

Causes and Implications of Persistent Atmospheric Carbon Dioxide Biases in Earth System Models

Forrest M. Hoffman

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

Environmental Science Division (EVS) Seminar
Argonne National Laboratory (ANL), Argonne, Illinois, USA
August 4, 2016

CLIMATE CHANGE
SCIENCE INSTITUTE

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 CO₂ observations be used to constrain future CO₂ 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?

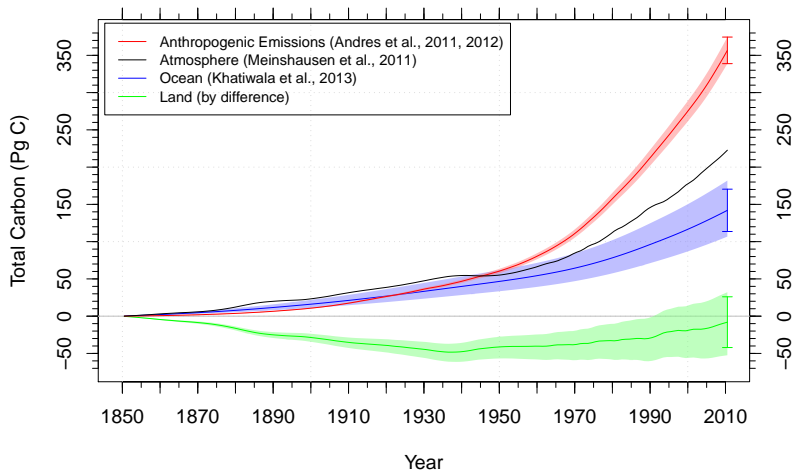
Question 2

Can contemporary atmospheric CO₂ observations be used to constrain future CO₂ projections?

Community Model Benchmarking

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

Observed Carbon Accumulation Since 1850



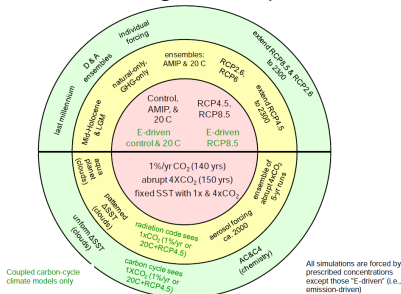
Observational estimates of anthropogenic carbon emissions (excluding land use change) and accumulation in atmosphere, ocean, and land reservoirs for 1850–2010. Atmosphere carbon is a fusion of Law Dome ice core CO_2 observations, the Keeling Mauna Loa record, and more recently the NOAA GMD global surface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

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

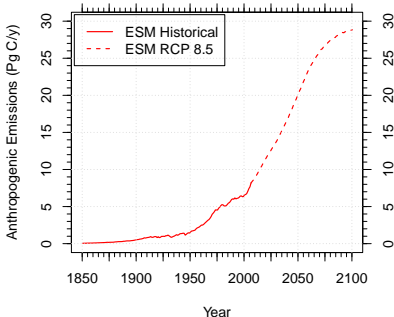
15 fully-prognostic ESMs that performed CMIP5 emissions-forced simulations

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



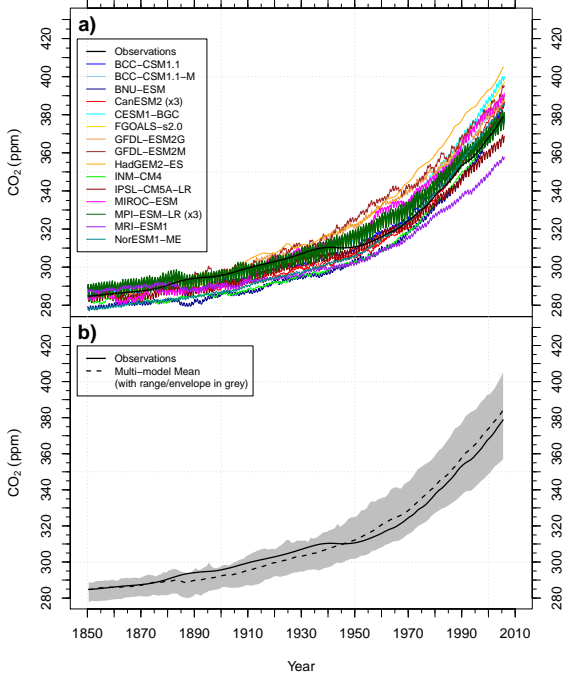
Emissions for Historical + RCP 8.5 Simulations



ESM Historical Atmospheric CO₂ Mole Fraction

(a) Most ESMs exhibited a high bias in predicted atmospheric CO₂ 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.

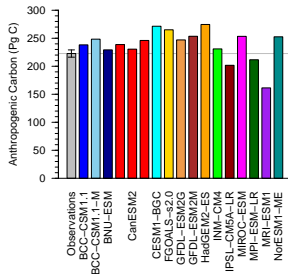


Model inventory comparison with Khatiwala et al. (2013)

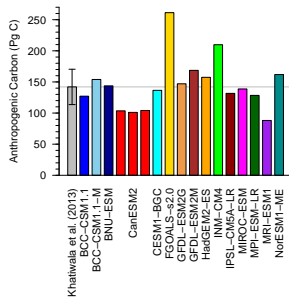
Once normalized by their atmospheric carbon inventories, most ESMs exhibited a low bias in anthropogenic ocean carbon accumulation through 2010.

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

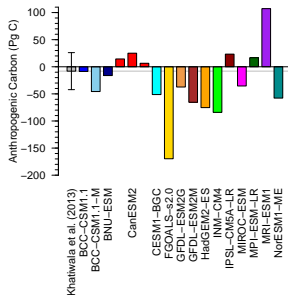
Atmosphere (1850–2010)



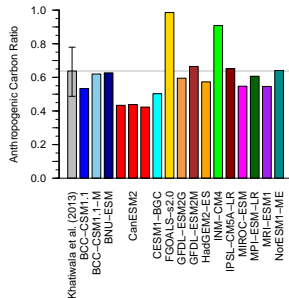
Ocean (1850–2010)



Land (1850–2010)



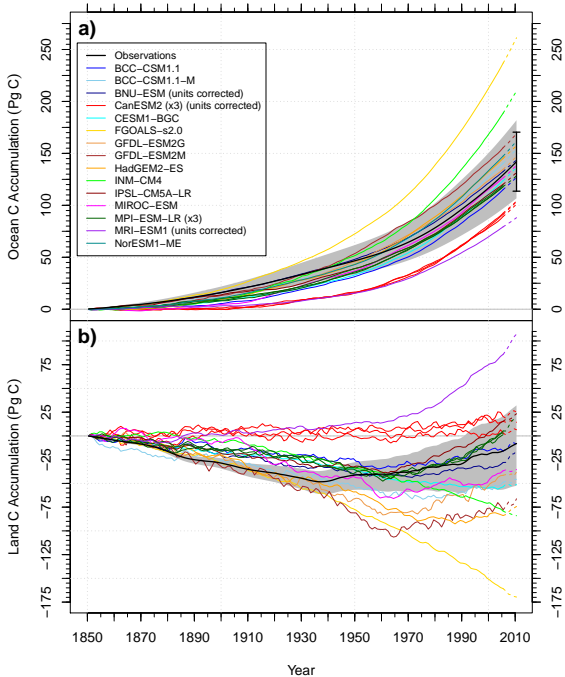
Ocean/Atmosphere (1850–2010)



ESM Historical Ocean and Land Carbon Accumulation

(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 CO₂ and land use change, ranging from a net source of 170 Pg C to a sink of 107 Pg C in 2010.



Question 1

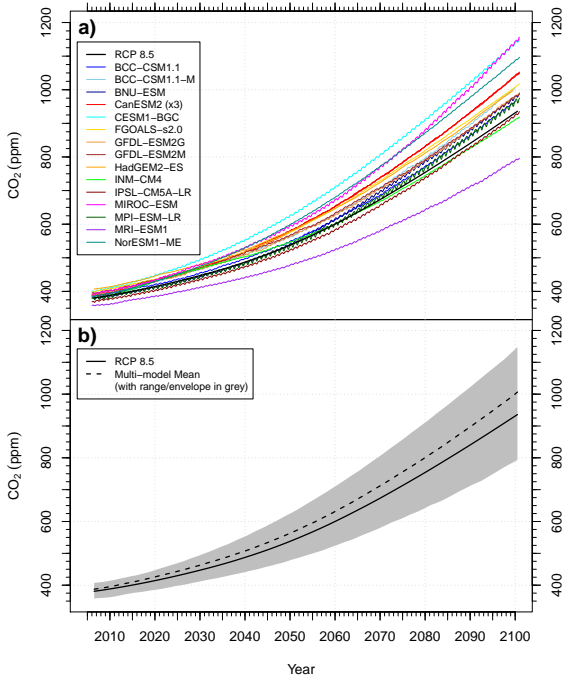
How well do Earth System Models (ESMs) simulate the observed distribution of anthropogenic carbon in atmosphere, ocean, and land reservoirs?

- ▶ Most ESMs exhibited a high bias in predicted atmospheric CO₂ mole fraction, ranging from 357–405 ppm in 2005.
- ▶ The multi-model mean atmospheric CO₂ 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 CO₂ and land use change, ranging from –170–107 Pg C in 2010.

ESM RCP 8.5 Atmospheric CO₂ Mole Fraction

Question 2

Can contemporary atmospheric CO₂ observations be used to constrain future CO₂ 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.

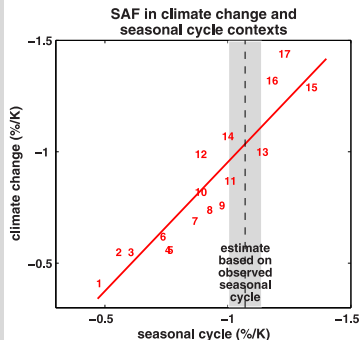
Reducing Uncertainties Using Observations

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.



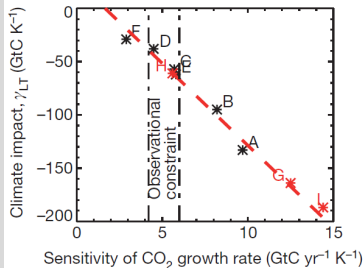
Reducing Uncertainties Using Observations

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 C⁴MIP models.

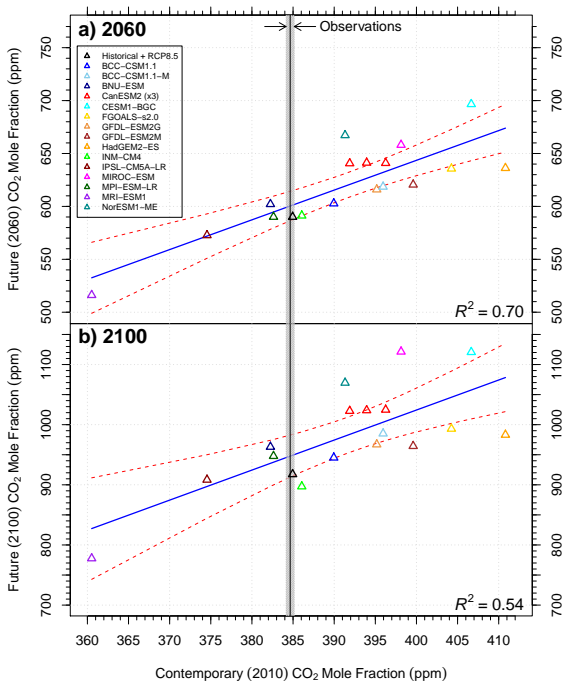


Future vs. Contemporary Atmospheric CO₂ Mole Fraction

I developed a new emergent constraint from carbon inventories.

A relationship exists between contemporary and future atmospheric CO₂ levels over decadal time scales because carbon model biases persist over decadal time scales.

Observed contemporary atmospheric CO₂ mole fraction is represented by the vertical line at 384.6 ± 0.5 ppm.

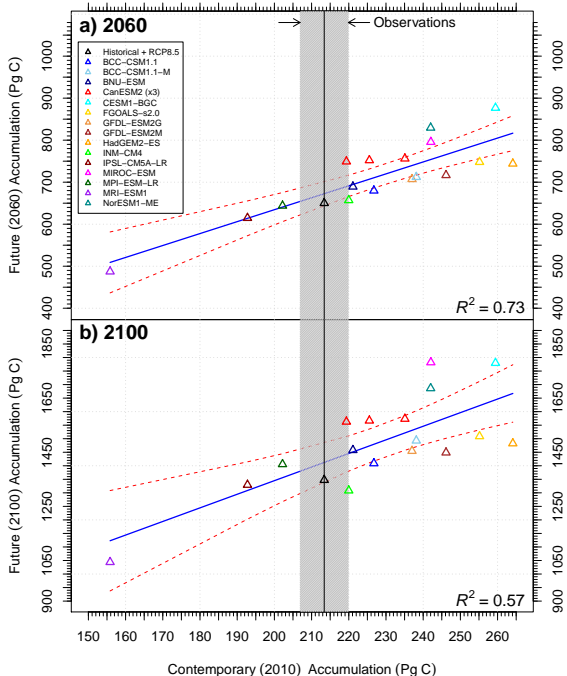


Future vs. Contemporary Atmospheric Accumulation

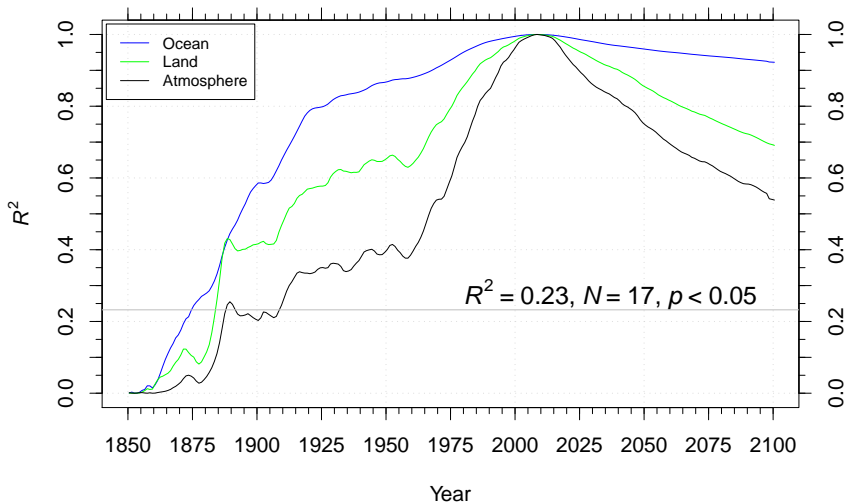
Removing pre-industrial CO₂ mole fraction biases from models, we found the relationship held, confirming the robustness of our result.

Observed contemporary anthropogenic atmospheric carbon inventory is represented by the vertical line at 213.4 ± 6.5 Pg C, which incorporates 1850 CO₂ mole fraction uncertainties.

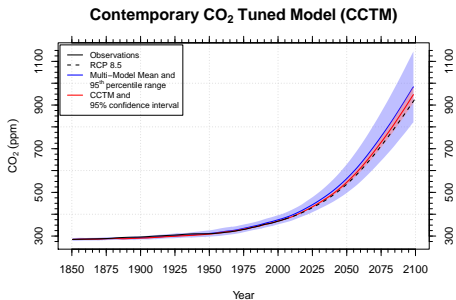
Adding uncertainties from fossil fuel emissions increased the uncertainty to ± 12.7 Pg C.



R^2 of Multi-model Bias Structure



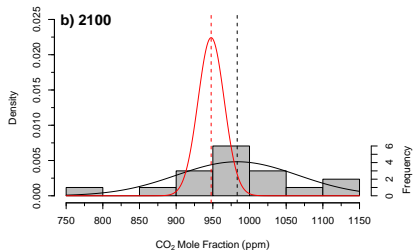
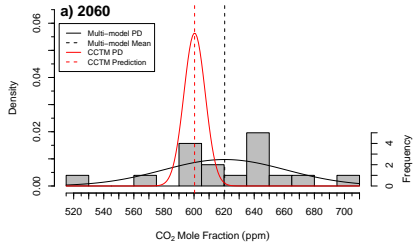
The coefficients of determination (R^2) for the multi-model bias structure relative to the set of CMIP5 model atmospheric CO₂ mole fractions (black), and oceanic (blue) and land (green) anthropogenic carbon inventories in 2010. Atmospheric CO₂ mole fractions are statistically significant for 1910–2100. Bias persistence was highest for the ocean, followed by land, and then by the atmosphere.



I used this regression to create a contemporary CO₂ tuned model (CCTM) estimate of the atmospheric CO₂ trajectory for the 21st century.

- ▶ Peak probability densities of CO₂ mole fraction predictions were lower for the CCTM than the multi-model means.
- ▶ The ranges of uncertainty were smaller by almost a factor of 6 at 2060 and almost a factor of 5 at 2100.

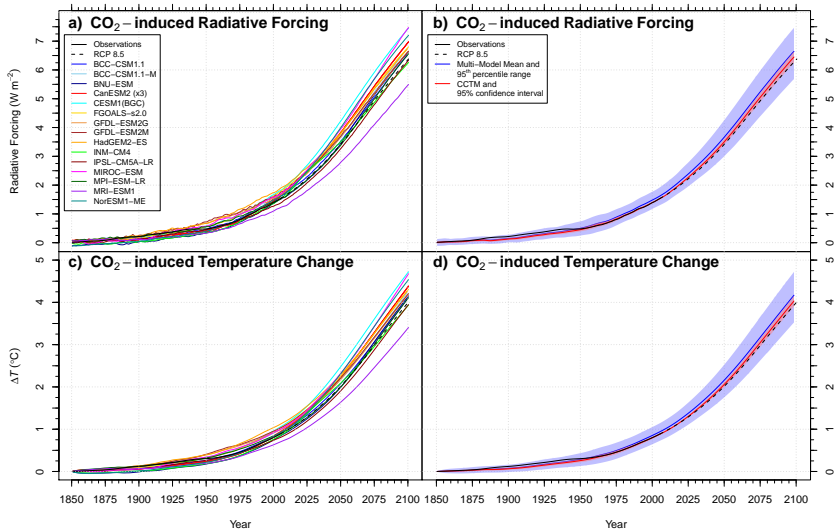
Probability Density of Atmospheric CO₂ Mole Fraction



Best estimate using Mauna Loa CO₂

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

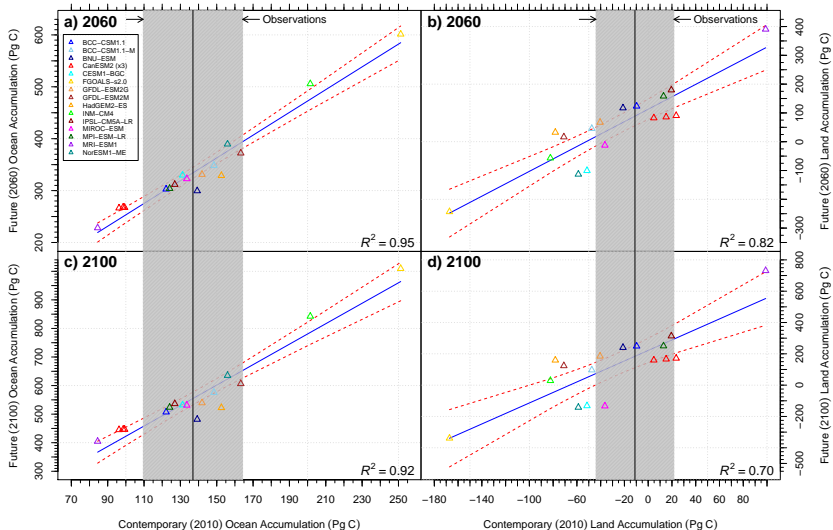
At 2100: 947 ± 35 ppm, 32 ppm below the multi-model mean



I calculated the CO₂ radiative forcing and used an impulse response function (tuned to the mean transient climate response of CMIP5 models) to equitably compute the resulting CO₂-induced temperature change (ΔT_{CO_2}) for models and the CCTM. The CO₂ biases for individual models contributed to ΔT_{CO_2} biases of -0.7°C to $+0.6^{\circ}\text{C}$ by 2100, relative to the CCTM estimate.

Future vs. Contemporary Ocean Accumulation

Future vs. Contemporary Land Accumulation



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 CO₂ observations to constrain future CO₂ 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 CO₂ levels to create a contemporary CO₂ 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 CO₂ from Mauna Loa and other monitoring stations.

Implications of Biases of CO₂ in ESMs

- ▶ Most of the model-to-model variability of CO₂ 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 CO₂ levels would be too low if estimated from the multi-model mean of ESMs.
- ▶ Modes could be improved through **extensive comparison with sustained observations** and **community model benchmarking**.

AGU PUBLICATIONS

Journal of Geophysical Research: Biogeosciences

RESEARCH ARTICLE
18 JULY 2013 10:00Z

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

Key Points
► The largest atmospheric CO₂ bias in Earth System Models (ESMs) is associated with the end of the observational record.
► The bias is associated with the end of the observational record.
► The bias is associated with the end of the observational record.

F. R. Hoffman¹, J. T. Forrester², K. R. Avner³, S. B. Brier⁴, F. Cadet⁵, D. P. C. Jones⁶, W. Rosenfeld⁷, S. Shalunov⁸, S. Loschnig⁹, S. Ghosh¹⁰, J. Sheehovska¹¹, K. C. Munn¹², J. Tjiputra¹³, K. M. Waliser¹⁴, and M. Wu¹⁵

¹Department of Earth System Science, University of California, Irvine, California, USA, ²Climate Change Science Institute and Program for Environmental and Earth System Science, National Institute of Environmental Health and Safety, U.S. Environmental Protection Agency, Research Triangle Park, North Carolina, USA, ³Canadian Centre for Climate Modelling and Analysis, Meteorological Service of Canada, Victoria, British Columbia, Canada, ⁴Centre for Global Change Science, University of Guelph, Guelph, Ontario, Canada, ⁵Centre for Global Change Science, University of Guelph, Guelph, Ontario, Canada, ⁶Centre for Global Change Science, University of Guelph, Guelph, Ontario, Canada, ⁷Centre for Global Change Science, University of Guelph, Guelph, Ontario, Canada, ⁸Centre for Global Change Science, University of Guelph, Guelph, Ontario, Canada, ⁹Centre for Global Change Science, University of Guelph, Guelph, Ontario, Canada, ¹⁰Centre for Global Change Science, University of Guelph, Guelph, Ontario, Canada, ¹¹Centre for Global Change Science, University of Guelph, Guelph, Ontario, Canada, ¹²Centre for Global Change Science, University of Guelph, Guelph, Ontario, Canada, ¹³Centre for Global Change Science, University of Guelph, Guelph, Ontario, Canada, ¹⁴Centre for Global Change Science, University of Guelph, Guelph, Ontario, Canada, ¹⁵Centre for Global Change Science, University of Guelph, Guelph, Ontario, Canada.

Correspondence to: F. R. Hoffman (hoffman@uci.edu)

Received 17 APRIL 2013
Accepted 17 JULY 2013
Published online 18 JULY 2013

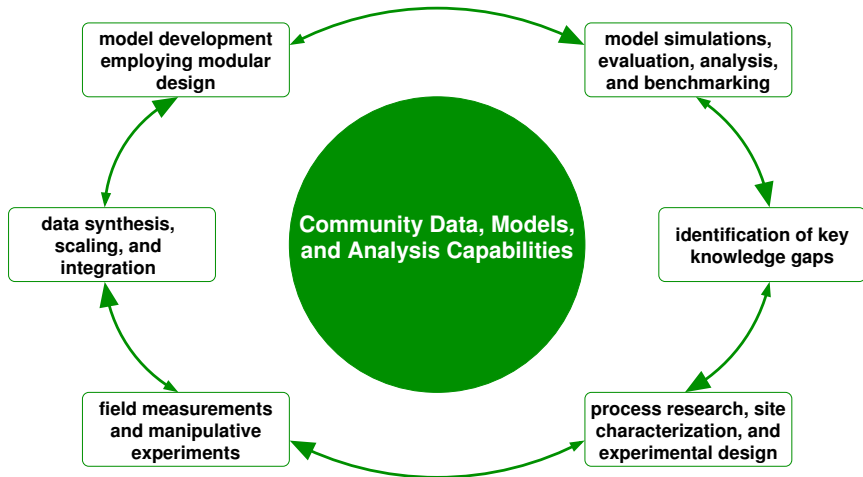
Abstract The strength of feedbacks between a changing climate and future CO₂ concentrations is uncertain and difficult to predict using Earth System Models (ESMs). We analyzed emission-driven simulations in which atmospheric CO₂ levels were constrained approximately to the historical (1950–2000) and future periods Representative Concentration Pathway (RCP) 4.5 for 2006–2100 (produced by 10 ESMs) for the first phase of the Coupled Model Intercomparison Project (CMIP5). Comparison of 100-year-averaged atmospheric CO₂ over the historical period with observations indicated that ESMs, on average, had a small positive bias in predictions of contemporary atmospheric CO₂. Model ocean carbon uptake in the 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 CO₂ levels in the end-of-the-model ensemble. We used this relationship to create a contemporary CO₂ level model (CCM) estimate of the atmospheric CO₂ history for the 21st century. The CCM predicted CO₂ estimates of 460, 540, and 640 ppm in 2006 and 847, 910, and 990 ppm in 2100, which were 21 ppm and 12 ppm below the multimodel mean during these two time periods. Using this surrogate approach, we estimated the likely range of future atmospheric CO₂ for each of the end-of-the-model CO₂ and indirect temperature increases for the RCP 4.5 scenario were considerably narrowed compared to estimates from the full ESM ensemble. Our analysis provided evidence that much of the end-of-the-model variation in projected CO₂ during the 21st century may not be known but instead may be observational error 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 comparing models to more closely match the long-term time series of CO₂ from Mauna Loa, our analysis suggests that uncertainties in future climate projections may be reduced.

1. Introduction

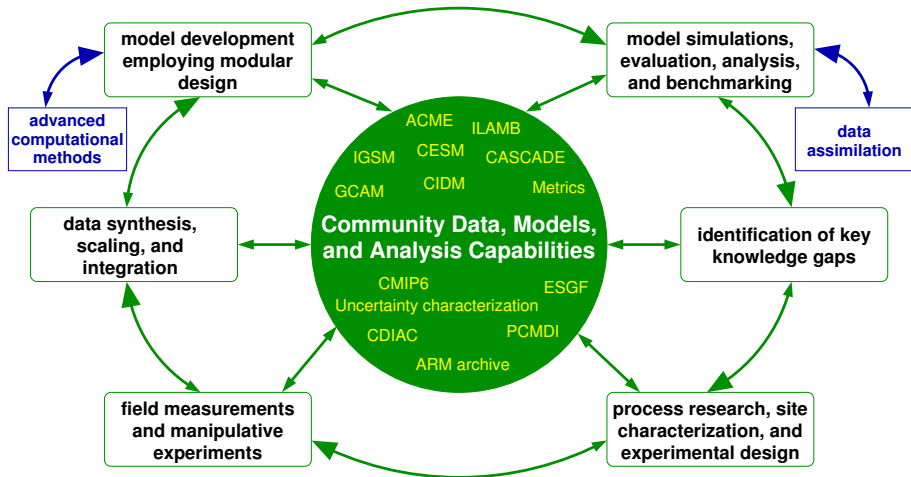
Anthropogenic emissions of radiatively active greenhouse gases into the atmosphere, especially carbon dioxide (CO₂), are rapidly changing the baseline of these gases and altering the Earth's climate (IPCC, 2007; Ramanath and G. Canabali, 2010). This perturbation of the global carbon cycle is expected 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 (Domenico et al., 2007). Understanding and predicting the strengths and direction of feedbacks is critically important

Hoffman, Forrest M., James T. Forrester, 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.

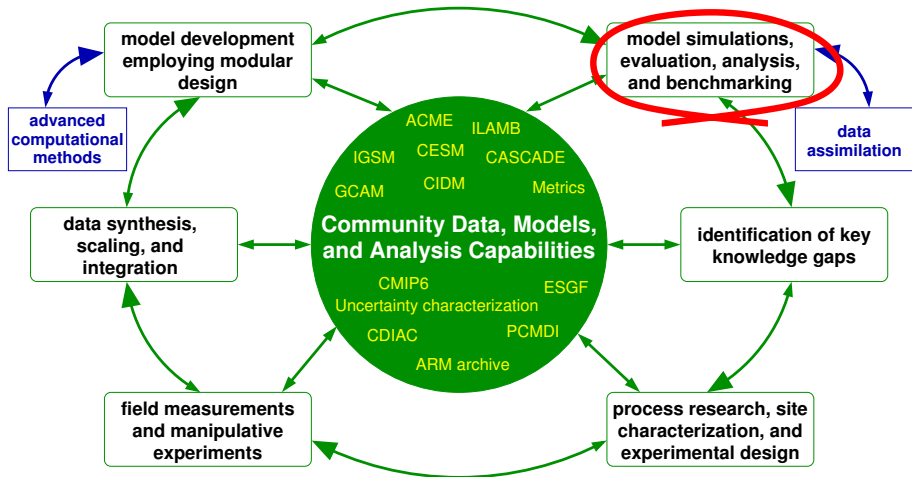
Model, Experiment, and Data Integration Strategy



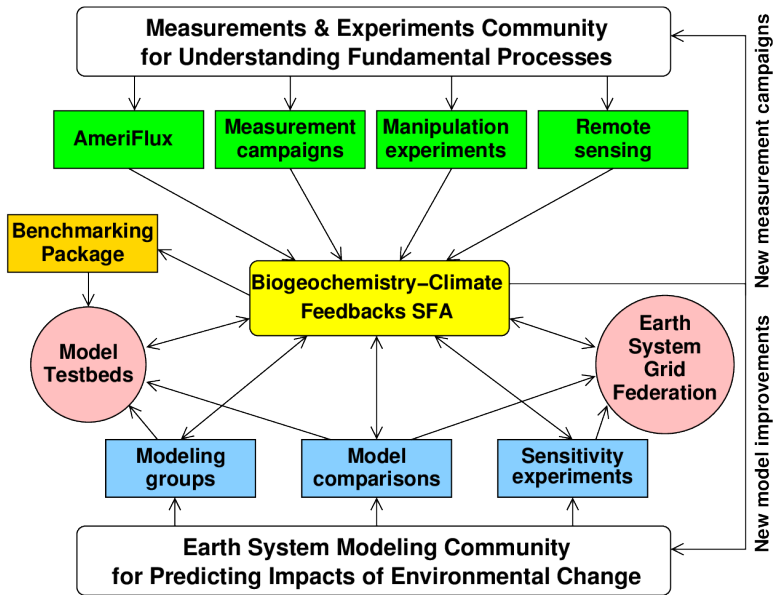
Model, Experiment, and Data Integration Strategy



Model, Experiment, and Data Integration Strategy

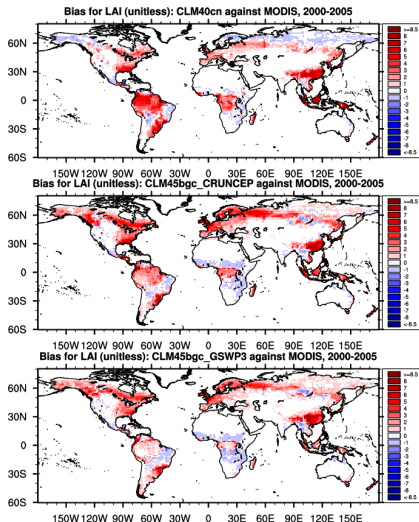


Biogeochemistry–Climate Feedbacks SFA Diagram



What is ILAMB?

- ▶ The **International Land Model Benchmarking (ILAMB)** project seeks to develop internationally accepted standards for land model evaluation.
- ▶ Model **benchmarking** can diagnose impacts of model development and guide synthesis efforts like IPCC.
- ▶ **Effective benchmarks** must draw upon a broad set of independent observations to evaluate model performance on multiple temporal and spatial scales.
- ▶ A free, **open source analysis and diagnostics software package** for community use will enhance model intercomparison projects.



Bias in mean annual leaf area index from comparison of three versions CLM with MODIS.



International Land Model Benchmarking (ILAMB) Meeting The Beckman Center, Irvine, CA, USA January 24-26, 2011



GLOBAL
IGBP
CHANGE
International
Geosphere-Biosphere
Programme

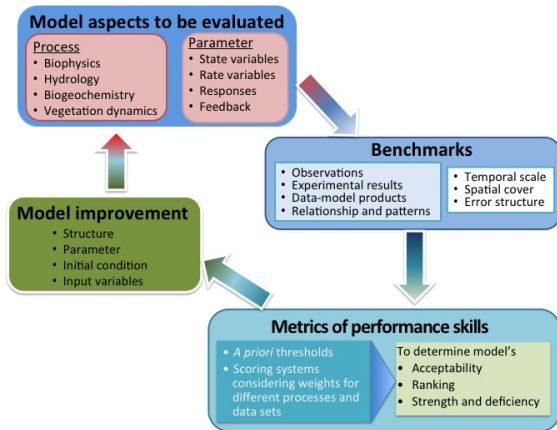


DEPARTMENT OF EARTH SYSTEM SCIENCE
SCHOOL OF PHYSICAL SCIENCES
UNIVERSITY OF CALIFORNIA • IRVINE

- ▶ 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.
- ▶ Methodology for model–data comparison and baseline standard for performance of land model process representations (Luo et al., 2012).

Benchmarking Methodology (Luo et al., 2012)

- ▶ Based on this methodology and prior work in C-LAMP, we developed a new model benchmarking package for ILAMB.
- ▶ Prototype is ready for use in NCL and a new version is under development using python.



(Luo et al., 2012)

ILAMB Prototype developed by Mingquan Mu at UCI

- ▶ Assesses 24 variables in 4 categories from ~45 datasets
 - ▶ aboveground live biomass, burned area, carbon dioxide, gross primary production, leaf area index, global net ecosystem carbon balance, net ecosystem exchange, ecosystem respiration, soil carbon
 - ▶ evapotranspiration, latent heat, terrestrial water storage anomaly
 - ▶ albedo, surface upward SW radiation, surface net SW radiation, surface upward LW radiation, surface net LW radiation, surface net radiation, sensible heat
 - ▶ surface air temperature, precipitation, surface relative humidity, surface downward SW radiation, surface downward LW radiation
- ▶ Graphics and scoring system
 - ▶ annual mean, bias, RMSE, seasonal cycle, spatial distribution, interannual coefficient of variation, spatial distribution, long-term trend
- ▶ Software is available at
<http://redwood.ess.uci.edu/mingquan/www/ILAMB/index.html>

ILAMB Prototype: Global Variables for 12 Models

Global Variables ([Info](#) for Weightings)

	MeanModel	ber-cmi-4-0a	BNU-ESM	CanESM2	CESM1-BGC	GDLM-ESM2G	HadGEM2-ES	inmcm4	IPSL-CM5A-LR	MIROC-ESM	MPI-ESM-LR	MRI-ESM1	NorESM1-ME
Aboveground Live Biomass	0.48	0.52	0.50	0.61	0.45	0.58	0.47	0.54	0.48	0.52	0.51	0.47	0.45
Burned Area	0.38	-	-	-	0.37	-	-	-	-	-	0.38	-	0.38
Carbon Dioxide	0.85	-	0.45	0.65	0.78	0.65	-	-	-	0.75	0.68	0.68	0.75
Coastal Primary Productivity	0.77	0.72	0.73	0.64	0.67	0.76	0.68	0.76	0.67	0.65	0.65	0.53	0.70
Land Area Index	0.66	0.66	0.61	0.60	0.53	0.45	0.55	0.60	0.66	0.62	0.68	0.63	0.50
Global Net Ecosystem Carbon Balance	0.58	-	0.38	0.37	0.38	0.18	-	0.46	0.25	0.38	0.42	0.37	0.40
Net Ecosystem Exchange	0.45	0.47	0.47	0.35	0.46	0.45	0.46	0.44	0.53	0.48	0.58	0.48	0.48
Ecosystem Respiration	0.75	0.72	0.72	0.65	0.67	0.71	0.66	0.78	0.67	0.68	0.68	0.47	0.64
Soil Carbon	0.55	0.56	0.42	0.56	0.38	0.51	0.51	0.53	0.57	0.53	0.41	0.53	0.39
Summary	0.64	0.59	0.54	0.54	0.55	0.53	0.59	0.57	0.57	0.58	0.54	0.51	0.55
Evapotranspiration	0.75	0.73	0.72	0.72	0.73	0.78	0.74	0.65	0.75	0.76	0.73	0.73	0.72
Latent Heat	0.86	0.76	0.77	0.77	0.78	0.74	0.77	0.72	0.77	0.75	0.76	0.78	0.76
Terrestrial Water Storage Anomaly	0.53	0.45	0.35	0.54	0.48	0.43	-	0.52	0.45	0.52	0.55	0.47	0.45
Summary	0.65	0.45	0.61	0.68	0.66	0.62	0.35	0.64	0.65	0.66	0.68	0.66	0.64
Albedo	0.72	0.71	0.63	0.71	0.73	0.65	0.74	0.67	0.71	0.67	0.73	0.64	0.72
Surface Upward SW Radiation	0.78	0.73	0.67	0.74	0.78	0.74	0.77	0.74	0.74	0.72	0.78	0.67	0.76
Surface Net SW Radiation	0.84	0.86	0.84	0.85	0.85	0.86	0.85	0.84	0.82	0.83	0.87	0.85	0.85
Surface Upward LW Radiation	0.96	0.91	0.91	0.91	0.92	0.91	0.92	0.89	0.96	0.91	0.92	0.92	0.92
Surface Net LW Radiation	0.81	0.82	0.81	0.79	0.82	0.81	0.83	0.75	0.78	0.78	0.81	0.82	0.81
Surface Net Radiation	0.78	0.75	0.76	0.80	0.80	0.80	0.79	0.74	0.77	0.76	0.88	0.78	0.80
Sensible Heat	0.76	0.65	0.70	0.71	0.75	0.65	0.75	0.66	0.65	0.65	0.65	0.72	0.72
Summary	0.75	0.78	0.75	0.78	0.80	0.78	0.80	0.75	0.76	0.76	0.75	0.77	0.75
Surface Air Temperature	0.87	0.87	0.85	0.85	0.88	0.85	0.87	0.85	0.87	0.85	0.88	0.88	0.87
Precipitation	0.78	0.67	0.64	0.67	0.76	0.68	0.72	0.68	0.68	0.68	0.76	0.65	0.63
Surface Relative Humidity	0.81	-	0.80	0.76	0.82	-	-	0.75	0.82	-	-	0.83	0.81
Surface Downward SW Radiation	0.86	0.88	0.87	0.87	0.88	0.87	0.87	0.87	0.83	0.86	0.88	0.86	0.88
Surface Downward LW Radiation	0.96	0.92	0.91	0.91	0.92	0.92	0.92	0.90	0.89	0.91	0.93	0.91	0.91
Summary	0.82	0.82	0.81	0.80	0.83	0.82	0.84	0.81	0.81	0.81	0.84	0.83	0.82
Overall	0.65	0.51	0.59	0.60	0.64	0.56	0.49	0.57	0.57	0.59	0.61	0.55	0.63

ILAMB Prototype: Global Variables for 12 Models

Global Variables ([Info](#) for Weightings)

	MeanModel	bcc-rcsm1-1-m	BNU-ESM	CanESM2	CESM1-BGC	GFDL-ESM2G	HadGE
Aboveground Live Biomass	0.68	0.52	0.50	0.61	0.65	0.58	0.6
Burned Area	0.38	-	-	-	0.37	-	-
Carbon Dioxide	0.85	-	0.65	0.65	0.78	0.65	-
Gross Primary Productivity	0.77	0.72	0.73	0.64	0.70	0.67	0.6
Leaf Area Index	0.66	0.66	0.41	0.60	0.53	0.49	0.5
Global Net Ecosystem Carbon Balance	0.58	-	0.38	0.27	0.38	0.18	-
Net Ecosystem Exchange	0.49	0.47	0.47	0.39	0.48	0.49	0.4
Ecosystem Respiration	0.75	0.72	0.72	0.65	0.67	0.71	0.6
Soil Carbon	0.55	0.50	0.42	0.56	0.38	0.51	0.5
Summary	0.64	0.59	0.54	0.54	0.55	0.53	0.5
Evapotranspiration	0.75	0.73	0.72	0.72	0.73	0.70	0.7
Latent Heat	0.80	0.76	0.77	0.77	0.78	0.74	0.7
Terrestrial Water Storage Anomaly	0.53	0.45	0.35	0.54	0.48	0.43	-
Summary	0.69	0.65	0.61	0.68	0.66	0.62	0.7
Albedo	0.72	0.71	0.61	0.71	0.73	0.69	0.7
Surface Upward SW Radiation	0.78	0.73	0.67	0.74	0.78	0.74	0.7
Surface Net SW	0.84	0.86	0.84	0.85	0.85	0.86	0.8

Scoring for Global GPP from Fluxnet-MTE

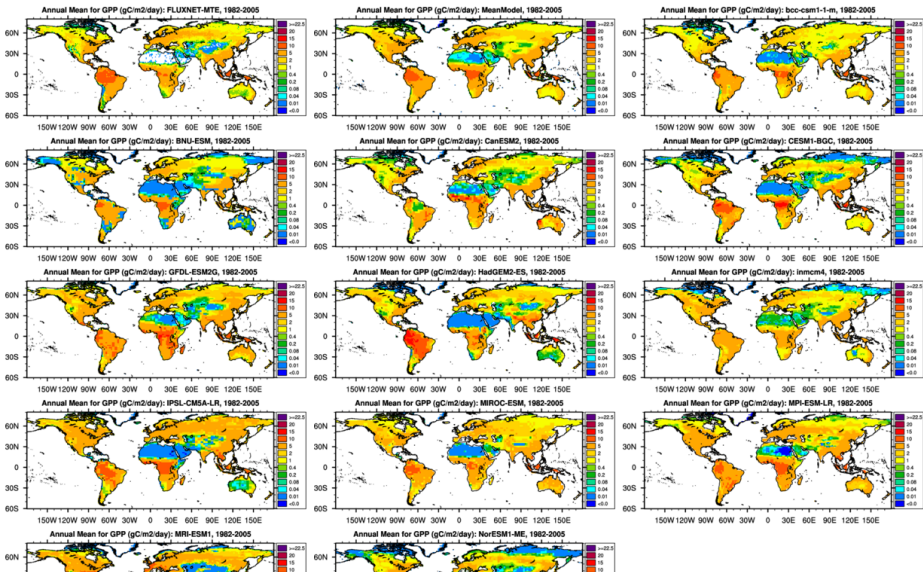
Diagnostic Summary for Gross Primary Productivity: Model vs. FLUXNET-MTE

	Global Patterns				Regional and Seasonal Patterns	Scoring (Info)				
	Annual Mean (PgC/yr)	Bias (PgC/yr)	RMSE (PgC/mon)	Phase Difference (months)	Regional Means	Global Bias	RMSE	Seasonal Cycle	Spatial Distribution	Overall
Benchmark Jung et al. (2009)	118.4	-	-	0.0	access to plots	-	-	-	-	-
MeanModel	145.3	26.9	4.7	0.6	access to plots	0.77	0.73	0.78	0.94	0.79
bcc-csm1-1-m	114.4	-4.0	6.0	-0.2	access to plots	0.72	0.64	0.80	0.89	0.74
BNU-ESM	102.0	-16.4	6.2	0.1	access to plots	0.69	0.66	0.78	0.84	0.73
CanESM2	129.2	10.8	7.3	0.8	access to plots	0.64	0.60	0.68	0.70	0.64
CESM1-BGC	130.3	11.9	5.8	0.5	access to plots	0.69	0.65	0.76	0.87	0.72
GFDL-ESM2G	175.1	56.7	9.8	0.5	access to plots	0.66	0.54	0.73	0.83	0.66
HadGEM2-ES	145.9	27.5	7.4	0.3	access to plots	0.65	0.58	0.78	0.79	0.68
inmcm4	111.4	-7.0	5.6	0.3	access to plots	0.71	0.66	0.78	0.83	0.73
IPSL-CM5A-LR	166.6	48.2	8.8	0.4	access to plots	0.63	0.56	0.77	0.84	0.67
MIROC-ESM	131.7	13.3	6.2	0.2	access to plots	0.72	0.66	0.74	0.86	0.73
MPI-ESM-LR	169.9	51.5	7.4	0.3	access to plots	0.67	0.62	0.70	0.89	0.70
MRI-ESM1	236.1	117.7	12.5	0.2	access to plots	0.45	0.43	0.79	0.59	0.54
NorESM1-ME	130.4	12.0	6.5	0.5	access to plots	0.66	0.62	0.76	0.84	0.70

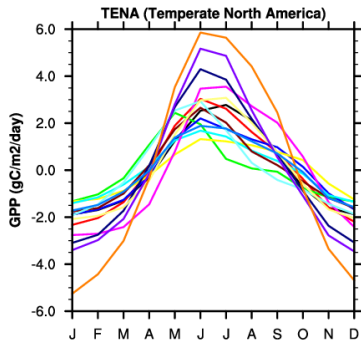
Notes: In calculating overall score, rmse score contributes double in comparison with all other scores.

Annual Mean Global GPP

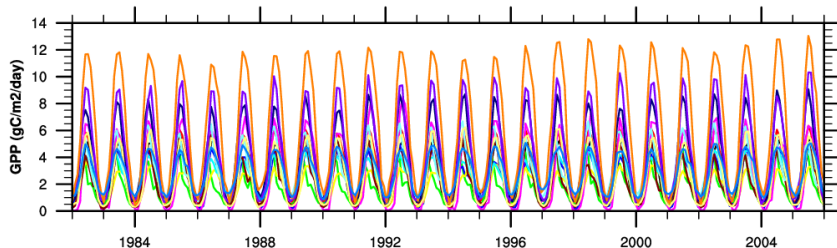
Models vs. FLUXNET-MTE



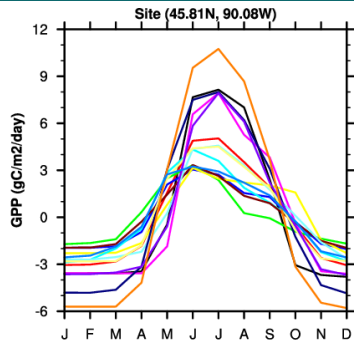
Seasonal Cycle of Regional GPP



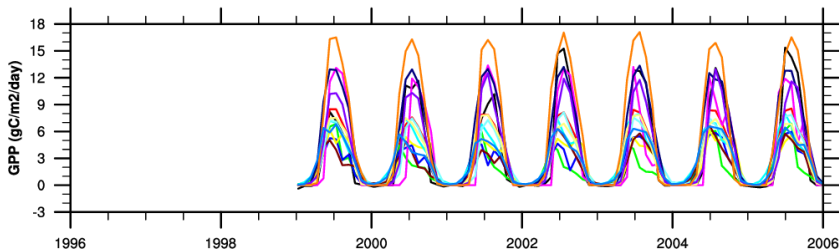
Model	Annual	Bias	RMSE
FLUXNET-MTE	2.36	-999.00	-999.00
MeanModel	2.99	0.63	0.74
bcc-csm1-1-m	1.82	-0.54	1.31
BNU-ESM	2.17	-0.19	0.62
CanESM2	1.76	-0.60	1.08
CESM1-BGC	2.45	0.08	0.78
GFDL-ESM2G	2.85	0.49	1.16
HadGEM2-ES	2.12	-0.24	0.72
inmcm4	3.06	0.70	1.20
IPSL-CM5A-LR	3.95	1.59	1.90
MIROC-ESM	2.48	0.12	0.35
MPI-ESM-LR	4.27	1.91	2.38
MRI-ESM1	6.13	3.76	4.46
NorESM1-ME	2.84	0.48	0.74



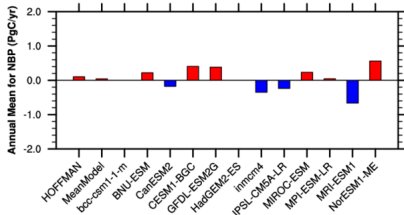
Seasonal Cycle of Site GPP



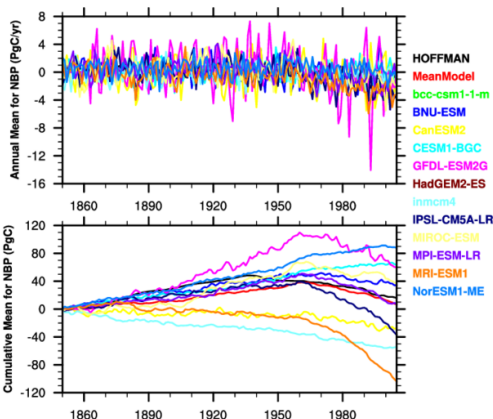
Model	Annual	Bias	RMSE
FLUXNET	3.62	-999.00	-999.00
MeanModel	3.11	-0.51	2.49
bcc-csm1-1-m	1.79	-1.83	4.42
BNU-ESM	2.00	-1.62	3.81
CanESM2	2.32	-1.30	3.69
CESM1-BGC	3.04	-0.58	3.19
GFDL-ESM2G	3.59	-0.03	2.87
HadGEM2-ES	2.06	-1.56	3.77
inmcm4	2.79	-0.83	2.75
IPSL-CM5A-LR	4.85	1.23	2.37
MIROC-ESM	2.81	-0.81	2.85
MPI-ESM-LR	3.68	0.06	1.72
MRI-ESM1	5.76	2.14	3.45
NorESM1-ME	2.69	-0.93	3.45



Global Net Ecosystem Carbon



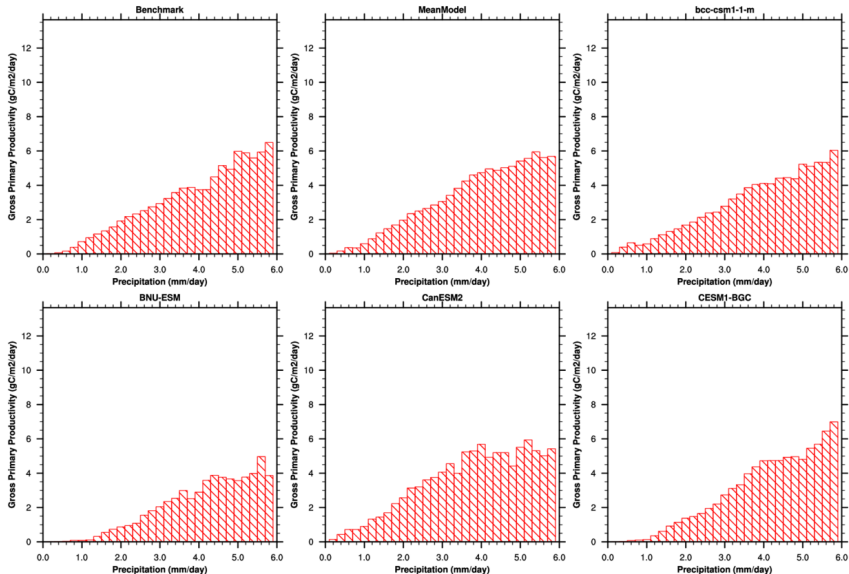
Global Net Ecosystem Carbon Balance



Long term carbon storage

Functional Relationships: GPP vs. Precipitation

Gross Primary Productivity vs. Precipitation



B. Root Mean Square Error Metric

For different variables, we use 2 different methods to calculate their global mean RMSE scores. For above ground biomass (biomass), burned area (burntarea), evapotranspiration (et), gross primary production (gpp), lead area index (lai), latent heat (le), net ecosystem exchange (nee), precipitation (pr), ecosystem respiration (reco), sensible heat (sh) and soil carbon (soilc), we use mass weighting (B3.1). For other variables, we use area weighting (B3.2).

$$M_i = 1 - \frac{RMSE_i}{\Phi_{obs,i}} \quad (B1)$$

$$M_i = e^{-RMSE_i / \epsilon} \quad (B2)$$

Mass weighting to calculate global mean RMSE score:

$$M = \frac{\sum_{i=1}^{ncells} M_i \times A_i \times |AM_{obs,i}|}{\sum_{i=1}^{ncells} A_i \times |AM_{obs,i}|} \quad (B3.1)$$

Area weighting to calculate global mean RMSE score:

$$M = \frac{\sum_{i=1}^{ncells} M_i \times A_i}{\sum_{i=1}^{ncells} A_i} \quad (B3.2)$$

We use Eqs. B1-2 and Eq. B3.1 or B3.2 to calculate root mean square error metric score M_i at grid cell or site i and its global mean M , respectively. Where $\Phi_{obs,i}$ is the root mean square for monthly mean annual cycle of the observation at grid cell i (for grid data) or site i (for site observation), and $RMSE_i$ is the root mean square error between model and observation. $AM_{obs,i}$ is annual mean of the observation at grid cell or site i . $|AM_{obs,i}|$ is to calculate its absolute value. A_i is the area for grid cell or site i . $ncells$ is the number of all land grid cells or sites where observation data is available. If the observation is site data, we set A_i equal to 1 (Ref: David Lawrence's personal Communication). This metric is used to compare magnitude and phase difference of the monthly mean annual cycle between the model and the observation.

C. Spatial Distribution Metric

$$M = \frac{4(1+R)}{(\sigma_j + 1) / \sigma_j)^2 (1 + R_j)} \quad (C)$$

2

We use Eq. C to calculate spatial distribution metric score M . R is the spatial correlation coefficient of the annual mean between model and observation. R_j is their ideal maximum correlation. Here, we set R_j equal to 1 for all models. σ_j is ratio for standard deviation of model to that of observation (Ref: Taylor, J. Geophys. Res., 106, 2001). This metric is used to compare magnitude and spatial pattern of annual mean of model with observation.

D. Seasonal Cycle Phase Metric

For different variables, we use 2 different methods to calculate their global mean phase scores. For above ground biomass (biomass), burned area (burntarea), evapotranspiration (et), gross primary production (gpp), lead area index (lai), latent heat (le), net ecosystem exchange (nee), precipitation (pr), ecosystem respiration (reco), sensible heat (sh) and soil carbon (soilc), we use mass weighting (D2.1). For other variables, we use area weighting (D2.2).

$$M_i = (1 + \cos \theta_i) / 2 \quad (D1)$$

Mass weighting to calculate global mean phase score:

$$M = \frac{\sum_{i=1}^{ncells} M_i \times A_i \times |AM_{obs,i}|}{\sum_{i=1}^{ncells} A_i \times |AM_{obs,i}|} \quad (D2.1)$$

Area weighting to calculate global mean phase score:

$$M = \frac{\sum_{i=1}^{ncells} M_i \times A_i}{\sum_{i=1}^{ncells} A_i} \quad (D2.2)$$

We use Eqs. D1 and D2.1 or D2.2 to calculate seasonal cycle phase metric score M_i at grid cell or site i and its global mean M , respectively. θ_i is the difference of the angle between the month of the maximum value for the model and that for the observation at grid cell i (for the grid data) or site i (for the site data). $AM_{obs,i}$ is annual mean of the observation at grid cell or site i . $|AM_{obs,i}|$ is to calculate its absolute value. A_i is the area for grid cell or site i . $ncells$ is the number of all land grid cells or sites where observation data is available. If the observation is site data, we set A_i equal to 1 (Ref: Prentice, et al., GBC, 25, 2011). This metric is used to compare phase difference of the monthly mean annual cycle between the model and the observation.

E. Interannual Variability Metric

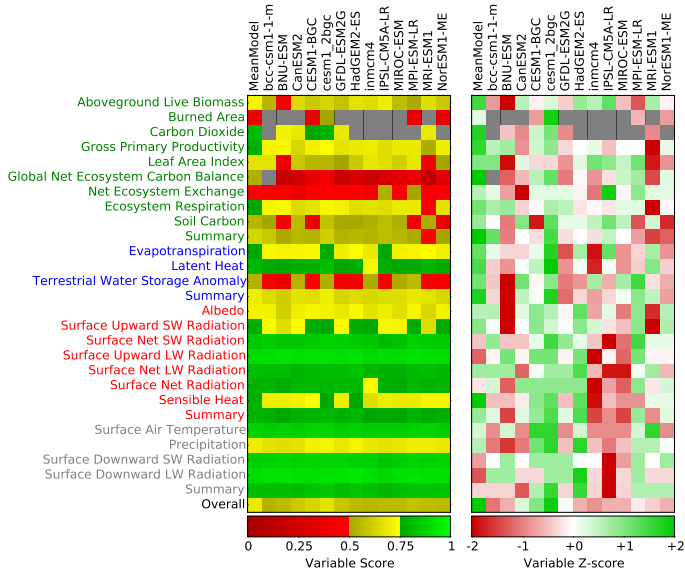
3

ILAMB Scoring Rules

Rules for scoring system

Score	Certainty of data	Scale appropriateness and coverage	Overall importance of constraint or process
1	Uncertainty estimates not available; significant methodological issues may influence data quality	Site level observations with limited regional coverage and/or short temporal duration	Observations that have limited influence on carbon cycle processes; includes some driver datasets and land surface measurements (e.g., Lin)
2	Uncertainty estimates not available; some methodological issues may influence data quality	Partial regional coverage; data sets providing up to 1 year of coverage	Driver observations or land surface measurements that have direct influence on carbon cycle processes (e.g., PPT, Tair, and Sin)
3	Uncertainty estimates not available; some peer-review evaluation of quality; minor methodological issues may remain	Regional coverage for at least 1 year; mismatches may exist between site-level and model grid cells	Biosphere process that contributes to carbon dynamics; data are a useful constraint for this specific process
4	Qualitative uncertainty information available from peer-review evaluations; methodology is well accepted	Important regional coverage; at least 1 year or more of observations	Important biosphere process regulating carbon cycle dynamics; data are moderately well-suited for constraining this process
5	Well defined and traceable uncertainty estimates; relatively low uncertainty estimates relative to range of model estimates; uncertainties less than $\pm 20\%$ at regional scales	Global scale in coverage; time series spanning multiple years; data products appropriate in scale for comparing directly with model grid cells	Critical process or constraint regulating climate-carbon or carbon-concentration feedbacks; data are well suited for discriminating among different model estimates

ILAMB Model Scoring by Variable



ILAMB Next Generation Layout

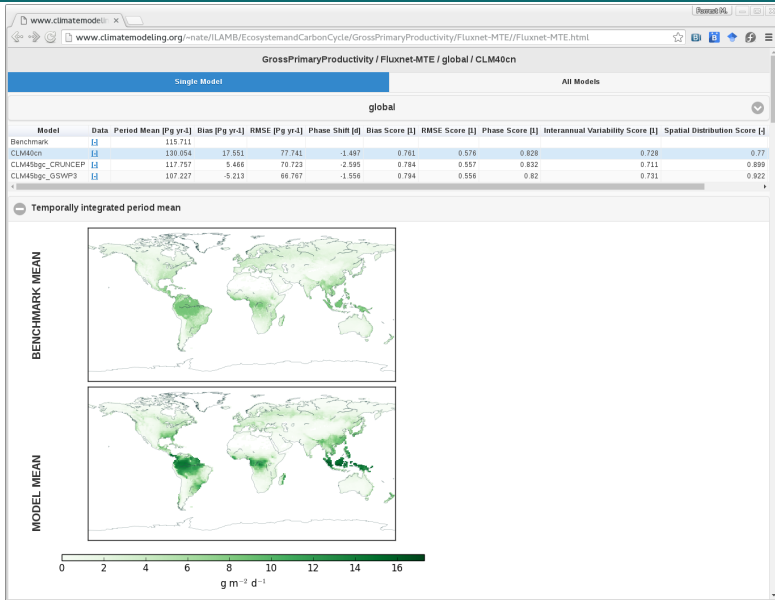
ILAMB Benchmark Results

Overview Results Table Model Comparisons

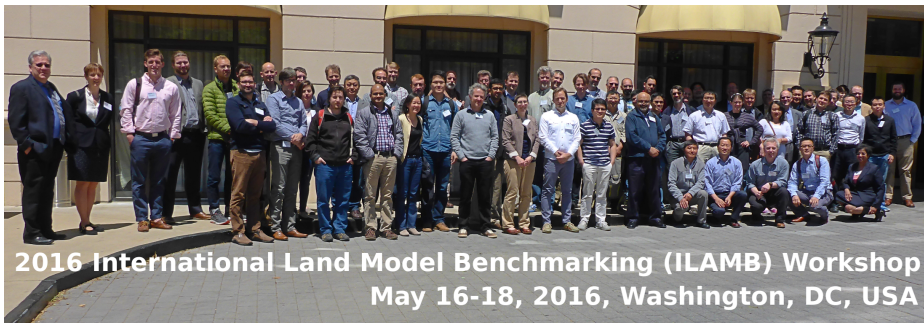
Columns...

	CLM40cn	CLM45bgc_CRUNCEP	CLM45bgc_GSWP3	
Biomass	0.40	0.40	0.41	▼
Burned Area	0.62	0.66	0.65	▼
Gross Primary Productivity	0.70	0.72	0.73	▲
Fluxnet (36.0%)	0.69	0.72	0.73	
Fluxnet-MTE (60.0%)	0.71	0.72	0.73	
Leaf Area Index	0.62	0.60	0.63	▼
Global Net Ecosystem Carbon Balance	0.17	0.23	0.20	▼
Net Ecosystem Exchange	0.55	0.55	0.55	▼
Ecosystem Respiration	0.67	0.70	0.72	▼
Soil Carbon	0.55	0.58	0.65	▼
Evapotranspiration	0.73	0.75	0.75	▼
Latent Heat	0.73	0.75	0.75	▼
Terrestrial Water Storage Anomaly	0.30	0.31	0.31	▼
Albedo	0.72	0.72	0.72	▼
Surface Upward SW Radiation	0.77	0.77	0.78	▼
Surface Net SW Radiation	0.80	0.80	0.81	▼
Surface Upward LW Radiation	0.81	0.81	0.82	▼
Surface Net LW Radiation	0.73	0.73	0.77	▼
Surface Net Radiation	0.77	0.77	0.78	▼
Sensible Heat	0.72	0.72	0.74	▼
Surface Air Temperature	0.83	0.83	0.84	▼
Precipitation	0.76	0.76	0.78	▼

ILAMB Next Generation Layout



Second U.S. ILAMB Workshop, May 2016



- ▶ **DOE sponsored** the second ILAMB Workshop in the U.S.
- ▶ **60+ researchers** from Australia, Japan, China, Germany, Sweden, Netherlands, U.K., and U.S. participated, representing **10 different major modeling centers**, and ~25 online attendees.
- ▶ **All workshop materials and reports “crowdsourced” online.**
- ▶ ILAMBv1 results and two tutorials for ILAMBv2 were presented.

Future ILAMB Development and Application

- ▶ Current ILAMB Prototype was applied to:
 - ▶ Model development of the Community Land Model (CLM)
 - ▶ CMIP5 Historical and esmHistorical simulations
 - ▶ ACME Land Model evaluation
- ▶ Within U.S. Department of Energy projects:
 - ▶ NGEE Arctic, NGEE Tropics, and SPRUCE are adopting the framework for evaluating process parameterizations & integrating field observations
 - ▶ ACME is developing metrics for evaluation of new land model features
 - ▶ BGC Feedbacks is developing the framework and benchmarking MIPs
- ▶ Future (and past) projects where we hope to apply ILAMB:
 - ▶ CMIP6, including C⁴MIP, LS3MIP, and LUMIP
 - ▶ TRENDY, MsTMIP, PLUME-MIP
 - ▶ NASA Permafrost Benchmark System (PBS) (Schaefer et al.)

Acknowledgments



U.S. DEPARTMENT OF
ENERGY

Office of Science



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.

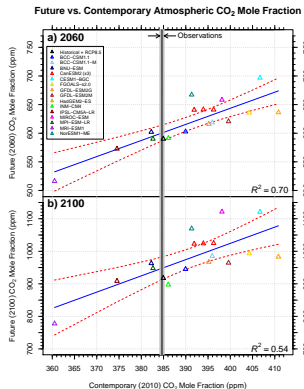
I wish to acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and thank the climate modeling groups for producing and making available their model output. For CMIP the U. S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

References

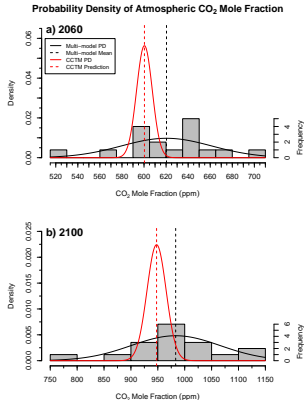
- R. J. Andres, J. S. Gregg, L. Losey, G. Marland, and T. A. Boden. Monthly, global emissions of carbon dioxide from fossil fuel consumption. *Tellus B*, 63(3):309–327, July 2011. doi:10.1111/j.1600-0889.2011.00530.x.
- R. J. Andres, T. A. Boden, F.-M. Bréon, P. Ciais, S. Davis, D. Erickson, J. S. Gregg, 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.
- 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 CO₂. *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 CO₂ projections from CMIP5 ESMS.



- Much of the model-to-model variation in projected CO₂ during the 21st century is tied to biases that existed during observational era.
- 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.
- Range of temperature increases at 2100 slightly reduced, from $5.1 \pm 2.2^\circ\text{C}$ for the full ensemble, to $5.0 \pm 1.9^\circ\text{C}$ after applying the emergent constraint.



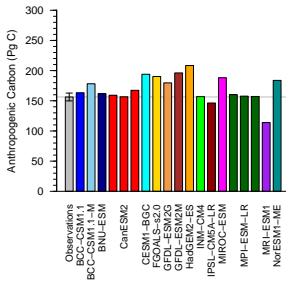
Best estimate using Mauna Loa CO₂

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

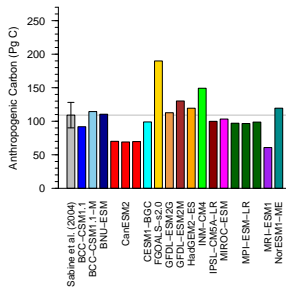
Hoffman, Forrest M., James T. Randerson, Vivek K. Arora, Qing Bao, Patricia Cadule, Duoying Ji, Chris D. Jones, Michio Kawamiya, Samar Khattiwala, Keith Lindsay, Atsushi Obata, Elena Shevliakova, Katharina D. Six, Jerry F. Tjiputra, Evgeny M. Volodin, and Tongwen Wu. February 2014. "Causes and Implications 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)

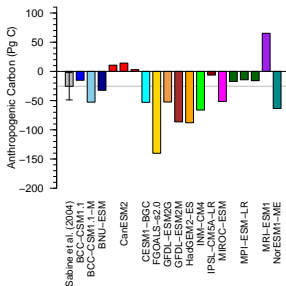
Atmosphere (1850–1994)



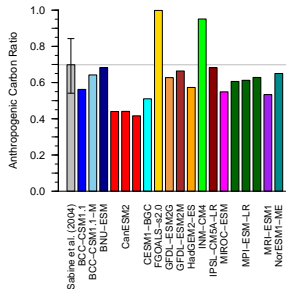
Ocean (1850–1994)



Land (1850–1994)



Ocean/Atmosphere (1850–1994)



Implications for CO₂, Radiative Forcing, and Temperature

Model	CO ₂ Mole			Radiative			Cumulative			ΔT		
	Fraction (ppm)			Forcing (W m ⁻²)			ΔT (°C)			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