Have Land Surface Carbon Cycle Models Improved Over Time?

Forrest M. Hoffman[1,2], Nathan Collier[1], Charles D. Koven[3], David M. Lawrence[4], Gretchen Keppel-Aleks[5], James T. Randerson[6], Mingquan Mu[6], William J. Riley[3], Qing Zhu[3], Jiafu Mao[1], Hyungjun Kim[7], J. Keith Moore[6], and Weiwei Fu[6]

[1]Oak Ridge National Laboratory (ORNL), [2]University of Tennessee Knoxville, [3]Lawrence Berkeley National Laboratory (LBNL), [4]National Center for Atmospheric Research (NCAR), [5]University of Michigan Ann Arbor, [6]University of California Irvine, and [7]University of Tokyo

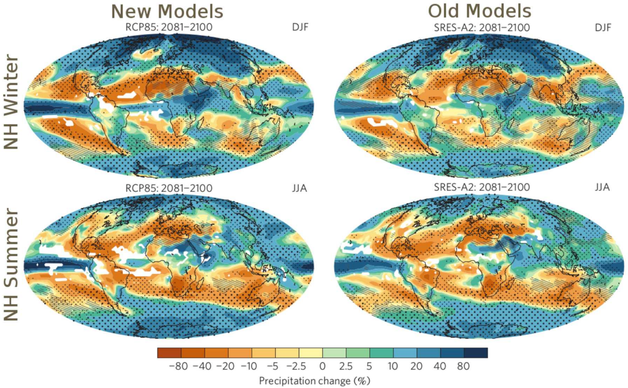
CESM Land Model Working Group Seminar

March 3, 2020

Problem: Model Uncertainty

Model uncertainty is one of the biggest challenges we face in Earth system science, yet comparatively little effort is devoted to fixing it (Carslaw et al., 2018)

- Model complexity is rapidly increasing as detailed process representations are added
- Evidence shows overall model uncertainty is reduced only slowly and is sometimes increased (Knutti and Sedláček, 2013)
- A balance must be struck between model "elaboration" and efforts to reduce model uncertainty



Patterns of precipitation change across two generations of models. Adapted from Knutti and Sedláček (2013).

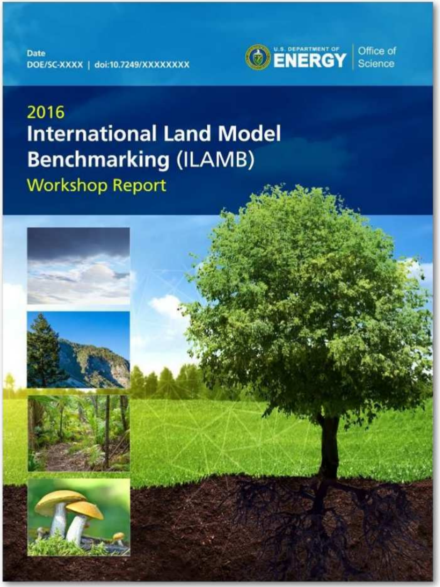
Why is Reducing Uncertainty a Challenge?

- Ecosystems have complex responses to a wide range of forcing factors in heterogeneous spatial environments, requiring a highly multivariate approach
- The focus is on adding complexity (e.g., more detailed representations of plant traits, photosynthesis, nutrient limitation, respiration), assuming more processes is better
- However, model uncertainty may increase, even as predictions of states and fluxes improves
- Rigorous confrontation of models with independent observations and large ensembles of simulations are required to reduce uncertainty
- Modeling centers have a limited capacity to conduct sensitivity experiments and systematically assess model fidelity, especially in fully coupled Earth system models
- Community-developed benchmarking tools are beginning to address part of the solution

International Land Model Benchmarking (ILAMB) Workshop
May 16–18, 2016, Washington, DC

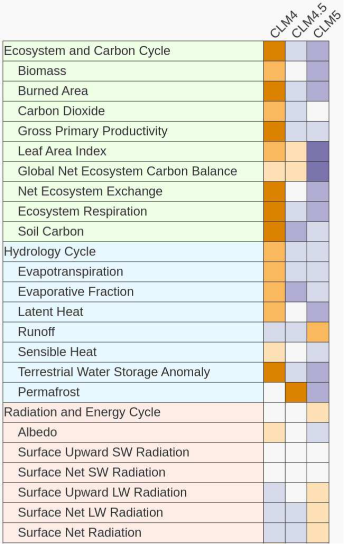
The **International Land Model Benchmarking (ILAMB)** community coordination activity was designed to

- Develop internationally accepted benchmarks
- Promote the use of these benchmarks
- Strengthen linkages between experimental, remote sensing, and modeling communities
- Support the design and development of open source benchmarking tools (Luo et al., 2012), like the **ILAMB Package** (Collier et al., 2018)

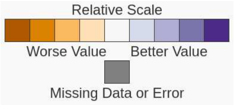


RUBISCO Argonne NATIONAL LABORATORY BROOKHAVEN NATIONAL LABORATORY BERKELEY LAB Los Alamos NATIONAL LABORATORY NCAR NATIONAL CENTER FOR ATMOSPHERIC RESEARCH OAK RIDGE National Laboratory UCI MICHIGAN

ILAMB Assesses Land Model Fidelity Across Three Generations



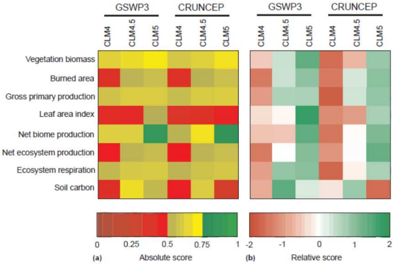
- CLM saw improvements in mechanistic treatment of hydrology, ecology, and land use with many more moving parts
- Simulations improved even with enhanced complexity
- Observational datasets are not always self-consistent
- Forcing uncertainty confounds assessment of model development (not shown)

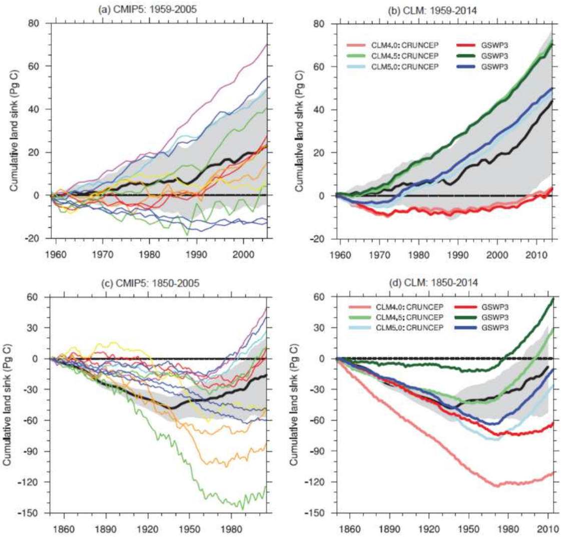


http://webext.cgd.ucar.edu/I20TR/_build_set1F/

(Lawrence et al., 2019)

Land Model Performance Depends Strongly on Forcing



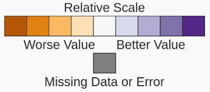


- Depending on the forcing used and the metric selected, different models may perform equally well
- ILAMB scores for CLM4, CLM4.5, and CLM5 forced with GSWP3 vs. CRUNCEP (above) and the cumulative land carbon sink for CMIP5 models vs. offline CLM (right).

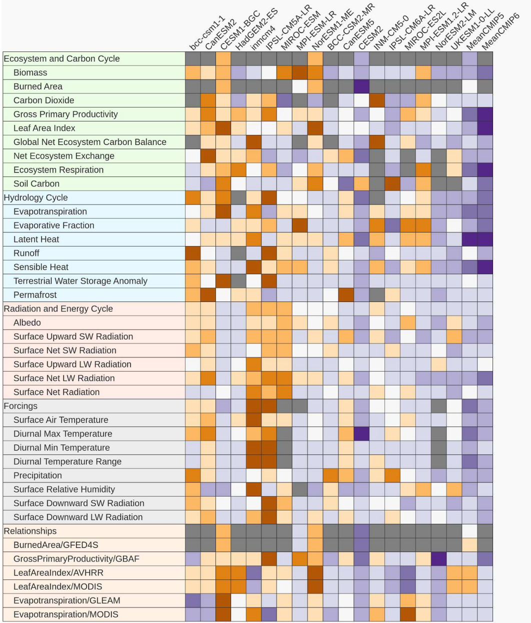
(Bonan et al., 2019)

CMIP5 vs. CMIP6 Models

- The CMIP6 suite of land models (right) has improved over the CMIP5 suite of land models (left)
- The multi-model mean for CMIP5 outperforms any single CMIP5 model
- The multi-model mean for CMIP6 outperforms any single CMIP6 model
- The multi-model mean CMIP6 land model is the “best model” overall
- Why did CMIP6 land models improve?



(Hoffman et al., in prep.)

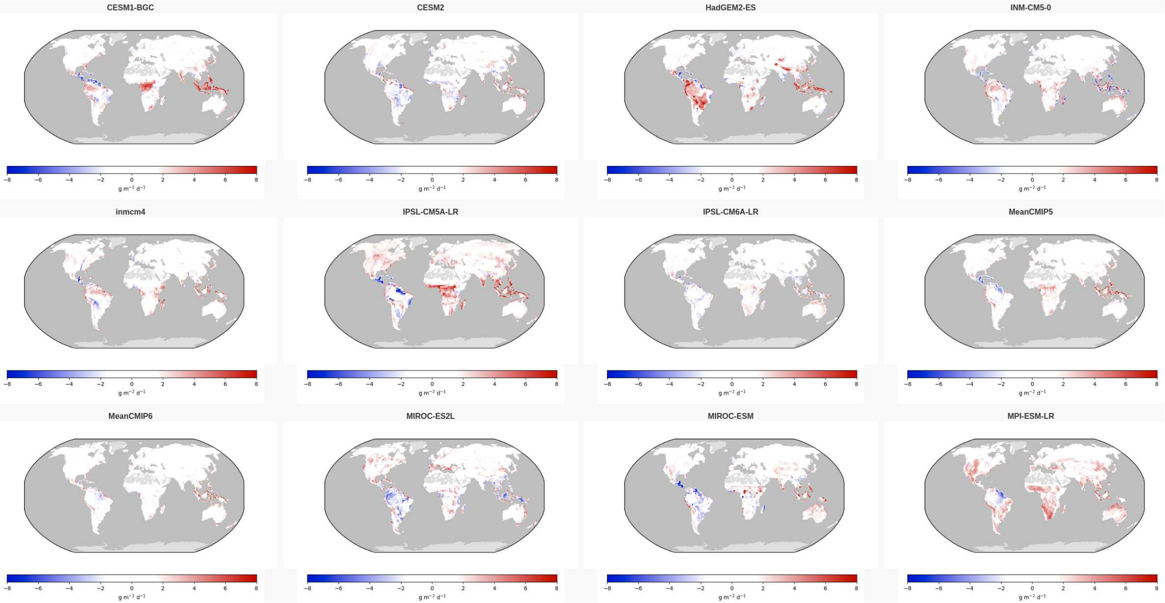


CMIP5 and CMIP6 Land Model Global GPP

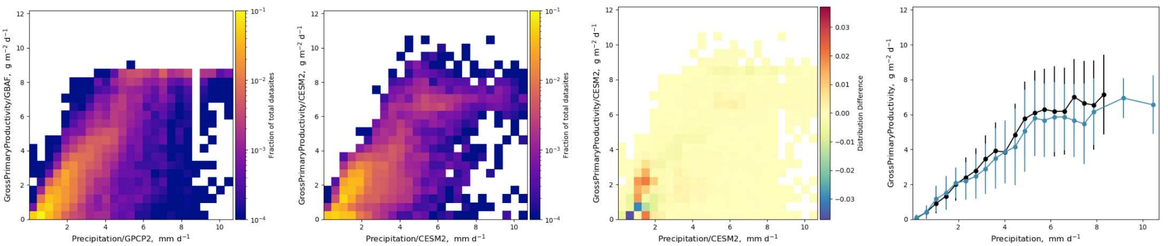
	Download Data	Period Mean (original grids) [Pg yr-1]	Model Period Mean (intersection) [Pg yr-1]	Model Period Mean (complement) [Pg yr-1]	Benchmark Period Mean (intersection) [Pg yr-1]	Benchmark Period Mean (complement) [Pg yr-1]	Bias [g m-2 d-1]	RMSE [g m-2 d-1]	Phase Shift [months]	Bias Score [1]	RMSE Score [1]	Seasonal Cycle Score [1]	Spatial Distribution Score [1]	Overall Score [1]
Benchmark	[-]	118.												
bcc-csm1-1	[-]	123.	114.	6.80	118.	0.0600	0.203	1.94	1.27	0.424	0.267	0.809	0.946	0.543
BCC-CSM2-MR	[-]	114.	109.	4.43	118.	0.501	-0.0721	1.68	1.28	0.433	0.326	0.796	0.941	0.564
CanESM2	[-]	129.	119.	7.33	118.		0.00724	2.27	2.10	0.364	0.345	0.658	0.832	0.509
CanESM5	[-]	141.	131.	8.05	118.		0.675	1.85	1.70	0.427	0.330	0.701	0.934	0.544
CESM1-BGC	[-]	129.	124.	4.32	118.	0.501	0.309	1.74	1.38	0.392	0.350	0.761	0.873	0.545
CESM2	[-]	110.	105.	4.21	118.	0.473	-0.0938	1.72	1.52	0.411	0.364	0.786	0.935	0.572
HadGEM2-ES	[-]	137.	132.	5.25	118.	0.686	0.533	2.24	1.25	0.366	0.265	0.781	0.848	0.505
INM-CM5-0	[-]	157.	147.	9.49	118.		0.629	1.81	1.32	0.413	0.340	0.796	0.925	0.563
inmcm4	[-]	136.	128.	8.25	113.	5.44	0.351	1.78	1.41	0.451	0.308	0.766	0.935	0.554
IPSL-CM5A-LR	[-]	165.	153.	9.00	118.	0.347	1.10	2.73	1.30	0.318	0.241	0.770	0.889	0.492
IPSL-CM6A-LR	[-]	116.	111.	4.25	118.	0.486	0.0566	1.45	1.32	0.498	0.364	0.751	0.960	0.587
MeanCMIP5	[-]	119.	114.	5.56	118.		0.505	1.39	1.18	0.465	0.423	0.780	0.961	0.610
MeanCMIP6	[-]	119.	114.	5.92	118.		0.160	1.14	1.13	0.526	0.457	0.798	0.964	0.640
MIROC-ES2L	[-]	116.	106.	7.96	118.	0.0975	-0.0380	1.89	2.11	0.379	0.320	0.622	0.920	0.512
MIROC-ESM	[-]	129.	121.	6.01	108.	10.1	0.308	2.06	1.40	0.425	0.322	0.749	0.918	0.547
MPI-ESM-LR	[-]	170.	162.	6.90	110.	8.62	1.22	2.37	1.43	0.378	0.291	0.699	0.926	0.517
MPI-ESM1.2-LR	[-]	141.	135.	5.42	110.	8.62	0.597	2.03	1.25	0.410	0.304	0.768	0.928	0.543
NorESM1-ME	[-]	129.	121.	6.29	118.		0.331	1.92	1.46	0.354	0.350	0.759	0.838	0.530
NorESM2-LM	[-]	107.	99.4	6.03	118.		-0.101	1.72	1.56	0.409	0.369	0.785	0.937	0.574
UKESM1-0-LL	[-]	127.	121.	5.41	118.	0.585	0.333	1.94	1.32	0.416	0.293	0.763	0.915	0.536

- Most models of the same lineage improved in various characteristics between CMIP5 and CMIP6
- The MeanCMIP5 and MeanCMIP6 models perform the best

(Hoffman et al., in prep.)


CESM1-BGC
CESM2
HadGEM2-ES
INM-CM5-0
inmcm4
IPSL-CM5A-LR
IPSL-CM6A-LR
MeanCMIP5
MeanCMIP6
MIROC-ES2L
MIROC-ESM
MPI-ESM-LR
−8 −6 −4 −2 0 2 4 6 8
g m−2 d−1

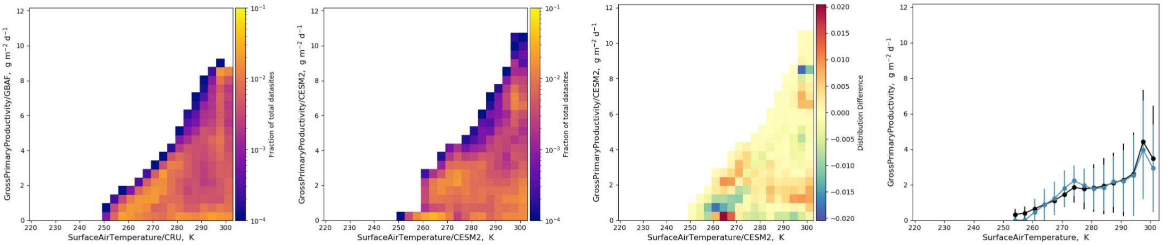
Precipitation/GPCP2



SurfaceDownwardSWRadiation/CERES

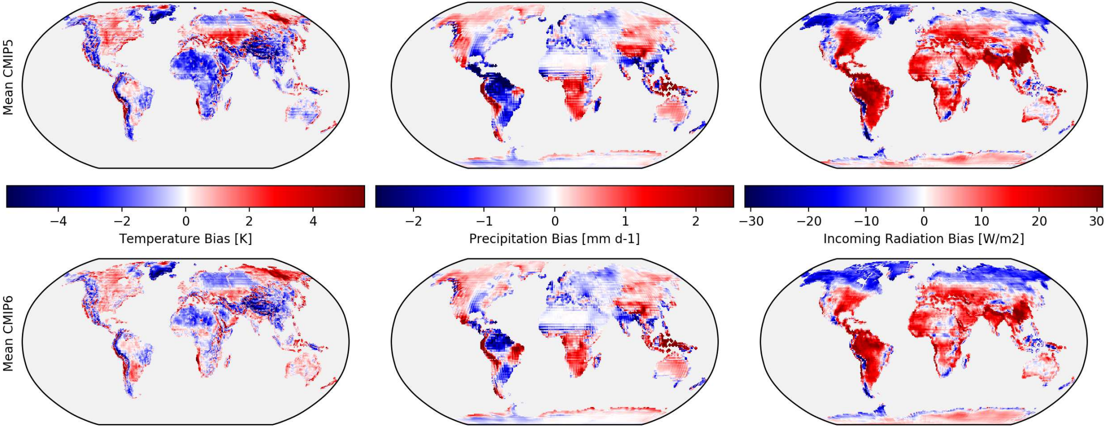
SurfaceNetSWRadiation/CERES

SurfaceAirTemperature/CRU



Reasons for Land Model Improvements

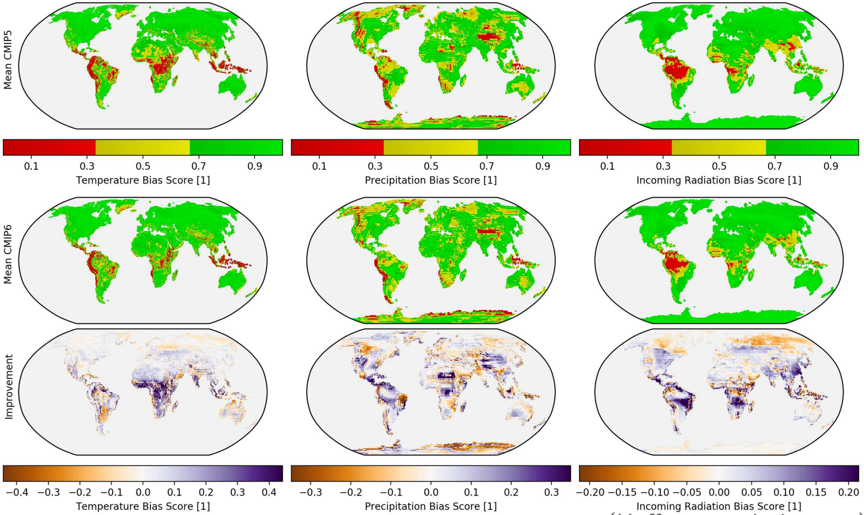
ESM improvements in climate forcings (temperature, precipitation, radiation) likely partially drove improvements exhibited by land carbon cycle models



(Hoffman et al., in prep.)

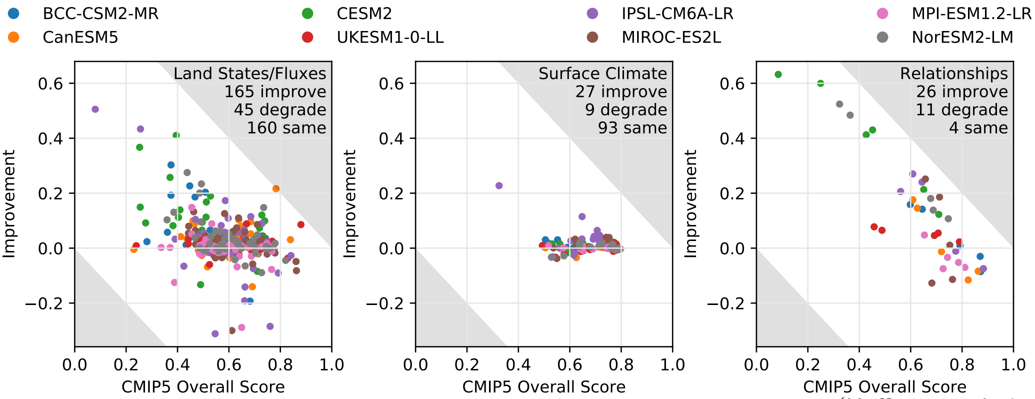
Reasons for Land Model Improvements

Differences in bias scores for temperature, precipitation, and incoming radiation were primarily positive, further indicating more realistic climate representation



(Hoffman et al., in prep.)

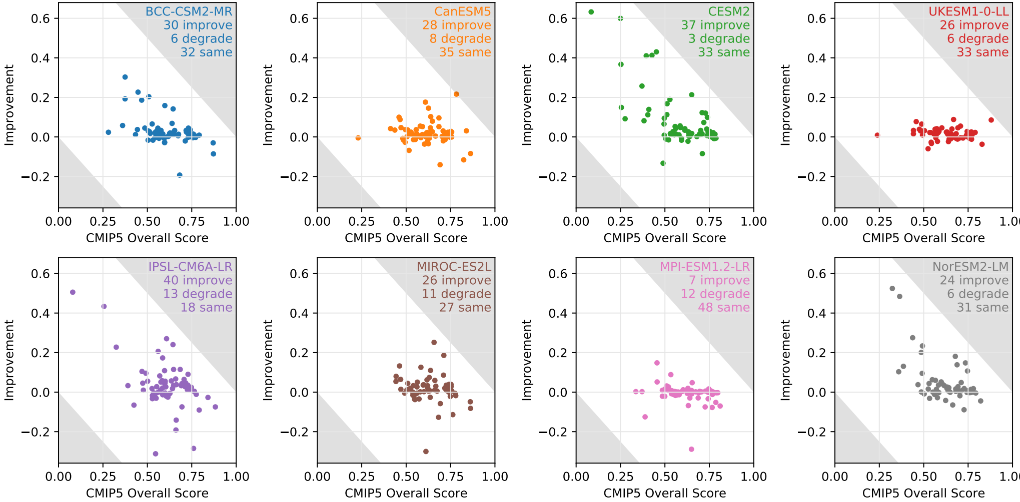
Reasons for Land Model Improvements



(Hoffman et al., in prep.)

While forcings got better, the largest improvements were in **variable-to-variable relationships**, *suggesting* that further land model development (increased complexity?) was also partially responsible for higher CMIP6 model scores

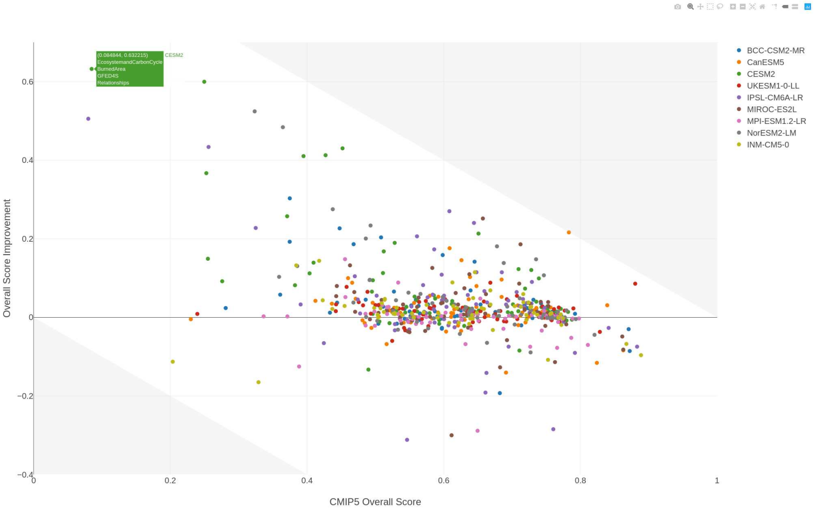
Improvements by Land Model



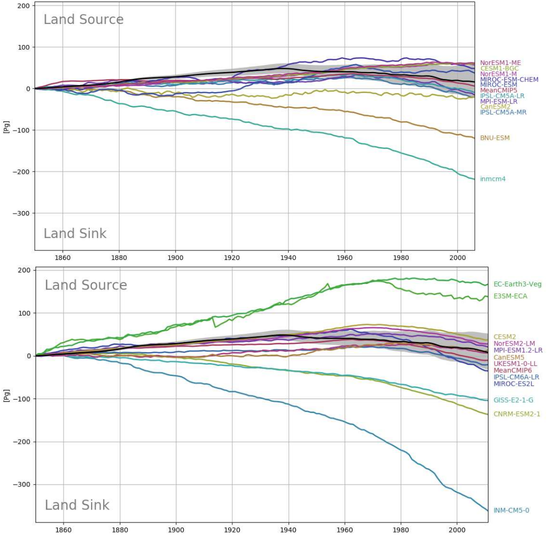
(Hoffman et al., in prep.)

Interactive Exploration of Multi-Model Performance

https://www.ilamb.org/CMIP5v6/historical/chart.html



Land Model Spread in Net Ecosystem Carbon Balance



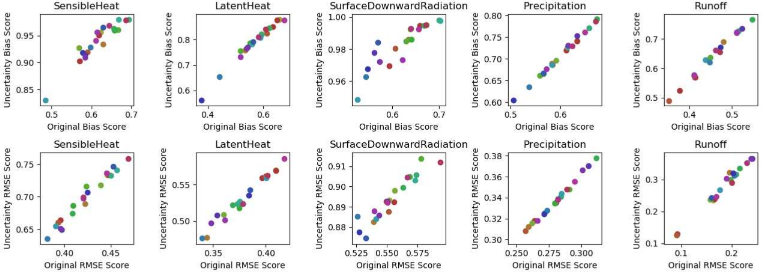
- The spread in the net ecosystem carbon balance increased between CMIP5 and CMIP6
  - CMIP5 at 2005: -215 Pg to 75 Pg $\rightarrow$ 290 Pg
  - CMIP6 at 2010: -360 Pg to 175 Pg $\rightarrow$ 535 Pg
- However, the range from most multi-generation models was reduced

(Hoffman et al., in prep.)

RUBISCO Argonne National Laboratory Brookhaven National Laboratory Berkeley Lab Los Alamos National Laboratory NCAR Oak Ridge National Laboratory UCI Michigan

Addressing Observational Uncertainty

- Few observational datasets provide complete uncertainties
- ILAMB uses multiple datasets for most variables and allows users to weight them according to a rubric of uncertainty, scale mismatch, etc.
- ILAMB can also use:
  - Full spatial/temporal uncertainties provided with data
  - Fixed, expert-derived uncertainty for a dataset
  - Uncertainties derived from combining multiple datasets

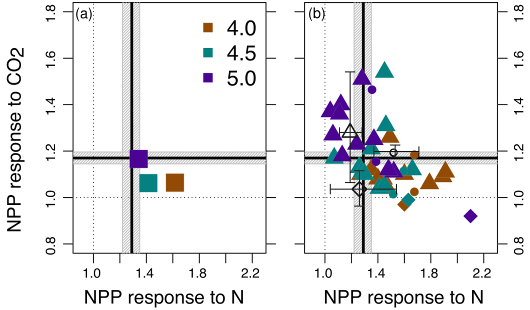


(Collier et al., in prep.)

- Experiments with CLASS self-consistent data (Hobeichi et al., 2020) demonstrates that while scores shift, including uncertainty rarely alters the rank ordering of models (figure)

Beyond Static Benchmarking

- To better support model development verification, we need to incorporate metrics from manipulative experiments
  - Simulated effect sizes of nitrogen versus CO_2 enrichment on rates of net primary production (NPP) calculated (a) globally or (b) for each plant functional type in CLM4, 4.5, and 5
  - Observational constraints for N response and CO_2 response are shown with vertical and horizontal polygons (mean $\pm 95\%$ confidence intervals)
  - In (b), observed (open symbols) and simulated (filled symbols) effect sizes of individual PFTs for woody vegetation, C_3 grasses, and C_4 grasses (triangles, circles, and diamonds, respectively)



(Wieder et al., 2019)

- Much more work is needed to foster land model ensemble simulations and benchmarking, including land model testbeds, diurnal and seasonal metrics, new synthesis datasets, . . .

Conclusions and Future Research

- Based on ILAMB model–data comparisons, CMIP6 land models improved over CMIP5 land models due to **1) improved climate forcing**; **2) improved process representation**
- Variable-to-variable relationships exhibited the largest improvements for some models
- Thus, CMIP6 models are more valuable for impact and adaptation/mitigation analysis
- Model improvements in mean states and fluxes may not result in reduced uncertainty or projected model spread
- Upon further examination, will improved multi-model performance result in reduced spread in feedback sensitivities, projected land carbon storage, and future climate change?
- Can we use ILAMB scores to weight contributions to multi-model means and thereby reduce contemporary biases, reduce future projected uncertainties, and alter expected mitigation targets?

Acknowledgments

This research was supported by the *Reducing Uncertainties in Biogeochemical Interactions through Synthesis and Computation Science Focus Area (RUBISCO SFA)*, which is sponsored by the Regional and Global Model Analysis (RGMA) activity of the Earth & Environmental Systems Modeling (EESM) Program in the Climate and Environmental Sciences Division (CESD) of the Office of Biological and Environmental Research (BER) in the US Department of Energy (DOE) Office of Science. Additional support was provided by the Laboratory Directed Research and Development Program of Oak Ridge National Laboratory, which is managed by UT-Battelle, LLC, for the US Department of Energy under contract DE-AC05-00OR22725.

We acknowledge the World Climate Research Programme (WCRP), which, through its Working Group on Coupled Modelling (WGCM), coordinated and promoted the Coupled Model Intercomparison Project phase 6 (CMIP6). We thank the climate modeling groups for producing and making available their model output, the Earth System Grid Federation (ESGF) for archiving the data and providing access, and the multiple funding agencies that support CMIP6 and ESGF. This research used resources of the National Energy Research Scientific Computing Center (NERSC), a DOE Office of Science User Facility supported by the Office of Science of the US Department of Energy under Contract No. DE-AC02-05CH11231. We thank DOE's RGMA activity, Data Management Program, and NERSC for making this coordinated CMIP6 analysis activity possible.

References

G. B. Bonan, D. L. Lombardozzi, W. R. Wieder, K. W. Oleson, D. M. Lawrence, F. M. Hoffman, and N. Collier. Model structure and climate data uncertainty in historical simulations of the terrestrial carbon cycle (1850–2014). *Global Biogeochem. Cycles*, 33(10):1310–1326, Oct. 2019. doi:10.1029/2019GB006175.

K. S. Carslaw, L. A. Lee, L. A. Regayre, and J. S. Johnson. Climate models are uncertain, but we can do something about it. *Eos Trans. AGU*, 99, Feb. 2018. doi:10.1029/2018EO093757.

N. Collier, F. M. Hoffman, D. M. Lawrence, G. Keppel-Aleks, C. D. Koven, W. J. Riley, M. Mu, and J. T. Randerson. The International Land Model Benchmarking (ILAMB) system: Design, theory, and implementation. *J. Adv. Model. Earth Sy.*, 10(11):2731–2754, Nov. 2018. doi:10.1029/2018MS001354.

S. Hobeichi, G. Abramowitz, and J. Evans. Conserving land–atmosphere synthesis suite (class). *J. Clim.*, 33(5):1821–1844, Mar. 2020. doi:10.1175/JCLI-D-19-0036.1.

R. Knutti and J. Sedláček. Robustness and uncertainties in the new CMIP5 climate model projections. *Nat. Clim. Change*, 3(4):369–373, Apr. 2013. doi:10.1038/nclimate1716.

D. M. Lawrence, R. A. Fisher, C. D. Koven, K. W. Oleson, S. C. Swenson, G. Bonan, N. Collier, B. Ghimire, L. van Kampenhout, D. Kennedy, E. Kluzek, P. J. Lawrence, F. Li, H. Li, D. Lombardozzi, W. J. Riley, W. J. Sacks, M. Shi, M. Vertenstein, W. R. Wieder, C. Xu, A. A. Ali, A. M. Badger, G. Bisht, M. van den Broeke, M. A. Brunke, S. P. Burns, J. Buzan, M. Clark, A. Craig, K. Dahlin, B. Drewniak, J. B. Fisher, M. Flanner, A. M. Fox, P. Gentine, F. M. Hoffman, G. Keppel-Aleks, R. Knox, S. Kumar, J. Lenaerts, L. R. Leung, W. H. Lipscomb, Y. Lu, A. Pandey, J. D. Pelletier, J. Perket, J. T. Randerson, D. M. Ricciuto, B. M. Sanderson, A. Slater, Z. M. Subin, J. Tang, R. Q. Thomas, M. Val Martin, and X. Zeng. The Community Land Model version 5: Description of new features, benchmarking, and impact of forcing uncertainty. *J. Adv. Model. Earth Sy.*, 11(12):4245–4287, Dec. 2019. doi:10.1029/2018MS001583.

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.

W. R. Wieder, D. M. Lawrence, R. A. Fisher, G. B. Bonan, S. J. Cheng, C. L. Goodale, A. S. Grandy, C. D. Koven, D. L. Lombardozzi, K. W. Oleson, and R. Q. Thomas. Beyond static benchmarking: Using experimental manipulations to evaluate land model assumptions. *Global Biogeochem. Cycles*, 33:1289–1309, Oct. 2019. doi:10.1029/2018GB006141.