Supplemental Materials: Evaluation of new and net community production estimates by multiple ship-based and autonomous observations in the Northeast Pacific Ocean

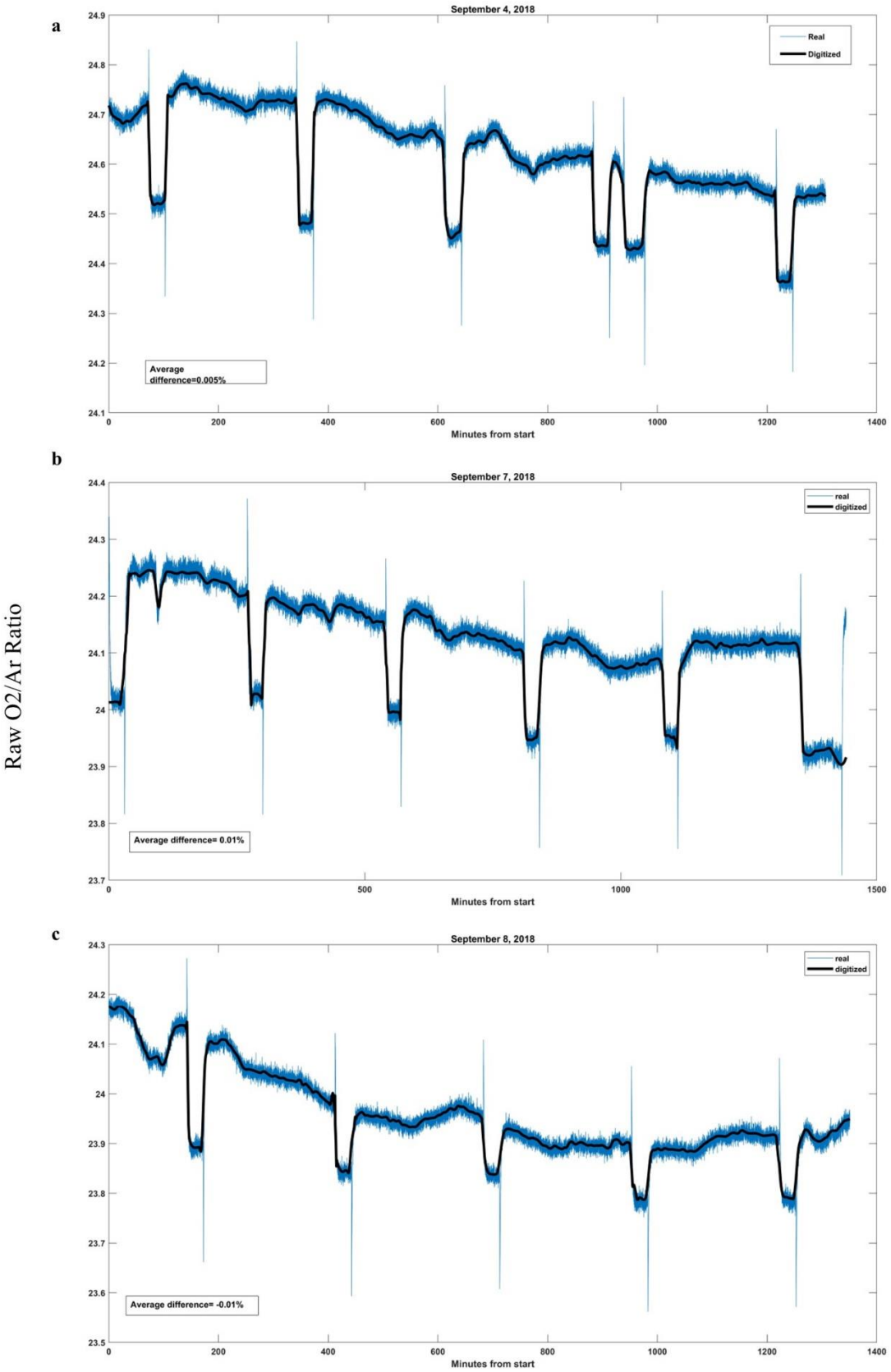
1 Methods

1.1 NCP estimates from continuous underway O_2/Ar ($NCP_{EIMS(P)}$, $NCP_{EIMS(S)}$)

Underway seawater was passed through a membrane equilibrator (3M Liqui-Cell MM Series Membrane Contractor G569 .75X1) at a flow rate of 100±5 ml/min using a magnetic gear pump (Ismatec Reglo-Z Digital). Dissolved gases in the seawater were equilibrated with headspace gas in the equilibrator cartridge and measured using a Pfeiffer Prisma 219 QMS 200 M1 quadrupole mass spectrometer. Air calibrations were performed every 3-4 hours for 30 minutes to measure the background O_2/Ar ratio in the air as a comparison point to the dissolved O_2/Ar in the seawater. EIMS data were processed using custom MATLAB pipelines (version R2019b, The MathWorks, Inc., Natick, Massachusetts, United States). Briefly, raw O_2/Ar data were paired with context data collected from the ships' underway CTD and averaged every 2 minutes. EIMS data were discarded when internal mass spectrometer pressure was out of normal operating range, when the ships' underway systems were shut off, when seawater flow to the equilibrator was below 80 or above 120 mL min^{-1}, and for the first 80 min after the instrument was restarted. Data 10 minutes after the switch from air to water and 4 minutes after the switch from water to air were also discarded to remove transition artefacts. O_2/Ar measurements collected from August 27 – 30, 2018 on the survey ship were potentially influenced by pressure buildup in the equilibrator due to inadequate cleaning and were also discarded. Due to errors with the EIMS nXT software, O_2/Ar data from the process ship taken August 14-31, 2018 were not saved. However, screenshots were taken of the graphed data on this ship every day throughout the cruise. The MATLAB script GRABIT (Jiro (2018). GRABIT (https://www.mathworks.com/matlabcentral/fileexchange/7173-grabit), MATLAB Central File Exchange. Retrieved October 2018) was used to digitize the plots for data that were not saved properly. Three plots from days where data were saved were digitized to test the accuracy of the method by comparing digitized data with real data. The average percent difference between the real data and the digitized data was calculated using MATLAB script LOWESS (Jeff Burkey (2020). LOWESS, Locally Weighted Scatterplot Smoothing for linear and non-linear data (enhanced) (https://www.mathworks.com/matlabcentral/fileexchange/22470-lowess-locally-weighted-scatterplot-smoothing-for-linear-and-non-linear-data-enhanced), MATLAB Central File Exchange. Retrieved Oct 2018) (Figure S1). The average percent difference between the digitized data and the EIMS collected data for all three days was less than or equal to 0.01%, well below the other sources of error for the O_2/Ar method. Digitized EIMS data were subsequently used for NCP analysis from the process ship covering the period of August 14-31, 2018.

Figure S1 – Agreement between real and digitized EIMS data

The blue curve in this figure shows the real O_2/Ar ratio as collected by the EIMS on the process ship on (a) September 4, 2018, (b) September 7, 2018, and (c) September 8, 2018. The black curve shows the EIMS data as digitized using the Matlab script GRABIT. The y-axis shows raw O_2/Ar ratio as read by the EIMS and the x-axis is time elapsed since the beginning of the plot in minutes.


a
September 4, 2018
Real
Digitized
Average difference=0.005%
Minutes from start
b
September 7, 2018
real
digitized
Average difference= 0.01%
Minutes from start
c
September 8, 2018
real
digitized
Average difference= -0.01%
Minutes from start
Raw O2/Ar Ratio

1.2 Discrete O_2/Ar samples

Discrete samples were collected for O_2/Ar analysis from both the underway seawater system near the EIMS and the rosette within 10 minutes of each other during rosette casts following methods outlined in Timmerman and Hamme (2021) and Emerson et al. (1999). Briefly, water was sucked through CO_2 flushed tubing into evacuated ~180-mL flasks containing dried $HgCl_2$ until half full. After weighing, the water was equilibrated with the headspace at constant temperature and removed. Headspace gases were purified through liquid nitrogen and analyzed against standards of similar composition on a MAT 253 at University of Victoria. Corrections were made for differential ionization caused by different oxygen concentrations in the sample and standards and for gases left behind in the water removed.

1.3 NCP estimates by oxygen budgets from Seaglider observations (NCP_{SG}, NCP_{G276}, NCP_{G469})

All optodes were factory calibrated prior to deployment, and oxygen values were recalculated using the modified Stern Volmer equation (Uchida *et al.*, 2018). A dynamic lag correction was applied to all glider data prior to analysis (Bittig and Körtzinger, 2017; Bittig *et al.*, 2014). For the Seaglider, the response time was adjusted to minimize the difference between ascents and dives, but the efficacy was difficult to evaluate for G276 and G469 because they only sampled on dives.

The Seaglider followed the EXPORTS campaign's Lagrangian Float (LF), allowing for cross calibration between the two assets. The two optodes carried by the LF (SBE63, Sea-bird Scientific; AA4330, Aanderaa Data Instruments AS) were calibrated during the cruise via Winkler titrations and air calibrations over the course of the deployment were used to correct for drift. The Seaglider was calibrated against the drift-corrected LF using a gain factor that was assumed to drift linearly in time since the time of deployment. The gain was calculated using the ratio of mixed layer saturation state of collocated profiles. The calculated drift was an approximately ~5% decrease per year.

Equation S1

$$m = m_{ref} + a(t + t_{ref}) = \frac{O_{2,sat}}{O_{2,sat}^{ref}}$$

where m is the gain factor, m_{ref} is the gain factor on the deployment date, t is the date, t_{ref} is the date of deployment (1 Aug 2018). To calibrate G276 and G469, oxygen saturation states in both the mixed layer and along a deep isotherm (2.9-3.1-°C, approximately 980 m) were compared against a reference asset to calculate the drifting gain factor as above, alongside a time-dependent slope and offset (b). Because of spatial heterogeneity, only data from when assets

were within 50 km of each other was used in the regressions. The Seaglider had the greatest temporal overlap with G276 and was used as a reference. Gain correction alone showed a downward drift of approximately 2% per year, while the time-dependent slope and intercept applied showed a downward drift of approximately 5% and intercept/offset increase of 280%. The corrected G276 data were used as a reference for G469, and the gain-only correction showed a downward drift of approximately 5% per year. The time-dependent slope and intercept showed a downward drift of approximately 9% and intercept/offset increase of 306%.

Equation S2

$$Corrected\ O_{2,sat}(t) = m(t) \cdot O_{2,sat} + b(t)$$

where m(t) is the slope and b(t) is the intercept. The slope and intercept regression provided the greatest agreement with the Seaglider mixed layer time series and was thus employed as a calibration. This choice was supported by the fact that the ratio of saturation states in the mixed layer decreased, but the deep ratio increased.

Uncertainties in this method include anything that is not accounted for in the mass balance equation, such as laterally advected oxygen. The measured oxygen concentration also impacts the accuracy and is sensitive to the time correction applied to the glider data. The magnitude of the diffusive and bubble gas fluxes were not well constrained. These uncertainties were estimated from the literature (Yang, Emerson and Bushinsky, 2017) and the overall error was estimated using a Monte Carlo simulation with 1000 runs based on the individual uncertainties of the input parameters (Table S1).

Table S1

Parameter	**Uncertainty**
O_2 Saturation	±0.1%
$K_{diffusive}$	±10%
$K_{complete}$ and $K_{partial}$	±25%
$F_{entrainment}$	±10%
ΔO_2:ΔC ratio	±0.1

1.4 NCP estimates from POC budgets from Wirewalker observations (NCP_{WW})

WW transmissivity was drift corrected by computing a clear water value. The median of the largest 10 points for each unique day of raw transmissivity (tr_{raw}) was computed to compute a clear water value (tr_{clear}) (personal comm. Meg Estapa). The clear water transmissivity (tr_{clear}) value was interpolated to the time-series of the profiling rate. Beam attenuation (c_p) was computed from transmissivity (tr_{raw}) with the following equations (where 0.002 is a factory blank value and 25 cm is the path length of the instrument):

Equation S3

$$tr = \frac{tr_{raw} - 0.002}{tr_{clear} - 0.002}$$

Equation S4

$$c_p = \log(tr) \times \frac{-1}{0.25}$$

The data was gridded to 1 m vertical bins computing the average and standard deviation, as well as median for each data bin.

R/V Revelle bottle particulate organic carbon bottle samples were then used to convert the WW 1-m binned median beam attenuation (c_p) to POC. Bottle samples were matched to WW measurements first finding all WW profiles within one hour of the POC bottle sample collection. Then the POC depth was used to find any WW observations at the closest depth (the minimum absolute value difference for bottle depth to WW depth). With a WW cast index and depth index corresponding to each POC sample, the average and standard deviation beam attenuation was computed for each bottle correspondent (i.e. if there are more than one cast this would be the average of the multiple cast c_p at that depth). The average latitude and longitude was also computed for each bottle-WW match and this was used to compute the distance between the matches with gsw_distance. POC bottles to WW c_p were linearly fit (WW c_p (1/m) = 0.027 * Bottle POC (umol/L) -0.009). Any distance greater than 5 km between platforms was removed from the fitting process.

1.5 NCP estimates from the DIC, NO_3^- and O_2 budgets from BGC-float observations ($NCP_{BGC(O2)}$, $NCP_{BGC(DIC)}$, NP_{BGC})

Additional methods for the BGC-Float are outlined in Huang et al. (2022).

1.6 NP estimates from ^{15}N Incubations (NP_{Inc})

Additional methods for the ^{15}N Incubations are outlined in Meyer et al. (2022).

1.7 NCP estimates from simulated in situ incubations (NCP_{Inc})

Additional methods for the dilution incubations are outlined in McNair *et al.* (2021).

2 Results

2.1 $NCP_{EIMS(P)}$, $NCP_{EIMS(S)}$

Respiration in the underway systems of ships by bacteria lining the pipe walls is a persistent problem in the research fleet (Juranek *et al.*, 2010). Discrete O_2/Ar samples collected from the ships' underway systems and rosette bottles indicate respiration in the lines of both ships, with an offset of -0.88±0.49% and -0.62±0.75% saturation (Figures S2 and S3) for the process and survey ship, respectively. EIMS values were corrected for these offsets before calculating NCP. The largest source of uncertainty in the EIMS-based NCP estimates is derived from uncertainty in gas transfer velocity estimated at ±20% (Bender *et al.*, 2011; Wanninkhof, 2014).

Figure S2 – O_2/Ar on the R/V *Roger Revelle* by EIMS and discrete sampling

The blue curve represents uncorrected O_2/Ar saturation percent from the EIMS. The red curve represents corrected EIMS values. The black dots are O_2/Ar samples from the underway system and the green points are O_2/Ar samples from the CTD.

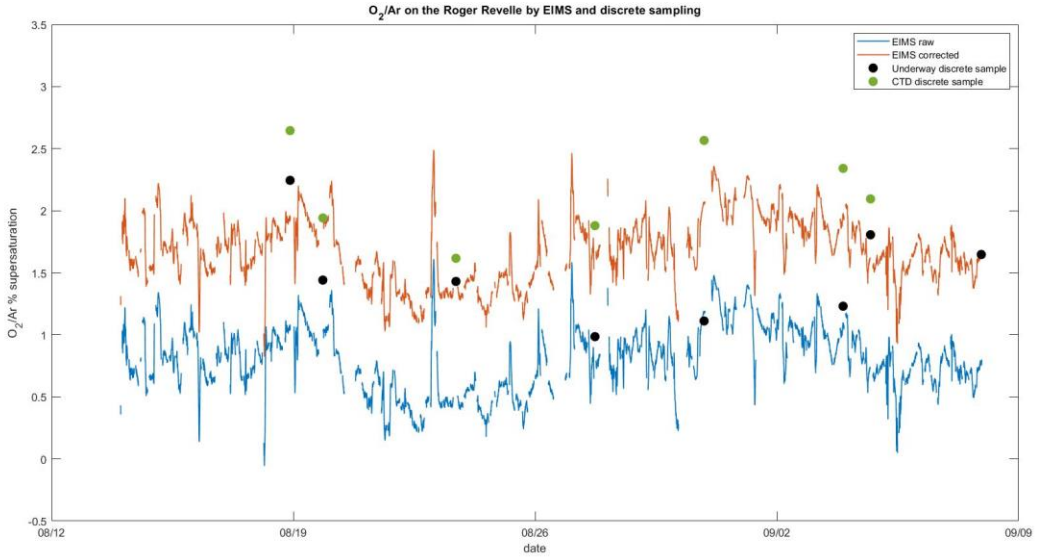
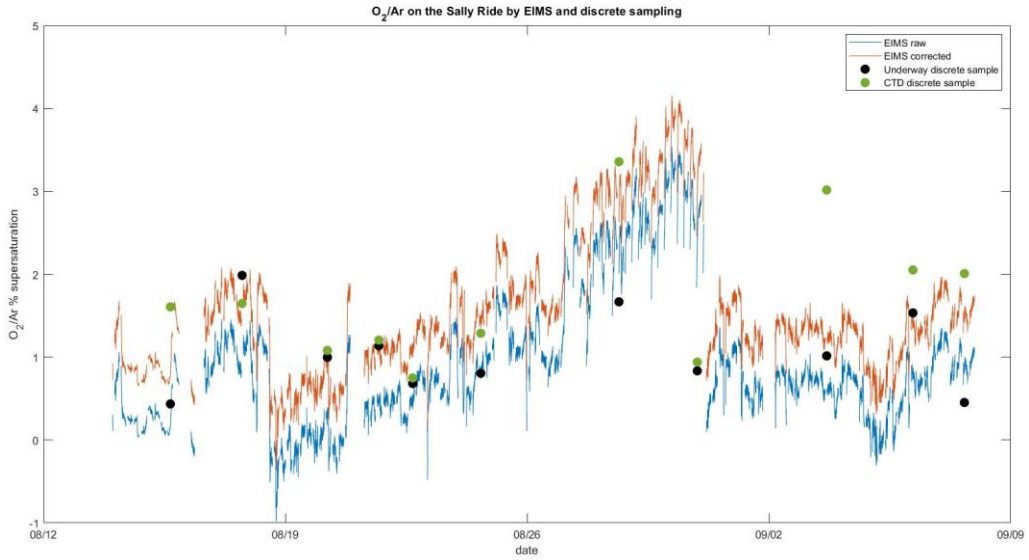


Figure S3 – O_2/Ar on the R/V *Sally Ride* by EIMS and discrete sampling



The blue curve represents uncorrected O_2/Ar saturation percent from the EIMS. The red curve represents corrected EIMS values. The black dots are O_2/Ar samples from the underway system and the green points are O_2/Ar samples from the CTD.

2.2 NCP_{WW}

The WW profiled for a total of 16 days over two deployments (Epochs 2 and 3). During each deployment, only 3 days that passed fitting criteria were used to compute diel GPP and CR. However, aside from days with large negative results, the additional 10 days were in general good agreement with the 6 days that passed fitting criteria. Uncertainties obtained from the (Barone *et al.*, 2019) fitting procedure were much smaller than from bootstrapping WW sampling to 6 hours profiles. When subsampling, there was no statistical difference between GPP and CR, meaning NCP was zero. However, an increasing magnitude of POC over the course of Epoch 2 paired with the slightly positive NCP, suggests that there was net growth during this time period.

2.3 $NCP_{BGC(O2)}$, $NCP_{BGC(DIC)}$, NP_{BGC}

Further information about the results from the BGC-Float can be found in Huang et al. submitted.

2.4 NP_{Inc}

In converting NO_3^- uptake-based new production estimates to NCP as proposed by Dugdale and Goering (1967), we assume steady state where the input of dissolved NO_3^- into the mixed layer and subsequent uptake by phytoplankton is balanced by the organic carbon export (Eppley and Peterson, 1979) from the mixed layer in the absence of nitrification. This assumption is supported by negligible nitrification in the mixed layer (A. Santoro, personal communication). Further information about the results of NP_{Inc} can be found in Meyer et al. submitted.

2.5 NCP_{Inc}

Further information about the results from the BGC-Float can be found in McNair et al. 2021.

2.6 Context data

On average, the mixed layer deepened over the course of the experiment (Figure S3). Additional data about the environmental conditions around Station P during the EXPORTS 2018 Field Deployment can be found in Siegel *et al.* (2021).

2.7 GAMM sensitivity testing

The base GAMM we built for this study is represented in Equation 14a. From the base model we added a number of additional elements and performed a number of sensitivity tests to determine if they improved the model fit. First, we added a temporal component to the base model following Equation 15 in the text (rewritten below for convenience):

Equation 14b

$$g(NCP\ or\ NP_i) = \alpha_i + \beta_i(type_i) + f(day_i) + \varepsilon_i$$

From this equation, we built two additional models, adding a term to remove autocorrelation between time t and t+1, and time t and t+2 (tAR1 and AR2, respectively). We then performed an ANOVA comparing the models that accounted for temporal autocorrelation with the model built in Equation 15 and found no evidence that removing temporal autocorrelation improved the model fit (p= 0.14,0.69 for tAR 1 and tAR2, respectively).

Next, we added a spatial component to the base model following Equation 16 in the text (rewritten below for convenience):

Equation 14b

$$g(NCP\ or\ NP_i) = \alpha_i + \beta_i(type_i) + f(distance_i) + \varepsilon_i$$

From this equation, we built two additional models, adding a term to remove autocorrelation between time t and t+1, and time t and t+2, (sAR1 and sAR2, respectively). We then performed an ANOVA comparing the models that accounted for temporal autocorrelation with the model built in Equation 15 and found no evidence that removing temporal autocorrelation improved the model fit (p= 0.08,0.87 for sAR1 and sAR2 respectively).

Next, we built a model to account for the effects of both space and time on the NCP or NP outcome.

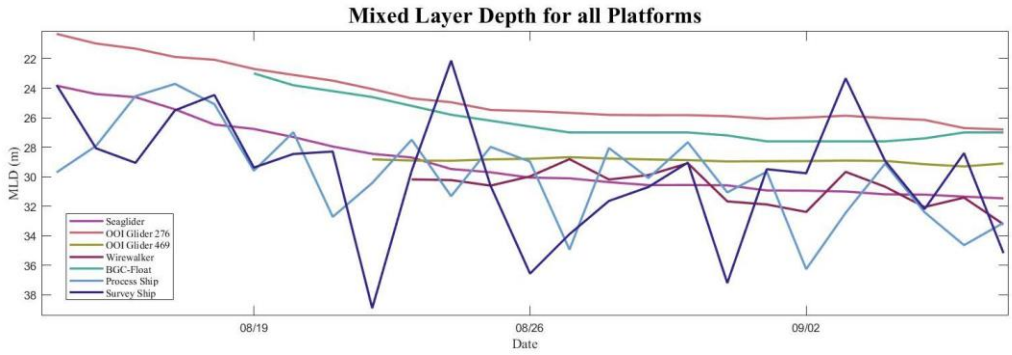
Equation 14d

$$g(NCP\ or\ NP_i) = \alpha_i + \beta_i(type_i) + f(day_i) + f(distance_i) + \varepsilon_i$$

The model presented in Equation 14d was compared to the base model using ANOVA and was shown to not be significantly different (p=0.22). We also tested models from Equation 14b and 14c without any autocorrelation structures against the base model using ANOVA. These were also not significantly different from the base model (p=0.44, 0.41 respectively). Following these sensitivity tests, only the model presented in the manuscript was significantly different from the base model.

Figure S4
Evolution of the mixed layer depth by platform from August 14 - September 7, 2018

This figure shows the MLD from each platform over the course of the cruise.



Supplementary References

Barone, B. *et al.* (2019) 'The estimation of gross oxygen production and community respiration from autonomous time-series measurements in the oligotrophic ocean', *Limnology and Oceanography: Methods*, 17(12), pp. 650–664. doi:10.1002/lom3.10340.

Bender, M.L. *et al.* (2011) 'Evaluating gas transfer velocity parameterizations using upper ocean radon distributions', *Journal of Geophysical Research: Oceans*, 116(C2). doi:10.1029/2009JC005805.

Bittig, H.C., Fiedler, B., Scholz, R., Krahmann, G. and Körtzinger, A., 2014. Time response of oxygen optodes on profiling platforms and its dependence on flow speed and temperature. *Limnology and Oceanography: Methods*, *12*(8), pp.617-636.

Bittig, H.C. and Körtzinger, A. (2017) 'Technical note: Update on response times, in-air measurements, and in situ drift for oxygen optodes on profiling platforms', *Ocean Science*, 13(1), pp. 1–11. doi:10.5194/os-13-1-2017.

Dugdale, R.C. and Goering, J.J., 1967. Uptake of new and regenerated forms of nitrogen in primary productivity 1. *Limnology and oceanography*, *12*(2), pp.196-206.

Emerson, S., Stump, C., Wilbur, D. and Quay, P., 1999. Accurate measurement of O2, N2, and Ar gases in water and the solubility of N2. *Marine Chemistry*, *64*(4), pp.337-347.

Eppley, R.W. and Peterson, B.J. (1979) 'Particulate organic matter flux and planktonic new production in the deep ocean', *Nature*, 282(5740), p. 677. doi:10.1038/282677a0.

Juranek, L.W. *et al.* (2010) 'Evidence of O2 consumption in underway seawater lines: Implications for air-sea O2 and CO2 fluxes', *Geophysical Research Letters*, 37(1). doi:10.1029/2009GL040423.

McNair, H.M. *et al.* (2021) 'Microzooplankton grazing constrains pathways of carbon export in the subarctic North Pacific', *Limnology and Oceanography*, 66(7), pp. 2697–2711. doi:10.1002/lno.11783.

Siegel, D.A. *et al.* (2021) 'An operational overview of the EXport Processes in the Ocean from RemoTe Sensing (EXPORTS) Northeast Pacific field deployment', *Elementa: Science of the Anthropocene*, 9(1). doi:10.1525/elementa.2020.00107.

Timmerman, A.H.V. and Hamme, R.C. (2021) 'Consistent Relationships Among Productivity Rate Methods in the NE Subarctic Pacific', *Global Biogeochemical Cycles*, 35(2), p. e2020GB006721. doi:10.1029/2020GB006721.

Uchida, T. *et al.* (2018) 'Eddy Iron Fluxes and Its Impact on Primary Production in the Southern Ocean', 2018, pp. OS34B-07.

Wanninkhof, R. (2014) 'Relationship between wind speed and gas exchange over the ocean revisited', *Limnology and Oceanography: Methods*, 12(6), pp. 351–362. doi:10.4319/lom.2014.12.351.

Yang, B., Emerson, S.R. and Bushinsky, S.M. (2017) 'Annual net community production in the subtropical Pacific Ocean from in situ oxygen measurements on profiling floats', *Global Biogeochemical Cycles*, 31(4), pp. 728–744. doi:10.1002/2016GB005545.