

# On evaluating ocean models with atmospheric potential oxygen

By TOBIAS NAEGLER<sup>1,2\*</sup>, PHILIPPE CIAIS<sup>1</sup>, JAMES C. ORR<sup>1</sup>, OLIVIER AUMONT<sup>3,4</sup>  
and CHRISTIAN RÖDENBECK<sup>5</sup>, <sup>1</sup>LSCE/IPSL, Laboratoire des Sciences du Climat et de l'Environnement,  
CEA-CNRS-UVSQ, Gif-sur-Yvette, France; <sup>2</sup>now at Institut für Umweltphysik, Heidelberg, Germany; <sup>3</sup>Laboratoire  
d'Océanographie Dynamique et de Climatologie, Paris, France; <sup>4</sup>now at L'Institut de Recherche pour le  
Développement, Brest, France; <sup>5</sup>Max-Planck Institut für Biogeochemie, Jena, Germany

(Manuscript received 18 May 2005; in final form 26 June 2006)

## ABSTRACT

We used observed and simulated atmospheric potential oxygen (APO) to evaluate simulated air-sea flux fields from 11 ocean global carbon cycle models. APO is defined in terms of atmospheric CO<sub>2</sub>, O<sub>2</sub> and N<sub>2</sub> so as not to depend on terrestrial photosynthesis and respiration. Hence, it is in principal suited to evaluate simulated air-sea fluxes of these gases. We forced two different atmospheric transport models, TM2 and TM3, with simulated air-sea fluxes from each of the 11 ocean models, and we compared resulting simulated latitudinal and seasonal variations in APO with observations. Differences between the two atmospheric transport models, which offer a first estimate of uncertainty due to atmospheric transport, are similar in magnitude to the average model-data differences and to the spread between the ocean models. Simulated annual mean meridional APO profiles qualitatively resemble the observations, although at individual stations there remain substantial differences between models and observations. The simulated amplitude of the seasonal APO variability was generally less than observed. We conclude that it is difficult to validate ocean models based on APO because shortcomings in atmospheric transport models and problems with data representativity cannot be distinguished from ocean model deficiencies.

## 1. Introduction

Ocean models have been evaluated by comparing observed and simulated distributions of tracers, such as CO<sub>2</sub>, <sup>14</sup>C and CFCs. Air-sea fluxes of gases simulated by ocean models may also be evaluated by imposing these fluxes in an atmospheric transport model and then comparing results with atmospheric observations. Yet atmospheric concentrations of most gases rarely depend only on air-sea fluxes. For example, atmospheric CO<sub>2</sub> and O<sub>2</sub> concentrations also depend on exchange of these gases between atmosphere and terrestrial biosphere (air-land exchange). However, the magnitude and interannual variability of air-land exchange of CO<sub>2</sub> remain controversial (Tans et al., 1990; Ciais et al., 1995; Le Quéré et al., 2000; Orr et al., 2001; Gurney et al., 2002). Furthermore, the seasonal cycle of atmospheric CO<sub>2</sub> largely results from air-land exchange. Recent model intercomparisons that have focused on the terrestrial biosphere produce considerable differences in the phase and amplitude of atmospheric CO<sub>2</sub> (Heimann et al., 1998). Furthermore, the terrestrial

rectifier effect for CO<sub>2</sub>, i.e. the covariance of atmospheric transport with the source and sink terms (Denning et al., 1995; Law et al., 1996) has proved difficult to quantify. Similar problems arise for oxygen, because CO<sub>2</sub> and O<sub>2</sub> fluxes are tightly coupled during photosynthesis and respiration. Thus, neither O<sub>2</sub> nor CO<sub>2</sub> atmospheric observations can be used individually to test an ocean model's corresponding air-sea fluxes. Uncertainties associated with the air-land exchange of these gases are too important.

Keeling and Shertz (1992) and Keeling et al. (1998) devised a scheme to deal with this problem by exploiting the close link between O<sub>2</sub> and CO<sub>2</sub> fluxes during air-land exchange. They assumed that for terrestrial biospheric processes, O<sub>2</sub> and CO<sub>2</sub> fluxes are coupled by a constant stoichiometric ratio (O<sub>2</sub>:C =  $-1.10 \pm 0.05$ , Severinghaus, 1995). They also considered that the analogous ratio for air-sea exchange (due to photosynthesis and respiration in the marine biosphere) was very different (O<sub>2</sub>:C =  $-1.4 \pm 0.2$ , Takahashi et al., 1985). Furthermore, CO<sub>2</sub> and O<sub>2</sub> are not as tightly coupled in the ocean as in the terrestrial biosphere, i.e. CO<sub>2</sub> is affected by ocean carbonate chemistry whereas O<sub>2</sub> is not. This different chemical behavior and the different terrestrial and marine stoichiometries allow us to separate the marine and terrestrial signals.

---

\*Corresponding author.  
e-mail: tobias.naegler@iup.uni-heidelberg.de  
DOI: 10.1111/j.1600-0889.2006.00197.x

Consequently, Keeling and Shertz (1992) and Keeling et al. (1998) defined a new tracer called ‘oceanic  $O_2/N_2$ ’ as  $O_2 + 1.1 \cdot CO_2$ , which is thus independent of terrestrial biospheric processes and also not affected by the terrestrial biospheric rectifier effect. This tracer is then mainly sensitive to air-sea gas exchange of  $O_2$ ,  $N_2$  and  $CO_2$  and it depends only slightly on fossil fuel combustion. Stephens et al. (1998) refined this concept, accounting for oxidation of  $CH_4$  and  $CO$ , and calling the new tracer ‘atmospheric potential oxygen’ (APO). Seasonal cycles of APO were used to test the seasonality of ocean model dynamics as well as the seasonality of the ocean biology by Six and Maier-Reimer (1996), Stephens et al. (1998) and Aumont (1998). Garcia and Keeling (2001) tested the seasonality of oceanic  $O_2$  fluxes derived from observations of the dissolved oxygen anomaly. Balkanski et al. (1999) used APO to test ocean primary production they had estimated from satellite data, whereas Gruber et al. (2001) tested their oceanic  $O_2$  fluxes derived from bulk data on the basis of a steady-state inverse modeling technique. Stephens et al. (1998) were the first to use annual mean APO gradients to test large-scale fluxes from ocean models. They concluded that shortcomings in the ocean models were responsible for the large differences in simulated versus observed APO. Conversely, Aumont (1998) found that uncertainties in atmospheric transport contributed substantially to the mismatch in APO, thereby questioning the approach of evaluating ocean models with observed APO.

Here, we have further tested the idea of validating ocean models with APO observations. For this purpose, we have simulated APO in two atmospheric transport models by driving each of them with monthly (if available, otherwise annual)  $CO_2$ ,  $O_2$  and  $N_2$  fluxes from 11 different ocean carbon cycle models (Table 1), all of which participated in the second phase of the Ocean Carbon-Cycle Model Intercomparison Project (OCMIP-2). Furthermore, we made analogous simulations but with data-

based flux fields of  $O_2$ ,  $N_2$  (Garcia and Keeling, 2001) and  $CO_2$  (Takahashi et al., 2002). We then compared simulated and observed APO, both in terms of annual mean gradients (OCMIP-2 models only) and seasonal cycles. Although our analysis is structurally similar to studies from Stephens et al. (1998), Gruber et al. (2001) and Battle et al. (2006), our use of multiple atmospheric transport models and multiple ocean models allows a more thorough analysis of uncertainties in atmospheric transport as well as an assessment of the sensitivity of simulated APO to a number of ocean model characteristics. Furthermore, the availability of new high quality APO data sets from Battle et al. (2006) and Tohjima et al. (2005) allows us to also focus on how interannual variability and data quality may degrade the ability to validate ocean models with atmospheric data.

## 2. APO: concept and controlling factors

Stephens et al. (1998) defined APO, in units of per meg, as

$$APO = \delta \left( \frac{O_2}{N_2} \right) + \left( \frac{1.1}{X_{O_2}} \right) \cdot TAC, \quad (1)$$

where TAC is total atmospheric carbon (in ppm) and  $X_{O_2}$  denotes the fraction of oxygen in air (0.2095) and serves in the conversion of units from ppm to per meg. Like APO, the ratio  $\delta \left( \frac{O_2}{N_2} \right)$  is also given in units of per meg:

$$\delta \left( \frac{O_2}{N_2} \right) = \left[ \frac{(O_2/N_2)_{sample}}{(O_2/N_2)_{reference}} - 1 \right] \times 10^6. \quad (2)$$

Although Stephens et al. (1998) define  $TAC = CO_2 + CH_4 + CO$ , we neglected the latter two terms in our simulations because atmospheric concentrations and air-sea fluxes of  $CH_4$  and  $CO$  are small when compared to  $CO_2$ . Thus  $TAC = CO_2$ .

Using an average stoichiometric  $O_2:CO_2$  ratio of  $-1.39$  for fossil fuel burning and of  $-1.1$  for photosynthesis and respiration

Table 1. Model characteristics and global-scale, air-to-sea fluxes of  $CO_2$ ,  $O_2$  and  $N_2$  for the reference year 1995. SD: seasonal dynamics, SB: seasonal biology, SIM: sea ice model, LSSM: lateral subgrid scale mixing, IS: isopycnal, HOR: horizontal, GM: Gent & McWilliams, NA: not available.

Model	Reference	SD	SB	SIM	LSSM	$CO_2$ flux PgC yr <sup>-1</sup>	$O_2$ flux Tmol yr <sup>-1</sup>	$N_2$ flux Tmol yr <sup>-1</sup>
AWI	Schlitzer (2000)	–	–	–	IS	2.8	3	98
CSIRO	Matear and Hirst (1999)	yes	yes	–	IS+GM	2.1	–219	32
IGCR	Yamanaka and Tajika (1996)	–	yes	–	HOR	2.5	–249	91
IPSL	Madec et al. (1998)	yes	yes	–	IS+GM	2.1	–198	127
LLNL	Duffy et al. (1997)	yes	yes	yes	IS+GM	2.1	–33	NA
MPIM	Maier-Reimer et al. (1993)	yes	yes	yes	HOR	2.5	–178	59
NCAR	Doney and Hecht (2002)	yes	yes	–	IS+GM	2.3	–250	103
PIUB	Stocker et al. (1992)	–	–	yes	HOR	2.3	–126	99
PRINCE	Gnanadesikan et al. (2002)	yes	yes	–	IS+GM	2.0	–25	70
SOC	Gordon (2000)	yes	yes	–	IS+GM	2.4	–128	NA
UL	Goosse et al. (1998)	yes	yes	yes	HOR	3.0	–140	103

(Severinghaus, 1995), APO (expressed in per meg) is calculated in our study as:

$$\begin{aligned}
 APO = & \left( \frac{1.0}{X_{O_2}} \right) \cdot (O_2^o - 1.39 \cdot CO_2^f) \\
 & - \left( \frac{1.0}{X_{N_2}} \right) \cdot (N_2^o) \\
 & + \left( \frac{1.1}{X_{O_2}} \right) \cdot (CO_2^o + CO_2^f), \quad (3)
 \end{aligned}$$

where  $O_2^o$ ,  $N_2^o$  and  $CO_2^o$  denote the simulated atmospheric concentrations of  $O_2$ ,  $N_2$  and  $CO_2$  due to exchange with the ocean and  $CO_2^f$  that due to fossil carbon emissions (all concentrations in ppm). As for  $X_{O_2}$ ,  $X_{N_2}$  is the fraction of nitrogen in air (0.7808) and allows for the conversion of ppm to per meg. The (globally constant)  $O_2:CO_2$  ratio for fossil fuel burning of  $-1.39$  is calculated from the Marland et al. (2005)  $CO_2$  emission data for the year 2000 and the average  $O_2:CO_2$  ratios of  $-1.95$  for gas fuel,  $-1.44$  for liquid fuel,  $-1.17$  for solid fuel,  $-1.98$  for gas flaring respectively 0.00 for cement production from Keeling (1988).

A detailed discussion of the factors controlling  $CO_2$ ,  $O_2$  and  $N_2$  air-sea fluxes can be found in Keeling et al. (1993), Stephens et al. (1998), Aumont (1998) and Keeling et al. (1998). Here, we provide only a brief summary.

Air-sea fluxes of gases, such as  $CO_2$ ,  $O_2$  and  $N_2$  are governed by both the piston velocity as well as the difference in partial pressure across the air-sea interface. The piston velocity is often parameterized as a function of wind speed (Wanninkhof, 1992), meaning there are high piston velocities over the subpolar oceans and low piston velocities in the tropics. Sea surface partial pressures of  $CO_2$  and  $O_2$  are controlled by the solubility pump (temperature and salinity effects) and the biological pump (photosynthesis and respiration effects as well as the  $CaCO_3$  counter-pump, see below). On the other hand for  $N_2$ , only the solubility pump is important; nitrogen fixation and denitrification are negligible. Additionally,  $\Delta pCO_2$  is subject to the increase of the atmospheric  $CO_2$  concentration due to human activities.

The effect of salinity on the solubility of  $CO_2$ ,  $O_2$  and  $N_2$  is relatively weak. Therefore, the solubility pump reflects mainly variations in sea surface temperature (SST) and thus heat fluxes. As water is transported poleward, SST decreases thereby increasing solubility of these three gases. Good examples of this transport are the Gulf Stream and the Kuroshio Current, both of which flow as western boundary currents transporting water from the tropics to the high northern latitudes. Thus, cooling of surface waters acts to increase oceanic uptake of  $CO_2$ ,  $O_2$  and  $N_2$ . Cooling also acts as a sink for APO because of  $N_2$ 's lower solubility and its minor contribution to changes in APO (eq. 3). Conversely, in the low latitudes, cold subsurface waters upwell due to equatorial divergence and they are warmed at the surface. Thus,  $CO_2$ ,  $O_2$  and  $N_2$  are lost from the ocean to the atmosphere, thereby increasing APO. The effect of the solubility pump is thus

to create an APO maximum at the equator and APO minima in the high latitudes.

Variations in  $CO_2$  and  $O_2$  due to the biological pump are superimposed on the thermal signal from the solubility pump. The biological pump produces opposing effects for both gases: During photosynthesis,  $CO_2$  is assimilated and  $O_2$  released; conversely, respiration produces  $CO_2$  and consumes  $O_2$ . Because the  $O_2:CO_2$  ratio for the marine biosphere of 1.4 (Takahashi et al., 1985) is higher than that for the terrestrial biosphere (1.1), marine photosynthesis increases APO whereas marine respiration decreases APO. Upwelling in the high latitudes explains why waters there are relatively enriched in  $CO_2$  and depleted in  $O_2$  (due to net respiration) (Keeling et al., 1993). On the other hand, surface waters converge and downwell in the subtropics carrying with them relatively high levels of  $O_2$  and low levels of  $CO_2$  (due to net photosynthesis).

Furthermore, the  $CaCO_3$  counter-pump (production of  $CaCO_3$  shells at the ocean surface then export and their dissolution at depth) increases  $pCO_2$  in the euphotic zone and reduces  $pCO_2$  in deep waters. Following Toggweiler and Samuels (1993), upwelling of deep waters in the Southern Ocean means that the  $CaCO_3$  counter-pump diminishes  $pCO_2$  in that region, but it increases  $pCO_2$  further north in low and mid-latitudes. Thus, the  $CaCO_3$  pump acts as an APO sink in the southern high latitudes (it decreases surface  $pCO_2$ ) and as an APO source at lower latitudes. The northern oceans may act similarly, although the high latitudes are less extensive and have less upwelling of deep water. In terms of APO then, the  $CaCO_3$  pump acts to reinforce the organic carbon pump in the high latitudes (APO sink) and in the subtropics (APO source); however, in the tropics these two components of the biological pump oppose one another. The combined biological pump thus creates an APO sink in the high latitudes and in the tropics and an APO source in the subtropical gyres.

Furthermore, the solubility pump and the combined biological pump appear to act in the same direction in the high latitudes (APO sink) and the subtropical gyres (APO source). Conversely, at the equator, these two pumps act in opposite directions: the solubility pump act as an APO source, while the biological pump acts as an APO sink. In general, the solubility pump drives  $CO_2$  and  $O_2$  fluxes in the same direction, with both affecting APO in the same sense. On the other hand,  $CO_2$  and  $O_2$  fluxes from the biological pump have opposite signs so that the net effect on APO is smaller. Hence simulated variations in APO are particularly sensitive to air-sea heat fluxes from the ocean models.

The air-to-sea flux of anthropogenic  $CO_2$ , particularly where it is largest, in the high latitudes and the tropics, tends to decrease the net sea-to-air flux of APO. More uptake of anthropogenic  $CO_2$  by the Southern Hemisphere, due in part to larger ocean surface area, creates an APO gradient between northern and Southern Hemispheres, with higher APO in the north.

### 3. Methods

#### 3.1. Modeled air-sea fluxes of $O_2$ , $N_2$ and $CO_2$

For air-sea fluxes of  $O_2$ ,  $N_2$  and  $CO_2$ , we used output from the 11 OCMIP ocean models summarized in Table 1. Details of each model are provided in references given in Table 1. All models are global and have coarse horizontal resolution, meaning they do not resolve eddies. Most of the models are primitive equation models. Exceptions are the MPIM model, which is a large-scale geostrophic model, AWI, which uses an adjoint technique to derive the circulation from hydrographic and geochemical data and the PIUB model, where the momentum equations are balanced between Coriolis forces, horizontal pressure gradients and zonal wind stress. The PIUB model is also a zonally averaged basin model. Models differ in their dynamical forcing, horizontal and vertical resolution, bottom topography, boundary layer schemes, surface boundary conditions, representation of the mixed layer, lateral sub-gridscale mixing, vertical discretization and advection schemes. Four models (LLNL, MPIM, PIUB and UL) are coupled to a sea-ice model. AWI and PIUB are annual mean models, the IGCR model employs annual mean ocean dynamics, but calculates ocean biology on a monthly timestep. Here we refer to CSIRO, IPSL, LLNL, MPIM, NCAR, PRINCE and SOC as the fully seasonal models, to IGCR as the semi-seasonal model and to the AWI and PIUB as the non-seasonal models.

Although the models differ in their physical characteristics, they use a common simple biogeochemical module. This module uses phosphorus as the only limiting nutrient; nitrogen and iron are not included. The OCMIP-2 approach to describing the phosphorus cycle is based on the ‘nutrient restoring’ approach of Najjar et al. (1992) and Anderson and Sarmiento (1995), but it also includes a semi-labile pool of DOC. Modeled surface phosphate concentrations are restored toward an observed monthly-mean climatology (Louanchi and Najjar, 2000). Production and consumption of oxygen and DIC are linked to the production and consumption of inorganic phosphorus via the constant Redfield ratios of  $r_{-O_2:P}=170$  and  $r_{C:P}=117$  (Anderson and Sarmiento, 1994). Details can be found in Najjar and Orr (1999).

The air-sea fluxes of  $O_2$  and  $CO_2$  from each ocean model were calculated using eq. 3 from Wanninkhof (1992), i.e. his wind speed-transfer coefficient for long-term winds, using a monthly mean climatology of satellite-derived winds (SSM/I, Goodberlet et al., 1989; Boutin and Etcheto, 1996). Schmidt numbers were calculated from simulated SST and the formulations from Keeling et al. (1998) for  $O_2$  and from Wanninkhof (1992) for  $CO_2$ . The piston velocity in grid cells containing sea ice was reduced by the factor  $(1 - I)$ , where  $I$  is the fraction of ice cover from monthly climatologies of Zwally et al. (1983) and Walsh (1978). We calculated the air-sea flux of  $N_2$  as

$$F_{N_2} = -\frac{\partial\beta_{N_2}}{\partial T} \cdot \frac{F_H}{c_p}, \quad (4)$$

where  $\partial\beta_{N_2}/\partial T$  is the Weiss (1970) solubility-temperature dependence for  $N_2$  using simulated SST and SSS,  $F_H$  is the modeled heat flux and  $c_p$  is the heat capacity of seawater (a globally uniform value of  $3.9 \cdot 10^6 \text{ J m}^{-3} \text{ K}^{-1}$ ). Here we assumed instantaneous air-sea equilibrium for nitrogen as well as heat fluxes. For the LLNL and SOC models, which did not provide heat fluxes, we calculated simulated APO without the  $N_2$  contribution.

#### 3.2. Observed air-sea fluxes of $O_2$ , $N_2$ and $CO_2$

For comparison with the OCMIP-derived results, we relied on air-sea fluxes of  $CO_2$ ,  $O_2$  and  $N_2$  based on oceanographic observations. For the  $CO_2$  flux, we started with the Takahashi et al. (2002) p $CO_2$  maps. Then we calculated the  $CO_2$  fluxes using marine boundary layer  $CO_2$  concentrations for 1995 from GlobalView (GlobalView, 2003) and the Wanninkhof (1992) quadratic gas transfer coefficient for long-term winds with monthly mean ECMWF wind fields for 1995.

For  $O_2$ , Garcia and Keeling (2001) calculated oxygen anomalies as the difference between measured and equilibrium concentrations of  $O_2$  in seawater at the measured temperature, salinity and sea level air pressure. To calculate the air-sea flux  $F_{O_2}$ , they also used Wanninkhof’s gas transfer coefficient for long-term winds, the Schmidt number from Keeling et al. (1998) and ECMWF monthly average winds. They calculated fluxes at the time and location of individual measurements and interpolated them in space and time by taking advantage of the correlation between  $F_{O_2}$  and heat flux anomalies. Thus, they defined a global monthly mean climatology of the air-sea  $O_2$  flux anomaly having an annual mean flux of zero at each grid point. For  $N_2$ , Garcia and Keeling calculated seasonal fluxes of nitrogen,  $F_{N_2}$ , as the product of ECMWF seasonal heat flux anomalies times the temperature derivative of the  $N_2$  solubility (Weiss, 1970). Because Garcia and Keeling (2001) provide only the seasonal component of the  $O_2$  and  $N_2$  fluxes, we used these estimates only in our analysis of seasonal variability. That is, we used the Garcia and Keeling (2001)  $O_2$  and  $N_2$  flux anomalies as boundary conditions to our atmospheric transport models in order to calculate the annual cycle of  $O_2/N_2$ . By combining those results with the annual cycle of atmospheric  $CO_2$ , derived from the Takahashi et al. (2002) p $CO_2$ , we were able to provide a data-based estimate of the annual cycle of APO, i.e. without recourse to an ocean model. We denote this fashion of computing APO as the GKT approach (for Garcia, Keeling and Takahashi).

#### 3.3. Fossil fuel fluxes

The annual mean  $CO_2$  emissions from fossil fuel combustion and cement manufacturing were compiled by Andres et al. (1996). When integrated spatially Andres’ maps for 1990 amount to a global emissions of  $5.8 \text{ Pg C yr}^{-1}$ . We scaled these maps, linearly, to reach the emissions rate of  $6.7 \text{ Pg C yr}^{-1}$  for year 2000 (Marland et al., 2005). We chose that year as the reference for

the fossil fuel emissions because it represents the middle of the period where data from Battle et al. (2006) are available. The uncertainty in the global emissions rate is about 10%, but regional uncertainties may be larger. For simplicity, we used a globally constant  $O_2:CO_2$  stoichiometric ratio of  $-1.39$ , a weighted average for the different fossil fuels and cement production (see Section 2).

### 3.4. Atmospheric transport

To put fluxes in terms of APO, we made simulations in two atmospheric models, TM2 and TM3. For boundary conditions we used the ocean model- and data-based air-sea fluxes of  $CO_2$ ,  $O_2$  and  $N_2$  as well as the perturbations in  $CO_2$  and  $O_2$  from fossil emissions. The two transport models provide a first estimate of uncertainty in simulated APO due to atmospheric transport. However, similarities between TM2 and TM3 mean that we under-predict the uncertainty due to atmospheric transport.

TM2 is an offline transport model (Heimann, 1995; Ramonet et al., 1996), which we have run using a  $2.5^\circ \times 2.5^\circ$  horizontal grid with 9 vertical sigma layers. Here TM2 was forced with the 12-hourly ECMWF reanalysis fields from 1995, using a 1-hour timestep. Tracers were advected with the slopes scheme from Russel and Lerner (1981). Subgrid-scale vertical transport is parameterized following the cumulus convection scheme of Tiedtke (1989) and stability-dependent vertical diffusion (Louis, 1979).

The TM3 transport model (Heimann and Körner, 2003) evolved from the TM2 model. We used a version of TM3 having a  $5^\circ \times 4^\circ$  horizontal grid and 19 hybrid (sigma-pressure) vertical layers. The same subgrid-scale vertical transport and advection schemes are used in TM3 as in TM2. For simulations with TM3, we used interannually varying NCEP winds from 1992 through 1995, i.e. for the 4 years of the model run.

Following the standard procedure for TransCom (Rayner and Law, 1995; Law et al., 1996), both TM2 and TM3 were spun up for 3 years. Then using fourth year results, we interpolated fields of  $O_2$ ,  $N_2$  and  $CO_2$  to the longitude and latitude of the observed station data. Generally, each model was subsampled in its surface layer. However, due to the coarse resolution of both models, the model topography in a grid box does not always match the real altitude of a station. In this case, the models were subsampled in the model layer representing the geopotential height of the station. To mimic maritime data selection at the two stations where only marine air masses are collected, we shifted model subsampling by  $2.5^\circ$  to the west for La Jolla as well as by  $2.5^\circ$  to both the south and to the east for Cape Grim.

### 3.5. Atmospheric observations

To evaluate simulated APO, we used the observational data sets from Battle et al. (2006) and Tohjima et al. (2005). Battle et al. (2006) published long-term average annual-mean APO for

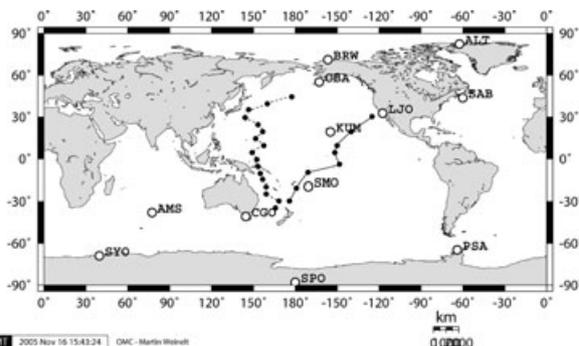


Fig. 1. Location of the APO measurement stations used in this study. Data from land stations (labeled circles) are taken from Battle et al. (2006) (ALT: Alert,  $82^\circ N$ ,  $62^\circ W$ , 210m altitude; BRW: Point Barrow,  $71^\circ N$ ,  $157^\circ W$ , 8m; CBA: Cold Bay,  $55^\circ N$ ,  $162^\circ W$ , 25m; SAB: Sable Island,  $44^\circ N$ ,  $60^\circ W$ , 5m; LJO: La Jolla,  $33^\circ N$ ,  $117^\circ W$ , 20m; KUM: Cape Kumukahi,  $19^\circ N$ ,  $154^\circ W$ , 40m; SMO: American Samoa,  $14^\circ S$ ,  $171^\circ W$ , 42m; AMS: Amsterdam Island,  $38^\circ N$ ,  $78^\circ E$ , 70m; CGO: Cape Grim,  $41^\circ S$ ,  $144^\circ E$ , 90m; PSA: Palmer Station,  $65^\circ S$ ,  $64^\circ W$ , 8m; SYO: Syowa Station,  $69^\circ N$ ,  $40^\circ E$ , 11m; SPO: South Pole,  $90^\circ S$ , 2810m). Approximate ship cruise tracks from Tohjima et al. (2005) in the Western Pacific and Battle et al. (2006) in the Eastern Pacific are indicated by the dashed and solid lines, respectively. The dots indicate the location where the model results are extracted to be compared with the data. The longitude of the dots for samples taken aboard of the Ka'imimoana (at  $4.5^\circ N$  and  $4.5^\circ S$ ) is the average longitude of the exact sampling locations from Battle et al. (2006).

a number of land stations and ship cruises in the Eastern Pacific (their figure 2, panel 4, annual mean APO derived from seasonal cycles) as well as APO time series from their land station network (see Fig. 1). From these time series, we calculated the average observed seasonal cycle of APO for each land station using a standard fitting routine with four harmonics (Thoning et al., 1989). Tohjima et al. (2005) published annual APO concentrations for 2002 and 2003 in the Western Pacific from ship transects (their figure 4).

## 4. Simulated ocean $CO_2$ , $O_2$ and $N_2$ fluxes

Simulated anthropogenic  $CO_2$  uptake in 1995 ranges from  $2.0 Pg C yr^{-1}$  for the PRINCE model to  $3.0 Pg C yr^{-1}$  for the UL model (Table 1). The model spread may be characterized as  $2.3 \pm 0.3 Pg C yr^{-1}$  (mean  $\pm 1\sigma$ ).

Simulated global ocean carbon uptake is not directly comparable to the global oceanic  $CO_2$  flux calculated from the  $pCO_2$  map compiled by Takahashi et al. (2002) (corrected value of  $1.64 Pg C$ , T. Takahashi, personal communication). By design, the OCMIP simulations did not include the river loop, which drives a natural loss of  $0.5 \pm 0.2 Pg C yr^{-1}$  from the ocean (A. Jacobson, Princeton University, personal communication). When we use that to correct the simulated anthropogenic  $CO_2$  flux ( $2.3 \pm 0.3 Pg C yr^{-1}$ ), the total simulated flux to the ocean becomes

$1.8 \pm 0.4 \text{ Pg C yr}^{-1}$ , slightly higher but still consistent within given uncertainties of the estimate from Takahashi et al. (2002).

The large-scale pattern of the total, annual mean air-sea  $\text{CO}_2$  flux (without the river loop) is similar among the OCMIP models:  $\text{CO}_2$  is released in the tropics, particularly from the equatorial Pacific,  $\text{CO}_2$  is absorbed in the northern Atlantic and the northwest Pacific and there are patchy patterns of both uptake and release south of  $40^\circ\text{S}$  (not shown). Regionally, the strength of sources and sinks differ considerably between models and between models and the data from Takahashi et al. (2002).

Most of the ocean models also show a similar pattern for the annual cycle of the air-sea  $\text{CO}_2$  flux: weak seasonality in the tropics, where there is outgassing throughout the year and pronounced seasonality in the higher latitudes (not shown). In most models, the amplitude of the annual cycle is larger than for the Takahashi et al. (2002) data-based fluxes. In the Northern Hemisphere, most models show  $\text{CO}_2$  uptake during winter and spring and  $\text{CO}_2$  release during summer and autumn everywhere north of  $10^\circ\text{--}20^\circ\text{N}$ . On the other hand, for the MPIM and NCAR models as well as the data-based estimates, there are two regimes in the Northern Hemisphere: in subtropical and temperate latitudes, the ocean loses  $\text{CO}_2$  in the summer and early autumn; north of  $50^\circ\text{N}$ , the oceans gains  $\text{CO}_2$  in summer and loses it in winter. South of  $30^\circ\text{S}$ , Takahashi et al. estimate that throughout the year the air-to-sea flux is positive or near zero; conversely, the LLNL and PRINCE models simulate stronger uptake and most of the other models simulate weak  $\text{CO}_2$  releases during the austral summer. In the Southern Hemisphere, the model that most resembles the Takahashi et al. fluxes is that from CSIRO.

The annual net air-to-sea  $\text{O}_2$  fluxes (Table 1) vary between  $+3 \text{ Tmol O}_2$  for the AWI model and  $-250 \text{ Tmol O}_2$  for NCAR. The mean value of all models is  $-121 \pm 83 \text{ Tmol } (\pm 1\sigma)$ . The large-scale spatial pattern of  $\text{O}_2$  sources and sinks (Fig. 2) shows some similar features in all the models (except for PIUB, which exhibits little structure). There are notable losses from the tropical ocean, particularly from the eastern equatorial Pacific and west of Africa. Furthermore, the models generally lose  $\text{O}_2$  to the atmosphere between  $40^\circ\text{S}$  and  $60^\circ\text{S}$ . The models simulate  $\text{O}_2$  uptake by the Southern Ocean, the North Atlantic and the Northwest Pacific.

Our comparison of the annual cycle of air-sea  $\text{O}_2$  fluxes offers a distinct advantage over the annual mean comparison above: it allows comparison of model results to data. Figure 3 shows the monthly mean oxygen flux anomalies for the OCMIP models and the observed oxygen fluxes from Garcia and Keeling (2001). Temporal patterns of simulated  $\text{O}_2$  uptake and release are similar. In the tropics, there is little seasonal variability. In the extratropics, there is stronger uptake during winter, when surface waters cool (solubility increases), biological  $\text{O}_2$  production is low and the mixed layer deepens, bringing  $\text{O}_2$ -depleted water to the surface. In contrast in summer,  $\text{O}_2$  is lost from the ocean when the mixed layer shallows (solubility de-

creases) and increased primary productivity results in net  $\text{O}_2$  production. The strength of the winter  $\text{O}_2$  sink varies among models. For example, in the temperate and sub-arctic Southern Hemisphere, SOC exhibits a strong oxygen sink during austral winter and a strong source during austral summer; the corresponding annual cycle in the LLNL and PRINCE models is much weaker. In the southern extratropics, the models with the strongest seasonality, SOC and UL, resemble most the Garcia and Keeling (2001) data-based  $\text{O}_2$  fluxes. In the Northern Hemisphere, however, all models predict weaker annual cycles than do Garcia and Keeling (2001).

## 5. Annual mean APO: results

Because the Battle et al. (2006) and Tohjima et al. (2005) data sets are given on different reference scales, they cannot be merged into a single data set. Therefore, Figs. 4 and 5 compare observed annual mean APO concentrations separately for each data set with simulated results from TM2 (panel A on each figure) and TM3 (panel B). Unfortunately, due to the lack of a station with a high quality long-term record that is far from APO sources and sinks, simulated and observed APO cannot be normalized to a common reference station. Thus for Figs. 4 and 5, we normalized APO by subtracting a constant offset—individually for each ocean model and the data—from the annual mean APO values at each station. For each model and the data, these offsets were chosen in a way that the sum of the annual mean APO values (the unweighted ‘global average’) equals zero.

All models as well as the observations show the same zonal structure in annual mean APO: There is a peak in APO close to or slightly south of the equator that is more pronounced in the eastern Pacific (Fig. 4) than in the western Pacific (Fig. 5) and there are lower APO concentrations in the higher latitudes. These annual-mean large-scale characteristics of APO have already been found in similar modeling studies (Stephens et al., 1998; Gruber et al., 2001; Battle et al., 2006) and are due to (1) the tropical sources of  $\text{O}_2$ ,  $\text{N}_2$  and  $\text{CO}_2$  that are most prominent in the eastern Pacific, (2) the strong APO sink in the Southern Ocean, (3) the corresponding yet weaker oceanic APO sink in the northern extratropics and (4) the fossil-fuel APO sink in mid-latitudes of the Northern Hemisphere. Although APO results from both transport models exhibit the same qualitative characteristics, the APO simulated in TM2 (panels A in Figs. 4 and 5) tends to be smoother than that for TM3 (panels B in the same figures), probably due to the coarser vertical resolution in TM2. Despite the overall qualitative agreement between models and observations, some features of the observed annual mean APO distribution are not captured by the simulations. The most notable differences are the observed high APO at SYO and the low APO at  $30^\circ\text{N}$  (Fig. 4) and the secondary APO maximum in the southern subtropics (Figs. 4 and 5).

A clearer comparison of model versus observed APO is possible in terms of differences between stations, an approach which

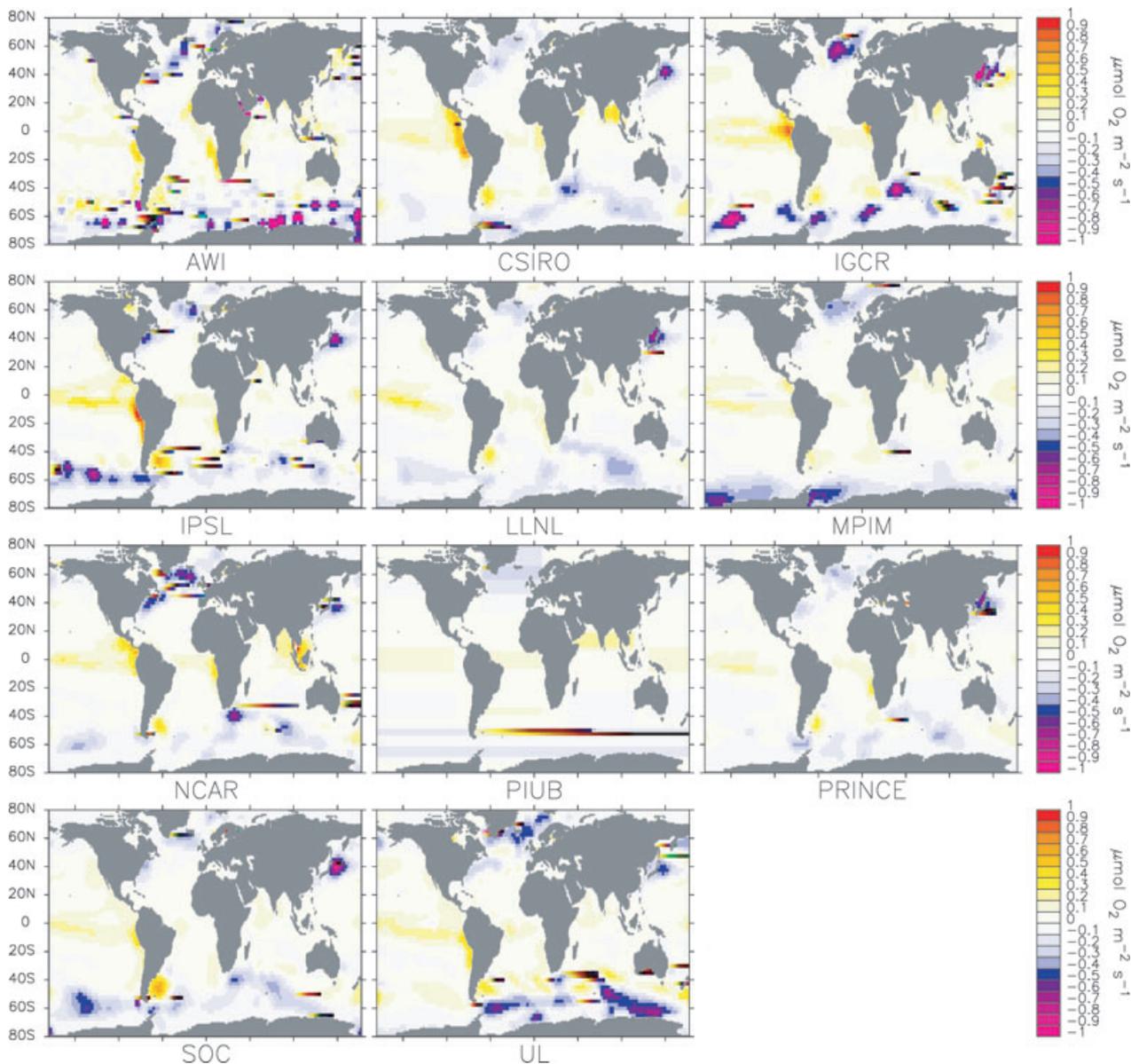


Fig. 2. Annual mean sea-to-air  $O_2$  flux as predicted by the OCMIP models. Units are in  $\mu\text{mol m}^{-2} \text{s}^{-1}$ .

is independent of the method of normalization. Thus, Fig. 6 compares observed versus modeled annual mean APO gradients between seven pairs of stations for simulations with both TM2 (sub-panels A) and TM3 (sub-panels B). Below, we describe results from each of these seven pairs of stations, in order from south to north. First, over Antarctica and the Southern Ocean (SPO-PSA) most ocean models predict relatively constant APO (Fig. 4). This homogeneous structure is more pronounced in TM3 than in TM2. Yet despite similar observed APO at SPO and PSA, there is a pronounced maximum of  $\approx 12$  per meg at SYO relative to the two neighboring stations. This varied structure is not captured by any of the models (Fig. 4). Second, across

the subpolar and temperate latitudes (PSA to CGO), TM2 and TM3 simulate contrasting APO gradients (panels 2a and 2b in Fig. 6). The TM3 simulations tend to overestimate the PSA-CGO gradient in APO, whereas the TM2 simulations generally underestimate it. Third, over the southern mid-latitudes to the southern subtropics (CGO to  $4.5^\circ\text{S}$ ), the ocean models generally capture observed trends, i.e. the APO maximum just south of the equator and the APO gradient between  $4.5^\circ\text{S}$  and CGO (Figs. 4 and 5 as well as panels 3a and 3b in Fig. 6). Yet the simulated equatorial bulge is less pronounced in TM2 than in TM3. Thus the APO simulated in TM2 tends to slightly underestimate the observed gradient between the southern subtropics

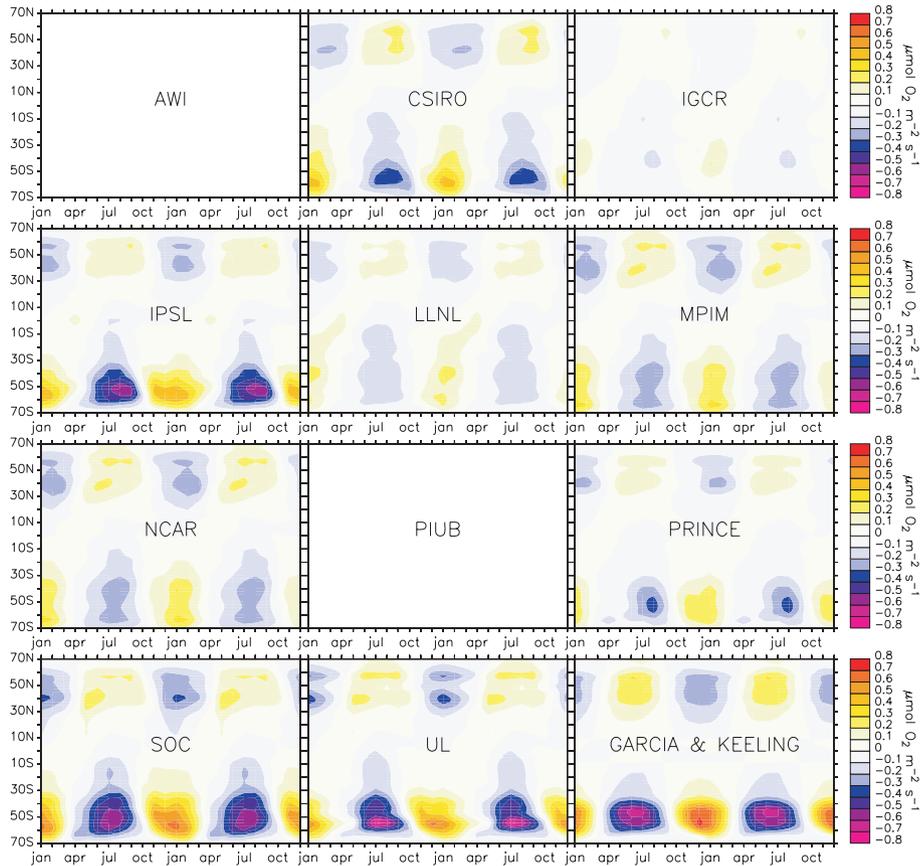


Fig. 3. Monthly anomalies of the sea-to-air  $O_2$  flux as estimated from data (Garcia and Keeling, 2001) and as simulated by the OCMIP models. Anomalies are given as the zonal mean minus annual zonal average sea-to-air  $O_2$  flux in  $\mu \text{ mol m}^{-2} \text{ s}^{-1}$  versus time. The annual mean models AWI and PIUB are not shown.

and mid-latitudes, whereas the TM3 results generally slightly overestimate it. These TM3 results contrast with those from the Battle et al. (2006) TM3 simulation, along with another set of air-sea fluxes, that underestimates the observed equatorial APO amplitude. We also find here that all OCMIP models predict a steady decrease from the tropical APO maximum toward CGO and thus do not capture the APO maxima observed at  $30^\circ\text{S}$  and AMS (Fig. 4). Fourth, in the northern tropics and subtropics, all models underestimate the observed gradient between  $4.5^\circ\text{N}$  and KUM (Fig. 6, panels 4a and 4b), which is consistent with results from Battle et al. (2006). Simulations with TM3 provide a more realistic gradient than do those with TM2. Although the gradient between  $4.5^\circ\text{S}$  and LJO is generally captured by most ocean models, none of those were able to match the observed low APO value at the shipboard station at  $30^\circ\text{N}$  (Fig. 4). Fifth, over the tropics between SMO and KUM, most ocean models properly simulate the observed gradient (Fig. 6, panels 5a and 5b). Sixth, across the northern mid to high latitudes (KUM-ALT), simulations with TM2 generally provide realistic gradients whereas those with TM3 slightly overestimate that observed gradient (panels 6a and 6b in Fig. 6). Once again then,

these TM3 results contrast with those from Battle et al. (2006), who simulated a KUM-ALT gradient that generally agreed with the data. Nonetheless, both TM2 and TM3, but especially the latter, exhibit a regional APO minimum at SAB and a regional APO maximum at CBA, neither of which has been observed (Fig. 6). In this case, our results generally agree with those from Battle et al. (2006), except we do not simulate a spurious APO minimum at LJO. Perhaps better agreement is obtained due to our shifted subsampling for LJO by  $2.5^\circ$  to the west (see Section 3.4). Finally, for the global north-south APO gradient (SPO-ALT), there are large differences between simulations: whereas the observed gradient is  $+2.6$  per meg, simulated differences range from  $+8.3$  per meg to  $-3.3$  per meg in TM2 ( $+8.1$  per meg to  $-6.6$  per meg in TM3) as shown in Fig. 6, panels 7a and b). For simulations with both TM2 and TM3, the CSIRO, MPIM, SOC and UL models predict SPO-ALT gradients that tend to agree with observations, the AWI and PIUB models predict strong negative SPO-ALT gradients and the IPSL, NCAR and PRINCE models strongly overestimate the observed gradient. Our range of simulated SPO-ALT gradients essentially bracket results from the modeling studies

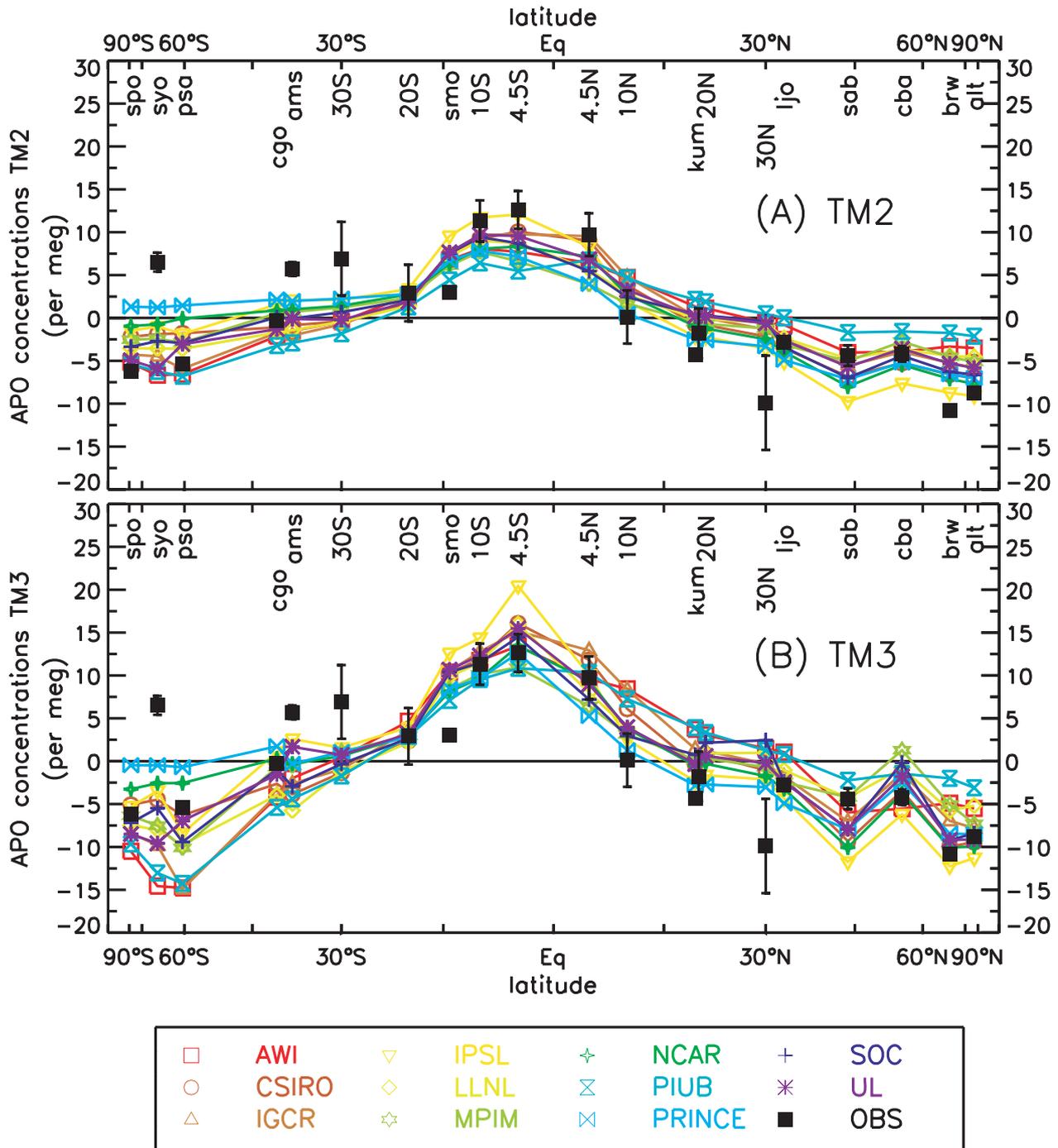


Fig. 4. Annual mean APO in per meg as observed (filled black squares, ship and land station data from Battle et al. (2006)) and simulated in (a) TM2 and (b) TM3 (colored symbols). See Fig. 1 for the exact location of the stations. APO was normalized so that the sum of all stations equals zero, i.e. for each model and for the data. The error bars indicate the uncertainties of the observed APO concentration as given by Battle et al. (2006).

from Stephens et al. (1998) (−1 to +4 per meg in TM2) and Battle et al. (2006) (+6 to +9 per meg in TM3). Only Gruber et al. (2001) simulate larger SPO-ALT gradients (using the GCTM atmospheric transport model) than given by our study.

### 6. Annual mean APO: discussion

Battle et al. (2006) have already pointed out that uncertainties in atmospheric transport complicate the validation of ocean models with APO observation. Here, we have found that simulated

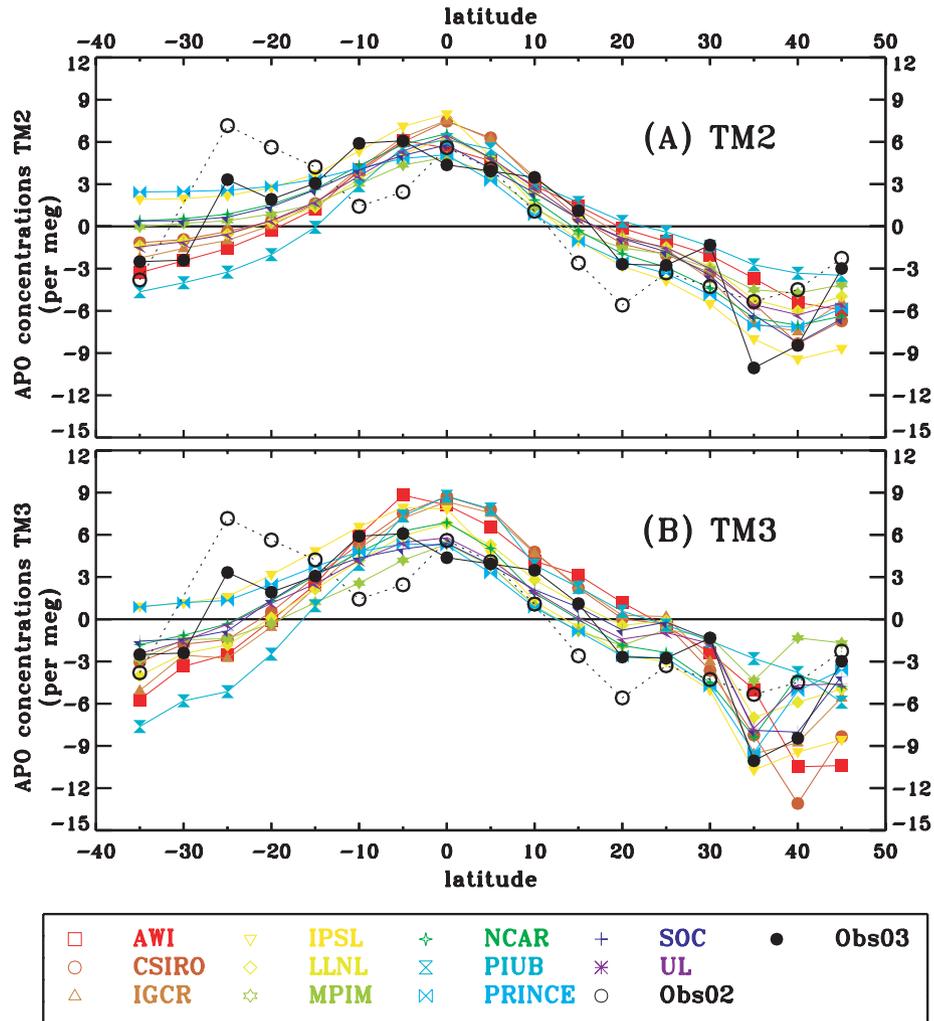


Fig. 5. Like Fig. 4, but for the ship transect data from the Western Pacific from Tohjima et al. (2005). The black open circles indicate the annual mean APO profile observed in 2002, whereas the filled black circles are the data for 2003. See Fig. 1 for the exact location of the stations.

APO gradients may depend as much on the atmospheric transport model as on the ocean model fluxes: Differences in APO gradients simulated in TM2 and TM3 with the same oceanic boundary conditions are generally similar or larger than the spread ( $\pm 1\sigma$ ) due to using different ocean model fluxes and are often larger than the deviations between observations and the models. The TM2-TM3 differences in simulated APO gradients do not show a distinct latitudinal pattern. However, they tend to be most pronounced in high latitudes and in the tropics compared to the subtropics and mid-latitudes. These large uncertainties due to atmospheric transport make it difficult to identify ocean model deficiencies from the data-model differences. For example, for the gradient between CGO and  $4.5^{\circ}\text{S}$ , with TM2 all ocean models overestimate the observed gradient (by up to 7.9 per meg), whereas with TM3 most ocean models underestimate this gradient (by up to 7.0 per meg) as shown in Fig. 6, panels 3a and 3b.

Furthermore, we used only two different atmospheric transport models in this study. Uncertainties due to atmospheric transport must be much larger. For example, for the north-south gradient in  $\text{CO}_2$  due to fossil fuel emissions, the TM2 and TM3 models fall in the middle of all the transport models that participated in the TransCom I study (Law et al., 1996; Bousquet, 1997), indicating that there is significant uncertainty in the inter-hemispheric exchange time and/or vertical mixing in the transport models. Thus, there is substantial uncertainty when just considering the fossil fuel component of the observed APO gradient between the northern and the Southern Hemisphere. Also the difference between TM2 and TM3 for the simulated terrestrial rectifier effect captures only about 50% of the full range of TransCom I models (Law et al., 1996; Bousquet, 1997), suggesting that TM2 and TM3 encompass a similar amount of uncertainty of the marine rectifier effect. Furthermore, the horizontal and vertical resolution as well as the parameterization of

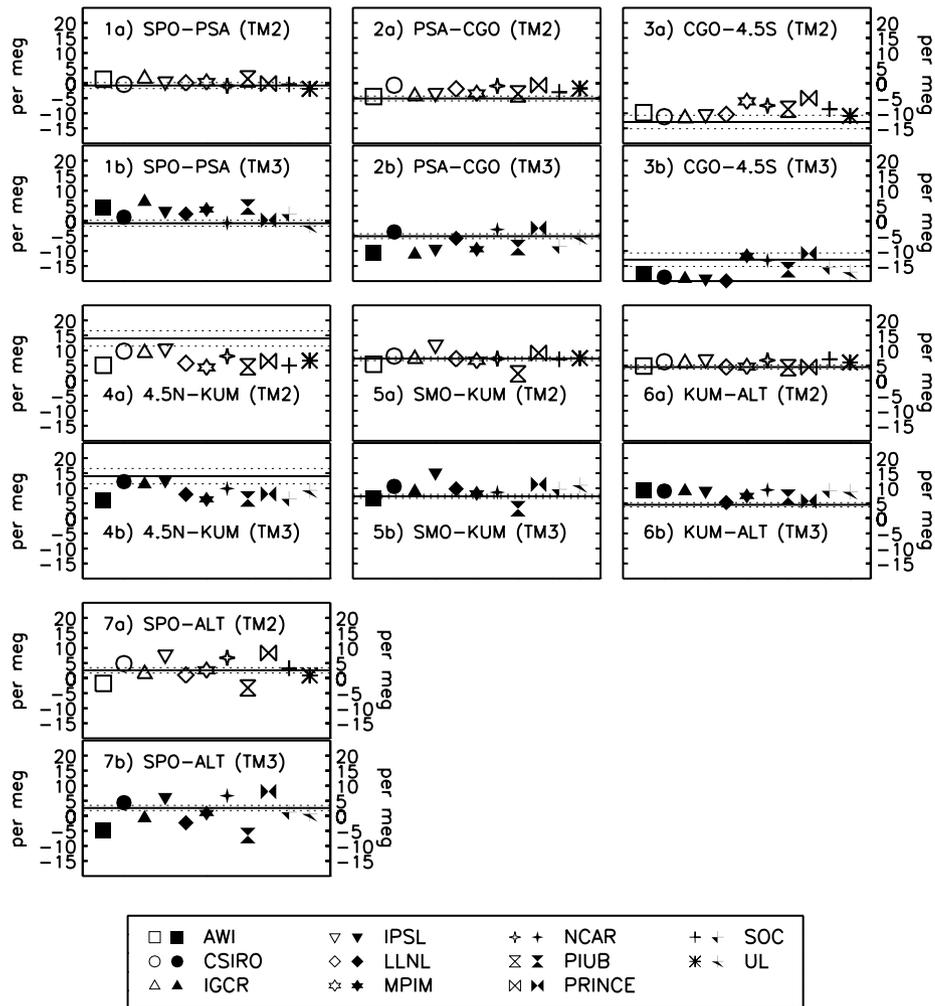


Fig. 6. APO differences between pairs of stations (station from the south minus that to the north) for the data (solid line, from Battle et al., 2006) and for the ocean models. TM2 results are shown as open symbols in sub-panels (a), the TM3 results as filled symbols in sub-panels (b). Observational  $\pm 1\sigma$ -error bars are indicated by the dashed lines and are calculated from the uncertainty of each annual mean APO concentration using standard error propagation.

subgrid-scale mixing processes in atmospheric transport models are far from perfect. Especially in the Southern Hemisphere, reanalyzed wind fields from both ECMWF and NCEP still suffer from data sparsity. Thus atmospheric transport models may well introduce systematic biases in the simulated APO concentrations that cannot be accounted for in our simple atmospheric transport model comparison. Thus even when TM2 and TM3 agree, we must be cautious when interpreting simulated and observed APO differences.

Differences between the observed and simulated global SPO-ALT gradient (Fig. 6, panels 7a and 7b) with some of the ocean models are qualitatively insensitive to the choice of the atmospheric transport model. Thus they may be due to ocean model deficiencies. Assuming that the deficiency in the SPO-ALT gradient Stephens et al. (1998) found in their study was indeed due to the ocean models, they provided four possible hypotheses:

- (1) poorly parameterized subgrid-scale eddy mixing,
- (2) poorly parameterized subgrid-scale vertical convection,
- (3) poorly simulated sea-ice formation in the Southern Ocean and
- (4) poorly represented thermohaline circulation in coarse-resolution ocean models.

Stephens et al. (1998) tested their first hypothesis with the LLOBM model, showing that including an isopycnal subgrid-scale eddy mixing scheme with the Gent and McWilliams parameterization does not substantially alter the simulated north–south APO gradient relative to a model with just the standard horizontal lateral mixing scheme. Aumont (1998) further showed that the discrepancies between simulated and observed APO gradients are relatively insensitive to the strength of convection in the Southern Ocean (Stephens’ 2nd hypothesis) as well as to different descriptions of sea ice formation and corresponding different rates of deep water formation (Stephens’ 3rd hypothesis).

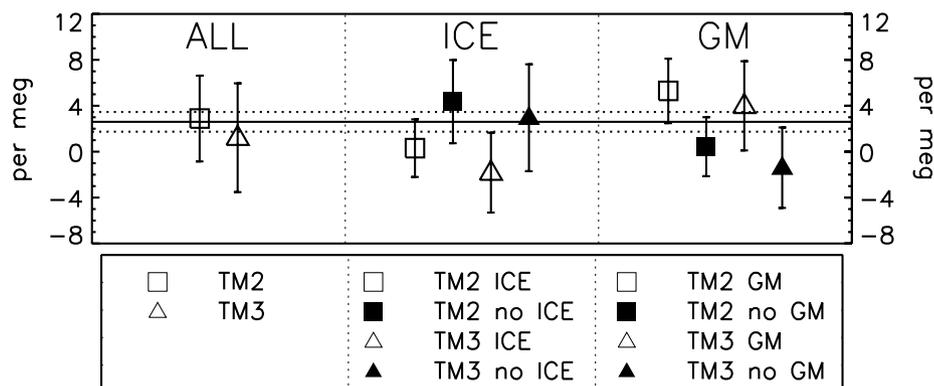


Fig. 7. Observed and average modeled APO gradients between SPO and ALT. The solid line indicates the observed gradient, the dashed lines the  $\pm 1\sigma$  uncertainty of the observed gradient. Panel ALL shows the simulated APO gradient averaged over *all* ocean models, transported either in TM2 (square) or TM3 (triangle). In panel ICE, we distinguish between the ocean models with an explicit sea-ice component module (LLNL, MPIM, PIUB and UL, open symbols) and the remaining models without a sea-ice module (filled symbols). Again, results are shown separately for the simulations in TM2 (squares) and TM3 (triangles). In panel GM, we show the average SPO-ALT gradient from ocean models with a Gent & McWilliams (GM) subgrid-scale mixing scheme (CSIRO, IPSL, LLNL, NCAR, PRINCE and SOC, open symbols) and from all models without GM (AWI, IGCR, MPIM, PIUB and UL, filled symbols). In all panels, the error bars indicate the spread of the group of ocean models discussed, i.e. the standard deviation of the simulated gradients.

Concerning Stephens' final hypothesis, Aumont (1998) also showed that ocean models with different tropical upwelling rates produced little difference in simulated APO gradients.

Further testing of these hypotheses is provided by our comparison of results from 11 ocean carbon cycle models. These models had different sea-ice formulations as well as different parameterizations of lateral subgrid-scale mixing. If the sea ice formulation were important to correctly simulate the global gradient between SPO and ALT, we would expect to see a substantial difference between the four OCMIP-2 models that had sea-ice modules (LLNL, MPIM, PIUB and UL) and the seven others that did not. Models with an explicit sea-ice module simulate SPO-ALT gradients that are on average 4–5 per meg lower than what is simulated by models without a sea-ice module. Surprisingly, TM3 yielded more realistic results when it was used in conjunction with air-sea fluxes from ocean models without an explicit sea-ice module rather than those that explicitly include sea-ice (Fig. 7, panel ICE). In TM2 though, neither group was particularly better in matching observed SPO-ALT APO gradients. Using an analogous approach, we also compared the APO results from the three models that had only a horizontal lateral mixing scheme (IGCR, MPIM and PIUB) to the six models that used an isopycnal scheme with the Gent and McWilliams parameterization (CSIRO, IPSL, LLNL, NCAR, PRINCE and SOC). One model, AWI, was excluded because it was an adjoint model with an isopycnal mixing scheme but not a GM parameterization. The modeled APO difference for each group of models is shown in Fig. 7 (panel GM). Simulated SPO-ALT gradients differ by about 5 per meg between both groups. However, there is no significant improvement of the global SPO-ALT gradient when employing the GM parameterization, neither in TM2 nor TM3: Models including a GM parameterization simulate

SPO-ALT gradients in APO which tend to be stronger than observed, whereas ocean models without the GM parameterization predict gradients that are too weak. Thus our APO-based analysis does not permit us to reveal features in ocean models that improve the simulation of large-scale  $O_2$  and  $CO_2$  fluxes.

Another potential problem when trying to compare models to observations is station representativeness. Most of the stations where APO is measured are located close to the ocean and might be subject to local oceanic sources of  $O_2$  or  $CO_2$  which are not resolved by the low-resolution ocean models. Furthermore, local meteorology at some stations might not be properly simulated by the atmospheric transport models. For example, the high observed annual mean APO concentration at SYO stands out relative to its neighboring stations (SPO, PSA) and it is not matched by any of the models. This deficiency suggests that local sources at SYO may mean that APO at SYO is not representative of APO concentrations typically found around 70°S. This suspicion is supported by the generally patchy  $O_2$  and  $CO_2$  source distribution in the Southern Ocean as simulated by the ocean models (see Fig. 2). Another possible explanation would be a strong local marine rectifier effect (i.e. a covariance of APO sinks and sources with atmospheric transport, Stephens et al. (1998)). However, such a rectifier term at SYO is not predicted by TM2 nor by TM3. In contrast to the high annual APO at SYO, the high APO seen around 30°S in the eastern Pacific and at AMS in the Indian Ocean (data from Battle et al., 2006, Fig. 4) appears to correspond with a similar maximum around 25°S in the western Pacific in the Tohjima et al. (2005) data set (Fig. 5), suggesting an APO source in the southern subtropics and adjacent mid-latitudes which is not predicted by any of the ocean models. In TM3, simulated APO overestimates observed values at CBA. There is also a tendency for TM3 to overestimate

the seasonal APO amplitude at the same station (Fig. 9). TM3 is known to predict a very strong marine rectifier effect in the Bering Strait region (Stephens et al., 1998 and T. Blaine, Scripps Institution of Oceanography, personal communication). Thus over the general area which includes CBA the regional rectifier effect in TM3 could also be too strong (Battle et al., 2006).

As a preliminary study, we first used the older APO data sets from Stephens et al. (1998) and Stephens (1999) that were based on shorter observation periods for  $O_2/N_2$  and  $CO_2$  and were also seriously affected by sampling problems (Manning, 2001; Battle et al., 2006). With that analysis, we found similar model-observation differences as Stephens et al. (1998) and Gruber et al. (2001). In particular, comparison to those observations suggested that simulated APO in high northern latitudes (ALT) was too large relative to high southern latitudes (PSA). Furthermore, in the sparse tropical APO data from Stephens et al. (1998) and Stephens (1999), there was no observational evidence of the APO maximum that was simulated in all the models. More recently, with the longer, fractionation-corrected and more comprehensive data set from Battle et al. (2006) along with the additional data from Tohjima et al. (2005), most of these data-model discrepancies have decreased or disappeared entirely. Hence this updated observational analysis reveals two other important issues when validating ocean models with APO observations: interannual variability and data quality. Whereas simulated APO in our study represents the climatological mean, the observed annual mean APO is based on a limited time series of observations with significant interannual variability (see our Fig. 5 for the Tohjima et al., 2005 data set and figure 6 in Battle et al., 2006). In other words, as pointed out by Battle et al. (2006) the agreement between models and recent data, collected between 1996 and 2003, is much better than agreement with older data from Stephens et al. (1998) and Stephens (1999) because ‘the gradient has evolved to match the model prediction’. Nonetheless, making high-quality oxygen measurements remains a challenge and there still remains a risk that some sample biases are not accounted for in the corrections applied to the data by Battle et al. (2006).

A number of other potential sources of error might contribute to the data-model mismatches. The representation of the common marine biogeochemical model (a diagnostic nutrient-restoring approach) used in the ocean models tested here was made as simple as possible during OCMIP-2 in order to facilitate model comparison. More realistic biogeochemical-ecosystem models are now in common use. Such models may produce different temporal and spatial patterns of biological productivity in the oceans, resulting in different  $CO_2$  and  $O_2$  flux patterns. Therefore, we may be relying on ocean model fluxes that have regional biases. One goal for future work would be to quantitatively assess the effects of including of more realistic ocean biogeochemical models. Other uncertainties that need to be assessed include the assumption of a globally constant Redfield

ratio, the air-sea gas exchange parameterization, the global average  $O_2:CO_2$  ratio for terrestrial photosynthesis and respiration (see new measurements from Seibt et al., 2004) and for fossil fuel emissions: Uncertainties in these  $O_2:CO_2$  ratios of the order of 5% translate into uncertainties in the simulated APO gradient between SPO and ALT of up to 3 per meg fossil-fuel component and up to 0.7 per meg for the terrestrial biosphere component, based on the annual mean  $CO_2$  gradients simulated in the TransCom I fossil fuel and neutral biosphere comparisons (Law et al., 1996). Furthermore, due to regionally different fossil fuel mixes and different  $O_2:CO_2$  ratios for different types of fossil fuel (Keeling, 1988),  $O_2:CO_2$  ratios of fossil emissions differ regionally from the global average value of  $-1.39$ . Whereas these deviations are small for most regions, China’s emissions have an average  $O_2:CO_2$  ratio of  $-1.1$  due to a large fraction of coal ( $O_2:CO_2 = -1.17$ ) and cement production ( $O_2:CO_2 = 0.00$ ). Thus, when assuming a globally constant  $O_2:CO_2$  ratio of  $-1.39$ , the APO fossil fuel component downwind of Chinese emissions will be overestimated. However, we assume that at these sites, the Chinese fossil  $CO_2$  emissions are already well diluted by background air. Thus the potential for significantly overestimating the fossil fuel component in the simulated APO appears small compared to other uncertainties in our study. The other stations used in this study are either located at remote sites having little regional fossil  $CO_2$  influence or at sites that are influenced by fossil  $CO_2$  emissions with  $O_2:CO_2$  ratios close to the global average. Therefore, we conclude that regional deviations from the global average stoichiometric fossil fuel  $O_2:CO_2$  ratio do not introduce significant errors in our study, a finding that is corroborated by a similar estimates from Battle et al. (2006) concerning their analysis. The change in the global average fossil fuel  $O_2:CO_2$  ratio due to changes in the global mix of fossil fuels is less than 1% in the 1990s and thus appears negligible compared to other factors contributing to uncertainties in simulated APO.

## 7. APO seasonal cycles: results

The AWI and PIUB models are both nonseasonal and thus the small the simulated seasonal variability in APO simulated with these two models is only due to seasonality in atmospheric transport (Fig. 8). Hence further discussion is focused only on the semi- and fully-seasonal models.

Observed seasonal APO amplitudes exhibit a distinct latitudinal pattern (Fig. 9a), with small amplitudes in the tropics and subtropics and larger amplitudes in the mid-to-high latitudes of both hemispheres. At the extratropical stations, the observed APO amplitudes are slightly smaller in the Northern Hemisphere than at corresponding latitudes in the south. The average of all models tends to slightly under-predict the seasonal APO amplitude at almost all stations (open circles in Fig. 9a). However, the spread between the ocean models (right error bars in Fig. 9a) is

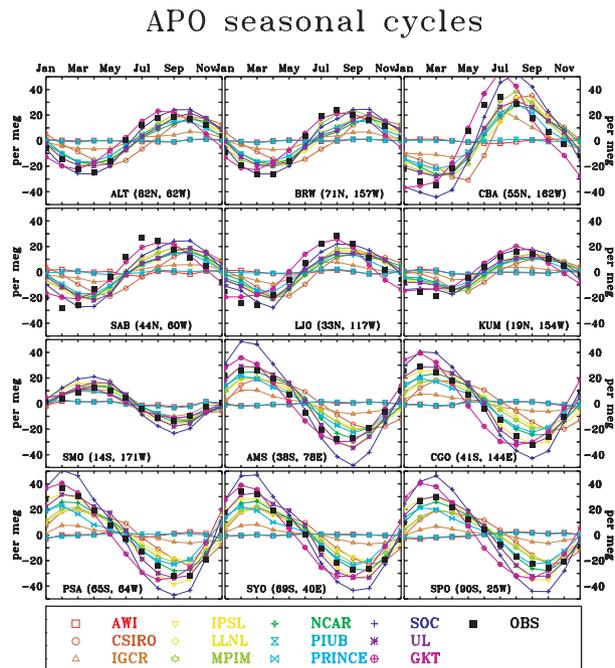


Fig. 8. Detrended seasonal cycle of APO for the data (filled black squares) and models (colored symbols, given as the TM2-TM3 average) at the SPO, SYO, CGO, AMS, SMO, KUM, LJO, NWR, CBA, BRW and ALT observation stations. Differences in the APO amplitude between TM2 and TM3 are discussed in Fig. 9. Phase differences between APO in TM2 and in TM3 are negligible. GKT denotes simulated APO based on observed  $O_2$  and  $N_2$  flux anomalies from Garcia and Keeling (2001) and  $CO_2$  flux anomalies calculated from data from Takahashi et al. (2002).

large, especially in the southern extratropics and at CBA. The APO simulated in TM2 generally exhibits smaller APO amplitudes than simulation results from TM3. The TM2–TM3 differences in the simulated APO amplitudes (left error bars in Fig. 9a) are generally small in the tropics and subtropics (10–20% of the average of the TM2 and TM3 amplitudes) but increase up to 40% toward the mid-to-high latitudes of both hemispheres. The average TM2–TM3 difference in the seasonal APO amplitude is generally larger than or equal to the mean bias between models and observations.

Figure 9b shows normalized APO amplitudes for the TM2 atmospheric transport model combined with each of the OCMIP models as well as data-based GKT fluxes. Normalized APO amplitudes were calculated by dividing the simulated APO amplitude for each model and each station by the observed amplitude at that respective station. Generally, seasonal amplitudes in APO simulated with TM2 tend to under-predict the observed values (i.e. the normalized amplitudes are always smaller than 1). Only the APO amplitudes from GKT generally match the observations. Although the SOC model's APO amplitudes match the observations in the Northern Hemisphere, they over-predict the APO seasonality in the south. In contrast, IPSL and UL perform

well in the Southern Hemisphere, but under-predict APO amplitudes in the north. For each ocean model, normalized APO amplitudes at SMO tend to be higher than elsewhere. All stations show a generally consistent rank of the models relative to their agreement with the data. Figure 9c shows corresponding results for TM3. Because seasonal APO amplitudes in TM3 are generally larger than those in TM2, especially in the Southern Hemisphere and at CBA, models which perform well in TM2 tend to over-predict APO amplitudes in TM3 and models which underestimated the APO seasonality in TM2 agree better with the observations when their fluxes are used instead with TM3.

The observed phasing of the seasonal cycle of APO is generally captured within a 1-month offset by the ocean models (Fig. 8). Differences between TM2 and TM3 in the simulated phase of APO are small. Simulated APO from the ocean model fluxes almost always lags the observed phase at all stations. At SAB, the simulated APO maximum lags the observations by two months for all ocean models. Conversely, the phasing of the APO minimum from the GKT fluxes is generally too early, on average by half a month.

## 8. APO seasonal cycles: discussion

Qualitatively, the latitudinal dependence of the observed APO amplitude is due to largely to seasonality of the  $O_2$  flux as well as the relative amount of ocean to land in a given station's latitudinal belt. The amplitude of the seasonal  $O_2$  flux in the southern extratropics is largest between  $40^\circ S$  and  $60^\circ S$  (Fig. 3), where we also find the largest seasonal APO amplitudes. South of  $60^\circ S$ , amplitudes are smaller because of the 'dilution' effect of the large Antarctic land mass. In the tropics and subtropics, the small seasonal amplitude in APO is consistent with small seasonal variability in air-sea  $O_2$  fluxes. As expected, the northern extratropics have a low seasonal amplitude in APO, relative to the Southern Hemisphere, because of the high land to ocean ratio. However, at one Northern Hemisphere station, CBA, there is a large amplitude, which is probably caused by a seasonal upwelling event and subsequent bloom that provokes an  $O_2$  efflux from the Bering Sea in summer (Battle et al., 2006).

At first glance, the relatively good agreement between simulated and observed APO seasonal cycles is surprising. The diagnostic nutrient restoring approach employed by the ocean models is now known to offer an inadequate representation of the dynamics for the ocean biological component (Six and Maier-Reimer, 1996). In particular in high latitudes, this parameterization does not capture the observed draw-down of surface  $pCO_2$  during the spring bloom in terms of the amplitude, phase, or even the sign of the  $pCO_2$  change. Thus, simulated seasonal variability of biologically driven  $CO_2$  and  $O_2$  fluxes is unrealistic. However, in regions where most of the seasonality in the  $O_2$  fluxes is driven by seasonally varying heat fluxes (and thus varying  $O_2$  solubility), there is relatively little concern.

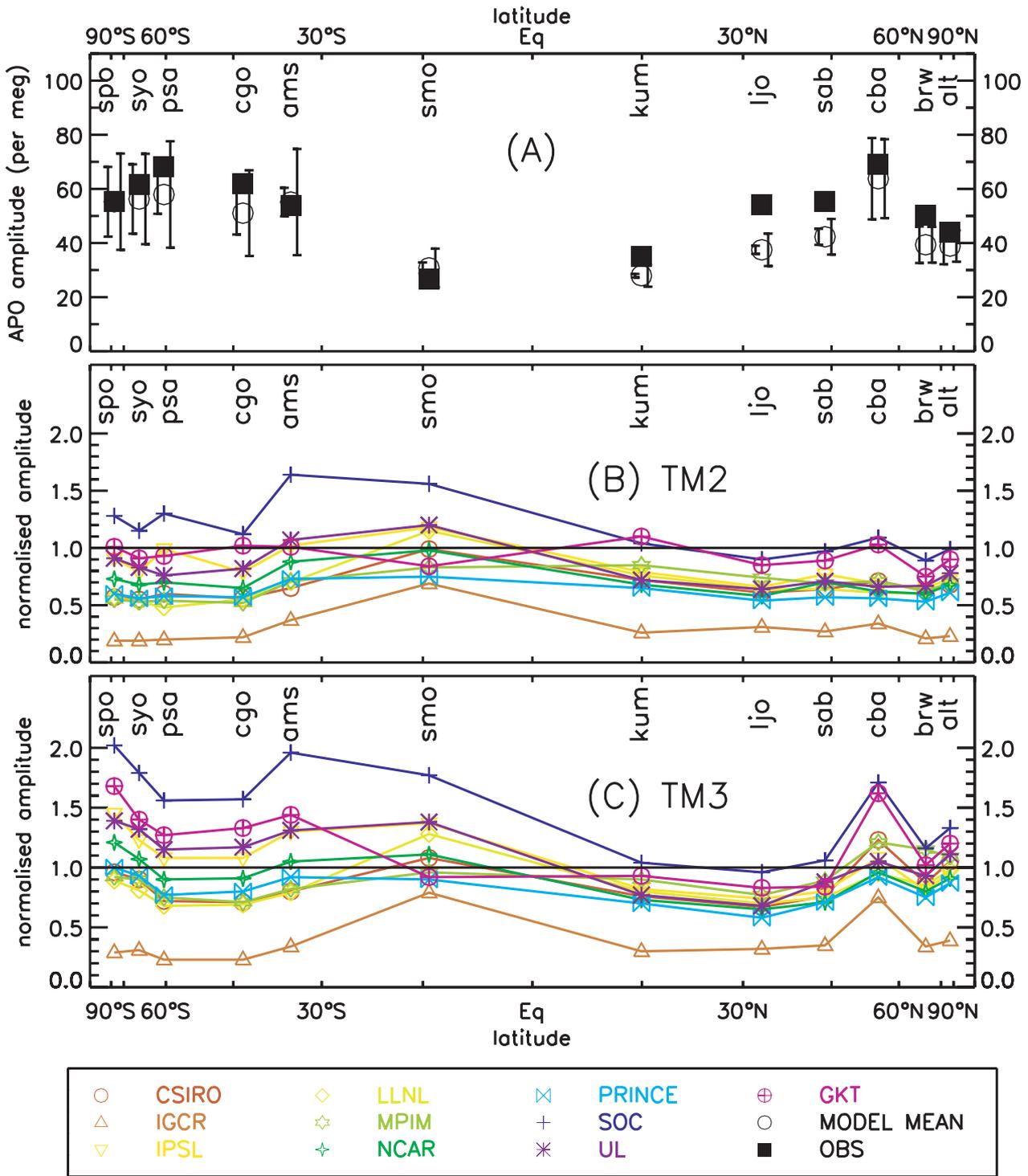


Fig. 9. Panel (a): Observed amplitude (filled squares) and average model amplitude (open circles, average over all fully seasonal ocean models and both transport models). The full length of the error bars on the left indicate the difference between the two atmospheric transport models. The error bars on the right indicate the spread amongst ocean models ( $\pm 1\sigma$ ) for the average of the amplitude from TM2 and TM3. Panel (b): Normalized APO amplitude (TM2 results) versus latitude. The normalized amplitude is calculated as the modeled peak-to-peak amplitude divided by the observed amplitude. Panel (c): As Panel (b), but for TM3.

In view of the large differences between APO seasonal amplitudes simulated with TM2 and TM3, it is difficult to thoroughly validate ocean models by comparing the simulated to the observed seasonal cycle in APO. Nonetheless, there are numerous possible reasons why most ocean model flux fields in combination with TM2 and many as well in combination with TM3, underestimate of the APO seasonal amplitude. First, inadequacies in the common ocean biogeochemical module used by the OCMIP-2 models may be partly responsible: As discussed above, the diagnostic nutrient restoring approach produces an inadequate seasonal  $O_2$  flux variability. Second, forcing the atmospheric transport models with only monthly mean fluxes of  $O_2$ ,  $CO_2$  and  $N_2$  may not be adequate. The same may also be true for the standard monthly mean dynamic forcing in the ocean models.

As another explanation, atmospheric transport models may overestimate the vertical mixing in the atmosphere, thereby resulting in a faster dilution of the seasonal flux signal into the troposphere and hence an underestimation of the seasonal amplitude of APO at the surface. Relative to TM2, the TM3 model's larger seasonal amplitude in APO in polar and subpolar regions may be partly explained by the higher vertical resolution in TM3 and the resulting weaker vertical mixing. Even with TM3 though, simulated seasonal amplitudes in APO remain smaller than observed for fluxes from many of the ocean models. Further improvements in vertical resolution and in the representation of subgrid-scale vertical mixing in atmospheric transport models could help to close the gap.

Although the three problems mentioned above (oversimplified ocean biogeochemistry, monthly restoring and atmospheric transport) may all contribute to model-data discrepancies, they cannot be invoked to explain model-model differences. Those must derive from different seasonal dynamics among the ocean models (e.g. mixed layer depth, upwelling and convection).

Simulated APO from data-based GKT fluxes are affected by uncertainties in atmospheric transport, as are the ocean model fluxes, but the GKT fluxes also have other shortcomings that might result in simulated APO that overestimates the seasonal amplitude, as seen in the TM3 results in the southern extratropics. For example, the data used to derive the GKT flux fields is severely limited in the Southern Ocean, particularly in winter. Also, Garcia and Keeling (2001) neglected possible seasonal variability of the coefficients in their regression of air-sea  $O_2$  flux anomalies with heat flux anomalies. Nonetheless the GKT-derived APO provides an invaluable data-based perspective that can help resolve the causes for differences between the models and the observed seasonal cycle of APO. For example, in contrast with the GKT-derived APO, all ocean models simulate an APO with larger seasonal amplitude at SMO than at neighboring stations. Thus it seems unlikely that the atmospheric transport models are responsible for this discrepancy. Even the semi-seasonal IGCRC model captures nearly 60% of the amplitude at SMO (in other ocean areas only 20–30%). This suggests that seasonality

in ocean biological new production is the driving force behind the APO amplitude around SMO. In terms of the phasing of the APO seasonal cycle, the OCMIP ocean model fluxes tend to produce a lagged response, by a few weeks, because simulated ocean phosphate, SST and salinity are restored to the observed fields with a 30-d time constant.

## 9. Summary and conclusions

We tested 11 ocean carbon cycle models, all of which participated in the Ocean Carbon-Cycle Model Intercomparison Project OCMIP, by using their air-sea fluxes of  $O_2$ ,  $N_2$  and  $CO_2$  as boundary conditions in two different 3-D atmospheric transport models, TM2 and TM3. We compared the resulting simulated APO gradients and seasonal cycles to those of observed APO. For most models, the meridional structure of the simulated annual mean APO concentrations resembles the observations qualitatively, i.e. the models predict an equatorial APO maximum and low APO values in the high latitudes of both hemispheres. However, the level of data-model agreement for large-scale gradients in APO depends on which ocean and atmospheric transport models are used. Despite qualitative agreement, there was consistent failure to properly simulate some small scale features in the observed APO. When using the ocean model flux fields with the TM2 atmospheric transport model, the global APO distribution was generally smoother than when they were used with the TM3 atmospheric transport model. The latter predicts steeper gradients between stations as well as larger seasonal APO amplitudes, probably due to the finer vertical resolution of TM3 compared to TM2. In the Northern Hemisphere, the simulated APO generally underestimates the observed amplitude. In the Southern Hemisphere with TM3, some ocean models simulate seasonal APO amplitudes that are larger than the observed values.

Despite qualitative agreement between modeled and observed annual mean APO concentrations, there are a number of difficulties that impede thorough validation of ocean carbon cycle models with APO observations. For example, atmospheric transport models and the reanalyzed winds used to force them are not without errors. We have made a preliminary attempt to address this problem by transporting the same set of ocean fluxes in two different atmospheric transport models with two different wind fields. The range of the results provides only a lower limit for the associated atmospheric transport uncertainty. However, in many cases, the atmospheric transport uncertainties are as large as or even larger than average data-model differences as well as the spread between different ocean models. Thus it remains difficult to assign data-model discrepancies to either ocean or atmospheric transport model deficiencies. Furthermore, we had to assume that APO observations at individual stations are directly comparable to the regional-scale APO concentrations as simulated in the low-resolution atmospheric models, but this might not always be the case. Additionally, interannual variability in APO is not negligible, making it difficult to validate

climatologically forced ocean models. For illustration, one can compare the older APO data set (sampling period ca. 1989–1998), originally from Stephens et al. (1998) and updated by Stephens (1999), with the more recent data set from Battle et al. (2006) (sampling period 1996–2003): When we used the old APO data, we found some severe mismatches between modeled and observed global APO gradients as found by Stephens et al. (1998), Aumont et al. (1999) and Gruber et al. (2001). However, with the more recent data set which is more comprehensive and corrected for fractionation, much of the data-model mismatch went away.

Remaining data-model discrepancies in the annual mean APO concentrations may be due to several causes: coarse ocean model resolution, inadequacies in simulated ocean transport and mixing, the overly simple diagnostic ocean biogeochemical model and uncertainties in the air-sea gas exchange parameterization as well as atmospheric transport. The tendency for simulated APO to underestimate the observed seasonal amplitude may be either in the atmospheric transport model or the ocean models or both. Atmospheric transport models may underestimate seasonal variability because they are forced with monthly mean ocean fluxes or because their vertical mixing may be too vigorous. Ocean models may underestimate seasonal variability because of deficiencies in dynamical or biogeochemical components. The phasing of the simulated APO derived from the ocean model fluxes generally lags that observed, but usually by less than a month. Such lags may be due to the ocean-model restoring of simulated phosphate, temperature and salinity toward observations.

Our analysis has shown that for the time being, APO observations do not offer a rigorous means to evaluate ocean models simply because their deficiencies cannot be distinguished from other causes of APO data-model discrepancies. To improve this situation, we have three recommendations. First, there is a need to account for interannual variability of the global carbon and oxygen cycles. This calls for still longer time series of observed APO as well as using newly available simulations that account for interannual variability in atmospheric and ocean dynamics as well as ocean biogeochemistry (air-sea O<sub>2</sub> and CO<sub>2</sub> fluxes). Second, there is a clear need to improve atmospheric transport models. Higher vertical and horizontal resolution will probably be required and vertical transport should be tested extensively and refined. For example, it would be helpful to compare simulated and observed vertical profiles of tracers in regions where it is considered that terrestrial and marine sources and sinks are relatively well known. Third, there is a need to further improve the network of APO measurement stations. Although sensitive to uncertainties in the O<sub>2</sub>:CO<sub>2</sub> ratio of photosynthesis and respiration, stations in the continental interior—far away from the oceanic sources and sinks of APO—are likely to be more representative of the large-scale distribution of APO relative to marine or coastal stations. Furthermore, existing and future APO measurement locations might be further optimized to be more representative of the spatial scale of APO concentrations as simulated

by the atmospheric models by closer analysis of meteorological data and the use of simulated retro-trajectories to pinpoint the origin of air masses.

## 10. Acknowledgements

We thank the OCMIP modelers R. Najjar, R. Schlitzer, R. Matear, A. Ishida, Y. Yamanaka, K. Caldeira, M. Wickett, E. Maier-Reimer, K. Lindsay, S. Doney, G.-K. Plattner, F. Joos, R. D. Slater, J. L. Sarmiento, A. Gnanadesikan, N. Gruber, I. Totterdell, A. Yool and A. Mouchet for having provided us with their model output. We also thank G.-K. Plattner for helping regrid of the PIUB results and P. Bousquet and P. Peylin for helping us with the TM2 model. We are grateful to M. Battle, S. Mikaloff Fletcher, M. L. Bender, R. Keeling, A. C. Manning, N. Gruber, P. P. Tans, M. B. Hendricks, D. T. Ho, C. Simonds, R. Mika, B. Paplawsky, Y. Tohjima, H. Mukai, T. Machida, Y. Nojiri and M. Gloor for sharing their data, as well as to R. Keeling and B. Stephens for helpful discussions and to two anonymous reviewers, whose comments considerably helped to improve the manuscript. We have also benefited from many discussions with M. Cuntz. Financial support came from IGBP/GAIM as well as the EU GOSAC Project (Contract ENV-CT97-0495). Computer time was provided by the French CEA (Grenoble).

## References

- Anderson, L. A. and Sarmiento, J. L. 1994. Redfield ratios of remineralization determined by nutrient data analysis. *Global Biogeochem. Cycl.* **8**, 65–80.
- Anderson, L. A. and Sarmiento, J. L. 1995. Global ocean phosphate and oxygen simulations. *Global Biogeochem. Cycl.* **9**, 621–636.
- Andres, R. J., Marland, G., Fung, I. and Matthews, E. 1996. A 1 × 1 distribution of carbon dioxide emissions from fossil fuel consumption and cement manufacture, 1950–1990. *Global Biogeochem. Cycl.* **10**(3), 419–429.
- Aumont, O. 1998. *Etude du cycle naturel du carbone dans un modèle 3D de l'océan mondial*. PhD thesis, Univ. Paris VI, Paris. June.
- Aumont, O., Orr, J. C., Monfray, P., Madec, G. and Maier-Reimer, E. 1999. Nutrient trapping in the equatorial Pacific: The ocean circulation solution. *Global Biogeochem. Cycl.* **13**(2), 351–369.
- Balkanski, Y., Monfray, P., Battle, M. and Heimann, M. 1999. Ocean primary production derived from satellite data: an evaluation with atmospheric oxygen measurements. *Global Biogeochem. Cycl.* **13**, 257–271.
- Battle, M., Fletcher, S. M., Bender, M., Keeling, R. F., Manning, A. C., and co-authors. 2006. Atmospheric potential oxygen: new observations and their implications for some atmospheric and oceanic models. *Global Biogeochem. Cycl.* **20** (GB1010), doi:10.1029/2005GB002534.
- Bousquet, P. 1997. *Optimisation des flux nets de CO<sub>2</sub>: assimilation des mesures atmosphériques de CO<sub>2</sub> et de <sup>13</sup>CO<sub>2</sub> dans un modèle de transport tridimensionnel*. PhD thesis, Université Paris VI, Paris.
- Boutin, J. and Etcheto, J. 1996. Consistency of Geosat, SSM/I and ERS1 global surface wind speeds; comparison with in-situ data. *J. Atmos. Oceanic Technol.* **13**, 183–197.

- Ciais, P., Tans, P. P., Trolier, M., White, J. W. and Francey, R. 1995. A large Northern Hemisphere terrestrial CO<sub>2</sub> sink indicated by the <sup>13</sup>C/<sup>12</sup>C ratio of atmospheric CO<sub>2</sub>. *Science* **269**, 1098–1102.
- Denning, A. S., Fung, I. Y. and Randall, D. A. 1995. Latitudinal gradient of atmospheric CO<sub>2</sub> due to seasonal exchange with land biota. *Nature* **376**, 240–243.
- Doney, S. and Hecht, M. 2002. Antarctic bottom water formation and deep water chlorofluorocarbon distributions in a global ocean climate model. *J. Phys. Oceanogr.* **32**(6), 1642–1666.
- Duffy, P. B., Caldeira, K., Selvaggi, J. and Hoffert, M. I. 1997. Effects of subgrid-scale mixing parameterizations on simulated distribution of natural <sup>14</sup>C, temperature and salinity in a three-dimensional ocean general circulation model. *J. Phys. Oceanogr.* **27**(4), 498–523.
- Garcia, H. and Keeling, R. 2001. On the global oxygen anomaly and air-sea flux. *J. Geophys. Res.* **106**(C12), 31 155–31 167.
- GlobalView, 2003. Cooperative Atmospheric Data Integration Project—Carbon Dioxide. Technical report, NOAA CMDL, Boulder, CO, USA, 2003. available via anonymous FTP: ftp.cmdl.noaa.gov, path: ccg/CO2/GLOBALVIEW.
- Gnanadesikan, A., Slater, R. D., Gruber, N. and Sarmiento, J. L. 2002. Oceanic vertical exchange and new production: a comparison between models and data. *Deep-Sea Research II* **2**, 363–401.
- Goodberlet, M. A., Swift, C. T. and Wilkerson, J. C. 1989. Remote sensing of ocean surface winds with the special sensor microwave/Imager. *J. Geophys. Res.* **94**, 14 571–15 555.
- Goosse, H., Deleersnijder, E., Fichefet, T. and England, M. H. 1998. Sensitivity of a global coupled ocean-sea ice model to the parameterization of vertical mixing. *J. Geophys. Res.* **104**(C6), 13 681–13 695.
- Gordon, C. 2000. The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dyn.* **16**, 147–168.
- Gruber, N., Gloor, E., Fan, S.-M. and Sarmiento, J. L. 2001. Air-sea flux of oxygen estimated from bulk data: implications for the marine and atmospheric oxygen cycles. *Global Biogeochem. Cycl.* **15**(4), 783–803.
- Gurney, K. R., Law, R. M., Denning, A. S., Rayner, P. J., Baker, D. and co-authors. 2002. Towards robust regional estimates of CO<sub>2</sub> sources and sinks using atmospheric transport models. *Nature* **415**, 626–630.
- Heimann, M. 1995. The global atmospheric tracer model TM2. Technical Report 10 (ISSN 0940-9327), Deutsches Klimarechenzentrum, Hamburg.
- Heimann, M. and Körner, S. 2003. The global atmospheric tracer model TM3. Technical report No 5, Max-Planck-Institut für Biogeochemie, Jena, Germany.
- Heimann, M., Esser, G., Haxeltine, A., Kaduk, J., Kicklighter, D. W., and co-authors. 1998. Evaluation of terrestrial carbon cycle models through simulations of the seasonal cycle of atmospheric CO<sub>2</sub>: first results of a model intercomparison study. *Global Biogeochem. Cycl.* **12**(1), 1–24.
- Keeling, R. F. 1988. *Development of an interferometric oxygen analyzer for precise measurements of the atmospheric O<sub>2</sub> mole fraction*. Phd thesis, Harvard University.
- Keeling, R. F. and Shertz, S. R. 1992. Seasonal and interannual variations in atmospheric oxygen and implications for the global carbon cycle. *Nature* **358**, 723–727.
- Keeling, R. F., Najjar, R. P., Bender, M. L. and Tans, P. P. 1993. What atmospheric oxygen measurements can tell us about the global carbon cycle. *Global Biogeochem. Cycl.* **7**(1), 37–67.
- Keeling, R. F., Stephens, B. B., Najjar, R. G., Doney, S. C., Archer, D. and co-authors. 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 Biogeochem. Cycl.* **12**, 141–163.
- Law, R. M., Rayner, P. J., Denning, A. S., Erickson, D., Piper, S. C., and co-authors. 1996. Variations in modelled atmospheric transport of carbon dioxide and the consequences for CO<sub>2</sub> inversions. *Global Biogeochem. Cycl.* **10**, 783–796.
- Le Quééré, C., Orr, J. C., Monfray, P., Aumont, O. and Madec, G. 2000. Interannual variability of the oceanic sink of CO<sub>2</sub> from 1979 to 1997. *Global Biogeochem. Cycl.* **14**(4), 1247–1265.
- Louanchi, F. and Najjar, R. G. 2000. A global monthly climatology of phosphate, nitrate and silicate in the upper ocean: spring-summer export production and shallow remineralization. *Global Biogeochem. Cycl.* **14**, 957–977.
- Louis, J. F. 1979. A parametric model of vertical eddy fluxes in the atmosphere. *Boundary Layer Meteorology* **17**, 187–202.
- Madec, G., Delecluse, P., Imbard, M. and Lévy, C. 1998. OPA Version 8.1-Ocean General Circulation model. Reference manual, IPSL, Paris, France.
- Maier-Reimer, E., Mikolajewicz, U. and Hasselmann, K. 1993. Mean circulation of the Hamburg LSG OGCM and its sensitivity on the thermohaline forcing. *J. Phys. Oceanogr.* **23**, 731–757.
- Manning, A. C. 2001. *Temporal variability of atmospheric oxygen from both continuous measurements and a flask sampling network: tools for studying the global carbon cycle*. Phd thesis, University of California, San Diego.
- Marland, G., Boden, T. A. and Andres, R. J. 2005. Global, regional and national CO<sub>2</sub> emissions. In: *Trends: A Compendium of Data on Global Change*. Carbon Dioxide Information Analysis Center, Oak Ridge, TN.
- Matear, R. J. and Hirst, A. C. 1999. Climate Change Feedback on the Future Oceanic CO<sub>2</sub> uptake. *Tellus*, **51B**, 722–733.
- Najjar, R. and Orr, J. C. 1999. *Biotic-HOWTO. Internal OCMIP Report*. LSCE/CEA Saclay, Gif-sur-Yvette, France, 15 pp.
- Najjar, R. G., Sarmiento, J. L. and Toggweiler, J. R. 1992. Downward transport and fate of organic matter in the ocean: simulations with a general circulation model. *Global Biogeochem. Cycl.* **6**(1), 45–76.
- Orr, J. C., Maier-Reimer, E., Mikolajewicz, U., Monfray, P., Sarmiento, J. L. and co-authors. 2001. Estimates of anthropogenic carbon uptake from four three-dimensional global ocean models. *Global Biogeochem. Cycl.* **15**, 43–59.
- Ramonet, M., Rouilly, J. C., Bousquet, P. and Monfray, P. 1996. Radon-222 measurements during the TROPOZ II campaign and comparison with a global atmospheric transport model. *J. Atmos. Chem.* **23**, 107–136.
- Rayner, P. J. and Law, R. M. 1995. A comparison of modelled responses to prescribed CO<sub>2</sub> sources. Tech. Paper 88, CSIRO Div. of Atmos. Res., Melbourne.
- Russell, G. L. and Lerner, J. A. 1981. A new finite-differencing scheme for the tracer transport equation. *J. Appl. Meteor.* **20**, 1483–1498.
- Schlitzer, R. 2000. Applying the adjoint method for global biogeochemical modeling. In: *Inverse Methods in Global Biogeochemical Cycles* (eds. Kasibhatla, P. Heimann, M. Hartley, D. Mahowald, N. Prinn R., and Rayner P.), Geophys. Monograph Series Vol. 114, Washington, DC, AGU.

- Seibt, U., Brand, W. A., Heimann, M., Lloyd, J., Severinghaus, J. P. and co-authors. 2004. Observations of O<sub>2</sub>:CO<sub>2</sub> exchange ratios during ecosystem gas exchange. *Global Biogeochem. Cycl.* **18**(GB4024), doi:2004GB002242.
- Severinghaus, J. P. 1995. *Studies of the terrestrial O<sub>2</sub> and carbon cycles in sand dune gases and in Biosphere 2*. PhD thesis, Columbia Univ.
- Six, K. D. and Maier-Reimer, E. 1996. Effects of plankton dynamics on seasonal carbon fluxes in an ocean general circulation model. *Global Biogeochem. Cycl.* **10**, 559–583.
- Stephens, B. B. 1999. *Field-based atmospheric oxygen measurements and the ocean carbon cycle*. PhD thesis, Univ. of California, San Diego.
- Stephens, B. B., Keeling, R. F., Heimann, M., Six, K. D., Murnane, R. and co-authors. 1998. Testing global ocean carbon cycle models using measurements of atmospheric O<sub>2</sub> and CO<sub>2</sub> concentration. *Global Biogeochem. Cycl.* **12**, 213–230.
- Stocker, T. F., Wright, D. G. and Mysak, L. A. 1992. A zonally averaged, coupled ocean-atmosphere model for paleoclimate studies. *J. Climate* **5**, 773–797.
- Takahashi, T., Broecker, W. S. and Langer, S. 1985. Redfield ratio based on chemical data from isopycnal surfaces. *J. Geophys. Res.* **90**, 6907–6924.
- Takahashi, T., Sutherland, S. C., Sweeney, C., Poisson, A., Metzl, N. and co-authors. 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 II* **49**, 1601–1622.
- Tans, P. P., Fung, I. Y. and Takahashi, T. 1990. Observational constraints on the global atmospheric CO<sub>2</sub> budget. *Science* **247**, 1431–1438.
- Thoning, K. W., Tans, P. P. and Komhyr, W. D. 1989. Atmospheric carbon dioxide at Mauna Loa Observatory, 2, Analysis of the NOAA/GMCC data, 1974–1985. *J. Geophys. Res.* **94**(D6), 8549–8565.
- Tiedtke, M. 1989. A comprehensive mass flux scheme for cumulus parameterization in large-scale models. *Mon. Wea. Rev.* **117**, 1779–1800.
- Toggweiler, J. R. and Samuels, B. 1993. New radiocarbon constraints on the upwelling of abyssal water to the ocean's surface. In: *The global carbon cycle* (ed. Heimann M.), pp. 333–366. Springer-Verlag, Berlin, Heidelberg.
- Tohjima, Y., Mukai, H., Machida, T., Nojiri, Y. and Gloor, M. 2005. First measurements of the latitudinal atmospheric O<sub>2</sub> and CO<sub>2</sub> distributions across the western Pacific. *Geophys. Res. Lett.* **32**, doi:10.1029/2005GL023311.
- Walsh, J. 1978. A data set on Northern Hemisphere sea ice extent, 1953–1976. Glaciological data, World Data Cent. A for Glaciol. [Snow and Ice], Boulder, Colorado. Report GD-2.
- Wanninkhof, R. 1992. Relationship between wind speed and gas exchange over the ocean. *J. Geophys. Res.* **97**(C5), 7373–7382.
- Weiss, R. F. 1970. The solubility of nitrogen, oxygen and argon in water and seawater. *Deep Sea Res.* **17**, 721–735.
- Yamanaka, Y. and Tajika, E. 1996. The role of the vertical fluxes of particulate organic matter and calcite in the oceanic carbon cycle: studies using an ocean biogeochemical general circulation model. *Global Biogeochem. Cycl.* **10**(2), 361–382.
- Zwally, H. J., Comiso, J., Parkinson, C., Campbell, W., Carsey, F. and co-authors. 1983. *Antarctic sea ice, 1973–1976: Satellite passive microwave observations*. NASA.