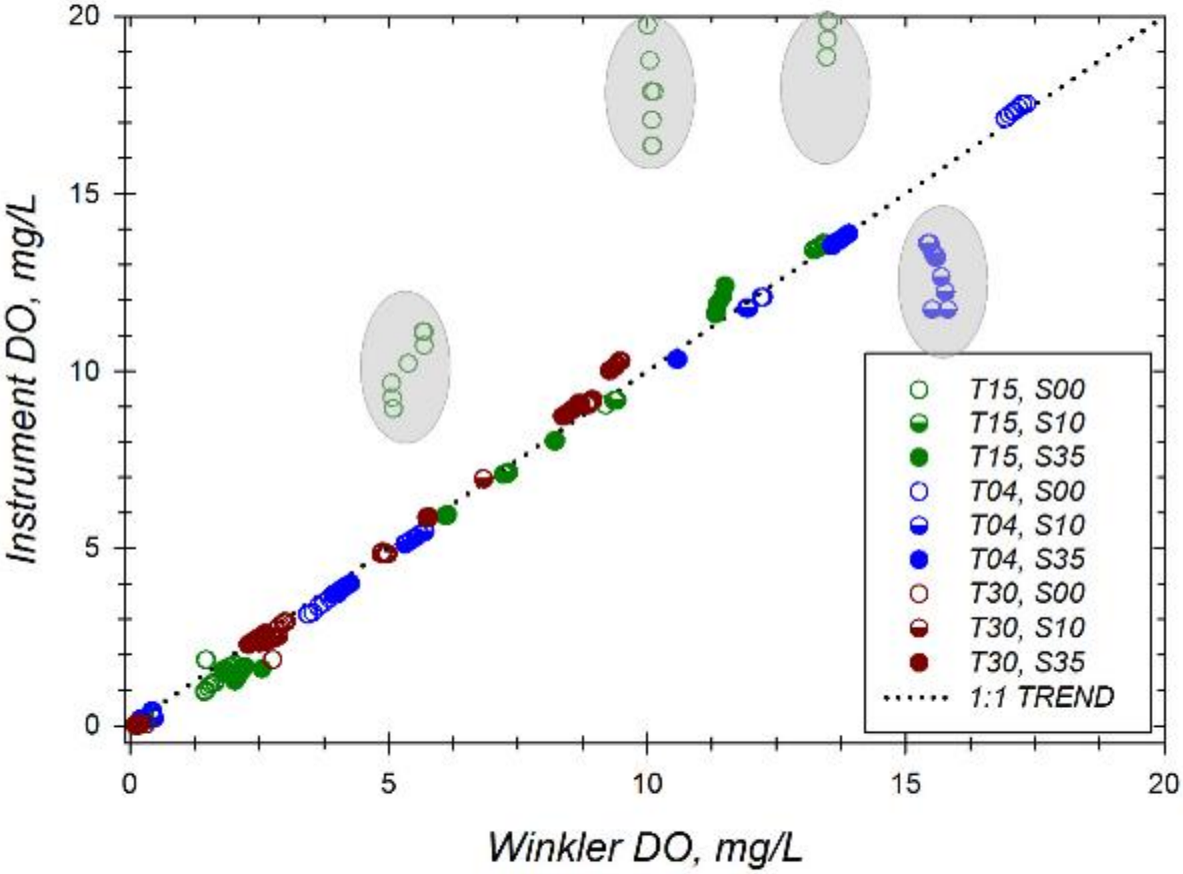

Figure V. Concordance of JFE AroW-USB (sn 0181) oxygen measures paired with discrete Winkler DO reference samples. Dotted line represents perfect agreement between the two measures.

Individual response slopes and intercepts of instrument versus reference DO for each of the laboratory trials are summarized in Table 6. Response slopes increased slightly with temperature when averaged across the three salinity trials. Means of the slopes were 0.996, 1.029, and 1.065 for 5, 15, and 30 °C respectively. Differences in slopes across salinity levels did not exhibit a consistent trend and were within 0.04 across the three levels.

Table 6. Summary of regression statistics for the AroW-USB versus reference sample response curves for each of the nine laboratory trials.

Test ID	R2	y Intercept	Slope
LT15S00*	0.996	-0.188	1.001
LT15S10	0.999	-0.226	1.021
LT15S35	0.995	-0.510	1.066
LT04S00	1.000	-0.280	1.021
LT04S10*	0.980	0.154	0.916
LT04S35	0.997	-0.274	1.050
LT30S00	0.997	-0.274	1.050
LT30S10	0.997	-0.313	1.093
LT30S35	1.000	-0.134	1.053

* Bubble contamination data excluded.

Results of the 56 day long-term stability and thermal stress challenge for the AroW-USB are shown in figure W. The instrument was maintained in a well circulated tank and oxygen content manipulated by alternately varying water temperature set point between 15 and 25 ^oC several times per week of deployment. The data completion result for the stability test was 100%. The time series of instrument readings at 15min intervals is plotted against discrete values for Winkler reference samples (*top panel*) along with the time series of the difference between instrument and reference measurements (*bottom panel*). The overall mean difference between measurements was -0.154 (s.d. = 0.319) mg/L. There was a minor drift in instrument accuracy over the deployment (slope = -0.006 mg/L/d; r^2=0.17) but the goodness of fit was low due to several outliers.

Results for the functional sensor response time assessment of the AroW-USB are shown in figure X. The top panel depicts the time series of 10s instrument reads during transfers between adjacent high (9.6 mg/L) and low (2.0 mg/L) DO water baths, maintained commonly at 15 ^oC. The bottom panel (*lower left*) depicts results fit with a 3 parameter exponential decay function: $DO_{rel} = DO_{relMin} + a\boldsymbol{e}^{-bt}$ and indicated τ calculated from fit. Data for low DO to high DO transitions (*lower right*) were treated similarly but normalized to steady state value in subsequent high DO tank and subsequently fit with an analogous 3 parameter exponential rise function: $DO_{rel} = DO_{relMin} + a(1\text{-}\boldsymbol{e}^{-bt})$ with indicated τ being directly calculated from fit. The calculated τ_{90} was 209 s during high to low transitions and 284 s for low to high transitions covering a DO range of approximately 8 mg/L at a constant 15 ^oC.

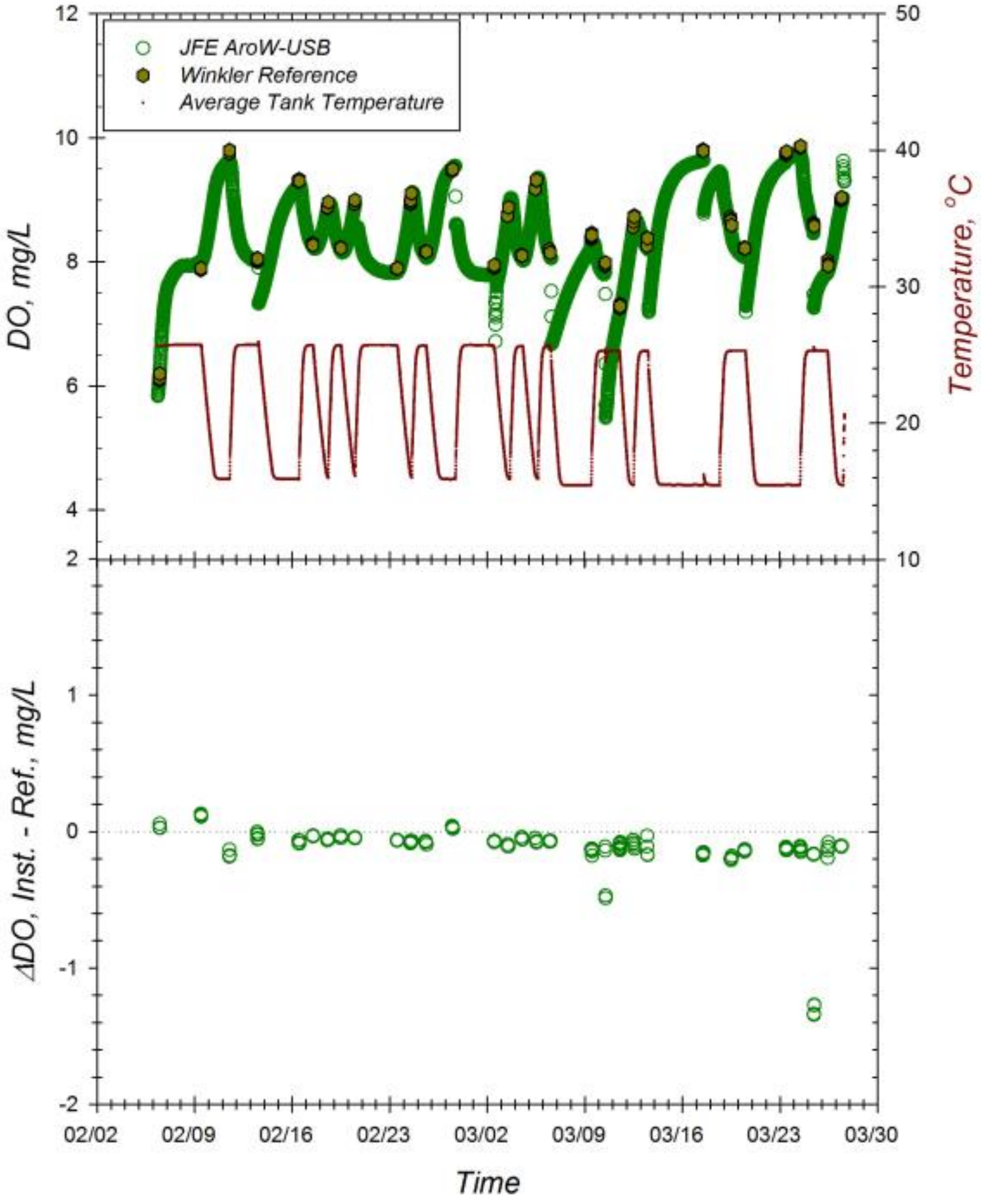

Figure W. Long-term stability and thermal stress challenge of the JFE AroW-USB (sn 0181). Instrument was maintained in well circulated tank and oxygen content manipulated by alternately varying water temperature set point between 15 and 25 °C several times per week of deployment. *Top Panel*: Time series of AROUSB instrument readings representing average of ten 1s burst reads at 15min intervals with discrete values for Winkler assay of grab samples taken at sensor level. *Bottom Panel*: Offset in paired DO readings of AROUSB and Winkler grab samples. Offsets beyond 2 mg/L (1 event) are excluded for clarity.

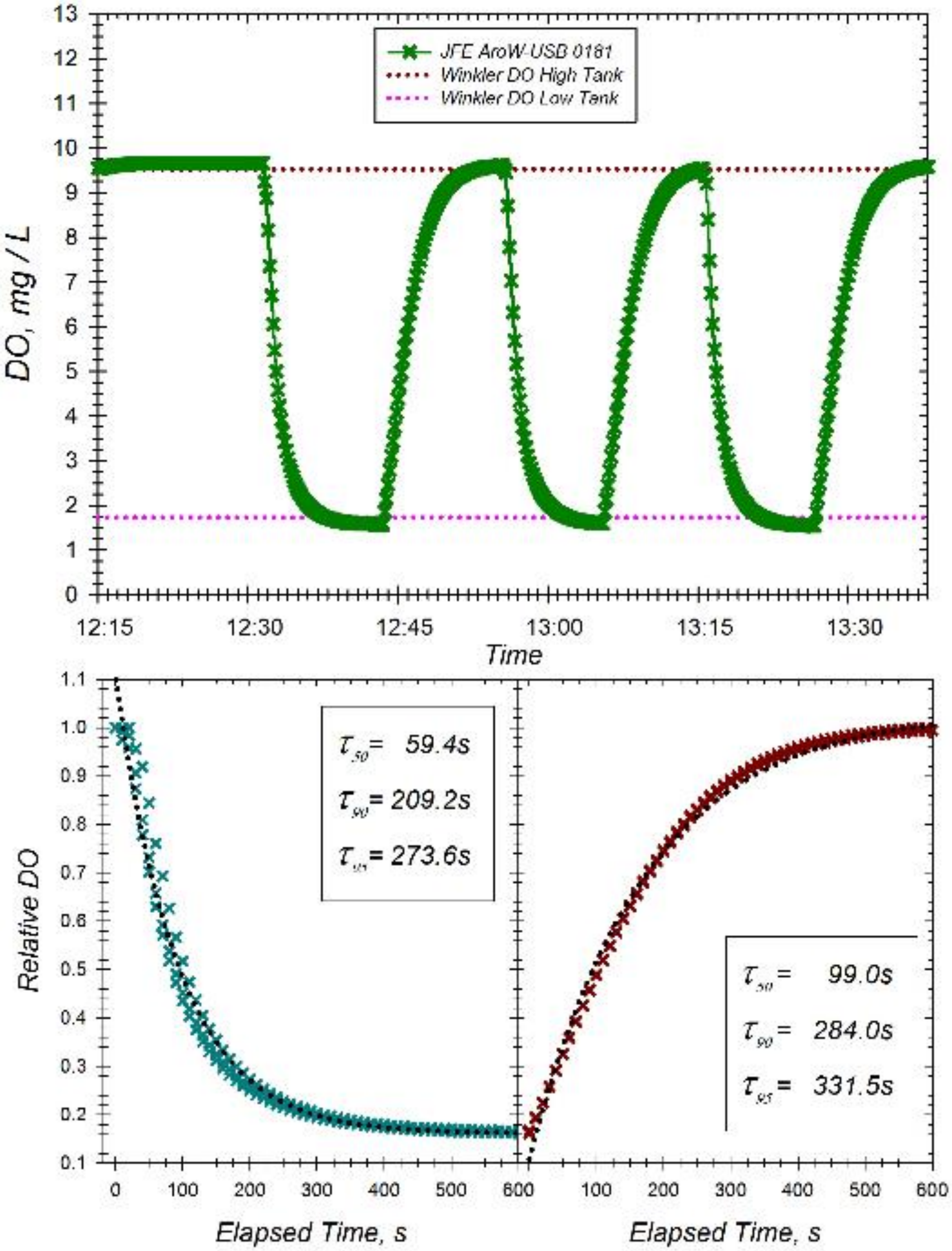

Figure X. Functional sensor response time for the JFE AroW-USB (sn 0181). *Top Panel*: Time series of 10s reads of instruments during transfers between adjacent high and low DO water baths, maintained commonly at 15 ºC. Indicated DO levels were maintained by air stone bubblers (high DO) or air+N2 (low DO) and verified by Winkler assays. *Bottom Panels*: All high DO to low DO transitions were normalized to last reading before condition switch for the indicated post transition time. Data was fit with a 3 parameter exponential decay function: $DO_{rel} = DO_{relMin} + a\boldsymbol{e}^{-bt}$ and indicated τ calculated from fit. Data for low DO to high DO transitions (*lower right*) were treated similarly but normalized to steady state value in subsequent high DO tank and subsequently fit with an analogous 3 parameter exponential rise function: $DO_{rel} = DO_{relMin} + a(1-\boldsymbol{e}^{-bt})$ with indicated τ being directly calculated from fit.

RESULTS of MOORED FIELD TESTS

Moored field tests were conducted to examine the performance of the JFE AroW-USB to consistently track natural changes in dissolved oxygen over extended deployment durations of 11-18 weeks. 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 AroW-USB was examined in three separate field deployment 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 7 The final reference temperature data was computed from the mean of two RBR thermistors and a SeaBird SBE that were mounted at the same sampling depth as the test instrument. Immediately before each deployment the JFE AroW-USB was exposed to a laboratory reference tank for 3-4 measurements to confirm good working status and provide an initial offset against its latest calibration (Table 8).

Table 7. Range and average for temperature, and salinity at each of the test sites during the sensor field deployments. Temperature and salinity were measured by at least 2 RBR temperature loggers and a SeaBird SBE 26 (or SBE26plus) mounted on the instrument rack for the duration of the deployment.

SITE (deployment period/duration)		**Temperature** °C)	**Salinity**
Keweenaw Waterway	Min.	0.04	0.0
9Jan – 22Apr	Max.	5.2	0.01
(n = 104 days)	Mean	0.7	0.01
Chesapeake Bay	Min.	19.3	4.7
20May – 5Aug	Max.	36.7	13.7
(n = 78 days)	Mean	25.6	10.9
Kaneohe Bay	Min.	23.1	27.3
24Sep – 21Jan	Max.	29.1	34.7
(n = 121 days)	Mean	25.8	33.4

Table 8. Results of the pre-deployment check for the AroW-USB for each deployment site. The DO of the reference solution is based on the mean of three Winkler titration analysis of the reference tank during the exposure. Ambient water was used for the reference solution at each deployment site.

Deployment Site	**Reference Solution**	**Measurement Temperature**	**Reference Solution DO (mg/L)**	**Instrument DO (mg/L)**
MTU	Site Water	11.7°C	10.193	10.35
CBL	Site Water	21.6°C	7.983	7.88
HIMB	Site Water	26.6°C	6.547	6.49

Moored Deployment at Michigan Tech Great Lakes Research Center

The 15 week deployment under ice took place from January 9 through April 22 in the Keweenaw Waterway adjacent to the Great Lakes Research Center in Houghton, MI. The deployment site is located at 47.12° N, 88.55° W, at the end of the pier at the Great Lakes Research Center docks. This site is located on the south side of the Keweenaw Waterway, and is connected to Lake Superior in both the NW and SE directions. The instrumentation rack was lowered off of the end of the pier with a ½ ton crane and rested on the bottom, under the ice, in 4.5m of water. A small shelter was constructed at the end of the pier to provide shelter during winter sampling efforts.

Photo 1. Aerial view of the Keweenaw Waterway (left) and dockside mooring deployment (right).

Time series results of ambient conditions for temperature and specific conductivity are given in figure 1. Temperature ranged from 0.04 - 5.3°C and specific conductivity from 87 - 137 µS/cm over the duration of the field test. The bottom panel displays the maximum difference recorded between all reference thermistors mounted at the same depth as the sensors sampling intakes as well as a meter above, at different locations across the mooring rack. The average temperature difference observed across the space of the mooring rack was 0.01°C with a maximum of 0.98°C. Differences between instrument and reference readings resulting from this variability should be minimized as the sampling bottle integrates across the mooring space.

Unexpected shifts between adjacent reference samples were noted on three occasions during the test. Upon inspection it was determined that these shifts occurred during changes in the batches of Winkler reagents. A correction to reference values was subsequently made based on the magnitude of change observed between the adjacent Winkler measurements after adjusting for ambient changes determined by the average of all seven DO sensors deployed on the mooring. Adjusted values are noted within each figure.

The AroW-USB operated successfully throughout the entire 15 week deployment and generated 9859 observations based on its 15 minute sampling interval, for a data completion result of 100%. Time series results of the AroW-USB and corresponding reference DO results are given in figure 2 (top panel). Ambient DO measured by the AroW-USB ranged from 8.669 to 15.076 mg/L compared to the range captured by reference samples of 10.918 to 14.007 mg/L. The sharp excursion observed the second week of April was real and was picked up by all instruments and correspondingly seen in specific conductance and temperature variability across the instrument rack (Fig. 1).

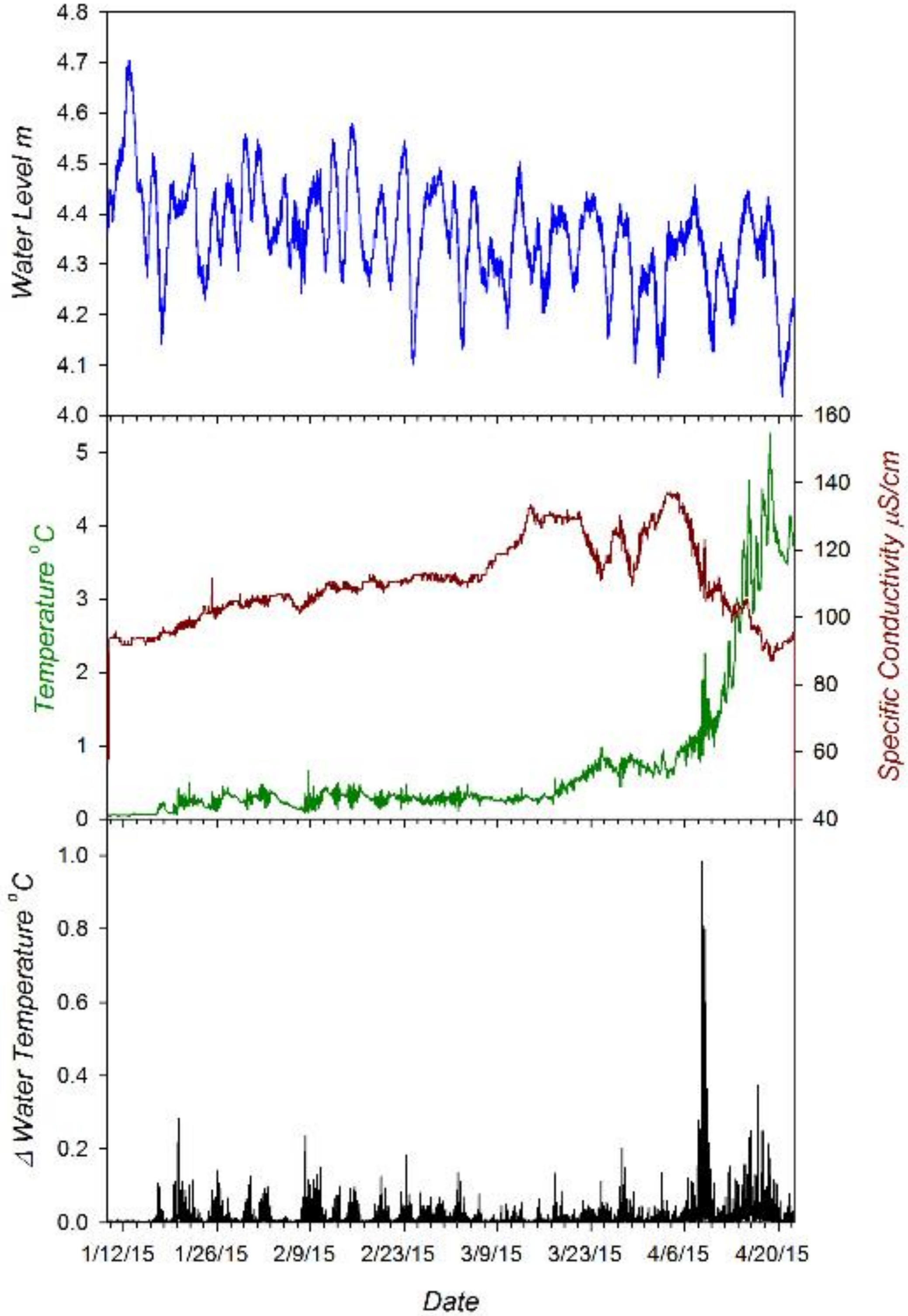

Figure 1. Environmental conditions encountered during deployment at the MTU site. Test sensor array deployed at 4.5 m fixed depth, variation in local water levels indicate active water flow around instruments (*Top Panel*). Variation in specific conductivity (red) and temperature (green) at depth of instrument sensor detected by an SBE 26 and two RBR Solo thermistors (*Middle Panel*). Temperature range determined from max-min temperatures detected by 4 RBR thermistors spanning instrument sensor array (*Bottom Panel*).

The time series of the difference between instrument and reference DO measurements for each matched pair (n=118 observations) is given in the bottom panel of figure 2. The average and standard deviation of the measurement difference over the total deployment was 0.170 ± 0.057 mg/L with a total range of 0.055 to 0.309 mg/L. The drift rate of instrument offset, estimated by linear regression (r^2=0.325), was -0.001 mg/L/d. This rate would include any biofouling effects as well as any electronic or calibration drift.

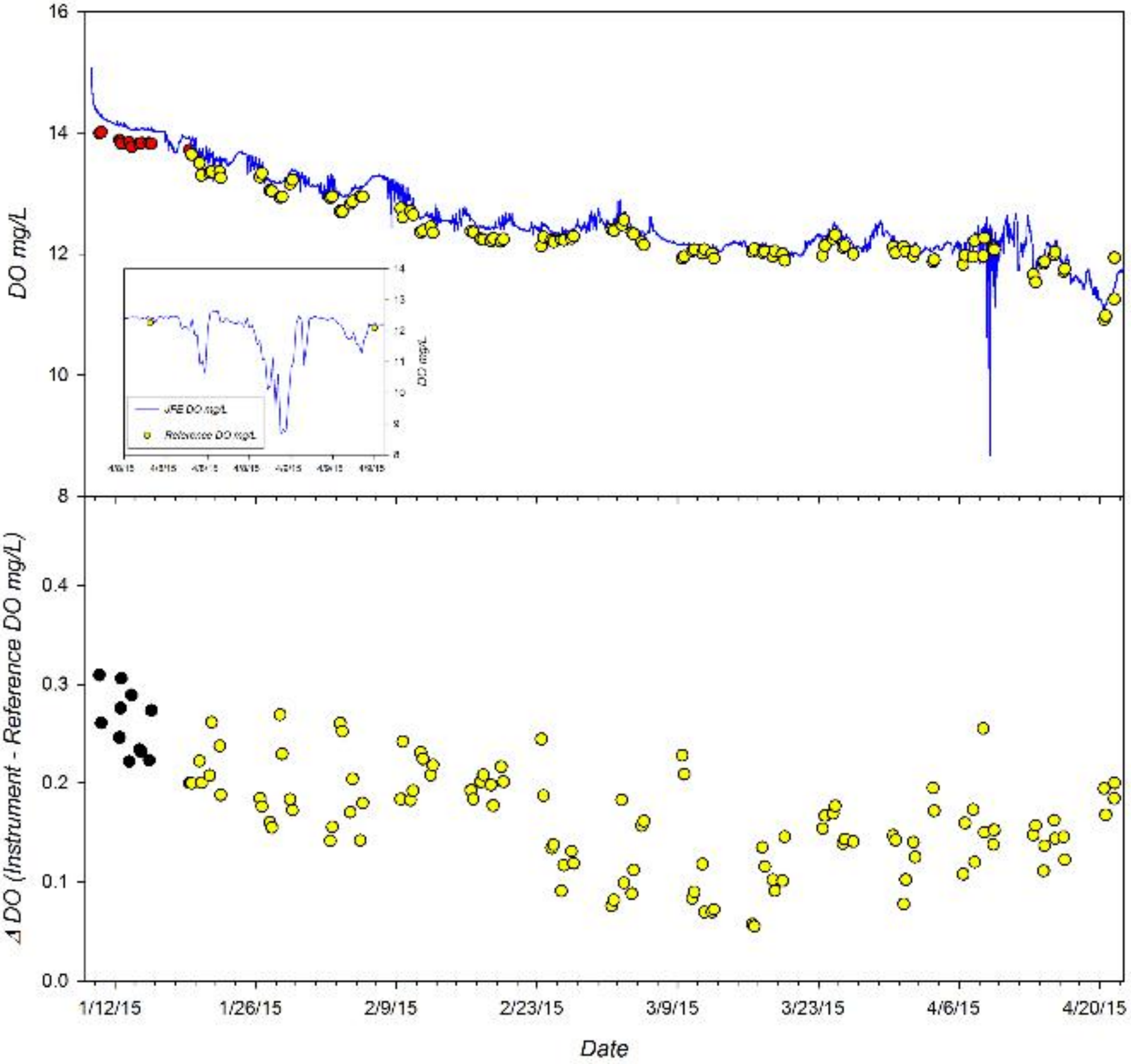

Figure 2. Time series of DO measured detected by a JFE AroW-USB deployed during the 15 week Great Lakes field trial. *Top Panel*: Continuous DO recordings from instrument (blue line) and DO of adjacent grab samples determined by Winkler titration (red circles; yellow circles represent adjusted reference values). *Bottom Panel*: The difference in measured DO relative to reference samples (Instrument DO mg/L – Ref DO mg/L) observed during deployment. *Insert:* Close up of excursion that occurred 4/8-4/9, bracketed by reference samples.

A cross-plot of the matched observations over the entire deployment is given in figure 3. Comparisons were limited to one month as a representation of a typical field deployment and to focus on initial instrument offset versus impacts of biofouling during an extended deployment. A linear regression of the data ($r^2 = 0.99$) produced a slope of 1.053 and intercept of -0.491.

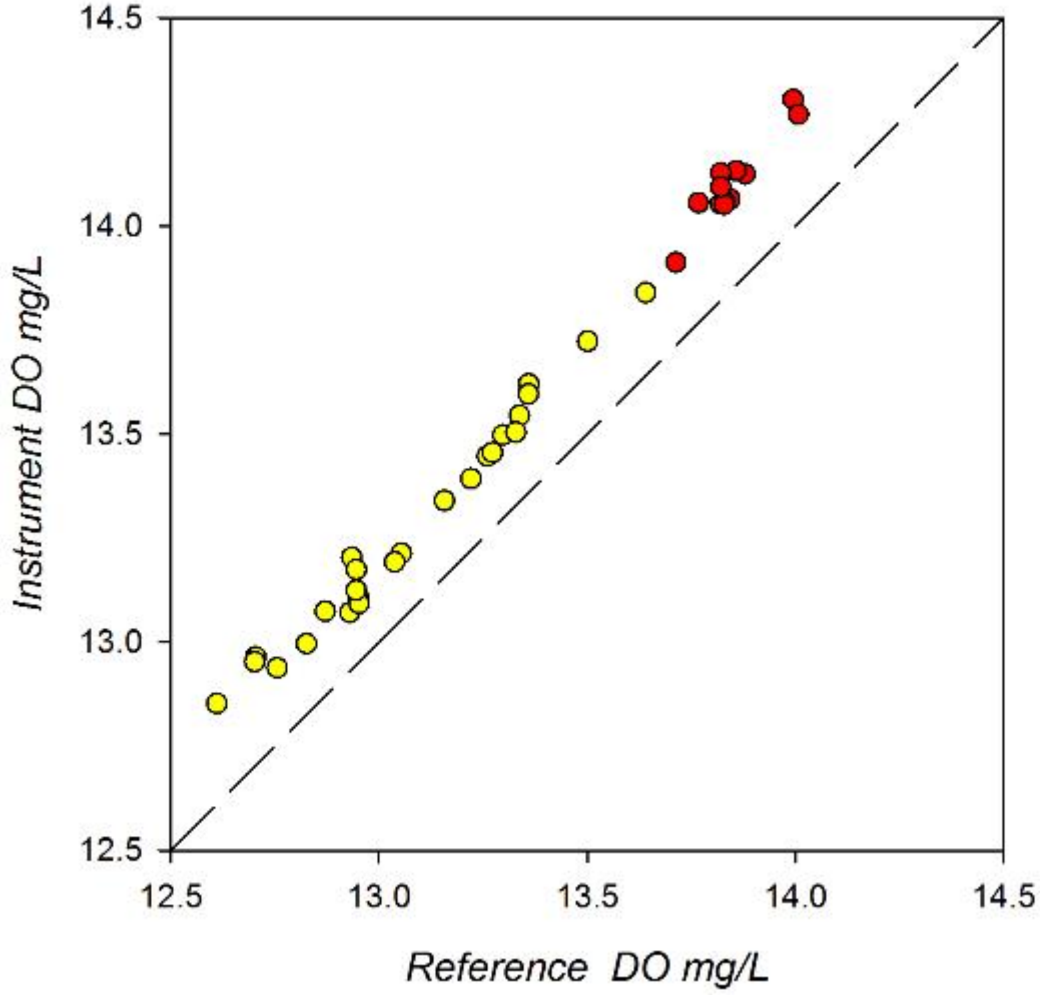

Figure 3. MTU GLRI field response plot of a JFE AROW-USB DO instrument compared to reference DO samples. The plotted line represents a 1:1 correspondence. (Yellow circles denote adjusted reference values).

Photos of test instrument before and after the field deployment to indicate potential impact of biofouling (Photo 2).

Photo 2. JFE AroW-USB prior to and following 15 week deployment under ice for the MTU field test.

Moored Deployment at Chesapeake Biological Laboratory (CBL)

An eleven week moored field test was conducted in Chesapeake Bay from May 20 to August 5, 2015. The deployment was located at 38.32°N, 76.45°W attached to the side of a pier at the mouth of the Patuxent River (Photo 3.) The site was brackish with an average water depth of the test site of 2.2 m.

Photo 3. 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 4. Temperature ranged from 19.2 to 30.1 and salinity from 4.7 to 13.7 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 0.21 ±0.25 ^{o}C, with a maximum of 3.26^{o}C. Differences between instrument and reference readings resulting from this variability should be minimized as the sampling bottle integrates across the mooring space.

The JFE AroW-USB operated successfully over the entire 11 week deployment from May 20 through August 4 and generated 7270 observations based on its 15 minute sampling interval, for a data completion result of 100%. Time series results of the AroW-USB and corresponding reference DO results are given in the top panel of figure 5. Ambient DO measured by the AroW-USB ranged from 2.610 to 14.510 mg/L compared to the range captured by the reference measurements of 4.370 to 10.858 mg/L. The bottom panel presents the time series of the difference between the AroW-USB and reference DO for each matched pair (n=142 observations out of a total of 142). The average and standard deviation of the measurement difference for the deployment was -0.056 ±0.131 mg/L, with the total range of differences between -0.375 to 0.392 mg/L. The drift rate of instrument offset, estimated by linear regression was -0.002 mg/L/d, but with a low goodness of fit due to high variability (r^2=0.16). This rate would include any biofouling effects as well as any electronic or calibration drift.

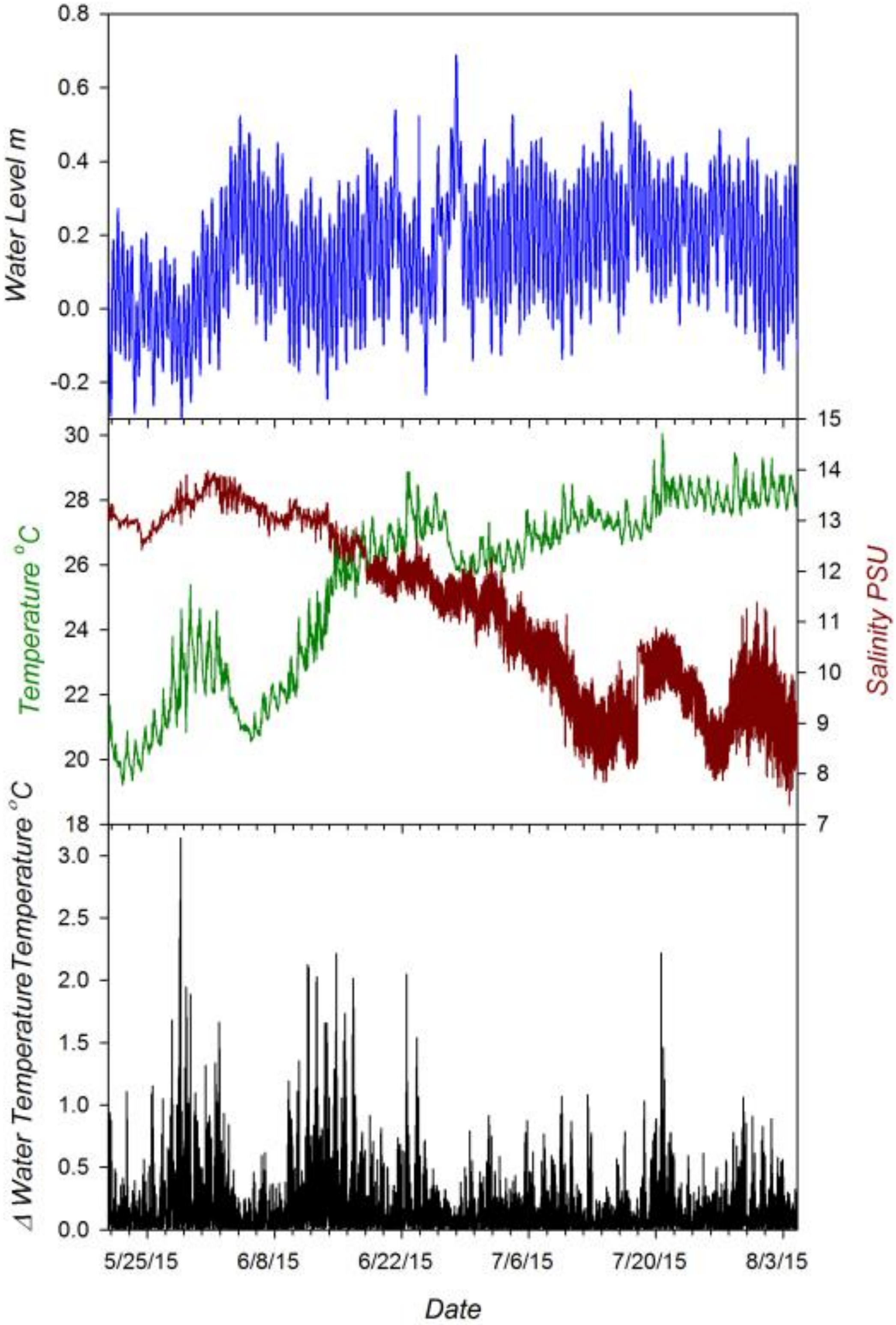

Figure 4. Environmental conditions encountered during the 11 week CBL floating dock deployment. Test sensor array deployed at 1 m fixed depth, variation in local tidal heights indicate active water flow around instrument (*Top Panel*). Variation in salinity (red) and temperature (green) at depth of instrument sensor detected by an SBE 26 and two RBR Solo thermistors (*Middle Panel*). Temperature range determined from max-min temperatures detected by RBR thermistors spanning instrument sensor array (*Bottom Panel*).

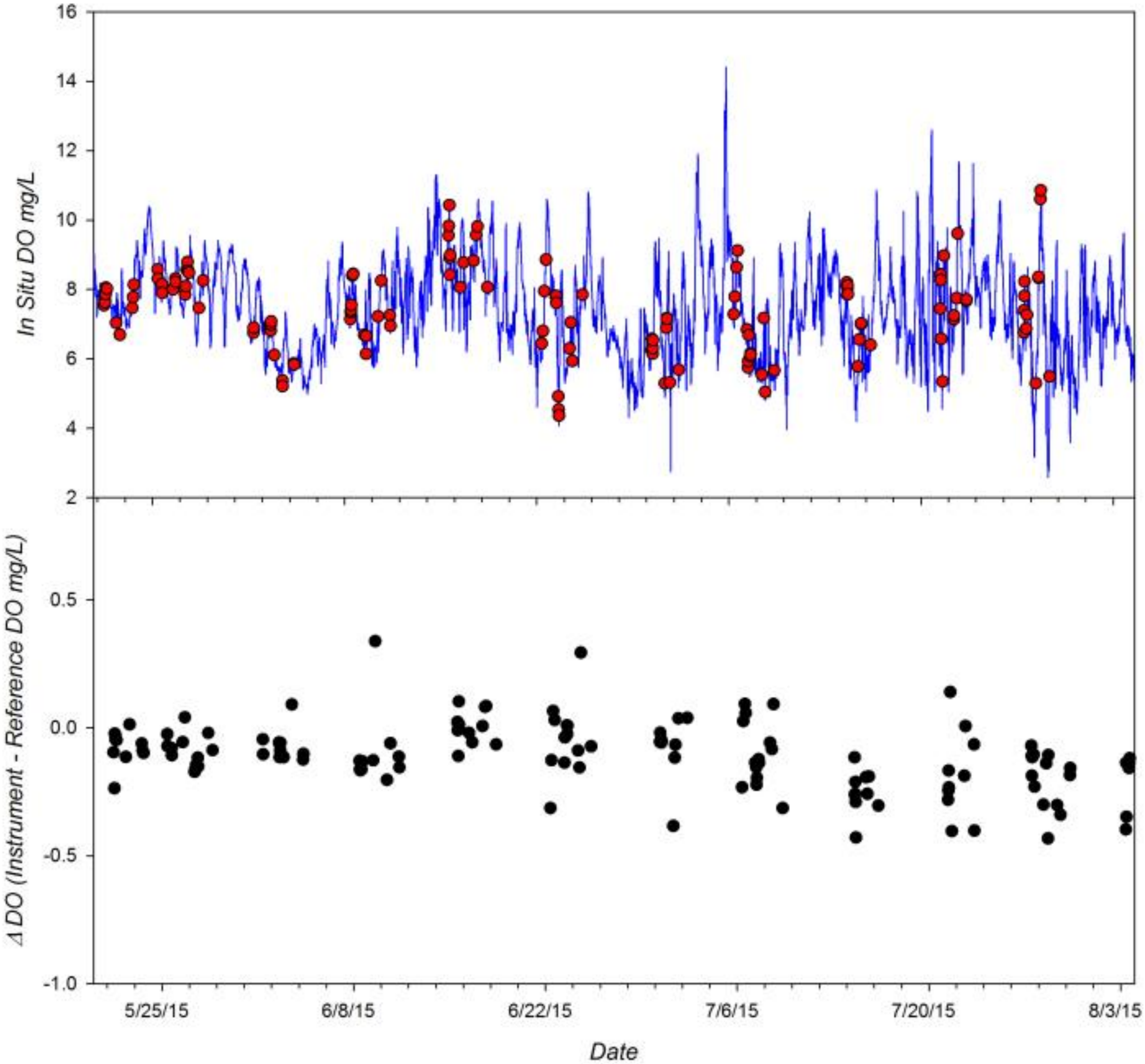

Figure 5. Time series of DO measured by the AroW-USB during the 12 week CBL field trial. *Top Panel*: Continuous DO recordings from instrument (blue line) and DO of adjacent grab samples determined by Winkler titration (red circles). *Bottom Panel*: The difference in measured DO relative to reference samples (Instrument DO mg/L – Ref DO mg/L) observed during deployment.

A cross-plot of the matched observations is given in figure 6. Comparisons were limited to one month as a representation of a typical field deployment and to focus on initial instrument offset versus impacts of biofouling during an extended deployment. A linear regression of the response curve ($r^2 = 0.99$) produced a slope of 1.014 and intercept of -0.177.

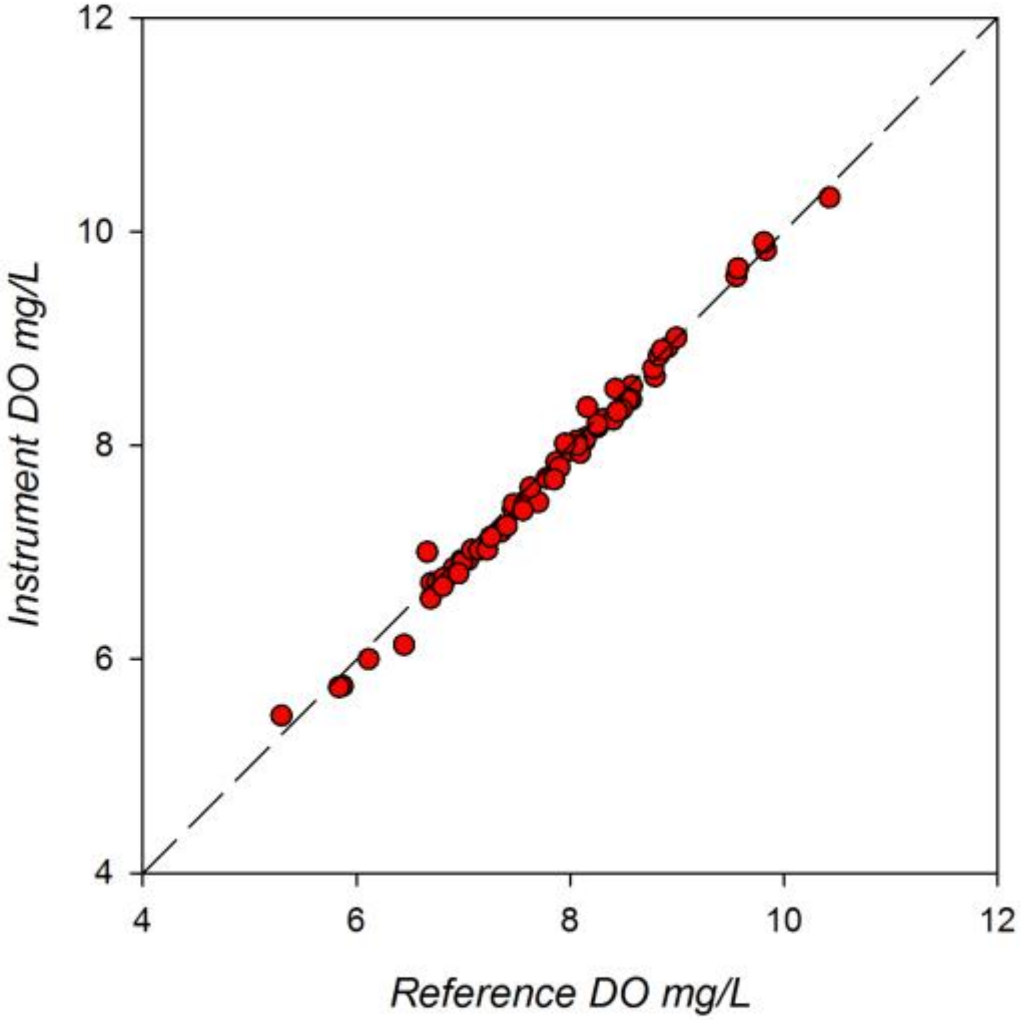

Figure 6. CBL field response plot a JFE AroW-USB compared to reference DO samples. The plotted line represents a 1:1 correspondence.

Photos of test instrument before and after the field deployment to indicate potential impact of biofouling (Photo 4).

Photo 4. The JFE AroW-USB instrument and following the 12 week CBL field trial.

Moored Deployment off Coconut Island in Kaneohe Bay, Hawaii

An 18 week long moored field test was conducted in Kaneohe Bay from September 22, 2015 to January 20, 2016. The deployment site was located at 21.43° N x 157.79° W, on the fringing reef flat surrounding Coconut Island (HIMB) in a depth of 3 meters (Photo 5). 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.

Photo 5. 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 7. Temperature at the sensor level ranged from 23.1 to 29.1 °C and salinity from 27.3 to 34.7 over the duration of the field test. The bottom panel displays the maximum difference recorded between all reference thermistors mounted at the same depth but located across the mooring rack. The average temperature difference observed across the space of the mooring rack was 0.15 ±0.17^{o}C, with a maximum of 1.23^{o}C. Differences between instrument and reference readings resulting from this variability should be minimized as the sampling bottle integrates across the mooring space.

The AroW-USB reported data throughout the entire deployment and generated 5653 observations based on its 30 minute sampling interval, but 785 observations fell outside of a natural ambient range (denoted by values more than ± 2 mg/L than the nearest reference sample) and were excluded from comparison against reference samples. The data completion result based on this accepted subset of data was 86%. Time series results of the AroW-USB and corresponding reference DO results are given in figure 8. Ambient DO measured by the AroW-USB ranged from 2.329 to 10.996 mg/L compared to the range captured by the reference measurements of 3.630 to 9.851 mg/L. The average and standard deviation of the differences between instrument and reference readings (limited to ± 2.0mg/L DO; n=75 of 129 potential observations) were 0.367 ± .637 mg/L, with a total range in the differences of -0.720 to 1.991 mg/L. The drift rate of instrument offset for the accepted subset of data, estimated by linear regression (r^2=0.74), was 0.165 mg/L/d. This rate would include any biofouling effects as well as any electronic or calibration drift.

.

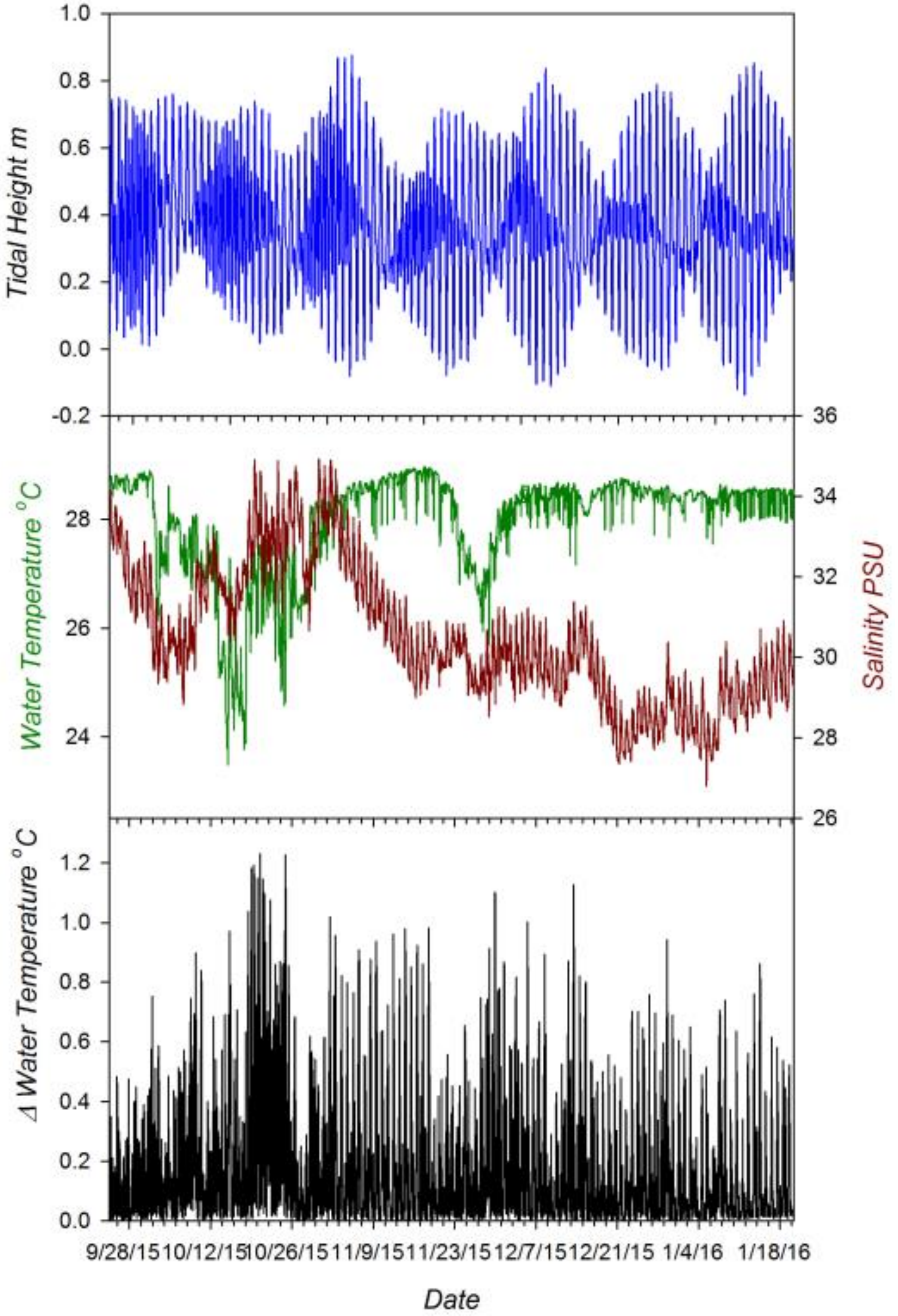

Figure 7. Environmental conditions encountered during the 4 month HIMB deployment on the fringing reef flat off Coconut Island Test sensor array deployed at 1 m fixed depth, variation in local tidal heights indicate active water flow around instrument (*Top Panel*). Variation in salinity (red) and temperature (green) 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 thermistors spanning instrument sensor array (*Bottom Panel*).

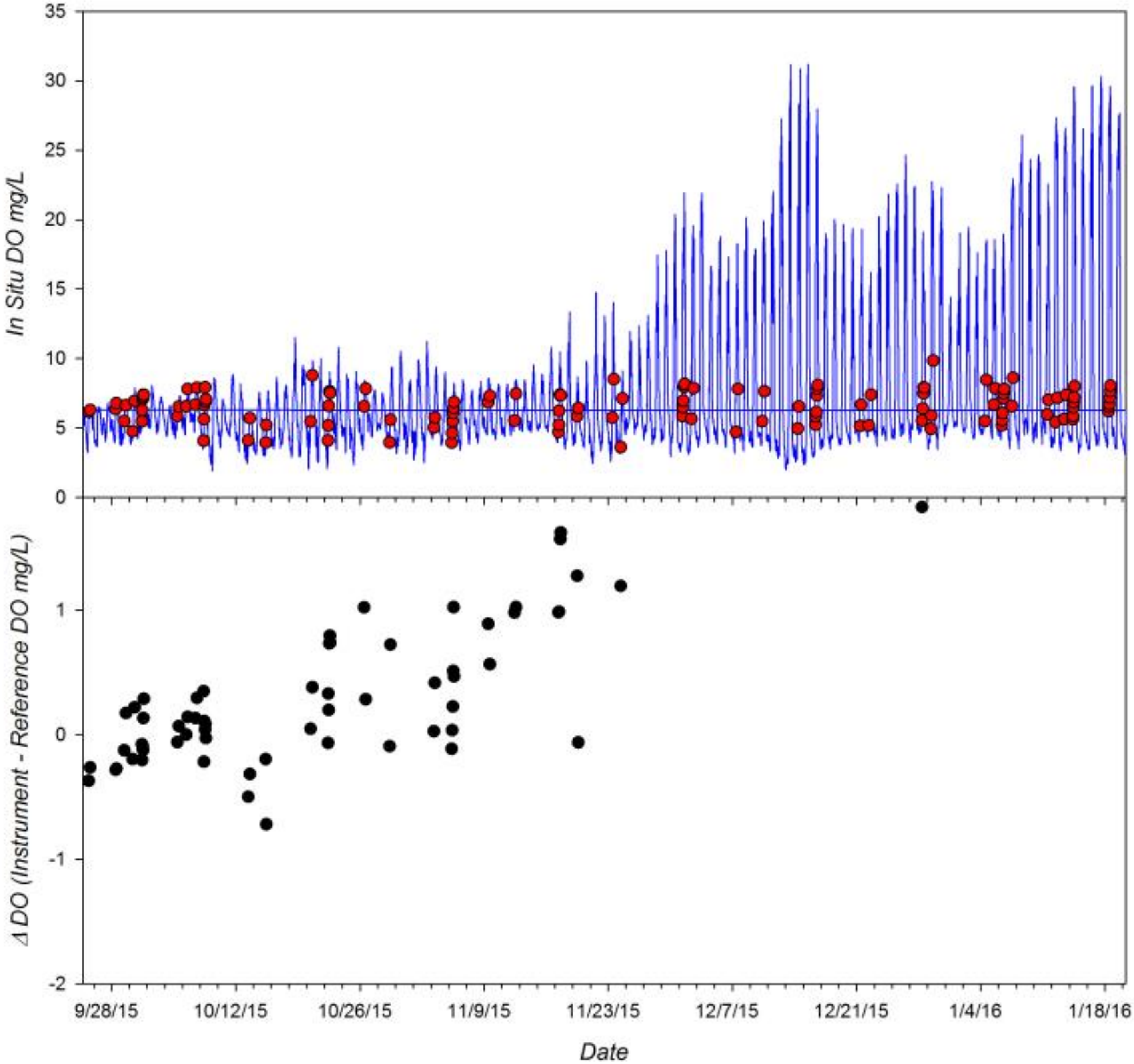

Figure 8. *Top panel:* Time series of DO measured by a JFE AroW-USB instrument deployed during the 18 week HIMB field trial. Continuous DO recordings from instrument (blue line) and DO of adjacent grab samples determined by Winkler Titrations (red circles.) *Bottom Panel:* Time series of the difference between the JFE AroW-USB and reference measurements (Instrument DO mg/L – Reference DO mg/L) for each matched pair in the range of ± 2 mg/L DO (n=75 of a total possible 129 observations).

A cross-plot of the matched observations for the first month of deployment is given in figure 9. Results reported after 11/25/16 were not included as they were deemed outside an acceptable range from true ambient concentrations. A linear regression of the subset data ($r^2 = 0.95$) had a slope of 1.12 and intercept of -0.514

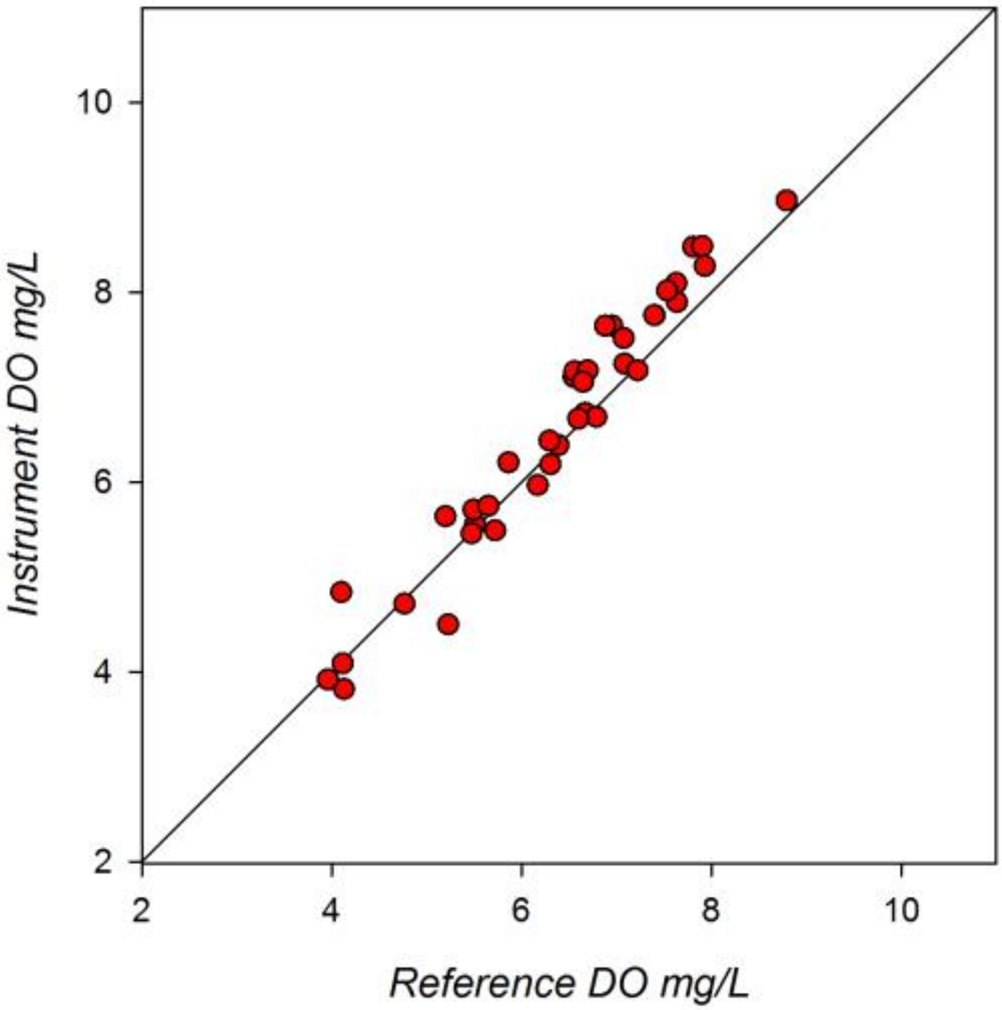

Figure 9. HIMB field response plot of a JFE AroW-USB instrument compared to reference DO samples determined by Winkler Titration. The plotted line represents a1:1 correspondence.

Photos of test instrument before and after the field deployment to indicate potential impact of biofouling (Photo 6).

Photo 6. The JFE AroW-USB prior to and following the 18 week HIMB field trial.

A global summary of instrument versus reference readings for all three deployment sites is plotted in figure 10. The AroW-USB response curves were highly linear across all three salinity ranges including freshwater, brackish water, and oceanic water. Good agreement between instrument and reference measurements was observed over a wide range of DO conditions varying between 4 to 14 mg/L. There was a slightly higher offset for the cold freshwater test in Houghton, MI, and the overall variability was slightly higher for the oceanic test in Kaneohe Bay. A linear regression of the composited data ($r^2 = 0.995$; $p<0.0001$)) had a slope of 1.038 and intercept of -0.290.

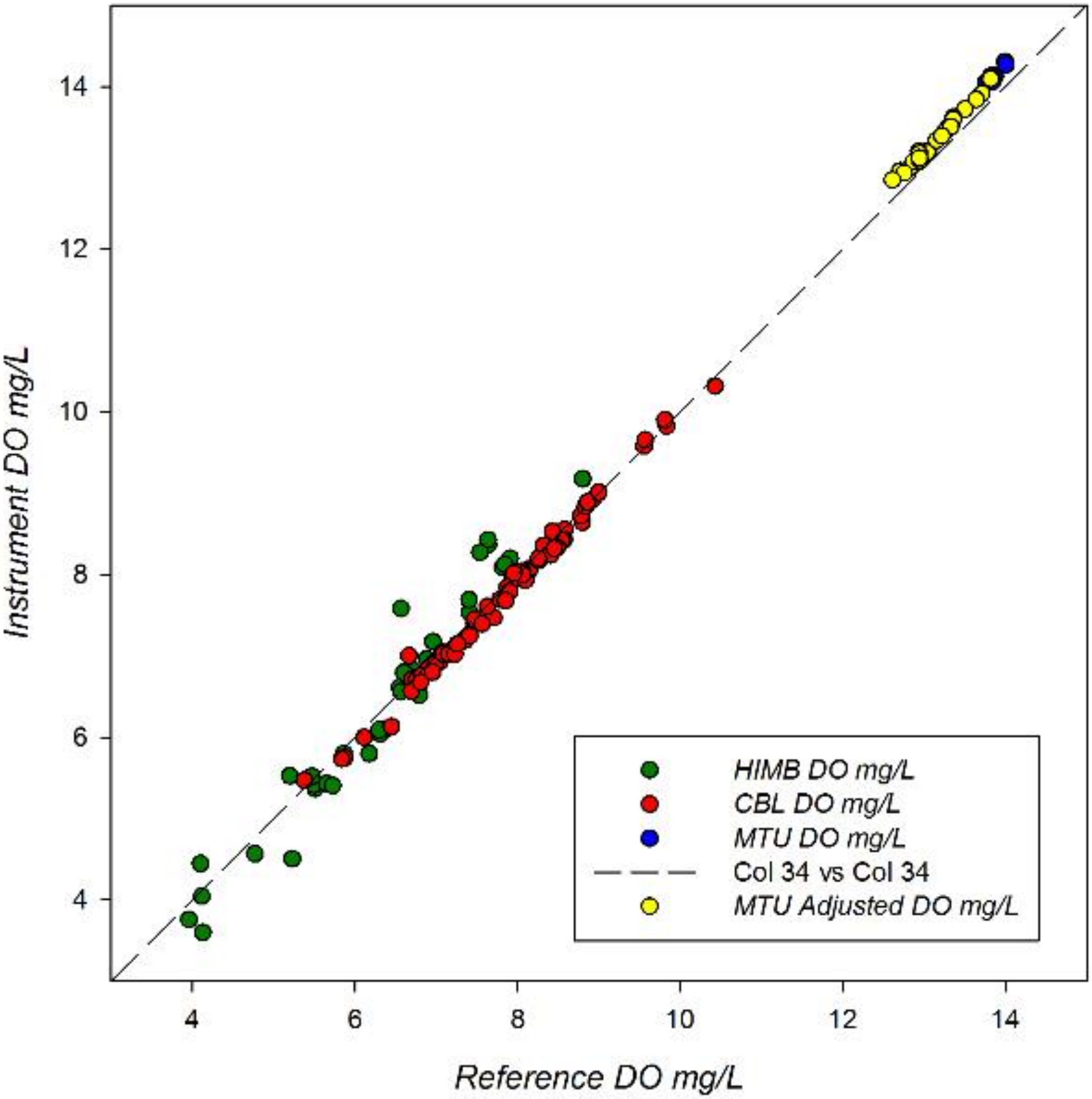

Figure 10. Global response plot for the AroW-USB instruments observed during the three ACT field trials. Black dotted line represents a 1:1 correspondence.

Profiling Deployment in the Great Lakes

Great Lakes profiling tests for the JFE AroUSB were performed aboard the R/V Laurentian at two separate locations in order to experience both normoxic and hypoxic hypolimnion (Photo 7). The normoxic site was located in Lake Michigan (43.184°N, 86.456°W) within a 64m deep water column, while the hypoxic site was in Muskegon Lake at 43.22°N, 86.30°W with a 24m deep water column.

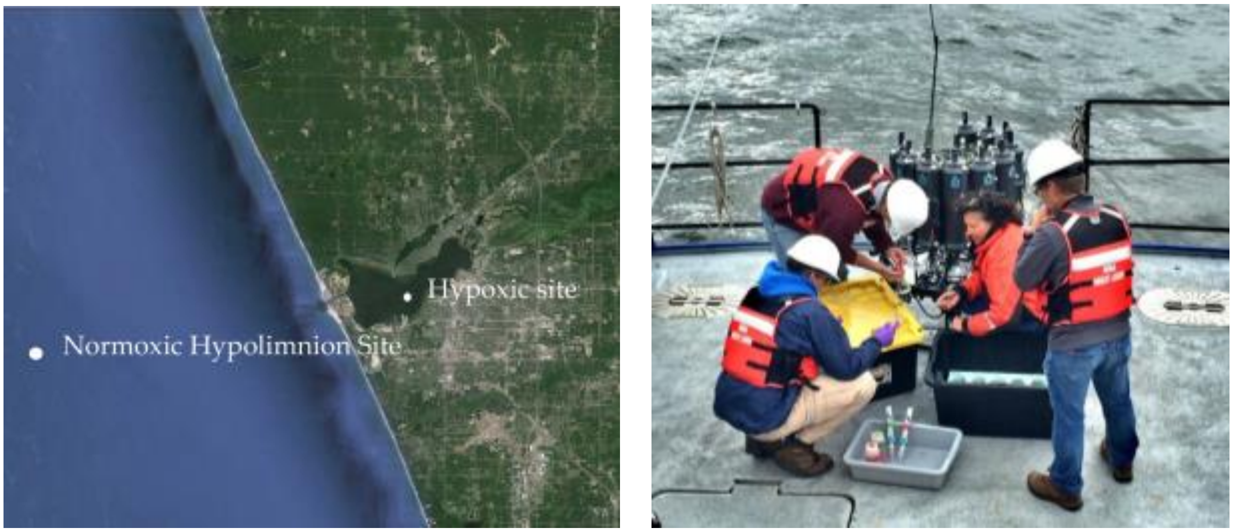

Photo 7. Aerial view of Lake Michigan and Muskegon Lake (left) and profile rig used for GL profiling

Two profiling trials were conducted at each location. The first trial involved equilibrating test instruments at the surface (3m) for ten minutes and then collecting three Niskin bottle samples at one minute intervals. Following the third sample, the rosette was quickly profiled into the hypolimnion where samples were collected immediately upon arrival and then each minute for the next 6 minutes (figure 11 and 13). The second trial was performed in the reverse direction where instruments were equilibrated at depth, three samples collected, and then profiled into the surface and sampled over the next 7 minutes. Note for Muskegon Lake cast 1 was discarded due to bottle misfires, so a third cast was performed to repeat the original pattern (Fig. 11).

Temperature and DO concentration profiles for Muskegon Lake and Lake Michigan are given in figures 12 and 14, respectively. In Muskegon Lake, the temperature ranged from 21.0 ^{o}C at the surface to 13.5 ^{o}C in the hypolimnion, with corresponding DO concentrations of 7.8 and 2.8 mg/L, respectively. In Lake Michigan, the temperature ranged from 21.0 ^{o}C at the surface to 4.1 ^{o}C in the hypolimnion, with corresponding DO concentrations of 8.6 and 12.6 mg/L, respectively.

Results for the JFE AroUSB for the two Muskegon Lake trials are shown in figure 15. The instrument exhibited a negative bias in the colder, low DO hypolimnion and a positive bias in the warm, high DO surface. Sensor equilibration time was slightly greater going from surface to hypolimnetic conditions (*right side, cast 3*). The range in measurement differences between instrument and reference was -0.42 to 0.34 mg/L for cast 2 and -0.75 to 0.27 mg/L for cast 3.

Results for the JFE AroUSB for the two Lake Michigan trials are shown in figure 16. The instrument exhibited a positive bias in both portions of the water column but the magnitude was higher in the cold high DO hypolimnion. Sensor equilibration time was similar between both trials, whether equilibrated at surface or depth. The range in measurement differences between instrument and reference was -0.16 to 0.53 mg/L for cast 1 and 0.18 to 0.50 mg/L for cast 2.

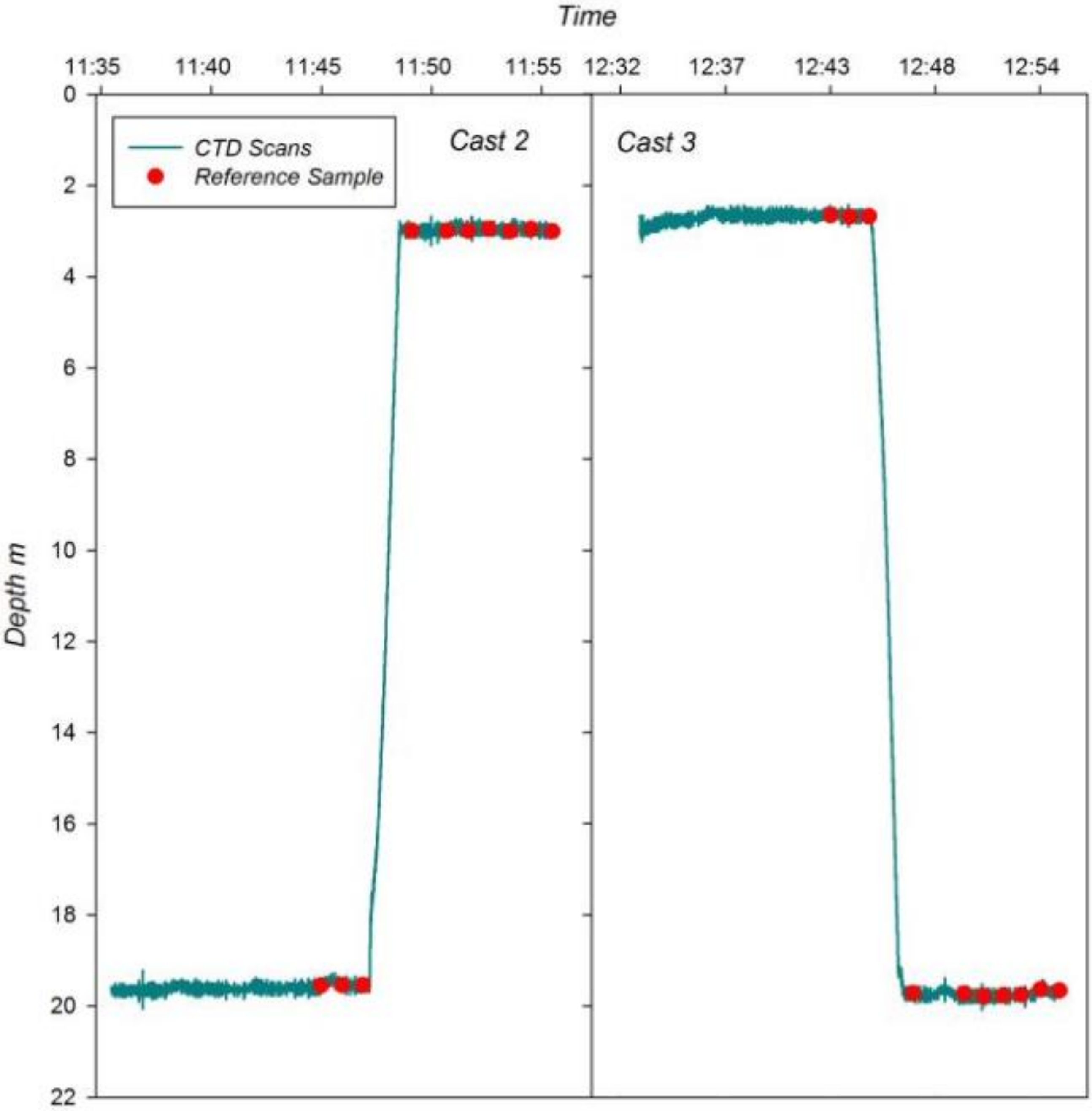

Figure 11. Time series of the rosette profiling and Niskin bottle sampling for the two profiling trials in Muskegon Lake.

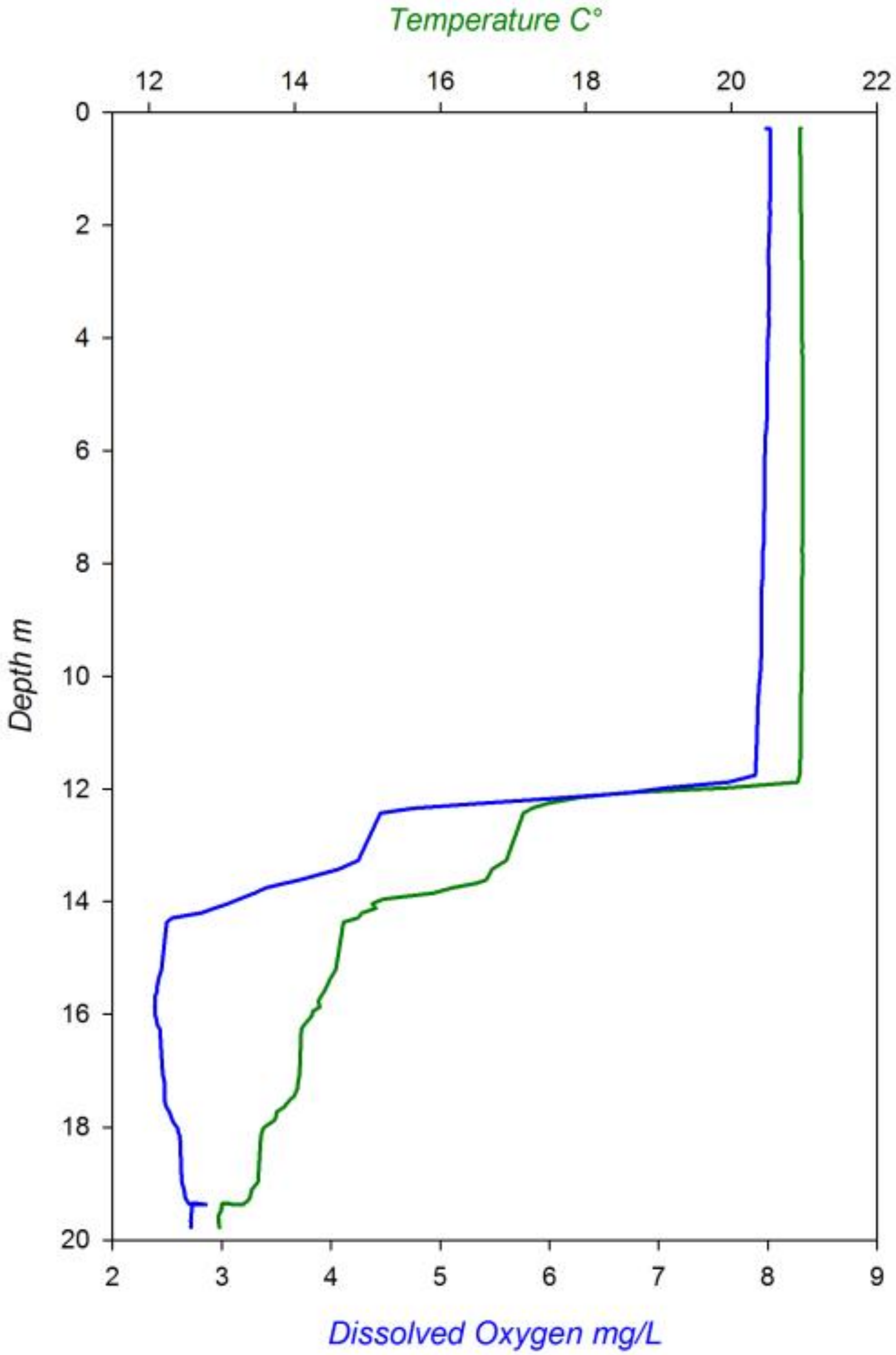

Figure 12. Temperature and DO concentration profiles for Muskegon Lake. Values were obtained from the average of multiple instruments sampling at 1s frequency. The hypolimnion was approximately 8m thick and contained depressed DO levels of less than 3 mg/L.

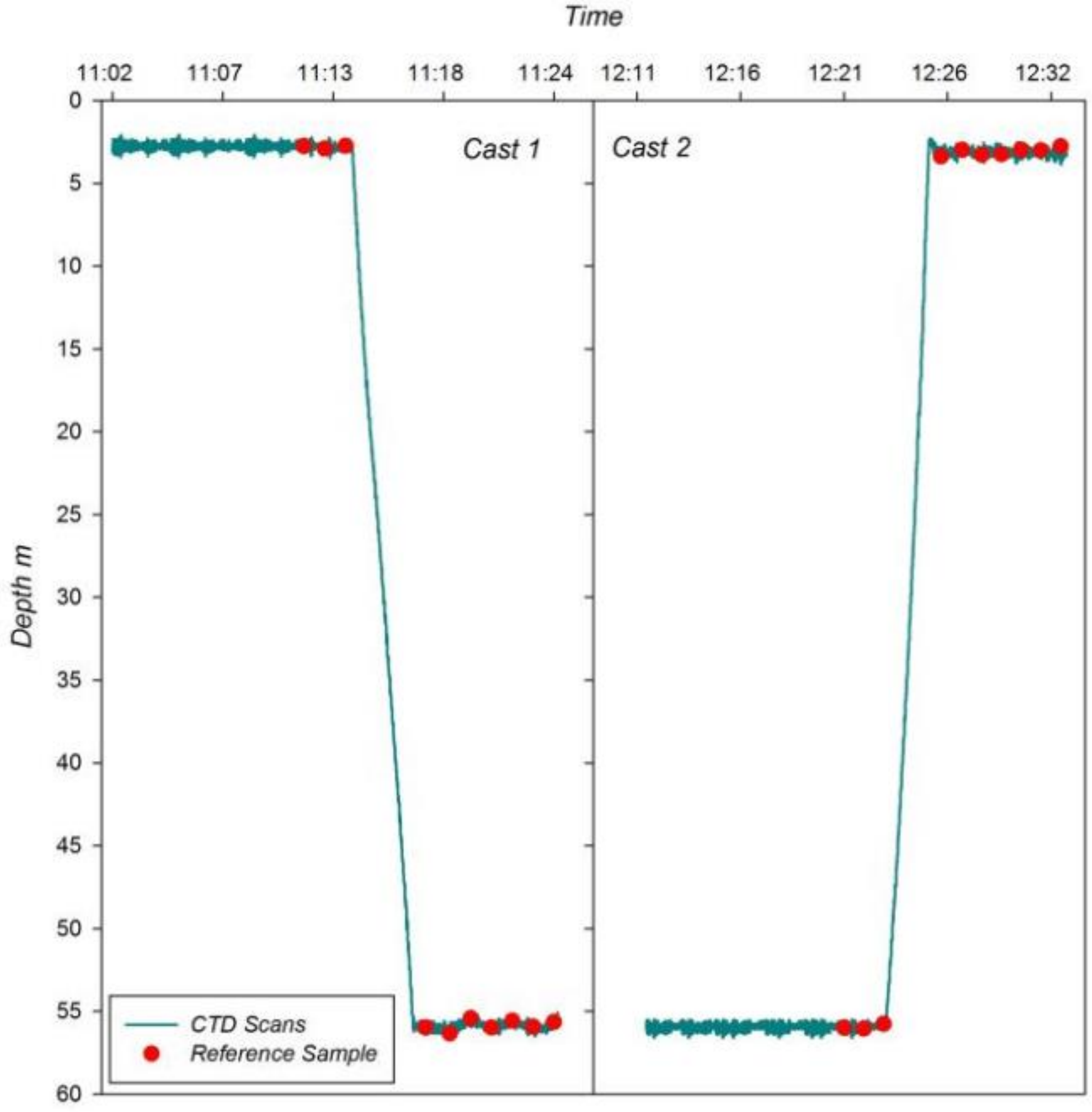

Figure 13. Time series of the rosette profiling and Niskin bottle sampling for the two profiling trials in Lake Michigan.

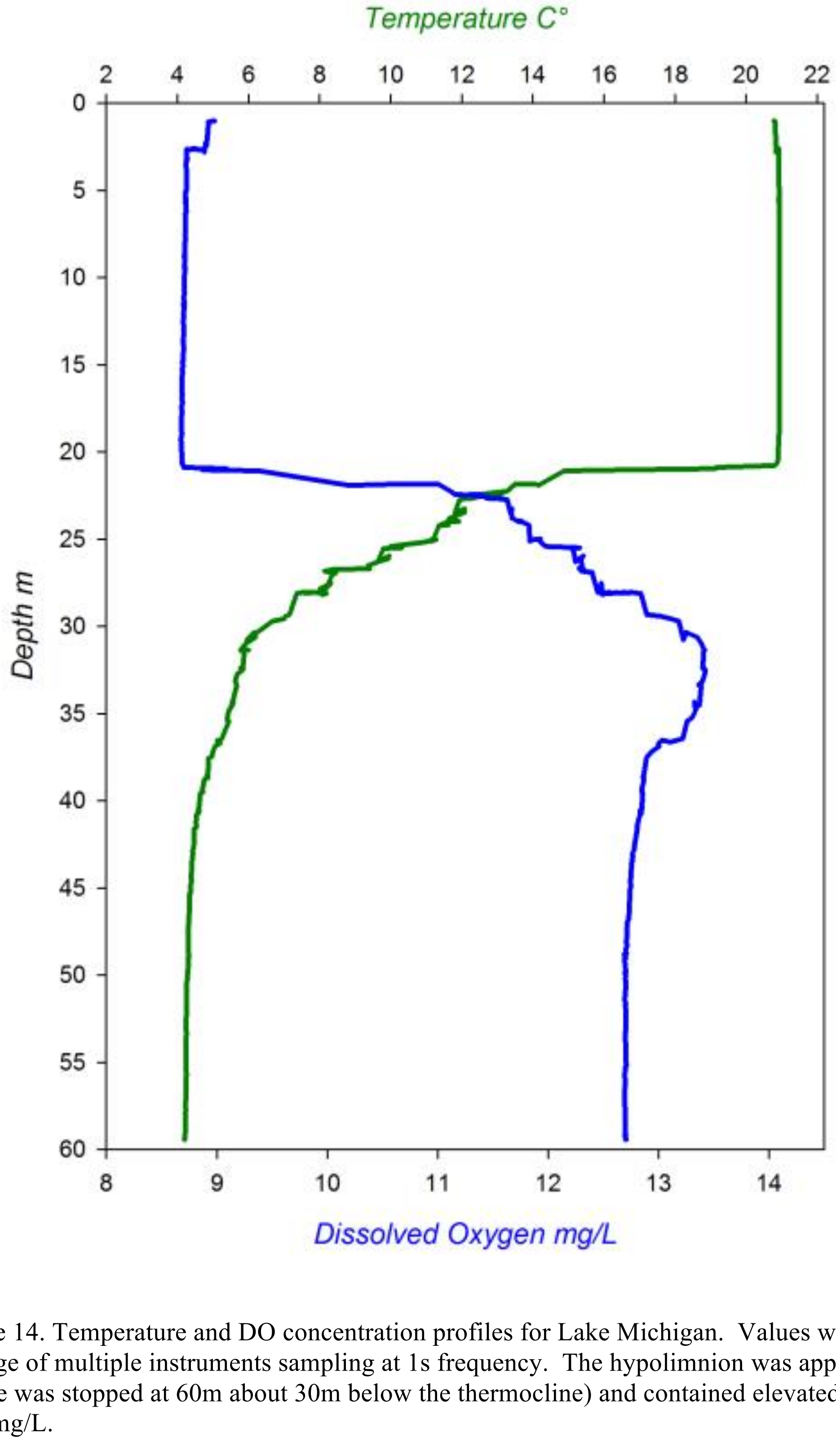

Figure 14. Temperature and DO concentration profiles for Lake Michigan. Values were obtained from the average of multiple instruments sampling at 1s frequency. The hypolimnion was approximately 40m thick (the rosette was stopped at 60m about 30m below the thermocline) and contained elevated DO levels of more than 12.6 mg/L.

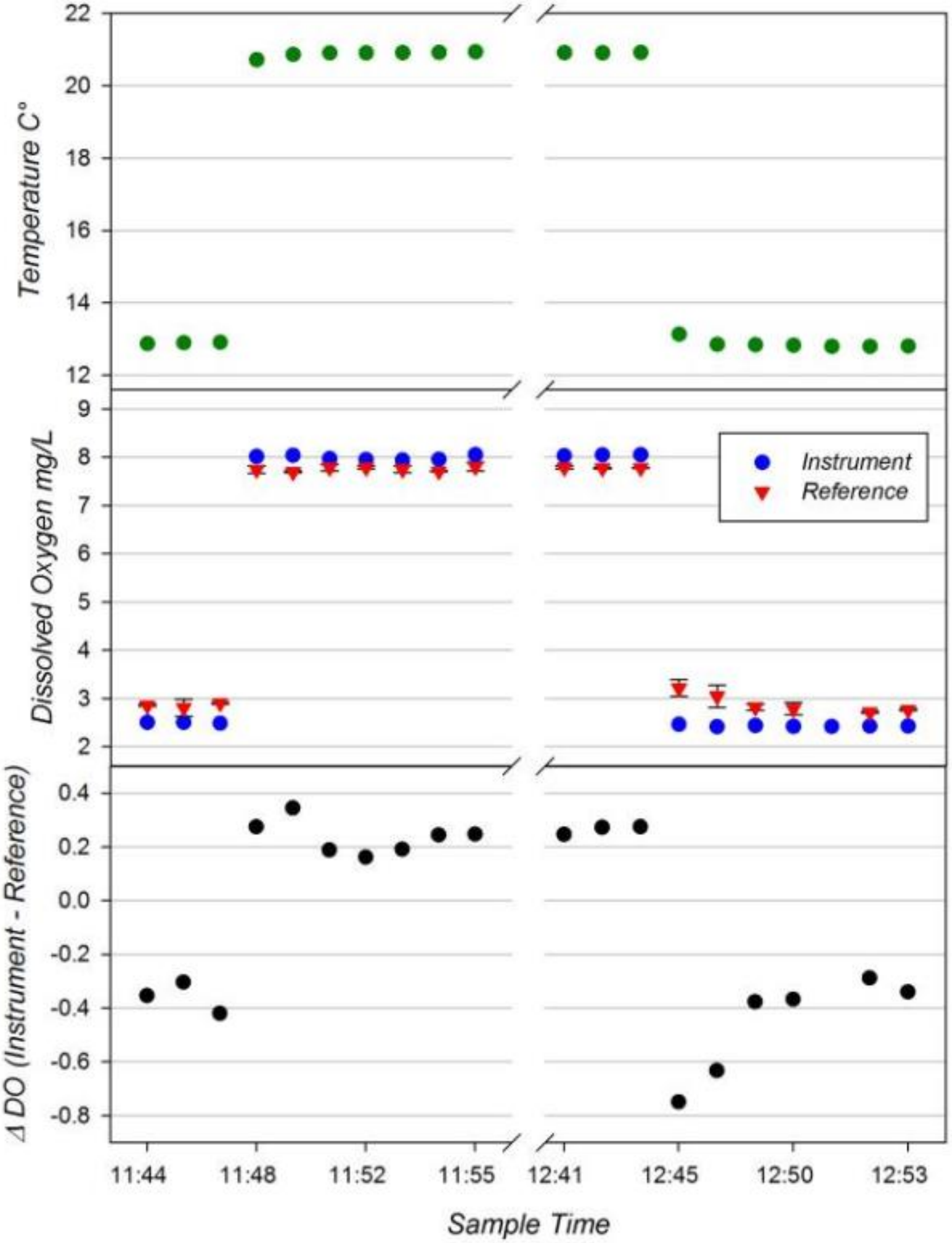

Figure 15. *Top panel:* Muskegon Lake temperature data at sample snap. *Middle panel:* DO recordings from instrument (blue dot) and DO of corresponding reference samples determined by Winkler Titrations (red triangles.) *Bottom panel:* Time series of the difference between the JFE AroUSB and reference measurements for each matched pair. Cast 2 and 3 are plotted on the same graph and separated by the axis break.

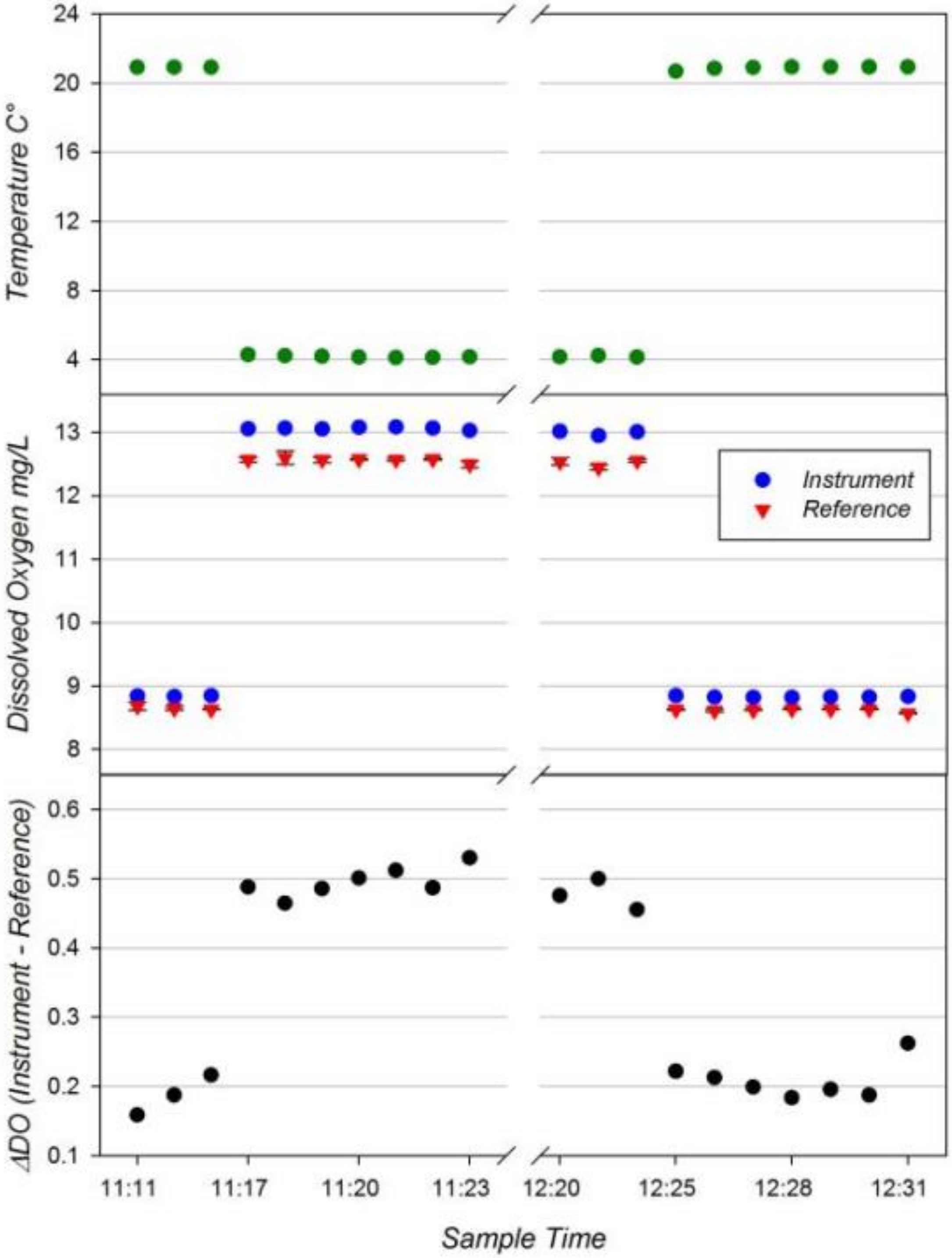

Figure 16. *Top panel:* Lake Michigan temperature data at sample snap. *Middle panel:* DO recordings from instrument (blue dot) and DO of adjacent sample snaps determined by Winkler Titrations (red triangles.) *Bottom panel:* Time series of the difference between the JFE AroUSB and reference measurements for each matched pair. Trial 1 and 2 are plotted on the same graph and separated by the axis break.

QUALITY ASSURANCE AND QUALITY CONTROL

All technical activities conducted by ACT comply with ACT's Quality Management System (QMS), which includes the policies, objectives, procedures, authority, and accountability needed to ensure quality in ACT's work processes, products, and services. The 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 ACT 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 the requirements of ISO/IEC 17025:2005(E), *General requirements for the competence of testing and calibration laboratories*; the American National Standards Institute (ANSI)/American Society for Quality (ASQ) E4-2004 *Quality Systems for Environmental Data and Technology Programs*; and U.S. Environmental Protection Agency, quality standards for environmental data collection, production, and use. An effective assessment program is an integral part of ACT's quality system. The ACT Quality Assurance (QA) Manager independently conducted two Technical Systems Audits (TSA) and data quality assessments of the reference sample data for the DO verification.

Quality Control Samples

Each site conducted weekly field duplicates which are presented below in Tables 9 – 11. The global average of the standard deviation among field duplicates for all field test sites was 0.03 ±0.07 (n=27), with 11 values exceeding our expected quality threshold of better than 0.013 mg/L DO. The average of the standard deviation among MTU field duplicates was .011 ±.014 (n=12), the average for CBL was .074 ±.006 (n=9) and the average for HIMB was .011 ±.012 (n=6). The higher variability at CBL likely reflected fine-scale heterogeneity in the water mass as was also noted by the greater variation in temperature across the mooring rack. In general, results attest to the representativeness of our sampling to water mass being analyzed by the test instruments and to consistent sample handling.

Table 9. Results of Field Duplicates (FD) for the Keweenaw Waterway, MI mooring test.

Date/Time	Rep	Temp	Spec Cond	DO	Std Dev	Mean	Std Dev	ABS Diff
1-14-15 10:30	FD1	.06	94	13.819	.005	13.819	0.00	0.000
	FD2			13.819	.002			
1-22-15 12:30	FD1	.31	99	12.981	.013	12.986	.007	.010
	FD2			12.991	.005			
1-29-15 16:00	FD1	.24	103	12.958	.041	12.947	.015	.021
	FD2			12.937	.013			
2-5-15 15:30	FD1	.21	106	12.671	.004	12.667	.006	.009
	FD2			12.662	.007			
2-19-15 15:30	FD1	.35	108	11.973	.008	11.974	0.000	.001
	FD2			11.974	.011			

2-26-15 15:30	FD1	.26	112	11.721	.011	11.755	.049	.069
	FD2			11.790	.076			
3-5-15 15:000	FD1	.34	112	11.615	.013	11.613	.002	.002
	FD2			11.612	.012			
3-12-15 10:15	FD1	.27	123	11.491	.028	11.477	.021	.029
	FD2			11.462	.007			
3-25-15 15:15	FD1	.72	118	11.474	.051	11.464	.015	.021
	FD2			11.453	.012			
4-3-15 10:00	FD1	.59	137	11.199	.012	11.201	.003	.004
	FD2			11.203	.001			
4-9-15 10:00	FD1	1.26	106	11.435	.008	11.430	.007	.010
	FD2			11.425	.009			
4-16-15 9:30	FD1	2.98	97	11.040	.005	11.042	.003	.004
	FD2			11.044	.006			

Table 10. Results of Field Duplicates (FD) for the Chesapeake Bay, MD mooring test.

Date/Time	Rep	Temp	Salinity	DO	Std Dev	Mean	Std Dev	ABS Diff
5-26-15 16:00	FD1	21.6	12.8	8.314	.011	8.269	.064	.090
	FD2			8.224	.004			
6-3-15 11:30	FD1	21.6	13.1	5.378	.003	5.297	.115	.163
	FD2			5.215	.015			
6-9-15 13:30	FD1	22.4	12.8	6.663	.0165	6.404	.366	.518
	FD2			6.145	.008			
6-17-15 9:30	FD1	26.6	12.1	8.827	.004	8.831	.006	.008
	FD2			8.835	.002			
6-24-15 11:30	FD1	27.5	11.0	7.051	.002	7.053	.003	.004
	FD2			7.055	.005			
7-7-15 14:00	FD1	26.9	9.9	6.157	.003	6.141	.023	.032
	FD2			6.125	.003			
7-15-15 9:30	FD1	27.3	8.7	5.781	.003	5.784	.005	.007

	FD2			5.788	.005			
7-22-15 9:30	FD1	28.0	9.2	7.151	.009	7.200	.069	.098
	FD2			7.25	.002			
7-28-15 13:30	FD1	29.13	8.3	8.336	.005	8.349	.019	.026
	FD2			8.362	.003			

Table 11. Results of Field Duplicates (FD) for the Kaneohe Bay, HI mooring test

Date/Time	Rep	Temp	Salinity	DO	Std Dev	Mean	Std Dev	ABS Diff
9-29-15 10:00	FD1	26.63	34.3	5.493	.006	5.501	.000	.016
	FD2			5.510	.007			
10-6-15 14:00	FD1	26.10	33.3	7.949	.007	7.880	.025	.139
	FD2			7.811	.014			
10-13-15 10:00	FD1	26.85	29.8	4.100	.016	4.114	.007	.027
	FD2			4.127	.006			
11-12-15 14:00	FD1	26.46	34.5	7.464	.029	7.472	.008	.016
	FD2			7.480	.017			
12-7-15 10:00	FD1	25.09	33.7	4.773	.011	4.750	.001	.046
	FD2			4.727	.012			
12-22-15 15:00	FD1	24.22	34.4	7.397	.025	7.399	.005	.005
	FD2			7.401	.032			

Technical System Audits

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 (*DO Sensor Verification Protocols*, PV14-01 21 October, 2014) and associated Standard Operating Procedures (SOPs). The ACT QA Manager conducted two TSAs over the course of the verification:

- The field tests at the CBL during July 29 – 30, 2015;
- The field tests at HIMB during September 28, 2015 – October 1, 2015.

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)*. The audits included a review of staff, test procedures (sample collection, sample analysis, data processing, etc.), facilities, and documentation.

The TSA's' findings for the two field tests were positive and indicated that these components of the DO sensor verification were being implemented in a manner consistent with the Test Protocols and SOPs. Minor deviations, such as schedule changes, were documented in laboratory records. None of the deviations in the Test Protocols had any effect on data quality for the verification and no corrective action was required. 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. Examination of maintenance and calibration logs provided evidence of recent and suitable calibration of sampling and analytical equipment.

Data Assessments

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. ACT's 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].

At the outset of the assessment phase, the data were verified and validated to evaluate whether the data have been generated according to the Test Protocols, satisfy acceptance criteria, and are appropriate and consistent with their intended use of evaluating the performance of the test sensors. Data verification evaluates the completeness, correctness, and consistency of the data sets against the requirements specified in the Test Protocols, measurement quality objectives (MQOs) in the ACT Quality Assurance Project Plan (QAPP), and any other analytical process requirements contained in SOPs. The ACT QA Manager reviewed the reference sample data sets from all field sites and 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 data on core parameters was traced in detail from raw data from field and laboratory logs and instrument readouts through data transcription or transference through data manipulation through data reduction to summary data, data calculations, and final reported data. Data validation established:

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

Data validation confirmed that ACT's sample measurement system performed in accordance with the performance goals specified in the ACT QAPP and the DO 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 validate that the data were collected in accordance with the verification's quality objectives.

A Data Quality Assessment (DQA), the third and final process of the overall data assessment, 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 DO sensors. The DQA determined that ACT's data quality objectives, described in Section 3.4 of the ACT QAPP, were achieved. This evidence 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 DO 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

Bittig, H.C.,Fiedler, B., Scholz, R., Krahmann, G., and Kortzinger, A. 2014. Time response of oxygen optodes on profiling platforms and its dependence on flow speed and temperature. Limnol. Oceanogr. Methods, 12: 617-636.

Carignan, R., Blais, A.M., and Vis, C. 1998. Measurement of primary production and community respiration in oligotrophic lakes using the Winkler method, Canada Journal of Fisheries and Aquatic Sciences 55:1078-1084.

Carritt, D.E., and Carpenter, J.H. 1966. Comparison and evaluation of currently employed modifications of the Winkler method for determining dissolved oxygen in seawater; a NASCO report. J. Mar. Res. 24: 286–318.

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. Erik Smith, Baruch Institute for Marine and Coastal Sciences, Ms. Janice Fulford, U.S. Geological Survey, Dr. Henry Bittig, Helmholtz Centre for Ocean Research Kiel, and Mr. Steve Ruberg, NOAA Great Lakes Environmental Research Laboratory 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.

August 20, 2016

Mario Tam

Date

Approved By: Dr. Mario Tamburri
ACT Executive Director

Thomas H. Johengen

August 20, 2016

Date

Approved By: Dr. Tom Johengen
ACT Chief Scientist

Earle N. Buckley

August 20, 2016

Date

Approved By: Dr. Earle Buckley
Quality Assurance Supervisor

In general, we consider the ARO-USB and AROW-USB performance that was presented in the report as satisfactory, since it shows how accurate RINKO sensors measurements are. However, there are some points that need to be addressed regarding the possible causes of the unwanted deviations presented by each instrument. These deviations may NOT be related to RINKO sensors capacity of accurately measuring DO concentrations, but rather related to the conditions in which samples were collected. We are commenting point by point below.

- **ARO-USB and AROW-USB Laboratory testing**: these tests have shown good agreement between DO oxygen measured by the instrument with a deviation under the accuracy stated by us (±2% of the full scale of the instrument, which in turn goes from 0 to 200% of dissolved oxygen saturation) in most of the cases. Nevertheless, there was a larger deviation pattern observed during all the tests, particularly when DO was saturated and/or supersaturated (e.g. see the last dataset from Figures D and U). These deviations are very likely to occur when there are pronounced slopes of DO concentration, which indicate that the concentration is changing fast and sometimes did not stabilized when the Winkler samples were collected (see the last dataset from Fig. B and Fig. O). Adding to that, there is also a possibility that during supersaturated condition, small bubbles that are not easily perceptible to the eye may be attached to the sensing foil and produced the difference between Winkler samples and the instruments. Note that small bubbles form very easily in supersaturated conditions and are very difficult to spot and remove without physical contact, which may explain the approximately constant pattern of deviation between Winkler method and RINKO instruments (e.g. see Fig. D, H, I, T and U). Thus, the deviations are very likely to be an experiment configuration issue sometimes involving stabilization or bubbles and their elimination.

- **ARO-USB Response Time Test:** the test was performed at 10s of sampling rate, which is not the highest sampling rate of the instrument. Also, the fitting adjustment that was used differs from our methodology to calculate the response time (we do not perform any fitting adjustment). Thus, the result does not correspond to the real response time of the ARO-USB.

- **AROW-USB Response Time Test:** this test also did not use the highest sampling rate of the instrument. The real response time need to be obtained using the instrument at its highest sampling rate, since AROW-USB is designed to measure long-term trends of DO concentration, and for that it makes use of a first-order filter. The filter is described as follows:

  $D_f = 0.912\ D_{n-1} + 0.088\ D_n$,

  where D_f is the filtered data, D_{n-1} is the immediate previous sampled data and D_n is the sampled data. The filter reduces high-frequency noise on the data and the immediate previous measure has a large effect on the filtered data. Therefore, using 10s, 20s or 30s of sampling rate will significantly increase the response time of the instrument, and although not mentioned in the report, the result does not correspond to the real response time of the instrument.

- **Moored deployment at Michigan Tech Great Lakes Research Center**: we have also some considerations on the data obtained from this field testing. The sensor seems to be working perfectly, registering all the DO variation along the whole period. However, there is a constant deviation between the AROW-USB results and Winkler titration samples. Although this constant deviation is inside the accuracy of the instrument, it is not easy to separate how much of this deviation is caused by the sensor drift or how much the issue reported with Winkler sampling (sampling corrected against a reference) influences the error magnitude.

- **Profiling deployment in the Great Lakes**: In Fig. 15, the difference between Winkler titration and ARO-USB is inside the accuracy range of the instrument, except two points where ΔDO reached values between -0.6 and -0.8 mg/L. The reason for this larger deviation between the methods may be related to errors associated with sample processing or the Winkler titration rather than the ARO-USB, since the variation on temperature observed around 12:45 – 12:50am is not enough to explain the variation presented by the Winkler titration on the DO concentration in the same period of time (15 minutes). In such a short period of time, we cannot explain the variation presented by Winkler titration and we are considering that the data provided by the ARO-USB are more likely to be close to the real DO concentration. In Fig, 16, although there is a larger deviation between the methods, all the values are inside the accuracy stated by us on the instrument specifications (2% of full scale).