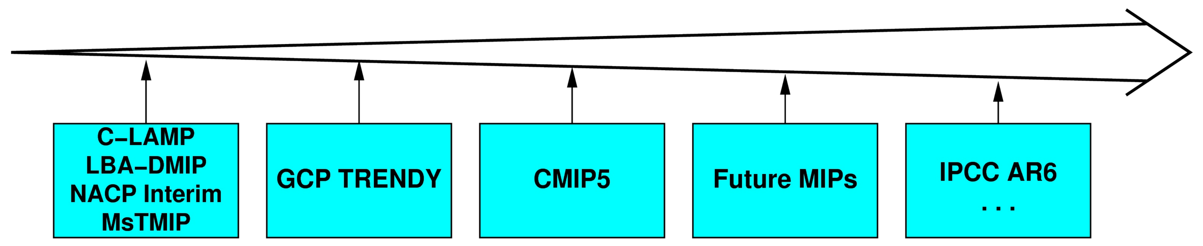
The International Land Model Benchmarking (ILAMB) Project

Forrest M. Hoffman[1,2] and James T. Randerson[2]

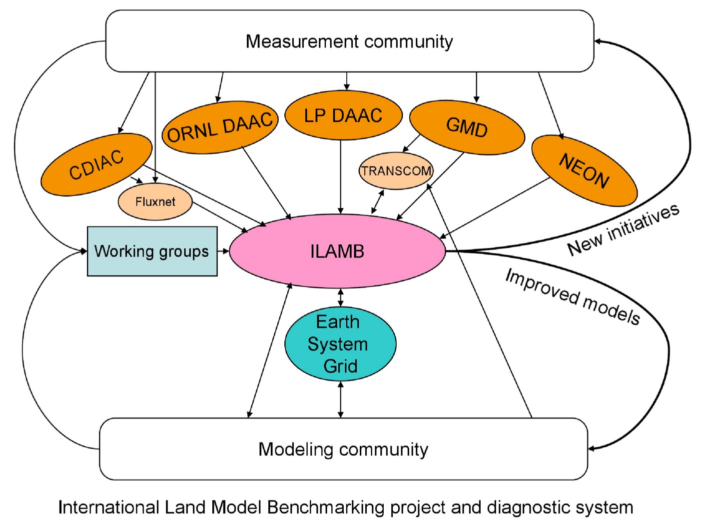
[1]Oak Ridge National Laboratory (ORNL) and [2]University of California-Irvine

Introduction

The objectives of the International Land Model Benchmarking (ILAMB) Project are to 1) develop community-accepted benchmarks for land model performance, 2) apply the benchmarks to global models, 3) support the design and development of a new open-source benchmarking osftware system, 4) and strengthen linkages between experimental, monitoring, remote sensing, and climate modeling communities in the design of new model tests and new measurement programs. Model benchmarks will quantitatively diagnose impacts of model development on carbon cycle and land surface processes, guide synthesis efforts (such as the IPCC), increase scrutiny of key data sets, identify gaps in observational programs, provide an application-specific set of criteria for model performance, and offer an optional weighting system for multi-model means based on performance on benchmarks.



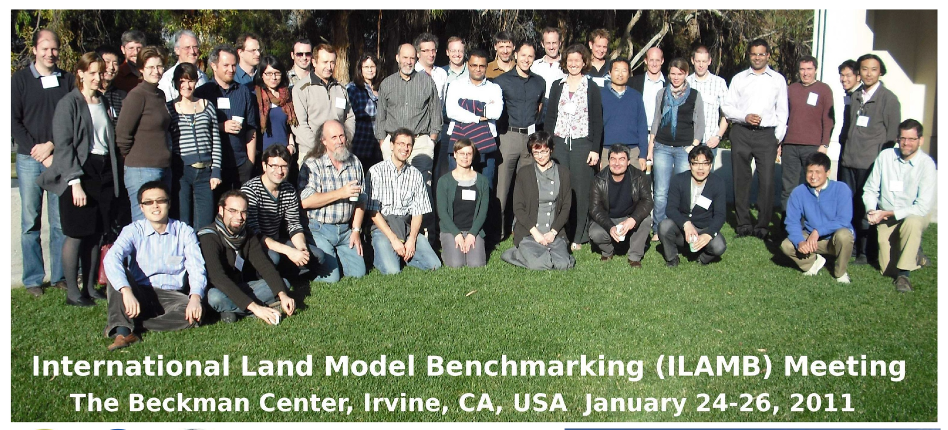
Many model intercomparison projects (MIPs) spend resources “reinventing the wheel” for variable naming conventions, model simulation protocols, and analysis and diagnostics software. The ILAMB software will offer each new MIP access to model-data comparison modules from past MIPs, improving MIP efficiency and allowing more frequent model evaluation activities.



By using the wide variety of measurements made, collected, and distributed by researchers and data centers, ILAMB identifies areas in which improvements can be made to models as well as identifying needs for new kinds of measurements. In addition, all the ILAMB model output will be distributed via the Earth System Grid (ESG), and model diagnostics will be available on the Web for use by the wider scientific community.

A benchmark is a quantitative test of model function, for which uncertainties in observations can be quantified. Acceptable peformance on benchmarks is a necessary but not sufficient condition for a fully functioning model. An effective benchmark is one that draws upon a broad set of independent observations to evaluate model performance on multiple temporal and spatial scales.

First ILAMB Community Meeting



The First ILAMB Meeting was co-organized by Forrest Hoffman, Chris Jones, Pierre Friedlingstein, and Jim Randerson. About 45 researchers participated from the United States, Canada, the United Kingdom, the Netherlands, France, Germany, Switzerland, China, Japan, and Australia.

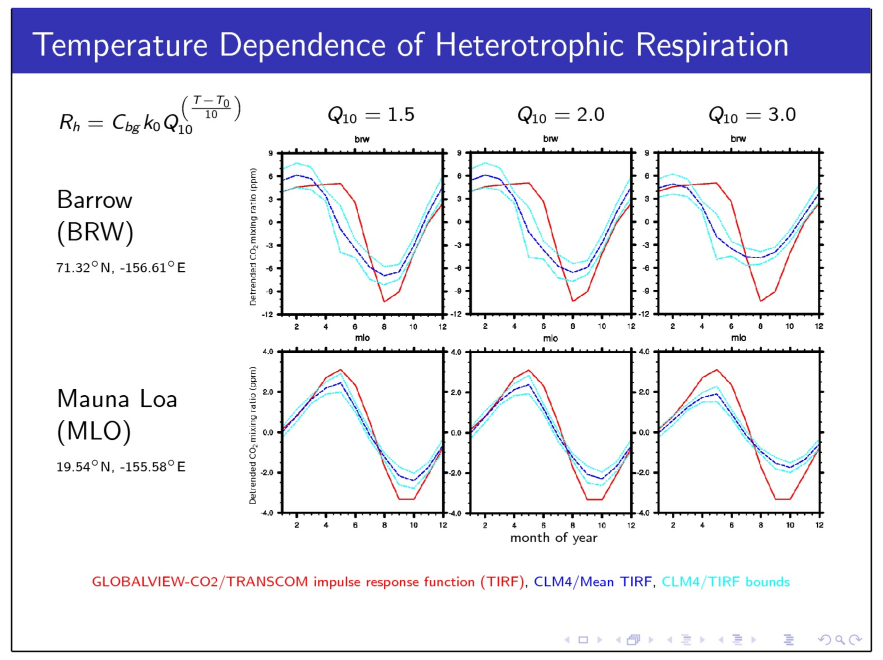
The goals of the ILAMB meeting were to

1) coordinate the design of the first set of benchmarks for global models,
2) coordinate the carbon cycle and land model evaluations for TRENDY and CMIP5 results,
3) develop an implementation plan for application of ILAMB benchmarks to TRENDY and CMIP5 output,
4) decide upon an approach for developing ILAMB software, and
5) decide upon a future schedule and means to secure funding.

Extensions to the Climate and Forecast (CF) metadata conventions for biogeochemistry data sets were also discussed. A proposal for these extensions will be developed by ILAMB participants.

Benchmark Example: Temperature Dependence of R_h

In many biogeochemistry models, the Q_{10} factor controls the dependence of heterotrophic respiration on temperature (see equation at right). Values of Q_{10} that are too large over-emphasize this temperature dependence and result in stronger respiration in mid-summer, producing a weakened seasonal cycle of CO_2.

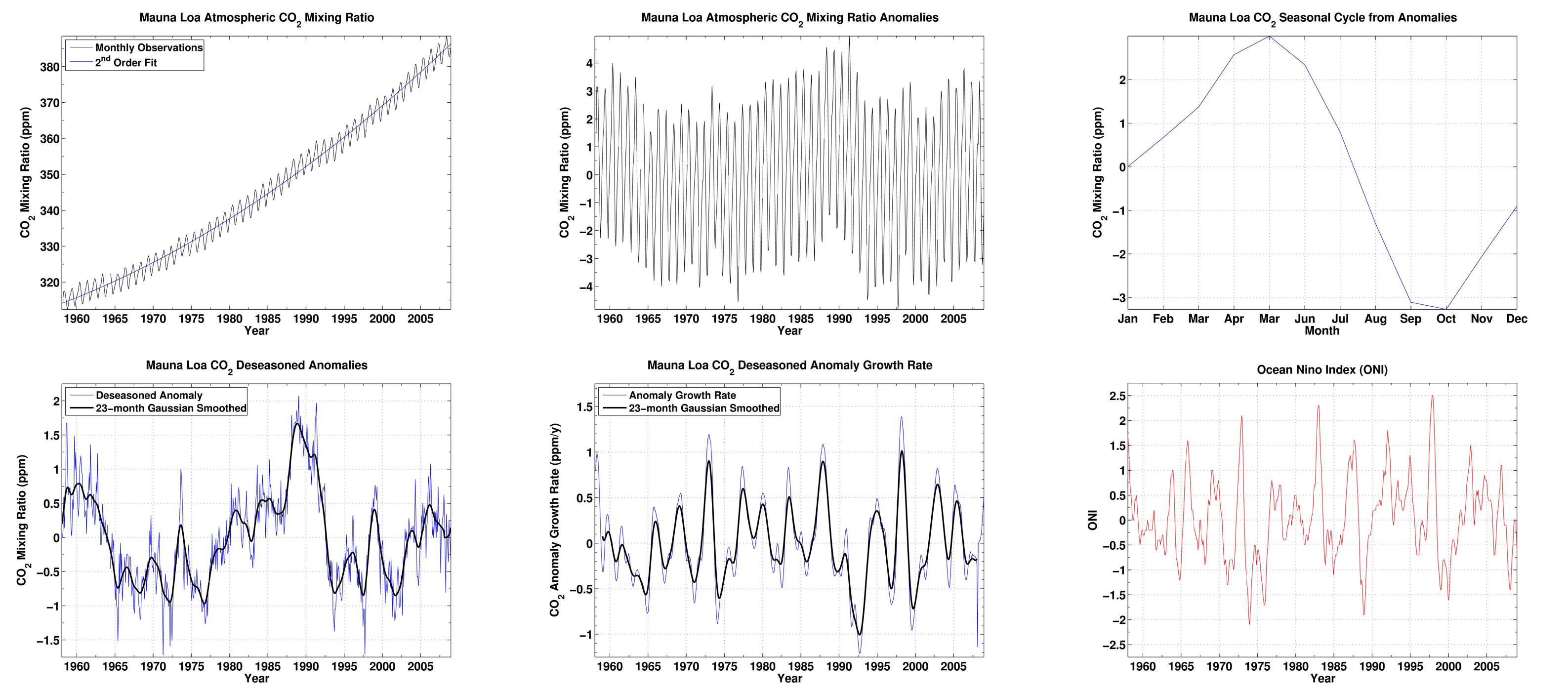


As a result, the observed annual cycle of CO_2 can be used to constrain this temperature dependence in models.

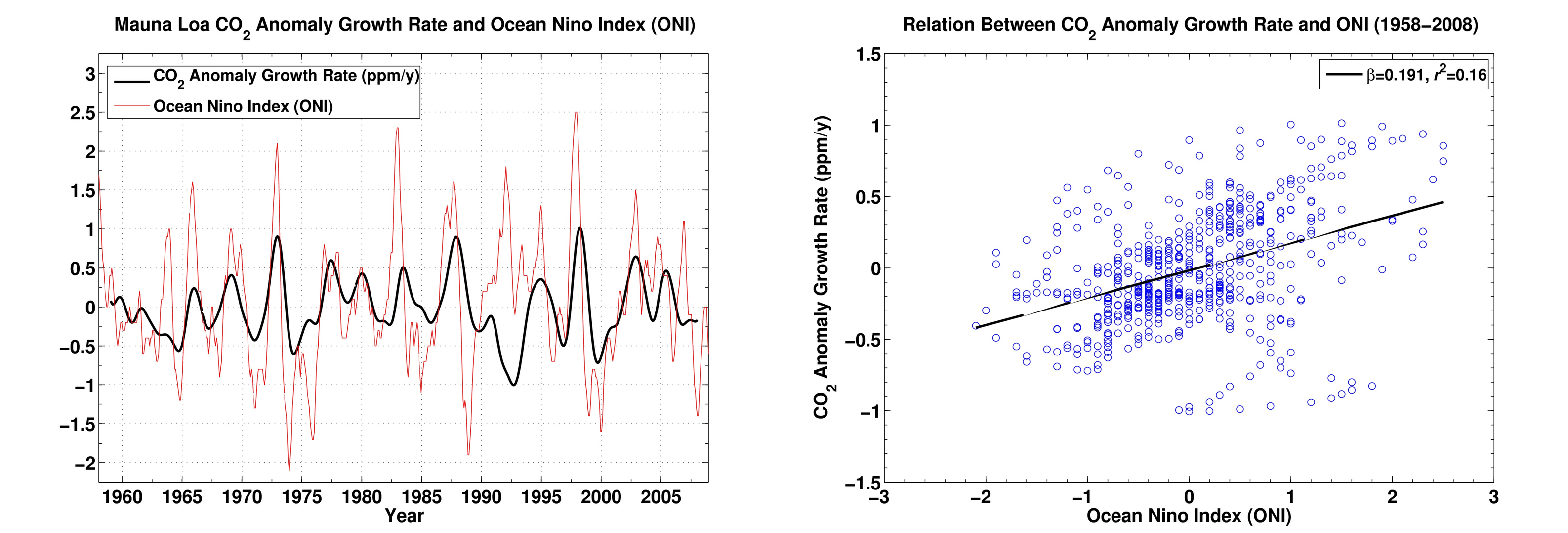
This constraint was used to tune CLM4 based on the C-LAMP comparions. See:

Randerson, James T., Forrest M. Hoffman, Peter E. Thornton, Natalie M. Mahowald, Keith Lindsay, Yen-Huei Lee, Cynthia D. Nevison, Scott C. Doney, Gordon Bonan, Reto Stöckli, Curtis Covey, Steven W. Running, and Inez Y. Fung. September 2009. “Systematic Assessment of Terrestrial Biogeochemistry in Coupled Climate-Carbon Models.” Global Change Biology, 15(9):2462–2484. doi:10.1111/j.1365-2486.2009.01912.x.

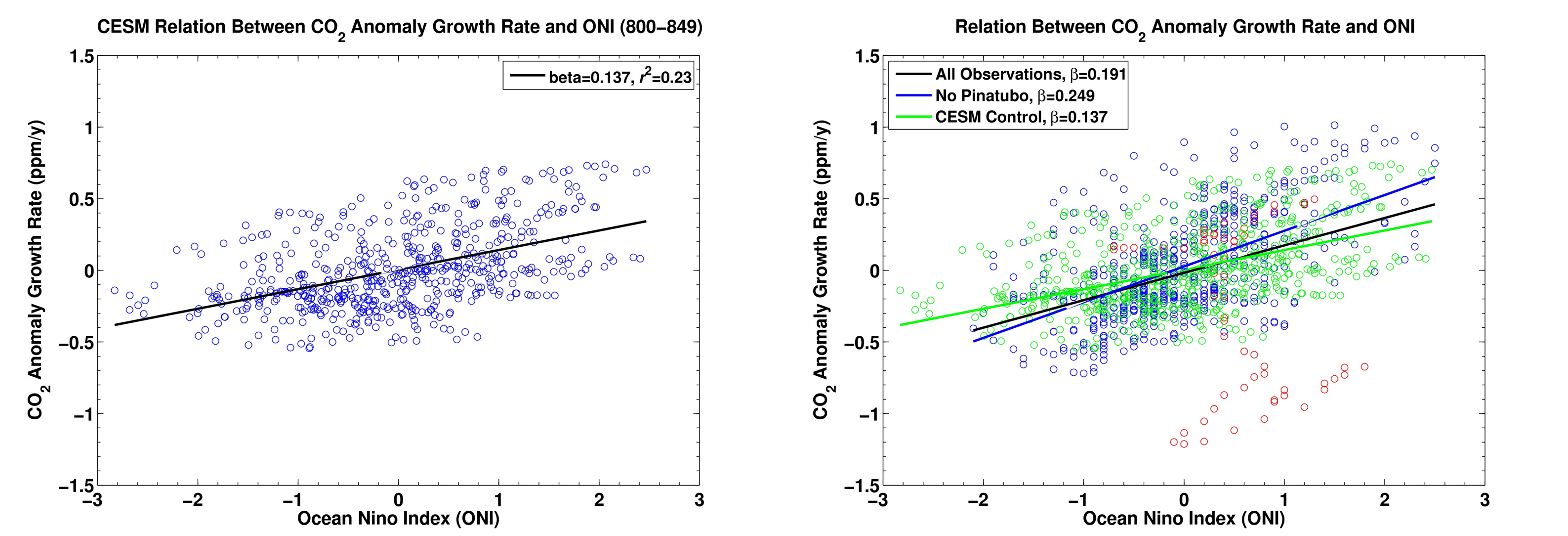
Benchmark Example: CO_2 Dependence on El Niño



The smoothed, deseasoned anomaly growth rate of the monthly Mauna Loa CO_2 record (calculated in the first five panels) can be compared to the Ocean Niño Index (ONI; shown in the sixth panel).



Atmospheric CO_2 concentration is strongly correlated with ONI except during periods of large volcanic eruptions (left). For example, Mt. Pinatubo erupted in June 1991. This relationship can be seen in the scatter plot at right.



The same relationship was extracted from a Community Earth System Model (CESM) control simulation (left). In the plot at right, the model relationship fit is plotted against that from observations with and without including the Mt. Pinatubo eruption period. From this plot, we can see that the CESM El Niño is a bit strong and that the 1991 to 1993 points from the Mt. Pinatubo period (shown in red) deviate significantly from the main body of the observations. The slope from the CESM control run is reasonably close to that of all the observations. This metric may be useful in constraining models and in providing a model performance benchmark.

ILAMB Meeting Summary

Five break-out groups met, one for each benchmark category, to identify cost function metrics and graphics.

Measurement and model uncertainty must be characterized and spatial scaling mismatch considered.

Key objectives are to use publicly available data and freely available software.

The R package will be used for generating statistical results and diagnostics.

Initial benchmarks will be implemented to evaluate the existing TRENDY and CMIP5 model results.

Initial ILAMB Benchmarks and Datasets

	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					
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

An initial set of benchmarks and available observational data sets identified by the break-out groups is shown in this table.

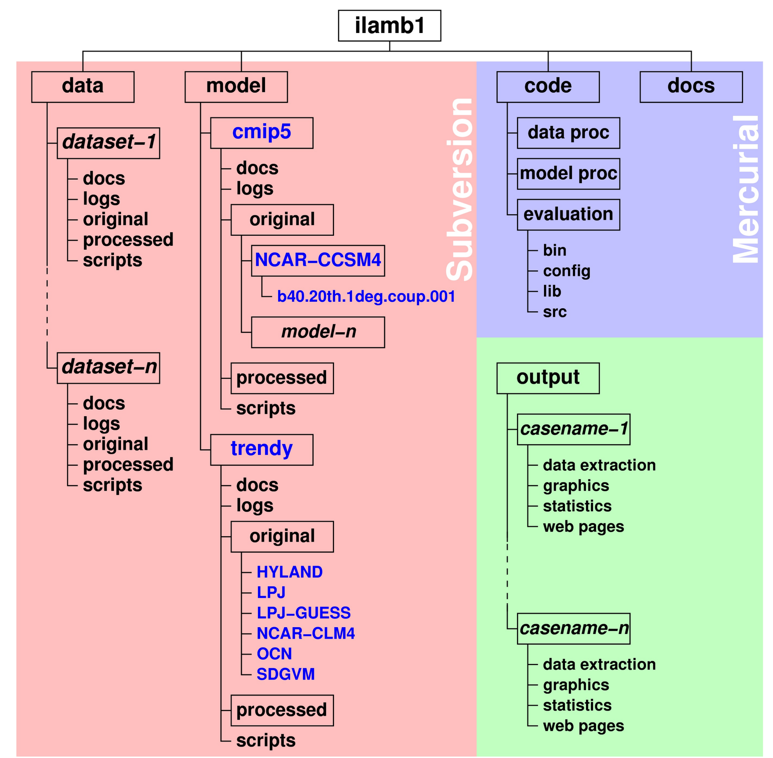
Depending upon the type of measurements available, the annual mean, seasonal cycle, interannual variability, and long-term trend of the model results will be assessed.

Observational data sets span scales from site/point *in situ* measurements to global remote sensing observations.

Next Steps for ILAMB Development

A team was identified to begin software architecture design.

A developmental hierarchy for data, model results, code, and documentation was created as shown at right.



Server-based and distributed version control systems will be used for tracking changes to data and code, respectively.

An ILAMB Town Hall meeting is planned for the AGU Fall Meeting in December.

See http://www.ilamb.org/ for more info.

Office of Science
U.S. DEPARTMENT OF ENERGY
Oak Ridge National Laboratory
Climate Change Science Institute

Acknowledgements: Research partially sponsored by the Climate and Environmental Sciences Division (CESD) of the Office of Biological and Environmental Research (OBER) within the U.S. Department of Energy's Office of Science (SC). 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.