CMIP5 Analysis and Model Benchmarking: Quantification and Reduction of Uncertainties Associated with Carbon Cycle–Climate System Feedbacks

Forrest M. Hoffman

Climate Change Science Institute (CCSI),
Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA

PLUME-MIP Workshop
East China Normal University, Shanghai, China
September 26, 2015

Climate Change Science Institute
Oak Ridge National Laboratory

Oak Ridge National Laboratory

Research Questions

Question 1

How well do Earth System Models (ESMs) simulate the observed distribution of anthropogenic carbon in atmosphere, ocean, and land reservoirs?

Research Questions

Question 1

How well do Earth System Models (ESMs) simulate the observed distribution of anthropogenic carbon in atmosphere, ocean, and land reservoirs?

Question 2

Can contemporary atmospheric CO_2 observations be used to constrain future CO_2 projections?

Research Questions

Question 1

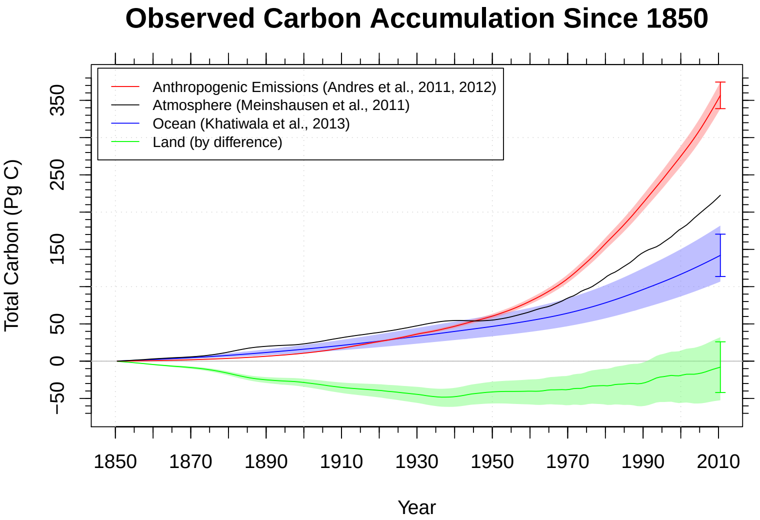
How well do Earth System Models (ESMs) simulate the observed distribution of anthropogenic carbon in atmosphere, ocean, and land reservoirs?

Question 2

Can contemporary atmospheric CO_2 observations be used to constrain future CO_2 projections?

Community Model Benchmarking

Systematic assessment of model fidelity, employing best-available observational data, can identify model weaknesses and inspire new measurements.



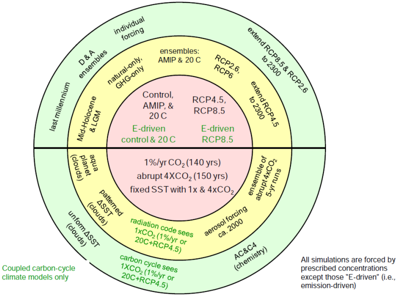
Observational estimates of anthropogenic carbon emissions (excluding land use change) and accumulation in atmosphere, ocean, and land reservoirs for 1850–2010. Atmosphere carbon is a fusion of Law Dome ice core CO_2 observations, the Keeling Mauna Loa record, and more recently the NOAA GMD global surface average, integrated for the purpose of forcing IPCC models. Total land flux is computed by mass balance as follows:

$$\Delta C_L = \sum_i F_i - \Delta C_A - \Delta C_O.$$

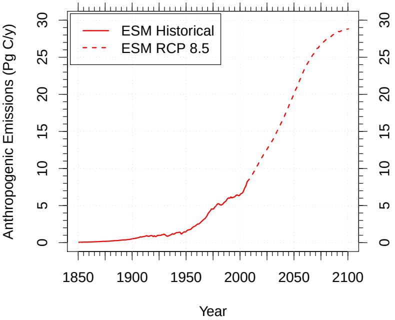
15 fully-prognostic ESMs that performed CMIP5 emissions-forced simulations

Model	Modeling Center
BCC-CSM1.1	Beijing Climate Center, China Meteorological Administration, CHINA
BCC-CSM1.1(m)	Beijing Climate Center, China Meteorological Administration, CHINA
BNU-ESM	Beijing Normal University, CHINA
CanESM2	Canadian Centre for Climate Modelling and Analysis, CANADA
CESM1-BGC	Community Earth System Model Contributors, NSF-DOE-NCAR, USA
FGOALS-s2.0	LASG, Institute of Atmospheric Physics, CAS, CHINA
GFDL-ESM2g	NOAA Geophysical Fluid Dynamics Laboratory, USA
GFDL-ESM2m	NOAA Geophysical Fluid Dynamics Laboratory, USA
HadGEM2-ES	Met Office Hadley Centre, UNITED KINGDOM
INM-CM4	Institute for Numerical Mathematics, RUSSIA
IPSL-CM5A-LR	Institut Pierre-Simon Laplace, FRANCE
MIROC-ESM	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Institute (University of Tokyo), and National Institute for Environmental Studies, JAPAN
MPI-ESM-LR	Max Planck Institute for Meteorology, GERMANY
MRI-ESM1	Meteorological Research Institute, JAPAN
NorESM1-ME	Norwegian Climate Centre, NORWAY

CMIP5 Long-Term Experiments

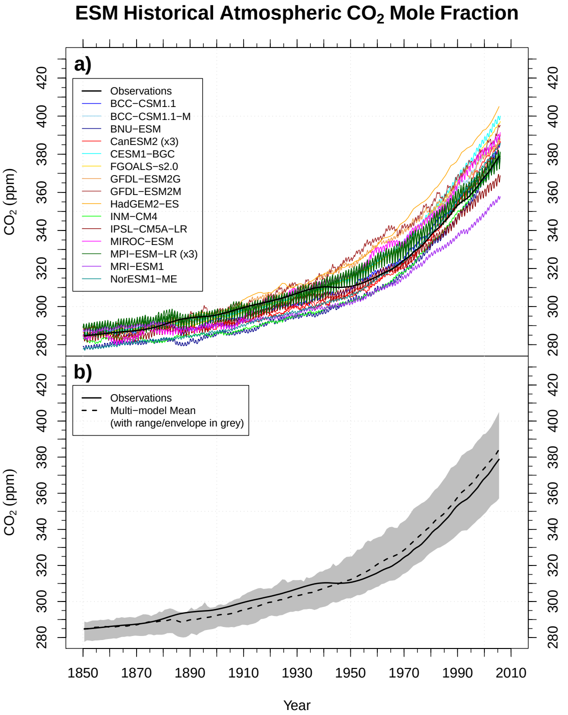


Emissions for Historical + RCP 8.5 Simulations



(a) Most ESMs exhibited a high bias in predicted atmospheric CO_2 mole fraction, which ranged from 357–405 ppm at the end of the historical period (1850–2005).

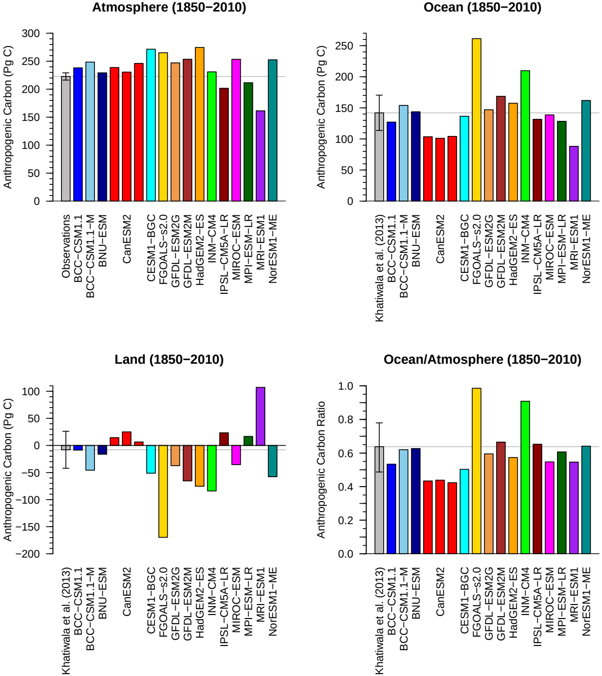
(b) The multi-model mean was biased high from 1946 throughout the 20th century, ending 5.6 ppm above the observed value of 378.8 ppm in 2005.



Model inventory comparison with Khatiwala et al. (2013)

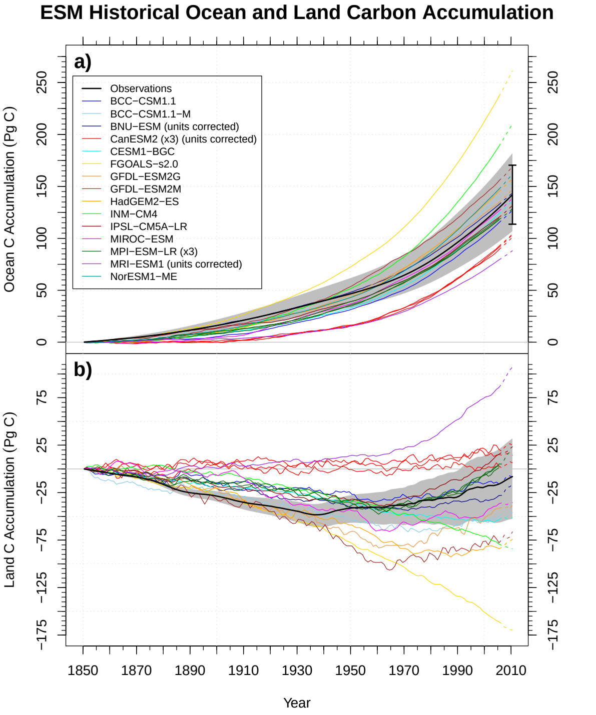
Once normalized by their atmospheric carbon inventories, most ESMs exhibited a low bias in anthropogenic ocean carbon accumulation through 2010.

The same pattern holds for the Sabine et al. (2004) inventory derived using the ΔC^* separation technique.



(a) Ocean inventory estimates had a fairly persistent ordering during the second half of the 20th century.

(b) ESMs exhibited a wide range of land carbon accumulation responses to increasing CO_2 and land use change, ranging from a net source of 170 Pg C to a sink of 107 Pg C in 2010.



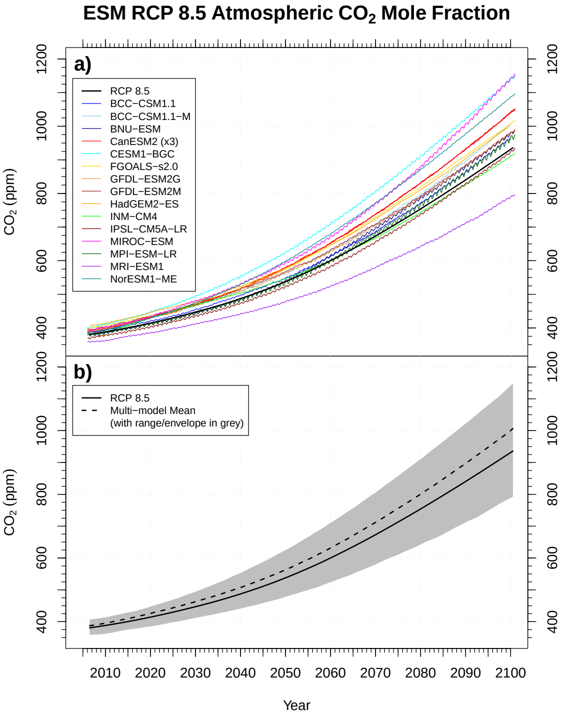
Question 1

How well do Earth System Models (ESMs) simulate the observed distribution of anthropogenic carbon in atmosphere, ocean, and land reservoirs?

- Most ESMs exhibited a high bias in predicted atmospheric CO_2 mole fraction, ranging from 357–405 ppm in 2005.
- The multi-model mean atmospheric CO_2 mole fraction was biased high from 1946 onward, ending 5.6 ppm above observations in 2005.
- Once normalized by atmospheric carbon accumulation, most ESMs exhibited a low bias in ocean accumulation in 2010.
- ESMs predicted a wide range of land carbon accumulation in response to increasing CO_2 and land use change, ranging from -170–107 Pg C in 2010.

Question 2

Can contemporary atmospheric CO_2 observations be used to constrain future CO_2 projections?



Reducing Uncertainties Using Observations

To reduce feedback uncertainties using contemporary observations,

1. there must be a relationship between contemporary variability and future trends on longer time scales within the model, and

Reducing Uncertainties Using Observations

To reduce feedback uncertainties using contemporary observations,

1. there must be a relationship between contemporary variability and future trends on longer time scales within the model, and
2. it must be possible to constrain contemporary variability in the model using observations.

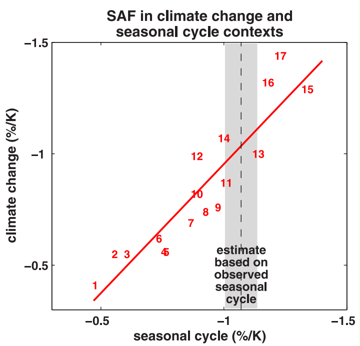
Reducing Uncertainties Using Observations

To reduce feedback uncertainties using contemporary observations,

1. there must be a relationship between contemporary variability and future trends on longer time scales within the model, and
2. it must be possible to constrain contemporary variability in the model using observations.

Example #1

Hall and Qu (2006) evaluated the strength of the springtime snow albedo feedback (SAF; $\Delta\alpha_s/\Delta T_s$) from 17 models used for the IPCC AR4 and compared them with the observed springtime SAF from ISCCP and ERA-40 reanalysis.



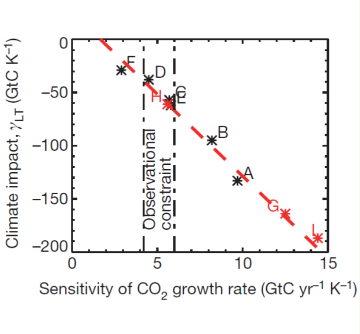
Reducing Uncertainties Using Observations

To reduce feedback uncertainties using contemporary observations,

1. there must be a relationship between contemporary variability and future trends on longer time scales within the model, and
2. it must be possible to constrain contemporary variability in the model using observations.

Example #2

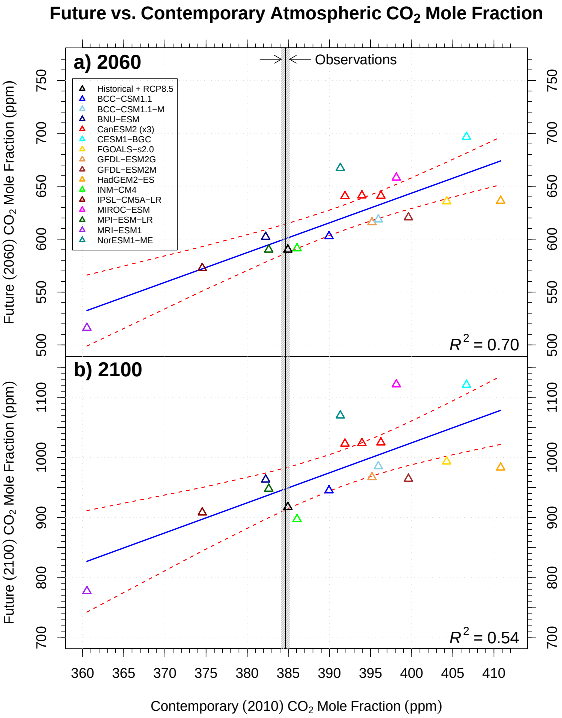
Cox et al. (2013) used the observed relationship between the CO_2 growth rate and tropical temperature as a constraint to reduce uncertainty in the land carbon storage sensitivity to climate change (γ_L) in the tropics using C^4MIP models.



I developed a new emergent constraint from carbon inventories.

A relationship exists between contemporary and future atmospheric CO_2 levels over decadal time scales because carbon model biases persist over decadal time scales.

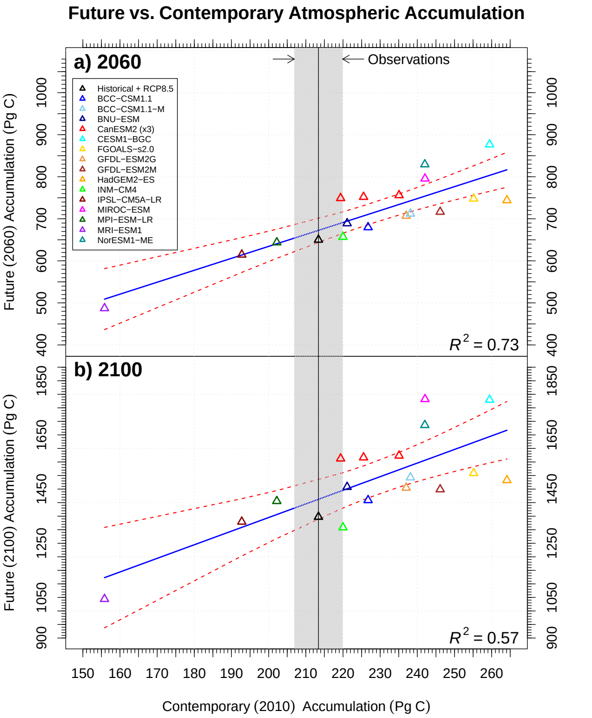
Observed contemporary atmospheric CO_2 mole fraction is represented by the vertical line at 384.6 ± 0.5 ppm.

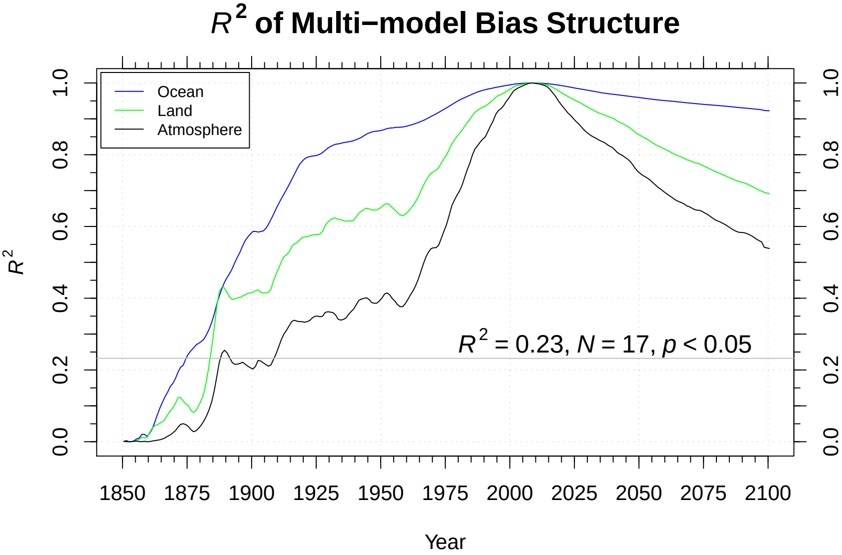


Removing pre-industrial CO_2 mole fraction biases from models, we found the relationship held, confirming the robustness of our result.

Observed contemporary anthropogenic atmospheric carbon inventory is represented by the vertical line at 213.4 ± 6.5 Pg C, which incorporates 1850 CO_2 mole fraction uncertainties.

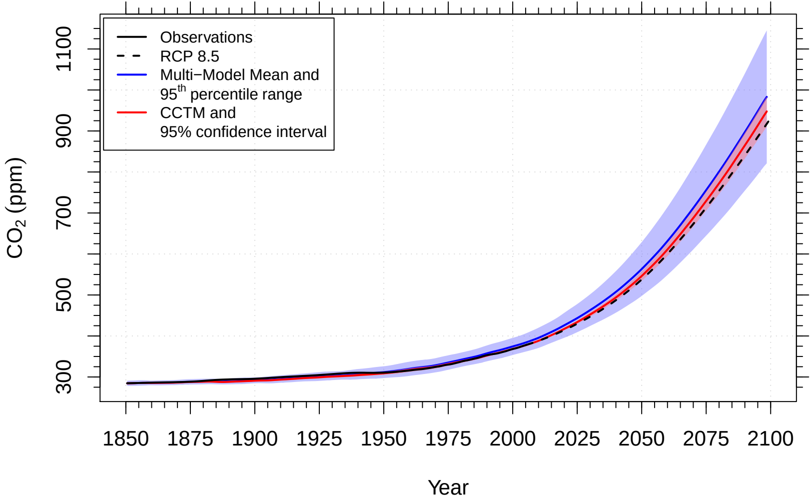
Adding uncertainties from fossil fuel emissions increased the uncertainty to ± 12.7 Pg C.





The coefficients of determination (R^2) for the multi-model bias structure relative to the set of CMIP5 model atmospheric CO_2 mole fractions (black), and oceanic (blue) and land (green) anthropogenic carbon inventories in 2010. Atmospheric CO_2 mole fractions are statistically significant for 1910–2100. Bias persistence was highest for the ocean, followed by land, and then by the atmosphere.

Contemporary CO_2 Tuned Model (CCTM)

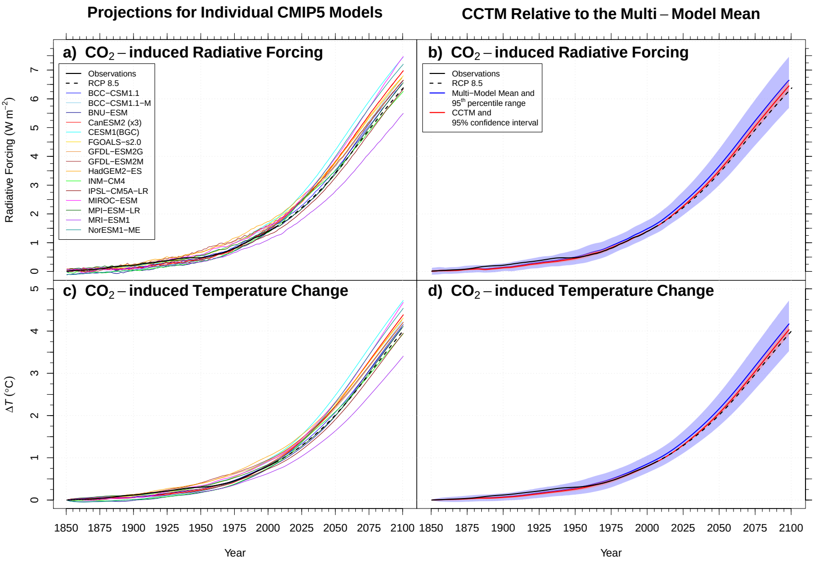


I used this regression to create a contemporary CO_2 tuned model (CCTM) estimate of the atmospheric CO_2 trajectory for the 21^{st} century.

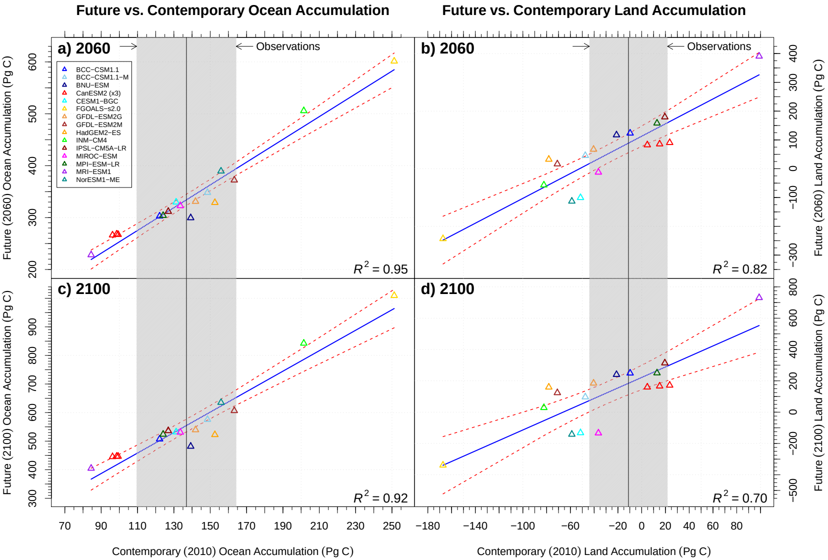
Best estimate developed using Mauna Loa CO_2 data:

At 2060: 600 ± 14 ppm, 21 ppm below the multi-model mean

At 2100: 947 ± 35 ppm, 32 ppm below the multi-model mean



I calculated the CO_2 radiative forcing and used an impulse response function (tuned to the mean transient climate response of CMIP5 models) to equitably compute the resulting CO_2-induced temperature change (ΔT_{CO_2}) for models and the CCTM. The CO_2 biases for individual models contributed to ΔT_{CO_2} biases of $-0.7°C$ to $+0.6°C$ by 2100, relative to the CCTM estimate.



I also developed a multi-model constraint on the evolution of ocean and land anthropogenic inventories. Since observational uncertainties are higher for ocean and land, uncertainties in future estimates cannot be reduced as much as for atmospheric CO_2.

Question 2

Can we use contemporary CO_2 observations to constrain future CO_2 projections?

- Yes.
- I developed a new emergent constraint from anthropogenic carbon inventories in atmosphere, ocean, and land reservoirs.
- Land and ocean processes contributing to contemporary carbon cycle biases persist over decadal timescales.
- I used the relationship between contemporary and future atmospheric CO_2 levels to create a contemporary CO_2 tuned model (CCTM) estimate for the 21st century.
  - At 2060: 600 ± 14 ppm, 21 ppm below the multi-model mean.
  - At 2100: 947 ± 35 ppm, 32 ppm below the multi-model mean.
- Uncertainties in future climate predictions may be reduced by improving models to match the long-term time series of CO_2 from Mauna Loa and other monitoring stations.

Implications of CO_2 Biases in ESMs

- Most of the model-to-model variability of CO_2 in the 21st century was traced to biases that existed at the end of the observational record.
- Future fossil fuel emissions targets designed to stabilize CO_2 levels would be too low if estimated from the multi-model mean of ESMs.
- Models could be improved through extensive comparison with observations and community model benchmarking.

AGU PUBLICATIONS
JGR
Journal of Geophysical Research: Biogeosciences

RESEARCH ARTICLE
10.1002/2013JG002381

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

F. M. Hoffman[1,2], J. T. Randerson[1], V. K. Arora[3], Q. Bao[4], P. Cadule[5], D. Ji[6], C. D. Jones[7], M. Kawamiya[8], S. Khatiwala[9], K. Lindsay[10], A. Obata[11], E. Shevliakova[12], K. D. Six[13], J. F. Tjiputra[14], E. M. Volodin[15], and T. Wu[16]

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

Key Points:
- We analyzed emission-driven simulations from 15 Earth System Models (ESMs)
- Most ESMs had a small positive bias in contemporary atmospheric CO_2 predictions
- We used a linear relationship to create a trajectory of future atmospheric CO_2

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

Abstract The strength of feedbacks between a changing climate and future CO_2 concentrations is uncertain and difficult to predict using Earth System Models (ESMs). We analyzed emission-driven simulations—in which atmospheric CO_2 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 CO_2 over the historical period with observations indicated that ESMs, on average, had a small positive bias in predictions of contemporary atmospheric CO_2. 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 CO_2 biases and future CO_2 levels for the multimodel ensemble. We used this relationship to create a contemporary CO_2 tuned model (CCTM) estimate of the atmospheric CO_2 trajectory for the 21st century. The CCTM yielded CO_2 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 CO_2, CO_2-induced radiative forcing, and CO_2-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 CO_2 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 CO_2 from Mauna Loa, our analysis suggests that uncertainties in future climate projections can be reduced.

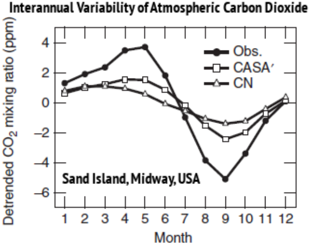
1. Introduction

Anthropogenic emissions of radiatively active greenhouse gases into the atmosphere, especially carbon dioxide (CO_2), 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 CO_2 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

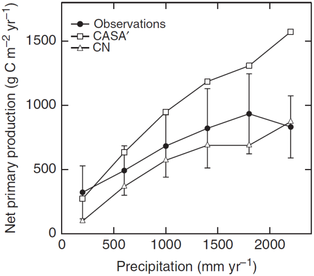
Hoffman, Forrest M., James T. Randerson, Vivek K. Arora, Qing Bao, Patricia Cadule, Duoying Ji, Chris D. Jones, Michio Kawamiya, Samar Khatiwala, Keith Lindsay, Atsushi Obata, Elena Shevliakova, Katharina D. Six, Jerry F. Tjiputra, Evgeny M. Volodin, and Tongwen Wu (2014), Causes and Implications of Persistent Atmospheric Carbon Dioxide Biases in Earth System Models, *J. Geophys. Res. Biogeosci.*, 119(2):141162, doi:10.1002/2013JG002381.

What is a Benchmark?

- A **Benchmark** is a quantitative test of model function achieved through comparison of model results with observational data.
- Acceptable performance on benchmarks **is a necessary but not sufficient condition** for a fully functioning model.
- **Functional benchmarks** offer tests of model responses to forcings and yield insights into ecosystem processes.
- Effective benchmarks must draw upon a broad set of independent observations to evaluate model performance on **multiple temporal and spatial scales**.



Models often fail to capture the amplitude of the seasonal cycle of atmospheric CO_2.



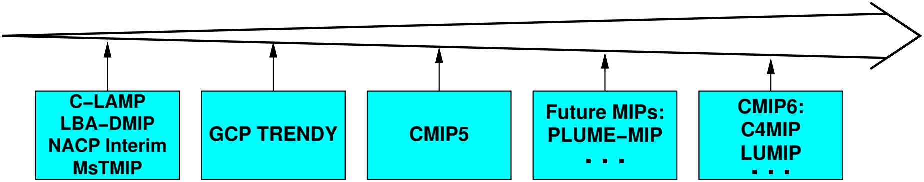
Models may reproduce correct responses over only a limited range of forcing variables.

(Randerson et al., 2009)

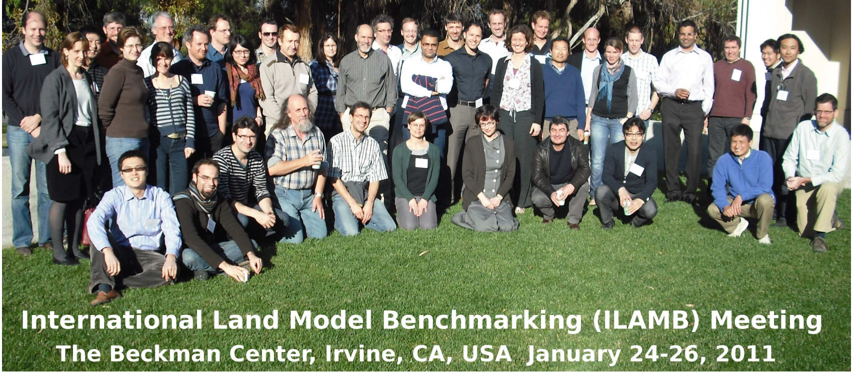
Why Benchmark?

- to demonstrate to the science community and public that the representation of coupled climate and biogeochemical cycles in Earth system models (ESMs) is improving;
- to quantitatively diagnose impacts of model development in related fields on carbon cycle processes;
- to guide synthesis efforts, such as the Intergovernmental Panel on Climate Change (IPCC), in the review of mechanisms of global change in models that are broadly consistent with available contemporary observations;
- to increase scrutiny of key datasets used for model evaluation;
- to identify gaps in existing observations needed for model validation;
- to accelerate incorporation of new measurements for rapid and widespread use in model assessment;
- to provide a quantitative, application-specific set of minimum criteria for participation in model intercomparison projects (MIPs).

An Open Source Benchmarking Software System

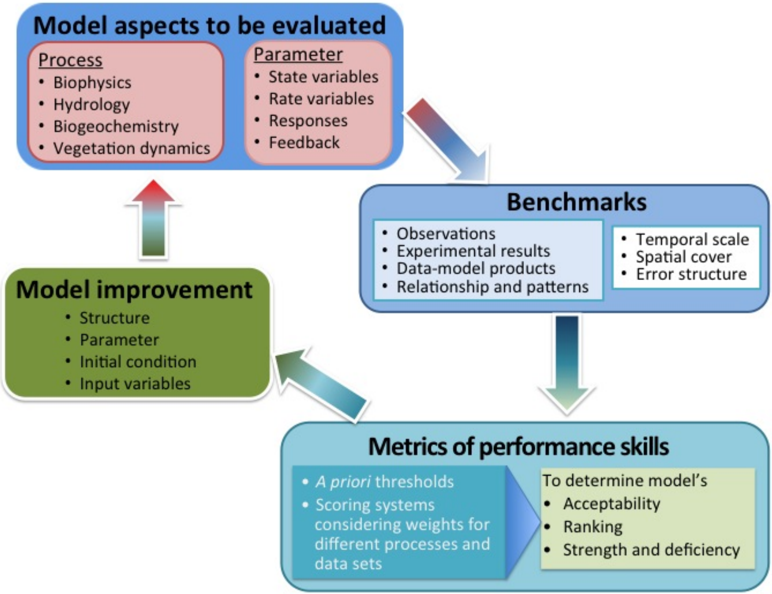


- Human capital costs of making rigorous model-data comparisons is considerable and constrains the scope of individual MIPs.
- Many MIPs spend resources "reinventing the wheel" in terms of variable naming conventions, model simulation protocols, and analysis software.
- **Need for ILAMB:** Each new MIP has access to the model-data comparison modules from past MIPs through ILAMB (*e.g.*, MIPs use one common modular software system). Standardized international naming conventions also increase MIP efficiency.



- We co-organized inaugural meeting and ∼45 researchers participated from the United States, Canada, the United Kingdom, the Netherlands, France, Germany, Switzerland, China, Japan, and Australia.
- **ILAMB Goals:** Develop internationally accepted benchmarks for model performance, advocate for design of open-source software system, and strengthen linkages between experimental, monitoring, remote sensing, and climate modeling communities. *Initial focus on CMIP5 models.*
- Provides methodology for model–data comparison and baseline standard for performance of land model process representations (Luo et al., 2012).

General Benchmarking Procedure



(Luo et al., 2012)

Example Benchmark Score Sheet from C-LAMP

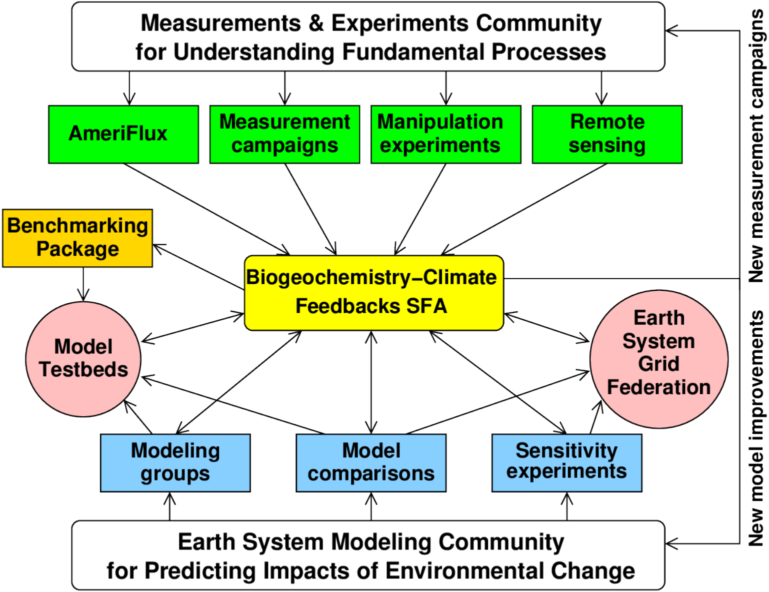
Models →

BGC Datasets ↓

Metric	Metric components	Uncertainty of obs.	Scaling mismatch	Total score	Sub-score	CASA′	CN
LAI	Matching MODIS observations			15.0		13.5	12.0
	• Phase (assessed using the month of maximum LAI)	Low	Low		6.0	5.1	4.2
	• Maximum (derived separately for major biome classes)	Moderate	Low		5.0	4.6	4.3
	• Mean (derived separately for major biome classes)	Moderate	Low		4.0	3.8	3.5
NPP	Comparisons with field observations and satellite products			10.0		8.0	8.2
	• Matching EMDI Net Primary Production observations	High	High		2.0	1.5	1.6
	• EMDI comparison, normalized by precipitation	Moderate	Moderate		4.0	3.0	3.4
	• Correlation with MODIS (r^2)	High	Low		2.0	1.6	1.4
	• Latitudinal profile comparison with MODIS (r^2)	High	Low		2.0	1.9	1.8
CO_2 annual cycle	Matching phase and amplitude at Globalview flash sites			15.0		10.4	7.7
	• 60°–90°N	Low	Low		6.0	4.1	2.8
	• 30°–60°N	Low	Low		6.0	4.2	3.2
	• 0°–30°N	Moderate	Low		3.0	2.1	1.7
Energy & CO_2 fluxes	Matching eddy covariance monthly mean observations			30.0		17.2	16.6
	• Net ecosystem exchange	Low	High		6.0	2.5	2.1
	• Gross primary production	Moderate	Moderate		6.0	3.4	3.5
	• Latent heat	Low	Moderate		9.0	6.4	6.4
	• Sensible heat	Low	Moderate		9.0	4.9	4.6
Transient dynamics	Evaluating model processes that regulate carbon exchange on decadal to century timescales			30.0		16.8	13.8
	• Aboveground live biomass within the Amazon Basin	Moderate	Moderate		10.0	5.3	5.0
	• Sensitivity of NPP to elevated levels of CO_2: comparison to temperate forest FACE sites	Low	Moderate		10.0	7.9	4.1
	• Interannual variability of global carbon fluxes: comparison with TRANSCOM	High	Low		5.0	3.6	3.0
	• Regional and global fire emissions: comparison to GFEDv2	High	Low		5.0	0.0	1.7
			Total:	**100.0**		**65.9**	**58.3**

(Randerson et al., 2009)

Biogeochemistry–Climate Feedbacks Scientific Focus Area



ILAMB Prototype Diagnostics System

An initial ILAMB prototype has been developed by Mingquan Mu at UCI.

- Current variables:

  Aboveground live biomass (North America FIA, tropical Saatchi et al.), Burned area (GFED3), CO_2 (NOAA GMD, Mauna Loa), Global net land flux (GCP), Gross primary production (Fluxnet-MTE), Leaf area index (AVHRR, MODIS), Net ecosystem exchange (Fluxnet), Respiration (Fluxnet), Soil C (HWSD, NCSCDv2), Evapotranspiration (LandFlux, GLEAM, MODIS), Latent heat (Fluxnet-MTE), Soil moisture (ESA), Terrestrial water storage change (GRACE), Precipitation (GPCP2), Albedo (MODIS, CERES), Surface up/down SW/LW radiation (CERES, WRMC.BSRN), Sensible heat (Fluxnet), Surface air temperature (CRU).

- Graphics and scoring systems:
  - Annual mean, Bias, RMSE, seasonal cycle, spatial distribution, interannual coeff. of variation and variability, long-term trend scores
  - Global maps, variable to variable, and time series comparisons

- Software:

  Freely distributed, designed to be user friendly and to enable easy addition of new variables

(Mu, Hoffman, Riley, Koven, Lawrence, Randerson)

ILAMB Prototype Layout: Global Variables

Global Variables (Info)

	MeanModel	bcc-csm1-1-m	BNU-ESM	CanESM2	CESM1-BGC	GFDL-ESM2G	HadGEM2
Aboveground Live Biomass	0.88	-	0.14	0.81	0.68	0.81	0.86
Burned Area	0.41	-	-	-	0.37	-	-
Carbon Dioxide	0.88	-	0.53	0.94	0.86	0.96	-
Global Net Land Flux	0.25	-	0.25	0.32	0.32	0.49	0.63
Gross Primary Production	0.80	0.74	0.74	0.74	0.77	0.72	0.75
Leaf Area Index	0.59	0.64	0.30	0.78	0.53	0.33	0.53
Net Ecosystem Exchange	0.36	0.29	0.19	0.16	0.28	0.64	0.28
Ecosystem Respiration	0.78	0.71	0.78	0.75	0.74	0.70	0.77
Soil Carbon	0.71	-	0.35	0.73	0.31	0.74	0.63
Summary	0.63	0.59	0.41	0.65	0.54	0.67	0.64
Evapotranspiration	0.75	0.83	0.74	0.82	0.73	0.76	0.77
Latent Heat	0.77	0.79	0.71	0.80	0.71	0.72	0.71
Soil Moisture	0.18	0.17	0.20	0.21	0.19	0.18	0.21
Terrestrial Water Storage Change	0.25	0.29	0.25	0.26	0.25	0.24	0.25
Precipitation	0.82	0.83	0.82	0.82	0.86	0.86	0.90
Summary	0.55	0.58	0.54	0.58	0.55	0.55	0.57
Albedo	0.76	0.74	0.75	0.77	0.80	0.76	0.79

ILAMB Prototype Layout: Variable to Variable

Variable to Variable Relationships

	Relationship	Benchmark	MeanModel	bcc-csm1-1-m	BNU-ESM	CanESM2	CESM1-BGC	GFDL-
Precipitation vs. Burned Area	function_bar	1	0.57	-	-	-	0.57	
Precipitation vs. Gross Primary Production	function_bar	1	0.91	0.92	0.93	0.50	0.93	0.
Surface Air Temperature vs. Burned Area	function_bar	1	0.00	-	-	-	0.08	
Surface Air Temperature vs. Gross Primary Production	function_bar	1	0.62	0.50	0.45	0.92	0.50	0.
Surface Downward SW Radiation vs. Gross Primary Production	function_bar	1	0.85	0.74	0.91	0.64	0.81	0.
Surface Net SW Radiation vs. Gross Primary Production	function_bar	1	0.72	0.72	0.72	0.86	0.79	0.
Overall			0.61	0.72	0.75	0.73	0.61	0.

ILAMB Prototype Layout: Time Series

Time Series Comparisons

	Benchmark
Burned Area	GFED3 [Giglio et al. (2010)]
Carbon Dioxide	NOAA.GMD [Dlugokencky et al. (2013)]
Gross Primary Production	FLUXNET [Lasslop et al. (2010)]
Net Ecosystem Exchange	FLUXNET [Lasslop et al. (2010)]
Ecosystem Respiration	FLUXNET [Lasslop et al. (2010)]
Surface Downward SW Radiation	WRMC.BSRN [Konig-Langl et al. (2013)]
Surface Upward SW Radiation	WRMC.BSRN [Konig-Langl et al. (2013)]
Surface Net SW Radiation	WRMC.BSRN [Konig-Langl et al. (2013)]
Surface Downward LW Radiation	WRMC.BSRN [Konig-Langl et al. (2013)]
Surface Upward LW Radiation	WRMC.BSRN [Konig-Langl et al. (2013)]
Surface Net LW Radiation	WRMC.BSRN [Konig-Langl et al. (2013)]

ILAMB Live Demo

Community Involvement Is Key to Success!

- Our international collaboration has made significant progress on development of metrics in the ILAMB prototype.
- Our **BGC-Feedbacks Project** is developing new model–data analysis studies for terrestrial and now marine biogeochemistry (see http://www.bgc-feedbacks.org/).
- We have proposed an **ILAMB Town Hall** at the upcoming American Geophysical Union (AGU) Fall Meeting in December.
- We are planning another community-wide meeting on model metrics and diagnostics in Washington, DC, USA in spring 2016.

International Land Model Benchmarking (ILAMB) Project
http://www.ilamb.org/

Contact: Forrest M. Hoffman <forrest@climatemodeling.org>

Acknowledgments

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 Oak Ridge Leadership Computing Facility (OLCF) at Oak Ridge National Laboratory (ORNL), which is managed by UT-Battelle, LLC, for the U. S. Department of Energy under Contract No. DE-AC05-00OR22725.

I wish to acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and thank the climate modeling groups 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

R. J. Andres, J. S. Gregg, L. Losey, G. Marland, and T. A. Boden. Monthly, global emissions of carbon dioxide from fossil fuel consumption. *Tellus B*, 63(3):309–327, July 2011. doi:10.1111/j.1600-0889.2011.00530.x.

R. J. Andres, T. A. Boden, F.-M. Bréon, P. Ciais, S. Davis, D. Erickson, J. S. Gregg, A. Jacobson, G. Marland, J. Miller, T. Oda, J. G. J. Olivier, M. R. Raupach, P. Rayner, and K. Treanton. A synthesis of carbon dioxide emissions from fossil-fuel combustion. *Biogeosci.*, 9(5):1845–1871, May 2012. doi:10.5194/bg-9-1845-2012.

P. M. Cox, D. Pearson, B. B. Booth, P. Friedlingstein, C. Huntingford, C. D. Jones, and C. M. Luke. Sensitivity of tropical carbon to climate change constrained by carbon dioxide variability. *Nature*, 494(7437):341–344, Feb. 2013. doi:10.1038/nature11882.

A. Hall and X. Qu. Using the current seasonal cycle to constrain snow albedo feedback in future climate change. *Geophys. Res. Lett.*, 33(3):L03502, Feb. 2006. doi:10.1029/2005GL025127.

F. M. Hoffman, 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. Causes and implications of persistent atmospheric carbon dioxide biases in Earth System Models. *J. Geophys. Res. Biogeosci.*, 119(2):141–162, Feb. 2014. doi:10.1002/2013JG002381.

S. Khatiwala, T. Tanhua, S. Mikaloff Fletcher, M. Gerber, S. C. Doney, H. D. Graven, N. Gruber, G. A. McKinley, A. Murata, A. F. Ríos, and C. L. Sabine. Global ocean storage of anthropogenic carbon. *Biogeosci.*, 10(4):2169–2191, Apr. 2013. doi:10.5194/bg-10-2169-2013.

Y. Q. Luo, J. T. Randerson, G. Abramowitz, C. Bacour, E. Blyth, N. Carvalhais, P. Ciais, D. Dalmonech, J. B. Fisher, R. Fisher, P. Friedlingstein, K. Hibbard, F. Hoffman, D. Huntzinger, C. D. Jones, C. Koven, D. Lawrence, D. J. Li, M. Mahecha, S. L. Niu, R. Norby, S. L. Piao, X. Qi, P. Peylin, I. C. Prentice, W. Riley, M. Reichstein, C. Schwalm, Y. P. Wang, J. Y. Xia, S. Zaehle, and X. H. Zhou. A framework for benchmarking land models. *Biogeosci.*, 9(10):3857–3874, Oct. 2012. doi:10.5194/bg-9-3857-2012.

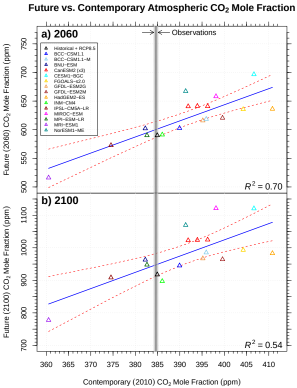
M. Meinshausen, S. Smith, K. Calvin, J. Daniel, M. Kainuma, J.-F. Lamarque, K. Matsumoto, S. Montzka, S. Raper, K. Riahi, A. Thomson, G. Velders, and D. P. van Vuuren. The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. *Clim. Change*, 109(1):213–241, Nov. 2011. doi:10.1007/s10584-011-0156-z.

J. T. Randerson, F. M. Hoffman, P. E. Thornton, N. M. Mahowald, K. Lindsay, Y.-H. Lee, C. D. Nevison, S. C. Doney, G. Bonan, R. Stöckli, C. Covey, S. W. Running, and I. Y. Fung. Systematic assessment of terrestrial biogeochemistry in coupled climate-carbon models. *Global Change Biol.*, 15(9):2462–2484, Sept. 2009. doi:10.1111/j.1365-2486.2009.01912.x.

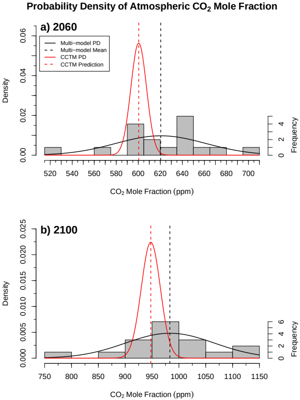
C. L. Sabine, R. A. Feely, N. Gruber, R. M. Key, K. Lee, J. L. Bullister, R. Wanninkhof, C. S. Wong, D. W. R. Wallace, B. Tilbrook, F. J. Millero, T.-H. Peng, A. Kozyr, T. Ono, and A. F. Rios. The oceanic sink for anthropogenic CO_2. *Science*, 305(5682):367–371, July 2004. doi:10.1126/science.1097403.

Emergent Constraint Developed from CMIP5 ESMs

An emergent constraint based on carbon inventories was applied to future atmospheric CO_2 projections from CMIP5 ESMs.



- Much of the model-to-model variation in projected CO_2 during the 21st century is tied to biases that existed during observational era.
- Model differences in the representation of concetration–carbon feedbacks and other slowly changing carbon cycle processes appear to be the primary driver of this variability.
- Range of temperature increases at 2100 slightly reduced, from $5.1 \pm 2.2°C$ for the full ensemble, to $5.0 \pm 1.9°C$ after applying the emergent constraint.

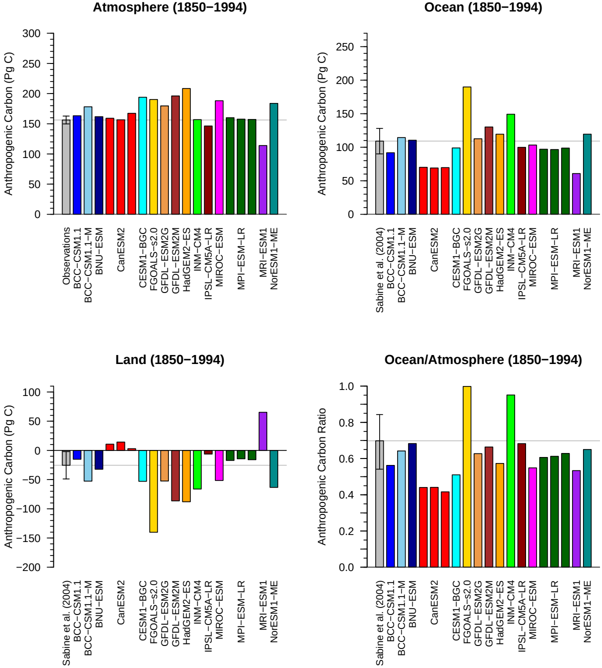


Best estimate using Mauna Loa CO_2

At 2060: 600 ± 14 ppm, 21 ppm below the multi-model mean
At 2100: 947 ± 35 ppm, 32 ppm below the multi-model mean

Hoffman, Forrest M., James T. Randerson, Vivek K. Arora, Qing Bao, Patricia Cadule, Duoying Ji, Chris D. Jones, Michio Kawamiya, Samar Khatiwala, Keith Lindsay, Atsushi Obata, Elena Shevliakova, Katharina D. Six, Jerry F. Tjiputra, Evgeny M. Volodin, and Tongwen Wu. February 2014. "Causes and Implications of Persistent Atmospheric Carbon Dioxide Biases in Earth System Models." *J. Geophys. Res. Biogeosci.*, 119(2):141–162. doi:10.1002/2013JG002381. *Most downloaded JGR-B paper for February 2014.*

Model inventory comparison with Sabine et al. (2004)



Implications for CO_2, Radiative Forcing, and Temperature

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
Multi-model 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