1 **Offset between profiling float and shipboard oxygen observations at depth imparts bias on**

2 **float pH and derived $p\text{CO}_2$**

3

4 Seth M. Bushinsky[1], Zachary Nachod[1], Andrea J. Fassbender[2], Veronica Tamsitt[3], Yuichiro

5 Takeshita[4], Nancy Williams

6

7 [1]Department of Oceanography, School of Earth and Space Science and Technology, University of

8 Hawai'i at Mānoa, Honolulu, HI, USA

9 [2]NOAA/OAR Pacific Marine Environmental Laboratory, Seattle, WA, USA

10 [3]University of South Florida, St. Petersburg, FL, USA

11 [4] Monterey Bay Aquarium Research Institute, Moss Landing, CA, USA

12

13

14 **Abstract**

15   Profiles of oxygen measurements from Argo profiling floats now vastly outnumber

16 shipboard profiles. Air calibration of a float's oxygen optode upon surfacing enables accurate

17 measurements in the upper ocean but does not necessarily provide similar accuracy at depth. In

18 this study we use a quality controlled shipboard dataset to show that, on average, the entire Argo

19 oxygen dataset is offset relative to shipboard measurements (float minus ship) at pressures of 1450

20 to 2000 db by -1.9 ± 4.7 µmol kg^{-1} (95% confidence interval around the mean of {-2.3, -1.5}) and

21 air calibrated floats are offset by -3.1 ± 5.3 µmol kg^{-1} (95% CI: {-3.7, -2.4}). The difference

22 between float and shipboard oxygen is most likely due to offsets in the float oxygen data and not

23 due to oxygen changes at these depths or biases in the shipboard dataset. In addition to posing

24 problems for the calculation of long-term ocean oxygen changes, these float oxygen offsets impact

25 the adjustment of float nitrate and pH measurements and therefore bias important derived

26 quantities such as the partial pressure of CO_2 ($p\text{CO}_2$) and dissolved inorganic carbon. Correcting

27 floats with air-calibrated oxygen for the float-ship oxygen offsets changes float pH by 3.2 ± 3.8

28 mpH and float-derived surface $p\text{CO}_2$ by -3.3 ± 4.1 µatm. This adjustment to float $p\text{CO}_2$ represents

29 half, or more, of the bias in float-derived $p\text{CO}_2$ reported in studies comparing float $p\text{CO}_2$ to

30 shipboard $p\text{CO}_2$ measurements.

31

32 **Plain Language Summary**

33   Oxygen has historically been measured using chemical titrations on water collected by

34 ships at sea. Over the past 20 years, sensors that measure oxygen have been deployed on robotic

35 profiling floats. Measurements by oxygen sensors on profiling floats now greatly exceed those

36 collected by ships. Here we compare all float oxygen data collected to shipboard measurements in

37 deep waters (1450 to 2000 m depth) where we do not expect oxygen to be changing in the ocean.

38 We find a difference between float and shipboard data. If left uncorrected, this difference would

39 give the false impression of a long-term oxygen change. This difference also impacts float-

40 measured pH and float-estimated carbon dioxide, both of which rely on float oxygen

41 measurements. Correcting oxygen, and therefore float pH and carbon dioxide, would largely

42 address a widely-studied bias in float measurements of these parameters.

43

44 **Key Points**

45   1. Air-calibrated float oxygen measurements are lower than shipboard data by 3.1 μmol kg^{-1}

46   at pressures of 1450 to 2000 db.

47   2. Correcting float oxygen for this offset would increase float pH by 3.2 mpH and lower float-

48   derived pCO$_2$ by -3.3 μatm.

49   3. This float oxygen offset would improve float and ship pCO$_2$ comparisons, removing most

50   or all of observed biases.

51

52

1. Introduction

Measurements of oxygen in the ocean are critical, as the ocean is losing oxygen due to warming, respiration, and stratification-induced changes to ventilation (Helm et al., 2011; Levin, 2018; Schmidtko et al., 2017; Stramma & Schmidtko, 2021). Measurements of ocean oxygen also can be used to understand the balance between photosynthesis and respiration in the surface and deep ocean (Bushinsky & Emerson, 2015; Emerson, 1987; Hennon et al., 2016; Yang et al., 2017). Historically, oxygen was primarily measured by Winkler titrations of water sampled from the ocean (Carpenter, 1965; Dickson, 1994). These titrations require reliable standards and trained operators to acheive high accuracy and precision. Winkler titrations have been supplemented by electrochemical and optical sensors attached to CTDs but have remained the gold standard of accurate oxygen measurement in the ocean.

Oxygen measurements on autonomous profiling floats are transforming our ability to observe the ocean at unprecedented levels of detail. The first Argo oxygen data was from floats deployed in the early 2000s equipped with Clark-cell electrodes (Figure 1) that provided fast response times but could drift rapidly and in unpredictable ways (Gruber et al., 2007). Oxygen optodes that measure the partial pressure of oxygen through an oxygen-sensitive luminophore have now become the standard oxygen sensor deployed on floats (Claustre et al., 2020; Gruber et al., 2009; Körtzinger et al., 2005; Tengberg et al., 2006) (Figure 1). Over 1800 Argo oxygen floats have been deployed as of 2023, with increasing numbers due to a mix of many small deployments by individual research groups and the development of a few projects deploying large numbers of floats (e.g., the Southern Ocean Carbon and Climate Observations and Modeling project, SOCCOM, (Johnson et al., 2017; Sarmiento et al., 2023); Global Ocean Biogeochemistry Array, GO-BGC (Roemmich et al., 2021; Schofield, O. A. et al., 2022)). Optodes drift significantly prior to deployment, but the advent of calibration using atmospheric oxygen has increased accuracy of surface data to better than 1% (Bittig, Körtzinger, et al., 2018; Bushinsky et al., 2016; D'Asaro & McNeil, 2013; Johnson et al., 2015). Floats calibrated using atmospheric data primarily use a single multiplicative gain correction based on the difference between float air measurements and calculated atmospheric oxygen that is then applied to the whole float profile (Figure 2). This approach assumes that the response of the oxygen optode at different oxygen levels and temperatures changes uniformly over time, and that the zero reading of the optode remains unchanged.
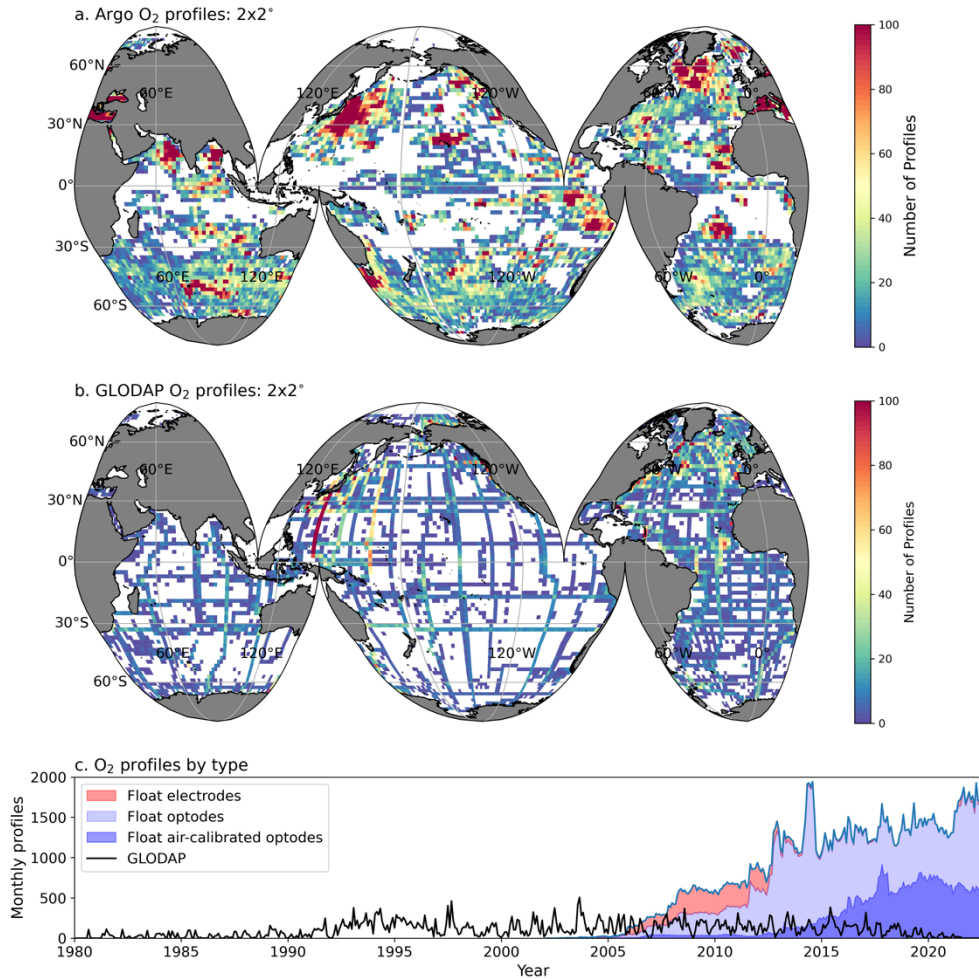
84

Figure 1. Observational density of float and ship oxygen and change in float sensor type over time. Float (a.) and GLODAP (b.) oxygen profile density on a $2 \times 2°$ grid. (c.) Total number of oxygen profiles per month, with the contribution of float sensor type shown (Clark cell electrodes (red) and optodes (blue), and the subset of optodes calibrated with in-air measurements (dark blue)), and total number of GLODAP profiles that contain oxygen per month (black line).

85   The Global Ocean Data Analysis Project (GLODAP; Key et al., 2015; Olsen et al., 2016,

86   2020) is an on-going international synthesis effort that evaluates surface-to-bottom ocean

87   biogeochemical bottle data for outliers and internal consistency, which has been crucial to

88   improving the accuracy and usability of shipboard data. A similar effort, the Surface Ocean CO_2

89   Atlas (SOCAT) collects and assesses surface pCO$_2$ measurements for accuracy prior to inclusion

90   in an annual data product (Bakker et al., 2016). These efforts have been instrumental in our

91   understanding of ocean biogeochemistry.   There is currently no comparable, on-going, post-

a.

O₂ in situ correction approaches:

Surface

1500 m

Raw

O₂ surface air adjustment

O₂ adjustment to multiple Winklers / depths

Raw

Gain only

Gain + offset

b.

NO₃⁻ and pH adjustment to deep estimates:

Surface

1500 m

Raw NO₃⁻

Offset adjustment

Raw pH

Estimates

Measured
Estimated
Adjusted to estimate
Derived

c.

$NO_3^-{}_{meas}$

$NO_3^-{}_{estimate} = f(T, S, O_2, Lat., Lon., press.)$

NO_3^-

$TA = f(Lat, Lon., press. T, S, O_2)$

pH_{meas}

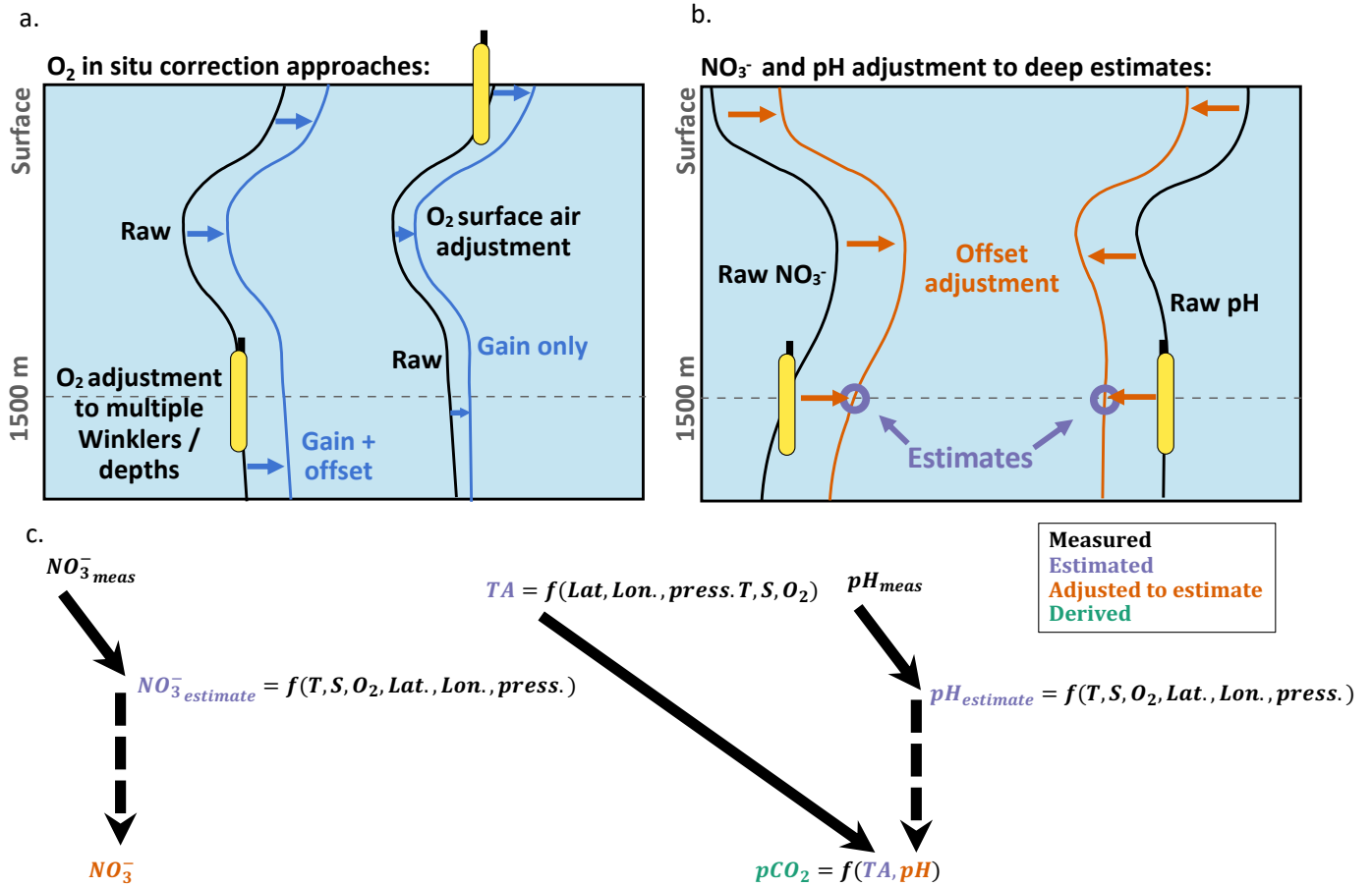$pH_{estimate} = f(T, S, O_2, Lat., Lon., press.)$

$pCO_2 = f(TA, pH)$

Figure 2. Schematic of the two primary in-situ oxygen calibration approaches and how the calibrated oxygen data are used to adjust float pH and nitrate measurements. (a.) Oxygen data collected on floats are typically calibrated using a gain and offset correction based on in-situ data, typically collected at float deployment, or a gain-only correction based on air-calibration measurements made upon float surfacing throughout its lifetime. (b.) Float nitrate and pH measurements are adjusted to algorithmic estimates of those parameters at 1500 db in a way that impacts the whole profile. (c.) Variable inputs, including oxygen, to the nitrate and pH algorithms, which are trained on shipboard data. Offsets between the float and ship oxygen will propagate through these algorithms into the adjusted nitrate and pH data.

92   deployment census of float biogeochemical data to assess consistency within the float dataset or

93   between float and shipboard data. Most studies that have used Argo oxygen data and assessed their

94   accuracy have focused on surface data or analyzed changes measured by individual floats, thereby

95   avoiding comparisons between different float measurements (Bushinsky & Emerson, 2015;

96   Johnson et al., 2015; Martz et al., 2008; Wolf et al., 2018). As the float array expands and

97   researchers begin using the entire dataset as a whole, or combining float and shipboard datasets to

98   create global interior ocean oxygen products (e.g., World Ocean Atlas (Garcia et al., 2010);

99   Gridded Ocean Biogeochemistry from Artificial Intelligence – Oxygen (Sharp et al., 2022)),

100 attention must be paid to the accuracy of float oxygen data at depth.

101   Largescale comparisons of float and ship datasets indicate that, to a first order, float oxygen

102 data have reached a comparable level of accuracy as ship-board measurements (Bushinsky et al.,

103 2017; Johnson et al., 2017; Maurer et al., 2021). However, two float-to-ship comparisons of deep

104 (1500-2000 m) oxygen data (Bushinsky et al., 2016; Drucker & Riser, 2016) indicate that surface

105 calibration may not be a sufficient adjustment for the entire float oxygen profile. There is some

106 indication that this could be due to inadequate calibration of the optode temperature response

107 (Bittig, Körtzinger, et al., 2018), but Bushinsky et al. (2016) re-calibrated the optode temperature

108 response for 11 floats in the lab prior to deployment, which did not address the difference in deep

109 data, so this approach does not seem to fully resolve the observed drift.

110   The accuracy of oxygen data throughout the water column is critical for understanding

111 long-term changes in ocean oxygen content. At depth, the accuracy of quality-controlled float

112 oxygen data is especially critical due to its use in adjusting other biogeochemical sensor data from

113 Argo floats, which has downstream impacts on derived carbonate system parameters. Nitrate and

114 pH sensors are now being deployed in large numbers (Claustre et al., 2020) and also require in-

115 situ adjustment to correct sensor drift (Maurer et al., 2021). The nitrate and pH measurements are

116 adjusted using estimates of these parameters at 1500 db that are derived from multiple linear

117 regression or neural network algorithms that were trained on the GLODAP shipboard database

118 (Bittig, Steinhoff, et al., 2018; Carter et al., 2016, 2021; Williams et al., 2016). These algorithmic

119 estimates use inputs of temperature, salinity, depth, latitude, longitude, and, importantly, oxygen.

120 Due to sensor response characteristics, adjustments to the nitrate and pH sensor data are applied

121 almost uniformly to the entire profile, so offsets at 1500 m directly impact surface measurements

122 (Maurer et al., 2021).

123   Float measurements of pH are widely used to estimate pCO_2, and other carbonate system

124 parameters, using an algorithm-based estimate of total alkalinity (TA) and a carbonate system

125 calculator. Many studies apply an additional adjustment to the quality controlled pH data prior to

126 estimating pCO_2 that is meant to correct for an empirical pH-dependent pH bias (Carter et al.,

127 2013, 2018; Williams et al., 2017). The accuracy of float pCO_2 estimates is critical, as studies

128 relying on float pCO_2 have identified significant differences between seasonal cycles of pCO_2, and

129 wintertime air-sea CO_2 fluxes from studies that rely on shipboard pCO$_2$ alone (e.g. Bushinsky et

130 al., 2019; Gray et al., 2018; Williams et al., 2016). Floats measure year-round, including during

131 winter months when rough weather makes shipboard observations rare, which provides immense

132 observational value, if accurate. A number of studies have directly compared float pCO$_2$ estimates

133 to shipboard observations (Bushinsky & Cerovečki, 2023; Coggins et al., 2023; Fay et al., 2018;

134 Gray et al., 2018; Williams et al., 2017), while other studies have indirectly compared float pCO$_2$

135 and pH through assessment of other carbonate system parameters or CO_2 fluxes (e.g., Long et al.,

136 2021; Mackay & Watson, 2021; Wu et al., 2022). A recent meta-analysis of float pCO$_2$ accuracy

137 found biases are likely between 6-9 µatm (float pCO$_2$ high), with direct comparisons yielding

138 lower biases (2-5 µatm) than indirect comparison (Wu & Qi, 2023).

139   Here we use crossover comparisons between ship and float data to quantify differences in

140 deep oxygen values and calculate the impact those offsets would have on adjusted and derived

141 parameters. We refer to oxygen "offset", without strict attribution of the source of the error. Given

142 the long-history and extensive use of Winkler titrations and the relatively short time-history of

143 oxygen sensors and float-mounted oxygen sensors, we assume that the GLODAP values are likely

144 correct and will present evidence later that supports this assumption.

145

146

147 **2. Methods**

148 *2.1 Float and shipboard datasets*

149   Biogeochemical float data were downloaded on June 20, 2023 from the Argo Data

150 Assembly Centers (DACs, closest snapshot: 2023-06-09 (Argo, 2023)) according to the list of

151 floats with oxygen, nitrate, or pH sensors in the Argo Global DACs Synthetic-Profile Index file.

152 1,839 Argo floats deployed between 2002 and 2023 were included in this dataset. Of these floats,

153 1,830 measured oxygen, of which 151,880 out of 270,654 profiles (56%) contained valid delayed

154 mode "ADJUSTED" oxygen data, indicating the data had been checked and/or corrected post

155 deployment (median difference between uncorrected and adjusted data was 11 µmol kg^{-1}, adjusted

156 data higher). 582 floats measured nitrate, with 43,280 delayed mode, adjusted profiles out of

157 67,505 total (64%). 492 floats measured pH, of which 19,137 profiles had delayed mode, adjusted

158 data out of 44,407 total (43%). Only delayed mode, adjusted float data flagged as "good" were

159 used for this analysis. Float oxygen sensor calibration type for floats with valid crossovers was

160   determined by reading the "SCIENTIFIC_CALIB_COMMENT" field (198 unique comments,

161   Table S1) in float "Sprof" files and were categorized as: air, non-air, and bad/no calibration.

162   Shipboard bottle measurements of salinity, temperature, oxygen, nitrate, total CO_2 (DIC),

163   or pH flagged as "good" from the GLODAP v2.2022 (Olsen et al., 2020) were used for comparison

164   with float observations. The GLODAP dataset was chosen because it includes a secondary quality

165   control and adjustment to shipboard data for overall accuracy and internal consistency. Shipboard

166   pH measurements in the GLODAP dataset include a variety of measurement approaches. Carter et

167   al. 2021 and 2018 "homogenized" the GLODAP pH dataset to be consistent with

168   spectrophotometric pH measurements using purified meta-cresol purple (Liu et al., 2011) prior to

169   training of the LIPHR (Locally Interpolated PH Regression)/ESPER (Empirical Seawater Property

170   Estimation Routine) algorithms. To recreate this homogenized dataset, we used ESPER to

171   calculate GLODAP pH on the total scale normalized to 25°C (pH_{25-T}) for every datapoint where a

172   "good" pH measurement was present. By using the algorithms at the same datapoints on which

173   they were trained, we recreated the homogenized dataset and used this data for the pH crossover

174   comparison.

175

176   *2.2 Crossover comparisons*

177   Criteria for crossover comparisons between float and shipboard measurements were

178   established using distance, pressure, potential density, and spice. For each float profile, we first

179   found all GLODAP bottle measurements within a 100 km radius. Float data from 1450 to 2000 db

180   were interpolated to a 1 db grid, with any gaps of over 125 db between successive float

181   measurements removed from the interpolated profiles. Potential density (σ_θ) and spiciness (τ)

182   relative to 0 db were calculated for both the float and shipboard data using the Gibbs SeaWater

183   Oceanographic Toolbox of TEOS-10 for Python (McDougall & Barker, 2011). For each GLODAP

184   bottle measurement between 1400 and 2100 db within the 100 km distance range, crossovers were

185   determined by looking for interpolated float data with differences of less than $\pm$ 0.005 kg m^{-3} σ_θ,

186     $\pm$ 0.005 τ, and $\pm$ 100 db from the GLODAP sample properties. These difference thresholds were

187   selected by analyzing levels of environmental oxygen noise in comparisons of individual floats

188   against themselves using a range of density, spiciness, and distance thresholds (Text S1, Figure

189   S1). Crossovers from any point in time were allowed. Mean property offsets (e.g., ΔC_{off}, for a

190   given property "C") were calculated for floats with at least 20 oxygen crossovers present and used

191   in the results shown here (Figure S2 for float crossover examples). On average, offsets in the 1450

192   – 2000 db range did not differ significantly as a function of depth or concentration (Figure S3).

193   While adjusting the filter criteria does impact the number of crossovers found for each float and

194   can impact the mean offset calculated for an individual float, the overall results presented in this

195   manuscript are relatively insensitive to the exact criteria used.

196

197   *2.3 Impact of oxygen offsets on derived parameters*

198   The impact of oxygen offsets (e.g., ΔC_{imp}, for a given property C) on float nitrate and pH

199   adjustments and float estimated pCO_2 and DIC ($\Delta NO_3^-{}_{imp}$, $\Delta pH_{25\text{-}T,imp}$, $\Delta pCO_{2,imp}$, and ΔDIC_{imp},

200   respectively) was determined for each float with valid oxygen and nitrate, or pH data and with a

201   valid GLODAP crossover comparison. As previously described, oxygen offsets impact estimated

202   pCO_2 and DIC through the effect of oxygen on the pH adjustment at 1500 m (Figure 2). While

203   oxygen is also used in the algorithmic estimation of total alkalinity, which is required for pCO_2

204   and DIC estimation, only a surface oxygen offset would impact the surface total alkalinity

205   estimate, and subsequently the surface pCO_2 and DIC estimates, which are the foci of this work.

206   Using temperature, salinity, and oxygen data at 1500 m from each profile as inputs to the

207   calibration algorithms, we calculated pH and nitrate, both with and without adjusting for the mean

208   float oxygen offset relative to GLODAP ($\Delta O_{2,off}$). We then calculated the differences in (impacts

209   on) pH and nitrate with and without the $\Delta O_{2,off}$ correction ($\Delta pH_{25\text{-}T,imp}$ and $\Delta NO_3^-{}_{imp}$) and applied

210   these differences to the adjusted, full float profiles. This approach allows us to determine the

211   impact a mean oxygen correction would have without attempting to replicate any step changes or

212   non-linear adjustments that may have been applied during the original data adjustment procedure

213   (Maurer et al., 2021). For nitrate, a uniform adjustment is applied to the entire nitrate profile. For

214   pH, the adjustment at 1500 m is scaled relative to the difference in temperature between each depth

215   and 1500 m, following the protocol used in the original pH measurement adjustment (Johnson et

216   al., 2023). $\Delta pCO_{2,imp}$ and ΔDIC_{imp} were calculated using CO2SYS first with the original pH as an

217   input and then with mean $\Delta pH_{25\text{-}T,imp}$ applied to the float profile and calculating the difference.

218   Floats equipped with pH and nitrate sensors have been deployed by many groups

219   throughout the world. The majority have been deployed as part of SOCCOM or GO-BGC and

220   adjusted by the data management teams of SOCCOM and the Monterey Bay Aquarium Research

221   Institute (MBARI). Following adjustment approaches used by SOCCOM/MBARI, total alkalinity

222   was calculated using LIARv2, pH using ESPER (ESPER-mixed, an average of a neural network-

223   based approach and an MLR; Carter et al., 2021) or LIPHR (Carter et al., 2018), and carbonate

224   system calculations using PyCO2SYS (v1.8.1; Humphreys et al., 2022, 2023). Both ESPER and

225   LIPHR were used because SOCCOM pH data at the time of download have been adjusted to one

226   of the two algorithms, depending on when the float was last active. Floats active prior to April

227   2023 were adjusted to LIPHR pH estimates, while floats active past this date are adjusted to

228   ESPER pH. Figures and results shown in the main text rely on ESPER-based adjustments while

229   the supplement includes results that rely on LIPHR-based adjustments, but average differences

230   between the two are of a second order.

231   One complication is the presence of a pH-dependent pH correction (Williams et al., 2017)

232   used by the SOCCOM/GO-BGC groups when calculating pCO$_2$. This correction accounts for the

233   difference between float ISFET measured pH, which has been shown to align with

234   spectrophotometric pH measurements (Takeshita et al., 2020), and pH values calculated from

235   measurements of TA and DIC (Carter et al., 2013; 2018). It is currently unclear how to best apply

236   a global correction similar to the one developed in Williams et al. (2017) and a recently published

237   paper by a working group focused on inter-consistency in carbonate system measurements

238   recommended removing this correction until a suitable correction for all float pH measurements

239   can be developed (Carter et al., 2023). We have avoided dealing with this issue by looking at the

240   difference between calculations with and without the oxygen offset impacts included. Any changes

241   in this pH-dependent bias correction will represent a different, additional impact to float pCO$_2$

242   estimates. However, the impacts of the oxygen offset should be very similar for estimated pCO$_2$

243   with or without this additional pH-dependent bias correction.

244

245

246   **3. Results**

247   For floats where oxygen offsets could be determined, 93% of the offsets were statistically

248   significant using a 1 sample t-test and a p-value threshold of 0.01 (Figures S2A and S2B for

249   examples with significant and non-significant crossovers). Our goal is to quantify the potential

250   impact of float oxygen offsets on other parameters, so we include all oxygen offset estimates,

251   including insignificant ones, when calculating the mean oxygen offset and all floats with pH in the

252   subsequent impact on derived parameters. Only including floats with a statistically significant

253   oxygen offset would overstate the magnitude of the mean dataset offset. However, if these
254   crossover comparisons were used in the future to correct float oxygen data it would be important
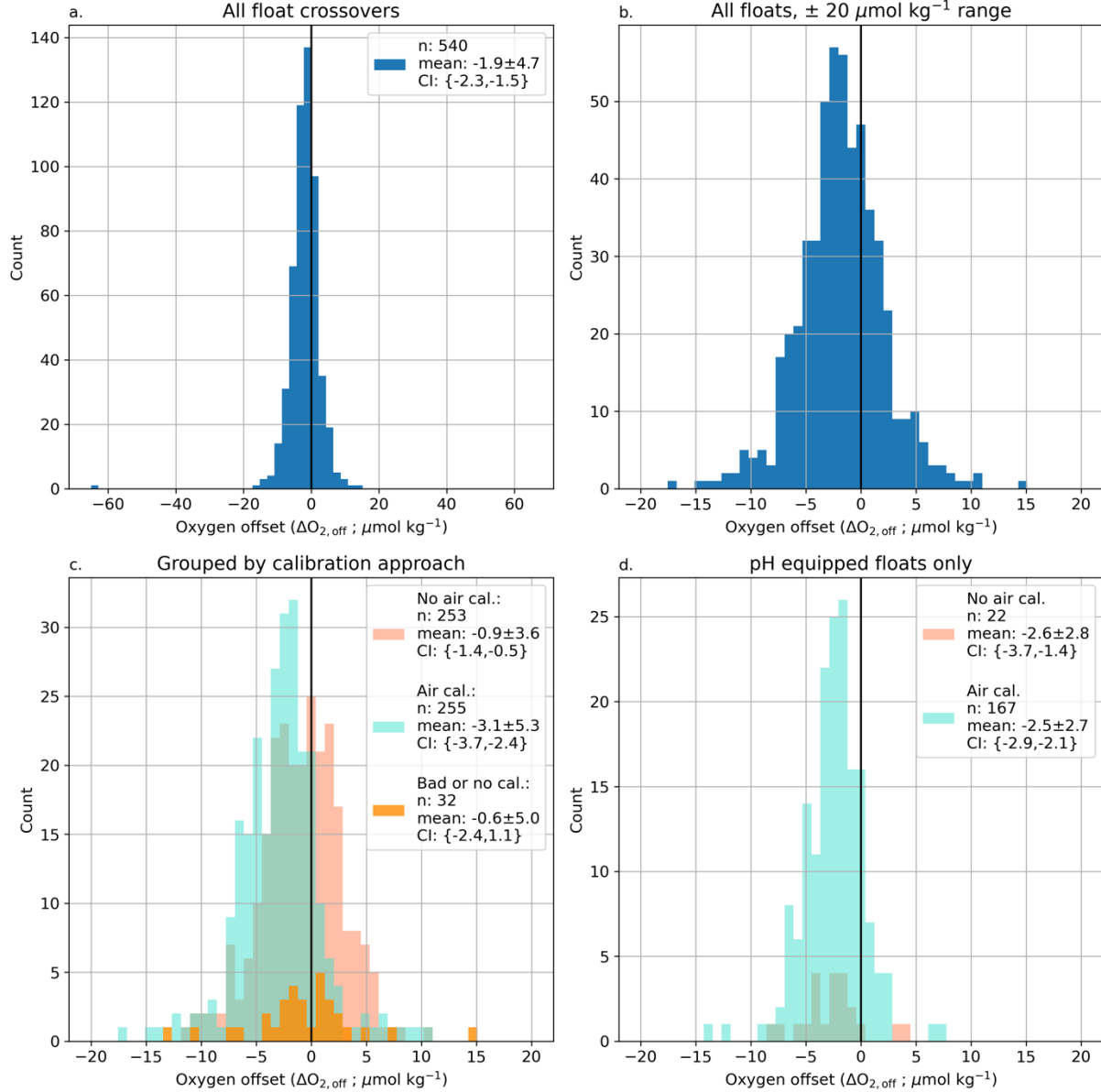255   to only adjust those floats with significant offsets so as not to over-adjust already good data.

Figure 3. Histograms of float oxygen offsets relative to shipboard measurements. Float minus ship mean offsets between 1450 and 2000 db for all floats (a) with crossovers to GLODAP data and within ± 0.005 kg m^{-3} σ_θ, ± 0.005 τ, and ± 100 db. (b) same as (a) but with a restricted x-axis range of ± 20 µmol kg^{-1}. (c) Offsets grouped by calibration type: non-air calibration listed (light red), air-calibrated (turquoise), and no calibration or bad calibration, though still marked as good data (orange). (d) same as (c) but only showing data from non-air and air calibrated floats equipped with pH sensors. Figure legends list the number of floats, mean offsets ± 1 SD, and 95% confidence intervals for each calibration category.

256   The mean oxygen offset for all floats is -1.9 ± 4.7 (1 SD) µmol kg^{-1} (n=540, float minus

257 GLODAP, Figure 3, Table S2), with a 95% confidence interval around the mean of -2.3 to -1.5

258 µmol kg^{-1}. Of the float oxygen sensors with valid crossovers, 253 had a non-air calibration method

259 listed (offset of -0.9 ± 3.6, 95% CI {-1.4, -0.5} µmol kg^{-1}), 255 were air calibrated (offset of -3.1

260 ± 5.3, 95% CI {-3.7, -2.4} µmol kg^{-1}), and 32 had bad calibration or no calibration listed, though

261 they were still flagged as good data (offset of -0.6 ± 5.0, 95% CI {-2.4, 1.1} µmol kg^{-1}) (Figure 3).

262 For floats equipped with pH, non-air calibrated floats had a mean oxygen offset of -2.6 ± 2.8 µmol

263 kg^{-1} (n=22, 95% CI {-3.7, -1.4} µmol kg^{-1}) and air-calibrated floats had a mean oxygen offset

264 of -2.5 ± 2.7 µmol kg^{-1} (n=167, 95% CI {-2.9, -2.1} µmol kg^{-1}).

265   Offsets were also calculated between GLODAP and float measurements of nitrate and pH

266 and between float estimates of DIC and GLODAP measurements (Figure 4, blue shaded

267 histograms). The mean difference between float and GLODAP nitrate measurements is 0.2 ± 0.4

268 µmol kg^{-1} ($\Delta NO_3^-{}_{off}$, n=232, 95% CI {0.1, 0.2} µmol kg^{-1}). The mean difference between float and

269 GLODAP pH (normalized to 25°C, total pH scale) is -4.5 ± 7.7 mpH ($\Delta pH_{25\text{-}T,off}$, n=113, 95% CI

270 {-5.9, -3.1} mpH). The mean difference between DIC estimated from float pH and LIARv2

271 alkalinity and GLODAP DIC measurements is 2.9 ± 6.1 µmol kg^{-1} (ΔDIC_{off}, n=154, 95% CI {1.9,

272 4.0} µmol kg^{-1}, full statistics for nitrate, pH$_{25\text{-}T}$, and DIC crossovers in Table S3).

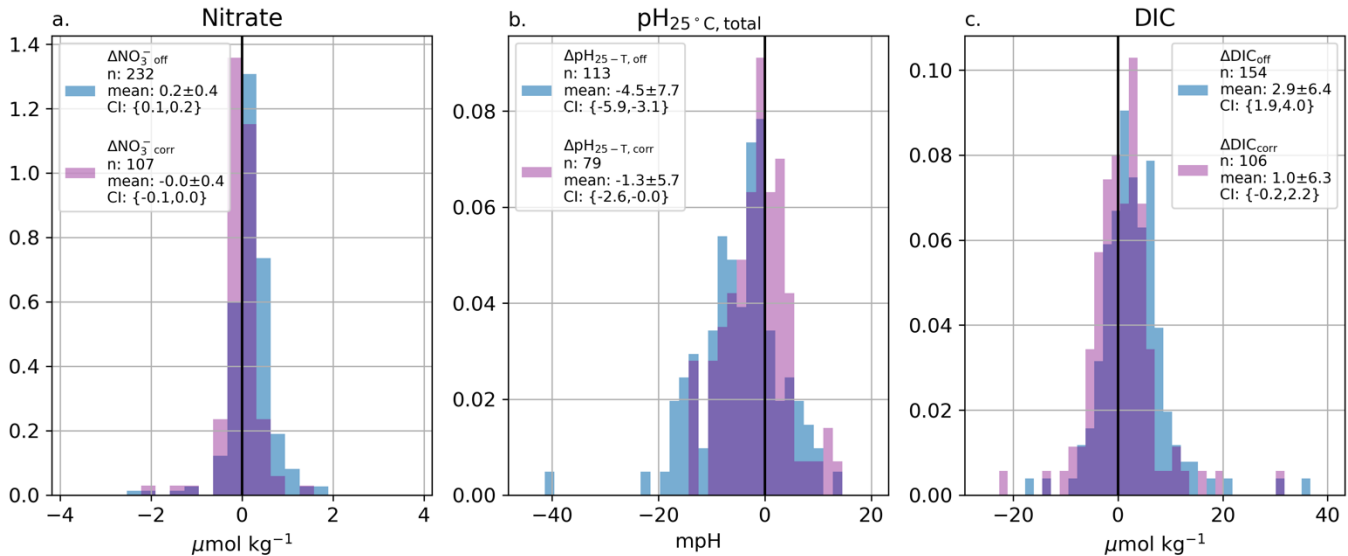**Figure 4. Frequency distributions of original nitrate, pH, and DIC offsets relative to GLODAP crossovers ($\Delta NO_3^-{}_{,off}$, ΔpH_{off}, ΔDIC_{off}, blue shaded histograms) and crossover comparisons after the impact of correcting for the oxygen offset has been applied ($\Delta NO_3^-{}_{,corr}$, ΔpH_{corr}, ΔDIC_{corr}, purple shaded histograms).** Correcting for the observed oxygen offset at depth fully corrects the nitrate offset relative to GLODAP and improves the pH and DIC crossover comparisons by 3.2 mpH and 1.9 µmol kg^{-1}, respectively. Full statistics in Tables S3 and S4.

273

274

4. Discussion

4.1 Oxygen offsets

The magnitude of the mean offset is larger for air-calibrated floats (-3.1 µmol kg^{-1}) than non-air calibrated floats (-0.9 µmol kg^{-1}). This likely reflects that some non-air calibrated floats are corrected using both deep and near-surface shipboard values or a full oxygen profile from a cast made at the deployment location (Figure 2) (Drucker & Riser, 2016; Takeshita et al., 2013), while air calibrated floats are primarily adjusted using a gain value derived from surface measurements of atmospheric oxygen over the float's lifetime (Bushinsky et al., 2016; Johnson et al., 2015, 2017; Maurer et al., 2021). The offsets shown in Figure 3 are small relative to the original correction from raw to "adjusted" oxygen (median difference between raw and adjusted float oxygen of -11 µmol kg^{-1}), but the mean offset for all comparisons other than the "Bad / no. cal" is significantly different from zero (95% confidence intervals, Figure 3 and Table S2) and, as we will discuss below, can have significant impacts on interpretation and use of this data.

The offset between float and shipboard data could reflect either that the float oxygen measurements are low or ship-based Winkler measurements are high. We will discuss both possibilities. It is important to recognize that we are considering approximately 12 different oxygen sensor models that utilize two different measurement principles made by three different manufacturers. The earliest oxygen sensors deployed on floats were initially Clark-cell electrodes (various SeaBird 43 models) while most current sensors are oxygen optodes made by multiple manufacturers, but with the same basic sensing approach and chemistry (Figure 1). Here we primarily focus on possible offsets for oxygen optodes rather than Clark cell electrodes, and on air-calibrated optodes specifically, as these are the current state of the art sensors and represent the bulk of floats deployed with pH sensors.

If the offset between float and shipboard oxygen is due to low-biased float oxygen measurements, three possible causes of the offsets are: (i) an in-situ difference in ocean oxygen content, (ii) a residual uncorrected pressure response, or (iii) a non-linear concentration-dependent drift in the sensor response. Basin-scale comparisons of float oxygen to shipboard measurements have indicated overall good float sensor accuracy in the surface ocean. A comparison of SOCCOM float oxygen data relative to GLODAP on pressure surfaces above and below the thermocline

304   indicated that the air-calibrated float data at depth may be low by 2-3 μmol kg^{-1} (Bushinsky et al.,

305   2017). Johnson et al. (2017) assessed SOCCOM float data relative to the respective float

306   deployment cruises, across the full depth range, and against GLODAP crossovers within 20 km

307   and below 300 m, finding that float oxygen was lower than GLODAP by 3.7 and 3.2 μmol kg^{-1} at

308   "Midrange", or approximately 250 μmol kg^{-1} [O$_2$]. Maurer et al. (2021) updated the Johnson et al.

309   (2017) GLODAP comparison, finding a 3.6 to 3.8 μmol kg^{-1} low oxygen offset.

310   While these earlier studies included shallower waters and compared data on pressure

311   surfaces instead of the combined σ_θ, τ, and pressure criteria of this study, all found similar

312   magnitude and direction of differences. Maurer et al. (2021) attribute the mean [O$_2$] difference to

313   a mean age difference of 18.6 years between the GLODAP and SOCCOM datasets due to the

314   linear declines in Southern Ocean interior oxygen concentrations found in Helm et al. (2011). In

315   Helm et al. (2011) the Southern Ocean has the largest region of oxygen change in the upper ocean

316   (100-1000 m) with the south Pacific and south Indian basins, seeing changes of up to -0.7 μmol l^{-1}

317   yr^{-1} between 1970 and 1992. However, deeper in the water column where we are conducting our

318   crossover comparison (1500 – 2000 m), changes are of a smaller magnitude and appear to primarily

319   be between 0 to -0.1 μmol l^{-1} yr^{-1} in the Southern Ocean and between -0.2 and +0.2 μmol l^{-1} yr^{-1}

320   across the global ocean. To rule out true oxygen changes as the source of offsets between the float

321   and GLODAP datasets, we can assess the GLODAP dataset for long-term oxygen changes and

322   compare that oxygen change to the magnitude of float oxygen offsets.

323   If true oxygen changes at depth are responsible for the offsets found in our present study,

324   we would expect the offset for a given float relative to GLODAP to be larger when compared to

325   older cruise data and decrease in magnitude (typically becoming less negative) the closer in time

326   between the float and GLODAP measurements. We therefore fit regressions to the float-GLODAP

327   offsets as a function of time, calculating a 95% confidence interval around the regression, to

328   determine if the regression ± CI included a zero offset at the midpoint time of the float deployment

329   (see Figure S2A/B for examples of the regression and CI). 66 float offsets (12%) could be a result

330   of observed change in oxygen concentration as determined by a trend in the GLODAP dataset.

331   However, in many of these cases, there is simply too much uncertainty in the GLODAP oxygen

332   regression with time, or little to no difference between the float and shipboard data. While actual

333   ocean oxygen change in the 1450 – 2000 db range remains a possible contributing factor to the

334   differences seen between the float and shipboard oxygen, it does not seem to be the main cause.

335   The optode pressure correction combines a sensor pressure response with the pressure

336   effect on oxygen solubility. Significant effort has been made to determine the appropriate pressure

337   correction between optode response and oxygen partial pressure (e.g., Bittig et al., 2015; Uchida

338   et al., 2008). We primarily focus on the mean oxygen offset for each float, averaging all crossover

339   data from 1450 – 2000 db. However, if we bin the data by depth rather than pressure, we do not

340   see a depth dependency in the float offsets in this depth range. While this may still exist, it does

341   not seem to be a first order factor in the deep oxygen offset (Figure S3). A related issue is the lag

342   in optode oxygen measurements due to the sensor response time (Bittig et al., 2014; Bittig &

343   Körtzinger, 2017). Oxygen gradients at these depths are small, so is not likely to be a first order

344   issue but may contribute to the offset in some regions.

345   Optodes have been shown to not change response at zero oxygen (Johnson et al., 2015) and

346   multiple calibrations over time have indicated relatively linear drift rates (Bittig & Körtzinger,

347   2015), lending support for the use of a gain correction across entire oxygen profiles. Bushinsky et

348   al. (2016) measured greater drift rates at lower oxygen concentration and postulated that the faster

349   drift rate at low oxygen concentrations represents a deformation of the oxygen calibration surface,

350   such that neither a gain nor an offset can be used to fully correct the range of oxygen measurements.

351   Drucker and Riser (2016) show data from eight floats indicating that a near-surface gain correction

352   leaves a negative offset at depth, similar to Bushinsky et al. (2016), but they did not provide an

353   explanation. Both Bushinsky et al. (2016) and Drucker and Riser (2016) compared float data to

354   bottle oxygen measurements from deployment casts, so oxygen change at depth did not play a

355   factor. While we are still uncertain as to the mechanism causing low float oxygen values at depth

356   relative to Winkler data, our current results from the entire oxygen Argo dataset indicate that this

357   does not seem to be an issue limited to a small number of floats and the possibility of non-linear

358   drift in the oxygen calibration surface remains.

359   An alternative to the float oxygen measurement being biased low, is that bottle

360   measurements of oxygen using Winkler titrations are biased high due to contamination by

361   atmospheric oxygen or impurities in reagents (Schmidtko et al., 2017). Sampling of oxygen and

362   subsequent Winkler titration involves careful isolation of the water from atmospheric

363   contamination. Incomplete flushing or trapped bubbles can add a significant bias to Winkler data

364   given that the low solubility of oxygen in water means that in equal volumes of air and water, the

365   air will hold 50 times more oxygen than the water. GLODAP data has been QC'd and adjusted for

366   internal consistency, but it is possible a bias is present in the deep data, especially if, on average,

367   this bias is present in either all shipboard observations or the shipboard observations used as a

368   reference for GLODAP adjustments.

369   Regardless of the source of the offset or mechanism for its existence, a systematic offset

370   between float and shipboard oxygen measurements presents a problem for the current use of

371   oxygen in float nitrate and pH parameter adjustment and derived parameter calculations. For all

372   mechanisms other than a true change in ocean oxygen content, these offsets also complicate

373   determination of long-term ocean oxygen changes. Given the mean difference in float and

374   GLODAP dataset ages and that float oxygen measurements are now made in far greater numbers

375   than shipboard data, a negative offset in the float data would appear in a combined data product as

376   an increase in the true ocean oxygen loss signal. For the rest of this discussion, we focus on the

377   impact of an oxygen offset on water properties derived from float measured data.

378

379   *4.2 Impact of oxygen offsets on pH*

380   As described earlier, float oxygen measurements at 1500 m are used to correct for drift in

381   the pH sensor data which is then used with TA estimates to derive $p\text{CO}_2$ (Williams et al., 2017).

382   We only consider pH-equipped floats with air-calibrated oxygen here, with equivalent figures and

383   tables for all pH-equipped floats provided in the supplement. The mean impact of correcting float

384   pH for observed oxygen offsets is 3.2 ± 3.8 mpH (n=119, 95% CI of 2.5 to 3.9 mpH, Figure 5,

385   Table 1). The impact of oxygen offsets on pH corrections can be understood through the

386   relationship between oxygen and inorganic carbon. A negative oxygen offset means that float

387   oxygen is lower than the expected value, so correcting the oxygen by a positive amount would

388   result in a corresponding reduction in the apparent remineralization signature of the water mass.

389   Higher oxygen and less remineralization would then imply lower dissolved inorganic carbon and
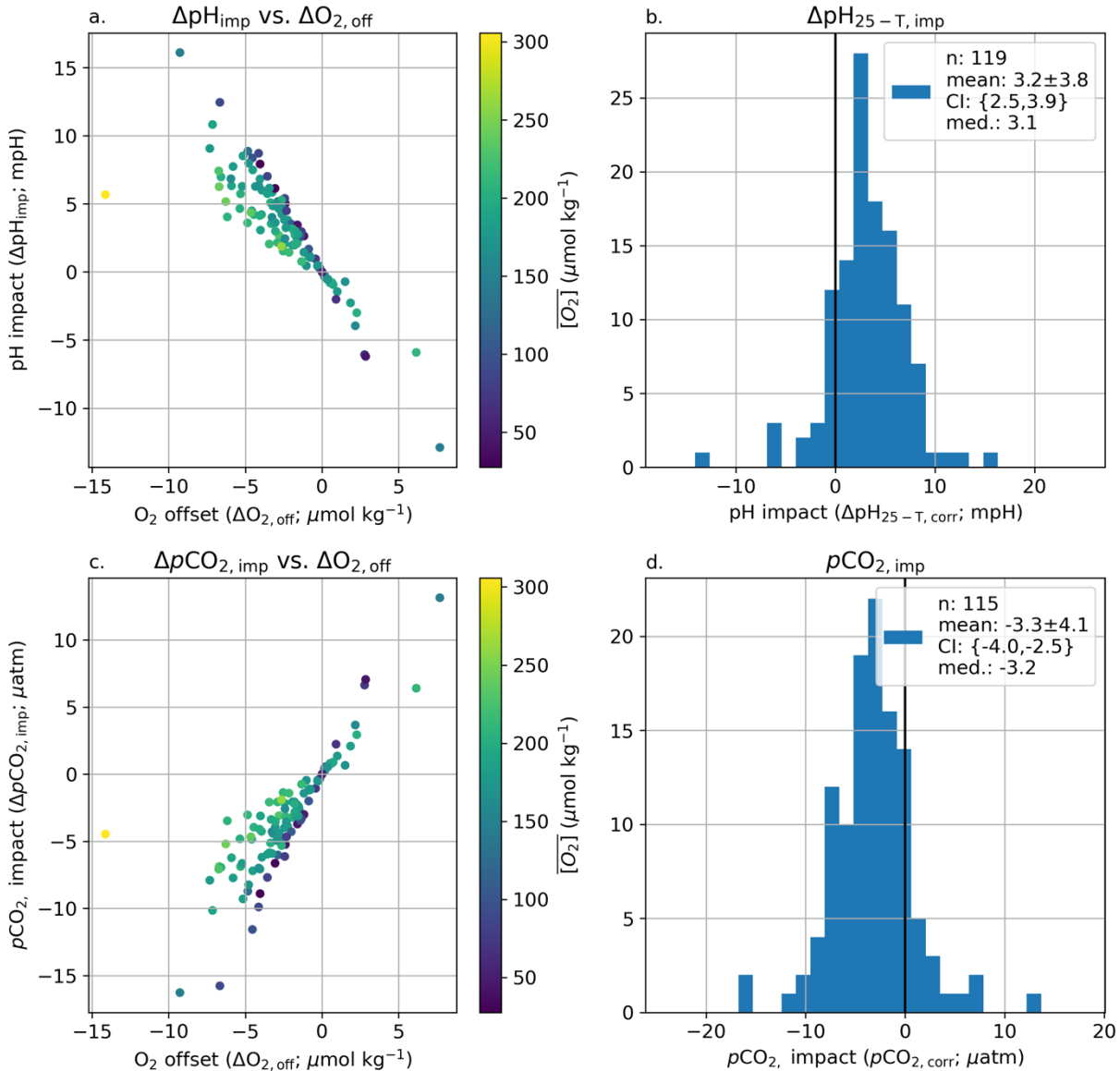
390   therefore higher pH.

391

392

393

394

395

396

397

398

Figure 5. 1500 m pH and surface pCO₂ impact from associated oxygen offset. The calculated impact of correcting pH (top panels) and pCO₂ (bottom panels) for observed oxygen offsets. (a) and (c) are scatter plots of the change in pH and pCO₂ for a given oxygen offset, with the points colored by the mean oxygen concentration of the float crossover comparisons. Correcting a float oxygen sensor that is low of correct would increase the float pH and decrease derived pCO₂. (b) and (d) are histograms of the pH and pCO₂ impacts, with zero marked with a black line and mean ± 1SD, 95% confidence intervals around the mean, and median values listed in the figure legends. pH impacts shown are at 1500 m, while pCO₂ impacts are from surface values. Results shown here are calculated using air-calibrated floats and the ESPER-mixed pH algorithm, equivalent results for LIPHR in Supplemental Figure S4 and for all pH-equipped floats in Figures S5 and S6.

399

400 **Table 1. Impact on pH at 1500 db and derived surface pCO_2 of observed oxygen offsets.**

	Count	Mean	SD	P value[2]	95.0% CI {low, high}	median	min	max
pH impact ($\Delta pH_{25\text{-}T,imp}$, mpH)[1]	119	3.2	3.8	<0.001	{2.5, 3.9}	3.1	-12.9	16.1
pCO_2 impact ($\Delta pCO_{2,imp}$, µatm)	115	-3.3	4.1	<0.001	{-4.0, -2.5}	-3.2	-16.3	13.1

401 [1]Impacts for pH calculated using ESPER-mixed. Impacts shown here are for air-calibrated floats

402 only. Equivalent numbers for LIPHR in Table S5.

403 [2]P-value testing the hypothesis that the mean impact is different from 0 at a 95% confidence

404 level.

405

406   The sensitivity of pH impact ($\Delta pH_{25\text{-}T,imp}$) to oxygen offset (represented as $\Delta pH_{25\text{-}T,imp}/\Delta O_{2,off}$) is between -0.4 and -2.3 mpH/(µmol kg^{-1}) (Figures 5, 6). The largest $\Delta pH_{25\text{-}T,imp}$ values

407

408 are observed in the Southern and Pacific Oceans, though the oxygen offsets do not show a similar

409 spatial pattern. Instead, there is a spatial pattern to the $\Delta pH_{25\text{-}T,imp}/\Delta O_{2,off}$ that corresponds to the

410 crossover oxygen concentration. The pH impact sensitivity (slope of the pH impact to O_2 offset)

411 is greatest at low oxygen concentrations (Figure 5), with the mean crossover oxygen concentration

412 explaining 61% of the variance in calculated $\Delta pH_{25\text{-}T,imp}/\Delta O_{2\,Off}$ (Figure 6).

413   This relationship can be understood by considering the changes in ocean chemistry as

414 respiration takes place in a parcel of water. For example, we can take a parcel of water with initial

415 TA and DIC of newly formed Subantarctic Mode Water from the southeast Pacific Ocean (Carter

416 et al., 2014). If we then calculate the impact of organic matter respiration on oxygen, DIC, and

417 TA, we can calculate the change in pH for every mole of oxygen respired. Initially, the $\Delta pH_{25\text{-}T,imp}/\Delta O_{2\,Off}$ is -2 mpH/(µmol kg^{-1}), but it drops to almost -2.8 mpH/(µmol kg^{-1}) after 250 µmol

418

419 kg^{-1} of oxygen have been respired (Figure S10), as the buffer capacity of the water is eroded with

420 increasing DIC; a situation analogous to surface ocean acidification. In the real ocean, regional

421 differences in ocean interior biogeochemistry, including the buffer capacity, will cause regional

422 differences in the sensitivity of pH to oxygen change. This likely accounts for larger spread in the

423 relationship between pH impact and oxygen offset for points in the Southern Ocean (Figure 6e),

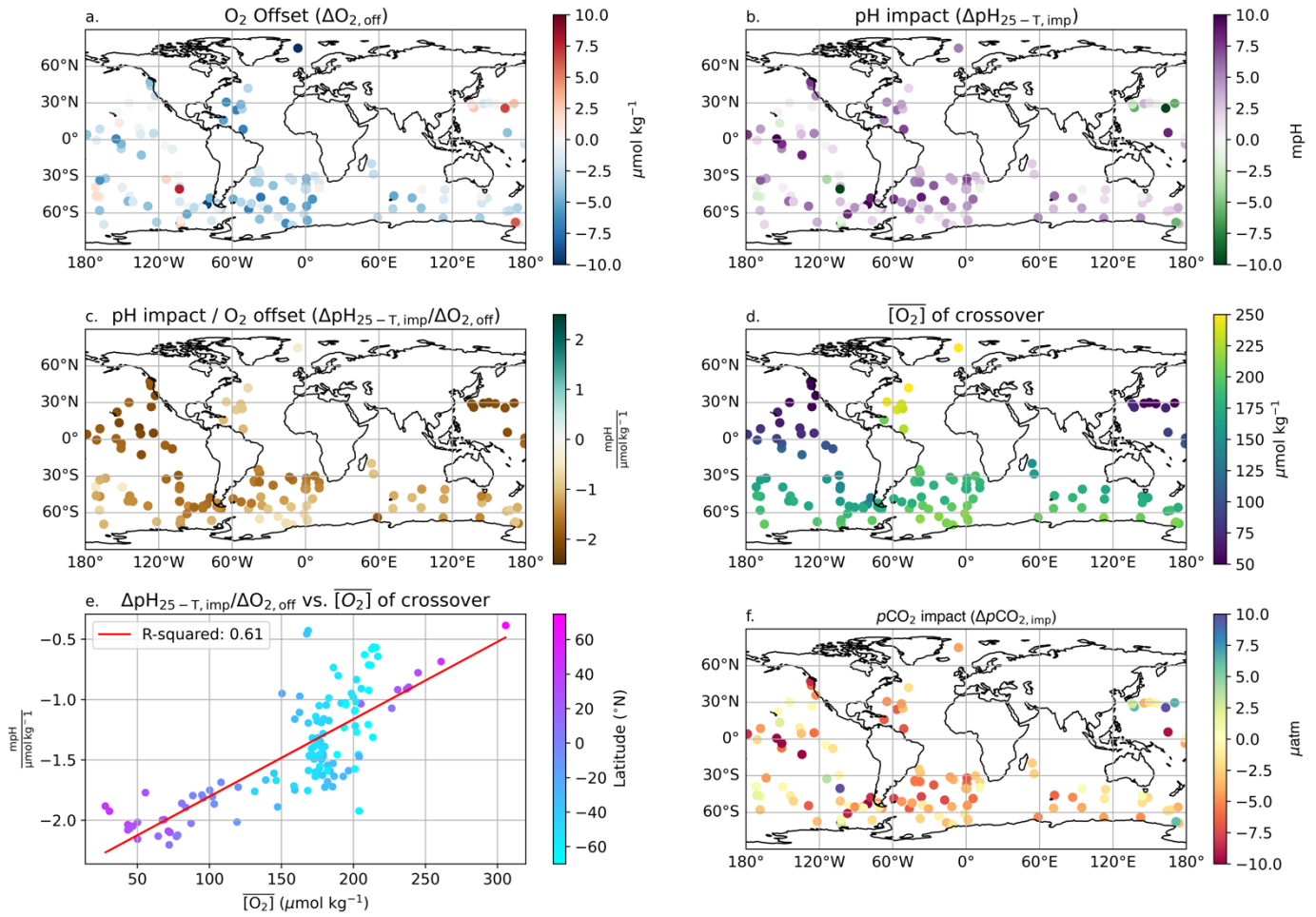424 where there is significant mixing of different water masses.

425

426

Figure 6. Relationships between oxygen offset, pH impact, mean oxygen concentration at crossover, and pCO_2 impact. Map of O_2 offset (a.) demonstrates no obvious spatial pattern in the magnitude of offsets while the impact of correcting for the O_2 offsets on pH (b.) tends to be greater in the Pacific and subpolar Southern Ocean. The ratio of the pH impact to O_2 offset ($\Delta pH_{25\text{-}T,imp}/\Delta O_{2,off}$, c.) is greatest in the north Pacific and subpolar Southern Ocean, reflecting the mean oxygen concentration at crossover ($\overline{[O_2]}$, d.). Plotting $\Delta pH_{25\text{-}T,imp}/O_{2,off}$ against $\overline{[O_2]}$ (e.) and calculating a linear fit (red line) indicates that variations in $\overline{[O_2]}$ explain ~ 61% of the variance in $\Delta pH_{imp}/\Delta O_{2,off}$. Much of the deviation from the linear fit occurs in the Southern Ocean (light blue points). This response leads to stronger pCO_2 impacts (f.) in the Pacific and subpolar Southern Ocean than in the North Atlantic or polar Southern Ocean. pH impacts shown are at 1500 m, while pCO_2 impacts are from surface values. Results shown here are calculated using floats with air-calibrated oxygen and the ESPER-mixed pH algorithm, equivalent results for LIPHR in Supplemental Figure S7 and for all pH-equipped floats in Figures S8 and S9.

427   *4.3 Impact of oxygen offsets on derived pCO₂*

428   The mean impact on float pCO_2 ($\Delta pCO_{2,imp}$) for correcting observed oxygen offsets is -3.3

429    $\pm$ 4.1 µatm (air-calibrated floats only, 95% CI of -4.0 to -2.5 µatm, Figure 5, Table 1). A reduction

430   of float derived $p\text{CO}_2$ by this magnitude would account for a large portion of the apparent bias in

431   float $p\text{CO}_2$ described in Wu and Qi (2023). For many of the studies with direct crossover

432   comparisons, a mean adjustment of -3.3 µatm would effectively eliminate the observed differences

433   between float and shipboard $p\text{CO}_2$. It should be noted that the range in impacts is much greater

434   than the mean (a range of $p\text{CO}_2$ impact between -16.3 and 13.1 µatm), with the impact on

435   individual floats and regions differing from the mean impact. Following pH, some of the greatest

436    $p\text{CO}_2$ impacts are found in the Southern Ocean and Pacific (Figure 6). Similar results are found if

437   LIPHR is used to calculate pH impacts instead of ESPER, though with a slightly greater mean

438    $p\text{CO}_2$ impact magnitude (-3.6 ± 4.7, 95% CI {-4.4, -2.7} µatm, Table S5) and differences for

439   individual floats.

440

441   *4.4 Impacts of oxygen offset on nitrate and derived DIC*

442   As described above, relative to GLODAP float nitrate is offset high (0.2 ± 0.4 µmol/kg),

443   pH is offset low (-4.5 ± 7.7 mpH), and DIC is offset high (2.9 ± 6.4 µmol/kg, Table S3). These

444   offsets are in approximately the correct ratios and directions for a biological signal, either real or

445   due to an offset in the oxygen measurement that is then propagated to pH and DIC. This indicates

446   that the oxygen offset is most likely not due to a problem with the Winkler titrations.

447   The calculated impact on nitrate of the oxygen offset ($\Delta\text{NO}_3^-{}_{imp}$) is small (mean -0.2 ± 0.2,

448   95% CI {-0.22, -0.14} µmol kg^{-1}, Figure S11, Table S6), in keeping with Redfield stoichiometry

449   of -16 N: 154 O_2 (Hedges et al., 2002) multiplied by an oxygen offset of ~2 µmol kg^{-1}. The impact

450   on DIC at 1500 m of correcting for oxygen offsets is -1 ± 1.2 µmol kg^{-1} (CI {-1.3, -0.8}, Figures

451   S12, S13, Table S7), which is the slightly smaller than multiplying the oxygen offset by a Redfield

452   ratio of -106 C: 154 O_2.

453   As a check that the mean impacts ($\Delta\text{NO}_3^-{}_{imp}$, $\Delta\text{pH}_{25\text{-}T,imp}$) calculated from the oxygen offset

454   would, in fact, improve the crossover comparisons with GLODAP data, we also re-ran the

455   crossover comparison after applying the $\Delta\text{O}_{2,off}$, $\Delta\text{NO}_3^-{}_{imp}$ and $\Delta\text{pH}_{25\text{-}T,imp}$ to the original data. The

456   resulting crossovers ($\Delta\text{NO}_3^-{}_{,corr}$, $\Delta\text{pH}_{25\text{-}T,corr}$, ΔDIC_{corr}, Figure 4, purple shaded histograms,

457   statistics in Table S4) indicate that, on average, correcting the nitrate for the oxygen offset

458   eliminates the entire nitrate offset, 3.2 mPH of the original -4.5 mPH pH offset, and 1.9 µmol kg^{-1}

459   of the original 2.9 µmol kg^{-1} DIC offset.

460   The improvement in, but not full correction of, float-derived DIC and measured pH relative
461   to the GLODAP crossovers provides an independent assessment that correcting the float data for
462   this deep O_2 offset does, in fact, improve the pH data and subsequent carbonate system derived
463   parameters. Furthermore, the fact that the nitrate bias appears to be fully corrected while the DIC,
464   pH, and pCO$_2$ impacts only partially correct for the differences between float values and crossover
465   or indirect comparisons gives an indication that the deep oxygen offset is not the only source of
466   bias in the derived float carbonate parameters. This could be due to additional biases in the float
467   pH, biases in the estimated TA, penetration of the ocean acidification signal to these depths,
468   internal consistency issues with the marine carbonate system, or some other factor. It appears that
469   correcting the oxygen may resolve a large fraction of the differences, but additional work remains
470   to fully separate pCO$_2$, pH, and DIC biases in the float data from true trends.

471   Float oxygen crossovers with shipboard data are not available for all floats, making it
472   difficult to determine the magnitude of an oxygen offset for all pH-equipped floats at this time.
473   One option to reduce the possibility of an oxygen-induced bias in pH and derived pCO$_2$ is to correct
474   float pH using an algorithm that does not use oxygen. Removing oxygen from the ESPER
475   algorithm yields a 4.2 ± 6.0 mpH and -4.3 ± 6.4 µatm pCO$_2$ impact, somewhat larger than the 3.2
476   ± 3.8 mpH and -3.3 ± 4.1 µatm impact we calculate from applying the observed oxygen offset
477   relative to GLODAP (Figure S14). Plotting the ESPER pH and pCO$_2$ impacts due to removing
478   oxygen or correcting for the observed offset against one another yields a cluster of points around
479   the 1:1 lines, though with considerable spread (Figure S15). This approach may indeed improve
480   derived parameters for floats with no GLODAP oxygen crossover for comparison, but is less
481   accurate than if an oxygen correction is possible.

482

483

5. Conclusions

485   Here we identify an offset between float and ship-board oxygen measurements between
486   1450 and 2000 db. The magnitude of this offset is significant for studies assessing long-term ocean
487   oxygen changes and for the use of float oxygen in adjusting pH and nitrate measurements and in
488   subsequent calculated derived carbonate system quantities.

489   Correcting oxygen offsets of -2.5 µmol kg^{-1} in pH-equipped, air-calibrated floats results in
490   mean pH changes of 3.2 mpH and pCO$_2$ changes of -3.3 µatm. These differences are of similar

491   direction and magnitude as many of the direct in-situ comparisons between float pCO_2 estimates

492   and underway pCO_2 measurements and therefore represent a first-order bias that needs to be

493   addressed in the biogeochemical float dataset. We do not definitively identify whether the oxygen

494   offset is present in the float or ship-board dataset, though the fact that correcting float oxygen

495   would improve the nitrate, DIC, and pH crossover comparisons to shipboard data are a strong

496   indication that the issue lies with the float observations.

497

498   Based on our findings we offer the following recommendations:

499   1. The oxygen offsets described here represent an empirical correction that must be

500   investigated to determine the underlying mechanism.

501   2. Float oxygen data at depth should be adjusted to historical shipboard data until our

502   mechanistic understanding of the causes behind the float – ship differences is sufficient

503   such that air calibration or other adjustments do not require a secondary correction.

504   3. Care should be used in combining float and ship data, with an understanding that small

505   biases may be present in either dataset that could impact the usability of the data

506   compilation to answer some scientific questions. The difference between the average age

507   of the float and GLODAP datasets and the shift to float profile numbers greatly exceeding

508   shipboard profiles in the 2000's mean that any offset between the datasets will appear as a

509   spurious change in ocean oxygen content. Additionally, individual floats may have

510   significantly greater magnitude biases than the mean and data should be assessed prior to

511   any use relying on absolute accuracy.

512   4. Standardization of float sensor calibration comments and equations will make future

513   studies of overall biogeochemical Argo float performance easier to perform.

514

Data availability

516   The "argo_synthetic-profile_index.txt" file used to determine biogeochemical Argo float WMO

517   numbers and data locations was downloaded from ftp.ifremer.fr/ifremer/argo/dac. GLODAP data

518   is available at https://glodap.info/index.php/merged-and-adjusted-data-product-v2-2022/. The

519   analysis and plotting code used for this manuscript are available at: 10.5281/zenodo.10866941

520   (Bushinsky et al., 2024).

521

522

539

540
541   **References**
542   Argo. (2000). Argo float data and metadata from Global Data Assembly Centre (Argo GDAC).

543   *SEANOE*. https://doi.org/10.17882/42182

544   Argo. (2023). Argo float data and metadata from Global Data Assembly Centre (Argo GDAC) -

545   Snapshot of Argo GDAC of June 09st 2023 [Data set]. SEANOE.

546   https://doi.org/10.17882/42182#103075

547   Bakker, D. C. E., Pfeil, B., Landa, C. S., Metzl, N., O'Brien, K. M., Olsen, A., et al. (2016). A

548   multi-decade record of high-quality fCO2 data in version 3 of the Surface Ocean CO2

549   Atlas (SOCAT). *Earth System Science Data*, *8*(2), 383–413. https://doi.org/10.5194/essd-

550   8-383-2016

551   Bittig, H. C., & Körtzinger, A. (2015). Tackling oxygen optode drift: Near-surface and in-air

552   oxygen optode measurements on a float provide an accurate in-situ reference. *Journal of*

553   *Atmospheric and Oceanic Technology*, *32*(8), 1536–1543.

554   https://doi.org/10.1175/JTECH-D-14-00162.1

555   Bittig, H. C., & Körtzinger, A. (2017). Technical note: Update on response times, in-air

556   measurements, and in situ drift for oxygen optodes on profiling platforms. *Ocean*

557   *Science*, *13*(1), 1–11. https://doi.org/10.5194/os-13-1-2017

558   Bittig, H. C., Fiedler, B., Scholz, R., Krahmann, G., & Körtzinger, A. (2014). Time response of

559   oxygen optodes on profiling platforms and its dependence on flow speed and

560   temperature. *Limnology and Oceanography: Methods*, *12*, 617–636.

561   https://doi.org/10.4319/lom.2014.12.617

562   Bittig, H. C., Fiedler, B., Fietzek, P., & Körtzinger, A. (2015). Pressure response of Aanderaa

563   and Sea-Bird oxygen optodes. *Journal of Atmospheric and Oceanic Technology*.

564   Bittig, H. C., Steinhoff, T., Claustre, H., Fiedler, B., Williams, N. L., Sauzède, R., et al. (2018).

565   An Alternative to Static Climatologies: Robust Estimation of Open Ocean CO2 Variables

566   and Nutrient Concentrations From T, S, and O2 Data Using Bayesian Neural Networks.

567   *Frontiers in Marine Science*, *5*(September), 328.

568   https://doi.org/10.3389/fmars.2018.00328

569   Bittig, H. C., Körtzinger, A., Neill, C., van Ooijen, E., Plant, J. N., Hahn, J., et al. (2018).

570   Oxygen Optode Sensors: Principle, Characterization, Calibration, and Application in the

571   Ocean. *Frontiers in Marine Science*, *4*. https://doi.org/10.3389/fmars.2017.00429

572   Bushinsky, S. M., & Cerovečki, I. (2023). Subantarctic Mode Water Biogeochemical Formation

573   Properties and Interannual Variability. *AGU Advances*, *4*(2), e2022AV000722.

574   https://doi.org/10.1029/2022AV000722

575   Bushinsky, S. M., & Emerson, S. (2015). Marine biological production from in situ oxygen

576   measurements on a profiling float in the subarctic Pacific Ocean. *Global Biogeochemical*

577   *Cycles*, *29*(12), 2050–2060. https://doi.org/10.1002/2015GB005251

578   Bushinsky, S. M., Emerson, S. R., Riser, S. C., & Swift, D. D. (2016). Accurate oxygen

579   measurements on modified Argo floats using in situ air calibrations. *Limnology and*

580   *Oceanography: Methods*, *14*(8), 491–505. https://doi.org/10.1002/lom3.10107

581   Bushinsky, S. M., Gray, A. R., Johnson, K. S., & Sarmiento, J. L. (2017). Oxygen in the

582   Southern Ocean From Argo Floats: Determination of Processes Driving Air-Sea Fluxes.

583   *Journal of Geophysical Research: Oceans*, *122*(11), 8661–8682.

584   https://doi.org/10.1002/2017JC012923

585   Bushinsky, S. M., Landschützer, P., Rödenbeck, C., Gray, A. R., Baker, D., Mazloff, M. R., et al.

586   (2019). Reassessing Southern Ocean Air-Sea CO 2 Flux Estimates With the Addition of

587   Biogeochemical Float Observations. *Global Biogeochemical Cycles*, *33*(11), 1370–1388.

588   https://doi.org/10.1029/2019GB006176

589   Bushinsky, S. M., Tamsitt, V., Nachod, Z., Fassbender, A., & Williams, N. (2024).

590   SethBushinsky/argo_deep_o2_bias: Offset between profiling float and shipboard oxygen

591   observations at depth imparts bias on float pH and derived pCO2 analysis and plotting

592   code (Version v0.1.1-submission). Retrieved from

593   https://doi.org/10.5281/zenodo.10866942

594   Carpenter, J. H. (1965). The Chesapeake Bay Institute Technique for the Winkler Dissolved

595   Oxygen Method. *Limnology and Oceanography*, *10*(1), 141–143.

596   https://doi.org/10.4319/lo.1965.10.1.0141

597   Carter, B. R., Radich, J. A., Doyle, H. L., & Dickson, A. G. (2013). An automated system for

598   spectrophotometric seawater pH measurements. *Limnology and Oceanography: Methods*,

599   *11*(1), 16–27. https://doi.org/10.4319/lom.2013.11.16

600   Carter, B. R., Talley, L. D., & Dickson, A. G. (2014). Mixing and remineralization in waters

601   detrained from the surface into Subantarctic Mode Water and Antarctic Intermediate

602   Water in the southeastern Pacific. *Journal of Geophysical Research: Oceans*, *119*(6),

603   4001–4028. https://doi.org/10.1002/2013JC009355

604   Carter, B. R., Williams, N. L., Gray, A. R., & Feely, R. A. (2016). Locally interpolated alkalinity

605   regression for global alkalinity estimation. *Limnology and Oceanography: Methods*,

606   (1998). https://doi.org/10.1002/lom3.10087

607   Carter, B. R., Feely, R. A., Williams, N. L., Dickson, A. G., Fong, M. B., & Takeshita, Y.

608   (2018). Updated methods for global locally interpolated estimation of alkalinity, pH, and

609   nitrate. *Limnology and Oceanography: Methods*, *16*(2), 119–131.

610   https://doi.org/10.1002/lom3.10232

611   Carter, B. R., Bittig, H. C., Fassbender, A. J., Sharp, J. D., Takeshita, Y., Xu, Y.-Y., et al.

612   (2021). New and updated global empirical seawater property estimation routines.

613   *Limnology and Oceanography: Methods*, *19*(12), 785–809.

614   https://doi.org/10.1002/lom3.10461

615   Carter, B. R., Sharp, J. D., Dickson, A. G., Álvarez, M., Fong, M. B., García-Ibáñez, M. I., et al.

616   (2023). Uncertainty sources for measurable ocean carbonate chemistry variables.

617   *Limnology and Oceanography*, lno.12477. https://doi.org/10.1002/lno.12477

618   Claustre, H., Johnson, K. S., & Takeshita, Y. (2020). Observing the Global Ocean with

619   Biogeochemical-Argo. *Annual Review of Marine Science*, *12*(1), 1–26.

620   https://doi.org/10.1146/annurev-marine-010419-010956

621   Coggins, A., Watson, A. J., Schuster, U., Mackay, N., King, B., McDonagh, E., & Poulton, A. J.

622   (2023). Surface ocean carbon budget in the 2017 south Georgia diatom bloom:

623   Observations and validation of profiling biogeochemical argo floats. *Deep Sea Research*

624   *Part II: Topical Studies in Oceanography*, *209*, 105275.

625   https://doi.org/10.1016/j.dsr2.2023.105275

626   D'Asaro, E. A., & McNeil, C. (2013). Calibration and stability of oxygen sensors on autonomous

627   floats. *Journal of Atmospheric and Oceanic Technology*, *30*(8), 1896–1906.

628   https://doi.org/10.1175/JTECH-D-12-00222.1

629   Dickson, A. G. (1994). Determination of dissolved oxygen in sea water by Winkler titration.

630   *WHP Operations and Methods*, (November), 1–13.

631   Drucker, R., & Riser, S. C. (2016). In situ phase-domain calibration of oxygen Optodes on

632   profiling floats. *Methods in Oceanography*, *17*, 296–318.

633   https://doi.org/10.1016/j.mio.2016.09.007

634   Emerson, S. (1987). Seasonal oxygen cycles and biological new production in surface waters of

635   the subarctic Pacific Ocean. *Journal of Geophysical Research*, *92*(C6), 6535–6544.

636   https://doi.org/10.1029/JC092iC06p06535

637   Fay, A. R., Lovenduski, N. S., Mckinley, G. A., Munro, D. R., Gray, A. R., Landschützer, P., et

638   al. (2018). Utilizing the Drake Passage Time-series to understand variability and change

639   in subpolar Southern Ocean pCO 2. *Biogeosciences Discussions*.

640   https://doi.org/10.5194/bg-2017-489

641   Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Baranova, O. K., Zweng, M. M., &

642   Johnson, D. R. (2010). *World Ocean Atlas 2009 Volume 3: Dissolved Oxygen, Apparent

643   Oxygen Utilization, and Oxygen Saturation. Noaa Atlas Nesdic* (Vol. 70, p. 344pp).

644   Washington, D. C.: U.S. Government Printing Office. Retrieved from

645   http://www.nodc.noaa.gov/OC5/indprod.html

646   Gray, A. R., Johnson, K. S., Bushinsky, S. M., Riser, S. C., Russell, J. L., Talley, L. D., et al.

647   (2018). Autonomous Biogeochemical Floats Detect Significant Carbon Dioxide

648   Outgassing in the High-Latitude Southern Ocean. *Geophysical Research Letters*, *45*(17),

649   9049–9057. https://doi.org/10.1029/2018GL078013

650   Gruber, N., Scott C. Doney, Steven R. Emerson3, Denis Gilbert4, Taiyo Kobayashi5, Arne

651   Körtzinger6, Gregory C. Johnson7, Kenneth S., & Johnson8, Stephen C. Riser3, and

652   Osvaldo Ulloa. (2007). THE ARGO-OXYGEN PROGRAM: A white paper to promote

653   the addition of oxygen sensors to the international Argo float program.

654   Gruber, N., Doney, S. C., Emerson, S. R., Gilbert, D., Kobayashi, T., Körtzinger, A., et al.

655   (2009). Adding oxygen to Argo: Developing a global in-situ observatory for ocean

656   deoxygenation and biogeochemistry. *Ocean Obs '09*.

657   Hedges, J. I., Baldock, J. a., Gélinas, Y., Lee, C., Peterson, M. L., & Wakeham, S. G. (2002).

658   The biochemical and elemental compositions of marine plankton: A NMR perspective.

659   *Marine Chemistry*, *78*(1), 47–63. https://doi.org/10.1016/S0304-4203(02)00009-9

660 Helm, K. P., Bindoff, N. L., & Church, J. A. (2011). Observed decreases in oxygen content of

661   the global ocean. *Geophysical Research Letters*, *38*(23), 1–6.

662   https://doi.org/10.1029/2011GL049513

663 Hennon, T. D., Riser, S. C., & Mecking, S. (2016). Profiling float-based observations of net

664   respiration beneath the mixed layer. *Global Biogeochemical Cycles*, *30*(6), 920–932.

665   https://doi.org/10.1002/2016GB005380

666 Humphreys, M. P., Lewis, E. R., Sharp, J. D., & Pierrot, D. (2022). PyCO2SYS v1.8: marine

667   carbonate system calculations in Python. *Geoscientific Model Development*, *15*(1), 15–

668   43. https://doi.org/10.5194/gmd-15-15-2022

669 Humphreys, M. P., Schiller, A. J., Sandborn, D., Gregor, L., Pierrot, D., van Heuven, S. M. A.

670   C., et al. (2023, January 19). PyCO2SYS: marine carbonate system calculations in

671   Python (Version v1.8.2). Zenodo. https://doi.org/10.5281/ZENODO.3744275

672 Johnson, K. S., Plant, J. N., Riser, S. C., & Gilbert, D. (2015). Air oxygen calibration of oxygen

673   optodes on a profiling float array. *Journal of Atmospheric and Oceanic Technology*,

674   *32*(11), 2160–2172. https://doi.org/10.1175/JTECH-D-15-0101.1

675 Johnson, K. S., Plant, J. N., Coletti, L. J., Jannasch, H. W., Sakamoto, C. M., Riser, S. C., et al.

676   (2017). Biogeochemical sensor performance in the SOCCOM profiling float array.

677   *Journal of Geophysical Research: Oceans*, *122*(8), 6416–6436.

678   https://doi.org/10.1002/2017JC012838

679 Johnson, K. S., Maurer, T. L., Plant, J. N., & Takeshita, Y. (2023). *BGC-Argo quality control*

680   *manual for pH* [Pdf]. [object Object]. https://doi.org/10.13155/97828

681   Key, R. M., Olsen, A., van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., et al. (2015). Global

682   Ocean Data Analysis Project, Version 2 (GLODAPv2). *ORNL/CDIAC-162, NDP-093*.

683   https://doi.org/10.3334/CDIAC/OTG.NDP093_GLODAPv2

684   Körtzinger, A., Schimanski, J., & Send, U. (2005). High Quality Oxygen Measurements from

685   Profiling Floats: A Promising New Technique. *Journal of Atmospheric and Oceanic*

686   *Technology*, *22*(3), 302–308. https://doi.org/10.1175/JTECH1701.1

687   Levin, L. A. (2018). Manifestation, drivers, and emergence of open ocean deoxygenation.

688   *Annual Review of Marine Science*, *10*, 229–260. https://doi.org/10.1146/annurev-marine-

689   121916-063359

690   Liu, X., Patsavas, M. C., & Byrne, R. H. (2011). Purification and Characterization of meta-

691   Cresol Purple for Spectrophotometric Seawater pH Measurements. *Environmental*

692   *Science & Technology*, *45*(11), 4862–4868. https://doi.org/10.1021/es200665d

693   Long, M. C., Stephens, B. B., McKain, K., Sweeney, C., Keeling, R. F., Kort, E. A., et al.

694   (2021). Strong Southern Ocean carbon uptake evident in airborne observations. *Science*,

695   *374*(6572), 1275–1280. https://doi.org/10.1126/science.abi4355

696   Mackay, N., & Watson, A. (2021). Winter air-sea CO 2 fluxes constructed from summer

697   observations of the Polar Southern Ocean suggest weak outgassing. *Journal of*

698   *Geophysical Research: Oceans*. https://doi.org/10.1029/2020jc016600

699   Martz, T. R., Riser, S. C., & Johnson, K. S. (2008). Ocean metabolism observed with oxygen

700   sensors on profiling floats in the South Pacific. *Limnology and Oceanography*, *53*(5, part

701   2), 2094–2111. https://doi.org/10.4319/lo.2008.53.5_part_2.2094

702   Maurer, T. L., Plant, J. N., & Johnson, K. S. (2021). Delayed-Mode Quality Control of Oxygen,

703   Nitrate, and pH Data on SOCCOM Biogeochemical Profiling Floats. *Frontiers in Marine*

704   *Science*, *8*(August), 1–20. https://doi.org/10.3389/fmars.2021.683207

705   McDougall, T. J., & Barker, P. M. (2011). Getting started with TEOS-10 and the Gibbs Seawater

706   (GSW) Oceanographic Toolbox. SCOR/IAPSO WG127. Retrieved from

707   https://www.teos-10.org/pubs/Getting_Started.pdf

708   Olsen, A., Key, R. M., Van Heuven, S., Lauvset, S. K., Velo, A., Lin, X., et al. (2016). The

709   global ocean data analysis project version 2 (GLODAPv2) - An internally consistent data

710   product for the world ocean. *Earth System Science Data*, *8*(2), 297–323.

711   https://doi.org/10.5194/essd-8-297-2016

712   Olsen, A., Lange, N., Key, R. M., Tanhua, T., Bittig, H. C., Kozyr, A., et al. (2020).

713   GLODAPv2.2020 - the second update of GLODAPv2. *Earth System Science Data*

714   *Discussions*. https://doi.org/10.5194/essd-2020-165

715   Roemmich, D., Talley, L., Zilberman, N., Osborne, E., Johnson, K., Barbero, L., et al. (2021).

716   The Technological, Scientific, and Sociological Revolution of Global Subsurface Ocean

717   Observing. *Oceanography*, 2–8. https://doi.org/10.5670/oceanog.2021.supplement.02-02

718   Sarmiento, J. L., Johnson, K. S., Arteaga, L. A., Bushinsky, S. M., Cullen, H. M., Gray, A. R., et

719   al. (2023). The Southern Ocean carbon and climate observations and modeling

720   (SOCCOM) project: A review. *Progress in Oceanography*, *219*, 103130.

721   https://doi.org/10.1016/j.pocean.2023.103130

722   Schmidtko, S., Stramma, L., & Visbeck, M. (2017). Decline in global oceanic oxygen content

723   during the past five decades. *Nature*, *542*(7641), 335–339.

724   https://doi.org/10.1038/nature21399

725   Schofield, O. A., Fassbender, A., Hood, M., Hill, K., & Johnson, K. (2022). A global ocean

726   biogeochemical observatory becomes a reality. *Eos*, *103*.

727   https://doi.org/10.1029/2022EO220149

728   Sharp, J. D., Fassbender, A. J., Carter, B. R., Johnson, G. C., Schultz, C., & Dunne, J. P. (2022).

729   *GOBAI-O$_2$: temporally and spatially resolved fields of ocean interior dissolved oxygen*

730   *over nearly two decades* (preprint). ESSD – Ocean/Chemical oceanography.

731   https://doi.org/10.5194/essd-2022-308

732   Stramma, L., & Schmidtko, S. (2021). Spatial and Temporal Variability of Oceanic Oxygen

733   Changes and Underlying Trends. *Atmosphere-Ocean*, *59*(2), 122–132.

734   https://doi.org/10.1080/07055900.2021.1905601

735   Takeshita, Y., Martz, T. R., Johnson, K. S., Plant, J. N., Gilbert, D., Riser, S. C., et al. (2013). A

736   climatology-based quality control procedure for profiling float oxygen data. *Journal of*

737   *Geophysical Research: Oceans*, *118*(10), 5640–5650. https://doi.org/10.1002/jgrc.20399

738   Takeshita, Y., Johnson, K. S., Coletti, L. J., Jannasch, H. W., Walz, P. M., & Warren, J. K.

739   (2020). Assessment of pH dependent errors in spectrophotometric pH measurements of

740   seawater. *Marine Chemistry*, *223*, 103801.

741   https://doi.org/10.1016/j.marchem.2020.103801

742   Tengberg, A., Hovdenes, J., Andersson, H. J., Brocandel, O., Diaz, R., Hebert, D., et al. (2006).

743   Evaluation of a lifetime-based optode to measure oxygen in aquatic systems. *Limnology*

744   *and Oceanography: Methods*, *4*(1964), 7–17. https://doi.org/10.4319/lom.2006.4.7

745   Uchida, H., Kawano, T., Kaneko, I., & Fukasawa, M. (2008). In Situ Calibration of Optode-

746   Based Oxygen Sensors. *Journal of Atmospheric and Oceanic Technology*, *25*(12), 2271–

747   2281. https://doi.org/10.1175/2008JTECHO549.1

748   Williams, N. L., Juranek, L. W., Johnson, K. S., Feely, R. A., Riser, S. C., Talley, L. D., et al.

749   (2016). Empirical algorithms to estimate water column pH in the Southern Ocean.

750   *Geophysical Research Letters*, *43*, 3415–3422. https://doi.org/10.1002/2016GL068539

751   Williams, N. L., Juranek, L. W., Feely, R. A., Johnson, K. S., Sarmiento, J. L., Talley, L. D., et

752   al. (2017). Calculating surface ocean pCO2 from biogeochemical Argo floats equipped

753   with pH: An uncertainty analysis. *Global Biogeochemical Cycles*, *31*(3), 591–604.

754   https://doi.org/10.1002/2016GB005541

755   Wolf, M. K., Hamme, R. C., Gilbert, D., Yashayaev, I., & Thierry, V. (2018). Oxygen Saturation

756   Surrounding Deep Water Formation Events in the Labrador Sea From Argo-O 2 Data.

757   *Global Biogeochemical Cycles*, *32*(4), 635–653. https://doi.org/10.1002/2017GB005829

758   Wu, Y., & Qi, D. (2023). The controversial Southern Ocean air-sea CO2 flux in the era of

759   autonomous ocean observations. *Science Bulletin*, S2095927323006151.

760   https://doi.org/10.1016/j.scib.2023.08.059

761   Wu, Y., Bakker, D. C. E., Achterberg, E. P., Silva, A. N., Pickup, D. D., Li, X., et al. (2022).

762   Integrated analysis of carbon dioxide and oxygen concentrations as a quality control of

763   ocean float data. *Communications Earth & Environment*, *3*(1), 92.

764   https://doi.org/10.1038/s43247-022-00421-w

765   Yang, B., Emerson, S. R., & Bushinsky, S. M. (2017). Annual net community production in the

766   subtropical Pacific Ocean from in situ oxygen measurements on profiling floats. *Global*

767   *Biogeochemical Cycles*, *31*(4), 728–744. https://doi.org/10.1002/2016GB005545

768