

Model-Data Integration: ILAMB, IOMB, and the Soil Carbon Dynamics Working Group

Forrest M. Hoffman (ORNL)

April 29, 2019



Office of Science

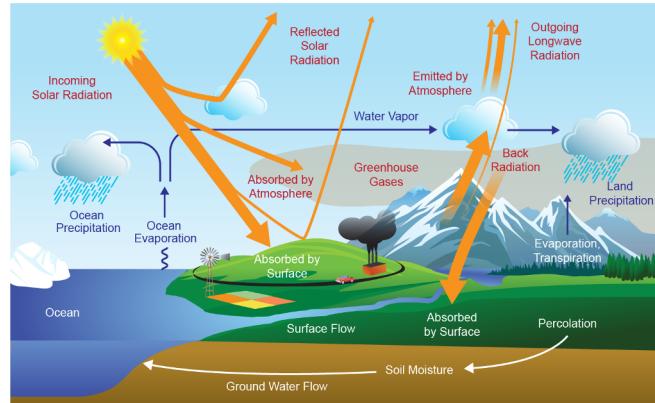
RUBISCO



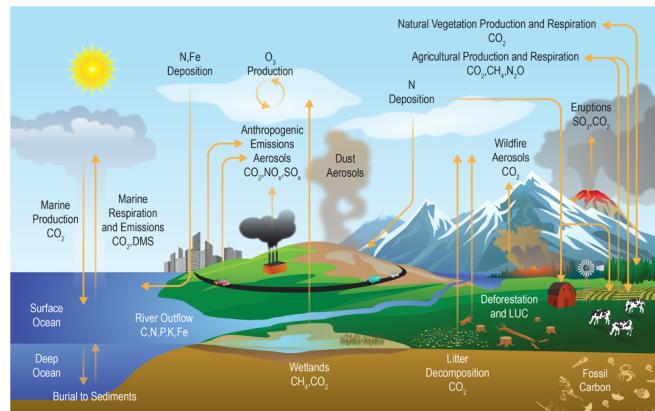
What is ILAMB?

A community coordination activity created to:

- **Develop internationally accepted benchmarks** for land model performance by drawing upon collaborative expertise
- **Promote the use of these benchmarks** for model intercomparison
- **Strengthen linkages between experimental, remote sensing, and Earth system modeling communities** in the design of new model tests and new measurement programs
- **Support the design and development of open source benchmarking tools** (Luo et al., 2012)



Energy and Water Cycles



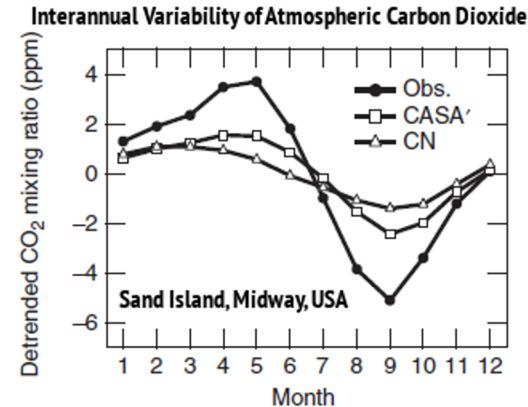
Carbon and Biogeochemical Cycles



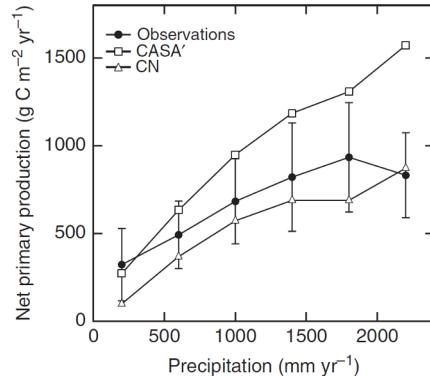


What is a Benchmark?

- A **benchmark** is a quantitative test of model function achieved through comparison of model results with observational data
- Acceptable performance on a benchmark **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 at multiple scales



Models often fail to capture the amplitude of the seasonal cycle of atmospheric CO₂



(Randerson et al., 2009)

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



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



- **First ILAMB Workshop** was held in Exeter, UK, on June 22-24, 2009
- **Second ILAMB Workshop** was held in Irvine, CA, USA, on January 24-26, 2011
 - ~45 researchers participated from the US, Canada, UK, Netherlands, France, Germany, Switzerland, China, Japan, and Australia
 - Developed methodology for model-data comparison and baseline standard for performance of land model process representations (Luo et al., 2012)





2016 International Land Model Benchmarking (ILAMB) Workshop

May 16–18, 2016, Washington, DC

Third ILAMB Workshop was held May 16–18, 2016

- Workshop Goals
 - Design of new metrics for model benchmarking
 - Model Intercomparison Project (MIP) evaluation needs
 - Model development, testbeds, and workflow processes
 - Observational data sets and needed measurements
- Workshop Attendance
 - 60+ participants from Australia, Japan, China, Germany, Sweden, Netherlands, UK, and US (10 modeling centers)
 - ~25 remote attendees at any time

Date
DOE/SC-XXXX | doi:10.7249/XXXXXXXX

U.S. DEPARTMENT OF
ENERGY | Office of
Science

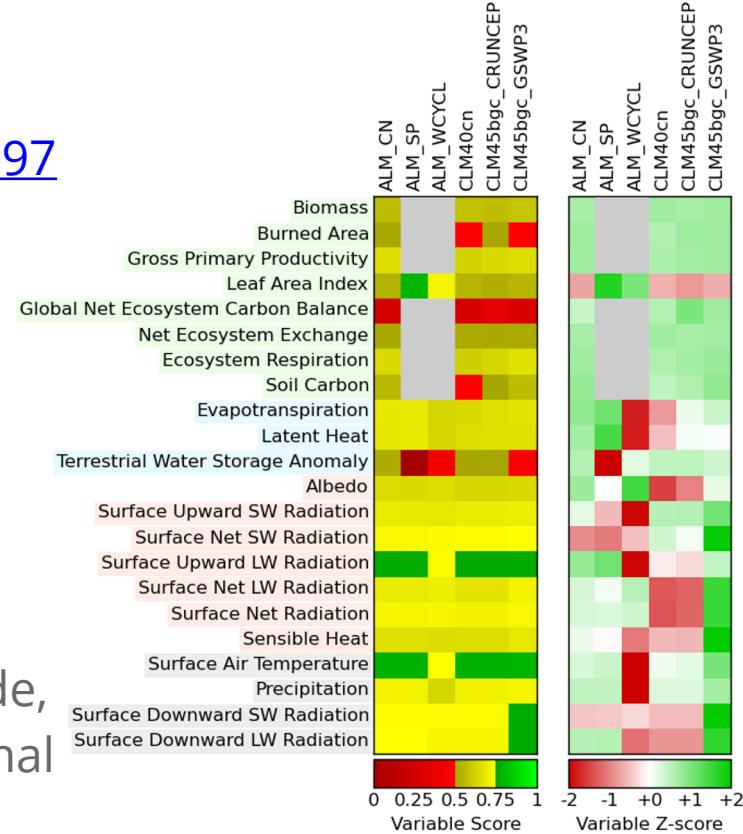
2016
International Land Model
Benchmarking (ILAMB)
Workshop Report





Development of ILAMB Packages

- **ILAMBv1** released at 2015 AGU Fall Meeting Town Hall, doi:[10.18139/ILAMB.v001.00/1251597](https://doi.org/10.18139/ILAMB.v001.00/1251597)
- **ILAMBv2** released at 2016 ILAMB Workshop, doi:[10.18139/ILAMB.v002.00/1251621](https://doi.org/10.18139/ILAMB.v002.00/1251621)
- Open Source software freely distributed
- Routinely used for E3SM and CESM evaluation during development
- Employed to evaluate CMIP5 models
- Models are scored based on statistical comparisons (bias, RMS error, phase, amplitude, spatial distribution, Taylor scores) and functional response metrics



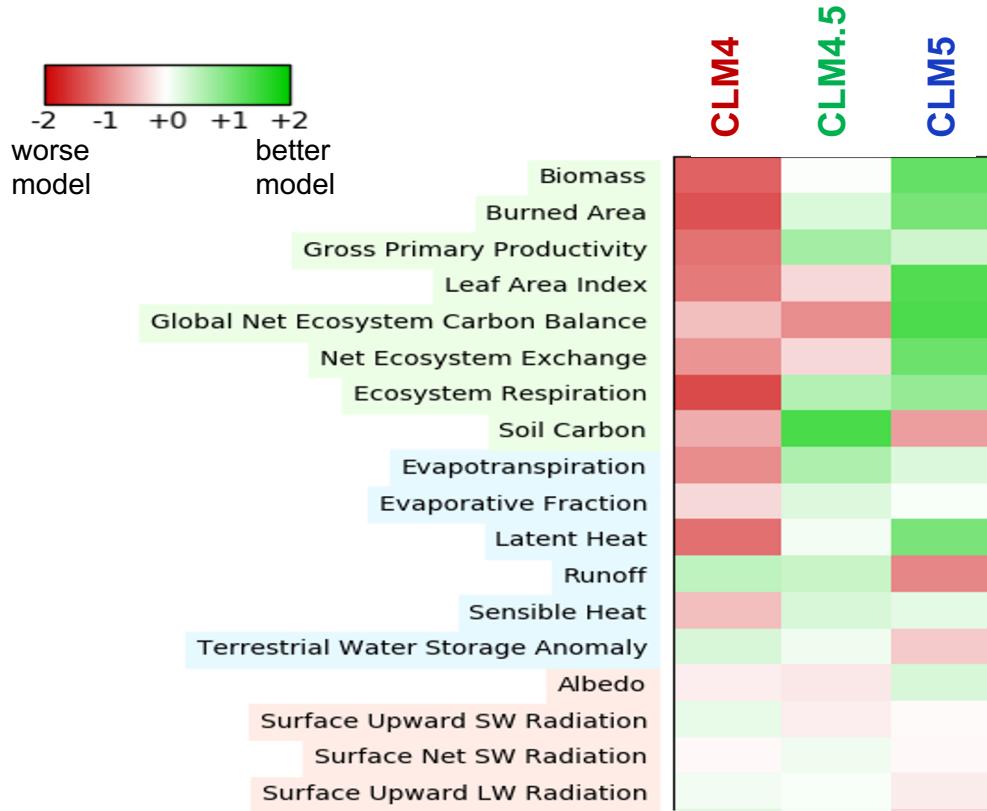


ILAMBv2 Package Current Variables

- **Biogeochemistry:** Aboveground live biomass (Contiguous US, Pan Tropical Forest), Burned area (GFED3), CO₂ (NOAA GMD, Mauna Loa), Gross primary production (Fluxnet, MTE), Leaf area index (AVHRR, MODIS), Global net land flux (GCP, Khatiwala/Hoffman), Net ecosystem exchange (Fluxnet, GBA), Ecosystem Respiration (Fluxnet, GBA), Soil C (HWSD, NCSCDv2)
- **Hydrology:** Evapotranspiration (GLEAM, MODIS), Latent heat (Fluxnet, MTE), Soil moisture (ESA), Terrestrial water storage anomaly (GRACE)
- **Energy:** Albedo (CERES, GEWEX, MODIS), Surface up SW/LW radiation (CERES, GEWEX.SRB, WRMC.BSRN), Sensible heat (Fluxnet, GBA)
- **Forcing:** Surface air temperature (CRU, Fluxnet), Precipitation (Fluxnet, GPCC, GPCP2), Surface down SW/LW radiation (Fluxnet, CERES, GEWEX.SRB, WRMC.BSRN)



ILAMB Assessing Several Generations of CLM



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

(Lawrence et al., in revision)



GrossPrimaryProductivity / GBAF / 1982-2008 / global / CESM1-BGC

Mean State

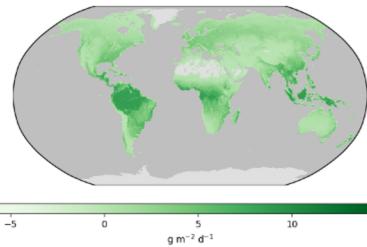
Relationships

All Models

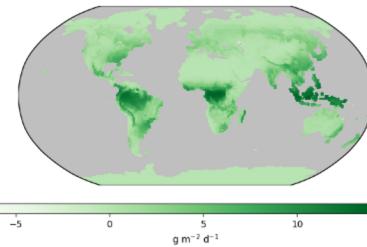
Data Information

Temporally integrated period mean

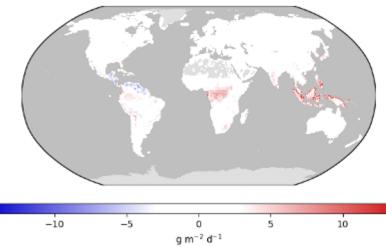
BENCHMARK MEAN



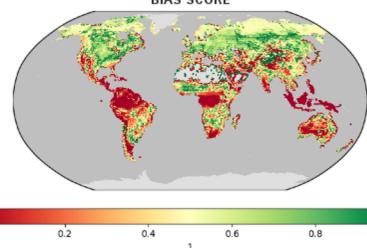
MODEL MEAN



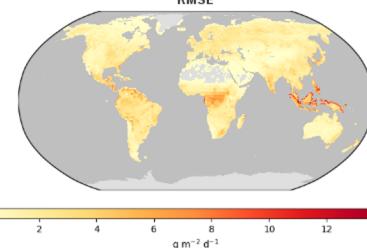
BIAS



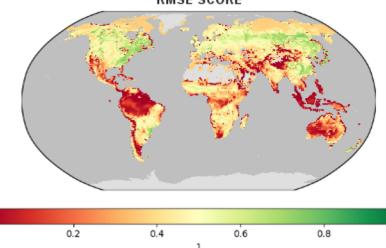
BIAS SCORE



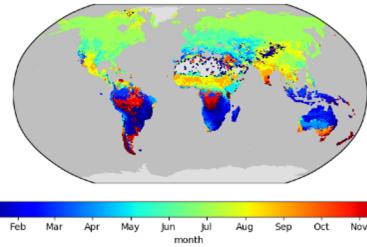
RMSE



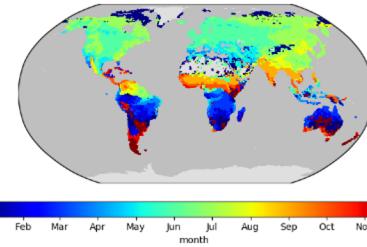
RMSE SCORE



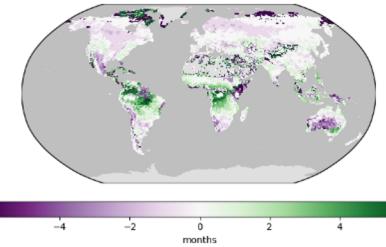
BENCHMARK MAX MONTH



MODEL MAX MONTH



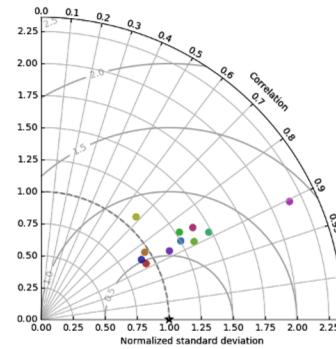
DIFFERENCE IN MAX MONTH



SEASONAL ANNUAL SCORE



SPATIAL TAYLOR DIAGRAM



MODEL COLORS

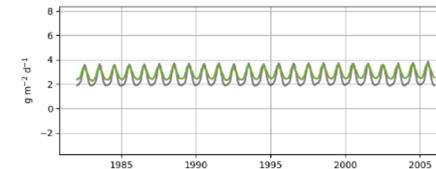
Benchmark	inmcm4
bcc-csm1-1-m	IPSL-CM5A-LR
BNU-ESM	MIROC-ESM
CanESM2	MPI-ESM-LR
CESM1-BGC	MRI-ESM1
GFDL-ESM2G	NorESM1-ME
HadGEM2-ES	

Spatially integrated regional mean

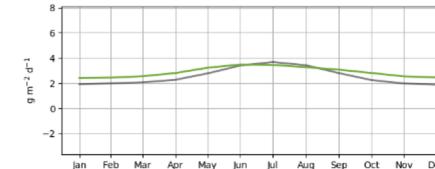
MODEL COLORS

Benchmark	inmcm4
bcc-csm1-1-m	IPSL-CM5A-LR
BNU-ESM	MIROC-ESM
CanESM2	MPI-ESM-LR
CESM1-BGC	MRI-ESM1
GFDL-ESM2G	NorESM1-ME
HadGEM2-ES	

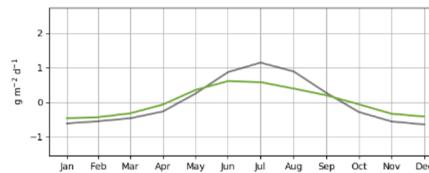
REGIONAL MEAN



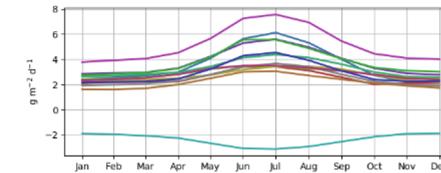
ANNUAL CYCLE



MONTHLY ANOMALY



ANNUAL CYCLE





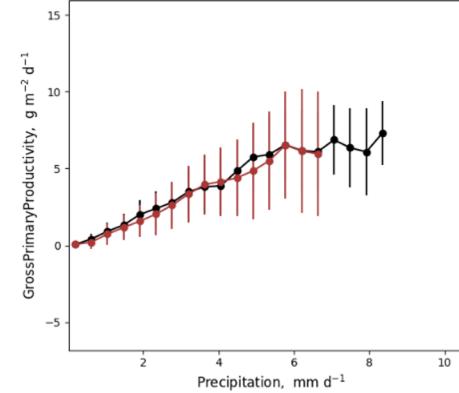
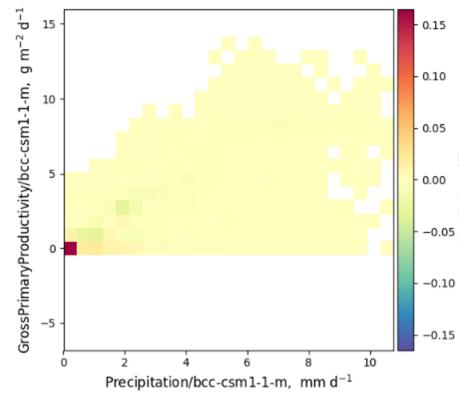
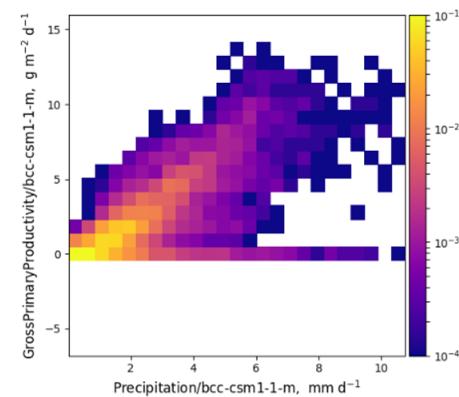
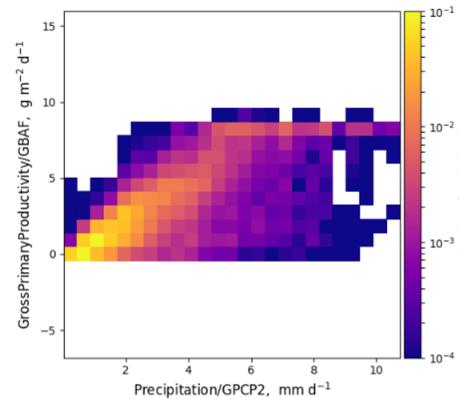
Mean State

Relationships

All Models

Data Information

Precipitation/GPCP2

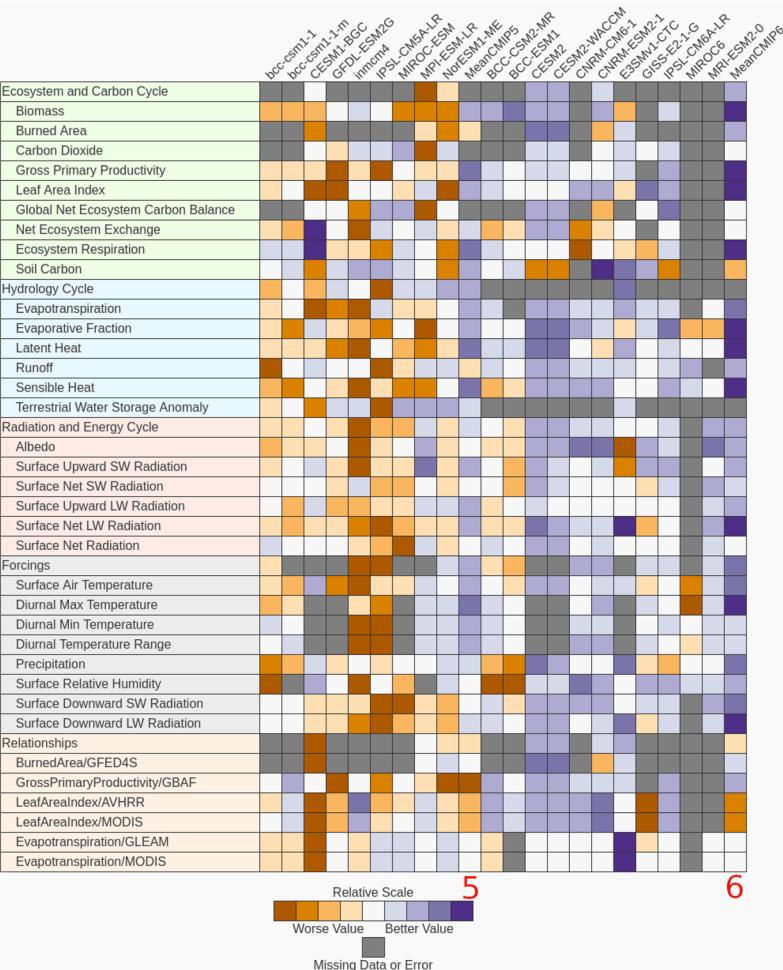




CMIP5 vs. CMIP6 Models

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

(Hoffman et al., in prep)

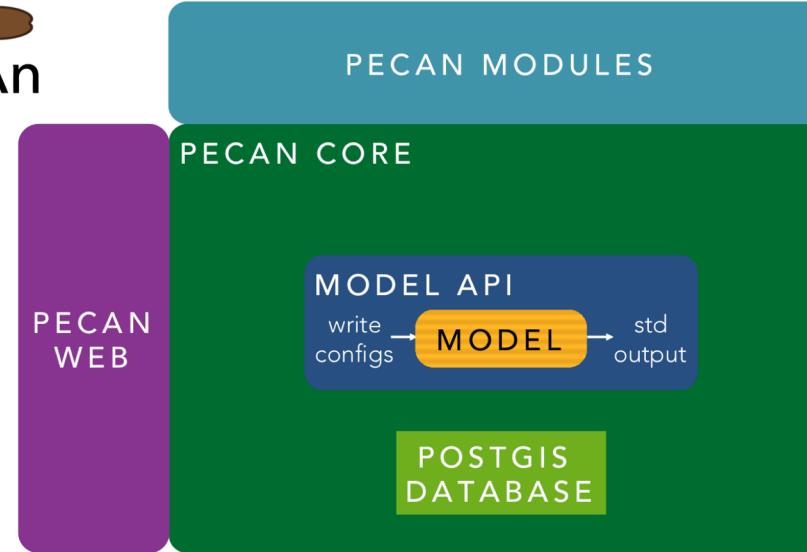




Predictive Ecosystem Analyzer (*PEcAn*)



BETYdb.org



Standardized inputs and outputs

Provenance: transparent & repeatable

Accessible interface

Reusable tools for execution, analysis, visualization



ILAMB + PEcAn Synergy

ILAMB brings:

- Benchmarking
- Diagnostics
- Evaluation, versioning
- Visualization

ILAMB

ILAMB gets:

- Multi-site, multi-model capability
- LSM to ESM scale
- UQ and DA
- Strong provenance

Synthetic virtual framework

APIs &
Tools

HPC

Data
“network”

Enabling comp. infrastructure

PEcAn brings:

- Multi-model
- Site, multi-site
- UQ, DA, other tools
- Provenance

PEcAn

PEcAn gets:

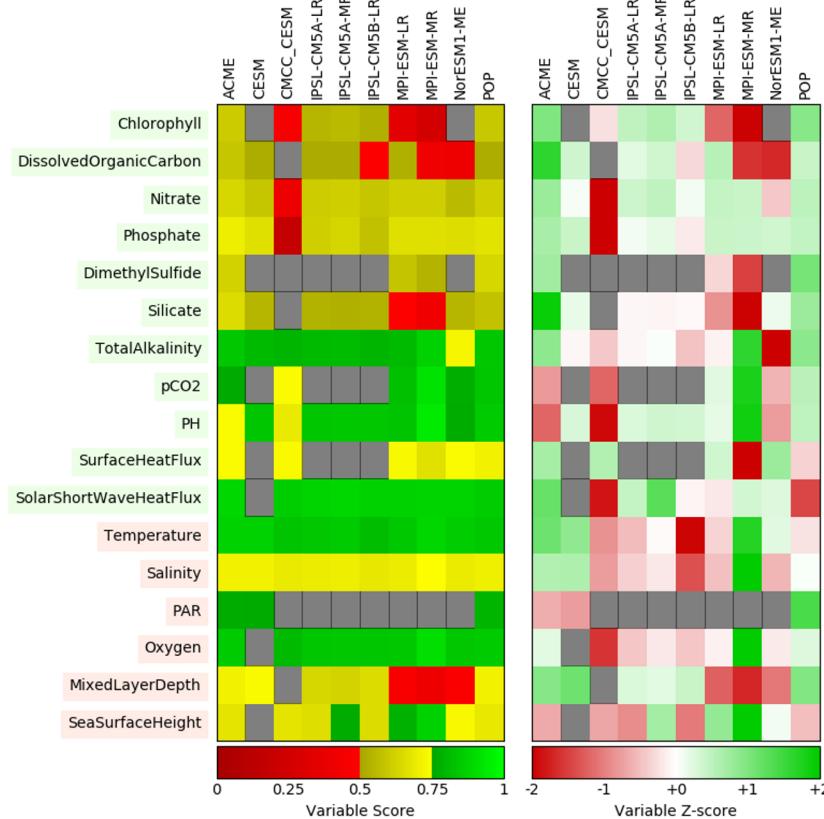
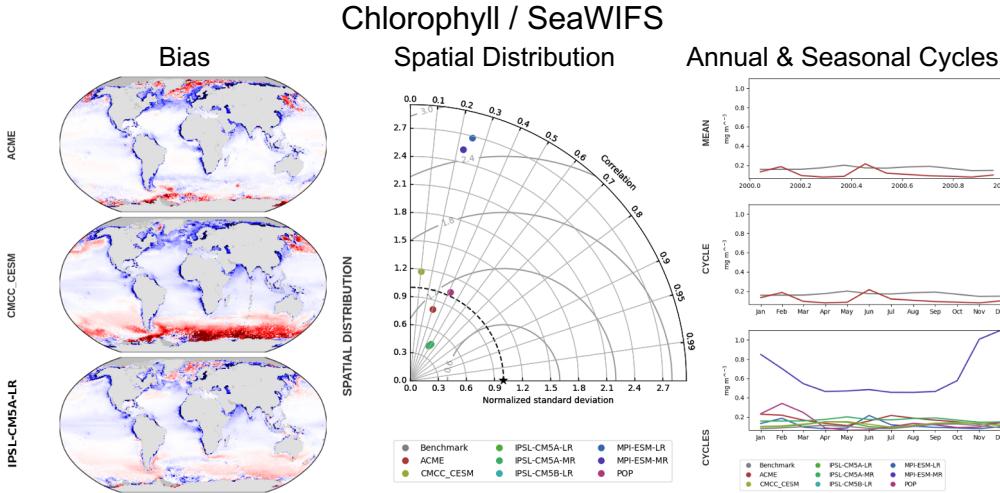
- Benchmarking
- Expanded model evaluation
- Visualization

See section E.4 in ILAMB report for more info



International Ocean Model Benchmarking (IOMB) Package

- Evaluates ocean biogeochemistry results compared with observations (global, point, ship tracks)
- Scores model performance across a wide range of independent benchmark data
- Leverages ILAMB code base, also runs in parallel
- Built on python and open standards
- Is also open source and will be released soon





Future ILAMB/IOMB Development and Application

- ILAMBv1 and ILAMBv2 were applied to:
 - CMIP5 Historical and esmHistorical simulations
 - Model verification during development of the ELM and CLM
- Within U.S. Department of Energy projects:
 - NGEE Arctic, NGEE Tropics, and SPRUCE contributing field data and metrics
 - E3SM using it to evaluate new land and ocean model features
- Ongoing model intercomparison projects: TRENDY, MsTMIP, and CMIP6
- Other groups are using and contributing to ILAMB:
 - NASA-funded Permafrost Benchmarking System
 - Community Surface Dynamics Modeling System (CSDMS) added into web modeling system
 - In-house model evaluation at various international modeling centers

Collier, Nathan, Forrest M. Hoffman, David M. Lawrence, Gretchen Keppel-Aleks, Charles D. Koven, William J. Riley, Mingquan Mu, and James T. Randerson (2018), The International Land Model Benchmarking (ILAMB) System: Design, Theory, and Implementation, *J. Adv. Model. Earth Syst.*, 10(11):2731–2754, doi:[10.1029/2018MS001354](https://doi.org/10.1029/2018MS001354).



For more information, come to the

Model Benchmarking and ILAMB Tutorial

*(Listed as TES Breakout D: ILAMB Soil C Demonstration and
Tutorial on the Agenda)*

Tuesday night at 7:00 p.m. – 8:30 p.m.

In Franklin Building 18/19



Soil Carbon Dynamics Working Group



U.S. DEPARTMENT OF
ENERGY

- Formed after community recommendation from the 2016 International Land Model Benchmarking (ILAMB) Workshop Report
- Objective is to apply data and models to improve predictive understanding
- June and September conference calls led to meeting at ORNL in October



Data to Knowledge

Synthesize existing data from collaborative networks, archives, and publications

Knowledge to Data

Perform simulations to test hypotheses and characterize model structural uncertainties

Predictive Understanding

Design functional relationship metrics to confront models and apply data-driven approaches to model formulation

Global Data Synthesis Theme

- Combine field observations from collaborative sampling networks and databases, including International Soil Carbon Network (ISCN) and published literature
- Quantify vertical distribution of SOM and responses to controlling mechanisms

Model-Data Integration Theme

- Develop consistent datasets for initializing, forcing, and benchmarking microbially explicit soil carbon models
- Characterize model structural uncertainty through software frameworks to understand controlling mechanisms

For more information, contact Forrest M. Hoffman <forrest@climatedevelopment.org> or Umakant Mishra <umishra@anl.gov>



For more information, come to the

Soil Carbon Dynamics Working Group Discussion

(Listed as the ILAMB BGC Meeting on the Agenda)

Tonight at 6:30 p.m. – 7:30 p.m.

In Franklin Building 15/16