LIMNOLOGY
and
OCEANOGRAPHY: METHODS

Limnol. Oceanogr.: Methods 10, 2012, 921–933

A novel electrochemical calibration setup for oxygen sensors and its use for the stability assessment of Aanderaa optodes

Henry C. Bittig, Björn Fiedler, Tobias Steinhoff, and Arne Körtzinger*
GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

Abstract

We present a laboratory calibration setup for the individual multi-point calibration of oxygen sensors. It is based on the electrochemical generation of oxygen in an electrolytic carrier solution. Under thorough control of the conditions, i.e., temperature, carrier solution flow rate, and electrolytic current, the amount of oxygen is strictly given by Faradays laws and can be controlled to within ± 0.5 µmol L^{-1} (2 SD). Whereas Winkler samples can be taken for referencing with a reproducibility between triplicates of 0.8 µmol L^{-1} (2 SD), the calibration setup can provide a Winkler-free way of referencing with an accuracy of ± 1.2 µmol L^{-1} (2 SD). Thus calibrated oxygen optodes have been deployed in the Southern Ocean and the Eastern Tropical Atlantic both in profiling and underway mode and confirm the validity of the laboratory calibrations to within few µmol L^{-1}. In two cases, the optodes drifted between deployments, which was easily identified using the calibration setup. The electrochemical calibration setup may thus facilitate accurate oxygen measurements on a large scale, and its small size makes it possible to configure as a mobile, sea-going, Winkler-free system for oxygen sensor calibrations.

Dissolved oxygen for long has been and still is one of the most common and most widely measured parameters of oceanography. Its observation has an unmatched history of more than a century. Stressing this, the laboratory principle for discrete samples remained essentially unchanged since Winkler (1888) and still is, with some improvements, the method of choice for reference measurements (Dickson 1995). Besides, dissolved oxygen has been termed the "oceanographer's canary bird" as it is influenced by major biogeochemical and physical processes (primary production, remineralization, air-sea gas exchange, and water mass ventilation) and thus represents a most sensitive key parameter in marine global change research (Körtzinger et al. 2004; Keeling et al. 2010). In addition, the broad interest in dissolved oxygen measurements is illustrated by the plethora of its measurement platforms: long time series moorings (Karl and Lukas 1996; Steinberg et al. 2001), repeat hydrography cruises (e.g., Stendardo et al. 2009) or underway measurements (Juranek et al. 2010), autonomous instruments (Körtzinger et al. 2005; Gruber et al. 2010), or during incubations and mesocosm studies (Robinson and Williams 2005).

Whereas there are sensors available to assist in such studies, they have to fulfill certain requirements of reliability, long-term stability, dynamic response, precision, and accuracy. Especially the latter is a critical issue. For example, the main uncertainty of a net community production estimate from in-situ oxygen and nitrogen gas measurements stems from the oxygen sensor calibration (Emerson et al. 2008). Their estimate of a surface mixed layer biological oxygen production of 4.8 ± 2.7 mol m^{-2} yr^{-1} at the Hawaii Ocean time series station is prone to a ± 2.5 mol m^{-2} yr^{-1} uncertainty due to an 0.5 % (approx. 1.2 µmol L^{-1}) error in the dissolved oxygen data input. For the application on Argo floats, an accuracy threshold of 5 µmol kg^{-1} has been defined for the data to be of useful quality to address scientific objectives, whereas the accuracy target for the desired data quality has been set to 1 µmol kg^{-1} (Gruber et al. 2010).

*Corresponding author: E-mail: *E-mail: hbittig@geomar.de

Acknowledgments

The authors want to thank the captains, crew, and scientists of R/V *Polarstern* ANT-XXVII/1 and ANT-XXVII/2 as well as R/V *Maria S. Merian* MSM 18/3. Many thanks go to Carolina Dufour (LEGI/CNRS, Université de Grenoble, Grenoble, France) for her patience with and dedication to the Winkler samples of ANT-XXVII/2, Andreas Pinck (GEOMAR, Kiel, Germany) for the design of the improved current source, Martina Lohmann (GEOMAR, Kiel, Germany) for measuring the Winkler samples of MSM 18/3, Jostein Hovdenes (AADI, Bergen, Norway) for support with the optodes, Andreas Schmuhl and Detlef Foge (AMT GmbH, Rostock, Germany) for helpful advice on their O_2 generator, and Sebastian Fessler (GEOMAR, Kiel, Germany) for assistance in the early stages of the project. Financial support by the following projects is gratefully acknowledged: OCEANET of the WGL Leibniz Association, O_2-Floats (KO 1717/3-1) and the SFB754 of the German Science Foundation (DFG), and the project SOPRAN (03F0611A and 03F0462A) of the German Research Ministry (BMBF).

DOI 10.4319/lom.2012.10.921

On the other hand, long-term stability of different sensor designs remains a critical issue. For optical oxygen sensors such as the Aanderaa optode, there is no evidence of drift during a deployment period (Tengberg et al. 2006). Between deployments, however, there are several observations that processes yet unidentified lead to a change in the sensor response, e.g., between factory calibration and in-situ data (Takeshita et al. 2010; Neill 2011 pers. comm.; this study). Whereas this is more a sensor issue, a dedicated calibration facility could improve the data quality through regular and accurate recalibrations. This emphasizes the need for a simple calibration setup.

The most common calibration approach is an in-situ calibration against Winkler samples of a colocated CTD cast. This can be done with high accuracy (Uchida et al. 2008), but is tedious for a larger number of sensors and logistically not always feasible. The main disadvantage is that the reference points for calibration are limited to the set of field conditions (oxygen content and temperature) encountered during the sampling time. Furthermore, they are superimposed by additional ambient effects like a pressure dependence. All data outside the parameter range provided by the field conditions during calibration are accessible only through extrapolation and thus less reliable. This is less of an issue for ship cruises with reference measurements throughout the entire cruise. It becomes more important for moored deployments with calibration opportunities typically only at the beginning and at the end of the deployment periods, or even worse for Argo-O_2 floats with a single deployment profile only.

The less popular approach is a multi-point laboratory calibration in which a set of reference points under controlled conditions are used for calibration. These should be so widely distributed as to cover all expected field conditions, and the field measurements are essentially interpolations between these reference points, which gives more confidence with regard to data quality. To adjust the temperature and the oxygen content, these variables have to be forced in a controlled manner. The former can be done by submerging the sensors in a thoroughly mixed, thermostated bath, whereas the latter can be accomplished by usage of gas cylinders of N_2 and O_2/N_2 mixtures and bubbling stones, which is done in all such setups known to the authors. As reference for the absolute oxygen content, Winkler samples or previously Winkler-calibrated sensors are used. This is crucial because a complete equilibration with the gas mixture requires both extended equilibration times and constant ambient pressure. Accuracies as high as 0.5 µmol L^{-1} can be achieved by such calibration setups (Neill 2011 pers. comm.). Here we present a different way to force the oxygen content by using electrochemistry instead of gas mixtures. This reduces the size of the setup significantly and enhances both the portability and ease of use.

The electrochemical approach is based on the electrolysis of aqueous solutions, where at the anode molecular oxygen is produced (Eq. 1).

$$\text{Anode} \quad \overset{\pm 0}{O_2} + 2\,H_2O + 4\,e^- \rightleftharpoons 4\,\overset{-II}{OH^-} \qquad E_0 = +0.401\ \text{V} \qquad (1)$$

$$\text{Cathode} \quad 2\,\overset{+I}{H_2O} + 2\,e^- \rightleftharpoons \overset{\pm 0}{H_2} + 2\,OH^- \qquad E_0 = -0.828\ \text{V} \qquad (2)$$

If the flow rate (V/t) through the electrolytic cell and the electrolytic current (Q/t) is set, the oxygen concentration of the electrolytic carrier solution is strictly given by Faradays laws (Eq. 3, 4), where n is the number of moles, z the number of electrons transferred, F the Faraday constant (96485 C mol^{-1}), I the electrolytic current, t the time, and c the volumetric concentration, respectively.

$$n \cdot z \cdot \mathrm{F} = I \cdot t \qquad (3)$$

$$\rightarrow \quad c = \frac{n}{V} = \frac{1}{z \cdot \mathrm{F}} \cdot \frac{I}{\frac{V}{t}} \qquad (4)$$

$$c \,/\, \mu\text{mol L}^{-1} = 155.5 \cdot \frac{I \,/\, \text{mA}}{\frac{V}{t} \,/\, \text{mL min}^{-1}}$$

For repeatability, the carrier solution has to be degassed, i.e., stripped of oxygen, before the electrolysis to ensure a common background between different runs. Thus, the carrier solution obtains a defined concentration of dissolved oxygen that can be used in a flow system-based calibration setup.

Materials and procedures

Materials

The calibration setup is based on the degassing of an electrolyte or carrier solution and the subsequent electrochemical in-situ production of dissolved oxygen. That solution is then adjusted in temperature and passed to the sensors for calibration.

The setup, shown schematically in Fig. 1, is designed as a flow system with an electrochemical oxygen generator (G200, AMT Analysenmesstechnik GmbH, Rostock/Germany) as the central element. A flow meter, a cryostat, a section to tap Winkler samples, and a pressure gauge were added as auxiliaries.

From the reservoir, the carrier solution, a 0.02 M sodium hydroxide solution, is transported through the flow system by means of a peristaltic pump (ISM829 Reglo Analog, Ismatec) provided by AMT. Downstream, the AMT generator contains a built-in degassing unit for the carrier solution and an electrolytic cell. Two separate circuits are used for the cathodic and anodic side. The degassing is based on maintaining a vacuum outside gas-permeable tubing through which the anodic carrier solution is passed and thus stripped of all dissolved gases. To ensure a stable electrolysis, the flow rate through the anode is controlled and the pump regulated by a high precision flow meter (miniCori-Flow M13, Bronkhorst Mättig GmbH) installed between the pump and the generator. Moreover, triplicate Winkler samples can be taken as references between the generator and custom-made flow-through cells for the oxygen sensors. Several flow-through cells and sensors

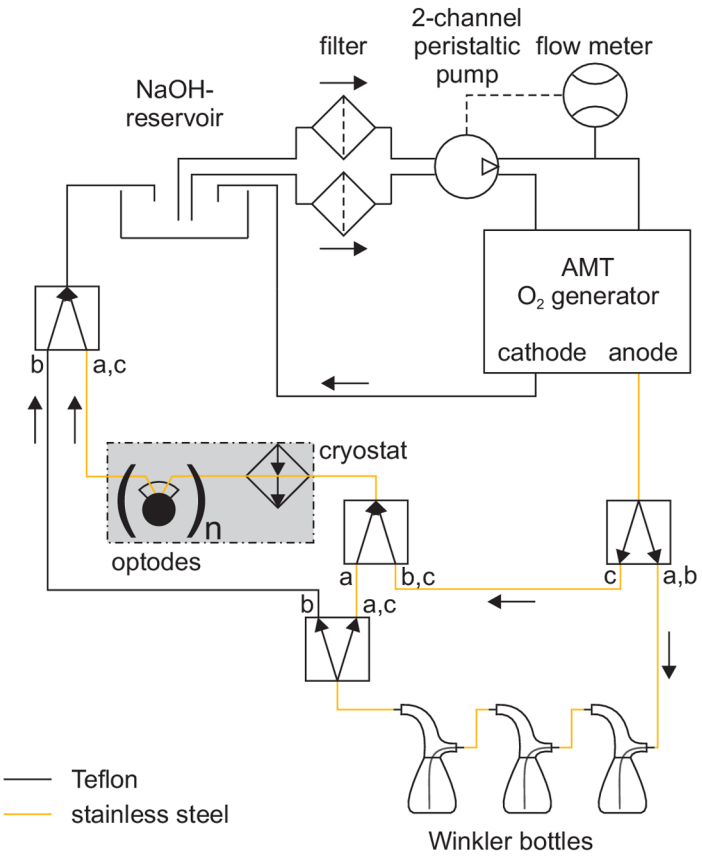

Fig. 1. Schematic of the calibration setup. The dash-dot encircled shaded gray area indicates the thermostated bath.

can be assembled in a row and calibrated simultaneously. They are completely submerged in a thermostated bath, in which the carrier solution has been brought to the same temperature. All the other parts of the system, including the generator itself, the carrier solution reservoir, and the Winkler bottles, are at room temperature. The tubing for the carrier solution downstream of the electrolytic cell is made of stainless steel, to exclude any air contamination. Valves can be used to bypass the oxygen sensors (option b in Fig. 1) or the Winkler bottles (option c in Fig. 1), respectively. A pressure sensor was added at the generator's degassing unit to monitor the residual vacuum pressure.

The flow rate through the generator is restricted between 10 mL min^{-1} and 12 mL min^{-1} to maintain both a homogeneous solution and complete dissolution of oxygen at the electrode. With an electrolysis current of 0 mA to 20 mA, oxygen concentrations between 0 µmol L^{-1} and 311 µmol L^{-1} (120% oxygen saturation at 25°C) can be achieved without limitations on distinct saturation levels. The temperature can be chosen freely within 1°C–36°C.

Procedures

A typical parameter set at constant generator settings (16 mA) is shown in Fig. 2. The strong dependence of the optodes phase signal on temperature is clearly visible. All oxygen data are based on these two raw parameters and depend both on an adequate functional model of sensor response and an adequate set of calibration coefficients (see Fig. 3).

For an eligible calibration reading, an arbitrary stability criterion of the drift in the oxygen concentration, smaller than 0.02 µmol L^{-1} min^{-1} over a period of 15 min has to be fulfilled. Under these conditions, only negligible gradients exist between generator exit, Winkler bottles, and sensor flow-through cells.

Standard procedures for Winkler samples require the bottles to be overflown by three times their volume before fixation (Dickson 1995). However, this is not feasible in a flow system with only 10 mL min^{-1} flow rate. Consequently, it has

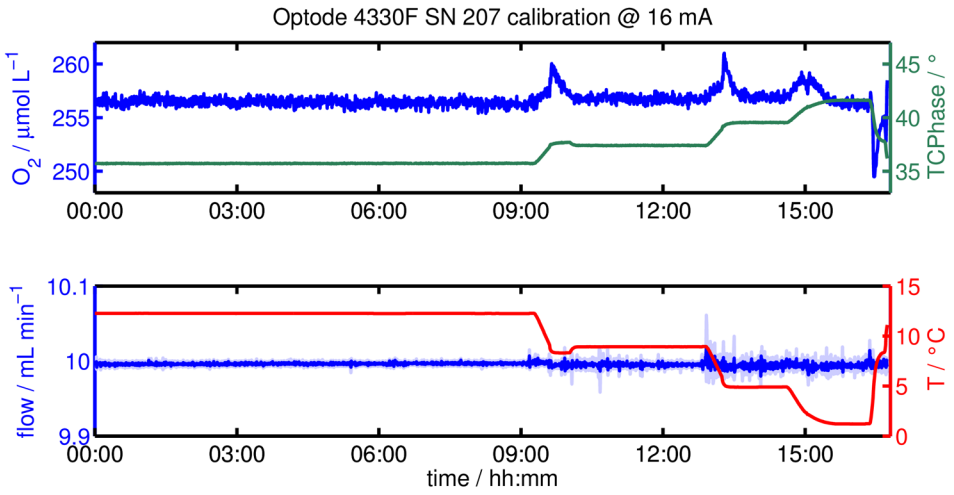

Fig. 2. Plot of optode and calibration setup parameters at constant generator settings and different thermostated bath temperatures. Upper panel: Oxygen concentration and optode phase signal. Lower panel: Carrier solution flow rate and temperature.

been adopted by flushing the bottles from bottom to top within the closed system using glass-made flow caps for the Winkler bottles (see schematic in Fig. 1). At sampling, after the stability criterion has been reached, the solution in the bottle neck possibly contaminated by atmospheric oxygen is then replaced by the solution from the Winkler flow cap above and thus contamination is minimized. At analysis, the pickled sample is acidified by twice the amount of sulphuric acid to account for the high pH of the carrier solution. If Winkler samples are taken at each calibration point, about 3 to 4 points can be done per working day.

The obtained data of temperature, sensor phase, and Winkler oxygen can then be fitted to any desired model of sensor oxygen response, an example of the Aanderaa optode oxygen response shown in Fig. 3. It is based on the Uchida et al. (2008) model. Unlike the original publication, the oxygen concentration is not used directly but is converted to partial pressure pO_2 and then used as fit parameter in the model.

The functional model (Eq. 6) is inspired by the Stern-Volmer equation (Eq. 5) substituting the lifetimes τ in the presence of oxygen and τ_0 in the absence of oxygen with the phase signals P and P_0, respectively. Additional polynomials are introduced to account for the temperature dependence of the Stern Volmer constant K_{SV} and zero phase signal (Eqs. 7 and 8) as well as to scale the phase signal again (Eq. 9).

$$\frac{\tau_0}{\tau} = 1 + K_{SV} \cdot c(O_2) \quad (5)$$

$$\frac{P_0}{P'} = 1 + K_{SV} \cdot pO_2 \quad (6)$$

$$K_{SV} = c_0 + c_1 \cdot T + c_2 \cdot T^2 \quad (7)$$

$$P_0 = c_3 + c_4 \cdot T \quad (8)$$

$$P' = c_5 + c_6 \cdot P \quad (9)$$

$$pO_2 = \frac{c(O_2)}{\alpha(O_2)} \quad (10)$$

Strictly, the Stern-Volmer equation is derived from molecular quenching kinetics that require the O_2 concentration inside the sensor membrane to be used (Eq. 5). Due to different solubilities, however, this is not the concentration of the ambient medium, i.e., sea water. Instead, the equilibrium between sensor membrane and environment is characterized by equal partial pressures pO_2, which is used for all calculations (Eq. 6).

The sensor membrane oxygen solubility as proportionality factor between partial pressure and concentration is thus included in the Stern Volmer constant K_{SV} by this approach. On the other hand, the Henry's law solubility constant $\alpha(O_2)$/µmol L^{-1} Pa^{-1} is used to convert the ambient sea water concentration to partial pressure (Eq. 10). The nonlinear response of the optode is clearly visible from the obtained data (Fig. 3).

The 0.02 M NaOH carrier solution may be considered as being nearly freshwater. For highest accuracy, a salinity correction should be applied to the partial pressure calculation, as saturation levels are affected by ionic interactions with the medium. However, sea water solubility (Garcia and Gordon 1992) is not per se applicable to a 0.02 M NaOH carrier solution with considerable different chemical composition. Oxygen solubility in different salt solutions was analyzed by Clegg and Brimblecombe (1990). Their parameterization gives a salinity correction factor for 0.02 M NaOH of 0.99158(6) between 1°C and 36°C, by which the (freshwater) solubility $\alpha(O_2)$ in Eq. 10 should be scaled. The factor relates to an effective sea water salinity of 1.25 psu–1.60 psu.

Performance assessment

Laboratory evaluation

The resolutions of the environmental variables, i.e., temperature, carrier solution flow rate, and electrolytic current, are 0.01°C, 0.001 mL min^{-1}, and 0.001 mA, respectively. They

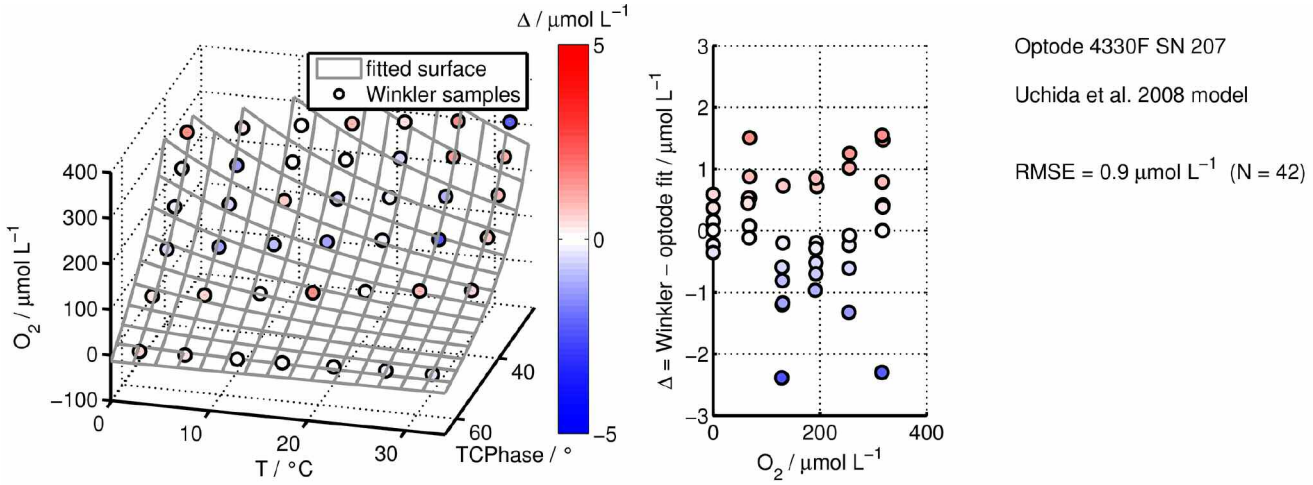

Fig. 3. Calibration response of an Aanderaa optode. The oxygen concentration is plotted against temperature and phase as independent variables (left panel). The Uchida calibration model fits the sensor's functional behavior by a Stern-Volmer-inspired, nonlinear approach. The color code and the middle panel give the absolute difference Δ between Winkler samples and fitted surface. Statistical figures for the model are also given in the right panel.

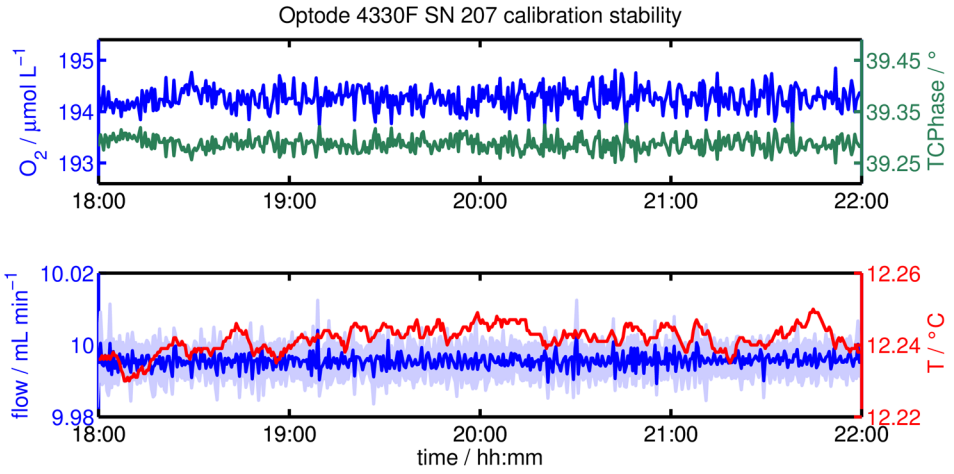

Fig. 4. Sample plot of a stable state of the calibration system. Upper panel: Oxygen concentration and phase signal as sensor data. Lower panel: Carrier solution flow rate and temperature as environmental conditions. Dark blue shows the flow rate averaged over the optode's sampling interval (30 s), whereas light blue includes the flow rate standard deviation within that interval.

can be constrained (2 SD) to within 0.007°C, 0.01 mL min^{-1}, and 0.01 mA, respectively, and the stability of the system is illustrated in Fig. 4. The variability in the environmental variables amounts to a theoretical uncertainty in the O_2 concentration of ± 0.5 µmol L^{-1} (2 SD). This is confirmed by the oxygen concentration observed by the sensors, which is stable to within the same range and mainly affected by fluctuations in the flow rate with a time lag of approx. 15 min.

The generator may be used to generate different oxygen concentrations only, while the Winkler samples provide absolute numbers for the sensor calibration. However, the comparison of the Winkler samples with the theoretical value from the generator settings (equation 4) indicates an alternative way of referencing.

The best fit between Winkler samples and generator settings is illustrated in Fig. 5. It can be seen that the slope is close enough to 1 to take 100% electrolysis efficiency as granted. In addition, their difference is shown in Fig. 6. This directly gives the accuracy and precision of the calibration setup without Winkler referencing, +4.7 µmol L^{-1} and ±1.2 µmol L^{-1}, respectively. At the same time, the offset observed is independent of the oxygen concentration and the temperature.

Assuming an incomplete degassing step, which is independent of the electrolytic current or thermostated bath temperature, the background for the electrolytic oxygen addition would increase uniformly. Therefore, a pressure sensor was installed temporarily at the degassing unit and a total residual pressure of 25 ± 2 mbar was observed. Starting from first principles, i.e., equilibrium with the atmosphere and the O_2 mole fraction in air, the degassing pressure equals a residual oxygen concentration of 6.6 ± 0.5 µmol L^{-1}. Without the pressure sensor installed, there is a reduced number of possible leaks. In consequence, the degassing pressure is presumable slightly lower and the 6.6 µmol L^{-1} should be considered as an upper

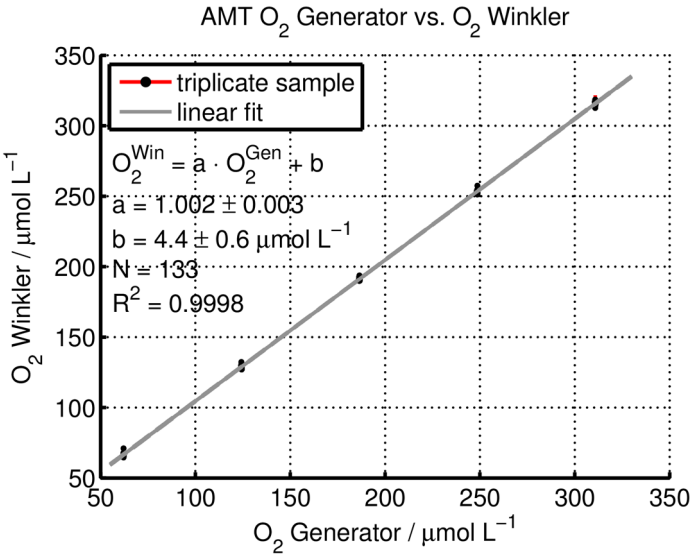

Fig. 5. Nominal generator oxygen concentration after Eq. 4 versus triplicate Winkler samples with standard deviation of the triplicates as red bars. The linear least-squares fit is indicated in gray.

bound for the first principles approach to explain the offset. The accuracy of the calibration setup without Winkler referencing is thus characterized by the repeatability of the degassing step and the level to which the residual oxygen content can be constrained.

Field evaluation

The calibration setup was assessed indirectly by the performance of oxygen optodes during the course of different field deployments. A first set of optodes was deployed in underway mode between Bremerhaven and Las Palmas and attached to the CTD in the Southern Ocean (locations given in Fig. 7a and 7b), whereas a second set of optodes was used in

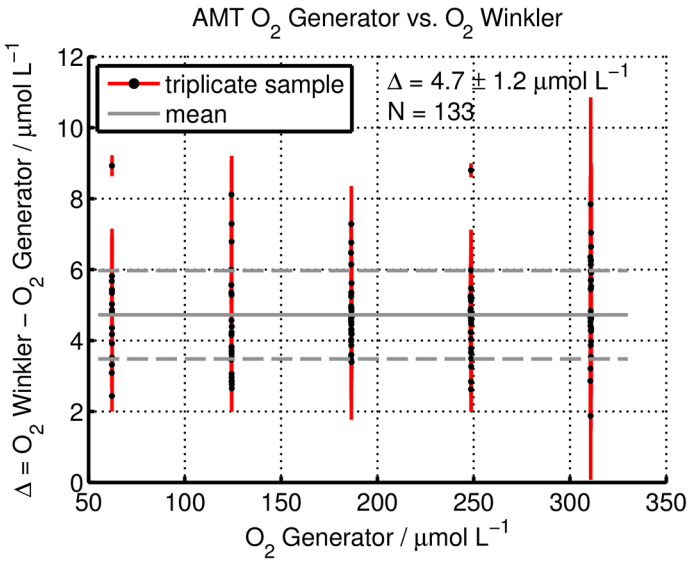

Fig. 6. Difference between triplicate Winkler samples and nominal generator oxygen concentration after Eq. 4 versus oxygen concentration. The mean of the residual Δ is marked in gray, and the standard deviation of the triplicates is indicated as red bars.

the Eastern Tropical Atlantic both in underway and CTD mode (Fig. 7c). In all cases, an individual multi-point calibration was performed before and after the cruises using the laboratory setup. Besides, sodium sulfite was used for the calibration of the zero oxygen level. The field data are based on Winkler bottle data sampled and analyzed according to standard procedures (Dickson 1995). While there is practically no published evidence of drift of optical sensors during deployments (Tengberg et al. 2006), the stability between deployments or between calibration and deployment is not granted (Takeshita et al. 2010; Neill 2011 pers. comm.; this study), for which the pre- and post-cruise laboratory calibrations should give sufficient indication.

At this point, a clear distinction must be done regarding accuracy statements for the oxygen sensors and the calibration setup. Any field evaluation relies on an adequate functional model of the sensor's oxygen response, e.g., the Uchida et al. (2008) model. Consequently, all field samples are compared with the combined performance of the calibration reference points and the functional model (e.g., Fig. 3).

Essentially, all sensor data depend on the proper conversion of the engineering raw data, i.e., phase shift and temperature, to the variables of interest, i.e., oxygen concentration. If the sensor data are excellent but the functional model or calibration parameters do not grasp the sensor's behavior, the derived data will be inaccurate. The same is true for the reverse extreme with an excellent functional model but blurred reference or raw data. Both effects are hard to distinguish and commonly merged under the term "sensor accuracy." On the other hand, the accuracy of the calibration setup itself, i.e., the qual-

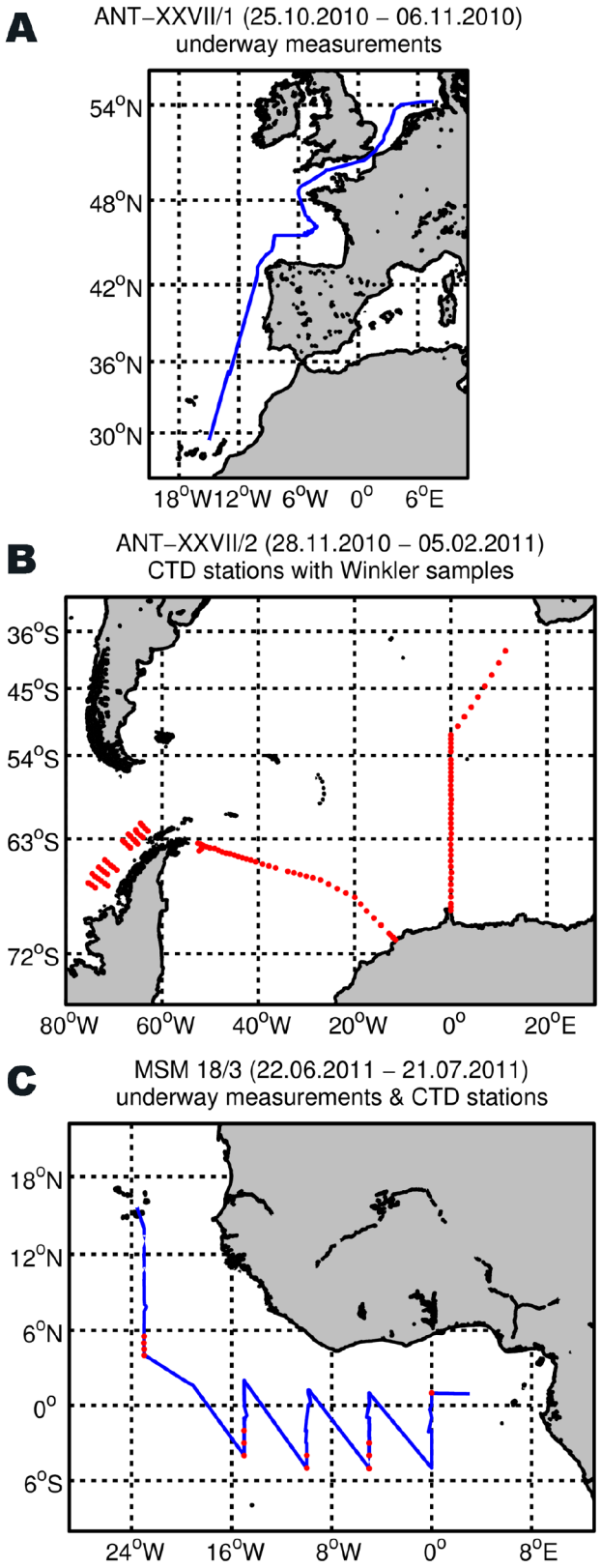

Fig. 7. Cruise plots of all cruises used for the field evaluation. Underway measurements are marked in blue and positions of CTD stations with Winkler bottle data are denoted as red dots.

ity of the calibration reference points, is independent of the sensor and thus independent of the sensor's functional model.

Thus, when using the setup to generate different conditions only, its accuracy is essentially the Winkler accuracy [0.8 µmol L^{-1} from triplicate Winkler samples (2 SD)], whereas for the Winkler-free mode of operation, the accuracy is represented by the degree to which the incomplete degassing of the carrier solution can be characterized, i.e., ± 1.2 µmol L^{-1} from comparison to 133 triplicate Winkler samples.

Still, any field application of (calibrated) oxygen sensors relies on the combination of both the calibration reference points and the functional model. The mean difference between sensor and Winkler reference data and its standard deviation gives a clear indication of their combined performance in the field and the sensor accuracy of interest. On the other hand, the root-mean-square error (RMSE) between laboratory Winkler samples and sensor data can be interpreted as the misfit between calibration reference points and sensor functional model, i.e., the accuracy of the laboratory calibration. The RMSE lies in the range of 0.9 µmol L^{-1} to 1.9 µmol L^{-1} for the optode calibrations discussed in the following paragraphs.

The first set of field data were obtained on R/V *Polarstern* during the cruises ANT-XXVII/1 and ANT-XXVII/2. Two Aanderaa oxygen optodes, a standard model 3830 and a fast response model 4330F, were calibrated before the cruises in October 2010 and recalibrated afterwards in July 2011. Whereas the initial calibration consisted of 29 points between 50 µmol L^{-1} and 315 µmol L^{-1} and between 1°C and 18°C, respectively, the post-cruise calibration was more extensive and contained 42 points between 0 µmol L^{-1} and 315 µmol L^{-1} and between 2°C and 32°C, respectively.

The comparison of both sets of laboratory data for both optodes is shown in Fig. 8a and 8b. The left panels show the shape of the fitted optode response function (Uchida et al. 2008) in gray and the 29 individual points (black circle) on which the initial calibration is based. The 42 points of the post-cruise calibration are distinguished on whether they fall within the calibrated range (yellow circle) or lie outside the initial calibration (purple circle). The statistics in the right panels are given for the points inside the calibrated range only. The color shading and the middle panels show the difference between initial calibration and Winkler data of the post-cruise calibration.

Good agreement was obtained between both data sets: The offset was found to be at the edge of the 95% confidence interval (2 SD) for the 3830 optode or indistinguishable for the 4330F optode, respectively. At the same time, it is obvious that the calibration becomes mediocre if predictions are made outside its range (e.g., between 20°C–32°C), even if it may perform well in distinct regions of the sample space (e.g., 50 µmol L^{-1}–200 µmol L^{-1}).

The first deployment was made in underway mode directly after the initial calibration between 25 Oct 2010 and 6 Nov 2010 on R/V *Polarstern* (ANT-XXVII/1). Only the 3830 model was used and 13 Winkler samples were taken between Bremerhaven and Las Palmas. Their results are shown in Fig. 9. The left panel gives the initial calibration (gray) with the 29 individual points (black circle) on which it is based. The underway Winkler samples are denoted by green circles, and the difference between calibrated optode reading and field Winkler samples is given both as color shading and in the middle panel. The optode's initial calibration is found to be at slightly higher oxygen concentrations than the Winkler samples. However, the offset of 1.9 ± 1.5 µmol L^{-1} is at the edge of significance and the laboratory calibration matches well to the field data.

A second, far more extensive evaluation was performed during the following cruise leg (R/V *Polarstern*, ANT-XXVII/2, 25 Nov 2010–5 Feb 2011, see Fig. 7b) with the sensors attached to the CTD. The CTD was stopped at each bottle stop, such that the sensor readings of temperature and salinity but not of oxygen were allowed to settle, before the Niskin bottle was closed. Following this procedure, a total of 2296 Winkler samples were taken at 122 stations, and the results are shown in Fig. 10 for both sensors.

Again, the color shading gives the difference between the initial laboratory calibration and the Winkler field data. In contrast to previous figures, a distinction is made between surface samples above and close to the thermocline ($p \leq 250$ dbar), marked with green circles, and samples below the thermocline ($p > 250$ dbar), marked with yellow circles.

The data obtained fall within a very narrow temperature range (0.5 ± 3.0°C for all Winkler samples), so that the temperature slope of the calibration cannot be validated by this data set. On the other hand, the maximum oxygen concentration during the laboratory calibration is limited by the maximum electrolysis current allowed by the generator. The issue becomes evident in the left panels of Fig. 10, where most deep samples are just at the edge of the calibrated range, whereas most surface samples are beyond. However, considering their location close to the limits of the calibrated range or slightly beyond, the initial calibration appears to be well suited for the field samples, as the offset Δ between sensor data and Winkler samples is only slightly exceeding its confidence limit (2 SD), and more importantly, there is no trend in the calibration bias visible (middle panel).

Due to issues with the dynamic response of the optodes, as well as increased variability, the scatter is enhanced in the surface region. Moreover, a pronounced response time effect is observed for the 3830 model with an approximately 3-fold response time compared to the 4330F model. Because the bottle stops are made during the upcast, the sensors lag behind the rising oxygen concentration toward the surface, which leads to a dynamically induced underestimation by the optodes, compared with the Winkler samples. This is clearly visible for the 3830 model as a negative bias in the green-circled samples in Fig. 10a, middle panel, and less of an issue for

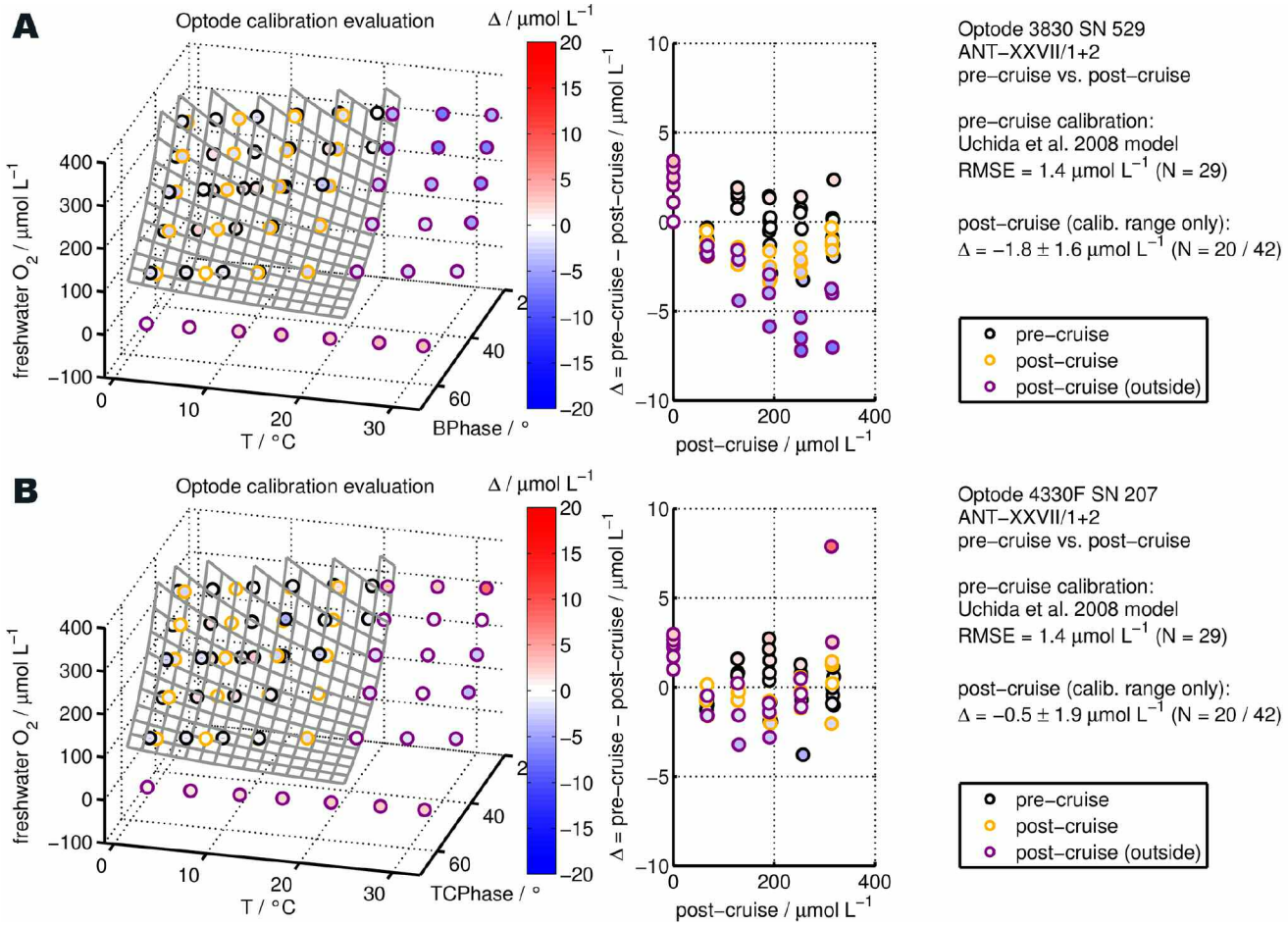

Fig. 8. Repeated calibration of optodes before and after the R/V *Polarstern* cruises ANT-XXVII/1 and ANT-XXVII/2. Left panel: Initial calibration points (black circle) and fitted optode response function (gray) with repeated calibration samples inside (yellow circle) and outside (purple circle) the calibrated range. Middle panel: Difference between pre-cruise calibrated optode reading and post-cruise Winkler samples. The color axis shows the same difference in both panels. Right panel: Statistical figures for both the initial and repeated calibration with 95% confidence interval.

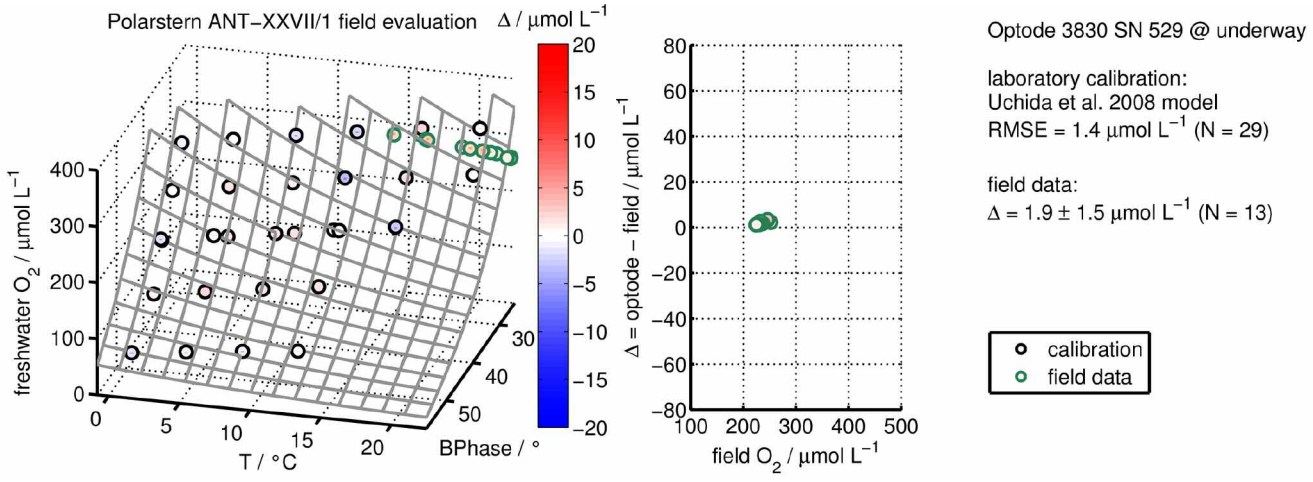

Fig. 9. Underway field evaluation of optode 3830 SN 529 during cruise ANT-XXVII/1. Left panel: Calibration points (black) and optode response as function of phase and temperature (gray) with field samples (green) mapped into the same, freshwater sample space of the calibration. Middle panel: Difference between optode reading and Winkler field samples. The color axis shows the same difference in both panels. Right panel: Statistical figures for both the calibration and the field samples.

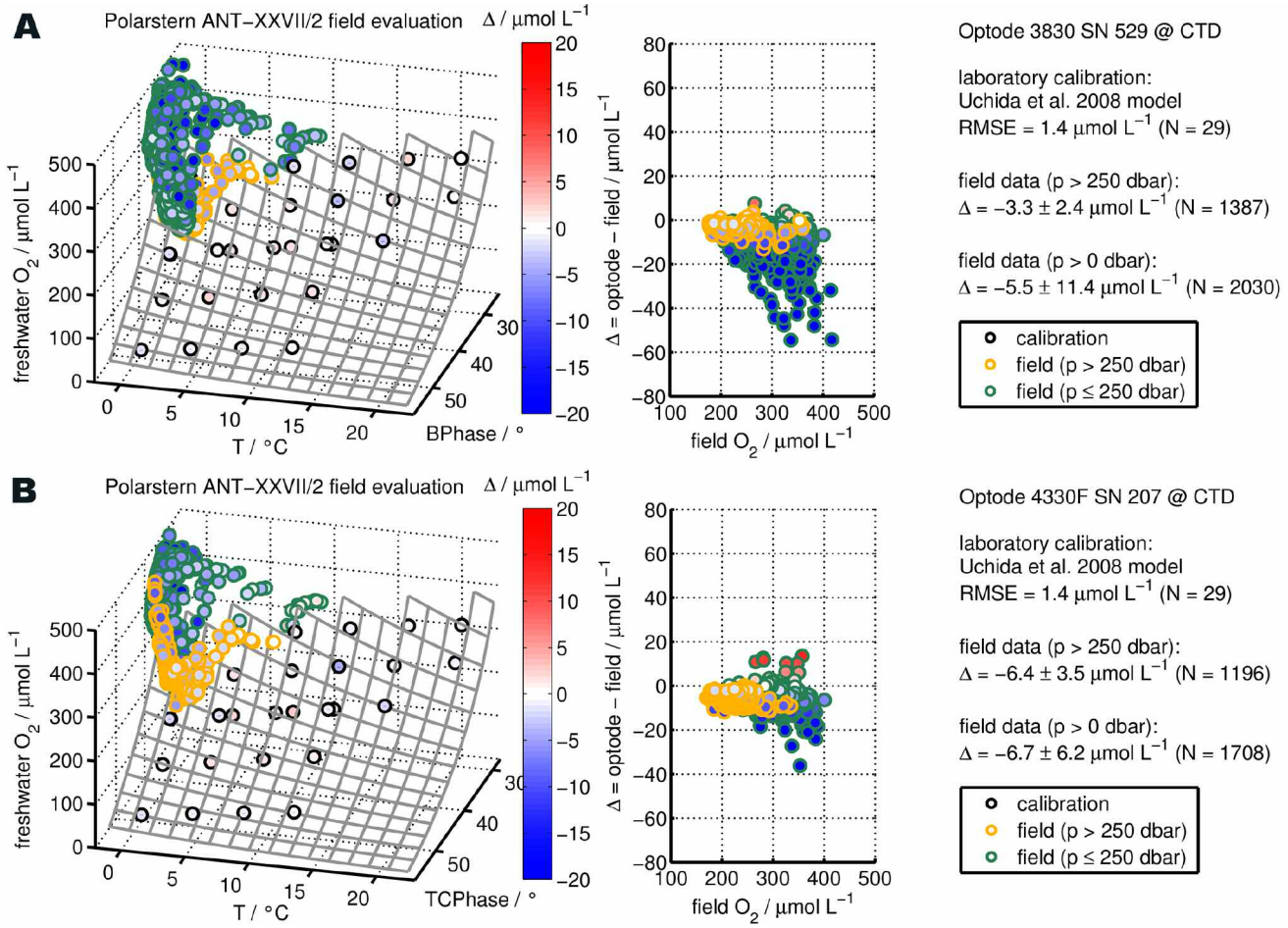

Fig. 10. CTD field evaluation during cruise ANT-XXVII/2. Left panel: Calibration points (black) and optode response as function of phase and temperature (gray) with field samples below the thermocline (yellow) and above the thermocline (green) mapped into the same freshwater sample space of the calibration. Middle panel: Difference between optode calibration and Winkler field samples for samples below the thermocline (yellow) and above the thermocline (green). The color axis shows the same difference in both panels. Right panel: Statistical figures for both the calibration and the field samples.

the fast response model 4330F (Fig. 10b). The latter is confirmed by the same offset Δ for the 4330F sensor for both the deep samples and all samples, including the surface gradient region, indicative of a fast enough sensor for that region.

A second set of field data were acquired on the R/V *Maria S. Merian* cruise MSM 18/3 (21 Jun 2011–21 Jul 2011) to the Eastern Tropical Atlantic. Two Aanderaa optodes, both a standard model 4330, were calibrated before the cruise in April 2011 and recalibrated afterward in December 2011. One was used for underway measurements whereas the other was attached to the CTD.

The initial calibration consisted of an extensive, 42-point calibration between 0 µmol L^{-1} and 315 µmol L^{-1} and 2°C and 32°C, respectively. In contrast, the post-cruise calibration was done with an improved setup as described in the last section, which features an electrolytic current source of up to 30 mA, and is thus capable of generating oxygen levels above 315 µmol L^{-1}. It consists of 42 points ranging from 0% to 130% oxygen saturation and 2°C to 32°C, respectively.

As illustrated in Fig. 11, there is a clear drift of the sensor's response between pre-cruise and post-cruise calibration. While the pre-cruise calibration RMSE is as low as 1.9 µmol L^{-1}, the observed difference between pre-cruise calibrated sensor readings and post-cruise calibration Winkler samples may be an order of magnitude higher. Moreover, both sensors possess a common deployment history (newly purchased and exposed to 2000 dbar several times) and show a comparable drift with a significant change in the sensor response. A similar drift behavior has been observed for other optodes (Neill 2011 pers. comm.). Because the match between field data and sensor data are better using the post-cruise calibration (not shown), the latter is chosen for the field evaluation of the calibration setup to decouple it from the unresolved sensor drift issue.

The optode in underway mode was evaluated against 59 Winkler samples during 26 days of continuous measurements (see Fig. 12).

Whereas there is a bias for the initial calibration in the order of 10 µmol L^{-1} (not shown), the post-cruise calibration gives a good match of –2.2 ± 3.3 µmol L^{-1}.

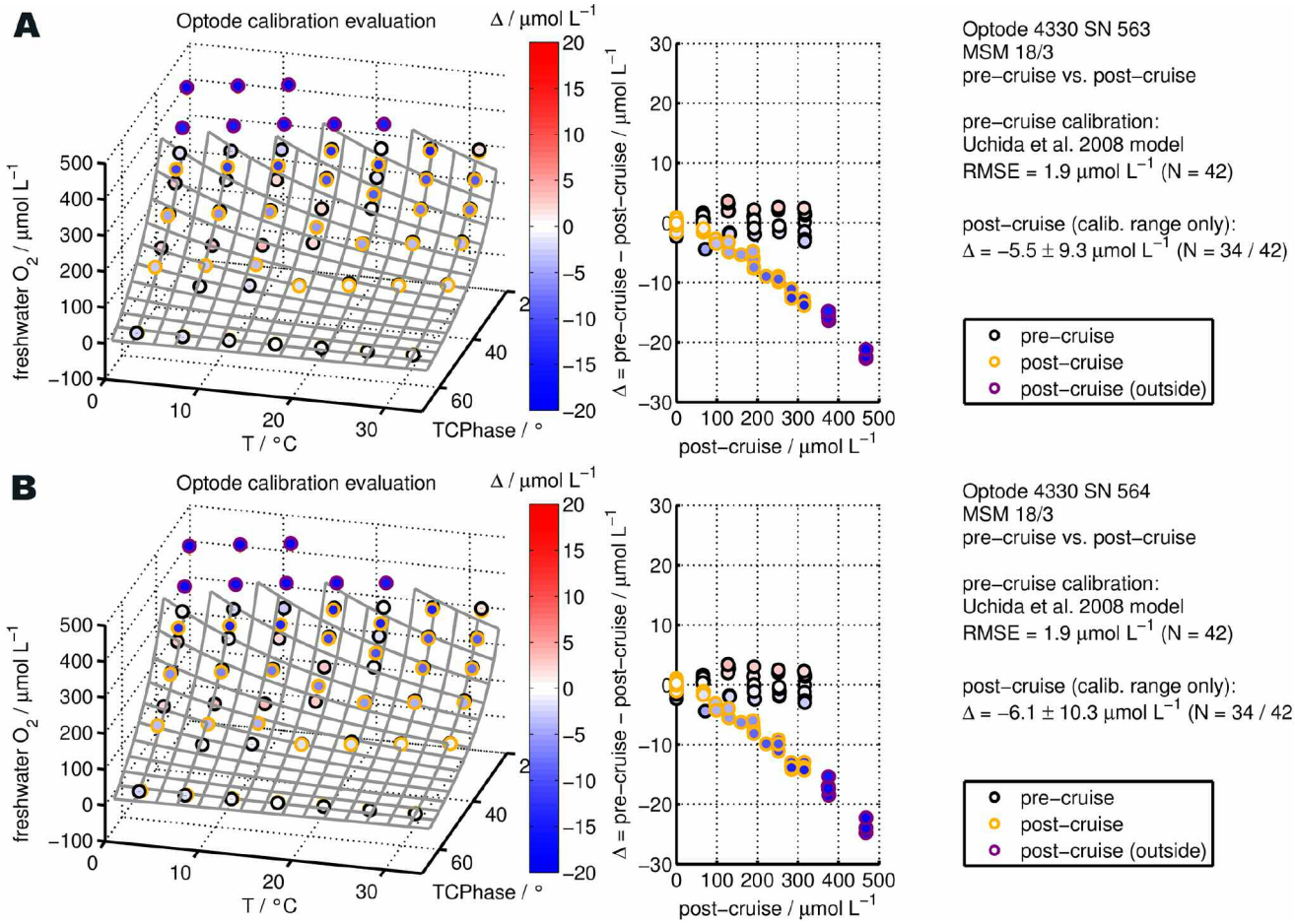

Fig. 11. Repeated calibration of optodes before and after the R/V *Maria S. Merian* cruise MSM 18/3. Left panel: Initial calibration points (black circle) and fitted optode response function (gray) with repeated calibration samples inside (yellow circle) and outside (purple circle) the calibrated range. Middle panel: Difference between pre-cruise calibrated optode reading and post-cruise Winkler samples. The color axis shows the same difference in both panels. Right panel: Statistical figures for both the initial and repeated calibration with 95% confidence interval.

During the cruise, the second optode was attached to the CTD at 13 stations with 282 Winkler samples available as reference. However, no separate bottle stops were performed during these casts and the Niskin bottles were fired in drive-by mode. The results are shown in Fig. 13 and the data distinguished between surface samples above or close to the thermocline (green circles) and deeper samples below 100 dbar (yellow circles) in analogy to the R/V *Polarstern* cruise.

In contrast to the R/V *Polarstern* cruise, the field data are spread both on a broad temperature and oxygen range and fall well within the calibration range of both the pre-cruise (not shown) and post-cruise calibration (left panel). However, there is a significantly higher scatter of the residuals (middle panel), which can be attributed to both the larger oxygen gradient and the sensor's dynamic response, the effect of which are amplified by the drive-by bottle fires. The calibration bias of –0.8 µmol L^{-1} for the deep samples and –3.2 µmol L^{-1} for all samples, respectively, is within or only slightly exceeds the laboratory calibration accuracy of 1.0 µmol L^{-1} and is well inside the field uncertainty. Additionally, there is no visible trend in the difference between optode reading and Winkler field samples, i.e., both the nonlinear temperature and oxygen behavior of the optode has been grasped by the laboratory calibration.

Discussion and summary

The flow-system based calibration setup with electrochemical O_2 generation proves to be well-suited for the individual multi-point calibration of oxygen sensors. Whereas the O_2 generator forces the oxygen content of the carrier solution, its flow rate needs to be constrained tightly in order to provide stable O_2 concentrations. By these means, different oxygen concentrations up to 315 µmol L^{-1} can be obtained at a high stability of within ± 0.5 µmol L^{-1} (2 SD). On the other hand, the temperatures of both the carrier solution and the sensors are thoroughly controlled as a prerequisite for reliable reference points for the sensor calibration.

Whereas triplicate Winkler samples with a typical reproducibility of 0.8 µmol L^{-1} (2 SD) can be taken for external ref-

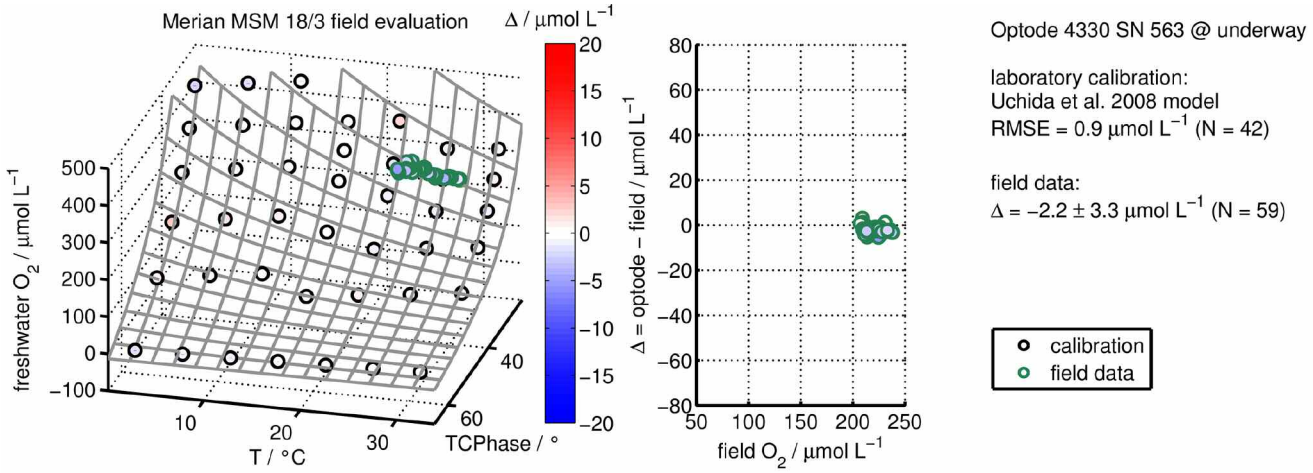

Fig. 12. Underway field evaluation of optode 4330 SN 563 during cruise MSM 18/3 with post-cruise calibration. Left panel: Calibration points (black) and optode response as function of phase and temperature (gray) with field samples (green) mapped into the same, freshwater sample space of the calibration. Middle panel: Difference between optode calibration and Winkler field samples. The color axis shows the same difference in both panels. Right panel: Statistical figures for both the calibration and the field samples.

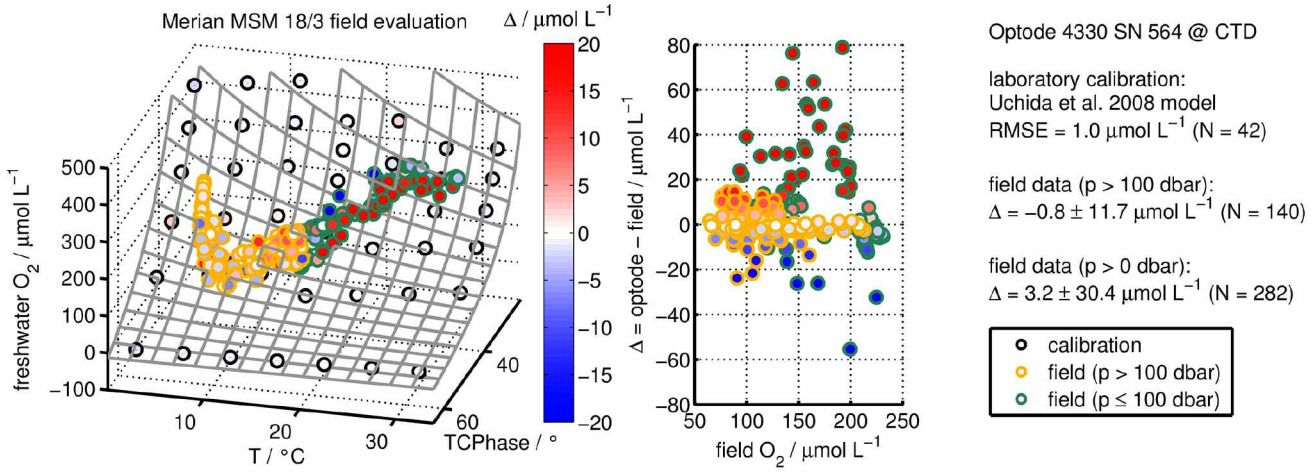

Fig. 13. CTD field evaluation of optode 4330 SN 564 during cruise MSM 18/3 with post-cruise calibration. Left panel: Calibration points (black) and optode response as function of phase and temperature (gray) with field samples below the thermocline (yellow) and above the thermocline (green) mapped into the same, freshwater sample space of the calibration. Middle panel: Difference between optode calibration and Winkler field samples for samples below the thermocline (yellow) and above the thermocline (green). The color axis shows the same difference in both panels. Right panel: Statistical figures for both the calibration and the field samples.

erence, the system may provide a Winkler-free way to calibrate oxygen sensors. As the environmental conditions can be controlled very accurately, the electrolytic current and carrier solution flow rate define the oxygen concentration to within an accuracy of ± 1.2 µmol L^{-1}. This illustrates the high repeatability of the system, albeit an incomplete degassing that causes a remaining offset of 4.7 µmol L^{-1} has to be taken into account. These figures have been obtained from the direct comparison of triplicate Winkler samples with the generator settings and are valid for the entire operation range, i.e., 1°C–36°C and 0 µmol L^{-1}–315 µmol L^{-1}, respectively.

However, the proper and accurate calibration of oxygen sensors is the main purpose. A good agreement has been found between the individual multi-point laboratory calibration of Aanderaa oxygen optodes and Winkler samples under various field conditions, both polar and tropical, and deployment modes, both profiling and underway. Whereas the polar deployments suffered from an imperfect match between the parameter range in calibration and field measurements for temperature and oxygen, the mismatch did not exceed –6.4 µmol L^{-1} for samples below the thermocline. Moreover, it is a constant offset to the otherwise well-grasped sensor's oxygen response, as indicated by the low scatter of ≤ 3.5 µmol L^{-1} (2 SD), and points toward issues in the fitting equations for the optode response at low temperatures. For the tropical deployments, all field samples are within the calibrated range.

There is a calibration bias of only –0.8 µmol L^{-1} for samples below the thermocline and both the temperature and the oxygen behavior of the optode are properly characterized by the laboratory calibration. However, a drift of the optodes between the pre- and post-cruise calibration has been observed for the tropical deployment, and only the post-cruise calibration has been used for the evaluation.

The laboratory calibrations showed RMSE values as measure of accuracy between 0.9 µmol L^{-1} and 1.9 µmol L^{-1} when combined with the Uchida et al. (2008) functional model of the optode's oxygen response. At the same time, the repeated calibrations with varying calibrated ranges indicate a good parameterization of the oxygen slope in the Uchida et al. (2008) model when being extrapolated (Fig. 11), whereas the temperature slope parameterization might have room for improvement (Fig. 8 and polar deployments mismatch).

During the course of the R/V *Maria S. Merian* field evaluation, the appeal of simple means for a repeated calibration has become obvious. From the pre- and post-cruise calibration, the otherwise elusive sensor drift is clearly identified. Thus, the interpretation of the field data can be based unambiguously on the more adequate calibration parameters.

While the observed sensor drift discredits the overall long-term stability of optical sensors, the two 4330 optodes changed their oxygen response in a very similar and distinct manner. Moreover, the post-cruise calibration fits well to the field data with several months in between. In fact, the time between deployment and post-cruise calibration is twice as long as the time between pre-cruise calibration and deployment, where most of the response change appeared. In consequence, optodes still represent the most stable oxygen sensors with a possibly noncontinuous drift related to its usage. In any case, the drift is not erratic and may be detected and corrected for.

The calibration setup presented here has the potential to make oxygen sensor calibrations less time and skill demanding and, more importantly, regular recalibrations feasible. Frequent snapshots of a sensor's oxygen response will be a crucial step towards an understanding of sensor drift between deployments and conditions that enhance or reduce this drift.

It should be noted that the laboratory calibration cannot be done in pure freshwater as the electrolytic medium by necessity contains ions, but its low salinity effect can be compensated for (Clegg and Brimblecombe 1990). The high pH, however, may not be a suitable environment for every kind of sensor and sensing material. Still, the calibration setup is not specific for a special sensor type, but any model that is compatible with high pH conditions can be used with a custom flow-through cell. The systems size does not exceed common bench-top instrumentation and, more importantly, it does not need separate gas cylinders or similar, difficult to handle equipment or consumables. The oxygen content is solely dependent on the conditions given by the setup and is independent of ambient humidity or atmospheric pressure, which are easily influenced by air conditioning in laboratories.

In addition, the calibration setup does not necessarily depend on external referencing, but offers a Winkler-free mode of operation. It is small and robust enough as to build a mobile, sea-going, and Winkler-free calibration setup for oxygen sensors. Moreover, the calibrations obtained by this laboratory setup proved to be valid under various field conditions and underline the versatility of the calibration setup. It thus represents a system capable to facilitate high accuracy automated dissolved oxygen measurements on a large scale by providing reliable and easy access to accurate individual multipoint sensor calibrations.

Comments and recommendations

The maximum electrolytic current of 20 mA provided by the oxygen generator proved to be insufficient as freshwater and saltwater (35 psu) oxygen saturation levels cannot be reached below 15°C and 5°C, respectively. Thus, a separate current source providing up to 30 mA was developed. This equals a concentration of 465 µmol L^{-1} or 105% and 133% saturation at 1°C in fresh- and saltwater, respectively, and should be adequate for most oceanographic purposes. Furthermore, the valves shown in Fig. 1 were replaced by electric isolation valves (100T3, Bio-Chem Fluidics), whereas the temperature, electrolytic current and flow rate regulation was integrated into the same LabVIEW routine as the sensor data logging. All this was done to eliminate sources of variability and to further improve the repeatability. These improvements were already in place for the R/V *Maria S. Merian* MSM 18/3 post-cruise calibration.

For the calibration setup described here, the total equipment costs amount to ca. 22000 Euro, while the running costs are basically the trained staff and consumables to measure the Winkler samples if desired.

Nomenclature

E_0	standard reduction potential / V
V	volume / mL
t	time / s
Q	charge / C
n	amount of substance / mol
z	number of electrons transferred, stochiometric factor
F	Faraday constant: 96485 C mol^{-1}
I	electric current / mA
$c(O_2)$	concentration of oxygen / µmol L^{-1}
pO_2	partial pressure of oxygen / Pa
$\alpha(O_2)$	Henry's law oxygen solubility constant, Bunsen coefficient / µmol L^{-1} Pa^{-1}
τ	fluorophore excited state lifetime in the presence of O_2 / s
τ_0	fluorophore excited state lifetime in the absence of O_2 / s

P	phase signal in the presence of O_2 / °
P_0	phase signal in the absence of O_2 (zero phase signal) / °
K_{SV}	Stern-Volmer constant of the fluorophore
$c_0...c_6$	calibration coefficients of the Uchida et al. (2008) model
p	hydrostatic pressure / dbar
AMT	AMT Analysenmesstechnik GmbH, Rostock/Germany
SD	standard deviation
RMSE	root-mean-square error

References

Clegg, S. L., and P. Brimblecombe. 1990. The solubility and activity coefficient of oxygen in salt solutions and brines. Geochim. Cosmichim. Acta 54:3315-3328 [doi:10.1016/0016-7037(90)90287-U].

Dickson, A. G. 1995. Determination of dissolved oxygen in sea water by Winkler titration. *In* WOCE operations manual, Part 3.1.3 operations & methods. WHP Office Report WHPO 91-1.

Emerson, S., C. Stump, and D. Nicholson. 2008. Net biological oxygen production in the ocean: Remote in situ measurements of O_2 and N_2 in surface waters. Global Biogeochem. Cycles 22:GB3023 [doi:10.1029/2007GB003095].

Garcia, 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].

Gruber, N., and others. 2010. Adding oxygen to Argo: Developing a global in-situ observatory for ocean deoxygenation and biogeochemistry. *In* J. Hall, D. E. Harrison, and D. Stammer [eds.], Proceedings of OceanObs'09: Sustained ocean observations and information for society. ESA Publication WPP-306, Venice, Italy, 21-25 Sept 2009, vol. 2.

Juranek, L. W., R. C. Hamme, J. Kaiser, R. Wanninkhof, and P. D. Quay. 2010. Evidence of O_2 consumption in underway seawater lines: Implications for air-sea O_2 and CO_2 fluxes. Geophys. Res. Lett. 37:L01601 [doi:10.1029/2009GL040423].

Karl, D. M., and R. Lukas. 1996. The Hawaii Ocean Time-series (HOT) program: Background, rationale and field implementation. Deep Sea Res. II 43:129-156 [doi:10.1016/0967-0645(96)00005-7].

Keeling, R. F., A. Körtzinger, and N. Gruber. 2010. Ocean deoxygenation in a warming world. Ann. Rev. Mar. Sci. 2:199-229 [doi:10.1146/annurev.marine.010908.163855].

Körtzinger, A., J. Schimanski, U. Send, and D. Wallace. 2004. The ocean takes a deep breath. Science 306:1337 [doi:10.1126/science.1102557].

———, ———, and ———. 2005. High quality oxygen measurements from profiling floats: A promising new technique. J. Atmos. Oceanic Technol. 22:302-308 [doi:10.1175/JTECH1701.1].

Robinson, C., and P. J. le B. Williams. 2005. Respiration and its measurement in surface marine waters, p. 147–180. *In* P. del Giorgio and P. J. le B. Williams [eds.] Respiration in aquatic ecosystems. Oxford Univ. Press [doi:10.1093/acprof:oso/9780198527084.003.0009].

Steinberg, D. K., C. A. Carlson, N. R. Bates, R. J. Johnson, A. F. Michaels, and A. H. Knap. 2001. Overview of the US JGOFS Bermuda Atlantic Time-series Study (BATS): a decade-scale look at ocean biology and biogeochemistry. Deep Sea Res. II 48:1405-1447 [doi:10.1016/S0967-0645(00)00148-X].

Stendardo, I., N. Gruber, and A. Körtzinger. 2009. CARINA oxygen data in the Atlantic Ocean. Earth Syst. Sci. Data 1:87-100 [doi:10.5194/essd-1-87-2009].

Takeshita, Y., T. R. Martz, K. S. Johnson, J. Plant, S. Riser, and D. Gilbert. 2010. Quality control and application of oxygen data from profiling floats. AGU Fall Meeting Abstracts.

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. Oceanic Technol. 25:2271-2281 [doi:10.1175/2008JTECHO549.1].

Winkler, L. W. 1888. Die Bestimmung des im Wasser gelösten Sauerstoffes. Ber. Dtsch. Chem. Ges. 21:2843-2854 [doi:10.1002/cber.188802102122].

Submitted 18 January 2012
Revised 16 August 2012
Accepted 20 September 2012