Variations in ocean deoxygenation across Earth System Models: Isolating the role of parametrized lateral mixing

A. Bahl\textsuperscript{1}, A. Gnanadesikan\textsuperscript{1}, and M.-A. Pradal

\textsuperscript{1}Department of Earth and Planetary Sciences, Johns Hopkins University, Baltimore, Maryland, USA

Corresponding Author: Anand Gnanadesikan
Olin 327 Department of Earth and Planetary Sciences
Johns Hopkins University
3400 N. Charles St.
Baltimore, MD

gnanades@jhu.edu
410-516-0722
Manuscript revised for Global Biogeochemical Cycles

Key points

1. An inter-model comparison reveals that oxygen deoxygenation is sensitive to the representation of lateral mixing.

2. Stronger lateral mixing results in greater deoxygenation, but largely in zones which have high oxygen in the base state.

3. Cross-model differences in oxygen fields are dominated by biological drawdown, but are driven by ventilation rather than consumption.


Abstract

Modern Earth System Models (ESMs) disagree on the impacts of anthropogenic global warming and the associated circulation changes on the distribution of oxygen and hypoxic waters. One reason for cross-model differences is the representation of lateral mesoscale eddy transport. As these eddies are smaller than the spatial scale of ESM ocean grids, their impact must be parameterized using a lateral mixing coefficient $A_{REDI}$. We varied $A_{REDI}$ over a factor of six within a single ESM, a range similar to that seen across modern ESMs, and also used two spatially dependent fields derived from altimetry, generating simulations with both preindustrial and increased CO$_2$. Global warming results in a decline in productivity and decrease in ventilation age in the tropics, increasing oxygen concentrations in the upper thermocline. In high latitudes, shallowing of deep convection reduces the supply of oxygen to the deep. The sum of the two processes depends on $A_{REDI}$, with an increase in hypoxic volume yet smaller total deoxygenation in the low mixing models, but a decrease in hypoxic volume yet larger total deoxygenation in the high mixing models. All models show decreases in suboxic volume which are largest in the low-mixing models. The variability across our model suite represents a substantial fraction of that seen in a subset of CMIP5 models and exhibits a similar decoupling between total deoxygenation and change in hypoxic volume. Uncertainty in lateral mixing remains an important contributor to uncertainty in projecting ocean deoxygenation.
Global warming is expected to change the amount of oxygen present in the oceans, both at the surface and at depth. Some of this is because warmer waters can hold less oxygen. But a bigger part of the expected changes is associated with biology. Because it makes surface waters lighter, global warming makes it more difficult for heavier deep waters rich in nutrients to come to the surface and drive biological productivity. This in turn means that less organic matter sinks into the deep ocean and rots. As rotting utilizes oxygen this means that there should be more oxygen at depth. But as oxygen-rich surface waters get lighter, it becomes harder for these waters to flow into the deep ocean and deliver oxygen to these depths, causing oxygen to fall. The balance these three processes determines whether increasing fractions of the deep ocean will become inhospitable to fish and other organisms. The Earth System Models used to project the impacts of climate change on oxygen predict different balances between these processes. In this paper we demonstrate that an important reason for this disagreement is uncertainty how to represent mixing associated with the oceanic “storms” known as mesoscale eddies.
1. Introduction

Global warming is expected to strongly influence the biogeochemical and physical processes responsible for the distribution of oxygen in the global oceans on decadal-to-century timescales (Vaquer-Sunyer and Duarte, 2008; Frölicher et al., 2009; Keeling et al., 2010; Stramma et al., 2010; Gnanadesikan et al., 2012; Cabrè et al., 2015; Shepherd et al., 2017; Oschlies et al., 2018). Changes to dissolved $O_2$ concentration under global warming are driven by three main processes: (i) warming of ocean waters reduces oxygen solubility; (ii) changes to large-scale circulation reduce vertical exchange and the supply of oxygen to the deep, which also (iii) reduces the supply of nutrients to the surface, thus driving down both productivity near the surface and the consumption of oxygen in the ocean interior (Praetorius et al., 2015). Recent trends in $O_2$ concentrations are summarized by Schmitko et al. (2017), who found a 2% decrease in global oxygen inventory since 1960.

A particular concern is that anthropogenic climate change will lead to the growth of oxygen minimum zones (OMZs) from which air-breathing fish species and many species of heterotrophic zooplankton are excluded (Hsia et al., 2013; Kwon et al., 2016; Schmittko et al., 2017). Such waters are also known as hypoxic. While the threshold for hypoxia depends on the organism involved (Vaquer-Sunyer & Duarte, 2008; Stramma et al., 2012; Mislan et al., 2017), we choose $O_2<88 \mu M (~2 ml/l)$ to define hypoxic waters, following previous work by Díaz & Rosenberg (2008) and Gnanadesikan et al. (2013). Such waters make up ~10% of the ocean’s total volume, the most extensive regions being found at mid-depths of the North Pacific, with smaller regions in the North Indian Ocean and the South Pacific Ocean (Bianchi et al., 2012).

Intensely hypoxic waters are associated with denitrification, the bacterial process that converts $NO_3^-$ into gaseous nitrogen, which reduces nitrate concentrations while increasing
phosphate concentrations. The resulting decreases in the \( N^* = \text{NO}_3^- - \text{16*PO}_4^- \) tracer (Gruber and Sarmiento, 1997) can be seen when oxygen drops below \( \sim 20 \mu\text{M} \). For purposes of this study (also following Gnanadesikan et al., 2013) we define such waters as suboxic. Currently, \( \sim 1\% \) of the ocean volume is suboxic (Bianchi et al., 2012). The denitrification found in such waters is important because it produces \( \text{N}_2\text{O} \), a powerful greenhouse gas, as a byproduct (Nevison et al. 2003, Freing et al., 2012) and serves to stabilize the ocean biosphere to perturbations in nitrogen supply (Tyrell, 1999).

ESMs have become invaluable tools for understanding the 20\textsuperscript{th} century evolution of the Earth’s physical climate on decadal to multi-decadal timescales and also for projecting changes in future climate. However, models still exhibit large inconsistencies in the simulation of dissolved oxygen (Ying & Yangchun, 2016) and projections of its change under global warming (Bopp et al., 2013). This is illustrated in Fig. 1, which shows oxygen profiles (left-hand column) and the volume of hypoxic water (\( \text{Mkm}^3/\text{deg latitude} \), right-hand column) across 6 of the CMIP5 models. A brief description of each model can be found in Table 1.

As previously noted (Bopp et al. 2013), comparing the preindustrial mean state with observations (Bianchi et al., 2012, symbols in Fig. 1a) shows that while all models agree reasonably well with the observations at the surface, significant divergence emerges below \( \sim 500\text{m} \). The models also simulate different magnitudes and patterns of hypoxic volume (Fig. 1b).

Many models show erroneous spatial distributions of oxygen with the northern and southern OMZ regions in the Pacific merging into a much shallower single southern tropical Pacific OMZ (Cabré et al., 2015). Although there is some relationship between the biases in the deep ocean concentrations and biases in hypoxic volume (IPSL-MR for example has excessively low concentrations and excessively high hypoxic volumes), the relationship is far from exact. For
example, the HadGEM and MPI models both produce relatively reasonable profiles of oxygen but HadGEM underestimates hypoxia, while MPI overestimates it. As previously noted by Ying & Yangchun (2016) GFDL-ESM2M and GFDL-ESM2G produce differences in oxygen concentrations, despite having the same biogeochemical, atmosphere, sea ice and land components but a different ocean physical component.

A climate change scenario in which atmospheric carbon dioxide is increased by 1%/yr for 70 years and then held constant for another 70 years was run for all six cases, with the changes in oxygen and hypoxia shown on the bottom panels of Fig. 1. Surprisingly, no simple relationships emerge between the mean state biases and the changes under global warming. Some models (NorESM and MPI), predict significant deoxygenation (Fig. 1c) showing a loss of ~10 µM on average, while others lose much less (HadGEM and ESM2M lose ~4 and ~2 µM respectively). More strikingly, even the sign of the change in hypoxic volume can differ between models (Fig. 1d). In particular, the three models with the most realistic profile of oxygen (HadGEM, MPI and ESM2M) show opposite sign changes in hypoxic volume (with ESM2M increasing hypoxic volume and MPI and HadGEM decreasing it under global warming).

Understanding the source of such differences is very difficult, as the ESMs in Fig. 1 differ in terms of climate sensitivity, representation of ocean biogeochemistry, and physical dynamics in the ocean.

In this paper, we focus on one process known to vary across models: lateral mesoscale eddy mixing, which following Redi (1982) is parameterized using a lateral turbulent diffusion coefficient $A_{REDI}$ (note: while such parametrization is unnecessary in very high-resolution models, such models are generally not run out for long enough times for the solution to reach quasi-equilibrium). This coefficient varies across Earth System Models with values ranging
from 200 m$^2$/s to 4000 m$^2$/s (Gnanadesikan et al., 2015b). While we do not expect $A_{\text{REDI}}$ to be 
the sole explanation for inter-model differences, previous work shows the impact different values 
of $A_{\text{REDI}}$ have on oxygen distributions (Gnanadesikan et al. 2013). The differences in $A_{\text{REDI}}$ 
across the CMIP5 models in Fig. 1 are shown in the fourth column of Table 1. Little clear 
relationship can be found with the differences seen in Fig. 1. This is consistent with Cabré et al. 
(2015) who found no correlation between the reported $A_{\text{REDI}}$ coefficients and Pacific OMZ 
volumes across CMIP5 models in the mean-state. However, by focusing on the impacts of 
changing $A_{\text{REDI}}$ within a single model (in which atmospheric parametrizations and ocean 
biogeochemistry are identical) we can better isolate the patterns of cross-model variation that 
could be due to uncertainties in this parameter and explain some of the reasons for why changes 
in oxygen and hypoxic volume appear to be decoupled.

In recent years, our group has examined multiple impacts of changing the $A_{\text{REDI}}$
coefficient within a single ESM. This work was motivated by Gnanadesikan et al. (2012) who 
found that in coarse-resolution ESMs, the mesoscale eddy transport scaled by this parameter is 
the dominant mechanism supplying oxygen to hypoxic zones. Gnanadesikan et al. (2013) ran a 
suite of simulations with different spatially constant values of $A_{\text{REDI}}$ and found that the value of 
$A_{\text{REDI}}$ was inversely correlated with the volume of hypoxic and suboxic waters. Pradal & 
Gnanadesikan (2014) found that increasing the mixing coefficient resulted in destabilizing high 
latitude haloclines and produced significant warming in both the Southern Ocean and North 
Pacific. Our team extended this work in Gnanadesikan et al. (2015b) to look at the impact of 
different mixing parameterizations on anthropogenic carbon uptake, also using a new spatially 
dependent distribution of mixing developed from altimetry measurements.
Until now, we have not analyzed the impact that uncertainty in the value of $A_{\text{REDI}}$ has on ocean deoxygenation, nor have we looked at the impact of including more realistic spatially-varying fields of $A_{\text{REDI}}$ on the distribution of oxygen. This paper analyzes and discusses the differences between future projections of ocean oxygen in the model suite introduced in Gnanadesikan et al. (2015b), with the aim of estimating how much of the difference seen between ESMs could be due to differences in mixing vs. differences in biogeochemical formulation or physical climate sensitivity.

We introduce our model in section 2, which includes a description of the physical and tracer cycling models used and a mathematical discussion of parametrized sub-grid scale turbulent mixing. Section 3 describes the experimental strategy and Section 4 describes the results of the simulations. Section 5 examines why the results differ across different models, highlighting both the role of biological utilization of oxygen and the importance of circulation in determining the pattern and magnitude of deoxygenation. Section 6 concludes this paper.

2. Model description and physical background

The model used in this study is a lower resolution version of the National Oceanic and Atmospheric Administration’s (NOAA) Geophysical Fluid Dynamics Laboratory (GFDL) ESM2M (Dunne et al. 2012, red lines in Fig.1) denoted as CM2Mc (Galbraith et al., 2011). CM2Mc numerically simulates atmosphere, land surface, sea ice, and ocean dynamics with an exchange grid system that allows for energy, momentum and tracers to pass between land, sea, ice and atmosphere every 3 hours. We summarize some of the salient features of this model below. For a more extensive description, the reader is referred to Galbraith et al. (2011).
2.1 Atmosphere module

The atmosphere module in CM2Mc is based on the GFDL Atmospheric Model, version 2 (AM2, GAMDT, 2005), and uses a finite volume dynamical core that is virtually identical to that used in the GFDL CM2.1, but at lower horizontal resolution. The atmosphere grid has a 3.75° longitudinal and a 3° latitudinal resolution with 24 vertical levels and a time step every 1 ½ hours. The atmospheric physics (aerosols, radiation, mixing schemes) are in essence identical to ESM2M.

2.2 Ocean model

The ocean model used is the Modular Ocean Model v4.1 (MOM4p1), code of Griffies (2009), simulating a dynamical core that uses a modified pressure coordinate, allowing for conservation of mass instead of volume. The model has time step every 3 hours and a grid with 28 vertical levels, a 3° longitudinal resolution and a varying latitudinal resolution with an average of about 1.5° and a minimum value of 0.6° in the equatorial latitudes. While this resolution is coarser than all of the CMIP models shown in Fig. 1, we will demonstrate that the quality of the simulation is comparable.

Biological cycling is accounted for with the use of a highly parameterized biogeochemical code referred to as BLING: Biogeochemistry with Light, Iron, Nutrients, and Gases (Galbraith et al., 2010). The version of BLING run here includes eight tracers. Six of these
are prognostic variables used for evolving the system: dissolved inorganic carbon, alkalinity,
micronutrient (nominally Fe), macronutrient (nominally some mix of NO$_3$ and PO$_4$), dissolved
organic material, and oxygen. These variables are advected and diffused by the same physical
processes that advect and diffuse temperature and salinity.

As described in Galbraith et al. (2010) nutrient concentrations and ambient light are used
to predict growth rate, from which a parameterized ecosystem (Dunne et al., 2005) is run to
compute estimates of large and small phytoplankton biomass and chlorophyll. These in turn are
used to compute the uptake of nutrients and distribution of light which then cause all the
biologically active fields to evolve. Oxygen is produced and consumed with a constant
stoichiometric ratio of 150:1, relative to phosphorus. The rate of remineralization of this sinking
biomass is suppressed in low-oxygen waters - an important consideration when analyzing the
oxygen budget and its changes under global warming.

Additionally, BLING is run with two “preformed” tracers (DIC$_{pre}$ and PO$_4$$_{pre}$) which are
set to the modeled values of DIC and PO$_4$ in the mixed layer. Below the mixed layer, these
tracers are advected and diffused in the same way as DIC and PO$_4$ but they have no source due to
remineralization (Ito & Follows, 2005). Thus, the difference between PO$_4$ and PO$_4$$_{pre}$ gives us
the amount of phosphate added (and implicitly oxygen consumed) by biological activity. We
term this tracer PO$_4$$_{remin}$=PO$_4$-PO$_4$$_{pre}$, the remineralized phosphate. Because globally averaged
total phosphate is constant, an increase in globally averaged remineralized phosphate must be
associated with the mix of waters feeding the deep ocean shifting towards ones with lower
preformed phosphate (Marinov et al., 2008). Note that our approach differs from that of
Frölicher et al., (2009) who track biologically driven changes in oxygen in terms of phosphate
alone- which may shift both as a result of shifting watermasses as well as oxygen utilization.
The model is also run with an ideal age tracer. Following Thiele and Sarmiento (1990), this tracer is set to zero at the first model point in the mixed layer and ages at a rate of 1 year/year thereafter. If one imagines a pipe running from the well-mixed surface layer to an interior point, the ideal age would be the time required to transit from the surface to that point, sometimes referred to as the ventilation time (England, 1995). In reality a range of pathways, both advective and diffusive, connect any interior point to the surface and the ideal age is an average of the transit times associated with these pathways.

2.3 Discussion of the turbulent diffusion coefficient, $A_{\text{REDI}}$

CM2Mc, like most ESMs run today, does not resolve spatial scales associated with mesoscale eddies, requiring that lateral transport of passive tracers associated with such tracer be parameterized. The turbulent flux of tracers along isopycnals is represented using a Fickian diffusion approximation, such that the flux of tracer $C$ in direction $s$ is given by (Redi, 1982):

\[ F_x^C = -A_{\text{REDI}} \frac{\partial C}{\partial s} \]  

Where $C$ is the concentration of the tracer. The resulting flux goes from high to low concentrations and is proportional to the size of the tracer gradient. $\partial s$ is taken along isopycnal surfaces within the ocean interior and horizontally in the mixed layer.

In addition to $A_{\text{REDI}}$, advection produced by mesoscale eddies along isopycnals is accounted for by using a shear-dependent coefficient Gent and McWilliams ($A_{\text{GM}}$) scheme (Gent and McWilliams, 1990). In regions where isopycnals are strongly tilted, growing eddies flux horizontal momentum downwards via form drag. This in turn results in an overturning circulation that flattens isopycnals. The resulting flux in the x and y-direction is:
\[
F_x^C = -C \times A_{GM} \frac{\partial S_{x,y}}{\partial z} \quad (2)
\]

Where \( A_{GM} \) is the diffusive coefficient. \( S_{x,y} \) are the slope of the isopycnal in the x and y-directions respectively. The slope is capped at a maximum value of \( S_x=0.01 \), to prevent unrealistically large velocities near mixed layers where slopes become infinite.

As is done in most of the CMIP5 models, \( A_{GM} \) is chosen to be proportional to the vertical shear. In our model we use the shear between depths of 100 m and 2000 m and choose a minimum coefficient of 200 m\(^2\)/s and maximum coefficient of 1400 m\(^2\)/s, so as to produce realistic hydrography and overturning in the Southern Ocean. As shown in Table 1, the different CMIP5 models use different minimum and maximum values, though the majority use a broadly similar shear-dependent formulation.

By contrast, there is no consensus across the CMIP5 models for how to specify the \( A_{REDI} \) coefficient. \( A_{REDI} \) is set equal to \( A_{GM} \) in some models (ESM2G, NorESM), to a constant value ranging from 500 m\(^2\)/s (HadGEM) to 1000 m\(^2\)/s (IPSL) in other models, and varies with the grid spacing in the MPI model, with maximum value of 400 m\(^2\)/s. While we do not show it here, as it does not have a comparable global warming simulation, the CMCC model of Fogli et al. (2009) uses a constant coefficient of 2000 m\(^2\)/s. While one might expect \( A_{REDI} = A_{GM} \) as both parametrize mesoscale eddy “mixing”, the two coefficients do parametrize different processes. Because many models vary \( A_{REDI} \) independently of \( A_{GM} \) and because we already know it affects oxygen minimum zones (Gnanadesikan et al., 2013) we focus on this one coefficient at this time.

3. Experimental design
A series of six integrations are made to examine the dependence of ocean biogeochemistry on different values of $A_{\text{REDI}}$. Four of the runs use values of $A_{\text{REDI}}$ that are constant in space and time: $A_{\text{REDI}} = 400, 800, 1200, \text{ and } 2400 \text{ m}^2/\text{s}$. These runs will be described as AREDI400 and AREDI800, or low-mixing, and AREDI1200 and AREDI2400, high-mixing. The other two runs use a pattern that varies in space but not in time.

The first of these, denoted as ABER2D, uses a two-dimensional field of $A_{\text{REDI}}$ inferred from sea-surface height measurements by Abernathey and Marshall (2013). ABER2D’s field of $A_{\text{REDI}}$ is as high as 10,000 m$^2$/s in high mixing areas, such as the edges of boundary currents in subtropical gyres, and decreases to very low values in lower mixing regions that are often far from boundary currents, such as the Southern Ocean (Abernathey and Marshall, 2013; see also Gnanadesikan et al., 2015a, fig. 1). In order to see the contrast in the spatial patterns of $A_{\text{GM}}$ vs. $A_{\text{REDI}}$ more clearly, we focus on a region of high mixing, the Northwest Pacific. As seen in Fig. 2a, the baroclinic growth parameterization predicts an $A_{\text{GM}}$ that is highest in the center of boundary currents and drops to very low values in nearby gyres. However, the Abernathey and Marshall (2013) estimate of $A_{\text{REDI}}$ shows higher values at the edge of the currents, and lower values in the center (see Klocker & Abernathey, 2014, for an explanation of why this occurs).

Finally, the last experiment uses a zonally averaged version of the ABER2D parameterization, denoted as ABERZONAL. This is used to determine whether the full 2-dimensional structure of the $A_{\text{REDI}}$ field is most important or simply the variation with latitude (which might be more constant in time and thus more easily applied to paleoclimatic simulations). Although these spatially varying fields of $A_{\text{REDI}}$ have been used in previous work in our group (Gnanadesikan et al., 2015a, 2015b, 2017a) they are substantially different than the
distributions of $A_{\text{REDI}}$ used in all the CMIP5 models and their impact on oxygen fields has not been previously documented.

Our control run (AREDI800) was initialized with temperature, salinity, phosphate, oxygen, and DIC from modern hydrographic data and spun up for 1500 years with orbital parameters, solar flux, atmospheric greenhouse gasses and aerosols held at “preindustrial” year 1860. Once the simulations reached year 1500, the additional five scenarios (AREDI400, AREDI1200, AREDI2400, ABER2D and ABERZONAL) were branched off the control. All six simulations where then continued for 500 years with constant solar fluxes, aerosols and greenhouse gases. At model year 1860 (360 years into the new control) additional runs are made in which the concentration of CO$_2$ is instantaneously doubled from the preindustrial value of 286 ppmv to a value of 572 ppmv for both CO$_2$ and radiation schemes. These simulations were run out for an additional 140 years. Instantaneously doubling atmospheric CO$_2$ allows us to estimate the forced component of climate change and how this component is reflected in ocean circulation. Instantaneously doubling in an ESM is an idealization that removes variation on the time scales associated with CO$_2$ growth. However, it also introduces a “shock” to the system that might not be reflected in historical simulations.

4. Results

4.1 Steady state differences and robustness

We begin by repeating the analysis in Fig.1 for our model suite. As was the case for the CMIP5 models, horizontally averaged oxygen is very similar at the surface, with differences starting to appear at 500 m. However, all six models simulate the general shape and mean value
throughout the water column with better skill than the IPSL and NorESM models. Amongst our models ABERZONAL and AREDI1200 best capture observations with mean oxygen concentrations (Table 2, column 3) relatively close to the observed value of 177.1 μM. All of our models are somewhat warmer than observations, with higher mixing being associated with more realistic lower temperatures (Table 2, column 2). As discussed in Pradal and Gnanadesikan (2014) this is due to higher mixing destabilizing high latitude haloclines, leading to more convection, which moves heat out of the ocean.

Although some part of the cross-model differences in oxygen are attributable to temperature, the fraction is relatively small. The 0.6°C decrease in temperature between AREDI400 and ARED2400 (Table 2, column 2) is only enough to produce a 1.4% change in solubility which, assuming sinking waters are in perfect equilibrium, would produce a 4 μM increase in oxygen concentration, a small fraction of the 31 μM we are trying to explain. The difference in remineralized phosphate (Table 2, column 5), by contrast, is 0.24 μM. Given the model stoichiometric ratio of 150:1, the resulting increase in oxygen would be expected to be 36 μM. The sum of the two processes gives a predicted change larger than seen in the models because we have neglected gas exchange. The waters sinking into the deep ocean are not, in fact fully equilibrated with the atmosphere (as also noted by Fröhlicher et al., 2009).

Hypoxic and suboxic volumes vary significantly across the model suite, (Table 2, columns 3 and 4). As previously noted in Gnanadesikan et al. (2013) low-mixing models, especially AREDI400, greatly overestimate hypoxic and suboxic volume compared to observations, while high-mixing models, specifically AREDI2400, greatly underestimate them. Despite having very large peak diffusivities, the ABER2D and ABERZONAL models tend to have less oxygen and more hypoxia than the AREDI2400 case. Note that there is a big difference
in suboxic volume between ABERZONAL and ABER2D, implying that the full two-
dimensional structure of the mixing coefficient matters for capturing these biogeochemically
important waters. As shown in Fig. S1 the hypoxic waters are centered in the tropics- with none
of the models capturing the hypoxia seen in the North Pacific. The models all fail to capture the
separation between the northern and southern OMZs.

Our model suite produces three-dimensional temperature, salinity, and oxygen fields
which are comparable in accuracy to those from the CMIP5 models. As shown in the Taylor
diagrams in Fig. 4, for oxygen our models lie between the worst-correlated model (NorESM with
a value of 0.65) and the best (MPI with a value of 0.92). AREDI1200, ABER2D, ABERZONAL,
and AREDI2400 produce almost identical correlations of 0.7. However, with the exception of
HadGEM2, CMIP5 models overestimate the global range, while our models underestimate it. In
terms of RMS error (distance from the 1,0 point) the AREDI400 simulation is comparable to the
best of the CMIP5 simulations examined here.

In both our model suite and in the CMIP5 models the models that produce the best
salinity fields (AREDI2400 and IPSL-MR respectively) are not the ones with the best spatial
distribution of oxygen. Similarly, the models with the best simulation of oxygen distribution do
not necessarily capture hypoxic and suboxic volume. The AREDI1200 model has less mean bias
than AREDI400 in oxygen and hypoxic volume (Table 2), but a pattern of variation that is less
correlated with observations. Similarly, MPI has a pattern of oxygen that is well-correlated and
an amplitude of variation that is too large, which is reflected in an overestimate of the hypoxic
volume. Which model is the “best” in the control case depends on which metric is used.
4.2 Deoxygenation under doubled atmospheric CO₂

The ocean’s response to doubled CO₂ also depends on the choice of AREDI. The data in Table 3 lists the changes in temperature, oxygen, hypoxic and suboxic volumes and remineralized phosphate across the models after 140 years of doubled CO₂. In contrast to the control, where models simulate a positive relationship between O₂ concentration and hypoxic volume, the size of the change in oxygen is not predictive of the size of the change in hypoxic volume. As can be seen from Fig 3c and d, the biggest absolute differences between models are seen at depths of ~300 m and at ~3000 m, which will motivate further analysis in Section 5.

Columns 3 and 4 in Table 3 show that the largest change in total O₂ concentration is seen in AREDI1200 (-9.4 μM) and the smallest in AREDI400 (-6.2 μM), a smaller range than in the CMIP5 models. However, low-mixing models increase the volume of hypoxic waters, while high-mixing models and spatially dependent models decrease the volume (Table 3, column 3, Fig. 3d) - with a similar spread to what is seen in CMIP5 models (Fig. 1d). While all models show a decrease in suboxic volume (Table 3, column 4), this decrease is largest for the low mixing models.

The reason for this interesting divergence in oxygen and hypoxic volume change must be understood in terms of the three-dimensional structure of oxygen change. Overlaying the zonal integral of hypoxic volume change with oxygen change (Fig. 5), we see tongues of lower oxygen penetrating into the deep ocean from high latitudes. When these tongues intersect regions where oxygen is sufficiently low (as in the tropics in AREDI400) we see an increase in hypoxic volume. As the mixing increases, the center of the tongue of oxygen decrease moves deeper and deeper into the water column away from regions which are on the edge of being hypoxic. In
tropical waters by contrast, all models show an increase in oxygen concentration. This result was previously seen in Gnanadesikan et al. (2012) in the higher-resolution CM2.1 model. The result is to shrink hypoxic waters in this depth and latitude range. In the high-mixing models (AREDI1200, AREDI2400, ABER2D and ABERZONAL) the tropical change dominates the change in hypoxic water volume.

5. Discussion

Understanding what drives cross-model differences in deoxygenation, both in the preindustrial control and under global warming requires further analysis. In particular, what is the relative importance of changing solubility versus changing biological utilization? We examine this question in section 5.1 by comparing how the two components change across models. We then discuss the differences in how the changes in biological utilization are driven, linking them to changes in ventilation associated with changes in stratification.

5.1 Understanding cross-model differences in deoxygenation

5.1.1 Temperature and solubility

The global average temperature change across the six models under global warming (Table 3, column 1) is essentially identical. The limited model spread comes as a major surprise given that the differing extents of sea ice in the preindustrial state (Pradal and Gnanadesikan, 2014) would be expected to produce different magnitudes of ice-albedo feedback. However, full analysis of this result is beyond the scope of this paper. For now, we note that the average change of 0.4°C neither explains the changes in oxygen nor its dependence on A_REDI. Such warming
would be expected to produce a 3.2 μM change in oxygen saturation, which only accounts for ¼ to ½ of the changes seen in the models.

This finding is consistent with Bopp et al. (2002) and Matear & Hirst (2003), who found that only ~ ¼ of the oxygen depletion in their global warming simulations could be attributed to solubility changes but differs from Frölicher et al. (2009) who found that the majority of deoxygenation was associated with changing solubility. Likewise, Schmidtko et al., (2017) attribute less than half of deep oxygen changes to temperature, although they did find strong solubility-induced changes in the surface layers. Similarly, we recognize that thermally-driven differences in deoxygenation could still be important on regional scales.

5.1.2 Changes in biological oxygen utilization

As was true for our control simulations, inter-model differences in deoxygenation under global warming are largely due to differences in biological drawdown. AREDI400 simulates a 0.033 μM increase in remineralized phosphate while AREDI2400 simulates a 0.056 μM increase. Given a 150:1 ratio of oxygen to phosphate, the 0.023 μM difference in the increase would be expected to produce a difference of 3.6 μM O₂ between the two models, similar in magnitude to the 3.2 μM change actually seen. If both the changes in temperature and remineralized phosphate were fully expressed in the AREDI1200 case (which sees the biggest change), we would expect to see a 11.6 μM decrease in oxygen concentration. However, in the presence of finite gas exchange the actual change of 9.4 μM (Table 3, column 2) is only 80% of the potential decrease.
Although changes in oxygen utilization explain the greater portion of the differences across both the control and global warming simulations, there are two possible reasons for such differences. The biological drawdown of oxygen can be thought of as ~150 times the inventory of remineralized phosphate. This in turn can be written as the flux of organic phosphorus to the deep ocean multiplied by its residence time. Changes and differences in the drawdown of oxygen can thus be driven by changes and difference in the flux of organic phosphorus to the deep, or in the age of deep waters. In the following section we explore which of these drivers is most important.

5.2 Understanding differences in biological utilization

5.2.1 Export production

Cross-model differences in export production across models are relatively small under both our control (Fig 6a, Tab.4) and global warming (Fig. 6b, Tab. 4) scenarios, showing that they cannot account for cross-model differences in oxygen inventory. Globally integrated export fluxes vary from 9.95 GtC/yr in AREDI400 to 11.1 GtC/yr in AREDI2400, well within the 9.8 +/-2 GtC/yr estimated by Dunne et al., (2007, symbols in Fig. 6a). The regional pattern of export is very similar across the model suite, with noticeable latitudinal shifts relative to observations. However, oxygen concentrations are 20% higher in AREDI2400 than in AREDI400, even though the export flux is 10% higher. The higher deep ocean oxygen in AREDI2400 must therefore be driven by greater physical exchange between the surface and deep ocean resulting in a larger oxygen supply to the deep.
CO₂ doubling results in a decrease in export across all model runs, with the global decline being fairly similar across all models (7-10%). The low mixing models, AREDI400 and AREDI800, experience the sharpest decline at 40°N. The largest range of changes is seen in the North Pacific, where AREDI1200 produces a decline of -0.20 Gt C/yr while ABERZONAL produces a decline of -0.05 Gt C/yr. While the consistent decline in export production across all latitudes would be expected to cause a decrease in remineralized phosphate, the opposite (with a concomitant decrease in oxygen) is actually seen.

5.2.2 Ventilation

Export production cannot explain even the sign of the differences in oxygen concentration across models. Nor can it explain the sign of (or different magnitudes of) the changes in oxygen concentration under global warming. We therefore turn to differences in ventilation as diagnosed by differences in age.

In Fig. 7, we compare changes of O₂ and ideal age under doubled CO₂ in each of the models at 300 m—one of the depths where large cross-model differences are seen in Fig. 3c. Oxygen decreases as age increases across each of the models, with AREDI800 simulating the strongest correlation of -0.84. All the models show oxygen increasing and age decreasing in the tropics. Under global warming, the upwelling of deep, dense, old waters in these regions drops more quickly than the wind-driven downward transport of “young” waters (Gnanadesikan et al., 2007, 2012), resulting in a decrease in age and biological storage of remineralized phosphate and thus an increase in oxygen.
Away from the tropics, we see clear differences between the models at 300 m in the western Pacific and the Southern Ocean. Although the size of the effect varies across models, all the models exhibit increases in age and decreases in oxygen in these regions. In the Northwest Pacific, the largest changes are found in AREDI400 and AREDI800, while in the Southern Ocean, the biggest changes are found in the ABERZONAL model.

At a depth of 3000 m (Fig. 8a) a similarly high correlation between changes in ideal age and oxygen is exhibited across the model suite. The high mixing models exhibit the strongest correlation with AREDI1200 at -0.98 and AREDI2400 at -0.97. Significant differences across models are seen in the Northwest Pacific with the high mixing models (AREDI1200, AREDI2400, ABER2D and ABERZONAL), all showing declines in oxygen concentration of more than 75 μM and increases in age by hundreds of years. The large changes in age in both the Southern Ocean and North Pacific point to a strong decline in vertical exchange, reasons for which we examine in the following subsection.

5.3 Explaining sensitivity of vertical exchange to lateral mixing

5.3.1 Stratification

As noted in Pradal and Gnanadesikan (2014), both differences in vertical exchange across different models and variability within models in subpolar regions can be linked to changes in salinity stratification. We might expect this to be true for our global warming cases as well. Increased atmospheric CO₂ concentration causes an acceleration of the hydrological cycle, with greater freshwater inputs to high latitudes. This results in surface freshening while deep waters remain relatively unchanged, inhibiting convection and overturning. Increasing stratification can
thus be part of a positive feedback cycle whereby fresh waters are less efficiently removed from convective regions (Stommel, 1961; Gnanadesikan et al., 2017b).

The strong linkage between stratification and vertical exchange is most clearly seen in the NW Pacific (120-180°W, 50-65°N). Salinity stratification is a good predictor of oxygen (Fig. 9a) and age (Fig. 9b) at 300 m both within and across model runs (changes in temperature stratification in this region actually move towards being destabilizing). All of these simulations have too much oxygen at 300 m relative to modern observations (symbol, Fig. 9a), including AREDI400 which has relatively realistic levels of stratification. It turns out that AREDI400 has overturning in this region which appears to be a little too strong and too deep (Gnanadesikan et al., 2015a, 2017b). This bias (which is also found in other ESMs) may represent a source of error in simulating modern changes in oxygen, but also may result in producing overly strong deoxygenation under global warming.

As seen in Fig. 5a, in the low-mixing models the decline in oxygen peaks at about 500 m, and is thus able to connect with low-latitude hypoxic regions, causing an expansion of hypoxic waters. The high mixing models, by contrast are able to maintain low stratification at 300 m and therefore see little change in oxygen at this depth (blue, green, pink and aqua point clouds in Fig. 9a). As a result, the decreases in tropical age and export production have a bigger impact on the volume of hypoxic waters, which decreases.

At 3000 m, there is a general decrease in oxygen concentration under global warming, but the largest drops are seen in the high mixing cases (green, blue, aqua and pink point clouds in Fig. 9c). This is because at 3000 m the low mixing cases start off stratified, and are essentially unventilated throughout (as seen from the stratification-age relationship in Fig. 9d). Increasing stratification thus has little impact on oxygen concentrations at depth (red and black point clouds...
in Fig. 9b). However, in the high mixing cases, there is an increase in salinity stratification over time which results in large drops in oxygen concentration at this depth (green and blue point clouds in Fig. 9b). In this region and depth, the increase in temperature stratification does play a role in stabilizing the water column, but it is an order of magnitude less important than salinity. In contrast to the results at 300 m, the observed oxygen and stratification (symbol, Fig. 9b) are closest to the initial state of the AREDI400 run (and overlap the global warming case for the AREDI800 run), and the relationship between the two is generally consistent with the inverse relationship between stratification and oxygen.

The dependence of age and oxygen on stratification in Fig. 9 are almost perfect mirror images. This suggests that their changes are both consequences of a locally driven decrease in ventilation. If changes in production were extremely important we would expect to see different shapes in the two point clouds. A similar correspondence is seen in the South Pacific sector of the Southern Ocean at 300 m (Fig. S2a and b), implying that a similar decrease in local ventilation is at work here. However, in the Weddell Sea the picture is quite different (Fig. S2 c and d) with no clear relationship between local stratification and either age or oxygen, although the two fields are closely related to each other. What appears to occur here is that all models see a shutoff in deep convection in this region, resulting in a long-term adjustment of both age and oxygen to deep ocean which is not complete at the end of the 140-year simulation. Within a single ESM, Yamamoto et al (2015) found similar long timescales of change, with a recovery of convection in the Weddell Sea and the associated increase in deep oxygen occurring 500 years after CO₂ was doubled.

5.3.2 Salt and fresh-water fluxes
In both the Pacific sector of the Southern Ocean and the Northwest Pacific, there are notable differences in how salinity stratification changes across the different models. These differences could be driven by the surface freshwater flux (the high-mixing models are warmer and so might be expected to transport more freshwater polewards) or by parameterized salt flux (the high-mixing models are more effective at supplying salt to high latitudes). As shown in Fig. 10a, the meridional freshwater fluxes in the control models are relatively similar across the model suite while the parameterized eddy salt fluxes (Fig. 10c) are quite different. As previously argued by Pradal & Gnanadesikan (2014) for the four constant AREDI simulations, it is the higher mixing supplying more salt to high northern and southern latitudes that drives destabilization of these regions and increased convection. It is worth noting that even though the ABER2D and ABERZONAL models have peak diffusivities much larger than AREDI2400, the associated salt fluxes (not previously reported) are bracketed by the AREDI1200 and AREDI2400 simulations.

Under global warming, the water fluxes show a pattern of increased export from the subtropics and increased convergence in the equatorial and subpolar regions. In northern latitudes, the largest signal is seen in the low mixing runs (AREDI400 and AREDI800, black and red lines respectively in Figure 10b), but the changes are relatively similar across models particularly in high latitudes.

The same cannot be said for the changes in salt flux (Fig. 10d). In the Southern Ocean, the poleward eddy salt flux decreases under global warming- representing a positive feedback on the salinity stratification. The decrease is anti-correlated with the mixing coefficient, (which as discussed in Gnanadesikan et al., (2015a,b) is low in subpolar latitudes for the ABER2D and ABERZONAL simulations) but is of the same sign for all of them. This may explain why all the
models show a cessation of deep Southern Ocean convection. In the northern subpolar oceans, the eddy salt flux increases under global warming with the largest increases seen in the higher mixing cases. In the northern subpolar gyre, the stronger negative feedback implied by eddy salt flux changes helps explain why the changes in salinity stratification and thus vertical exchange are weaker in the higher mixing simulations.

6. Conclusions

We have shown that the value of the lateral diffusive mixing coefficient $A_{REDI}$ can have a significant impact on the pattern of ocean deoxygenation and changes in the amount of hypoxic and suboxic waters. Cross-model differences in global mean deoxygenation are largely determined by the differences in biological drawdown which, consistent with previous work dominates at depth (Bopp et al., 2002; Bopp et al., 2013; Cocco et al., 2013; Cabrè et al., 2015; Keller et al., 2016). Rather than being driven by differences in export (and thus oxygen utilization rate) the differences across models are determined by how long the utilization of oxygen is allowed to accumulate. This in turn is controlled by deep convection in high latitudes, which is very sensitive to the value of $A_{REDI}$. Interestingly, this implies that the preformed nutrient content of the ocean as a whole must decrease. While the general findings of this study are consistent with current understanding (Shepherd et al., 2017), our quantification of the impact of lateral mixing represents a new result.

Some of the lessons of our sensitivity study can be applied to the CMIP5 models. Repeating the analysis in Fig. 5 for these models (Fig. S3), we find that they also show tongues of oxygen decrease emanating from high latitudes and oxygen increases at mid-depths in the
tropics. Similar to our model suite, the CMIP5 models show a significant variation in the depths to which these tongues of decreasing oxygen penetrate, the magnitude of the peak decrease and the extent to which they project onto regions that are close to being hypoxic. Our results suggest that these cross-model differences are strongly affected by differences in deep convection. Further work is needed however, to understand the physical balances that lead to such differences.

The pattern of tropical oxygen increase and decrease in hypoxic volume is less consistent across the CMIP5 models than it is in our model suite. This may be because models handle feedbacks between oxygen and nutrient cycling differently (particularly denitrification and nitrogen fixation, as discussed by Tyrell, 1999). Although the variation in $A_{\text{REDI}}$ across the CMIP5 models has the potential to produce differences of similar magnitude to those seen in Figs. 1 and S3, it is unlikely to be the sole driver.

As has been previously noted, the observed increase in hypoxic volumes in recent decades (Schmidtko et al., 2017) is not consistently found under global warming in the models that most reliably reproduce the oxygen profile and distribution (Oschlies et al., 2018). This could imply that the observed changes are due to natural variability connected with the Pacific Decadal Oscillation (Czechel et al., 2012; Ito & Deutsch, 2013; Duteil et al., 2018). Alternatively, our results suggest that models which have too vigorous and stable deep convection in the North Pacific (whether because they have $A_{\text{REDI}}$ coefficients that are too large or for some other reason) will also fail to produce the observed increase in hypoxia.

Our results with the ABER2D model suggest that implementing a more realistic spatial distribution of $A_{\text{REDI}}$ is possible without significantly degrading the simulation of oxygen. Although it would have been convenient if the zonal variation of this coefficient (represented by
ABERZONAL) completely captured the impacts of a two-dimensional mixing coefficient, this
does not appear to be the case. Even more work is needed to constrain the actual three-
dimensional distribution of lateral mixing and to understand its changes over time. Changes in
the mixing of oxygen over time by have been invoked to drive changes in regions such as the
tropical Atlantic (Brandt et al., 2010). Because we do not allow $A_{REDI}$ to vary in time or in the
vertical, our results should be seen at setting bounds on what a dynamic theory of mixing might
be able to accomplish. Our results also suggest that mixing coefficients in some regions,
particularly convective regions and the edges of oxygen minimum zones, may be more important
than other regions (such as the centers of the subtropical gyres). Dynamical theories for linking
mixing coefficients to stratification have been proposed for idealized scenarios such as zonally
uniform channels with a flat bottom, but have not been implemented in realistic climate models
with realistic bathymetry. Our results emphasize how important such work is for accurate
projection of oceanic biogeochemical cycling.

Acknowledgements: This work was supported by the National Oceanographic and Atmospheric
Administration under grant NA16OAR4310174 and the National Science Foundation Frontiers
in Earth Systems Dynamics Program under grant EAR-1135382 and by the JHU Institute for
Data-Intensive Engineering and Sciences. CMIP5 model output and the observed oxygen
datasets are publicly available. Output from our model suite is available at the JHU Dataverse
site doi:10.7281/T1/IWOVKD,
References


https://doi.org/10.1002/2013MS000273

doi:10.1038/nature15753


http://dx.doi.org/10.1098/rsta.2017.0240

http://doi.org/10.1029/2009JC005976


<table>
<thead>
<tr>
<th>Model</th>
<th>Institute</th>
<th>Ocean Biogeochemical Model</th>
<th>$A_{REDI} \text{ (m}^2\text{s}^{-1})$</th>
<th>$A_{GM} \text{ (m}^2\text{s}^{-1})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESM2G</td>
<td>Geophysical Fluid Dynamics Laboratory, NOAA, USA</td>
<td>TOPAZ</td>
<td>BCG closure (50-900)</td>
<td>BCG closure (100-900)</td>
</tr>
<tr>
<td>ESM2M</td>
<td>Geophysical Fluid Dynamics Laboratory, NOAA, USA</td>
<td>TOPAZ</td>
<td>600</td>
<td>BCG closure (100-800)</td>
</tr>
<tr>
<td>HadGEM2</td>
<td>Met Office Hadley Centre, UK</td>
<td>diat-HadOCC</td>
<td>500</td>
<td>BCG closure (150-2,000)</td>
</tr>
<tr>
<td>IPSL-CM5A-MR</td>
<td>Institute Pierre Simon Laplace, France</td>
<td>PISCES</td>
<td>1,000</td>
<td>BCG closure</td>
</tr>
<tr>
<td>MPI-ESM-MR</td>
<td>Max Planck Research</td>
<td>HAMOCC5</td>
<td>$2.5 \times 10^{-3} \cdot \Delta x, y$, Nominal 400</td>
<td>$1.6 \times 10^{-3} \cdot \Delta x, y$, Nominal 250</td>
</tr>
<tr>
<td>NorESM1-ME</td>
<td>Norwegian Climate Centre, Norway</td>
<td>MICOM</td>
<td>Equal to near-surface $A_{GM}$</td>
<td>BCG closure (0-1500)</td>
</tr>
</tbody>
</table>

**Table 1.** Summary of CMIP5 models shown Fig. 1, including biological model, value of $A_{REDI}$ coefficient and value of $A_{GM}$ coefficient. BCG denotes closures that scale with the baroclinic growth rate (effectively the vertical shear); in cases where minimum and maximum values are imposed, published values are given in parentheses.
<table>
<thead>
<tr>
<th>$A_{\text{REDI}}$ (m²/s)</th>
<th>Global average Temp (°C)</th>
<th>O₂ concentration (μM/kg)</th>
<th>Volume of O₂&lt;88 μM (Mkm³)</th>
<th>Volume of O₂&lt;20 μM (Mkm³)</th>
<th>Remineralized Phosphate (PO₄-PO₄₉ₑ₉, μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Obs.</td>
<td>3.660</td>
<td>177.1</td>
<td>150.1</td>
<td>17.70</td>
<td>N/A</td>
</tr>
<tr>
<td>400</td>
<td>4.846</td>
<td>154.7</td>
<td>182.6</td>
<td>51.28</td>
<td>0.905</td>
</tr>
<tr>
<td>800</td>
<td>4.604</td>
<td>161.4</td>
<td>141.8</td>
<td>31.39</td>
<td>0.844</td>
</tr>
<tr>
<td>1200</td>
<td>4.379</td>
<td>173.5</td>
<td>102.1</td>
<td>20.10</td>
<td>0.755</td>
</tr>
<tr>
<td>2400</td>
<td>4.244</td>
<td>185.5</td>
<td>42.71</td>
<td>5.450</td>
<td>0.661</td>
</tr>
<tr>
<td>2D</td>
<td>4.604</td>
<td>168.3</td>
<td>87.0</td>
<td>13.3</td>
<td>0.775</td>
</tr>
<tr>
<td>Zonal</td>
<td>4.448</td>
<td>176.7</td>
<td>53.9</td>
<td>4.84</td>
<td>0.750</td>
</tr>
</tbody>
</table>

**Table 2.** Globally averaged temperature, globally averaged O₂ concentration, globally integrated volume of hypoxic waters (O₂<88 μM) and suboxic waters (O₂<20 μM), and globally averaged remineralized phosphate concentration for six parameterizations of A₁REDI run in CM2Mc. Model results are from a 100-year average of the climatological run. There are no observations of preformed phosphate available.
### Table 3

<table>
<thead>
<tr>
<th>$A_{\text{REDI}}$ (m$^2$/s)</th>
<th>Temp change (°C)</th>
<th>$O_2$ concentration change (μM)</th>
<th>Volume of $O_2&lt;88$ μM change (Mkm$^3$)</th>
<th>Volume of $O_2&lt;20$ μM change (Mkm$^3$)</th>
<th>Remineralized Phosphate (PO$_4^{-}$PO$_4^{-}\text{pre}$) change (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>400</td>
<td>0.43</td>
<td>-6.2</td>
<td>8.3</td>
<td>-1.8</td>
<td>0.033</td>
</tr>
<tr>
<td>800</td>
<td>0.42</td>
<td>-7.7</td>
<td>3.3</td>
<td>-2.3</td>
<td>0.046</td>
</tr>
<tr>
<td>1200</td>
<td>0.42</td>
<td>-9.4</td>
<td>-4.0</td>
<td>-2.7</td>
<td>0.056</td>
</tr>
<tr>
<td>2400</td>
<td>0.39</td>
<td>-8.6</td>
<td>-2.4</td>
<td>-1.1</td>
<td>0.050</td>
</tr>
<tr>
<td>2D</td>
<td>0.42</td>
<td>-8.8</td>
<td>-1.5</td>
<td>-1.8</td>
<td>0.051</td>
</tr>
<tr>
<td>Zonal</td>
<td>0.41</td>
<td>-7.6</td>
<td>-2.0</td>
<td>-1.1</td>
<td>0.041</td>
</tr>
</tbody>
</table>

Global average changes in temperature, oxygen and remineralized phosphate and globally integrated changes in hypoxic and suboxic volume 140 years after instantaneous CO$_2$ doubling.
<table>
<thead>
<tr>
<th></th>
<th>Global</th>
<th>Atlantic 30N-65N</th>
<th>Pacific 30N-65N</th>
<th>Southern Ocean&lt;30S</th>
<th>Tropics 30S-30N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satellite</td>
<td>9.8 +/- 20%</td>
<td>1.1</td>
<td>1.5</td>
<td>2.4</td>
<td>3.9</td>
</tr>
<tr>
<td>AREDI400</td>
<td>9.95 (-0.73)</td>
<td>1.31 (-0.14)</td>
<td>1.14 (-0.12)</td>
<td>2.69 (-0.09)</td>
<td>5.01 (-0.45)</td>
</tr>
<tr>
<td>AREDI800</td>
<td>10.4 (-0.78)</td>
<td>1.31 (-0.09)</td>
<td>1.43 (-0.22)</td>
<td>2.86 (-0.14)</td>
<td>5.02 (-0.39)</td>
</tr>
<tr>
<td>AEDI1200</td>
<td>10.7 (-0.73)</td>
<td>1.26 (-0.04)</td>
<td>1.56 (-0.20)</td>
<td>2.96 (-0.15)</td>
<td>5.12 (-0.39)</td>
</tr>
<tr>
<td>AREDI2400</td>
<td>11.1 (-0.73)</td>
<td>1.28 (-0.08)</td>
<td>1.65 (-0.09)</td>
<td>3.25 (-0.17)</td>
<td>5.08 (-0.44)</td>
</tr>
<tr>
<td>ABER2D</td>
<td>10.9 (-0.78)</td>
<td>1.25 (-0.06)</td>
<td>1.62 (-0.18)</td>
<td>3.14 (-0.20)</td>
<td>5.13 (-0.41)</td>
</tr>
<tr>
<td>ABERZONAL</td>
<td>10.9 (-0.67)</td>
<td>1.26 (-0.13)</td>
<td>1.60 (-0.05)</td>
<td>3.04 (-0.16)</td>
<td>5.21 (-0.37)</td>
</tr>
</tbody>
</table>

**Table 4.** Export production in model suite and observation in Gt C/yr. Model results shown using 100-year average climatology and across a depth of 100m. Changes for years 40-140 after doubling are shown in parentheses.
**Figure Captions**

**Fig. 1.** Oxygen in six CMIP5 models: top panels show the preindustrial control runs compared to observations (symbol). bottom panels show changes after a perturbation simulation where atmospheric CO2 is increased at a rate of 1%/yr for 70 years and then held constant for another 70 years (denoted as “1% to doubling”). (A) Horizontally averaged oxygen in control simulations with observations shown by the symbols. (B) Hypoxic volume in Mkm$^3$/deg in control simulation with observations shown by the symbols. (C) Change in oxygen concentration at end of the 1% to doubling simulations. (D) Change in hypoxic volume at the end of the 1% to doubling simulations.

**Fig. 2.** Eddy mixing coefficients in CM2Mc for the Northwest Pacific with currents at 200 m overlaid. (A) $A_{GM}$ predicted from the vertical shear between 200 and 1500 m. (B) $A_{REDI}$ values predicted from altimetry done by Abernathey & Marshall (2013) used in ABER2D simulation.

**Fig. 3.** Same as Fig. 1 but for CM2Mc model suite.

**Fig. 4.** Taylor diagram evaluating fit to observations for temperature (T), salinity (S) and oxygen (O) concentrations relative to World Ocean Atlas 2009. (A) In CM2Mc models for a century average from the 400-500$^{th}$ year of the model spin-up. (B) For the CMIP5 models in Fig. 1. The radial coordinate shows the normalized standard deviation, angle shows correlation. A model perfectly matching observations would lie in coordinate [1,0]. The further from this point, the higher the RMS error.
Fig. 5. Zonally integrated change in hypoxic area (colors, Mkm$^2$/deg) and oxygen concentration (contours, μM) after 140 years of doubled CO$_2$ in CM2Mc model suite. (A) AREDI400 (B) AREDI800, (C) AREDI1200 (D) AREDI2400 (E) ABER2D (F) ABERZONAL.

Fig. 6. Zonally integrated particle export flux across 100 m the CM2Mc model suite in Gt C/deg/yr) compared to observations (symbols) from Dunne et al. (2007). (A) Control model. (B) Change over years 40-140 after instantaneous doubling of CO$_2$.

Fig. 7. Changes in the distribution of oxygen concentration (colors, μM) and age (contours, yr) at 300 m 140 years after doubling CO$_2$ across our suite of simulations. Contour range is from -100 yr by 20 yr. The low mixing cases (AREDI400 and AREDI800, panels A and B) show the largest depletion in the North Pacific at this depth. The two spatially varying cases, (ABER2D and ABERZONAL, panels E and F) simulate large changes in the Southern Ocean. Changes in the higher mixing models (AREDI1200 and AREDI2400, panels C and D) are generally smaller in high latitudes and larger in low latitudes.

Fig. 8. Changes in oxygen concentration (μM) and age at 3000 m 140 years after doubling CO$_2$ across suite of simulations. Contour range is doubled at this depth from -300 to 300 yr by 30. The high-mixing models (C, D) experience a significant decline in oxygen concentration in the Northwest Pacific, while the low-mixing cases (A, B) experience no change/increases at 3000 m depth in this region. The spatially variable mixing runs (E, F) produce similar results to high-mixing models.
Fig. 9. Oxygen and age in the NW Pacific at two depths as a function of salinity stratification across the CM2Mc model suite. Symbols in left-hand plots show observations. (A) Oxygen vs. salinity stratification relative to 300 m. (B) Ideal vs. salinity stratification relative to 300 m. (C and D) mimic top two panels but at 3000 m.

Fig. 10. Northward fluxes of freshwater (top row) and salt carried by parametrized eddies (bottom row). Left column shows control conditions. Right column shows the change under global warming. Note that a downward slope to the right of the freshwater flux implies freshening (implicitly stabilizing the water column), while for the salt flux it implies salinification (implicitly destabilizing the water column).
(A) Global mean $O_2$, CMIP5 Control

(B) Volume of hypoxic waters, CMIP5 Control

(C) $\Delta O_2$, CMIP5 Model, 1pct Runs

(D) $\Delta$Volume of hypoxic waters, CMIP5 1Pct
Figure 2.
(A) Global mean $O_2$, JHU runs

(B) Hypoxic volume

(C) ΔGlobal mean $O_2$, JHU runs

(D) ΔHypoxic volume
Figure 5.
(A) Particle export flux (100m)

(B) Change in Export Production 40–140 yrs after CO₂ doubling
Figure 7.
Figure 8.
(A) $\Delta O_2$ and $\Delta Age$ AREDI400, 3000m

(B) $\Delta O_2$ and $\Delta Age$ AREDI800, 3000m

(C) $\Delta O_2$ and $\Delta Age$ AREDI1200, 3000m

(D) $\Delta O_2$ and $\Delta Age$ AREDI2400, 3000m

(E) $\Delta O_2$ and $\Delta Age$ ABER2D, 3000m

(F) $\Delta O_2$ and $\Delta Age$ ABERZONAL, 3000m
Figure 10.