Geophysical Research Letters

Supporting Information for

Empirical Algorithms to Estimate Water Column pH in the Southern Ocean

N. L. Williams[1*], L. W. Juranek[1], K. S. Johnson[2], R. A. Feely[3], S. C. Riser[4], L. D. Talley[5], J. L. Russell[6], J. L. Sarmiento[7], and R. Wanninkhof[8]

[1]College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, Oregon, USA
[2]Monterey Bay Aquarium Research Institute, Moss Landing, California, USA
[3]Pacific Marine Environmental Laboratory, National Oceanic and Atmospheric Administration, Seattle, Washington, USA
[4]School of Oceanography, University of Washington, Seattle, Washington, USA
[5]Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA
[6]Department of Geosciences, University of Arizona, Tucson, Arizona, USA
[7]Program in Atmospheric and Oceanic Sciences, Princeton University, Princeton, New Jersey, USA
[8]Atlantic Oceanographic and Meteorological Laboratory, National Oceanic and Atmospheric Administration, Miami, Florida, USA

*Corresponding author: Nancy L. Williams (nancy.williams@oregonstate.edu)

Contents of this file

Text S1
Figures S1 to S2
Tables S1 to S4

Text S1.

After the pH algorithms were proven useful for sensor performance evaluation and adjustment we trained two additional algorithms specifically for this purpose: pH^{Ox}_{Deep} and N^{Ox}_{Deep}. These deep algorithms were trained using bottle data south of 45° S between 1000 and 2100 m depth from the S4P 2011 and P16S 2014 repeat hydrographic lines. Excluding shallower data from these deep algorithms slightly decreased errors and bias in the residuals at 1500 m (depth where the sensor is compared to the algorithm) relative to the algorithms trained using data between 100 and 2100 m. The algorithm-predicted pH (or nitrate) is compared with the float-measured pH (or nitrate) at 1500 m for each float through time. Discrepancies in sensor pH generally occurred as slow drifts away from the algorithm-predicted pH. The sensor reference potential was adjusted using a minimum number of linear trends in the data, determined as described by *Owens* *and Wong* [2009] for the adjustment of profiling float salinity data, to force the sensor pH trend at 1500 m to match the algorithm pH trend. The adjusted reference potential determined at 1500 m was then used to compute pH on the remainder of each profile. The coefficients for these deep algorithms are found in Supplemental Tables S3 and S4. Because these deep algorithms were trained for the sole purpose of sensor adjustment they will not accurately predict pH or nitrate outside of the 1000 to 2100 m range. See *Johnson et al.* [2016] for a more detailed explanation of how MLR algorithms are used to adjust float sensor data. Any adjustments made to the SOCCOM float sensor data are documented in the QC (Quality Control) change log available at http://soccom.princeton.edu/soccomviz.php.

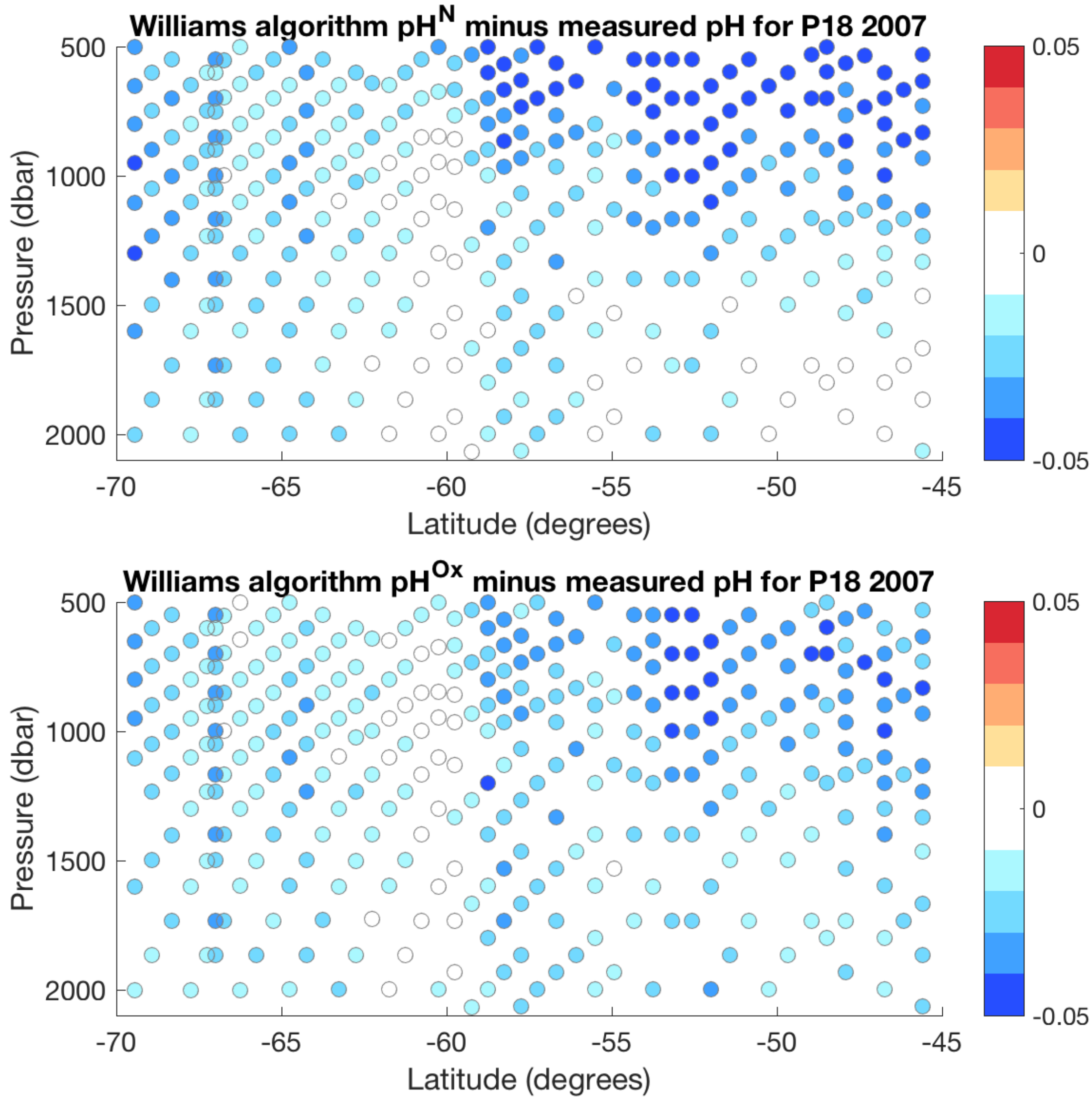

Figure S1. The difference between predicted and measured pH using the pH^{N} algorithm (top) and the pH^{Ox} algorithm (bottom) along P18S 2007, which was not used to train the algorithms.

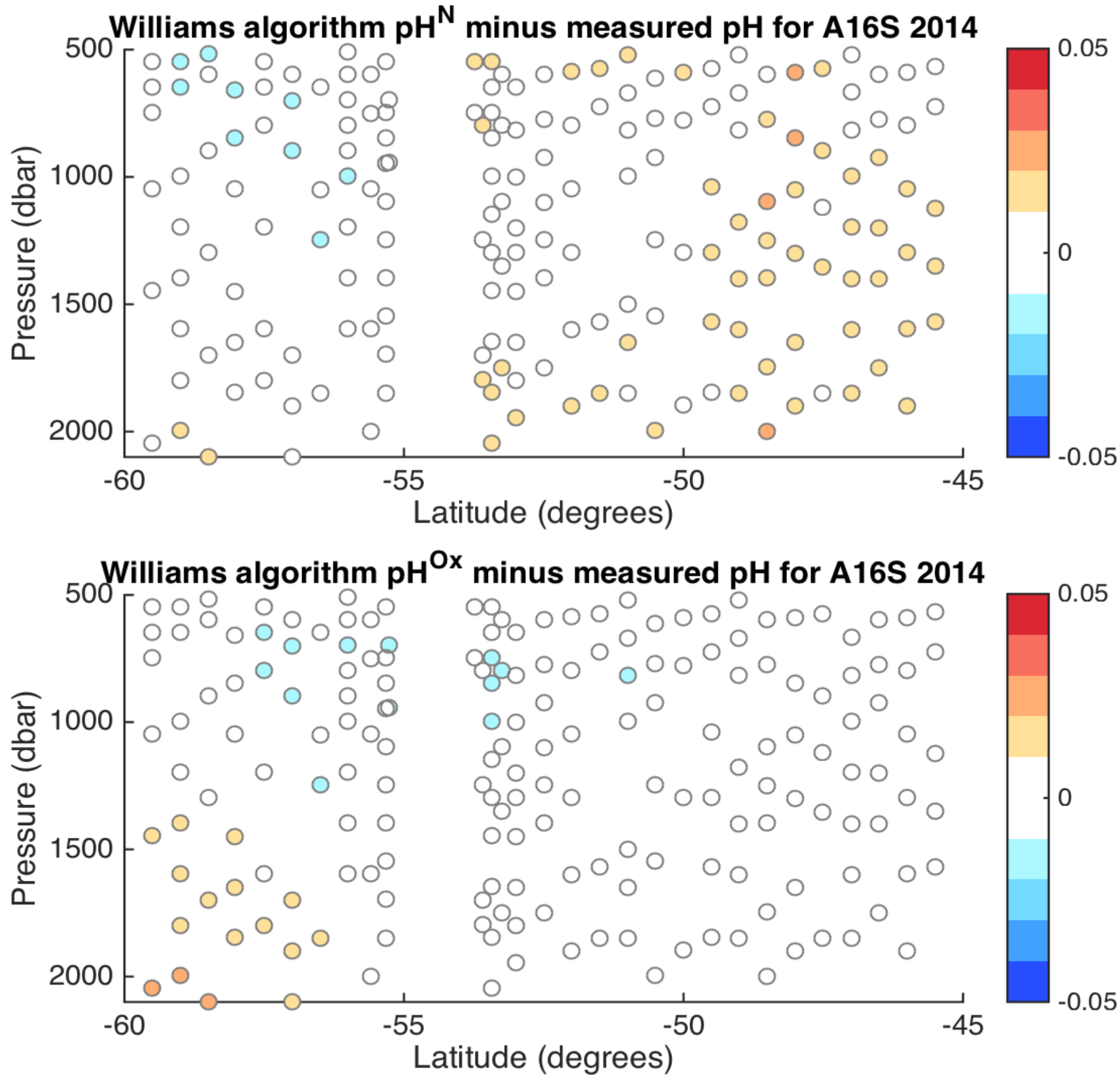

Figure S2. The difference between predicted and measured pH using the pH^N algorithm (top) and the pH^{Ox} algorithm (bottom) along GO-SHIP line A16S 2014 in the South Atlantic, which was not used to train the algorithms.

Table S1. Biogeochemical sensor information				
Parameter	Sensor Type	Accuracy	propagation to pH^N	propagation to pH^{Ox}
Nitrate	ISUS/SUNA	0.5 µmol kg^{-1}	0.004	
Oxygen	optode	2 µmol kg^{-1}		0.003
pH	Deep-Sea DuraFET	0.01		

Table S2. Summary of MLR fit coefficients and statistics[a]

Fit name	R^2	RMS Error	Parameter (units)	VIF	Coefficient ± STD Error
pH^N	0.982	0.010			$\beta_0 = 9.243 \pm 2.947 \times 10^{-2}$
			Salinity	1.7	$\beta_1 = -2.501 \times 10^{-2} \pm 9.025 \times 10^{-4}$
			Temperature (°C)	1.4	$\beta_2 = -1.691 \times 10^{-2} \pm 1.273 \times 10^{-4}$
			Pressure (dbar)	1.5	$\beta_3 = -2.235 \times 10^{-5} \pm 4.027 \times 10^{-7}$
			Nitrate (µmol kg^{-1})	1.7	$\beta_4 = -1.373 \times 10^{-2} \pm 6.724 \times 10^{-5}$
pH^{Ox}	0.965	0.008			$\beta_0 = 2.479 \pm 4.694 \times 10^{-2}$
			Salinity	1.8	$\beta_1 = 1.475 \times 10^{-1} \pm 1.324 \times 10^{-3}$
			Temperature (°C)	1.0	$\beta_2 = 2.371 \times 10^{-3} \pm 7.422 \times 10^{-5}$
			Pressure (dbar)	1.3	$\beta_3 = -3.054 \times 10^{-5} \pm 2.921 \times 10^{-7}$
			Oxygen (µmol kg^{-1})	1.7	$\beta_4 = 1.667 \times 10^{-3} \pm 7.421 \times 10^{-6}$

[a]For a sample with Salinity = 34.470, Temperature = 2.966 °C, Pressure = 1599 dbar, Nitrate = 34.22 µmol kg^{-1}, and Oxygen =180.0 µmol kg^{-1}, pH^N = 7.825 and pH^{Ox} = 7.822.

Table S3. pH^{Ox}_{Deep} coefficients and statistics[a]				
R^2	RMS Error	Parameter (units)	VIF	Coefficient ± STD Error
				$\beta_0 = 1.380 \pm 1.269 \times 10-1$
		Oxygen (μmol kg^{-1})	1.2	$\beta_1 = 1.802 \times 10^{-3} \pm 1.775 \times 10^{-5}$
0.98	0.004	Salinity	1.8	$\beta_2 = 1.786 \times 10^{-1} \pm 3.593 \times 10^{-3}$
		Temperature (°C)	2.0	$\beta_3 = 7.482 \times 10^{-3} \pm 4.242 \times 10^{-4}$
		Pressure (dbar)	1.1	$\beta_4 = -3.966 \times 10^{-5} \pm 5.756 \times 10^{-7}$

[a]This algorithm is optimized for sensor adjustment and was trained using only data between 1000 and 2100 m. This algorithm will not accurately predict pH outside of that depth range.

Table S4. N^{Ox}_{Deep} coefficients and statistics[a]				
R^2	RMS Error	Parameter (units)	VIF	Coefficient ± STD Error
				$\beta_0 = 5.44 \times 10^2 \pm 8.80$
		Sigma Theta	3.0	$\beta_1 = -1.14 \times 10 \pm 3.73$
		Oxygen (µmol kg^{-1})	1.9	$\beta_2 = -1.08 \times 10^{-1} \pm 2.14 \times 10^{-3}$
0.85	0.3	Salinity	2.5	$\beta_3 = -4.92 \pm 3.18$
		Temperature (°C)	3.0	$\beta_4 = -2.69 \pm 2.89 \times 10^{-1}$
		Pressure (dbar)	1.2	$\beta_5 = 3.14 \times 10^{-4} \pm 3.06 \times 10^{-5}$

[a]This algorithm is optimized for sensor adjustment and was trained using only data between 1000 and 2100 m. This algorithm will not accurately predict nitrate outside of that depth range.