



HAL
open science

Benchmarking coupled climate-carbon models against long-term atmospheric CO₂ measurements

P. Cadule, P. Friedlingstein, L. Bopp, S. Sitch, C. Jones, Philippe Ciais, S. Piao, P. Peylin

► **To cite this version:**

P. Cadule, P. Friedlingstein, L. Bopp, S. Sitch, C. Jones, et al.. Benchmarking coupled climate-carbon models against long-term atmospheric CO₂ measurements. *Global Biogeochemical Cycles*, 2010, 24 (2), pp.n/a-n/a. 10.1029/2009GB003556 . hal-02927120

HAL Id: hal-02927120

<https://hal.science/hal-02927120>

Submitted on 28 Oct 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Benchmarking coupled climate-carbon models against long-term atmospheric CO₂ measurements

P. Cadule,^{1,2} P. Friedlingstein,^{1,3} L. Bopp,¹ S. Sitch,^{4,5} C. D. Jones,⁶ P. Ciais,¹ S. L. Piao,⁷ and P. Peylin^{1,8}

Received 1 May 2009; revised 27 November 2009; accepted 31 December 2009; published 26 June 2010.

[1] We evaluated three global models of the coupled carbon-climate system against atmospheric CO₂ concentration measured at a network of stations. These three models, HadCM3LC, IPSL-CM2-C, and IPSL-CM4-LOOP, participated in the C⁴MIP experiment and in various other simulations of the future climate impacts on the land and ocean carbon cycle. A new set of performance metrics is defined and applied to quantify each model's ability to reproduce the global growth rate, the seasonal cycle, the El Niño–Southern Oscillation (ENSO)–forced interannual variability of atmospheric CO₂, and the sensitivity to climatic variations. Knowing that the uncertainty on the amplitude, in 2100, of the climate-carbon feedback is mainly due to the uncertainty of the response of the terrestrial biosphere to the climate change, our new metrics primarily target the evaluation of the land parameterization of the carbon cycle. The modeled fluxes are prescribed to the same global atmospheric transport model LMDZ4, and the simulated concentrations are compared to available observations. We found that the IPSL-CM4-LOOP model is best able to reproduce the phase and amplitude of the atmospheric CO₂ seasonal cycle in the Northern Hemisphere, while the other two models generally underestimate the seasonal amplitude. This points to some shortcomings in describing the vegetation phenology and heterotrophic respiration response to climate. We also found that IPSL-CM2-C produces a climate-driven abnormal source of CO₂ to the atmosphere in response to El Niño anomalies. Here a good model performance rests upon a realistic simulation of ENSO-type climate variability and the subsequent tropical carbon cycle response. The three climate models underestimate the sea surface temperature warm anomaly during an El Niño, but HadCM3LC does best in reproducing the interannual CO₂ variability. More efforts are needed to further develop metrics for assessing the sensitivity of the carbon cycle to climate change, and this work should now be extended to assess ocean carbon models against observations.

Citation: Cadule, P., P. Friedlingstein, L. Bopp, S. Sitch, C. D. Jones, P. Ciais, S. L. Piao, and P. Peylin (2010), Benchmarking coupled climate-carbon models against long-term atmospheric CO₂ measurements, *Global Biogeochem. Cycles*, 24, GB2016, doi:10.1029/2009GB003556.

1. Introduction

[2] Increasing atmospheric CO₂ concentration will be the most important driver of climate change throughout the 21st

century. The atmospheric CO₂ concentration has already increased by around 100 ppm (+30%) compared to 1860s preindustrial levels [Trenberth *et al.*, 2007]. This increase has both a direct and an indirect effect on the carbon cycle, the latter being due to climate change induced by the concentration change. Global models of the coupled climate-carbon system have shown that the indirect climatic effect induces a reduction of both land and ocean capacity to absorb atmospheric CO₂ [Cox *et al.*, 2000; Friedlingstein *et al.*, 2001; Dufresne *et al.*, 2002; Friedlingstein *et al.*, 2003; Jones *et al.*, 2003; Zeng *et al.*, 2004; Thompson *et al.*, 2004; Govindasamy *et al.*, 2005; Matthews *et al.*, 2005]. Consequently, these reduced sinks lead to an additional amount of CO₂ remaining in the atmosphere, which ranges between 20 and 220 ppm by 2100 following the SRES-A2 scenario, as shown by Friedlingstein *et al.* [2006] in the Coupled

¹UMR CEA, IPSL, LSCE, UVSQ, CNRS, Gif-sur-Yvette, France.

²IPSL, UPMC, Paris, France.

³QUEST, Department of Earth Sciences, University of Bristol, Bristol, UK.

⁴Met Office, JCHMR, Wallingford, UK.

⁵School of Geography, University of Leeds, Leeds, UK.

⁶Met Office, Hadley Centre, Exeter, UK.

⁷Department of Ecology, Peking University, Beijing, China.

⁸Laboratoire de Biogéochimie et Ecologie des Milieux Continentaux, BIOEMCO, Thiverval-Grignon, France.

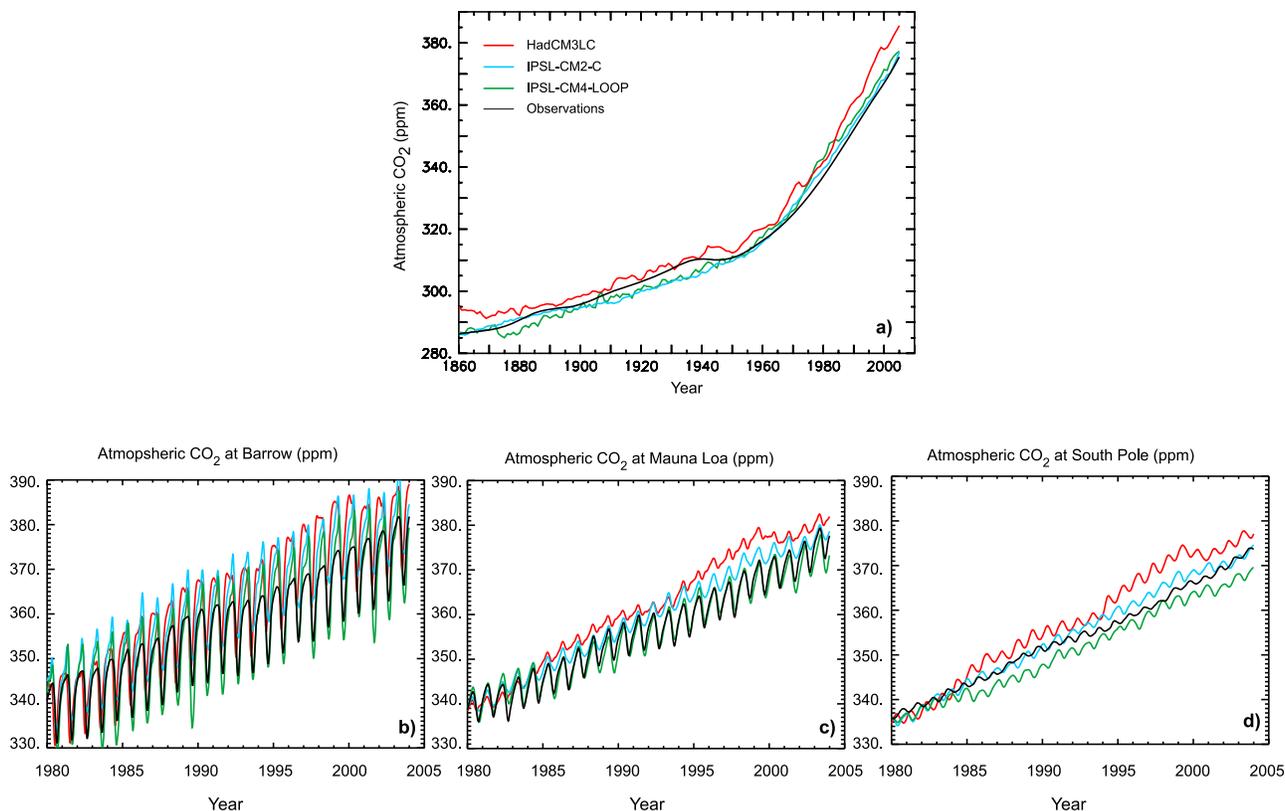


Figure 1. (a) Global average atmospheric CO₂ concentration simulated by the three coupled models: HadCM3LC (red), IPSL-CM2-C (blue), and IPSL-CM4-LOOP (green) over the historical period. Observed global CO₂ concentration, from combined ice core and atmospheric measurements, is shown in black. (b) Atmospheric CO₂ concentration at Point Barrow (BRW), Alaska, observed (black) and simulated by the atmospheric transport of the surface fluxes from the three coupled models. Color scheme is as in Figure 1a. (c and d) Same as Figure 1b for Mauna Loa (MLO), Hawaii, and South Pole (SPO), respectively.

Carbon Cycle Climate Model Intercomparison Project (C⁴MIP). This positive feedback leads to an increase of global warming by up to 0.9°C [Cadule *et al.*, 2009]. In models, the carbon cycle response to climate change consists of (1) a decline of tropical forests and a widespread, climate-driven, loss of soil carbon; and (2) a decrease of CO₂ uptake by the oceans, caused by decreased solubility and by a shrinking volume of the surface mixed layer in contact with the atmosphere. Further, Raddatz *et al.* [2007] show that the C⁴MIP models unanimously agree on a key role of tropical regions in controlling the global carbon-climate positive feedback. However, the models wildly diverge concerning the amount of carbon released by the tropics. Furthermore, the C⁴MIP models also diverge in their response to climate change at midlatitudes and high latitudes. Indeed, climate warming has opposing effects on terrestrial ecosystems: on the one hand, an increased productivity will increase carbon storage in biomass and soils, but on the other hand, warmer and wetter conditions will accelerate the decomposition of litter and soil organic carbon. The balance between these opposing effects thus determines the magnitude, and even the sign, of carbon cycle feedbacks in midlatitude and high-latitude ecosystems.

[3] The large range of C⁴MIP models' results, regarding the climate-carbon feedback amplitude, reflects divergences in their representations of basic carbon cycle processes and their interactions [Sitch *et al.*, 2008; Le Quéré *et al.*, 2005; Heimann and Reichstein, 2008]. Nevertheless, all these C⁴MIP models are able to simulate a 20th century CO₂ increase broadly consistent with the historical record, with an atmospheric CO₂ concentration of 380 ± 14 ppm in 2005, as compared to the observed value of 379 ppm [Trenberth *et al.*, 2007]. Note that, as per the C⁴MIP protocol, various radiative forcings (such as non-CO₂ GHGs, aerosols, and volcanoes) are not taken into account even if studies have shown the importance of these atmospheric agents on the climate [Mitchell *et al.*, 1995; Meehl *et al.*, 2007] and the carbon cycle [Jones *et al.*, 2003]. The HadCM3LC [Cox *et al.*, 2000], IPSL-CM2-C [Dufresne *et al.*, 2002], and IPSL-CM4-LOOP coupled models, more deeply analyzed in this paper, produce an atmospheric CO₂ concentration of 386, 377, and 377 ppm, respectively, in 2005 (Figure 1a).

[4] Simulating the cumulated change in atmospheric CO₂ during the 20th century is thus a necessary but not sufficient condition to validate a coupled climate-carbon cycle model. It is in fact not even a very stringent constraint on the future

Table 1. Selected CO₂ Measurement Stations

Stations	Code	Position	Altitude (m)	
Alert	ALT	Nunavut, Canada	82°45'N 62°52'W	210
Amsterdam Island	AMS	France	37°95'S 77°53'E	150
Terceira Island	AZR	Azores, Portugal	38°77'N 27°38'W	30
Barrow	BRW	Alaska, United States	71°32'N 156°60'W	94
Cape Grim	CGO	Tasmania, Australia	40°68'S 144°68'E	11
Cape Kumukahi	KUM	Hawaii, United States	19°52'N 154°82'W	3
Mace Head	MHD	Galway, Ireland	53°33'N 09°90'W	25
Mauna Loa	MLO	Hawaii, United States	19°52'N 155°58'W	3397
Niwot Ridge	NWR	Colorado, United States	40°05'N 105°58'W	3475
Schauinsland	SCH	Germany	47°55'N 07°55'W	1205
American Samoa	SMO	Tutuila, American Samoa	14°24'N 170°57'W	42
South Pole	SPO	Antarctica	89°98'S 24°80'W	2810

CO₂ projections [Melnikov and O'Neill, 2006]. It is therefore important both to identify observational constraints which will reduce the uncertainty in the range of future model projections and to improve the understanding of carbon cycle mechanisms.

[5] In this study, we propose a set of performance metrics, based on various characteristics of atmospheric CO₂, to quantify the ability of climate-carbon coupled models to reproduce key processes of climate-carbon projections. We apply these new metrics to three C⁴MIP models: HadCM3LC, IPSL-CM2-C, and IPSL-CM4-LOOP. The method consists in indirectly evaluating land and ocean modeled carbon fluxes against observations of atmospheric CO₂ at various stations. This is achieved by calculating the atmospheric CO₂ field derived from each model flux distribution (section 3.3). The model evaluation is performed at three different time scales which give different constraints on the carbon cycle. First, we analyze the long-term trend of atmospheric CO₂, which provides information on the model capability to simulate realistic land and ocean carbon sinks over the historical period. Second, we evaluate the modeled atmospheric CO₂ seasonal cycle, which, particularly at Northern Hemisphere atmospheric CO₂ stations, constrains the model's simulation of the continental fluxes seasonal activity: vegetation growth in spring and summer and vegetation decay in autumn. Third, we analyze the interannual variability of the atmospheric CO₂ as a constraint on the model capability to simulate realistic El Niño–Southern Oscillation (ENSO) climate patterns and impacts on land and ocean carbon fluxes. For the seasonal and interannual variability, we first evaluate the model capability to represent the CO₂ signal, and then we evaluate the sensitivity of the atmospheric CO₂ to climatic fluctuations for these two time scales. Statistical methods are used to analyze model-data misfits (section 3.5) and to define a new metric to quantify the various models performances (section 3.6). Results are presented in section 4. The analysis and discussion of models' traits and shortcomings are presented in section 5, and conclusions are drawn.

2. Models' Description

2.1. Introduction

[6] The three models (HadCM3LC, IPSL-CM2-C, and IPSL-CM4-LOOP) used in this study appear in the IPCC

Fourth Assessment Report [Meehl et al., 2007]. These models also participated in the C⁴MIP project [Friedlingstein et al., 2006]. A description of the three models is given in Appendix A.

2.2. Simulation Scenario

[7] For this study, we used the C⁴MIP simulations [Friedlingstein et al., 2006]. The three models were forced by the same anthropogenic CO₂ emissions prescribed from historical data, for the 1860–2000 period [Marland et al., 2005; Houghton and Hackler, 2002] and from the SRES-A2 scenario, for the 2000–2003 period [Nakicenovic et al., 2000]. The other greenhouse gases (GHGs) and the anthropogenic aerosols were set to preindustrial values. The models did not account for any change in solar irradiance or natural aerosols due to volcanic eruptions. Emissions of CO₂ to the atmosphere due to land use change were prescribed within the emissions scenario, but the direct effect of land use changes on vegetation distribution, as well as changes in nitrogen deposition, were excluded in these simulations. The simulated terrestrial biosphere and ocean carbon fluxes thus only respond to the increase of atmospheric CO₂ concentration and to climate change which itself is only due to the increase of atmospheric CO₂ concentration.

3. Methodology

3.1. Atmospheric Site Selection

[8] We use the global cooperative data product GLOBALVIEW-CO₂ which contains 277 extended CO₂ records derived from observations made by different measurement networks, covering the whole globe for the period 1979–2008 [GLOBALVIEW-CO₂, 2008]. Here we choose 12 monitoring sites representative of high, middle, and low latitudes of each hemisphere (Table 1), and we focus over the 1979–2003 period.

3.2. Data Filtering Technique

[9] Our goal is the evaluation of models at different time scales from intra-annual to interannual and longer term. We use a curve-fitting procedure based on a polynome (degree 2), four harmonics, and a digital filtering technique relying on fast Fourier transform and low-pass filters [Thoning et al., 1989] to decompose atmospheric CO₂ flask observations, sampled at discrete weekly intervals, into three components:

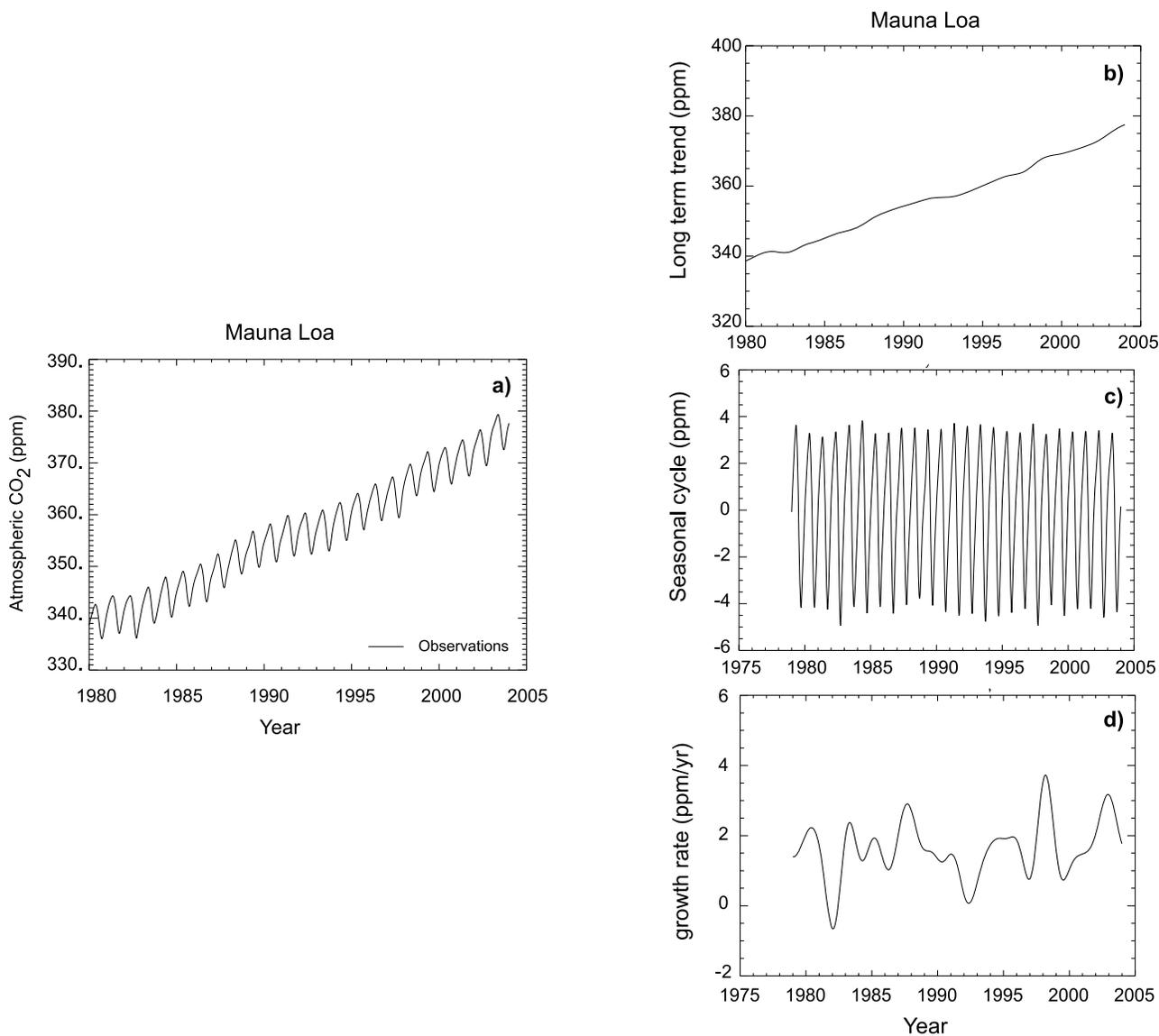


Figure 2. Decomposition of the (a) atmospheric CO₂ signal into its (b) long-term, (c) seasonal, and (d) interannual components as done by *Thoning et al.* [1989].

(1) a long-term trend (TR) component, (2) a smoothed seasonal cycle (SC), and (3) an interannually variable (IAV) component (Figure 2).

3.3. Atmospheric Transport

[10] Atmospheric CO₂ concentration variations at a given station result from transport acting on surface fluxes. In order to evaluate model fluxes against measured CO₂ concentrations, it would be ideal to describe, in the same coupled modeling framework, the spatial and temporal variations of the atmospheric CO₂ field. Most models do not, however, compute this field. We thus used the global 3-D atmospheric transport model LMDZ4 [*Hourdin et al.*, 2006] to transport land and ocean carbon flux maps from HadCM3LC, IPSL-CM2-C and IPSL-CM4-LOOP. Even though HadCM3LC does simulate atmospheric transport of CO₂, we opted to

perform a separate transport of its fluxes for this study for consistency with the other models. Any reported difference between the modeled CO₂ concentration fields will be due to the simulated carbon fluxes and not to atmospheric transport. Likewise any errors in the transport are now common to each model. LMDZ4 was nudged toward interannually varying reanalyzed meteorological wind fields from ECMWF, which ensure maximum consistency in the comparison of modeled and observed CO₂ concentrations. Using the transport field from each OAGCM would have lead to additional source of biases in this benchmarking exercise. Nevertheless, a comparison of CO₂ concentrations (not shown here) showed that HadCM3LC-simulated CO₂ agreed very closely with that transported by LMDZ4. Anthropogenic emissions from fossil fuel combustion [*Marland et al.*, 2005] and deforestation [*Houghton and Hackler*, 2002] were also transported

Table 2. The 22 Aggregate Regions as Defined in the TransCom-3 Project

Latitudinal Grouping	Regions
Northern Land	Boreal North America
	Boreal Asia
	Temperate North America
	Temperate Asia
Northern Oceans	Europe
	Northern Ocean
	North Atlantic
Tropical Land	North Pacific
	Northern Africa
	Tropical America
Tropical Oceans	Tropical Asia
	West Pacific
	East Pacific
	Tropical Atlantic
Southern Land	Tropical Indian
	Australia
	Southern Africa
Southern Oceans	Temperate South America
	South Atlantic
	South Pacific
	South Indian
	Southern Ocean

within LMDZ4 and added to each tracer before comparison to the observations.

3.4. Regions of Influence

[11] The notion of region of influence is defined to quantify the relative contribution of several key regions of the globe in terms of surface flux contributions to the temporal CO₂ concentration variations at few stations. Using the LMDZ4 model, we transported the carbon fluxes obtained from atmospheric inversions (P. Peylin et al., unpublished results, 2009) for 22 regions (11 terrestrial and 11 oceanic; see Table 2) as defined by the TransCom-3 project [Baker et al., 2006] and performed a principal component analysis (PCA) of the monthly simulated concentrations. The principal components account for the variability of the data, and the coefficients can be interpreted as the contribution of each region to the CO₂ concentration measured at each station. These contributions vary depending on the time scale (seasonal, interannual, or decadal) [Randerson et al., 1997; Heimann et al., 1998; Nevison et al., 2008] and also on the spatial distribution of the terrestrial and oceanic carbon fluxes. Here we only use the PCA, at seasonal time scale, for the estimation of the sensitivity of the atmospheric CO₂ concentration to temperature. The relative contribution of each region is given in Appendix C at the 12 stations, for the seasonal cycle. Table C1 enables us to highlight the regions that primarily influence the air masses at the different stations.

[12] For the seasonal cycle (Table C1), ALT and BRW are mainly influenced by arctic and high northern latitude terrestrial regions, BRW being, in addition, influenced by Northern Ocean. MHD and AZR are mostly influenced by boreal North America, Europe, and boreal Asia and to a lesser extent by temperate North America. SCH and NWR are both primarily influenced by Europe while NWR is also influenced by temperate North America and high northern

latitudes. KUM and MLO are influenced by the midlatitudes and high latitudes of the Northern Hemisphere, KUM being more dominated by the boreal regions than MLO. SMO is mainly influenced by Africa, while AMS and CGO are mainly influenced by the tropics. For SPO, and more generally for the four stations of the Southern Hemisphere, the 22 regions contribute much more evenly than for the eight stations of the Northern Hemisphere. Finally, Table C1 also shows that the contribution of the oceanic regions is very low compared to that of the terrestrial regions.

3.5. Model Data Statistical Analysis

[13] We use statistical tests, the Pearson correlation (PEARSON), the normalized standard deviation (NSD), and a simple normalized model to data deviation measure (MOD) to perform comparisons between models and observations. These tests, described in Appendix B, stand as the base for quantifying models performances.

3.6. Multi-Time-Scale Evaluation Traits

[14] The study of the influence of land and ocean carbon fluxes on the atmospheric CO₂ concentration, both in terms of seasonal cycle and of interannual variability, is well documented [Pearman and Hyson, 1981; Cleveland et al., 1983; Keeling et al., 1984; Bacastow et al., 1985; Peterson et al., 1986; Thompson et al., 1986; Enting, 1987; Thoning et al., 1989; Keeling et al., 1996; Jones et al., 2001; Buermann et al., 2007]. We define here for each time scale (seasonal, interannual, and longer term) a series of traits. A trait is a quantity which can be calculated both with modeled and observed CO₂ time series, providing a basis for model evaluation. We use the above statistical tests to quantify the discrepancy between models and observations for each trait.

3.6.1. Long-Term Trend

[15] We evaluate two characteristic traits of the modeled long-term CO₂ trend. The long-term CO₂ trend (TR) is itself a trait which provides information on the decadal mean carbon budget, assuming perfect atmospheric transport (i.e., if decadal trends in transport exist, they will be perfectly captured by the LMDZ4 model). For each of the models and for the observations, this trait is composed of 25 yearly averages over the 1979–2003 period. We also calculated the long-term north–south gradient (TR-NSG) of atmospheric CO₂ concentration as a trait. This trait was calculated over the 1979–2003 period by subtracting CO₂ at South Pole (SPO) from CO₂ of each other station and provides information on the spatial (north–south) distribution of the land and ocean carbon sinks. We apply the NSD test to the TR trait of the 12 stations data set, and to the TR-NSG trait of the 11 stations data set as SPO is used as reference station for this latter trait.

[16] We apply the NSD test to the TR trait of the 12 stations data set and to the TR-NSG trait of the 11 stations data set as SPO is used as reference station for this latter trait.

3.6.2. Seasonal Cycle

[17] We evaluate three characteristic traits of the modeled CO₂ seasonal cycle. The first and second traits are the amplitude (SC-a) and phase (SC-p), respectively, of the climatologic average seasonal cycle at each station, which reflects the dynamics of the carbon cycle during the course

of the year. The third trait is the trend in the peak-to-peak amplitude of the seasonal cycle (SC-ptp). The trend in the strength of the seasonal pulse of CO₂ provides information on the trend in the land carbon sink. These three traits are defined over the 1979–2003 period. We apply the NSD and PEARSON tests to the two climatologic average seasonal cycle traits to evaluate the amplitude and phase, respectively. We also apply the MOD test to the peak-to-peak trait.

3.6.3. Interannual Variability

[18] We evaluate two characteristic traits of the interannual variability of atmospheric CO₂ growth rate. The first trait concerns the positive anomalies of the IAV (IAV+), and the second concerns the negative anomalies of the IAV (IAV−). To obtain the IAV traits, we calculated the first principal component of the empirical orthogonal function (EOF) of the sea surface temperature (SST) for each model. We selected the boxes in which the SST anomaly is maximum, which are the Niño 1.2 (90°W–80°W, 10°S–0) box for IPSL-CM2-C and the Niño3 (150°W–90°W, 5°S–5°N) box for both HadCM3LC and IPSL-CM4-LOOP. We then searched for the optimal positive time lag between SST anomalies, in the respective boxes, and anomalies of the CO₂ growth rate at each station. The optimal lag (of n months) is defined when the correlation between SST and the CO₂ growth rate, shifted by n months, is maximum. We then extracted positive (and negative) SST anomalies, and, using the lag defined at each station, retrieved the corresponding atmospheric CO₂ anomalies. We define positive (negative, respectively) SST anomalies as the periods for which the SST value is greater than $+0.7 \times$ standard deviation, (lower than $-0.7 \times$ standard deviation, respectively) for more than three consecutive months. For each positive (negative, respectively) SST anomaly period, only the maximum (minimum, respectively) is retained, and the corresponding of the CO₂ growth rate anomalies is retrieved. Note that in order to facilitate the computation and analysis of the IAV related traits, the Pinatubo period was removed from the observations. Using the MOD test, we then compare, at each station and for both positive and negative SST phases, the averaged maximum and minimum CO₂ growth rates to the observed ones. We used the SST anomalies in the Niño 3.4 (170°W–120°W, 5°S–5°N) box for the observations [Trenberth, 1997]. The SST data (observations and models) were smoothed with a 5 month running mean.

3.6.4. Sensitivity of Atmospheric CO₂ to Climate Fluctuations

[19] We evaluate three traits for the sensitivity of atmospheric CO₂ to climate fluctuations. The first trait is the sensitivity of atmospheric CO₂ concentration to temperature for the seasonal cycle and is referred to as γ_{SC} . To obtain this trait, we use the climatologic seasonal cycle of atmospheric CO₂ concentration at each station over the 1979–2003 period. We then calculate the climatologic seasonal cycle of the temperature simulated by the models, over the 1979–2003 period, in each of the 11 terrestrial regions (Table 2) and average these values using weights expressing the contribution of each terrestrial region to a specific atmospheric CO₂ station (see section 3.4 and Appendix C). The use of the 11 terrestrial regions, instead of the 22 (ter-

restrial and oceanic) regions, is driven by the relatively small contribution of the oceanic regions, as indicated in section 3.4 and Appendix C.

[20] For the observations, the same averaging technique was applied to the observed temperature given by the CRU [Mitchell and Jones, 2005]. We then calculate γ_{SC} for the 12 climatologic months using (1), where C_m and T_m are the CO₂ concentration and temperature, respectively, for month m , for models or observations; r is the index for the different regions (Table 2); and α is the weighting factor given by the PCA (Table C1).

$$\gamma - SC_m = \frac{C_m - C_{m-1}}{\frac{\sum_r \alpha_r \times T_{r,m} - \sum_r \alpha_r \times T_{r,m-1}}{\sum_r \alpha_r}} \quad (1)$$

[21] The second trait relates to the sensitivity of the atmospheric CO₂ growth rate to the SST anomalies in the Niño 3 SST box and is referred to as γ_{IAV} . To obtain this trait, we make use of the extracted SST and CO₂ growth rate anomalies as explained in section 3.6.3. However, knowing that the study only covered 25 years, the number of points corresponding to the positive and negative phases of the ENSO is not sufficient to perform a separate evaluation of these two phases. As per the common definition, the positive and negative phases of the ENSO correspond to positive and negative anomalies, respectively, of SST in the specific Niño box. The two sets of points were considered as a whole, and the evaluation was performed on this combined set. We calculate the regression line for the combined sets (positive and negative phases). Further, as a result of a positive ENSO phase, precipitation decreases, and the temperature increases, in the tropics. Conversely, the response to a negative phase of the ENSO is an increase of precipitation in the tropics, but combined with a nonsignificant change of temperature. Therefore, the response of the terrestrial biosphere to the different phases of the ENSO may not be symmetric, and this explains that the regression line of the combined sets does not have a null intercept. As a result, we evaluate both the slope and intercept of the regression line as part of a single trait. We apply the PEARSON test to γ_{SC} and the MOD test to both the slope and intercept of the regression lines for γ_{IAV} and average the results of these two MOD tests. Table 3 summarizes the selected traits and associated performed tests.

3.7. Metric Definition

[22] The traits defined above help evaluate models against observations in a more quantitative way. Here we propose to associate a mark to each statistical test result in order to combine the different statistical tests into an overall performance mark associated to each of the models. The PEARSON test can return any value in the range $[-1, 1]$. We decompose that range into the following set of sub-ranges $\{-1, [-1, -0.8], \dots, [0.6, 0.8], [0.8, 1]\}$ to which we associate the following marks $\{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$. The NSD test can return any value within

Table 3. Chosen Traits and Associated Statistical Tests^a

	PEARSON	NSD	MOD
TR		X	
TR-NSG		X	
SC-a		X	
SC-p	X		
SC-ptp			X
IAV+			X
IAV-			X
γ_{SC}	X		
γ_{IAV}			X

^aMOD, model-to-data deviation measure. NSD, normalized standard deviation.

the range $[0, +\infty]$ with a nominal value of 1. To be able to associate a mark to a specific NSD test (trait, station), we propose to infer a validity range around the observed CO_2 value by using observation derived estimates of land and ocean carbon fluxes from atmospheric inversions (P. Peylin et al., unpublished results, 2009). The valid range around any specific observation is arbitrarily defined by transporting the inversion fluxes multiplied by 0.5 and 2, respectively. For these two fluxes, we calculate the NSD for the specific traits at each station. This gives, for any trait using NSD, a lower and an upper bound of NSD at each station. This range is then divided in regular subranges to which we associate the following marks $\{0.1, 0.2, 0.4, 0.6, 0.8, 1, 0.8, 0.6, 0.4, 0.2, 0.1\}$. A mark of zero is then given for any model that falls out of the upper or lower bounds of NSD. Similarly to the NSD test, we consider, for the MOD test, that a model with a value either 2 times greater than the observations or 0.5 times lower than the observations is out of the valid range. This defines the acceptable range of the MOD test results which is thus $[-0.33, 0.33]$. We decompose that range into the following set of subranges $\{[-0.33, -0.25], [-0.25, -0.20], \dots, [0.20, 0.25], [0.25, 0.33]\}$ and which are associated to the following marks $\{0.1, 0.2, 0.4, 0.6, 0.8, 1, 0.8, 0.6, 0.4, 0.2, 0.1\}$. For test results outside of the valid range, the associated mark is 0.

[23] For each station s and each characteristic trait t , we obtain a test result, transformed into a mark, symbolized as $M_{s,t}$. The performance mark of a given model, for a given trait, is thus:

$$P_t = \frac{\sum_s M_{s,t}}{nstations} \quad (2)$$

where $nstations$ is the number of stations considered. In the general case, $nstations$ is equal to 12, except for the TR-NSG test, for which it is equal to 11.

[24] Further, we average the marks for the CO_2 evaluation by separately averaging (without any weight) the two long-term trend-related marks, the three seasonal cycle related marks, the two marks on interannual variability, and the two marks of the sensitivity traits.

4. Results

[25] In this section, we present the traits described in section 3.6 for all models along with the relevant observa-

tion field. Although we generally only show figures for three major atmospheric CO_2 stations (Point Barrow, Alaska, latitude 71.32°N , longitude 156.62°W , BRW; Mauna Loa, Hawaii, latitude 19.54°N , longitude 155.58°W , MLO; and South Pole, latitude 89.98°S , longitude 24.8°W , SPO), the analysis was performed for the 12 selected stations. We then give the marks of each model for every trait (averaged over the 12 stations). Table 4 and Figure 3 summarize the different performance marks obtained by the models. A general discussion on the models capability to reproduce these traits will be proposed in section 5.

[26] The atmospheric CO_2 time evolution simulated at BRW, MLO, and SPO for the three models is shown on Figures 1b, 1c, and 1d, respectively. Figure 4 shows the observed (solid black) and simulated (HadCM3LC in solid red, IPSL-CM2-C in solid blue and IPSL-CM4-LOOP in solid green) long-term trend (TR, trait 1) at the three illustrative stations (Figures 4a, 4b, and 4c for BRW, MLO, and SPO, respectively). HadCM3LC has a general tendency to overestimate the long-term trend across all three stations just as for IPSL-CM2-C, but to a lesser degree for the latter. IPSL-CM4-LOOP shows a much better agreement with the observations while, in opposition to HadCM3LC and IPSL-CM2-C, underestimating the trend. The 12 stations averaged mark for the long-term trend (trait 1, TR) is 0.60, 0.80, and 0.85 for HadCM3LC, IPSL-CM2-C, and IPSL-CM4-LOOP, respectively.

[27] The evolution of the north-south gradient (here displayed for MLO) is shown in Figure 5. We only analyze the reproduction of its long-term trend (i.e., the dotted line in Figure 5); neither the mean value nor the year-to-year variability of the gradient is used in the mark here. At MLO, the trend simulated by IPSL-CM4-LOOP is lower than the observed, while the other two models show a better agreement with observations, although this behavior is not robust across the 12 stations. The 12 stations averaged mark for the trend in the north-south gradient (trait 2, TR-NSG) is 0.31, 0.39, and 0.53 for HadCM3LC, IPSL-CM2-C, and IPSL-CM4-LOOP, respectively.

[28] Note that the score for the long-term trend mostly highlights the differences existing between the different

Table 4. Performance Marks for HadCM3LC, IPSL-CM2-C, and IPSL-CM4-LOOP

	HadCM3LC	IPSL-CM2-C	IPSL-CM4-LOOP
<i>CO₂ Evaluation</i>			
TR	0.60	0.80	0.85
TR-NSG	0.31	0.39	0.53
SC-a	0.57	0.84	0.88
SC-p	0.13	0.23	0.66
SC-ptp	0.12	0.18	0.03
IAV+	0.64	0.00	0.33
IAV-	0.27	0.15	0.00
<i>Sensitivity Evaluation</i>			
γ_{SC}	0.35	0.69	0.98
γ_{IAV}	0.60	0.01	0.13
Total	0.40	0.37	0.49

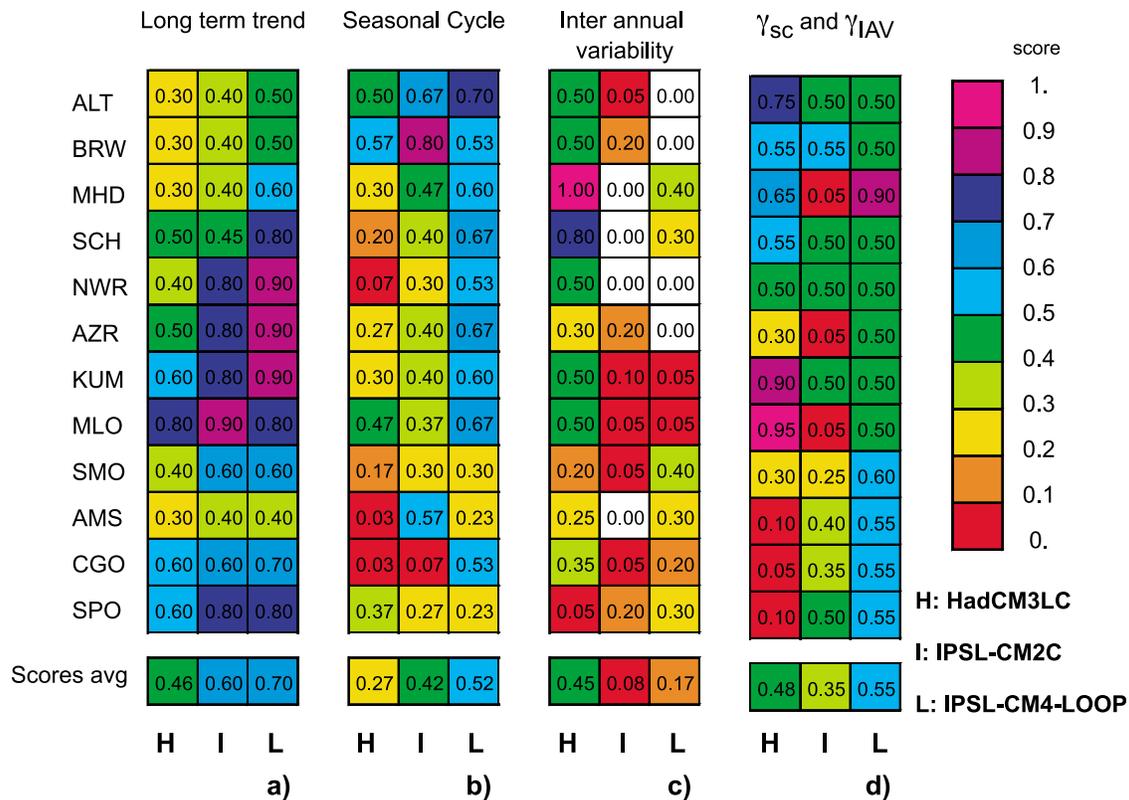


Figure 3. Matrix displaying the three models’ (H, HadCM3LC; I, IPSL-CM2-LC; and L, IPSL-CM4-LOOP) scores (see color bar) at all stations. Long-term trend score is the average of two traits: the trend in growth rate and the trend in interhemispheric gradient. Seasonal cycle score is the average of three traits: climatic average traits (phase and amplitude) and the trend in the peak-to-peak amplitude. Interannual variability score is the average of two traits: the El Niño and La Niña CO₂ variabilities. The “ γ_{SC} and γ_{IAV} ” is the average of two traits: (seasonal and interannual) CO₂-temperature sensitivities.

stations and models for the TR-NSG trait, as the decadal trend (TR trait) is very similar if not identical across the different stations for a given model.

[29] The simulations of the climatologic average CO₂ seasonal cycle are displayed for BRW, MLO, and SPO (Figure 6). HadCM3LC phase is 1 to 2 months too early at BRW and almost out of phase (6 months) at SPO. At MLO, the model’s phase is better, but its amplitude is more than a factor of 2 too low. The two IPSL models (in particular, IPSL-CM4-LOOP) show a rather good agreement both in terms of amplitude and phase (especially at BRW for IPSL-CM2-C and at MLO for IPSL-CM4-LOOP), although they are still out-phased by 3 months at SPO. Reasons for the HadCM3LC behavior are detailed in section 5. The 12 stations’ averaged marks for the phase and amplitude of the climatologic average CO₂ seasonal cycle (traits 3, SC-a and 4, SC-p) are 0.57 and 0.13 for HadCM3LC, 0.84 and 0.23 for IPSL-CM2-C, and 0.88 and 0.66 for IPSL-CM4-LOOP, respectively.

[30] Next, we evaluated the long-term trend in the peak-to-peak amplitude of the seasonal cycle (Figures 7 and 8). This trend was reported at several stations [Keeling *et al.*, 1996; Randerson *et al.*, 1997, 1999; Piao *et al.*, 2008], and we thus performed the analysis on the 12 stations.

Figure 8 shows the change in the peak-to-peak amplitude relative to 1979 at BRW, MLO, and SPO. IPSL-CM4-LOOP has a tendency to overestimate the trend at the three stations, while HadCM3LC has a very realistic trend at SPO and to a lesser degree at MLO, but a larger one than the observations at BRW. IPSL-CM2-C is in fair agreement with the observations at BRW.

[31] The 12 stations’ averaged mark for the trend in the peak-to-peak amplitude (trait 5, SC-ptp) is 0.12, 0.18, and 0.03 for HadCM3LC, IPSL-CM2-C, and IPSL-CM4-LOOP, respectively. Traits 6 (IAV+) and 7 (IAV-) are the reproduction of the positive and negative phases, respectively, of the ENSO variability. Figure 9 shows the observed and simulated interannual anomalies of the CO₂ growth rate at MLO. The amplitude of both the positive and negative anomalies is much better simulated by the HadCM3LC model than by the two IPSL models, with a generally very good agreement with the observations in the Northern Hemisphere. In particular, the IPSL-CM2-C model has a much lower CO₂ interannual variability than the observations. The 12 stations averaged mark for positive and negative phase of ENSO variability are 0.64 and 0.27 for HadCM3LC, 0.00 and 0.15 for IPSL-CM2-C, and 0.33 and 0.00 for IPSL-CM4-LOOP, respectively.

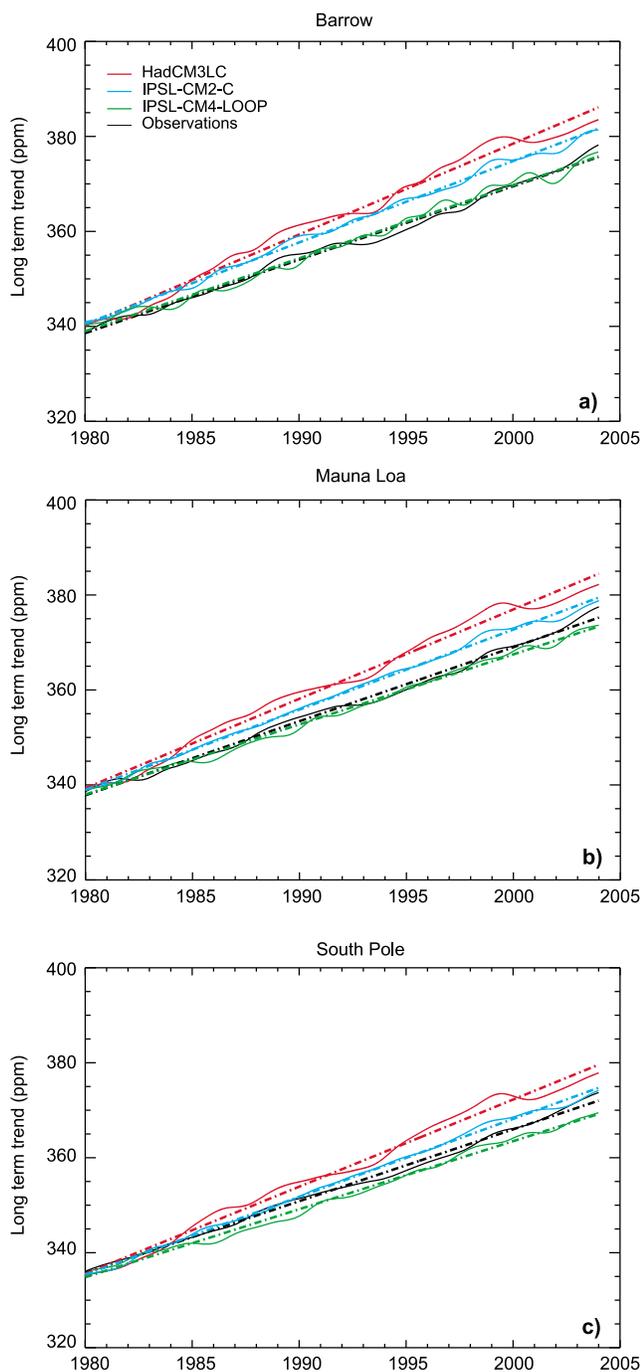


Figure 4. Observed and simulated long-term trend of atmospheric CO₂ at the three illustrative stations (BRW, MLO, and SPO). The deseasonalized annual CO₂ concentration is shown with a solid line, while its averaged value over the 1979–2003 period is shown with a dotted line. Modeled CO₂ is set to the observation value for January 1979.

[32] We then look at the links between seasonal cycle of temperature and atmospheric CO₂, in order to test the model climate-carbon sensitivity at the seasonal time scale (trait 8). Figure 10 shows models' and observations' seasonal evolution in an atmospheric CO₂ versus surface temperature diagram. As described in section 3.6.4, a weighted mean of the surface temperatures was calculated over the continental regions which controls the seasonal variability of the atmospheric CO₂ station considered. At BRW and MLO, the HadCM3LC model produces a much too narrow ellipse, indicating a too weak atmospheric CO₂ seasonal amplitude, despite a realistic temperature range over the seasons. At MLO, IPSL-CM4-LOOP shows an ellipse in very good agreement with the observations, although there is a clear cold bias in the simulated temperature. IPSL-CM2-C has a too low CO₂ pulse despite a larger than observed temperature range (particularly at MLO). The 12 stations' averaged mark for the trend in seasonal CO₂-temperature sensitivity (trait 8) is 0.35, 0.69, and 0.98 for HadCM3LC, IPSL-CM2-C and IPSL-CM4-LOOP, respectively.

[33] Finally, we analyzed the CO₂-temperature sensitivity on interannual time scale for the combined positive and negative phase of ENSO variability (Figure 11 for MLO). HadCM3LC model has a CO₂-temperature regression line intercept which is nearest to the observations compared to the other two models, and while IPSL-CM2-C shows a regression line slope more faithful to the observations, HadCM3LC overall test result outperforms the other models. The 12 stations' averaged marks for the CO₂-temperature sensitivity in response to ENSO variability (trait 9) are 0.60, 0.01, and 0.13 for HadCM3LC, IPSL-CM2-C, and IPSL-CM4-LOOP, respectively.

5. Analysis and Discussion

[34] In this paper we have attempted to define some observed traits of atmospheric CO₂ which we would expect

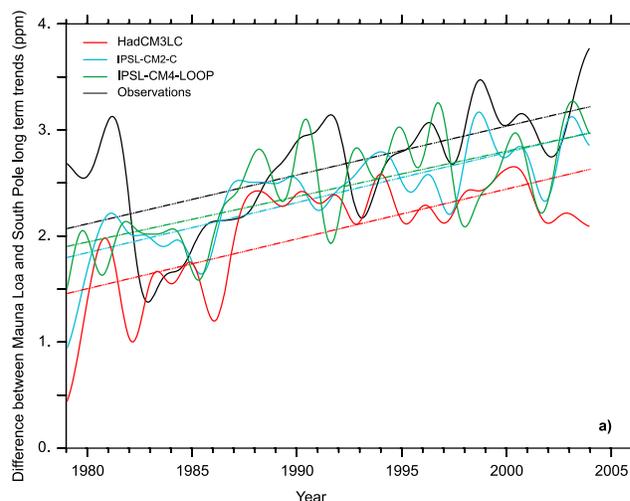


Figure 5. Time evolution of the interhemispheric gradient of atmospheric CO₂, defined here as the difference between the deseasonalized annual CO₂ concentration at MLO and SPO.

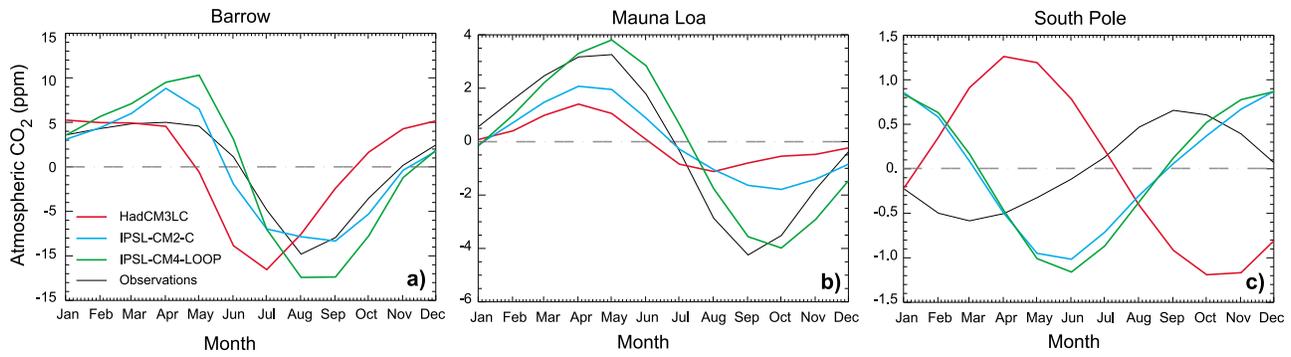


Figure 6. Simulated and observed climatologic averaged CO₂ seasonal cycle at BRW, MLO, and SPO.

coupled climate-carbon cycle earth system models (ESMs) to be able to capture and some metrics to quantitatively measure model performance against them. We have chosen to use a common transport model in order to assess inter-

ESM differences in simulated CO₂ fluxes, but the method could be applied to ESM-simulated CO₂ where available. It is not our intent to suggest that CO₂ should be a sole measure for model evaluation. Many other metrics are

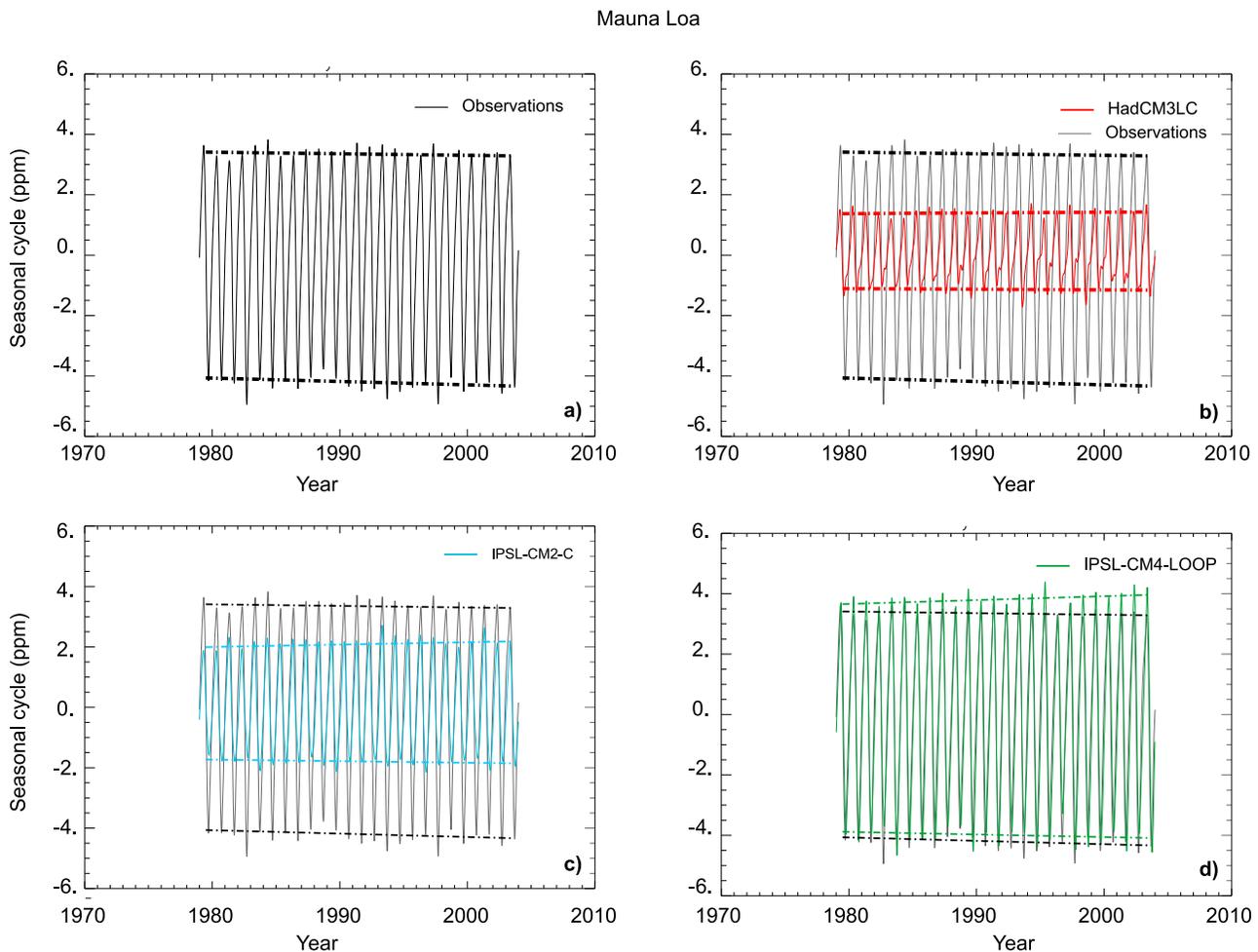


Figure 7. (a) Change in the observed amplitude of the atmospheric CO₂ seasonal cycle at MLO. (b, c, and d) Change in the amplitude of the seasonal cycle at MLO for the three models. The observed data are also shown in grey.

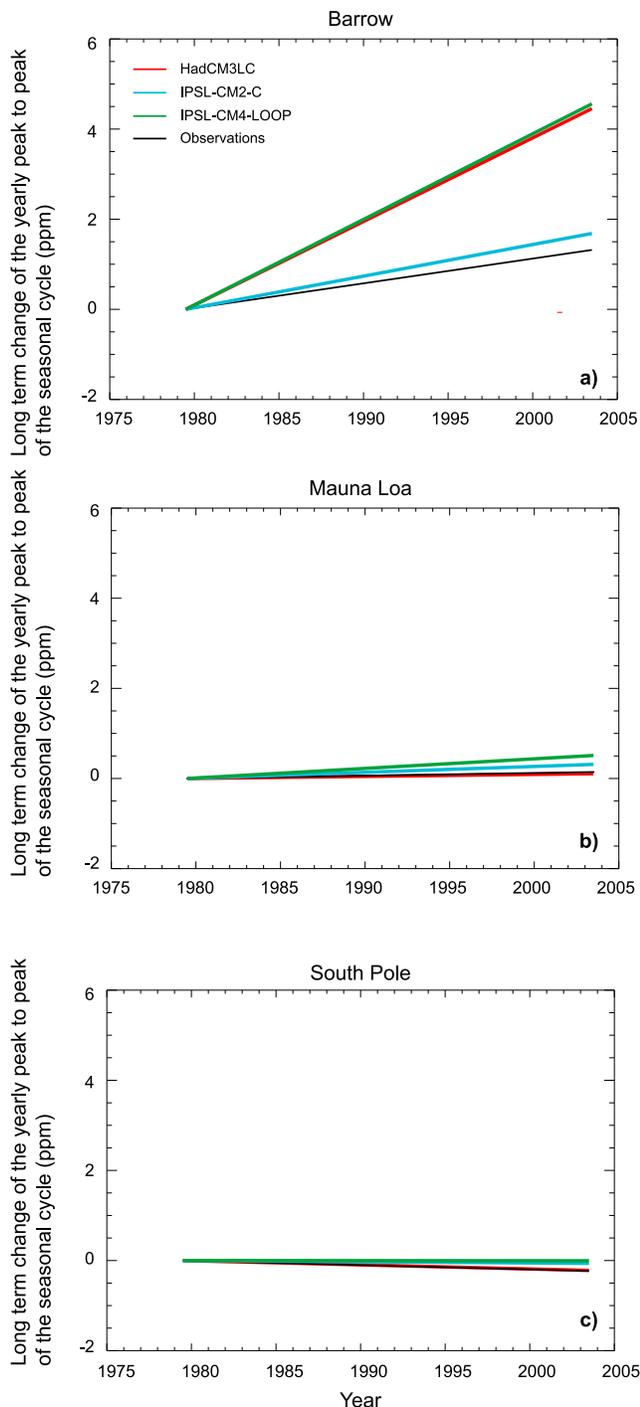


Figure 8. Observed and simulated linear trend of the peak-to-peak amplitude of the seasonal cycle of CO₂ at the three illustrative stations for (a) BRW, (b) MLO, and (c) SPO.

possible, such as high-frequency site-level carbon and energy fluxes from eddy covariance measurements, vegetation cover, structure, and carbon storage or phenological behavior, such as leaf area observed by remote sensing. Such metrics have been proposed and evaluated elsewhere [Randerson *et al.*, 2009]. Here we have focused on CO₂ as a large-scale integrator of both terrestrial and ocean CO₂

fluxes with valuable information on a wide range of time scales.

5.1. Methodology

[35] It should be noted that the evaluation was performed with coupled models which does not enable us to specifically highlight either the role of the climate or the role of the carbon cycle in the difference that lies between the models and the observations. This study could be augmented by an identical study where the carbon cycle models are run off-line (i.e., forced by observed climate). With an identical climate it would be possible to explicitly identify the modeled carbon cycle deficiencies. Further, this study only provides indirect information on net simulated fluxes but does not enable us to evaluate the detailed specific processes playing a role in the carbon cycle at different time scales nor does it enable us to evaluate their sensitivity to climate over the 1979–2003 period. The follow-up study will thus incorporate a finer-grain evaluation using data issued from measurements at flux towers.

5.2. Missing Forcings and Processes

5.2.1. Missing Forcings

[36] There are of course limitations to our approach, at least partly due to the experimental design of the simulations available to us here and partly due to limitations in the models themselves. The C⁴MIP intercomparison study [Friedlingstein *et al.*, 2006] used a deliberately simplified experimental design as its aim was to compare results across a range of coupled climate-carbon cycle models. Those simulations were not designed to do any more than approximate the 20th century. They neglected non-CO₂ GHGs, aerosols, and natural forcings which, by chance, roughly cancel in the 20th century global mean [e.g., Solomon *et al.*, 2007, Figure SPM.2]. Hence they get about the right level of global warming, but at least partly for the wrong reasons. Although this makes the climate-carbon cycle model experiment setup simpler, it does imply that spatial pattern of climate change and hence carbon cycle changes may not be realistic and hence limit the power of the model evaluation against the observations.

[37] Limitations arising from the lack of natural and non-CO₂ anthropogenic forcings mean the simulations do not capture the detailed spatial and temporal pattern of 20th century climate and CO₂. Jones *et al.* [2003] and Cadule *et al.* [2009] showed the impact of different forcings, finding anthropogenic aerosols especially important giving rise to a different time evolution, different N-S gradient, and different land: ocean uptake split by present day and into the future. The response of the global carbon cycle to the eruption of Mt. Pinatubo in 1991 (absent in our experiments and deliberately excluded from our analysis) also makes direct comparison with the TransCom period of 1992–1996 [Gurney *et al.*, 2004] quantitatively difficult [Jones and Cox, 2001; Lucht *et al.*, 2002]. The models assessed here, in common with the other C⁴MIP models, also lack some relevant and possibly significant carbon cycle processes which are important in determining present-day carbon sinks and variability. As such, they may not be expected to exactly recreate the observed record.

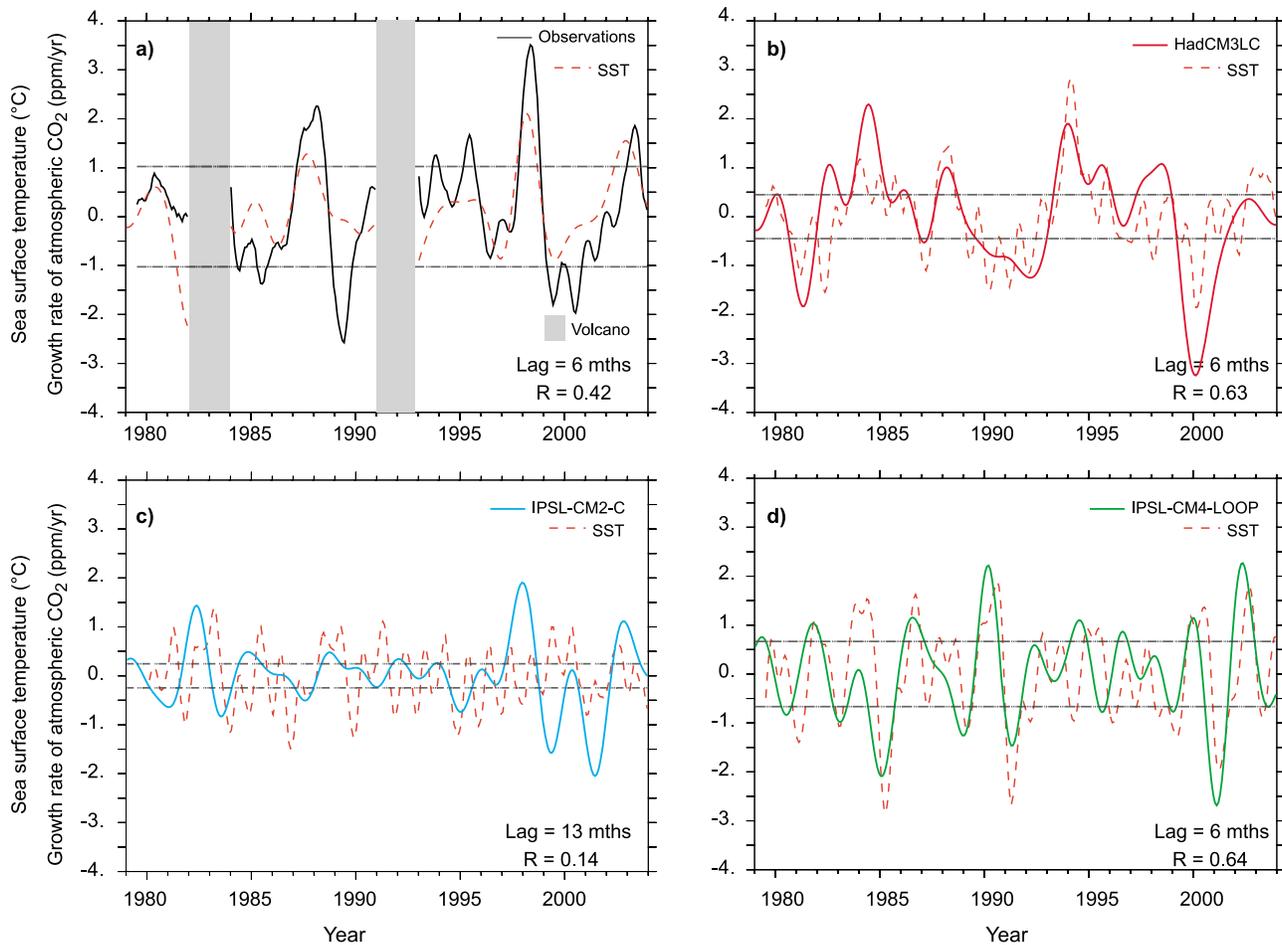


Figure 9. Interannual variability of atmospheric CO₂ growth rate at Mauna Loa (solid line) and sea surface temperature (SST) anomalies in the Niño three index region (Niño 1.2 for IPSL-CM2-C; see Figure 11) for the observations and the three models. The horizontal solid lines represent $0.7 \times \sigma$ of the SST variability. This threshold is used to define El Niño (positive anomalies) and La Niña (negative anomalies).

5.2.2. Key Missing Processes

[38] 1. Nutrient cycling: Nutrient limitation is likely to affect the large-scale behavior of terrestrial carbon sinks due to both remineralization of nitrogen in soils under climate warming and anthropogenic nitrogen deposition from industrial activity and fertilizer application [Magnani *et al.*, 2007; Hungate *et al.*, 2003; Sokolov *et al.*, 2008].

[39] 2. Land use change and management: Over the 20th century, land use effects are important for both regional and global carbon balance [McGuire *et al.*, 2001; Brovkin *et al.*, 2004; Bala *et al.*, 1985] and for land physical properties and their implication on the climate [Bonan *et al.*, 1997; Betts, 2001; Govindasamy *et al.*, 2001; Davin *et al.*, 2007]. Although the C⁴MIP experiments included prescribed CO₂ emissions from land use change [Houghton and Hackler, 2002], the land surface in the models was not directly disturbed, meaning aspects of land use change, such as regrowth of forest on abandoned crop areas, were not represented. There was also no account taken of potentially important fluxes due to agriculture and land management [Bondeau *et al.*, 2007; Friend *et al.*, 2007].

[40] 3. Changes in direct:diffuse light: Diffuse light can penetrate the forest canopy more effectively than direct light and stimulate photosynthesis in lower leaf layers [Alton *et al.*, 2007; Still *et al.*, 2009]. Anthropogenic aerosols have reduced the amount of sunlight reaching the Earth's surface in recent decades [Stanhill and Cohen, 2001; Liepert, 2002; Niyogi *et al.*, 2004], which apart from the effect on climate will directly affect vegetation productivity through changes in the direct:diffuse partitioning of solar radiation [Knobl and Baldocchi, 2008; Mercado *et al.*, 2009]. Volcanic aerosols may also have played a similar role [Angert *et al.*, 2004].

[41] 4. Ozone: Ozone is known to negatively impact regional plant production and terrestrial carbon cycle [Felzer *et al.*, 2004; Sitch *et al.*, 2007], but is also believed to impact the ocean carbon fluxes [Andrew *et al.*, 2009].

[42] 5. Fire: Some, but not all, models include a mechanistic representation of natural fire which affects both the long term and IAV of terrestrial carbon fluxes [Langenfelds *et al.*, 2002; van der Werf *et al.*, 2006] and the seasonal cycle at tropical monitoring stations [Wittenberg *et al.*,

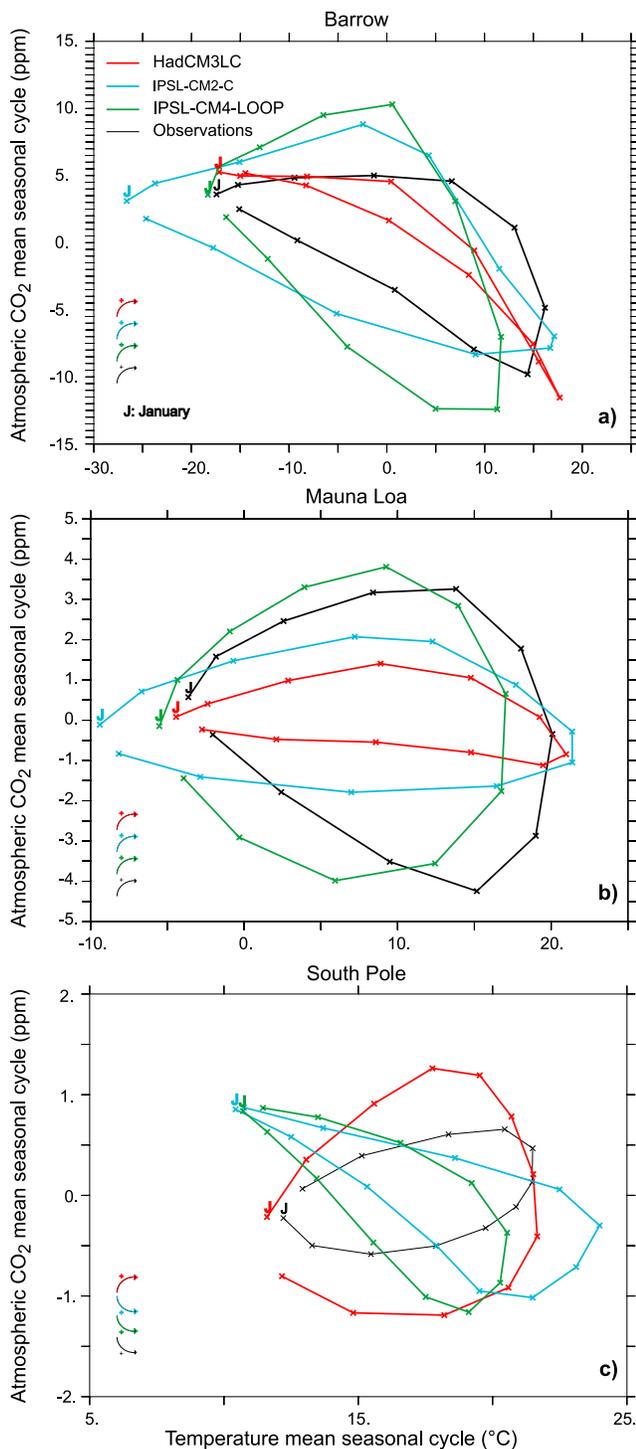


Figure 10. CO₂-temperature diagram showing the seasonal course of atmospheric CO₂ at the three illustrative stations for (a) BRW, (b) MLO, and (c) SPO as a function of surface temperature averaged over the continental region upwind of the station. The letter “J” represents the month of January; the small arrows in the bottom left corners represent the direction (clockwise or anticlockwise) of the seasonal cycle.

1998]. Models without fire (such as HadCM3LC and IPSL-CM2-C) which capture the observed IAV sensitivity to ENSO variability may be doing so for the wrong reasons. Even models which do include fire activity generally do not represent anthropogenic fires, which may themselves be linked to climate variations and contribute to the IAV of carbon fluxes [van der Werf *et al.*, 2009; Field *et al.*, 2009].

[43] 6. Dynamic vegetation: Some, but not all, models include simulation of the fractional vegetation cover of the land surface. Those that do thus include an additional process which may affect the future carbon balance of the biosphere, due to changes in biome cover. But in doing so they may have errors in the control (preindustrial) vegetation simulation which subsequently affect their ability to simulate the CO₂ traits examined here. For example, the seasonal cycle of CO₂ flux may be sensitive to the simulated land cover.

5.3. Seasonal Cycle

[44] The poor performance of HadCM3LC and IPSL-CM2-C (Figure 12) at reproducing the seasonal cycle traits warrants further investigation. The models simulate an early drawdown of CO₂ at all monitoring stations, in particular for those at high northern latitudes. In addition, HadCM3LC and IPSL-CM2-C underestimate the amplitude of the seasonal cycle of CO₂ at Mauna Loa. Subsequent analyses of the regional net and gross carbon fluxes and those at individual flux towers identified several causes. The early drawdown in CO₂ is due to an advanced onset of the growing season with simulated bud burst and leaf growth for deciduous vegetation occurring too early in the spring. The simulated seasonal net flux of CO₂ over North America is out of phase with inversion estimates, and indeed, during the second half of the summer the net carbon flux changes from a sink to a source of atmospheric CO₂. This is largely responsible for the low-amplitude seasonal cycle simulation for HadCM3LC at Mauna Loa. Several factors compound to give this result. From the gross fluxes it is evident that the model simulates an early onset of the growing season as compared with inversion estimates and data from temperate North American flux sites. By the middle of the summer the limited soil water supply is exhausted, leading to a reduction in plant production, indicative of a plant drought response. Also, HadCM3LC simulates higher seasonal amplitude variations in heterotrophic respiration than the IPSL models, with a large midsummer peak over temperate North America. This is likely due to too cold soil temperature in winter leading to too small winter respiration rates, especially in snow-covered regions, rather than the respiration sensitivity to temperature.

[45] HadGEM2-ES is the next generation Earth System model of the Hadley Centre [Collins *et al.*, 2008]. It uses the MOSES-TRIFFID land surface scheme of HadCM3LC, with improved representations of canopy photosynthesis and light interception with the adoption of a multilayer approach [Mercado *et al.*, 2007], and the RothC soil biogeochemistry scheme [Jenkinson, 1990; Jones *et al.*, 2005]. In light of our seasonal cycle evaluation of HadCM3LC we have made several important modifications to HadGEM2-ES. We now delay the onset of the growing season using a 5°C growing

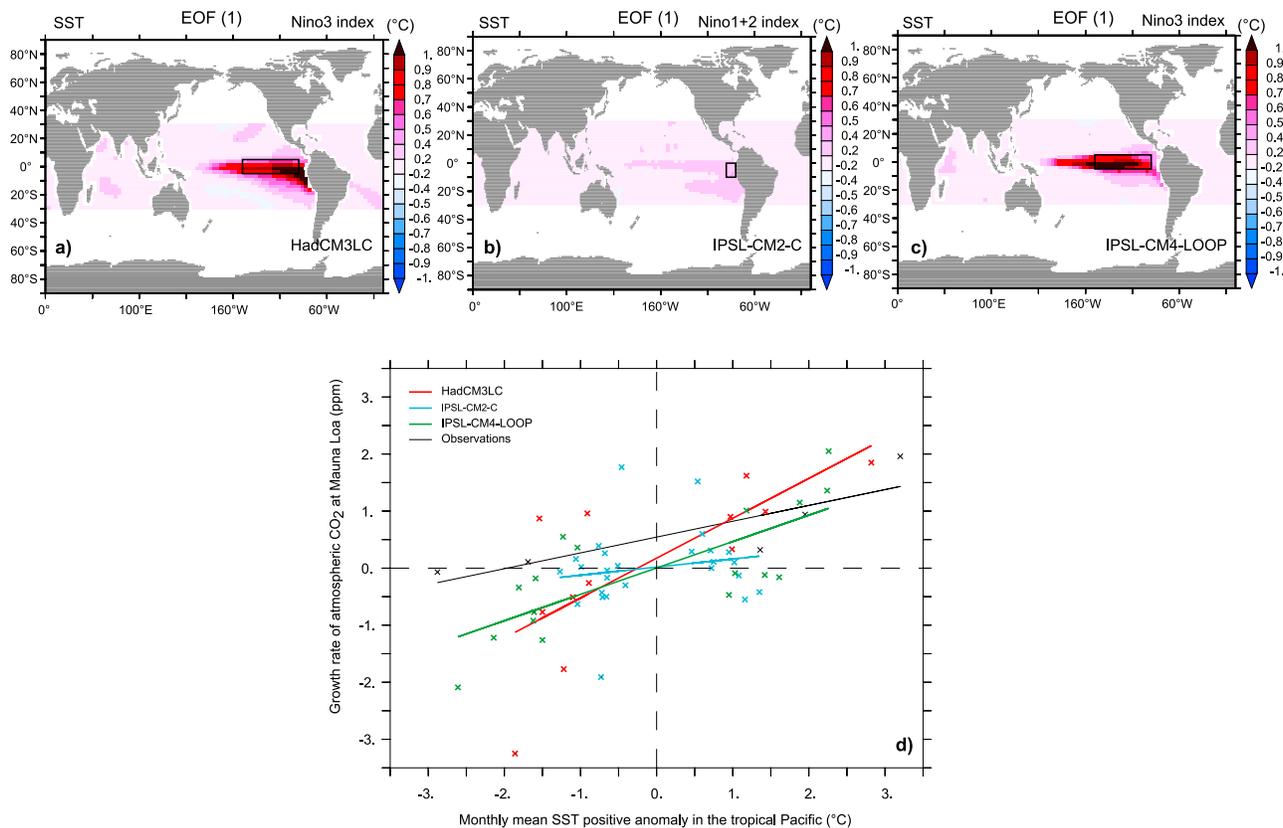


Figure 11. (a) First empirical orthogonal function (EOF) of HadCM3LC model SST variability showing the spatial pattern of the ENSO variability. The box in the Eastern Tropical Pacific represents the Niño three index region for which SST anomalies are calculated. (b) Same as in Figure 11a for IPSL-CM2-C except that the Niño 1.2 region was used here because of the too low variability of the model. (c) Same as in Figure 11a for IPSL-CM4-LOOP. (d) CO₂-temperature diagram showing the positive and negative anomalies of atmospheric CO₂ growth rate as a function of anomalies of Eastern Tropical Pacific SST.

degree base for the deciduous vegetation phenology, as used in the LPJ-DGVM [Sitch *et al.*, 2003]. The winter cold bias in top soil temperature under snow [Wiltshire, 2006] was causing too low respiration rates in winter and too high in summer. In addition, the majority of soil organic carbon (between 56% and 71%) resides in the top 40 cm for grasslands, shrublands, and forest ecosystems [Jobbágy and Jackson, 2000]. Therefore it is more realistic to use the simulated soil temperature of the second layer (from 10–35 cm depth) to drive microbial respiration, thus leading to an improved (reduced) seasonal amplitude variation in heterotrophic respiration. Addressing the physical cold bias by adopting an arguably more appropriate deeper soil temperature rather than reducing the respiration sensitivity to temperature avoids changing our future sensitivity to climate change for the wrong reason. Combined, these changes contribute to the superior simulation of the CO₂ seasonal cycle at the global network of monitoring stations, and at flux sites, by HadGEM2-ES compared with HadCM3LC [Collins *et al.*, 2008].

[46] At the various stations influenced by air masses from the Northern Hemisphere, IPSL-CM4-LOOP simulates both the phase and amplitude of the CO₂ climatologic seasonal

cycle in agreement with the observations. However, this model overestimates the change in amplitude of the CO₂ climatologic seasonal cycle at these same stations, conferring on IPSL-CM4-LOOP a very bad score for the SC peak-to-peak trait. While several factors can explain this, it should be noted that IPSL-CM4-LOOP atmospheric model simulates a cold bias in the high latitudes of the Northern Hemisphere which may have an effect on the sensitivity, to temperature increase, of the terrestrial biosphere. Furthermore, when the land carbon cycle is activated, the cold bias is amplified by the parameterization of the albedo of the fallen leaves which is greater than the one of the soil, inducing a reduction of the radiative energy absorbed by the surface. As an example, over the 1982–2002 period, north to 50°N, when compared to the Global Inventory Monitoring and Modeling Studies (GIMMS) data [Tucker *et al.*, 2004], IPSL-CM4-LOOP overestimates the increase of the Leaf Area Index (LAI) monthly trend, by 30% in average over the growing season (March to August) and by 160% in maximum (in July for GIMMS data and in August for IPSL-CM4-LOOP).

[47] Preliminary studies with IPSL-ESM, the new IPSL Earth system model, indicate that the integration of clouds

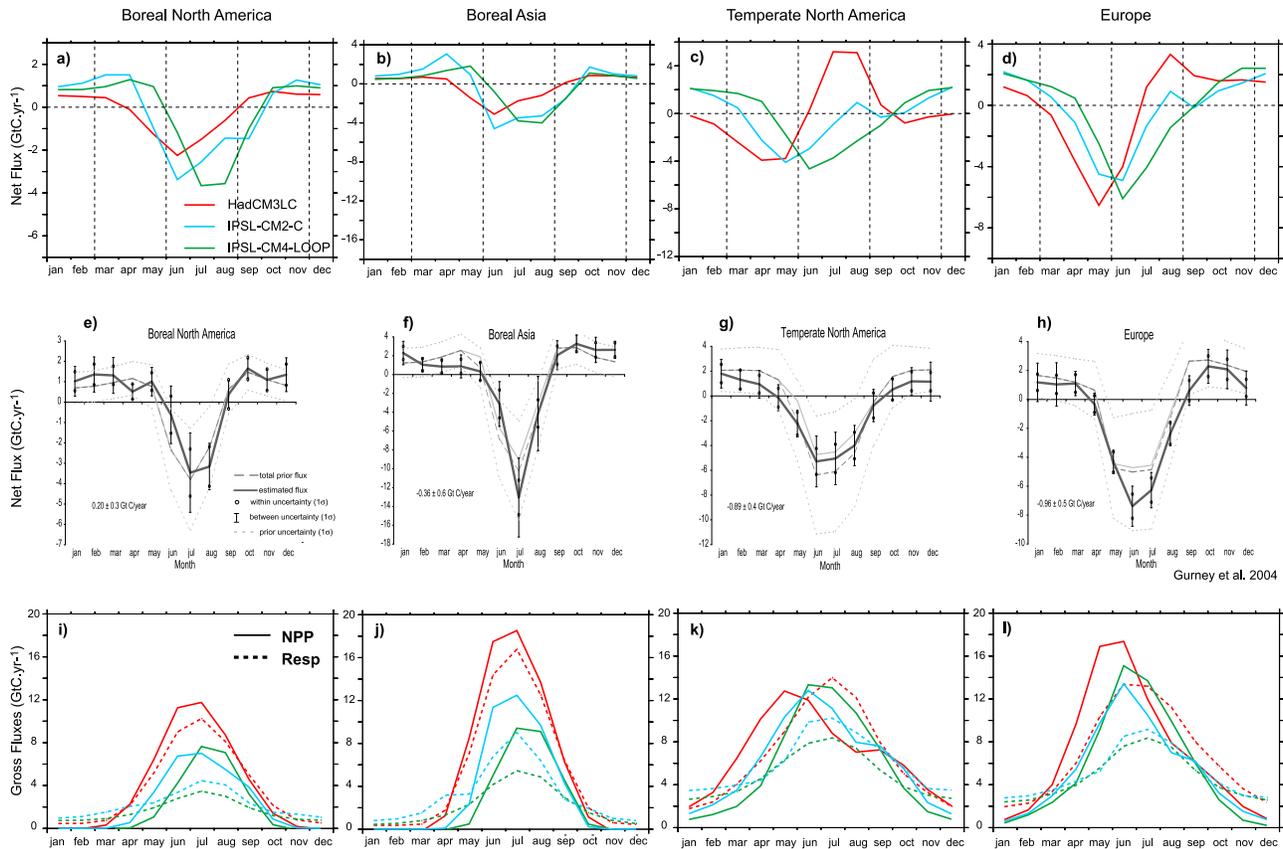


Figure 12. (a to d) Net CO₂ flux averaged over four regions of the Northern Hemisphere ((Figure 12a) boreal North America, (Figure 12b) boreal Asia, (Figure 12c) temperate North America, and (Figure 12d) Europe) for the three models. Positive values denote a source of CO₂ to the atmosphere. (e to h) Net CO₂ flux (PgC yr⁻¹) averaged over the same regions, as derived from TransCom-3 atmospheric CO₂ inversions [Gurney *et al.*, 2004]. The error bar represents the intermodel uncertainty. (i to l) Net Primary Productivity (NPP, solid line) and heterotrophic respiration (Resp, dashed line) averaged over the same four regions for the three models. All fluxes are in PgC yr⁻¹.

microphysics and increase in the spatial horizontal resolution of the atmospheric model both augment the global surface temperature and thus reduce the cold bias.

5.4. Interannual Variability

[48] As for the seasonal cycle, the interannual variability of atmospheric CO₂ growth rate is dominated by the variability in the terrestrial net carbon fluxes [Keeling *et al.*, 1996; Battle *et al.*, 2000; Bousquet *et al.*, 2000; Le Quére *et al.*, 2003; Baker *et al.*, 2006]. Several studies have shown the ENSO-induced climate variations are primarily responsible for carbon anomalies in the tropics, but also in temperate and boreal regions [Jones *et al.*, 2001; Patra *et al.*, 2005; van der Werf *et al.*, 2004]. The response of tropical fluxes to El Niño Southern Oscillation climate anomalies is the most documented [Gurney *et al.*, 2003, 2004; Rodenbeck *et al.*, 2003; Baker *et al.*, 2006]. ENSO temperature and precipitation anomalies cause changes in plant productivity, ecosystem respiration, and biomass burning. Positive anomalies in the atmospheric CO₂ growth rate are often due to warmer and dryer El Niño events, with

carbon being released by the terrestrial biosphere. Negative anomalies correspond to the wetter and cooler La Niña events.

[49] For a given model a bad score to the interannual variability evaluation traits may correspond to a deficiency of the model to reproduce ENSO variability in terms of ocean-atmosphere coupling dynamics, climate sensitivity, and teleconnections, and carbon cycle response to climate local variability.

[50] The simulated SST anomaly over the specific Niño box for the ENSO positives phases is on average 1.48 ± 0.77 , 0.85 ± 0.28 , and $1.57 \pm 0.52^\circ\text{C}$ for HadCM3LC, IPSL-CM2-C, and IPSL-CM4-LOOP, respectively, compared to the average observed anomaly of $2.48 \pm 0.94^\circ\text{C}$. Conversely, for the negative phase of the ENSO, the average SST anomaly is -1.22 ± 0.38 , -0.75 ± 0.24 , and $-1.72 \pm 0.45^\circ\text{C}$ for HadCM3LC, IPSL-CM2-C, and IPSL-CM4-LOOP, respectively, compared to the average observed anomaly of $-1.97 \pm 0.60^\circ\text{C}$. Both HadCM3LC and IPSL-CM4-LOOP thus show ENSO SST dynamics relatively similar to the observations, while IPSL-CM2-C largely

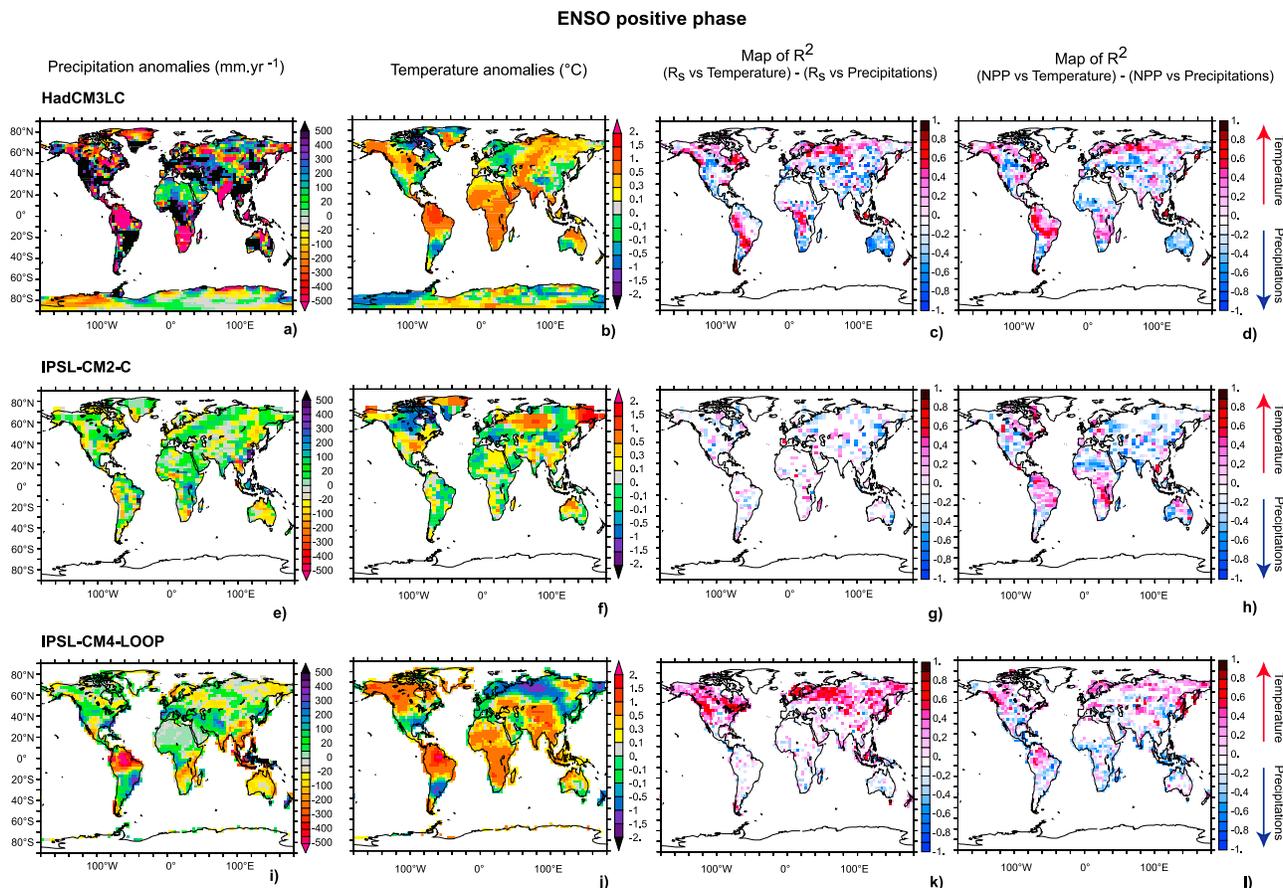


Figure 13. Modeled El Niño variability. (a) Spatial distribution of continental precipitation anomalies (mm yr^{-1}) simulated by the HadCM3LC model for an El Niño composite calculated as the average of the El Niño events defined in Figure 8. (b) Same as Figure 13a for surface temperature. (c) El Niño time scale correlation between heterogeneous respiration (RESP) and climate, defined as the difference between the RESP-temperature coefficient of determination and the RESP-precipitation coefficient of determination. Positive (red) values mean that heterogeneous respiration follows temperature during El Niño events, while negative (blue) values mean that heterogeneous respiration follows precipitation during El Niño events. White regions are either regions which do not show significant correlation between respiration and climate, or regions where precipitation and temperature are equally important at driving respiration anomalies. (d) Same as in Figure 13c but for NPP. (e to h and i to l) same as Figures 13a–13d but for the IPSL-CM2-C and IPSL-CM4-LOOP models, respectively.

underestimates it, because of a coarser ocean spatial resolution compared to HadCM3LC and IPSL-CM4-LOOP, as suggested by Meehl *et al.* [1993] and Knutson *et al.* [1997]. To confirm the capability of both HadCM3LC and IPSL-CM4-LOOP to reproduce the ENSO variability, the maximum entropy power spectra of the SST, from the equatorial Pacific, were calculated. HadCM3LC and IPSL-CM4-LOOP simulate a maximum of the power spectra between 2 and 7 years over 300 years of the model control run [see Jones *et al.*, 2001, Figure 5] and between 2 and 7 years over 250 years of the model control run, respectively, which is the typical range of El Niño frequency [Randall *et al.*, 2007].

5.4.1. ENSO Positive Phase

[51] Sich *et al.* [2008] have shown that during the positive phase of the ENSO, the sensitivity of the terrestrial

biosphere simulated by TRIFFID and ORCHIDEE is in agreement, at global scale, with the observations, when the models are forced by the observed climate. Furthermore, Jones *et al.* [2001] have shown that, at global scale, the carbon cycle simulated by HadCM3LC responds to the simulated climate anomalies in agreement with the observations (Figure 13). Note that as the decomposition and oxidation of carbon soil (R_s) and Net Primary Productivity (NPP) response is mostly dominated by the temperature anomalies, which are in agreement with the observations, this explains the good performance of HadCM3LC for the γ_{LAV} trait even if the HadCM3LC overestimates the precipitation anomalies.

[52] During the positive phase of ENSO, IPSL-CM2-C does not reproduce the observed temperature and precipita-

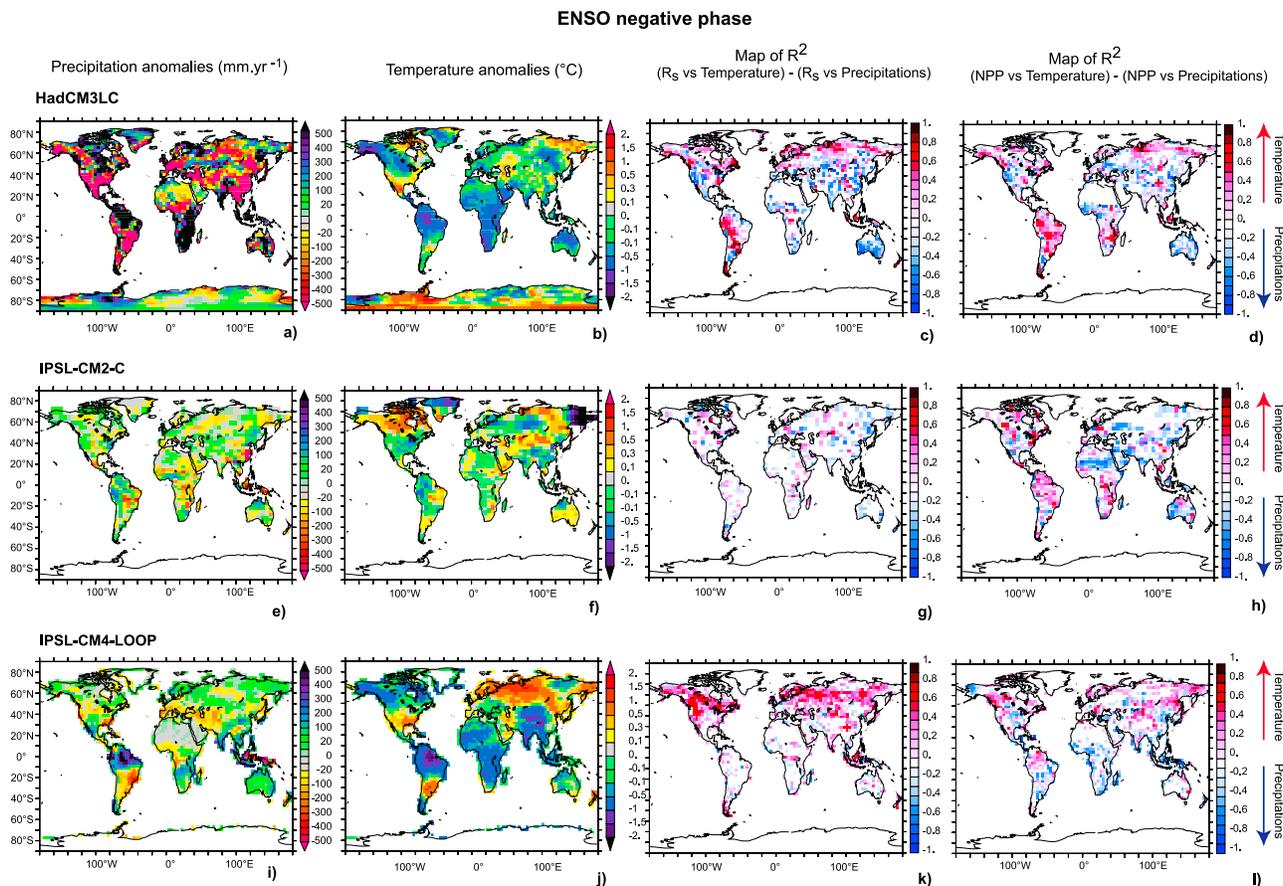


Figure 14. Same as Figure 13, but for La Niña events.

tion patterns (Figure 13), which explains the poor performance of IPSL-CM2-C for the interannual variability traits.

[53] Both the temperature and precipitation anomalies pattern simulated by IPSL-CM4-LOOP are in general good agreement with the observations, except for precipitation in South Africa and tropical Asia where the model simulates a reduction that is opposite to the observations (Figure 13). As precipitation anomalies control component carbon fluxes, both for NPP and for decomposition and oxidation of carbon soil and fire disturbances (R_t), in these regions, IPSL-CM4-LOOP overestimates the effect of carbon fluxes from South Africa and tropical Asia to the positive anomaly of the CO_2 growth rate. This explains the poor performance of IPSL-CM4-LOOP in terms of interannual variability, notably at MLO.

5.4.2. ENSO Negative Phase

[54] The precipitation anomalies induced by the negative phase of ENSO are well reproduced by HadCM3LC (except for Africa), while the temperature anomalies are clearly misrepresented (Figure 14): HadCM3LC simulates a cooling which spans the whole globe with a greater intensity in North America and in the tropics. As a consequence, this cooling induces an increase in land carbon storage and explains the

abnormal simulated negative net carbon flux anomaly which explains the difference, between HadCM3LC and the observations, of sensitivity of the atmospheric CO_2 to the interannual variability, which impacts negatively the performance of HadCM3LC for the γ_{IAV} trait.

[55] As for the positive phase of ENSO, IPSL-CM2-C does not reproduce the observed patterns of variability in temperature and precipitation during the negative phase of ENSO (Figure 14). As stated above, these deficiencies arise from the poor representation of ENSO dynamics in this model and explain the poor performance of IPSL-CM2-C for the interannual traits.

[56] During the negative phase of the ENSO, IPSL-CM4-LOOP simulates an increase of precipitation, in agreement with the observations over the Amazonian rainforest but in opposition to the observations over a large part of central Africa and over tropical Asia. As HadCM3LC and in opposition to observations, IPSL-CM4-LOOP simulates a cooling which spans over the whole tropical region reaching -2°C in the Amazonian basin (Figure 14). IPSL-CM4-LOOP deficiencies to simulate precipitation and temperature anomalies lead to an increase in NPP in the South America region and to a decrease of R_t in tropical Asia and to a lesser

extent in Africa. Both this increase of NPP and decrease of Rt could be at the origin of IPSL-CM4-LOOP poor performance for the γ_{IAV} trait.

5.5. From Seasonal and Interannual Variability to Climate Change

[57] It is not yet clear to what extent, if at all, these metrics will constrain future model projections, but their consistent use to evaluate late 20th century ESM simulations will provide valuable guidance on model performance and avenues for future model development and improvement. An ultimate goal would be to link such observed quantities with future behavior in ESMs and hence be able to derive constraints, e.g., through weighting of individual members within an ensemble of future projections. Analysis of multimodel ensembles that exist already (e.g., from C⁴MIP [Friedlingstein et al., 2006]) or are planned (e.g., for IPCC AR5 [Hibbard et al., 2007]) and perturbed parameter ensembles of single models (e.g., with HadCM3C (B. Booth et al., Global warming uncertainties due to carbon cycle feedbacks exceed those due to CO₂ emissions, manuscript in preparation, 2010)) will enable progress toward this goal.

[58] A perturbed-parameter ensemble of HadCM3LC climate simulations (QUMP; Murphy et al. [2004]) showed a significant shift in the frequency distribution of climate sensitivity when each member was weighted according to a Climate Prediction Index: a quantitative measure of model skill akin to those proposed in this study. However, how do we know that better simulation of the mean state gives a better simulation of sensitivity to forcing? Whilst it is undeniably useful to use observations to evaluate models, identify deficiencies, and guide development and improvement, it is not yet clear if such quantitative use of metrics provides a reliable constraint on future projections. There are also methodological problems of “double counting”: it may not be valid to use observations in such a way if the same data have been used for model development. Some independent data may have to be withheld for the metrics-based evaluation.

[59] Model-observation comparisons can, though, be a valuable guide to process understanding and thus provide a guide to which aspects of model performance are reliable and which processes or regions are more critical for future projections. For example, a principal component analysis showed that IAV is predominantly determined by fluxes from the tropics, whereas SC was determined predominantly by northern extratropical land fluxes. Hence our evaluation of IAV may provide a constraint on tropical behavior more so than on extratropical behavior, and the SC may constrain the high-latitude behavior. Raddatz et al. [2007] showed that future feedback in the C⁴MIP models was strongly controlled by tropical behavior on land, so we might suppose that IAV analysis gives some constraint on those future projections. Yet, very large amounts of carbon are stored in boreal forests with high-latitude soil carbon changes especially important for future climate-carbon cycle feedbacks, so the SC behavior is clearly crucial as well.

[60] For future changes to be constrained by observable quantities we must be confident they are driven by the same,

or at least similar, processes. For example, our analysis in terms of temperature sensitivity of carbon fluxes implicitly assumes that other climate variables change with temperature in a fixed pattern: clearly, changes in other environmental conditions (precipitation, light levels, etc.) affect carbon fluxes as well, but we can use temperature, T , as a proxy for the wider “climate” if the patterns of change scale simply with T . If future changes in other climate variables follow the same scaling with T , then the observed behavior may well act as a strong constraint on the future.

[61] For the seasonal cycle, especially in the high latitudes where we have identified it as most important, carbon fluxes are strongly linked with light availability. In the future we expect high latitudes to warm faster than other regions of the world but we do not expect changes in the seasonal cycle of day length. Also for precipitation, P , we do not necessarily expect the seasonal $P:T$ relationship to hold for climate change of P and T : over a seasonal cycle, higher T in summer is generally associated with lower P , but under climate change, higher T may be accompanied by increased P .

[62] Similarly for IAV-, can we be confident that general climate patterns in future will be related to T in the same way as they are on interannual time scales? To some extent this may be true as some, although not all, GCMs predict future climate change to resemble El Niño-like patterns of climate change [Cox et al., 2004]. But even so, we do not know if the long-term response of terrestrial carbon fluxes to given environmental conditions will be the same as to short-term variability. Possible acclimation of processes such as soil decomposition is still hotly debated [Giardina and Ryan, 2000; Davidson et al., 2000; Knorr et al., 2005; Davidson and Janssens, 2006].

[63] Over the tropical ocean, IAV due to ENSO is driven predominantly by upwelling of carbon-rich deep water [Jones et al., 2001; Schneider et al., 2008] with solubility effects of SST variations secondary. Hence future SST patterns may play less of a crucial role in ocean CO₂ fluxes than changes in global ocean circulation of dissolved inorganic carbon and nutrients.

[64] More research is required to determine which observed quantities are related (at least in the models) to key future behavior. Then we may be able to derive useful constraints.

5.6. Land Carbon Cycle Versus Ocean Carbon Cycle

[65] As detailed above, the analysis we present here is mainly focused on the land carbon cycle. This is largely due to the intrinsic nature of atmospheric CO₂ variability on seasonal to interannual time scales, for which the land carbon fluxes clearly dominate. As such, the metrics we develop here would mainly provide constraints on model performance for the land carbon cycle component of ESMs.

[66] Similarly to the land carbon fluxes, the ocean carbon fluxes will contribute both to the evolution of atmospheric CO₂ over the next decades and to the magnitude of the climate-carbon feedback. The analysis of the multimodel ensemble from C⁴MIP also clearly demonstrates the large uncertainty associated with the ocean carbon flux response to climate change [Friedlingstein et al., 2006] and the need

for a more thorough evaluation of the ocean carbon cycle component of ESMs.

6. Conclusion

[67] We have defined a new set of performance metrics to test coupled carbon-climate models against atmospheric CO₂ observations. Land fluxes are primarily tested. Three state-of-the-art global models were evaluated using these new metrics, which contain multiple time scale constraints based upon the (1) long-term CO₂ trend, i.e., the global carbon budget, (2) seasonal cycle of CO₂, (3) interannual variability of the CO₂ growth rate forced by ENSO climate anomalies, and (4) sensitivity to climatic variations. Time-varying fluxes from the HadCM3LC, IPSL-CM2-C, and IPSL-CM4-LOOP models of the C⁴MIP ensemble are used as input of the global LMDZ4 transport model, and the modeled CO₂ field is compared to station observations. For the seasonal variability analysis, an inversion of the LMDZ4 transport is applied to determine the regional footprint of CO₂ at each station. Using many stations and different statistical measures of the model-data misfit at multiple time scales, we define a single mark, which allows us to compare models with each other, and against data, at a glance.

[68] We found that the IPSL-CM4-LOOP model is best able to reproduce the phase and the amplitude of atmospheric CO₂ at northern stations. The two other models generally underestimate the seasonal amplitude. This points out their shortcomings in describing the vegetation phenology and the heterotrophic respiration response to climate. Such misfit of the seasonal cycle is worrisome because the future response of the land carbon cycle to climate change is certainly highly seasonal and has a different sign between spring and autumn [Piao *et al.*, 2008]. One way to overcome this difficulty is to improve both the seasonal cycle of climate in climate models and the carbon cycle response to it, for instance by using satellite data to constrain phenological processes [Botta *et al.*, 2000] and eddy covariance networks to constrain the timing of respiration versus photosynthesis [e.g., Kucharik *et al.*, 2006].

[69] We also found that IPSL-CM2-C produces a climate-driven abnormal source of CO₂ to the atmosphere during El Niño-like anomalies and, conversely, an abnormal CO₂ sink during La Niña-like anomalies. Nevertheless it is encouraging to see that this fundamental carbon cycle response to climate is properly encapsulated in the three models; the magnitude of such response differs among models. The three climate models all underestimate the SST warm anomaly during El Niño and, to a lesser extent, the cold anomalies during La Niña. ENSO-driven changes in precipitation regimes are also not so well captured, except by HadCM3LC, a climate bias which translates into a CO₂ flux bias. Overall, HadCM3LC does best in reproducing the interannual CO₂ variability in response ENSO forcing.

[70] This example shows that multiple time scales are needed to evaluate models and that a model which does best on seasonal scales may not necessarily be outperforming others on interannual time scales. The advantage of defining a single metric, although the procedure is a bit rigid and

looks tedious to implement, is that it allows for testing of future structural improvements of models and inclusion of new processes in the same rigorously defined framework. Here we have shown that the signal of atmospheric CO₂ concentration, despite being a coarse-scale and process-integrated signal, has a great usefulness in falsifying models against observations.

[71] More efforts are needed to further improve such metrics for assessing the sensitivity of the carbon cycle to climate change. A next logical step for benchmarking terrestrial carbon cycle models would be to use eddy covariance observations of flux-climate relationships [Reichstein *et al.*, 2007; Ito *et al.*, 2008], global satellite information [Running, 2006], and manipulative experiments of warming or altered precipitation at ecosystem scale [Luo *et al.*, 2008; Gerten *et al.*, 2008]. Further, the definition of metrics should be extended to assess ocean carbon cycle models against observations, using, for instance, ocean flux observations [Takahashi *et al.*, 2009], ocean interior inventories [Sabine *et al.*, 2004], and atmospheric tracers such as atmospheric potential oxygen [Naegler *et al.*, 2007]. Finally, the availability of observational data over longer time series is essential in view of benchmarking models over the longest possible periods which is necessary, even if not sufficient, to build confidence in future projections.

Appendix A: Model Description

A1. HadCM3LC

[72] HadCM3LC [Cox *et al.*, 2000] is based on the third Hadley Centre coupled ocean-atmosphere model, HadCM3 [Gordon *et al.*, 2000], coupled to an ocean carbon cycle model HadOCC and a dynamic global vegetation model TRIFFID. The ocean carbon cycle model, HadOCC [Palmer and Totterdell, 2001], simulates the movements of carbon within the ocean system, including exchange of CO₂ with the atmosphere, the circulation of dissolved inorganic carbon within the ocean, and the cycling of carbon by the marine biota. It explicitly represents ocean carbonate chemistry, and the biological component of HadOCC is an explicit ecosystem model consisting of the four components: nutrient (assumed to be nitrate), phytoplankton, zooplankton, and (sinking) detritus. The terrestrial carbon cycle model, TRIFFID [Cox, 2001], defines the state of the terrestrial biosphere in terms of the soil carbon and the structure and coverage of five plant functional types (broadleaf tree, needleleaf tree, C₃ grass, C₄ grass, and shrub) within each model grid box. The areal coverage, leaf area index, and canopy height of each type are updated based on a carbon balance approach, in which vegetation change is driven by net carbon fluxes calculated within the MOSES-2 tiled land surface scheme [Essery *et al.*, 2003]. Carbon fluxes for each of the vegetation types are derived using the coupled photosynthesis-stomatal conductance model developed by Cox *et al.* [1998] which utilizes existing models of leaf-level photosynthesis in C₃ and C₄ plants [Collatz *et al.*, 1991, 1992]. Net primary productivity is allocated into the growth of existing vegetation biomass and to expansion of the vegetated area in each grid cell. Leaf

Table C1. Principal Component Analysis at the 12 Stations for the Seasonal Cycle Based on Peylin's Inversions^a

Station Code	Terrestrial Regions											Percent Variance
	Boreal North America	Temperate North America	Boreal Asia	Temperate Asia	Europe	Northern Africa	Tropical America	Tropical Asia	Australia	Southern Africa	Temperate South America	
	ALT	22	8	31	2	21	4	1	0	1	1	
BRW	25	5	29	1	14	2	1	0	0	1	0	
MHD	27	12	20	3	26	5	1	0	0	2	1	
SCH	17	15	13	3	37	6	1	1	0	3	1	
NWR	11	22	19	5	25	1	3	3	0	3	1	
AZR	28	14	20	4	24	3	0	1	0	3	1	
KUM	13	14	24	4	22	2	4	3	0	2	0	
MLO	11	15	18	4	19	8	6	5	0	1	2	
SMO	4	6	5	3	10	20	4	1	2	21	13	
AMS	3	6	4	2	8	18	9	1	1	19	5	
CGO	3	5	4	2	7	16	8	1	9	17	6	
SPO	4	7	5	3	9	21	10	1	0	22	5	

Station Code	Oceanic Regions											Percent Variance
	Northern Ocean	North Atlantic	North Pacific	West Pacific	East Pacific	Tropical Atlantic	Tropical Indian	South Atlantic	South Pacific	South Indian	Southern Ocean	
ALT	0	0	0	0	0	0	0	0	0	0	0	96
BRW	20	0	0	0	0	0	0	0	0	0	0	94
MHD	1	0	0	0	0	0	0	0	0	0	0	93
SCH	0	1	1	0	0	0	0	0	0	0	0	88
NWR	1	2	2	0	0	0	0	0	0	0	0	85
AZR	1	0	0	0	0	0	0	0	0	0	0	92
KUM	1	5	5	1	0	0	0	0	0	0	0	93
MLO	1	4	4	1	0	0	1	0	0	0	0	93
SMO	0	0	0	4	2	0	1	1	1	0	3	58
AMS	0	1	1	0	0	0	1	2	2	11	6	52
CGO	0	1	1	0	0	0	1	2	3	5	8	60
SPO	0	1	1	1	0	0	1	2	1	0	8	63

^aTerrestrial biosphere and ocean, which were all transported with LMDZ4 and filtered according to *Thoning et al.* [1989] (see section 3.2). The codes for the stations are expanded in Table 1. Peylin's inversions are for terrestrial biosphere and ocean. At each station, the contribution from all regions (see Table 2 for details on the TransCom-3 regions) sums up to 100%. The third subtable shows the representativeness of the 1st mode of the PCA.

phenology (bud burst and leaf drop) is updated, using accumulated temperature-dependent leaf turnover rates.

A2. IPSL-CM2-C

[73] IPSL-CM2-C [*Dufresne et al.*, 2002] shows results of the coupling between the IPSL ocean-atmosphere general circulation model IPSL-CM2 [*Khodri et al.*, 2001] and the two carbon cycle models, IPSL-OCCM1 [*Aumont et al.*, 1999] for the ocean part, and the Scheme for Large-Scale Atmosphere Vegetation Exchange (SLAVE) model for the terrestrial part [*Friedlingstein et al.*, 1995a, 1995b]. SLAVE is driven by surface air temperature, precipitation, and solar radiation and calculates NPP following a light use efficiency formulation [*Field et al.*, 1995] that is a function of temperature and water stress. NPP increases with CO₂ under a Michaelis-Menten beta factor formulation [*Gifford*, 1992], which has a global value of 0.5, in the upper range of experimental data [*DeLucia et al.*, 1999], although nitrogen limitation and deposition as well as vegetation dynamics and land use changes are ignored in this study. The ocean carbon model, IPSL-OCCM1 [*Aumont et al.*, 1999; *Le Quéré*, 1999], based on the HAMOCC3 biogeochemical scheme [*Maier-Reimer*, 1993], is driven by monthly mean global fields of oceanic circulation, temperature, salinity, and surface

fields of winds, sea ice, and water fluxes all issued from the Ocean-Atmosphere General Circulation Model (OAGCM).

A3. IPSL-CM4-LOOP

[74] IPSL-CM4-LOOP results of the coupling between the IPSL coupled ocean-atmosphere general circulation model IPSL-CM4 [*Marti et al.*, 2005] and the two carbon cycle models: the Pelagic Interactions Scheme for Carbon and Ecosystems Studies (PISCES) biogeochemical model [*Aumont et al.*, 2003] for the ocean part and the Organizing Carbon and Hydrology in Dynamic Ecosystems (ORCHIDEE) model for the terrestrial part [*Krinner et al.*, 2005]. PISCES includes three nutrients, two phytoplanktons, two zooplanktons, one detritus, and semilabile dissolved organic matter. It explicitly represents the collimation of phytoplankton growth by light and three nutrients: phosphate, iron, and silicate. The phytoplankton reservoir is split in two size fractions corresponding to nanophytoplankton and diatoms. Two sizes of zooplankton (microplankton and mesozooplankton) are also explicitly considered. ORCHIDEE is a dynamic global vegetation model (DGVM) which calculates energy and hydrology budgets, carbon assimilation, allocation and decomposition, and vegetation competition. In these simulations, the natural and agricultural vegetation distributions were prescribed. ORCHIDEE distinguishes

13 plant functional types (PFTs). In every grid point, different PFTs can coexist.

Appendix B: Model Data Statistical Analysis

[75] For the following, x refers to the modeled CO_2 , and o refers to the observed CO_2 .

B1. Pearson Correlation

[76] The Pearson correlation is the ratio of the sample covariance of two variables to the nonnull product of their standard deviations. In the following formula, x refers to the model value, and o to the observations.

$$\Gamma_{xo} = \frac{\sigma_{xo}}{\sigma_x \sigma_o} = \frac{\frac{1}{N} \times \sum_{i=1}^N (x_i - \bar{x}_o)(o_i - \bar{o})}{\sqrt{\frac{1}{N} \times \sum_{i=1}^N (x_i - \bar{x}_o)^2} \times \sqrt{\frac{1}{N} \times \sum_{i=1}^N (o_i - \bar{o})^2}} \quad (\text{B1})$$

[77] The Pearson test is used to evaluate if models and observations values are linearly correlated. Note that the Pearson correlation is subject to outlying data and is bounded by -1 and $+1$.

B2. Normalized Standard Deviation

[78] The normalized standard deviation (NSD) is the ratio of model to observed standard deviations. In the following formula, x refers to the model value, and o refers to the observations.

$$\frac{S}{S_{obs}} = \frac{\sqrt{\frac{1}{N} \times \sum_{i=1}^N (x_i - \bar{x}_o)^2}}{\sqrt{\frac{1}{N} \times \sum_{i=1}^N (o_i - \bar{o})^2}} \quad (\text{B2})$$

[79] A value of $\text{NSD} = 1$ means that the model perfectly captures the magnitude of observed variability. A value lower than 1 corresponds to underestimating the magnitude of observed variability, while a value greater than 1 corresponds to overestimating it.

B3. Models to Observations Deviation

[80] When comparing a single value between models and observations, we use the simple following formula, where x refers to the model value and o refers to the observations.

$$\text{MOD} = \frac{x - o}{x + o} \quad (\text{B3})$$

Appendix C: Contribution to the Seasonal Cycle of Atmospheric CO_2 at Observations Stations

[81] Table C1 shows the principal component analysis at the 12 stations for the seasonal cycle based on Peylin's inversions.

[82] **Acknowledgments.** This study was supported by the European project ENSEMBLES (GOCE-CT-2003-505539) and CarboOcean (511176 (GOCE)). We thank the people who performed the C^4MIP simulations with IPSL-CM2-C. We thank the Centre de Calcul Recherche et Technologie (CEA) and the IT department of the Laboratoire des Sciences du Climat et de l'Environnement. The work of C.D.J. and S.A.S. was supported by the Joint DECC, Defra, and MoD Integrated Climate Programme: DECC/Defra (GA01101), MoD (CBC/2B/0417_Annex C5).

References

- Alton, P., P. North, and S. Los (2007), The impact of diffuse sunlight on canopy light-use efficiency, gross photosynthetic product and net ecosystem exchange in three forest biomes, *Global Change Biol.*, *13*(4), 776–787.
- Andrew, L., A. F. Codron, L. Bopp, N. Metz, P. Cadule, A. Tagliabue, and J. Le Sommer (2009), Stratospheric ozone depletion reduces ocean carbon uptake and enhances ocean acidification, *Geophys. Res. Lett.*, *36*, L12606, doi:10.1029/2009GL038227.
- Angert, A., S. Biraud, C. Bonfils, W. Buermann, and I. Fung (2004), CO_2 seasonality indicates origins of post-Pinatubo sink, *Geophys. Res. Lett.*, *31*, L11103, doi:10.1029/2004GL019760.
- Aumont, O., J. Orr, P. Monfray, G. Madec, and E. Maier-Reimer (1999), Nutrient trapping in the equatorial Pacific: The ocean circulation solution, *Global Biogeochem. Cycles*, *13*(2), 351–369.
- Aumont, O., E. Maier-Reimer, S. Blain, and P. Monfray (2003), An ecosystem model of the global ocean including Fe, Si, P colimitations, *Global Biogeochem. Cycles*, *17*(2), 1060, doi:10.1029/2001GB001745.
- Bacastow, R., C. Keeling, and T. Whorf (1985), Seasonal amplitude increase in atmospheric CO_2 concentration at Mauna Loa, Hawaii, 1959–1982, *J. Geophys. Res.*, *90*(D6), 10,529–10,540.
- Baker, D., et al. (2006), TransCom 3 inversion intercomparison: Impact of transport model errors on the interannual variability of regional CO_2 fluxes, 1988–2003, *Global Biogeochem. Cycles*, *20*, GB1002, doi:10.1029/2004GB002439.
- Bala, B., K. Caldeira, M. Wickett, T. J. Phillips, D. B. Lobell, C. Delire, and A. Mirin (1985), Combined climate and carbon-cycle effects of large-scale deforestation, *Proc. Natl. Acad. Sci.*, *104*(16), 6550–6555, doi:10.1073/pnas.0608998104.
- Battle, M., M. Bender, P. Tans, J. White, J. Ellis, T. Conway, and R. Francey (2000), Global carbon sinks and their variability, inferred from atmospheric O_2 and $\delta^{13}\text{C}$, *Science*, *287*, 2467–2470.
- Betts, R. A. (2001), Biogeophysical impacts of land use on present-day climate: Near-surface temperature change and radiative forcing, *Atmos. Sci. Lett.*, *2*(1–4), 39–51, doi:1006/asle.2001.0023.
- Bonan, G. B., K. J. Davis, D. Baldocchi, D. Fitzjarrald, and H. Neumann (1997), Comparison of the NCAR LSM1 land surface model with BOREAS aspen and jack pine tower fluxes, *J. Geophys. Res.*, *102*(D24), 29,065–29,075, doi:10.1029/96JD03095.
- Bondeau, A., et al. (2007), Modelling the role of agriculture for the 20th century global terrestrial carbon balance, *Global Change Biol.*, *13*, 679–706.
- Botta, A., N. Viovy, P. Ciais, P. Friedlingstein, and P. Monfray (2000), A global prognostic scheme of leaf onset using satellite data, *Global Change Biol.*, *6*, 709–725.
- Bousquet, P., P. Peylin, P. Ciais, C. Le Quééré, P. Friedlingstein, and P. Tans (2000), Regional changes of CO_2 fluxes over land and oceans since 1980, *Science*, *290*, 1342–1346, doi:10.1126/science.290.5495.1342.
- Brovkin, V., S. Sitch, W. von Bloh, M. Claussen, E. Bauer, and W. Cramer (2004), Role of land cover changes for atmospheric CO_2 increase and climate change during the last 150 years, *Global Change Biol.*, *10*, 1–14, doi:10.1111/j.1365-2486.2004.00812.x.
- Buermann, W., B. R. Lintner, C. D. Koven, A. Angert, J. E. Pinzon, C. J. Tucker, and I. Y. Fung (2007), The changing carbon cycle at Mauna Loa Observatory, *Proc. Natl. Acad. Sci.*, *104*, 4249–4254.
- Cadule, P., L. Bopp, and P. Friedlingstein (2009), A revised estimate of the processes contributing to global warming due to climate-carbon feedback, *Geophys. Res. Lett.*, *36*, L14705, doi:10.1029/2009GL038681.
- Cleveland, E. R., R. K. Johnson, and R. W. Mandigo (1983), Index selection and feed intake restriction in swine: I. Effect on rate and composition of growth, *J. Animal Sci.*, *56*, 560–569.
- Collatz, G. J., J. T. Ball, C. Grivet, and J. A. Berry (1991), Physiological and environmental regulation of stomatal conductance, photosynthesis, and transpiration: A model that includes a laminar boundary layer, *Agric. For. Meteorol.*, *54*, 107–136.

- Collatz, G. J., M. Ribas-Carbo, and J. A. Berry (1992), Coupled photosynthesis-stomatal conductance model of leaves for C_4 plants, *J. Aust. Plant Physiol.*, *19*, 519–538.
- Collins, W. J., et al. (2008), Evaluation of the HadGEM2 model, *Hadley Cent. Tech. Note 74*, Met Off., U. K.
- Cox, P. (2001), Description of the “TRIFFID” dynamic global vegetation model, *Hadley Cent. Tech. Note 24*, Met Off., U. K.
- Cox, P., C. Huntingford, and R. Harding (1998), A canopy conductance and photosynthesis model for use in a GCM land surface scheme, *J. Hydrol.*, *212/213*, 79–94.
- Cox, P., R. Betts, C. Jones, S. Spall, and I. Totterdell (2000), Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model, *Nature*, *408*, 184–187.
- Cox, P., R. A. Betts, M. Collins, C. Harris, C. Huntingford, and C. Jones (2004), Amazonian forest dieback under climate-carbon cycle projections for the 21st century, *Theoret. Appl. Climatol.*, *78*, 137–156.
- Davidson, E., and I. Janssens (2006), Temperature sensitivity of soil carbon decomposition and feedbacks to climate change, *Nature*, *440*, 165–173.
- Davidson, E. A., S. E. Trumbore, and R. Amundson (2000), Soil warming and organic carbon content, *Nature*, *408*, 789–790.
- Davin, E., N. de Noblet-Ducoudré, and P. Friedlingstein (2007), Impact of land cover change on surface climate: Relevance of the radiative forcing concept, *Geophys. Res. Lett.*, *34*, L13702, doi:10.1029/2007GL029678.
- DeLucia, E., et al. (1999), Net primary production of a forest ecosystem with experimental CO_2 enrichment, *Science*, *284*, 1177–1179.
- Dufresne, J.-L., P. Friedlingstein, M. Berthelot, L. Bopp, P. Ciais, L. Fairhead, H. Le Treut, and P. Monfray (2002), On the magnitude of positive feedback between future climate change and the carbon cycle, *Geophys. Res. Lett.*, *29*(10), 1405, doi:10.1029/2001GL013777.
- Enting, I. (1987), The interannual variation in the seasonal cycle of carbon dioxide concentration at Mauna Loa, *J. Geophys. Res.*, *92*(D5), 5497–5504.
- Essery, R., M. Best, R. Betts, P. Cox, and C. Taylor (2003), Explicit representation of subgrid heterogeneity in a GCM land surface scheme, *J. Hydrometeorol.*, *4*, 530–543.
- Felzer, B., D. Kicklighter, J. Melillo, C. Wang, Q. Zhang, and R. Prin (2004), Effects of ozone on net primary production and carbon sequestration in the conterminous United States using a biogeochemistry model, *Tellus, Ser. B*, *56*, 230–248.
- Field, C. B., J. Randerson, and C. Malmstrom (1995), Global net primary production: Combining ecology and remote sensing, *Remote Sens. Environ.*, *51*, 74–88.
- Field, R., G. van der Werf, and S. S. Shen (2009), Human amplification of drought-induced biomass burning in Indonesia since 1960, *Nat. Geosci.*, *2*, 185–188, doi:10.1038/ngeo443.
- Friedlingstein, P., I. Y. Fung, E. A. Holland, J. G. John, G. P. Brasseur, D. J. Erickson, and D. S. Schimel (1995a), On the contribution of the biospheric CO_2 fertilization to the missing sink, *Global Biogeochem. Cycles*, *9*, 541–556.
- Friedlingstein, P., I. C. Prentice, I. Y. Fung, J. G. John, and G. P. Brasseur (1995b), Carbon-biosphere-climate interactions in the last glacial maximum climate, *J. Geophys. Res.*, *100*, 7203–7221.
- Friedlingstein, P., L. Bopp, P. Ciais, L. Dufresne, J.-L. Fairhead, H. Le Treut, P. Monfray, and J. Orr (2001), Positive feedback between future climate change and the carbon cycle, *Geophys. Res. Lett.*, *28*(8), 1543–1546.
- Friedlingstein, P., L. Bopp, P. Ciais, J.-L. Dufresne, L. Fairhead, H. Le Treut, P. Monfray, and J. Orr (2003), How positive is the feedback between future climate change and the carbon cycle?, *Tellus, Ser. B*, *55*, 692–700.
- Friedlingstein, P., et al. (2006), Climat-carbon cycle feedback analysis: Results from the C^4 MIP model intercomparison, *J. Clim.*, *19*, 3337–3353.
- Friend, A., et al. (2007), FLUXNET and modelling the global carbon cycle, *Global Change Biol.*, *13*, 610–633, doi:10.1111/j.1365-2486.2006.01223.x.
- Gerten, D., et al. (2008), Modelled effects of precipitation on ecosystem carbon and water dynamics in different climatic zones, *Global Change Biol.*, *14*(10), 2365–2379, doi:10.1111/j.1365-2486.2008.01651.x.
- Giardina, C., and M. Ryan (2000), Evidence that decomposition rates of organic carbon in mineral soil do not vary with temperature, *Nature*, *404*, 858–861.
- Gifford, R. (1992), Implications of the globally increasing atmospheric CO_2 concentration and temperature for the Australian terrestrial carbon budget: Integration using a simple-model, *Aust. J. Bot.*, *40*, 527–543.
- GLOBALVIEW-CO2 (2008), *Cooperative Atmospheric Data Integration Project: Carbon Dioxide*, CD-ROM, ESRL, NOAA, Boulder, Colo., (Also available via anonymous ftp to ftp.cmdl.noaa.gov; path:Ccg/co2/GLOBALVIEW)
- Gordon, C., C. Cooper, C. A. Senior, H. Banks, J. M. Gregory, T. C. Johns, J. F. B. Mitchell, and R. A. Wood (2000), The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustment, *Clim. Dyn.*, *16*, 147–168.
- Govindasamy, B., P. B. Duffy, and K. Caldeira (2001), Land use changes and Northern Hemisphere cooling, *Geophys. Res. Lett.*, *28*(2), 291–294.
- Govindasamy, B., S. Thompson, A. Mirin, M. Wickett, K. Caldeira, C. Delire, and P. B. Duffy (2005), Increase of carbon cycle feedback with climate sensitivity: Results from a coupled climate and carbon cycle model, *Tellus, Ser. B*, *57*, 153–163.
- Gurney, K., et al. (2003), Transcom 3 CO_2 inversion intercomparison: 1. Annual mean control results and sensitivity to transport and prior flux information, *Tellus, Ser. B*, *55*(2), 555–579.
- Gurney, K., et al. (2004), Transcom 3 inversion intercomparison: Model mean results for the estimation of seasonal carbon sources and sinks, *Global Biogeochem. Cycles*, *18*, GB1010, doi:10.1029/2003GB002111.
- Heimann, M., and M. Reichstein (2008), Terrestrial ecosystem carbon dynamics and climate feedbacks, *Nature*, *451*, 289–292.
- Heimann, M., et al. (1998), Evaluation of terrestrial carbon cycle models through simulations of the seasonal cycle of atmospheric CO_2 : First results of a model intercomparison study, *Global Biogeochem. Cycles*, *12*(1), 1–24.
- Hibbard, K., G. Meehl, P. Cox, and P. Friedlingstein (2007), A strategy for climate change stabilization experiments, *Eos Trans. AGU*, *88*(20), doi:10.1029/2007EO200002.
- Houghton, R., and J. Hackler (2002), Carbon flux to the atmosphere from land-use changes, in *Trends: A Compendium of Data on Global Change*, Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl. Lab., U.S. Dept. of Energy, Oak Ridge, Tenn.
- Hourdin, F., et al. (2006), The LMDZ4 general circulation model climate performance and sensitivity to parameterized physics with emphasis on tropical convection, *Clim. Dyn.*, *19*(15), 3445–3482.
- Hungate, B., J. S. Dukes, M. R. Shaw, Y. Luo, and C. B. Field (2003), Nitrogen and climate change, *Science*, 1512–1513.
- Ito, A., et al. (2008), Can we reconcile differences in estimates of carbon fluxes from land-use change and forestry for the 1990s?, *Atmos. Chem. Phys.*, *8*(12), 3291–3310.
- Jenkinson, D. S. (1990), The turnover of organic-carbon and nitrogen in soil, *Philos. Trans. R. Soc. London*, *329*, 361–368.
- Jobbágy, E. G., and R. B. Jackson (2000), The vertical distribution of soil organic carbon and its relation to climate and vegetation, *Ecol. Appl.*, *10*(2), 423–436.
- Jones, C., and P. Cox (2001), Modelling the volcanic signal in the atmospheric CO_2 record, *Global Biogeochem. Cycles*, *15*, 453–466.
- Jones, C., M. Collins, P. Cox, and S. Spall (2001), The carbon cycle response to ENSO: A coupled climate-carbon cycle model study, *J. Clim.*, *14*, 4113–4129.
- Jones, C., P. Cox, R. Essery, and D. Roberts (2003), Strong carbon cycle feedbacks in a climate model with interactive CO_2 and sulphate aerosols, *Geophys. Res. Lett.*, *30*(9), 1479, doi:10.1029/2003GL018667.
- Jones, C., C. McConnell, K. Coleman, P. Cox, P. Falloon, D. Jenkinson, and D. Powlson (2005), Global climate change and soil carbon stocks: predictions from two contrasting models for the turnover of organic carbon in soil, *Global Change Biol.*, *11*, 154–166, doi:10.1111/j.1365-2486.2004.00885.x.
- Keeling, C. D., A. F. Carter, and W. G. Mook (1984), Seasonal, latitudinal, and secular variations in the abundance and isotopic ratios of atmospheric CO_2 . II: Results from oceanographic cruises in the tropical Pacific Ocean, *J. Geophys. Res.*, *88*, 4615–4628.
- Keeling, C. D., J. F. S. Chin, and T. P. Whorf (1996), Increased activity of northern vegetation inferred from atmospheric CO_2 measurements, *Nature*, *382*, 146–149.
- Khodri, M., Y. Leclainche, G. Ramstein, P. Braconnot, O. Marti, and E. Cortijo (2001), Simulating the amplification of orbital forcing by ocean feedbacks in the last glaciation, *Nature*, *410*, 570–574.
- Knobl, A., and D. D. Baldocchi (2008), Effects of diffuse radiation on canopy gas exchange processes in a forest ecosystem, *J. Geophys. Res.*, *113*, G02023, doi:10.1029/2007JG000663.
- Knorr, W., I. C. Prentice, J. C. House, and E. A. Holland (2005), Long-term sensitivity of soil carbon turnover to warming, *Nature*, *433*, 298–301.
- Knutson, T. R., S. Manabe, and D. Gu (1997), Simulated ENSO in a global coupled ocean-atmosphere model: Multidecadal amplitude modulation and CO_2 -sensitivity, *J. Clim.*, *10*, 138–161.

- Krinner, G., N. Viovy, N. de Noblet-Ducoudré, J. Ogée, J. Polcher, P. Friedlingstein, P. Ciais, S. Sitch, and I. C. Prentice (2005), A dynamic global vegetation model for studies of the coupled atmosphere-biosphere system, *Global Biogeochem. Cycles*, *19*, GB1015, doi:10.1029/2003GB002199.
- Kucharik, C. J., C. Barford, M. El Maayar, S. C. Wofsy, R. K. Monson, and D. D. Baldocchi (2006), Evaluation of a Dynamic Global Vegetation Model (DGVM) at the forest stand-level: Vegetation structure, phenology, and seasonal and inter-annual CO₂ and H₂O vapor exchange at three AmeriFlux study sites, *Ecol. Modell.*, *196*, 1–31.
- Langenfelds, R. L., R. J. Francey, B. C. Pak, L. P. Steele, J. Lloyd, C. M. Trudinger, and C. E. Allison (2002), Interannual growth rate variations of atmospheric CO₂ and its $\delta^{13}\text{C}$, H₂, CH₄, and CO between 1992 and 1999 linked to biomass burning, *Global Biogeochem. Cycles*, *16*(3), 1048, doi:10.1029/2001GB001466.
- Le Quéré, C. (1999), Variabilité du carbone océanique de 1979 1997: Modélisation et évaluation, Ph.D. thesis, Univ. Paris VI, Paris.
- Le Quéré, C., et al. (2003), Two decades of ocean CO₂ sink and variability, *Tellus, Ser. B*, *55*, 649–656.
- Le Quéré, C., et al. (2005), Ecosystem dynamics based on plankton functional types for global ocean biogeochemistry models, *Global Change Biol.*, *11*, 2016–2040, doi:10.1111/j.1365-2486.2005.01004.x.
- Liepert, B. G. (2002), Observed reductions in surface solar radiation in the United States and worldwide from 1961 to 1990, *Geophys. Res. Lett.*, *29*(10), 1421, doi:10.1029/2002GL014910.
- Lucht, W., I. C. Prentice, R. B. Myneni, S. Sitch, P. Friedlingstein, W. Cramer, P. Bousquet, W. Buermann, and B. Smith (2002), Climatic control of the high-latitude vegetation greening trend and pinatubo effect, *Science*, *296*, 1687–1689.
- Luo, Y., et al. (2008), Modeled interactive effects of precipitation, temperature, and [CO₂] on ecosystem carbon and water dynamics in different climatic zones, *Global Change Biol.*, *14*(9), 1986–1999, doi:10.1111/j.1365-2486.2008.01629.x.
- Magnani, F., et al. (2007), The human footprint in the carbon cycle of temperate and boreal forests, *Nature*, *447*, 848–852.
- Maier-Reimer, E. (1993), Geochemical cycles in an ocean general circulation model. preindustrial tracer distributions, *Global Biogeochem. Cycles*, *7*(3), 645–677.
- Marland, G., T. Boden, and R. J. Andres (2005), Global, regional, and national CO₂ emissions, in *Trends: A Compendium of Data on Global Change*, Carbon Dioxide Inf. Anal. Cent., Oak Ridge Natl. Lab., U.S. Dept. of Energy, Oak Ridge, Tenn.
- Marti, O., et al. (2005), *The New IPSL Climate System Model: IPSL-CM4*, Rep. 26, Inst. Pierre-Simon Laplace, Paris.
- Matthews, H. D., A. J. Weaver, and K. J. Meissner (2005), Terrestrial carbon cycle dynamics under recent and future climate, *J. Clim.*, *18*, 1609–1628.
- McGuire, A. D., et al. (2001), Carbon balance of the terrestrial biosphere in the twentieth century: Analyses of CO₂, climate and land use effects with four process-based ecosystem models, *Global Biogeochem. Cycles*, *15*(1), 183–206.
- Meehl, G. A., G. Branstator, and W. Washington (1993), Tropical Pacific interannual variability and CO₂ climate change, *J. Clim.*, *6*, 42–63.
- Meehl, G., et al. (2007), Global climate projections, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., pp. 747–845, Cambridge Univ. Press, Cambridge, U. K.
- Melnikov, N., and B. O'Neill (2006), Learning about the carbon cycle from global budget data, *Geophys. Res. Lett.*, *33*, L02705, doi:10.1029/2005GL023935.
- Mercado, L. M., C. Huntingford, J. H. C. Gash, P. M. Cox, and J. Venkata (2007), Improving the representation of radiation interception and photosynthesis for climate model applications, *Tellus, Ser. B*, *59*(3), 553–565.
- Mercado, L. M., N. Bellouin, S. Sitch, O. Boucher, C. Huntingford, and P. Cox (2009), Impact of changes in diffuse radiation on the global land carbon sink, *Nature*, *458*, 1014–1017, doi:10.1038/nature07949.
- Mitchell, J., T. C. Johns, J. Gregory, and S. Tett (1995), Climate response to increasing levels of greenhouse gases and sulphate aerosols, *Nature*, *376*, 501–504.
- Mitchell, T. D., and P. D. Jones (2005), An improved method of constructing a database of monthly climate observations and associated high-resolution grids, *Int. J. Climatol.*, *25*, 693–712.
- Murphy, J. M., D. M. H. Sexton, D. N. Barnett, G. S. Jones, M. J. Webb, M. Collins, and D. A. Stainforth (2004), Quantification of modelling uncertainties in a large ensemble of climate change simulations, *Nature*, *430*, 768–772, doi:10.1038/nature0277.
- Naegler, T., P. Ciais, J. C. Orr, O. Aumont, and C. Rodenbeck (2007), On evaluating ocean models with atmospheric potential oxygen, *Tellus, Ser. B*, *59*, 138–156, doi:10.1111/j.1600-0889.2006.00197.x.
- Nakicenovic, N., et al. (2000), *Special Report on Emissions Scenarios: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change*, edited by N. Nakicenovic and R. Swart, 599 pp., Cambridge Univ. Press, Cambridge, U. K. (Available at <http://www.grida.no/climate/ipcc/emission/index.htm>)
- Nevison, C., N. Mahowald, S. Doney, I. Lima, G. van der Werf, J. Randerson, D. Baker, P. Kasibhatla, and G. McKinley (2008), Contribution of ocean, fossil fuel, land biosphere and biomass burning carbon fluxes to seasonal and interannual variability in atmospheric CO₂, *J. Geophys. Res.*, *113*, G01010, doi:10.1029/2007JG000408.
- Niyogi, D., et al. (2004), Direct observations of the effects of aerosol loading on net ecosystem CO₂ exchanges over different landscapes, *Geophys. Res. Lett.*, *31*, L20506, doi:10.1029/2004GL020915.
- Palmer, J., and I. Totterdell (2001), Production and export in a global ocean ecosystem model, *Deep Sea Res.*, *48*, 1169–1198.
- Patra, P. K., S. Maksyutov, and T. Nakazawa (2005), Analysis of atmospheric CO₂ growth rates at Mauna Loa using inverse model derived CO₂ fluxes, *Tellus, Ser. B*, *57*, 357–365.
- Pearman, G. I., and P. Hyson (1981), The annual variation of atmospheric CO₂ concentration observed in the Northern Hemisphere, *J. Geophys. Res.*, *86*, 9839–9843.
- Peterson, J., W. Komhyr, L. Waterman, G. R. H. K. Thoning, and T. Conway (1986), Atmospheric CO₂ variations at Barrow, Alaska, *J. Atmos. Chem.*, *4*, 491–510.
- Piao, S., et al. (2008), Net carbon dioxide losses of northern ecosystems in response to autumn warming, *Nature*, *451*, 49–52.
- Raddatz, T. J., C. J. Reick, W. Knorr, J. Kattge, E. Roeckner, R. Schnur, K.-G. Schnitzler, P. Wetzel, and J. Jungclaus (2007), Will the tropical land biosphere dominate the climate-carbon cycle feedback during the twenty-first century?, *Clim. Dyn.*, *29*, 565–574.
- Randall, D., et al. (2007), Climate models and their evaluation, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., pp. 590–662, Cambridge Univ. Press, Cambridge, U. K.
- Randerson, J., M. Thompson, T. Conway, I. Fung, and C. Field (1997), The contribution of terrestrial sources and sinks to trends in the seasonal cycle of atmospheric carbon dioxide, *Global Biogeochem. Cycles*, *11*, 535–560.
- Randerson, J., C. Field, I. Fung, and P. Tans (1999), Increases in early season ecosystem uptake explain recent changes in the seasonal cycle of atmospheric CO₂ at high northern latitudes, *Geophys. Res. Lett.*, *26*, 2765–2769.
- Randerson, J. T., et al. (2009), Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models, *Global Change Biol.*, *15*(10), 2462–2484.
- Reichstein, M., et al. (2007), Reduction of ecosystem productivity and respiration during the European summer 2003 climate anomaly: A joint flux tower, remote sensing and modelling analysis, *Global Change Biol.*, *13*(3), 634–651, doi:10.1111/j.1365-2486.2006.01224.x
- Rodenbeck, C., S. Houweling, M. Gloor, and M. Heimann (2003), CO₂ flux history 1982–2001 inferred from atmospheric data using a global inversion of atmospheric transport, *Atmos. Chem. Phys.*, *3*, 1919–1964.
- Running, S. (2006), Is global warming causing more, larger wildfires?, *Science*, *313*(5789), 927–928, doi:10.1126/science.1130370.
- Sabine, C., et al. (2004), The oceanic sink for anthropogenic CO₂, *Science*, *305*(5682), doi:10.1126/science.1097403.
- Schneider, B., et al. (2008), Climate-induced interannual variability of marine primary and export production in three global coupled climate carbon cycle models, *Biogeosciences*, *5*, 597–614.
- Sitch, S., et al. (2003), Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model, *Global Change Biol.*, *9*, 357–373.
- Sitch, S., P. M. Cox, W. J. Collins, and C. Huntingford (2007), Indirect radiative forcing of climate change through ozone effects on the land-carbon sink, *Nature*, *448*, 791–794.
- Sitch, S., et al. (2008), Evaluation of the terrestrial carbon cycle, future plant geography and climate-carbon cycle feedbacks using five Dynamic Global Vegetation Models (DGVMs), *Global Change Biol.*, *25*, doi:10.1111/j.1365-2486.2008.01626.x.
- Sokolov, A., D. Kicklighter, J. Melillo, B. Felzer, C. Schlosser, and T. Cronin (2008), Consequences of considering carbon-nitrogen interactions on the feedbacks between climate and the terrestrial carbon cycle, *J. Clim.*, *21*, 3776–3796.

- Solomon, S., et al. (2007), Technical summary, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., pp. 129–234, Cambridge Univ. Press, Cambridge, U. K.
- Stanhill, G., and S. Cohen (2001), Global dimming: A review of the evidence for a widespread and significant reduction in global radiation with discussion of its probable causes and possible agricultural consequences, *Agric. For. Meteorol.*, *107*, 155–278.
- Still, C. J., et al. (2009), The influence of clouds and diffuse radiation on ecosystem, atmosphere CO₂ and CO¹⁸O exchanges, *J. Geophys. Res.*, *114*, G01018, doi:10.1029/2008JG000675.
- Takahashi, T., et al. (2009), Climatological mean and decadal change in surface ocean pCO₂, and net sea-air CO₂ flux over the global oceans, *Deep Sea Res. II*, *56*, 554–577, doi:10.1016/j.dsr2.2008.12.009.
- Thompson, M., I. Enting, G. Pearman, and I. Hyson (1986), Interannual variation of atmospheric CO₂ concentration, *J. Atmos. Chem.*, *4*, 125–155.
- Thompson, S. L., B. Govindasamy, A. Mirin, K. Caldeira, C. Delire, J. Milovich, M. Wickett, and D. Erickson (2004), Quantifying the effects of CO₂-fertilized vegetation on future global climate and carbon dynamics, *Geophys. Res. Lett.*, *31*, L23211, doi:10.1029/2004GL021239.
- Thoning, K., P. Tans, and W. Komhyr (1989), Atmospheric carbon dioxide at Mauna Loa observatory: 2. Analysis of the NOAA GMCC data, *J. Geophys. Res.*, *94*(D6), 8549–8565.
- Trenberth, K. (1997), The definition of El Nio, *Bull. Am. Meteorol. Soc.*, *78*, 2771–2777.
- Trenberth, K., et al. (2007), Observations: Surface and atmospheric climate change, in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by S. Solomon et al., pp. 235–336, Cambridge Univ. Press, Cambridge, U. K.
- Tucker, C. J., J. E. Pinzon, and M. E. Brown (2004), Global Inventory Modeling and Mapping Studies, NA94apr15b.n11-V1g, 2.0, 15 April 1994, <http://ftp.glcf.umd.edu/data/gimms/>, Global Land Cover Facil., Univ. of Md., College Park, Md.
- van der Werf, G., J. Randerson, G. Collatz, L. Giglio, P. Kasibhatla, A. J. Arellano, S. Olsen, and E. Kasischke (2004), Continental-scale partitioning of fire emissions during the 1997 to 2001 El Nio/La Nia period, *Atmos. Chem. Phys.*, *3*(5654), 73–76, doi:10.1126/science.1090753.
- van der Werf, G. R., J. T. Randerson, L. Giglio, G. J. Collatz, P. S. Kasibhatla, and A. F. Arellano (2006), Interannual variability in global biomass burning emissions from 1997 to 2004, *Atmos. Chem. Phys.*, *6*, 3423–3441.
- van der Werf, G., D. Morton, R. DeFries, J. Olivier, P. Kasibhatla, R. Jackson, G. Collatz, and J. Randerson (2009), CO₂ emissions from forest loss, *Nat. Geosci.*, *2*, 737–738, doi:10.1038/ngeo671.
- Wiltshire, A. (2006), Modelling the surface energetics of patchy arctic tundra snowcover, Ph.D. thesis, Univ. of Durham, Durham, U. K.
- Wittenberg, U., M. Heimann, G. Esser, A. D. McGuire, and W. Sauf (1998), On the influence of biomass burning on the seasonal CO₂ signal as observed at monitoring stations, *Global Biogeochem. Cycles*, *12*(3), 531–535.
- Zeng, N., H. Qian, E. Munoz, and R. Iacono (2004), How strong is carbon cycle-climate feedback under global warming?, *Geophys. Res. Lett.*, *31*, L20203, doi:10.1029/2004GL020904.
- L. Bopp, P. Cadule, P. Ciais, P. Friedlingstein, and P. Peylin, UMR CEA, IPSL, LSCE, UVSQ, CNRS, Bât. 712, CEA, L'Orme des Merisiers, F-91191 Gif-sur-Yvette, France. (patricia.cadule@lsce.ipsl.fr)
- C. D. Jones, Met Office, Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK.
- S. L. Piao, Department of Ecology, Peking University, Beijing 100871, China.
- S. Sitch, Met Office, JCHMR, Maclean Building, Crowmarsh-Gifford, Wallingford OX10 8BB, UK.