Computational Approaches for Model, Experiment, and Data Integration Supporting Site Characterization and Model Evaluation

Forrest M. Hoffman[1,2], Jitendra Kumar[1], Zachary Langford[3], Stan D. Wullschleger[1], Damian Maddalena[4], William W. Hargrove[5], Mingquan Mu[2], William J. Riley[6], and James T. Randerson[2]

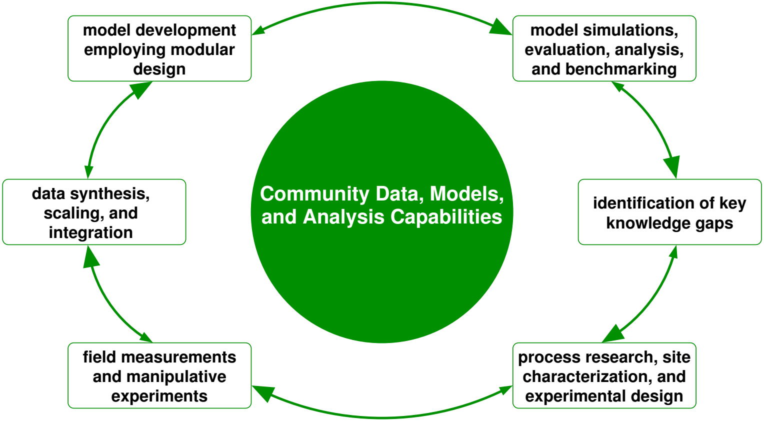
[1]Oak Ridge National Laboratory, [2]University of California–Irvine, [3]University of Tennessee–Knoxville, [4]University of North Carolina–Wilmington, [5]USDA Forest Service, and [6]Lawrence Berkeley National Laboratory

May 28, 2015
UCAR Events Center, Boulder, Colorado, USA

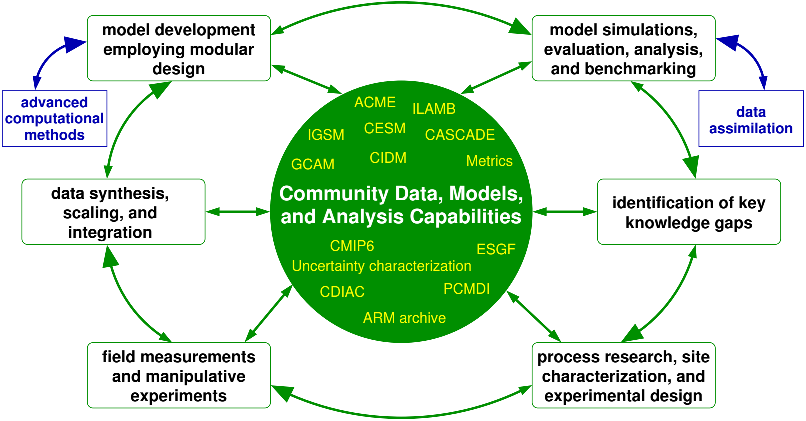
Community Surface Dynamics Modeling System (CSDMS) Annual Meeting 2015
Data meet Models, Models meet Data

CSDMS
COMMUNITY SURFACE DYNAMICS MODELING SYSTEM

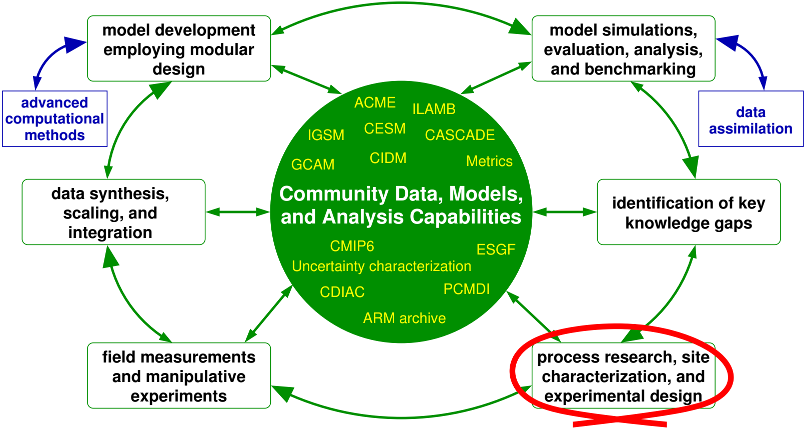
Model, Experiment, and Data Integration Strategy



Model, Experiment, and Data Integration Strategy

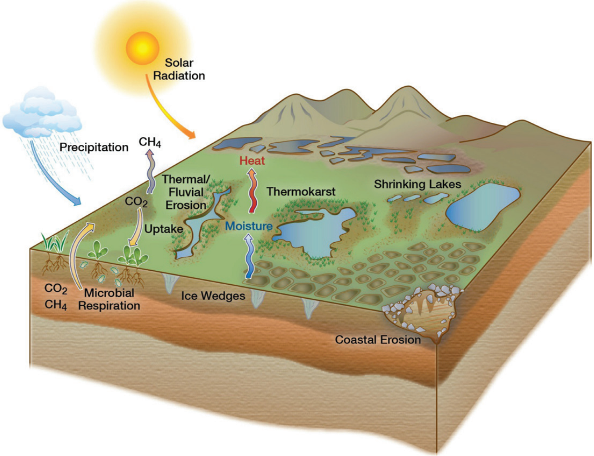


Model, Experiment, and Data Integration Strategy



Next-Generation Ecosystem Experiments (NGEE Arctic)

http://ngee.ornl.gov/

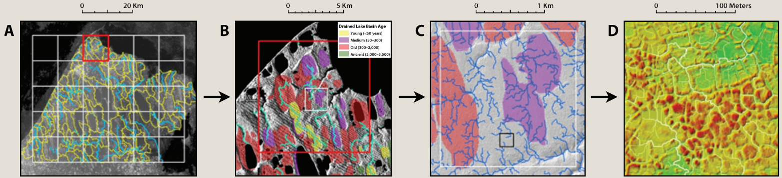


The Next-Generation Ecosystem Experiments (NGEE Arctic) project is supported by the Office of Biological and Environmental Research in the DOE Office of Science.

Integrating Across Scales

- NGEE Arctic process studies and observations are strongly linked to model development and application for improving process representation, initialization, calibration, and evaluation.
- A hierarchy of models will be deployed at fine, intermediate, and climate scales to connect observations to models and models to each other in a quantitative up-scaling and down-scaling framework.

Hydrologic and Geomorphic Features at Multiple Scales. *At the scale of (A) a high-resolution ESM, (B) a single ESM grid cell, (C) a 2 × 2 km domain of high-resolution Light Detection and Ranging (LiDAR) topographic data, and (D) polygonal ground. Yellow outlines in panel A show geomorphologically stable hydrologic basins, connected by stream channels (blue). Colored regions in panels B and C show multiple drained thaw lake basins within a single 10 × 10 km grid cell (B) or a 2 × 2 km domain (C), with progressively more detailed representation of stream channels (blue). Colors in panel D represent higher (red) to lower (green) surface elevations for a fine-scale subregion, with very fine drainage features (white). [Los Alamos National Laboratory, University of Alaska Fairbanks, and University of Texas at El Paso]*



Quantitative Sampling Network Design

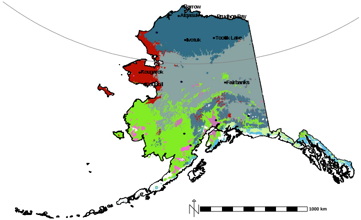
- Resource and logistical constraints limit the frequency and extent of observations, necessitating the development of a systematic sampling strategy that objectively represents environmental variability at the desired spatial scale.
- Required is a methodology that provides a quantitative framework for informing site selection and determining the representativeness of measurements.
- Multivariate spatiotemporal clustering (MSTC) was applied at the landscape scale (4 km^2) for the State of Alaska to demonstrate its utility for representativeness and scaling.
- An extension of the method applied by Hargrove and Hoffman for design of National Science Foundation's (NSF's) National Ecological Observatory Network (NEON) domains.

Data Layers

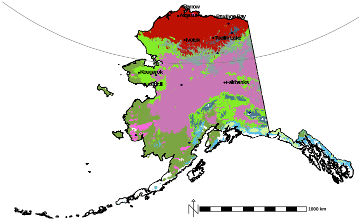
Table: 37 characteristics averaged for the present (2000–2009) and the future (2090–2099).

Description	Number/Name	Units	Source
Monthly mean air temperature	12	$^\circ$C	GCM
Monthly mean precipitation	12	mm	GCM
Day of freeze	mean	day of year	GCM
	standard deviation	days	
Day of thaw	mean	day of year	GCM
	standard deviation	days	
Length of growing season	mean	days	GCM
	standard deviation	days	
Maximum active layer thickness	1	m	GIPL
Warming effect of snow	1	$^\circ$C	GIPL
Mean annual ground temperature at bottom of active layer	1	$^\circ$C	GIPL
Mean annual ground surface temperature	1	$^\circ$C	GIPL
Thermal offset	1	$^\circ$C	GIPL
Limnicity	1	%	NHD
Elevation	1	m	SRTM

10 Alaska Ecoregions, Present and Future



2000–2009



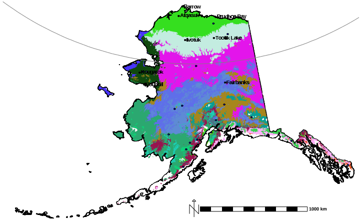
2090–2099

(Hoffman et al., 2013)

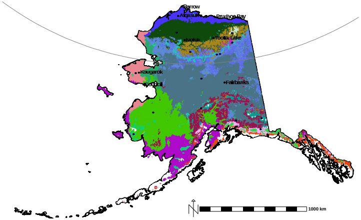
Since the random colors are the same in both maps, a change in color represents an environmental change between the present and the future.

At this level of division, the conditions in the large boreal forest become compressed onto the Brooks Range and the conditions on the Seward Peninsula “migrate” to the North Slope.

20 Alaska Ecoregions, Present and Future



2000–2009



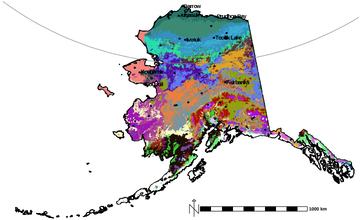
2090–2099

(Hoffman et al., 2013)

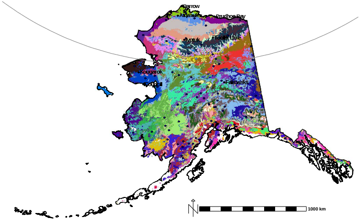
Since the random colors are the same in both maps, a change in color represents an environmental change between the present and the future.

At this level of division, the two primary regions of the Seward Peninsula and that of the northern boreal forest replace the two regions on the North Slope almost entirely.

50 and 100 Alaska Ecoregions, Present



$k = 50$, 2000–2009



$k = 100$, 2000–2009

(Hoffman et al., 2013)

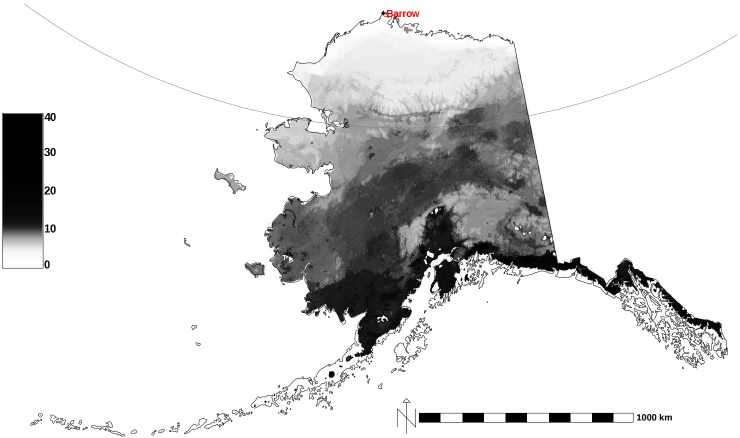
Since the random colors are the same in both maps, a change in color represents an environmental change between the present and the future.

At high levels of division, some regions vanish between the present and future while other region representing new combinations of environmental conditions come into existence.

NGEE Arctic Site Representativeness

- This representativeness analysis uses the standardized *n*-dimensional data space formed from all input data layers.
- In this data space, the Euclidean distance between a sampling location (like Barrow) and every other point is calculated.
- These data space distances are then used to generate grayscale maps showing the similarity, or lack thereof, of every location to the sampling location.
- In the subsequent maps, white areas are well represented by the sampling location or network, while dark and black areas as poorly represented by the sampling location or network.
- This analysis assumes that the climate surrogates maintain their predictive power and that no significant biological adaptation occurs in the future.

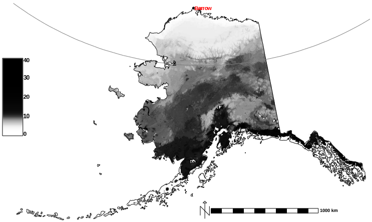
Present Representativeness of Barrow or “Barrow-ness”



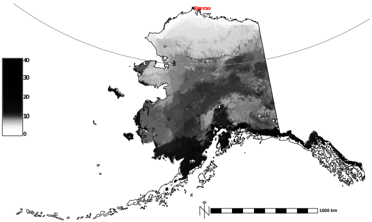
(Hoffman et al., 2013)

Light-colored regions are well represented and dark-colored regions are poorly represented by the sampling location listed in **red**.

Present vs. Future Barrow-ness



2000–2009

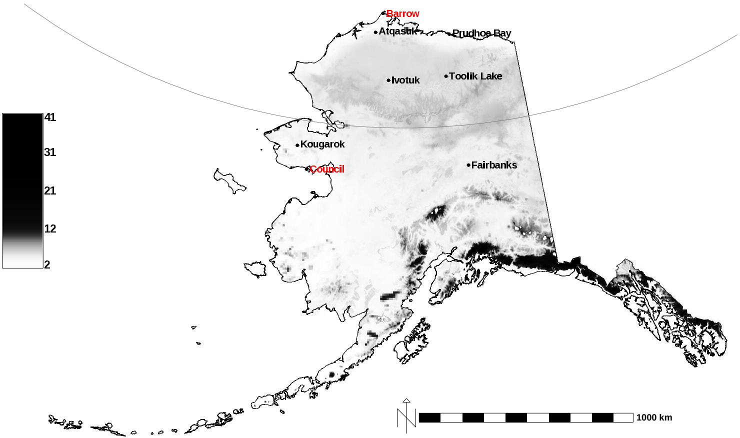


2090–2099

(Hoffman et al., 2013)

As environmental conditions change, due primarily to increasing temperatures, climate gradients shift and the representativeness of Barrow will be reduced in the future.

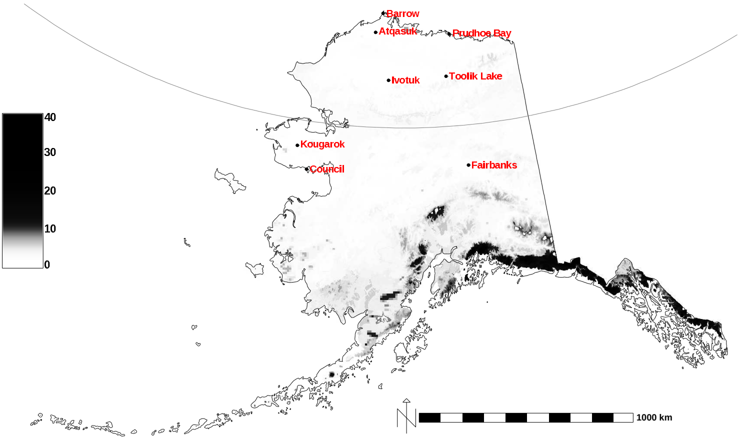
Network Representativeness: Barrow + Council



(Hoffman et al., 2013)

Light-colored regions are well represented and dark-colored regions are poorly represented by the sampling location listed in **red**.

Network Representativeness: All 8 Sites



(Hoffman et al., 2013)

Light-colored regions are well represented and dark-colored regions are poorly represented by the sampling location listed in **red**.

State Space Dissimilarities: 8 Sites, Present (2000–2009)

Table: Site state space dissimilarities for the present (2000–2009).

Sites	Council	Atqasuk	Ivotuk	Toolik Lake	Kougarok	Prudhoe Bay	Fairbanks
Barrow	9.13	4.53	5.90	5.87	7.98	3.57	12.16
Council		8.69	6.37	7.00	2.28	8.15	5.05
Atqasuk			5.18	5.23	7.79	1.74	10.66
Ivotuk				1.81	5.83	4.48	7.90
Toolik Lake					6.47	4.65	8.70
Kougarok						7.25	5.57
Prudhoe Bay							10.38

(Hoffman et al., 2013)

State Space Dissimilarities: 8 Sites, Present and Future

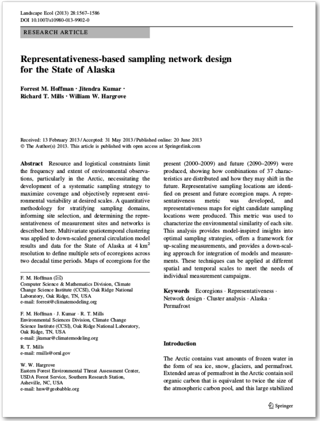
Table: Site state space dissimilarities between the present (2000–2009) and the future (2090–2099).

		Future (2090–2099)							
	Sites	Barrow	Council	Atqasuk	Ivotuk	Toolik Lake	Kougarok	Prudhoe Bay	Fairbanks
Present (2000–2009)	Barrow	3.31	9.67	4.63	6.05	5.75	9.02	3.69	11.67
	Council	8.38	1.65	8.10	5.91	6.87	3.10	7.45	5.38
	Atqasuk	6.01	9.33	2.42	5.46	5.26	8.97	2.63	10.13
	Ivotuk	7.06	7.17	5.83	1.53	2.05	7.25	4.87	7.40
	Toolik Lake	7.19	7.67	6.07	2.48	1.25	7.70	5.23	8.16
	Kougarok	7.29	3.05	6.92	5.57	6.31	2.51	6.54	5.75
	Prudhoe Bay	5.29	8.80	3.07	4.75	4.69	8.48	1.94	9.81
	Fairbanks	12.02	5.49	10.36	7.83	8.74	6.24	10.10	1.96

(Hoffman et al., 2013)

Representativeness: A Quantitative Approach for Scaling

- MSTC provides a quantitative framework for stratifying sampling domains, informing site selection, and determining representativeness of measurements.
- Representativeness analysis provides a systematic approach for up-scaling point measurements to larger domains.


Landscape Ecol (2013) 28:1567–1586
DOI 10.1007/s10980-013-9902-0

RESEARCH ARTICLE

Representativeness-based sampling network design for the State of Alaska

Forrest M. Hoffman · Jitendra Kumar · Richard T. Mills · William W. Hargrove

Received: 13 February 2013 / Accepted: 31 May 2013 / Published online: 20 June 2013

Abstract Resource and logistical constraints limit the frequency and extent of environmental observations, particularly in the Arctic, necessitating the development of a systematic sampling strategy to maximize coverage and objectively represent environmental variability at desired scales. A quantitative methodology for stratifying sampling domains, informing site selection, and determining the representativeness of measurement sites and networks is described here. Multivariate spatiotemporal clustering was applied to down-scaled general circulation model results and data for the State of Alaska at 4 km^2 resolution to define multiple sets of ecoregions across two decadal time periods. Maps of ecoregions for the present (2000–2009) and future (2090–2099) were produced, showing how combinations of 37 characteristics are distributed and how they may shift in the future. Representative sampling locations are identified on present and future ecoregion maps. A representativeness metric was developed, and representativeness maps for eight candidate sampling locations were produced. This metric was used to characterize the environmental similarity of each site. This analysis provides model-inspired insights into optimal sampling strategies, offers a framework for up-scaling measurements, and provides a down-scaling approach for integration of models and measurements. These techniques can be applied at different spatial and temporal scales to meet the needs of individual measurement campaigns.

Keywords Ecoregions · Representativeness · Network design · Cluster analysis · Alaska · Permafrost

F. M. Hoffman (✉)
Computer Science & Mathematics Division, Climate Change Science Institute (CCSI), Oak Ridge National Laboratory, Oak Ridge, TN, USA
e-mail: forrest@climatemodeling.org

F. M. Hoffman · J. Kumar · R. T. Mills
Environmental Sciences Division, Climate Change Science Institute (CCSI), Oak Ridge National Laboratory, Oak Ridge, TN, USA
e-mail: jkumar@climatemodeling.org

R. T. Mills
e-mail: rmills@ornl.gov

W. W. Hargrove
Eastern Forest Environmental Threat Assessment Center, USDA Forest Service, Southern Research Station, Asheville, NC, USA
e-mail: hnw@geobabble.org

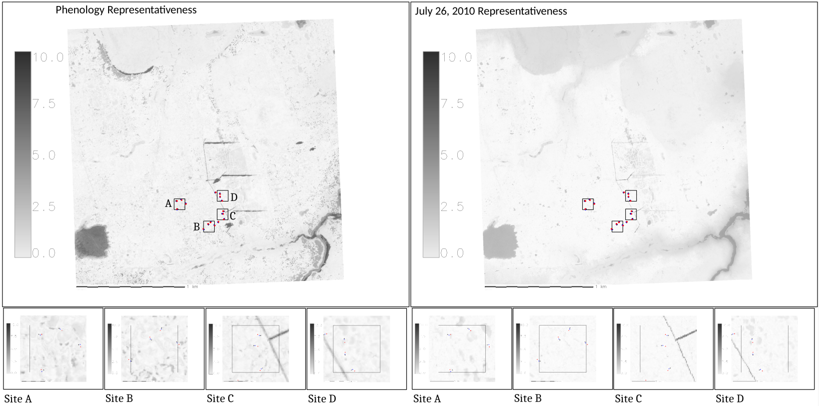
Introduction

The Arctic contains vast amounts of frozen water in the form of sea ice, snow, glaciers, and permafrost. Extended areas of permafrost in the Arctic contain soil organic carbon that is equivalent to twice the size of the atmospheric carbon pool, and this large stabilized

Hoffman, F. M., J. Kumar, R. T. Mills, and W. W. Hargrove (2013), "Representativeness-Based Sampling Network Design for the State of Alaska." *Landscape Ecol.*, 28(8):1567–1586. doi:10.1007/s10980-013-9902-0.

Received US-IALE's 2014 Outstanding Paper in Landscape Ecology Award!

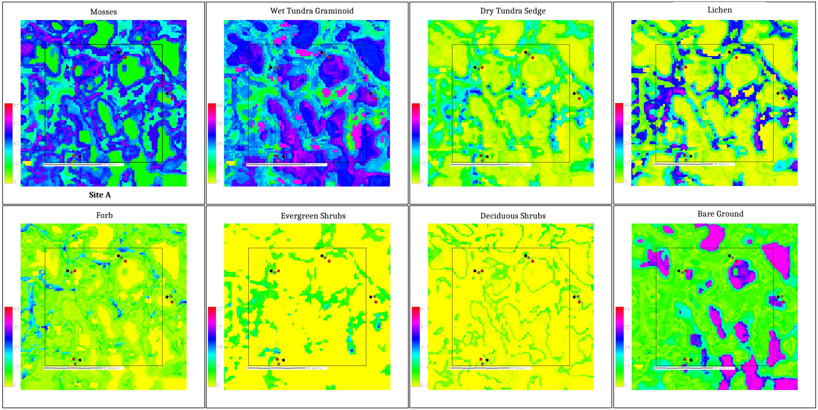
Barrow Environmental Observatory (BEO)



(Langford et al., in prep)

Representativeness map for vegetation sampling points in A, B, C, and D sampling area with phenology (left) and without (right), based on WorldView2 satellite images for the year 2010 and LiDAR data.

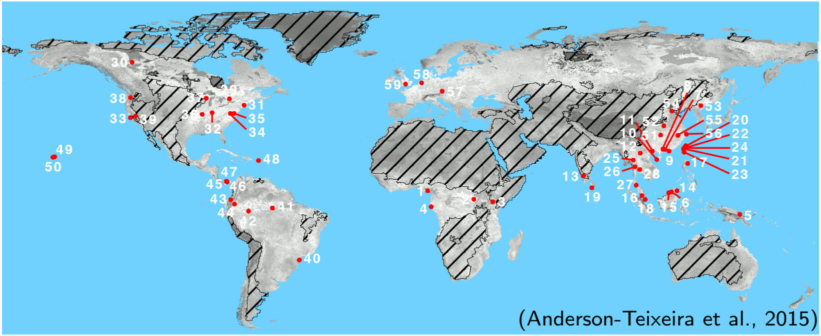
Barrow Environmental Observatory (BEO)



(Langford et al., in prep)

Example plant functional type (PFT) distributions scaled up from vegetation sampling locations.

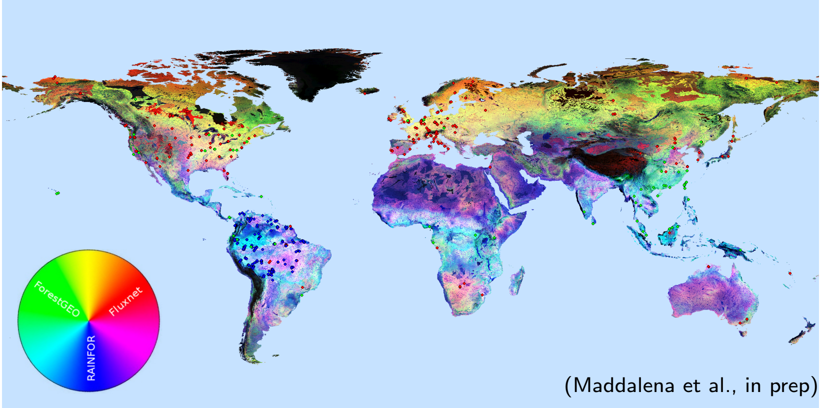
ForestGEO Network Global Representativeness



(Anderson-Teixeira et al., 2015)

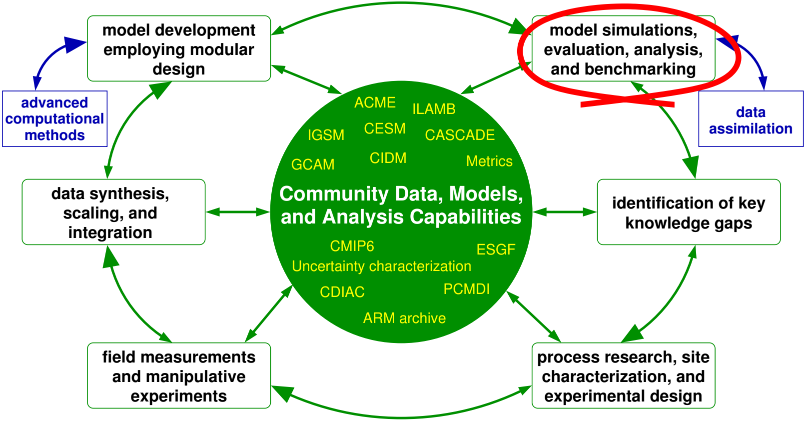
Map illustrating ForestGEO network representation of 17 bioclimatic, edaphic, and topographic conditions globally. Light-colored regions are well represented and dark-colored regions are poorly represented by the ForestGEO sampling network. Stippling covers non-forest areas.

Triple-Network Global Representativeness



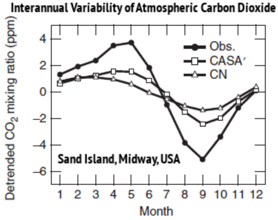
Map indicates which sampling network offers the most representative coverage at any location. Every location is made up of a combination of three primary colors from Fluxnet (red), ForestGEO (green), and RAINFOR (blue).

Model, Experiment, and Data Integration Strategy

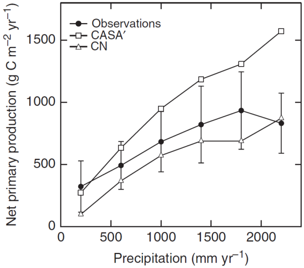


What is a Benchmark?

- 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.



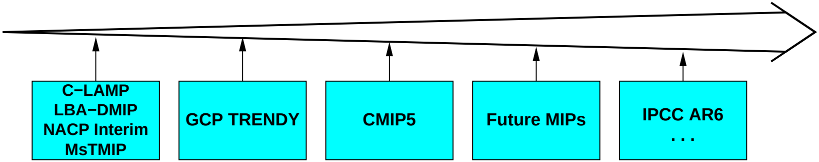
Models may reproduce correct responses over only a limited range of forcing variables.

(Randerson et al., 2009)

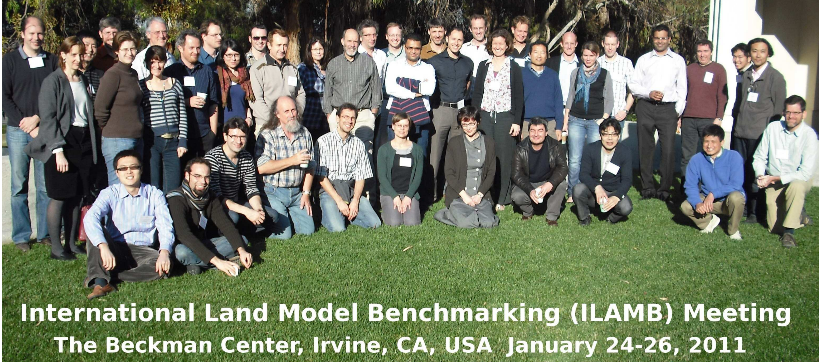
Why Benchmark?

- 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

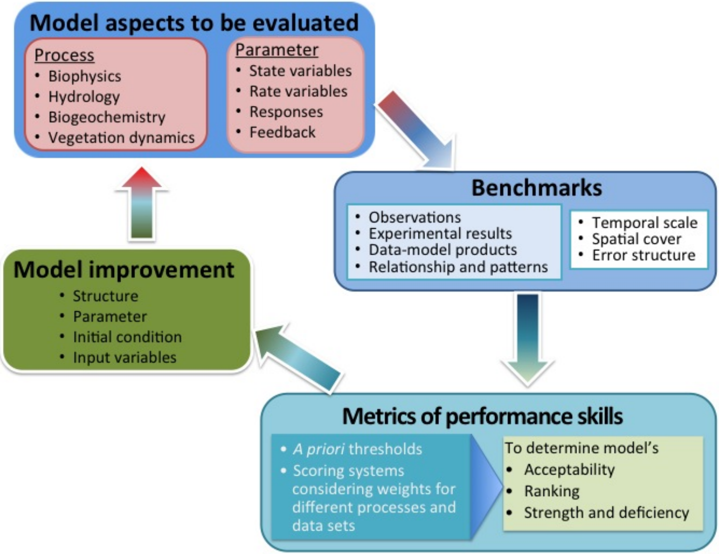


- Human capital costs of making rigorous model-data comparisons is considerable and constrains the scope of individual MIPs.
- Many MIPs spend resources "reinventing the wheel" in terms of variable naming conventions, model simulation protocols, and analysis software.
- **Need for ILAMB:** Each new MIP 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)

ILAMB 1.0 Benchmarks

	Annual Mean	Seasonal Cycle	Interannual Variability	Trend	Data Source
Atmospheric CO_2					
Flask/conc. + transport		✓	✓	✓	NOAA, SIO, CSIRO
TCCON + transport		✓	✓	✓	Caltech
Fluxnet					
GPP, NEE, TER, LE, H, RN	✓	✓	✓		Fluxnet, MAST-DC
Gridded: GPP	✓	✓	?		MPI-BGC
Hydrology/Energy					
runoff ratio (R/P) ~~river flow~~	✓		✓		GRDC, Dai, GFDL
global runoff/ocean balance	✓				Syed/Famiglietti
albedo (multi-band)		✓	✓		MODIS, CERES
soil moisture		✓	✓		de Jeur, SMAP
column water		✓	✓		GRACE
snow cover	✓	✓	✓	✓	AVHRR, GlobSnow
snow depth/SWE	✓	✓	✓	✓	CMC (N. America)
T_{air} & P	✓	✓	✓	✓	CRU, GPCP and TRMM
Gridded: LE, H	✓	✓			MPI-BGC, dedicated ET
Ecosystem Processes & State					
soil C, N	✓				HWSD, MPI-BGC
litter C, N	✓				LIDET
soil respiration	✓	✓	✓	✓	Bond-Lamberty
FAPAR	✓	✓			MODIS, SeaWIFS
biomass & change	✓			✓	Saatchi, Pan, Blackard
canopy height	✓				Lefsky, Fisher
NPP	✓				EMDI, Luyssaert
Vegetation Dynamics					
fire — burned area	✓	✓	✓		GFED3
wood harvest	✓			✓	Hurtt
land cover	✓				MODIS PFT fraction

Example Benchmark Score Sheet from C-LAMP

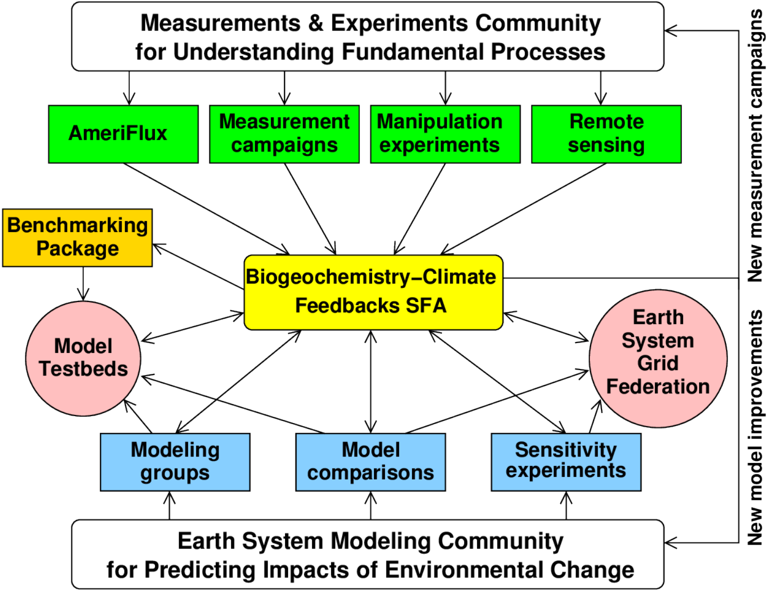
Models

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]() for Weightings)

	MeanModel	bcc-csm1-1-m	BNU-ESM	CanESM2	CESM1-BGC	GFDL-ESM2G	HadGEM2-ES	inmcm4	IPSL-CM5A-LR	MIROC-ESM	MPI-ESM-LR	MRI-ESM1	NorESM1-M
Aboveground Live Biomass	0.69	0.55	0.43	0.68	0.64	0.57	0.69	0.62	0.71	0.57	0.55	0.62	0.64
Gross Primary Productivity	0.77	0.73	0.74	0.65	0.72	0.65	0.71	0.70	0.66	0.68	0.67	0.53	0.71
Burned Area	0.56	-	-	-	0.56	-	-	-	-	-	0.55	-	0.56
Carbon Dioxide	0.94	-	0.88	0.88	0.90	0.96	-	-	-	-	-	0.88	-
Leaf Area Index	0.65	0.63	0.45	0.64	0.57	0.49	0.61	0.67	0.65	0.58	0.65	0.48	0.53
Global Net Ecosystem Carbon Balance	0.52	-	0.18	0.25	0.36	0.20	0.30	0.22	0.28	0.35	0.38	0.16	0.33
Net Ecosystem Exchange	0.50	0.50	0.47	0.41	0.50	0.44	0.50	0.45	0.52	0.48	0.48	0.50	0.49
Ecosystem Respiration	0.74	0.72	0.72	0.68	0.68	0.71	0.69	0.65	0.65	0.66	0.66	0.44	0.68
Soil Carbon	0.51	0.49	0.43	0.57	0.37	0.53	0.50	0.51	0.53	0.51	0.42	0.52	0.38
Summary	0.65	0.60	0.53	0.59	0.59	0.57	0.56	0.53	0.56	0.54	0.54	0.52	0.54
Evapotranspiration	0.75	0.74	0.72	0.73	0.72	0.71	0.73	0.69	0.75	0.69	0.73	0.74	0.71
Latent Heat	0.75	0.73	0.70	0.72	0.71	0.69	0.73	0.68	0.75	0.69	0.72	0.74	0.70
Terrestrial Water Storage Anomaly	0.61	0.54	0.44	0.60	0.59	0.55	0.51	0.58	0.56	0.58	0.61	0.58	0.57
Summary	0.70	0.67	0.62	0.68	0.67	0.65	0.66	0.65	0.69	0.65	0.69	0.69	0.66
Albedo	0.74	0.72	0.64	0.72	0.73	0.71	0.76	0.69	0.71	0.68	0.74	0.67	0.71
Surface Upward SW Radiation	0.79	0.74	0.69	0.76	0.78	0.77	0.78	0.75	0.75	0.73	0.76	0.69	0.76
Surface Net SW Radiation	0.86	0.87	0.85	0.86	0.87	0.87	0.87	0.85	0.83	0.84	0.88	0.86	0.87
Surface Upward LW Radiation	0.90	0.92	0.92	0.91	0.92	0.92	0.92	0.89	0.90	0.90	0.92	0.91	0.92
Surface Net LW Radiation	0.82	0.82	0.81	0.80	0.81	0.82	0.83	0.80	0.78	0.79	0.81	0.82	0.81
Surface Net Radiation	0.80	0.80	0.77	0.80	0.81	0.81	0.80	0.75	0.79	0.77	0.82	0.79	0.81
Sensible Heat	0.71	0.64	0.66	0.65	0.69	0.66	0.68	0.63	0.69	0.65	0.68	0.64	0.70
Summary	0.79	0.77	0.75	0.77	0.79	0.78	0.79	0.75	0.77	0.75	0.79	0.76	0.79
Surface Air Temperature	0.89	0.88	0.89	0.88	0.90	0.88	0.89	0.86	0.88	0.88	0.90	0.90	0.88
Precipitation	0.77	0.75	0.72	0.75	0.77	0.74	0.79	0.75	0.74	0.74	0.78	0.77	0.74
Surface Downward SW Radiation	0.88	0.89	0.89	0.89	0.89	0.89	0.89	0.88	0.84	0.88	0.89	0.88	0.89
Surface Downward LW Radiation	0.91	0.92	0.92	0.92	0.92	0.92	0.93	0.91	0.89	0.91	0.92	0.92	0.92

ILAMB Prototype Layout: Global Variables

Global Variables (Info for Weightings)

	MeanModel	bcc-csm1-1-m	BNU-ESM	CanESM2	CESM1-BGC	GFDL-ESM2G	HadGE
Aboveground Live Biomass	0.69	0.55	0.43	0.68	0.64	0.57	0.
Gross Primary Productivity	0.77	0.73	0.74	0.65	0.72	0.65	0.
Burned Area	0.56	-	-	-	0.56	-	-
Carbon Dioxide	0.94	-	0.88	0.88	0.90	0.96	-
Leaf Area Index	0.65	0.63	0.45	0.64	0.57	0.49	0.
Global Net Ecosystem Carbon Balance	0.52	-	0.18	0.25	0.36	0.20	0.
Net Ecosystem Exchange	0.50	0.50	0.47	0.41	0.50	0.44	0.
Ecosystem Respiration	0.74	0.72	0.72	0.68	0.68	0.71	0.
Soil Carbon	0.51	0.49	0.43	0.57	0.37	0.53	0.
Summary	0.65	0.60	0.53	0.59	0.59	0.57	0.
Evapotranspiration	0.75	0.74	0.72	0.73	0.72	0.71	0.
Latent Heat	0.75	0.73	0.70	0.72	0.71	0.69	0.
Terrestrial Water Storage Anomaly	0.61	0.54	0.44	0.60	0.59	0.55	0.
Summary	0.70	0.67	0.62	0.68	0.67	0.65	0.
Albedo	0.74	0.72	0.64	0.72	0.73	0.71	0.

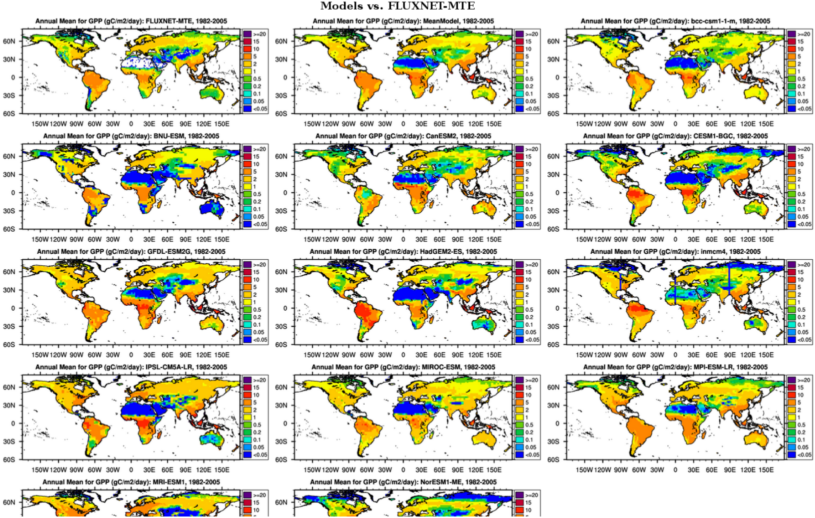
ILAMB: Gross Primary Productivity (GPP) Diagnostics

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

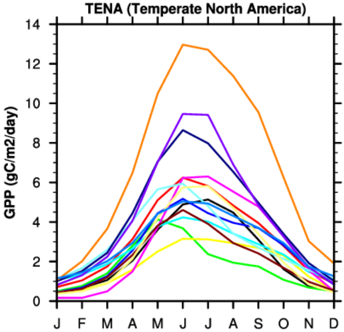
	Global Patterns				Regional Patterns	Scoring (Info)					
	Annual Mean (PgC/yr)	Bias (PgC/yr)	RMSE (PgC/mon)	Phase Difference (months)	Regional Mean	Global Bias	RMSE	Seasonal Cycle	Spatial Distribution	Interannual Variability	Overall
Benchmark [Jung et al. (2010)]	118.4	-	-	0.0	access to plots	-	-	-	-	-	-
MeanModel	170.6	52.2	4.7	0.2	access to plots	0.75	0.70	0.85	0.94	-	0.79
bcc-csm1-1-m	122.4	4.0	6.0	-0.3	access to plots	0.72	0.65	0.82	0.90	-	0.75
BNU-ESM	145.4	27.0	6.0	-0.3	access to plots	0.71	0.66	0.82	0.86	-	0.74
CanESM2	129.7	11.3	7.5	0.6	access to plots	0.66	0.60	0.70	0.70	-	0.65
CESM1-BGC	147.5	29.1	5.9	0.0	access to plots	0.71	0.66	0.81	0.85	-	0.74
GFDL-ESM2G	222.5	104.1	10.0	0.3	access to plots	0.62	0.51	0.81	0.82	-	0.65
HadGEM2-ES	166.0	47.6	7.2	0.2	access to plots	0.70	0.62	0.84	0.82	-	0.72
inmcm4	136.9	18.5	5.7	-0.1	access to plots	0.72	0.64	0.80	0.91	-	0.74
IPSL-CM5A-LR	223.0	104.6	9.3	0.1	access to plots	0.60	0.52	0.82	0.84	-	0.66
MIROC-ESM	129.2	10.8	6.3	0.0	access to plots	0.68	0.61	0.81	0.85	-	0.71
MPI-ESM-LR	169.7	51.3	7.7	0.2	access to plots	0.60	0.55	0.79	0.88	-	0.67
MRI-ESM1	289.6	171.1	14.2	0.1	access to plots	0.50	0.45	0.81	0.53	-	0.55
NorESM1-ME	165.8	47.4	6.7	-0.1	access to plots	0.68	0.64	0.80	0.81	-	0.71

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

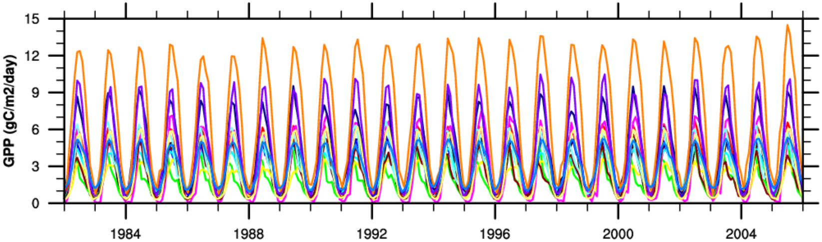
ILAMB: GPP Annual Mean Diagnostics



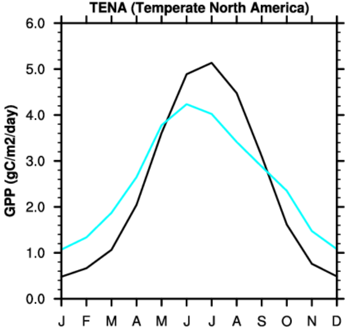
ILAMB: GPP Temperate North America Diagnostics



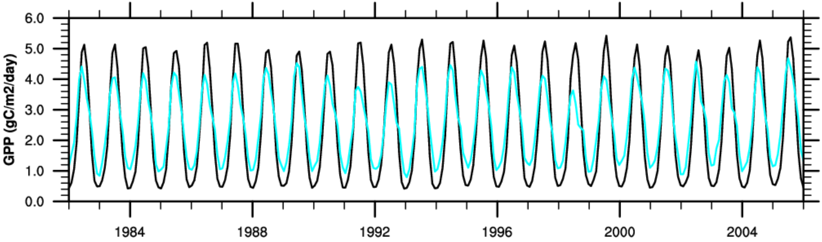
Model	Annual	Bias	RMSE
FLUXNET-MTE	2.36	-999.00	-999.00
MeanModel	3.15	0.79	0.88
bcc-csm1-1-m	1.80	-0.56	1.24
BNU-ESM	2.67	0.31	0.62
CanESM2	1.83	-0.53	0.99
CESM1-BGC	2.51	0.15	0.71
GFDL-ESM2G	2.85	0.49	0.98
HadGEM2-ES	2.12	-0.24	0.62
inmcm4	3.06	0.70	1.15
IPSL-CM5A-LR	4.24	1.88	2.14
MIROC-ESM	2.58	0.22	0.40
MPI-ESM-LR	4.33	1.97	2.41
MRI-ESM1	6.79	4.43	5.14
NorESM1-ME	2.99	0.63	0.77



ILAMB: CESM1-BGC GPP for Temperate North America



Model	Annual	Bias	RMSE
FLUXNET-MTE	2.36	-999.00	-999.00
MeanModel	3.15	0.79	0.88
bcc-csm1-1-m	1.80	-0.56	1.24
BNU-ESM	2.67	0.31	0.62
CanESM2	1.83	-0.53	0.99
CESM1-BGC	2.51	0.15	0.71
GFDL-ESM2G	2.85	0.49	0.98
HadGEM2-ES	2.12	-0.24	0.62
inmcm4	3.06	0.70	1.15
IPSL-CM5A-LR	4.24	1.88	2.14
MIROC-ESM	2.58	0.22	0.40
MPI-ESM-LR	4.33	1.97	2.41
MRI-ESM1	6.79	4.43	5.14
NorESM1-ME	2.99	0.63	0.77



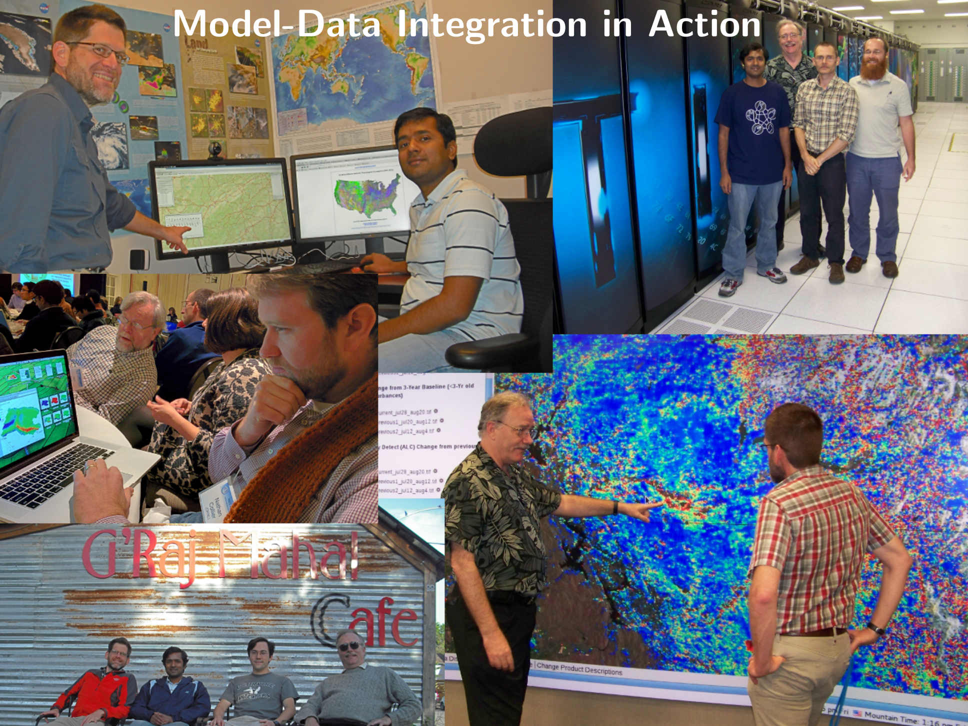
Next Steps for ILAMB Development

- ILAMB information is available at `http://www.ilamb.org/`
- The ILAMB prototype, based on the NCAR Command Language (NCL), is available at `http://redwood.ess.uci.edu/mingquan/www/ILAMB/`
- A next generation ILAMB system is under development in Python (NumPy, SciPy, Matplotlib + Basemap)
- Development of the next generation system is ongoing using a GitHub repository with documentation in Sphinx
- A community meeting is being planned for this winter
- Information about the DOE-sponsored Biogeochemistry–Climate Feedbacks Project is available at `http://www.bgc-feedbacks.org/`

Take Home Message

- **Modelers:** Confront models with data. *Just like voting, do this early and often!*
  - Make model evaluation tools and data free and open, facilitating community contributions. *It takes a village!*
  - Design model experiments and analyses to identify weaknesses and inspire new measurements.
- **Data Gatherers:** Make data available early and characterize and report <u>all</u> measurement uncertainties.
  - Confront the environment with new sensors, drones, and aerial and space-based instrumentation to answer key questions about mechanisms.
  - Conduct measurements to improve our understanding of processes and inform model development.
- **Integrated Assessors:** Creatively employ multi-model projections and use results of model evaluation as a lens through which to view predictions of the future.

Model-Data Integration in Action



References

F. M. Hoffman, J. Kumar, R. T. Mills, and W. W. Hargrove. Representativeness-based sampling network design for the State of Alaska. *Landscape Ecol.*, 28(8):1567–1586, Oct. 2013. doi: 10.1007/s10980-013-9902-0.

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.

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.

Acknowledgements

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, the U. S. Department of Agriculture Forest Service, and the National Science Foundation (AGS-1048890). This research used resources of the National Center for Computational Sciences (NCCS) at Oak Ridge National Laboratory (ORNL), which is managed by UT-Battelle, LLC, for the U. S. Department of Energy under Contract No. DE-AC05-00OR22725.