Evaluation of the Southern Ocean O_2/Ar-based NCP estimates in a model framework

Bror F. Jonsson,[1] Scott C. Doney,[2] John Dunne,[3] and Michael Bender[1]

[1] The sea-air biological O_2 flux assessed from measurements of surface O_2 supersaturation in excess of Ar supersaturation ("O_2 bioflux") is increasingly being used to constrain net community production (NCP) in the upper ocean mixed layer. In making these calculations, one generally assumes that NCP is at steady state, mixed layer depth is constant, and there is no O_2 exchange across the base of the mixed layer. The object of this paper is to evaluate the magnitude of errors introduced by violations of these assumptions. Therefore, we examine the differences between the sea-air biological O_2 flux and NCP in the Southern Ocean mixed layer as calculated using two ocean biogeochemistry general circulation models. In this approach, NCP is considered a known entity in the prognostic model, whereas O_2 bioflux is estimated using the model-predicted O_2/Ar ratio to compute the mixed layer biological O_2 saturation and the gas transfer velocity to calculate flux. We find that the simulated biological O_2 flux gives an accurate picture of the regional-scale patterns and trends in model NCP. However, on local scales, violations of the assumptions behind the O_2/Ar method lead to significant, non-uniform differences between model NCP and biological O_2 flux. These errors arise from two main sources. First, venting of biological O_2 to the atmosphere can be misaligned from NCP in both time and space. Second, vertical fluxes of oxygen across the base of the mixed layer complicate the relationship between NCP and the biological O_2 flux. Our calculations show that low values of O_2 bioflux correctly register that NCP is also low ($<10\,\mathrm{mmol\,m^{-2}\,day^{-1}}$), but fractional errors are large when rates are this low. Values between 10 and $40\,\mathrm{mmol\,m^{-2}\,day^{-1}}$ in areas with intermediate mixed layer depths of 30 to 50 m have the smallest absolute and relative errors. Areas with O_2 bioflux higher than $30\,\mathrm{mmol\,m^{-2}\,day^{-1}}$ and mixed layers deeper than 40 m tend to underestimate NCP by up to $20\,\mathrm{mmol\,m^{-2}\,day^{-1}}$. Excluding time periods when mixed layer biological O_2 is undersaturated, O_2 bioflux underestimates time-averaged NCP by 5%–15%. If these time periods are included, O_2 bioflux underestimates mixed layer NCP by 20%–35% in the Southern Ocean. The higher error estimate is relevant if one wants to estimate seasonal NCP since a significant amount of biological production takes place when mixed layer biological O_2 is undersaturated.

1. Introduction

[2] Surface ocean ecosystems and their associated carbon fluxes are a fundamental part of the global carbon cycle. Understanding these fluxes, however, is a daunting challenge since marine ecosystems are governed by complex interactions among living organisms, nutrients, and physical properties. This complexity drives a need for estimates of biological production with high temporal and spatial coverage. To address such needs, various algorithms using remotely sensed proxies for phytoplankton biomass have been developed to estimate primary production on seasonal timescales with high spatial resolutions [*Behrenfeld and Falkowski*, 1997; *Carr et al.*, 2006; *Westberry et al.*, 2008].

[3] Also critical to the carbon cycle is the ability to map the distribution of net community production (NCP), a term that represents the rate of photosynthesis minus the rate of respiration [*Wiliams*, 1993]. NCP can be constrained by observations in a number of ways, typically the time evolution of carbon, nutrients, or oxygen concentrations. Such methods include the direct tracking of organic particle stocks [*Langdon and Marra*, 1995; *Marra et al.*, 1993] and incubations in light

[1]Department of Geosciences, Guyot Hall, Princeton University, Princeton, New Jersey, USA.
[2]Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA.
[3]Geophysical Fluid Dynamics Laboratory (GFDL), Princeton, New Jersey, USA.

Corresponding author: B. F. Jonsson, Princeton University, Guyot Hall, Princeton, NJ, 08544-1004, USA. (bjonsson@princeton.edu)

bottles [*Williams and Purdie*, 1991]. A promising method to constrain the NCP range in a system and give insight into controlling processes is Lagrangian particle tracking of satellite Chl [*Joensson et al.*, 2011]. Finally, simultaneous measurement of oxygen and argon supersaturation [*Craig and Hayward*, 1987; *Kaiser et al.*, 2005] (hereafter referred to as $\Delta O_2/Ar$ method) can be used as a constraint of NCP. Each of these methods has made important contributions to our understanding of upper ocean carbon fluxes, but each also has serious limitations. As a consequence of such issues, all measurements of NCP have substantial uncertainties but nevertheless are accurate enough to provide a foundation for our understanding of variations of this property in time and space, and of the mediating processes.

[4] This paper focuses on the $\Delta O_2/Ar$ method, which is based on the fact that positive net community production tends toward surface water O_2 supersaturation. In an ideal steady-state system, the degree of O_2 supersaturation is balanced by a flux of O_2 across the sea-air interface without influence from the subsurface, making it possible to estimate NCP only using O_2 supersaturation and gas transfer velocity. This relationship is, in reality, much more complicated since O_2 supersaturation also depends on factors such as the local atmospheric pressure, bubble entrainment, changes in water temperature [*Spitzer and Jenkins*, 1989], and ventilation of the low-O_2 subsurface. With the exception of ventilation, it is possible to compensate for these physical processes by referencing to argon supersaturation [e.g., *Craig and Hayward*, 1987; *Spitzer and Jenkins*, 1989]. The flux of oxygen across the sea-air interface can then be calculated from the combined O_2 and Ar supersaturations and the gas transfer velocity. We will call this property "O_2 bioflux" to distinguish it from the total sea-air exchange of O_2. The $\Delta O_2/Ar$ method is now in broad use as a way to estimate NCP, partly because the method can be used on seawater sampled underway without the need for dedicated ship time or extensive attention during the cruise [*Cassar et al.*, 2009; *Kaiser et al.*, 2005].

[5] In recent years, a number of studies have estimated NCP from O_2/Ar measurements and used the results to understand factors controlling this property. Outside of the Southern Ocean, this work includes studies at the Hawaii Ocean Time Series and the Bermuda Atlantic Time Series study sites [e.g., *Nicholson et al.*, 2012, and references therein]. This work documented the seasonal cycle of net community production, annual carbon fluxes, and the relation to vertical mixing in the water column. Work in the subantarctic North Atlantic [*Quay et al.*, 2012] also documented the seasonal cycle and annual carbon fluxes. Continuous underway studies in the equatorial Pacific [*Stanley et al.*, 2010] showed that, in the western half of the basin, NCP was low and the zonal gradient was small. In the Southern Ocean, *Hendricks et al.* [2004] and *Reuer et al.* [2007] measured O_2/Ar in discrete samples collected on cruises and showed that NCP was lower in the northern reaches of the Southern Ocean than in the center of the basin and that one could have very high NCP where diatoms were absent due to silica limitation. *Huang et al.* [2012] measured NCP in discrete samples in the western Antarctic Peninsula region and found a very strong link to irradiance and chlorophyll.

[6] Analytical uncertainties, and uncertainties in the gas transfer velocity, lead to errors in bioflux. The measurement precision of the O_2/Ar ratio is 0.1%–0.2% [*Cassar et al.*, 2009]. The corresponding analytical error in NCP, for temperatures ranging from 0°C to 30°C and gas transfer velocity of $3\,\mathrm{m\,day^{-1}}$, is only $1–2\,\mathrm{mmol\,m^{-2}\,day^{-1}}$. The error associated with the gas transfer velocity is harder to quantify. The root mean square difference between gas transfer velocity inferred from ^{222}Rn measurements and values estimated by weight-averaging recent wind speeds [*Reuer et al.*, 2007; *Bender et al.*, 2011] is $\pm40\%$ [*Bender et al.*, 2011]. However, this number overestimates errors in the parameterizations because it includes errors due to uncertainties in wind speeds reconstructed for the 1970s, as well as errors in the ^{222}Rn estimates themselves. Short-term experiments have tended to give gas transfer velocities generally agreeing with the estimate derived from the global bomb ^{14}C budget [*Sweeney et al.*, 2007; *Ho et al.*, 2011], and average ^{222}Rn estimates from the GEOSECS program are in excellent agreement with these parameterizations [*Bender et al.*, 2011].

[7] As generally applied, $\Delta O_2/Ar$ data constrain NCP in the mixed layer based on several assumptions: constant NCP, constant mixed layer depth, and no O_2 exchange across the base of the mixed layer. The aim of this study is to estimate the uncertainty in calculated values of NCP due to violations of these assumptions. Our strategy is to use the output from two general circulation models with included ocean biogeochemistry and sample them in an analogous fashion to real-life observations. This approach can be classified as an "observing system simulation experiment." We consider NCP to be a known entity from the models and derive the $\Delta O_2/Ar$ ratio from the models' O_2 and Ar concentrations. The sea-air biological O_2 flux (O_2 bioflux) can then be calculated using the model's surface winds for the gas transfer velocity. The difference between known NCP from the model and our calculated O_2 bioflux gives an estimate of how well the $\Delta O_2/Ar$ method can estimate NCP. We note that comparing model NCP and O_2 bioflux becomes more complicated if substantial amounts of oxygen are transported across the base of the mixed layer. Such fluxes can arise locally from seasonal entrainment of supersaturated or undersaturated waters from the seasonal thermocline or from large-scale upwelling of undersaturated subsurface waters. The models are still fully internally consistent, but discrepancies between O_2 bioflux and NCP now depend on model-specific descriptions of oxygen transport.

[8] In this study, we will primarily focus on the Southern Ocean poleward of 40°S. The Southern Ocean is nitrogen and phosphorus replete, silicate depleted in its northern part in the summer but sufficient elsewhere, and iron limited over most of its area. There are a number of reasons to focus on this basin. Upper ocean carbon fluxes in the Southern Ocean shape nutrient fluxes to surface waters of much of the extra-polar ocean through the mixing of southern-sourced mode and intermediate waters into the euphotic zone [*Sarmiento and Gruber*, 2006]. These fluxes also influence the sea surface pCO₂ field in the Southern Ocean, and hence the rate of CO_2 uptake in this critical region. Science and supply ships ply the Southern Ocean in the spring and summer, offering unparalleled resources for biogeochemical studies of the mixed layer. However, our calculations are based on global models, and we will extend our approach to other areas in the future.

[9] To examine differences between O$_2$ bioflux and NCP due to time dependence, we start by using a simple box model to explore the temporal evolution of ΔO$_2$/Ar and O$_2$ bioflux. We will then turn to two full-ocean biogeochemistry general circulation models, the Geophysical Fluid Dynamics Laboratory's Tracers Of Phytoplankton with Allometric Zooplankton version 1 (TOPAZ) [*Dunne et al.*, 2010] and the National Center for Atmospheric Research (NCAR) Community Climate System Model with BioGeochemistry (BGCCSM) [*Doney et al.*, 2009a; *Moore and Doney*, 2004], in order to assess the errors in O$_2$ bioflux as a measure of NCP when the full range of processes included in these models is invoked.

2. Methods

2.1. Calculations of Biological Oxygen Flux and Definitions of Terms

[10] Oxygen bioflux is calculated by assuming that O$_2$ supersaturation in the surface ocean has two components, a physical component due to local atmospheric pressure, temperature changes, and bubble entrainment [e.g., *Spitzer and Jenkins*, 1989], and a biological component due to recent net community production in the mixed layer. The physical component is removed by dividing Ar supersaturation from O$_2$ supersaturation, leaving "biological O$_2$ supersaturation":

$$\Delta(O_2/Ar) = \frac{\Delta(O_2/Ar)_{sample}}{\Delta(O_2/Ar)_{sat}} - 1 \ , \tag{1}$$

where $(O_2/Ar)_{sample}$ is the measured seawater concentration ratio, $(O_2/Ar)_{sat}$ is the saturation concentration ratio, and $\Delta(O_2/Ar)$ is the biological O$_2$ supersaturation, normally expressed in percent. "O$_2$ bioflux" (mmol m^{-2} day^{-1}) is then calculated as

$$O_2\text{bioflux} \ = \ \Delta(O_2/Ar) \cdot k \cdot [O_2]_{sat} \cdot \rho, \tag{2}$$

where k is the gas transfer velocity, $[O_2]_{sat}$ is equilibrium concentration of O$_2$, and ρ is the density in the surface water [*Reuer et al.*, 2007].

[11] O$_2$ bioflux itself actually has two components. The first, which we want to access, is recent net community production in the mixed layer. The second is associated with O$_2$ being generated or consumed just below the base of the mixed layer and subsequently influencing the mixed layer supersaturation. This added perturbation due to O$_2$ transport complicates the use of O$_2$ bioflux to characterize NCP, and this paper will evaluate the magnitude of this effect. Because negative ΔO$_2$/Ar can be associated with upwelling of undersaturated waters as well as with net respiration, we calculate O$_2$ bioflux only when biological O$_2$ is supersaturated in the mixed layer, similar to the practice commonly used to evaluate field O$_2$/Ar data.

[12] We calculate gas transfer velocities using the same parameterization as the models, but with a 60 day weighted average based on each day's proximity in time to the sampling day, and the fraction of the mixed layer flushed in each subsequent interval until sampling [*Bender et al.*, 2011; *Reuer et al.*, 2007]. This weighting is similar to an exponential decay with an e-folding time, or effective timescale, of about 10 days for representative winds and mixed layer depths. The 60-day time period is long enough to allow

for almost all gas in the mixed layer to have exchanged with the atmosphere. This approach is generally adopted by studies where the ΔO$_2$/Ar method is used [e.g., *Cassar et al.*, 2011; *Hamme et al.*, 2012]

[13] Since the ΔO$_2$/Ar method is based on the inventory of O$_2$ in the mixed layer, it is in effect a proxy for NCP time averaged over the gas exchange timescale. This timescale t can, in a slowly changing system, be estimated from the residence time of mixed layer O$_2$; assuming no mixing with deeper waters:

$$\tau = \frac{\text{MLD}}{k} \tag{3}$$

where MLD is the mixed layer depth and k is the gas transfer velocity. The black lines in Figure 1 show that τ depends on MLD and wind speed in a nonlinear fashion. Plausible combinations of MLD and wind speed can generate residence times from 1 to 100 days. The green and red points in Figure 1 represent a random subset of Southern Ocean MLD and wind speed combinations from the two models. Long residence times in Figure 1 are unrealistic because they are calculated from events with low winds that are not sustained for the extent of the predicted residence times. Based on a Southern Ocean mean τ of 9.2 days in BGCCSM and 11.1 days in TOPAZ, we have chosen 10 days as a value for the period over which O$_2$ bioflux represents NCP. BGCCSM value is comparable to the 10-day e-folding timescale of the earlier described 60-day weighting of gas transfer velocities. We therefore compare O$_2$ bioflux with NCP integrated over the mixed layer and averaged over the 10 days prior to sampling. This term is NCP$_{10day}$.

2.2. Box Model

[14] We use a one-box model to explore the time evolution of ΔO$_2$/Ar and biological O$_2$ fluxes in an idealized surface ocean. The model is simplified by assuming that the contribution from physical processes to O$_2$ supersaturation is zero and [Ar] = [Ar]$_{sat}$. This means that O$_2$ bioflux
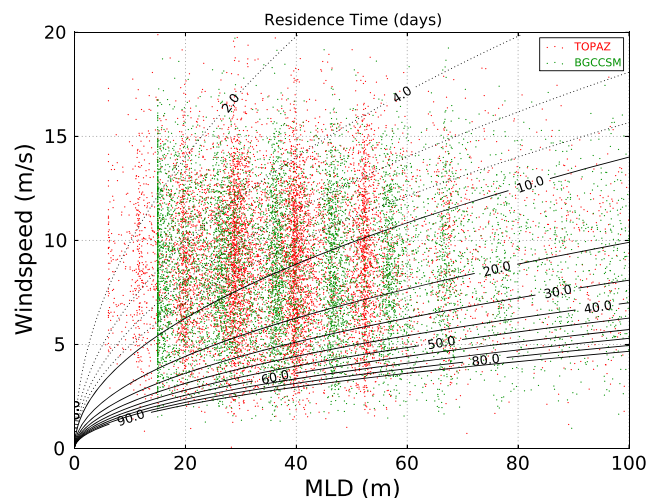

Figure 1. Residence time of oxygen in the mixed layer, based on mixed layer depth and wind speed, for random model grid points where O$_2$/Ar > 0 from TOPAZ (red) and BGCCSM (green) dots, respectively. Black isolines show different residence times in days.

is identical to the total sea-air flux of O$_2$. We keep the O$_2$ bioflux connotation for simplicity in comparisons with other results, even if the distinction between O$_2$ bioflux and the O$_2$ sea-air flux is in this case unnecessary. The model is governed by the following relationships:

$$O_2\text{bioflux} \;=\; k\big([O_2] - [O_2]_{sat}\big) \qquad (4)$$

$$\frac{\delta O_2}{\delta t} = \text{NCP} - O_2\text{bioflux}. \qquad (5)$$

[15] The source of O$_2$ is net community production, and the sink is gas exchange. Salinity is set to 33 psu and temperature to 5°C. Wind speed, mixed layer depth, and the time course of NCP are all prescribed as illustrated in the relevant figures. Gas transfer velocity (k) in the one-box model is calculated from the wind speed parameterization of *Wanninkhof* [1992]. The simulations are integrated with a 1-h time step for 365 days. The model calculates the supersaturation of O$_2$, which in this case is analogous to ΔO_2/Ar since no physical processes are included. We use the resulting time series of O$_2$ supersaturation to calculate the sea-air biological O$_2$ flux (i.e., O$_2$ bioflux), which we finally compare with the prescribed mixed layer NCP.

2.3. Biogeochemical Global Circulation Models

[16] The TOPAZ biogeochemical model [*Dunne et al.*, 2005] uses the method of moments (MOM)-4 Modular Ocean Model [*Griffies et al.*, 2005] with a 1° horizontal B grid with higher meridional resolution near the equator (to 1/3°). The vertical distribution of grid cells consists of 50 layers in z coordinates with the first 22 layers from the surface each being 10-m thick. MOM-4 includes representation of the k-profile parameterization (KPP) mixed layer diffusivity scheme [*Large et al.*, 1994], Bryan-Lewis deeper vertical mixing, neutral physics (re-orientation of diffusivities along isopycnal surfaces), Gent-McWilliams isopycnal thickness diffusion [*Gent and McWilliams*, 1990], bottom topography represented with partial cells, isotropic and anisotropic friction, and a multiple-dimensional flux limiting tracer advection scheme using the third-order Sweby flux limiter. Wind stress, freshwater flux, and incoming shortwave and longwave radiation are prescribed as boundary conditions from the reanalysis effort of the ECMWF and NCAR Common Ocean-ice Reference Experiments (CORE). TOPAZ considers the cycling of carbon, nitrogen, phosphorus, oxygen, iron, silicon, alkalinity, and lithogenic material through the phytoplankton community; zooplankton grazing and detritus cycling are implicit.

[17] Community Climate System Model with BioGeochemistry (BGCCSM) is based on the Los Alamos National Laboratory Parallel Ocean Program [*Smith and Gent*, 2004]. In our application, the grid is symmetric in the Southern Hemisphere with a zonal resolution of 3.6°. Meridional resolution decreases from 1.8° at mid-latitudes to about 0.8° at high and low latitudes. The surface layer is 12-m thick; there are in total five layers to 111 m, and 25 layers to the bottom. This model also invokes the Gent-McWilliams isopycnal mixing and the KPP upper ocean model and is forced by CORE boundary conditions. The ecosystem module [*Moore and Doney*, 2004] is linked with an ocean biogeochemistry module [*Doney et al.*, 2009b] with full oxygen and carbonate system thermodynamics.

[18] The air-sea fluxes of O$_2$ and CO$_2$ in both models are computed using prescribed atmospheric conditions (surface pressure, mole fraction), model-predicted surface water concentrations, NCEP surface winds, and the quadratic dependence of the gas exchange coefficient on wind speed model of [*Wanninkhof*, 1992]. Argon was added as a prognostic tracer to the simulations in both models in an analogous fashion to O$_2$; i.e., O$_2$ and Ar solubility are similarly determined using model temperature and salinity, and Ar uses the same gas exchange-wind speed model as O$_2$, but adjusted for the correct solubility and Schmidt number.
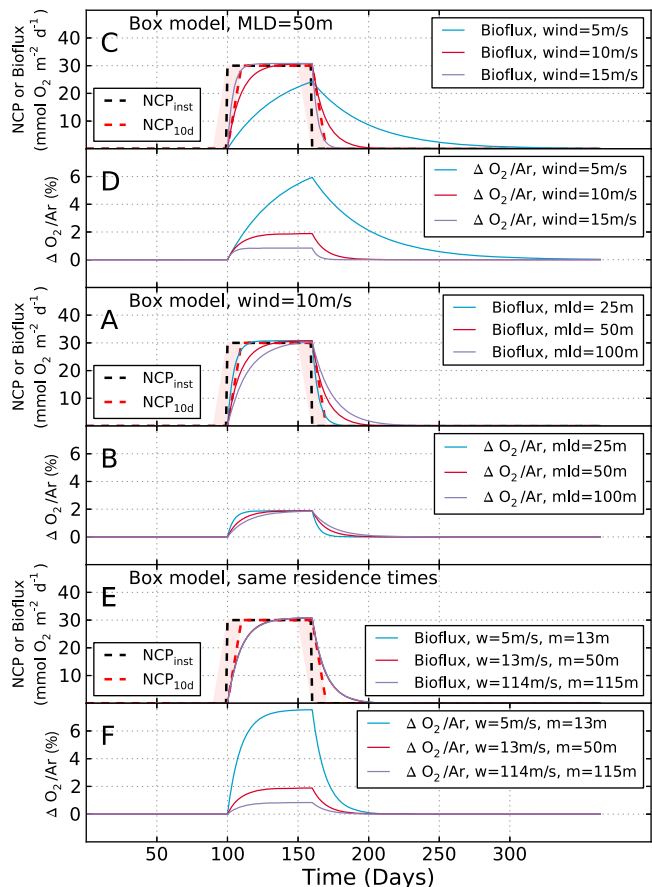

Figure 2. (a) O$_2$ bioflux calculated by a box model. NCP is prescribed as a rectangular pulse of 30 mol O$_2$ m^{-2} day^{-1} between days 100 and 160 (black dotted line). Red dotted line indicates 10-day backward average of NCP, and solid lines indicate the resulting O$_2$ bioflux for different wind speeds. The pink shading indicates the timescale over which NCP$_{10day}$ is averaged. (b) ΔO_2/Ar for the corresponding mixed layers. The wind speeds of 5, 10, and 15 m/s correspond to gas transfer velocities of 1.3, 5.1, and 11.5 m/day, respectively. (c, d) Analogous results to Figures 2a and 2b, but for varying mixed layer depths instead of wind speeds. The fixed wind speed of 10 m/s corresponds to a gas transfer velocity of 5.1 m^{-2} day^{-1}. (e, f) The evolution of O$_2$ bioflux and ΔO_2/Ar for different combinations of wind and mixed layer depth with the same residence time. Note that that the lines for O$_2$ bioflux in Figure 2e are on top of each other.

3. Results

[19] We begin our analysis by integrating a suite of experiments forcing the box model with a 60-day pulse of NCP at day 100 of a year-long simulation, and otherwise setting NCP to zero. This rectangular NCP pulse of $30 \, \text{mmol} \, \text{m}^{-2} \, \text{day}^{-1}$ is an unrealistic representation of biological production but illustrates well the fundamental interactions between NCP and air-sea gas exchange. In the first three cases, we prescribe a constant mixed layer of 50 m and different wind speeds of 5, 10, and $15 \, \text{m s}^{-1}$, respectively. The resulting bioflux and $\Delta O_2/Ar$ (Figures 2a and 2b) lag both NCP_{inst} and NCP_{10day}, leading to differences between O_2 bioflux and NCP immediately after rapid NCP changes. The case with a wind speed of $5 \, \text{m s}^{-1}$ shows a particularly large error, whereas O_2 bioflux tracks NCP_{10day} well when the wind speed is $15 \, \text{m s}^{-1}$. In the next three cases, we prescribe a constant wind speed of $10 \, \text{m s}^{-1}$ and three different mixed layer depths of 25, 50, and 100 m, respectively (Figures 2c and 2d). The agreement between O_2 bioflux and NCP_{10day} in these three cases is comparable to that for the intermediate and strong wind cases with 50 m MLD.

[20] It is likely that the main controlling factor for lags between O_2 bioflux and NCP is the gas residence time (τ). We test this by conducting three further experiments with different combinations of wind speed and mixed layer depth that have the same residence time. The combinations of wind and mixed -layer depths are wind $= 5 \, \text{m s}^{-1}$, MLD $= 12 \, \text{m}$; wind $= 10 \, \text{m s}^{-1}$, MLD $= 50 \, \text{m}$; and wind $= 15 \, \text{m s}^{-1}$, MLD $= 114 \, \text{m}$. These combinations generate O_2 biofluxes that are similar to the point where the lines are indistinguishable from each other (Figure 2e). This is, however, not the case with $\Delta O_2/Ar$ where different wind speeds but the same residence time generate different results (as seen in Figure 2f).

[21] Our next step is to test if smooth changes in NCP over the growing season can generate lags between O_2 bioflux and NCP and if the l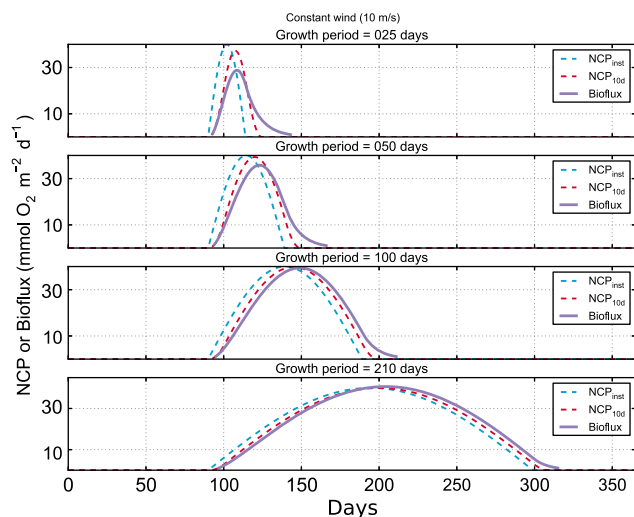ength of the growing season affects possible lags. We use a more realistic sinusoidal shape of the NCP pulse in this case, as shown in Figure 3. The different panels show seasonal NCP curves with durations of the growing period (NCP > 0) of 25, 50, 100, and 210 days, respectively, and the resulting O_2 bioflux from our box model. We find that O_2 bioflux lags NCP more when the growing season is shorter due to steeper dNCP/dt, with increasingly longer tails of O_2 bioflux after NCP falls to zero.

[22] To evaluate the combined effects of wind, NCP, and MLD, we integrate the box model with prescribed fields from TOPAZ and BGCCSM. Time series of NCP and winds for each grid cell in the models are used together with a fixed MLD. This set-up isolates the interactions of mixed layer biology and gas exchange on $\Delta O_2/Ar$ and O_2 bioflux, neglecting for the moment effects generated from entrainment by a dynamic mixed layer. The results from these simulations are presented in Figure 4. Each panel shows cumulative histograms of the fractional error between O_2 bioflux and NCP_{10day}, binned by O_2 bioflux as normal probability plots. With this kind of y-axis, data that belong to a normal distribution will be represented as a straight line [*Glover et al.*, 2011].
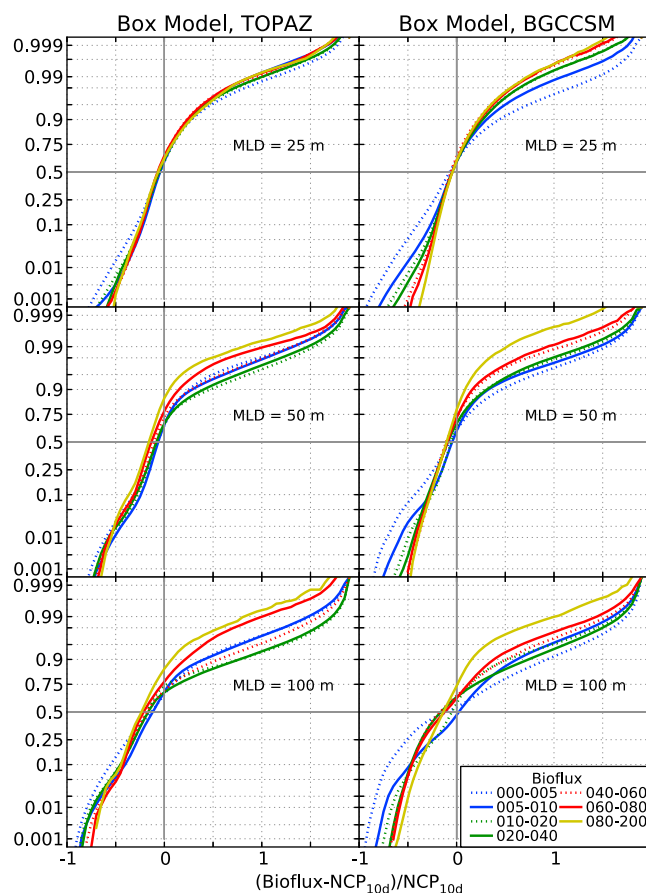

Figure 3. NCP and O_2 bioflux for different lengths of growing season calculated using a simple box model. Black lines show instant NCP, red line 10-day backward mean, and blue line O_2 bioflux. The pink shading indicates the timescale over which NCP_{10day} is averaged. Mixed layer depth is 50 m, and wind speed is 10 m/s ($\kappa = 5$ m/day).

Figure 4. Cumulative histograms showing the fractional difference of O_2 bioflux from NCP_{10day} in a simple box model, using a metric defined as $(O_2 \text{ bioflux-}NCP_{10day})/NCP_{10day}$. Each panel represents simulations with different mixed layer depths. The values of the metric are binned by O_2 bioflux, and each bin is represented by one line. A key describing which line correlates to which bin can be found in the bottom right panel.

[23] Each line on the plot represents the cumulative distribution for a range of O$_2$ bioflux values (as described in the lower right panel) calculated by the box model using a specific MLD and prescribed wind and NCP from one of the two models. If O$_2$ bioflux were a perfect proxy for NCP$_{10day}$, this line would be vertical and cross the origin at zero (all values of NCP-O$_2$ bioflux in the cumulative histogram would be zero). This is not the case for any of the combinations of time series and parameters tested by the box model. Instead, there are both random and systematic errors. The slope of the lines indicates the magnitude of the random errors, which can be diagnosed from the points where the line crosses .25 and 0.75 on the y axis. These lines show that the majority of points have errors below 15%, except for low O$_2$ bioflux values combined with deep mixed layers. At the upper and lower bounds, the distributions exhibit large negative and positive errors with relatively low probabilities. Further, for large positive fraction errors, the curves tend to deviate to the right from a straight line, indicating that the errors are non-Gaussian with long tails. In other words, a small number of points have very large errors.

[24] Systematic errors are shown where the line crosses 0.0 on the x axis: it crosses above 0.5 on the y axis, meaning that more than half of the O$_2$ bioflux values underestimate NCP$_{10day}$. It is clear from Figure 4 that random errors are larger for deeper mixed layers, which is consistent with the findings in Figure 2. There is also a tendency for deeper mixed layers to have a larger systematic error (the lines do not cross the y axis at origin). This is primarily an artifact in the analysis method due to the fact that O$_2$ bioflux does not "catch up" to NCP$_{10day}$ at the end of the time series. Different ranges of O$_2$ bioflux are similar for shallow MLD but tend to diverge when MLD is deep, with lower ranges of O$_2$ bioflux having larger errors. The main exception is the 0–5 mmol m^{-2} day^{-1} bin. At these low values of O$_2$ bioflux, there are many more cases of O$_2$ bioflux being 50% of NCP or less. As we will see elsewhere, low values of NCP are associated with low values of bioflux, but there are large percentage errors in the magnitude of small fluxes.

[25] We now turn the analysis to the full 3-D models. We use daily mean fields from TOPAZ and BGCCSM to calculate ΔO$_2$/Ar and O$_2$ bioflux. All calculations are done on each grid cell for each day before any averaging is conducted. Figure 5 compares O$_2$ bioflux with 10-day averaged values of mixed layer NCP for both TOPAZ and BGCCSM at model date December 16. We note that there are large differences in the distribution of Southern Ocean NCP between the two models, which we will discuss elsewhere. This issue does not concern us here as we are only interested in comparing O$_2$ bioflux and NCP calculated using the same internally consistent model ocean biogeochemistry and physics.

[26] For both models, O$_2$ bioflux captures the general pattern of NCP. In TOPAZ, there is a band of high NCP along the northern rim of the Southern Ocean in the Indian sector, another band in the center of the Southern Ocean in the Atlantic and Indian sectors, and high NCP along the coast. These features are also present in O$_2$ bioflux, although the magnitude of the variations is attenuated and the distribution is smoothed. In the Australian sector and in the region west of the Drake Passage, NCP and O$_2$ bioflux values are generally <10 mmol m^{-2} day^{-1}. There is a plume of high NCP downstream of South Georgia Island that is not clearly recorded in O$_2$ bioflux.
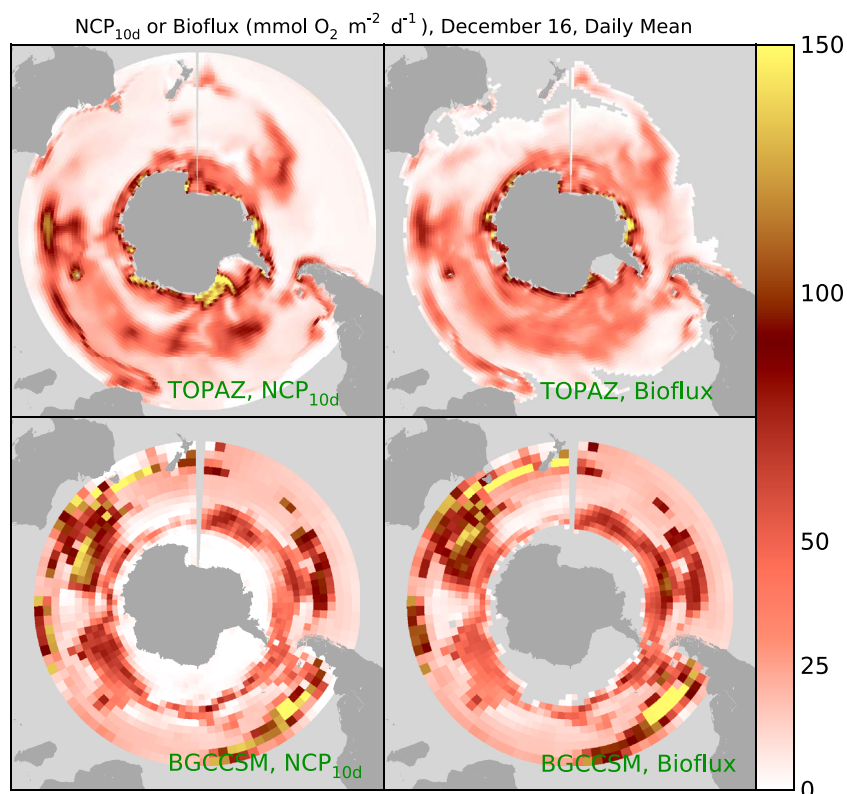

Figure 5. Mean of NCP$_{10day}$ and O$_2$ bioflux for 1 day (December 16) in TOPAZ and BGCCSM.
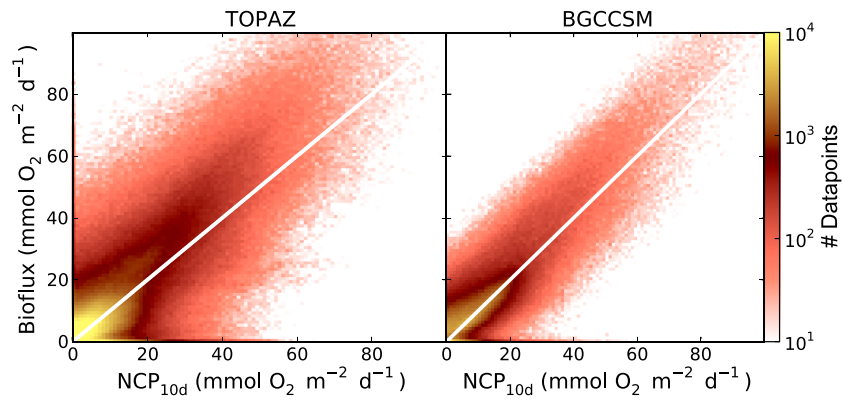
Figure 6. Oxygen bioflux vs NCP_{10day} for all data points in TOPAZ and BGCCSM. Values are binned, and colors represent number of data points that falls within each bin.
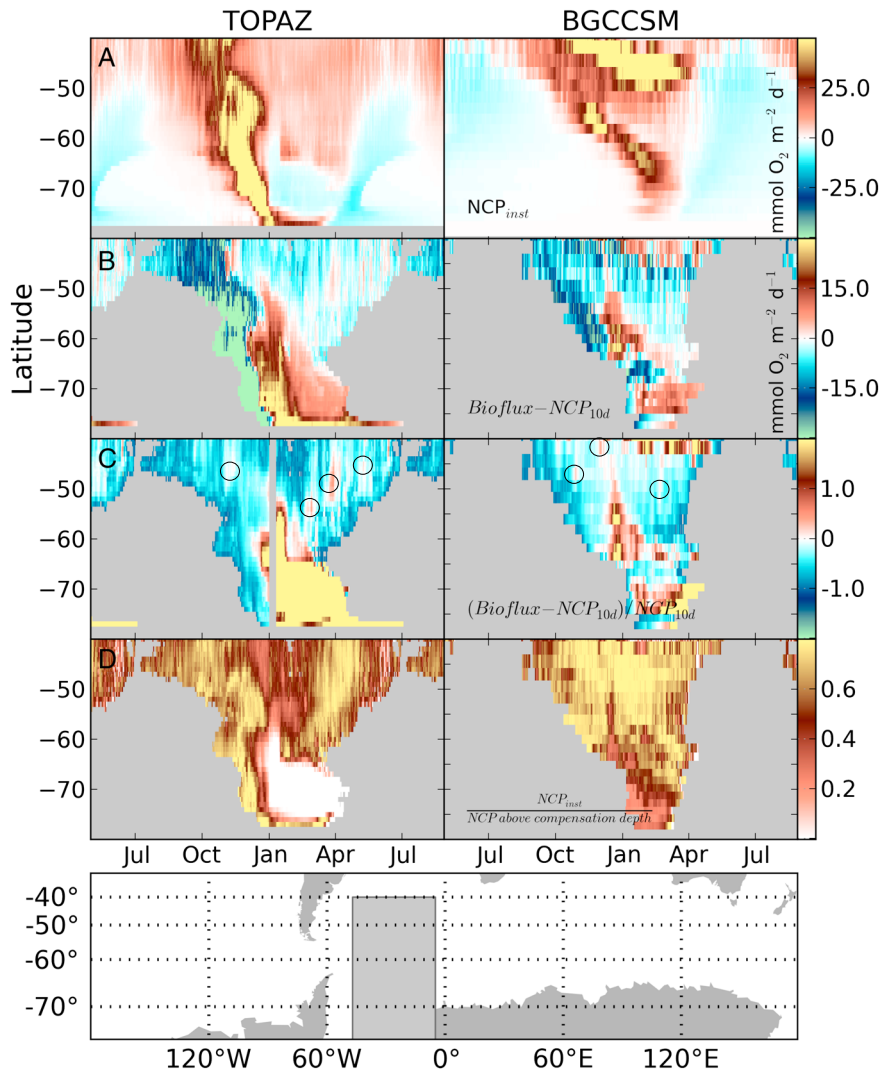

Figure 7. Hofmøller plots of (a) NCP_{inst}, (b, c) O₂ bioflux-NCP_{10day})/NCP_{10day}, and (d) the fraction of NCP occurring in the mixed layer in a box covering the Atlantic sector of the Southern Ocean (see map at the bottom of the figure). All values are zonal medians within the box. In Figures 7c and 7d, the grey areas are points where O₂ bioflux is <0. Circles in Figure 7c indicate event-like spikes in positive bioflux.

[27] In BGCCSM, NCP is high in the Australian sector, to the east and the west of the southern part of South America, and in a circumpolar band that passes just north of the Weddell Sea and the Ross Sea. These features are all well resolved in O$_2$ bioflux. Low values of NCP near the coast are represented by oxygen-undersaturated waters for which O$_2$ bioflux is undefined. In summary, in the context of our two models, O$_2$ bioflux gives an accurate picture of the geographical variation in NCP, although the pattern is smoothed considerably in TOPAZ.

[28] Differences between O$_2$ bioflux and mixed layer NCP$_{10day}$ become more apparent at the local grid-point scale. In Figure 6, a two-dimensional histogram of O$_2$ bioflux is plotted vs. 10-day averaged NCP for all the grid points analyzed from the TOPAZ and CCSM models. In BGCCSM, the two properties covary but with scatter about the 1:1 line, while in TOPAZ, O$_2$ bioflux generally plots above the 1:1 line. This divergence from the 1:1 line could be due to subsurface biological production, as discussed later on. When NCP $< \sim$10 mmol m^{-2} day^{-1}, O$_2$ bioflux
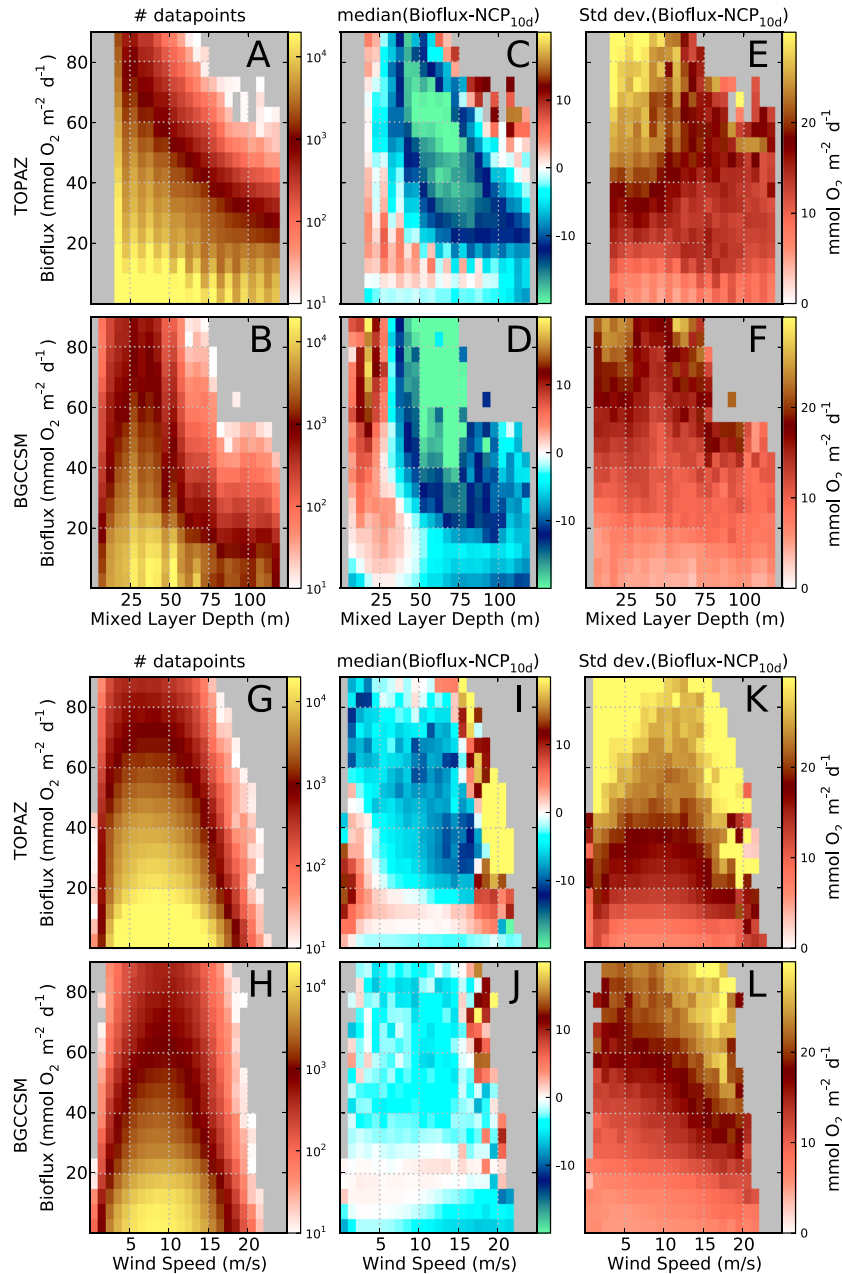

Figure 8. Statistical characteristics of NCP$_{10day}$ when binned by O$_2$ bioflux and mixed layer depths. The bin dimensions are 5 mmol m^{-2} day^{-1} and 5 m, respectively. (a, b) The number of data points in each bin for TOPAZ and BGCCSM, respectively. (c, d) The median difference between O$_2$ bioflux and NCP$_{10day}$ in each bin. (e, f) The corresponding standard deviation. (g, h, i, j, k, l) Analogous to Figures 8a–8f, but with wind speed on the x axis instead of mixed layer depth.

values are also low, but fractional errors are large. In TOPAZ, for example, O_2 bioflux reaches as high as 50 mmol M^{-2} day^{-1} in some samples when NCP is ~0, but such anomalies are 3 orders of magnitude less frequent than samples where NCP and O_2 bioflux are both close to 0.

[29] To diagnose the source of the scatter between O_2 bioflux and NCP_{10day} in Figure 6, we select a subregion of the model domains for an extended analysis. This region extends northward from the Weddell Sea and is shown as a grey box in the bottom panel of Figure 7. Although this is the only area presented here, we have conducted the following analyses on several similar regions with analogous results. All panels in Figure 7 are constructed by calculating a zonal average for each day and latitudinal row of grid cells. The resulting time series are presented as Hofmøller diagrams with latitude on the y axis and time on the x axis. The color of each point represents the value of a property at the appropriate latitude and time. These figures indicate, unsurprisingly, a strong latitudinal difference in biological production with shorter growth seasons in high latitudes. In both models, slightly negative NCP values prevail over wide reaches of the ocean in the fall and winter, reflecting low rates of net heterotrophy. TOPAZ also tends to have an earlier, shorter, and more intense growing season than CCSM. In TOPAZ, waters are slightly heterotrophic in January–April at high latitudes (Figure 7). In both models, O_2 bioflux systematically underestimates NCP_{10day} in most of the region. There are two exceptions: event-based spikes of high O_2 bioflux not corresponding to high NCP_{10day} (shown in Figure 7c as intermittent areas of red and emphasized with circles) and, at high

latitudes, values of O_2 bioflux that consistently overestimate NCP_{10day}, especially in TOPAZ.

[30] While O_2 bioflux on average underestimates NCP, local conditions may lead to O_2 bioflux values greater than NCP, as in the areas marked by red, brown, and yellow colors in Figure 7b. Anomalously high values of O_2 bioflux are matched by low ratios of mixed layer NCP divided by NCP integrated down to the compensation depth (Figure 7c). In these areas, the mixed layer is especially shallow, and a large fraction of NCP occurs below the mixed layer, as shown in Figure 7d. Mixing can then transport this biological O_2 into the mixed layer to be lost to the atmosphere. O_2 bioflux thus registers NCP not only in the mixed layer but to some degree in the underlying net autotrophic waters as well.

[31] To examine the general pattern of O_2 bioflux deviations with respect to NCP, we first exclude all grid points where O_2 bioflux < 0, similar to analysis methods for field $\Delta O_2/Ar$ data. Then, for each model, we bin NCP_{10day} values by O_2 bioflux and MLD. We can then examine how the differences between O_2 bioflux and NCP_{10day} are correlated to different levels of O_2 bioflux and MLD in the two models. Figures 8a and 8b show the number of data points in each bin, which are 5 mmol m^{-2} day^{-1} and 10 m in size, respectively. Most values fall within the bins representing low values of O_2 bioflux and shallow mixed layer depths. The differences between the models can likely be explained by different grids and mixed layer dynamics—TOPAZ has higher horizontal and vertical resolution, but BGCCSM can generate more shallow mixed layers. Figures 8c and 8d represent the median of the difference between O_2 bioflux
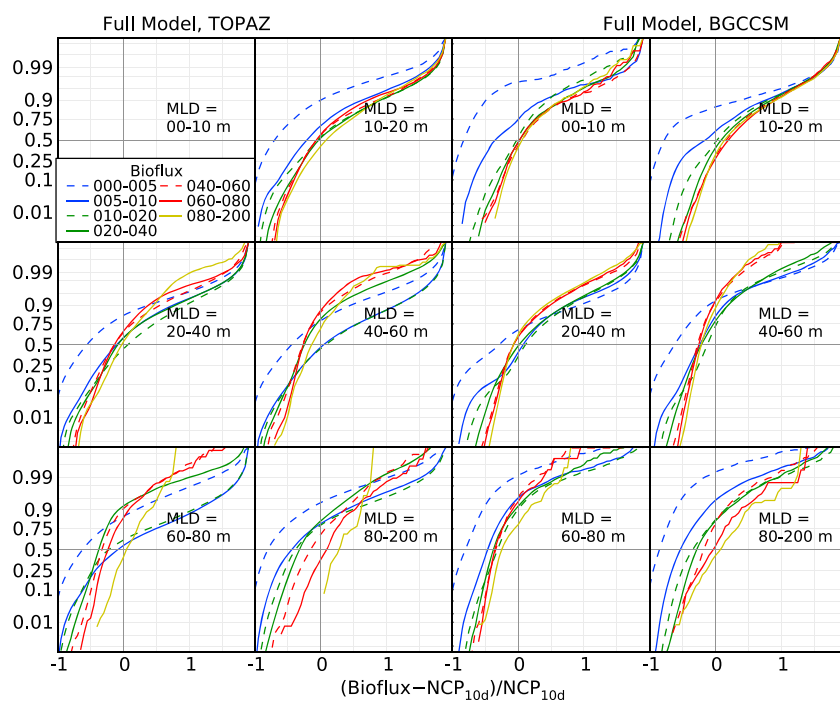

Figure 9. Cumulative histograms showing the fractional difference of O_2 bioflux from NCP_{10day} in TOPAZ and BGCCM, using a metric defined as $(O_2$ bioflux-$NCP_{10day})/NCP_{10day}$. Each panel shows histograms in a given range of mixed layer depths. Within each panel, the metrics are then binned by O_2 bioflux (mmol m^{-2} day^{-1}), represented by different lines. A key describing which line correlates to which bin can be found in the upper left panel.

and NCP_{10day}. We report median instead of mean to exclude biases from extreme outliers. In BGCCSM, O_2 bioflux systematically overestimates NCP_{10day} when MLD is shallower than 30 m and underestimates NCP_{10day} when MLD > 50 m.

[32] The fractional differences are particularly large for low values of O_2 bioflux. In the range of $0–10 \, m^{-2} \, day^{-1}$, the median difference between O_2 bioflux and NCP_{10day} is generally as large as the measured value of bioflux. The pattern in TOPAZ is qualitatively similar with some quantitative differences: O_2 bioflux underestimates NCP_{10day} for most bins where mixed layer depths are larger than 50 m. O_2 bioflux is larger than $20 \, mmol \, m^{-2} \, day^{-1}$ here as well, but the overestimation of NCP_{10day} extends to much deeper mixed layers when O_2 bioflux is low. This effect is probably due to the regions in the model where biological production occurs mainly below the base of the mixed layer, as seen in Figure 7. TOPAZ also differs when O_2 bioflux is high in shallow mixed layers. For such samples, the underestimation of NCP_{10day} is much larger in TOPAZ than in BGCCSM,

perhaps because the two models have different grids and mixed layer dynamics. Finally, TOPAZ shows bins with high O_2 bioflux and deep mixed layers where NCP_{10day} is overestimated. Note from Figure 8a that the number of data points with these conditions is quite small. The grid cells affiliated with these bins are located on the border between regions of high NCP/shallow mixed layers and low NCP/deep mixed layers. At times, supersaturated waters are advected from the former regions to the latter and generate conditions with low NCP, deep mixed layers, and high O_2 bioflux. This process is probably not a factor in BGCSSM due to its lower horizontal resolution. The standard deviation of the errors in TOPAZ (Figure 8f) is much higher than that of the errors in BGCCSM (Figure 8e). Poorer agreement in TOPAZ might reflect the shorter growing season and rapid NCP variations.

[33] We know from Figure 2 that lags between O_2 bioflux and NCP are correlated with the residence time of O_2 in the mixed layer. Figures 8g–8l show how the wind relates to bioflux in an analogous fashion to Figures 8a–8f.
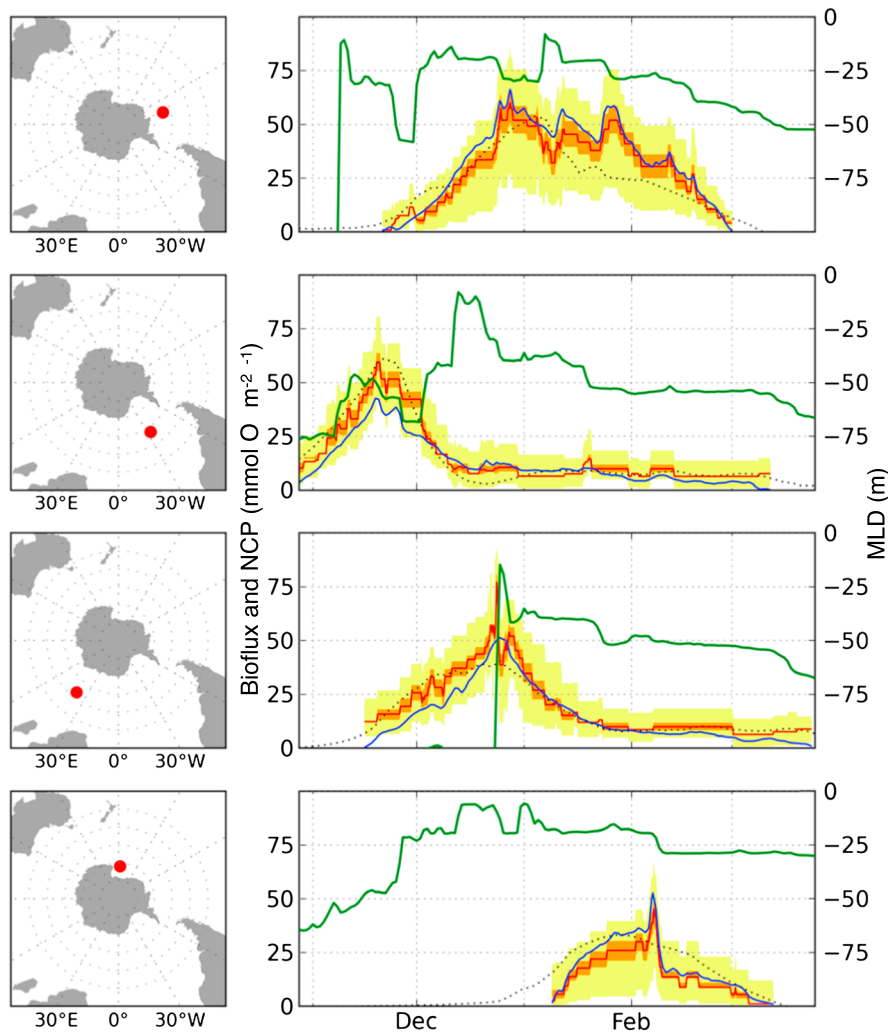

Figure 10. MLD and measures of production at four sites vs. time, according to BGCCSM. Green lines give time history of mixed layer depth. Blue lines are time series of model O_2 bioflux, and dotted black lines are time series of NCP. The red line is the median value of NCP_{10day} for all grid points with the same values of MLD and O_2 bioflux. The Orange band encompasses 68% of NCP_{10day} values for all model grid points with the model values of MLD and bioflux. The yellow band encompasses 90% of these points.

The patterns of median errors in O_2 bioflux in Figures 8i and 8j are similar to what we would have expected when comparing them to Figures 8c and 8d. O_2 bioflux overestimates NCP at high wind speeds and underestimates NCP at low winds. The patterns, however, are not as pronounced as those with mixed layer depths. We believe this result to be due to the winds' more complicated relationship with the observed errors, which affects not only residence times but also the depth of the mixed layer and how much O_2 bioflux a given $\Delta O_2/Ar$ will generate.

[34] Figure 9 displays the cumulative histograms for the fractional difference of O_2 bioflux from NCP_{10day} for the full 3-D models in an analogous fashion to Figure 4. The results presented in this figure show more complex patterns than those seen in Figure 8, though similarities exist. The primary result is that in the MLD range of 20–60 m, which is common in the summertime Southern Ocean, bioflux estimates of NCP have a typical error of less than 15%, except for an aliasing (see below). When MLD is >60 m, NCP values

tend to be low [e.g., *Cassar et al.*, 2011], and fractional differences between bioflux and NCP are larger. The results that differ will be discussed in detail later in the text. It is clear that the lowest ranges of O_2 bioflux ($<10\,\mathrm{mmol\,m^{-2}\,day^{-1}}$) have large fractional errors, but such low values of bioflux are qualitatively consistent with low values of NCP. A comparison of Figures 4 and 9 suggests that most of the random errors (the shape of the lines) can be explained by the box model, but that the full 3-D models generate much larger systematic errors (seen in the line's crossing of the y axis) when MLD > 60 m.

[35] Finally, we use the knowledge of how O_2 bioflux interacts with NCP and MLD for a more finely grained error estimate. Using the binned values of NCP_{10day} described in Figure 8, we can create a histogram from all model NCP_{10day} values with a certain combination of O_2 bioflux and MLD. Such histograms can then be used to illustrate what systematic errors and variability models suggest for each set of in situ O_2 bioflux and MLD values.
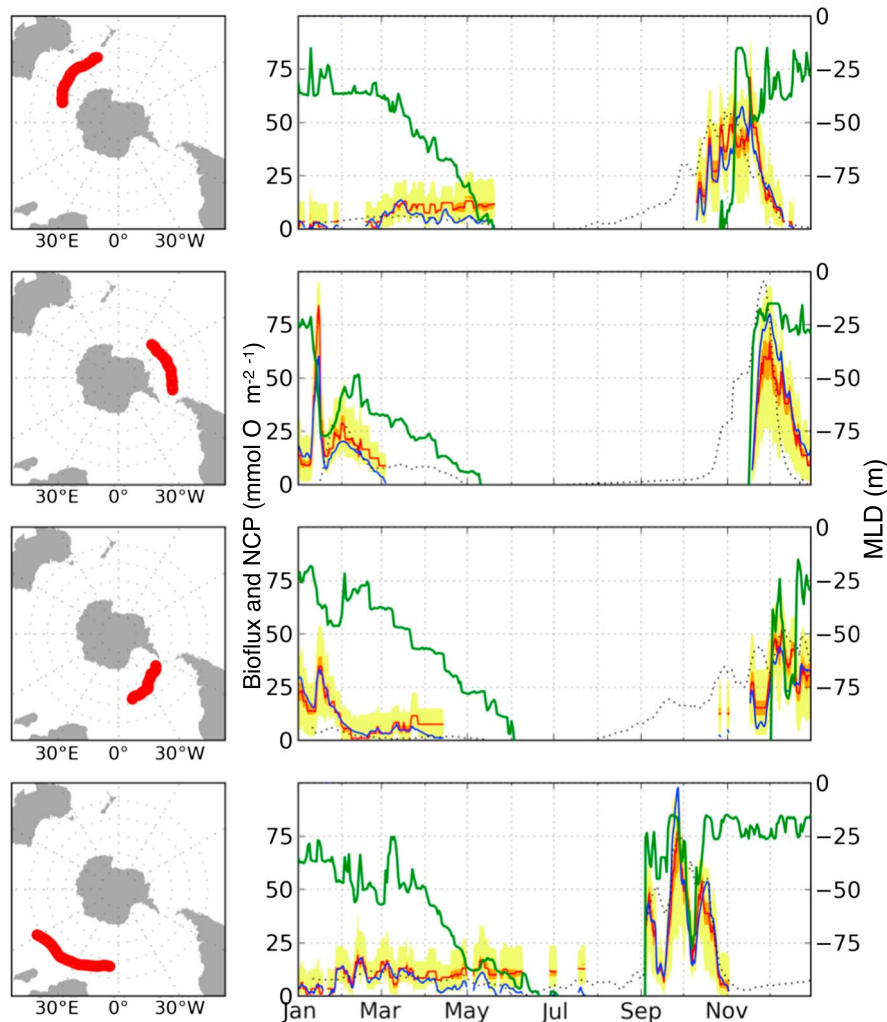

Figure 11. MLD and measures of production along four Lagrangian trajectories vs. time, according to TOPAZ. Green lines give time history of mixed layer depth. Blue lines are time series of model O_2 bioflux, and dotted black lines are time series of NCP. The red line is the median value of NCP_{10day} for all grid points with the same values of MLD and O_2 bioflux. The orange band encompasses 68% of NCP_{10day} values for all model grid points with the model values of MLD and O_2 bioflux. The yellow band encompasses 90% of these points.

[36] Figures 10 and 11 show examples of such histograms for a number of simulated time series of NCP_{10day} taken from the models. The blue curves represent O₂ bioflux for the time series calculated as described in the text, and the dotted black curves represent NCP averaged over the 10-day period prior to sampling. The red lines, orange bands, and yellow bands show the distribution of model NCP values when mixed layer depth and O₂ bioflux correspond to the plotted values. The red line is the mean value of NCP under these conditions. The orange band encompasses 68% of the samples; the yellow band encompasses 90%. Figure 10 shows the evolution of O₂ bioflux and NCP_{10day} for four stationary points (Eulerian time series) in BGCCSM. Consider the bottom panel, corresponding to a site at the northwestern edge of the Ross Sea. During the last part of January, NCP_{10day} (dotted black line) is $32\ mmol\,m^{-2}\,day^{-1}$, and O₂ bioflux (solid blue line) is essentially identical. The red line shows that, for all grid boxes with the diagnosed O₂ bioflux and mixed layer depth (~15 m), the mean value of NCP is $25\ mmol\,m^{-2}\,day^{-1}$. The orange band shows that 68% of all grid cells with the same values of O₂ bioflux and MLD have NCP_{10day} values within the range $21\text{–}29\ mmol\,m^{-2}\,day^{-1}$, and 90% fall within the much broader range of $4\text{–}38\ mmol\,m^{-2}\,day^{-1}$. These ranges are the error ranges encountered if we were to attempt to correct O₂ bioflux to NCP assuming that any grid box with the observed O₂ bioflux and MLD values gives an equally valid measure of NCP. Figure 11 compares O₂ bioflux to NCP in TOPAZ for particles advected in a Lagrangian frame by model velocity fields also from TOPAZ.

[37] Results in Figures 10 and 11 show that bioflux generally gives a good representation of the magnitude of NCP and its time course over the growing season. However, a subset of the results will have large errors. Results in Figure 9 show that bioflux most accurately represents NCP at typical summertime MLD between 20 and 60 m. When MLD is shallower, NCP may be underestimated because of vertical transport. When MLD is >60 m, NCP is generally low. Bioflux values are also low, but fractional differences between NCP and bioflux may be large.

4. Discussion

[38] Our evaluation of sea-air biological oxygen flux (O₂ bioflux) as a proxy for NCP shows that O₂ bioflux generally gives a valuable picture of the large-scale geographical and temporal variation in NCP. However, there are significant and variable discrepancies in local scales between these two terms. There are two main sources of errors: misalignment in time and space between oceanic biological production and atmospheric ventilation of excess O₂ to the atmosphere, and systematic errors from horizontal or vertical transport of O₂ across the base of the mixed layer.

[39] The misalignment of O₂ bioflux and NCP_{10day} originates from O₂ bioflux lagging NCP_{10day} when biological production is varying, as shown in our box model experiments. The effect of such temporal lags is more pronounced for deeper mixed layers, as seen in Figure 4. These temporal lags can also generate a spatial decorrelation under certain circumstances. Waters in regions of high biological production during conditions of low winds can be advected away and ventilated during high wind events at a different location. This spatial decorrelation explains the somewhat puzzling pattern of O₂ bioflux overestimating NCP in regions with high O₂ bioflux and deep mixed layers, as seen in the upper right corner of Figure 8c. These areas are located on the border between regions of high NCP/shallow mixed layers and low NCP/deep mixed layers. Mixing of waters between the two regions generates conditions with low NCP, deep mixed layers, and high O₂ bioflux.

[40] Figure 4 shows that these misalignments lead to large variability in how well local O₂ bioflux estimates reflect NCP_{10day} when daily data from individual grid cells are compared. The lag between O₂ bioflux and NCP_{10day} is, in general, symmetric, which means that the underestimation of NCP in the early part of the growing season is compensated by overestimation at the end. Wintertime upwelling, however, creates highly undersaturated conditions when NCP begins to turn positive, which has to be compensated for before O₂ bioflux can register any biological production. Any oxygen produced during this time period will be neglected by the ΔO₂/Ar method and not compensated for at the end of the growing season. Upwelling of undersaturated water in autumn can also lead to O₂ bioflux underestimating NCP.

[41] The other main source of error is transport of oxygen to and from the mixed layer via resolved advective transport and subgrid diffusive processes. These fluxes can be identified by changes in mixed layer ΔO₂/Ar supersaturation not explained by O₂ bioflux or NCP and are responsible for O₂ bioflux systematically misrepresenting NCP_{10day}. Our analyses indicate that entrainment is only generating a small part of these changes in mixed layer oxygen inventory. Instead, diapycnal transport seems to play a much larger role in two ways. First, oxygen supersaturation can build up seasonally below the mixed layer when the euphotic zone exceeds the mixed layer depth, as is almost always the case. Subsurface oxygen supersaturation can, in some locations and seasons, greatly exceed observed values in the mixed layer, which are damped by gas exchange. This excess of thermocline biological oxygen can be transported vertically to the mixed layer via diapycnal diffusion, resulting in a positive aliasing of mixed layer O₂ supersaturation and O₂ bioflux [e.g., *Nicholson et al.*, 2012].

[42] Second, the large-scale circulation of the Southern Ocean involves the upwelling of low-oxygen thermocline and deep waters into the Southern Ocean surface layer. A net oxygen influx results from the ventilation of these upwelled waters, which are carrying the oxygen debit from subsurface respiration in the low latitudes [*Gruber et al.*, 2001]. Most ventilation has been expected to take place in wintertime, but the process seems to continue into the summer by these mechanisms. Ventilation through stratified waters is generally slow enough that the mixed layer remains biologically supersaturated. However, ventilation still attenuates the supersaturation of O₂ due to NCP.

[43] These results underscore the validity of established conventions to reject negative O₂ bioflux values used to estimate net community production [e.g. *Reuer et al.*, 2007]. It is possible that O₂ bioflux is less than zero due to net heterotrophy, but it is clear that other processes can influence and even create such conditions as well. As a matter of fact, we find that net community production is positive 30% of the time when O₂ bioflux is negative in the BGCCSM

model. This suggests that O$_2$ bioflux values below zero are poor indicators for ecosystem heterotrophy in the Southern Ocean without further knowledge about the physical conditions in the area around the observations.

[44] It is clear from our results that the ΔO$_2$/Ar method suffers from a systematic negative bias when used to estimate NCP in the Southern Ocean. We estimate this error to be about 6% in BGCCSM and 15% in TOPAZ when we only include days with positive ΔO$_2$/Ar, which is consistent with the conventions for use of this method. The need to remove negative ΔO$_2$/Ar values has a limited effect when the method is used over short timescales or for individual observations, but can create significant mismatches when estimating NCP integrated over the growing season. It is well known that deep wintertime mixing in parts of the Southern Ocean creates O$_2$-undersaturated springtime surface waters. The mixed layer O$_2$ will eventually be saturated from NCP, but none of the biological production during undersaturated conditions is registered by the ΔO$_2$/ Ar method. To evaluate this error, we compare the integrated summer NCP with what the ΔO$_2$/Ar method predicts. The result is that O$_2$ bioflux underestimates seasonal NCP by 36% for TOPAZ and by 21% for BGCCSM. This discrepancy between O$_2$ bioflux and NCP also implies a net influx of O$_2$ from the atmosphere into the Southern Ocean and subsequent export via ocean advection as reported by *Gruber* [2001].

[45] To evaluate the relative influence of different processes affecting the ΔO$_2$/Ar method, we compare results from the box model with analogous data from BGCCSM. The box model is run in two different configurations. In the first, time-dependent wind speed, NCP, and MLD are prescribed to equal values from BGCCSM. The varying mixed layer depth is only implemented for gas residence time calculations and does not change the content of O$_2$ in the mixed layer. This can be interpreted as "deeper waters" in the box model with the same O$_2$ concentration as in the mixed layer and where no vertical transport of O$_2$ occurs. While somewhat idealized, this configuration helps us distinguish between errors from temporal lags and errors from vertical processes. In the second box model simulation, we initialize the box model O$_2$ concentrations at the start of the growing season with undersaturated, wintertime values from the BGCCSM. We compare the two box model simulations with the full model by calculating the root mean squared deviation (RMSD) between O$_2$ bioflux and NCP (shown in Figure 12) for all grid cells and all days when O$_2$ bioflux is above zero. The first box model simulation has an RMSD of 3.3 mmol O$_2$ m^{-2} day^{-1} for the full region, which is about half of the errors in the full model run with both vertical and horizontal transport processes included (RMSD = 7.4 mmol O$_2$ m^{-2} day^{-1}). The box model simulation initialized with winter undersaturation has an RMSD to 3.9, suggesting that the deep wintertime O$_2$/Ar undersaturation has only a limited influence on errors in the ΔO$_2$/Ar method during the entire growing season. The median bias for the case with undersaturation is also small, with O$_2$ bioflux underestimating NCP$_{10day}$ by 0.1 mmol O$_2$ m^{-2} day^{-1} for the full region and summer season. The corresponding value for the box model with no winter undersaturation is 0 and 1.0 mmol O$_2$ m^{-2} day^{-1} for the full General Circulation model.
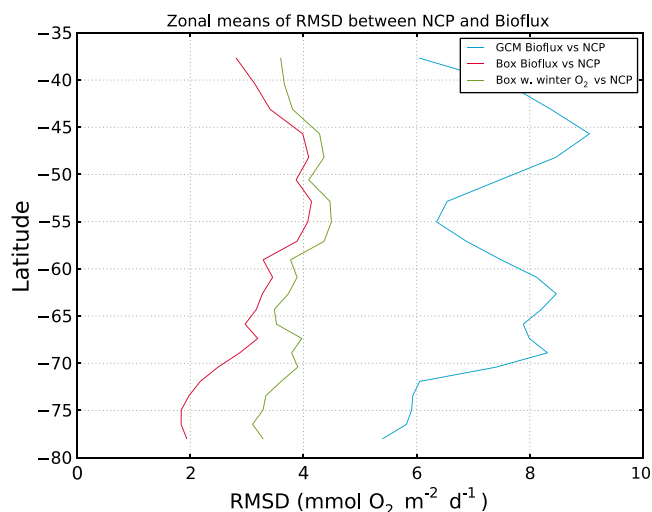

Figure 12. Zonally averaged RMSD between NCP$_{10day}$ and O$_2$ bioflux in the BGCCSM model. The black line is based on O$_2$ bioflux diagnosed from the full GCM. The blue line is based on O$_2$ bioflux from the box model where realistic NCP, winds, and mixed layer depths from the GCM have been used. The green line is based on O$_2$ bioflux from an analogous box model run as the blue line with the exception that the model is initiated with an O$_2$/Ar undersaturation that mimics winter conditions. NCP$_{10day}$ is identical in the three cases.

[46] It is also possible to evaluate the relative importance of misalignment versus vertical fluxes by comparing Figures 4 and 9. These two figures provide analogous error diagnostics for the box model (which does not have vertical dynamics) and the full 3-D models. The cumulative histograms in Figure 9 show slopes similar to the ones in Figure 4 when MLD is >50 m and O$_2$ bioflux is larger than 10 mmol m^{-2} day^{-1}, confirming that temporal misalignment alone can explain a large part of the random errors between O$_2$ bioflux and NCP$_{10day}$. Cumulative histograms representing shallow mixed layers and low levels of O$_2$ bioflux differ between the two figures, suggesting that processes other than misalignment are at play in those regimes. Systematic errors seem to mainly be a function of vertical fluxes in the full model since the results in the two figures diverge, in accordance with expectations.

[47] Figure 8 shows that errors in O$_2$ bioflux are correlated with MLD and the magnitude of O$_2$ bioflux. However, other factors might be as important in understanding these issues. Processes like turbulence, eddy occurrence, fronts, and stirring are probably critical as they most likely control the biological production and affect the vertical flux of oxygen from the mixed layer. Small-scale variability in wind and pressure can also be an important factor, as wind affects O$_2$ bioflux in many different ways (see Appendix and 1-D modeling work in *Keeling et al.* [1998]). Since these small-scale processes are not resolved in TOPAZ or BGCCSM, we cannot assess their quantitative role in the present modeling context.

[48] While this analysis is aimed at seasonally varying NCP, it is important to note that this effect would also occur when NCP changes rapidly in response to transient events such as mixed layer deepening associated with passage of storms, and in connection to fronts and eddies. Finally, the temporal and spatial variations of gas exchange transfer

Table 1. Minus One Sigma (−σ) and Plus One Sigma (+σ) Limits of the Confidence Interval, and the Mean Bias (mmb) Between NCP and O₂ Bioflux for Different Ranges of O₂ Bioflux and Mixed Layer Depths in the BGCSSM Model[a]

MLD (m)	O₂ Bioflux (mmol m⁻² day⁻¹)																			
	0–10				10–25				25–50				50–75				75–100			
	−σ	mmb	+σ	#	−σ	mmb	+σ	#	−σ	mmb	+σ	#	−σ	mmb	+σ	#	−σ	mmb	+σ	#
0–10	+2.8	+5.6	+7.6	4.1e+3	−2.3	−0.1	±1.2	3.3e+3	−7.3	−3.1	+1.1	1.7e+3	−11.8	−5.2	+1.8	8.3e+2	−15.5	−6.3	+3.8	4.0e+2
10–25	−1.8	+1.6	+2.8	4.0e+4	−4.8	−2.1	0	4.6e+4	−9.6	−5.5	−0.5	3.0e+4	−18.3	−11.5	−4.2	1.1e+4	−24.6	−14.4	−2.2	5.3e+3
25–50	−1.3	+0.5	+0.9	1.3e+5	−3.6	−1.0	+0.2	1.3e+5	−3.6	+1.3	+6.9	6.4e+4	−1.7	+3.5	+11.3	2.6e+4	−0.7	+5.2	+14.4	1.3e+4
50–75	+2.8	+4.9	+5.4	9.1e+4	+1.6	+5.1	+5.3	4.8e+4	+9.7	+14.3	+18.2	1.5e+4	15.6	+22.1	+28.3	4.3e+3	+16.4	+24.4	+33.8	1.2e+3
75–100	+4.6	+7.7	+9.0	1.9e+4	+3.0	+8.0	+11.5	1.2e+4	+8.6	+13.1	+16.6	3.4e+3	+8.6	+18.2	+25.8	3.3e+2	+0.3	+19.6	+22.3	5.2e+1
0–100	−0.3	+2.6	+3.6	2.9e+5	−2.7	+0.5	+1.7	2.4e+5	−4.3	+1.5	+7.0	1.1e+5	−6.3	+1.4	+10.4	4.3e+4	−6.7	+0.8	+12.4	2.0e+4

[a]Number of values (#) used to calculate the means and confidence intervals are also presented in the table. Positive values of mmb indicate that O₂ bioflux underestimates NCP. Confidence intervals are generated from all values of model NCP that fall within each bin of O₂ bioflux and mixed layer depth ranges.

velocities on a global scale described in *Keeling et al.* [1998] could also lead to biases in the $\Delta O_2/Ar$ method discussed here.

5. Conclusions

[49] We conclude with some guidance on how the results from this study can be used in evaluating real-world observations. It should be noted that the values mentioned here are solely based on BGCSSM and TOPAZ and could potentially be different if other models were used. We also note that small-scale processes such as eddies are not resolved in the models.

[50] We find that observations of O₂ bioflux below $10 \, \mathrm{mmol \, m^{-2} \, day^{-1}}$ have high fractional errors. Therefore, values in this range for the Southern Ocean should be interpreted with caution. However, low bioflux values generally indicate low NCP.

[51] There is a strong correlation between errors in O₂ bioflux estimated from NCP and the depth of the mixed layer, which complicates comparisons between O₂ bioflux and MLD. After evaluating all combinations of random and systematic errors found in the models, we suggest a rule of thumb that O₂ bioflux overestimates NCP by ~10–15 $\mathrm{mmol \, m^{-2} \, day^{-1}}$ in areas where the mixed layer is shallower than 30 m and underestimates NCP by ~10–15 $\mathrm{mmol \, m^{-2} \, day^{-1}}$ when the depth of the mixed layer is deeper than 40 m. These values are based on a combination of mean errors and confidence intervals. These values cannot be used to scale observations. We show these relationships in more detail in Table 1, where mean errors and 95% confidence intervals are presented for different ranges of mixed layer depths and O₂ biofluxes. This table can be used to estimate error bars for observations, given that previously mentioned caveats are taken into consideration.

[52] The fact that lags between NCP and O₂ bioflux depend on the temporal gradients of NCP suggests that it is important to know both the period and the amplitude over which biological production varies in order to estimate this effect in the real world.

[53] It is highly useful to have hydrographic observations together with any $\Delta O_2/Ar$ measurements since the physical state of the ocean influences how well O₂ bioflux can estimate NCP. The $\Delta O_2/Ar$ method is particularly sensitive to conditions where the mixed layer varies over short temporal or spatial scales, especially if this variability potentially correlates with changes in biological production [*Hamme et al.*, 2012]. As such, we consider the error estimates provided here to be lower limits, given that small-scale processes such as eddies are not resolved in the models. One example could be a frontal crossing where the mixed layer shallows as the front is approached and deepens after the crossing. While vertical processes associated with the front have the potential to generate elevated biological production, we show that correlations between O₂ bioflux and mixed layer depth reflect important artifacts, making the observed patterns difficult to interpret. Overall, the more complex the circulation, the more uncertain are observational estimates of O₂ bioflux.

References

Behrenfeld, M. J., and P. G. Falkowski (1997), A consumer's guide to phytoplankton primary productivity models, *Limnol. Oceanogr.*, *42*(7), 1479–1491.

Bender, M. L., S. Kinter, N. Cassar, and R. Wanninkhof (2011), Evaluating gas transfer velocity parameterizations using upper ocean radon distributions, *J. Geophys. Res.*, *116*(C2), C02010, doi:10.1029/2009JC005805.

Carr, M.-E. et al. (2006), A comparison of global estimates of marine primary production from ocean color, *Deep-Sea Res. Pt Ii*, *53*(5-7), 741–770, doi:10.1016/j.dsr2.2006.01.028.

Cassar, N., B. A. Barnett, M. L. Bender, J. Kaiser, R. C. Hamme, and B. Tilbrook (2009), Continuous high-frequency dissolved O$_2$/Ar measurements by equilibrator inlet mass spectrometry, *Anal. Chem.*, *81*(5), 1855–1864, doi:10.1021/ac802300u.

Cassar, N., P. J. DiFiore, B. A. Barnett, M. L. Bender, A. R. Bowie, B. Tilbrook, K. Petrou, K. J. Westwood, S. W. Wright, and D. Lefevre (2011), The influence of iron and light on net community production in the Subantarctic and Polar Frontal Zones, *Biogeosci.*, *8*(2), 227–237, doi:10.5194/bg-8-227-2011.

Craig, H., and T. Hayward (1987), Oxygen supersaturation in the ocean: Biological versus physical contributions, *Science*, *235*(4785), 199–202.

Doney, S. C., I. Lima, R. A. Feely, D. M. Glover, K. Lindsay, N. Mahowald, J. K. Moore, and R. Wanninkhof (2009a), Mechanisms governing interannual variability in upper-ocean inorganic carbon system and air-sea CO$_2$ fluxes: Physical climate and atmospheric dust, *Deep-Sea Res. Pt Ii*, *56*, 640–655, doi:10.1016/j.dsr2.2008.12.006.

Doney, S. C., I. Lima, J. K. Moore, K. Lindsay, M. J. Behrenfeld, T. K. Westberry, N. Mahowald, D. M. Glover, and T. Takahashi (2009b), Skill metrics for confronting global upper ocean ecosystem-biogeochemistry models against field and remote sensing data, *J. Marine Syst.*, *76*(1-2), 95–112, doi:10.1016/j.jmarsys.2008.05.015.

Dunne, J., R. A. Armstrong, A. Gnanadesikan, and J. Sarmiento (2005), Empirical and mechanistic models for the particle export ratio, *Global Biogeochem. Cycles*, *19*(4), GB4026, doi:10.1029/2004GB002390.

Dunne, J. P., A. Gnanadesikan, J. L. Sarmiento, and R. D. Slater (2010), Technical description of the prototype version (v0) of Tracers Of Phytoplankton with Allometric Zooplankton (TOPAZ) ocean biogeochemical model as used in the Princeton IFMIP model, *Biogeosci.*, *7*, 3593.

Gent, P., and J. McWilliams (1990), Isopycnal mixing in ocean circulation models, *J. Phys. Oceanogr.*, *20*(1), 150–155.

Glover, D. M., W. J. Jenkins, and S. C. Doney (2011), Modeling Methods for Marine Science, Cambridge University Press Cambrige, U. K.

Griffies, S. et al. (2005), Formulation of an ocean model for global climate simulations, *Ocean Sci.*, *1*, 45–79. [online] Available from: http://www.ocean-sci.net/1/45/2005/os-1-45-2005.html

Gruber, N., M. Gloor, S.-M. Fan, and J. L. Sarmiento (2001), Air-sea flux of oxygen estimated from bulk data: Implications for the marine and atmospheric oxygen cycles, *Global Biogeochem. Cycles*, *15*, 783–803.

Hamme, R. C. et al. (2012), Dissolved O$_2$/Ar and other methods reveal rapid changes in productivity during a Lagrangian experiment in the Southern Ocean, *J. Geophys. Res-Oceans*, *117*, doi:10.1029/2011JC007046.

Hendricks, M. B., M. L. Bender, and B. A. Barnett (2004), Net and gross O$_2$ production in the Southern Ocean from measurements of biological O$_2$ saturation and its triple isotope composition, *Deep Sea Res. I*, *51*(11), 1541–1561, doi:10.1016/j.dsr.2004.06.006.

Ho, D. T., R. Wanninkhof, P. Schlosser, D. S. Ullman, D. Hebert, and K. F. Sullivan (2011), Toward a universal relationship between wind speed and gas exchange: Gas transfer velocities measured with ^{3}He/SF$_6$ during the Southern Ocean Gas Exchange Experiment, *J. Geophys. Res. Oceans*, *116*, C00F04, doi:10.1029/2010JC006854.

Huang, K., H. W. Ducklow, M. Vernet, N. Cassar, and M. L. Bender (2012), Export production and its regulating factors in the West Antarctica Peninsula region of the Southern Ocean, *Global Biogeochem. Cycles*, *26*, doi:10.1029/2010GB004028.

Joensson, B. F., J. E. Salisbury, and A. Mahadevan (2011), Large variability in continental shelf production of phytoplankton carbon revealed by satellite, *Biogeosci.*, *8*(5), 1213–1223, doi:10.5194/bg-8-1213-2011.

Kaiser, J., M. Reuer, B. Barnett, and M. Bender (2005), Marine productivity estimates from continuous O$_2$/Ar ratio measurements by membrane inlet mass spectrometry, *Geophys. Res. Lett.*, *32*(19), L19605, doi:10.1029/2005GL023459.

Keeling, R., B. Stephens, R. Najjar, S. Doney, D. Archer, and M. Heimann (1998), Seasonal variations in the atmospheric O$_2$/N$_2$ ratio in relation to the kinetics of air-sea gas exchange, *Global Biogeochem. Cycles*, *12*(1), 141–163.

Langdon, C., J. Marra, and C. Knudson (1995), Measurements of net and gross O$_2$ production, dark O$_2$ respiration, and ^{14}C assimilation at the Marine Light-Mixed Layers site (59°N, 21°W) in the Northeast Atlantic Ocean, *J. Geophys Res. Oceans*, *100*(C4), 6645–6653.

Large, W., J. McWilliams, and S. Doney (1994), Oceanic vertical mixing: A review and a model with a nonlocal boundary-layer parameterization, *Rev. Geophys.*, *32*(4), 363–403.

Marra, J., W. Chamberlin, and C. Knudson (1993), Proportionality between in situ carbon assimilation and biooptical measures of primary production in the Gulf of Maine in summer, *Limnol. Oceanogr.*, *38*(1), 232–238.

Moore, J., and S. Doney (2004), Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model, Global Biogeochemical Cycles, 18(4), doi:10.1029/2004GB002220

Nicholson, D. P., R. H. R. Stanley, E. Barkan, D. M. Karl, B. Luz, P. D. Quay, and S. C. Doney (2012), Evaluating triple oxygen isotope estimates of gross primary production at the Hawaii Ocean Time-Series and Bermuda Atlantic Time-Series Study sites, *J. Geophys. Res. Oceans*, *117*, C05012, doi:10.1029/2010JC006856.

Quay, P., J. Stutsman, and T. Steinhoff (2012), Primary production and carbon export rates across the subpolar N. Atlantic Ocean basin based on triple oxygen isotope and dissolved O2 and Ar gas measurements, *Global Biogeochem. Cycles*, *26*, doi:10.1029/2010GB004003.

Reuer, M. K., B. A. Barnett, M. L. Bender, P. G. Falkowski, and M. B. Hendricks (2007), New estimates of Southern Ocean biological production rates from O$_2$/Ar ratios and the triple isotope composition of O$_2$, *Deep Sea Res.*, *54*(6), 951–974, doi:10.1016/j.dsr.2007.02.007.

Sarmiento, J. L., and N. Gruber (2006), Ocean Biogeochemical Dynamics, Princeton University Press, Princeton, N. J.

Smith, R., and P. Gent (2004), Anisotropic Gent-McWilliams parameterization for ocean models, *J. Phys. Oceanogr.*, *34*(11), 2541–2564.

Spitzer, W., and W. Jenkins (1989), Rates of vertical mixing, gas-exchange and new production: Estimates from seasonal gas cycles in the upper ocean near Bermuda, *J. Mar. Res.*, *47*(1), 169–196.

Stanley, R. H. R., J. B. Kirkpatrick, N. Cassar, B. A. Barnett, and M. L. Bender (2010), Net community production and gross primary production rates in the western equatorial Pacific, *Global Biogeochem. Cycles*, *24*, doi:10.1029/2009GB003651.

Sweeney, C., E. Gloor, A. R. Jacobson, R. M. Key, G. McKinley, J. L. Sarmiento, and R. Wanninkhof (2007), Constraining global air-sea gas exchange for CO$_2$ with recent bomb ^{14}C measurements, *Global Biogeochem. Cycles*, *21*(2), GB2015, doi:10.1029/2006GB002784.

Wanninkhof, R. (1992), Relationship between wind-speed and gas-exchange over the ocean, *J. Geophys. Res.*, *97*(C5), 7373–7382.

Westberry, T., M. J. Behrenfeld, D. A. Siegel, and E. S. Boss (2008), Carbon-based primary productivity modeling with vertically resolved photoacclimation, *Global Biogeochem. Cycles*, *22*(2), GB2024, doi:10.1029/2007GB003078.

Wiliams, P. J. L. (1993), On the definition of plankton production terms, *ICES Marine Sci. Symp.*, *197*, 9–19.

Williams, P., and D. Purdie (1991), In vitro and in situ derived rates of gross production, net community production and respiration of oxygen in the oligotrophic subtropical gyre of the North Pacific Ocean, *Deep Sea Res.*, *38*(7), 891–910, doi:10.1016/0198-0149(91)90024-A.