LIMNOLOGY
and
OCEANOGRAPHY: METHODS

ASLO

Limnol. Oceanogr.: Methods 00, 2016, 00–00
© 2016 Association for the Sciences of Limnology and Oceanography
doi: 10.1002/lom3.10107

Accurate oxygen measurements on modified Argo floats using in situ air calibrations

Seth M. Bushinsky,[†]* Steven R. Emerson, Stephen C. Riser, Dana D. Swift
[1]School of Oceanography, University of Washington, Seattle, Washington

Abstract

Oxygen is an important tracer for biological processes in the ocean. Measuring changes in oxygen over annual cycles provides information about photosynthesis and respiration and their impact on the carbon cycle. Long-term, accurate oxygen measurements over wide areas are needed to determine changes in ocean oxygen content and oxygen deficient zones. Oxygen sensors have been increasingly mounted on Argo floats that profile between 2000 m and the surface. Most of these measurements are currently too inaccurate to calculate the air-sea gas flux, which is the dominant flux of oxygen in the surface ocean and typically driven by surface oxygen supersaturation states of only several percent. In this study, we present data from 17 Aanderaa oxygen optodes mounted on 11 Argo floats modified to make atmospheric measurements for calibration. Optodes measure oxygen equally well in air and water, allowing the use of atmospheric oxygen to perform on-going, in situ calibrations throughout the float lifetime. We find that it is necessary to make atmospheric measurements at night, that raising optodes higher into the air reduces variance in measurements, and that multiple measurements each time a float surfaces provide the best calibration data. Initial optode calibration on deployment has an average uncertainty of $\pm 0.1\%$ ($1\ \sigma$) and drift can be calculated to $\pm 0.1\%\ yr^{-1}$. Measurable drift was determined in 10–12 optodes out of the 14 that were deployed for ~ 2 yr. The maximum drift rate measured was $-0.5\%\ yr^{-1}$, which is large enough to strongly impact calculations of air-sea oxygen fluxes.

Oxygen is produced by photosynthesis in a stoichiometric ratio to the amount of organic carbon fixed and consumed when biological matter is respired. Any organic carbon that is produced in the upper ocean and respired at depth will be matched by an equivalent flux of oxygen out of the upper ocean. The movement of organic carbon from the surface ocean to the deep ocean is known as the biological carbon pump, a key component of the carbon cycle (Volk and Hoffert 1985).

Biological carbon export from the surface ocean is an important part of the carbon cycle but is difficult to measure due to the relatively small fraction of organic matter that leaves the surface ocean and the variability of export across space and time. Estimates of biological carbon export based on remote sensing data have sought to fill in gaps in our understanding of the magnitude and variability of this flux to the deep ocean (Westberry et al. 2008; Laws et al. 2011; Siegel et al. 2014). These estimates rely on in situ validation measurements that are often limited to discrete times or fixed locations. Seasonal and interannual variability of carbon production, consumption, and export requires annual measurements, which further complicates in situ validation (Emerson 2014).

Oxygen mass balance estimates of carbon export leverage the stoichiometric relationship between oxygen and organic carbon to use oxygen as a tracer for carbon export (Emerson et al. 1997, 2008). Emerson et al. (2008) concluded that a $\pm$ 0.5% error in mooring $[O_2]$ measurements produces a $\pm$ 50% error in net biological production calculated using a mass balance method. The method's sensitivity to oxygen accuracy is due to the fact that oxygen supersaturation (the relative difference between the amount of oxygen the water can hold at equilibrium with the atmosphere and the actual oxygen content) varies by only several percent over the course of a year in most areas of the ocean. Much of the excess oxygen produced during photosynthesis drives a flux into the atmosphere that is a function of the degree of oxygen supersaturation. Therefore, to accurately estimate the biological carbon flux, high accuracy oxygen measurements or precise measurements of the difference in oxygen partial pressure between the surface ocean and atmosphere are needed.

Widespread, accurate oxygen measurements are also needed to determine the extent of oxygen minimum zones

*Correspondence: sb17@princeton.edu

[†]Present address: Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey

(OMZs), interior regions of the ocean depleted in oxygen due to degradation of large amounts of organic matter and limited ventilation (Kamykowski and Zentara 1990; Helly and Levin 2004). Expansion of oxygen minimum zones has been documented over the past several decades and shown to cause changes in ecosystem composition (Diaz and Rosenberg 2008). OMZs are influenced by seasonal currents in addition to long term trends (Stramma et al. 2008), making it important to have annual measurements over large areas, which is difficult using intermittent repeat hydrography cruises. The accuracy of oxygen titrations also decreases with lower oxygen concentrations, limiting a researcher's ability to effectively determine temporal trends.

To date, most long-term studies of oxygen changes and fluxes have taken the form of either repeat hydrography comparisons (Stramma et al. 2008, 2010; Stendardo and Gruber 2012) or time series stations [for example Ocean Station Papa (OSP) in the Alaska Gyre (Whitney et al. 2007)]. These are invaluable to our understanding of oxygen dynamics and long-term changes, but are either limited in time resolution (repeat hydrography) or space (time series locations).

The Argo program has exponentially increased available temperature and salinity measurements through the use of profiling floats distributed throughout the ocean (Roemmich et al. 2009). While several hundred Argo floats have been deployed with oxygen sensors of various types (Gruber et al. 2009), these sensors have been insufficiently accurate to be used for determination of air-sea gas exchange. Takeshita et al. (2013) post-calibrated all currently functioning oxygen sensors using a combination of WOCE climatology at depth and a climatological seasonal cycle around saturation in the near surface. This procedure improves accuracy to $\pm 3\%$ of surface oxygen measurements, but is not accurate enough to resolve long-term oxygen trends or make accurate estimates of air-sea oxygen flux.

Aanderaa oxygen optodes, which are typically thought to be stable in the field, have a known decrease in sensitivity prior to deployment of $\sim$ 4–5% yr^{-1} (D'Asaro and McNeil 2013). Once deployed on Argo floats, optodes spend the vast majority of their time at 1000 m depth and 2–5°C, which likely accounts for their increased stability in the field. However, given the logistical challenges in recovering deployed floats, few floats are re-visited after a significant amount of time so it is unlikely that a drift of $< 1\%$ yr^{-1} would be readily observed. To accurately estimate air-sea oxygen flux, drift of even a few tenths of a percent per year can become important.

To improve the accuracy of oxygen optodes on Argo floats and make them useful for studies involving air-sea fluxes or low-oxygen measurements, in situ calibrations are needed. Körtzinger et al. (2005) proposed using in situ atmospheric measurements to calibrate Argo oxygen floats. This method has been tested in a few smaller scale studies. Bushinsky and Emerson (2013) used measurements from air pumped down to a submerged housing to calibrate an optode in a calibration system designed for moorings. Emerson and Bushinsky (2014) compared 1 yr of an air-calibrated Argo oxygen float to data from an independently calibrated surface mooring at Ocean Station Papa and found good agreement between calibrations (updated data from this float is presented in this study). Fiedler et al. (2013) conducted several 1 and 2 month deployments off the Cape Verde Islands testing in situ air calibrations with an accuracy of ± 2 μmol kg^{-1}, while Bittig and Körtzinger (2015) performed a 1.5 yr experiment off the West African coast and were able to use air measurements for calibration by implementing a correction method for wave splashing of the optode. Both the Fiedler et al. and Bittig and Körtzinger studies used floats equipped with external air bladders that could raise the oxygen optodes higher out of the water in an attempt to achieve a cleaner air sample. Johnson et al. (2015) utilized single air measurements from profiling floats to determine an average gain correction for optode measurements and found no evidence for long-term drift greater than $\pm 1\%$ after deployment.

In this study, we describe our efforts to develop a method for in situ calibration of oxygen optodes on profiling floats that improves on the accuracy of previous attempts. Optodes were calibrated in the laboratory against Winkler measurements to characterize optode behavior and sensitivity. Next, optodes deployed on modified Argo floats (hereafter referred to as Special-Oxygen-Sensor Argo floats, or SOS-Argo) at Ocean Station Papa and in the Kuroshio Extension (KE) were calibrated multiple times on deployment by ship-based Winkler calibration casts. Finally, the Argo floats described in this study sampled the atmosphere repeatedly at each surfacing to calibrate the sensor using atmospheric pO_2, allowing determination of deployment calibration and long-term drift.

Materials and procedures

Lab calibrations

Water calibration

Optodes were calibrated in the lab prior to deployment according to the general procedure described in Bushinsky and Emerson (2013). Calibrations were performed from 1°C to 21°C (a minimum of five temperatures across this range) and from 0% to 125% oxygen saturation. Optodes were submerged in a well mixed, temperature controlled water bath. Saturations above zero were reached by bubbling nitrogen or oxygen gas to strip oxygen out or add oxygen to the water. Diffusion across the top surface of the water bath was limited by floating pieces of foam. Calibration experiments used artificial seawater with salinity between 32 and 35 prepared according to Kester et al. (1967). Replicate Winkler measurements were made at each oxygen and temperature combination (approximately 30–45 points per optode).

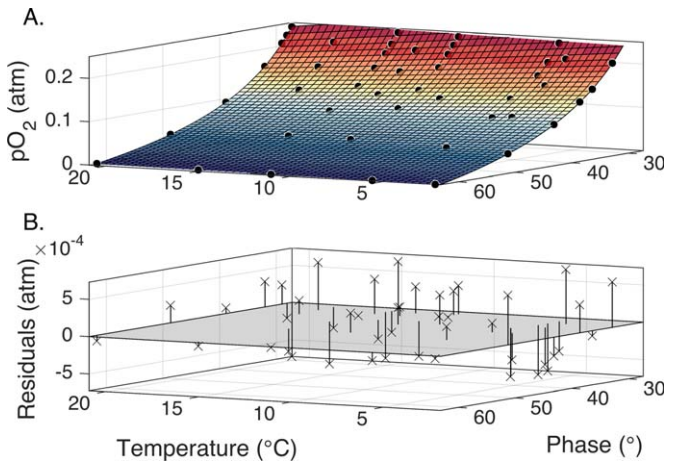

Fig. 1. Example calibration surface fit and residuals for optode 863, deployed on float 8387 in the Kuroshio Extension. Optode measured temperature and phase were matched to Winkler replicates (black circles, A). Calibration coefficients were determined according to Uchida et al. (2008), shown as the colored surface, with colors representing levels of equal pO_2. Residual pO_2 was determined for each Winkler replicate (B). Calibration points that do not meet measurement criteria were removed (see text): four data points were removed from this calibration because replicate Winkler measurements were greater than 0.2%, 0 points were removed because the optode phase measurement drifted by more than 0.05% during Winkler sampling, and 2 points were removed because residuals fell outside of 2 s.d. of the mean surface. A total of 6 out of 52 calibration points removed before the final coefficient determination, which was typical of most optode calibrations.

Zero oxygen measurements were carried out across all temperatures by saturating a beaker with sodium sulfite (Na_2SO_3, $\sim$ 5 g/500 mL seawater) to strip out oxygen. The beaker was placed inside a temperature controlled cooler and $[O_2]$ in the beaker was initially lowered by bubbling with N_2 gas.

Winkler titrations yield oxygen concentration ($[O_2]$), while Aanderaa optodes theoretically respond to the partial pressure of oxygen (pO_2) in water (Demas et al. 1999; Stokes and Somero 1999; Tengberg et al. 2003). The concentration of oxygen, $[O_2]$, is related to its pO_2, through the Henry's Law coefficient, $K_{H,C}$:

$$[O_2] = K_{H,O2} \times pO_2 \quad (1)$$

where $K_{H,O2}$ is calculated from the saturation concentration of oxygen at a given temperature and salinity (García and Gordon 1992). Winkler measurements were matched with optode readings taken in between the start and end of Winkler sampling. Winkler replicates were used for the calibrations only if the percent standard deviation of the replicate was less than $\pm$ 0.2% in the higher concentration range. The precision of Winkler titrations is a function of the smallest increment of sodium thiosulfate titrant added to the solution, which required that the $\pm$ 0.2% threshold be increased to $\pm$ 1% for oxygen values at 20% of saturation. Optode data were removed if the measurements changed by more than 0.05% during the Winkler sampling period.

Table 1. Summary of lab calibration fits and removed points.

	Lab calibration			
SN	**Initial**	**Removed**	**Final**	**RMSE (%)**
OSP				
895	60	35	25	0.49
896	59	30	29	0.71
KE				
842	52	9	43	0.20
852	52	7	45	0.20
854	31	2	29	0.30
861	31	3	28	0.21
863	52	6	46	0.21
865	52	8	44	0.18
866	52	14	38	0.27
890	31	2	29	0.33
891	32	3	29	0.35
892	31	2	29	0.27
894	31	1	30	0.32
897	52	8	44	0.20
898	32	4	28	0.45
938	52	7	45	0.27
941	52	10	42	0.37
Mean				0.31

Laboratory optode calibrations against Winkler measurements for optodes deployed on floats at Ocean Station Papa (OSP) and in the Kuroshio Extension (KE). Each optode is an Aanderaa 4330 listed by serial number (SN). The number of Winkler replicate and optode response pairings across all oxygen saturations and temperatures are indicated as those initially collected (Initial), the number of points removed (Removed), and the final number of points (Final) used for the calibration. Calibration coefficients and surfaces were fit to the data for each optode as in Fig. 1. Root mean square error (RMSE) is calculated from the residuals of the Winkler replicates minus the optode oxygen calculated using the surface fit. The reason for the larger RMSE for the OSP floats is discussed in the text.

If both optode and Winkler samples met the above criteria, raw optode response (phase, units of degrees) was plotted against temperature and Winkler-measured oxygen to calculate calibration coefficients (Fig. 1). Calibration surfaces were fit to Winkler-derived pO_2 using the Stern-Volmer equation from Uchida et al. (2008). Any data point with a residual (difference between the Winkler pO_2 and the calculated surface) greater than 2 standard deviations from the mean of the absolute value of the residuals was removed and the calibration coefficients were recalculated. See Table 1 for total calibration points used for each optode. Calibrations were performed over the course of 1–2 weeks to avoid optode drift during the duration of the run.

Air response

Optode air response for all KE optodes was measured in the lab using methods similar to those in Bushinsky and

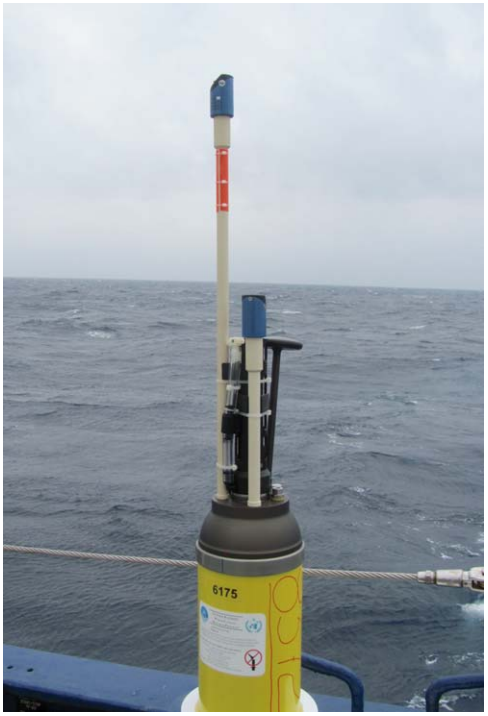

Fig. 2. An SOS-Argo float in the dual optode configuration bound for deployment in the Kuroshio Extension. Both optode stalks are supported with cable ties and brackets against the CTD mast. Five floats deployed in the KE had this configuration of two optodes, with the higher optode at 61 cm above the float end cap and one optode at 25 cm above the end cap. The other five floats only had one optode in the upper position. The SOS-Argo float deployed at OSP had two optodes, with the lower optode only 10 cm above the end cap.

Emerson (2013). Humidified air was pumped into a fully opaque, submerged chamber containing up to four optodes and held at several temperatures between 2°C and 21°C, while the pressure was measured using a Paroscientific Model 223A-101 pressure transducer [accuracy ± 0.01% (Wearn and Larson 1982)]. Atmospheric pO_2 was calculated according to:

$$pO_2^{atmos} = \left(P_{Atm} - pH_2O\left(\frac{RH}{100}\right)\right)X_{O2} \quad (2)$$

where pO_2^{atmos} is the partial pressure of oxygen expected in the atmosphere, P_{atm} is the total air pressure measured by the Paroscientific pressure sensor, pH_2O is the water vapor pressure (Zeebe and Wolf-Gladrow 2001), RH is the relative humidity in percent (assumed to be 100% for lab measurements), and X_{O2} is the mole fraction of oxygen in the atmosphere [$X_{O2} = 0.20946$ (Glueckauf 1951)]. Air response was measured within 2 weeks of water calibrations.

Float modifications

SOS-Argo floats are modified Apex profiling floats constructed at the University of Washington and designed to accommodate an air calibration period while staying within the operational requirements for the Argo program. All floats had an optode on a 61 cm stalk above the end cap (Fig. 2) and 6 of the floats had a second optode at either ~ 10 cm [close to standard Argo optode height (Körtzinger et al. 2005)] or 25 cm height. The variable heights were used to determine the necessary elevation above the sea surface to reduce the influence of waves on the sensor during air measurements. Stalks were milled from polyether ether ketone (PEEK) tubes, with an outside diameter of 20.3 mm and a through hole of 4 mm. Stalks were threaded into the float end caps with an O-ring seal. To maintain stability despite the increased mass higher on the float's vertical axis, float hulls were made from carbon fiber instead of aluminum to lower the center of gravity.

Float firmware was modified to allow air measurements after surfacing. On surfacing, each float entered a telemetry phase, during which the GPS location and data files were transmitted. After finishing telemetry, but prior to initiating a dive, the optode(s) were sampled in the atmosphere according to an adjustable interval and duration. Data from the surface period was stored in a separately transmitted file.

Field data

Float deployments

The initial deployment at Ocean Station Papa (float S/N 8397) was designed to test physical and firmware modifications. The float was deployed at 50.0 N, 144.8 W from the CCGS John P. Tully on 02 June 2012 and has been operating since on a 5–10 d cycle (Fig. 3). Float 8397 has completed 227 profiles as of August 2015. Four comparison hydrocasts with Winkler oxygen measurements were performed near the float. Argo floats sample as they rise to the surface, making timing of comparison casts and matching Winkler and optode measurements difficult. For the initial deployment period, the float was programed to profile as quickly as possible, with approximately 18–19 h between profiles. The first three calibration casts were between 1.7 and 5 km away from the location of the matching float profile and were sampled between approximately 0.5 h ahead of profile termination and 2.5 h after. For the fourth calibration cast, the ship waited for the float to surface and transmit its location before moving in to sample, 0.7 km away and 0.75 h after profile termination. The fourth cast was only to 20 m depth to further reduce the time difference between Winkler samples and optode measurements.

In February and March 2013 10 floats were deployed from the R/V *Melville* at roughly equal spacing (alternating single and dual optode configurations) between 30 N, 146 E, and 40 N, 150 E in the Kuroshio Extension region (Fig. 3). Deployments were performed as part of a University of Washington student cruise. KE floats have completed between 182 and 194 profiles as of August 2015.

Comparison casts were performed immediately prior to deployment on the northern cruise track, and tracked for a

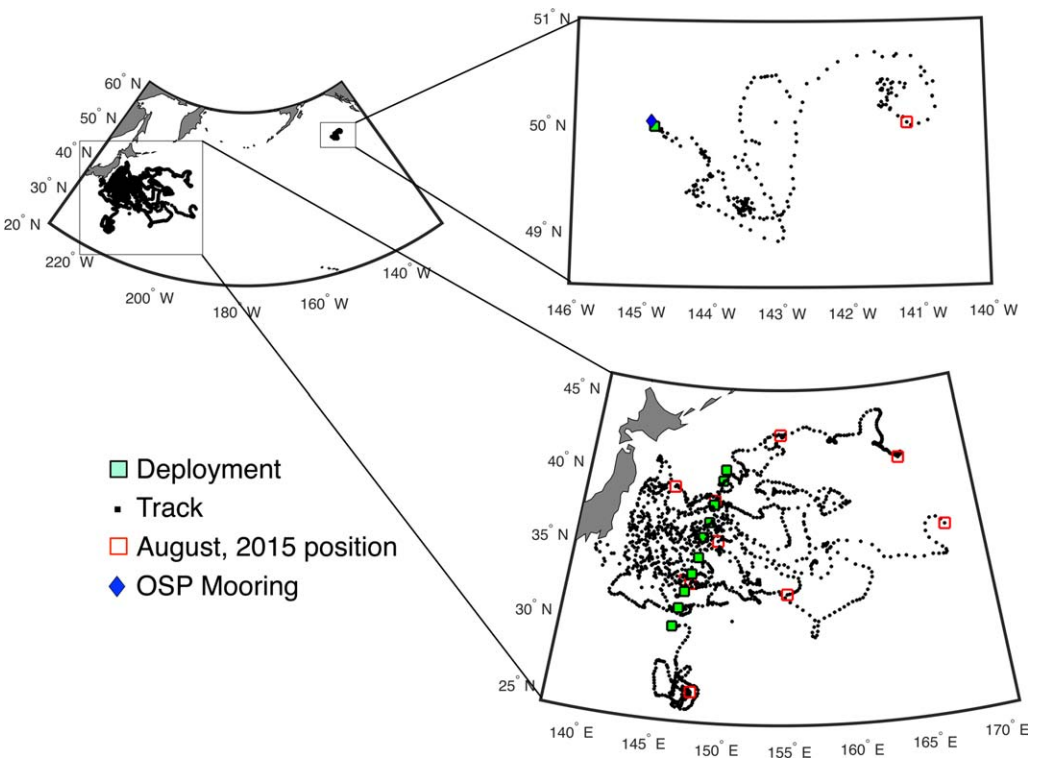

Fig. 3. SOS-Argo float deployment locations and tracks. Float deployment locations are indicated by green squares, drift tracks by black points, and location as of August, 2015 by red squares. The Kuroshio Extension float deployments were from February to March 2013. The OSP float (upper right) was deployed at 50.0 N, 144.8 W on 02 June 2012, near the OSP mooring.

second comparison cast along the return, southern cruise track. Surfacing locations were predicted based on drift track and floats were instructed to surface only once the R/V *Melville* was in the area of the float. Float surfacing could only be scheduled at least one profile in advance, so ship position and float surfacing were predicted 1–3 d in advance. Prior to expected surfacing, the Melville was positioned ~ 9 km away from predicted surfacing location. Each float was allowed to surface, a position was received, and the Melville approached the location but stood off to allow the float to complete its telemetry cycle and a surface air sampling period. After the float began its dive the ship moved to its location and performed a CTD comparison cast. This procedure shortened the time between float data collection and calibration casts to ~ 2–4 h. For all hydrocasts, replicate Winkler samples were taken at 24 depths down to 2000 m, with more samples near the surface.

In situ air calibration

Air measurements were made after every float surfacing. Initial float profiles for both OSP and KE were on a minimum profile cycle (~ 18 h) to maximize communication opportunities. Subsequent profiles for OSP were moved to a 5–10 d cycle and altered to surface only during nighttime. KE floats were instructed to only surface at night throughout their deployment. Occasionally, float surfacing time drifted into the daytime and had to be reset, resulting in periods without nighttime measurements.

On surfacing, each optode measured atmospheric oxygen (pO_2^{optode}) every 2 min for 1 h. For calibration of float measurements, true atmospheric oxygen (pO_2^{atmos}) was calculated from Eq. 2 using optode temperature and National Centers for Environmental Prediction (NCEP) reanalysis pressure and relative humidity data interpolated to the location and time of float surface air periods. The NCEP relative humidity product is an estimate for the *RH* at 10 m above sea surface, which was then extrapolated to optode height in the same manner as Bittig and Körtzinger (2015).

Assessment

Lab results

Water calibrations

Root mean square errors (RMSE) of calibration surface fits to lab Winkler data averaged ± 0.27% of atmospheric pO_2 for the KE optodes and ± 0.60% for the OSP optodes (Table 1). OSP optodes were calibrated first, over 6 weeks, which allowed sensor drift to impact the calibration surface fit. An average of 6 out of 42 points were removed for the KE floats according to the Winkler and optode criteria outlined in the water calibration procedure section. Optodes are theoretically a good instrument with which to measure low oxygen

water. The response of the platinum porphyrine foil is more sensitive at low partial pressures of oxygen than at high values. One way to check how well our calibration surfaces are explaining optode response at low oxygen is to observe the impact of adding or removing zero oxygen measurements (made in water with oxygen stripped by sodium sulfite) from the surfaces. The difference between optode output in near anoxic water using calibration coefficients that include and do not include zero oxygen data is ± 1 μmol/kg. This gives good confidence that our low oxygen Winklers are adequately constraining the calibration surface at low oxygen values. For calculation of oxygen concentrations above 20% saturation, addition of zero oxygen points to the calibration surface had a smaller impact than the error in the calibration surface.

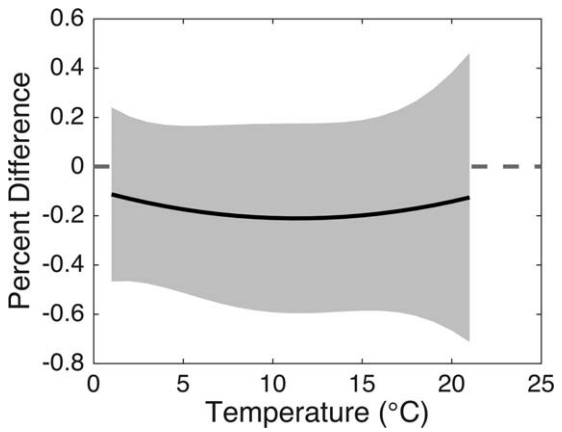

Fig. 4. Percent difference between optode pO_2 air measurements and calculated atmospheric pO_2 according to Eq. 2 in laboratory experiments with 15 Aanderaa optodes. The average of all optodes is shown with the black line and the ± 1 s.d. is in gray. The dashed gray line indicates zero percent difference.

Air response

To assess the ability of optodes to measure pO_2 in air, optode air measurements were made multiple times from 1°C to 21°C and compared with that expected based on the air pressure and temperature (Eq. 2). The mean offset of optode response between measured and expected pO_2^{atmos} was −0.2% ± 0.4% for the 15 KE optodes (Fig. 4). Thus, there was no statistically significant difference between optode response in air and water. This contrasts with our previous work, in which we showed an offset in air measurements that varied with temperature (Bushinsky and Emerson 2013, Fig. 8). A likely cause of the previously seen offset was drift over the ~ 8 months between optode calibration and air response experiments.

Field results

In situ calibration against air

Averaged OSP float pO_2 measurements from each air period are plotted in Fig. 5 with pO_2^{atmos} calculated from Eq. 2 using optode temperature and NCEP relative humidity and pressure. Both the upper (blue) and lower (red) optode measurements calculated using lab calibrations are lower than predicted pO_2^{atmos} (black) due to sensor drift prior to deployment. Lower optode atmospheric measurements at OSP have a larger variance during each sample period (an average of ± 0.35% s.d.) than upper measurements (± 0.23%, Table 2). Lower OSP optode measurements are also higher than upper, although this could be due to a lower rate of drift prior to deployment. The average standard deviation of

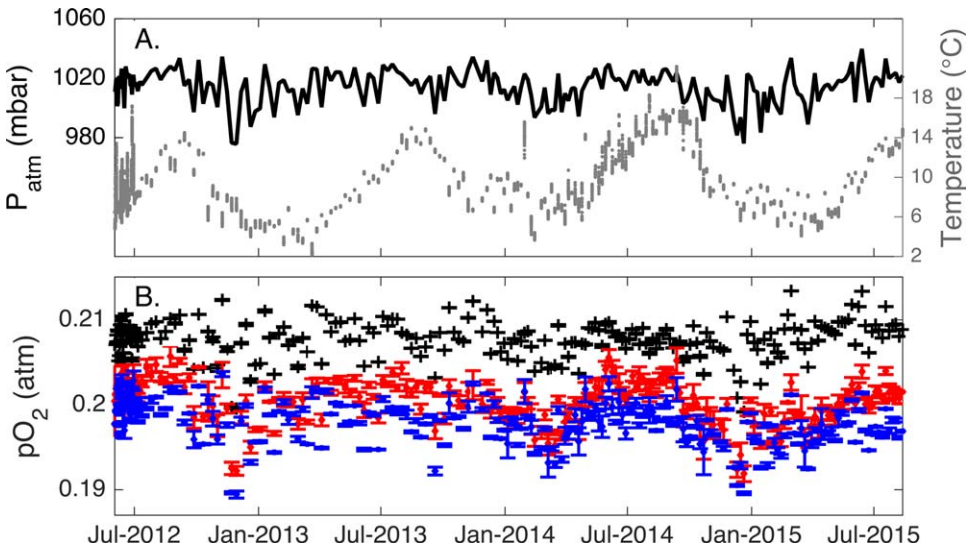

Fig. 5. Atmospheric conditions and upper and lower optode pO_2 for OSP float. (A) Atmospheric pressure (black) and optode temperature in air (gray). (B) Black symbols are atmospheric pO_2 calculated according to Eq. 2, using NCEP determined sea level pressure and relative humidity. During each air period, measurements were made every 2 min for ~ 1 h. Blue and red symbols are upper and lower optode air period averaged pO_2, respectively. The mean standard deviation for all air period averages (error bars, ± 1 s.d.) is ± 0.23% for the upper optode and ± 0.35% for the lower.

Table 2. Mean standard deviation for upper and lower optodes.

	Mean air period σ	
Float SN	**61 cm**	**25 cm**
8397*	0.23	0.35
8372	0.38	0.91
7661	0.44	0.92
8381	0.26	0.49
8382	0.32	0.70
8387	0.30	0.80

Mean standard deviation for upper and lower optodes on all dual optode floats in this study. Each air period, consisting of ~ 30 atmospheric measurements, is averaged with the mean standard deviation shown here. Mean s.d. for lower optodes was 2x that of upper optodes. *The lower optode on float 8397 was mounted on a 10 cm stalk rather than 25 cm for the others.

air period averages for all dual optode floats was ± 0.69% for lower optodes and ± 0.32% for upper optodes. This difference is likely due to increased splashing by waves, allowing surface oxygen supersaturation to influence lower optode air measurements.

To compare expected pO_2^{atmos} with measured pO_2^{optode} we calculate the percent offset between optode measurements and air pO_2 ($dpO_2 = \left[(pO_2^{optode} - pO_2^{atmos})/pO_2^{atmos}\right] \times 100$) with pO_2^{atmos} calculated as in Eq. 2. When plotted against the local time of day, mean dpO_2 and the variance at OSP increase to a maximum at local noon (Fig. 6) and a minimum variance during nighttime hours. This diurnal change correlates best with time of day, rather than measured temperature, estimated relative humidity, or other parameters. We therefore suspect this is due to light interference with the optical sensor and scheduled all floats to surface at night. Any periods during which float profiles ended during daytime were removed from analysis of air measurements and only nighttime (22:00–05:00 local) measurements were used for atmospheric calibrations.

To assess the need for multiple air measurements after each surfacing, nighttime air period means were calculated from increasing numbers of air measurements taken every 2 min during the 1 h air period. The mean of each air period was subtracted from the means made with fewer air measurements and the standard deviation of these mean offsets for a representative optode (SN 854 on float 8381) is plotted in Fig. 7. For all optodes, the standard deviation of the mean difference between the first measurement in an air period and that air period's average is ± 0.61%. Increasing the number of air measurements to 10 reduced the difference from the air period mean to ± 0.17% and after 20 measurements the difference in means was ± 0.08%.

Variance in air period averages was highest during periods of changing atmospheric conditions. We thus filtered atmospheric data by removing air measurements in which: the standard deviation of dpO_2 changed faster than 0.2% over 10 min, P_{atm} changed by more than 0.1% over the air sampling period, or air temperature changed by more than 1°C over the air sampling period. The initial 10 min of measurements were removed to allow water to drip off of the sensor and reduce the effect of possible evaporative cooling.

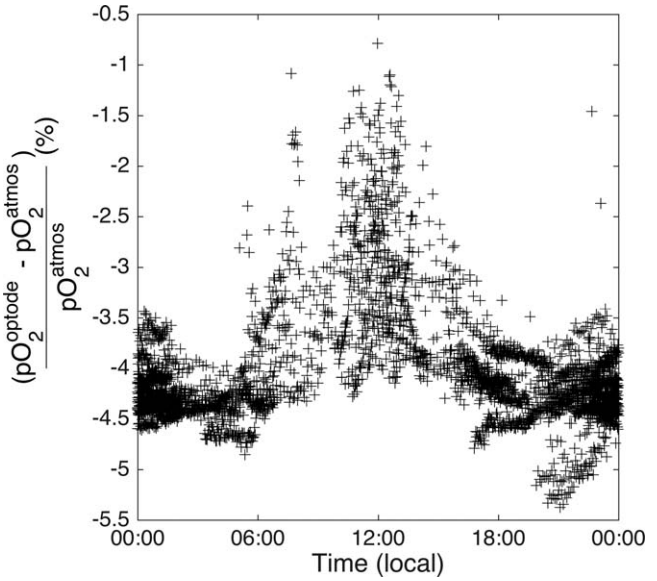

Fig. 6. Uncorrected optode air measurements vs. local time of day. Air measurements from the OSP upper optode during the first 6 months of deployment are plotted as the percent difference between optode and atmospheric pO_2 determined using NCEP atmospheric data and Eq. 2. Local noon corresponds to the highest mean and variance in optode air measurements. After determination of this relationship surface measurements were shifted to nighttime only.

Daytime (gray), and nighttime unfiltered (blue) and filtered (red) data for the OSP float are plotted in Fig. 8. Mean standard deviations for unfiltered air periods were 2–3.5x larger than filtered air periods. As mentioned above, air periods that occurred during daylight had large standard deviations (error bars in Fig. 8) and were biased high. During the period from December 2013 to October 2014 the OSP float began surfacing during the early morning instead of the middle of the night and this anomaly can be seen in the data.

In general, lower optode air measurements display a seasonal cycle that indicates higher values in the summer when surface waters are supersaturated and is likely the result of splashing by waves (see Bittig and Körtzinger 2015). This is especially evident in the unfiltered data, but is still partially visible in the filtered dataset. Overall, filtered air periods have the smallest standard deviations for both the upper and lower optodes. The remaining filtered air period averages are used to determine the initial calibration of pO_2 on deployment and any subsequent drift.

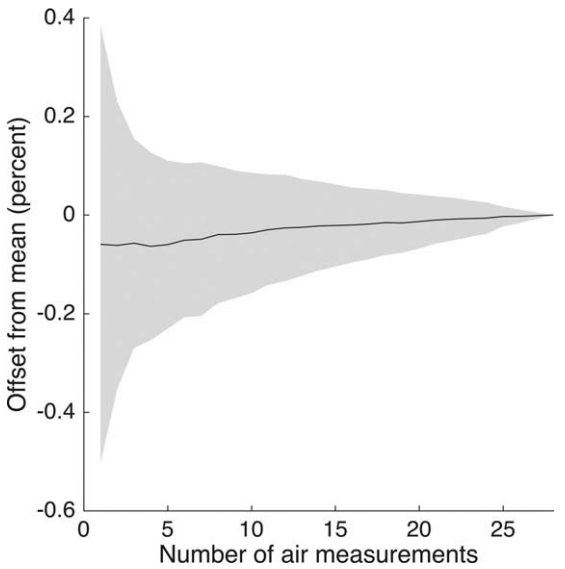

Fig. 7. Optode air period mean offsets with increasing numbers of atmospheric measurements. After each profile, the optode, mounted on a 61 cm stalk, made measurements every 2 min for ~ 1 h. For each surface air period, the mean percent difference from atmospheric pO_2 is calculated with all points ($n = 28$) and subtracted from means calculated with increasing numbers of data points. The black line is the mean and gray area the standard deviation of 171 nighttime air periods for a representative optode on a float deployed in the Kuroshio Extension (SN 8381). Averaging more air measurements reduces the mean standard deviation of the air period offsets. For 14 optodes, the mean standard deviation of the first measurement is ± 0.61%. This decreases to ± 0.17% after 10 measurements and ± 0.08% after 20 measurements. By definition, at 28 measurements, the mean standard deviation is 0.

The black lines in Fig. 8 are linear regressions against filtered air period dpO_2 for the OSP optodes. The *y*-intercept of the linear regression at the time of float deployment represents the sensor drift between the time of optode calibration and deployment. The slope of the linear regression is the rate of drift during the deployment. Confidence intervals plotted around the regression are ± 1 (gray area) and 2 (blue area) standard deviations. Air period averages were used to calculate linear regressions instead of all individual air measurements. The OSP optodes measured pO_2 at the *y*-intercept (at the time of deployment) that was lower than expected from atmospheric calibration by −4.2% ± 0.1% (Upper) and −3.0% ± 0.1% (Lower, ± 1 σ) (Table 3). Note that the individual optode differences from expected are not the same and indicate optode-specific drift. The Winkler comparison cast for this float shows a difference between optode and Winkler of −4.5% ± 0.2% (upper) and −3.8% ± 0.2% (lower), which differs from the air calibration by 0.3% and 0.8%, respectively (Table 4).

Some of the in situ temperatures from the KE floats exceeded the maximum calibration temperature of 21°C. Because the calibration surface above this temperature was uncertain, measurements in this temperature range were removed from the regression to atmospheric data. Air measurements above 21°C for floats 5328, 7665, 8372, 8375, 8382, 8387, and 8394 are thus not plotted in Fig. 9. Using the linear regression to remaining atmospheric data for each optode, we can determine a high temperature (> 21°C) correction for those optodes by calculating the difference between actual air measurements and the linear fit to the filtered atmospheric data. This allows an oxygen calculation for surface waters at all measured temperatures.

Linear regressions for all KE floats were made against filtered air periods and are plotted with the regression lines and confidence intervals in Fig. 9. Linear regressions are used because we are interested in correcting long-term drift only and do not think seasonal fluctuations are part of that drift. Regressions are made only on whole years of data to avoid seasonal biasing from incomplete years. The differences between Winkler comparison measurements and initial in situ air calibrations at float deployment are plotted with each optode's atmospheric measurements in Fig. 9 (yellow squares). The average difference for all optodes between Winkler measured optode offset and dpO_2 from atmospheric calculations is −0.5% ± 0.7% (Winkler minus dpO_2, Table 4). The time and space differences between float profile measurements and Winkler comparison samples is a potential problem for calibration of floats using surface measurements on deployment. Comparison of temperature and salinity plots for float profiles and ship comparison casts indicated that five of the KE comparison casts sampled water with different properties than the matched float profiles (data not shown).

Any significant optode drift during deployment will show as a data trend away from the initial intercept in Fig. 9 (dashed gray lines). To determine whether a given trend was measurable, confidence intervals around the linear regressions in Figs. 8 and 9 were calculated as:

$$\mathrm{CI}_{\%} = t_{\left(1-\frac{q}{100}\right)[\nu]} \sqrt{\frac{s^2_{Y \cdot X}}{\sum x^2}} \tag{3}$$

where q is the percent confidence interval, t is the Students *t*-distribution critical value derived from a table of *t*-values (Rohlf and Sokal 2003) for the given confidence interval and degrees of freedom, ν (number of measurements, $n - 2$), $\sum x^2$ is the sum of the square differences of x (time in days of each air period) and the mean time, and $s^2_{Y \cdot X}$, the unexplained variance (Sokal and Rohlf 2003), which is calculated from:

$$s^2_{Y \cdot X} = \frac{\sum d^2_{Y \cdot X}}{\nu} \tag{4}$$

where $\sum d^2_{Y \cdot X}$ is the unexplained sum of squares, or the sum of the square difference between each air period average and the prediction for that time from the linear regression. For a

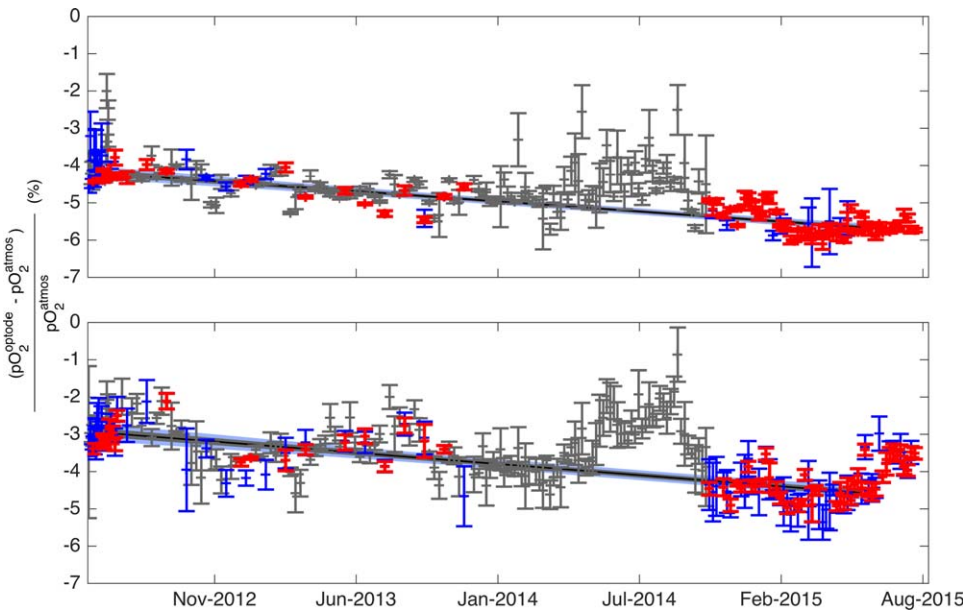

Fig. 8. Comparison between optode air measurement and calculated pO_2 over 3 yr of the OSP deployment for the upper (top) and lower (bottom) optodes. Percent difference between optode pO_2 calculated using laboratory calibrations and atmospheric pO_2 from Eq. 2 is plotted for all air periods (gray), nighttime (blue), and nighttime that passed the filter described in Assessment: Field results (red). Error bars represent ± 1 s.d. around the mean for each air period. Solid black lines are the linear regression to the filtered (red) measurements. The pronounced seasonal cycle in the lower optode is likely due to splashing from waves or spray. During the period from December 2013 to October 2014 the floats were surfacing during daytime, which is the reason for the departure from the trend and the large standard deviation during this time. Linear regressions to the filtered data are black lines, bounded by ± 1 σ (gray areas) and ± 2 σ (blue areas) confidence intervals calculated according to Eqs. 3 and 4.

time series with consistent variance, longer time periods over which measurements are taken (which increases $\sum x^2$) and more measurements (n, which increases v) will result in smaller confidence intervals.

Of the 13 optodes with long enough time series to check drift, 10 displayed measurable drift at 95% confidence intervals (2 σ, blue areas) and 12 were measurable at 68% (1 σ, gray areas). Both of the optodes on the OSP float, which has the longest record at 3 yr, are drifting at a rate of −0.5% yr^{-1}. The average in situ drift rate for all optodes is −0.1% yr^{-1}, with a range in most drift rates of −0.5% to 0.2% yr^{-1}. The one sensor outside of this range, the lower optode on float 8387 with a drift rate of 0.5% yr^{-1}, has a relatively large spread in the air data.

Discussion

In situ atmospheric calibration

The OSP float was deployed 8 km from the OSP mooring site (50°N × 145°W, Fig. 3). Mounted on the mooring bridle was an optode calibrated according to the same lab procedure described in Materials and procedures and subsequently corrected to Winkler measurements taken from the CCGS Tully on deployment in June 2012 and again in August 2012 and June 2013. In June 2013, a new mooring was deployed with an oxygen optode that was also calibrated to Winkler measurements on deployment and in August 2013 before an electrical malfunction in December 2013. A third mooring deployment in June 2014 provides another 5 months of Winkler calibrated data.

The continued proximity of the OSP float to the surface mooring provides a long-term source of comparison. The 2.5 yr mooring time series of supersaturation (black line) is plotted in Fig. 10 with the atmospherically calibrated float data from each profile (blue line and x's). The mean difference between the mooring optode ΔO_2 calibrated to in situ Winkler measurements and float optode ΔO_2 calibrated to air measurements is ± 0.2%. This comparison is updated from Emerson and Bushinsky (2014) when only 1 yr of float data was available. While we would not expect exact agreement between float and mooring given their spatial separation, there is no long-term bias between these two independently calibrated oxygen measurements.

In situ air calibrations depend on one's ability to accurately measure atmospheric oxygen. Splashing water, mismatches between measured pO_2 and pressure data, rapidly changing temperature, or any other factor that will impact dpO_2 can degrade the quality of the calibration. We approached these problems by raising the optode higher out of the water, making many measurements during each air period, and then filtering the data to remove anomalous data. Raising the optodes to 61 cm reduced the average s.d. for air periods by half (Table 2). Lower optodes exhibited stronger seasonal cycles in dpO_2 than upper (Fig. 9) and all

Table 3. Linear regressions to unfiltered and filtered air period averages.

				Unfiltered air period averages			Filtered air period averages			Differences (unfiltered – filtered)	
UW Float #	WMO #	Optode SN	Position (Upper/ Lower)	*n*	Intercept (% ± σ)	Slope (% yr^{-1} ± σ)	*n*	Intercept (% ± σ)	Slope (% yr^{-1} ± σ)	Intercept difference	Slope difference
8937	5903743	896	U	81	−4.1 ± 0.1	−0.5 ± 0.0	63	−4.2 ± 0.1	−0.5 ± 0.0	0.1	0.0
		895	L	81	−3.0 ± 0.1	−0.6 ± 0.0	48	−3.0 ± 0.1	−0.5 ± 0.1	0.0	0.0
5328	5904029	852	U	148	−1.6 ± 0.1	−0.3 ± 0.1	43	−1.6 ± 0.1	−0.4 ± 0.1	0.0	0.1
8372	5904026	941	U	147	−3.0 ± 0.1	−0.2 ± 0.1	65	−3.2 ± 0.1	−0.2 ± 0.1	0.3	0.0
		938	L	147	−3.9 ± 0.1	−0.3 ± 0.1	16	−3.7 ± 0.1	−0.2 ± 0.1	−0.1	−0.1
8375	5904030	898	U	137	−3.4 ± 0.2	−0.4 ± 0.2	75	−3.2 ± 0.2	−0.3 ± 0.2	−0.1	−0.1
8387	5904027	897	U	139	−2.4 ± 0.1	−0.1 ± 0.1	84	−2.7 ± 0.1	0.1 ± 0.1	0.3	−0.2
		863	L	139	−4.2 ± 0.1	0.0 ± 0.1	26	−4.3 ± 0.2	0.5 ± 0.2	0.1	−0.4
8394	5904031	861	U	139	−3.3 ± 0.2	0.3 ± 0.1	75	−2.6 ± 0.1	0.0 ± 0.1	−0.7	0.3
7661	5904024	894	U	27	−1.9 ± 0.2						
		892	L	27	−4.3 ± 0.2						
7657	5904032	842	U	21	−3.7 ± 0.2						
8382	5904025	890	U	142	−5.6 ± 0.2	0.1 ± 0.2	63	−5.0 ± 0.2	0.2 ± 0.2	−0.7	−0.1
		891	L	142	−3.7 ± 0.1	0.0 ± 0.1	13	−3.1 ± 0.2	−0.2 ± 0.2	−0.6	0.2
7665	5904033	866	U	145	−2.6 ± 0.1	−0.1 ± 0.1	85	−2.7 ± 0.1	0.0 ± 0.1	0.1	−0.1
8381	5904028	854	U	138	−2.8 ± 0.1	0.2 ± 0.1	116	−2.8 ± 0.1	0.2 ± 0.1	0.0	0.0
		865	L	138	−3.4 ± 0.1	0.1 ± 0.1	46	−2.8 ± 0.2	−0.3 ±0.1	−0.6	0.4
Mean					±0.1	−0.13 ± 0.1		±0.1	−0.12 ± 0.1		

Comparison of intercept and slopes calculated for linear regressions to unfiltered and filtered air period averages. For each optode, UW float number, World Meteorological Organization (WMO) number, serial number (SN) of optode, position on the float (upper are on a 61 cm stalk, lower on a 25 cm stalk), *n* (the number of air period averages used in the regression), and the intercept and slope ± 1 σ are listed. Unfiltered data include all air periods for a float, while filtered data removed air periods during which atmospheric conditions were rapidly changing as described in the text (Assessment: Field results). The differences between unfiltered and filtered intercepts and slopes are also calculated. Floats 7661 and 7657 did not operate for long enough to filter data or calculate drift rates. No means are calculated for the intercepts as these represent optode specific drift from lab calibrations. Optodes with more time in between calibration and deployment will drift more. Pre-deployment drift rates are plotted in Fig. 11.

of the KE lower optodes were taller than typical optode deployment height. Since there is no reason to expect the mean around the seasonal cycle in dpO_2 is the true value, a simple averaging will not suffice for determining the correct offset. It is probable that this cycle is biased by differences in wave energy, wind speed, and temperature changes throughout the year.

Bittig and Körtzinger (2015) adjust for wave splashing by assuming the optode is measuring a mix of water and air and adjust the data such that there is no difference between water and air measurements at 100% saturation. Implementing this method with nighttime data from the first year of data collected by the lower optode on the OSP float yields an initial offset of −2.8% ± 0.4%, comparable to the offset found in this study of −3.0% ± 0.1% (Table 3). The difference in confidence intervals may be due to the difference in time periods used for the offset calculation. However, the seasonal cycle in dpO_2 cannot be explained for all optodes by splashing. For example, the upper and lower optodes of float 8381 follow anti-correlated seasonal cycles (Fig. 9). The lower optode dpO_2 maxima occur during the summer, when splashing by supersaturated water could explain elevated optode air measurements, but the upper optode measured minimum values at these times.

Optode drift evaluation

The percent change in optode response calculated relative to air calibration on deployment for all optodes in this study is plotted in Fig. 11A. Factory calibrations (black *x*'s), lab calibrations (blue circles), and air calibrations illustrate a roughly linear drift prior to deployment. The linear relationship shows that the first order problem with the factory calibration is drift, rather than inaccuracy in the calibration. Plotting all optodes with multiple calibrations and normalizing to the [O_2] calculated using the air calibration on deployment yields an average drift of −7.2% ± 1.5% yr^{-1} at near saturation values. This drift rate is comparable to the initial, post factory calibration drift rate illustrated in D'Asaro and McNeil (2013) of ∼ 6.3% over the first year.

Table 4. Comparison of air calibration to Winkler measurements.

Float SN	Optode SN	Position (Upper/Lower)	Air calibration intercept (% ± σ)	Winkler Cal. (% ± σ)	Difference (Winkler – air calibration)
8937	896	U	−4.2 ± 0.1	−4.5 ± 0.2	−0.3
	895	L	−3.0 ± 0.1	−3.8 ± 0.2	−0.8
5328	852	U	−1.6 ± 0.1	−2.3 ± 0.2	−0.7
8372	941	U	−3.2 ± 0.1	−3.5 ± 0.1	−0.2
	938	L	−3.7 ± 0.1	−4.2 ± 0.1	−0.5
8375	898	U	−3.2 ± 0.2	−4.6 ± 0.2	−1.4
8387	897	U	−2.7 ± 0.1	−2.8 ± 0.4	−0.2
	863	L	−4.3 ± 0.2	−4.4 ± 0.4	−0.1
8394	861	U	−2.6 ± 0.1	−3.1 ± 1.0	−0.6
8382	890	U	−5.0 ± 0.2	−4.2 ± 0.3	0.8
	891	L	−3.1 ± 0.2	−2.4 ± 0.3	0.7
7665	866	U	−2.7 ± 0.1	−3.5 ± 0.0	−0.8
				Mean difference:	−0.5 ± 0.7

Initial air calibration of optodes compared with Winkler measurements from nearby hydrocasts. Air calibrations are based on the intercept at time of deployment of the linear regression to filtered atmospheric data in Fig. 9. Winkler calibration is the mean and standard deviation of surface measurements from deployment and re-visitation hydrocasts. Winklers for float 8381 were not matched well in space or time and therefore not included in this comparison.

Post deployment drift is more than an order of magnitude lower than pre-deployment drift. The 10 optodes that showed measurable drift at a 95% confidence interval averaged drift of approximately −0.2% yr^{-1}. Of these, seven optodes displayed an average decrease in sensitivity of −0.4% yr^{-1} and three optodes increased in sensitivity by an average of 0.3% yr^{-1}. All float optodes spend the majority of their time at ∼ 1000 db in the dark, with temperatures that never vary more than a few degrees. Surface conditions for floats with optodes displaying similar rates of drift are very different. Float 8397 is in the northeast Pacific subarctic gyre where a strong seasonal cycle in light and temperature causes seasonal cycles in surface supersaturation of ∼ 6–10%. Float 5328 is in the western side of the subtropical gyre where surface waters are ∼ 15°C warmer and the seasonal amplitude of oxygen supersaturation is only ∼ 4%. All three of the optodes on those two floats drifted by an average of −0.5% yr^{-1}. Most of the other surface conditions experienced by the floats fall in between these two extremes, suggesting optode-specific differences in drift rates rather than differences due to environmental conditions. Interestingly, all optodes presented in this article were equipped with sensing foils cut from the same sheet. This should maximize the similarity of the sensing foils.

This is the first time drift has been measured in deployed Aanderaa optodes. Previous studies that found no drift in optodes deployed on profiling floats had a few significant differences from the current study. Takeshita et al. (2013) calibrated already deployed optodes using climatological data that could not resolve changes of less than several percent. Even at the highest rates of drift measured in this study, it is unlikely that any of these optodes will display a post-deployment drift of greater than 3% over their expected float lifetimes of 4–5 yr. Johnson et al. (2015) applied an air calibration technique similar to what is done here but using only one air measurement each time a float surfaces. This introduces additional variance to the air measurements (Fig. 7) that would make these small, but important rates of drift difficult to detect.

For comparison with the Johnson et al. (2015) method of using one atmospheric measurement after each surfacing, we subsampled our unfiltered data and recalculated drift intercepts and slopes for each optode. For four out of five lower optodes in this study, using only the first measurement of each air period yielded an absolute difference in initial offset of 0.2–0.8% and for three out of five lower optodes drift slope differences of 0.5–0.7% were observed. For all upper optodes, no significant difference was observed for initial offset or drift slope when calculated from initial air period measurements or averages of the entire air period. This illustrates the value in raising the optodes above the air-water interface when possible and the importance of multiple air measurements when not possible.

Tengberg et al. (2006) and Bittig and Körtzinger (2015) each tested one optode on a profiling float for over a year and found no drift at confidence levels of ± 0.7 and 1%, respectively. In our study it appears that drift rate, and possibly direction, are optode or float specific and that some indicate measurable drift while others do not.

Over the 4–5 yr expected lifetime of these floats, drift rates of −0.5% yr^{-1} are important for gas flux calculations.

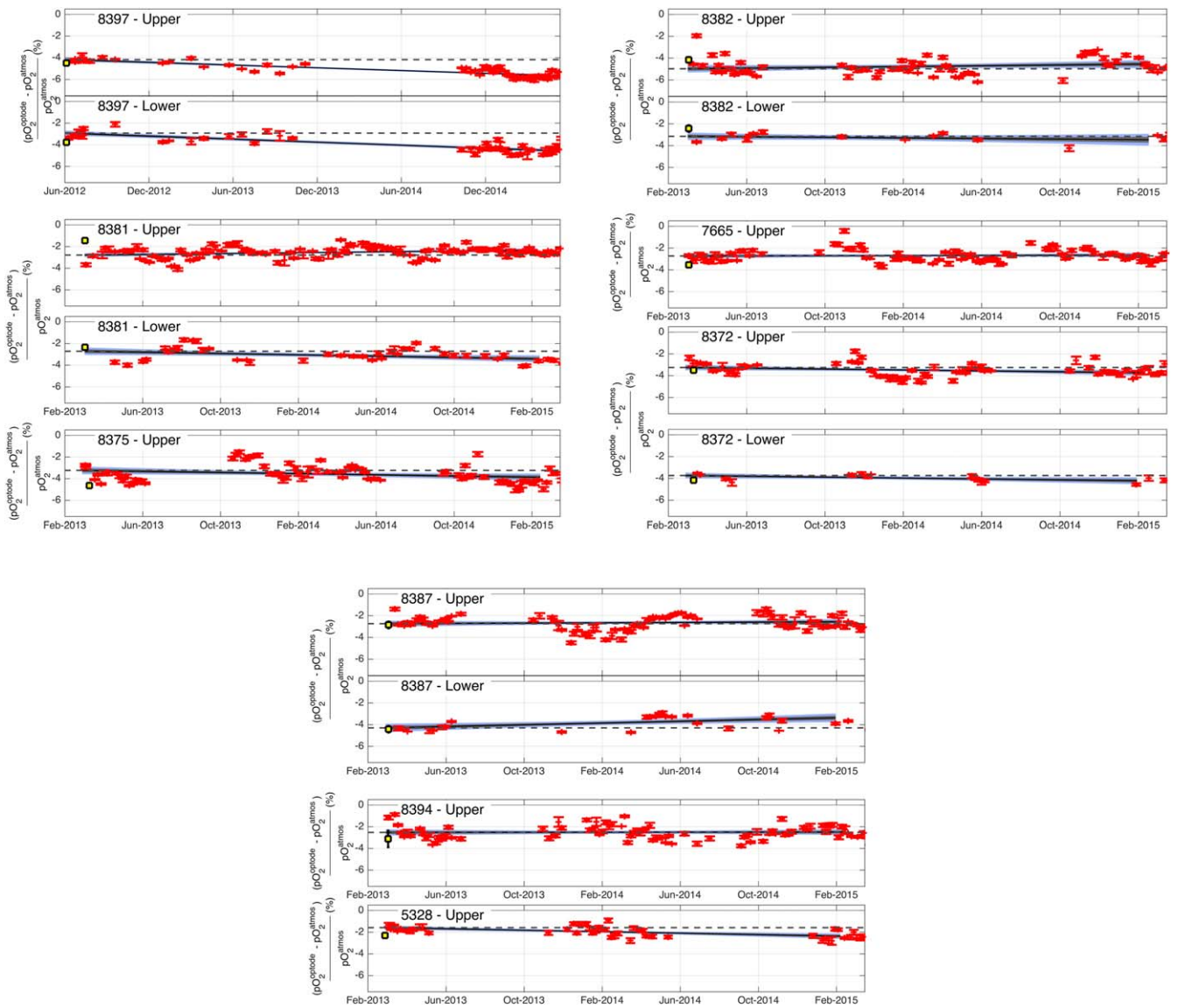

Fig. 9. Filtered air measurements from 14 optodes over ~ 2 annual cycles. Red symbols are the percent difference between optode and atmospheric pO_2 calculated from Eq. 2. Error bars represent ± 1 s.d. of the mean for each air period. Data from upper, and lower if present, optodes are plotted for all floats and have been filtered according to the criteria outlined in Assessment: Field results. Solid black lines are linear regressions to dpO_2, with ± 1 σ and ± 2 σ confidence intervals in gray and blue areas, respectively. Dashed black lines are the intercepts of the linear regressions calculated at time of deployment, which represents the initial in situ calibration of the optode on deployment and the drift since laboratory calibration (Table 3). Data trends away from the initial intercept lines indicates drift. For 10 of the 14 optodes, drift was measurable at ± 2 σ and for 12 of 14 optodes drift was measurable at ± 1 σ. Yellow squares are the difference between Winkler and optode measurements matched for calibration casts (Table 4). Note that the time scale for float 8397, deployed at OSP, starts 1 yr earlier than the rest of the floats. A modified version of the 8397 – Upper plot is found in Bushinsky and Emerson (2015).

Assuming typical summertime conditions at OSP and an oxygen supersaturation of 4%, a −0.5% yr^{-1} reduction in measured saturation would reduce the oxygen flux out of the ocean by ~ 12% after 1 yr and ~ 50% after 4 yr (Liang et al. 2013; Wanninkhof 2014). Because of the asymmetry of O_2 supersaturation, this effect is also not uniform throughout the year. Wintertime errors in the oxygen flux would be even greater; assuming a typical surface undersaturation of −1% ΔO_2, a −0.5% error results in a 44% increase in the oxygen flux to the ocean while a −2% error results in an flux overestimate of 175%.

Low oxygen measurements

Determining oxygen accurately at low concentrations is important for measuring and understanding oxygen minimum zones. The cutoff for severe hypoxia is ~ 22 μmol kg^{-1},

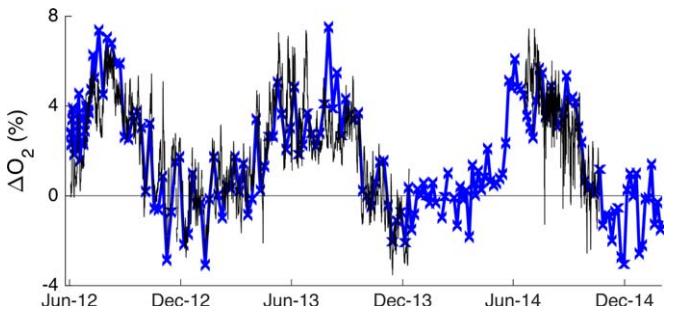

Fig. 10. Surface ocean oxygen supersaturation in the subarctic Pacific. O_2 concentrations were determined from optodes on a mooring at OSP (black) and an Argo float (blue). Mooring data are corrected to O_2 concentration determined by Winkler titrations sampled on deployment and subsequently in August 2012, February 2012, June 2013, and August 2013. Float data were calibrated to air measurements for both initial offset and subsequent drift.

or ~ 10% saturation (Hofmann et al. 2011) which underscores the need to understand optode response and drift in this region of the calibration curve.

To monitor and correct for drift at low oxygen values, it is necessary to know how the calibration surface changes with drift. If drift is a change in sensitivity (a gain correction) that is a uniform percent of measured pO_2, then calibration points at near saturation values can be used to adjust low oxygen values as well. The best evidence we have for how the optodes drift at low $[O_2]$ comes from the ship comparison casts in the Kuroshio Extension, where Winkler samples in the oxygen minimum can be compared with the optode drift corrected to the atmospheric calibrations.

The oxygen minimum in this region is between 1000 m and 1400 m deep and has a concentration of ~ [40] μmol kg^{-1}. Drift calculated from the response change between factory calibrations, lab calibrations, and in situ Winkler measurements yields a change in optode response of $-23\% \pm 3.2\%$ yr^{-1}, in the oxygen minimum, far higher than drift measured in the surface ocean at saturation (Fig. 11). The drift rate is expressed as a percentage of the measured concentration in the same manner as a gain correction. Optode data were adjusted for a pressure effect of 3.2%/1000 db (Uchida et al. 2008) prior to comparison with in situ Winkler samples. While Winkler titrations are less accurate at low $[O_2]$, this rate of drift is dependent on multiple Winkler replicates averaged within the oxygen minimum and larger than could be explained by titration error.

Expressed in a concentration offset, the drift rate for the KE floats plotted in Fig. 11A,B is -19.3 ± 4.2 μmol kg^{-1} yr^{-1} at saturation and -7 ± 0.7 μmol kg^{-1} yr^{-1} in the oxygen minimum, implying that neither an offset nor a gain correction can be applied linearly to the entire calibration surface.

For example, optode 854, the upper optode deployed on float 8381, drifted by -6.7% yr^{-1} for near saturation waters based on the atmospheric calibration at the time of deployment. Using the factory calibration, this optode would have measured ~ 20 μmol kg^{-1} below true in surface waters on deployment. A gain correction based on the atmospheric calibration correctly adjusts the surface measurements, but under corrects the concentration in the oxygen minimum. A -6.7% yr^{-1} correction applied to the factory calibration value of 27.6 μmol kg^{-1} brings the concentration to 29.9 μmol kg^{-1}, 6.5 μmol kg^{-1} low of the Winkler measured 36.4 μmol kg^{-1} at 1400 m depth.

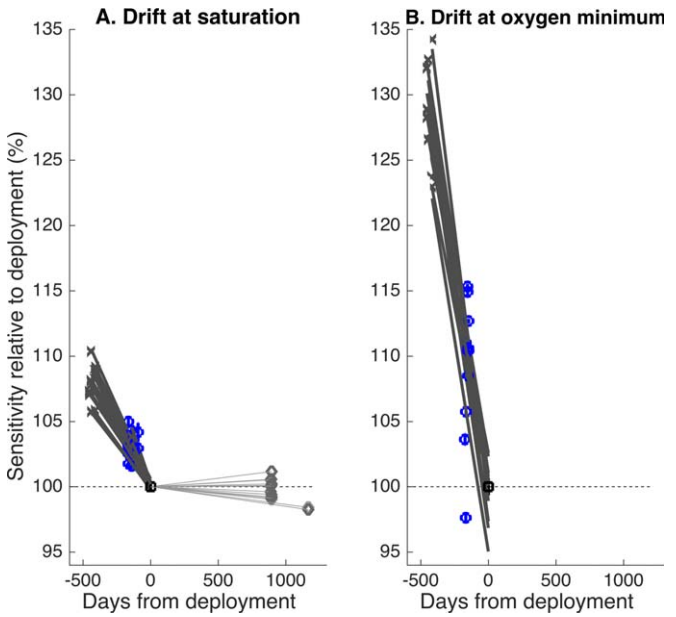

Fig. 11. The change in optode sensitivity vs. time in saturated surface waters (saturation $[O_2]$ of ~ 275 μmol kg^{-1} shown) and at the O_2 minimum (ave. $[O_2]$ 40 μmol kg^{-1}). Percent change in optode sensitivity relative to that calculated at deployment is on the *Y*-axis. Near atmospheric measurements (A) are calculated using: factory calibration (gray *x*'s), lab calibrations (blue circles), in situ atmospheric calibrations (black square), and in situ drift rates (gray and red diamonds). Oxygen minimum values are the percent differences of the factory calibrations (gray *x*'s) and laboratory calibrations (blue circles) plotted relative to Winkler-determined O_2 values from one of the deployment casts (black square). Optode measurements at depth were corrected for pressure effects [3.2%/1000 db, Uchida et al. (2008)], which include the effect of pressure on both the sensor and oxygen solubility. Black lines in both plots are the linear fit to pre-deployment values with a slope of $-7.2\% \pm 1.5\%$ yr^{-1} at saturation and $-23\% \pm 3.2\%$ yr^{-1} in the oxygen minimum. Gray lines (A) are the drift rate post-deployment in the surface ocean.

Johnson et al. (2015) argues that a simple gain correction can be used with atmospheric measurements for the entire water column, based on the minimal difference between factory calibrated optode measurements in oxygen deficient zones and zero. Bittig and Körtzinger (2015) provide more evidence for linear drift across all oxygen saturations by performing multiple calibrations on an optode through time. It is possible that while optode behavior in the absence of oxygen does not change, the oxygen calibration surface changes

shape during drift such that low oxygen response changes more rapidly than near saturation values. Bittig and Körtzinger (2015) do observe an "oxygen independent offset" near zero but that offset is positive and does not match the oxygen minimum results shown here. More research into the drift behavior of optodes in low oxygen water is needed.

Comments and recommendations

Air measurements provide a method for accurate in situ calibration of oxygen optodes. We have demonstrated that they can be used to determine both initial offset and subsequent drift to a few tenths of a percent. Raising the optodes further out of the water provides a cleaner signal with less noise from splashing waves, making it easier to characterize long-term drift. Optodes drift rapidly prior to deployment, and we have shown here evidence for continued drift after deployment in 10 out of 14 optodes using a 95% confidence interval and 12 out of 14 optodes using a 68% confidence interval. The observed drift in some cases is great enough that it will significantly impact air-sea oxygen flux calculations from these floats. It is unknown whether all optodes continue to drift post deployment, but the observed drift in the majority of optodes in this study argues strongly for the need to monitor in situ drift rates to calculate air-sea gas fluxes. Optodes show promise for low oxygen measurements, but drift needs to be better understood and categorized to explain the differences seen here between drift at saturation and drift at low oxygen values.

To summarize, we recommend the following procedures for future in situ atmospheric calibrations on both newly deployed and (where possible) existing Argo oxygen floats:

1. Atmospheric measurements should be made at night or only nighttime data should be used for air calibrations
2. Repeated measurements should be made at each air period, both because the initial measurement may not be representative of an average and multiple measurements allow filtering to remove highly variable air periods
3. Filtering atmospheric data can reduce variance in air measurements and improve accuracy if air periods are biased during certain times of the year
4. Raising optodes higher out of the water will reduce the variance in air periods

The following uncertainties must still be dealt with:

1. Optode calibrations must be performed across the entire temperature range expected for atmospheric conditions
2. A simple gain correction does not appear to adequately correct low oxygen data

References

Bittig, H. C., and A. Körtzinger. 2015. Tackling oxygen optode drift: Near-surface and in-air oxygen optode measurements on a float provide an accurate in-situ reference. J. Atmos. Ocean. Technol. **32**: 1536–1543. doi:10.1175/JTECH-D-14-00162.1

Bushinsky, S. M., and S. Emerson. 2013. A method for in-situ calibration of Aanderaa oxygen sensors on surface moorings. Mar. Chem. **155**: 22–28. doi:10.1016/j.marchem.2013.05.001

D'Asaro, E. A., and C. McNeil. 2013. Calibration and stability of oxygen sensors on autonomous floats. J. Atmos. Ocean. Technol. **30**: 1896–1906. doi:10.1175/JTECH-D-12-00222.1

Demas, J. N., B. A. DeGraff, and P. B. Coleman. 1999. Oxygen sensors based on luminescence quenching. Anal. Chem. News Featur. **December** 1: 793–800. doi:10.1021/ac9908546

Diaz, R. J., and R. Rosenberg. 2008. Spreading dead zones and consequences for marine ecosystems. Science **321**: 926–929. doi:10.1126/science.1156401

Emerson, S. 2014. Annual net community production and the biological carbon flux in the ocean. Global Biogeochem. Cycles **28**: 1–12. doi:10.1002/2013GB004680

Emerson, S., P. Quay, D. M. Karl, C. Winn, and L. M. Tupas. 1997. Experimental determination of the organic carbon flux from open-ocean surface waters. Nature **389**: 951–954. doi:10.1038/40111

Emerson, S., C. Stump, and D. Nicholson. 2008. Net biological oxygen production in the ocean: Remote in situ measurements of O2 and N2 in surface waters. Global Biogeochem. Cycles **22**: GB3023. doi:10.1029/2007GB003095

Emerson, S. R., and S. M. Bushinsky. 2014. Oxygen concentrations and biological fluxes in the open ocean. Oceanography **27**: 168–171. doi:10.5670/oceanog.2014.20

Fiedler, B., P. Fietzek, N. Vieira, P. Silva, H. C. Bittig, and A. Körtzinger. 2013. In situ CO2 and O2 measurements on a profiling float. J. Atmos. Ocean. Technol. **30**: 112–126. doi:10.1175/JTECH-D-12-00043.1

García, H. E., and L. I. Gordon. 1992. Oxygen solubility in seawater: Better fitting equations. Limnol. Oceanogr. **37**: 1307–1312. doi:10.4319/lo.1992.37.6.1307

Glueckauf, E. 1951. The composition of atmospheric air, p. 3–10. *In* Thomas F. Malone [ed]. Compendium of meteorology. American Meteorological Society.

Gruber, N., and others. 2009. Adding oxygen to Argo: Developing a global in-situ observatory for ocean deoxygenation and biogeochemistry. *In* Proceedings of Ocean Obs '09: Sustained Ocean Observations and Information for Society.

Helly, J. J., and L. A. Levin. 2004. Global distribution of naturally occurring marine hypoxia on continental margins. Deep-Sea Res. Part I Oceanogr. Res. Pap. **51**: 1159–1168. doi:10.1016/j.dsr.2004.03.009

Hofmann, A. F., E. T. Peltzer, P. M. Walz, and P. G. Brewer. 2011. Hypoxia by degrees: Establishing definitions for a changing ocean. Deep-Sea Res. Part I Oceanogr. Res. Pap. **58**: 1212–1226. doi:10.1016/j.dsr.2011.09.004

Johnson, K. S., Plant, J. N., Riser, S. C., and Gilbert, D. (2015). Air oxygen calibration of oxygen optodes on a

profiling float array. Journal of Atmospheric and Oceanic Technology, **32**: 2160–2172. doi:10.1175/JTECH-D-15-0101.1

Kamykowski, D., and S.-J. Zentara. 1990. Hypoxia in the world ocean as recorded in the historical data set. Deep-Sea Res. Part I Oceanogr. Res. Pap. **37**: 1861–1874. doi:10.1016/0198-0149(90)90082-7

Kester, D. R., Duedall, I. W., Connors, D. N., and Pytkowicz, R. M. 1967. Preparation of Artificial Seawater. Limnology and Oceanography, **12**: 176–179.

Körtzinger, A., J. Schimanski, and U. Send. 2005. High quality oxygen measurements from profiling floats: A promising new technique. J. Atmos. Ocean. Technol. **22**: 302–308. doi:10.1175/JTECH1701.1

Laws, E. A., E. D'Sa, and P. Naik. 2011. Simple equations to estimate ratios of new or export production to total production from satellite-derived estimates of sea surface temperature and primary production. Limnol. Oceanogr.: Methods **9**: 593–601. doi:10.4319/lom.2011.9.593

Liang, J.-H., C. Deutsch, J. C. McWilliams, B. Baschek, P. P. Sullivan, and D. Chiba. 2013. Parameterizing bubble-mediated air-sea gas exchange and its effect on ocean ventilation. Global Biogeochem. Cycles **27**: 894–905. doi:10.1002/gbc.20080

Roemmich, D., and others. 2009. The Argo Program: Observing the global ocean with profiling floats. Oceanography **22**: 34–43. doi:10.5670/oceanog.2009.36

Rohlf, F. J., and R. R. Sokal. 2003. Statistical tables, 3rd ed. W. H. Freeman and Company.

Siegel, D. A., K. O. Buesseler, S. C. Doney, S. F. Sailley, M. J. Behrenfeld, and P. W. Boyd. 2014. Global assessment of ocean carbon export by combining satellite observations and food-web models. Global Biogeochem. Cycles **28**: 181–196. doi:10.1002/2013GB004743

Sokal, R. R., and F. J. Rohlf. 2003. Biometry, 3rd ed. W. H. Freeman and Company.

Stendardo, I., and N. Gruber. 2012. Oxygen trends over five decades in the North Atlantic. J. Geophys. Res. **117**: C11004. doi:10.1029/2012JC007909

Stokes, M. D., and G. N. Somero. 1999. An optical oxygen sensor and reaction vessel for high-pressure applications. Limnol. Oceanogr. **44**: 189–195. doi:10.4319/lo.1999.44.1.0189

Stramma, L., P. Brandt, J. Schafstall, F. Schott, J. Fischer, and A. Körtzinger. 2008. Oxygen minimum zone in the North Atlantic south and east of the Cape Verde Islands. J. Geophys. Res. **113**: C04014. doi:10.1029/2007JC004369

Stramma, L., S. Schmidtko, L. A. Levin, and G. C. Johnson. 2010. Ocean oxygen minima expansions and their biological impacts. Deep-Sea Res. Part I Oceanogr. Res. Pap. **57**: 587–595. doi:10.1016/j.dsr.2010.01.005

Takeshita, Y., and others. 2013. A climatology-based quality control procedure for profiling float oxygen data. J. Geophys. Res. Oceans **118**: 5640–5650. doi:10.1002/jgrc.20399

Tengberg, A., and others. 2003. Optodes to measure oxygen in the aquatic environment. Sea Technol. **44**: 10–15.

Tengberg, A., and others. 2006. Evaluation of a lifetime-based optode to measure oxygen in aquatic systems. Limnol. Oceanogr.: Methods **4**: 7–17. doi:10.4319/lom.2006.4.7

Uchida, H., T. Kawano, I. Kaneko, and M. Fukasawa. 2008. In situ calibration of optode-based oxygen sensors. J. Atmos. Ocean. Technol. **25**: 2271–2281. doi:10.1175/2008JTECHO549.1

Volk, T., and M. I. Hoffert. 1985. Ocean carbon pumps: Analysis of relative strengths and efficiencies in ocean-driven atmospheric CO_2 changes. *In* E. T. Sundquist and W. S. Broecker [eds.] The carbon cycle and atmospheric CO_2: Natural variations archean to present. American Geophysical Union.

Wanninkhof, R. 2014. Relationship between wind speed and gas exchange over the ocean revisited. Limnol. Oceanogr.: Methods **12**: 351–362. doi:10.4319/lom.2014.12.351

Wearn, R. B., and N. G. Larson. 1982. Measurements of the sensitivities and drift of Digiquartz pressure sensors. Deep-Sea Res. Part A Oceanogr. Res. Pap. **29**: 111–134. doi:10.1016/0198-0149(82)90064-4

Westberry, T., M. J. Behrenfeld, D. A. Siegel, and E. Boss. 2008. Carbon-based primary productivity modeling with vertically resolved photoacclimation. Global Biogeochem. Cycles **22**: 1–18. doi:10.1029/2007GB003078

Whitney, F. A., H. J. Freeland, and M. Robert. 2007. Persistently declining oxygen levels in the interior waters of the eastern subarctic Pacific. Prog. Oceanogr. **75**: 179–199. doi:10.1016/j.pocean.2007.08.007

Zeebe, R. E., and D. A. Wolf-Gladrow. 2001. CO_2 in seawater: Equilibrium, kinetics, isotopes, Elsevier, Oceanography Series.

Acknowledgments

We thank Diane Perry and Charles Stump for their work on lab and field calibration of optodes prior to deployment. Rick Rupan and Dale Ripley constructed the floats at the UW float laboratory. Float deployments were made possible by the crews of the CCGS John P. Tully and R/V *Melville*, assistance from Marie Robert and Doug Yelland of the Fisheries and Oceans Canada Institute of Ocean Sciences, Michael Craig and Jennifer Keene of the NOAA-Pacific Marine Environmental Laboratory, and the students of the UW senior thesis class on board the Melville. OSP mooring deployments were possible through collaboration with Meghan Cronin and Keith Ronnholm of NOAA PMEL. This work was supported by a National Science Foundation grant OCE-1129112 and a National Science Foundation Graduate Research Fellowship.

Submitted 13 November 2015
Revised 29 January 2016
Accepted 14 March 2016

Associate editor: Mike DeGrandpre