



RESEARCH ARTICLE

10.1002/2013JG002381

Key Points:

- We analyzed emission-driven simulations from 15 Earth System Models (ESMs)
- Most ESMs had a small positive bias in contemporary atmospheric CO₂ predictions
- We used a linear relationship to create a trajectory of future atmospheric CO₂

Supporting Information:

- Readme
- Supplementary Materials.

Correspondence to:

F. M. Hoffman,
forrest@climatemodeling.org

Citation:

Hoffman, F. M., J. T. Randerson, V. K. Arora, Q. Bao, P. Cadule, D. Ji, C. D. Jones, M. Kawamiya, S. Khatiwala, K. Lindsay, A. Obata, E. Shevliakova, K. D. Six, J. F. Tjiputra, E. M. Volodin, and T. Wu (2014), Causes and implications of persistent atmospheric carbon dioxide biases in Earth System Models, *J. Geophys. Res. Biogeosci.*, 119, 141–162, doi:10.1002/2013JG002381.

Received 13 MAY 2013

Accepted 17 DEC 2013

Accepted article online 19 DEC 2013

Published online 15 FEB 2014

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

Causes and implications of persistent atmospheric carbon dioxide biases in Earth System Models

F. M. Hoffman^{1,2}, J. T. Randerson¹, V. K. Arora³, Q. Bao⁴, P. Cadule⁵, D. Ji⁶, C. D. Jones⁷, M. Kawamiya⁸, S. Khatiwala⁹, K. Lindsay¹⁰, A. Obata¹¹, E. Shevliakova¹², K. D. Six¹³, J. F. Tjiputra¹⁴, E. M. Volodin¹⁵, and T. Wu¹⁶

¹Department of Earth System Science, University of California, Irvine, California, USA, ²Climate Change Science Institute and Computational Earth Sciences Group, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA, ³Canadian Center for Climate Modeling and Analysis, Meteorological Service of Canada, Victoria, British Columbia, Canada, ⁴State Key Laboratory of Numerical Modeling for Atmospheric Sciences and Geophysical Fluid Dynamics, Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China, ⁵Institut Pierre Simon Laplace, Laboratoire des Sciences du Climat et de l'Environnement, Gif sur Yvette CEDEX, France, ⁶State Key Laboratory of Earth Surface Processes and Resource Ecology, College of Global Change and Earth System Science, Beijing Normal University, Beijing, China, ⁷Hadley Center, U.K. Met Office, Exeter, UK, ⁸Research Institute for Global Change, Japan Agency for Marine-Earth Science and Technology, Kanagawa, Japan, ⁹Lamont Doherty Earth Observatory, Columbia University, Palisades, New York, USA, ¹⁰Climate and Global Dynamics Division, National Center for Atmospheric Research, Boulder, Colorado, USA, ¹¹Meteorological Research Institute, Japan Meteorological Agency, Ibaraki, Japan, ¹²Geophysical Fluid Dynamics Laboratory, Princeton University, Princeton, New Jersey, USA, ¹³Max Planck Institute for Meteorology, Hamburg, Germany, ¹⁴Uni Climate, Uni Research, Bergen, Norway, ¹⁵Institute of Numerical Mathematics, Russian Academy of Science, Moscow, Russia, ¹⁶Climate System Modeling Division, Beijing Climate Center, China Meteorological Administration, Beijing, China

Abstract The strength of feedbacks between a changing climate and future CO₂ concentrations is uncertain and difficult to predict using Earth System Models (ESMs). We analyzed emission-driven simulations—in which atmospheric CO₂ levels were computed prognostically—for historical (1850–2005) and future periods (Representative Concentration Pathway (RCP) 8.5 for 2006–2100) produced by 15 ESMs for the Fifth Phase of the Coupled Model Intercomparison Project (CMIP5). Comparison of ESM prognostic atmospheric CO₂ over the historical period with observations indicated that ESMs, on average, had a small positive bias in predictions of contemporary atmospheric CO₂. Weak ocean carbon uptake in many ESMs contributed to this bias, based on comparisons with observations of ocean and atmospheric anthropogenic carbon inventories. We found a significant linear relationship between contemporary atmospheric CO₂ biases and future CO₂ levels for the multimodel ensemble. We used this relationship to create a contemporary CO₂ tuned model (CCTM) estimate of the atmospheric CO₂ trajectory for the 21st century. The CCTM yielded CO₂ estimates of 600 ± 14 ppm at 2060 and 947 ± 35 ppm at 2100, which were 21 ppm and 32 ppm below the multimodel mean during these two time periods. Using this emergent constraint approach, the likely ranges of future atmospheric CO₂, CO₂-induced radiative forcing, and CO₂-induced temperature increases for the RCP 8.5 scenario were considerably narrowed compared to estimates from the full ESM ensemble. Our analysis provided evidence that much of the model-to-model variation in projected CO₂ during the 21st century was tied to biases that existed during the observational era and that model differences in the representation of concentration-carbon feedbacks and other slowly changing carbon cycle processes appear to be the primary driver of this variability. By improving models to more closely match the long-term time series of CO₂ from Mauna Loa, our analysis suggests that uncertainties in future climate projections can be reduced.

1. Introduction

Anthropogenic emissions of radiatively active greenhouse gases into the atmosphere, especially carbon dioxide (CO₂), are rapidly increasing the burden of these gases and altering the Earth's climate [IPCC, 2007; Raupach and J. G. Canadell, 2010]. This perturbation of the global carbon cycle is expected to induce feedbacks from the terrestrial biosphere and oceans on future CO₂ concentrations and the climate system. These climate-carbon cycle feedbacks are highly uncertain, difficult to predict, and potentially large [Denman et al., 2007]. Understanding and predicting the strength and direction of feedbacks is critically important

for estimating future atmospheric CO₂ concentrations and, therefore, accurately predicting the effects and extent of climate change.

Models of Earth's climate system are used to predict responses to human and natural forcings into the future, while hindcasts are used to judge the ability of individual models to reproduce observed patterns. Current generation Earth System Models (ESMs) attempt to capture the complex interactions and feedbacks between climate, terrestrial and ocean ecosystems, and human activities. Scenarios describing alternative prospective future socioeconomic, technological, and environmental conditions are used to generate a consistent set of chemical, biological, and land use data to drive ESMs [Moss *et al.*, 2010]. The results from such ESM simulations are valuable for diagnosing the magnitude of mitigation efforts required to stabilize CO₂ levels in the atmosphere under various scenarios, taking into account carbon cycle responses and feedbacks. Traditionally, such models were provided with a trajectory of CO₂ and other greenhouse gases consistent with scenario assumptions about population, energy resources and consumption, and agricultural policies and practices. Recently, as improvements to the representation of biogeochemical processes on land and in the ocean and better atmospheric chemistry have been added to ESMs, scenario-derived emissions of radiatively active gases, consistent with plausible natural and anthropogenic influences, are used to force ESMs. Concentration-driven simulation results are frequently analyzed to evaluate the mean carbon stocks and fluxes and to constrain biosphere processes and feedbacks in land and ocean models [Friedlingstein *et al.*, 2006; Arora *et al.*, 2013; Anav *et al.*, 2013]. They also provide the opportunity to estimate emissions scenarios consistent with a specific trajectory of atmospheric CO₂ [Jones *et al.*, 2013]. Emission-driven simulations, in contrast, provide the opportunity to assess the implications of biases resulting from uncertainties associated with ecosystem processes and feedbacks as the effects of those uncertainties propagate through the coupled ESM.

Friedlingstein *et al.* [2003, 2006] developed a framework for analysis of climate–carbon cycle feedbacks and applied it to 11 coupled climate–carbon cycle atmosphere–ocean general circulation models for the Coupled Climate–Carbon Cycle Model Intercomparison Project (C⁴MIP). Friedlingstein *et al.* [2003, 2006] introduced model sensitivities of land and ocean carbon sinks to climate (γ_L and γ_O , respectively) and to atmospheric CO₂ concentration (β_L and β_O , respectively) as metrics of climate–carbon and concentration–carbon feedbacks, respectively, to complement the overall climate sensitivity to atmospheric CO₂ parameter, α , in common use. In their study, Friedlingstein *et al.* [2006] found that the model sensitivities of the land carbon sinks to climate (γ_L) varied by almost a factor of 9 and to concentration (β_L) by almost a factor of 14. Moreover, the models varied by a factor of almost 8 in their gain (g) of the climate–carbon cycle feedback. Arora *et al.* [2013] performed a similar analysis for nine ESMs participating in the Fifth Phase of the Coupled Model Intercomparison Project (CMIP5) [Taylor *et al.*, 2012] and found that γ_L varied by almost a factor of 6, β_L varied by almost a factor of 7, and the emissions-derived gain (g_E) varied by more than a factor of 6. The emissions-derived gain (g_E) is analogous to the Friedlingstein *et al.* [2006] gain (g) for concentration-forced simulations. The multimodel mean feedback parameters, and their standard deviations, were lower in the nine CMIP5 models than in the C⁴MIP models, with γ_L being 26% weaker and β_L being 32% weaker. These differences may be partially explained by differences in the future emissions scenarios used in the two studies. Nevertheless, these results point to very large uncertainties in the response of terrestrial biosphere models to climate change and rising CO₂ concentrations and in the overall strength of the feedbacks they predict. While the framework developed by Friedlingstein *et al.* [2006] is useful for evaluating the overall strength of feedback responses within a given model and for comparing concentration and climate sensitivities between models, it provides no indication about the likelihood of any model being correct. In addition, multiple factors contribute to the apparent strength of the β_L , β_O , γ_L , and γ_O sensitivities, and the concentration and climate sensitivities interact with each other nonlinearly through biological and chemical processes [Gregory *et al.*, 2009].

In the studies described above characterizing carbon cycle feedback processes, no comparisons were made to observations. This is the next crucial step for reducing uncertainties associated with future scenarios of global climate change. Recent research has made initial steps in this direction. Cox *et al.* [2013] used the observed relationship between the growth rate of atmospheric CO₂ and tropical temperature as a constraint to reduce predicted uncertainty in the land carbon storage sensitivity to climate change (γ_L) in the tropics in C⁴MIP models. Similarly, Gillett *et al.* [2013] used the ratio of warming to cumulative emissions of CO₂ to estimate a transient response to cumulative emissions from observations for comparison with 12 CMIP5 models. Such innovative use of contemporary measurements to constrain carbon cycle responses

to climate change is important for reducing the range of uncertainty in future climate change projections [Randerson, 2013]. Moreover, comparisons with data sets derived from the synthesis of measurements collected over a wide geospatial range can provide constraints on individual processes and on carbon cycle responses that are sensitive to initial conditions. *Todd-Brown et al.* [2013] compared soil carbon stocks from 11 CMIP5 models with the Harmonized World Soil Database and the Northern Circumpolar Soil Carbon Database. Despite reasonable global-scale agreement with these observations, most ESMs failed to reproduce grid-scale soil carbon variations, suggesting that key processes may be missing in the majority of ESMs.

The goal of this paper is to identify long-term CO₂ biases in emission-driven simulation results produced by ESMs participating in CMIP5 and describe the causes and implications of those biases for future climate projections during the middle and latter half of the 21st century. In our analysis, we developed a new approach using contemporary inventory observations and structural information about feedbacks within the CMIP5 models to constrain future CO₂ predictions and to reduce uncertainties associated with the range of possible CO₂ mole fractions consistent with the Representative Concentration Pathway (RCP) 8.5 emissions scenario.

2. Methods

2.1. Model Descriptions

We analyzed historical and future emission-driven simulation results produced using ESMs for CMIP5. The historical simulations, referred to as experiment 5.2 or esmHistorical [Taylor et al., 2012], were forced with spatially distributed CO₂ emissions reconstructed from fossil fuel consumption estimates [Andres et al., 2011] for the period 1850–2005. The future simulations, referred to as experiment 5.3 or esmrcp85 [Taylor et al., 2012], were forced with projected CO₂ emissions for the period 2006–2100, following the scenario described by the Representative Concentration Pathway (RCP) 8.5 [Moss et al., 2010]. Model output was obtained primarily from the Earth System Grid Federation, an international network of distributed climate data servers [Williams et al., 2011].

Simulation results were produced by fully coupled ESMs with interactive terrestrial and marine biogeochemistry models, which feature climate–carbon cycle feedback mechanisms. Since the simulations were forced with CO₂ emissions, these models prognostically computed global atmospheric CO₂ mole fractions, which represent an integration of physical, chemical, and biological processes on Earth and their interactions and feedbacks with the climate system. The ESMs employed different aerosol emissions, land use change processes, and process parameterizations, leading to a range of different aerosol and greenhouse gas concentrations, radiative forcings, and climate interactions. The ability of models to accurately reproduce the observed atmospheric CO₂ mole fraction trajectory over the historical period provides a broad indication of model fidelity, a necessary but not sufficient condition for credible ESM performance. Each of the models that generated output used in this study is listed in Table 1.

2.2. Model Output

Monthly output of prognostic atmospheric CO₂ and surface ocean CO₂ flux from emission-driven ESM simulations were analyzed to evaluate the evolution of the carbon cycle over the twentieth and 21st centuries. Atmospheric CO₂ was obtained either as the total atmospheric mass of CO₂ and converted to mole fraction or as the atmospheric CO₂ mole fraction at every atmosphere model layer. In the latter case, the global mole fraction was calculated as the area-weighted mean of CO₂ in the lowest atmosphere level. Surface ocean CO₂ flux was integrated spatially to determine global carbon uptake and further integrated over time to estimate the global change in ocean carbon inventory. While the net terrestrial CO₂ flux was available for some models in the form of net biospheric productivity, here annual land carbon uptake was calculated as the difference between the prescribed annual anthropogenic emissions and the sum of the annual change in atmosphere and ocean carbon inventories. Therefore, the change in land carbon storage for a given year was estimated as

$$\Delta C_L = \sum_i F_i - \Delta C_A - \Delta C_O, \quad (1)$$

where F_i was the total anthropogenic fossil carbon emissions from all sources i (fossil fuel burning and cement production) for that year, ΔC_A was the change in atmospheric carbon storage for that year, and ΔC_O was the change in ocean carbon storage for that year. A single trajectory of annual anthropogenic carbon emissions, derived from the experimental forcing, was used in calculations

Table 1. Models That Generated Output Used in This Study^a

Model	Modeling Center (or Group)	Component Models and Resolutions				Sea Ice
		Atmosphere	Land	Ocean		
BCC-CSM1.1 [Wu et al., 2013]	Beijing Climate Center, China Meteorological Administration, China	AGCM2.1 (2.875°×2.875°, L26)	BCC_AVIM1.0 (2.875°×2.875°)	MOM4_L40 (1°×(1- $\frac{1}{3}$)°, L40)	SIS (1°×(1- $\frac{1}{3}$)°)	
BCC-CSM1.1(m) [Wu et al., 2013]	Beijing Climate Center, China Meteorological Administration, China	AGCM2.2 (1.125°×1.125°, L26)	BCC_AVIM1.0 (1.125°×1.125°)	MOM4_L40 (1°×(1- $\frac{1}{3}$)°, L40)	SIS (1°×(1- $\frac{1}{3}$)°)	
BNU-ESM ^{ah} [Dai et al., 2003, 2004; College of Global Change and Earth System Science, 2012]	Beijing Normal University, China	CAM3.5 (2.875°×2.875°, L26)	CoLM3 and BNUDGVM (C/N) (2.875°×2.875°, L10)	MOM4p1 and IBGC (1°×(1- $\frac{1}{3}$)°, L50)	CICE4.1 (1°×(1- $\frac{1}{3}$)°)	
CanESM2 ^b [Arora et al., 2011]	Canadian Center for Climate Modeling and Analysis, Canada	CanAM4 (2.81°×2.81°, L35)	CLASS2.7 and CTEM (2.81°×2.81°)	CanOM4 and CMOG (1.41°×0.94°, L40)	CanOM4 (2.81°×2.81°)	
CESM1-BGC ^h [Hurrell et al., 2013; Keppel-Aleks et al., 2013; Long et al., 2013]	Community Earth System Model Contributors, NSF-DOE- NCAR, USA	CAM4 (0.9°×1.25°, L30)	CLM4 (0.9°×1.25°)	POP2 and NPZD (1°×(1- $\frac{1}{3}$)°, L60)	CICE4 (1°×(1- $\frac{1}{3}$)°)	
FGOALS-s2.0 ^c [Bao et al., 2013; Liu et al., 2012; Lin et al., 2013]	LASG, Institute of Atmospheric Physics, CAS, China	SAMIL2.4.7 (1.67°×2.81°, L26)	CLM3 and VEGAS2.0 (1.67°×2.81°)	LICOM2.0 (1°×(1- $\frac{1}{2}$)°, L30)	CSIM5 (1°×(1- $\frac{1}{2}$)°)	

^a Atmospheric CO₂ required unit correction.

^b Ocean carbon flux required unit correction.

^c FGOALS-s2 model provided no ocean carbon fluxes.

^d GFDL-ESM2g and GFDL-ESM2m output available beginning January 1861.

^e HadGEM2-ES output available for December 1859 through November 2099; annual atmospheric CO₂ obtained directly from Hadley Center.

^f IPSL-CM5A-LR monthly atmospheric CO₂ obtained directly from IPSL.

^g MPI-ESM-LR provided three esmHistorical realizations and one esmrcp85 realization.

^h Atmospheric CO₂ mole fraction was computed from three-dimensional output.

Table 1. Continued.

Model	Modeling Center (or Group)	Component Models and Resolutions				Sea Ice
		Atmosphere	Land	Ocean		
GFDL-ESM2g, GFDL-ESM2m ^d [Dunne et al., 2012; 2013]	NOAA Geophysical Fluid Dynamics Laboratory, USA	AM2 (2° × 2.5°, L24)	LM3 (2° × 2.5°)	MOM4 (1° × (1- $\frac{1}{3}$)°, L50)	SIS (1° × (1- $\frac{1}{3}$)°)	
HadGEM2-ES ^e [Collins et al., 2011; Jones et al., 2011]	Met Office Hadley Center, UK	HadGAM2 and UKCA (1.25° × 1.875°, L38)	MOSES2 and TRIFFID (1.25° × 1.875°)	HadGOM2 and diat-HadOCC (1° × (1- $\frac{1}{3}$)°, L40)	HadGOM2 (1° × (1- $\frac{1}{3}$)°)	
INM-CM4 ^{ab} [Volodin et al., 2010]	Institute for Numerical Mathematics, Russia	(2° × 1.5°, L21)	(2° × 1.5°)	(1° × 0.5°, L40)	(1° × 0.5°)	
IPSL-CM5A-LR ^f [Dufresne et al., 2013]	Institut Pierre Simon Laplace, France	LM24 (3.75° × 1.9°, L39)	ORCHIDEE (3.75° × 1.9°)	ORCA2 and PISCES (2° × (2- $\frac{1}{2}$)°, L31)	LIM2 (2° × (2- $\frac{1}{2}$)°)	
MIROC-ESM ^h [Watanabe et al., 2011; Oschlies, 2001]	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (University of Tokyo), and National Institute for Environmental Studies, Japan	MIROC-AGCM and SPRINTARS (2.875° × 2.875°, L80)	MATSIRO and SEIB-DGVM (2.875° × 2.875°, L6)	COCO3.4 and NPZD (1.5° × 1°, L44)	COCO3.4 (1.5° × 1°)	
MPI-ESM-LR ^g [Maier-Reimer et al., 2005; Raddatz et al., 2007; Brovkin et al., 2009]	Max Planck Institute for Meteorology, Germany	ECHAM6 (2.81° × 2.81°, L47)	JSBACH (2.81° × 2.81°)	MPIOM and HAMOCC (1.5° × 1.5°, L40)	MPIOM (1.5° × 1.5°)	
MRI-ESM1 [Yukimoto et al., 2011; Nakano et al., 2011; Yukimoto et al., 2012; Obata and Shibata, 2012]	Meteorological Research Institute, Japan	GSMUV (0.75° × 0.75°, L48)	HAL and MRI-LCCM2 (0.75° × 0.75°)	MRI.COM3 (1° × 0.5°, L51)	MRI.COM3 (1° × 0.5°)	
NorESM1-ME [Berntsen et al., 2013; Iversen et al., 2013; Tjiputra et al., 2013]	Norwegian Climate Center, Norway	CAM4-Oslo (1.9° × 2.5°, L26)	CLM4 (1.9° × 2.5°)	MICOM and HAMOCC (1° × (1- $\frac{1}{3}$)°, L53)	CICE4 (1° × (1- $\frac{1}{3}$)°)	

for all model results. Carbon fluxes due to land use change were included implicitly in ΔC_L and were not included explicitly in total fossil carbon emissions $\sum_i F_i$. We assumed F_i in each model followed the historical and RCP 8.5 time series on the Potsdam Institute for Climate Impact Research site (<http://www.pik-potsdam.de/~mmalte/rcps/index.htm>). We also assumed the individual ESMs were at steady state at the beginning of the historical simulation (i.e., that drift in the control simulation was minimal).

2.3. Comparing CMIP5 ESMs With Long-Term Carbon Cycle Observations

Atmospheric CO₂ mole fraction observations used for comparison with model projections of atmospheric CO₂ over the historical period were the same as those used to force the corresponding concentration-driven simulations, which were not analyzed here. Compiled by Tom Wigley and Malte Meinshausen, these “end-of-year CO₂ concentrations” consist of a combination of 75 year smoothed Law Dome ice core data [Etheridge *et al.*, 1996] up to 1832, 20 year smoothed Law Dome ice core data for 1823–1958, the Keeling Mauna Loa record, with 0.59 ppm subtracted (which is the Mauna Loa mean minus the NOAA global mean over 1982–1986) for 1959–1981, and the NOAA global mean value for 1982–2008 [Conway *et al.*, 1994]. Development of these and related forcing data for preindustrial control, twentieth century, and RCP simulations are described by Meinshausen *et al.* [2011].

Sabine *et al.* [2004] used inorganic carbon measurements from the World Ocean Circulation Experiment and the Joint Global Ocean Flux Study, both conducted in the 1990s, and the tracer-based ΔC^* separation technique to estimate the global oceanic anthropogenic CO₂ sink for the period 1800–1994. Their ocean inventory estimate of 118 ± 19 Pg C accounts for approximately 48% of the total emissions from fossil fuel burning and cement production. They subtracted this ocean inventory estimate and the change in atmospheric inventory over the same period of 165 ± 4 Pg C from the estimate of cumulative emissions of 244 ± 20 Pg C to obtain a cumulative terrestrial biosphere source of 39 ± 28 Pg C.

More recently, Khatiwala *et al.* [2009] applied a Green’s function model for ocean tracer transport, estimated from tracer and salinity data using a maximum entropy deconvolution technique, to simulate the time evolution of the ocean inventory and uptake rate of anthropogenic CO₂ for the period 1765–2008. They estimated the ocean inventory and uptake rate in 2008 to be 140 ± 25 Pg C and 2.3 ± 0.6 Pg C yr⁻¹, respectively. When they adjusted the estimate to include the Arctic Ocean and marginal seas not represented in the Global Ocean Data Analysis Project (GLODAP) database, the global inventory increased by approximately 11 Pg C. Using annual estimates of anthropogenic emissions and the atmospheric inventory, including uncertainties, they produced a trajectory for the terrestrial carbon budget, indicating that the terrestrial biosphere was a source of anthropogenic CO₂ until the 1940s, after which it became a sink. Tans [2009] performed a similar mass balance calculation using an empirical pulse response function constrained by the integrated ocean uptake in 1994 [Sabine *et al.*, 2004], the 1993–2002 uptake rate centered on late 1997 from atmospheric oxygen measurements [Manning and Keeling, 2006], and the 1995–2000 uptake rate estimate from an ocean inverse model [Gruber *et al.*, 2009]. Deriving net terrestrial emissions as a residual, instead of including land use emissions explicitly due to their large uncertainty, Tans [2009] also found that net terrestrial emissions were positive before 1940 and were negative thereafter, making their cumulative contribution in 2008 small.

Khatiwala *et al.* [2013] produced a newly updated global ocean anthropogenic carbon sink trajectory through 2010 using the Green’s function model. A cumulative sum of this ocean uptake provided an ocean anthropogenic carbon inventory estimate for 2010 of 150 ± 26 Pg C. Adding a partial estimate for accumulation in marginal seas and coastal areas from Lee *et al.* [2011] of 8.6 ± 0.6 Pg C yielded a more spatially comprehensive estimate of 160 ± 26 Pg C. Since the Lee *et al.* [2011] estimate was a lower bound, the upper bound was constrained using multiple Community Climate System Model-based simulations, resulting in a range for the inventory outside the GLODAP region of 9–14 Pg C. However, Khatiwala *et al.* [2013] ultimately computed a “best estimate” inventory for the GLODAP region in 2010 of 143 Pg C by averaging results from three different inversion methods, including the Green’s function model. Using the above range for marginal seas and coastal areas, they provide a 2010 global best estimate inventory of 152–157 Pg C. Selecting the midpoint value yields a final estimate of 155 Pg C with an uncertainty of $\pm 20\%$. Here we scaled the Green’s function time series to obtain the 155 Pg C best estimate for 2010 and to account for marginal seas and coastal areas.

The Sabine *et al.* [2004] and Khatiwala *et al.* [2013] data-based estimates with uncertainties provide valuable global constraints on model carbon cycle processes and feedbacks. However, these inventory estimates must be further adjusted to the 1850 equilibrium starting date of model simulations. The Sabine *et al.* [2004] ocean inventory estimate for 1994 was adjusted by subtracting the difference between the 1850 and 1800 ocean inventory estimates from the scaled Khatiwala *et al.* [2013] 1765–2010 time series, yielding 109 ± 19 Pg C. Similarly, the scaled Khatiwala *et al.* [2013] ocean inventory best estimate for 2010 was adjusted by subtracting the 1850 value, yielding 141 ± 38 Pg C. Using the adjusted trajectory of ocean uptake and applying equation (1) on the time series through 2010, we calculated total carbon accumulation in the ocean and on land from 1850 to 2010 with their uncertainties based on Khatiwala *et al.* [2013] uncertainty estimates (Figure S1). Land and ocean carbon sinks computed using this approach were consistent with combined estimates reported by Ballantyne *et al.* [2012]; however, here the net land flux included land use emissions.

2.4. A Framework for Constraining Future Trends

One approach for reducing uncertainties using contemporary observations is to identify relationships between contemporary variability and future trends within the models and constrain the contemporary variability using observations. This strategy was employed by Hall and X. Qu [2006], who evaluated the strength of the springtime snow albedo feedback ($\Delta\alpha_s/\Delta T_s$) from 17 models used for the IPCC Fourth Assessment Report and compared them with the observed springtime snow albedo feedback from the International Satellite Cloud Climatology Project and ERA-40 reanalysis data. They found a linear relationship between model predictions of seasonal and 21st century snow albedo feedbacks. Hall and X. Qu [2006] assumed that this relationship, which represents consistency in the structure of these models, accurately reflects functional behavior in nature and used observational estimates of the contemporary seasonal cycle snow albedo feedback to constrain the longer-term snow albedo feedback that occurs in the models during the 21st century. More recently, this approach was applied by Cox *et al.* [2013] to the carbon cycle. In this latter study, the authors were able to show that the long-term climate sensitivity of tropical carbon fluxes was related to this same sensitivity on interannual timescales. By using contemporary observations, they were able to narrow the likely range of future model scenarios, showing that the likelihood of forest dieback was probably overestimated in earlier work.

As described below, we found a similar linear relationship over decadal timescales between contemporary and future atmospheric CO₂ mole fractions in CMIP5 emission-driven simulations. Specifically, models that had higher positive biases in atmospheric CO₂ mole fraction by the end of the observational era in 2010 tended to predict higher atmospheric CO₂ levels during the 21st century for the RCP 8.5 scenario than models that more closely matched the observations. We used this relationship, and implicitly the collection of CMIP5 models, to construct a hypothetical model that was tuned to contemporary observations, hereafter referred to as the contemporary CO₂ tuned model, or CCTM.

The CCTM estimate was obtained using the following approach. First, we computed a linear regression between atmospheric CO₂ at each future year (y axis) and atmospheric CO₂ during 2010 (x axis), defined as the 5 year mean for 2006–2010. In particular, we repeatedly applied the regression formula

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad (2)$$

where x_i was the 2010 CO₂ mole fraction, y_i was the future CO₂ mole fraction, and ϵ_i was an error term for every model $i = 1, \dots, n$. β_0 and β_1 were two parameters, representing the y -intercept and the slope of the resulting line, respectively. Here, $n = 17$, representing the 17 separate simulations from 15 models from the CMIP5 collection. For every future interval, the error term ϵ_i was minimized using ordinary least squares to yield a linear regression model,

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i, \quad (3)$$

where \hat{y}_i was the predicted future CO₂ mole fraction from the linear regression model that minimizes the residual, $e_i = y_i - \hat{y}_i$. The least squares estimates for the parameters were calculated using a standard algebraic approach. As shown in the results section below, these regressions were statistically significant through 2100, although uncertainties increased through time. Second, we estimated the intersection of this regression with a vertical line representing NOAA Global Monitoring Division (GMD) observations in 2010

(384.6 ± 0.5 ppm; a 5 year mean centered on 2008) Conway *et al.*, [1994]. This intercept at each time interval and the 95% confidence limits on the intercept comprised our CCTM estimate.

The CCTM estimate allowed us to inquire what might be the impact of tuning ESMs to capture the observed recent trajectory of global atmospheric CO₂. This approach takes advantage of the collection of CMIP5 models—including the wide range of sensitivities of gross land and ocean carbon fluxes to elevated CO₂ and climate changes, residence time distributions of carbon in ocean and land reservoirs, and feedbacks—to create an estimate with a zero bias at the end of the observed record. As such, it can be thought of as a “black box” approach to representing the carbon cycle. It was useful in developing approaches for analyzing ESM uncertainties because of the long-term bias persistence observed for this set of models.

We also developed a similar multimodel constraint on the evolution of ocean and land cumulative flux (inventory) time series to better understand why atmospheric CO₂ biases were so persistent. As described in the results, observational uncertainties were considerably higher for the ocean and land inventories, and, as a consequence, it was not possible to reduce uncertainties in future estimates by the same amount as for atmospheric carbon dioxide.

2.5. Calculating Climate Implications of CO₂ Biases

Individual models directly calculate radiative forcing and surface temperature responses to anthropogenic CO₂ and therefore have different climate sensitivities (α). For our analysis, we chose to use a standard method for approximating radiative forcing and subsequent temperature changes to equitably assess the climate implications of CO₂ biases across all models. We adopted the method described by Boucher and Reddy [2008], who employed an impulse response function (IRF) to describe the evolution of atmospheric CO₂ and global surface temperature. First, we followed Ramaswamy *et al.* [2001] to approximate radiative forcing due to anthropogenic CO₂ at time t as

$$\text{RF}(t) = m \ln \left(\frac{[\text{CO}_2](t)}{[\text{CO}_2](t_0)} \right), \quad (4)$$

where m was set equal to 5.35 W m^{-2} and $[\text{CO}_2](t_0)$ was defined as 284 ppm in the year 1850. Second, we calculated CO₂-corrected predictions of future surface temperature following the method of Boucher and Reddy, [2008, Appendix A]. This method approximates the delayed response of surface temperature to radiative forcing as a sum of two exponentials with adjustment times of 8.4 and 410 years. Coefficients c_1 and d_1 in the exponentials [Boucher and Reddy, 2008 Appendix A] were multiplied by 0.895 to obtain a transient climate response of 1.9 K per doubling of CO₂ as reported by Gillett *et al.* [2013] for the mean of the CMIP5 ESMs. For details of the surface temperature correction, please see supporting information. While using the IRF from another ESM might alter the mean temperature change per unit of radiative forcing presented here, it would not change the order among models. We note that this calculated temperature change accounts only for CO₂-driven climate change and does not include observed cooling due to aerosols or contributions from other greenhouse gases like methane (CH₄), nitrous oxide (N₂O), and chlorofluorocarbons (CFCs).

2.6. Quantifying Uncertainty

Sabine *et al.* [2004] provided uncertainty estimates for their estimate of the ocean anthropogenic carbon inventory. Khatiwala *et al.* [2013] used the GLODAP/WOA05 databases to generate global estimates of historical anthropogenic CO₂ ocean uptake, and they propagated uncertainties from these databases through their Green's function model to provide uncertainties for these uptake estimates. We used these uncertainties in quadrature to provide an uncertainty range for the Khatiwala *et al.* [2013] inventory and propagated them through equation (1) to provide estimates of uncertainty for land carbon accumulation. For the linear regression models used here to construct the CCTM estimate, 95% confidence intervals were calculated and propagated into estimates of atmospheric CO₂, radiative forcing, and temperature change. For purposes of uncertainty comparison, the 95% confidence interval ranges for the CCTM were compared with the 95th percentile of the range for the multimodel distribution, assuming a normal distribution.

3. Results

3.1. Contemporary Biases

Comparison of ESM prognostic atmospheric CO₂ mole fraction over the historical period with observations indicated that ESMs, on average, had a high bias in their predictions of contemporary atmospheric CO₂ (Figure 1a). For the multimodel mean, this high bias was persistent from 1946 throughout the twentieth

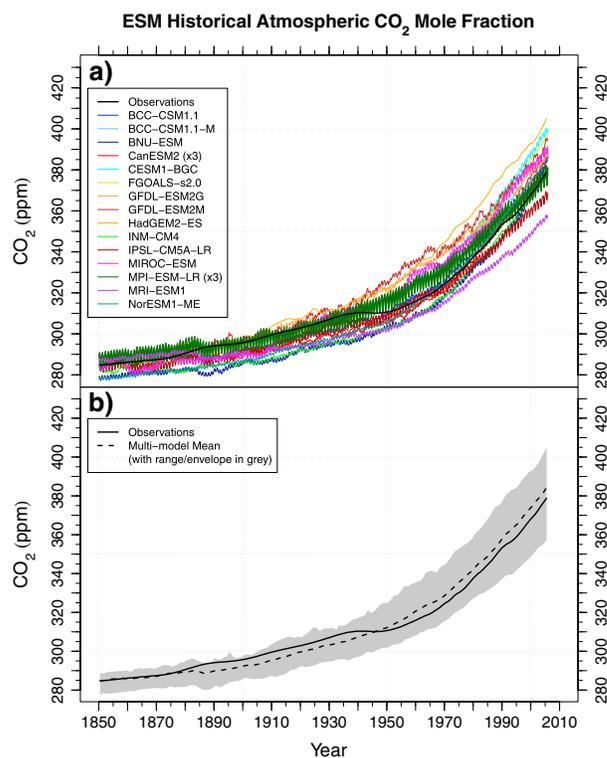


Figure 1. (a) Most ESMTs exhibit a high bias in atmospheric carbon dioxide (CO_2) mole fraction. The predicted atmospheric CO_2 mole fraction for the 19 historical simulations shown here ranges from 357 to 405 ppm at the end of the CMIP5 historical period (1850–2005). (b) The multimodel mean is biased high from 1946 throughout the remainder of the twentieth century, ending 5.6 ppm above observations in 2005.

Terrestrial biosphere models within ESMTs also had a wide range of responses, with both positive and negative net carbon accumulation throughout the twentieth century (Figures 3 and S2). Terrestrial and ocean carbon accumulation compensated for one another ($R = -0.91$, Figure S3), reducing the bias in predicted atmospheric CO_2 . This compensation effect was exemplified by the INM-CM4 model, which had the correct atmospheric CO_2 in 2005, but had strong ocean uptake that was balanced by weak land carbon uptake. During the second half of the twentieth century, the land carbon sink was persistent with high rates during the 1990s and 2000s (Table 2). Thought to be due to changes in human land use (i.e., reduced deforestation, new afforestation, and secondary regrowth of previously cleared land), wildfire suppression [Girod *et al.*, 2007; Hurtt *et al.*, 2002], and enhanced forest growth due to rising atmospheric CO_2 levels and higher rates of nitrogen deposition [Pan *et al.*, 2011; Phillips *et al.*, 2009], this growing land sink reinforced rising ocean uptake rates and resulted in a doubling of global carbon uptake between 1960 and 2010 [Ballantyne *et al.*,

century (Figure 1b). By the end of the historical model simulation period (2005), the multimodel mean was 5.6 ppm above observations and the models ranged from 21.7 ppm below to 26.2 ppm above the observed CO_2 mole fraction of 378.8 ppm. Of the 19 historical simulations from 15 ESMTs included in this analysis, only two predicted a CO_2 mole fraction well below observations in 2005. By 2010, near the end of the observational record, the multimodel mean was 7.9 ppm higher than the global mean CO_2 mole fraction reported by NOAA GMD [Conway *et al.*, 1994]. This bias was probably a conservative estimate of the true multimodel mean bias because fossil fuel emissions from the RCP 8.5 scenario during 2006–2010 (8.6 Pg C yr^{-1}) were slightly lower than the observed emissions (8.7 Pg C yr^{-1}) [Peters *et al.*, 2013; Le Quéré *et al.*, 2013].

3.2. Causes of the Contemporary Bias

Most ESMTs exhibited a small or moderate low bias in ocean carbon accumulation from 1870 to 1930 when compared with adjusted estimates from Khatiwala *et al.* [2013], but most ESMTs were contained within the envelope of observational uncertainty after 1930 (Figure 2a). Ocean carbon accumulation ranged from 88 to 261 Pg C, with a multi-

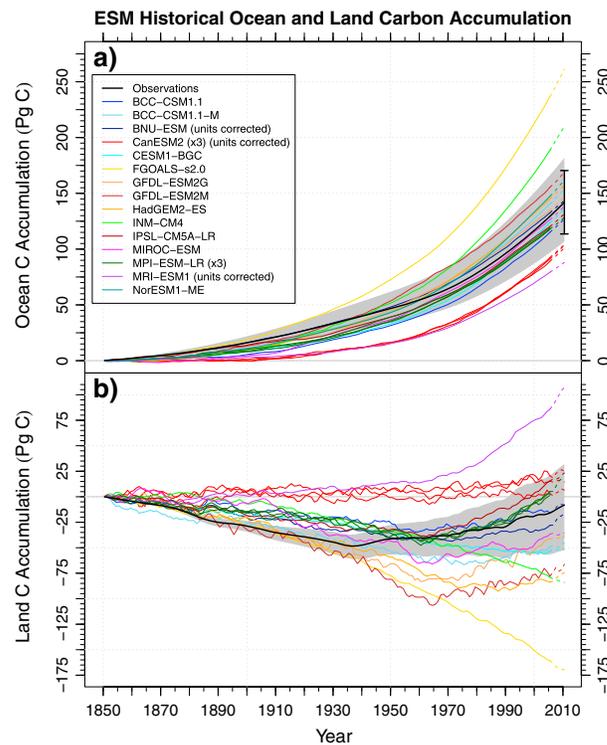


Figure 2. (a) Ocean and (b) land anthropogenic carbon inventories from CMIP5 models compared to estimates from *Khatiwalal et al.* [2013]. Most ESMs exhibit a low bias in ocean anthropogenic carbon accumulation from 1870 to 1930 as compared with adjusted estimates from *Khatiwalal et al.* [2013]. While some models enter the envelope of observational uncertainty later in the twentieth century, this was often a consequence of the increasing high bias in atmospheric CO₂ mole fractions. ESMs had a wide range of land carbon accumulation responses to increasing atmospheric CO₂ and land use change, ranging from a cumulative source of 170 Pg C to a cumulative sink of 107 Pg C in 2010. In these figures, solid colored lines represent historical simulation results and the extending dashed line segments represent the first 5 years of the RCP 8.5 simulations. The shaded polygon represents the uncertainties surrounding the adjusted observational estimates of ocean and land carbon accumulation, and the error bars correspond to the ±20% uncertainty in the *Khatiwalal et al.* [2013] best estimate of ocean carbon accumulation for 2010.

aspects of the climate system, unless compensated for by biases in concentration-carbon feedbacks (β_l and β_o) or climate sensitivities (α). Atmospheric CO₂ mole fraction projections out to 2100 under the RCP 8.5 scenario for all ESMs are shown in Figure S4. Corresponding anthropogenic carbon inventories for the ocean and land out to 2100 are shown in Figure S5.

3.4. Persistence of Biases Into the Future

To explore the persistence of atmospheric CO₂ biases beyond the present, we examined the relationship between 5 year mean contemporary and future atmospheric CO₂ mole fractions from ESMs. Figure 4a reveals a strong linear relationship between the predicted sizes of contemporary and future atmospheric CO₂ biases in 2060 with a coefficient of determination $R^2 = 0.70$. This correlation declined to $R^2 = 0.54$ in 2100 (Figure 4b) probably as a consequence of varying climate-carbon cycle feedbacks taking effect in different models. Because model biases in atmospheric CO₂ mole fraction are persistent, biases at year 1850 affect biases at year 2010. To investigate the impact of different model baselines, we also examined the relationship between the 5 year mean contemporary and future anthropogenic atmospheric carbon inventory in 2060 (Figure S6a) and 2100 (Figure S6b), taking into account uncertainties from measurements of nineteenth century CO₂ and fossil emissions. This alternative metric slightly changed the ordering of models and strengthened the coefficient of determination, further confirming the robustness of the bias persistence relationship. To explore the value of a tuned model with no CO₂ bias at the end of the historical period, we compared the CCTM estimate described in section 2 with the set of CMIP5 model predictions and the

2012]. Although the multimodel mean distribution of land sinks closely matched the observations, individual model estimates varied widely. BCC-CSM1.1-M, CESM1-BGC, FGOALS-s2.0, GFDL-ESM2M, HadGEM2-ES, INM-CM4, and NorESM1-ME tended to underestimate land sinks, whereas CanESM2 and MRI-ESM1 tended to overestimate them (Figure 2b).

3.3. Implications of Contemporary Atmospheric CO₂ Biases in CMIP5 Models

High atmospheric CO₂ biases produced radiative forcing during the latter half of the twentieth century that was too large in the affected ESMs (Table 3). For the year 2010, the multimodel mean atmospheric CO₂ mole fraction was 7.9 ppm above observations, corresponding to a radiative forcing that was 0.10 W m⁻² higher than that obtained from the observed atmospheric CO₂ mole fraction. The integrated effect of the radiative forcing bias from the multimodel mean during the nineteenth and twentieth centuries led to CO₂-induced temperature change that was 0.06°C higher by 2010 than an estimate derived from the observed CO₂ trajectory. Across all ESMs, the temperature change bias for 2010 ranged from -0.20°C to 0.24°C. Because land and ocean carbon uptake rates are likely to be reduced with climate warming (negative γ_l and γ_o), these temperature biases have the potential to further reinforce atmospheric CO₂ biases in the 21st century, leading to persistent and divergent biases into the future for many

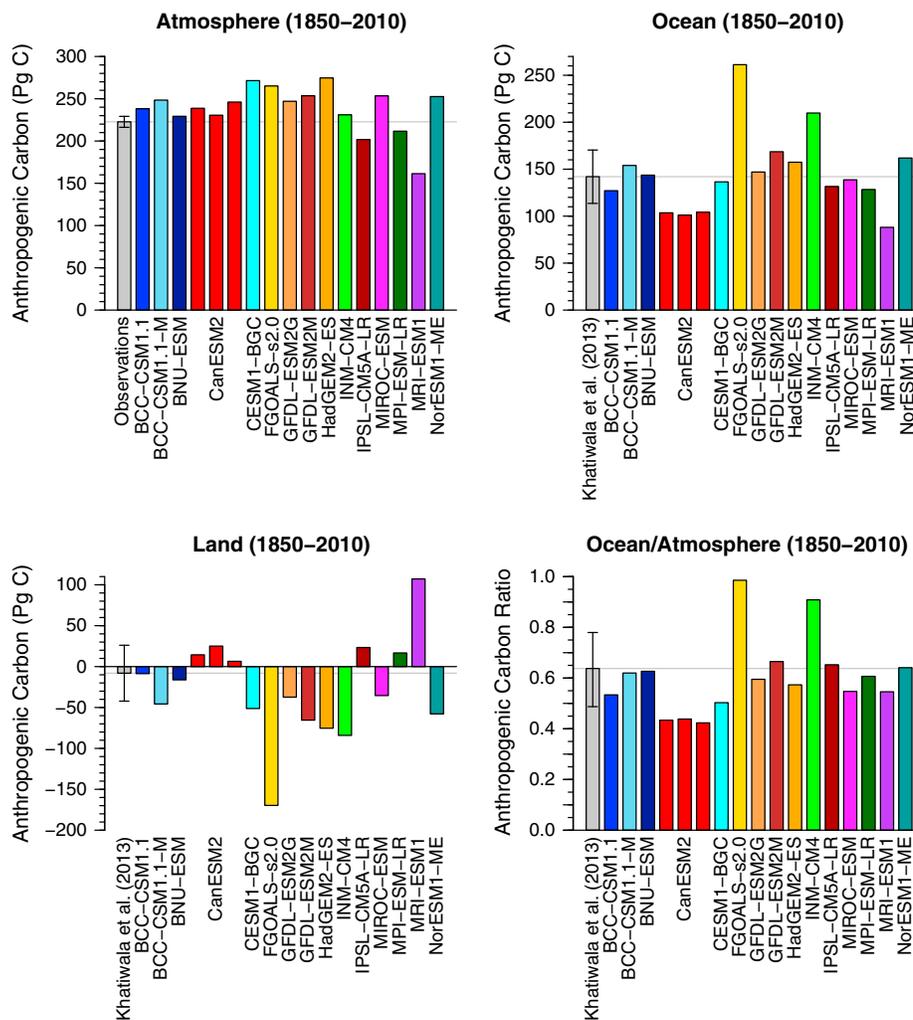


Figure 3. Reconstructed atmospheric CO₂ levels and observationally based estimates of ocean carbon uptake from Khaitiwala et al. [2013] provide constraints on carbon inventories in the ocean, and on land when combined with fossil fuel and atmospheric CO₂ observations. While ocean carbon accumulation appears adequate in some model results, ocean carbon accumulation in most ESMs shows a low bias once normalized by atmospheric accumulation (lower right panel).

RCP 8.5 CO₂ mole fraction trajectory. Figure 5 shows the coefficients of determination (R^2) of the CCTM atmospheric CO₂ mole fraction trajectory, as well as for the trajectories for ocean and land carbon accumulation when the same method is applied for those reservoirs. All of the coefficients of determination peak at one for the contemporary tuning year (2008, the center of the 2006–2010 averaging period), as expected, and decrease on either side, into the past and future. Statistical significance ($p < 0.05$) was maintained with $N = 17$ model results for R^2 values above 0.23 (i.e., after about 1910 and through 2100 for atmospheric CO₂). The resulting atmospheric CO₂ trajectory for 1850–2100 is shown as the red line in Figure 6.

The CCTM estimate suggests that for a tuned model, future atmospheric CO₂ in 2060 under the RCP 8.5 scenario would be 600 ± 14 ppm (including the 95% confidence interval of the estimate). In contrast, the multimodel mean atmospheric CO₂ mole fraction in 2060 was 621 ± 80 ppm, which was above and outside the confidence interval for the CCTM estimate (Figure 7a). Individual model predictions spanned a range from 516 to 697 ppm in 2060. The spread of the CCTM was considerably smaller than that of the multimodel 95th percentile distribution spread. In 2100, the CCTM estimate yielded an atmospheric CO₂ mole fraction of 948 ± 35 ppm, while the multimodel mean prediction was 980 ± 161 ppm (Figure 7b). The CanESM2, CESM1-BGC, MIROC-ESM, and NorESM1-ME models predicted atmospheric CO₂ mole fractions greater than 1000 ppm by 2100. In terms of anthropogenic atmospheric carbon accumulation, the CCTM estimate in 2060 under the RCP 8.5 scenario was 672 ± 28 Pg C (including the 95% confidence interval of the esti-

Table 2. Comparison of Observationally Based Estimates of Decadal Atmosphere, Ocean, and Land Uptake Rates of Anthropogenic CO₂ With CMIP5 Model Predictions^a

Model	1960–1969				1970–1979				1980–1989				1990–1999				2000–2010			
	FFE	Atm	Ocn	Land																
Observations	3.0	1.8	1.0	0.2	4.6	2.7	1.4	0.6	5.5	3.4	1.8	0.3	6.2	3.2	2.0	1.0	7.7	4.3	2.4	1.1
Multimodel Mean		1.7	1.3	0.0		2.8	1.5	0.3		3.3	1.9	0.2		3.5	2.2	0.6		4.4	2.5	0.8
BCC-CSM1.1		1.2	1.0	0.8		3.1	1.4	0.0		3.4	1.8	0.3		3.5	2.0	0.8		4.7	2.3	0.7
BCC-CSM1.1-M		1.8	1.3	-0.2		3.0	1.7	-0.1		3.1	2.0	0.4		3.6	2.2	0.4		4.3	2.6	0.8
BNU-ESM		1.8	1.1	-0.0		2.8	1.5	0.3		3.2	1.7	0.6		4.3	2.0	-0.1		4.0	2.1	1.7
CanESM2 r1		1.5	1.0	0.4		3.2	1.3	0.2		4.0	1.6	-0.1		3.6	1.7	0.9		5.7	2.3	-0.3
CanESM2 r2		2.1	0.9	-0.1		2.7	1.2	0.7		3.9	1.7	-0.2		3.5	1.7	1.0		5.3	2.2	0.3
CanESM2 r3		1.7	1.0	0.3		3.4	1.1	0.1		3.1	1.6	0.8		4.9	2.0	-0.7		4.9	2.2	0.6
CESM1-BGC		2.4	1.2	-0.6		3.2	1.5	-0.1		4.0	1.9	-0.4		4.2	2.2	-0.2		4.9	2.4	0.3
FGOALS-s2.0		2.1	2.1	-1.3		3.5	2.8	-1.6		3.6	3.4	-1.5		3.9	4.0	-1.6		4.9	4.6	-1.8
GFDL-ESM2G		1.7	1.4	-0.1		2.1	1.4	1.0		3.0	1.7	0.8		3.3	2.1	0.9		4.1	2.2	1.4
GFDL-ESM2M		1.4	1.5	-0.0		2.0	1.6	1.0		3.4	1.9	0.2		3.0	2.2	1.0		4.3	2.5	0.9
HadGEM2-ES		2.7	1.6	-1.3		2.9	1.8	-0.1		3.8	2.0	-0.4		3.5	2.3	0.4		4.3	2.4	1.0
INM-CM4		2.1	1.7	-0.8		3.1	2.4	-0.9		3.5	2.8	-0.9		3.6	3.3	-0.7		5.0	4.0	-1.2
IPSL-CM5A-LR		1.2	1.1	0.6		2.3	1.3	1.0		2.7	1.8	1.0		3.0	1.9	1.3		3.8	2.0	1.9
MIROC-ESM		0.7	1.3	1.0		2.9	1.4	0.4		3.4	1.7	0.4		4.0	2.2	-0.1		4.2	2.2	1.3
MPI-ESM-LR r1		1.3	1.2	0.5		2.7	1.4	0.6		3.0	1.7	0.8		2.7	1.9	1.6		3.5	2.0	2.2
MPI-ESM-LR r2		1.6	1.2	0.1		2.2	1.3	1.0		3.3	1.7	0.4		2.5	1.9	1.8		—	—	—
MPI-ESM-LR r3		1.0	1.1	0.9		2.8	1.3	0.5		3.2	1.7	0.6		2.4	1.9	2.0		—	—	—
MRI-ESM1		1.4	0.8	0.7		2.4	1.2	1.0		2.1	1.3	2.0		2.4	1.5	2.3		2.9	1.8	3.0
NorESM1-ME		2.0	1.4	-0.5		3.0	1.7	-0.2		3.6	2.1	-0.3		4.0	2.5	-0.3		4.4	2.7	0.9

^aOnly models that provided both atmospheric CO₂ and ocean carbon fluxes are included in the multimodel means. Positive values represent additions to the atmosphere, ocean, or land reservoirs, and negative values represent losses. All units are Pg C yr⁻¹.

mate). The multimodel mean anthropogenic atmospheric carbon accumulation in 2060 was 715 ± 173 Pg C, which was above and outside the confidence interval for the CCTM estimate (Figure S7a). In 2100, the CCTM estimate yielded an anthropogenic atmospheric carbon accumulation of 1412 ± 72 Pg C, while the multimodel mean prediction was 1488 ± 347 Pg C (Figure S7b).

To assess the mechanisms causing the strong relationship between contemporary and future atmospheric CO₂ levels among the models, we also developed CCTM-like estimates for the individual ocean and land inventories (Figure S8). This analysis revealed that the ordering of ocean inventories among the models was more persistent into the future than for land inventories, but for both components, statistically significant multimodel relationships existed between contemporary (2010) and future values through the end of the 21st century (Figure 5). However, because uncertainties in ocean and land inventories were larger, constraints offered by contemporary observations were considerably weaker than for atmospheric CO₂, in terms of the future evolution of these inventory components (Figure S9).

3.5. Implications of a Persistent Atmospheric CO₂ Bias

To explore the climate implications of the persistent atmospheric CO₂ biases described above, we compared the radiative forcing (equation (4)) and the resulting temperature change (equations S1 and S2) for the CCTM estimate and the set of CMIP5 model predictions. Figure 8a shows the radiative forcing due only to CO₂ calculated for each of the CMIP5 models. The model range was 5.4–7.4 W m⁻² at year 2100 for RCP 8.5. Figure 8b shows the multimodel mean radiative forcing compared to the radiative forcing for the CCTM estimate. As with the CO₂ comparison described above, the spread of the CCTM was considerably smaller than that of the multimodel 95th percentile distribution spread. In 2100, the CCTM estimate yielded a radiative forcing of 6.4 ± 0.2 W m⁻², while the multimodel mean prediction was 6.6 ± 0.9 W m⁻². Figure 8c shows the corresponding cumulative temperature change due to this CO₂ radiative forcing for each of the CMIP5 models. The temperature increase for the models ranged from 3.3°C to 4.6°C. Figure 8d shows the corresponding multimodel mean cumulative temperature change compared to the CCTM estimate. In 2100, the CCTM estimate yielded a cumulative temperature increase from the CO₂-induced radiative forcing of 4.0 ± 0.1 °C, while the multimodel mean prediction was 4.2 ± 0.6 °C.

The CO₂ mole fraction, CO₂-induced radiative forcing, and CO₂-induced cumulative temperature change for each of the CMIP5 models are shown in Table 3 for the years 2010, 2060, and 2100. In addition, the last three columns of the table show the temperature change bias between the models and the CCTM estimate.

Table 3. Atmospheric CO₂ Mole Fraction, Radiative Forcing, and Resulting Temperature Changes for Each of the CMIP5 ESMs, the Multimodel Mean, the CCTM Estimate, and the Combination of Observed and RCP 8.5 Projection for the Years 2010, 2060, and 2100^a

Model	CO ₂ Mole Fraction (ppm)			Radiative Forcing (W m ⁻²)			Cumulative ΔT (°C)			ΔT Bias (°C)		
	2010	2060	2100	2010	2060	2100	2010	2060	2100	2010	2060	2100
BCC-CSM1.1	390	603	945	1.70	4.03	6.43	0.97	2.39	4.02	0.03	0.02	-0.01
BCC-CSM1.1-M	396	619	985	1.78	4.16	6.65	1.04	2.49	4.16	0.10	0.12	0.13
BNU-ESM	382	602	963	1.59	4.02	6.53	0.90	2.33	4.07	-0.04	-0.04	0.04
CanESM2 r1	394	641	1024	1.75	4.36	6.86	0.98	2.58	4.30	0.04	0.21	0.27
CanESM2 r2	392	641	1023	1.72	4.35	6.85	0.98	2.57	4.30	0.04	0.20	0.27
CanESM2 r3	396	641	1025	1.78	4.35	6.87	1.01	2.58	4.30	0.07	0.21	0.27
CESM1-BGC	407	697	1121	1.92	4.80	7.34	1.12	2.85	4.64	0.18	0.48	0.61
FGOALS-s2.0	404	636	993	1.89	4.31	6.70	1.09	2.57	4.23	0.15	0.20	0.20
GFDL-ESM2G	395	616	967	1.77	4.14	6.56	1.04	2.49	4.12	0.10	0.12	0.09
GFDL-ESM2M	400	621	964	1.83	4.18	6.54	1.09	2.52	4.13	0.15	0.15	0.10
HadGEM2-ES	411	636	983	1.98	4.31	6.64	1.18	2.60	4.20	0.24	0.23	0.17
INM-CM4	386	591	897	1.64	3.92	6.15	0.92	2.36	3.86	-0.02	-0.01	-0.17
IPSL-CM5A-LR	375	573	908	1.48	3.75	6.22	0.86	2.21	3.87	-0.08	-0.16	-0.16
MIROC-ESM	398	658	1121	1.81	4.50	7.35	1.06	2.67	4.58	0.12	0.30	0.55
MPI-ESM-LR r1	383	590	948	1.60	3.91	6.45	0.95	2.31	4.03	0.01	-0.06	0.00
MRI-ESM1	361	516	778	1.28	3.20	5.39	0.74	1.89	3.33	-0.20	-0.48	-0.70
NorESM1-ME	391	667	1070	1.72	4.57	7.09	0.98	2.68	4.46	0.04	0.31	0.43
Multimodel Mean	392	621	980	1.72	4.18	6.63	1.00	2.48	4.17	0.06	0.11	0.14
CCTM Estimate	385	600	948	1.62	4.01	6.45	0.94	2.37	4.03	—	—	—
Historical + RCP 8.5	385	590	917	1.63	3.91	6.27	0.94	2.32	3.93	0.00	-0.05	-0.10

^aValues are 5 year means for the time periods 2006–2010, 2056–2060, and 2096–2100.

In 2010, the temperature bias of the multimodel mean was 0.06°C (ranging from -0.20°C to 0.24°C), and this bias increased to 0.11°C in 2060. Individual model results showed that some biases increased, some decreased, and others remained the same between 2010 and 2060. The MRI-ESM1 and CESM1-BGC models had the largest temperature biases in 2060, at -0.48°C and 0.48°C, respectively, while the INM-CM4 and MPI-ESM-LR models had the smallest temperature biases in 2060, at -0.02°C and 0.01°C, respectively. By 2100, the multimodel mean temperature bias had increased to 0.14°C. The MRI-ESM1 and CESM1-BGC models had the largest temperature biases in 2100, at -0.70°C and 0.61°C, respectively. The temperature biases for individual models were significant and increased with time during the 21st century. The original RCP 8.5 atmospheric CO₂ mole fraction trajectory resulted in a -31 ppm mole fraction bias and a -0.10°C temperature bias from the CCTM estimate at 2100. This result suggests a small inconsistency between the RCP 8.5 specification of the CO₂ mole fraction trajectory and the corresponding fossil fuel emissions trajectory. The RCP 8.5 trajectory was derived from the MESSAGE-MACRO integrated assessment model [Riahi *et al.*, 2011], which incorporates the MAGICC/SCENGEN (version 4.1) coupled gas-cycle/climate model [Wigley, 2003] that includes a net positive carbon cycle feedback but lacks explicit representation of many ecosystem processes that influence climate-carbon and concentration-carbon feedbacks. Prior to its use in deriving the RCP 8.5 trajectory, parameters in the carbon cycle model of MAGICC/SCENGEN (version 5.3) [Wigley, 2008] were changed to give concentration projections consistent with the results from the C⁴MIP activity [Friedlingstein *et al.*, 2006].

It is important to note in the context of the results described above that model-to-model variations in atmospheric CO₂ trajectories documented here contributed to only a small amount of the model-to-model variation in surface air temperature changes. This is because many of the models in the ensemble had different representations of aerosol processes, including forcings and feedbacks, and because the models had widely varying climate sensitivities (e.g., Gillett *et al.* [2013]). Specifically, the multimodel mean estimate of temperature change from the beginning of the simulations was 3.1 ± 1.3°C at 2060 and 5.1 ± 2.2° at 2100. When we adjusted each model temperature estimate for the impact of CO₂ biases using the CO₂-induced temperature biases shown in Table 3, the multimodel mean changed slightly to 3.0 ± 1.2°C at 2060 and 5.0 ± 1.9°C at 2100.

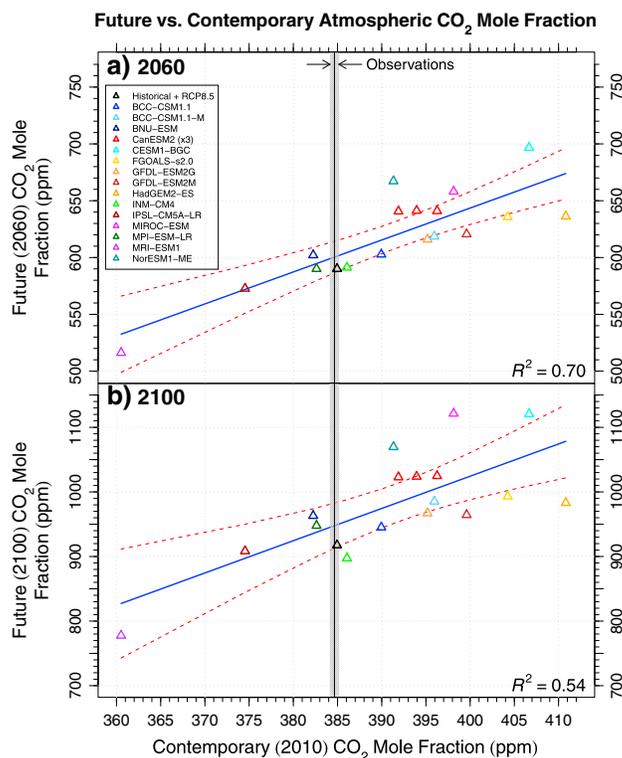


Figure 4. (a) Future (2060) versus contemporary (2010) atmospheric CO₂ mole fraction fit for CMIP5 emissions-forced simulations of RCP 8.5 and (b) future (2100) versus contemporary (2010) atmospheric CO₂ mole fraction for the same set of model simulations. The observed atmospheric CO₂ mole fraction is represented by the vertical line at 384.6 ppm with an uncertainty range (± 0.5 ppm) shown in gray. The linear regression model is represented by the blue line surrounded by red dashed lines indicating a 95% confidence interval. While a point is plotted for the historical observed atmospheric CO₂ and the RCP 8.5 concentration trajectory derived from a reduced form model without explicit feedbacks, that point is not included in the linear regression.

4. Discussion

4.1. Why Do Carbon Cycle Biases Persist on Decadal Timescales?

In our analysis, we found that the ordering among model predictions of atmospheric CO₂ persisted for several decades. Models that had the highest positive biases near the end of the observational record in 2010 were more likely to have higher positive biases in earlier decades, during the latter half of the twentieth century (Figures 1 and 8). Similarly, this same set of models also had the highest set of future atmospheric CO₂ projections during the middle and latter half of the 21st century in response to RCP 8.5 emissions (Figure 4). Many structural model elements probably contributed to this bias and ordering persistence, including pro-

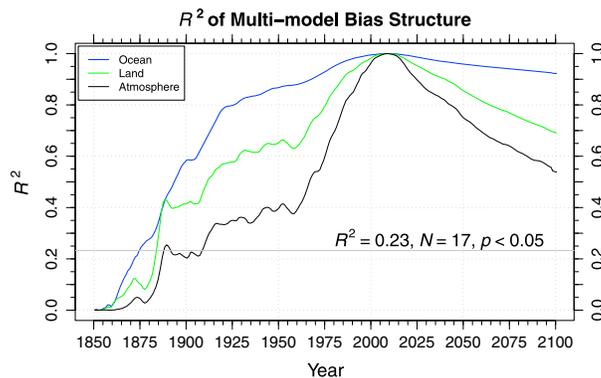


Figure 5. The coefficients of determination (R^2) for the multimodel bias structure, from which the contemporary CO₂ tuned model (CCTM) was derived, relative to the set of CMIP5 model atmospheric CO₂ mole fractions (black) and oceanic (blue) and land (green) anthropogenic carbon inventories in 2010, defined as the 5 year mean for the period 2006–2010.

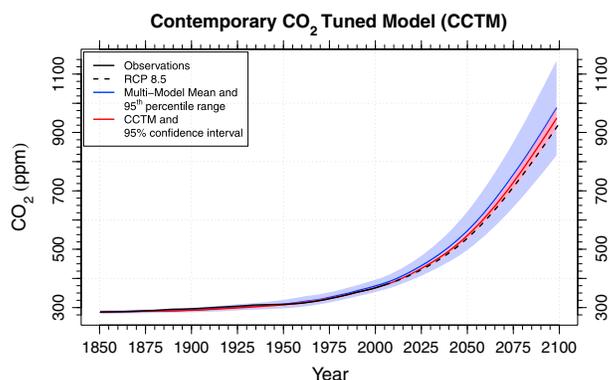


Figure 6. The contemporary CO₂ tuned model (CCTM) atmospheric CO₂ mole fraction estimate compared to the CMIP5 multimodel mean trajectory. The pink range surrounding the CCTM represents the 95% confidence interval from the linear model around the contemporary observation projected onto the y axis of historical or future CO₂ mole fractions for every year. The blue line represents the multimodel mean CO₂ trajectory, and the blue range indicates the 95th percentile of the range for the multimodel standard deviation, assuming a normal distribution (1.96 σ).

cesses that influence the strength of concentration-carbon feedbacks. One important example includes the representation of ocean mixing processes that regulate the formation of intermediate and deep waters in the ocean. Past work from analysis of 13 simulations from the second phase of the Ocean Carbon Cycle Model Intercomparison Project indicated that climate models often underestimate this overturning in the Southern Ocean [Doney *et al.*, 2004; Matsumoto *et al.*, 2004; Dutay *et al.*, 2002]. In addition, Russell *et al.* [2006] performed an intercomparison of the Southern Ocean circulation in CMIP3 control simulations and found that the maximum wind stress in the Southern Hemisphere, nominally associated with the Antarctic Circumpolar Current, was located too far equatorward in most models. In ESMs, such deficiencies in model

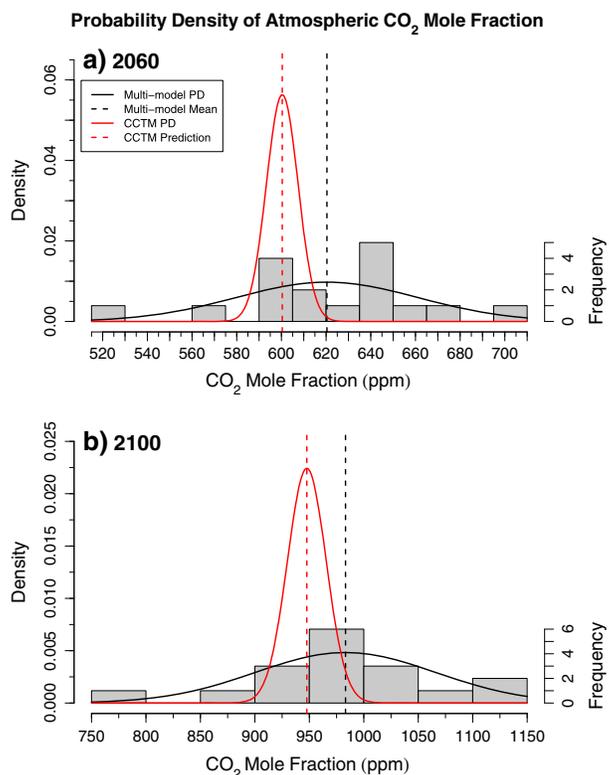


Figure 7. The probability density of CO₂ mole fraction predictions from the CCTM peaks lower than the probability density for multi-model mean for (a) 2060 and (b) 2100. In addition, the width of the probability density is much smaller for the CCTM, by almost a factor of 6 at 2060 and almost a factor of 5 at 2100, indicating a significant reduction in the range of uncertainty for the CCTM prediction.

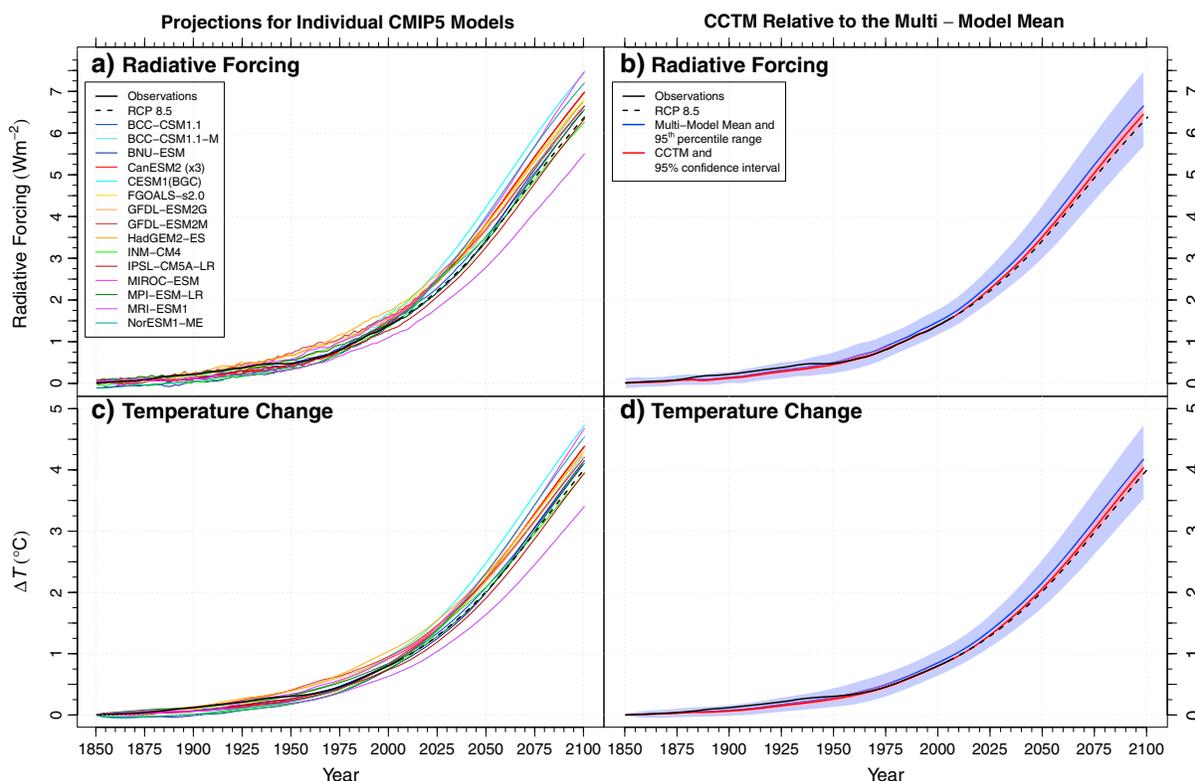


Figure 8. (a) and (c) CO_2 -induced radiative forcing and temperature change computed from the prognostic atmospheric CO_2 mole fraction for each of the CMIP5 models. (b) and (d) Corresponding radiative forcing and temperature change for the multimodel mean and contemporary CO_2 tuned model (CCTM). The pink range surrounding the CCTM represents the uncertainty propagated from the 95% confidence interval from the linear model for the CCTM atmospheric CO_2 trajectory. The blue range surrounding the multimodel mean represents the uncertainty propagated from the 95th percentile of the range for the standard deviation of the multimodel mean atmospheric CO_2 trajectory.

structure and large-scale circulation have the potential to limit CO_2 uptake by the oceans and are likely to contribute to a persistent atmospheric CO_2 bias over time because many of the physical processes regulating mixing are unlikely to change rapidly. Biases in atmospheric CO_2 caused by this type of mechanism likely grow through time as the atmospheric CO_2 growth rate accelerates and transport of carbon out of the mixed layer becomes an increasing bottleneck to net ocean carbon uptake. Our finding that many models underestimated the ocean anthropogenic carbon inventory (Figures 3 and S2) is consistent with other studies indicating some ocean models exhibit weak meridional overturning circulation [Downes et al., 2011; Sallée et al., 2013]. However, additional research is needed to understand the causes of model-to-model variations in ocean carbon uptake for the CMIP5 models.

On land, similar deficiencies in model structure have the potential to contribute to persistent multidecadal biases in carbon fluxes. Key regulators of carbon uptake on land in response to elevated levels of atmospheric CO_2 include, for example, the response of gross primary production (GPP) to CO_2 concentration, the allocation of GPP to longer lived woody pools, and subsequent increases in soil organic matter pools [Thompson et al., 1996; Luo et al., 2006]. Carboxylation parameterizations of Rubisco often follow the form of a modified Michaelis-Menten equation [Farquhar et al., 1980] and vary considerably among models. Models that have lower estimates of the maximum carboxylation rate in different biomes, in response to nitrogen limitation (e.g., Thornton et al. [2007]) or other factors, are likely to have smaller CO_2 -driven increases in GPP by the end of the twentieth or 21st centuries. Similarly, models that have reduced allocation of GPP to wood pools will also have lower rates of carbon uptake, given the same trajectory of GPP increases. Since in many models, the maximum carboxylation rate is either held constant or unlikely to rapidly change in response to changing environmental conditions, this parameterization can induce a long-term bias in carbon fluxes. The same argument applies to allocation submodels: although many plant allocation models are dynamic [Friedlingstein et al., 1999; Arora and Boer, 2005; Litton et al., 2007] and respond to regional variations in light availability, soil moisture, and other environmental controls, many aspects of these models are unlikely

to change rapidly during the twentieth and 21st centuries, allowing flux biases to persist in response to monotonic increases in atmospheric CO₂.

Other land model structural components not associated with concentration-carbon feedbacks also can contribute to long-term flux biases. For example, land use carbon emissions are an important component of the terrestrial carbon budget and are highly uncertain [Hansen *et al.*, 2010; Houghton *et al.*, 2012; Baccini *et al.*, 2012; Andres *et al.*, 2012; Harris *et al.*, 2012]. Model estimates of this flux can be biased if, for example, the representations of aboveground and belowground carbon pools within the model do not capture observed patterns. As a consequence, carbon losses for a given rate of land clearing may be too high or too low, with a bias that is persistent if rates of land clearing change gradually from one decade to the next. Similarly, climate-carbon feedbacks, including, for example, the response of heterotrophic respiration to temperature [Davidson and Janssens, 2006] could also contribute to long-term biases. Nevertheless, for the CMIP5 models, their contribution during the latter half of the twentieth century and first half of the 21st century might be expected to be smaller than other drivers, given that temperature and other changes in climate increase through time [Arora *et al.*, 2013].

The overall success of contemporary atmospheric CO₂ observations in constraining future CO₂ levels (e.g., Figure 7) is probably related to several factors. First, the atmospheric anthropogenic carbon inventory is known relatively well, in contrast to the much larger uncertainties associated with the ocean and land inventories. Second, concentration-carbon feedbacks appear to contribute more to the intermodel variations of future (2100) atmospheric CO₂ level projections than climate-carbon feedbacks [Arora *et al.*, 2013]. In this context, the rapid rise of atmospheric CO₂ observed over the last few decades provides an important direct test of the combined set of ocean and land concentration-carbon mechanisms operating within the models, and, as described above, any biases today are likely amplified as the growth rate of CO₂ accelerates. Although temperature and other climate changes also occurred during this period, the magnitude of these changes was much smaller as compared to what is expected during the middle and latter part of the 21st century. As a consequence, the variations in atmospheric CO₂ estimates among models resulting from climate-carbon feedbacks were likely relatively small for the contemporary period (e.g., Arora *et al.* [2013]). The growing importance of climate-carbon feedbacks probably contributes to the increasing uncertainty in our CCTM estimate during the latter part of the 21st century (Figure 6).

4.2. What Is the Value of Improving Carbon Cycle Processes to Match Contemporary CO₂?

One of the goals of the integrated assessment modeling community in developing the different representative concentration pathways (RCPs) was to enable ESMs to compute possible emissions scenarios consistent with a particular atmospheric CO₂ trajectory [van Vuuren *et al.*, 2011]. This is valuable, for example, in identifying the magnitude of required mitigation efforts to stabilize CO₂ levels in the atmosphere at a particular mole fraction, taking into account carbon cycle responses and feedbacks (e.g., Jones *et al.* [2013]).

Our analysis has several implications for the interpretation of future compatible emissions time series derived from the set of ESMs participating in CMIP5. First, the compatible emissions time series derived from the multimodel mean of concentration-forced simulations during the 21st century is likely to be too low. This assertion is based on the observation that (1) fossil fuel emissions would have to be reduced below observations to eliminate the high bias found in the multimodel mean during the last few decades (Figure 1) and (2) our finding that biases observed today were significantly correlated with future atmospheric CO₂ projections because of parameterizations of slowly changing carbon cycle processes.

Second, the range of variation in compatible emission estimates among individual models during the remainder of the 21st century has a large component that can be avoided for any given concentration-forced scenario by reducing or eliminating biases in contemporary atmospheric CO₂. Specifically, if each model were individually optimized to eliminate biases in atmospheric CO₂ during the last few decades, the range of compatible emissions projections during the 21st century would be considerably compressed. In our analysis, we investigated the potential magnitude of this uncertainty reduction by using the entire set of CMIP5 ESMs to construct a tuned model (CCTM). Projections from the CCTM provided almost a sixfold reduction in uncertainty of atmospheric CO₂ levels at 2060 and nearly a fivefold reduction at 2100. As previously noted, the range of model projections diverges through time during the 21st century, as climate-carbon cycle feedbacks strengthen. However, even by the end of the century, a significant component of the variation among models can be attributed to biases that exist today. This result is consistent with results from Arora *et al.* [2013] that show much of the model-to-model variation in carbon cycle

estimates is driven by concentration-carbon feedbacks and only to a lesser degree by variation in climate-carbon feedbacks.

Considering the carbon cycle a “black box” from the perspective of climate change impacts on other aspects of the Earth system, there is significant value in model development efforts to eliminate biases in atmospheric CO₂ that occur by the end of the observational record. By doing so for the set of simulations evaluated here, high biases in radiative forcing and global temperature increases could be reduced in many of the models (Figure 8). Improved estimates of CO₂-induced climate change, in turn, would reduce uncertainties related to rates of snow and ice melt [Flanner *et al.*, 2009] and other processes contributing to climate feedbacks [Hall and Qu, 2006; Davidson and Janssens, 2006; Zaehle *et al.*, 2010; Koven *et al.*, 2011]. Benefits would also exist for developing more precise estimates of changes in ocean surface chemistry [Caldeira and Wickett, 2003; Doney *et al.*, 2009a] and ocean circulation [Downes *et al.*, 2011; Sallée *et al.*, 2013], and better estimates of climate change impacts on agriculture and other aspects of human society [Lobell *et al.*, 2011].

An interesting question then emerges regarding how best to reduce these biases within individual models and for the set of ESMs as a whole contributing to future climate assessments. Many structural elements of the models may be improved through extensive comparison of ESMs with observations and the development of community-wide benchmarking and evaluation systems such as the International Land Model Benchmarking (ILAMB) project [Luo *et al.*, 2012] and equivalent ocean projects [Doney *et al.*, 2009b]. These efforts are underway, and significant advances are expected over the next several years. Biases also may be reduced by having closer coordination among different ESM development teams and allocating more time to evaluating coupled transient ESM simulations during the nineteenth and twentieth centuries. More specifically, given that constraints on some long-term flux components are uncertain, modeling teams may need to optimize several sets of parameters to achieve a more realistic integrated carbon simulation. For example, adjustments to parameterizations of subgrid scale mesoscale eddy mixing can improve many aspects of physical ocean system [Gent and McWilliams, 1990; Danabasoglu and Marshall, 2007; Danabasoglu *et al.*, 2008, Gent, 2011] but may have unintended consequences for ocean carbon uptake. At a minimum, more quantification and analysis of these trade-offs is needed, and ocean carbon benchmarks need to be fully considered when modifications are made to ocean model physics.

On land, uncertainties in land use histories and responses of carbon storage to elevated CO₂ and other changing resources provide additional opportunities for making model adjustments that can improve the fidelity of the model's overall atmospheric CO₂ trajectory but not conflict with available data constraints. Ecosystem manipulation experiments and observations also are needed to improve our understanding of ecosystem processes and their representation in models. In addition, a robust set of Earth system observations is needed to quantify climate change impacts on terrestrial carbon sinks and carbon dynamics associated with land use change.

5. Conclusions

The trajectories of atmospheric CO₂ mole fraction for 19 historical and 17 future emission-driven simulation results produced for CMIP5 by 15 fully coupled ESMs were analyzed. Comparison of ESM prognostic atmospheric CO₂ over the historical period with observations indicated that ESMs, on average, had a high bias in their predictions of contemporary CO₂ levels. Comparison with observationally based estimates of anthropogenic carbon inventories in the ocean indicated that this bias was driven by weak to nominal ocean carbon uptake in many ESMs and that terrestrial and ocean carbon accumulation often compensated for one another within individual models, reducing the bias in predicted atmospheric CO₂. We found a linear relationship over decadal timescales between contemporary and future atmospheric CO₂ mole fractions and used this relationship to construct a model of the atmospheric CO₂ trajectory tuned to contemporary observations, which we called the CCTM. CCTM estimates of atmospheric CO₂ were 21 ppm lower than the multimodel mean at 2060 and 32 ppm lower at 2100. Using an impulse response function, we approximated radiative forcing and temperature changes resulting from ESM, CCTM, and observed CO₂ trajectories. Comparison of temperature changes from ESMs with the CCTM estimate indicated a small positive multimodel mean bias during the 21st century. Individual model results exhibited a much larger range of CO₂-induced temperature change, from 1.9°C to 2.9°C in 2060 and from 3.3°C to 4.6°C in 2100, demonstrating the net

effect and significant climate implications associated with the large model spread in carbon accumulation in ocean and land reservoirs.

Atmospheric CO₂ biases persist in models for decades because parameterizations of biological and physical processes related to carbon accumulation on land and in the ocean do not allow the system to change rapidly. Many of the biases associated with concentration-carbon feedbacks (i.e., Arora *et al.* [2013]) likely increase through time in the RCP 8.5 scenario as the atmospheric CO₂ growth rate accelerates. Because of the high atmospheric CO₂ bias exhibited by ESMs for the contemporary period, future fossil fuel emissions trajectories designed to stabilize atmospheric CO₂ levels, sometimes called “allowable” emissions, would be too low if estimated from the multimodel mean. We have shown that a significant component of the variation of atmospheric CO₂ levels among models during the 21st century was linked to biases in their predictions of contemporary atmospheric CO₂. This suggests improving the agreement of individual models with the contemporary atmospheric CO₂ record could reduce the magnitude of future CO₂ biases in many models and narrow the range of predicted radiative forcing and CO₂-induced global temperature increases. To reduce biases in individual models, a rigorous campaign of extensive and multifaceted evaluation—directed at improving model structure and optimizing model parameters through comparison with contemporary observations—must be performed. Community-based benchmarking and model evaluation systems, such as ILAMB, tighter coordination among ESM development teams, and optimization of model parameters using all available observational constraints has the potential to both reduce model biases and significantly decrease the multimodel spread of carbon cycle predictions for future development scenarios and mitigation.

Acknowledgments

The authors wish to thank R. J. Stouffer for his thoughtful comments on an early draft of the paper, which lead to improvements in the discussion of our results. In addition, the authors gratefully acknowledge the careful reviews of T. A. Boden, R. J. Norby, and S. Denning. This research was sponsored by the Climate and Environmental Sciences Division (CESD) of the Biological and Environmental Research (BER) Program in the U. S. Department of Energy Office of Science and the National Science Foundation (AGS-1048890). This research used resources of the National Center for Computational Sciences (NCCS) at Oak Ridge National Laboratory (ORNL), which is managed by UT-Battelle, LLC, for the U. S. Department of Energy under contract DE-AC05-00OR22725. CDJ was supported by the Joint DECC/Defra Met Office Hadley Center Climate Program (GA01101). This is a contribution to the BIOFEEDBACK project of the Center for Climate Dynamics (SKD) at the Bjerknes Center for Climate Research. The National Center for Atmospheric Research is sponsored by the National Science Foundation. We acknowledge the World Climate Research Program's Working Group on Coupled Modeling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Table 1 of this paper) for producing and making available their model output. For CMIP, the U. S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

References

- Anav, A., P. Friedlingstein, M. Kidston, L. Bopp, P. Ciais, P. Cox, C. Jones, M. Jung, R. Myneni, and Z. Zhu (2013), Evaluating the land and ocean components of the global carbon cycle in the CMIP5 Earth System Models, *J. Clim.*, *26*(18), 6801–6843, doi:10.1175/JCLI-D-12-00417.1.
- Andres, R. J., J. S. Gregg, L. Losey, G. Marland, and T. A. Boden (2011), Monthly, global emissions of carbon dioxide from fossil fuel consumption, *Tellus B.*, *63*(3), 309–327, doi:10.1111/j.1600-0889.2011.00530.x.
- Andres, R. J., et al. (2012), A synthesis of carbon dioxide emissions from fossil-fuel combustion, *Biogeosciences*, *9*(5), 1845–1871, doi:10.5194/bg-9-1845-2012.
- Arora, V. K., and G. J. Boer (2005), A parameterization of leaf phenology for the terrestrial ecosystem component of climate models, *Global Change Biol.*, *11*(1), 39–59, doi:10.1111/j.1365-2486.2004.00890.x.
- Arora, V. K., J. F. Scinocca, G. J. Boer, J. R. Christian, K. L. Denman, G. M. Flato, V. V. Kharin, W. G. Lee, and W. J. Merryfield (2011), Carbon emission limits required to satisfy future representative concentration pathways of greenhouse gases, *Geophys. Res. Lett.*, *38*, L05805, doi:10.1029/2010GL046270.
- Arora, V. K., et al. (2013), Carbon-concentration and carbon-climate feedbacks in CMIP5 Earth System Models, *J. Clim.*, *26*(15), 5289–5314, doi:10.1175/JCLI-D-12-00494.1.
- Baccini, A., et al. (2012), Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps, *Nat. Clim. Change*, *2*(3), 182–185, doi:10.1038/nclimate1354.
- Ballantyne, A. P., C. B. Alden, J. B. Miller, P. P. Tans, and J. W. C. White (2012), Increase in observed net carbon dioxide uptake by land and oceans during the past 50 years, *Nature*, *488*(7409), 70–72, doi:10.1038/nature11299.
- Bao, Q., et al. (2013), The Flexible Global Ocean-Atmosphere-Land System model, spectral version 2: FGOALS-s2, *Adv. Atmos. Sci.*, *30*(3), 561–576, doi:10.1007/s00376-012-2113-9.
- Bentsen, M., et al. (2013), The Norwegian Earth System Model, NorESM1-M—Part 1: Description and basic evaluation of the physical climate, *Geosci. Model Dev.*, *6*(3), 687–720, doi:10.5194/gmd-6-687-2013.
- Boucher, O., and M. S. Reddy (2008), Climate trade-off between black carbon and carbon dioxide emissions, *Energ. Policy*, *36*(1), 193–200, doi:10.1016/j.enpol.2007.08.039.
- Brovkin, V., T. Raddatz, C. H. Reick, M. Claussen, and V. Gayler (2009), Global biogeophysical interactions between forest and climate, *Geophys. Res. Lett.*, *36*, L07405, doi:10.1029/2009GL037543.
- Caldeira, K., and M. E. Wickett (2003), Oceanography: Anthropogenic carbon and ocean pH, *Nature*, *425*(6956), 365, doi:10.1038/425365a.
- College of Global Change and Earth System Science (2012), Beijing Normal University-Earth System Model. http://esg.bnu.edu.cn/BNU_ESM_webs/htmls/index.html.
- Collins, W. J., et al. (2011), Development and evaluation of an Earth-System model—HadGEM2, *Geosci. Model Dev.*, *4*(4), 1051–1075, doi:10.5194/gmd-4-1051-2011.
- Conway, T. J., P. P. Tans, L. S. Waterman, K. W. Thoning, D. R. Kitzis, K. A. Masarie, and N. Zhang (1994), Evidence for interannual variability of the carbon cycle from the National Oceanic and Atmospheric Administration/Climate Monitoring and Diagnostics Laboratory Global Air Sampling Network, *J. Geophys. Res.*, *99*(D11), 22,831–22,855, doi:10.1029/94JD01951.
- Cox, P. M., D. Pearson, B. B. Booth, P. Friedlingstein, C. Huntingford, C. D. Jones, and C. M. Luke (2013), Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability, *Nature*, *494*(7437), 341–344, doi:10.1038/nature11882.
- Dai, Y., et al. (2003), The common land model, *Bull. Am. Meteorol. Soc.*, *84*(8), 1013–1023, doi:10.1175/BAMS-84-8-1013.
- Dai, Y., R. E. Dickinson, and Y.-P. Wang (2004), A two-big-leaf model for canopy temperature, photosynthesis, and stomatal conductance, *J. Clim.*, *17*(12), 2281–2299, doi:10.1175/1520-0442(2004)017<2281:ATMFTC>2.0.CO;2.
- Danabasoglu, G., and J. Marshall (2007), Effects of vertical variations of thickness diffusivity in an ocean general circulation model, *Ocean Model.*, *18*(2), 122–141, doi:10.1016/j.ocemod.2007.03.006.
- Danabasoglu, G., R. Ferrari, and J. C. McWilliams (2008), Sensitivity of an ocean general circulation model to a parameterization of near-surface eddy fluxes, *J. Clim.*, *21*(6), doi:10.1175/2007JCLI1508.1.

- Davidson, E. A., and I. A. Janssens (2006), Temperature sensitivity of soil carbon decomposition and feedbacks to climate change, *Nature*, *440*(7081), 165–173, doi:10.1038/nature04514.
- Denman, K. L., et al. (2007), Couplings between changes in the climate system and biogeochemistry, 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. 499–588, Cambridge Univ. Press, Cambridge, United Kingdom and New York, NY, USA.
- Doney, S. C., et al. (2004), Evaluating global ocean carbon models: The importance of realistic physics, *Global Biogeochem. Cycles*, *18*, GB3017, doi:10.1029/2003GB002150.
- Doney, S. C., V. J. Fabry, R. A. Feely, and J. A. Kleypas (2009a), Ocean acidification: The other CO₂ problem, *Annu. Rev. Mar. Sci.*, *1*(1), 169–192, doi:10.1146/annurev.marine.010908.163834.
- Doney, S. C., I. Lima, J. K. Moore, K. Lindsay, M. J. Behrenfeld, T. K. Westberry, N. Mahowald, D. M. Glover, and T. Takahashi (2009b), Skill metrics for confronting global upper ocean ecosystem-biogeochemistry models against field and remote sensing data, *J. Mar. Syst.*, *76*(1–2), 95–112, doi:10.1016/j.jmarsys.2008.05.015.
- Downes, S. M., A. Gnanadesikan, S. M. Griffies, and J. L. Sarmiento (2011), Water mass exchange in the Southern Ocean in coupled climate models, *J. Phys. Oceanogr.*, *41*(9), 1756–1771, doi:10.1175/2011JPO4586.1.
- Dufresne, J. L., et al. (2013), Climate change projections using the IPSL-CM5 Earth System Model: From CMIP3 to CMIP5, *Clim. Dyn.*, *40*(9–10), 2123–2165, doi:10.1007/s00382-012-1636-1.
- Dunne, J. P., et al. (2012), GFDL's ESM2 global coupled climate-carbon Earth System Models. Part I: Physical formulation and baseline simulation characteristics, *J. Clim.*, *25*(19), 6646–6665, doi:10.1175/JCLI-D-11-00560.1.
- Dunne, J. P., et al. (2013), GFDL's ESM2 global coupled climate-carbon Earth System Models. Part II: Carbon system formulation and baseline simulation characteristics, *J. Clim.*, *26*(7), 2247–2267, doi:10.1175/JCLI-D-12-00150.1.
- Dutay, J.-C., et al. (2002), Evaluation of ocean model ventilation with CFC-11: Comparison of 13 global ocean models, *Ocean Model.*, *4*(2), 89–120, doi:10.1016/S1463-5003(01)00013-0.
- Etheridge, D. M., L. P. Steele, R. L. Langenfelds, R. J. Francey, J.-M. Barnola, and V. I. Morgan (1996), Natural and anthropogenic changes in atmospheric CO₂ over the last 1000 years from air in Antarctic ice and firn, *J. Geophys. Res.*, *101*(D2), 4115–4128, doi:10.1029/95JD03410.
- Farquhar, G. D., S. Caemmerer, and J. A. Berry (1980), A biochemical model of photosynthetic CO₂ assimilation in leaves of C₃ species, *Planta*, *149*(1), 78–90, doi:10.1007/BF00386231.
- Flanner, M. G., C. S. Zender, P. G. Hess, N. M. Mahowald, T. H. Painter, V. Ramanathan, and P. J. Rasch (2009), Springtime warming and reduced snow cover from carbonaceous particles, *Atmos. Chem. Phys.*, *9*(7), 2481–2497, doi:10.5194/acp-9-2481-2009.
- Friedlingstein, P., G. Joel, C. B. Field, and I. Y. Fung (1999), Toward an allocation scheme for global terrestrial carbon models, *Global Change Biol.*, *5*(7), 755–770, doi:10.1046/j.1365-2486.1999.00269.x.
- Friedlingstein, P., J.-L. Dufresne, P. M. Cox, and P. Rayner (2003), How positive is the feedback between climate change and the carbon cycle?, *Tellus*, *55B*(2), 692–700, doi:10.1034/j.1600-0889.2003.01461.x.
- Friedlingstein, P., et al. (2006), Climate-carbon cycle feedback analysis, results from the C⁴ MIP model intercomparison, *J. Clim.*, *19*(14), 3373–3353, doi:10.1175/JCLI3800.1.
- Gent, P. R. (2011), The Gent–McWilliams parameterization: 20/20 hindsight, *Ocean Model.*, *39*(1–2), 2–9, doi:10.1016/j.ocemod.2010.08.002.
- Gent, P. R., and J. C. McWilliams (1990), Isopycnal mixing in ocean circulation models, *J. Phys. Oceanogr.*, *20*(1), 150–155, doi:10.1175/1520-0485(1990)020<0150:IMIOCM>2.0.CO;2.
- Gillett, N. P., V. K. Arora, D. Matthews, and M. R. Allen (2013), Constraining the ratio of global warming to cumulative CO₂ emissions using CMIP5 simulations, *J. Clim.*, *26*(18), 6844–6858, doi:10.1175/JCLI-D-12-00476.1.
- Girod, C. M., G. C. Hurtt, S. Frolking, J. D. Aber, and A. W. King (2007), The tension between fire risk and carbon storage: Evaluating U.S. carbon and fire management strategies through ecosystem models, *Earth Interact.*, *11*(2), 1–33, doi:10.1175/EI188.1.
- Gregory, J. M., C. D. Jones, P. Cadule, and P. Friedlingstein (2009), Quantifying carbon cycle feedbacks, *J. Clim.*, *22*(19), 5232–5250, doi:10.1175/2009JCLI2949.1.
- Gruber, N., et al. (2009), Oceanic sources, sinks, and transport of atmospheric CO₂, *Global Biogeochem. Cycles*, *23*, GB1005, doi:10.1029/2008GB003349.
- Hall, A., and X. Qu (2006), Using the current seasonal cycle to constrain snow albedo feedback in future climate change, *Geophys. Res. Lett.*, *33*(3), L03502, doi:10.1029/2005GL025127.
- Hansen, M. C., S. V. Stehman, and P. V. Potapov (2010), Quantification of global gross forest cover loss, *Proc. Nat. Acad. Sci.*, *107*(19), 8650–8655, doi:10.1073/pnas.0912668107.
- Harris, N. L., S. Brown, S. C. Hagen, S. S. Saatchi, S. Petrova, W. Salas, M. C. Hansen, P. V. Potapov, and A. Lotsch (2012), Baseline map of carbon emissions from deforestation in tropical regions, *Science*, *336*(6088), 1573–1576, doi:10.1126/science.1217962.
- Houghton, R. A., G. R. van der Werf, R. S. De Fries, M. C. Hansen, J. I. House, C. Le Quéré, J. Pongratz, and N. Ramankutty (2012), Chapter G2: Carbon emissions from land use and land-cover change, *Biogeosci. Discuss.*, *9*(1), 835–878, doi:10.5194/bgd-9-835-2012.
- Hurrell, J. W., et al. (2013), The Community Earth System Model: A framework for collaborative research, *Bull. Am. Meteorol. Soc.*, *94*, 1339–1360, doi:10.1175/BAMS-D-12-00121.
- Hurt, G. C., S. W. Pacala, P. R. Moorcroft, J. Caspersen, E. Shevliakova, R. A. Houghton, and B. Moore (2002), Projecting the future of the U.S. carbon sink, *Proc. Nat. Acad. Sci.*, *99*(3), 1389–1394, doi:10.1073/pnas.012249999.
- IPCC (2007), Summary for policymakers, 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., Cambridge Univ. Press, Cambridge, United Kingdom and New York, NY, USA.
- Iversen, T., et al. (2013), The Norwegian Earth System Model, NorESM1-M—Part 2: Climate response and scenario projections, *Geosci. Model Dev.*, *6*(2), 389–415, doi:10.5194/gmd-6-389-2013.
- Jones, C., et al. (2013), 21st century compatible CO₂ emissions and airborne fraction simulated by CMIP5 Earth System Models under four representative concentration pathways, *J. Clim.*, *26*(13), 4398–4413, doi:10.1175/JCLI-D-12-00554.1.
- Jones, C. D., et al. (2011), The HadGEM2-ES implementation of CMIP5 centennial simulations, *Geosci. Model Dev.*, *4*(3), 543–570, doi:10.5194/gmd-4-543-2011.
- Keppel-Aleks, G., et al. (2013), Atmospheric carbon dioxide variability in the community Earth System Model: Evaluation and transient dynamics during the 20th and 21st centuries, *J. Clim.*, *26*(13), 4447–4475, doi:10.1175/JCLI-D-12-00589.1.
- Khatiwal, S., F. Primeau, and T. Hall (2009), Reconstruction of the history of anthropogenic CO₂ concentrations in the ocean, *Nature*, *462*(7271), 346–349, doi:10.1038/nature08526.
- Khatiwal, S., et al. (2013), Global ocean storage of anthropogenic carbon, *Biogeosciences*, *10*(4), 2169–2191, doi:10.5194/bg-10-2169-2013.

- Koven, C. D., B. Ringeval, P. Friedlingstein, P. Ciais, P. Cadule, D. Khvorostyanov, G. Krinner, and C. Tarnocai (2011), Permafrost carbon-climate feedbacks accelerate global warming, *Proc. Nat. Acad. Sci.*, 108(36), 14,769–14,774, doi:10.1073/pnas.1103910108.
- Le Quéré, C., et al. (2013), The global carbon budget 1959–2011, *Earth Syst. Sci. Data*, 5(1), 165–185, doi:10.5194/essd-5-165-2013.
- Lee, K., C. L. Sabine, T. Tanhua, T.-W. Kim, R. A. Feely, and H.-C. Kim (2011), Roles of marginal seas in absorbing and storing fossil fuel CO₂, *Energy Environ. Sci.*, 4(4), 1133–1146, doi:10.1039/C0EE00663G.
- Lin, P., Y. Yu, and H. Liu (2013), Long-term stability and oceanic mean state simulated by the coupled model FGOALS-s2, *Adv. Atmos. Sci.*, 30(1), 175–192, doi:10.1007/s00376-012-2042-7.
- Litton, C. M., J. W. Raich, and M. G. Ryan (2007), Carbon allocation in forest ecosystems, *Global Change Biol.*, 13(10), 2089–2109, doi:10.1111/j.1365-2486.2007.01420.x.
- Liu, H., P. Lin, Y. Yu, and X. Zhang (2012), The baseline evaluation of LASG/IAP climate system ocean model (LICOM) version 2, *Acta Meteor. Sinica*, 26(3), 318–329, doi:10.1007/s13351-012-0305-y.
- Lobell, D. B., W. Schlenker, and J. Costa-Roberts (2011), Climate trends and global crop production since 1980, *Science*, 333(6042), 616–620, doi:10.1126/science.1204531.
- Long, M. C., K. Lindsay, S. Peacock, J. Keith Moore, and S. C. Doney (2013), Twentieth-century oceanic carbon uptake and storage in CESM1(BGC), *J. Clim.*, 26(18), 6775–6800, doi:10.1175/JCLI-D-12-00184.1.
- Luo, Y., D. Hui, and D. Zhang (2006), Elevated CO₂ stimulates net accumulations of carbon and nitrogen in land ecosystems: A meta-analysis, *Ecology*, 87(1), 53–63, doi:10.1890/04-1724.
- Luo, Y. Q., et al. (2012), A framework for benchmarking land models, *Biogeosciences*, 9(10), 3857–3874, doi:10.5194/bg-9-3857-2012.
- Maier-Reimer, E., I. Kriest, J. Segschneider, and P. Wetzol, (2005), The HAMBURG Ocean Carbon Cycle model HAMOCC 5.1—Technical description, Release 1.1, *Tech. rep.*, Max-Planck Institute for Meteorology.
- Manning, A. C., and R. F. Keeling (2006), Global oceanic and land biotic carbon sinks from the Scripps atmospheric oxygen flask sampling network, *Tellus*, 58(2), 95–116, doi:10.1111/j.1600-0889.2006.00175.x.
- Matsumoto, K., et al. (2004), Evaluation of ocean carbon cycle models with data-based metrics, *Geophys. Res. Lett.*, 31(7), L07303, doi:10.1029/2003GL018970.
- Meinshausen, M., et al. (2011), The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, *Clim. Change*, 109(1), 213–241, doi:10.1007/s10584-011-0156-z.
- Moss, R. H., et al. (2010), The next generation of scenarios for climate change research and assessment, *Nature*, 463(7282), 747–756, doi:10.1038/nature08823.
- Nakano, H., H. Tsujino, M. Hirabara, T. Yasuda, T. Motoi, M. Ishii, and G. Yamanaka (2011), Uptake mechanism of anthropogenic CO₂ in the Kuroshio Extension region in an ocean general circulation model, *J. Oceanogr.*, 67(6), 765–783, doi:10.1007/s10872-011-0075-7.
- Obata, A., and K. Shibata (2012), Damage of land biosphere due to intense warming by 1000-fold rapid increase in atmospheric methane: Estimation with a climate–carbon cycle model, *J. Clim.*, 25(24), 8524–8541, doi:10.1175/JCLI-D-11-00533.1.
- Oschlies, A. (2001), Model-derived estimates of new production: New results point towards lower values, *Deep Sea Research Part II: Topical Studies in Oceanography*, 48(10), 2173–2197, doi:10.1016/S0967-0645(00)00184-3.
- Pan, Y., et al. (2011), A large and persistent carbon sink in the world's forests, *Science*, 333(6045), 988–993, doi:10.1126/science.1201609.
- Peters, G. P., R. M. Andrew, T. Boden, J. G. Canadell, P. Ciais, C. Le Quéré, G. Marland, M. R. Raupach, and C. Wilson (2013), The challenge to keep global warming below 2°C, *Nat. Clim. Change*, 3(1), 4–6, doi:10.1038/nclimate1783.
- Phillips, O. L., et al. (2009), Drought sensitivity of the Amazon rainforest, *Science*, 323(5919), 1344–1347, doi:10.1126/science.1164033.
- Raddatz, T., C. Reick, W. Knorr, J. Kattge, E. Roeckner, R. Schnur, K.-G. Schnitzler, P. Wetzol, and J. Jungclaus (2007), Will the tropical land biosphere dominate the climate–carbon cycle feedback during the twenty-first century?, *Clim. Dyn.*, 29(6), 565–574, doi:10.1007/s00382-007-0247-8.
- Ramaswamy, V., O. Boucher, J. Haigh, D. Hauglustaine, J. Haywood, G. Myhre, T. Nakajima, G. Y. Shi, and S. Solomon (2001), Radiative forcing of climate change, in *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, edited by J. T. Houghton et al., pp. 350–416, Cambridge Univ. Press, Cambridge, United Kingdom and New York, NY, USA.
- Randerson, J. T. (2013), Climate science: Global warming and tropical carbon, *Nature*, 494(7437), 319–320, doi:10.1038/nature11949.
- Raupach, M. R., and J. G. Canadell (2010), Carbon and the anthropocene, *Curr. Opin. Environ. Sustainability*, 2(4), 210–218, doi:10.1016/j.cosust.2010.04.003.
- Riahi, K., S. Rao, V. Krey, C. Cho, V. Chirkov, G. Fischer, G. Kindermann, N. Nakicenovic, and P. Rafaj (2011), RCP 8.5—A scenario of comparatively high greenhouse gas emissions, *Clim. Change*, 109(1), 33–57, doi:10.1007/s10584-011-0149-y.
- Russell, J. L., R. J. Stouffer, and K. W. Dixon (2006), Intercomparison of the Southern Ocean circulations in IPCC coupled model control simulations, *J. Clim.*, 19(18), 4560–4575, doi:10.1175/JCLI3869.1.
- Sabine, C. L., et al. (2004), The oceanic sink for anthropogenic CO₂, *Science*, 305(5682), 367–371, doi:10.1126/science.1097403.
- Sallée, J.-B., E. Shuckburgh, N. Bruneau, A. J. S. Meijers, T. J. Bracegirdle, Z. Wang, and T. Roy (2013), Assessment of Southern Ocean water mass circulation and characteristics in CMIP5 models: Historical bias and forcing response, *J. Geophys. Res. Oceans*, 118, 1830–1844, doi:10.1002/jgrc.20135.
- Tans, P. (2009), An accounting of the observed increase in oceanic and atmospheric CO₂ and an outlook for the future, *Oceanography*, 22(4), 26–35, doi:10.5670/oceanog.2009.94.
- Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An overview of CMIP5 and the experiment design, *Bull. Am. Meteorol. Soc.*, 93(4), 485–498, doi:10.1175/BAMS-D-11-00094.1.
- Thompson, M. V., J. T. Randerson, C. M. Malmström, and C. B. Field (1996), Change in net primary production and heterotrophic respiration: How much is necessary to sustain the terrestrial carbon sink?, *Global Biogeochem. Cycles*, 10(4), 711–726, doi:10.1029/96GB01667.
- Thornton, P. E., J.-F. Lamarque, N. A. Rosenbloom, and N. M. Mahowald (2007), Influence of carbon-nitrogen cycle coupling on land model response to CO₂ fertilization and climate variability, *Global Biogeochem. Cycles*, 21, GB4018, doi:10.1029/2006GB002868.
- Tjiputra, J. F., C. Roelandt, M. Bentsen, D. M. Lawrence, T. Lorentzen, J. Schwinger, Ø. Seland, and C. Heinze (2013), Evaluation of the carbon cycle components in the Norwegian Earth System Model (NorESM), *Geosci. Model Dev.*, 6(2), 301–325, doi:10.5194/gmd-6-301-2013.
- Todd-Brown, K. E. O., J. T. Randerson, W. M. Post, F. M. Hoffman, C. Tarnocai, E. A. G. Schuur, and S. D. Allison (2013), Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations, *Biogeosciences*, 10(3), 1717–1736, doi:10.5194/bg-10-1717-2013.
- van Vuuren, D. et al. (2011), The representative concentration pathways: An overview, *Clim. Change*, 109(1), 5–31, doi:10.1007/s10584-011-0148-z.

- Volodin, E., N. Dianskii, and A. Gusev (2010), Simulating present-day climate with the INMCM4.0 coupled model of the atmospheric and oceanic general circulations, *Izv. Atmos. Ocean. Phys.*, *46*(4), 414–431, doi:10.1134/S000143381004002X.
- Watanabe, S., et al. (2011), MIROC-ESM 2010: Model description and basic results of CMIP5-20c3m experiments, *Geosci. Model Dev.*, *4*(4), 845–872, doi:10.5194/gmd-4-845-2011.
- Wigley, T. M. L. (2003), *MAGICC/SCENGEN 4.1, User Manual*, National Center for Atmospheric Research, Boulder, CO 80307.
- Wigley, T. M. L. (2008), *MAGICC/SCENGEN 5.3 (version 2), User Manual*, National Center for Atmospheric Research, Boulder, CO 80307.
- Williams, D. N., et al. (2011), The Earth System Grid Federation: Software framework supporting CMIP5 data analysis and dissemination, *CLIVAR Exchanges*, *16*(2), 40–42.
- Wu, T., et al. (2013), Global carbon budgets simulated by the Beijing Climate Center Climate System Model for the last century, *J. Geophys. Res. Atmos.*, *118*(10), 4326–4347, doi:10.1002/jgrd.50320.
- Yukimoto, S., et al. (2011), Meteorological Research Institute-Earth System Model version 1 (MRI-ESM1)—Model description, *Tech. Rep. 64*, Meteorological Research Institute.
- Yukimoto, S., Y. Adachi, and M. Hosaka (2012), A new global climate model of the Meteorological Research Institute: MRI-CGCM3—Model description and basic performance, *J. Meteorol. Soc. Jpn.*, *90*(A), 23–64, doi:10.2151/jmsj.2012-A02.
- Zaehle, S., P. Friedlingstein, and A. D. Friend (2010), Terrestrial nitrogen feedbacks may accelerate future climate change, *Geophys. Res. Lett.*, *37*, L01401, doi:10.1029/2009GL041345.