

# An Atmospheric Constraint on the Seasonal Air-Sea Exchange of Oxygen and Heat in the Extratropics

Eric J. Morgan<sup>1</sup> , Manfredi Manizza<sup>1</sup> , Ralph F. Keeling<sup>1</sup> , Laure Resplandy<sup>2</sup> , Sara E. Mikaloff-Fletcher<sup>3</sup> , Cynthia D. Nevison<sup>4</sup> , Yuming Jin<sup>1</sup> , Jonathan D. Bent<sup>1,5,6</sup> , Olivier Aumont<sup>7</sup> , Scott C. Doney<sup>8</sup> , John P. Dunne<sup>9</sup> , Jasmin John<sup>9</sup> , Ivan D. Lima<sup>10</sup> , Matthew C. Long<sup>5</sup> , and Keith B. Rodgers<sup>11,12</sup> 

<sup>1</sup>Geosciences Research Division, Scripps Institution of Oceanography, University of California, San Diego, CA, USA, <sup>2</sup>Geosciences Department and High Meadows Environmental Institute, Princeton University, Princeton, NJ, USA, <sup>3</sup>National Institute of Water and Atmospheric Research, Wellington, New Zealand, <sup>4</sup>INSTAAR University of Colorado, Boulder, CO, USA, <sup>5</sup>National Center of Atmospheric Research, Boulder, CO, USA, <sup>6</sup>Now at Picarro, Inc., Santa Clara, CA, USA, <sup>7</sup>LOCEAN, Paris, France, <sup>8</sup>Department of Environmental Sciences, University of Virginia, Charlottesville, VA, USA, <sup>9</sup>NOAA/OAR Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA, <sup>10</sup>Woods Hole Oceanographic Institution, Woods Hole, MA, USA, <sup>11</sup>Center for Climate Physics, Institute for Basic Science, Busan, South Korea, <sup>12</sup>Pusan National University, Busan, South Korea

## Key Points:

- Atmospheric O<sub>2</sub> and Ar can constrain the hemispheric air-sea oxygen to heat flux ratio
- This constraint is insensitive to uncertainties in atmospheric transport
- It is also a simple metric for testing ocean biogeochemical models

## Supporting Information:

Supporting Information may be found in the online version of this article.

## Correspondence to:

E. J. Morgan,  
[ejmorgan@ucsd.edu](mailto:ejmorgan@ucsd.edu)

## Citation:

Morgan, E. J., Manizza, M., Keeling, R. F., Resplandy, L., Mikaloff-Fletcher, S. E., Nevison, C. D., et al. (2021). An atmospheric constraint on the seasonal air-sea exchange of oxygen and heat in the extratropics. *Journal of Geophysical Research: Oceans*, 126, e2021JC017510. <https://doi.org/10.1029/2021JC017510>

Received 23 APR 2021  
 Accepted 20 JUL 2021

**Abstract** The air-sea exchange of oxygen (O<sub>2</sub>) is driven by changes in solubility, biological activity, and circulation. The total air-sea exchange of O<sub>2</sub> has been shown to be closely related to the air-sea exchange of heat on seasonal timescales, with the ratio of the seasonal flux of O<sub>2</sub> to heat varying with latitude, being higher in the extratropics and lower in the subtropics. This O<sub>2</sub>/heat ratio is both a fundamental biogeochemical property of air-sea exchange and a convenient metric for testing earth system models. Current estimates of the O<sub>2</sub>/heat flux ratio rely on sparse observations of dissolved O<sub>2</sub>, leaving it fairly unconstrained. From a model ensemble we show that the ratio of the seasonal amplitude of two atmospheric tracers, atmospheric potential oxygen (APO) and the argon-to-nitrogen ratio (Ar/O<sub>2</sub>), exhibits a close relationship to the O<sub>2</sub>/heat ratio of the extratropics (40–70°). The amplitude ratio,  $A_{APO}/A_{ArN_2}$ , is relatively constant within the extratropics of each hemisphere due to the zonal mixing of the atmosphere.  $A_{APO}/A_{ArN_2}$  is not sensitive to atmospheric transport, as most of the observed spatial variability in the seasonal amplitude of  $\delta APO$  is compensated by similar variations in  $\delta(Ar/N_2)$ . From the relationship between O<sub>2</sub>/heat and  $A_{APO}/A_{ArN_2}$  in the model ensemble, we determine that the atmospheric observations suggest hemispherically distinct O<sub>2</sub>/heat flux ratios of  $3.3 \pm 0.3$  and  $4.7 \pm 0.8$  nmol J<sup>-1</sup> between 40 and 70° in the Northern and Southern Hemispheres respectively, providing a useful constraint for O<sub>2</sub> and heat air-sea fluxes in earth system models and observation-based data products.

**Plain Language Summary** Typically, the surface of the ocean releases oxygen to the atmosphere during summer and takes it up during winter. This cycle is driven by circulation, biology (photosynthesis and respiration), and the seasonal cycle in water temperature, which changes the solubility of oxygen in surface water. We have used measurements of two atmospheric tracers, one which tracks oxygen and one which tracks heat, to estimate the amount of oxygen taken up or released by a change in ocean heat content. By looking at ocean models and atmospheric observations of the two atmospheric tracers, we find that the oxygen exchange between the ocean and atmosphere in the Southern Hemisphere is more responsive to changes in heat content than in the Northern Hemisphere. These hemispheric metrics are useful tests of how ocean models simulate some biological and physical processes.

## 1. Introduction

In extratropical latitudes a prominent seasonal cycle exists in the exchange of oxygen (O<sub>2</sub>) between the ocean surface and the atmosphere, progressing from net evasion in spring/summer to net invasion in fall/winter (Emerson, 1987; Jenkins & Goldman, 1985). This seasonality is partly driven by variations in ocean photosynthesis/respiration, with associated changes in the stratification of the surface layer

(Keeling, Stephens, et al., 1998; Najjar & Keeling, 1997). During the spring/summer months, O<sub>2</sub> is produced via photosynthesis in the upper ocean and released into the atmosphere. During fall/winter, rates of photosynthesis are low, and vertical mixing brings O<sub>2</sub>-depleted waters to the surface, causing O<sub>2</sub> to be taken up from the atmosphere, a process known as ventilation. Superimposed on these biologically driven exchanges of O<sub>2</sub> are seasonal changes in O<sub>2</sub> solubility, which can also drive air-sea fluxes (Keeling et al., 1993; Najjar & Keeling, 2000). Since dissolved O<sub>2</sub> is less soluble in warm water than in cold water, warming of the ocean surface releases O<sub>2</sub> to the atmosphere while cooling causes O<sub>2</sub> uptake, and in consequence O<sub>2</sub> fluxes due to thermal effects have similar phasing to those due to biological effects, as both processes cause O<sub>2</sub> to be released to the atmosphere when the ocean surface is warming, and cause uptake when it is cooling.

The total seasonal O<sub>2</sub> flux, combining these biological and thermal effects, has been shown to be closely correlated to the net air-sea heat flux when averaged over broad latitude bands and monthly time scales (Garcia & Keeling, 2001; Keeling & Garcia, 2002; Manizza et al., 2012). Keeling and Garcia (2002) showed that seasonal anomalies of heat and O<sub>2</sub> fluxes, estimated from bulk parameterizations, were sufficiently well correlated with each other throughout the seasonal cycle to allow resolving the ratio of the O<sub>2</sub> flux to the heat flux for spatial domains extending over several degrees of latitude and longitude. These ratios spanned a range of  $\sim 1.5 \text{ nmol J}^{-1}$  in subtropical latitudes to  $\sim 5 \text{ nmol J}^{-1}$  in the subpolar latitudes of both hemispheres, with slightly higher O<sub>2</sub>/heat ratios in the Southern Hemisphere extratropics than in the Northern, an asymmetry, which suggests a hemispheric-scale difference in the biological response to ocean warming. The accuracy of these estimates is limited, however, by uncertainty in the bulk parameterization (e.g., exchange velocities) and sparse measurements of dissolved O<sub>2</sub> in surface waters (Garcia & Keeling, 2001).

Earth system models predict that continued ocean heat uptake over the next century will drive significant loss of oceanic O<sub>2</sub>, resulting in widespread deoxygenation with cascading effects on marine ecology and carbon cycling (Bopp et al., 2013; Doney, 2010; Gruber, 2011; Keeling et al., 2010). Since many earth system models underestimate current rates of deoxygenation (Oschlies et al., 2018; Schmidtko et al., 2017), and the magnitude of interannual variations (Eddebbar et al., 2017; Long et al., 2016; Oschlies et al., 2018; Rödenbeck et al., 2008), an important test of these models is to verify that they duplicate the known response of biology to warming and stratification on a range of time scales, including on seasonal time scales. As suggested by Keeling et al. (2010), there is likely a broad similarity between the O<sub>2</sub> flux/heat flux ratios expected from long-term warming and the corresponding ratios from seasonal warming.

Here, we develop an improved estimate of the ratio of the seasonal air-sea flux of O<sub>2</sub> and heat, based on measurements of two atmospheric tracers, atmospheric potential oxygen (APO) and the argon-to-nitrogen ratio, Ar/N<sub>2</sub>. APO is effectively the sum of atmospheric O<sub>2</sub> and CO<sub>2</sub> concentrations and varies seasonally mainly due to ocean atmosphere exchanges of O<sub>2</sub>; seasonal exchange of O<sub>2</sub> with the land biosphere are canceled in APO by the addition of CO<sub>2</sub> (Stephens et al., 1998; Battle et al., 2006) and seasonal variations of APO pertaining to fossil fuel burning are small (Steinbach et al., 2011). The seasonal cycle of Ar/N<sub>2</sub> is diagnostic of air-sea heat fluxes, since the atmospheric abundances of both Ar and N<sub>2</sub> are not significantly impacted by biology (Battle et al., 2003; Blaine, 2005). Both gases are released from the ocean surface when it is warmed and taken up when it is cooled, resulting in a positive anomaly in the atmosphere under warming conditions and a negative anomaly in the atmosphere under cooling conditions. The contribution of N<sub>2</sub> air-sea fluxes to atmospheric Ar/N<sub>2</sub> is smaller than that of Ar fluxes because the solubility of N<sub>2</sub> has a lower temperature dependence than does Ar (Hamme & Emerson, 2004). Changes in the atmospheric abundance of N<sub>2</sub> lead to small changes in Ar/N<sub>2</sub> due to the larger abundance of N<sub>2</sub> in the atmosphere (Blaine, 2005). Owing to the rapid zonal mixing of the atmosphere, the seasonal cycles in both  $\delta\text{APO}$  and Ar/N<sub>2</sub> integrate signals from the entire zonal band (Cassar et al., 2008; Keeling, Stephens, et al., 1998).

As the seasonal amplitude of APO is primarily driven by the seasonal air-sea exchange of O<sub>2</sub>, and the seasonal amplitude of Ar/N<sub>2</sub> is driven primarily by solubility effects, we explore the idea that the ratio of the seasonal amplitude of both tracers,  $A_{\text{APO}}/A_{\text{ArN}_2}$ , would exhibit a close relationship to the seasonal O<sub>2</sub>/heat ratio. Previously, both APO and Ar/N<sub>2</sub> have been used to test model simulations of ocean biogeochemistry (Battle et al., 2003, 2006; Bent, 2014; Blaine, 2005; Cassar et al., 2008; Rodgers et al., 2014;

Stephens et al., 1998). Our approach differs in that we focus on the ratio of the two, which provides a link with O<sub>2</sub>/heat flux ratios and, as we show, has the advantage of being less sensitive to atmospheric transport.

The purpose of this study is to use APO and Ar/N<sub>2</sub> measurements to establish hemispheric-scale constraints on the seasonal O<sub>2</sub>/heat flux ratios, and demonstrate their usefulness in testing ocean biogeochemical models. We present updated observations of the cycles in APO and Ar/N<sub>2</sub> from surface flask time series, and test, through atmospheric transport model simulations, the degree to which the amplitude ratio,  $A_{APO}/A_{ArN_2}$ , is sensitive to uncertainties in atmospheric transport. Drawing on a suite of five ocean biogeochemical models, we use  $A_{APO}/A_{ArN_2}$  to evaluate their performance and to constrain the average seasonal O<sub>2</sub>/heat flux ratio of each hemisphere, thus providing two observational metrics for testing ocean biogeochemical models:  $A_{APO}/A_{ArN_2}$  and the seasonal O<sub>2</sub>/heat ratio. We also compare the observed amplitude ratio to the ratio implied by observation-based climatologies of the seasonal extratropical air-sea O<sub>2</sub> and heat fluxes, thus providing a basis for improving upon prior estimates (e.g., by Garcia & Keeling 2001), and producing an additional test of the performance of ocean models.

## 2. Methods

### 2.1. Atmospheric Measurements

The procedures for collecting and analyzing flasks for CO<sub>2</sub>, O<sub>2</sub>/N<sub>2</sub>, and Ar/N<sub>2</sub> at Scripps Institution of Oceanography (SIO) have been described previously in detail (Keeling, Manning, et al., 1998; Keeling et al., 2004, 2007). In brief, cryogenically dried samples are collected at ~1 atm pressure in 5-L glass flasks, sealed with Viton O-rings. Samples are collected in triplicate approximately every 2 weeks and returned to SIO for analysis. At the Cape Kumukahi site (see below), samples are collected approximately once per week and have been sampled in duplicate since 2010. O<sub>2</sub>/N<sub>2</sub> ratios are measured with a custom-built interferometer, CO<sub>2</sub> concentrations are measured with a Siemens nondispersive infrared (NDIR) CO<sub>2</sub> analyzer (Keeling, Manning, et al., 1998), and Ar/N<sub>2</sub> ratios are measured on an ISOPRIME mass spectrometer (Keeling et al., 2004). Instruments are calibrated daily with two high-pressure tanks. These calibrations are corrected for long-term drift through regular analysis of a suite 18 high-pressure tanks, which serve as primary references (Keeling et al., 2007).

Measurements of O<sub>2</sub>/N<sub>2</sub> are expressed as relative differences between a sample and reference according to:

$$\delta(O_2 / N_2) = \left( \frac{(O_2 / N_2)_{sample}}{(O_2 / N_2)_{reference}} - 1 \right) \times 10^6 \quad (1)$$

where  $\delta(O_2/N_2)$  is expressed in “per meg” units. Ar/N<sub>2</sub> ratios are similarly reported as  $\delta(Ar/N_2)$ :

$$\delta(Ar / N_2) = \left( \frac{(Ar / N_2)_{sample}}{(Ar / N_2)_{reference}} - 1 \right) \times 10^6 \quad (2)$$

Values of O<sub>2</sub>/N<sub>2</sub> are reported on the SIO 2017 O<sub>2</sub>/N<sub>2</sub> Scale, Ar/N<sub>2</sub> on the SIO Ar/N<sub>2</sub> Scale, and CO<sub>2</sub> on the SIO VH344-2020 CO<sub>2</sub> Scale, all as defined on April 13, 2021. Data from the Scripps O<sub>2</sub> Program are freely available from the program website at <https://scripps02.ucsd.edu/>.

We compute the tracer APO as:

$$\delta APO = \delta(O_2 / N_2) + \frac{1.05}{X_{O_2}}(CO_2 - 350) \quad (3)$$

where 1.05 is a nominal value for the long-term average O<sub>2</sub>:C molar exchange ratio of the land biosphere (Keeling & Manning, 2014), CO<sub>2</sub> is the dry air CO<sub>2</sub> mole fraction in ppm, and 350 is an arbitrary reference CO<sub>2</sub> concentration.  $X_{O_2}$  is the mole fraction of oxygen in the atmosphere as defined by the Scripps O<sub>2</sub> Program, 0.2094, and converts from mole fraction to per meg units (Manning & Keeling, 2006).  $\delta APO$  is in units of per meg, but can also be expressed as a molar amount, that is,  $APO = O_2 + 1.05 CO_2$ .

**Table 1**  
Summary of  $\delta APO$  and  $\delta(Ar/N_2)$  Measurements and Related Results, January 1, 2005–June 1, 2020

	Palmer	Cape Grim	Kumukahi	La Jolla	Cold Bay
	PSA	CGO	KUM <sup>a</sup>	LJO	CBA
Latitude (°N)	−64.77	−40.68	19.52	32.87	55.12
Longitude (°E)	−64.05	114.69	−154.81	−117.26	−162.43
$A_{APO}$ <sup>b</sup> (per meg)	65.2 ± 1.1	56.6 ± 1.1	35.0 ± 0.8	61.6 ± 1.4	70.1 ± 2.1
$A_{ArN_2}$ <sup>c</sup> (per meg)	18.6 ± 1.1	12.3 ± 1.2	11.6 ± 0.7	20.0 ± 1.0	21.8 ± 1.2
$A_{APO}/A_{ArN_2}$ <sup>d</sup>	3.5 ± 0.2	4.6 ± 0.5	3.0 ± 0.2	3.1 ± 0.2	3.2 ± 0.2
Slope <sup>e</sup> (J nmol <sup>−1</sup> )	0.94 ± 0.2	0.85 ± 0.2	0.67 ± 0.1	0.73 ± 0.1	1.0 ± 0.2
Intercept <sup>f</sup>	−0.26 ± 0.6	0.1 ± 0.6	1.1 ± 0.5	0.81 ± 0.5	−0.22 ± 0.8
$\gamma_{40-70}$ <sup>g</sup> (nmol J <sup>−1</sup> )	4.0 ± 0.2	5.4 ± 0.4	2.9 ± 0.2	3.1 ± 0.2	3.5 ± 0.2
	Southern Hemisphere		Northern Hemisphere		
$\gamma_{40-70}$ <sup>h</sup> (nmol J <sup>−1</sup> )	4.7 ± 0.8		3.3 ± 0.3		

<sup>a</sup>Includes data from substitute sampling sites KHP and LEL (see Methods). <sup>b</sup>Amplitude of the first harmonic of the average seasonal cycle in  $\delta APO$ . <sup>c</sup>Amplitude of the first harmonic of the average seasonal cycle in  $\delta(Ar/N_2)$ . <sup>d</sup>Amplitude ratio of the first harmonics of the average seasonal cycle in  $\delta APO$  and  $\delta(Ar/N_2)$ , in per meg/per meg. <sup>e</sup>Slope of regression of  $A_{APO}/A_{ArN_2}$  and  $\gamma_{40-70}$  for each site. <sup>f</sup>Intercept of regression of  $A_{APO}/A_{ArN_2}$  and  $\gamma_{40-70}$  for each site. Unitless in per meg/per meg. <sup>g</sup>Inferred  $\gamma_{40-70}$  based on observed  $A_{APO}/A_{ArN_2}$ , determined by solving the regression of modeled  $A_{APO}/A_{ArN_2}$  and  $\gamma_{40-70}$  (the  $O_2$ /heat flux ratio), for each site. <sup>h</sup>Hemispheric average of inferred  $\gamma_{40-70}$ . For the Northern Hemisphere this is the average of LJO and CBA only. Abbreviations: APO, atmospheric potential oxygen; KHP, Kahakai Park; LEL, Lelewi Point; PSA, Palmer Station.

We focus here on results from five stations in the Scripps  $O_2$  network, as summarized in Table 1: Palmer Station (PSA) and Cape Grim (CGO) in the Southern Hemisphere, and Cold Bay (CBA), La Jolla (LJO), and Cape Kumukahi (KUM) in the Northern Hemisphere. We exclude results from two additional nontropical stations in the network (South Pole and Alert Station) which are subject to strong winter-time temperature inversions, which we believe may impact  $Ar/N_2$  due to thermal fractionation (Adachi et al., 2006). We also exclude results from Samoa and Mauna Loa, where the  $Ar/N_2$  cycles are much weaker and provide less useful model constraints. We include results from KUM (19.54°N), which appears to have well-resolved cycles despite being in the tropics. Since June 2018, sampling at KUM has been discontinued due to the eruption of Kīlauea. The KUM record has been continued at two substitute sampling sites, Lelewi Point (LEL; 19.73°N; 155.01°W) and Kahakai Park (KHP; 19.56°N, 154.89°W). Data from LEL and KHP are considered to be part of the KUM record, and are referred to with station code KUM.

## 2.2. Determination of Seasonal Amplitudes in Atmospheric Data

To extract the seasonal component of each station's observed time series, we fit a second-order polynomial and two harmonics. Here, we are interested primarily in the amplitude of the annual harmonic, since the second harmonic of the  $\delta(Ar/N_2)$  time series is not as well resolved. Uncertainties were propagated within a Monte Carlo framework, by bootstrapping the curve fitting procedure 1,000 times, with the average first harmonic of all iterations taken as the seasonal amplitude of  $\delta APO$  ( $A_{APO}$ ) and  $\delta(Ar/N_2)$  ( $A_{ArN_2}$ ). For each iteration of the bootstrap the dataset is sampled  $n$  times with replacement, where  $n$  is the number of data points in the time series: 359, 305, 595, 300, and 315 for PSA, CGO, KUM, LJO, and CBA, respectively. Values are given in per meg, rendering the amplitude ratio ( $A_{APO}/A_{ArN_2}$ ) unitless in (per meg)/(per meg). The uncertainty of  $A_{APO}$  or  $A_{ArN_2}$  is reported as one standard deviation of the ensemble.

### 2.3. Ocean Biogeochemical Models

We compare atmospheric data to simulated atmospheric cycles using climatological air-sea gas fluxes calculated from a suite of five global ocean biogeochemical general circulation models (OBMs): (a) CESM, (b) CCSM, (c) MOM4, (d) NEMO-Control (NEMO-C), and (e) NEMO-WSTIR (NEMO-W). Each of these OBMs consists of a global ocean general circulation model (including an interactive sea-ice model) that embeds an ocean biogeochemical component. Each biogeochemical model has an ecosystem module where biological production of organic matter is based on availability of light and multiple nutrients (nitrogen, phosphorus, silicate, iron). For our study, we rely on the modeled estimates of the monthly air-sea fluxes of O<sub>2</sub>, CO<sub>2</sub> and heat. All models selected for this study are forced with different reanalyzed atmospheric products that include air temperature, wind-speed, solar radiation, and precipitation (e.g., CORE-II [Large & Yeager, 2009]). The atmospheric forcing imposed on these ocean models varies year by year according to climate variability (Large & Yeager, 2009). Modeled flux fields were averaged over the years 1990–2009 to produce a single monthly climatology.

#### 2.3.1. CESM and CCSM

The CCSM3 and CESM1 runs refer to forced runs of the ocean component of the NCAR Community Earth System Model. Full descriptions of the CCSM and CESM runs are given in Collins et al. (2006) and Long et al. (2013), respectively. In CESM, biogeochemistry was represented by the Biogeochemical Elemental Cycling (BEC) model (Moore et al., 2004, 2013) within the Parallel Ocean Program (POP) v2 general circulation model (Smith et al., 2010). CESM was forced with the Coordinated Ocean-Ice Reference Experiments CORE) v2, spun up with four 60-year cycles. CCSM was run with a related but distinct version of BEC and POP v1.4.3 and forced with CORE v2 (Large & Yeager, 2009). The ecological-biogeochemical simulation was spun up for several hundred years, prior to initiating the interannual varying forcing, using a repeat annual cycle of physical forcing and dust deposition. The model ecosystem components converge to a repeat annual cycle within a few years of spinup. The full interannual variability in physics is initiated in a model year equivalent to calendar year 1958 and then integrated forward through 2009.

#### 2.3.2. MOM4

MOM4 is a version of NOAA Geophysical Fluid Dynamics Laboratory's Modular Ocean Model (Griffies et al., 2005), where the TOPAZv0 biogeochemistry model (Dunne et al., 2013) is embedded. A technical description of TOPAZv0 is available in Dunne et al. (2010), as used in (Sarmiento et al., 2010). MOM4 implements a tripolar grid to explicitly resolve the Arctic Ocean circulation at a nominal 1° resolution, and includes various physical parameterizations including fresh water input to surface ocean, neutral and subgrid-scale diffusion, and an explicit surface mixed layer (Griffies et al., 2005). MOM4 was forced with CORE.v1. The simulation was spun up for 300 years with five cycles of the period 1948–2007.

#### 2.3.3. NEMO

NEMO-C and NEMO-W both use versions of the NEMO (Nucleus for European Modeling of the Oceans) model with versions of the NEMO-OPA (Océan PARallélisé) physical oceanography model (Madec, 2016), and versions of the PISCES biogeochemistry model (Aumont & Bopp, 2006) with the ORCA2 tripolar gridding scheme, which produces self-consistent circumpolar flows. NEMO-C and NEMO-W simulations were run by Rodgers et al. (2014) using NEMO 3.2 and the biogeochemistry component PISCES (Aumont & Bopp, 2006), and using the DRAKKAR upper-ocean forcing set 4.1 (Brodeau et al., 2010) based on ERA-40 wind reanalysis (Uppala et al., 2005).

NEMO-W is essentially a sensitivity test of NEMO-C, as they only differ in that NEMO-W contains a wind stirring parameterization to better match observational constraints in the Southern Ocean. The ad hoc parameterization increases summer mixed layer depths by introducing additional kinetic energy into the upper ocean as a parameterization of shear-induced turbulence. As discussed by Rodgers et al. (2014), this

parameterization has significant consequences for the air-sea exchange of O<sub>2</sub>, CO<sub>2</sub>, and heat. We include NEMO-W because it serves as a case study for how  $A_{APO}/A_{ArN_2}$  can be used to evaluate ocean biogeochemical models.

#### 2.4. Simulated N<sub>2</sub> and Ar Fluxes

Since Ar and N<sub>2</sub> are not carried as tracers in any of the ocean models, fluxes of these gases are estimated from ocean surface heat fluxes, using the formula of Keeling and Shertz (1992), as modified by Jin et al. (2007):

$$F = -\frac{1}{1.3} \frac{\partial S}{\partial T} \frac{Q}{C_p} \quad (4)$$

Here,  $\partial S / \partial T$  (mol kg<sup>-1</sup> °C<sup>-1</sup>) is the temperature derivative of the solubility (Weiss, 1970), and  $Q$  is the net air-sea heat flux (J m<sup>-2</sup> s<sup>-1</sup>).  $C_p$  is the heat capacity of seawater and is a function of salinity and temperature (Jamieson et al., 1969), but for simplicity is treated as a constant, 3993 J kg<sup>-1</sup> °C<sup>-1</sup>, since for a difference in  $T$  of 30°C the contribution to the change in  $F$  is only 0.4%.  $\partial S / \partial T$  depends significantly on temperature, so use of Equation 4 requires both heat fluxes and sea surface temperatures.

The factor of 1 / 1.3 in Equation 4, which we refer to as the “Jin correction factor,” was introduced by Jin et al. (2007), and accounts for incomplete equilibration and other processes which can decouple heat and gas fluxes. We do not expect the value to be strictly constant, since it must depend on factors which vary in space and time, such as mixed layer depth. Instead of attempting to refine Equation 4 to allow for such variations, we assign the Jin correction factor a mean of 1/1.3 = 0.77 with a 1 $\sigma$  uncertainty of  $\pm 0.08$ , the latter value being drawn from Table 2 of Manizza et al. (2012). Jin et al. (2007) also included a 2-week delay of N<sub>2</sub> fluxes compared to heat fluxes, which we neglect because we focus here on amplitudes rather than phasing.

We apply Equation 4 to compute N<sub>2</sub> fluxes, and then compute Ar fluxes by assuming a constant proportionality,  $E$ , between Ar and N<sub>2</sub> fluxes. This is justified by the near constancy of the ratio of the temperature dependence of the solubilities (Manizza et al., 2012), from which  $E$  is obtained:

$$E = \frac{\partial S_{Ar} / \partial T}{\partial S_{N_2} / \partial T} \quad (5)$$

The value of 1 /  $E$  (with 1 $\sigma$  uncertainty) is taken as  $34.00 \pm 0.06$ . Since the value of  $E$  is well constrained through empirical methods and supported by additional modeling experiments (Manizza et al., 2012), but varies slightly with temperature, we took the uncertainty as one standard deviation of the ratio of the solubilities over the temperature range of 5–20°C. Scaling Ar fluxes directly from N<sub>2</sub> fluxes allows changes in atmospheric Ar/N<sub>2</sub> ratio to be computed directly from changes in N<sub>2</sub>, without needing to carry Ar as a separate tracer in the atmospheric transport model. We neglect the role of N<sub>2</sub> fluxes due to nitrogen fixation since they are negligible when compared to the magnitude of seasonal fluxes due to heat exchange (Blaine, 2005).

#### 2.5. O<sub>2</sub>/Heat Flux Ratios

We define the seasonal O<sub>2</sub> flux to heat flux ratio,  $\gamma$ , as

$$\gamma = \frac{A_{FO_2}}{A_{FQ}} \quad (6)$$

where  $A_{FQ}$  is the seasonal amplitude of the net air-sea heat flux in J m<sup>-2</sup> s<sup>-1</sup>, and  $A_{FO_2}$  is the seasonal amplitude of the net air-sea O<sub>2</sub> flux in nmol m<sup>-2</sup> s<sup>-1</sup>, yielding an O<sub>2</sub>/heat ratio ( $\gamma$ ) in nmol J<sup>-1</sup>. O<sub>2</sub> and heat fluxes were each averaged over specific latitude bands to produce time series of heat and O<sub>2</sub> flux with a monthly resolution. These time series (with  $n = 12$ ) were then subjected to the same algorithms used to process the  $\delta$ APO and  $\delta$ (Ar/N<sub>2</sub>) time series (see Section 2.2), yielding average seasonal amplitudes by bootstrapping a two-harmonic fit 1,000 times. During this Monte Carlo simulation we also perturb each variable with uncertainty where appropriate in order to propagate the uncertainty in the harmonic fits through to  $\gamma$ . The reported

value of  $A_Q$  or  $A_{FO_2}$  is the ensemble average of the first harmonic, with the final uncertainty taken as one standard deviation of the same. We compute gamma values for 10° latitude bands from 90°S to 90°N, which we refer to as  $\gamma_i$ , and for the single band from 40° to 70° in each hemisphere, which we designate as  $\gamma_{40-70}$ .

## 2.6. Semi-Empirical Fluxes

In addition to modeled air-sea fluxes, we also compared atmospheric data to a simulation based on climatologies of bulk parameterized air-sea fluxes of O<sub>2</sub>, N<sub>2</sub>, and CO<sub>2</sub>. Since these are observationally based but not direct observations, we refer to them collectively as “semi-empirical.” O<sub>2</sub> fluxes are from the climatology of Garcia and Keeling (2001), and are based on interpolated observations of dissolved O<sub>2</sub>, temperature, and salinity from the years 1960–1998. The Garcia and Keeling (2001) fluxes were computed using the Wanninkhof (1992) gas exchange formula (W92), which has been subsequently shown to overestimate air-sea fluxes (Naegler, 2009; Wanninkhof, 2014). To correct for this bias, all fluxes have been scaled downward by a factor ( $\pm 1 \sigma$ ) of  $0.82 \pm 0.05$ . This adjustment is supported by Bent (2014), who found that the fluxes from Garcia and Keeling (2001) overstated observations of column-average APO by ~20%.

Semi-empirical N<sub>2</sub> fluxes were calculated from Equation 4 by scaling heat fluxes, sea-surface temperatures, and ice coverage from the Estimating the Circulation and Climate of the Ocean project, version 4 (ECCOV4) model (Forget et al., 2015). This product is a climatological version of the heat fluxes derived from ocean state estimates generated through an ocean general circulation model (MITgcm; Marshall et al., 1997) not only forced by re-analyzed products, but also constrained by available in-situ observations of temperature and salinity via the use of adjoint techniques generated by ECCOV4. Ar fluxes were generated from these N<sub>2</sub> fluxes using a constant scaling, as described above. Heat fluxes were averaged to a monthly climatology covering the period 1992–2012.

Semi-empirical CO<sub>2</sub> fluxes were obtained from the climatology of Takahashi et al. (2009), accessed from the National Centers for Environmental Information on November 14, 2018, with observations from 1970 to 2007.

## 2.7. Atmospheric Transport Modeling

All ocean model air-sea fluxes of O<sub>2</sub>, N<sub>2</sub>, and CO<sub>2</sub> were transported with a single atmospheric transport model, TM3 (Transport Model 3, see Heimann & Körner, 2003). Modeled fluxes with an irregular native grid (NEMO-W, NEMO-C, CCSM, CESM) were regridded in two steps, first to a regular latitude/longitude grid using inverse distance weighting interpolation, and then to the TM3 grid. For NEMO-W and NEMO-C CO<sub>2</sub>, for the first, regridding it was necessary to use a more sophisticated kriging method, due to artifacts created by the inverse distance weighting interpolation. All functions were taken from the R package “EFDR.” MOM4 and the semi-empirical fluxes were only regridded in a single step.

TM3 was run with a horizontal resolution of 4° × 5° (latitude × longitude) and 19 vertical levels. O<sub>2</sub>, CO<sub>2</sub>, and N<sub>2</sub> were treated as if they were trace gases, thus carrying anomalies relative to an arbitrary background concentration in ppm equivalent (1 ppm =  $\mu \text{ mol mol}^{-1}$ ), ignoring the changes in the total moles of dry air caused by the fluxes. The runs were carried out for six years, repeating transport fields for the year 1995 from NCEP (Kalnay et al., 1996) and discarding the first four years to avoid spin-up artifacts. To compare TM3 simulations to observations at the sampling stations, we extracted monthly average output at the lowest vertical level at the grid cell of the station.

For each OBM, the results of a multiyear ocean simulation (1990–2009) were detrended at each grid cell to produce a climatological average seasonal cycle with a monthly timestep. TM3 simulates anomalies as a mole fraction in parts per million (ppm) of dry air. From these, we computed:

$$\delta(O_2 / N_2) = \frac{\Delta O_2}{X_{O_2}} - \frac{\Delta N_2}{X_{N_2}} \quad (7)$$

$$\delta APO = \frac{\Delta O_2}{X_{O_2}} - \frac{\Delta N_2}{X_{N_2}} + \frac{1.05}{X_{O_2}} CO_2 \quad (8)$$

$$\delta(Ar / N_2) = \frac{E \Delta N_2}{X_{Ar}} - \frac{\Delta N_2}{X_{N_2}} \quad (9)$$

which yields  $\delta(O_2/N_2)$  and  $\delta(Ar/N_2)$  in per meg units.  $\Delta O_2$ ,  $\Delta CO_2$ , and  $\Delta N_2$  are the atmospheric concentrations of these tracers in ppm.  $X_{O_2}$ ,  $X_{N_2}$ , and  $X_{Ar}$  are constant reference mole fractions of  $O_2$ ,  $N_2$ , and Ar, and equal to 0.2094, 0.7808, and 0.00934, respectively.  $E$  is equal to 1/34, as discussed above. Though the marine  $CO_2$  component of APO typically contributes only a few per meg to the amplitude of the  $\delta APO$  cycle (Nevison et al., 2008), we include it here since the contribution could be larger in model simulations with large seasonal  $CO_2$  fluxes.

### 2.8. Uncertainty Propagation for Modeled Atmospheric Tracers

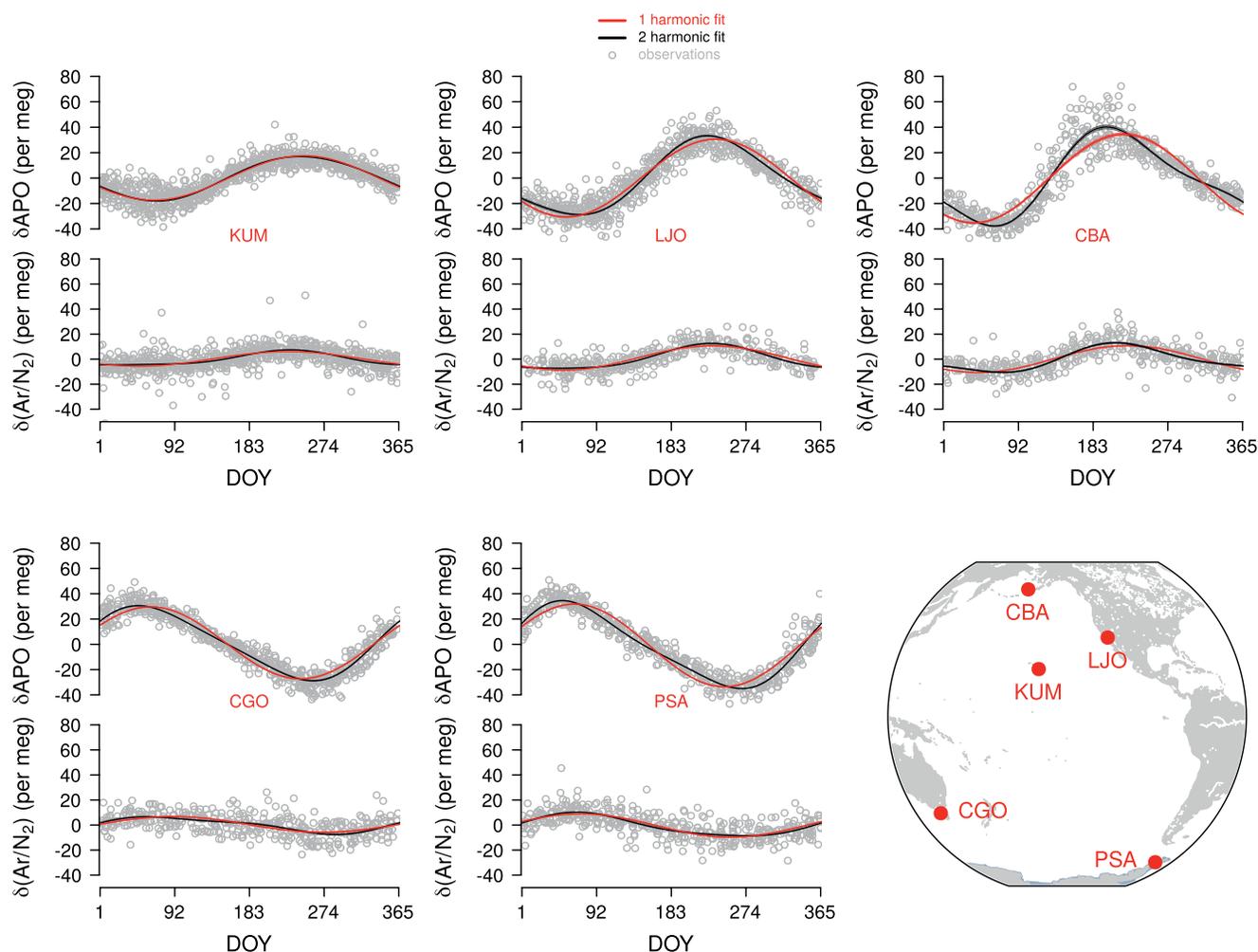
Simulated station time series from TM3 forward runs of the OBMs were treated in the same way as the observations, though with added sources of uncertainty. In each iteration of the Monte Carlo, in addition to bootstrapping the time series (sampling the average monthly time series of the corresponding grid cell with replacement 1,000 times, each time series having 12 values), we included uncertainty in the Jin et al. (2007) correction factor and in the value of  $E$ . For the semi-empirical  $O_2$  atmospheric fields, a further term was included to represent uncertainty in the gas exchange velocity correction.

We also included uncertainty due to atmospheric transport in the Monte Carlo framework. This applied to simulated  $A_{APO}$ ,  $A_{ArN_2}$ , and  $A_{APO}/A_{ArN_2}$  from the forward runs of the OBM fluxes (and the semi-empirical run). To estimate this uncertainty we drew upon results from a previous study, referred to here as ‘‘APO TransCom’’ (Blaine, 2005). This was a transport model intercomparison which used a set of monthly average  $O_2$  and  $N_2$  fluxes to compare seasonal cycles in APO simulated across a suite of atmospheric transport models. In Blaine (2005),  $O_2$  fluxes were from Garcia and Keeling (2001) without rescaling for the gas exchange velocity parameterization.  $N_2$  fluxes were calculated from heat fluxes from (Gibson et al., 1999) and sea-surface temperatures from Shea et al. (1992), using Equation 4, but omitting the Jin correction factor of 1/1.3. Since  $CO_2$  fluxes were not included by Blaine (2005), we computed values of  $A_{O_2N_2}$ ,  $A_{ArN_2}$ , and  $A_{O_2N_2}/A_{ArN_2}$  at 10 sites, including PSA, CGO, KUM, LJO, and CBA, treating  $A_{O_2N_2}$  as a proxy for  $A_{APO}$ . We used Equations 7 and 9 to obtain  $\delta(O_2/N_2)$  and  $\delta(Ar/N_2)$  time series at each station, and then took the first harmonic as the seasonal amplitude, as described above. The uncertainty in atmospheric transport at each site was taken as one standard deviation of the eight modeled values of  $A_{O_2N_2}$ ,  $A_{ArN_2}$ , and  $A_{O_2N_2}/A_{ArN_2}$ , and propagated through the Monte Carlo framework, using  $A_{O_2N_2}$  as representative of  $A_{APO}$ . The results are discussed in Section 3.2. All reported values of  $A_{O_2N_2}/A_{ArN_2}$  in this work from OBMs or the semi-empirical product include this uncertainty.

### 2.9. Determination of $O_2$ /Heat Flux Ratios Implied From Observed $A_{APO}/A_{ArN_2}$

To find the seasonal  $O_2$ /heat flux ratio,  $\gamma_{40-70}$ , implied from the observed values of  $A_{APO}/A_{ArN_2}$ , we performed a linear regression of  $\gamma_{40-70}$  and  $A_{APO}/A_{ArN_2}$  from the suite of ocean biogeochemical models. Since the model ensemble is small, we use both NEMO-W and NEMO-C (the former being a sensitivity test of the latter), as well as the semi-empirical flux product in the regression. We then solved these regression equations for  $\gamma_{40-70}$ , using observed  $A_{APO}/A_{ArN_2}$ . Simulated  $A_{APO}/A_{ArN_2}$  values were taken from each station's location in the atmospheric field produced by the forward run of the OBM, and simulated  $\gamma_{40-70}$  values were taken from the hemisphere in which the station lies.

We then performed a standard major axis (SMA) regression (Legendre & Legendre, 2012) 1,000 times, each time randomly perturbing  $A_{APO}/A_{ArN_2}$  and  $\gamma_{40-70}$  within their propagated  $1\sigma$  uncertainties. Final hemispheric averages of  $\gamma_{40-70}$  were determined as part of the Monte Carlo, using PSA and CGO for the Southern Hemisphere and LJO and CBA for the Northern Hemisphere.



**Figure 1.** Average seasonal cycles of observed, detrended  $\delta$ APO(atmospheric potential oxygen) and  $\delta(\text{Ar}/\text{N}_2)$  at select sites from the Scripps Institution of Oceanography (SIO)  $\text{O}_2$  Program, plotted as a function of the day of the year (DOY) of sampling. Each point represents the average of a flask pair or trio. Red lines show a single harmonic fit and black lines show a two-harmonic fit. In the bottom right panel, site locations are shown. The observations are from the period January 1, 2005–June 1, 2020.

### 3. Results and Discussion

#### 3.1. Observed Seasonal Amplitudes

Figure 1 shows the average seasonal cycles of  $\delta$ APO and  $\delta(\text{Ar}/\text{N}_2)$  as observed at five surface stations, with one and two harmonic fits to the data. A summary of the fitted parameters for the annual harmonic (fundamental) is given in Table 1. The  $\delta$ APO and  $\delta(\text{Ar}/\text{N}_2)$  cycles are similar to those reported by the Scripps  $\text{O}_2$  Program for earlier periods (Keeling, Stephens, et al., 1998; Keeling et al., 2004). The seasonal cycles are clearly defined in both tracers and exhibit little year-to-year variability. The scattered outliers visible in both records were retained because they had no associated analytical flags. In the CBA  $\delta$ APO record, summertime synoptic events are visible, though they have only a minor influence on the annual harmonic fit.

The cycles in  $\delta(\text{Ar}/\text{N}_2)$  and  $\delta$ APO have similar phasing, with the maxima of the annual harmonic generally aligned to within  $\sim 1$  week. At two sites, we find larger phase differences: at KUM, where  $\delta$ APO lags by  $20 \pm 4$  days, and at CGO, where  $\delta$ APO leads  $\delta(\text{Ar}/\text{N}_2)$  by  $27 \pm 6$  days. With two harmonics the phase differences at CGO and KUM both reduce to 11 days. An independent  $\delta(\text{Ar}/\text{N}_2)$  dataset of observations at CGO from the Princeton University program resolved the  $\delta(\text{Ar}/\text{N}_2)$  cycle with an annual maximum several weeks earlier than found here (Cassar et al., 2008; Nevison, Keeling, et al., 2012), and thus in better phase alignment

with  $\delta A_{PO}$ . While better agreement may exist in the phasing of the higher harmonics, the  $\delta(Ar/N_2)$  data in either record is not precise enough to reliably resolve them, and we do not interpret these smaller phase offsets as a significant challenge to the close relationship between the seasonal cycles of the two tracers.

Depending on the station, the amplitude ratios ( $A_{APO}/A_{ArN_2}$ ) based on the annual harmonic are between 3.0 and 4.6, with  $1\sigma$  uncertainties of 5%–10% (Table 1). Despite the large latitudinal variation in the magnitude of  $A_{APO}$  and  $A_{ArN_2}$ , the three Northern Hemisphere sites are all lower in  $A_{APO}/A_{ArN_2}$  than the two Southern Hemisphere sites. This suggests that  $A_{APO}/A_{ArN_2}$  is characteristic of a large portion of the atmosphere in both hemispheres. The difference between PSA and CGO is larger when compared to the difference between the three Northern Hemisphere stations, indicating that there may be larger latitudinal gradients in  $A_{APO}/A_{ArN_2}$  in the Southern Hemisphere, or that PSA or CGO may have a higher degree of regional influence. The  $\delta(Ar/N_2)$  in the Princeton University record was influenced by sampling artifacts which tended to exaggerate the amplitude of the seasonal cycle, which was noticed when sample collection was changed to an autosampling device (Cassar et al., 2008). This artifact apparently reduced  $A_{APO}/A_{ArN_2}$ , since it had a small to negligible impact on  $A_{APO}$ . While a full comparison of the two records is outside of the scope of this work, we note that the autosampler data from Macquarie Island (55.5°S; 158.95°E) and CGO between 2003 and 2010 produced nearly identical amplitude ratios to our PSA and CGO data (respectively) for the 2005–2020 period:  $3.6 \pm 0.7$  for Macquarie Island and  $4.2 \pm 0.5$  for CGO (from data courtesy of M. Bender, Princeton University).

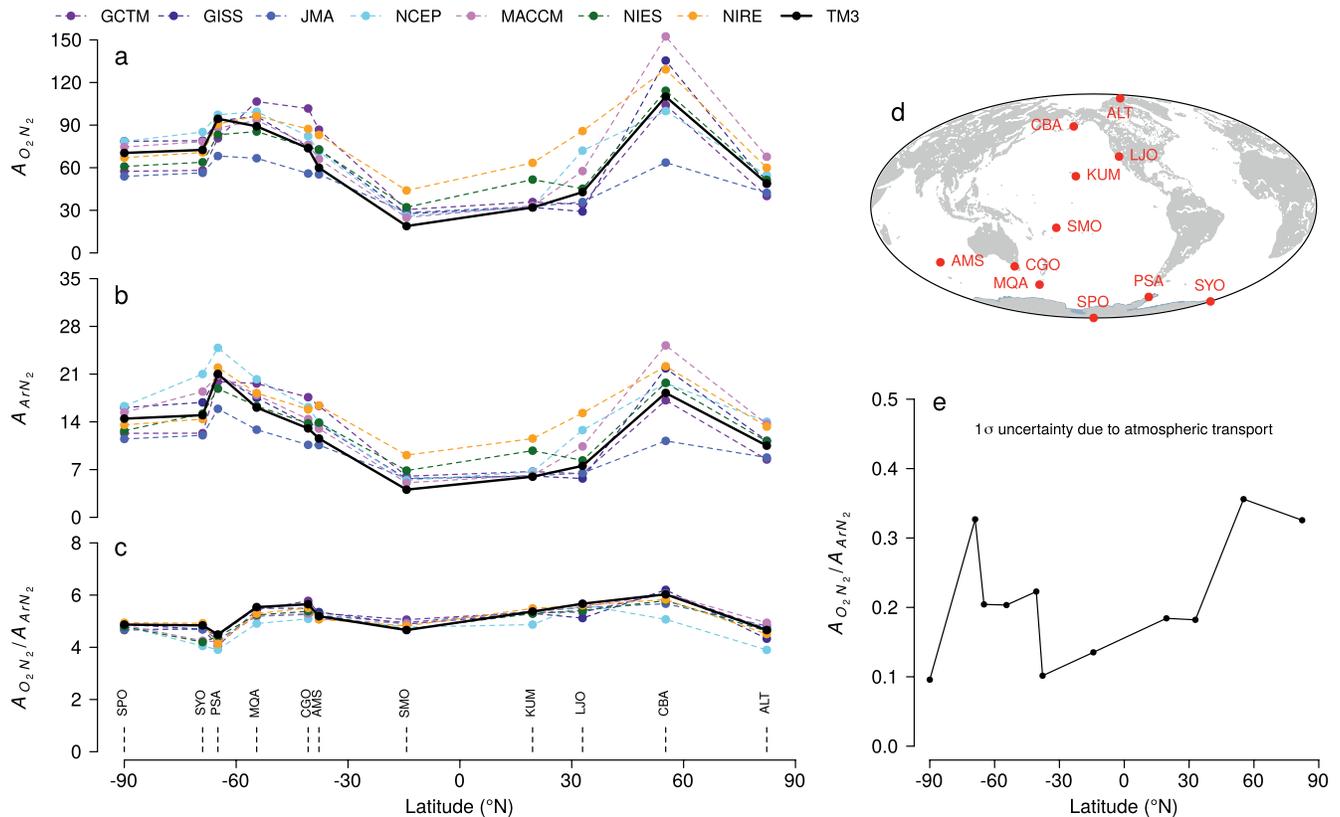
### 3.2. Insensitivity of Simulated $A_{APO}/A_{ArN_2}$ to Atmospheric Transport Modeling

As mentioned in the Introduction, we expect that model simulations of  $A_{APO}/A_{ArN_2}$  will be relatively insensitive to atmospheric circulation because of compensating influences on  $A_{APO}$  and  $A_{ArN_2}$ . We test this expectation using results from the APO TransCom study, which compared  $A_{O_2N_2}$ ,  $A_{ArN_2}$ , and  $A_{O_2N_2}/A_{ArN_2}$  from a suite of atmospheric transport models, all using the same air-sea fluxes of  $O_2$  and  $N_2$  (Blaine, 2005; Nevison, Baker, et al., 2012). For comparison purposes, we also transported these fluxes with TM3. Since the contribution of oceanic  $CO_2$  to APO is small, we consider  $A_{O_2N_2}$  to be sufficiently representative of  $A_{APO}$ . For reference, the contribution of oceanic  $CO_2$  to APO in the model ensemble used throughout this study (distinct from the APO TransCom study) was 13% on average.

As shown in Figure 2,  $A_{O_2N_2}$  and  $A_{ArN_2}$  vary typically by  $\pm 20\%$  ( $1\sigma$ ) between models, with the largest model-to-model differences at LJO and CBA. Models with high  $A_{O_2N_2}$  tend also to have high  $A_{ArN_2}$ , leading to variations in  $A_{O_2N_2}/A_{ArN_2}$  of  $\pm 4\%$  on average. The residual variability in  $A_{O_2N_2}/A_{ArN_2}$  across models results presumably because the spatiotemporal pattern of the  $O_2$  air-sea fluxes are not the same as the corresponding patterns for  $N_2$  and Ar, and in consequence these differences interact with differences in horizontal or vertical atmospheric transport to yield small model-to-model differences in  $A_{O_2N_2}/A_{ArN_2}$ . These results suggest that our selection of TM3 as a transport model contributes only around a  $\pm 4\%$  uncertainty to predictions of  $A_{APO}/A_{ArN_2}$ , making this a precise metric for challenging ocean models. Although this conclusion may be somewhat sensitive to the version of the  $N_2$  and  $O_2$  fluxes that were used by TransCom, we note that using fluxes that incorporated the revisions of Naegler (2009) and Jin et al. (2007), would have impacted all models nearly identically, and would not impact the level of model-to-model agreement.

### 3.3. Relationship Between Flux Ratios and Atmospheric Ratios in Model Simulations

The OBMs and the semi-empirical product predict sharp latitudinal gradients in the magnitude of the seasonal amplitude of both  $O_2$  and heat air-sea fluxes, with the largest seasonality seen between ca. 30° and 70° in both hemispheres (Figure 3). The OBM flux fields varied in their degree of sea ice masking and polar cutoffs, so larger divergence and disagreement near the poles is not surprising. While the models agree in their overall latitudinal distribution of the seasonal amplitude of  $O_2$  fluxes, significant differences can be found at the regional level (Figure S1). CESM, for instance, simulates stronger seasonality generally; it particularly predicts higher seasonality in the Labrador Sea, North Atlantic, and region east of the Drake Passage. MOM4, in contrast, predicts a large area of strong seasonality in the Atlantic sector of the Southern Ocean. The semi-empirical product, which is based on spatial interpolation, is nearly homogeneous within each zonal band. In comparison to  $O_2$  fluxes, OBM heat fluxes are more similar in their spatial patterns,

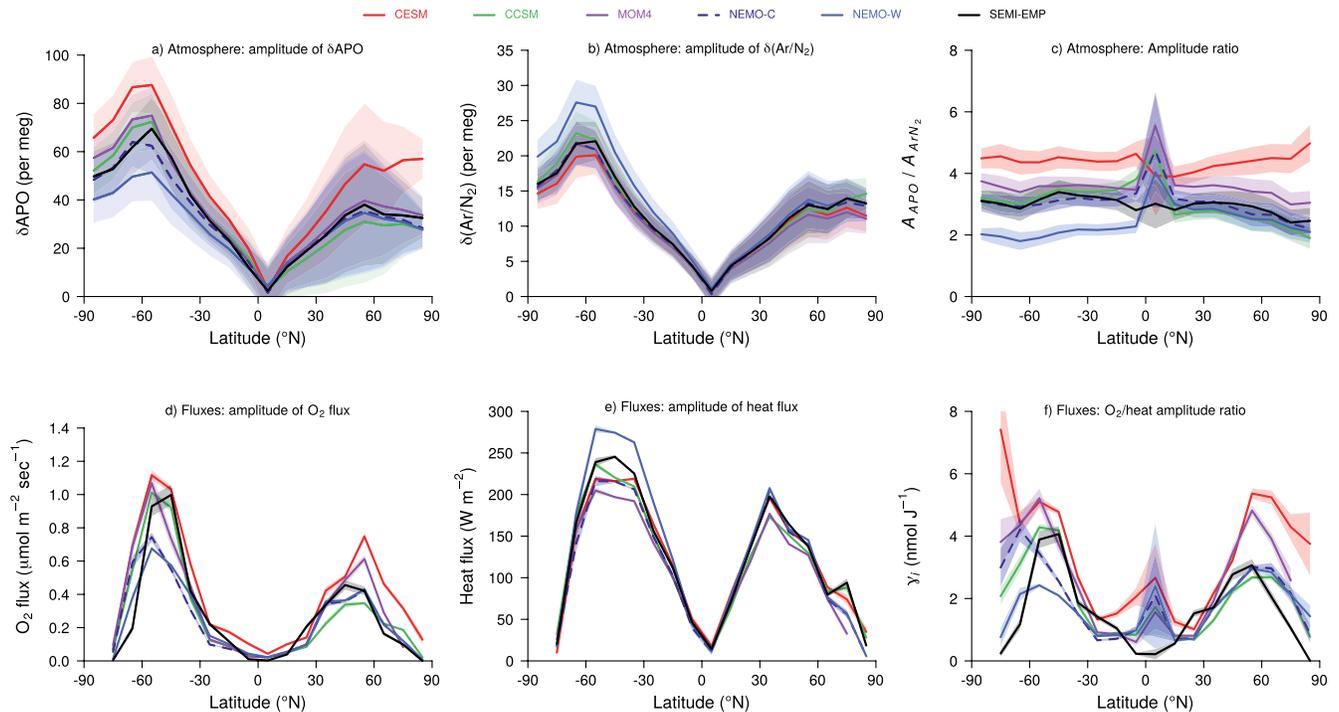


**Figure 2.** Amplitudes of  $\delta(\text{O}_2/\text{N}_2)$  (panel a;  $A_{\text{O}_2/\text{N}_2}$ ),  $\delta(\text{Ar}/\text{N}_2)$  (panel b;  $A_{\text{Ar}/\text{N}_2}$ ), and their ratio  $A_{\text{O}_2/\text{N}_2}/A_{\text{Ar}/\text{N}_2}$  (panel c) as a function of latitude from the atmospheric potential oxygen (APO) TransCom experiment. Each model transported the same  $\text{O}_2$  and  $\text{N}_2$  air-sea fluxes. Each point represents the output from a transport model at a surface station, colored by transport model. Sampling locations are given in panel (d), and are also indicated as a dashed vertical line with label in panel (c). The resulting  $1\sigma$  uncertainty in per meg/per meg due to atmospheric transport is given as a function of latitude in panel (e). The results indicate that while transport causes significant variability at sampling locations, comparing the ratio of two species largely cancels model differences. Note these fluxes are from a different model intercomparison than the one that forms the basis of this study, and as a result have not been corrected for the W92 gas exchange parameterization, and do not include  $\text{CO}_2$  (see text).

differing mostly in the extent of key hotspots at high latitudes (Figure S2). These latitudinal distributions in  $F_{\text{O}_2}$  and  $F_Q$  are imparted to the atmosphere in  $A_{\text{APO}}$  and  $A_{\text{Ar}/\text{N}_2}$ , albeit with a significant amount of smoothing due to atmospheric mixing (Figures S3 and S4).

Analogously to  $A_{\text{APO}}/A_{\text{Ar}/\text{N}_2}$ , we can compute the ratio of the amplitudes of the air-sea fluxes of  $\text{O}_2$  and heat,  $\gamma$  (see Equation 6). The use of a single metric to define an entire season (and a large spatial area) is supported by the strong relationship between  $F_{\text{O}_2}$  and  $F_Q$  in the OBMs and semi-empirical fluxes (Figures 4 and S5). The seasonal  $\text{O}_2$  to heat flux ratio, averaged in 10 degree bands ( $\gamma_i$ ), varies strongly with latitude.  $\gamma_i$  in all models is lower in the tropics and highest between ca.  $50^\circ$ – $60^\circ$ . In contrast to  $\gamma_i$ , however,  $A_{\text{APO}}/A_{\text{Ar}/\text{N}_2}$  is fairly constant within each hemisphere (Figures 3 and S6), due to atmospheric mixing and the dominant contribution to the atmospheric burden of both tracers by the regions with the highest air-sea fluxes.

The model ensemble can also inform on a key point, which is to identify the range of latitudes in each hemisphere which controls the fairly constant value of  $A_{\text{APO}}/A_{\text{Ar}/\text{N}_2}$  observed at most stations. To address this question we computed values of  $\gamma$  for a range of latitudes, and found that the zonal band between  $40^\circ$  and  $70^\circ$  ( $\gamma_{40-70}$ ) in each hemisphere demonstrated a strong correlation with  $A_{\text{APO}}/A_{\text{Ar}/\text{N}_2}$ , both for the hemispheric averages and individual stations. This latitude range is also the general bounds of the zonal bands where fluxes (and the seasonality) of  $\text{O}_2$  and heat fluxes are the strongest. The relationship is shown in Figure 5. Models with higher  $A_{\text{APO}}/A_{\text{Ar}/\text{N}_2}$  ratios have higher  $\gamma_{40-70}$ , with the results clustering along a line that roughly runs through the origin. Depending on the stations, these lines have slopes of  $\sim 0.6$ – $1 \text{ J}^\circ\text{nmol}^{-1}$ .



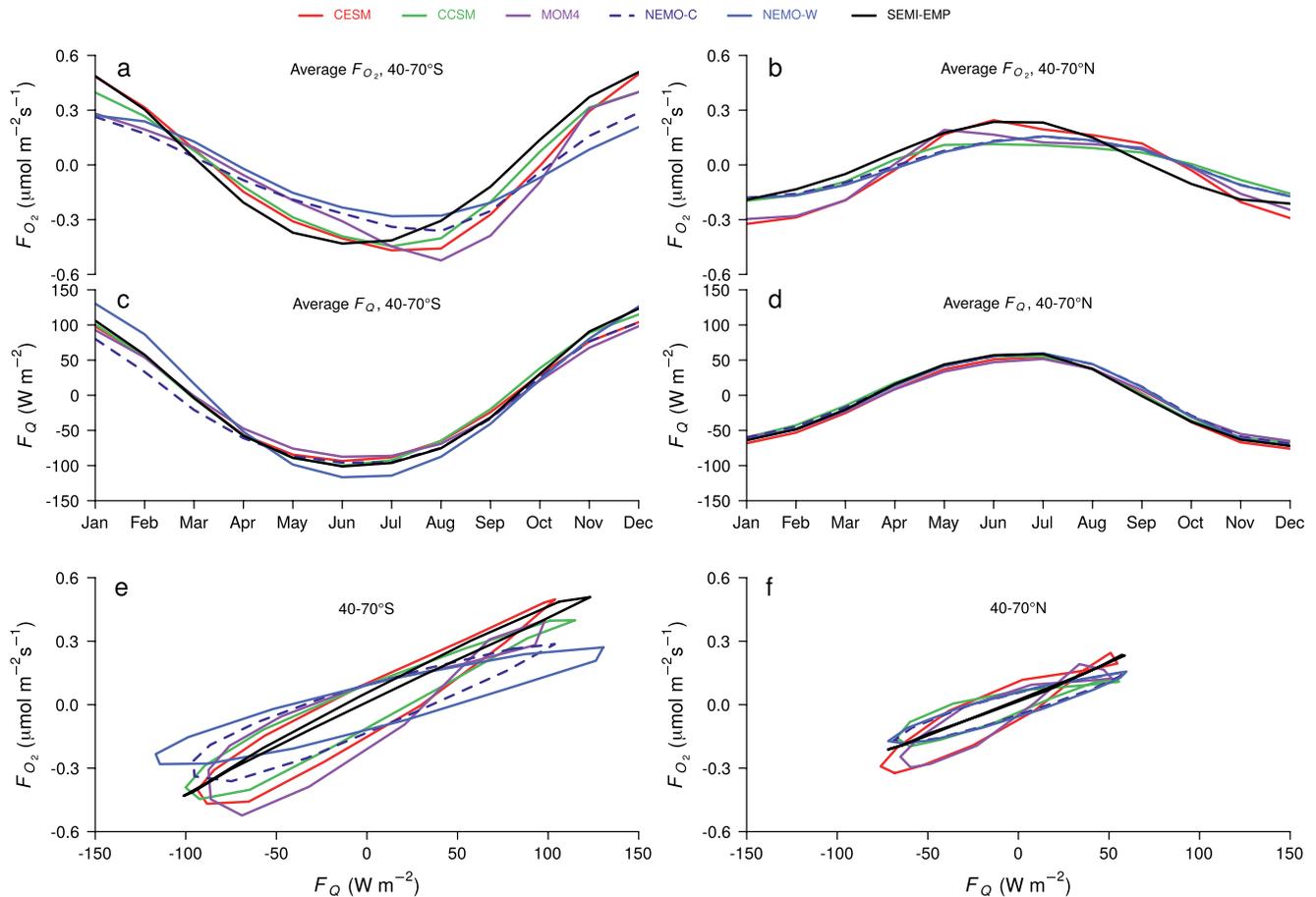
**Figure 3.** Comparisons of modeled and semi-empirical seasonal amplitudes of  $\delta$ APO (atmospheric potential oxygen) (a),  $\delta(\text{Ar}/\text{N}_2)$  (b), and  $A_{\text{APO}}/A_{\text{ArN}_2}$  (c),  $\text{O}_2$  air-sea fluxes (d), air-sea net heat fluxes (e), and the ratio of the two,  $\gamma_i$  (f). Data are all binned by  $10^\circ$  of latitude and are averaged over the zonal band.  $1\sigma$  uncertainties are shown as semi-transparent bands or error bars. For the atmospheric tracers this includes uncertainty in atmospheric transport.

### 3.4. Implied $\gamma_{40-70}$ From Observed $A_{\text{APO}}/A_{\text{ArN}_2}$

The robust linear relationship between  $\gamma_{40-70}$  and  $A_{\text{APO}}/A_{\text{ArN}_2}$  in the OBMs (Figure 5) affords the opportunity to calculate the value of  $\gamma_{40-70}$  implied by observed  $A_{\text{APO}}/A_{\text{ArN}_2}$ . Solving the regression equations for each station yields implied  $\gamma_{40-70}$  values of 4.0–5.4  $\text{nmol J}^{-1}$  in the Southern Hemisphere and 2.9–3.5  $\text{nmol J}^{-1}$  in the Northern Hemisphere (Table 1). The reported uncertainty in these estimates accounts for the impact of atmospheric transport (see Section 3.2) as well as plausible variations in heat and  $\text{O}_2$  flux patterns, using the spread across models.

Averaging the stations within each hemisphere, we obtain an average  $\gamma_{40-70}$  of  $4.7 \pm 0.8 \text{ nmol J}^{-1}$  for the Southern Hemisphere and  $3.2 \pm 0.5 \text{ nmol J}^{-1}$  for the Northern Hemisphere. In reporting the Northern Hemisphere average in Table 1, we chose to omit KUM since it is significantly closer to the tropics. This choice makes only a small difference: excluding KUM yields  $3.3 \pm 0.4 \text{ nmol J}^{-1}$  for the Northern Hemisphere.

Since the relationship employed here is that of a tracer/tracer to a flux/flux, it does not require that the models accurately simulate the large number of physical and biogeochemical drivers which control the distributions of these fluxes and tracers in actuality, only that the temporal and spatial correlation between  $\text{O}_2$ /heat and Ar/heat is as strong in nature as it is in the models. Small scatter can be seen in the relationship between  $\gamma_{40-70}$  and  $A_{\text{APO}}/A_{\text{ArN}_2}$ . This scatter is stronger at CBA, CGO, and PSA than at KUM and LJO (Figure 5). Some of the scatter is due to small local or regional variations in  $A_{\text{APO}}/A_{\text{ArN}_2}$ , which is not apparent when the entire zonal band between 40 and  $70^\circ$  is used to calculate  $A_{\text{APO}}/A_{\text{ArN}_2}$  (data not shown). That is, if a station is close to regions with large seasonal  $\text{O}_2$  and heat fluxes which have a different  $\text{O}_2$ /heat ratio than that of the integrated 40– $70^\circ$  band, the local influence can have a small effect on  $A_{\text{APO}}/A_{\text{ArN}_2}$ . To capture the hemispheric scale ratio, the ideal station is one that is far from regions of strong localized fluxes. The remaining scatter appears to be due to using APO as a proxy for  $\text{O}_2$ , since the former includes small contributions from  $\text{CO}_2$  and  $\text{N}_2$ . We also note that the semi-empirical fluxes yield values of  $\gamma_{40-70}$  and  $A_{\text{APO}}/A_{\text{ArN}_2}$  which are in good agreement with the linear relationship suggested by the model ensemble. We interpret this as being supportive of our approach, since the semi-empirical fluxes are observationally based and independent of the OBMs.

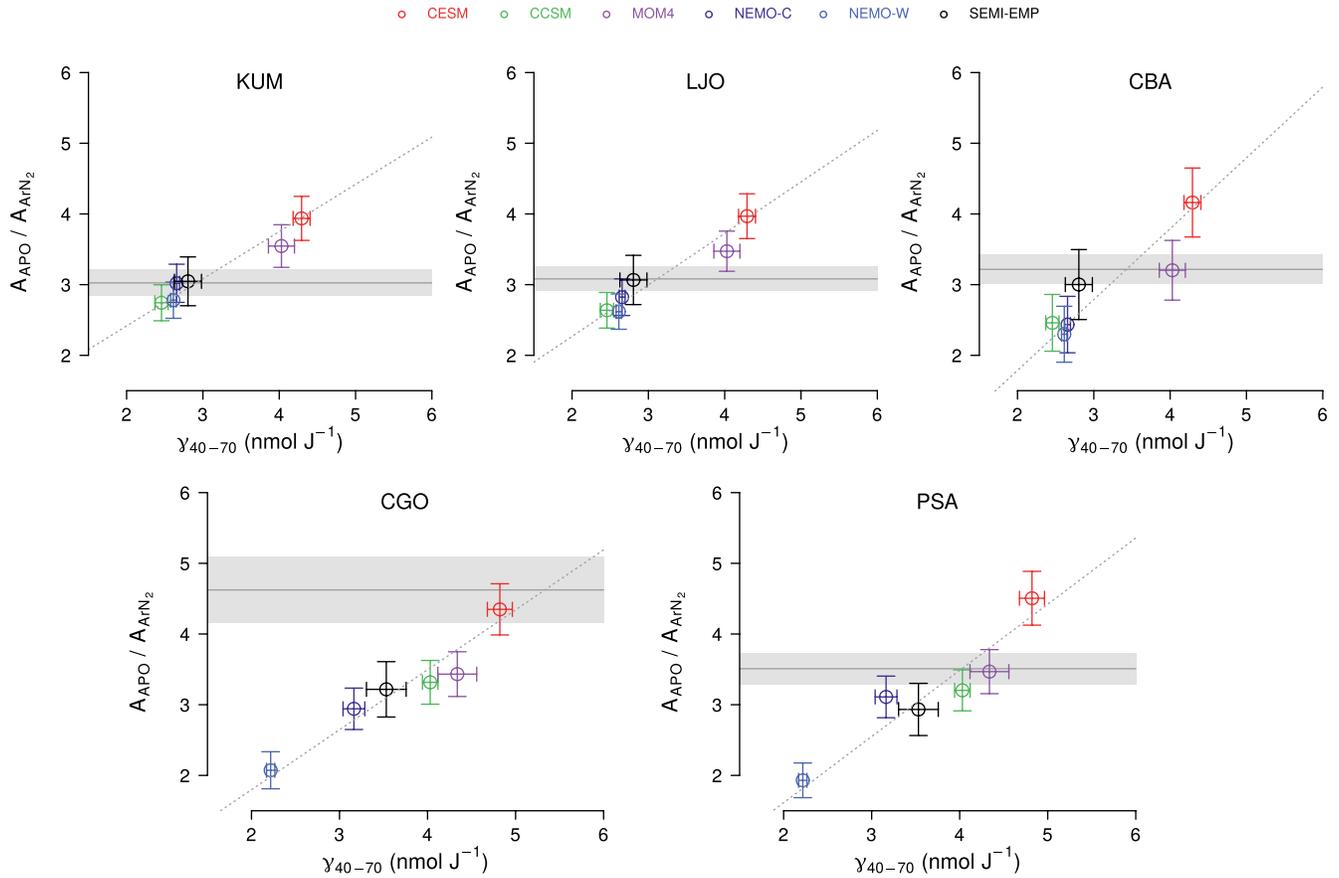


**Figure 4.** Average seasonal cycles of simulated air-sea fluxes of  $O_2$  ( $F_{O_2}$ ) in the extratropics of each hemisphere (a) and (b) and heat ( $F_Q$ ; panels (c–d) from five ocean models and the semi-empirical flux product. Fluxes were averaged over  $40\text{--}70^\circ$  in each hemisphere at a monthly time step. In the panels (e) and (f),  $F_{O_2}$  is shown as a function of  $F_Q$  in each hemisphere.

### 3.4.1. Caveats and Potential Biases in Implied $\gamma_{40-70}$

We present constraints on  $\gamma_{40-70}$  with several caveats. One important consideration is the extent to which the SIO flask sampling locations are zonally representative, since this determines how well  $A_{APO}/A_{ArN_2}$  at these sites serves to constrain the average seasonal flux of the hemisphere. Since our sites are mostly located in the Pacific, they might only serve to represent the flux signal from the Pacific sector of each hemisphere. To address this question we performed a sensitivity test by running the fluxes from the Atlantic sector only ( $70^\circ\text{W}$  to  $20^\circ\text{E}$ , pole to pole) in the semi-empirical flux product forward in TM3 (all other fluxes were masked). At the Northern Hemisphere stations, the Atlantic contribution to  $A_{APO}$  and  $A_{ArN_2}$  is between 23% and 29% at both LJO and KUM, but only 14%–15% at CBA. As the Atlantic sector contributes  $50 \pm 5\%$  of the total hemispheric flux, our results from the Northern Hemisphere stations have a clear Pacific bias, giving 2 to 3 times more weight to fluxes from the Pacific than the Atlantic sector of the globe. At CGO and PSA in the Southern Hemisphere, the Atlantic fluxes contribute 22%–30% to  $A_{APO}$  and  $A_{ArN_2}$  while contributing  $23 \pm 3\%$  of the global flux each month. The results from the Southern Hemisphere stations thus yield estimates of  $A_{APO}/A_{ArN_2}$  and  $\gamma_{40-70}$  that are more uniformly representative of the full zone.

Our constraint is sensitive to the Jin et al. (2007) correction factor. A change of 0.08 (the  $1\sigma$  uncertainty of the factor) would lead to a  $\sim 0.4 \text{ nmol J}^{-1}$  change in  $\gamma_{40-70}$ . Having no basis upon which to reevaluate the correction factor, we have kept it in place with its uncertainty.



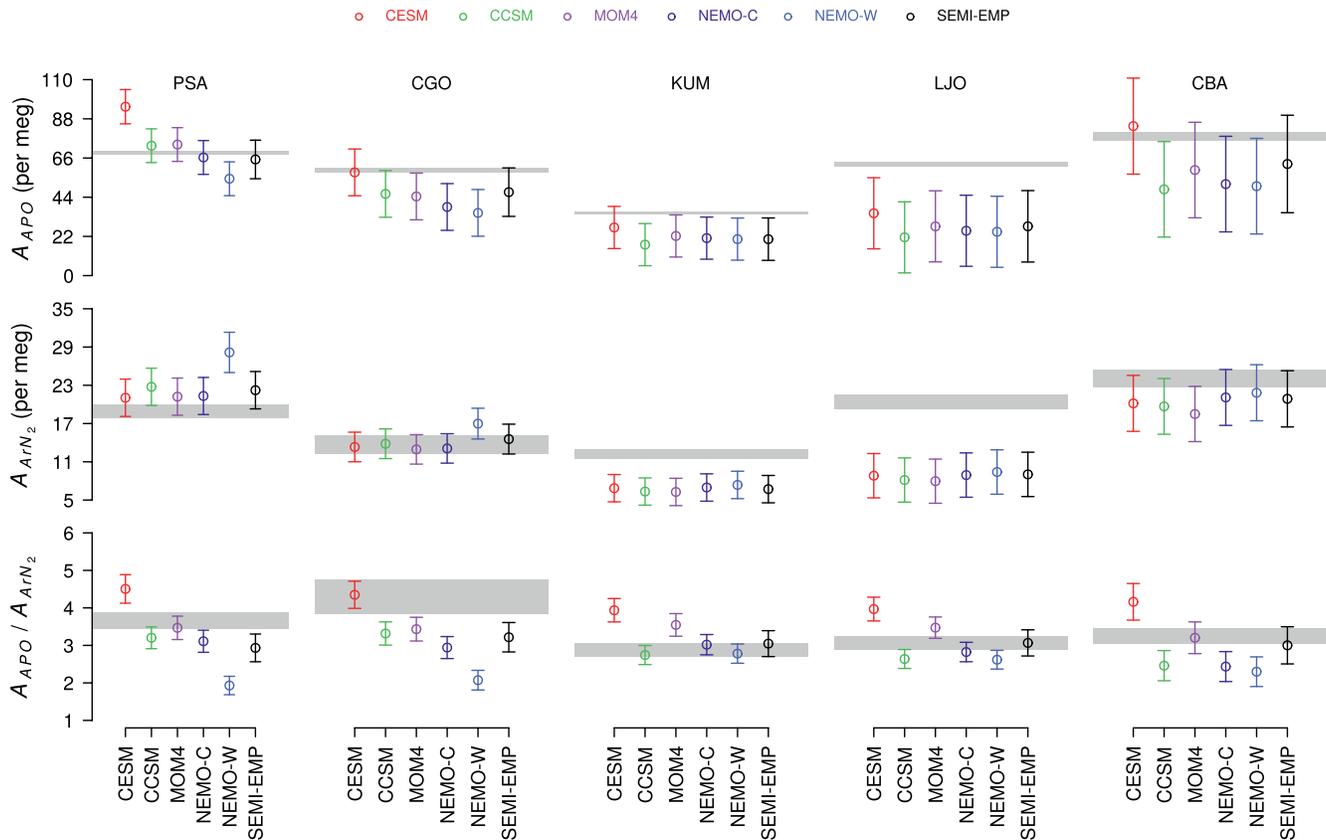
**Figure 5.** Relationship between the modeled  $O_2$ /heat flux ratio ( $\gamma_{40-70}$ ) and the ratio of the seasonal amplitudes of  $\delta APO$  (atmospheric potential oxygen) and  $\delta(Ar/N_2)$  ( $A_{APO}/A_{ArN_2}$ ) for five selected surface stations. Also depicted is the semi-empirical run. Each point represents the climatological average ratio of the seasonal amplitude of the fluxes or atmospheric tracer. The average value of  $A_{APO}/A_{ArN_2}$  at each surface station is plotted against the value of  $\gamma_{40-70}$  from its respective hemisphere. Horizontal gray bars show the  $1\sigma$  uncertainty of the observed value of  $A_{APO}/A_{ArN_2}$  for each station, with the value as the black horizontal line. A linear regression is shown for each individual station as a dashed line. Error bars represent  $1\sigma$  uncertainties.

Air-sea exchange of heat and inert gases can be decoupled by bubble effects (Emerson & Bushinsky, 2016), surface effects limiting the air-sea exchange of gases, such as surfactants or skin temperatures (e.g., Pereira et al., 2018), mixing effects, which, through the nonlinearity of gas solubility, can create oversaturation (Dietze & Oschlies, 2005), radiation penetrating below the mixed layer (e.g., Spitzer & Jenkins, 1989), and the formation of sea ice, which could cause a transfer of heat without a concordant exchange of gas. Our approach is only sensitive to those processes which would have an appreciable impact of the seasonal exchange of inert gases, integrated over the 40–70° latitude band in each hemisphere. A modeling experiment in which Ar and  $N_2$  were carried as explicit tracers would avoid such issues and likely improve on the constraint we offer by reducing these sources of uncertainty.

### 3.5. Evaluation of Model Performance

A comparison of the atmospheric observations to the TM3 forward run output for the OBMs at each of the five surface stations is shown in Figure 6. Generally the models tend to overpredict  $A_{ArN_2}$  in the Southern Hemisphere and underpredict in the Northern Hemisphere, with more spread in the Southern Hemisphere than in the Northern.

The model predictions of  $A_{APO}$  are more varied. The ensemble tends to underpredict in both hemispheres, but does bracket the observations at PSA, and to a lesser extent at CGO and CBA. CESM is an outlier in that it typically predicts larger  $A_{APO}$  than other models in both hemispheres. CESM performs better at CGO than



**Figure 6.** Comparisons of modeled  $A_{APO}$ ,  $A_{ArN_2}$ , and  $A_{APO}/A_{ArN_2}$  across models and the semi-empirical forward run. Each point represents the average value of the seasonal amplitude or amplitude ratio at a surface station, given as the title of each column. The observed value of the amplitude or amplitude ratio is shown as a gray bar, the boundaries of which denote the average  $\pm 1\sigma$ . Error bars contain all sources of uncertainty discussed in the text (expressed as  $1\sigma$ , including uncertainty due to atmospheric transport).

at PSA, implying that it overpredicts the amplitude ratio due to a too-vigorous seasonal exchange of APO at high latitudes in the Southern Ocean.

The models generally perform the poorest in representing  $A_{APO}$  at LJO and KUM, where predicted  $A_{APO}$  is too low by 20–30 per meg. At these two sites the spread between models is also lower. A similar difficulty in reproducing  $\delta APO$  amplitudes at LJO in TM3 using semi-empirical fluxes was noted in earlier work (Garcia & Keeling, 2001; Keeling, Stephens, et al., 1998). This difficulty is perhaps not unexpected given the relatively coarse resolution of the  $4^\circ \times 5^\circ$  grid with which TM3 was run. Contributing to the error may also be an inability to accurately simulate the finer-scale aspects of atmospheric circulation along the North American west coast, such as the along-shore jet, which would typically bring air from higher latitudes to LJO.

The models generally underestimate  $A_{APO}/A_{ArN_2}$  at the Southern Hemisphere sites but center on the observations in the Northern Hemisphere. Each model simulates fairly consistent  $A_{APO}/A_{ArN_2}$  within each hemisphere, despite the strong variations of  $A_{APO}$  and  $A_{ArN_2}$  with latitude. This consistency is seen even at LJO and KUM, where the models underpredict  $A_{APO}$  and  $A_{ArN_2}$ , but produce ratios which are close to the observations, as at the other Northern Hemisphere sites. The systematic bias at LJO and KUM demonstrates an advantage of using  $A_{APO}/A_{ArN_2}$  as an ocean model benchmark, since the underestimation of each individual quantity largely cancels out in the ratio of the two, eliminating bias from what is likely a transport error. That is, the model skill in representing  $O_2$  and heat fluxes is likely greater than it appears if we evaluated  $A_{APO}$  or  $A_{ArN_2}$  alone.

Further inferences can be made in regards to model skill in simulating  $A_{APO}/A_{ArN_2}$ . For instance, at CBA, the agreement is generally good across models for  $A_{ArN_2}$  but poor for  $A_{APO}$  (with the exception of MOM4),

resulting in poorer performance in simulating  $A_{APO}/A_{ArN_2}$ . MOM4 has a higher seasonality of  $F_{O_2}$  than NEMO-C, NEMO-W, and CCSM in these regions, but lower than CESM. This suggests that these models could be accurately representing air-sea heat fluxes, but poorly simulating biological processes in some regions of the Northern Hemisphere high latitudes.

### 3.5.1. A Case Study: Sensitivity of Amplitude Ratios to Mixed Layer Depths in the Southern Ocean

As a demonstration of how the  $A_{APO}/A_{ArN_2}$  metric can be applied to evaluate model simulations, we compare the performance of NEMO-C and NEMO-W. NEMO-W is a sensitivity perturbation of NEMO-C, in which the turbulence kinetic energy (TKE) parameterization of NEMO-C was given an ad hoc parameterization to bring the summer mixed layer of the Southern Ocean into better agreement with observations (Rodgers et al., 2014). NEMO-W yields deeper mixed layers in both winter and summer compared to NEMO-C (Rodgers et al., 2014), and while the summer mixed layers are in better agreement with observations in the Southern Ocean, winter mixed layers are too deep (Rodgers et al., 2014).

The NEMO-W wind stirring parameterization reduces the seasonal exchange of  $O_2$  in comparison to NEMO-C, as indicated by  $A_{APO}$  at PSA and CGO (Figure 6). Seasonal heat storage and loss, however, is increased, as indicated by the larger values of  $A_{APO}$  at the same sites. The effects compound in the amplitude ratio, as evident in the significantly lower values of  $A_{APO}/A_{ArN_2}$  in NEMO-W. The NEMO-W simulation also yields a lower  $A_{APO}/A_{ArN_2}$  in the Southern Hemisphere than in the Northern. NEMO-C already predicts a too-low  $O_2$ /heat ratio generally, relative to the observational constraint, and tuning Southern Ocean mixed layers to provide a better physical state only increases the discrepancy.

### 3.6. Implications for Hemispheric Asymmetry and Carbon

The atmospheric observations point to a higher value of  $\gamma_{40-70}$  in the Southern Hemisphere than in the Northern. This hemispheric asymmetry is also apparent in the semi-empirical flux product (Garcia & Keeling, 2001). We note that hydrographic observations (from the surface through the thermocline) are supportive of a hemispheric difference, since the North Atlantic Ocean has a distinctly lower  $O_2$  to heat relationship than much of the rest of the global ocean, as assessed through the tracer  $O_2^*$  (Keeling & Garcia, 2002). The Atlantic makes a large contribution to the Northern Hemisphere  $\gamma_{40-70}$ , and is likely a major factor in the observed interhemispheric difference (see Section 3.4.1). While our study does not inform on the causes of an interhemispheric asymmetry in  $\gamma_{40-70}$ , we note that such a hemispheric difference runs counter to what might have been expected from first principles on the basis of iron limitation in the SH (Cassar et al., 2007; Moore & Braucher, 2008), which is expected to limit NCP and hence  $O_2$  outgassing in austral summer. The observed interhemispheric difference is also surprising when considering the results of the NEMO-C and NEMO-W comparison, which indicated that deeper mixed layers would yield lower  $\gamma_{40-70}$ , not higher (as discussed above). Ventilation of deeper water masses with low  $O_2$  can contribute significantly to the winter uptake of  $O_2$  by the surface ocean, and we speculate that winter ventilation may be more significant to  $\gamma_{40-70}$  than limitations on summertime biological fluxes of  $O_2$ .

CCSM and NEMO-C simulate the hemispheric asymmetry in the  $O_2$ /heat ratio over the key latitudes of  $40^\circ$ - $70^\circ$  with some skill, while the rest of the ensemble predicts a greater degree of parity (Figure 3c). Such model-data discrepancies in the  $O_2$ /heat ratio could have implications for carbon. As shown by Keeling et al. (1993), the  $O_2$ /heat flux ratio is mechanistically linked to the seasonal correlation between DIC and temperature in surface waters, which in turn influences the seasonal cycle in  $pCO_2$ . If DIC concentrations were constant across seasons,  $pCO_2$  would tend to be higher in summer than winter, due to the temperature dependence of  $CO_2$  solubility. At lower latitudes, this solubility forcing dominates the changes in DIC (Weiss et al., 1982). At higher latitudes, the solubility cycle is strongly countered by the photosynthesis/ventilation cycle which causes surface-water DIC to be higher in winter than summer, just as it causes  $O_2$  outgassing in summer and ingassing in winter (Takahashi et al., 2002). A model that overestimates the  $O_2$ /heat ratio will therefore also likely overestimate the surface water excess of DIC in winter compared to summer, and thus misrepresent the seasonal cycle in  $pCO_2$ . It has also been shown that ocean models that have too strong a biological carbon pump in the Southern Ocean (Nevison et al., 2016) tend to overestimate the seasonal amplitude of  $\delta APO$ , and too large a  $CO_2$  sink ( $> 1 \text{ PgC yr}^{-1}$ ) in the Southern Ocean when

compared to present-day estimates (Takahashi et al., 2009). An important follow-on study would be to reconcile model/data discrepancies in both  $O_2$ /heat and  $pCO_2$ .

### 3.7. Evaluation of Semi-Empirical Fluxes

In Figure 6, we also show the results of the TM3 forward runs of the semi-empirical fluxes. The semi-empirical APO flux product is a combination of three separate datasets: the Garcia and Keeling (2001)  $O_2$  flux climatology, the Takahashi et al. (2009)  $CO_2$  flux climatology, and the ECCOV4 heat flux climatology. Simulated  $A_{APO}$  agrees quite well with observations at the two Southern Hemisphere sites (PSA and CGO), and at CBA, the most poleward site in the Northern Hemisphere. The values of  $A_{ArN_2}$  agree well with observed  $A_{ArN_2}$  at CBA, but less well at PSA and CGO, indicating either poor representation of air-sea heat fluxes in Southern Hemisphere extratropical latitudes, or that the Jin et al. (2007) correction is less suited to this region. Equation 4 is an approximation which neglects important processes that contribute to abiotic air-sea gas exchange (Dietze & Oschlies, 2005), and the Jin et al. (2007) correction was implemented simply to bring heat-derived fluxes more generally into better agreement with modeled solubility fluxes. Accordingly, we would not expect it to be equally effective everywhere.

When evaluated on its performance in simulating  $A_{APO}/A_{ArN_2}$ , we find that the semi-empirical flux product is in excellent agreement with the observations in the Northern Hemisphere but in poorer agreement in the Southern Hemisphere, particularly at CGO. At this site predicted  $A_{APO}$  is too low by  $\sim 10$  per meg, and the peak of the annual harmonic is 2 weeks too early. Given that the semi-empirical flux product is in better agreement with respect to  $A_{ArN_2}$  than to  $A_{APO}$ , and also that it is in better agreement in the amplitude ratio at PSA than at CGO, the results suggest that the Garcia and Keeling (2001) air-sea fluxes of  $O_2$  may be particularly deficient in the lower latitudes of the Southern Ocean.

Since the semi-empirical product allows for the computation of  $\gamma_{40-70}$  directly, we can compare these to the values of  $\gamma_{40-70}$  implied from the observations (as discussed in Section 3.3). The direct ratio of the semi-empirical product gives  $3.5 \pm 0.2$  and  $2.8 \pm 0.2$   $nmol J^{-1}$  in the Southern and Northern Hemispheres, respectively. These values are significantly lower (particularly in the Southern Hemisphere) than the corresponding atmospheric-based estimates of  $4.7 \pm 0.8$   $nmol J^{-1}$  and  $3.3 \pm 0.4$   $nmol J^{-1}$  (Table 1). While the semi-empirical approach has the advantage of resolving the  $O_2$ /heat flux ratio with finer north-south resolution ( $\pm 10^\circ$  latitude), the atmospheric-based metric is likely more reliable for large-scale averages because it is less dependent on uncertain parameterizations and sparse coverage.

## 4. Summary and Conclusions

We have developed two empirical metrics for evaluating ocean biogeochemical models, one based on the ratio of the amplitude of the seasonal cycles of  $\delta APO$  and  $\delta(Ar/N_2)$  (i.e.,  $A_{APO}/A_{ArN_2}$ ) and the other based on zonally averaged ratios of the seasonal cycles in  $O_2$  flux and heat flux (i.e.,  $\gamma_{40-70}$ ). Both metrics serve to test the strength of the seasonal cycle in the air-sea flux of  $O_2$  relative to heat, but with different scales of spatial representation. The  $A_{APO}/A_{ArN_2}$  metric independently constrains  $\gamma_{40-70}$  within the Northern and Southern Hemispheres separately, but has only limited ability to resolve patterns at a finer scale because of the relatively rapid atmospheric mixing within each hemisphere.

Building on the study of Garcia and Keeling (2001), who found a strong relationship between seasonal extratropical air-sea  $O_2$  and heat fluxes, we have presented an observationally based constraint on the seasonal hemispheric  $O_2$ /heat flux ratio ( $\gamma_{40-70}$ ). This constraint is determined from the amplitude ratio ( $A_{APO}/A_{ArN_2}$ ) of the seasonal cycles in  $\delta APO$  and  $\delta(Ar/N_2)$ . This metric is largely insensitive to uncertainties in atmospheric transport and mixing, owing to compensating influences on  $A_{APO}$  and  $A_{ArN_2}$ , making it a potentially robust measure of the average ratio in each hemisphere. The atmospheric data yield  $\gamma_{40-70}$  of  $4.7 \pm 0.8$   $nmol J^{-1}$  for the SH and  $3.3 \pm 0.3$   $nmol J^{-1}$  for the NH. When compared to predictions from a suite of general ocean circulation models, we find that the models generally underestimate  $A_{APO}/A_{ArN_2}$  and hence  $\gamma_{40-70}$ , particularly in the Southern Hemisphere. Our work advances the application of  $\delta(O_2/N_2)$  and  $\delta(Ar/N_2)$  measurements for evaluating ocean models, by testing the relative strength of the seasonal cycles in air-sea fluxes of carbon, oxygen, and heat at extratropical latitudes.

## Data Availability Statement

SIO O<sub>2</sub> Program data is available from <https://scripps2.ucsd.edu/>. Air-sea flux data from the model ensemble is available from <https://doi.org/10.5281/zenodo.4716840>.

## Acknowledgments

The recent atmospheric measurements of the Scripps O<sub>2</sub> program have been supported via funding from the NSF and the National Oceanographic and Atmospheric Administration (NOAA) under grants 1304270 and OAR-CIPO-2015-2004269. M. Manizza and R. F. Keeling thank NSF for financial support via the OCE-1130976 grant. M. Manizza thanks additional financial support from NSF via the ARRA OCE-0850350 grant. S. C. Doney acknowledges support from NSF PLR-1440435. Keith Rodgers acknowledges support from IBS-R028-D1. Gael Forget and the ECCO group kindly provided the ECCOV4 heat fluxes. The authors thank group members Stephen Walker, Sara Afshar, Shane Clark, Bill Paplawsky for support in the flask analyses and data workup. They thank staff from the U.S. Antarctic program for flask collections at Palmer Station, staff from the U.S. Weather Service for flask collections at Cold Bay, Alaska, and staff from the National Oceanographic and Atmospheric Administration for flask collections in Hawaii and Samoa. Comments from Graeme MacGilchrist helped to improve the manuscript. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the author(s) and do not necessarily reflect the views of the supporting agencies.

## References

- Adachi, Y., Kawamura, L., & Keeling, R. (2006). Diffusive separation of the lower atmosphere. *Science*, *311*(5766), 1429. <https://doi.org/10.1126/science.1121312>
- Aumont, O., & Bopp, L. (2006). Globalizing results from ocean in situ iron fertilization studies. *Global Biogeochemical Cycles*, *20*(2), GB2017. <https://doi.org/10.1029/2005GB002591>
- Battle, M., Bender, M., Hendricks, M., Ho, D., Mika, R., McKinley, G., et al. (2003). Measurements and models of the atmospheric Ar/N<sub>2</sub> ratio. *Geophysical Research Letters*, *30*(15), 1786. <https://doi.org/10.1029/2003GL017411>
- Battle, M., Fletcher, S., Bender, M., Keeling, R., Manning, A., Gruber, N., et al. (2006). Atmospheric potential oxygen: New observations and their implications for some atmospheric and oceanic models. *Global Biogeochemical Cycles*, *20*(1), GB1010. <https://doi.org/10.1029/2005GB002534>
- Bent, J. (2014). *Airborne oxygen measurements over the Southern Ocean as an integrated constraint of seasonal biogeochemical processes*. San Diego: University of California.
- Blaine, T. (2005). *Continuous measurements of atmospheric Ar/N<sub>2</sub> as a tracer of air-sea heat flux: Models, methods, and data (unpublished doctoral dissertation)*. San Diego: University of California.
- Bopp, L., Resplandy, L., Orr, J. C., Doney, S. C., Dunne, J. P., Gehlen, M., et al. (2013). Multiple stressors of ocean ecosystems in the 21st century: Projections with CMIP5 models. *Biogeosciences*, *10*(10), 6225–6245. <https://doi.org/10.5194/bg-10-6225-2013>
- Brodeau, L., Barnier, B., Treguier, A.-M., Penduff, T., & Gulev, S. (2010). An ERA40-based atmospheric forcing for global ocean circulation models. *Ocean Modelling*, *31*(3), 88–104. <https://doi.org/10.1016/j.ocemod.2009.10.005>
- Cassar, N., Bender, M., Barnett, B., Fan, S., Moxim, W., Levy, H., & Tilbrook, B. (2007). The Southern Ocean biological response to Aeolian iron deposition. *Science*, *317*(5841), 1067–1070. <https://doi.org/10.1126/science.1144602>
- Cassar, N., McKinley, G., Bender, M., Mika, R., & Battle, M. (2008). An improved comparison of atmospheric Ar/N<sub>2</sub> time series and paired ocean-atmosphere model predictions. *Journal of Geophysical Research*, *113*, D21122. <https://doi.org/10.1029/2008JD009817>
- Collins, W., Bitz, C., Blackmon, M., Bonan, G., Bretherton, C., Carton, J., et al. (2006). The community climate system model version 3 (CCSM3). *Journal of Climate*, *19*(11), 2122–2143. <https://doi.org/10.1175/JCLI3761.1>
- Dietze, H., & Oschlies, A. (2005). On the correlation between air-sea heat flux and abiotically induced oxygen gas exchange in a circulation model of the North Atlantic. *Journal of Geophysical Research*, *110*, C09016. <https://doi.org/10.1029/2004JC002453>
- Doney, S. (2010). The growing human footprint on coastal and open-ocean biogeochemistry. *Science*, *328*(5985), 1512–1516. <https://doi.org/10.1126/science.1185198>
- Dunne, J., Gnanadesikan, A., Sarmiento, J., Slater, R., & Hiscock, M. R. (2010). Supplement: Technical description of the prototype version (v0) of Tracers Of Phytoplankton with Allometric Zooplankton (TOPAZ) ocean biogeochemical model as used in the Princeton IFMIP\* model. *Biogeosciences*, *7*, 3593–3624. <https://doi.org/10.5194/bg-7-3593-2010>
- Dunne, J., John, J., Shevliakova, E., Stouffer, R., Krasting, J., Malyshev, S., et al. (2013). GFDL's ESM2 global coupled climate-carbon earth system models. Part II: Carbon system formulation and baseline simulation characteristics. *Journal of Climate*, *26*(7), 2247–2267. <https://doi.org/10.1175/JCLI-D-12-00150.1>
- Eddebbbar, Y., Long, M. C., Resplandy, L., Rödenbeck, C., Rodgers, K., Manizza, M., & Keeling, R. (2017). Impacts of ENSO on air-sea oxygen exchange: Observations and mechanisms. *Global Biogeochemical Cycles*, *31*, 901–921. <https://doi.org/10.1002/2017GB005630>
- Emerson, S. (1987). Seasonal oxygen cycles and biological new production in surface waters of the subarctic Pacific Ocean. *Journal of Geophysical Research*, *92*, 6535–6544. <https://doi.org/10.1029/JC092iC06p06535>
- Emerson, S., & Bushinsky, S. (2016). The role of bubbles during air-sea gas exchange. *Journal of Geophysical Research: Oceans*, *121*(6), 4360–4376. <https://doi.org/10.1002/2016JC011744>
- Forget, G., Campin, J.-M., Heimbach, P., Hill, C. N., Ponte, R. M., & Wunsch, C. (2015). ECCO version 4: An integrated framework for non-linear inverse modeling and global ocean state estimation. *Geoscientific Model Development*, *8*(10), 3071–3104. <https://doi.org/10.5194/gmd-8-3071-2015>
- Garcia, H., & Keeling, R. (2001). On the global oxygen anomaly and air-sea flux. *Journal of Geophysical Research*, *106*(C12), 31155–31166. <https://doi.org/10.1029/1999JC000200>
- Gibson, J., Kållberg, P., Uppala, S., Hernandez, A., Nomura, A., & Serrano, E. (1999). *ECMWF Re-analysis report series: ERA-15 description, version 2*. European Centre for Medium-Range Weather Forecasts.
- Griffies, S., Gnanadesikan, A., Dixon, K., Dunne, J., Gerdes, R., Harrison, M., et al. (2005). Formulation of an ocean model for global climate simulations. *Ocean Science*, *1*(1), 45–79. <https://doi.org/10.5194/os-1-45-2005>
- Gruber, N. (2011). Warming up, turning sour, losing breath: Ocean biogeochemistry under global change. *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, *369*, 1980–1996. <https://doi.org/10.1098/rsta.2011.0003>
- Hamme, R., & Emerson, S. (2004). The solubility of neon, nitrogen and argon in distilled water and seawater. *Deep-Sea Research I*, *51*(11), 1517–1528. <https://doi.org/10.1016/j.dsr.2004.06.009>
- Heimann, M., & Körner, S. (2003). *The global atmospheric tracer model TM3: Model description and user's manual, release 3.8a*. Max Planck Institute for Biogeochemistry.
- Jamieson, D., Tudhope, J., Morris, R., & Cartwright, G. (1969). Physical properties of sea water solutions: Heat capacity. *Desalination*, *7*(1), 23–30. [https://doi.org/10.1016/S0011-9164\(00\)80271-4](https://doi.org/10.1016/S0011-9164(00)80271-4)
- Jenkins, W., & Goldman, J. (1985). Seasonal oxygen cycling and primary production in the Sargasso Sea. *Journal of Marine Research*, *43*, 465–491. <https://doi.org/10.1357/002224085788438702>
- Jin, X., Najjar, R. G., Louanchi, F., & Doney, S. C. (2007). A modeling study of the seasonal oxygen budget of the global ocean. *Journal of Geophysical Research*, *112*(C5), C05017. <https://doi.org/10.1029/2006JC003731>
- Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., et al. (1996). The NCEP/NCAR 40-year reanalysis project. *Bulletin of the American Meteorological Society*, *77*(3), 437–471. [https://doi.org/10.1175/1520-0477\(1996\)077<0437:TNYRP>2.0.CO;2](https://doi.org/10.1175/1520-0477(1996)077<0437:TNYRP>2.0.CO;2)

- Keeling, R., Blaine, T., Paplawsky, B., Katz, L., Atwood, C., & Brockwell, T. (2004). Measurement of changes in atmospheric Ar/N<sub>2</sub> ratio using a rapid-switching, single-capillary mass spectrometer system. *Tellus B: Chemical and Physical Meteorology*, 56B(4), 322–338. <https://doi.org/10.3402/tellusb.v56i4.16453>
- Keeling, R., & Garcia, H. (2002). The change in oceanic O<sub>2</sub> inventory associated with recent global warming. *Proceedings of the National Academy of Sciences*, 99(12), 7848–7853. <https://doi.org/10.1073/pnas.122154899>
- Keeling, R., Körtzinger, A., & Gruber, N. (2010). Ocean deoxygenation in a warming world. *Annual Review of Marine Science*, 2(1), 199–229. <https://doi.org/10.1146/annurev.marine.010908.163855>
- Keeling, R., & Manning, A. (2014). Studies of recent changes in atmospheric O<sub>2</sub> content. *Treatise on Geochemistry* (2nd ed., pp. 385–404). Elsevier Ltd. <https://doi.org/10.1016/B978-0-08-095975-7.00420-4>
- Keeling, R., Manning, A., Paplawsky, W., & Cox, A. (2007). On the long-term stability of reference gases for atmospheric O<sub>2</sub>/N<sub>2</sub> and CO<sub>2</sub> measurements. *Tellus B: Chemical and Physical Meteorology*, 59B, 3–14. <https://doi.org/10.1111/j.1600-0889.2006.00228.x>
- Keeling, R., Manning, A. C., McEvoy, E., & Shertz, S. (1998). Methods for measuring changes in atmospheric O<sub>2</sub> concentration and their application in southern hemisphere air. *Journal of Geophysical Research*, 103(D3), 3381–3397. <https://doi.org/10.1029/97JD02537>
- Keeling, R., Najjar, R., Bender, M., & Tans, P. (1993). What atmospheric oxygen measurements can tell us about the global carbon cycle. *Global Biogeochemical Cycles*, 7(1), 37–67. <https://doi.org/10.1029/92GB02733>
- Keeling, R., & Shertz, S. (1992). Seasonal and interannual variations in atmospheric oxygen and implications for the global carbon cycle. *Nature*, 358, 723–727. <https://doi.org/10.1038/358723a0>
- Keeling, R., Stephens, B., Najjar, R., Doney, S., Archer, D., & Heimann, M. (1998). Seasonal variations in the atmospheric O<sub>2</sub>/N<sub>2</sub> ratio in relation to the kinetics of air-sea gas exchange. *Global Biogeochemical Cycles*, 12(1), 141–163. <https://doi.org/10.1029/97GB02339>
- Large, W., & Yeager, S. (2009). The global climatology of an interannually varying air-sea flux data set. *Climate Dynamics*, 33(2), 341–364. <https://doi.org/10.1007/s00382-008-0441-3>
- Legendre, P., & Legendre, L. (2012). *Numerical ecology* (3rd ed.). Elsevier.
- Long, M., Deutsch, C., & Ito, T. (2016). Finding forced trends in oceanic oxygen. *Global Biogeochemical Cycles*, 30(2), 381–397. <https://doi.org/10.1002/2015GB005310>
- Long, M., Lindsay, K., Peacock, S., Moore, J., & Doney, S. (2013). Twentieth-century oceanic carbon uptake and storage in CESM1(BGC). *Journal of Climate*, 26(18), 6775–6800. <https://doi.org/10.1175/JCLI-D-12-00184.1>
- Madec, G. (2016). *Technical report: NEMO ocean engine. Note du Pôle de modélisation de l'Institut Pierre-Simon Laplace No 27*. NEMO.
- Manizza, M., Keeling, R., & Nevison, C. (2012). On the processes controlling the seasonal cycles of the air-sea fluxes of O<sub>2</sub> and N<sub>2</sub>O: A modelling study. *Tellus B: Chemical and Physical Meteorology*, 64, 1–18. <https://doi.org/10.3402/tellusb.v64i0.18429>
- Manning, A., & Keeling, R. (2006). Global oceanic and land biotic carbon sinks from the Scripps atmospheric oxygen flask sampling network. *Tellus Series B Chemical and Physical Meteorology*, 58B, 95–116. <https://doi.org/10.1111/j.1600-0889.2006.00175.x>
- Marshall, J., Adcroft, A., Hill, C., Perelman, L., & Heisey, C. (1997). A finite-volume, incompressible Navier Stokes model for studies of the ocean on parallel computers. *Journal of Geophysical Research: Oceans*, 102(C3), 5753–5766. <https://doi.org/10.1029/96JC02775>
- Moore, J., & Braucher, O. (2008). Sedimentary and mineral dust sources of dissolved iron to the world ocean. *Biogeosciences*, 5(3), 631–656. <https://doi.org/10.5194/bg-5-631-2008>
- Moore, J., Doney, S., & Lindsay, K. (2004). Upper ocean ecosystem dynamics and iron cycling in a global three-dimensional model. *Global Biogeochemical Cycles*, 18(4). <https://doi.org/10.1029/2004GB002220>
- Moore, J., Lindsay, K., Doney, S., Long, M., & Misumi, K. (2013). Marine ecosystem dynamics and biogeochemical cycling in the community earth system model [CESM1(BGC)]: Comparison of the 1990s with the 2090s under the RCP4.5 and RCP8.5 scenarios. *Journal of Climate*, 26(23), 9291–9312. <https://doi.org/10.1175/JCLI-D-12-00566.1>
- Naegler, T. (2009). Reconciliation of excess <sup>14</sup>C-constrained global CO<sub>2</sub> piston velocity estimates. *Tellus Series B Chemical and Physical Meteorology*, 61B(2), 372–384. <https://doi.org/10.1111/j.1600-0889.2008.00408.x>
- Najjar, R., & Keeling, R. (1997). Analysis of the mean annual cycle of the dissolved oxygen anomaly in the World Ocean. *Journal of Marine Research*, 55(1), 117–151. <https://doi.org/10.1357/0022240973224481>
- Najjar, R., & Keeling, R. (2000). Mean annual cycle of the air-sea oxygen flux: A global view. *Global Biogeochemical Cycles*, 14(2), 573–584. <https://doi.org/10.1029/1999GB900086>
- Nevison, C., Baker, D., & Gurney, K. (2012a). A methodology for estimating seasonal cycles of atmospheric CO<sub>2</sub> resulting from terrestrial net ecosystem exchange (NEE) fluxes using the Transcom T3L2 pulse-response functions. *Geoscientific Model Development Discussions*, 5, 2789–2809. <https://doi.org/10.5194/gmdd-5-2789-2012>
- Nevison, C., Keeling, R., Kahru, M., Manizza, M., Mitchell, B., & Cassar, N. (2012b). Estimating net community production in the Southern Ocean based on atmospheric potential oxygen and satellite ocean color data. *Global Biogeochemical Cycles*, 26(1). <https://doi.org/10.1029/2011GB004040>
- Nevison, C., Mahowald, N., Doney, S., Lima, I., & Cassar, N. (2008). Impact of variable air-sea O<sub>2</sub> and CO<sub>2</sub> fluxes on atmospheric potential oxygen (APO) and land-ocean carbon sink partitioning. *Biogeosciences*, 5(3), 875–889. <https://doi.org/10.5194/bg-5-875-2008>
- Nevison, C., Manizza, M., Keeling, R., Stephens, B., Bent, J., Dunne, J., et al. (2016). Evaluating CMIP5 ocean biogeochemistry and Southern Ocean carbon uptake using atmospheric potential oxygen: Present-day performance and future projection. *Geophysical Research Letters*, 43(5), 2077–2085. <https://doi.org/10.1002/2015GL067584>
- Oschlies, A., Brandt, P., Stramma, L., & Schmidtko, S. (2018). Drivers and mechanisms of ocean deoxygenation. *Nature Geoscience*, 11(7), 467–473. <https://doi.org/10.1038/s41561-018-0152-2>
- Pereira, R., Ashton, I., Sabbaghzadeh, B., Shutler, J., & Upstill-Goddard, R. (2018). Reduced air-sea CO<sub>2</sub> exchange in the Atlantic Ocean due to biological surfactants. *Nature Geoscience*, 11, 492–496. <https://doi.org/10.1038/s41561-018-0136-2>
- Rödenbeck, C., Quéré, C. L., Heimann, M., & Keeling, R. (2008). Interannual variability in oceanic biogeochemical processes inferred by inversion of atmospheric O<sub>2</sub>/N<sub>2</sub> and CO<sub>2</sub> data. *Tellus Series B Chemical and Physical Meteorology*, 60B(5), 685–705. <https://doi.org/10.1111/j.1600-0889.2008.00375.x>
- Rodgers, K., Aumont, O., Mikaloff Fletcher, S., Plancherel, Y., Bopp, L., de Boyer Montégut, C., et al. (2014). Strong sensitivity of Southern Ocean carbon uptake and nutrient cycling to wind stirring. *Biogeosciences*, 11(15), 4077–4098. <https://doi.org/10.5194/bg-11-4077-2014>
- Sarmiento, J., Slater, R., Dunne, J., Gnanadesikan, A., & Hiscock, M. R. (2010). Efficiency of small scale carbon mitigation by patch iron fertilization. *Biogeosciences*, 7(11), 3593–3624. <https://doi.org/10.5194/bg-7-3593-2010>
- Schmidtko, S., Stramma, L., & Visbeck, M. (2017). Decline in global oceanic oxygen content during the past five decades. *Nature*, 542(7641), 335–339. <https://doi.org/10.1038/nature21399>

- Shea, D., Trenberth, K., & Reynolds, R. (1992). A global monthly sea surface temperature climatology. *Journal of Climate*, 5(9), 987–1001. [https://doi.org/10.1175/1520-0442\(1992\)005<0987:AGMSST>2.0.CO;2](https://doi.org/10.1175/1520-0442(1992)005<0987:AGMSST>2.0.CO;2)
- Smith, R., Jones, P., Briegleb, B., Bryan, F., Danabasoglu, G., Dennis, J., & Yeager, S. (2010). *The parallel ocean program (POP) reference manual: Ocean component of the community climate system model (CCSM) and community earth system model (CESM) (No. LAUR-10-01853)*. Los Alamos National Laboratory.
- Spitzer, W., & Jenkins, W. (1989). Rates of vertical mixing, gas exchange and new production: Estimates from seasonal gas cycles in the upper ocean near Bermuda. *Journal of Marine Research*, 47(1), 169–196. <https://doi.org/10.1357/002224089785076370>
- Steinbach, J., Gerbig, C., Rödenbeck, C., Karstens, U., Minejima, C., & Mukai, H. (2011). The CO<sub>2</sub> release and oxygen uptake from fossil fuel emission estimate (coffee) dataset: Effects from varying oxidative ratios. *Atmospheric Chemistry and Physics*, 11(14), 6855–6870. <https://doi.org/10.5194/acp-11-6855-2011>
- Stephens, B., Keeling, R., Heimann, M., Six, K., Murnane, R., & Caldeira, K. (1998). Testing global ocean carbon cycle models using measurements of atmospheric and concentration. *Global Biogeochemical Cycles*, 12(2), 213–230. <https://doi.org/10.1029/97GB03500>
- Takahashi, T., Sutherland, S., Sweeney, C., Poisson, A., Metzl, N., Tilbrook, B., et al. (2002). Global sea-air CO<sub>2</sub> flux based on climatological surface ocean pCO<sub>2</sub>, and seasonal biological and temperature effects. *Deep Sea Research Part II: Topical Studies in Oceanography*, 49(9), 1601–1622. [https://doi.org/10.1016/S0967-0645\(02\)00003-6](https://doi.org/10.1016/S0967-0645(02)00003-6)
- Takahashi, T., Sutherland, S., Wanninkhof, R., Sweeney, C., Feely, R., Chipman, D., et al. (2009). Climatological mean and decadal change in surface ocean pCO<sub>2</sub>, and net sea-air CO<sub>2</sub> flux over the global oceans. *Deep Sea Research Part II: Topical Studies in Oceanography*, 56(8), 554–577. <https://doi.org/10.1016/j.dsr2.2008.12.009>
- Uppala, S., Kllberg, P., Simmons, A., Andrae, U., Bechtold, V. D. C., Fiorino, M., et al. (2005). The ERA-40 re-analysis. *Quarterly Journal of the Royal Meteorological Society*, 131(612), 2961–3012. <https://doi.org/10.1256/qj.04.176>
- Wanninkhof, R. (1992). Relationship between wind speed and gas exchange. *Journal of Geophysical Research*, 97(92), 7373–7382. <https://doi.org/10.1029/92JC00188>
- Wanninkhof, R. (2014). Relationship between wind speed and gas exchange over the ocean revisited. *Limnology and Oceanography: Methods*, 12, 351–362. <https://doi.org/10.4319/lom.2014.12.351>
- Weiss, R. (1970). The solubility of nitrogen, oxygen and argon in water and seawater. *Deep-Sea Research*, 17(4), 721–735. [https://doi.org/10.1016/0011-7471\(70\)90037-9](https://doi.org/10.1016/0011-7471(70)90037-9)
- Weiss, R., Jahnke, R., & Keeling, C. (1982). Seasonal effects of temperature and salinity on the partial pressure of CO<sub>2</sub> in seawater. *Nature*, 300, 511–513. <https://doi.org/10.1038/300511a0>