



Evaluating Southern Ocean biological production in two ocean biogeochemical models on daily to seasonal timescales using satellite chlorophyll and O₂/Ar observations

B. F. Jonsson¹, S. Doney³, J. Dunne², and M. L. Bender¹

¹Department of Geosciences, Princeton University, Princeton, New Jersey, USA

²Geophysical Fluid Dynamics Laboratory, NOAA, Princeton, New Jersey, USA

³Marine Chemistry and Geochemistry Department, Woods Hole Oceanographic Institution, Woods Hole, Massachusetts, USA

Correspondence to: B. F. Jonsson (bjonsson@princeton.edu)

Received: 3 April 2014 – Published in Biogeosciences Discuss.: 20 June 2014

Revised: 15 December 2014 – Accepted: 30 December 2014 – Published: 4 February 2015

Abstract. We assess the ability of ocean biogeochemical models to represent seasonal structures in biomass and net community production (NCP) in the Southern Ocean. Two models are compared to observations on daily to seasonal timescales in four different sections of the region. We use daily satellite fields of chlorophyll (Chl) as a proxy for biomass and in situ observations of O₂ and Ar supersaturation ($\Delta\text{O}_2/\text{Ar}$) to estimate NCP. $\Delta\text{O}_2/\text{Ar}$ is converted to the flux of biologically generated O₂ from sea to air (O₂ bioflux). All data are aggregated to a climatological year with a daily resolution. To account for potential regional differences within the Southern Ocean, we conduct separate analyses of sections south of South Africa, around the Drake Passage, south of Australia, and south of New Zealand.

We find that the models simulate the upper range of Chl concentrations well, underestimate spring levels significantly, and show differences in skill between early and late parts of the growing season. While there is a great deal of scatter in the bioflux observations in general, the four sectors each have distinct patterns that the models pick up. Neither model exhibits a significant distinction between the Australian and New Zealand sectors and between the Drake Passage and African sectors. South of 60° S, the models fail to predict the observed extent of biological O₂ undersaturation. We suggest that this shortcoming may be due either to problems with the ecosystem dynamics or problems with the vertical transport of oxygen.

1 Introduction

Recent years have seen an intense effort to better understand the global biogeochemical cycle. Scientific cruises organized by programs such as CLIVAR, WOCE, and GEOTRACES have generated a wealth of information about physical and chemical tracers in the global oceans, most of which have been aggregated to climatologies (e.g., Garcia et al., 2010). These field data, together with satellite observations of phytoplankton biomass, have helped in assessing the mean state and variability of the global marine ecosystems. Concurrently, a number of global ocean general circulation models (OGCMs) with added functionality to simulate biogeochemical processes have been developed, mainly to study trends and variabilities in earth's climate and the global carbon cycle.

This type of global climate model is predominately constructed and tuned to simulate decadal large-scale processes such as the global wind-driven and thermohaline circulation or the distribution of major water masses. However, an increased interest in biophysical interactions on smaller temporal and spatial scales, together with the recent ability to run global eddy-resolving OGCMs, has raised the question of how well coarser global climate models perform in daily to seasonal time domains. A model which has a high skill in estimating the basin-wide annual mean phytoplankton biomass might not correctly simulate process-level details in the seasonal cycle of biological production, such as the onset of spring blooms.

With this study, we aim to do such an evaluation by comparing two state-of-the-art OGCMs with observations of Chl and biological production. Since our main focus is the seasonal cycle and regional variability, we use two properties with high temporal and spatial resolution – satellite-derived Chl and in situ estimations of net community production (NCP) based on the $\Delta\text{O}_2/\text{Ar}$ method. Chl and NCP are particularly well suited because they represent the “end product” of bottom-up-driven biogeochemical models and are both of general interest in climate change studies. Satellite-derived net primary production (NPP) is not included in this study since it is closely correlated to Chl on the timescales of interest. We focus on the Southern Ocean south of 40°S since a large number of $\Delta\text{O}_2/\text{Ar}$ measurements have been collected from this area (Huang et al., 2012; Reuer et al., 2007; Casar et al., 2011). While satellite-derived Chl is a well-known and widely used property, $\Delta\text{O}_2/\text{Ar}$ based NCP will require a more detailed introduction.

The $\Delta\text{O}_2/\text{Ar}$ method was developed to estimate oceanic NCP by measuring the saturation of O_2 and Ar in the mixed layer (e.g., Spitzer and Jenkins, 1989). O_2 supersaturation occurs from both biological O_2 production and three physical processes: warming, changes in air pressure, and bubble entrainment. It is possible to decouple the biological component from the physical component by using the saturation of Ar, since Ar has similar physical properties as O_2 but is biologically inert. Following Craig and Hayward (1987), we define O_2 supersaturation relative to Ar as

$$\Delta\text{O}_2/\text{Ar} = \frac{(\text{O}_2/\text{Ar})_{\text{sample}}}{(\text{O}_2/\text{Ar})_{\text{eq}}} - 1. \quad (1)$$

This term is equivalent to the biological O_2 supersaturation. Knowing $\Delta\text{O}_2/\text{Ar}$, one can approximate the loss of biological O_2 via gas transfer across the air–sea interface (hereafter denoted O_2 bioflux) via the relationship

$$\text{O}_2 \text{ bioflux} \approx \kappa \cdot \Delta\text{O}_2/\text{Ar} \cdot \text{O}_2\text{eq}, \quad (2)$$

where κ is gas transfer velocity based on Wanninkhof (1992) and O_2eq the concentration of O_2 at equilibrium. It is possible to use either $\Delta\text{O}_2/\text{Ar}$ or O_2 bioflux when comparing observations and models. The relative advantages of the two properties are further discussed in Appendix A; we will use O_2 bioflux in this study.

One key challenge for OGCMs is to simulate the vertical exchange of water and tracers. Most physical processes that generate vertical mixing or advection act on length scales several orders of magnitude smaller than the model resolution. Instead, the models use different types of parameterizations to mimic the transport between surface waters and the deep ocean (Large et al., 1997; Doney et al., 2007). The results are evaluated by comparing model simulations with observed global distributions of transient tracers (e.g., radiocarbon or chlorofluorocarbons), especially in the deep sea. The boundary layer parameterizations embedded in OGCMs

are also often tested in 1-D against high-frequency observations (e.g., Large et al., 1994). It is, however, less clear how well vertical processes in the mixed layer and thermocline are represented by these parameterizations. One reason is that, until now, there have been few good methods for evaluating vertical transports on these spatial and temporal scales.

In this study we explore the feasibility of using O_2 bioflux to evaluate how well vertical processes are resolved on shorter timescales. This is possible due to the fact that the actual O_2 balance of the mixed layer is

$$d\text{O}_2/dt = \text{NCP} + (-\text{FO}_2 + D_{\text{in}})/h_{\text{ml}}, \quad (3)$$

where $d\text{O}_2/dt$ is the time rate of change of dissolved oxygen in units of $\text{mmol m}^{-3} \text{day}^{-1}$, FO_2 is the sea to air gas exchange flux, D_{in} is net input (or loss) of O_2 to the mixed layer from ocean physics (i.e., mixing and advection), and h_{ml} is the mixed layer depth (Ducklow and Doney, 2013). Assuming steady state, this equation can be rewritten as a negative relationship between O_2 bioflux and the net downward vertical transport of O_2 :

$$\text{NCP} = \text{O}_2 \text{ bioflux} + \text{O}_2 \text{ vertical flux}. \quad (4)$$

O_2 bioflux can hence be used to give a combined evaluation of how well models simulate upper ocean biogeochemical rate processes, sea–air O_2 fluxes, and vertical mixing. Such a combined assessment is valuable since the interaction between physics and biology is a key source for variability on short timescales (days to months).

We will first compare regional and seasonal patterns in Chl and O_2 bioflux between observations and models, continue with exploring how O_2 bioflux can show discrepancies in how the models simulate vertical transports, and finally discuss how the results could help to identify mechanisms that contribute to mismatches between observations and models.

2 Methods

The main focus of the study is to compare model output with in situ observations and satellite-derived properties. The data from each source have different temporal and spatial resolution and span different ranges in time. To compensate for these discrepancies, we regrid in situ observations and satellite fields to a model grid with roughly 1° resolution at equal intervals in time. We also combine observations from different years but the same day of the year (e.g., 1 January 2003, 1 January 2004, 1 January 2005) to a climatological year with a daily time resolution.

2.1 Satellite data

We use remotely observed chlorophyll concentrations from MODIS (Moderate Resolution Imaging Spectroradiometer)/Aqua on the Level-3 $9\text{ km} \times 9\text{ km}$ grid. Daily satellite images from 2003 to 2010 were aggregated to the model grids

and averaged to a day-of-the-year climatology. Satellite data, particularly in the Southern Ocean, suffers from a high frequency of days where clouds, light conditions, sea ice, or other problems disqualify the observations. The relative frequency of days with valid information in our data set is between 5 and 15 % on the original grid and between 20 and 50 % on the aggregated model grid.

2.2 In situ observations

We use $\Delta O_2/Ar$ observations from 19 Southern Ocean cruises between 1999 and 2009, all occurring during the austral summer. The geographical locations of the respective ship tracks are presented in Fig. 1. The measurements were conducted by two different methods: water was collected in bottles and analyzed in the lab (sampling locations shown in blue in Fig. 1) on 16 cruises, and $\Delta O_2/Ar$ was measured directly using a ship-borne flow-through system on 3 cruises (shown as red in Fig. 1). The measurements are clustered in space and time reflecting tracks of the ships of opportunity used in this study. We use these sampling clusters in our analysis as natural areas to compare and contrast different parts of the Southern Ocean. A more detailed description of the sampling strategies, measurement methods, and data sources can be found in Reuer et al. (2007) and Cassar et al. (2009, 2011).

Bioflux is calculated as the product of the biological O_2 supersaturation and the gas transfer velocity. The latter is determined using the Wanninkhof (1992) parameterization expressing gas transfer velocity in terms of a quadratic function of wind speed and the Schmidt number. We do the calculation using daily averages of the NESDIS wind product with a 0.5° resolution based on data from the QuikSCAT (Quick Scatterometer) satellite (Ebuchi et al., 2002). The gas transfer velocity for each $\Delta O_2/Ar$ measurement is calculated from the daily mean local wind speeds during the 60 days preceding collection. A time-weighted value for the gas transfer velocity is calculated based on the fraction of the mixed layer flushed in each subsequent interval until sampling (Reuer et al., 2007; Bender et al., 2011). The resulting gas transfer velocity is then used in Eq. (3) to calculate O_2 bioflux from $\Delta O_2/Ar$ supersaturation. A detailed analysis of possible uncertainties associated with the method can be found in Jonsson et al. (2013).

Finally we use mixed layer depth (MLD) as a baseline diagnostic of how well vertical physical processes in the surface ocean are resolved. A total of about 75 000 vertical profiles from Argo floats between 2001 and 2012 are used to estimate in situ MLDs for the different regions. The observed MLD is defined as the depth where density is 0.03 kg m^{-3} higher than at the most shallow observation.

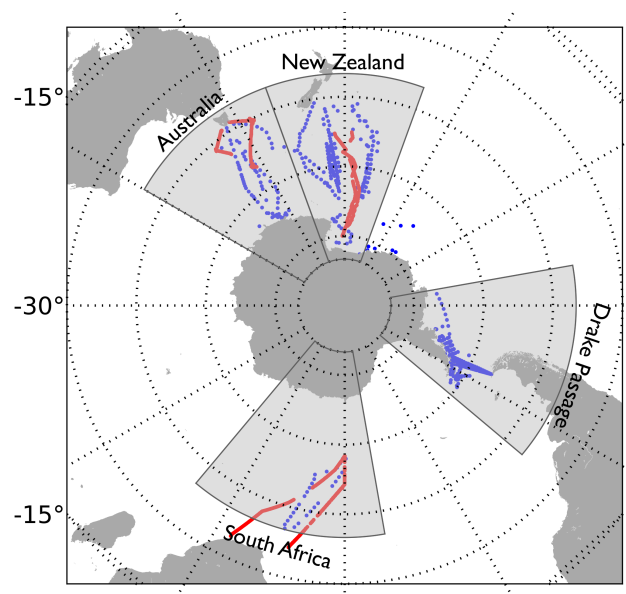


Figure 1. Map of O_2 observations used in the study. Grey shadings signifies the four regions we focus on in this study. Dots indicate discreet sampling locations, while red lines indicate continuous sampling.

2.3 Models

The observations are compared with output from the ocean components of two Intergovernmental Panel on Climate Change (IPCC)-class ocean biogeochemical models. The TOPAZ (Tracers in the Ocean with Allometric Zooplankton) ocean model is based on version 4 of the modular ocean model (MOM4; Griffies et al., 2005) with a vertical z coordinate and a horizontal B-grid with a tripolar coordinate system (North America, Siberia, and Antarctica) to resolve the Arctic. The model has a nominally 1° horizontal resolution globally, with higher meridional resolution near the equator (to $1/3^\circ$). There are 50 vertical layers; resolution is 10 m in the upper 200 m, and coarser below. MOM4 includes a representation of the K-profile parameterization (KPP) planetary boundary layer scheme (Large et al., 1994), Bryan–Lewis deeper vertical mixing, 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. For these studies, the ocean model is forced by prescribed boundary conditions from the reanalysis effort of the ECMWF (European Centre for Medium-Range Weather Forecasts) and NCAR (National Center for Atmospheric Research) Common Ocean-ice Reference Experiments (CORE).

BGCCSM (BioGeochemistry Community Climate System Model) is based on the Los Alamos National Laboratory Parallel Ocean Program (POP) (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 midlatitudes to about 0.8° at high and low latitudes. The surface layer is 12 m thick; in total, there are 5 layers to 111 m and 25 layers to the bottom. This model invokes Gent–McWilliams' isopycnal mixing and the KPP upper ocean model and is also forced by prescribed boundary conditions from CORE.

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 (National Centers for Environmental Prediction) surface winds, and the quadratic dependence of the gas exchange coefficient on wind speed (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 parameterization as O_2 .

Both models include complex biogeochemistry/ecosystem components with macro- and micronutrients, organic matter, and three phytoplankton functional groups: small phytoplankton, large phytoplankton, and diazotrophs. Both invoke co-limitation by iron, light, and nitrogen, mediated in part by their influence on the Chl:C ratio, and slower rates of photosynthesis at lower temperatures. Diazotrophs have high N:P ratios and low photosynthetic efficiencies (TOPAZ) or high iron requirements (TOPAZ, BGCCSM). NCP is calculated as the difference between production and consumption of carbon by the different functional groups.

It is possible to use either $\Delta O_2/Ar$ or O_2 bioflux for comparing observations with models. One could argue that $\Delta O_2/Ar$ is a more robust property since it is an observed quantity, whereas O_2 bioflux is a derived product that depends on the wind field. On the other hand, model $\Delta O_2/Ar$ depends on the choice of wind forcing, whereas O_2 bioflux is assumed to be directly linked to NCP. Our tests show that both methods give similar accuracy, and we choose to use O_2 bioflux since its units are more appropriate for the current study (see Appendix A for further discussion).

3 Results

First we test the feasibility of aggregating data from different years into a single climatological year. This approach is useful only if the difference between seasons is significantly larger than the interannual variability. We test the assumption by comparing how model O_2 bioflux changes from one date to another in a climatological year against the standard deviation at the same dates between a number of years. As a test case, fields from BGCCSM were used to generate climatologies for the days 15 November and 15 December by averaging data for the entire Southern Ocean from four consecutive model years. Our results show that the mean difference is $19.8 \text{ mmol } O_2 \text{ m}^2 \text{ d}^{-1}$ between the two different

dates, whereas the standard deviation over 4 years on the respective dates is only $8.7 \text{ mmol } O_2 \text{ m}^{-2} \text{ d}^{-1}$. It is hardly an unexpected result that spring values of O_2 bioflux differ from summer since the Southern Ocean is a high-latitude region. This result encourages us to aggregate observations and model data from different years into a climatological year.

Next, we compare the model simulations with observations. Chlorophyll simulated by BGCCSM and TOPAZ is related to satellite observations retrieved by the MODIS/Aqua mission from 1 January 2003 to 31 December 2010. Each daily satellite image is reprojected from a native $9 \text{ km} \times 9 \text{ km}$ resolution to the TOPAZ $1^\circ \times 1^\circ$ grid, and the fields are aggregated to a climatological year, as mentioned earlier. The resulting data sets are analyzed in four geographical sectors, shown as shaded areas in Fig. 1. We zonally average the data in each sector to a Hofmøller diagram with latitude on the y axis, time on the x axis, and daily zonal Chl averages as colors (Figs. 2–5). It should be noted that one problem with using satellite-retrieved chlorophyll is a systematic lack of data during winter due to low-light conditions and sea-ice cover in combination with the satellite's track. These periods are shown as gray areas in the figures. Figure 6 shows a more detailed representation of the data at 60° south.

Figures 7–10 present model NCP and bioflux vs. our observations. Panel a in each figure shows the temporal evolution of NCP vs. latitude in the two models, and panel b the corresponding model O_2 bioflux. The sampling locations are indicated on the model plots by gray circles. Panel c, finally, presents O_2 bioflux from the observations shown in Fig. 1. As mentioned earlier, all measurements that fall on the same day of the year and grid cell are combined to one mean value, which is indicated by the color of the dot. Note that the aggregated values of observed O_2 bioflux presented together with TOPAZ are somewhat different from the ones connected to BGCCSM, since the two model grids are different.

A general pattern arises where the models simulate upper ranges of Chl well but underestimate spring levels significantly. Each model also shows differences in skill between early and late parts of the growing season, performing better during the spring and early summer. While there is a great deal of scatter in the bioflux observations, the four sectors each show distinct patterns, whereas the models exhibit little distinction between the Australian and New Zealand sectors and between the Drake Passage and African sectors. Next, we compare the simulated fields and data for each of the study regions in more detail.

3.1 Drake passage

The seasonal change in satellite-derived chlorophyll for the Drake Passage sector is shown in the top panel of Fig. 2. It is possible to discern a seasonal cycle, even though we lack winter and early spring data in higher latitudes. Towards the south, the earliest retrieved concentrations are between three and five times or more lower than maximum values at the late

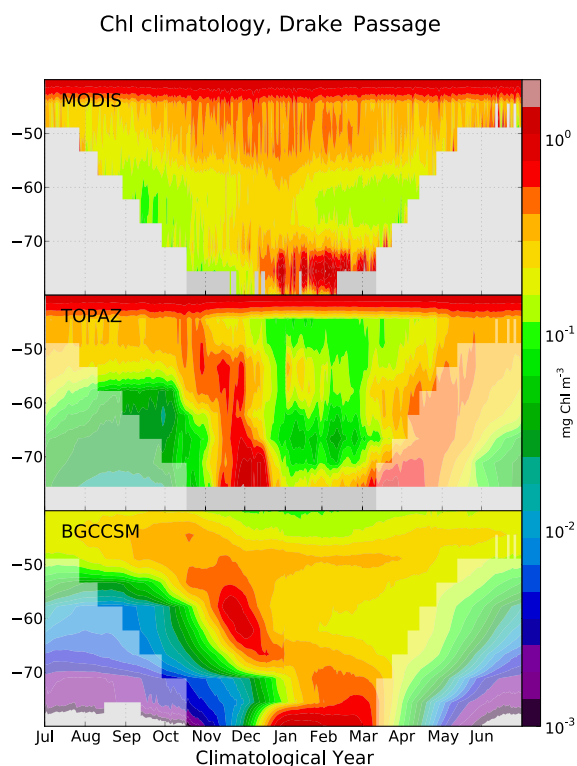


Figure 2. Hofmøller plots of satellite Chl (top panel), Chl simulated by TOPAZ (middle panel), and Chl simulated by BGCCSM (bottom panel) in Drake Passage. All panels are zonal medians within the box. All data that fall on a specific grid cell on a specific day of the year are averaged to one value.

spring/summer peak. Both the magnitude and timing of maximum summer Chl concentrations vary with latitude, with the strongest and latest blooms occurring in high latitudes. [Chl] reaches 1 mg m^{-3} south of 70° S , decreases to 0.5 mg m^{-3} between 60 and 70° S , and increases again north of 50° S . The growing season, as inferred from the period of elevated summer Chl, is about 3 months south of 70° S and generally lengthens going to the north. The bands of very high Chl north of 45° S and south of 70° S are likely due to the transport of sedimentary iron from coastal waters generating elevated biological production.

Panels b and c present Chl climatologies from TOPAZ and BGCCSM. The grey shadings indicate where satellite coverage is missing. It is clear that the seasonal chlorophyll progression in both models diverges from observations, whereas the magnitude of peak Chl concentrations is simulated quite well. TOPAZ underestimates Chl concentrations somewhat early in the season and generates an intense spring bloom extending over the entire sector from 75 to 40° S . (“Bloom” as used here indicates a transient period of high [Chl] lasting no more than 1 month.) In the observations, high [Chl], once established, continues throughout most of the summer, and concentrations $> 1 \text{ mg m}^{-3}$ are limited to the region south

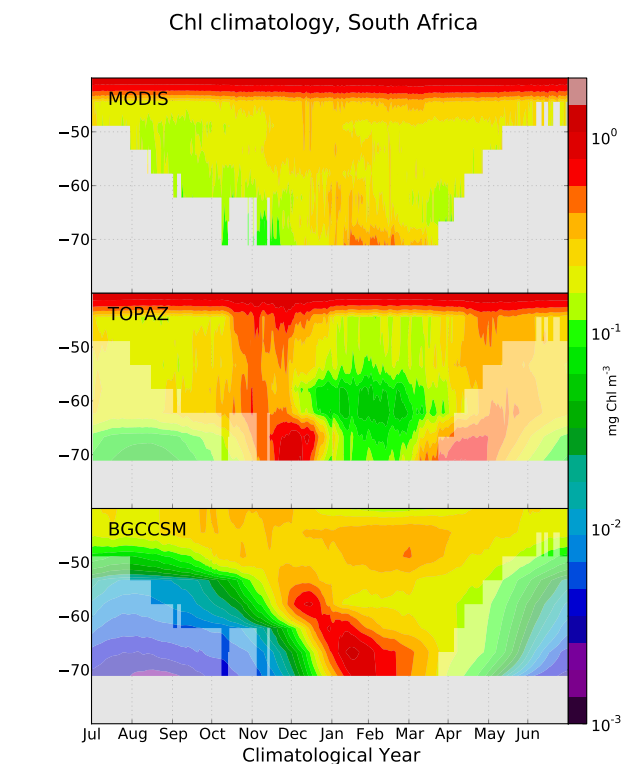


Figure 3. Hofmøller plots of satellite Chl (top panel), Chl simulated by TOPAZ (middle panel), and Chl simulated by BGCCSM (bottom panel) south of South Africa. All panels are zonal medians within the box. All data that fall on a specific grid cell on a specific day of the year are averaged to one value.

of 70° S . In these southern regions, the onset of the TOPAZ bloom comes 1–2 months earlier than in the observations. The spring bloom in TOPAZ collapses after about 1 month with concentrations of Chl that are too low over the austral summer as a result. Finally, the season ends with TOPAZ generating a weaker fall bloom during part of the season without satellite coverage. South of 50° S , BGCCSM begins the season with much lower Chl concentrations than both TOPAZ and satellite observations. The model generates a much more intense bloom between 60 and 70° S than observations suggest but has good skill in predicting the timing and magnitudes south of 70° S . Both models suggest that during the bloom, biomass is significantly higher in the frontal regions compared to observations.

O_2 bioflux observations in the Drake Passage region originates both from transects crossing the Drake Passage and nearby cruises in the Palmer Long-Term Ecological Research program annual survey west of the Antarctic Peninsula. Our results are thus partly influenced by coastal processes outside the models’ domain. In the Drake Passage sector, TOPAZ and BGCCSM show similar patterns in NCP (panel a) north of 65° S . South of 65° S , TOPAZ simulates intensive NCP over the summer, whereas BGCCSM has close to zero net

Chl climatology, Australia

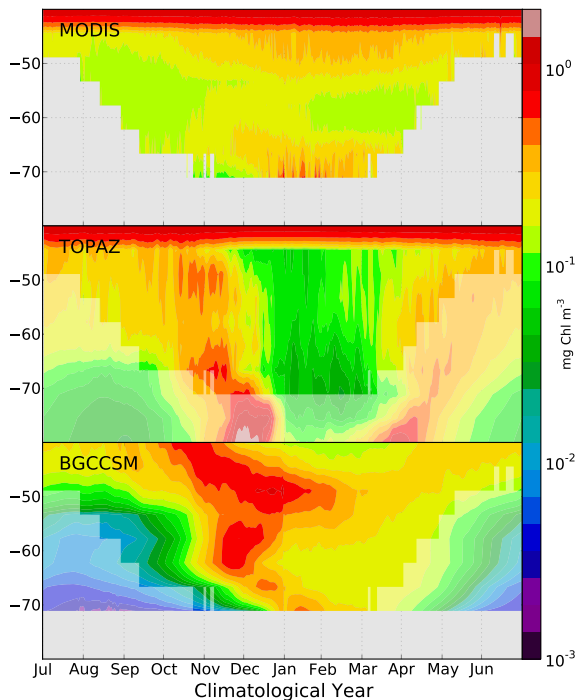


Figure 4. Hofmøller plots of satellite Chl (top panel), Chl simulated by TOPAZ (middle panel), and Chl simulated by BGCCSM (bottom panel) south of Australia. All panels are zonal medians within the box. All data that fall on a specific grid cell on a specific day of the year are averaged to one value.

production. Both models have predominantly negative O_2 biofluxes south of $50^\circ S$ in October (panel b) and a progressive change to positive fluxes from north to south with time. The negative values of O_2 bioflux in TOPAZ are about twice as large as they are in BGCCSM. In TOPAZ, O_2 bioflux is positive throughout the domain in January, whereas in BGCCSM there is a southerly region of negative flux throughout the season.

Panel c of Fig. 7 shows observations of O_2 bioflux in the region. The pattern of negative fluxes south of $60^\circ S$ in November and December corresponds qualitatively with the models. There is a wealth of observations south of $60^\circ S$ in January and February, but they show significant variability. Some observations suggest biological production rates of up to $50 \text{ mmol m}^{-2} \text{ d}^{-1}$, which is similar to the levels predicted by TOPAZ. Others correspond to the results from BGCCSM with O_2 bioflux estimates near or below 0. The mean observed O_2 bioflux values in January and February south of $65^\circ S$ are $8 \pm 20 \text{ mmol } O_2 \text{ m}^{-2} \text{ d}^{-1}$. The corresponding values for BGCCSM are 9 ± 13 , and for TOPAZ, 16 ± 27 . The observations in April suggest undersaturated conditions that correspond with BGCCSM but not with TOPAZ.

Chl climatology, New Zealand

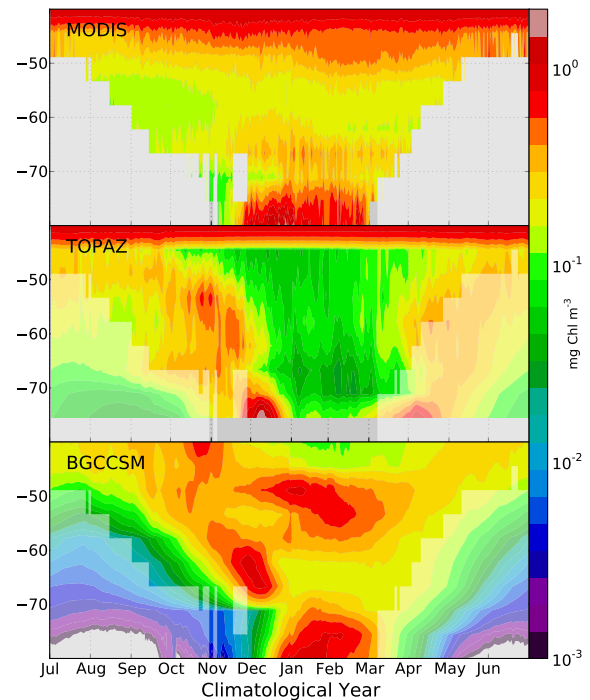


Figure 5. Hofmøller plots of satellite Chl (top panel), Chl simulated by TOPAZ (middle panel), and Chl simulated by BGCCSM (bottom panel) south of New Zealand. All panels are zonal medians within the box. All data that fall on a specific grid cell on a specific day of the year are averaged to one value.

3.2 South Africa

The observations in the South Africa sector (Fig. 3, top panel) show similar patterns to the Drake Passage sector for satellite-derived Chl, with the main exception of a weaker bloom north of $50^\circ S$. Note the lack of observations south of $70^\circ S$ where of Chl concentrations may be very high. TOPAZ (panel b) begins the season with somewhat higher concentrations of Chl than observed and, with time, generates a significantly stronger bloom. As in the Drake Passage, the TOPAZ bloom crashes midsummer, after which model Chl concentrations are significantly lower than observations. BGCCSM starts the season with much lower Chl concentrations than the satellite data and generates an exaggerated bloom as well, especially south of $55^\circ S$.

NCP and O_2 bioflux climatologies (Fig. 8) show that TOPAZ has an earlier and more intense band of positive NCP than BGCCSM (panel a). BGCCSM has a slow southward progression of the positive NCP band with NCP reaching $20 \text{ mmol m}^{-2} \text{ d}^{-1}$ at $40^\circ S$ in October and at $65^\circ S$ in February. TOPAZ, on the other hand, has a more uniform pattern, in which most of the region reaches these levels of NCP by mid-November. The two models have O_2 bioflux

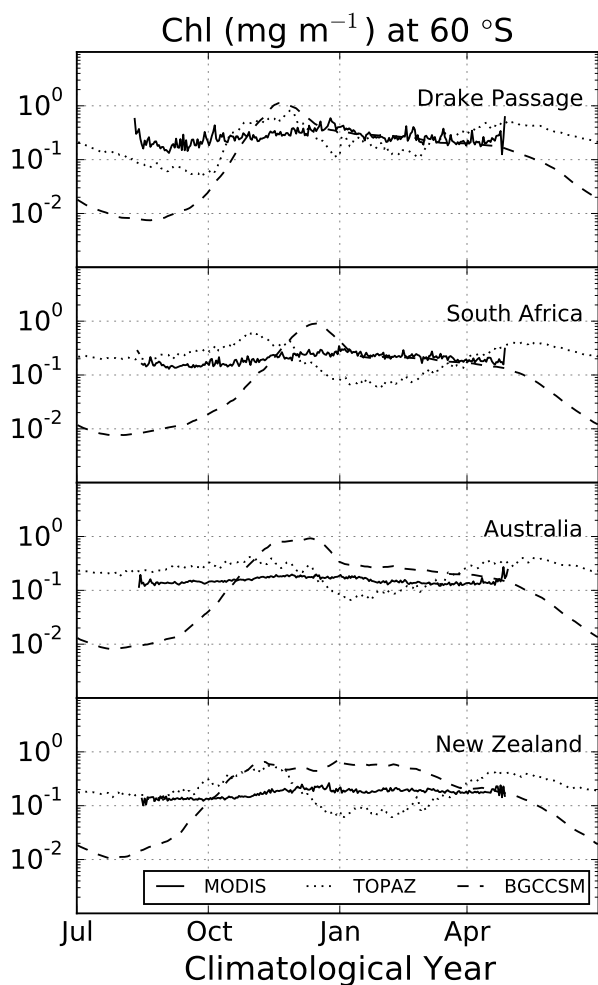


Figure 6. Detail of the Chl concentrations in mg m^{-3} from Figs. 2–5. The lines show a slice of each original panel at 60°S . All data that fall on a specific grid cell on a specific day of the year are averaged to one value.

patterns (shown in panel b) that vary in a fashion similar to that of the Drake Passage region. Both models start with a negative O_2 bioflux: $< -30 \text{ m}^2 \text{ d}^{-1}$ in TOPAZ and $\sim -10 \text{ mmol m}^{-2} \text{ d}^{-1}$ in much of the domain in BGCCSM. This negative bioflux switches to positive values faster in TOPAZ than in BGCCSM, most likely due to the earlier onset of biological production. Finally, BGCCSM has a slightly earlier and stronger switch to negative O_2 bioflux at the end of the growing season than does TOPAZ.

We have O_2 bioflux observations from four crossings in the South Africa sector (Fig. 8) distributed in time from December to late March. The November–December transect has a pattern of weakly positive O_2 bioflux in the north that turns negative below 50°S . Both models show a somewhat different pattern with positive O_2 bioflux south of 50°S as well. While both models show positive values, the observed January transect has predominantly negative biofluxes. Finally,

the two transects in March show a pattern with high positive O_2 bioflux in the north and significant negative bioflux at 55°S . BGCCSM has a better skill in recreating these patterns than does TOPAZ.

3.3 Australia and New Zealand regions

Chl in the Australian and New Zealand sectors (Figs. 4 and 5) follow the general pattern of Drake Passage and South Africa. The main exception is BGCCSM generating spring blooms further north and in a spottier pattern than in the sectors discussed earlier. In this model, the New Zealand sector has four distinct areas/periods of high Chl concentrations: in January–February south of 70°S , November–December between 60 and 70°S , January between 45 and 55°S , and October–November north of 45°S . This peculiar pattern could be due to interactions with hydrology in the frontal regions. The Australian sector shows similar patterns for the latter three areas, whereas the region south of 70°S lacks data. TOPAZ shows indications of being out of phase with both MODIS and BGCCSM after the beginning of December, whereas BGCCSM captures the seasonal cycle in MODIS better.

NCP climatologies from the Australian sector (Fig. 9, top panel) show that TOPAZ generates short intense periods of positive NCP early in the growing season. In the southern reach of the domain, NCP stays high into the fall, whereas elsewhere summer and fall NCP values are low ($10 \text{ mmol m}^{-2} \text{ d}^{-1}$). BGCCSM, on the other hand, has a much longer and more intense period of positive NCP than in other parts of the Southern Ocean. In this model, spring conditions south of 45 – 50°S are dominated by weaker negative O_2 biofluxes when compared to other regions in both models (panel b). BGCCSM switches to negative bioflux at the end of the summer, whereas O_2 bioflux in TOPAZ stays positive throughout March.

The observations in this part of the ocean differ from those simulated by TOPAZ. For example, almost all measurements north of 55°S report higher O_2 biofluxes, $> 25 \text{ mmol m}^2 \text{ d}^{-1}$, in February. These field data also exhibit low or even negative biofluxes at 65°S where TOPAZ predicts strong biological production and high O_2 bioflux from mid-November to mid-February. BGCCSM is more in line with observed patterns, with the main exception of an early onset of negative O_2 bioflux in the fall at 40 – 50°S in the model where the observations still suggest strong positive biofluxes. In general, between 50 and 60°S , TOPAZ and BGCCSM both simulate a strong spring bloom, whereas observations show a simpler pattern of sustained high production throughout the latter part of spring and summer. Model NCP and O_2 bioflux in the New Zealand region are similar to values in the Australian sector (Fig. 10). The O_2 bioflux observations in panel c show considerable scatter but in general are marked by high values north of the polar front at 60°S , with low or negative values of bioflux to the south at most times. Both models capture the

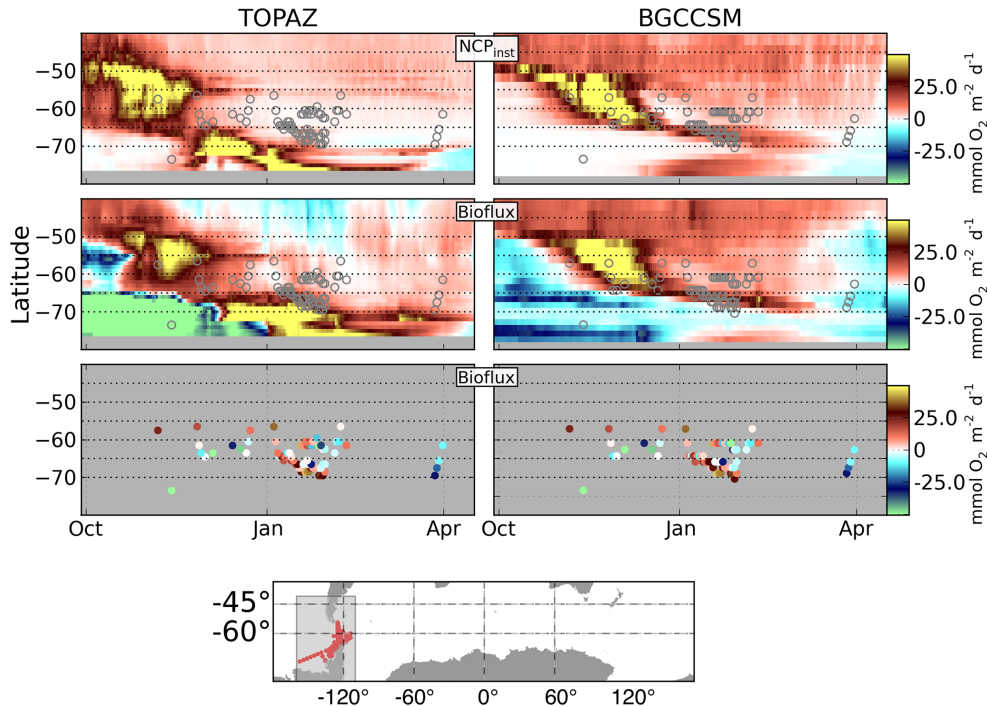


Figure 7. Hofmøller plots of model NCP ($\text{mmol m}^{-2} \text{d}^{-1}$, top panel), model O_2 bioflux ($\text{mmol m}^{-2} \text{d}^{-1}$, middle panel), and observed O_2 bioflux ($\text{mmol m}^{-2} \text{d}^{-1}$, bottom panel) in Drake Passage (locations of observations are presented as red points on the map). All model values are zonal medians within the box. All observations that fall on a specific grid cell on a specific day of the year are averaged to one value.

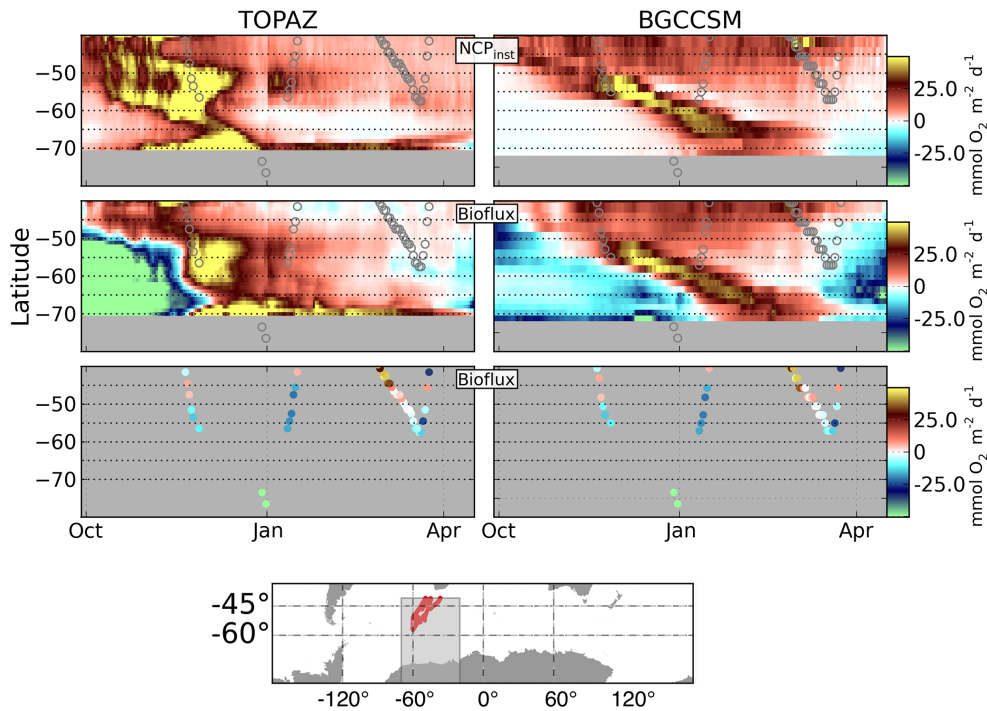


Figure 8. Hofmøller plots of model NCP ($\text{mmol m}^{-2} \text{d}^{-1}$, top panel), model O_2 bioflux ($\text{mmol m}^{-2} \text{d}^{-1}$, middle panel), and observed O_2 bioflux ($\text{mmol m}^{-2} \text{d}^{-1}$, bottom panel) in the Southern Ocean south of South Africa (locations of observations are presented as red points on the map). All model values are zonal medians within the box. All observations that fall on a specific grid cell on a specific day of the year are averaged to one value.

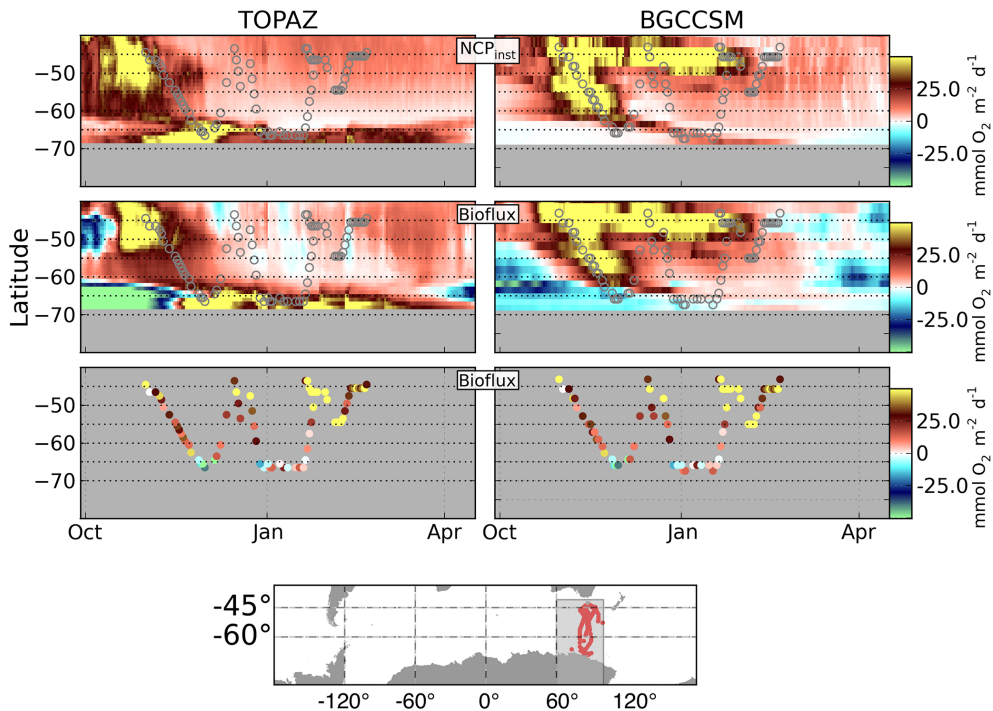


Figure 9. Hofmøller plots of model NCP ($\text{mmol m}^{-2} \text{d}^{-1}$, top panel), model O_2 bioflux ($\text{mmol m}^{-2} \text{d}^{-1}$, middle panel), and observed O_2 bioflux ($\text{mmol m}^{-2} \text{d}^{-1}$, bottom panel) in the Southern Ocean south of Australia (locations of observations are presented as red points on the map). All model values are zonal medians within the box. All observations that fall on a specific grid cell on a specific day of the year are averaged to one value.

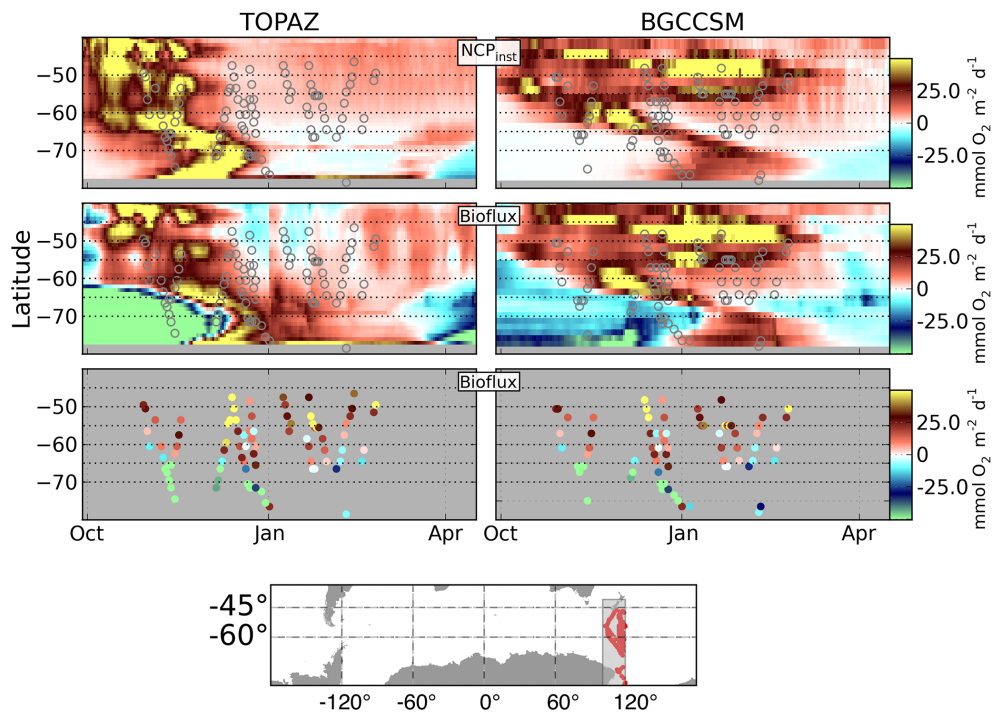


Figure 10. Hofmøller plots of model NCP ($\text{mmol m}^{-2} \text{d}^{-1}$, top panel), model O_2 bioflux ($\text{mmol m}^{-2} \text{d}^{-1}$, middle panel), and observed O_2 bioflux ($\text{mmol m}^{-2} \text{d}^{-1}$, bottom panel) in the Southern Ocean south of New Zealand (locations of observations are presented as red points on the map). All model values are zonal medians within the box. All observations that fall on a specific grid cell on a specific day of the year are averaged to one value.

highly negative values of bioflux early in the growing season south of 60–65° S and occasionally negative values of bioflux later in the growing season. BGCCSM simulates our observations of sustained high NCP north of the polar front from November through February.

3.4 Cross-regional, Southern Ocean analysis

The data–model comparisons of O₂ bioflux for the different regions have certain patterns in common. In both models, the spring period of strong positive flux starts later in high latitudes. In general, TOPAZ tends to have an earlier, shorter, and more intense period of high Chl concentrations than does BGCCSM. Observed values of O₂ bioflux are much more variable than simulated values. This is expected, given the relatively smooth and coarse fields of the models, and likely reflects the absence of mesoscale processes in the models.

Our next step is to compare the seasonal range in O₂ bioflux values between observations and models as a function of latitude (Fig. 11). We aggregate the observations to days of the year and model grid cells as described above but compare the resulting values with corresponding individual data points in the models matched by both location and collection time. Such comparison using individual data points suffers more from small-scale spatial mismatches than the zonal model averages used earlier, but it allows us to better compare the seasonal range of O₂ bioflux values between models and observations. Figure 11 shows O₂ bioflux vs. latitude for the two models and observations in each of the previously defined regions. We find that BGCCSM generally predicts the meridional variability of ranges in O₂ bioflux, suggesting that processes constraining NCP are simulated well. The seasonal maximum of model O₂ bioflux might not occur at the same time in the models as in the real world, but the magnitude of the maximum values seems to be fairly well predicted. Equatorward of about 60° S, the models also tend to capture range and meridional structure of negative O₂ bioflux, which reflects both low wintertime NCP and physical transport. In contrast, the models do not capture well the observed strong negative O₂ bioflux at high latitudes. TOPAZ also tends to exaggerate high positive O₂ bioflux in some areas, such as 60–70° S in the Drake Passage and New Zealand regions, whereas O₂ bioflux is underestimated in TOPAZ between 40 and 50° S in the Australian and New Zealand regions.

We compare observed and simulated O₂ bioflux values for the same locations and times. The scatterplot in Fig. 12 compares O₂ bioflux simulated by TOPAZ (blue) and BGCCSM (red) with observations binned to the same days of the year on the respective model's grid. It is clear from this figure that both models show a low correlation with the observations ($r^2 = 0.024$ for TOPAZ and $r^2 = 0.23$ for BGCCSM). This low correlation is expected: lags in time or displacements in space can generate large differences between the models and observations even if the fundamental processes are simulated

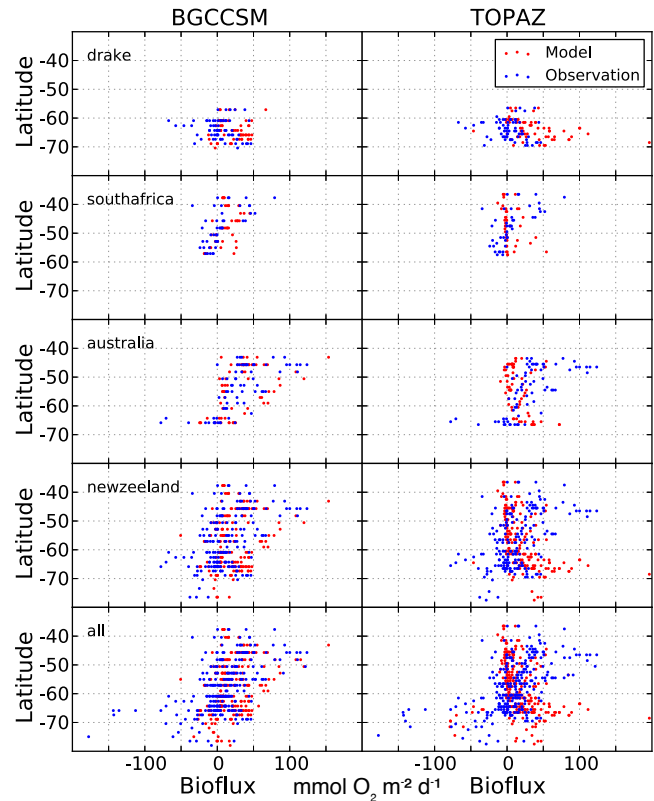


Figure 11. Scatterplots of observed (blue) and model (red) O₂ bioflux ($\text{mmol m}^{-2} \text{d}^{-1}$) versus latitude in four regions of the Southern Ocean. All observations that fall on a specific grid cell at a specific year day are averaged to one value. Models are subsampled at the location and year day of the observations.

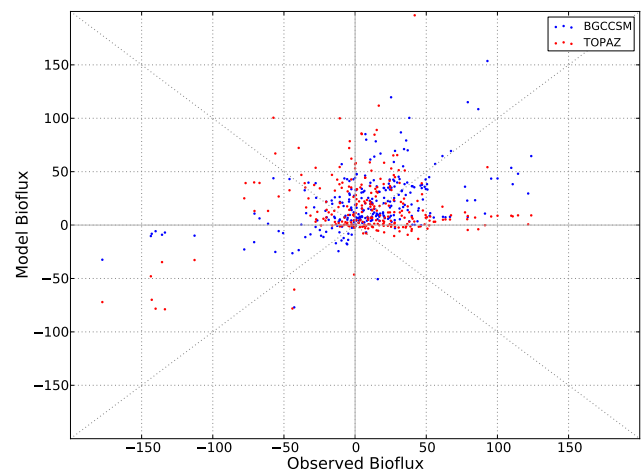


Figure 12. Scatterplots of model versus observed O₂ bioflux ($\text{mmol m}^{-2} \text{d}^{-1}$) for the observational sampling sites shown in Fig. 1. All observations that fall on a specific grid cell at a specific day of the year are averaged to one value. Models are subsampled at the location and day of the year of the observations.

with high skill (Doney et al., 2009). Further, the field data contain mesoscale variability that cannot be captured in the models (even if the models were eddy-resolving, the details of the simulated turbulent fields would differ from observed). It is also clear that the distributions of model-data residuals (difference from a diagonal 1 : 1 line) are asymmetrical. The upper-right quadrant, where both observations (obs) and model data (mod) are positive, has considerable scatter about the 1 : 1 line but no apparent bias. The lower-right quadrant (negative model, positive observation) is mainly empty, showing that the model rarely predicts undersaturation when the observations report supersaturation. An exception is the clustering of TOPAZ NCP values around 0, a consequence of low production simulated by that model after an intense bloom. The upper-left quadrant is heavily populated, showing that the models frequently simulate positive values of bioflux when the ocean is in fact undersaturated. Consistent with this pattern, the results in the lower-left quadrant show that, when data and models agree that bioflux is negative, the observations are more negative than the models. The overall picture is that both models have a positive bias in predicting bioflux, mainly due to fewer negative values compared to the observations (Fig. 12). For all data points the model bias is statistically significant for both a paired t test (BGCCSM: $n = 271$, $t = -4.803$, $p = 0.000$; TOPAZ: $n = 273$, $t = -2.870$, $p = 0.004$) and a one-tailed binomial test (BGCCSM: pos = 167, tot = 271, $p = 0.00005$; TOPAZ: pos = 151, tot = 273, $p = 0.035$). Neither model shows a statistically significant bias when only positive values are considered.

3.5 Mixed layer depths

Finally, we create MLD climatologies for the different regions by integrating Argo and model MLDs using methods mentioned earlier (Fig. 13). Both models are able to simulate the general trends rather well, with deep mixed layers in winter, shoaling in spring and shallow mixed layers during the summer. Regional structures are also similar in general, with the main exception of deep winter mixing extending too far south in TOPAZ, even if this is somewhat inconclusive due to lack of observations. The two main differences between models and observations are that the spring shoaling tend to be later and more gradual in the observations. There is also much smaller-scale variability in the Argo climatology, which could be explained by sparse data but also by the models having a mixed layer dynamics that are too smooth.

4 Discussion

When comparing model and satellite climatologies of Chl concentrations and O₂ bioflux, we find both similarities and significant differences. The models are able to predict spring and summertime maximum levels of Chl and O₂ bioflux

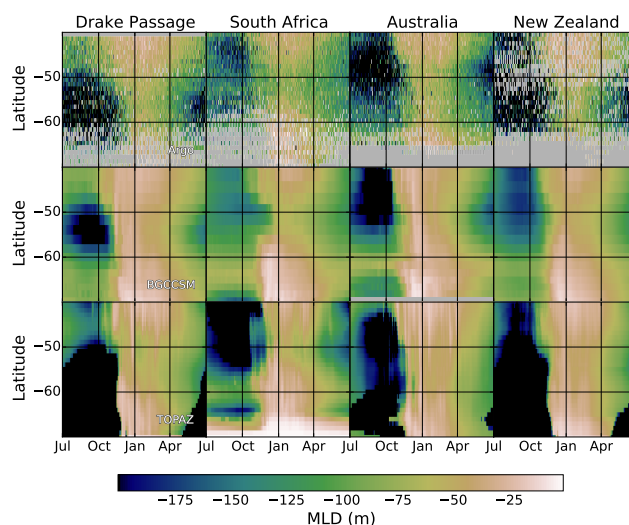


Figure 13. Climatological Hofmøller plots of MLD (m) in different regions of the Southern Ocean. All model values are zonal medians within the box. All observations that fall on a specific grid cell on a specific year day are averaged to one value.

well, but levels are much too low during the winter and early spring. Such underestimations are particularly important in the case of Chl, which by nature has a lognormal distribution (Campbell, 1995) and is hence skewed towards low values. One explanation for this behavior is a combination of model grazing and/or phytoplankton mortality being too strong in the winter, as reported for the BGCCSM in the subpolar North Atlantic by Behrenfeld et al. (2013). These patterns could also be explained by the two models simulating vertical export of phytoplankton that is too weak during summer and export that is too strong during winter. Johnson et al. (2013) have shown that MODIS/Aqua generally underestimate the dynamical range of Chl in the Southern Ocean, which would suggest that these differences might be even larger.

Another general pattern is that the models tend to simulate the increases in Chl during spring and early summer rather well but show highly diverging behavior later in the season. In TOPAZ, the ecosystem tends to crash in early January with much too low biomass as an effect, whereas BGCCSM has patches where far too much Chl is produced. Work by Hashioka et al. (2013) suggest that such differences in skill over the season could be explained by changes in processes that control the ecosystem. The spring bloom onset is thought to be controlled mainly by physical factors such as light, temperature, and vertical stratification, whereas the summer peak magnitude is thought to be controlled by nutrient availability, grazing, mortality, and other ecological and biogeochemical factors. Finally, the decrease in Chl concentrations after a summer peak mainly depends on ecosystem dynamics such as grazing, succession, and other interactions between organisms, as well as vertical export of particulate organic matter. Phytoplankton production is in general better resolved by the

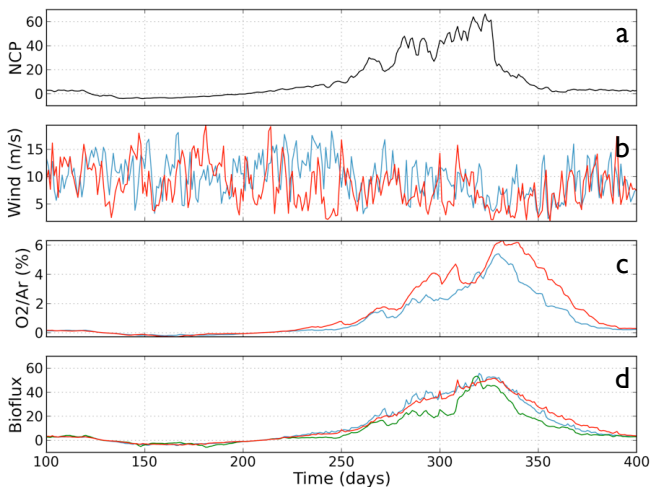


Figure 14. Panel (a) shows mixed-layer integrated NCP from the TOPAZ model at 61.5° S, 159.5° W over a 400-day period. Panel (b) shows two wind time series for the same time. The blue line represents the NCEP/CORE reanalysis winds from the same location as the model NCP data, and the red line represents Quikscat-satellite-derived winds from the same location. The two lines in panel (c) represent the resulting $\Delta\text{O}_2/\text{Ar}$ supersaturation from a box model simulation based on the time series in panels (a) and (b). Panel (d) shows O_2 bioflux calculated from the time series in panels (b) and (c). The red line is based on reanalysis winds and $\Delta\text{O}_2/\text{Ar}$, the blue line is based on Quikscat winds and $\Delta\text{O}_2/\text{Ar}$, and the green line is based on Quikscat winds and $\Delta\text{O}_2/\text{Ar}$ calculated from reanalysis winds.

models, whereas heterotrophic processes and vertical transports are challenging to implement well. These processes are often stochastic in their behavior, and there is a lack of observation to parameterize them accurately.

With respect to models and observations, the misfit in spring Chl biomass is likely explained by the differences in mixed layer dynamics between models and observations. Consistent winter mixed layers that are too deep followed by a rapid shoaling without much small-scale variability, as observed in the models, would lead winter biomass that is too low and exaggerated spring blooms, again as observed. Different factors, such as the timing and progression of spring-time mixed-layer shoaling, diapycnal mixing, and mesoscale or submesoscale processes, can strongly affect the onset of the spring bloom since the availability of light and nutrients in the mixed layer will be impacted. Problems with MLD dynamics have been shown by earlier studies in climate models similar to TOPAZ and BGCCSM (e.g., Fox-Kemper et al., 2011, and references therein).

Both models show significant skill in simulating how the ranges of O_2 bioflux vary meridionally in the Southern Ocean north of 60° S. The fact that the models provide such reasonable results suggests that the models constrain ecosystem processes in the region rather well. The models show, however, large regional variations that tend to follow the patterns

discussed earlier in Chl concentrations, both when compared with each other and with observations. We also find different and varying patterns in the seasonal cycle of biological net community production, with TOPAZ often having a shorter, more intense growing season than BGCCSM. The fact that ecosystem processes seem to be well simulated but the models still have regional and timing issues could suggest problems with the physical model, such as how well lateral currents are represented or how mixed-layer and thermocline dynamics are parameterized.

One important difference between models and observations is that models fail to predict observed events of negative O_2 bioflux in waters south of 60° S. We suggest two possible explanations for why the models lack these events: problems with the ecosystem dynamics and problems with the vertical transport of oxygen. It is possible that observed summertime undersaturation is generated by net heterotrophy (negative NCP) temporally decoupled from earlier biological production. O_2 supersaturation from periods of positive NCP would then be lost via air–sea exchange before the start of net respiration. It is also necessary for particulate organic carbon to remain in the mixed layer for long enough to be respired, and hence particle export has to be limited. Such events have been observed during the GasEx III experiment (Hamme et al., 2012) but are not generated by either model. Both TOPAZ and BGCCSM simulate NCP to evolve more smoothly than observations across time and space, even if significant variability is seen (Jonsson et al., 2013).

Our second explanation as to why models fail to capture summertime mixed layer O_2 undersaturation is that they underestimate rates and characteristics of vertical mixing. Several studies have shown the potential for entrainment and submesoscale processes to transport waters between the thermocline and the mixed layer on short timescales. Such events would introduce O_2 undersaturated waters into the mixed layer and generate the type of conditions we find in our observations. OGCMs such as TOPAZ and BGCCSM lack the spatial resolution to resolve these kinds of processes and hence the ability to generate the undersaturated conditions we observe.

We cannot conclusively disprove either explanation for the models' failure to produce negative O_2 bioflux in summer, but both the mismatches in MLD dynamics and the specific patterns in O_2 undersaturation strongly suggest that vertical transport is the dominating factor. This explanation is also supported by other studies, such as Lachkar et al. (2007) and Sallee and Rintoul (2011). Both find much stronger subduction rates south of 60° S when comparing high-resolution models to ones with lower resolution. The areas of increased subduction described by Sallee and Rintoul correspond very well with the latitude bands where we observe undersaturated conditions not generated by the models. (Below 60° S for the New Zealand, Australia, and Drake Passage sections; between 60 and 40° S for the South African section.) Other studies furthermore suggest that such physical processes captured by higher spatial resolutions tend to be episodic in

nature, making them prime candidates to generate the undersaturated O_2 conditions seen in our observations. One might argue that the lower frequency of negative O_2 bioflux in the models is simply due to model overestimation of NCP. Such a conclusion is not consistent with the models' accurate simulations of the seasonal range of positive O_2 bioflux (cf. Fig. 12). The models appear to selectively overestimate O_2 bioflux when the observations of O_2 bioflux are negative.

5 Conclusions

The fact that both TOPAZ and BGCCSM simulate the ranges of NCP and peak summer Chl concentrations well suggests that the models are able to parameterize at least physiological ecosystem processes depending on mechanistic relationships rather well. We find three main problems in the models: errors in the timing and initial biomass levels of the spring bloom, regional displacements of high biomass and NCP, and failure to generate observed extents of negative O_2 bioflux. The combined picture of how models and observations compare, together with results from other studies, suggests that the main reasons for these errors are how vertical physical processes are simulated.

Appendix A

To evaluate $\Delta O_2/Ar$ or O_2 bioflux for comparing observations with models, we use a box model that calculates the evolution of $\Delta O_2/Ar$ and O_2 bioflux from prescribed time series of NCP and wind. The box is a 50 m deep water column with no vertical or horizontal advection or mixing. O_2 air–sea exchange is simulated using O_2 saturation and the Wanninkhof (1992) gas transfer velocity parameterization; the exact setup is described in Jonsson et al. (2013). We conduct three experiments with the box model using a time series of NCP from TOPAZ at 160° W, 61° S (Fig. 14a). In the first experiment, represented by the blue lines in Fig. 14b–d, we use CORE-2 atmospheric reanalysis winds with a $2^\circ \times 2^\circ$ resolution from the same position as the NCP observations (panel b). The resulting $\Delta O_2/Ar$ supersaturation and O_2 bioflux are presented in panels c and d, respectively. In the second experiment, which is presented with red lines in panels b–d, we instead use Quikscat-satellite-derived wind data with a $0.25^\circ \times 0.25^\circ$ resolution from the same location as the NCP to simulate the effect of a different wind product. When comparing the two cases, we find that O_2 bioflux is a more robust indicator of the underlying ecological behavior (NRMSD = 5.0 %) than $\Delta O_2/Ar$ (NRMSD = 10.7 %), which we find more sensitive to short-term differences in the wind forcing applied to the model.

This effect, however, is only true if O_2 bioflux is calculated using the same wind history as that used to generate $\Delta O_2/Ar$ in the box. In the real world, the forcing is the true wind stress. However, O_2 bioflux is calculated with imperfect wind estimates from satellites or atmospheric reanalysis. The observed $\Delta O_2/Ar$ values are fixed and cannot be adjusted to compensate for the errors in the winds used to calculate O_2 bioflux. We test the effect of wind errors by using the reanalysis winds to drive the box model and QuikSCAT winds for the O_2 bioflux calculation. The resulting O_2 bioflux of such an experiment, shown as a green line in panel d, has a similar error (NRMSD = 11.3 %) to that of $\Delta O_2/Ar$ in panel c. Finally, we compare the relative effect of imperfect wind estimates between $\Delta O_2/Ar$ and O_2 bioflux. Time series of model NCP and wind from 500 locations in the Southern Ocean are used to force our box model in an analogous fashion to the experiment presented in Fig. 2d in Jonsson et al. (2013). We find that O_2 bioflux shows similar errors to $\Delta O_2/Ar$, suggesting that both properties are useful for comparing the models and observations. We choose to use O_2 bioflux since the units of flux are more appropriate for the current study.

Acknowledgements. This work was supported in part by funding from the National Aeronautic and Space Administration (NASA NNX08AF12G) and the National Science Foundation (NSF OPP-0823101).

Edited by: E. Boss

References

- Behrenfeld, M. J., Doney, S. C., Lima, I. D., Boss, E. S., and Siegel, D. A.: Annual cycles of ecological disturbance and recovery underlying the subarctic Atlantic spring plankton bloom, *Global Biogeochem. Cy.*, 27, 526–540, 2013.
- Bender, M. L., Kinter, S., Cassar, N., and Wanninkhof, R.: Evaluating gas transfer velocity parameterizations using upper ocean radon distributions, *J. Geophys. Res.*, 116, C02010, doi:10.1029/2009JC005805, 2011.
- Campbell, J. W.: The lognormal distribution as a model for bio-optical variability in the sea, *J. Geophys. Res.*, 100, 13237–13254, doi:10.1029/95JC00458, 1995.
- Cassar, N., Barnett, B. A., Bender, M. L., Kaiser, J., Hamme, R. C., and Tilbrook, B.: Continuous High-Frequency Dissolved O₂/Ar Measurements by Equilibrator Inlet Mass Spectrometry, *Anal. Chem.*, 81, 1855–1864, 2009.
- Cassar, N., DiFiore, P. J., Barnett, B. A., Bender, M. L., Bowie, A. R., Tilbrook, B., Petrou, K., Westwood, K. J., Wright, S. W., and Lefevre, D.: The influence of iron and light on net community production in the Subantarctic and Polar Frontal Zones, *Biogeosciences*, 8, 227–237, doi:10.5194/bg-8-227-2011, 2011.
- Craig, H. C. and Hayward, T.: Oxygen supersaturation in the ocean – biological versus physical contributions, *Science*, 235, 199–202, 1987.
- Doney, S. C., Yeager, S., Danabasoglu, G., Large, W. G., and McWilliams, J. C.: Mechanisms governing interannual variability of upper-ocean temperature in a global ocean hindcast simulation, *J. Phys. Oceanogr.*, 37, 1918–1938, 2007.
- Doney, S. C., Stow, C. A., Jolliff, J., McGillicuddy, D. J., Allen, J. I., Friedrichs, M. A. M., Rose, K. A., and Wallhead, P.: Skill assessment for coupled biological/physical models of marine systems, *J. Marine Syst.*, 76, 4–15, 2009.
- Ducklow, H. W. and Doney, S. C.: What is the metabolic state of the oligotrophic ocean? A debate, *Annual Review of Marine Science*, 5, 525–533, doi:10.1146/annurev-marine-121211-172331, 2013.
- Ebuchi, N., Graber, H. C., and Caruso, M. J.: Evaluation of wind vectors observed by QuikSCAT/SeaWinds using ocean buoy data, *J. Atmos. Ocean. Tech.*, 19, 2049–2062, 2002.
- Fox-Kemper, B., Danabasoglu, G., Ferrari, R., Griffies, S. M., Hallberg, R. W., Holland, M. G., Maltrud, M. E., Peacock, S., and Samuels, B. L.: Parameterization of mixed layer eddies. III: Implementation and impact in global ocean climate simulations, *Ocean Model.*, 39, 61–78, 2011.
- Garcia, H. E., Locarnini, R. A., Boyer, T. P., Antonov, J. I., Baranova, O. K., Zweng, M. M., and Johnson, D. R.: World Ocean Atlas, Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation, U.S. Government Printing Office, Washington D.C., 2009th Edn., 2010.
- Gent, P. R. and McWilliams, J. C.: Isopycnal Mixing in Ocean Circulation Models, *J. Phys. Oceanogr.*, 20, 150–155, 1990.
- Griffies, S. M., Gnanadesikan, A., Dixon, K., Dunne, J. P., Gerdes, R., Harrison, M., Rosati, A., Russell, J., Samuels, B. L., and Spelman, M.: Formulation of an ocean model for global climate simulations, *Ocean Sci.*, 1, 45–79, 2005, <http://www.ocean-sci.net/1/45/2005/>.
- Hamme, R. C., Cassar, N., Lance, V. P., Vaillancourt, R. D., Bender, M. L., Strutton, P. G., Moore, T. S., DeGrandpre, M. D., Sabine, C. L., Ho, D. T., and Hargreaves, B. R.: Dissolved O₂/Ar and other methods reveal rapid changes in productivity during a Lagrangian experiment in the Southern Ocean, *J. Geophys. Res.*, 117, C00F12, doi:10.1029/2011JC007046, 2012.
- Hashioka, T., Vogt, M., Yamanaka, Y., Le Qu  r  r, C., Buitenhuis, E. T., Aita, M. N., Alvain, S., Bopp, L., Hirata, T., Lima, I., Sailley, S., and Doney, S. C.: Phytoplankton competition during the spring bloom in four plankton functional type models, *Biogeosciences*, 10, 6833–6850, doi:10.5194/bg-10-6833-2013, 2013.
- Huang, K., Ducklow, H. W., Vernet, M., Cassar, N., and Bender, M. L.: Export production and its regulating factors in the West Antarctica Peninsula region of the Southern Ocean, *Global Biogeochem. Cy.*, 26, GB2005, doi:10.1029/2010GB004028, 2012.
- Johnson, R., Strutton, P. G., Wright, S. W., McMin, A., and Meiners, K. M.: Three improved satellite chlorophyll algorithms for the Southern Ocean, *J. Geophys. Res.-Oceans*, 118, 3694–3703, 2013.
- Jonsson, B. F., Doney, S. C., Dunne, J., and Bender, M.: Evaluation of the Southern Ocean O₂/Ar-based NCP estimates in a model framework, *J. Geophys. Res.-Biogeo.*, 118, 385–399, 2013.
- Lachkar, Z., Orr, J. C., Dutay, J.-C., and Delecluse, P.: Effects of mesoscale eddies on global ocean distributions of CFC-11, CO₂, and $\Delta^{14}\text{C}$, *Ocean Sci.*, 3, 461–482, doi:10.5194/os-3-461-2007, 2007.
- Large, W., Doney, S. C., and McWilliams, J. C.: Oceanic Vertical Mixing – a Review and a Model With a Nonlocal Boundary-Layer Parameterization, *Rev. Geophys.*, 32, 363–403, 1994.
- Large, W. G., Doney, S. C., Danabasoglu, G., and McWilliams, J. C.: Sensitivity to surface forcing and boundary layer mixing in a global ocean model: Annual-mean climatology, *J. Phys. Oceanogr.*, 27, 2418–2447, 1997.
- Reuer, M. K., Barnett, B. A., Bender, M. L., Falkowski, P. G., and Hendricks, M. B.: New Estimates of Southern Ocean Biological Production Rates from O₂/Ar Ratios and the Triple Isotope Composition of O₂, *Deep-Sea Res. Pt. I*, 54, 951–974, 2007.
- Sallee, J.-B. and Rintoul, S. R.: Parameterization of eddy-induced subduction in the Southern Ocean surface-layer, *Ocean Model.*, 39, 146–153, 2011.
- Smith, R. and Gent, P. R.: Anisotropic Gent-McWilliams parameterization for ocean models, *J. Phys. Oceanogr.*, 34, 2541–2564, 2004.
- Spitzer, W. S. and Jenkins, W. J.: Rates of vertical mixing, gas-exchange and new production – estimates from seasonal gas cycles in the upper ocean near Bermuda, *J. Mar. Res.*, 47, 169–196, 1989.
- Wanninkhof, R.: Relationship between wind speed and gas exchange over the ocean, *J. Geophys. Res.*, 97, 7373–7382, 1992.