Integrating Earth Science Research through Model, Experiment, and Data Synthesis

Forrest M. Hoffman[1,2], Jitendra Kumar[1], William W. Hargrove[3], Zachary Langford[1], Damian Maddalena[1], William J. Riley[4], and James T. Randerson[2]

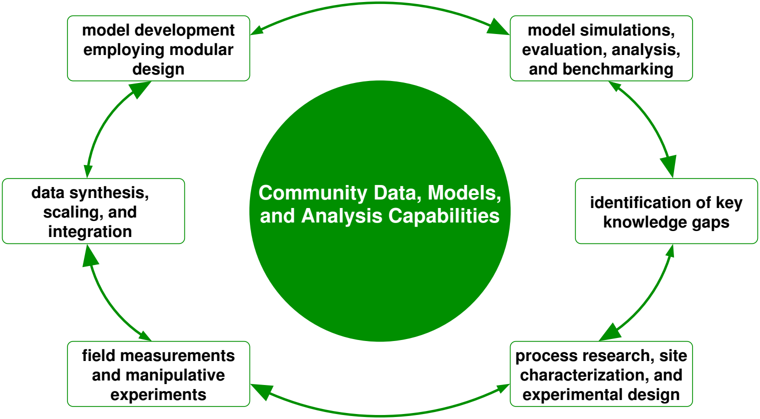
[1]Oak Ridge National Laboratory, [2]University of California Irvine, [3]USDA Forest Service, and [4]Lawrence Berkeley National Laboratory

October 29, 2014

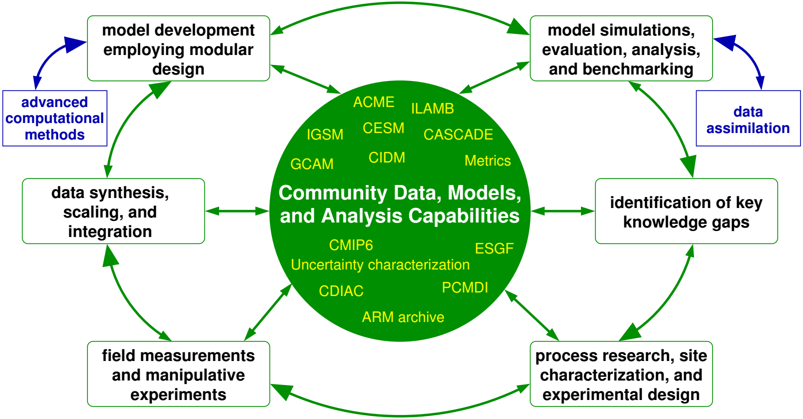
Green Ocean Amazon Joint Principal Investigators Meeting

Woodrow Wilson International Center for Scholars, Washington, DC, USA

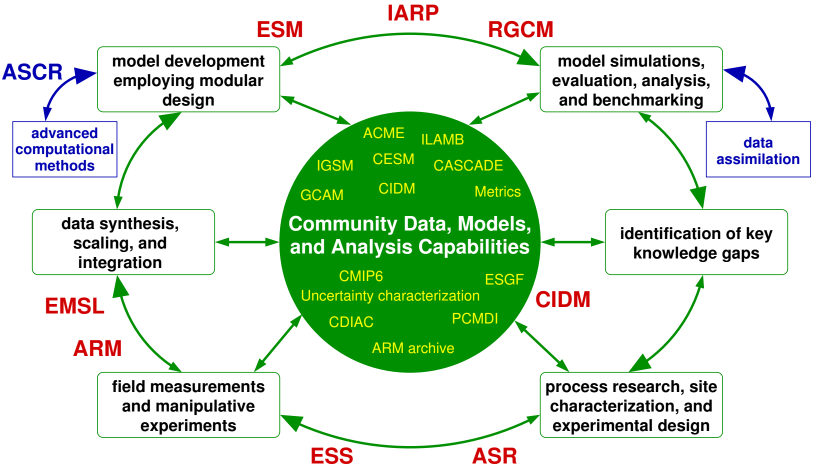
Model, Experiment, and Data Integration Strategy



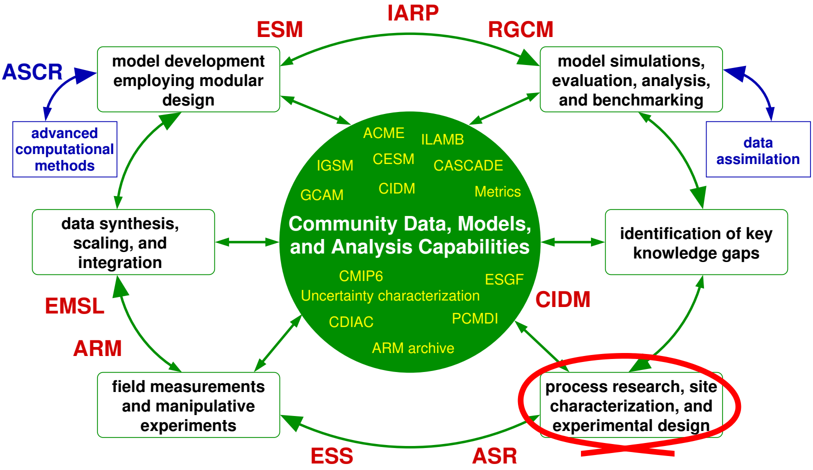
Model, Experiment, and Data Integration Strategy



Model, Experiment, and Data Integration Strategy



Model, Experiment, and Data Integration Strategy



Quantitative Sampling Network Design

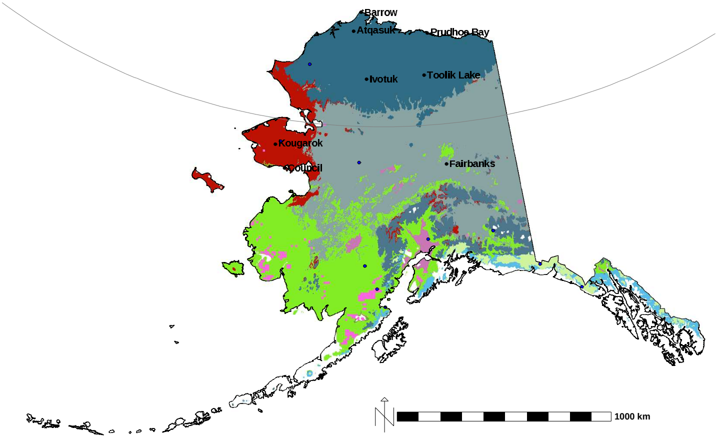
- Resource and logistical constraints limit the frequency and extent of observations, necessitating the development of a systematic sampling strategy that objectively represents environmental variability at the desired spatial scale.
- Required is a methodology that provides a quantitative framework for informing site selection and determining the representativeness of measurements.
- Multivariate spatiotemporal clustering (MSTC) was applied at the landscape scale ($4\ \mathrm{km}^2$) 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).

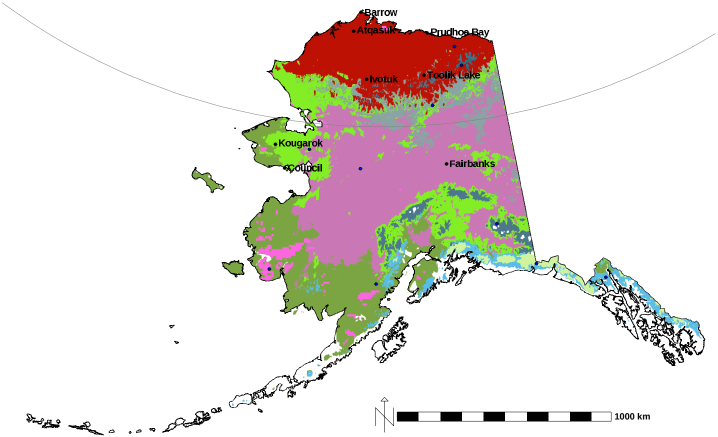
Description	Number/Name	Units	Source
Monthly mean air temperature	12	$^\circ$C	GCM
Monthly mean precipitation	12	mm	GCM
Day of freeze	mean	day of year	GCM
	standard deviation	days	
Day of thaw	mean	day of year	GCM
	standard deviation	days	
Length of growing season	mean	days	GCM
	standard deviation	days	
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 (2000–2009)



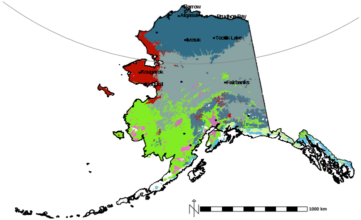
Each ecoregion is a different random color. Blue filled circles mark locations most representative of mean conditions of each region.

10 Alaska Ecoregions (2090–2099)

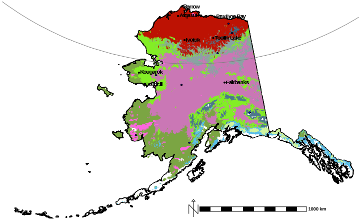


Each ecoregion is a different random color. Blue filled circles mark locations most representative of mean conditions of each region.

10 Alaska Ecoregions, Present and Future



2000–2009



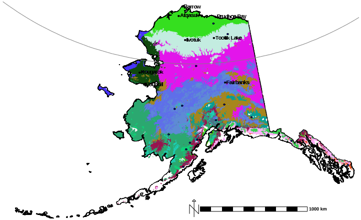
2090–2099

(Hoffman et al., 2013)

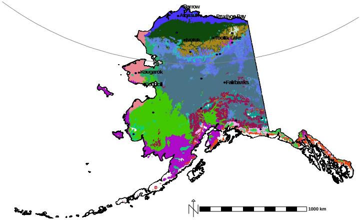
Since the random colors are the same in both maps, a change in color represents an environmental change between the present and the future.

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)

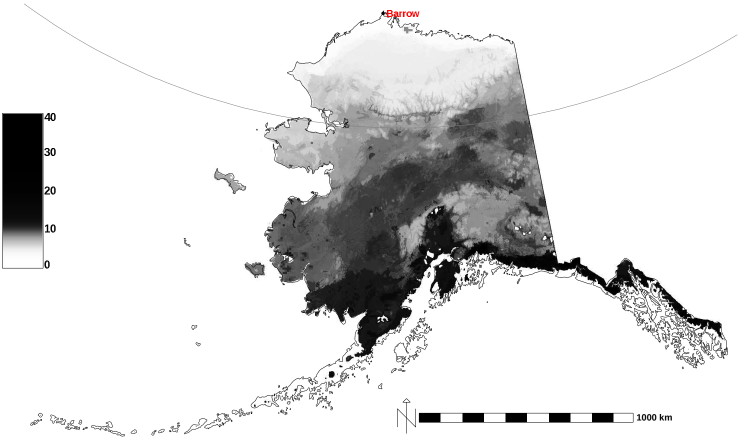
Since the random colors are the same in both maps, a change in color represents an environmental change between the present and the future.

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.

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.

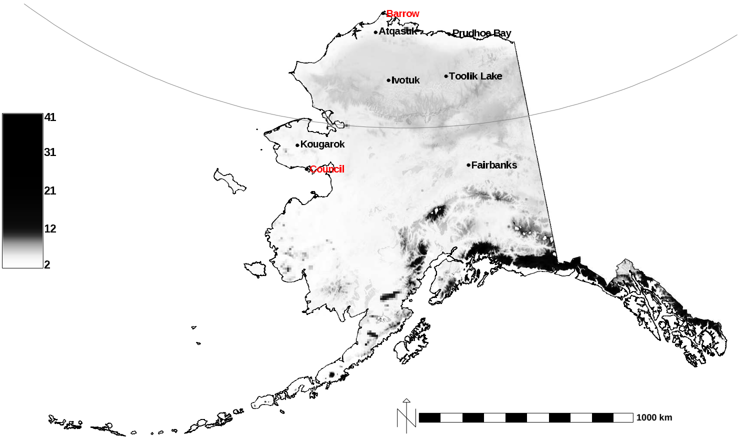
Present Representativeness of Barrow or “Barrow-ness”



(Hoffman et al., 2013)

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

Network Representativeness: Barrow + Council



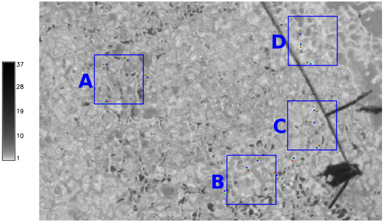
(Hoffman et al., 2013)

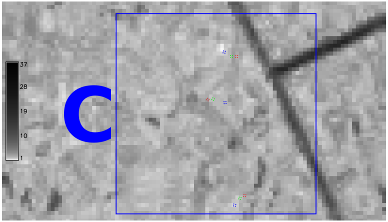
Light-colored regions are well represented and dark-colored regions are poorly represented by the sampling location listed in **red**.

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.
- Methodology is independent of resolution, thus can be applied from site/plot scale to landscape/climate scale.
- It can be extended to include finer spatiotemporal scales, more geophysical characteristics, and remote sensing data.
- Paper describing the methodology:
  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 2014 Outstanding Paper in Landscape Ecology Award!

Barrow Environmental Observatory (BEO)

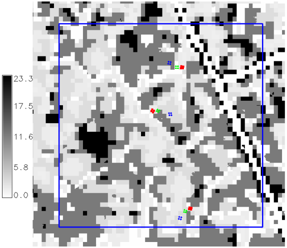




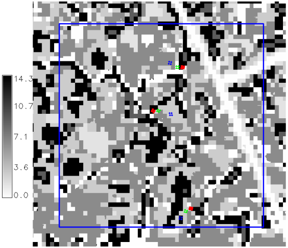
(Kumar et al., in prep)

Representativeness map for vegetation sampling points for A, B, C, and D sampling area (left) and zoomed in on the C samping area (right) developed from WorldView2 satellite images for the year 2010 and LiDAR data.

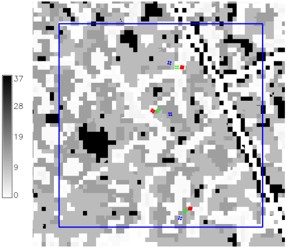
Vegetation sampling locations represent polygon troughs (red), edges (green), and centers (blue).



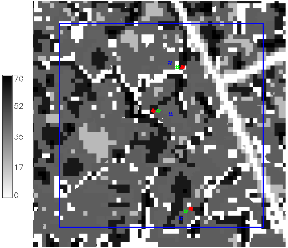
(a) dry tundra gramanoid



(b) forb



(c) lichen

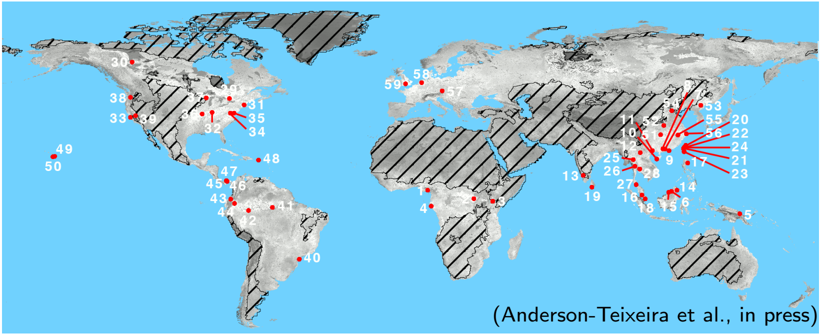


(d) moss

(Kumar et al., in prep)

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

ForestGEO Network Global Representativeness

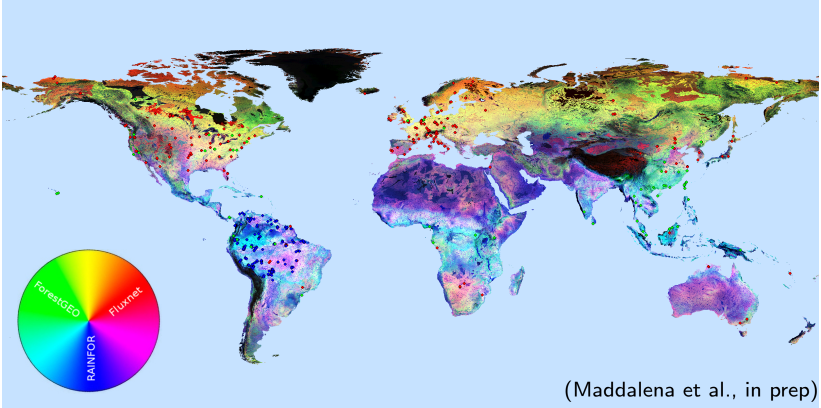


(Anderson-Teixeira et al., in press)

Light-colored regions are well represented and dark-colored regions are poorly represented by the ForestGEO sampling network.

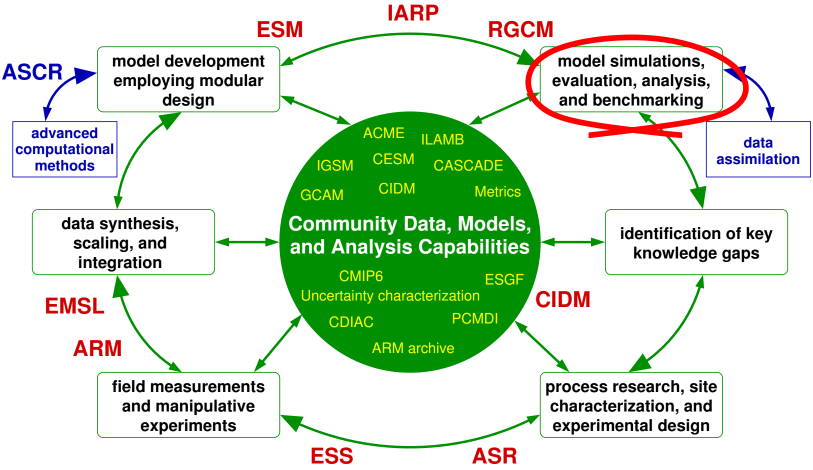
Animation of the time evolution of the ForestGEO network: https://climate.ornl.gov/∼jkumar/share/forestGEOall_years.gif

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).

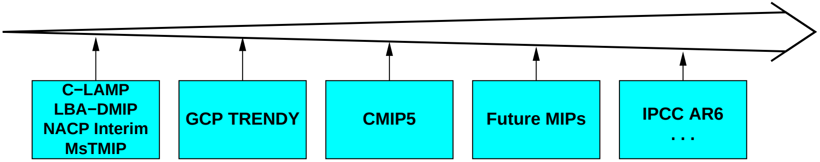
Model, Experiment, and Data Integration Strategy



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);

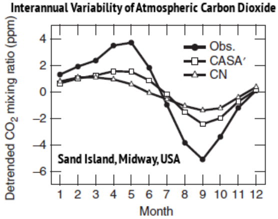
An Open Source Benchmarking Software System



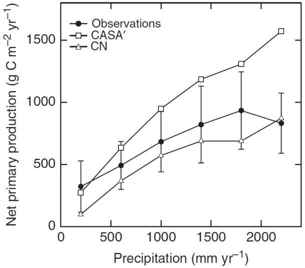
- Human capital costs of making rigorous model-data comparisons is considerable and constrains the scope of individual MIPs.
- Many MIPs spend resources "reinventing the wheel" in terms of variable naming conventions, model simulation protocols, and analysis software.
- **Need for ILAMB:** Each new MIP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

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**.

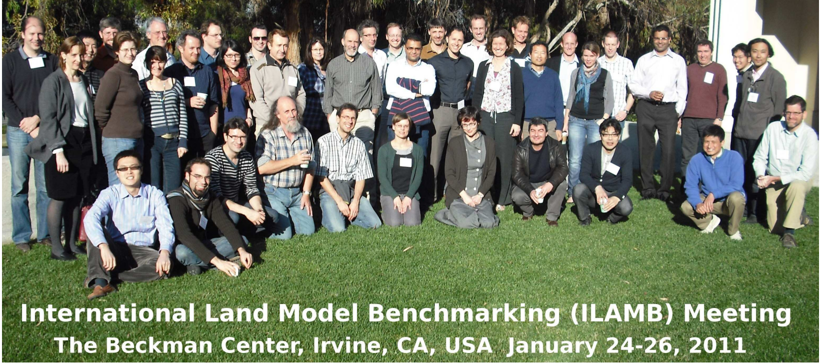


Models often fail to capture the amplitude of the seasonal cycle of atmospheric CO_2.



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

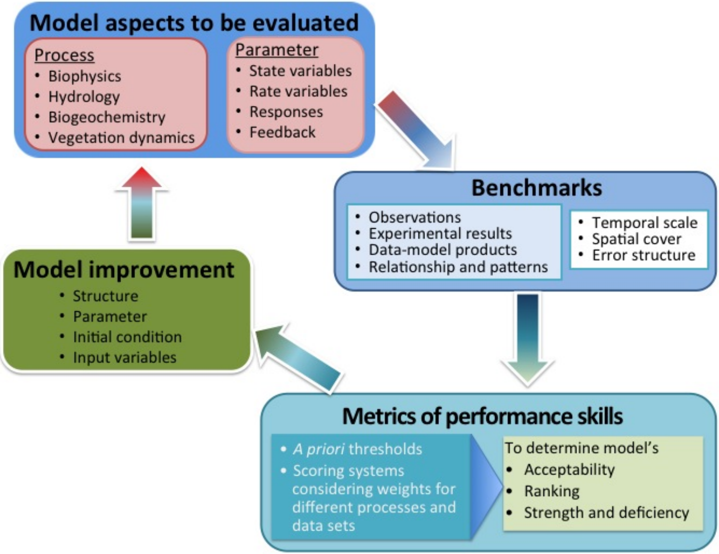
(Randerson et al., 2009)



NASA AIMES GLOBAL IGBP CHANGE International Geosphere-Biosphere Programme Global Carbon Project OAK RIDGE National Laboratory 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. *Initial focus on CMIP5 models.*
- Provides methodology for model–data comparison and baseline standard for performance of land model process representations (Luo et al., 2012).

General Benchmarking Procedure



(Luo et al., 2012)

Example Benchmark Score Sheet from C-LAMP

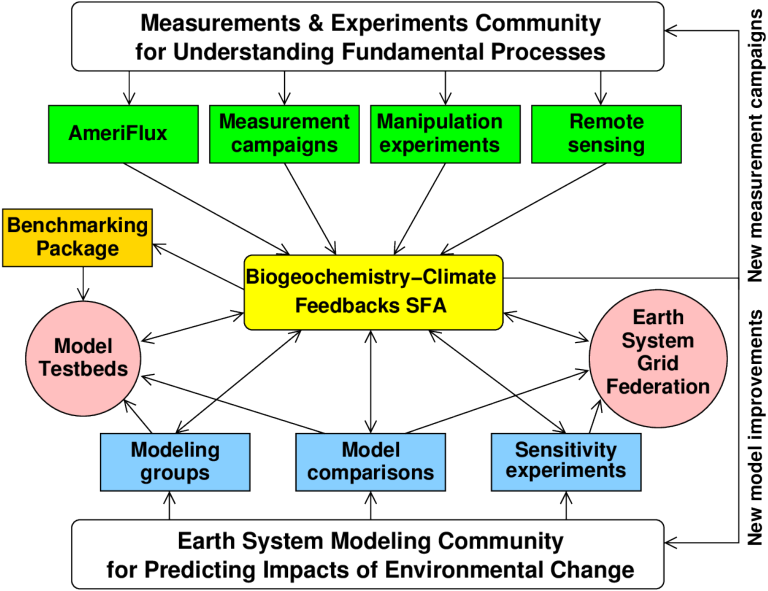
Models →

BGC Datasets ↓

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

(Randerson et al., 2009)

Biogeochemistry–Climate Feedbacks Scientific Focus Area



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.

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

FOREST SERVICE U S DEPARTMENT OF AGRICULTURE

NSF

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.

We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we 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.