WGCM and AIMES

October 21, 2011

- C4MIP (Pierre Friedlingstein)
- MAREMIP (Yasu Yamanaka/Scott Doney)
- iLAMB (Jim Randerson/Forest Hoffman)

International Land Model Benchmarking (ILAMB) Project

Goals:

- Develop benchmarks for land model performance, with a focus on carbon cycle, ecosystem, surface energy, and hydrological processes. The benchmarks should be designed and accepted by the community.

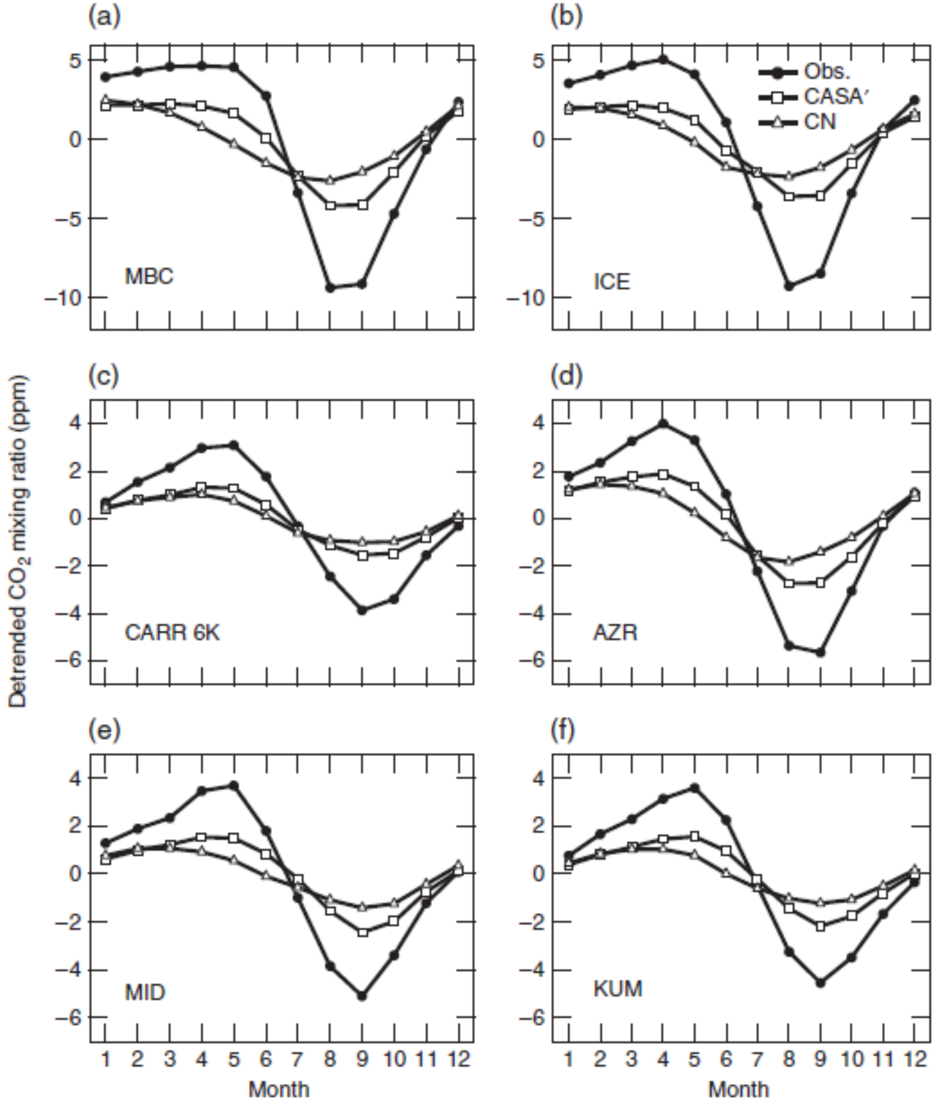
- Apply these benchmarks to global models

- Support the design and development of a new, open-source, benchmarking software system for either diagnostic or model intercomparison purposes

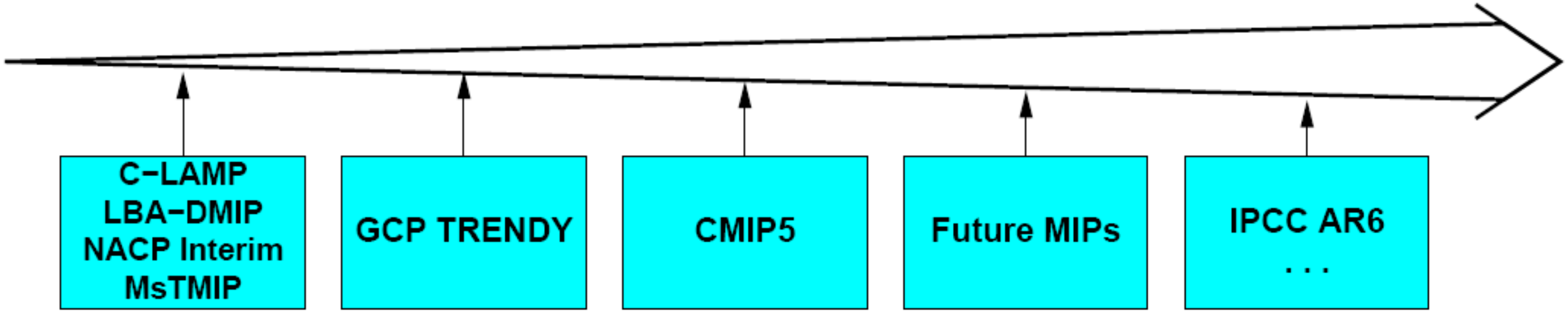
- Strengthen linkages between experimental, monitoring, remote sensing, and climate modeling communities in the design of new model tests and new measurement programs

What is a benchmark?

1. A quantitative test of model function, for which the uncertainties associated with the observations can be quantified
2. Acceptable performance on benchmarks ***is a necessary but not sufficient condition*** for a fully functioning model
3. Since all datasets have strengths and weaknesses, an effective benchmark may be one that draws upon a broad set of independent observations to evaluate model performance on multiple temporal and spatial scales



An Open Source Benchmarking Framework



- Human capital costs of making rigorous model-data comparisons is considerable and constrains the scope of individual MIPs.
- Many MIPs spend resources "reinventing the wheel" in terms of variable naming conventions, model simulation protocols, and analysis software.
- **Need for ILAMB:** Each new MIP has access to the model-data comparison modules from past MIPs through ILAMB (*e.g.*, MIPs use one common modular software system). Standardized international naming conventions also increase MIP efficiency.

- Meeting Co-organized by Forrest Hoffman (UC-Irvine and ORNL), Chris Jones (UK Met Office Hadley Centre), Pierre Friedlingstein (U. Exeter), and Jim Randerson (UC-Irvine).
- About 45 researchers participated from the United States, Canada, the United Kingdom, the Netherlands, France, Germany, Switzerland, China, Japan, and Australia.

Meeting Summary

- Five break-out groups met, one for each benchmark category, to identify cost function metrics and graphics.
- Measurement and model uncertainty must be characterized and spatial scaling mismatch considered for evaluation.

- Key objectives are to use publicly available data and freely available software.
- The R package will be used for generating statistical results and diagnostics.
- Five initial benchmarks will be implemented to evaluate existing TRENDY and CMIP5 model results.

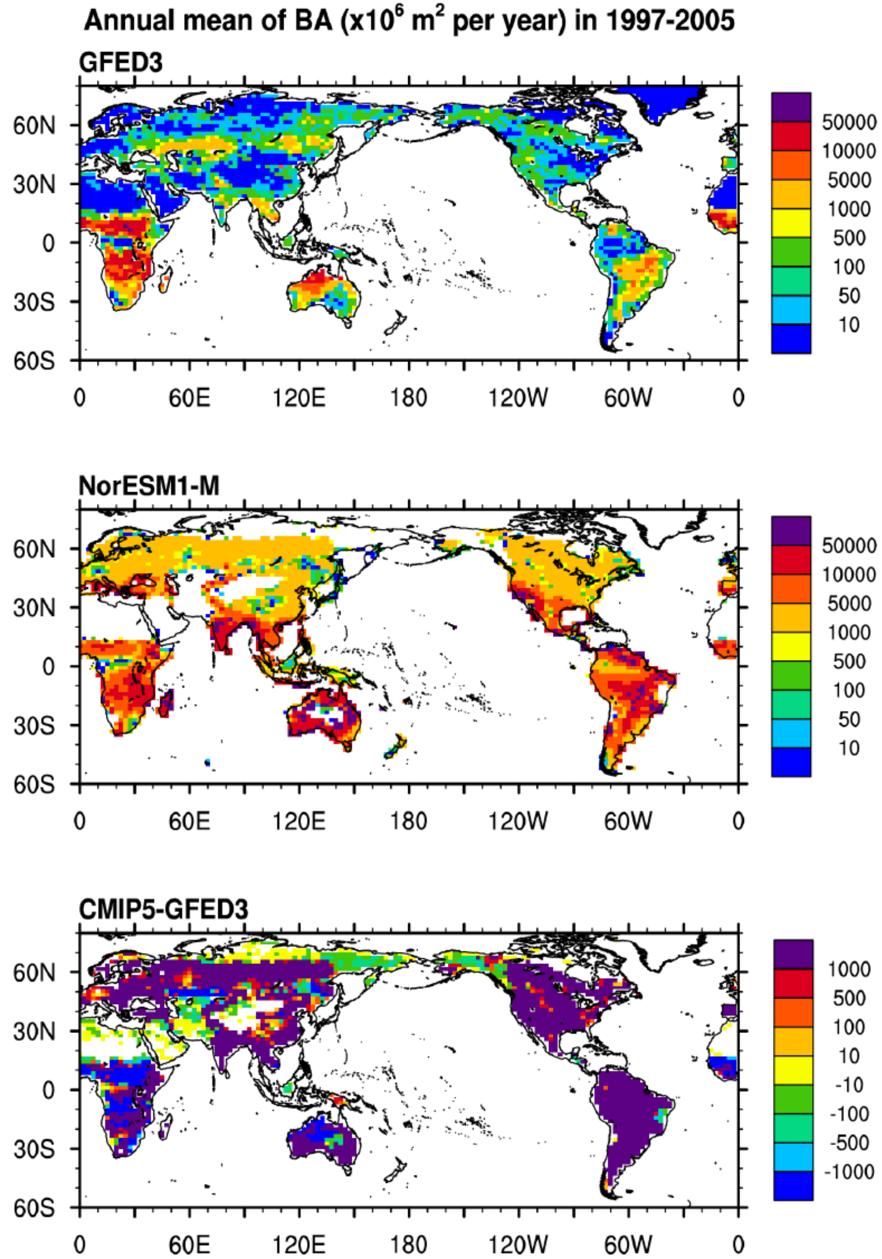
ILAMB 1.0 benchmark is now under development
Satellite and ground-based obs. will contribute to many components

	Annual Mean	Seasonal Cycle	Interannual Variability	Trend	Data Source
Atmospheric CO_2					
Flask/conc. + transport		✓	✓	✓	NOAA, SIO, CSIRO
TCCON + transport		✓	✓	✓	Caltech
Fluxnet					
GPP, NEE, TER, LE, H, RN	✓	✓	✓		Fluxnet, MAST-DC
Gridded: GPP	✓	✓	?		MPI-BGC
Hydrology/Energy					
river flow	✓		✓		GRDC, Dai, GFDL
global runoff/ocean balance	✓				Syed/Famiglietti
albedo (multi-band)		✓	✓		MODIS, CERES
soil moisture		✓	✓		de Jeur, SMAP
column water		✓	✓		GRACE
snow cover	✓	✓	✓	✓	AVHRR, GlobSnow
snow depth/SWE	✓	✓	✓	✓	CMC (N. America)
T_{air} & P	✓	✓	✓	✓	CRU, GPCP and TRMM
Gridded: LE, H	✓	✓			MPI-BGC, dedicated ET
Ecosystem Processes & State					
soil C, N	✓				HWSD, MPI-BGC
litter C, N	✓				LIDET
soil respiration	✓	✓	✓	✓	Bond-Lamberty
FAPAR	✓	✓			MODIS, SeaWIFS
biomass & change	✓			✓	Saatchi, Pan, Blackard
canopy height	✓				Lefsky, Fisher
NPP	✓				EMDI, Luyssaert
Vegetation Dynamics					
fire — burned area	✓	✓	✓		GFED3
wood harvest	✓			✓	Hurtt
land cover	✓				MODIS PFT fraction

CMIP5 for the IPCC 5th Assessment is fundamentally different from C4MIP: historical transient model simulations forced with observed trajectories of atmospheric composition and land use are now available

Country	Model	Historical	ESM Historical
China	BCC-CSM1.1	✓	✓
Canada	CanESM2	✓	✓
United States	CCSM4	✓	
France	CNRM-CM5	✓	
Australia	CSIRO-Mk3.6-0	✓	
United States	GISS-E2-R	✓	
United Kingdom	HadCM3	✓	
United Kingdom	HadGEM2-CC	✓	
United Kingdom	HadGEM2-ES	✓	
Russia	INMCM4	✓	✓
France	IPSL-CM5A-LR	✓	
Japan	MIROC4h	✓	
Japan	MIROC5	✓	
Japan	MRI-CGCM3	✓	
Norway	NorESM1-M	✓	

Example ILAMB 1.0 benchmark for fires

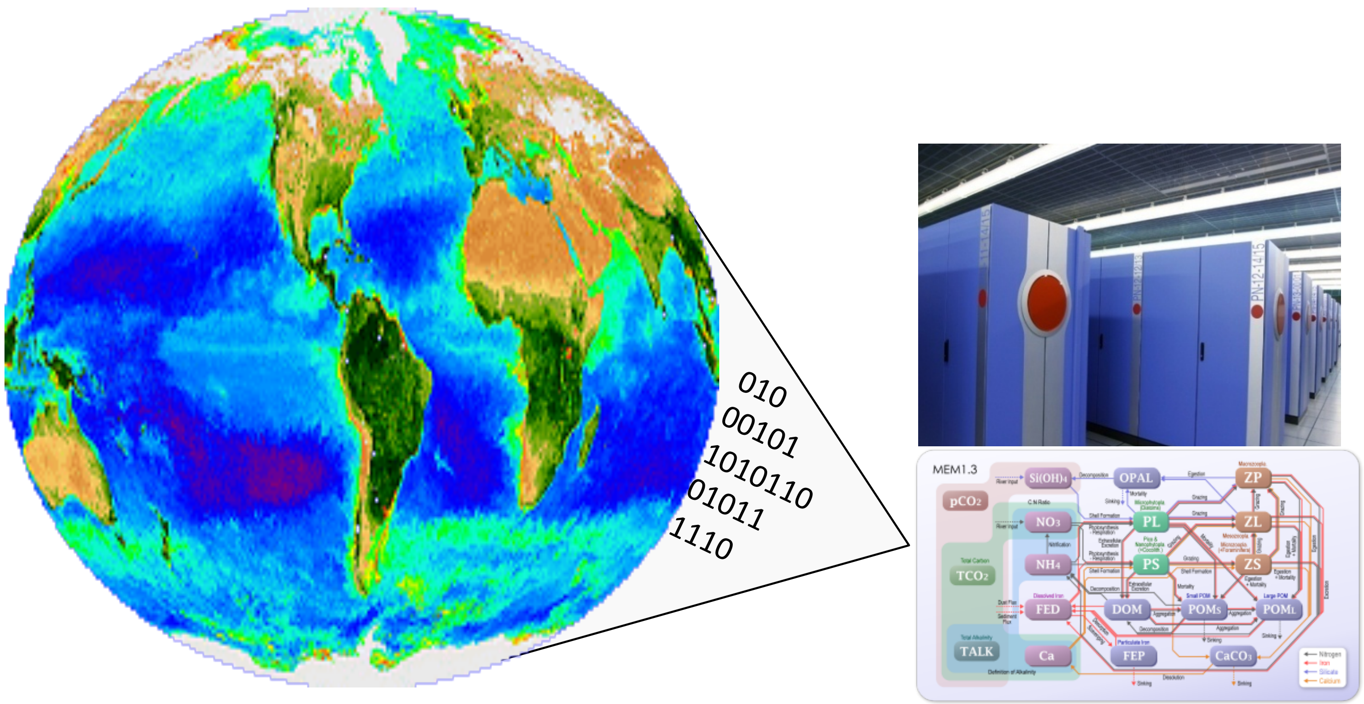


Next Steps

- A team was identified to begin implementing 5–6 benchmarks in existing model results from TRENDY and now CMIP5.
- A draft document proposing additional new netCDF Climate and Forecast (CF) conventions, beyond those created for CMIP5, is available for comment.
- Monthly conference calls started in September.
- A development Wiki is coming soon.
- ILAMB Side Meeting at AGU Fall Meeting on Monday night.
- Next ILAMB meeting in Beijing, China, in early 2012.

International Land Model Benchmarking (ILAMB) Project
http://www.ilamb.org/

Marine Ecosystem Model Inter-comparison Project (MAREMIP)



T. Hirata on behalf of Y. Yamanaka
and
Many other MAREMIP participants

20-22 Nov., Boulder, AIMES SSC meeting

Carbon stock of marine biota is small.
But carbon flux due to marine biota is large!

Different phytoplankton have different functional roles in biogeochemical cycles

Taxonomic group	*Biogeochemical Function*	*Size class*
▪Diatom	C, Si	Micro
▪Dinoflagellates	C, DMSp	Micro
▪Prymnesiophyte (Haptophytes)	C, ($CaCO_3$, DMSp)	Nano
▪Cyanobacteria	C, (N_2)	Pico

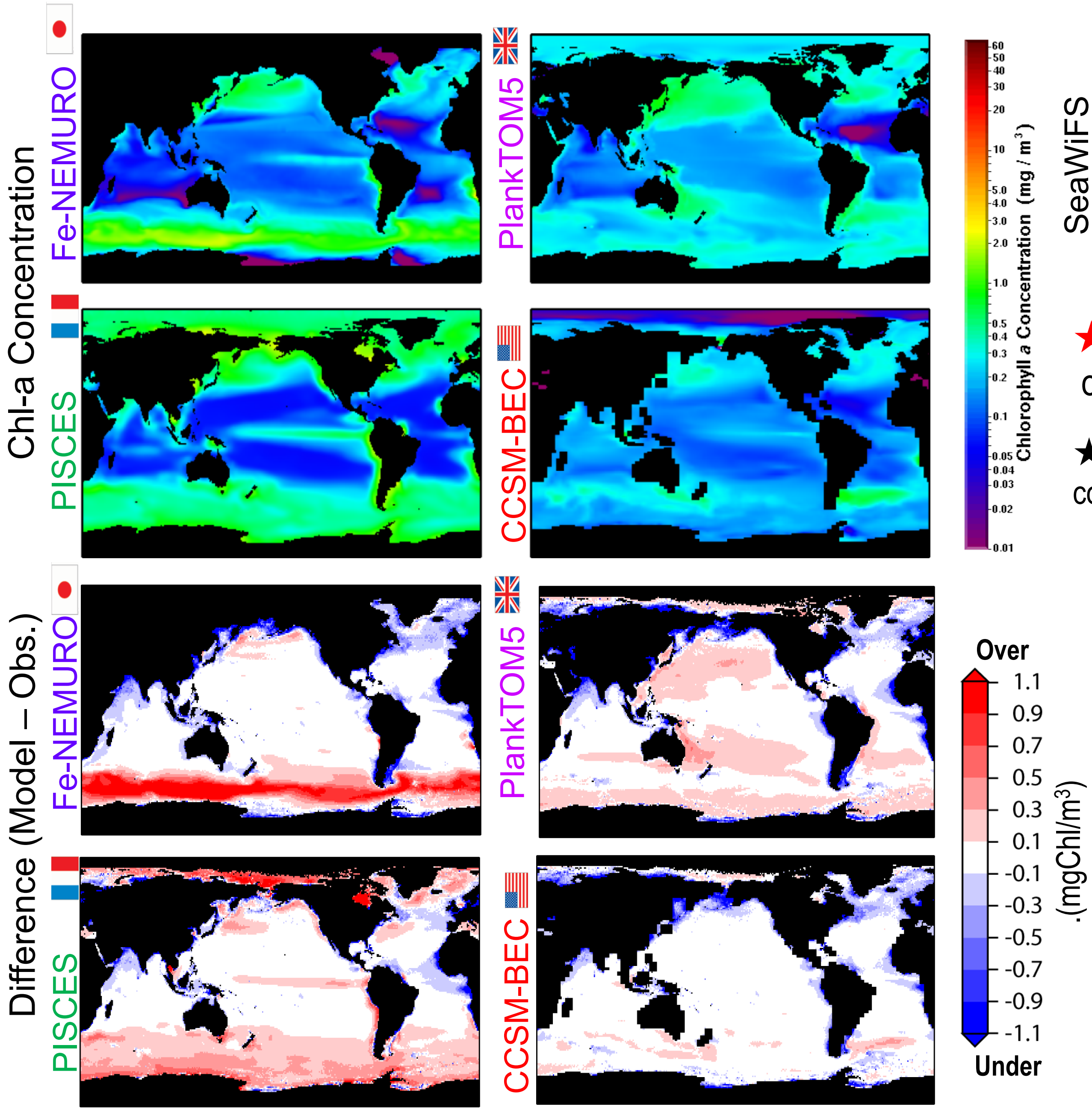
MARine Ecosystem Model Intercomparison Project (MAREMIP)

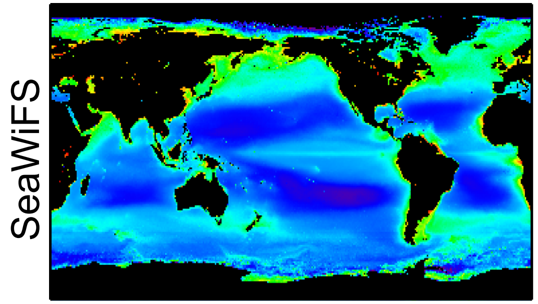
- Ocean biogeochemistry is strongly influenced by the specific activity of various types of plankton. In an effort to improve the representation of marine ecosystems, ocean biogeochemistry models have evolved to include a growing number of organisms aggregated according to their functionality into "Plankton Functional Types" (PFTs) …
- MAREMIP aims to foster the development of models based on phytoplankton functional types (PFTs) in order to progress towards the resolution of important scientific questions regarding marine biogeochemistry and climte.
- MAREMIP will also help build a community around marine ecosystem models, promote the interactions between modellers and observationalists and the development of targeted observations, and
- Synthesize the existing data for the evolution of marine ecosystem models.

Model development + science + capacity building + data synthesis

Provided by Dr. Hashioka

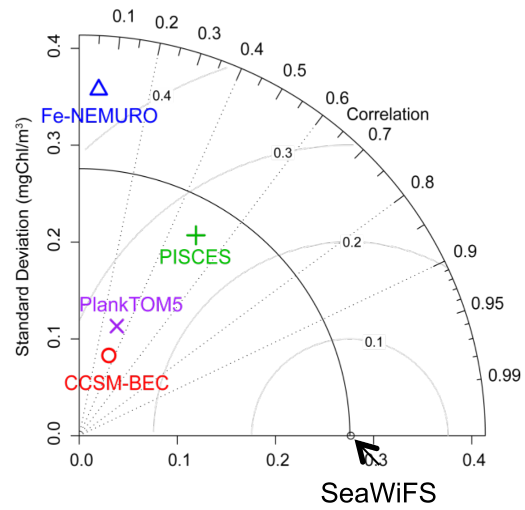
Last year… we conducted hindcast experiments to compare total phytoplankton population in the ocean… (i.e. Phase 0)





★ All models reasonably captured the observed global pattern of Chl-*a* conc.

★Fe-NEMURO overestimated the Chl-a conc. in the Southern Ocean.



Taylor Diagram
Horizontal Distribution of Chl-a
(n<360x180)

(2) Membership changed - others

Expanded community members since Phase 0

Phase 0 members
- GBR：UEA
- JPN： JAMSTEC/Hokkaido Univ
- USA：WHOI
- FRA：IPSL

\+

New members
- USA：GFDL
- USA：NASA
- JPN： MRI
- GBR：PML
- GBR: SOC
- ITA：INGV

\+

Observationists
Satellite
In situ

\+

In situ data management

Project Meeting in Plymouth, UK
Date: Sun 26 June 2011

✓International Workshop (21)
(Europe, Japan, USA, Australia)

✓ Modellers (old & new), satellite algorithm developers, in situ observationalists, and data manager…

a.m. Presentations of Research
Scientific achievements from modellers

p.m. Discussions on MAREMIP activities
- ✓ Data synthesis
- ✓ Towards AR5 contribution (e.g. time schedule of future projection runs, community paper etc)
- ✓ Setting a protocol of model intercomparison for **future projection** experiments (under RCP8.5)

C. Quéré (TCCCR, UK), Y. Yamanaka(HU, JPN), S. Doney (WHOI, USA), L. Bopp (CNRS, FRA), J. Dunne (NOAA, USA)
M. Vogt (ETH, SUI), T. Hashioka (HU, JPN), S. Sailly (WHOI, USA), I. Allen (PML, UK), T. Hirata (HU, JPN), W. Gregg (NASA, USA), M. Kidston (ISPL, FRA), H. Nakano (MRI, JPN), A. Yool (SOC, UK), S. Itoh (ORI, JPN)
S. Alvain (UL, France), B. Brewin (PML, UK), E. Buitenhuis (UEA, UK), M. Friedrichs (VIMS, USA), N. Hardman-Mountford (PML, UK), R. Matear (CSIRO, AUS), J. Peloquin (ETH, SUI), M. Pahlow (IH, GER), S. Pesant (MARUM, GER)
R. Rivkin (MU, CAN), N. Stephens (PML, UK), M. Vichi (INGV, ITA), Y. W. Luo (WHOI, USA)

Protocol for future projection runs (with RCP 8.5):

Output variables

Variable/long name	Short name	
3-D Physical		
P1 - Depth resolved PAR/light intensity	ETOT	
2-D Physical		
P1 - Surface radiation (PAR)	SPAR	
3-D Biogeochemical – from OCMIP		
P1 - Phytoplankton PFT biomass	PHY1 - PHYn	
P1 - Zooplankton PFT biomass	ZOO1 – ZOOn	
P1 - Bacterial PFT biomass	BAC	
P1 - Phytoplankton PFT chlorophyll	PHYCHL1 - PHYCHLn	
P1 - Total Chlorophyll	TOTCHL	
P1 – Total phytoplankton biomass	TOTPHY	
P1 – Total zooplankton biomass	TOTZOO	
P1 – Net primary production for each PFT	NPPPHY1 – NPPPHYn	
P1 – Sinking flux of POC/Export	EPOC	
P1 – Remineralisation rate	REMIN	
3-D Biogeochemical – MAREMIP addition		
P2 – pH	pH	-
P2 - Omega aragonite	OMEGAA	-
P2 - Omega calcite	OMEGAC	
P2 - Calcite concentration		
P2 - Aragonite concentration		
? P2 - Carbon fluxes between PFTi → PFTj		
? P2 - Carbon fluxes between PFTi → POC/DOC		
? P2 - Grazing of zPFTi on pPFTj		
? P2 - Fraction of POC from PFTi		
? P2 - Fraction of DOC from PFTj		
? P2 - Nutrient limitation terms/PFT growth rates		

3-D Biogeochemical from PFT definition

1. N fixers

P2 - Nitrogen fixation rate
P2 - Denitrification rate
P2 - Anammox rate

2. DMS producers

P2 - DMS production rate
P2 - DMS concentration
P2 - DMSP concentration

3. Silicifiers

P2 - Silicate production
(P2) - Export of Silicate etc. (included above)

4. Calcifiers

P2 - Carbonate production
P2 - Calcification rate
(P1) - Export of calcite (included above)
P2 - Carbonate dissolution

5. Picophytoplankton

\-

6. Bacteria

(P1) - Remineralisation (included above)

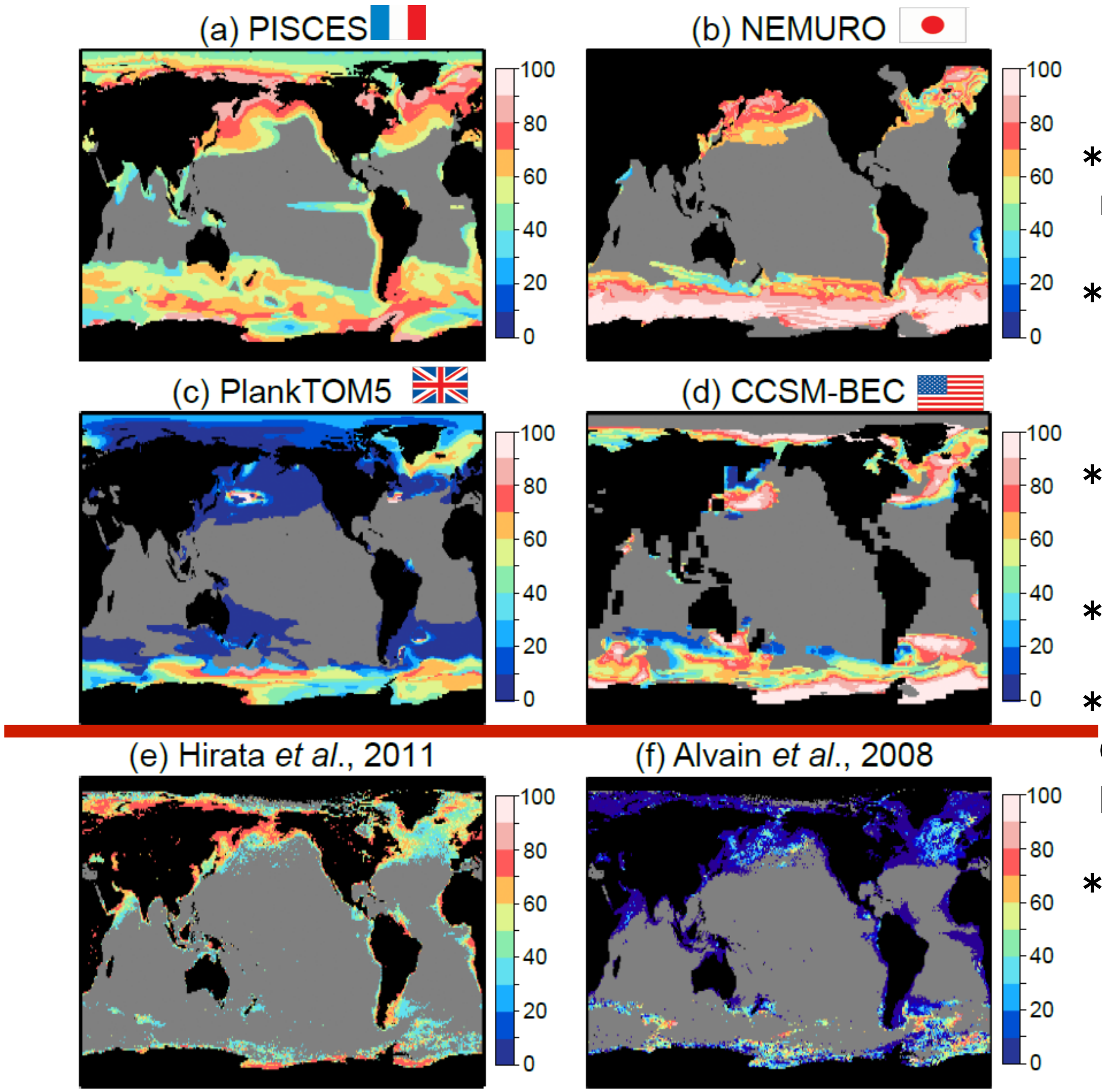
7-9. Micro-/Meso-/Macrozooplankton

P2 - Grazing rates
P2 - DOC/POC production rates

2-D Biogeochemical – MAREMIP addition

Variable/long name	Short name
P2 - Surface pH	SpH
P2 - Surface Omega aragonite	SOMEGAA
P2 - Surface Omega calcite	SOMEGAC
P2 - Depth of aragonite saturation horizon	DOMEGASAT
P2 - Depth of calcite saturation horizon	DOMEGASAT

Simulate the present day to 2100 (monthly, but also daily for the first and last 10 years)



* Comparison of percentage of diatom at the maximum chl-a conc. during spring bloom.

* In PISCES, NEMURO and CCSM-BEC, percentage of diatoms high (60 to 80%). These results are close to Hirata et al. (2011).

* In PlankTOM5, the percentage is high in SO and NA. These results are close to Alvain et al. (2008)

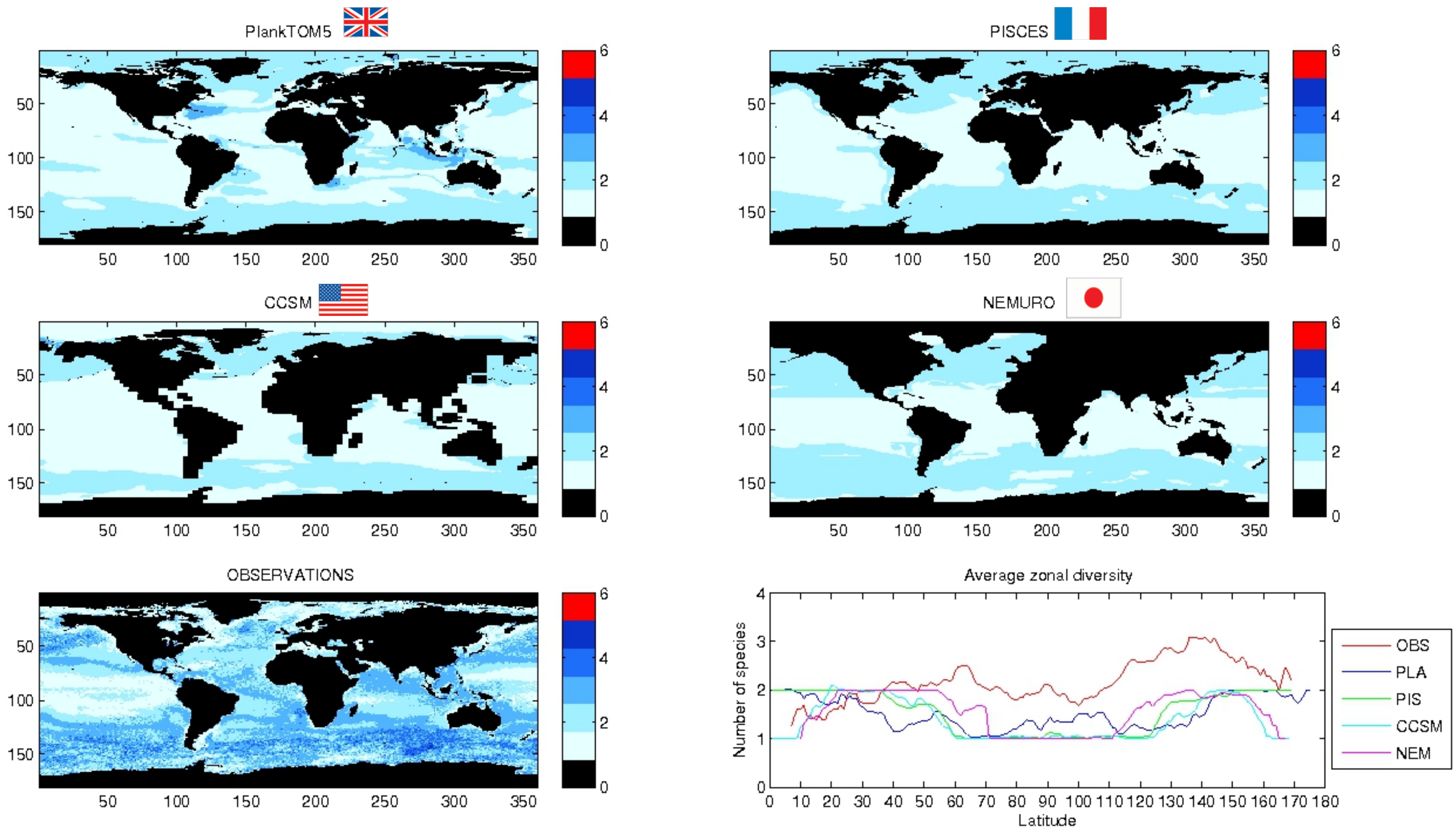
* Even in observations there are large differences

* Obs. : Hirata et al. (2011) shows dominance of diatom in many regions. In particular, the percentage is high (around 80%) in NP and SO

* Obs. : Alvain et al. (2008) shows dominance of diatoms in SO. In the northern hemisphere, there are not many regions of diatom dominance.

Number of PFTs appeared

(analogue of "diversity" concept, but not all groups in real oceans were considered in the models)



Lower at lower latitudes, higher at higher latitude

Provided by Dr. Vogt

2011 MAREMIP outputs as a community

1. Hashioka et al., submitted

2. Vogt et al., submitted

3. Synthesis data to be published in “Earth System Science Data (ESSD)”

(in situ data for validation of models, not model outputs)

Backup Slides

Why Benchmark?
Ordering from least to most controversial!

1. Show the broader science community and the public that the representation of ecosystems and the carbon cycle in climate models is improving
2. In Earth System models, provide a means to quantitatively diagnose impacts of model development in related fields on carbon cycle and land surface processes
3. Guide synthesis efforts (such as the IPCC) towards the review of mechanisms of global change in models that are broadly consistent with available contemporary observations
4. Increase scrutiny of key datasets used for model evaluation
5. Identify gaps in existing observations needed for model validation
6. Provide a quantitative, application-specific set of minimum criteria for participation in model intercomparison projects (MIPs)
7. Provide an optional weighting system for multi-model mean estimates of future changes in the carbon cycle

Example of an early carbon benchmarking system: C-LAMP

Models →

BGC & ecosystem datasets ↓

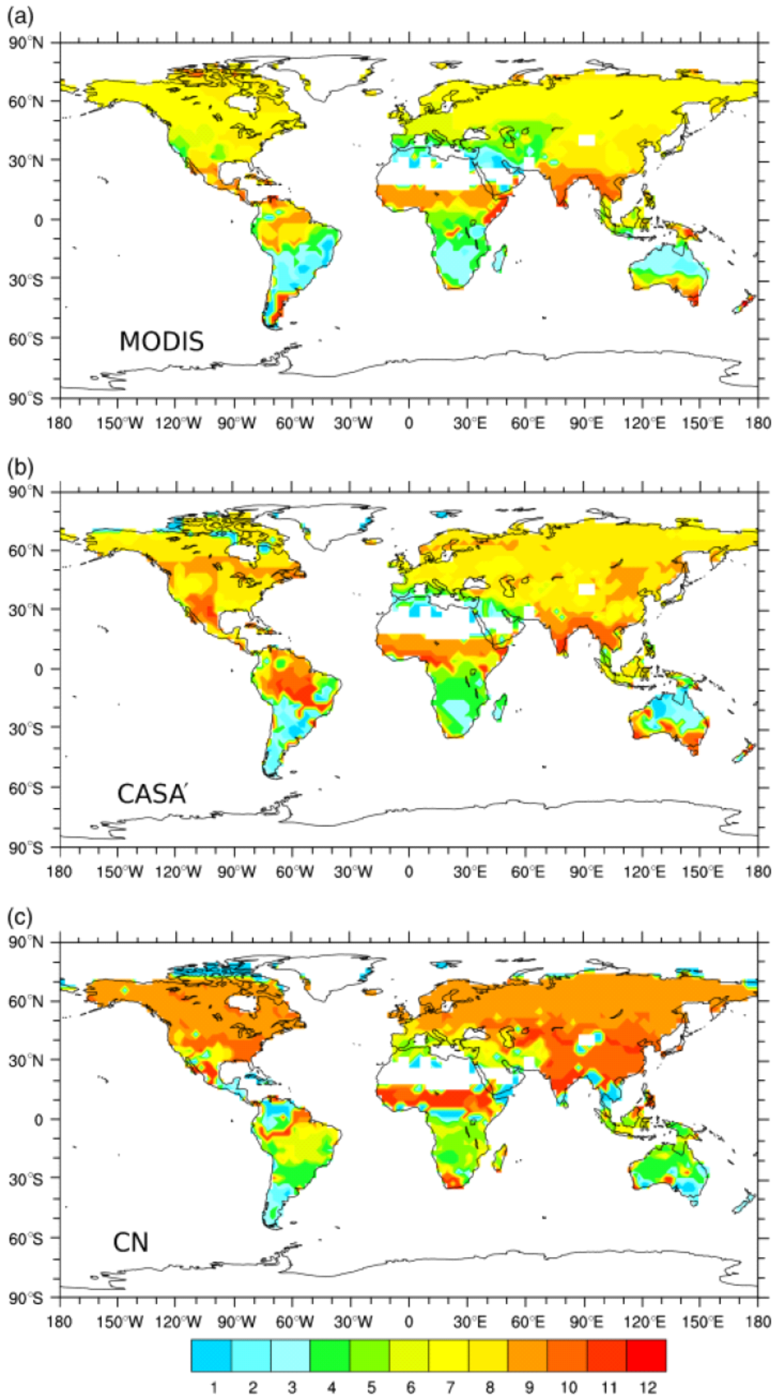
Table 3 Summary of model evaluation datasets and analysis

Metric	Metric components	Uncertainty level of obs.	Scaling mismatch	Total score	Sub-score	CASA′	CN
LAI	Matching MODIS observations			15		13.5	12.0
	– Phase (assessed using the month of maximum LAI)	Low	Low		6	5.1	4.2
	– Maximum (derived separately for major biome classes)	Moderate	Low		5	4.6	4.3
	– Mean (derived separately for major biome classes)	Moderate	Low		4	3.8	3.5
NPP	Comparisons with field observations and satellite products			10		8.0	8.2
	- Matching EMDI Net Primary Production observations	High	High		2	1.5	1.6
	- EMDI comparison, normalized by precipitation	Moderate	Moderate		4	3.0	3.4
	- Correlation with MODIS	High	Low		2	1.6	1.4
	- Latitudinal profile comparison with MODIS	High	Low		2	1.9	1.8
CO_2 annual cycle	Matching the phase and amplitude at Globalview flask sites			15		10.4	7.7
	60–90°N	Low	Low		6	4.1	2.8
	30–60°N	Low	Low		6	4.2	3.2
	0–30°N	Moderate	Low		3	2.1	1.7
Energy & CO_2 fluxes	Matching eddy covariance monthly mean observations			30		17.2	16.6
	- Net ecosystem exchange	Low	High		6	2.5	2.1
	- Gross primary production	Moderate	Moderate		6	3.4	3.5
	- Latent heat	Low	Moderate		9	6.4	6.4
	- Sensible heat	Low	Moderate		9	4.9	4.7
Transient dynamics	Evaluating model processes that regulate carbon exchange on decadal to centennial timescales			30		16.7	13.8
	- Aboveground live biomass within the Amazon Basin	Moderate	Moderate		10	5.3	5.0
	- Sensitivity of NPP to elevated levels of CO_2: comparison to temperate forest FACE sites	Low	Moderate		10	7.9	4.1
	- Interannual variability of global carbon fluxes: comparison with TRANSCOM	High	Low		5	3.6	3.0
	- Regional and global fire emissions: comparison to GFEDv2	High	Low		5	0.0	1.7
	Total:				100	65.7	58.4

CASA, Carnegie-Ames-Stanford Approach; CN, carbon–nitrogen; NPP, net primary production; FACE, free air carbon dioxide enrichment; MODIS, MODerate Resolution Imaging Spectroradiometer; EMDI, Ecosystem Model Data Intercomparison; LAI, leaf area index; GFEDv2, Global Fire Emissions Database Version 2.

Randerson et al. (2009) GCB

Lessons from C-LAMP



Month of peak leaf area

Use of the most robust aspects of observations important for reducing model uncertainties

Randerson et al. (2009) GCB