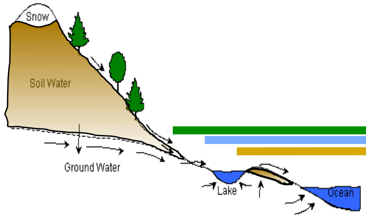
International LAnd Model Benchmarking (ILAMB) Project

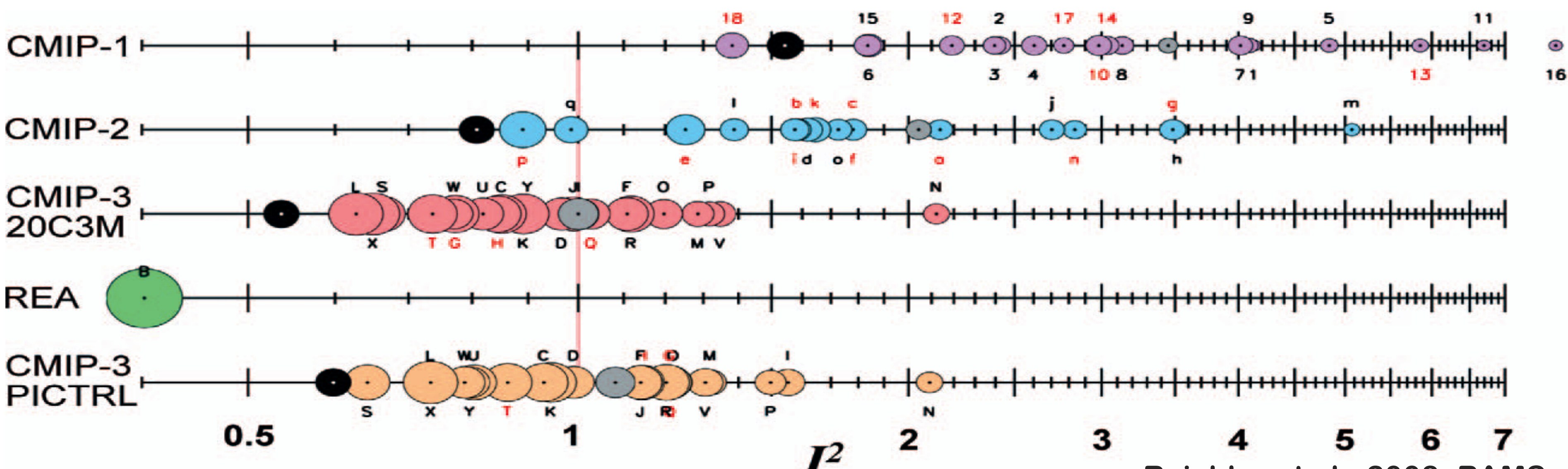
Jim Randerson, Forrest Hoffman, Bill Riley, Dave Lawrence, **Mingquan Mu**, Charlie Koven, Gretchen Keppel Aleks, Nate Collier

U.S. DEPARTMENT OF ENERGY

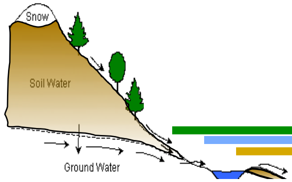


ILAMB Goals

- Develop benchmarks/metrics for land model performance, with emphasis on breadth (carbon cycle, ecosystem, surface energy, and hydrological processes)
- Support the design and development of a new, open-source, benchmarking software system for diagnostics and MIPs
- Strengthen linkages between experimental, monitoring, remote sensing, and climate modeling communities in design of model tests and new measurement programs



Reichler et al., 2008, BAMS



International LAnd Model Benchmarking (ILAMB) project
Variables and Visualization within ILAMB

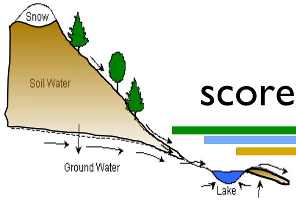
- Currently integrates analysis of 25 variables in 4 categories from ~60 datasets
  - Above ground live biomass, burned area, carbon dioxide, gross primary production, leaf area index, global net ecosystem carbon balance, net ecosystem exchange, ecosystem respiration, soil carbon
  - evapotranspiration, latent heat, sensible heat, runoff, evaporative fraction, terrestrial water storage anomaly
  - albedo, surface upward SW radiation, surface net SW radiation, surface upward LW radiation, surface net LW radiation, surface net radiation
  - surface air temperature, precipitation, surface relative humidity, surface downward SW radiation, surface downward LW radiation
- Graphics and scoring system
  - annual mean, bias, relative bias, RMSE, seasonal cycle phase, spatial distribution, interannual variability, variable-to-variable
  - Global maps, time series plots averaged over specific regions, individual measurement sites, functional relationships

Global Variables

Global Variables (Info for Weightings)

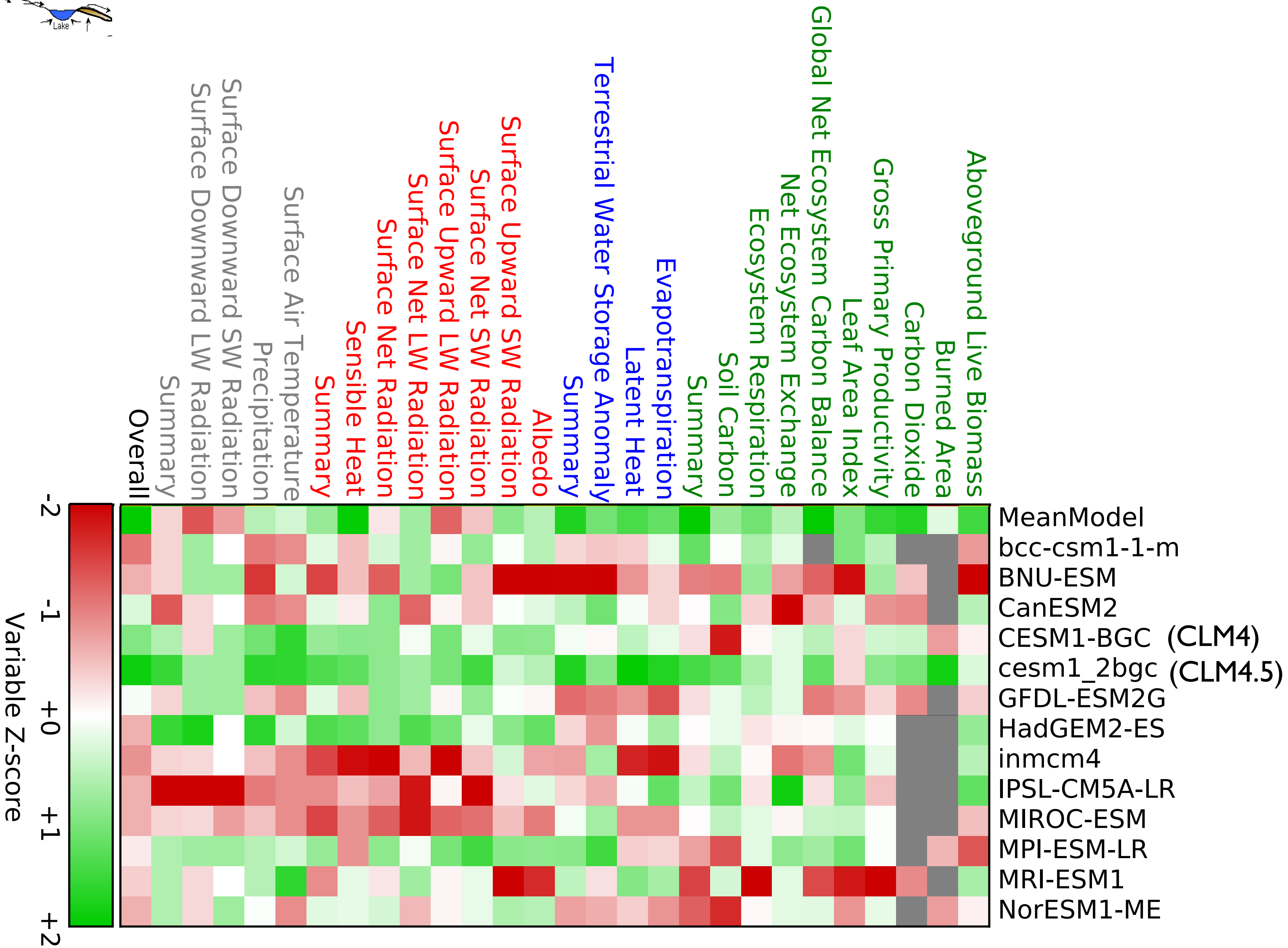
	MeanModel	bcc-csm1-1-m	BNU-ESM	CanESM2	CESM1-BGC	GFDL-ESM2G	HadGEM2-ES	inmcm4	IPSL-CM5A-LR	MIROC-ESM	MPI-ESM-LR	MRI-ESM1	NorESM1-ME
Live Biomass Carbon	0.73	0.68	0.33	0.65	0.60	0.62	0.72	0.50	0.56	0.62	0.58	0.56	0.57
Burned Area	0.38	-	-	-	0.37	-	-	-	-	-	0.38	-	0.38
Carbon Dioxide	0.85	-	0.65	0.65	0.78	0.65	-	-	-	0.79	0.68	0.68	0.75
Gross Primary Productivity	0.77	0.72	0.73	0.64	0.70	0.67	0.68	0.70	0.67	0.69	0.69	0.53	0.70
Leaf Area Index	0.66	0.66	0.41	0.60	0.53	0.49	0.59	0.68	0.66	0.62	0.68	0.43	0.50
Global Net Ecosystem Carbon Balance	0.58	-	0.38	0.27	0.38	0.18	-	0.46	0.25	0.38	0.42	0.27	0.40
Net Ecosystem Exchange	0.49	0.47	0.47	0.39	0.48	0.49	0.46	0.44	0.53	0.48	0.50	0.48	0.48
Ecosystem Respiration	0.75	0.72	0.72	0.65	0.67	0.71	0.66	0.70	0.67	0.68	0.68	0.47	0.66
Soil Carbon	0.55	0.50	0.42	0.56	0.38	0.51	0.51	0.53	0.57	0.53	0.41	0.53	0.39
Summary	0.64	0.62	0.51	0.55	0.55	0.54	0.60	0.56	0.55	0.59	0.55	0.50	0.54
Evapotranspiration	0.75	0.73	0.72	0.72	0.73	0.70	0.74	0.69	0.75	0.70	0.73	0.73	0.72
Evaporative Fraction	0.84	0.76	0.77	0.81	0.81	0.75	0.81	0.81	0.72	0.75	0.75	0.80	0.79
Latent Heat	0.80	0.76	0.77	0.77	0.78	0.74	0.77	0.72	0.77	0.75	0.76	0.78	0.76
Runoff	0.61	0.59	0.60	0.58	0.64	0.59	-	0.62	0.57	0.56	0.66	0.70	0.62
Sensible Heat	0.76	0.69	0.70	0.71	0.75	0.69	0.75	0.66	0.69	0.69	0.69	0.72	0.72
Terrestrial Water Storage Anomaly	0.38	0.37	0.36	0.38	0.38	0.38	-	0.38	0.37	0.38	0.38	0.38	0.38
Summary	0.68	0.65	0.65	0.66	0.67	0.64	0.77	0.64	0.64	0.63	0.66	0.68	0.66
Albedo	0.72	0.71	0.61	0.71	0.73	0.69	0.74	0.67	0.71	0.67	0.73	0.64	0.72
Surface Upward SW Radiation	0.77	0.74	0.67	0.74	0.78	0.74	0.77	0.74	0.73	0.72	0.78	0.67	0.76
Surface Net SW Radiation	0.84	0.86	0.84	0.85	0.86	0.86	0.86	0.84	0.82	0.83	0.87	0.85	0.85
Surface Upward LW Radiation	0.89	0.91	0.91	0.91	0.92	0.91	0.92	0.89	0.90	0.91	0.92	0.91	0.91
Surface Net LW Radiation	0.81	0.82	0.81	0.79	0.81	0.81	0.83	0.80	0.78	0.78	0.81	0.81	0.81
Surface Net Radiation	0.78	0.79	0.76	0.80	0.80	0.81	0.80	0.74	0.77	0.77	0.81	0.78	0.80
Summary	0.80	0.80	0.77	0.80	0.81	0.80	0.82	0.77	0.78	0.78	0.82	0.78	0.81
Surface Air Temperature	0.87	0.87	0.85	0.85	0.88	0.85	0.87	0.85	0.87	0.85	0.88	0.88	0.87
Precipitation	0.71	0.69	0.67	0.69	0.72	0.69	0.73	0.69	0.69	0.69	0.72	0.70	0.70
Surface Relative Humidity	0.81	-	0.80	0.76	0.82	-	-	0.79	0.82	-	-	0.83	0.81
Surface Downward SW Radiation	0.86	0.88	0.87	0.87	0.88	0.87	0.87	0.87	0.83	0.86	0.88	0.86	0.88
Surface Downward LW Radiation	0.89	0.92	0.91	0.91	0.92	0.92	0.92	0.90	0.89	0.91	0.93	0.91	0.91
Summary	0.82	0.83	0.81	0.80	0.83	0.82	0.84	0.81	0.81	0.82	0.84	0.83	0.82
Overall	0.69	0.54	0.59	0.61	0.64	0.57	0.48	0.58	0.57	0.59	0.61	0.59	0.63

Notes: 4 Categories are divided: Ecosystem and Carbon Cycle, Hydrology and Turbulent Flux, Radiation and Energy Cycle, and Forcings.

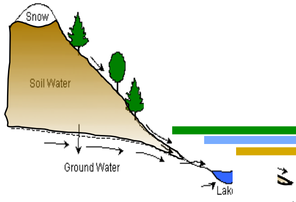


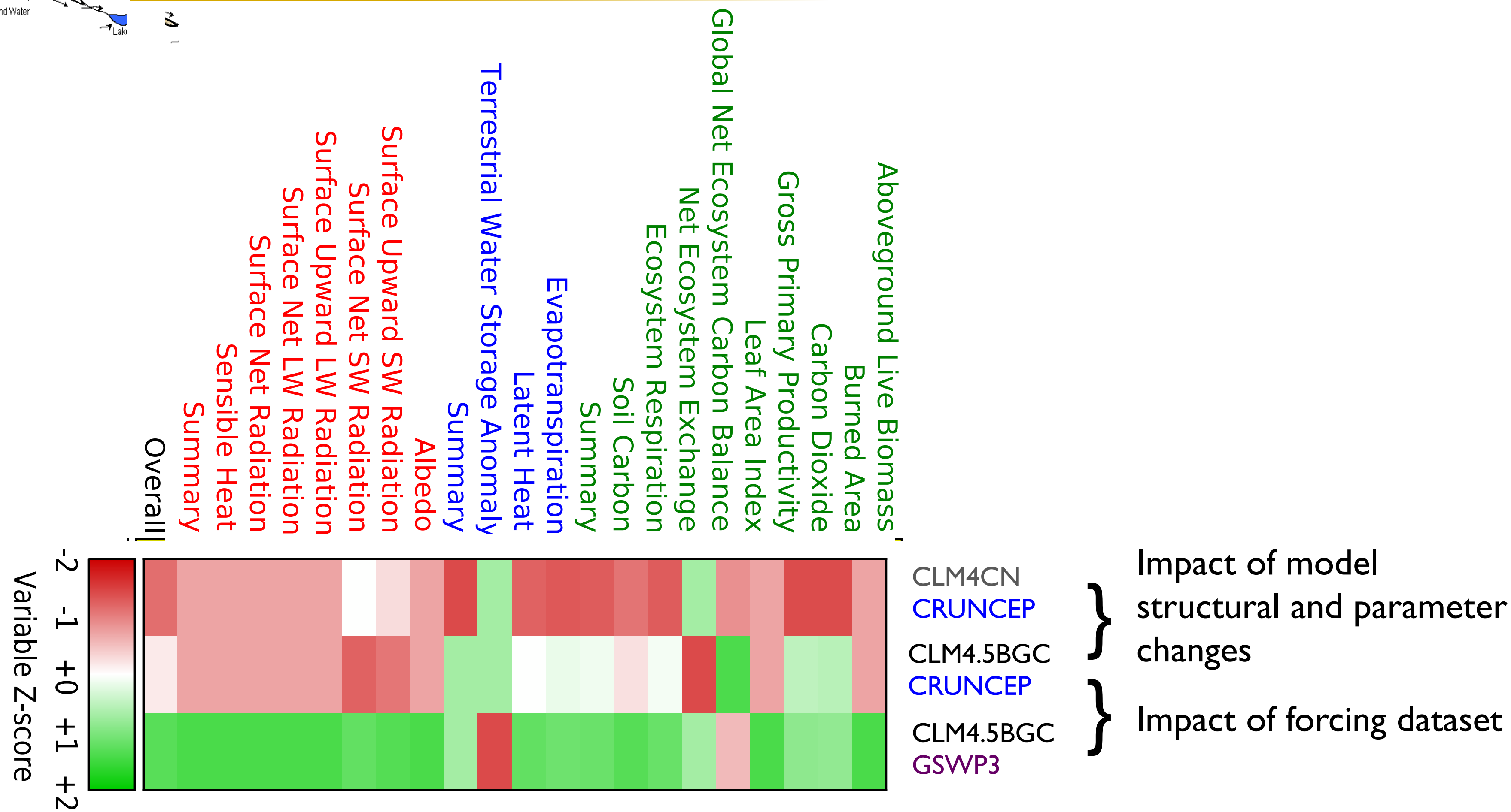
International LAnd Model Benchmarking (ILAMB) package

scores for RMSE, interannual variability, pattern correlation, variable-to-variable comparisons, +

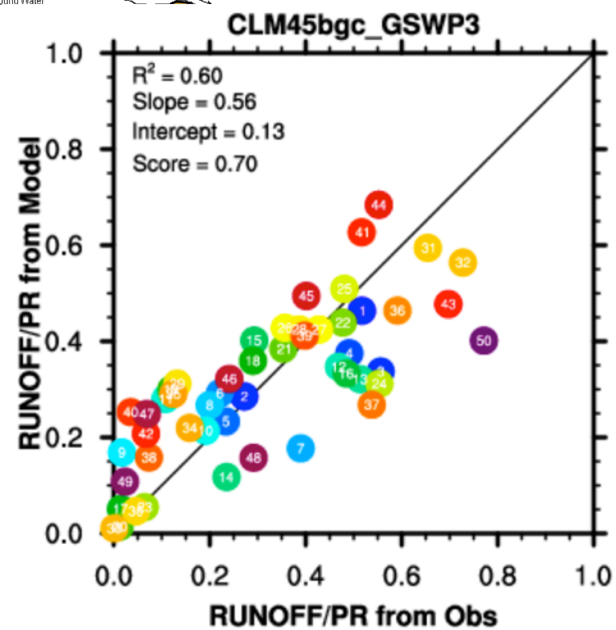


Utilizing ILAMB in model development and assessment process

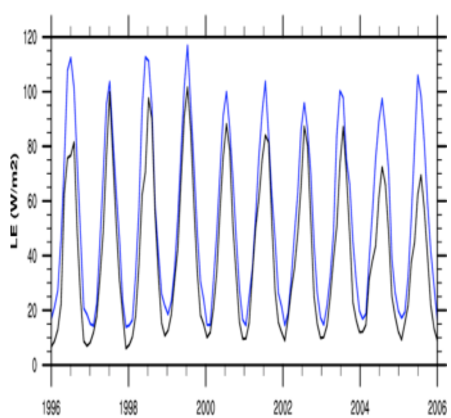




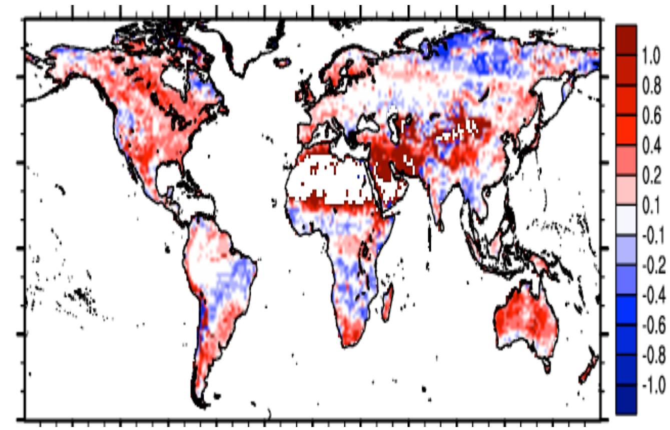
Examples of ILAMB metrics / plots



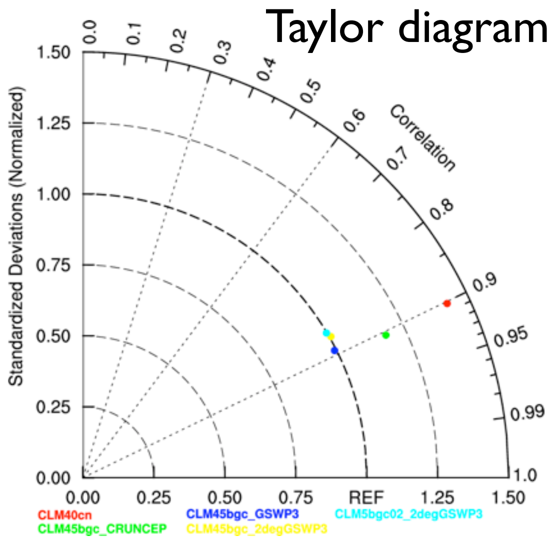
Tower Site



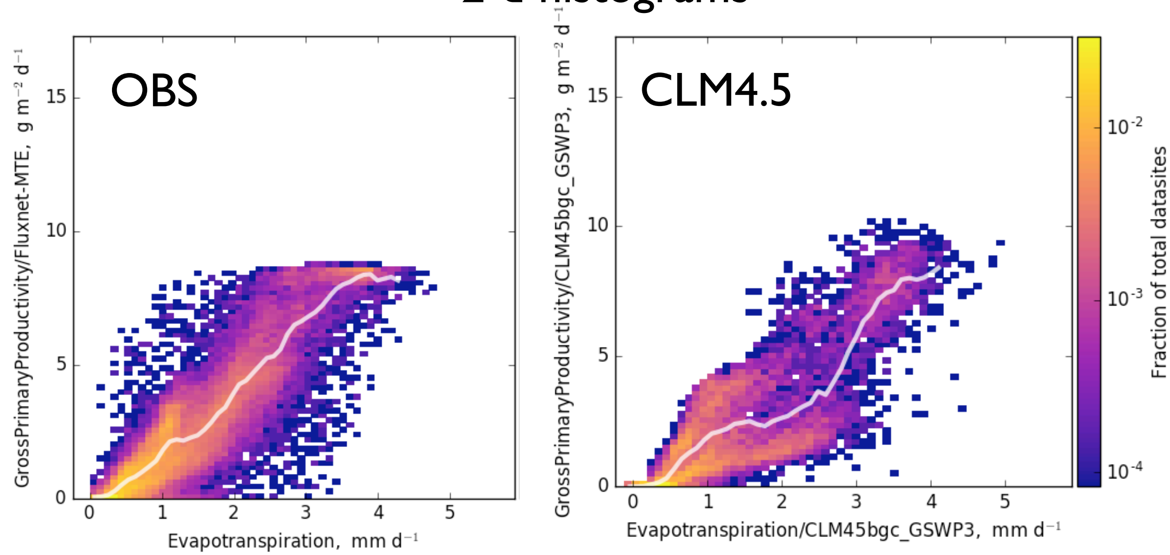
Global bias, relative bias, RMSE



Taylor diagram



2-d histograms



ILAMB versions 1 and 2 are available

- Version 1 – written in NCL
  - http://redwood.ess.uci.edu/mingquan/www/ILAMB/index.html
  - Tuned and vetted versions working with CMIP5 historical, CMIP5 esmHistorical, and CLM development branches

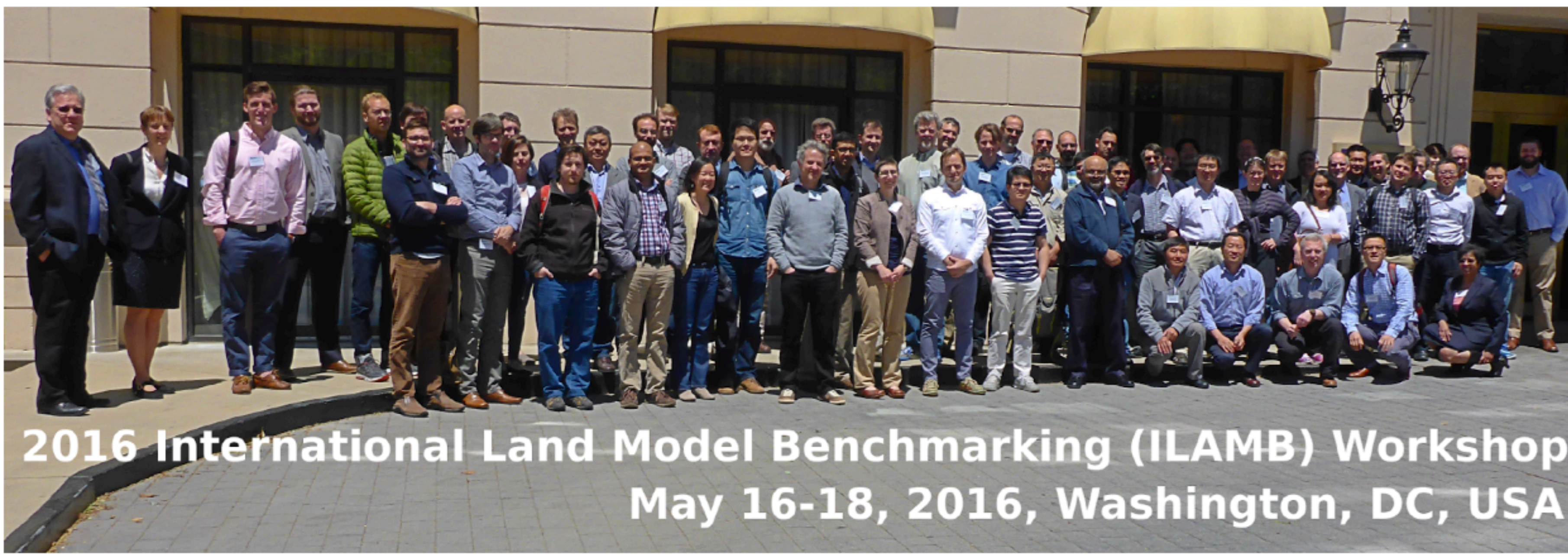
- Version 2 – written in Python and is parallel
  - Hosted in a git repository: https://bitbucket.org/ncollier/ilamb
  - Tutorial:: http://climate.ornl.gov/~ncf/ILAMB/docs/index.html
  - Sample output: http://www.climatemodeling.org/~nate/ILAMB/index.html

- Both versions have the following features:
  - constructed with a modular structure, so that new models, variables or benchmarks can be easily added
  - High quality output files (encapsulated postscript files) can be used directly for publications or proposals.

2016 International Land Model Benchmarking (ILAMB) Workshop

Report of an international workshop held in Washington, DC, USA, May 16–18, 2016

Supported by the US Department of Energy Office of Science, Biological and Environmental Research

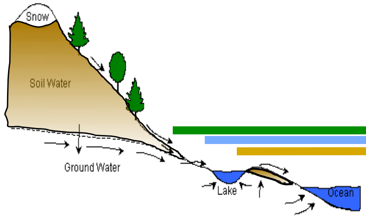


This report is dedicated to

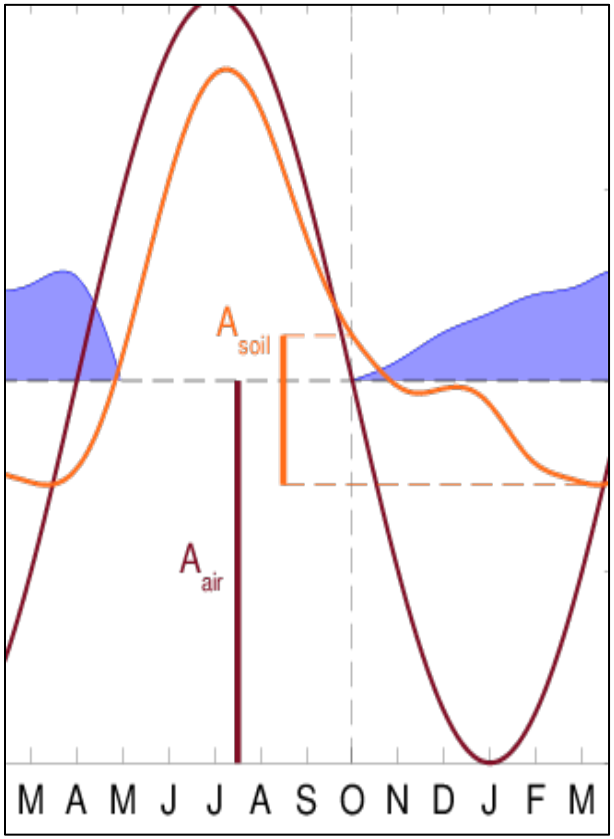
Dr. Andrew G. Slater

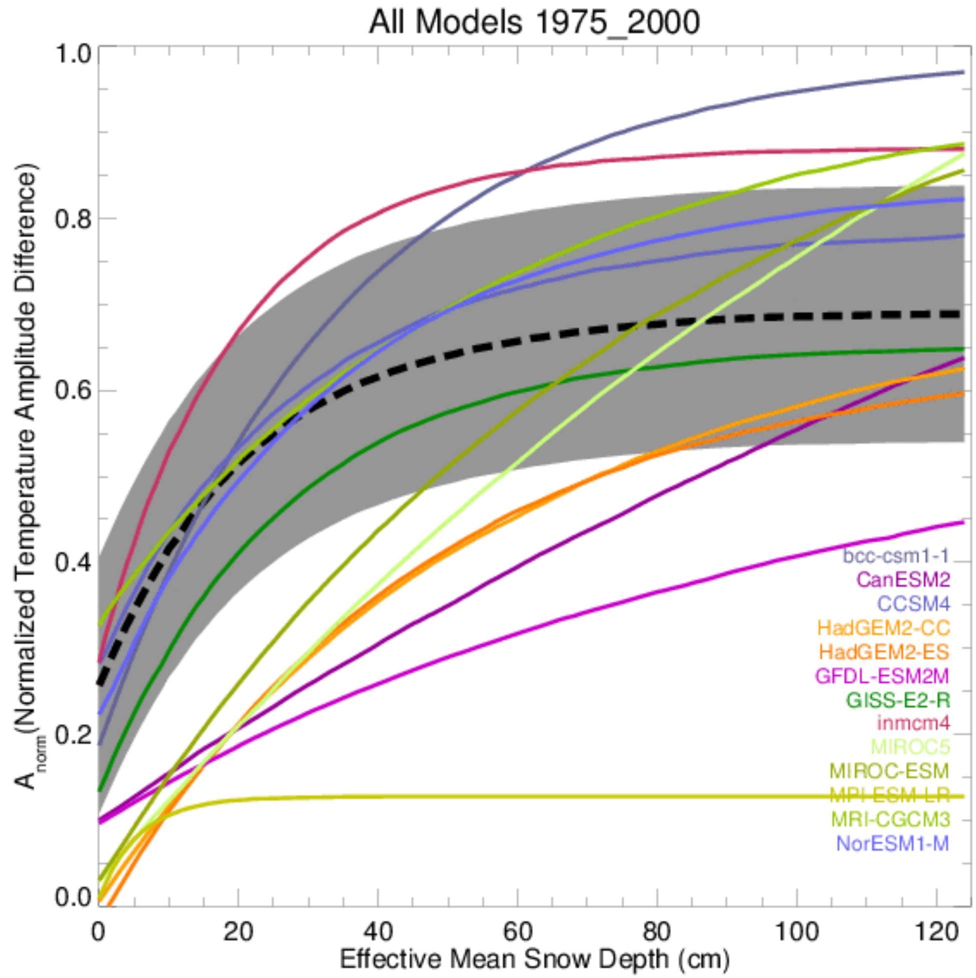
— A brilliant scientist, workshop contributor, and friend to many.

Drew left this world before his time in September 2016.

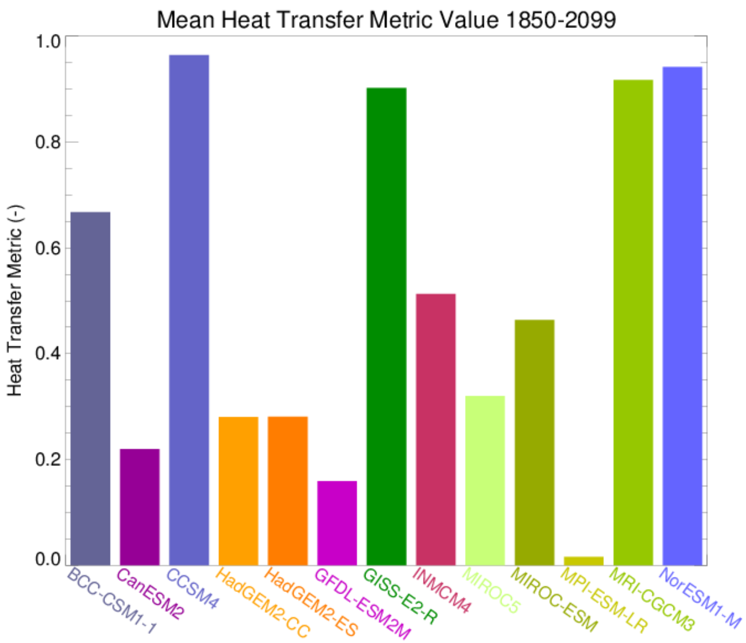


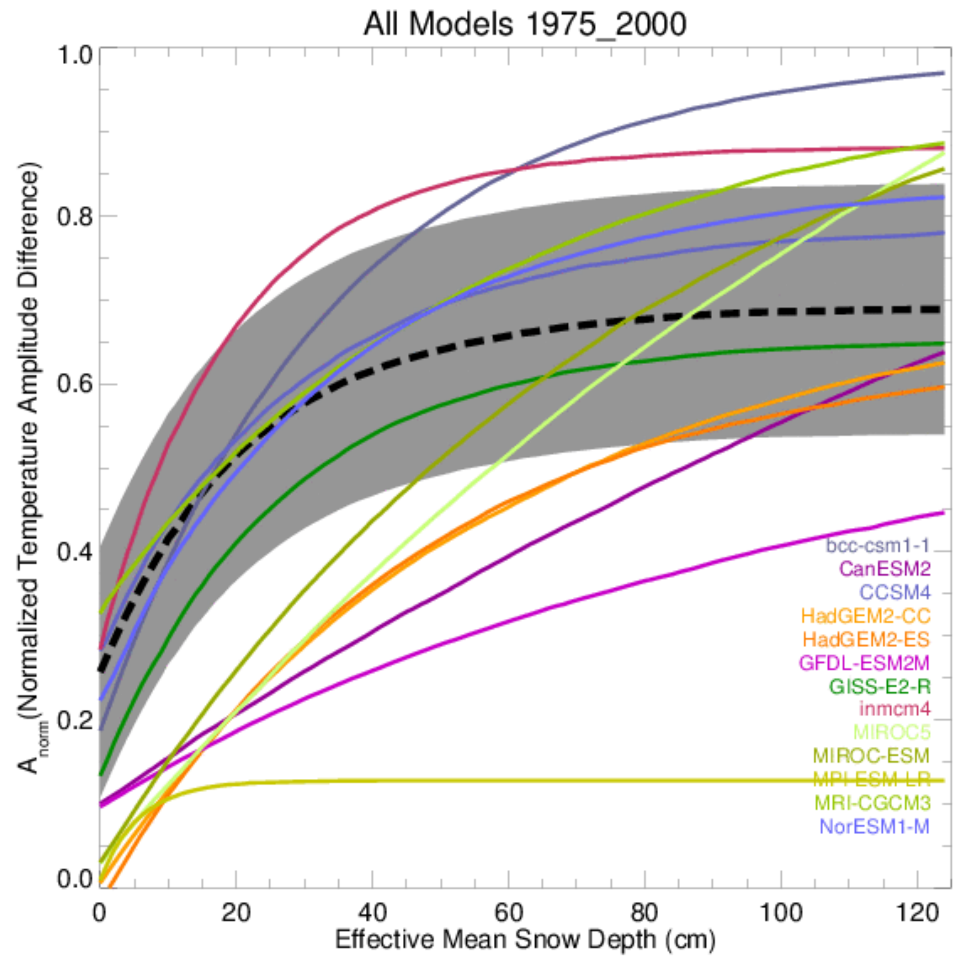
Process-oriented metrics: heat transfer through snow





Process-oriented metrics: heat transfer through snow

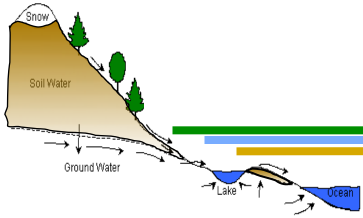




ILAMB Workshop Report

KEY RECOMMENDATIONS

- Well-established aspects of model assessment should be a routine component of the model development process that over time becomes increasingly comprehensive.
- Evaluation tools should include testing the predictive power of models under a changing climate.
- Benchmarking packages should span a wide range of spatial and temporal scales and extents.
- Integration of a diversity of evaluation tools into a common workflow framework could lead to new insights into climate processes and phenomena.
- Evaluation and benchmarking systems should be open source and freely distributed to leverage the work of many modeling teams and to minimize redundancy.
- Benchmarking tools should be integrated with data repositories that support standardized access through an applications programming interface.



Synergies across land model assessment activities

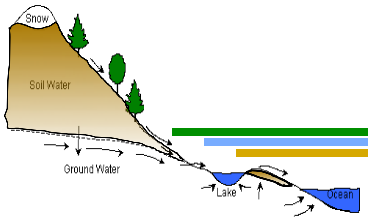
- Coordination of these distinct and international land model benchmarking/ assessment activities is challenging due to the diversity of approaches and the complexities of the international funding environment.
- Over the longer term, it may be possible and beneficial to integrate existing land diagnostics packages (ILAMB, PALS, LVT, etc) under a loosely coordinated framework (potentially similar to ESMValTool).
- Benefit of a reduction of effort related to the overhead of benchmarking (e.g., workflow processes such as reading in, processing, and reformatting model and observational data), allowing more effort and funding to be devoted to metrics development.
- Should explore potential for a joint benchmarking analysis project, wherein each of the existing packages is applied to a set of multi-model output that would enable direct comparison and evaluation of how each package uniquely contributes to our understanding of model strengths and weaknesses.

Summary

- ILAMB may be a useful tool for model development and assessment
  - Along with tower site simulations, other diagnostics packages, scientific insight and intuition, case studies, etc.
- Provides quick and comprehensive comparison against growing set of observations and metrics
- Future development of ILAMB to enhance utility in model development
  - Emergent constraints
  - Land-atmosphere coupling metrics
  - Experimental manipulations (N-addition, rainfall exclusion, etc)
  - Develop and integrate arctic and tropical ecosystems modules
  - Prepare for CMIP6

Potential metrics for inclusion in a comprehensive land benchmarking/metric system

- Large-scale state and flux estimates
  - LH, SH, total water storage, albedo, river discharge, SCF, LAI, soil and veg C stocks, GPP, NEE, ER, burnt area, permafrost distribution, T_{2m}, P, …
  - RMSE, spatial pattern corr, interannual variance, annual cycle phase, trends
- Functional relationships and emergent properties
  - soil moisture – ET, soil moisture – runoff, precip – GPP, stomatal response to VPD, precip – burnt area, transient carbon storage trajectory, runoff ratio, spring albedo transition
- Experimental manipulation (testing model functional responses)
  - Nitrogen additions, FACE, artificial warming, rainfall exclusion, ecosystem response to disturbance



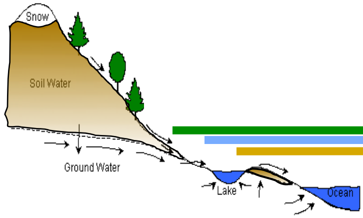
Summary

- ILAMB useful tool for model development and assessment
  - Along with tower site simulations, other diagnostics packages, scientific insight and intuition, case studies, etc.
- Provides quick and comprehensive comparison against growing set of observations and metrics
- Future development of ILAMB to enhance utility in model development
  - Parallelization
  - Compare against years outside observational period (e.g. 1850 control)

Utilizing ILAMB in model development and assessment process

Global Variables (Info for Weightings)

	CLM45bgc_2degGSWP3	CLM5bgc01_2degGSWP3
Aboveground Live Biomass	0.71	0.64
Burned Area	0.51	0.42
Gross Primary Productivity	0.75	0.72
Leaf Area Index	0.57	0.58
Global Net Ecosystem Carbon Balance	0.47	0.45
Net Ecosystem Exchange	0.49	0.51
Ecosystem Respiration	0.73	0.70
Soil Carbon	0.56	0.58
Summary	0.60	0.58

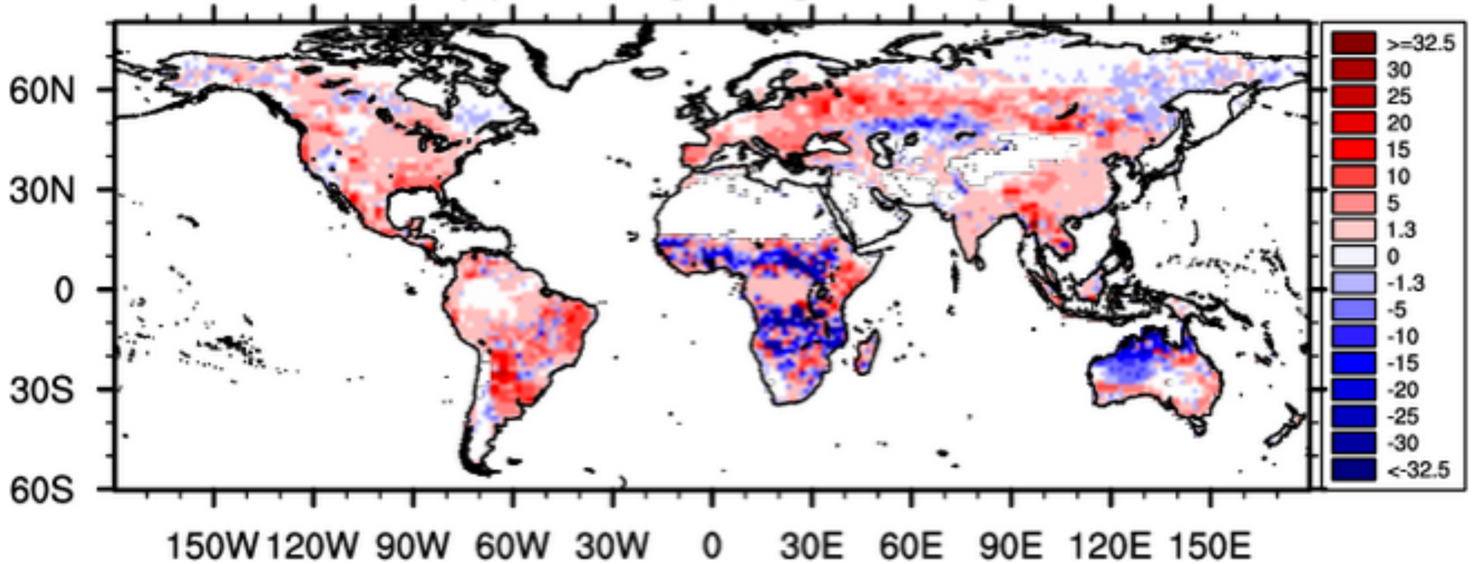


Diagnostic Summary for Burned Area: Model vs. GFED3

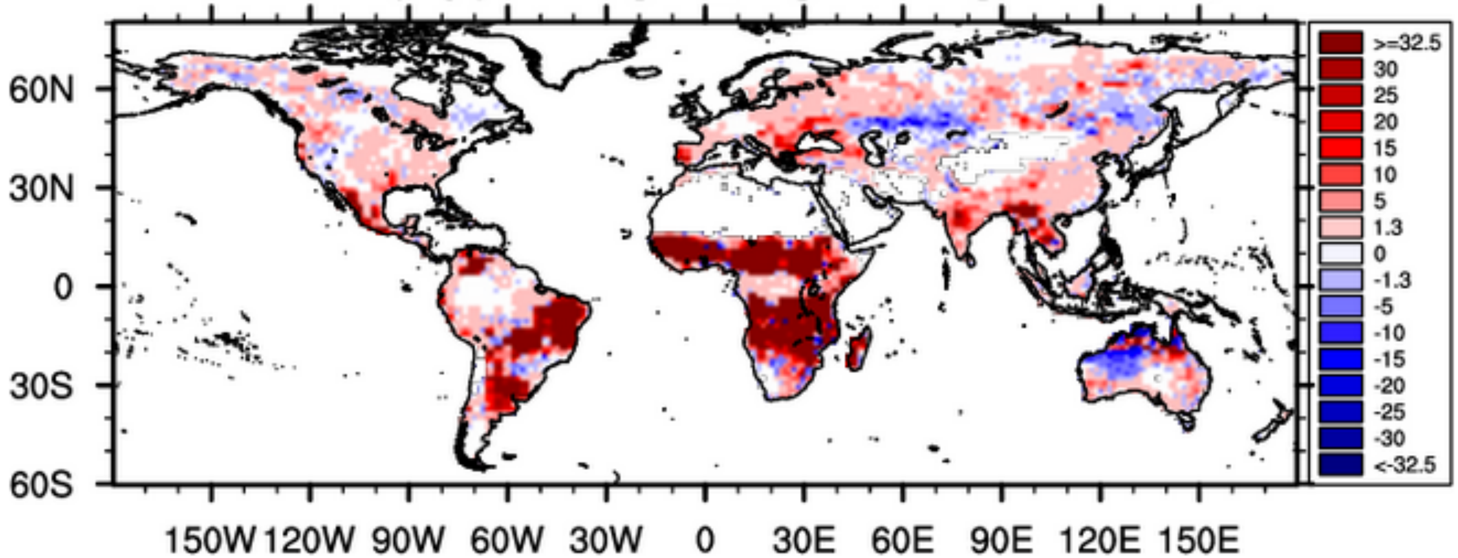
	Global Patterns				Regional and Seasonal Patterns	Scoring (Info)					
	Annual Mean (Mha/yr)	Bias (Mha/yr)	RMSE (Mha/mon)	Phase Difference (months)	Regional Means	Global Bias	RMSE	Seasonal Cycle	Spatial Distribution	Interannual Variability	Overall
Benchmark [Giglio et al. (2010)]	362.8	-	-	0.0	access to **plots**	-	-	-	-	-	-
CLM45bgc_2degGSWP3	378.8	16.1	85.5	1.6	access to **plots**	0.52	0.40	0.72	0.48	0.53	0.51
CLM5bgc01_2degGSWP3	1578.9	1216.1	208.9	0.5	access to **plots**	0.32	0.27	0.86	0.26	0.52	0.42

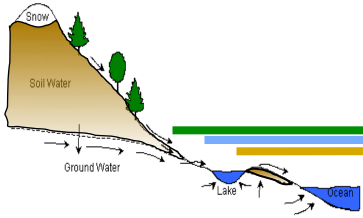
Notes: In calculating overall score, rmse score contributes double in comparison with all other scores.

Bias for BURNTAREA (%/yr): CLM45bgc_2degGSWP3 against GFED3, 1997-2010



Bias for BURNTAREA (%/yr): CLM5bgc01_2degGSWP3 against GFED3, 1997-2010

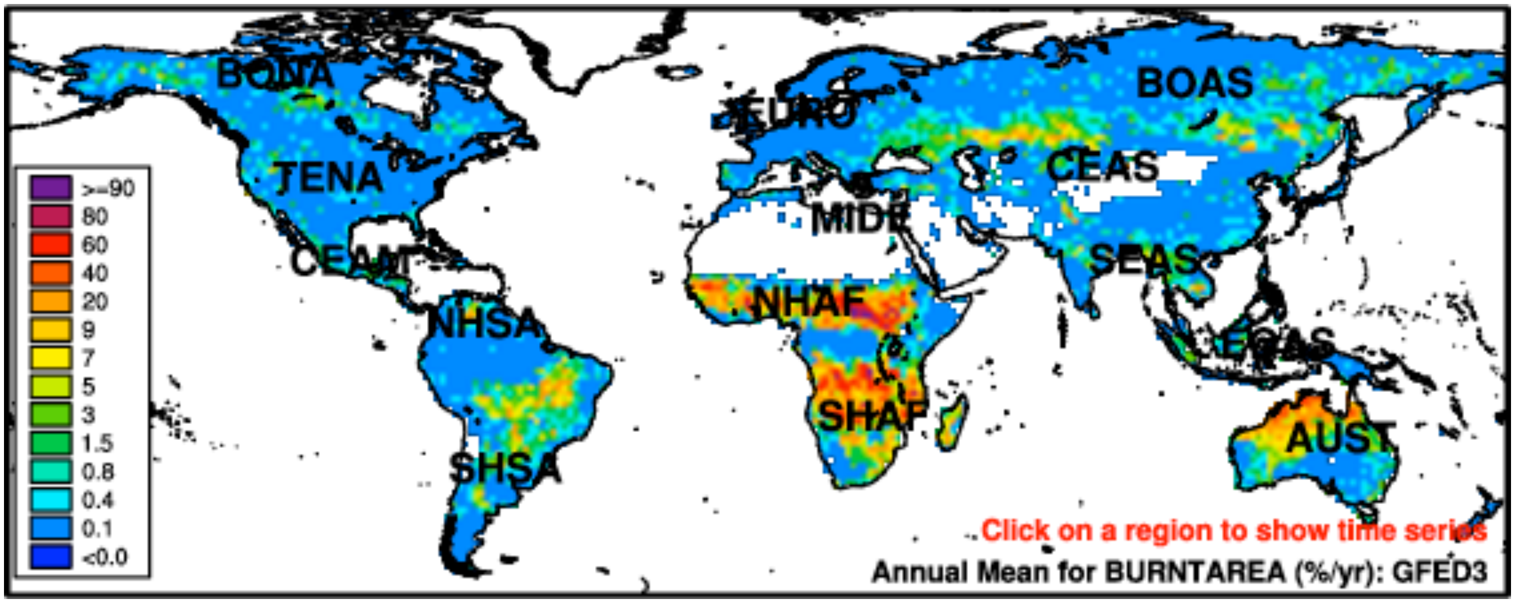


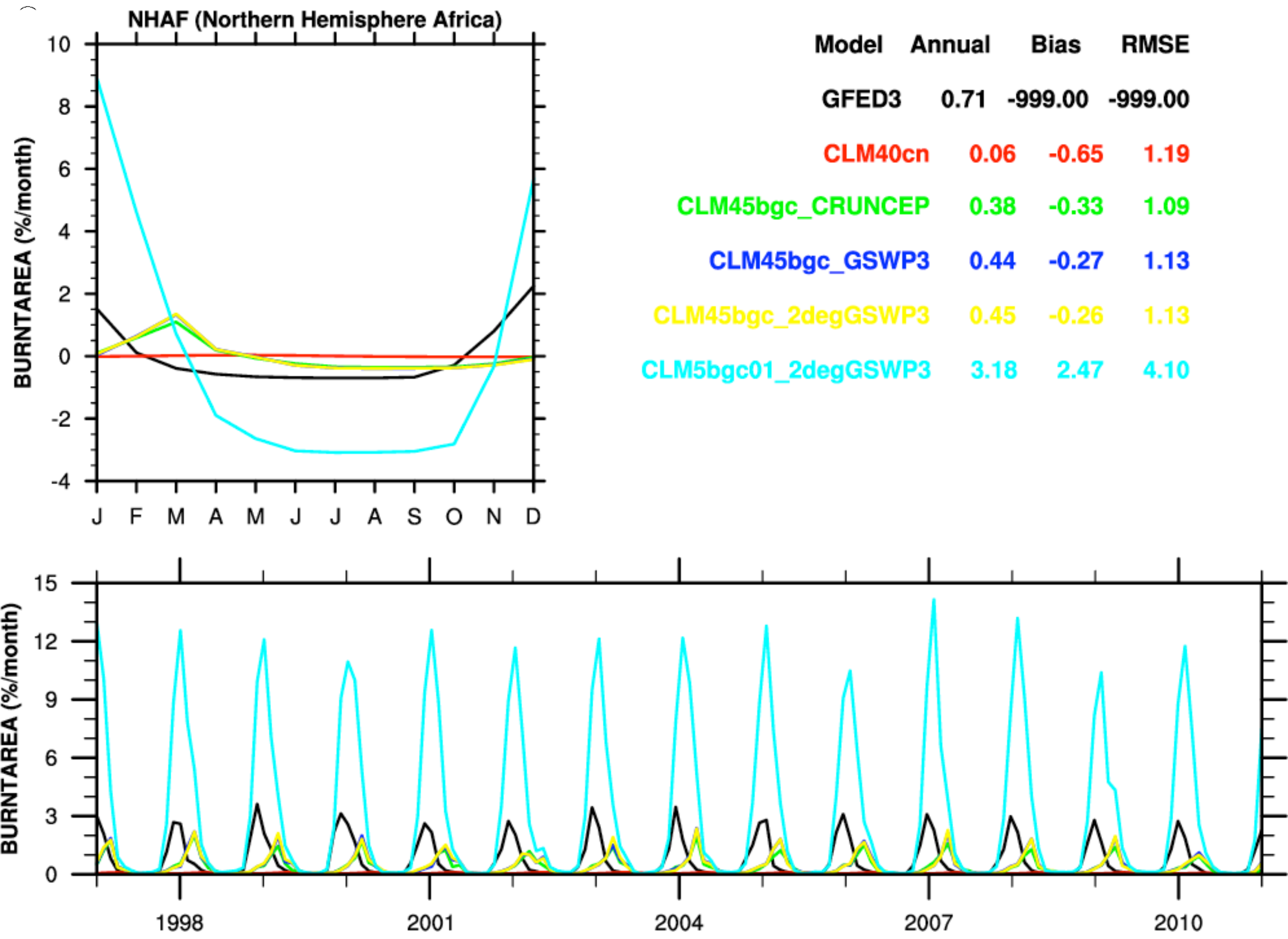


Diagnostic Summary for Burned Area: Model vs. GFED3

	Global Patterns				Regional and Seasonal Patterns	Scoring (Info)					
	Annual Mean (Mha/yr)	Bias (Mha/yr)	RMSE (Mha/mon)	Phase Difference (months)	Regional Means	Global Bias	RMSE	Seasonal Cycle	Spatial Distribution	Interannual Variability	Overall
Benchmark [Giglio et al. (2010)]	362.8	-	-	0.0	access to plots	-	-	-	-	-	-
CLM45bgc_2degGSWP3	378.8	16.1	85.5	1.6	access to plots	0.52	0.40	0.72	0.48	0.53	0.51
CLM5bgc01_2degGSWP3	1578.9	1216.1	208.9	0.5	access to plots	0.32	0.27	0.86	0.26	0.52	0.42

Notes: In calculating overall score, rmse score contributes double in comparison with all other scores.

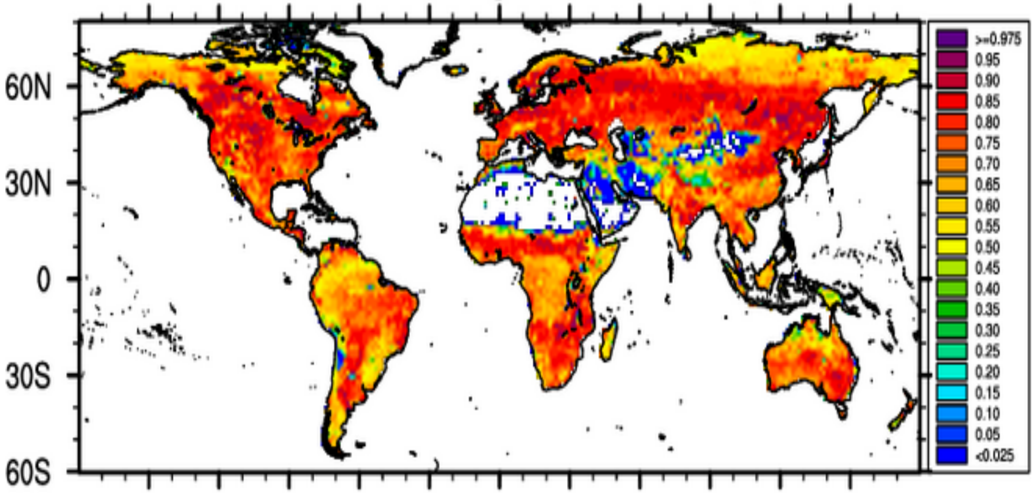

BONA
TENA
CEAM
NHSA
SHSA
EURO
MIDE
NHAF
SHAF
BOAS
CEAS
SEAS
EQAS
AUST
>=90
80
60
40
20
9
7
5
3
1.5
0.8
0.4
0.1
<0.0
Click on a region to show time series
Annual Mean for BURNTAREA (%/yr): GFED3



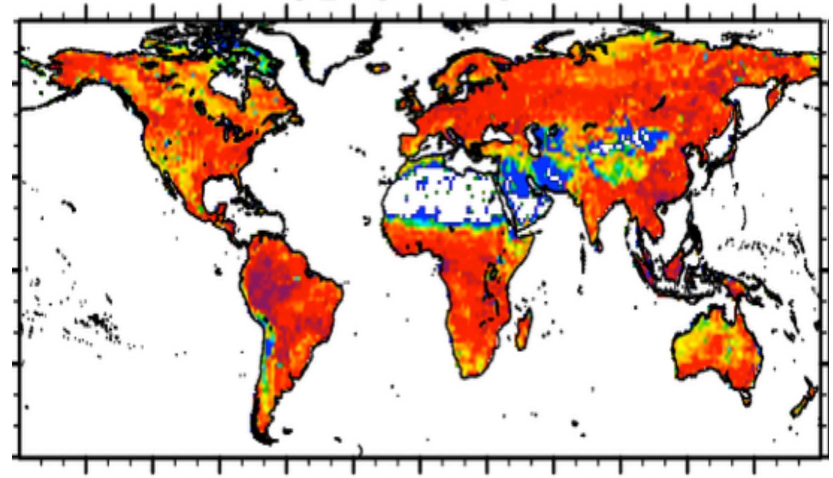
NHAF (Northern Hemisphere Africa)
BURNTAREA (%/month)
10
8
6
4
2
0
-2
-4
J F M A M J J A S O N D
Model Annual Bias RMSE
GFED3 0.71 -999.00 -999.00
CLM40cn 0.06 -0.65 1.19
CLM45bgc_CRUNCEP 0.38 -0.33 1.09
CLM45bgc_GSWP3 0.44 -0.27 1.13
CLM45bgc_2degGSWP3 0.45 -0.26 1.13
CLM5bgc01_2degGSWP3 3.18 2.47 4.10
BURNTAREA (%/month)
15
12
9
6
3
0
1998
2001
2004
2007
2010

RMSE Score for Latent Heat: FLUXNET-MTE

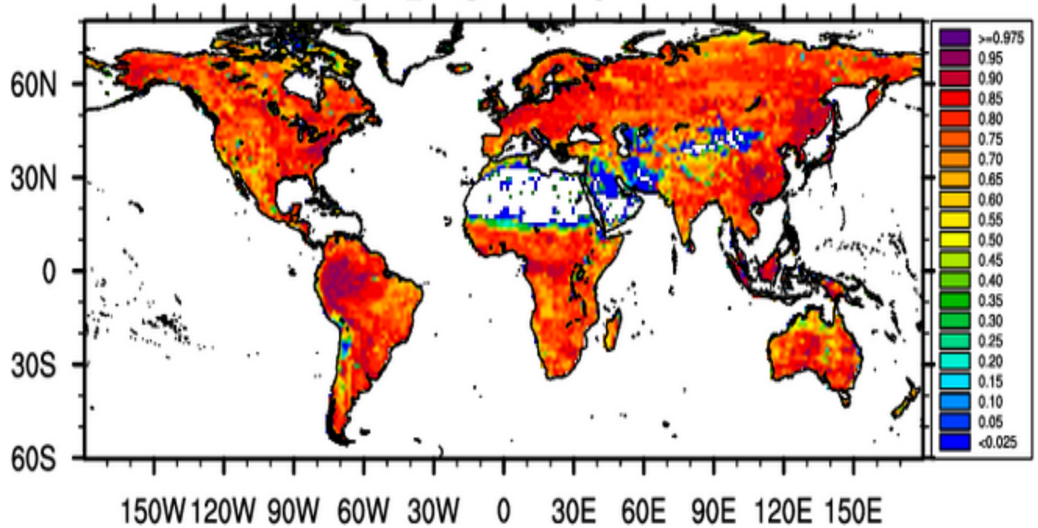
RMSE Score for LE: CLM40cn against FLUXNET-MTE, 1982-2008

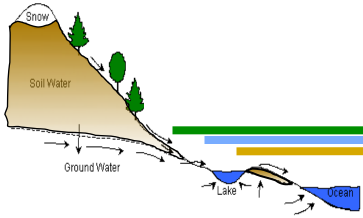


;E Score for LE: CLM45bgc_2degGSWP3 against FLUXNET-MTE, 1982-2

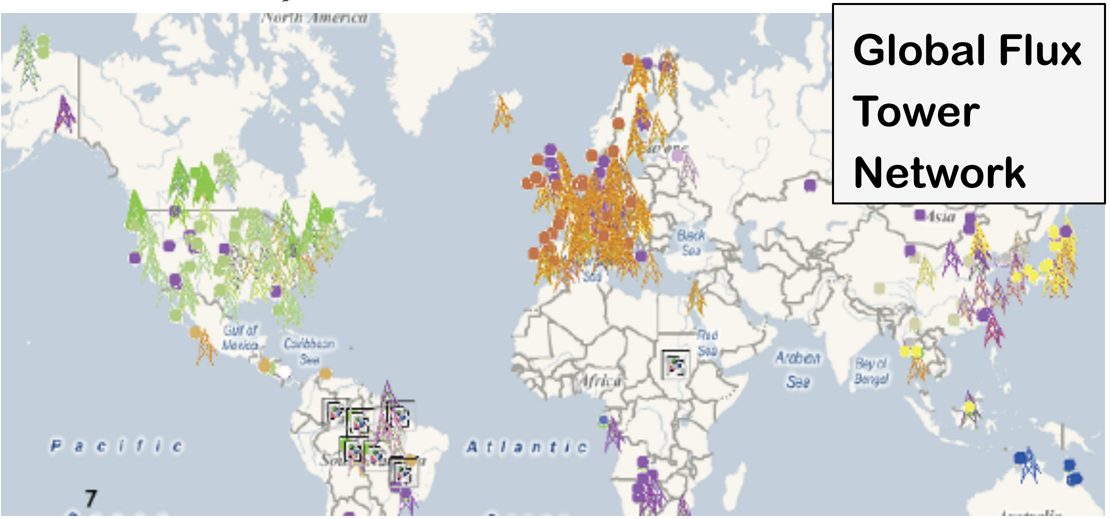


RMSE Score for LE: CLM5bgc01_2degGSWP3 against FLUXNET-MTE, 1982-2008

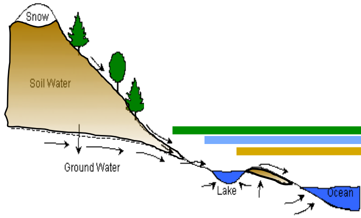




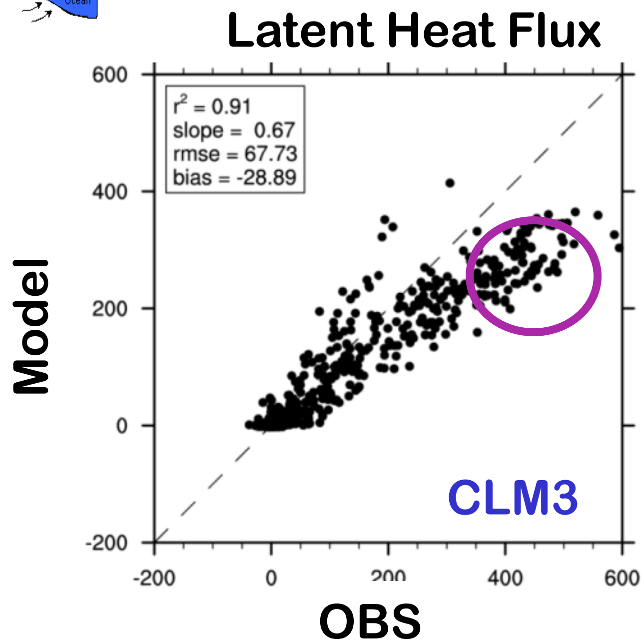
Evaluating and Improving the model with Tower Flux data

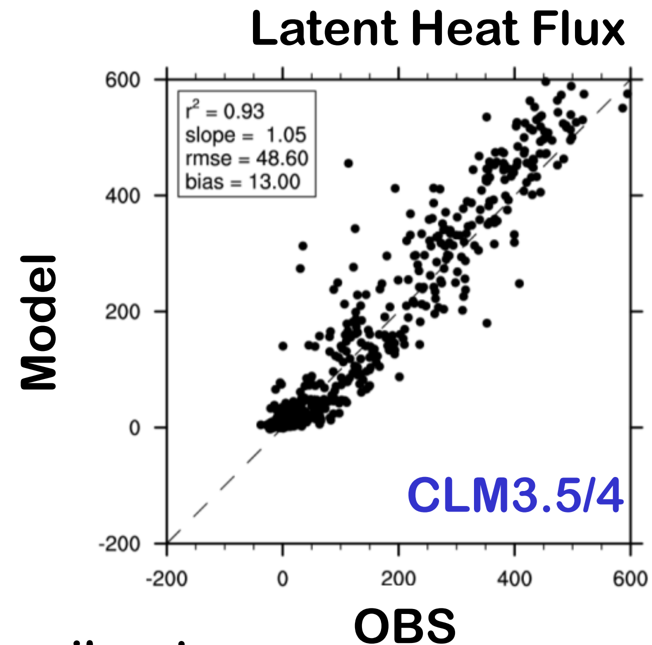


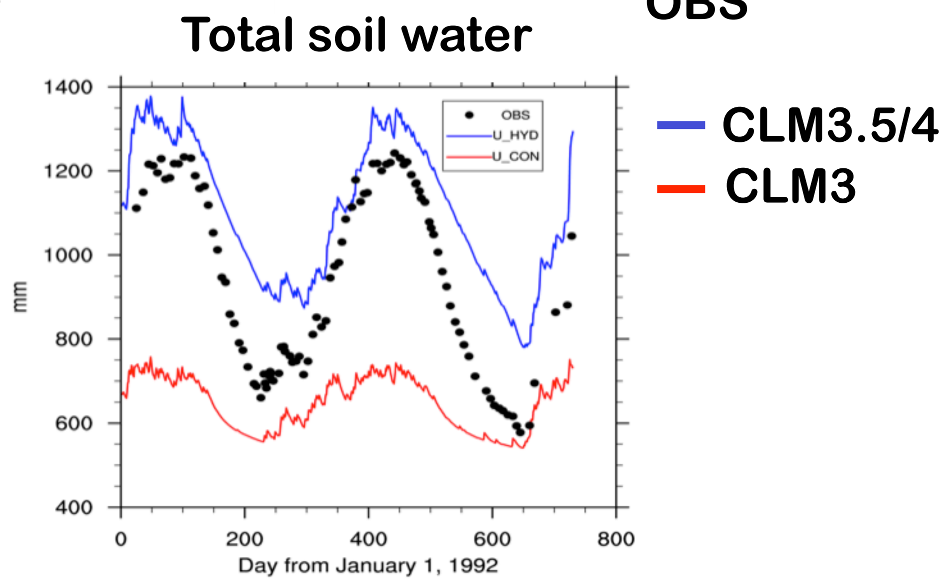
Global Flux Tower Network

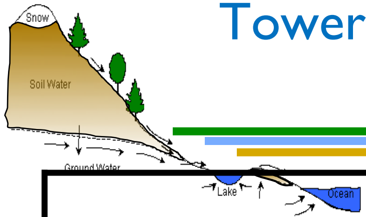


Abracos tower site (Amazon)









Tower flux statistics (15 sites incl. tropical, boreal, mediterannean, alpine, temperate; hourly)

	Latent Heat Flux		Sensible Heat Flux	
	r	RMSE (W/m^2)	r	RMSE (W/m^2)
CLM3	0.54	72	0.73	91
CLM3.5	0.80	50	0.79	65
CLM4SP	0.80	48	0.84	58