MsTMIP Model Intercomparison Framework

Christopher R Schwalm

DataONE – ILAMB Working Group:

Exploration, Visualization, and Analysis (EVA)

November 16, 2011

Outline

1. Big science questions
2. Traits
3. Metrics
4. Benchmarks
5. Model development
6. Challenges

Science → Benchmarking

Tier 1: Big Science Questions

What are the most dominant controls on land-atmosphere carbon exchange?

Tier 2: Key Outcomes

Quantify how land cover, CO_2 fertilization, nitrogen deposition, and climate modulate interannual variability of terrestrial carbon sink.

Tier 3: Science-Validation Interface

Models as integrators of biophysical/biogeochemical "knowledge". Assessment of robustness needed to address key outcomes.

Tier 4: Benchmarking

Do models exhibit the "correct" magnitude of carbon fluxes and pools?

Science → Benchmarking

1. Science *leads* benchmarking
2. Benchmarking supports scientific outcomes
3. Not all benchmarks needed for all science questions
4. Buy-in from broader biogeoscience community

Traits + Metrics

- Long-term mean/trend
  - Spatial correlation & bias: calculated across all pixels
  - Latitudinal correlation & bias : zonal mean vector
- Mean monthly cycle
  - Phase offset: difference in maximal month
  - Amplitude: map of maximal value
- Interannual variability
  - Sensitivity to climatic drivers
  - Spatial and latitudinal metrics

Benchmarks

Category	Benchmarks	Model Variable Name	Time
Gridded	Gross Primary Productivity (MODIS)	GPP	2000-2010
	Net Primary Productivity (MODIS)	NPP	2000-2010
	Fire Emissions (GFED)	Fire_flux	1997-2009
	Leaf Area Index (MODIS)	LAI	2000-2010
	Evapotranspiration (MODIS)	Evap	2000-2010
	Shortwave Albedo (MODIS)	SW_albedo	2000-2010
	Gross Primary Productivity	GPP	1982-2008
	Net Ecosystem Exchange	NEE	1982-2008
	Total Respiration	TotalResp	1982-2008
	Sensible Heat	Qh	1982-2008
	Latent Heat	Qle	1982-2008
	Biomass (AGB, total)	AbvGrndWood, TotLivBiom (tropics)	2000
	Biomass (AGB only)	AbvGrndWood (CONUS forest)	2001
	Biomass (AGB, total)	AbvGrndWood, TotLivBiom	2000
	Total Soil Carbon	TotSoilCarb	1950-1995 (mean only)
	Average Layer Soil Moisture (AMSR-E)	SoilMoist	2002-2009
	Total Snow Depth	SnowDepth	1998-2010
	Snow Water Equivalent	SWE	2000-2010
Site	Gross Primary Productivity	GPP	1991-2006 (varies by site)
	Net Ecosystem Exchange	NEE	1991-2006 (varies by site)
	Total Respiration	TotalResp	1991-2006 (varies by site)
	Sensible Heat	Qh	1991-2006 (varies by site)
	Latent Heat	Qle	1991-2006 (varies by site)
	Net Primary Productivity	NPP	1937-2000 (single year per site)
	Net Primary Productivity	NPP	1897-2006 (single year per site)
	Leaf Area Index	LAI	1932-2000 (single year per site)
	Soil Moisture	SoilMoist	1952-2010 (varies by site)
Composite	NH Snow Extent	Area where SnowDepth > 0	1967-2010
	Cropland NPP (CONUS)	NPP from croplands only	1990-2010
	Atmospheric CO2	NEE -> [CO2] by latitudinal zone	Varies by location

Benchmarks

Category	Benchmarks	Model Variable Name	Time
Gridded	Gross Primary Productivity (MODIS)	GPP	2000-2010
	Net Primary Productivity (MODIS)	NPP	2000-2010
	Fire Emissions (GFED)	Fire_flux	1997-2009
	Leaf Area Index (MODIS)	LAI	2000-2010
	Evapotranspiration (MODIS)	Evap	2000-2010
	Shortwave Albedo (MODIS)	SW_albedo	2000-2010
	Gross Primary Productivity	GPP	1982-2008
	Net Ecosystem Exchange	NEE	1982-2008
	Total Respiration	TotalResp	1982-2008
	Sensible Heat	Qh	1982-2008
	Latent Heat	Qle	1982-2008
	Biomass (AGB, total)	AbvGrndWood, TotLivBiom (tropics)	2000
	Biomass (AGB only)	AbvGrndWood (CONUS forest)	2001
	Biomass (AGB, total)	AbvGrndWood, TotLivBiom	2000
	Total Soil Carbon	TotSoilCarb	1950-1995 (mean only)
	Average Layer Soil Moisture (AMSR-E)	SoilMoist	2002-2009
	Total Snow Depth	SnowDepth	1998-2010
	Snow Water Equivalent	SWE	2000-2010
Site	Gross Primary Productivity	GPP	1991-2006 (varies by site)
	Net Ecosystem Exchange	NEE	1991-2006 (varies by site)
	Total Respiration	TotalResp	1991-2006 (varies by site)
	Sensible Heat	Qh	1991-2006 (varies by site)
	Latent Heat	Qle	1991-2006 (varies by site)
	Net Primary Productivity	NPP	1937-2000 (single year per site)
	Net Primary Productivity	NPP	1897-2006 (single year per site)
	Leaf Area Index	LAI	1932-2000 (single year per site)
	Soil Moisture	SoilMoist	1952-2010 (varies by site)
Composite	NH Snow Extent	Area where SnowDepth > 0	1967-2010
	Cropland NPP (CONUS)	NPP from croplands only	1990-2010
	Atmospheric CO2	NEE -> [CO2] by latitudinal zone	Varies by location

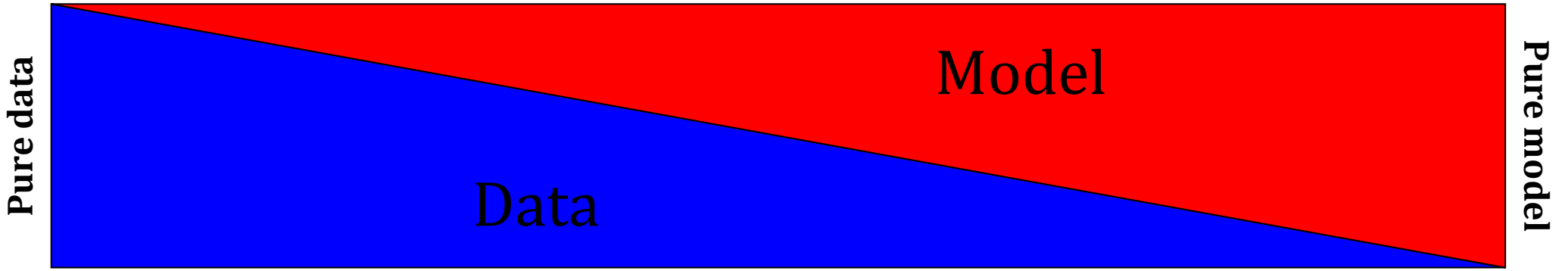
Benchmarks

Category	Benchmarks	Model Variable Name	Time
Gridded	Gross Primary Productivity (MODIS)	GPP	2000-2010
	Net Primary Productivity (MODIS)	NPP	2000-2010
	Fire Emissions (GFED)	Fire_flux	1997-2009
	Leaf Area Index (MODIS)	LAI	2000-2010
	Evapotranspiration (MODIS)	Evap	2000-2010
	Shortwave Albedo (MODIS)	SW_albedo	2000-2010
	Gross Primary Productivity	GPP	1982-2008
	Net Ecosystem Exchange	NEE	1982-2008
	Total Respiration	TotalResp	1982-2008
	Sensible Heat	Qh	1982-2008
	Latent Heat	Qle	1982-2008
	Biomass (AGB, total)	AbvGrndWood, TotLivBiom (tropics)	2000
	Biomass (AGB only)	AbvGrndWood (CONUS forest)	2001
	Biomass (AGB, total)	AbvGrndWood, TotLivBiom	2000
	Total Soil Carbon	TotSoilCarb	1950-1995 (mean only)
	Average Layer Soil Moisture (AMSR-E)	SoilMoist	2002-2009
	Total Snow Depth	SnowDepth	1998-2010
	Snow Water Equivalent	SWE	2000-2010
Site	Gross Primary Productivity	GPP	1991-2006 (varies by site)
	Net Ecosystem Exchange	NEE	1991-2006 (varies by site)
	Total Respiration	TotalResp	1991-2006 (varies by site)
	Sensible Heat	Qh	1991-2006 (varies by site)
	Latent Heat	Qle	1991-2006 (varies by site)
	Net Primary Productivity	NPP	1937-2000 (single year per site)
	Net Primary Productivity	NPP	1897-2006 (single year per site)
	Leaf Area Index	LAI	1932-2000 (single year per site)
	Soil Moisture	SoilMoist	1952-2010 (varies by site)
Composite	NH Snow Extent	Area where SnowDepth > 0	1967-2010
	Cropland NPP (CONUS)	NPP from croplands only	1990-2010
	Atmospheric CO2	NEE -> [CO2] by latitudinal zone	Varies by location

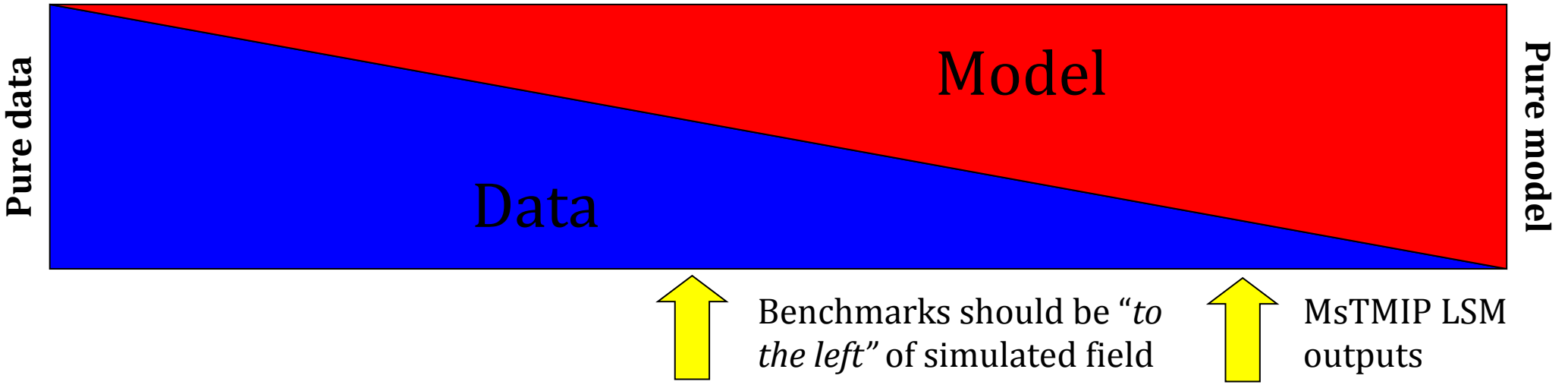
Benchmarks

Category	Benchmarks	Model Variable Name	Time
Gridded	Gross Primary Productivity (MODIS)	GPP	2000-2010
	Net Primary Productivity (MODIS)	NPP	2000-2010
	Fire Emissions (GFED)	Fire_flux	1997-2009
	Leaf Area Index (MODIS)	LAI	2000-2010
	Evapotranspiration (MODIS)	Evap	2000-2010
	Shortwave Albedo (MODIS)	SW_albedo	2000-2010
	Gross Primary Productivity	GPP	1982-2008
	Net Ecosystem Exchange	NEE	1982-2008
	Total Respiration	TotalResp	1982-2008
	Sensible Heat	Qh	1982-2008
	Latent Heat	Qle	1982-2008
	Biomass (AGB, total)	AbvGrndWood, TotLivBiom (tropics)	2000
	Biomass (AGB only)	AbvGrndWood (CONUS forest)	2001
	Biomass (AGB, total)	AbvGrndWood, TotLivBiom	2000
	Total Soil Carbon	TotSoilCarb	1950-1995 (mean only)
	Average Layer Soil Moisture (AMSR-E)	SoilMoist	2002-2009
	Total Snow Depth	SnowDepth	1998-2010
	Snow Water Equivalent	SWE	2000-2010
Site	Gross Primary Productivity	GPP	1991-2006 (varies by site)
	Net Ecosystem Exchange	NEE	1991-2006 (varies by site)
	Total Respiration	TotalResp	1991-2006 (varies by site)
	Sensible Heat	Qh	1991-2006 (varies by site)
	Latent Heat	Qle	1991-2006 (varies by site)
	Net Primary Productivity	NPP	1937-2000 (single year per site)
	Net Primary Productivity	NPP	1897-2006 (single year per site)
	Leaf Area Index	LAI	1932-2000 (single year per site)
	Soil Moisture	SoilMoist	1952-2010 (varies by site)
Composite	NH Snow Extent	Area where SnowDepth > 0	1967-2010
	Cropland NPP (CONUS)	NPP from croplands only	1990-2010
	Atmospheric CO2	NEE -> [CO2] by latitudinal zone	Varies by location

Defined Reference



Defined Reference



1. Not all benchmarks created equal
2. Scale mismatch
3. Degree of *modeled-ness* [is MODIS GPP useable?]
4. Uncertainty [points to intervals]
5. Right answer → wrong reason

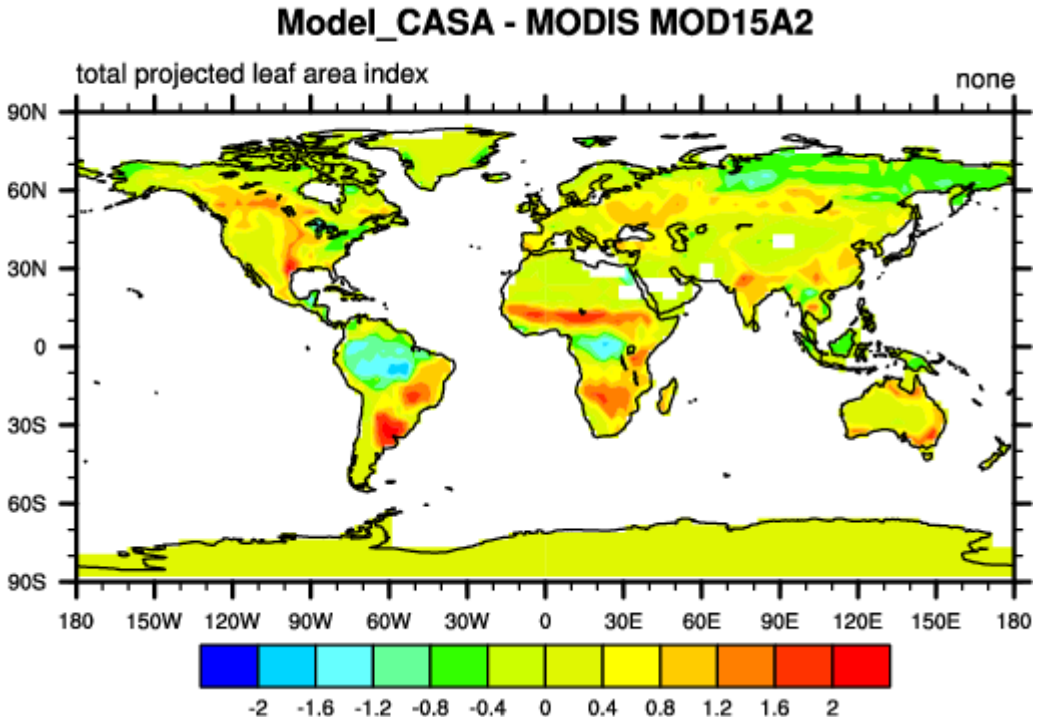
Scoring

- Across MIPs vs. within MIPs
- Weights (Cadule et al., 2009; Randerson et al., 2009; Reichler & Kim, 2008)
- Normalization
  - Gridded GPP (IAV) from upscaled FLUXNET
  - Assume 15 models using “best” simulation [BG1]
  - Step 1: z-score transform
  - Step 2: Using sigma across all 15 models calculate z-score for “perfect model”
  - Step 3: Average z-scores for bias and correlation for final score (including “perfect model”)

Scoring

- Extendable to multiple realizations/benchmarks with arbitrary weighting
- Weighting also useable for across MIP scoring: Weighted bias and correlation (C-LAMP).
- Non-parametric ranking also available
- Tabular summary
- Visualization: Taylor diagram, difference maps, 3D skill surfaces (with χ^2)

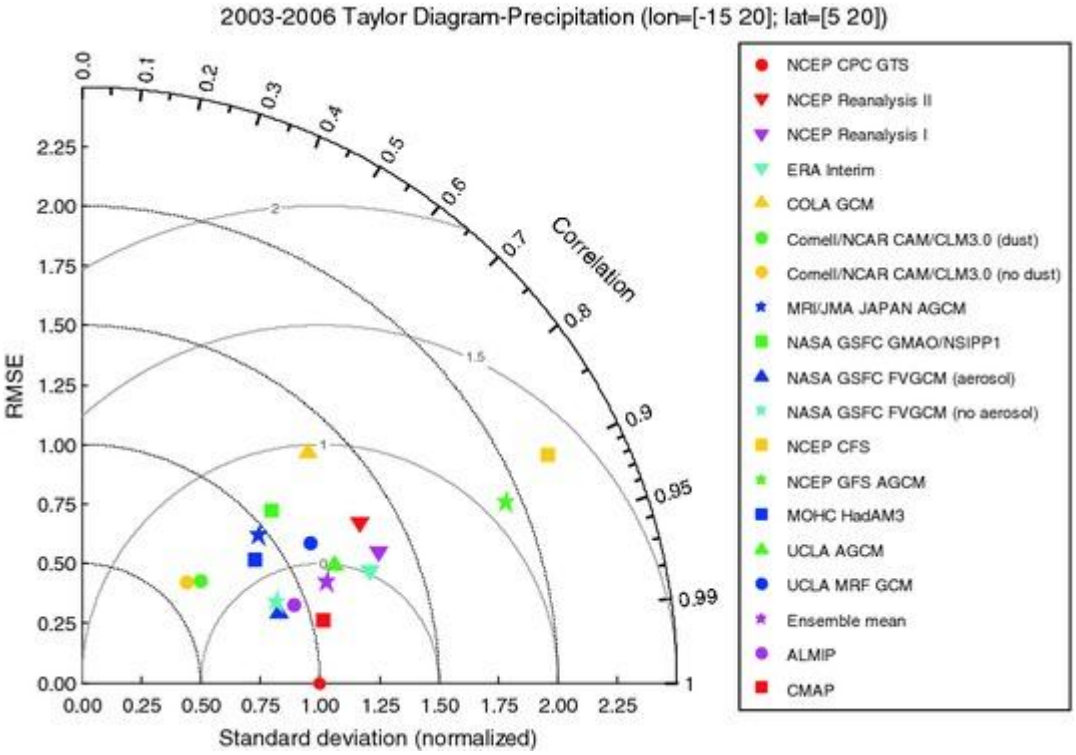
Visualizations



Difference between CASA and MODIS LAI.

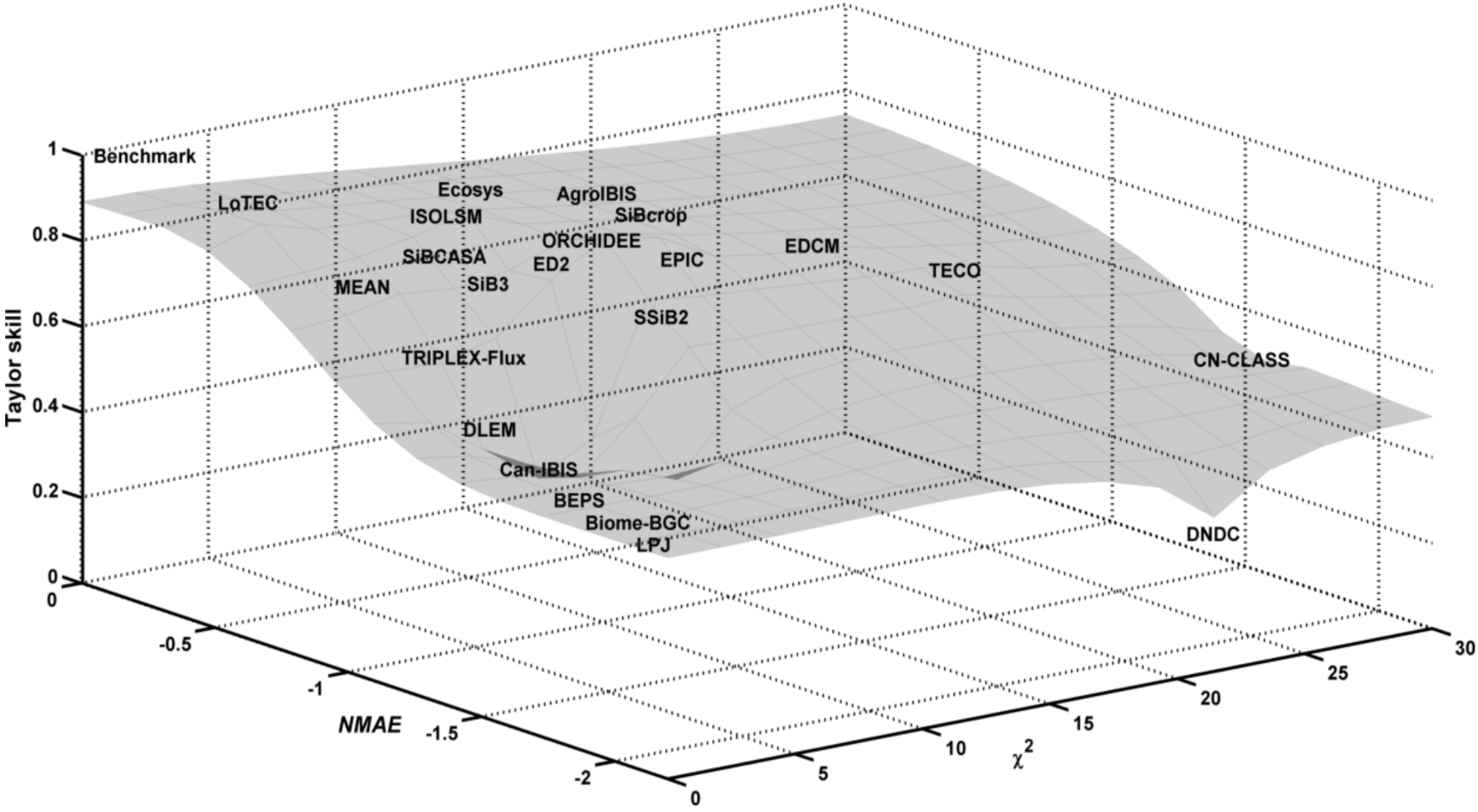
Source: http://www.climatemodeling.org/c-lamp/results/diagnostics/CASA/lai/global_Mean_model_vs_ob.png

Visualizations



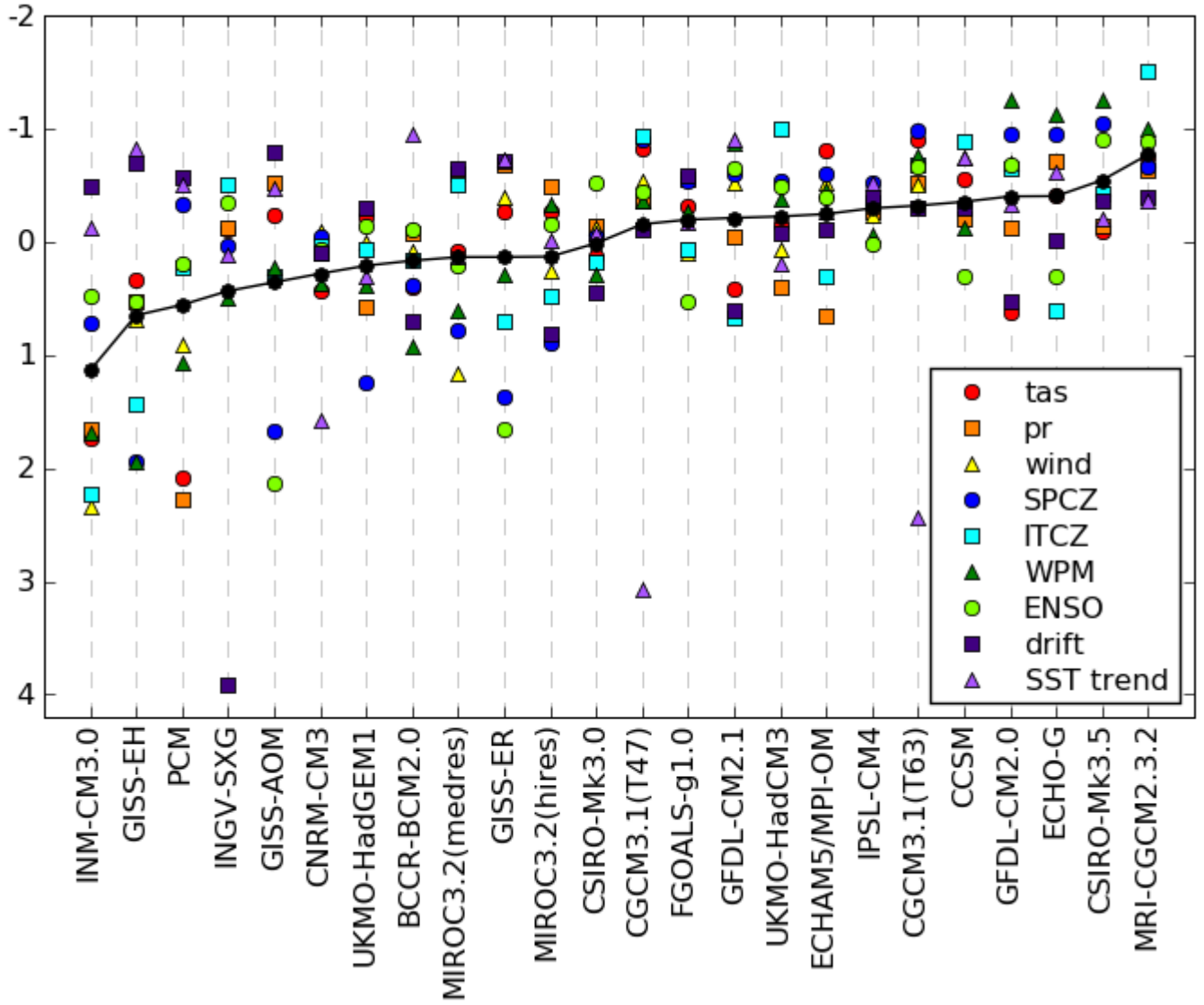
Taylor diagram displaying statistical comparisons of 12 model runs' estimates with observation of the West African mean precipitation pattern for May to October 2003–2006. Source: Xue et al. (2010) Climate Dynamics

Visualizations



3D model skill surface for 22 terrestrial biosphere models from the NACP Site Synthesis. Skill metrics are Taylor skill (S), normalized mean absolute error (NMAE), and reduced χ^2 statistic. Better model-data agreement corresponds to the upper left corner. Benchmark represents perfect model-data agreement: S = 1, NMAE = 0 and χ^2 = 1. Source: Schwalm et al. (2010) JGR

Visualizations

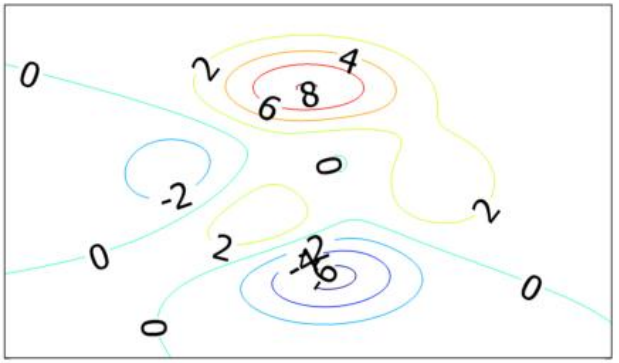


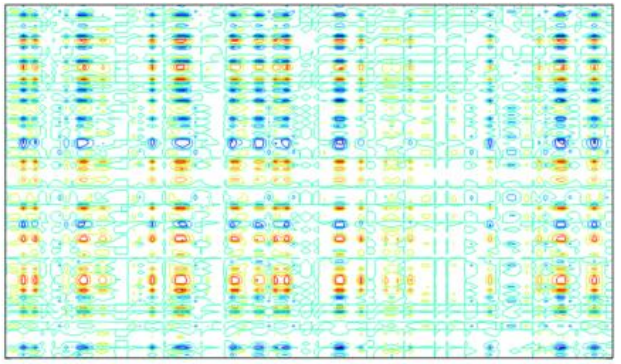
Source: Irving et al. (2010) Climate Research

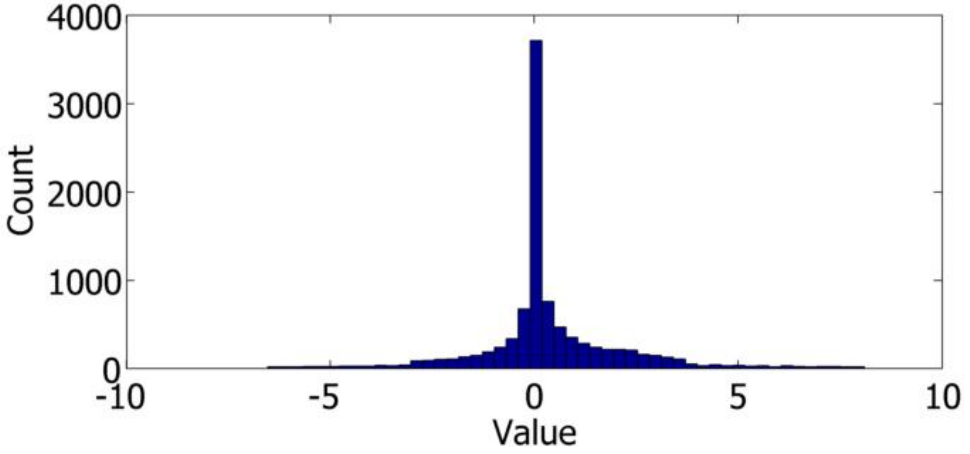
Challenges

- What datasets/metrics/outputs are needed for model development?
- "Combined" benchmarks
  - NEE normalized by soil moisture
  - beta factor for NPP [FACE]
  - Sensitivities/feedbacks [snow-albedo]
  - Teleconnective patterns [ENSO]
- Model skill as function of time
- Spatial statistics

Spatial Statistics

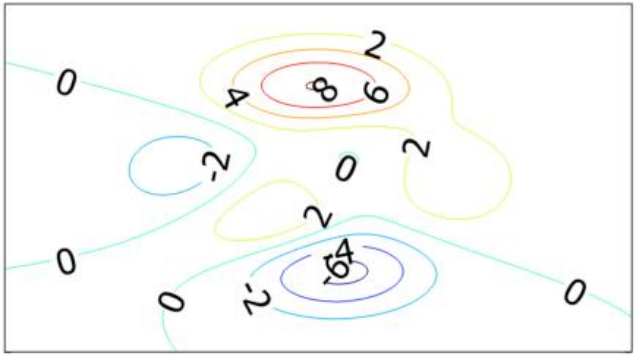


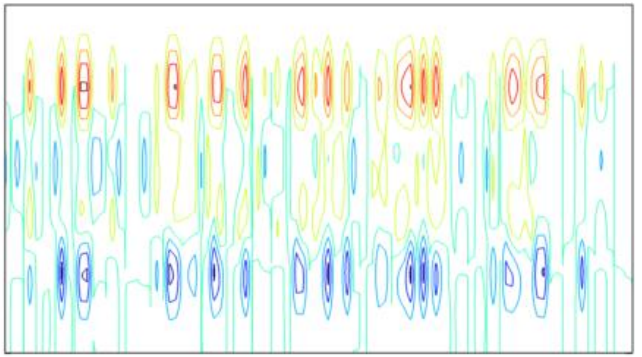


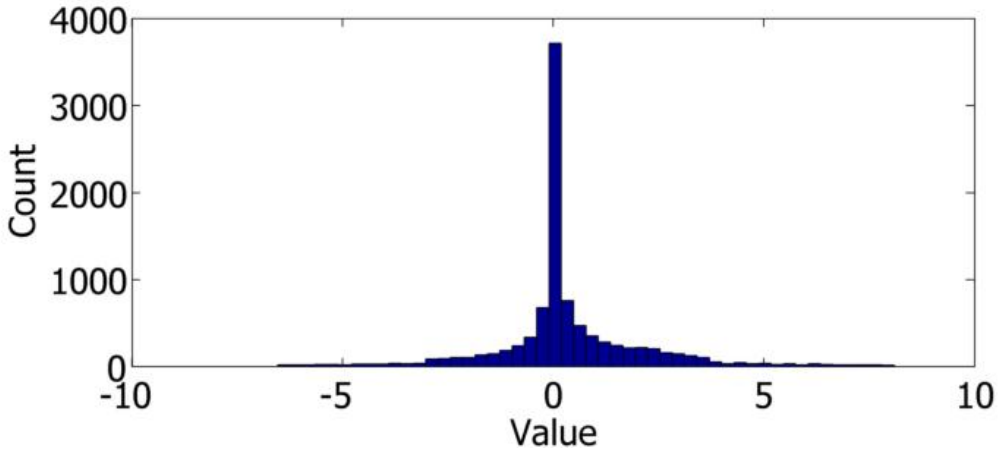


Exact same data – same histogram
Metric using *grid cells as replicates* will give same answer despite underlying difference in spatial texturing

Spatial Statistics







Exact same data – same histogram

Metric assessing *latitudinal gradient* will give same answer despite underlying difference in spatial texturing