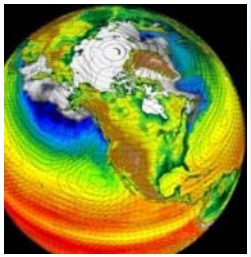
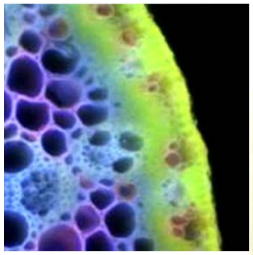
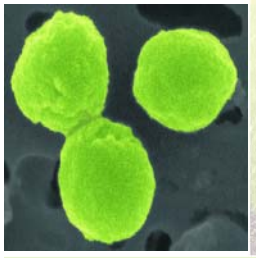


BERAC Briefing

2016 International Land Model Benchmarking (ILAMB) Workshop

May 16–18, 2016, Washington, DC, USA

October 27, 2016

Renu Joseph & Dorothy Koch

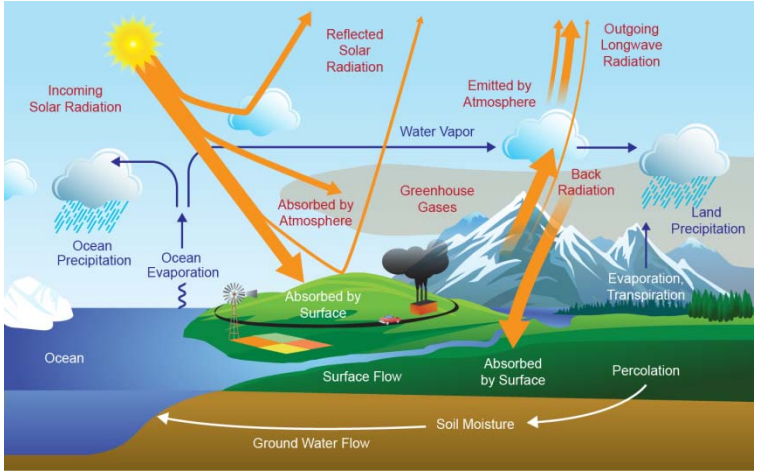
Workshop Co-Chairs: Forrest M. Hoffman, William J. Riley, James T. Randerson, Gretchen Keppel-Aleks, and David M. Lawrence

International Land Model Benchmarking (ILAMB)

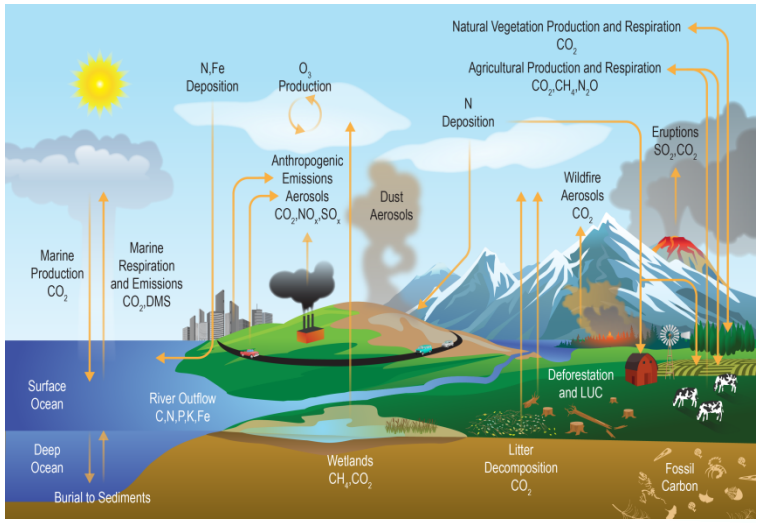
What is ILAMB?

A community coordination activity created to:

1. **Develop internationally accepted benchmarks** for land model performance by drawing upon collaborative expertise
2. **Promote the use of these benchmarks** for model intercomparison
3. **Strengthen linkages between experimental, remote sensing, and climate modeling communities** in the design of new model tests and new measurement programs
4. **Support the design and development of open source benchmarking tools**.



Energy and Water Cycles

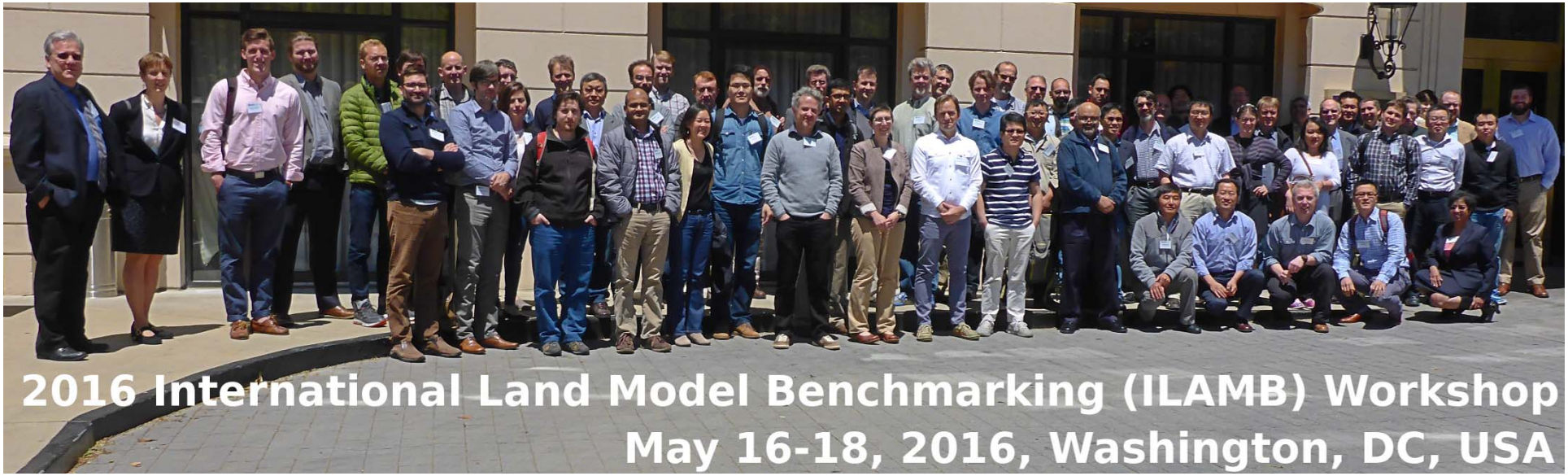


Carbon and Biogeochemical Cycles

Second US ILAMB Workshop, May 16–18, 2016

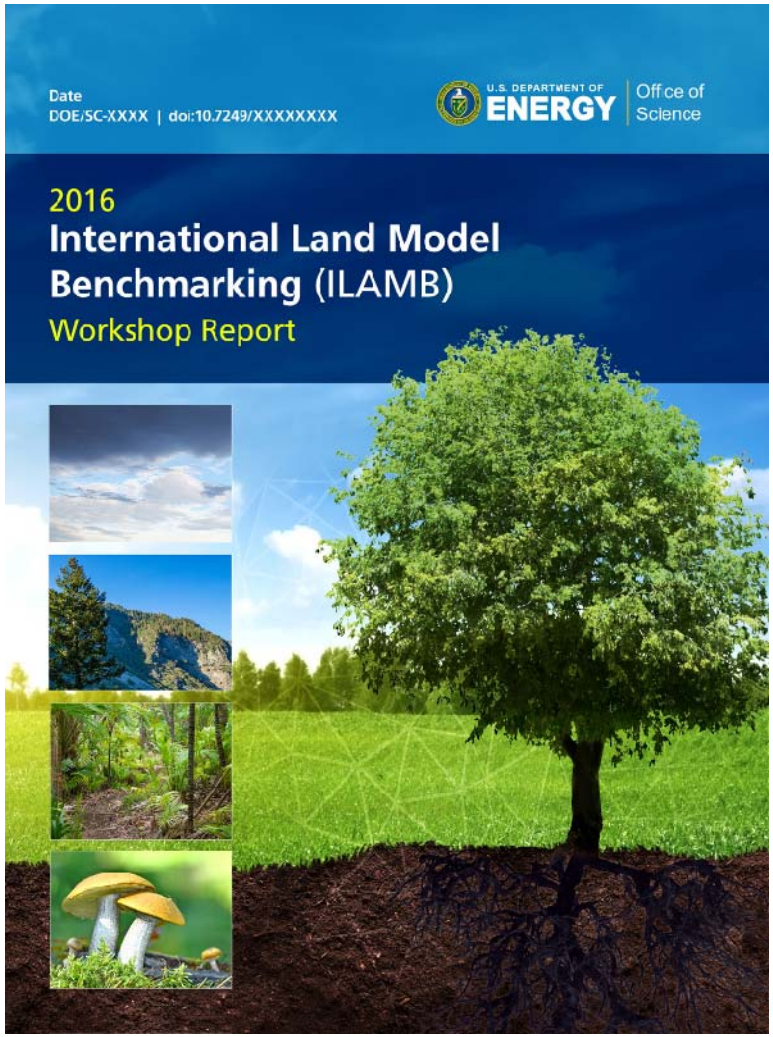
Goals of the workshop

- Understand benchmarking needs for model development and model analysis (including multi-model efforts)
- Engage the research community in defining scientific priorities for
  - Design of new metrics for model benchmarking
  - Model Intercomparison Project (MIP) evaluation needs
  - Model development, testbeds, and workflow practices
  - Observational data sets and needed measurements



Highlights of the Workshop

- **60+ participants** from Australia, Japan, China, Germany, Sweden, Netherlands, UK, and US
- **10 modeling centers** represented
- **~25 online attendees** at any time
- Report highlights
  - current state of the science
  - challenges and opportunities for benchmarking
  - model development needs
  - field and laboratory measurement priorities



Hoffman, F. M., C. D. Koven, G. Keppel-Aleks, D. M. Lawrence, W. J. Riley, J. T. Randerson, A. Ahlström, G. Abramowitz, D. D. Baldocchi, M. Best, B. Bond-Lamberty, M. De Kauwe, A. S. Denning, A. Desai, V. Eyring, R. Fisher, P. J. Gleckler, M. Huang, G. Hugelius, A. K. Jain, N. Y. Kiang, H. Kim, R. D. Koster, S. V. Kumar, H. Li, Y. Luo, J. Mao, N. G. McDowell, U. Mishra, P. Moorcroft, G. S. H. Pau, D. M. Ricciuto, K. Schaefer, C. R. Schwalm, S. Serbin, E. Shevliakova, A. G. Slater, J. Tang, M. Williams, J. Xia, C. Xu, R. Joseph, and D. Koch (2016), *International Land Model Benchmarking (ILAMB) 2016 Workshop Report*, DOE/SC-XXXX, U.S. Department of Energy, Office of Science, Germantown, Maryland, USA, XXX pp., doi:10.7249/XXXXXXX.

Crowdsourcing and Social Media at the Workshop

- Videoconferencing used for all plenary sessions
- All slides and meeting notes crowdsourced in Google Slides and Google Docs
- Twitter used for ideas, comments & questions
- White papers written and reviewed through crowdsourcing
- Technology reduced gender, racial, and cultural imbalances and travel costs and emissions

White Paper Synthesis and Workshop Outcomes

Model Intercomparison Projects (MIPS)

- CMIP6 DECK
- Coupled Climate– Carbon Cycle (C4MIP)
- Land Surface, Snow, and Soil Moisture (LS3MIP)
- Multi-scale Synthesis & Terrestrial (MsTMIP)
- Processes Linked to Uncertainties Modeling Ecosystems (PLUME-MIP)

Integrating and Cross-cutting Themes

- Process-specific experiments
- Metrics from extreme events
- Design of new perturbation experiments
- High latitude processes
- Tropical processes
- Remote sensing
- Eddy covariance flux networks

Major Processes

- Ecosystem processes and states
- Hydrology
- Atmospheric CO_2
- Soil carbon and nutrient biogeochemistry
- Surface fluxes
- Vegetation dynamics

Identified Benchmarks for

- Process understanding
- Quantified feedbacks
- Reduced uncertainties
- Improved model projections

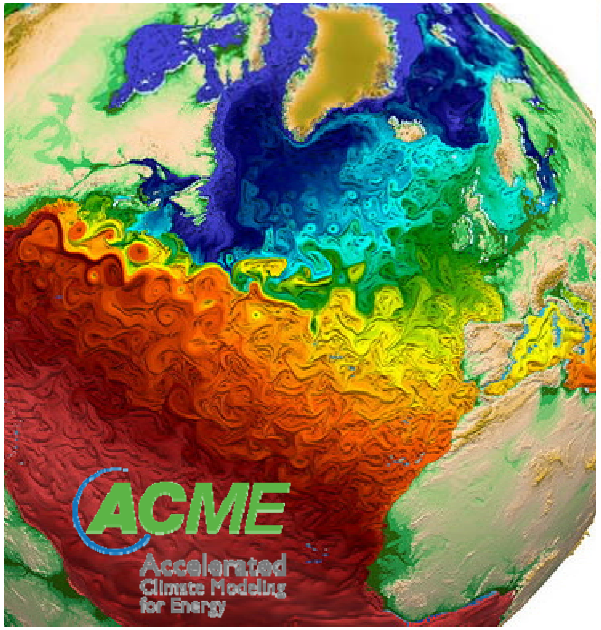
Benchmarking Challenges for Model Development

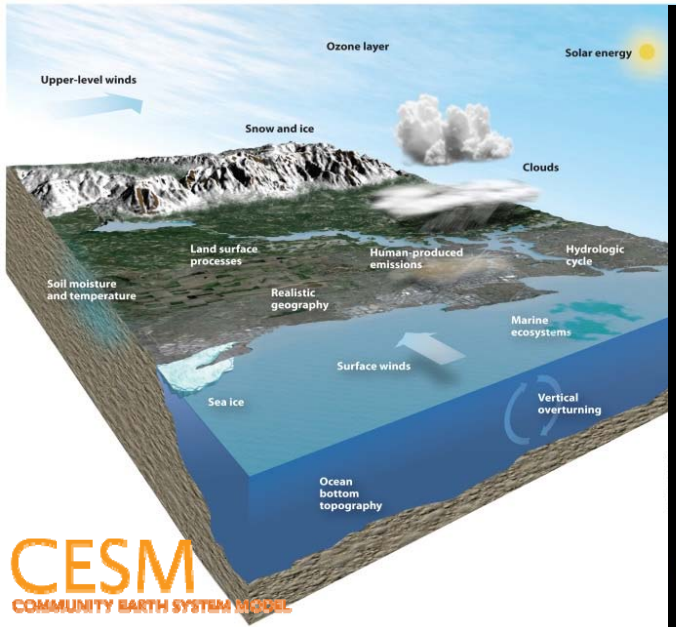
Developing metrics that make appropriate use of observational data remains **a scientific challenge** because of

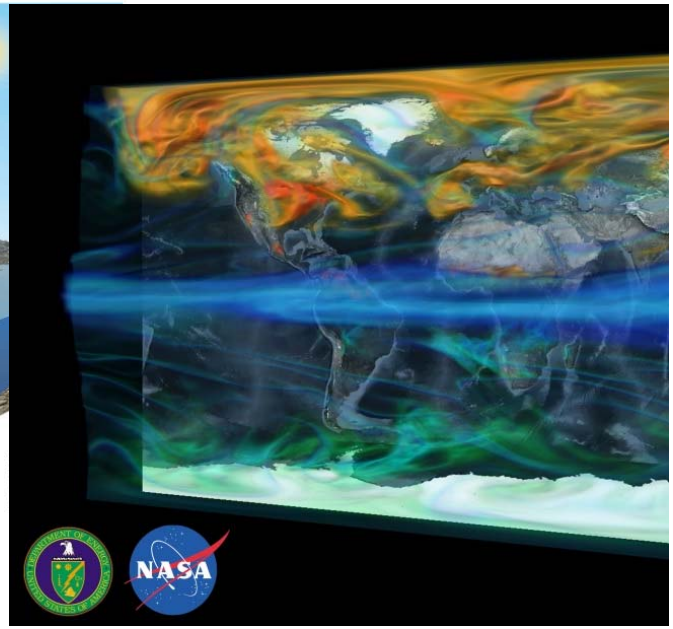
- **spatial and temporal mismatch** between models and measurements,
- **poorly characterized uncertainties** in observational data products,
- **biases** in reanalysis and forcing data,
- **model simplifications**, and structural and parametric **uncertainties**

Model testing requires

- Ability to test **alternate model formulations and develop in-situ diagnostics**
- **Land Model Testbed (LMT) capability** should be incorporated into routine model development testing (e.g., nightly/weekly automated testing)

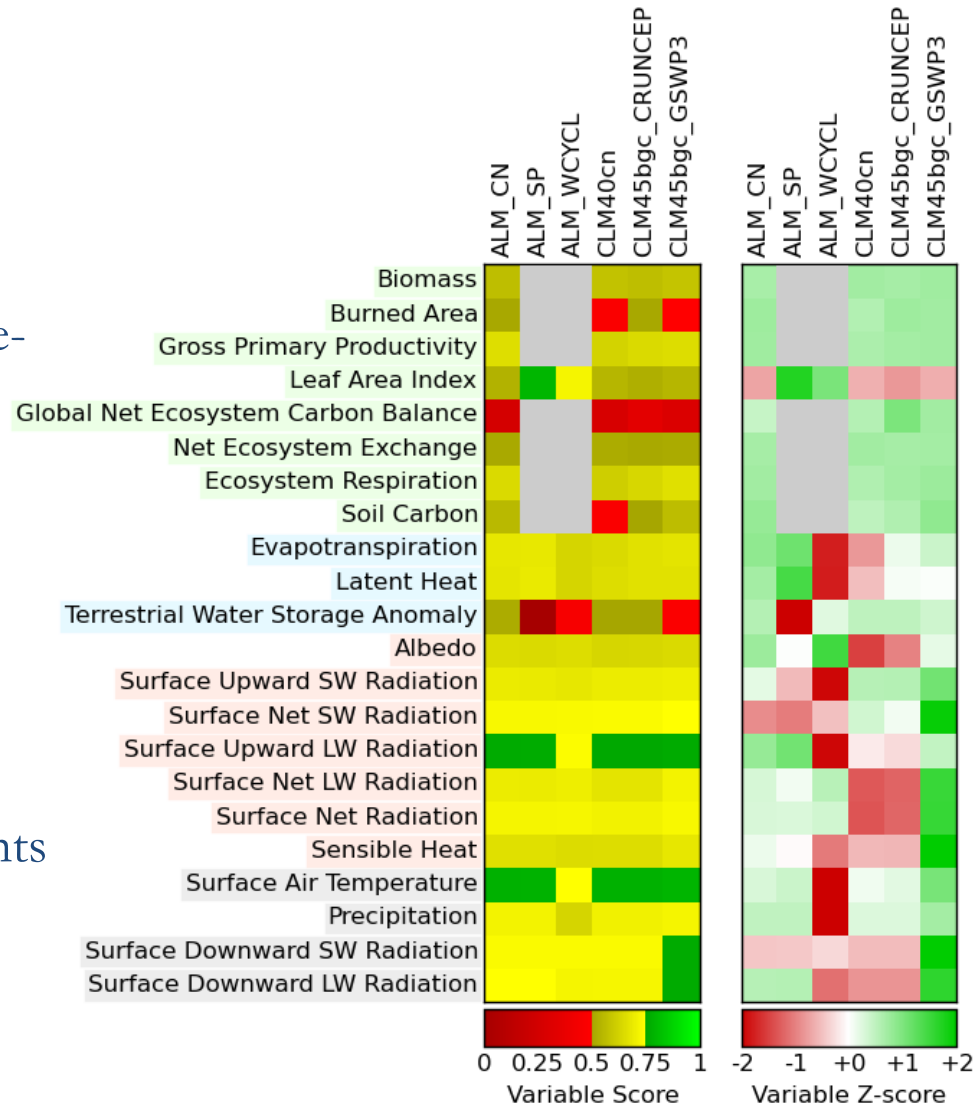






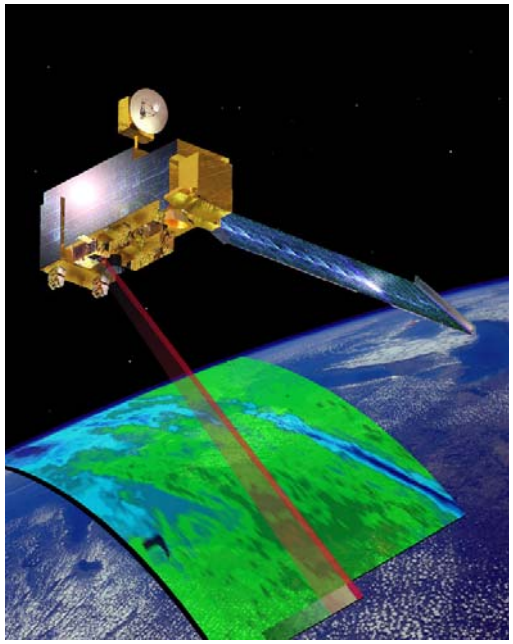
Benchmarking Approaches

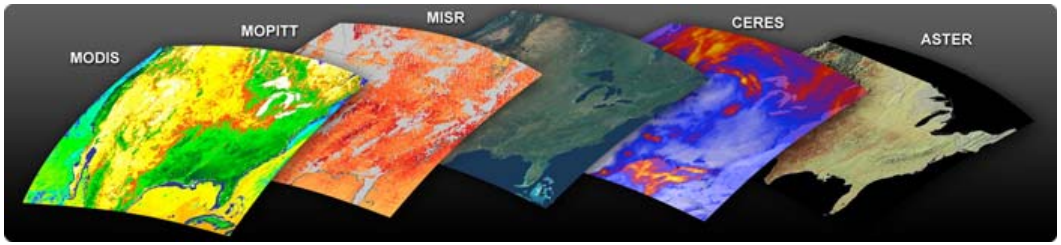
- Statistical comparisons (bias, root-mean-square error (RMSE), phase, amplitude, spatial distribution, Taylor diagrams and scores)
- **Functional response** metrics or variable-to-variable comparisons
- **Emergent constraints**
- Reduced complexity models and traceability analysis
- Formal **uncertainty quantification** methods
- Meta-analyses of perturbation experiments



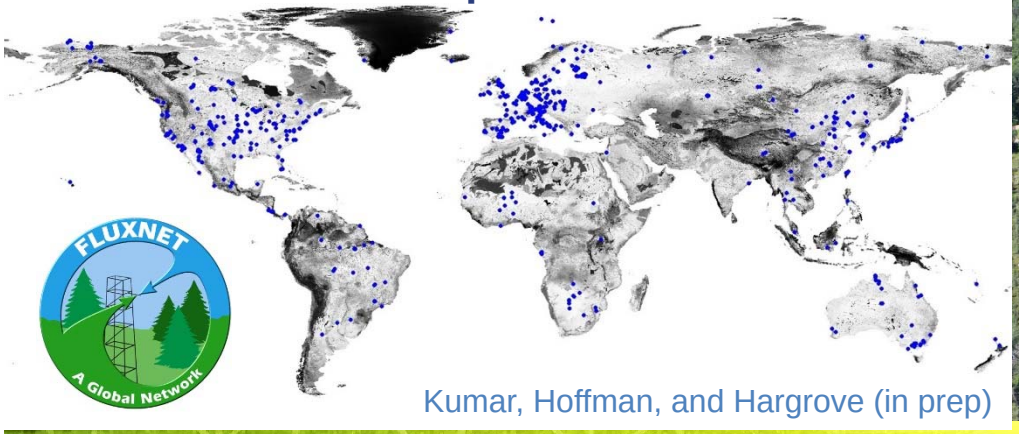
Benchmarking Challenges and Priorities

- **Super site benchmarks** for AmeriFlux and FLUXNET
- **Benchmarks for soil carbon** turnover, distribution, transport
- **Metrics for extreme events** & response of ecosystems
- **Data for vegetation** recruitment, growth, mortality, phenology, canopy structure
- Benchmarks for critical **high latitude & tropical ecosystems**
- Leverage **field projects & remote sensing methods**

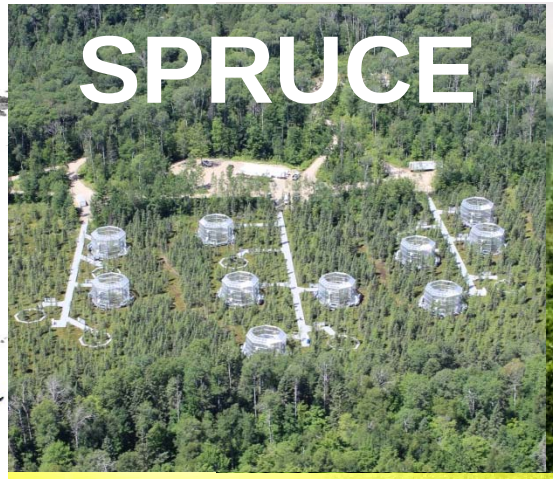




FLUXNET Representativeness



Kumar, Hoffman, and Hargrove (in prep)



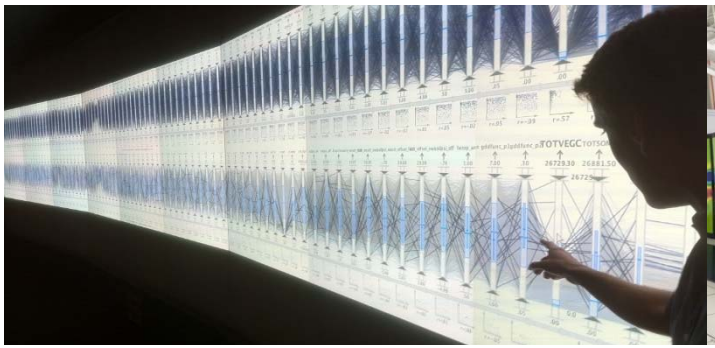
Next Generation Benchmarking to Evaluate Models

Uncertainty Quantification (UQ) Frameworks

- Develop a hierarchy of approaches that integrate and report carbon cycle model diagnostics as a matrix of flows and turnover times to attribute responses to specific ecosystem components
- Apply Bayesian UQ approaches to identify model uncertainties
- Use UQ results to guide data collection activities and target process representation improvements

Computational Needs and Requirements

- Scalable algorithms and machine learning techniques should be developed for evaluating and benchmarking high resolution and long time series ESM results
- Scientific computing facilities should strike a balance between resources for compute-intensive vs. data-intensive applications as they plan expansion to exascale computing
- **Once developed, tools can be integrated with data repositories and archives**



Conclusion and Next Steps

- Successfully brought together the international community to identify scientific challenges for future research for model development and analysis
- To address specific benchmarking priorities, form small targeted working groups should be formed to research and publish
    - Super site benchmarks for AmeriFlux and FLUXNET
    - Benchmarks for soil carbon turnover, distribution, transport
    - Metrics for extreme events & response of ecosystems
    - Data for vegetation recruitment, growth, mortality, phenology, canopy structure
    - Benchmarks for critical high latitude & tropical ecosystems
    - Leverage field projects & remote sensing methods
- A top priority is supporting CMIP6 activities with additional ILAMB development for automated analysis and model–data intercomparison
- Support expanded use of Earth system models

Thank you!