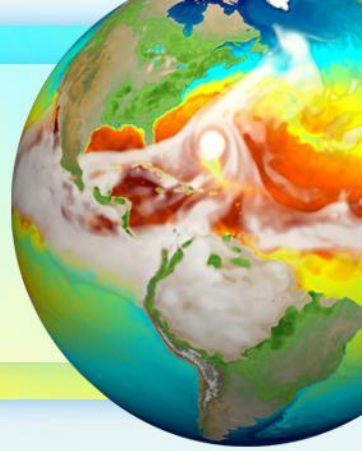


Diagnostics for the v2 Model

Forrest M. Hoffman (for Kate Calvin) and Susannah Burrows

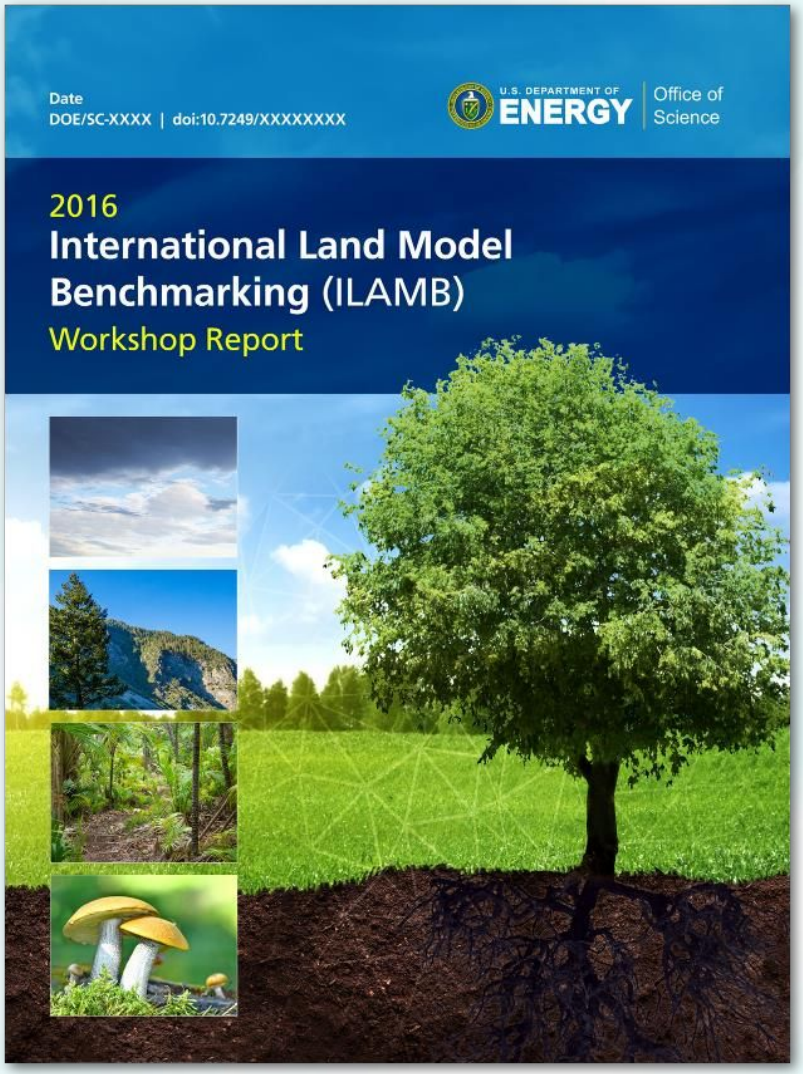
Overview of ILAMB/IOMB

The International Land Model Benchmarking (ILAMB) project is a community coordination activity to

- Develop internationally accepted benchmarks
- Promote the use of these benchmarks
- Strengthen linkages between experimental, remote sensing, and climate modeling communities
- Support the design and development of open source benchmarking tools

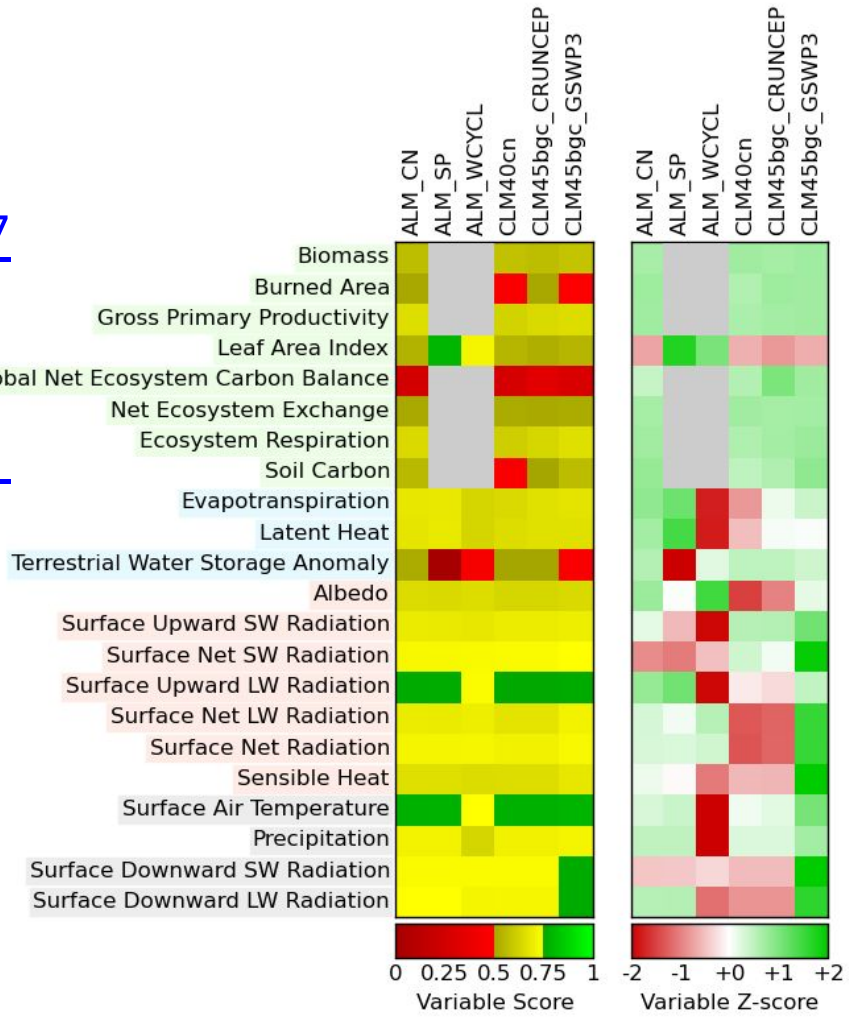
The ILAMB Package is an Open Source toolkit for evaluating land biogeochemistry models through comparisons with observations.

The International Ocean Model Benchmarking (IOMB) Package leverages the ILAMB code base for marine biogeochemistry.



International Land Model Benchmarking (ILAMB)

- **ILAMBv1** released at 2015 AGU Fall Meeting Town Hall, doi:10.18139/ILAMB.v001.00/1251597
- **ILAMBv2** released at 2016 ILAMB Workshop, doi:10.18139/ILAMB.v002.00/1251621
- 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)
- Functional response metrics



Rigorous model benchmarking with ILAMB

Objective: To provide a platform for objectively and systematically benchmarking terrestrial biogeochemistry & land surface models.

Approach: We developed an open source benchmarking software package that generates graphical diagnostics and scores model performance based on comparisons with observational data.

Results/Impacts: We used a suite of in situ, remote sensing, and reanalysis data sets in a Python package developed to evaluate model fidelity. Described is the benchmarking philosophy and mathematical methodology embodied in the ILAMB package, which is already in use in international modeling centers.

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

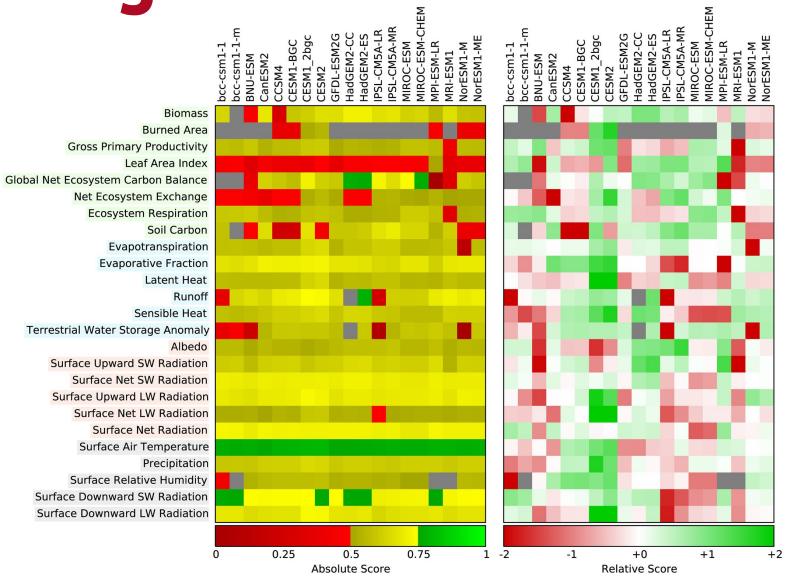


Figure: ILAMB scores land models (columns) across a variety of variables (rows).

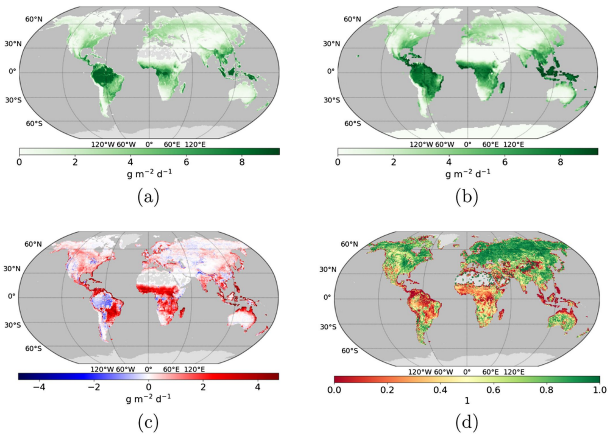


Figure: Example model–data comparison for gross primary production (GPP).

Evaluating Uncertainties in Marine Biogeochemical Models

Objective: To demonstrate the application of International Ocean Model Benchmarking (IOMB) package in evaluating marine biogeochemical models.

Approach: Used IOMB to analyze the skill of Earth system models (ESMs) in predicting marine biogeochemical variables and surface ocean concentrations and sea–air fluxes of dimethylsulfide (DMS).

Results/Impacts: Models over-predict surface concentrations of DMS in the eastern tropical Pacific by a factor of two and concentrations in the marine boundary layer by a factor of three. This could lead to excessive cooling in models due to the aerosol indirect effect.

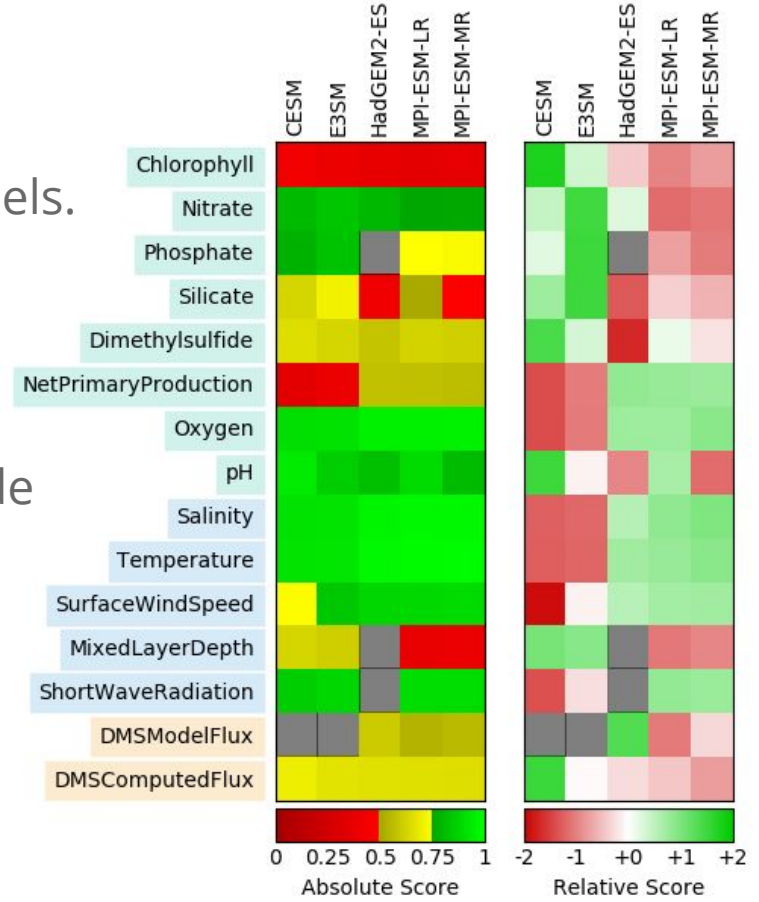
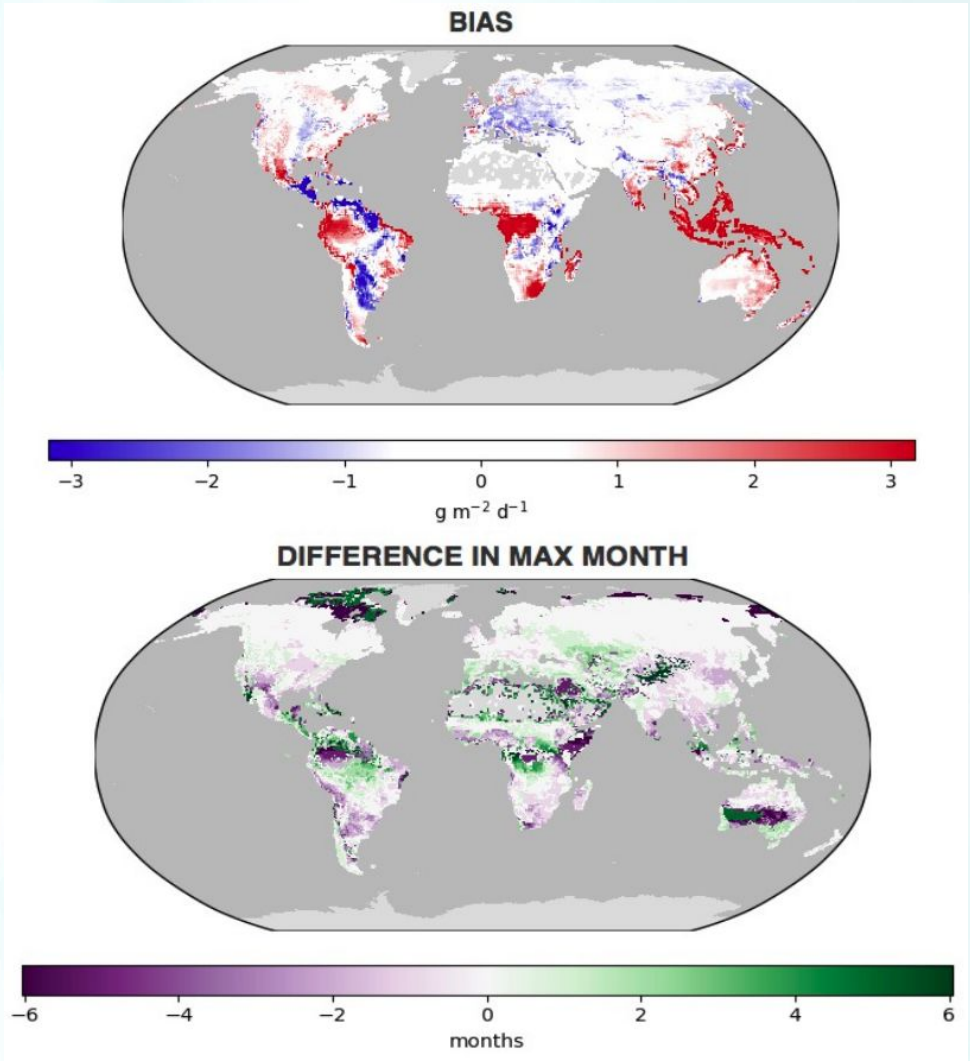
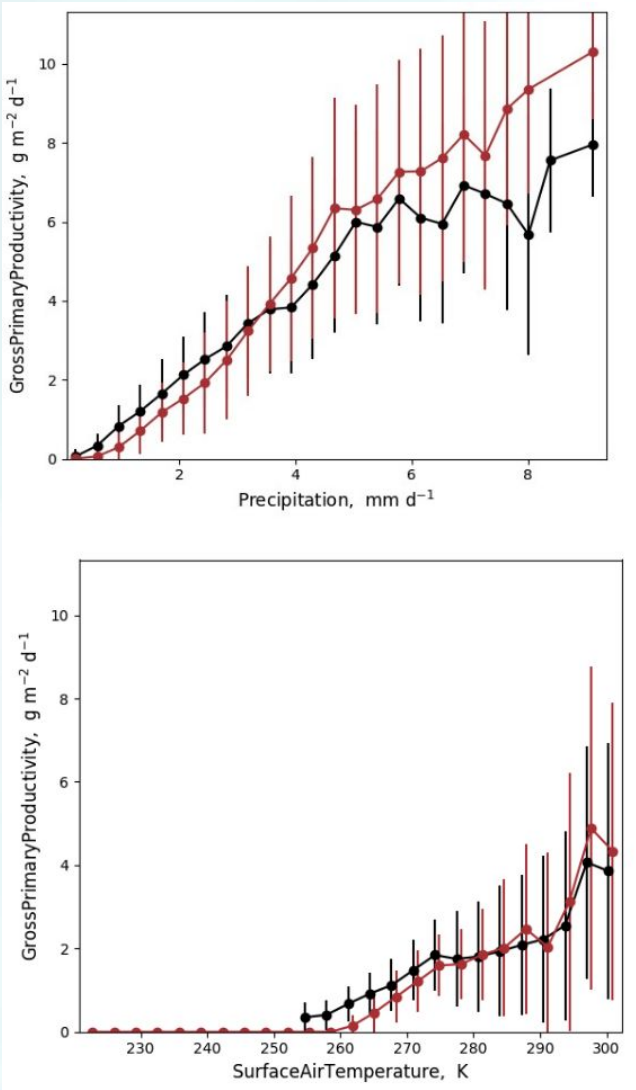


Figure: Absolute (left) and relative (right) scores for model performance of five ESM ocean models compared with observational data.

Ogunro, Oluwaseun O., **Scott M. Elliott**, Oliver W. Wingenter, Clara Deal, **Weiwei Fu**, **Nathan Collier**, **Forrest M. Hoffman** (2018), Evaluating Uncertainties in Marine Biogeochemical Models: Benchmarking Aerosol Precursors, *Atmos.*, 9(5), 184, doi:10.3390/atmos9050184.

E3SM GPP Bias and Relationships





Assessing Changes in the CMIP5 vs. CMIP6 Ensembles

Objective: To determine if the CMIP6 suite of models deliver better performance for land than the CMIP5 suite over the contemporary period.

Results/Impacts: Initial analysis of available models indicates a marked improvement in the multi-model mean from CMIP5 to CMIP6.

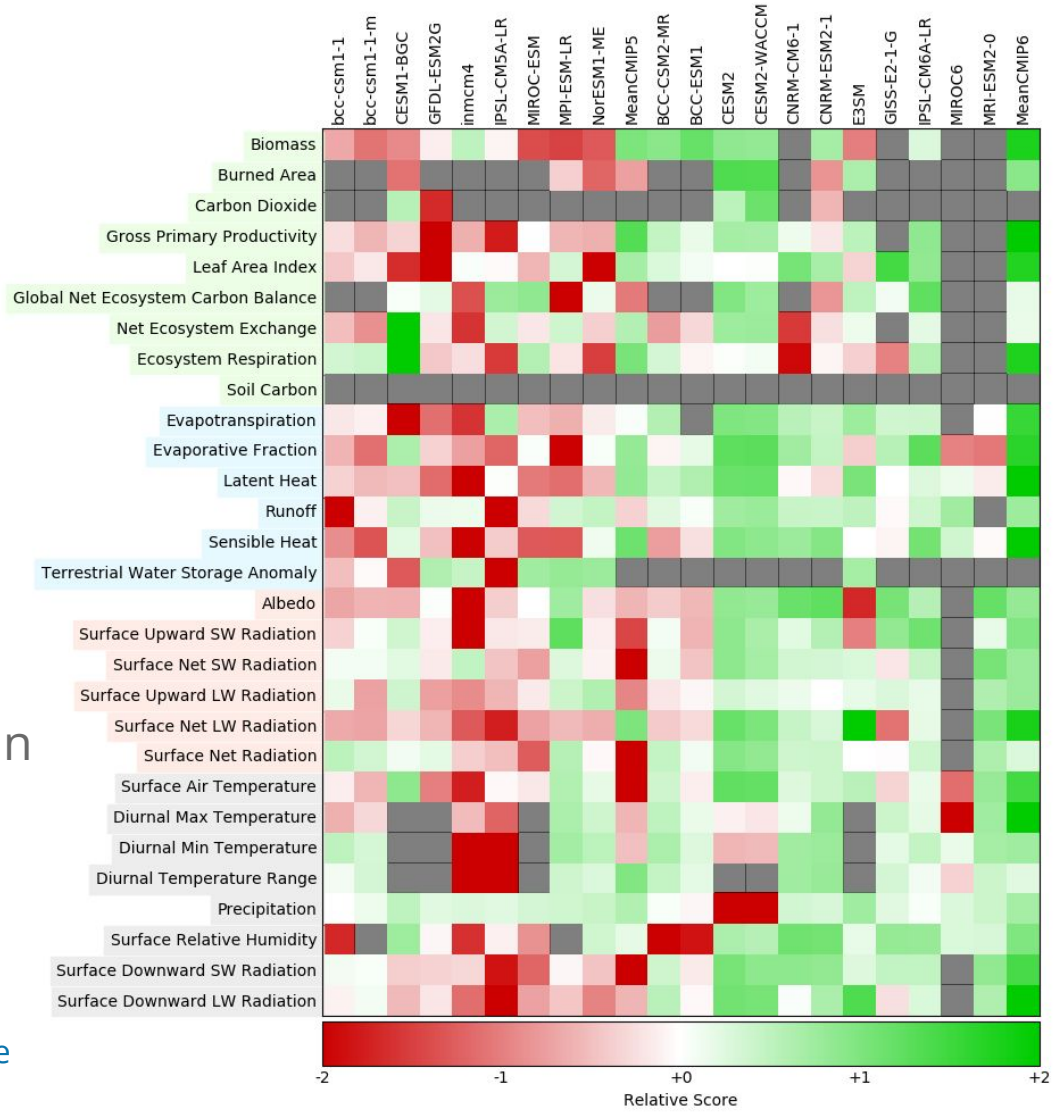


Figure: Relative scores for land model performance of CMIP5 models and CMIP5 multi-model mean (left) vs. the CMIP6 models and CMIP6 multi-model mean (right).

V1 Diagnostics that need to be hardened for convenience / reproducibility

- All sea-ice diagnostics
- More ocean BGC?
- Beta, gamma?
- Nutrient limitation diagnostics?
- Others?

New v2 model features that will need diagnostics

- FATES
- Diurnal FLUXNET site eval
- Crop model
- More biomass, GPP, LAI, and energy datasets
- Phenology (phenocams?)
- Prognostic CO_2
- GCAM
- …others?

Discussion

- How feasible is it to integrate the diagnostics we need, what is the plan?
- Are there important diagnostics we have overlooked?