Offset between profiling float and shipboard oxygen observations at depth imparts bias on float pH and derived pCO$_2$

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

[1]Department of Oceanography, School of Earth and Space Science and Technology, University of Hawaiʻi at Mānoa, Honolulu, HI, USA

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

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

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

Contents of this file

Introduction

Figures in this supplement include additional detail on the crossover analysis and alternative algorithms described in the main text.

Text S1. **Offsets between float data and GLODAP measurements.**

To help determine the optimal criteria to select crossovers with shipboard observations, we analyzed the relationship between different selection criteria and changes in float oxygen data. For each float with oxygen data, the profile closest in time to the mean float lifetime was selected as a reference profile. For each float oxygen measurement between 1500 and 2000 db, differences in measured oxygen were calculated between the reference profile and all other float profiles using a range of distance, potential density (σ_θ) and spiciness (τ) criteria. Potential density and spiciness were best able to reduce the "environmental" data noise while still retaining adequate numbers of crossovers (Figure S1).

Crossovers relative to GLODAP were analyzed to determine if differences are significant and whether they could be due to true change in ocean oxygen (Figures S2A/B). Offsets show no average significant relationship with pressure Figure S3.

Oxygen calibration approaches were grouped into "air", "non-air", and "no/bad calibration" according to the calibration comment (Table S1).

Text S2. Results for alternative pH algorithms and float selection.

Two pH algorithms are used in this manuscript. Results for ESPER (Carter et al., 2021) are presented in the main text, while results using LIPHR are included here (Figures S4, S6, S7, S9, S13). Additionally, crossovers and impacts presented in the main text focus on air calibrated floats. Results for all floats are presented here (Figures S5, S6, S8, S9).

Oxygen offset relative to nitrate impact and nitrate impact histograms (Figure S11). DIC is derived from float pH and estimated TA. Oxygen impacts on derived DIC are presented here for surface and 1500 db estimates, and for both LIPHR and ESPER pH algorithms (Figures S12 and S13).

We also tested the impact of removing oxygen entirely from the ESPER algorithm (Figure S14). This would allow a mode of correcting float pH when no crossover is available with which to assess float oxygen sensor accuracy. The impact of removing float oxygen vs. correcting float oxygen on pH and pCO$_2$ is shown in Figure S15.

Text S3. Example calculation of relationship between respiration and sensitivity of pH to incremental changes in DIC.

The pH impact of a given change in DIC will vary according to the current carbonate chemistry of a given water mass. To test this, we took water properties from newly formed Subantarctic Mode Water (Carter et al., 2014) and calculated the DIC and TA changes due to organic matter respiration in units of oxygen. We assumed a stoichiometry of DIC changes of -106 C: 154 O$_2$ and TA changes of -16 TA: 106 DIC. For each mol of O$_2$ consumed during respiration, we calculated a change in pH per mol O$_2$ (Figure S10). While initial starting chemistries would be different in different water masses and the fraction of carbonates produced will also influence the carbonate system, this provides an example of the changing relationship between pH impact and oxygen offset as a function of accumulated respiration signal.
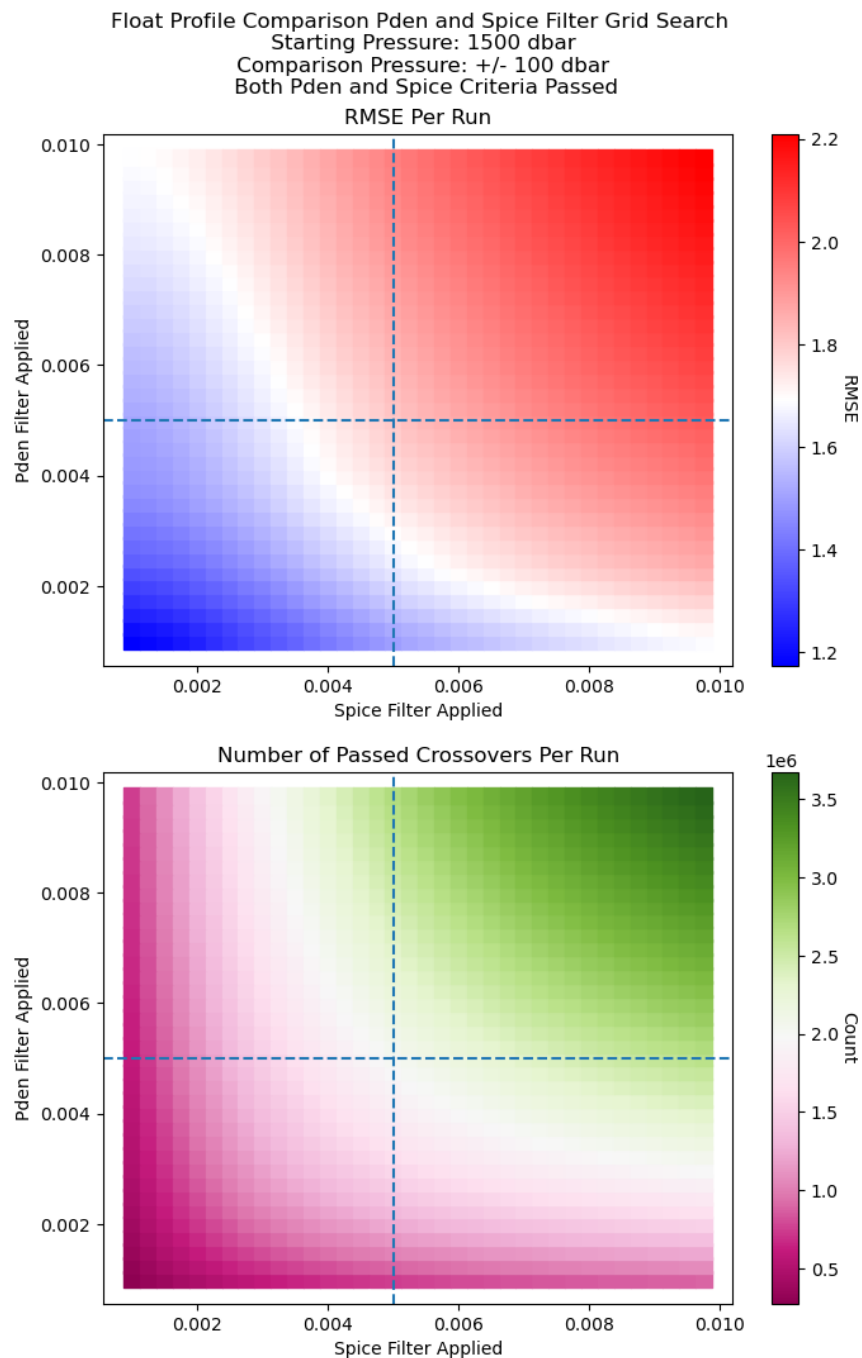
Figure S1. RMSE (μmol kg^{-1}) and amount of data remaining for different combinations of density and spice filters for float oxygen data. The "filter applied" level indicates the allowable differences (±) between measurements to include them as a crossover.
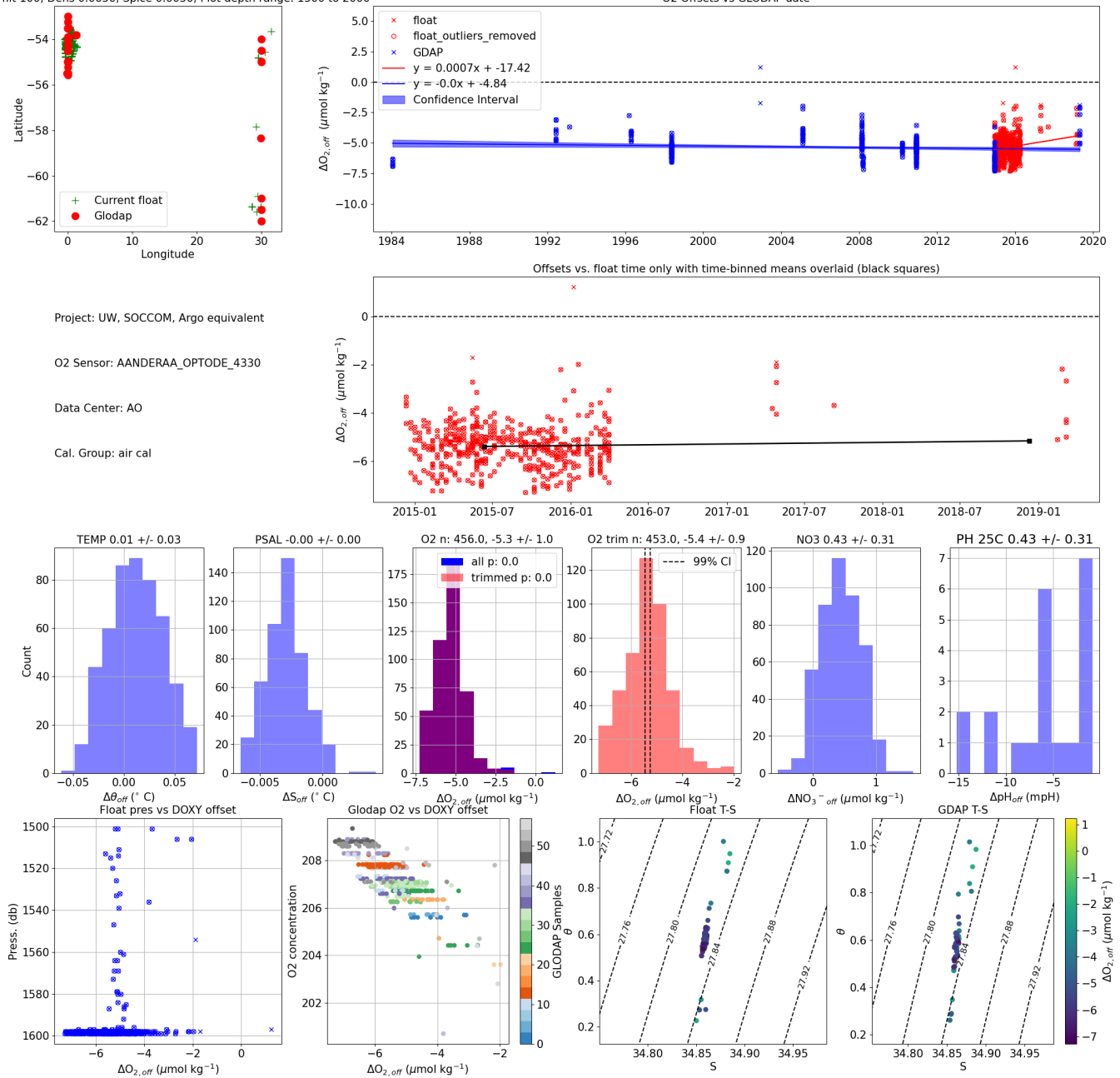
Figure S2A. Example crossover comparison between float 5904469 and GLODAP observations of oxygen, nitrate, and pH. From top left to bottom right: float and ship crossover locations; oxygen offsets (float minus ship) relative to GLODAP measurement times (blue x's) and float measurement times (red x's); oxygen offsets relative to float measurement times only (red x's); histograms of offsets for temperature, salinity, oxygen, oxygen with outliers removed (trimmed), nitrate, pH at 25C on the total scale; oxygen offset vs. float pressure; oxygen offset vs. GLODAP oxygen concentration, colored to indicate individual GLODAP measurements; oxygen offsets on T-S diagrams for float and GLODAP data. Regression of oxygen offsets relative to GLODAP data (blue line and shading) indicate that the offset for this float is not likely due to true changes in ocean oxygen concentrations. Confidence intervals for the mean oxygen offset with

outliers removed do not overlap with zero, indicating that this float oxygen is likely different from the GLODAP data.
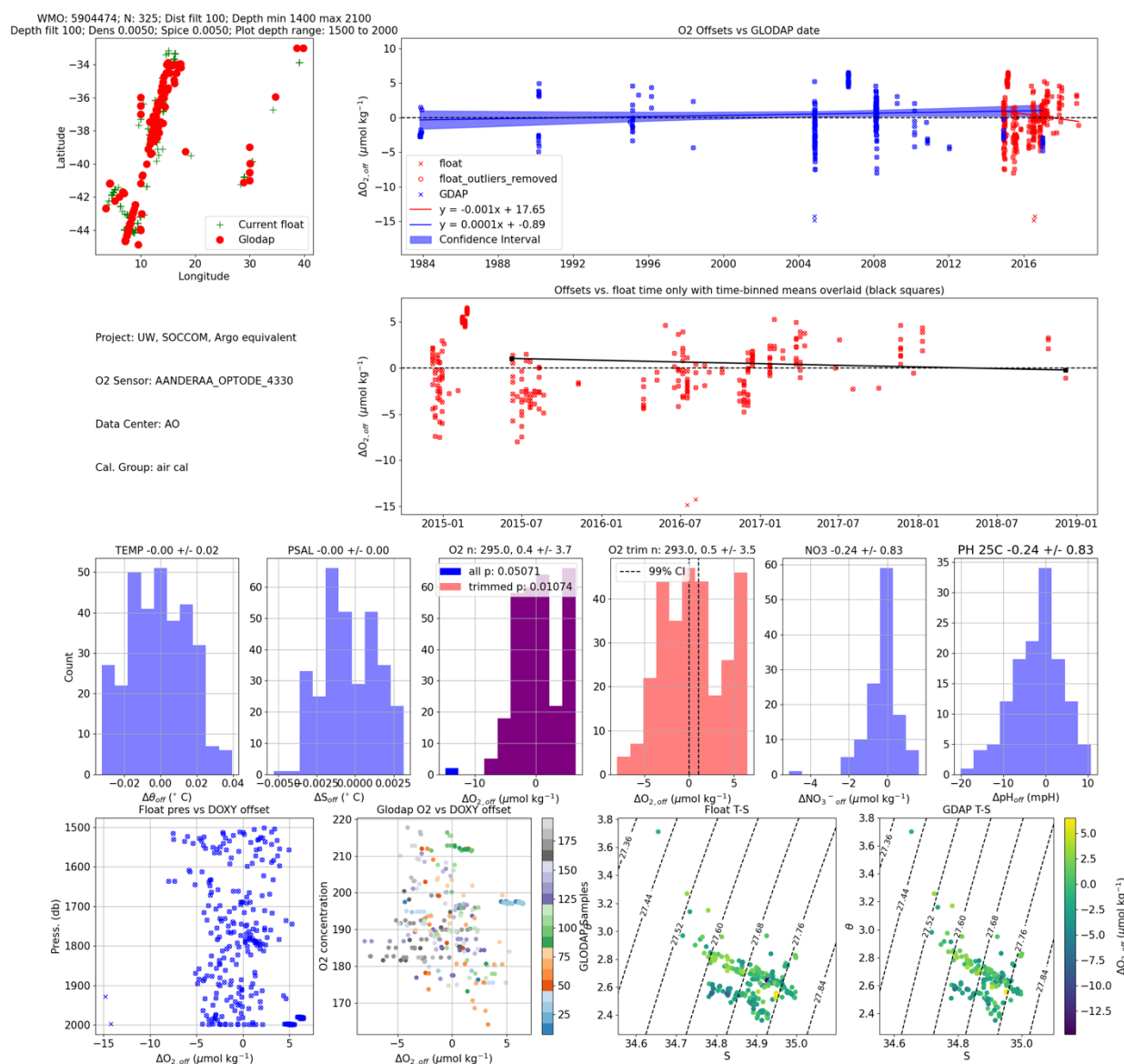

Figure S2B. Same as Figure S2A, but for float 5904474. Regression of oxygen offsets relative to GLODAP data (blue line and shading) indicate that the offset for this float may be due to true

changes in ocean oxygen concentrations. Confidence intervals for the oxygen offset with outliers removed overlap with zero, indicating that this difference is not significant.
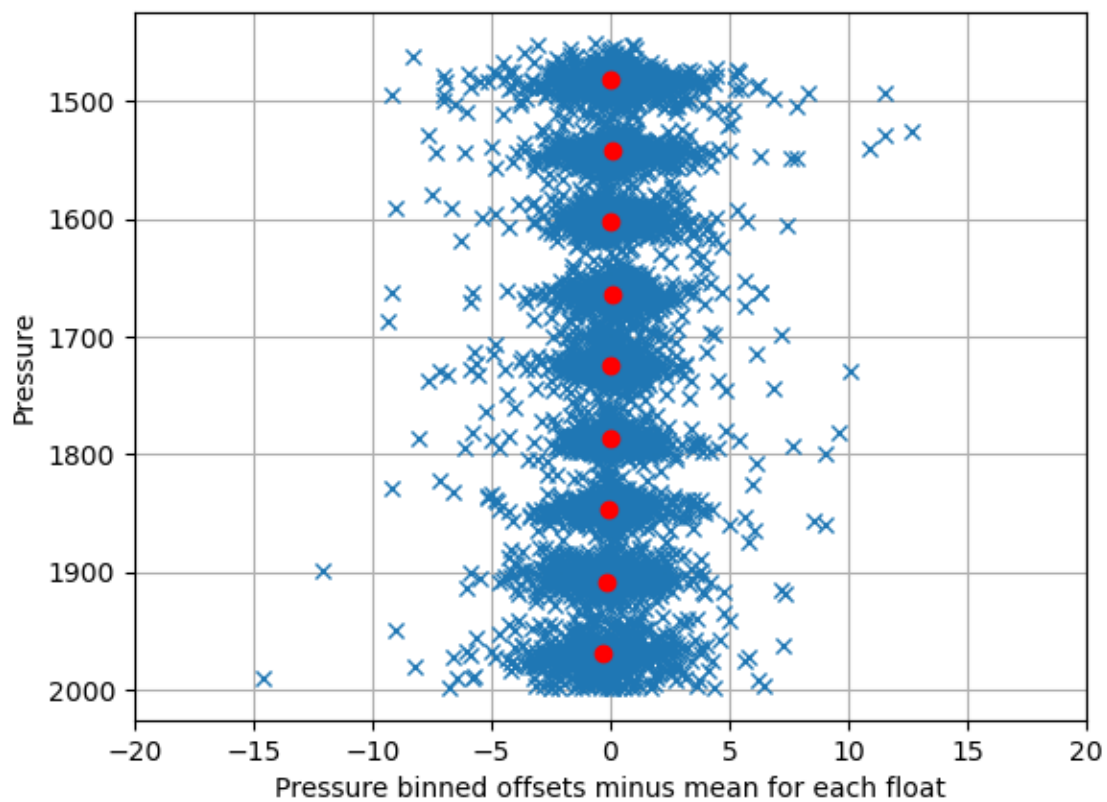
Figure S3. Float oxygen offsets (μmol kg^{-1}) binned by pressure (db), with mean float offset removed. Red dots indicate 50 db bin means.
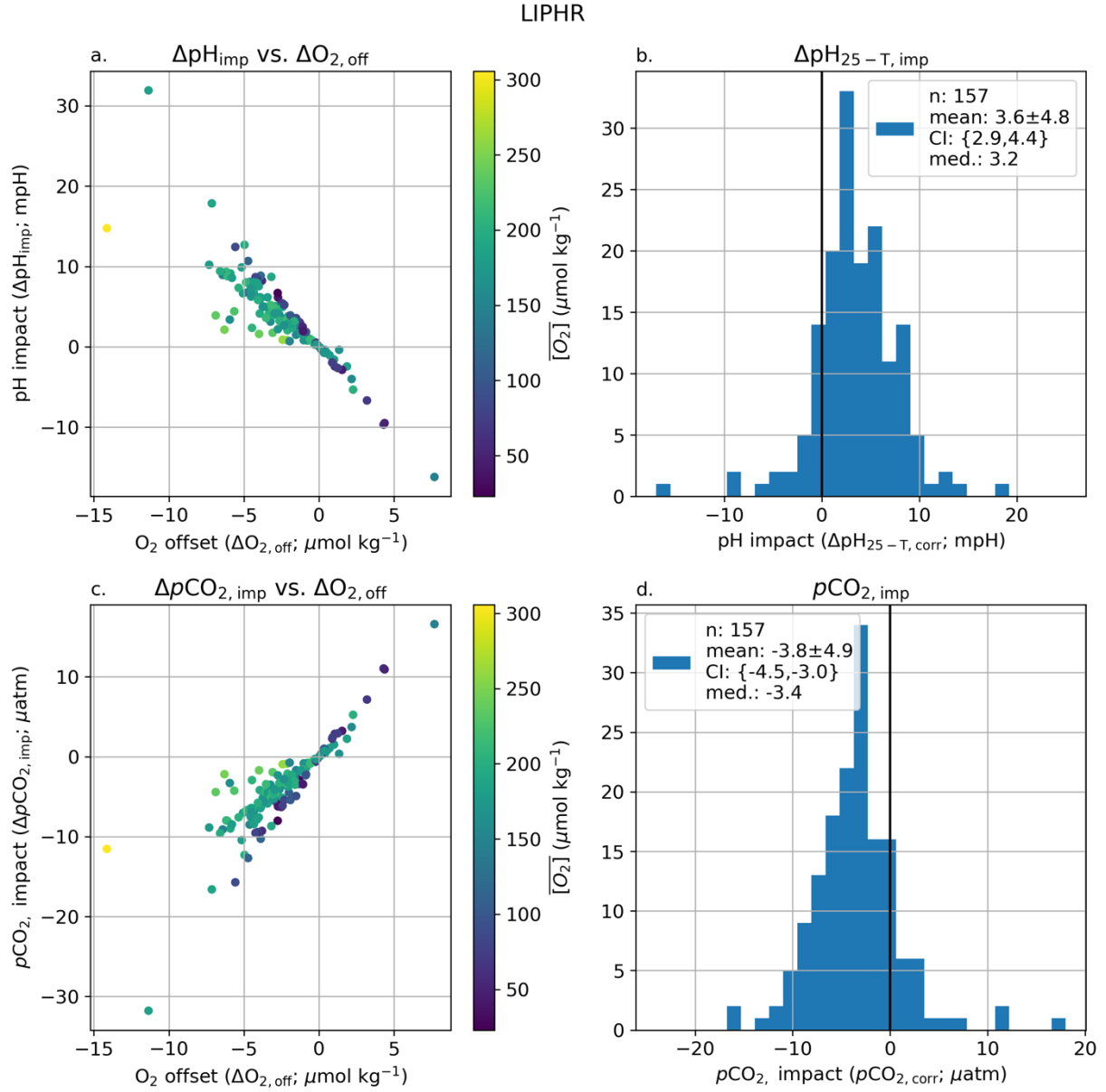
Figure S4. Equivalent to manuscript Figure 5 but using the LIPHR pH MLR algorithm instead of the ESPER mixed (neural network/MLR average). Data shown are for air calibrated floats only. Full statistics in Table S4. Figures S5 and S6 show results for all pH equipped floats.
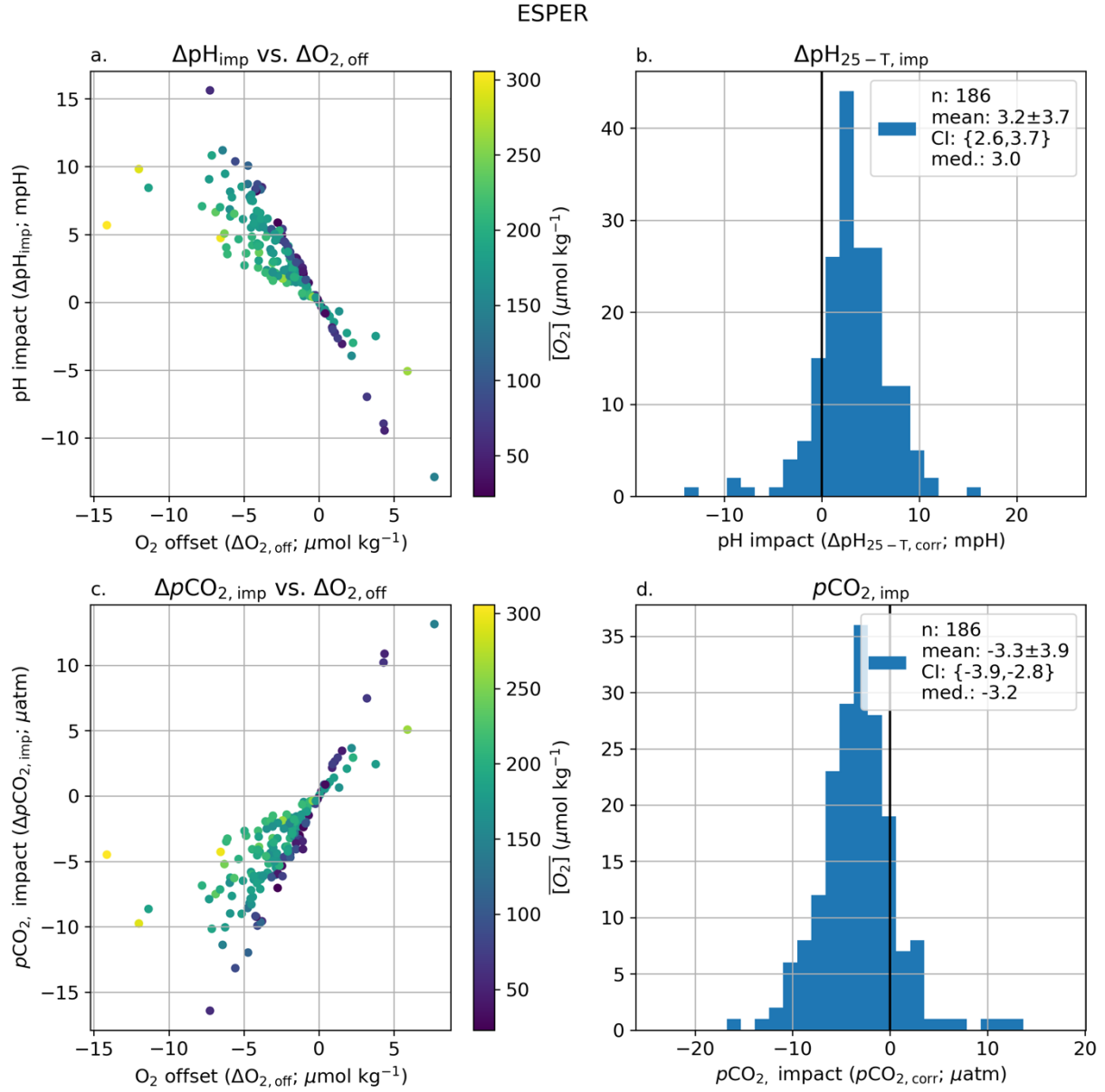
Figure S5. Impact scatter / histograms on pH and pCO₂ for all floats (not just air) – ESPER.

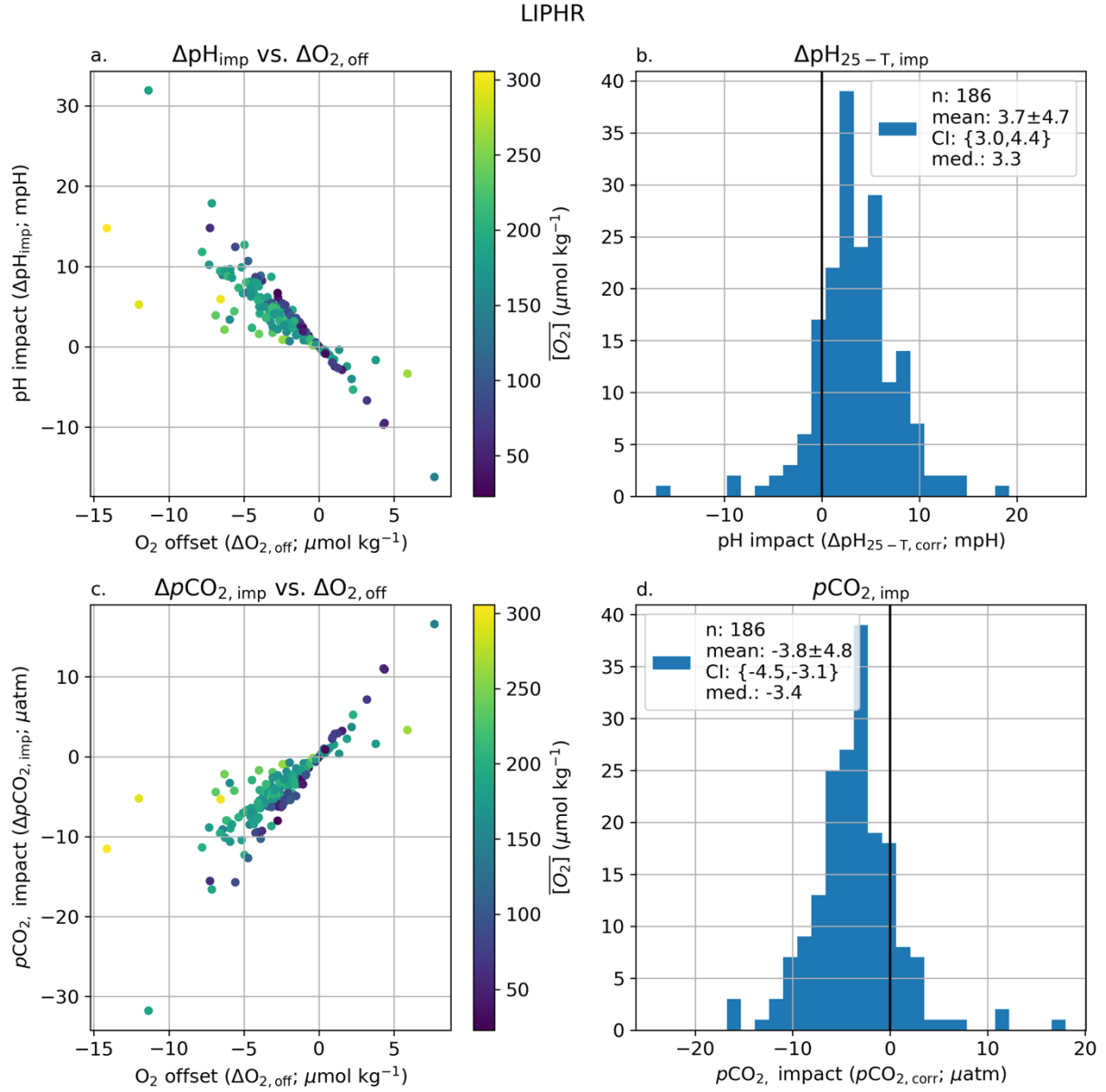**Figure S6. Impact scatter / histograms on pH and pCO₂ for all floats (not just air) – LIPHR.**
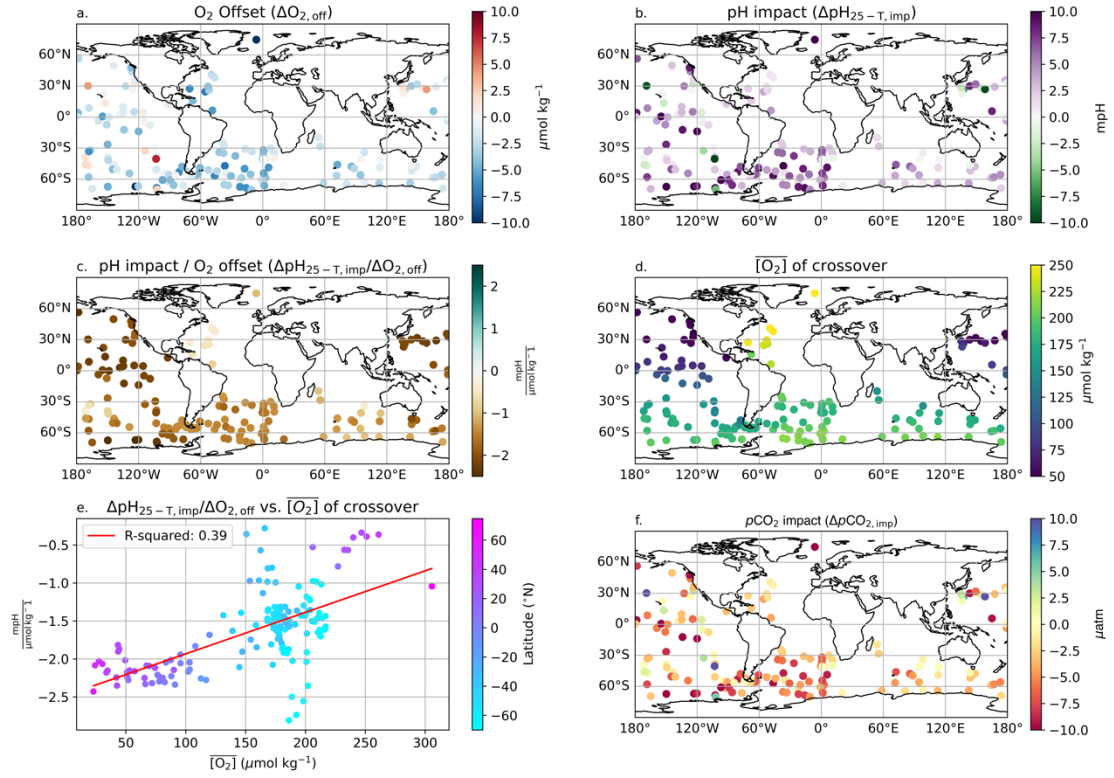
Figure S7. Same as manuscript figure 6 but using the LIPHR pH algorithm.
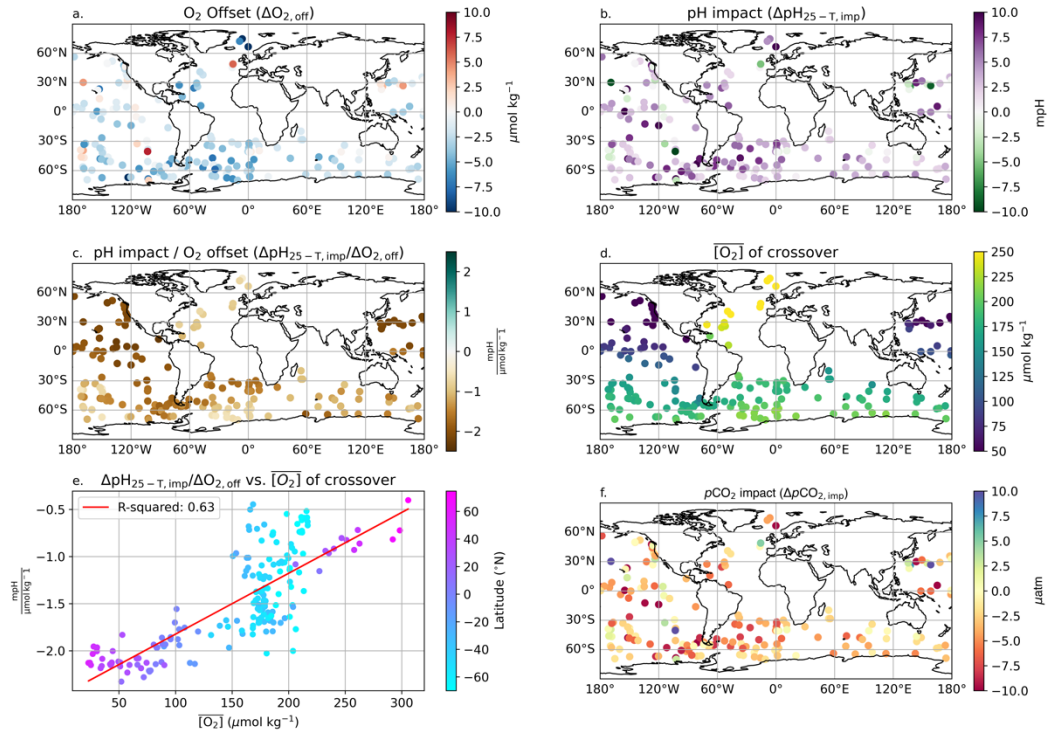
Figure S8. Same as manuscript figure 6, but showing all floats equipped with pH.
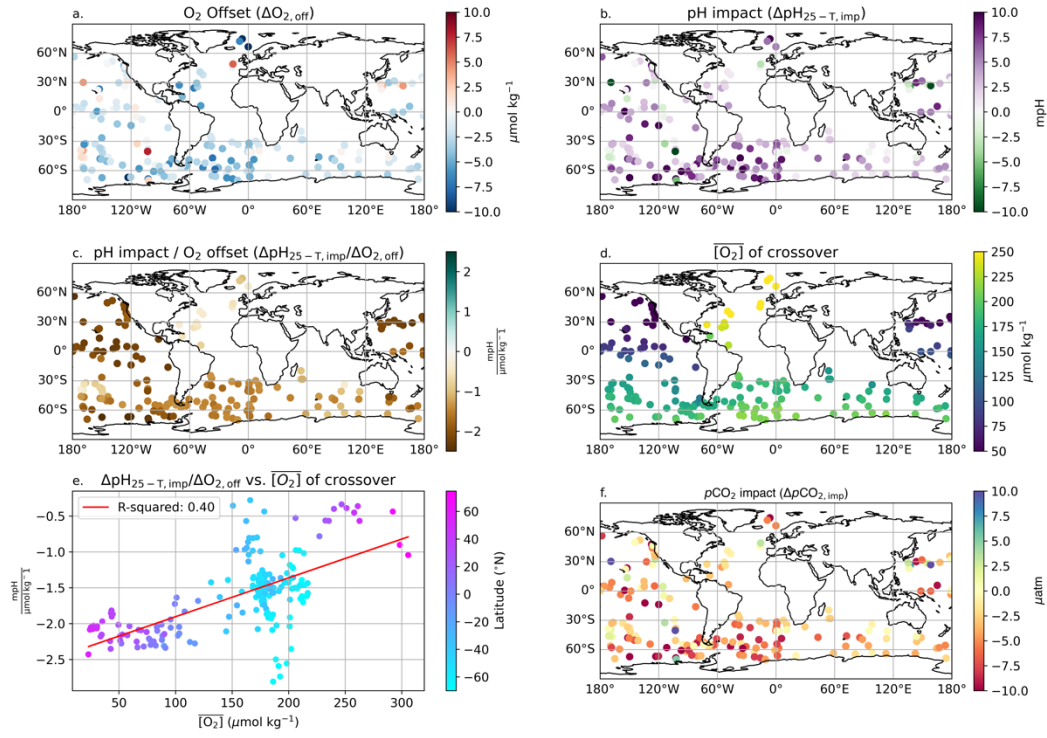
Figure S9. Same as manuscript figure 5, but showing all floats equipped with pH and using the LIPHR pH algorithm.
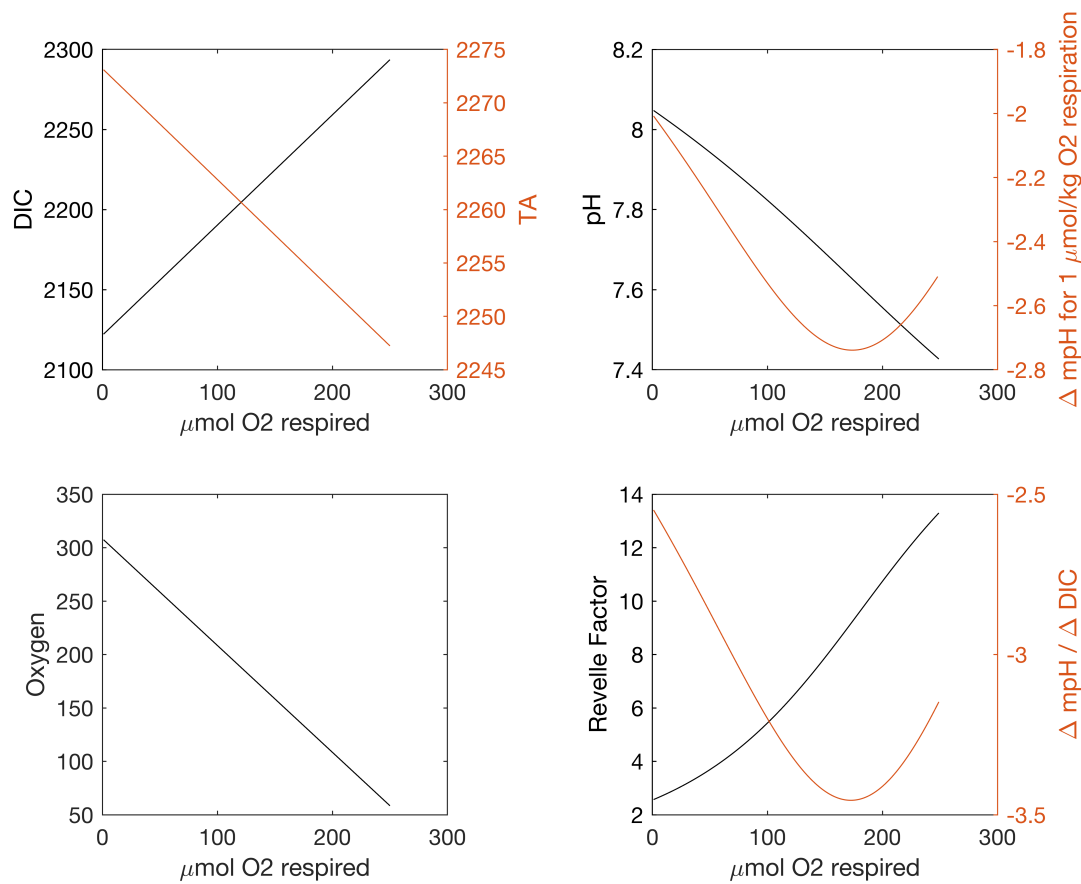
Figure S10. Example of changes in carbonate chemistry responses with accumulated respiration. Changes in DIC, TA, pH, ΔmpH/(Δμmol kg^{-1} O$_2$ respired), Revelle Factor, and ΔmpH/ΔDIC for a parcel of water starting with properties of newly formed Subantarctic Mode Water from the southeast Pacific. For mol of oxygen respired by the respiration of organic matter, each parameter is calculated to demonstrate the sensitivity of pH to a given change in oxygen.
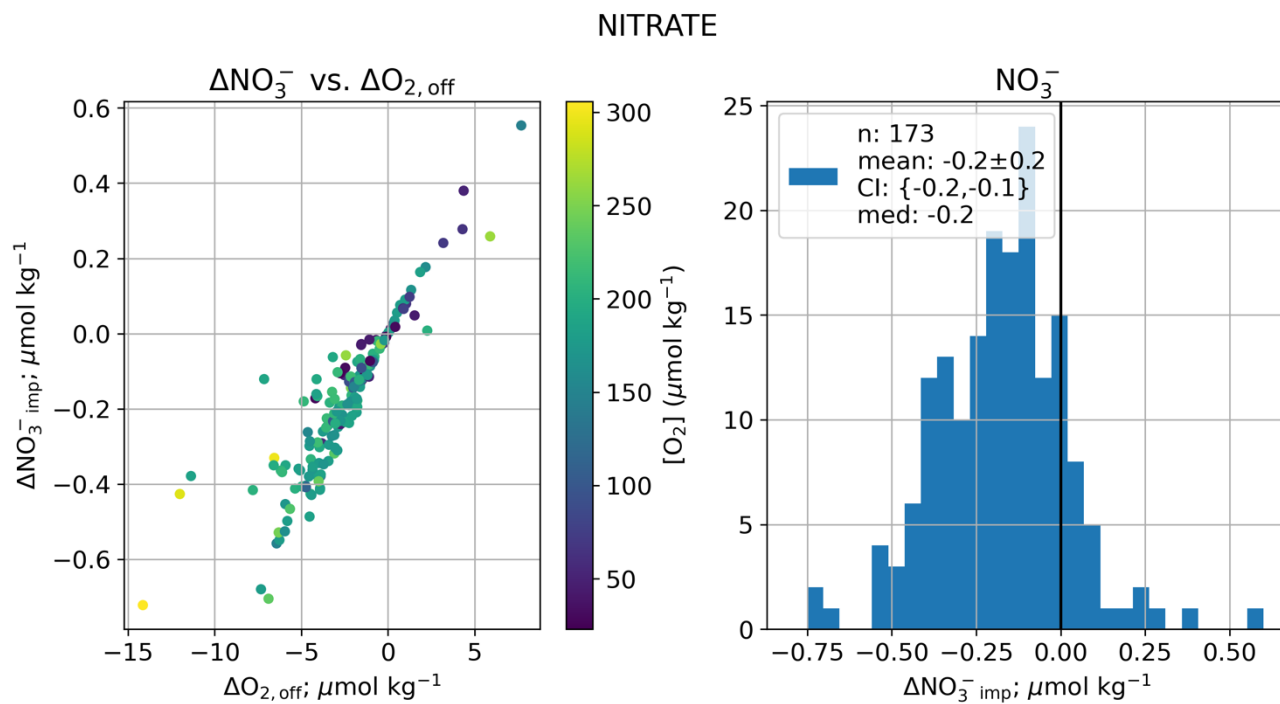
Figure S11. Impact of correcting O2 on nitrate.
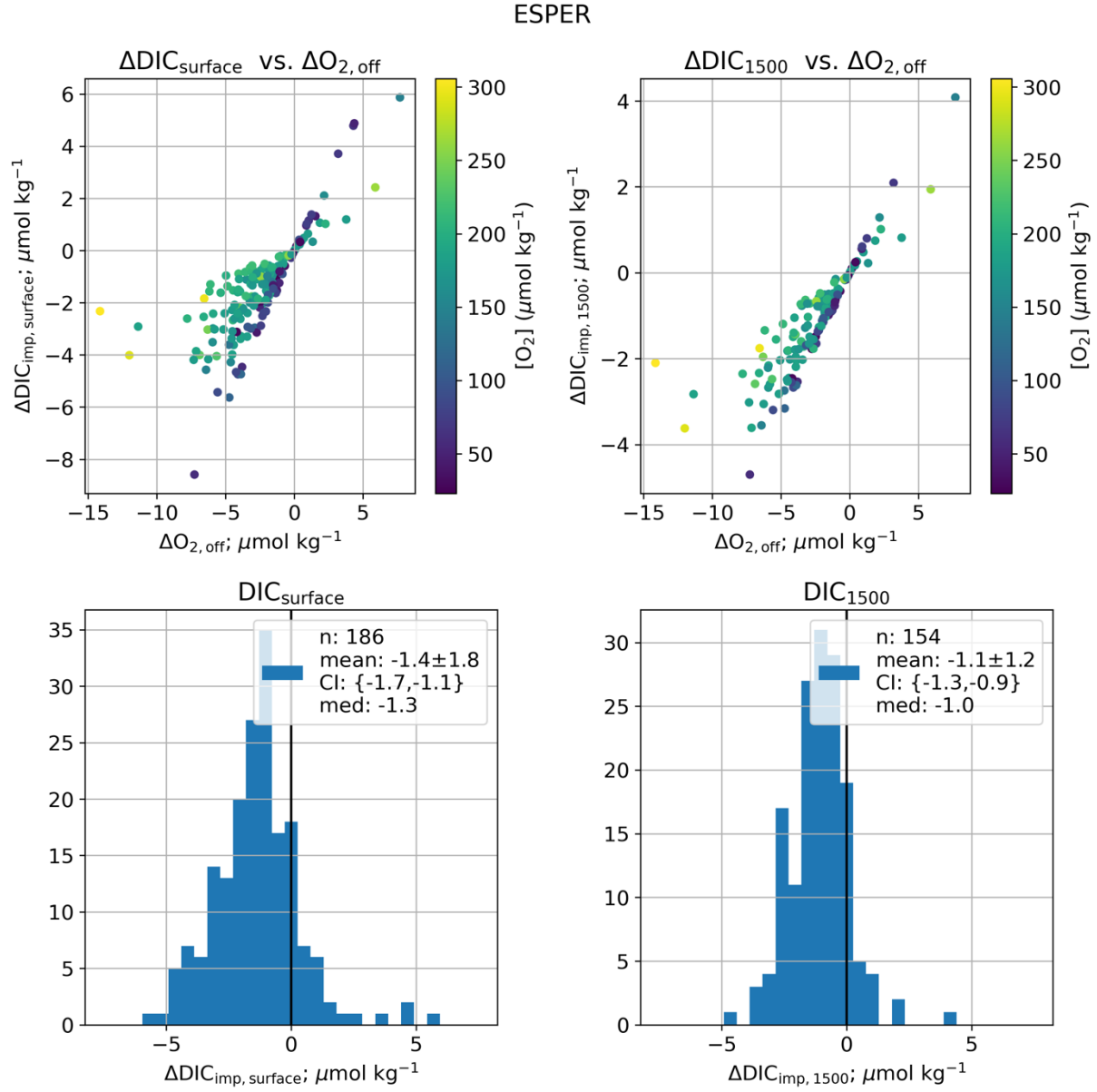
Figure S12. ESPER DIC impact. LIPHR DIC impact in Figure S13.

Figure S13. Impact of correcting O2 on DIC LPIHR

Figure S14. Impact of removing oxygen from the ESPER algorithm used to correct float pH.

Figure S15. pH and pCO_2 impacts calculated from observed oxygen offsets relative to GLODAP (x axes) vs. those calculated from removing oxygen from the ESPER algorithm (y axes).

Table S1. Oxygen sensor calibration comments

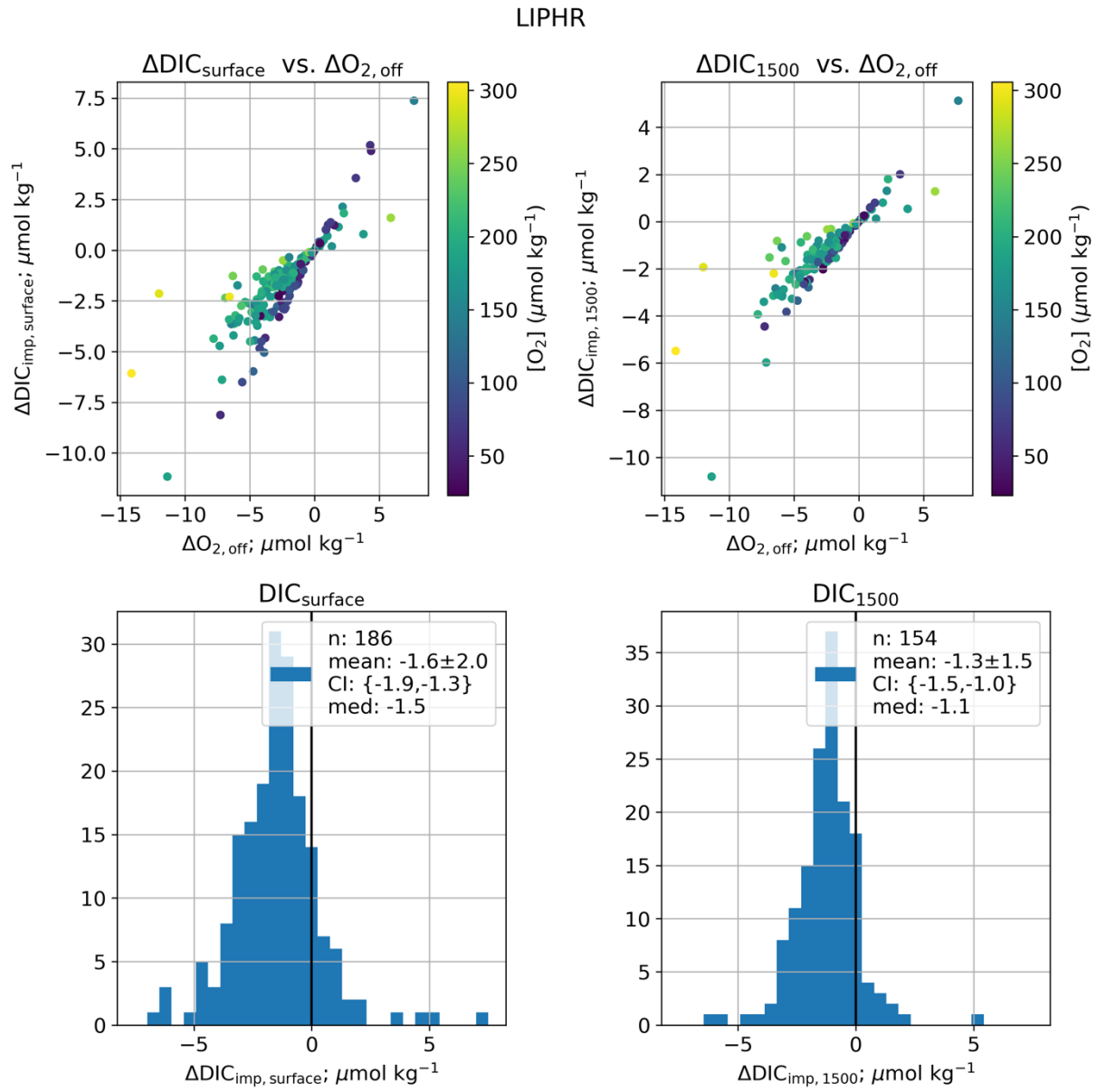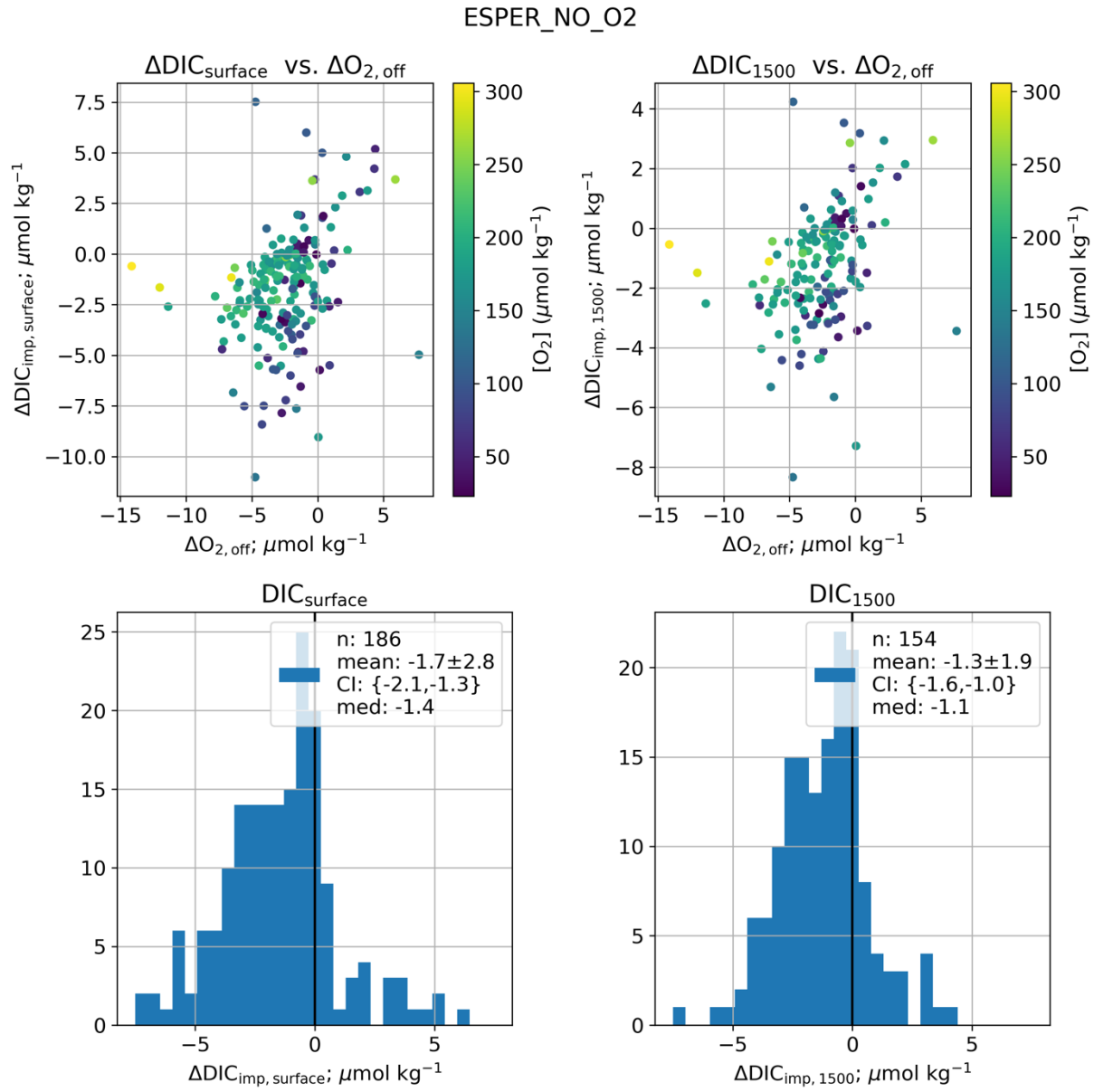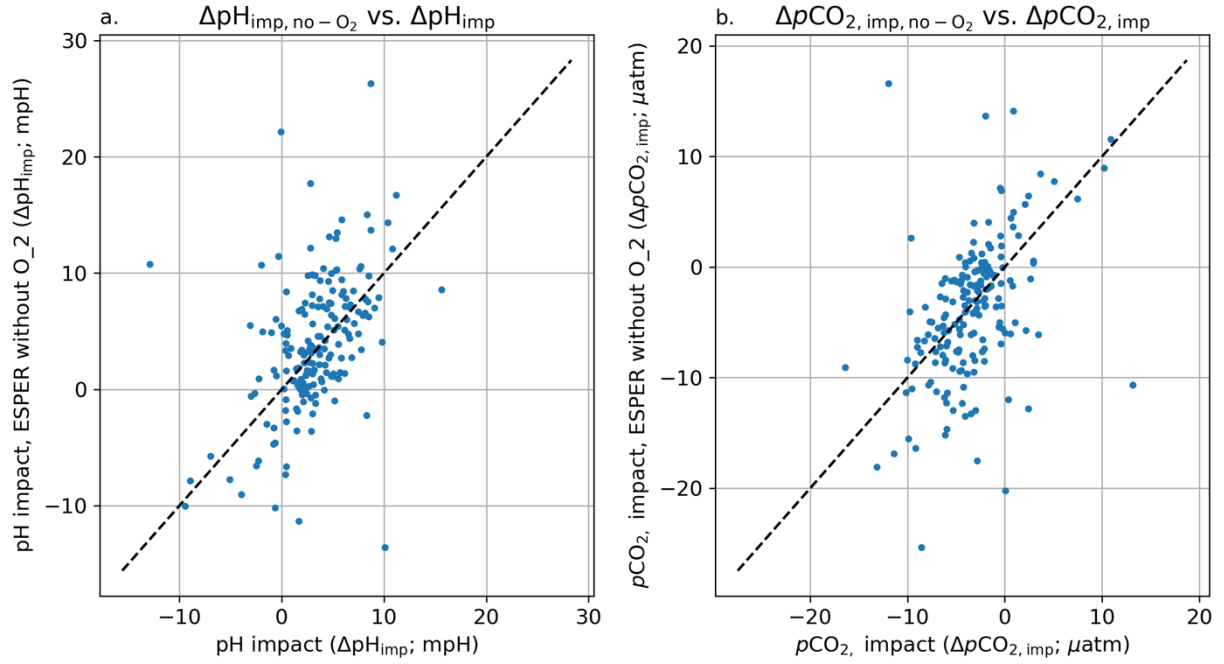| #	Calibration Comment ([SCIENTIFIC_CALIB_COMMENT] field in Sprof.nc files)	Number of WMOs w/ Comment[1]	Calibration category
1	[Blank]	500	no cal/bad
2	SVU Foil calibration coeficients were used. G determined from float measurements in air. See Johnson et al.,2015,doi:10.1175/JTECH-D-15-0101.1	302	air cal
3	Polynomial calibration coeficients were used. G determined by surface measurement comparison to WOA 2018.See Takeshita et al.2013,doi:10.1002/jgrc.20399	126	no air cal
4	DOXY_ADJUSTED is computed from an adjustment of in water PSAT or PPOX float data at surface by comparison to woaPSAT climatology or woaPPOX{woaPSAT,floatTEMP,floatPSAL} at 1 atm, DOXY_ADJUSTED_ERROR is computed from a PPOX_ERROR of 10.0 mbar +1mb/year	107	no air cal
5	Bad data; not adjustable	95	no cal/bad
6	Polynomial calibration coeficients were used. G determined from float measurements in air. See Johnson et al.,2015,doi:10.1175/JTECH-D-15-0101.1	95	air cal
7	Polynomial calibration coeficients were used. G determined by surface measurement comparison to World Ocean Atlas 2009.See Takeshita et al.2013,doi:10.1002/jgrc.20399	94	no air cal
8	DOXY_ADJUSTED is estimated from an adjustment of in water PPOX float data at surface by comparison to WOA2018 PPOX {woaPSAT, floatTEMP, floatPSAL} at 1 atm; DOXY_ADJUSTED_ERROR recomputed from a PPOX_DOXY_ERROR = 10 mbar increasing by 1 mbar per year	92	no air cal
9	DOXY_ADJUSTED is computed from an adjustment of in water PSAT or PPOX float data at surface by comparison to woaPSAT climatology or WOA PPOX in using woaPSAT and floatTEMP and PSAL at 1 atm, DOXY_ADJUSTED_ERROR is computed from a PPOX_ERROR of 10 mbar	77	no air cal
10	none	74	no cal/bad

#	Calibration Comment ([SCIENTIFIC_CALIB_COMMENT] field in Sprof.nc files)	Number of WMOs w/ Comment[1]	Calibration category
11	Polynomial calibration coeficients were used. G determined by surface measurement comparison to World Ocean Atlas 2018.See Takeshita et al.2013,doi:10.1002/jgrc.20399	73	no air cal
12	DOXY_ADJUSTED corrected using continuous in-air measurements as in Johnson et al. (2015)	71	air cal
13	DOXY_ADJUSTED corrected based on the WOA 2018 climatology as described in Johnson et al. (2015)	69	no air cal
14	SVU Foil calibration coeficients were used. G determined by surface measurement comparison to World Ocean Atlas 2018.See Takeshita et al.2013,doi:10.1002/jgrc.20399	63	no air cal
15	Percent saturation corrected as a linear function of PSAT modified from Takeshita et al (2013); PSAT converted from DOXY and DOXY_ADJUSTED converted from PSAT_ADJUSTED; oxygen corrected as a linear function of DOXY by comparison to a climatology (CARS09)	60	no air cal
16	no adjustment is performed because of issues in CTD	56	no cal/bad
17	DOXY_ADJUSTED computed using Stern-Volmer equation with coeffs refit from foil calibration data & WOD data as in Drucker & Riser (2016). The quoted error was computed via comparisons with WOA09 data, interpolated to float location, depth, and season.	47	no air cal
18	DOXY_ADJUSTED is computed from an adjustment of in water PSAT or PPOX float data at surface by comparison to woaPSAT climatology or woaPPOX{woaPSAT,floatTEMP,floatPSAL} at 1 atm, DOXY_ADJUSTED_ERROR is computed from a PPOX_ERROR of 10 mbar +1mb/year	46	no air cal
19	optode multi calibration, adjusted with median of the annual WOA climato at surface	36	no air cal
20	DOXY_QCs are modified during visual check.	33	no air cal

#	Calibration Comment ([SCIENTIFIC_CALIB_COMMENT] field in Sprof.nc files)	Number of WMOs w/ Comment[1]	Calibration category
21	Partial pressure corrected as a linear function of PPOX using in-air measurements as in https://doi.org/10.3389/fmars.2021.683207, PPOX converted from DOXY and DOXY_ADJUSTED converted from PPOX_ADJUSTED, ERROR calculated as 2*std(SLOPE)*205mbar+2mbar	31	air cal
22	1-point multiplicative correction using WOD at 1800 dbar. The quoted error was computed via comparisons with monthly or annual climatology data, interpolated to float location, depth, and season, from WOA09.	26	no air cal
23	optode multi calibration, adjusted with median of the monthly WOA climato at surface	26	no air cal
24	RT adjustment by comparison of float surface O2sat to WOA2018 surface O2sat using automatically generated gain from MBARI (jplant@mbari.org)	24	no air cal
25	RT adjustment by comparison of float surface O2sat to WOA18 surface O2sat; DOXY adjusted by gain-only factor on DOXY	22	no air cal
26	GAIN determined from WOA2013 O2sat along the five initial float cycles	21	no air cal
27	DOXY_ADJUSTED is estimated from an adjustment of in water PSAT or PPOX float data at surface by comparison to WOA PSAT climatology or WOA PPOX in using PSATWOA and TEMP and PSALfloat at 1 atm, DOXY_ADJUSTED_ERROR is estimated from a PPOX_ERROR of 10 mbar.	19	no air cal
28	not applicable	19	no cal/bad
29	Data are unadjustable owing to dissolved oxygen sensor problem.	17	no cal/bad
30	G determined by surface measurement comparison to World Ocean Atlas 2009. See Takeshita et al.2013,doi:10.1002/jgrc.20399	17	no air cal
31	RT adjustment based on extrapolation from previous DMQC adjustment; DOXY adjusted by gain-only factor on DOXY	17	no air cal
32	DOXY_ADJUSTED is estimated from an adjustment of in air PPOX float data by comparison to NCEP reanalysis,	16	air cal

#	Calibration Comment ([SCIENTIFIC_CALIB_COMMENT] field in Sprof.nc files)	Number of WMOs w/ Comment[1]	Calibration category
	DOXY_ADJUSTED_ERROR is recomputed from a PPOX_ERROR = 10 mbar		
33	Adjustment via comparison of float surface O2sat to WOA2018 surface O2sat using automatically generated gain from MBARI (jplant@mbari.org). All profiles individually inspected and compared against WOA2018 for accuracy	15	no air cal
34	Adjusted PPOX at zero for the minimum and estimate the gain with Monthly WOA2018 (OMZ)	14	no air cal
35	G determined from float measurements in air. See Johnson et al.,2015,doi:10.1175/JTECH-D-15-0101.1	13	air cal
36	Percent saturation corrected as a linear function of PSAT; Comparison to a single reference profile (isobaric match as in Takeshita et al. (2013)) on cycle 0; PSAT converted from DOXY and DOXY_ADJUSTED converted from PSAT_ADJUSTED	13	no air cal
37	Percent saturation corrected as a linear function of PSAT; Comparison to a single reference profile (isobaric match as in Takeshita et al. (2013)) on cycle 1; PSAT converted from DOXY and DOXY_ADJUSTED converted from PSAT_ADJUSTED	13	no air cal
38	Adjusted on WOA monthly climatology at surface	12	no air cal
39	Adjusted with SAGEO2 with in-air measurements (Johnson et al., 2015)	12	air cal
40	No QC available for DOXY	10	no cal/bad
41	RT adjustment by comparison of float surface O2sat to WOA13 surface O2sat; DOXY adjusted by gain-only factor on DOXY	10	no air cal
42-235	194 additional comments represented in 9 or fewer floats	267	

[1]Total number of WMOs per comment is greater than the total number of floats analyzed for this study because comments are listed per profile and some floats have multiple comments included throughout their deployment.

Table S2. Statistics for oxygen offsets between 1450 and 2000 db shown in Figure 3.

Float cal. type	Count	Mean	SD	p_value	95.0% CI low	95.0% CI high	Median	Min.	Max.
All	660	-1.9	3.9	<0.001	-2.2	-1.6	-1.9	-18.1	15.0
No air cal.	305	-1.0	3.9	<0.001	-1.4	-0.5	-0.9	-18.1	11.0
Air cal.	307	-3.0	3.4	<0.001	-3.4	-2.7	-2.7	-17.8	7.7
No cal.	48	-1.3	5.2	0.092	-2.8	0.2	-1.1	-16.7	15.0
pH - No air cal.	40	-2.7	3.8	<0.001	-3.9	-1.5	-2.4	-18.1	5.9
pH - Air cal.	207	-2.4	2.7	<0.001	-2.8	-2.0	-2.3	-14.1	7.7

All columns other than count and p value have units of μmol kg^{-1}.

Table S3. Statistics for Nitrate, pH 25C total, and DIC crossovers between 1450 and 2000 db.

Parameter	count	mean	std	p_value	95.0% CI low	95.0% CI high	median	min	max
NITRATE_ADJUSTED	264	0.2	0.4	<0.001	0.1	0.2	0.2	-2.3	1.6
pH_25C_TOTAL_ADJUSTED	106	-5.9	7.7	<0.001	-7.3	-4.4	-5.4	-40.0	14.1
DIC	148	3.6	7.0	<0.001	2.5	4.7	2.9	-15.0	56.0

All columns other than count and p value have units of μmol kg^{-1} or mpH.

Table S4. Statistics for Nitrate, pH 25C total, and DIC crossovers between 1450 and 2000 db after oxygen correction impacts have been applied.

Parameter	count	mean	std	p_value	95.0% CI low	95.0% CI high	median	min	max
NITRATE_ADJUSTED	140	-0.0	0.3	0.757	-0.1	0.0	-0.0	-1.4	1.3
pH_25C_TOTAL_ADJUSTED	82	-2.6	5.5	<0.001	-3.8	-1.4	-2.4	-20.5	11.6
DIC	114	1.9	5.9	0.001	0.8	3.0	1.6	-19.3	32.8

All columns other than count and p value have units of μmol kg^{-1} or mpH.

Table S5. Stats for pH and pCO$_2$ impact LIPHR:

	Count	Mean	SD	p_value	95.0% CI low	95.0% CI high	median	min	max
pH_impact	119	3.5	4.5	<0.001	2.7	4.3	3.4	-16.2	17.9
pCO2_impact	115	-3.6	4.7	<0.001	-4.4	-2.7	-3.6	-18.5	16.6

Table S6. Stats for NO3 impact between 1450 and 2000 db.

Parameter	count	mean	SD	p_value	95.0% CI low	95.0% CI high	median	min	max
no3_impact	110	-0.2	0.2	<0.001	-0.22	-0.14	-0.2	-0.8	0.6

Table S7. Stats for DIC impact ESPER and LIPHR

ESPER

Parameter	count	mean	SD	p_value	95.0% CI low	95.0% CI high	median	min	max
dic_impact_surf	115	-1.4	1.8	<0.001	-1.73	-1.08	-1.4	-6.6	5.9
dic_impact_1500	115	-1.0	1.2	<0.001	-1.27	-0.81	-1.0	-5.1	4.1

LIPHR

Parameter	count	mean	SD	p_value	95.0% CI low	95.0% CI high	median	min	max
dic_impact_surf	115	-1.5	2.0	<0.001	-1.86	-1.14	-1.5	-7.6	7.4
dic_impact_1500	115	-1.1	1.5	<0.001	-1.42	-0.88	-1.1	-6.0	5.1

Carter, B. R., Talley, L. D., & Dickson, A. G. (2014). Mixing and remineralization in waters detrained from the surface into Subantarctic Mode Water and Antarctic Intermediate Water in the southeastern Pacific. *Journal of Geophysical Research: Oceans*, *119*(6), 4001–4028. https://doi.org/10.1002/2013JC009355

Carter, B. R., Bittig, H. C., Fassbender, A. J., Sharp, J. D., Takeshita, Y., Xu, Y.-Y., et al. (2021). New and updated global empirical seawater property estimation routines. *Limnology and Oceanography: Methods*, *19*(12), 785–809. https://doi.org/10.1002/lom3.10461