A data assimilation system for Land Surface Models – information from fluxes, phenology and biomass

David Moore, Francesc Montané, Ross Alexander, Valerie Trouet, Flurin Babst, Ave Arellano, Natasha MacBean, Amy Hudson *University of Arizona*

Tim Hoar, NCAR, Andrew Fox *University of Arizona/NCAR,*

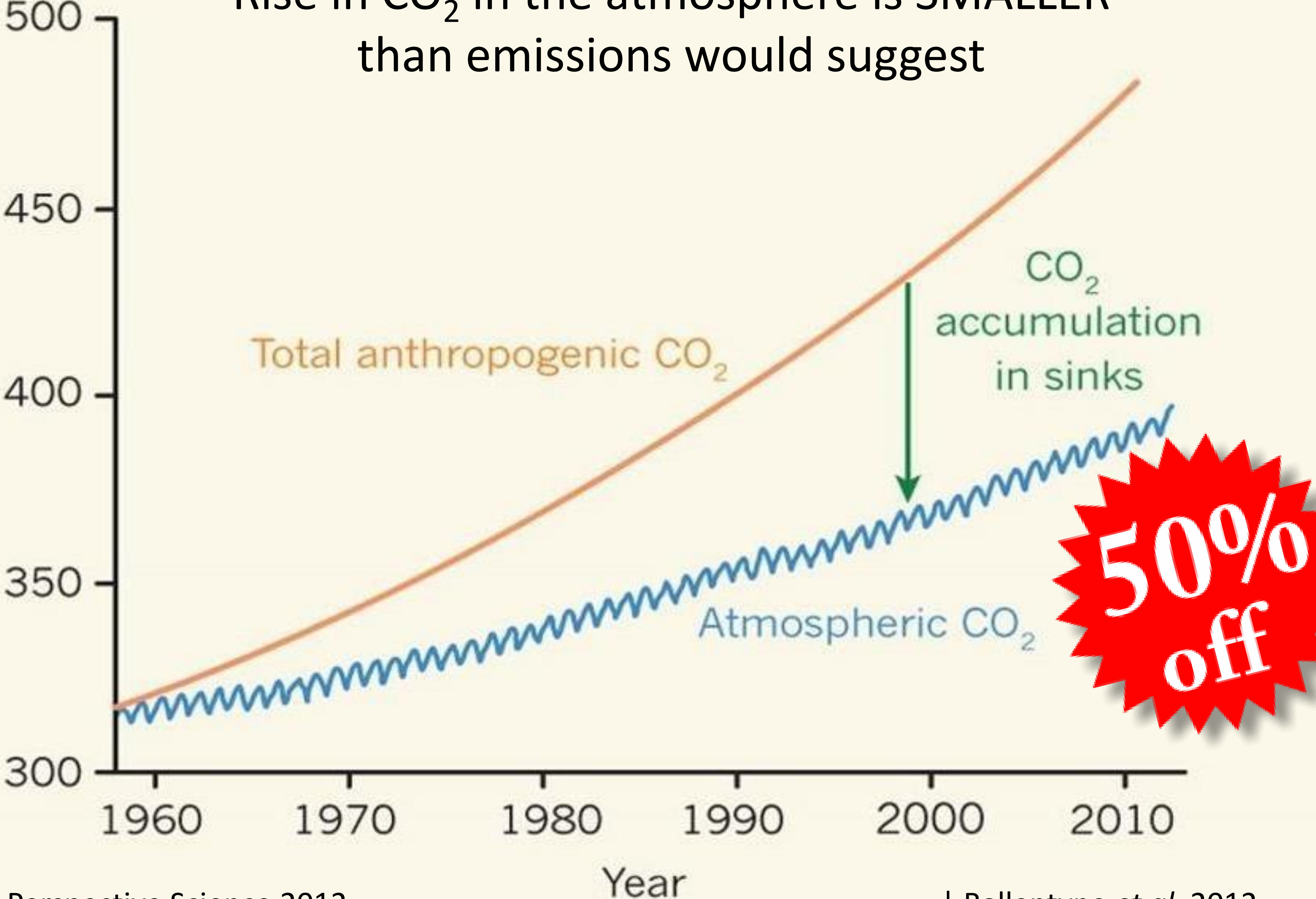
Andrew Richardson, Harvard University Min Chen, Carnegie Institute

U.S. DEPARTMENT OF ENERGY

Regional Climate Dynamics

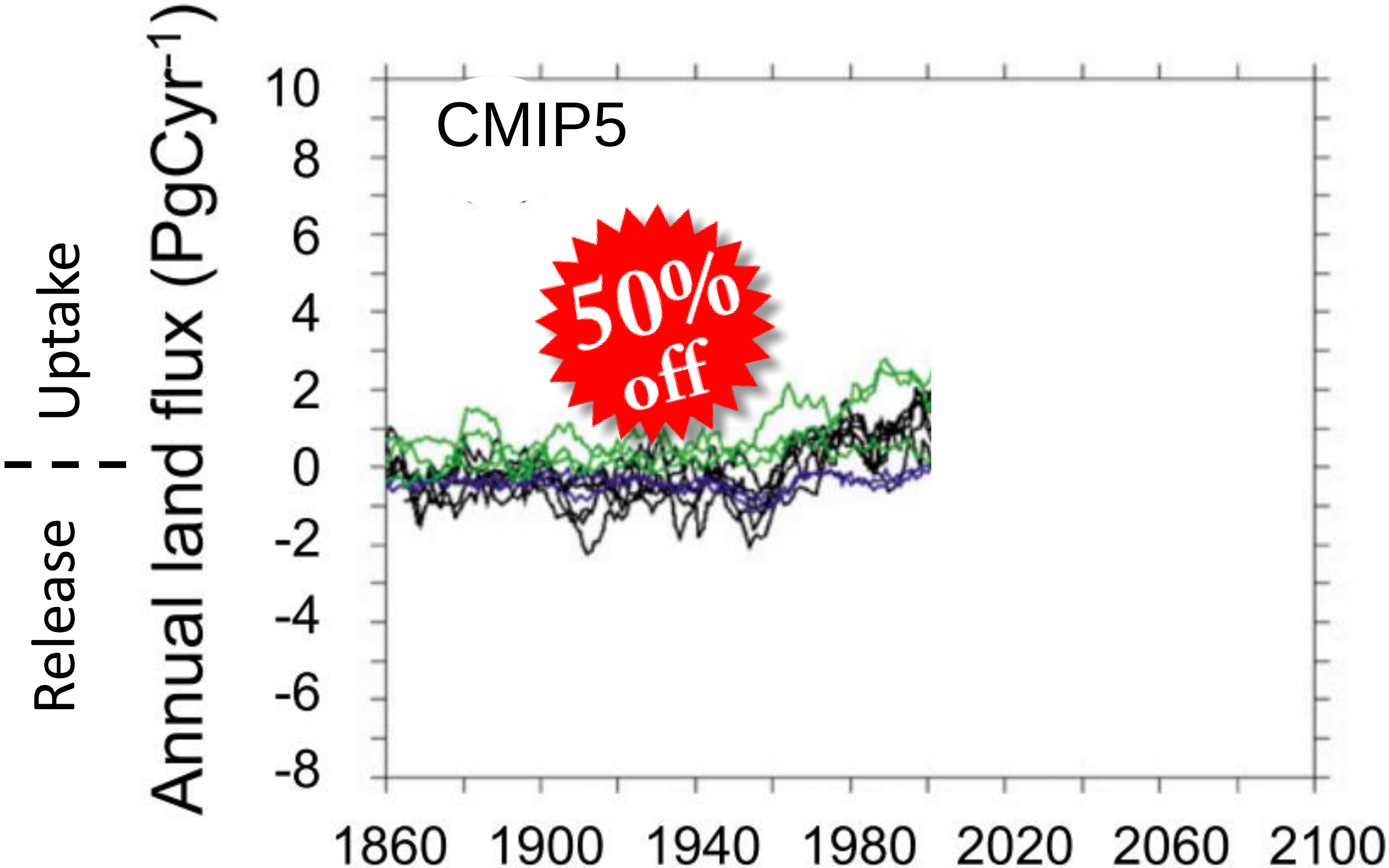
@djpmoore

Rise in CO_2 in the atmosphere is SMALLER than emissions would suggest



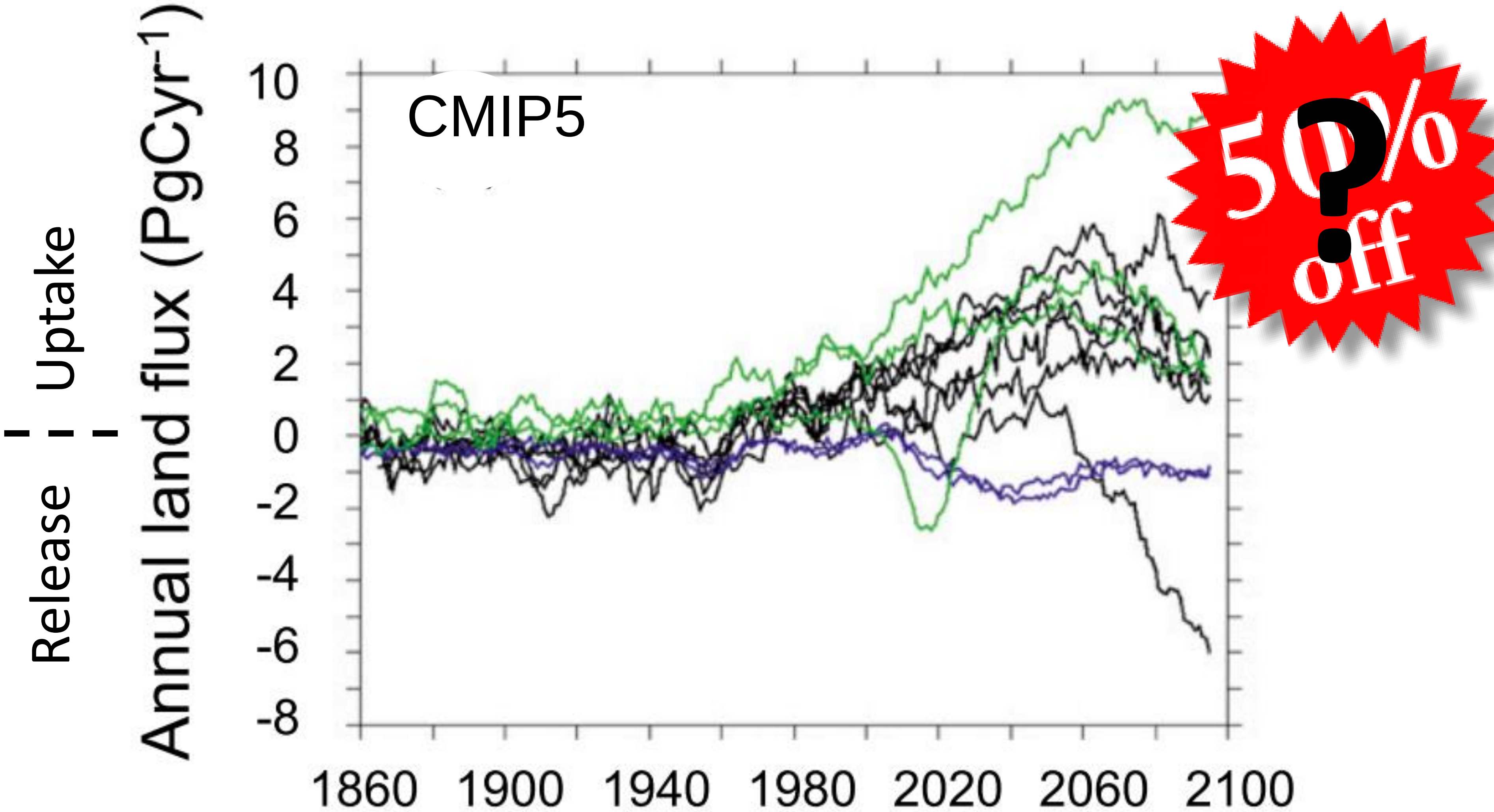
Perspective Science 2012 | Ballentyne *et al*. 2012

Models all generally agree



11 Land Surface models from different Earth System Models

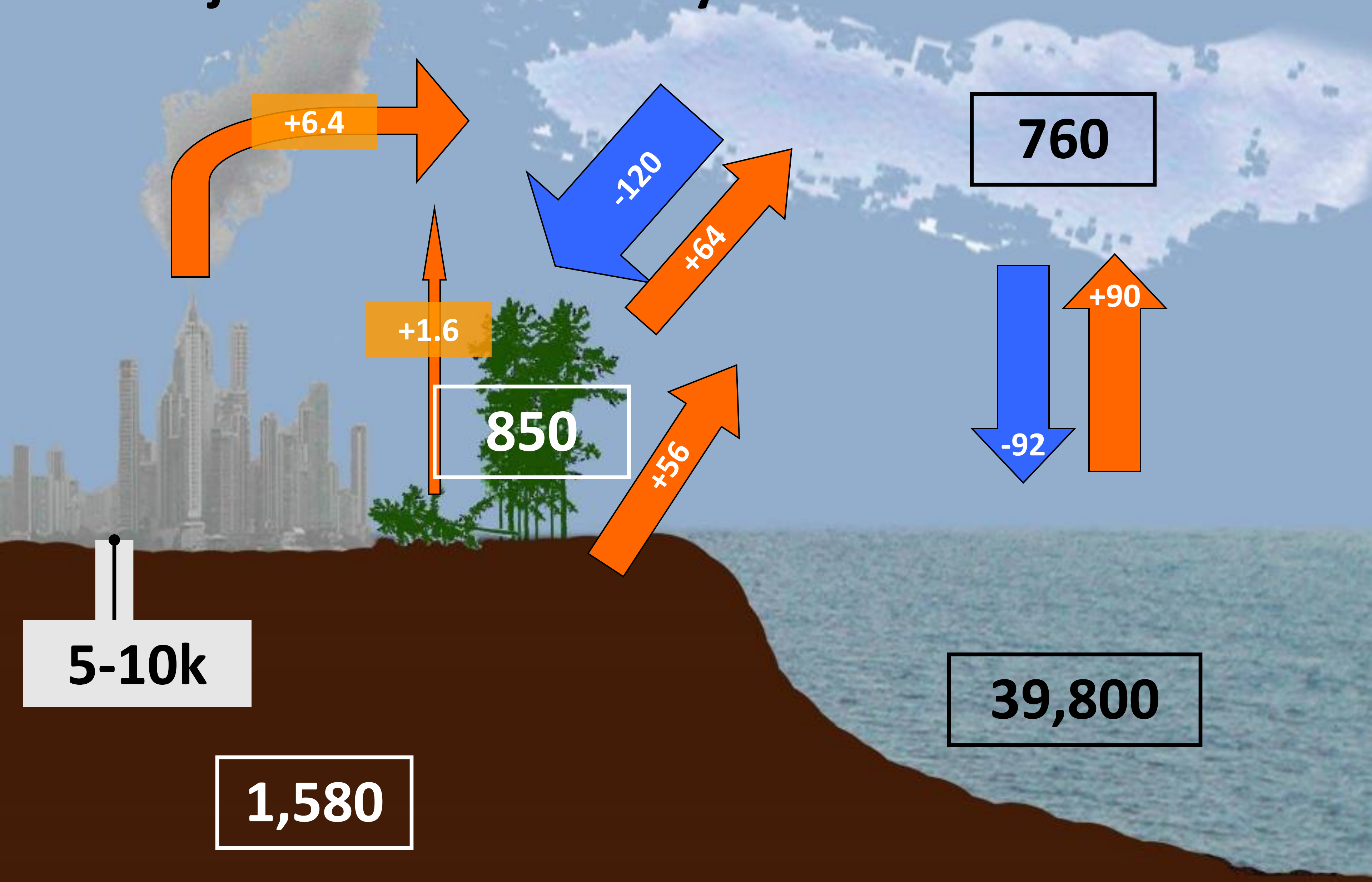
Spaghetti Carbon-Era*



11 Land Surface models from different Earth System Models

**Pun courtesy of Dr Sarah J Ivory … pers comm*

Major fluxes of the C cycle



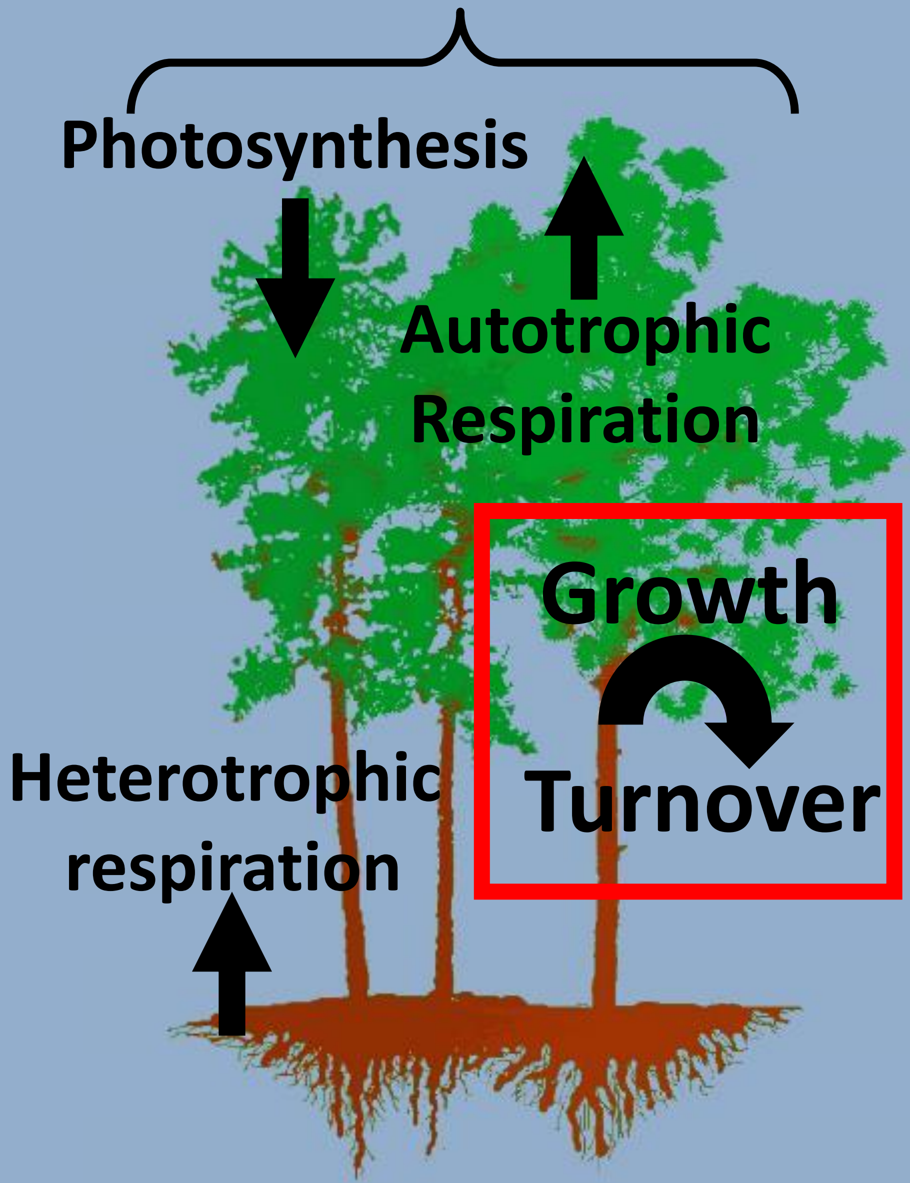
Values: Schlesinger 1997

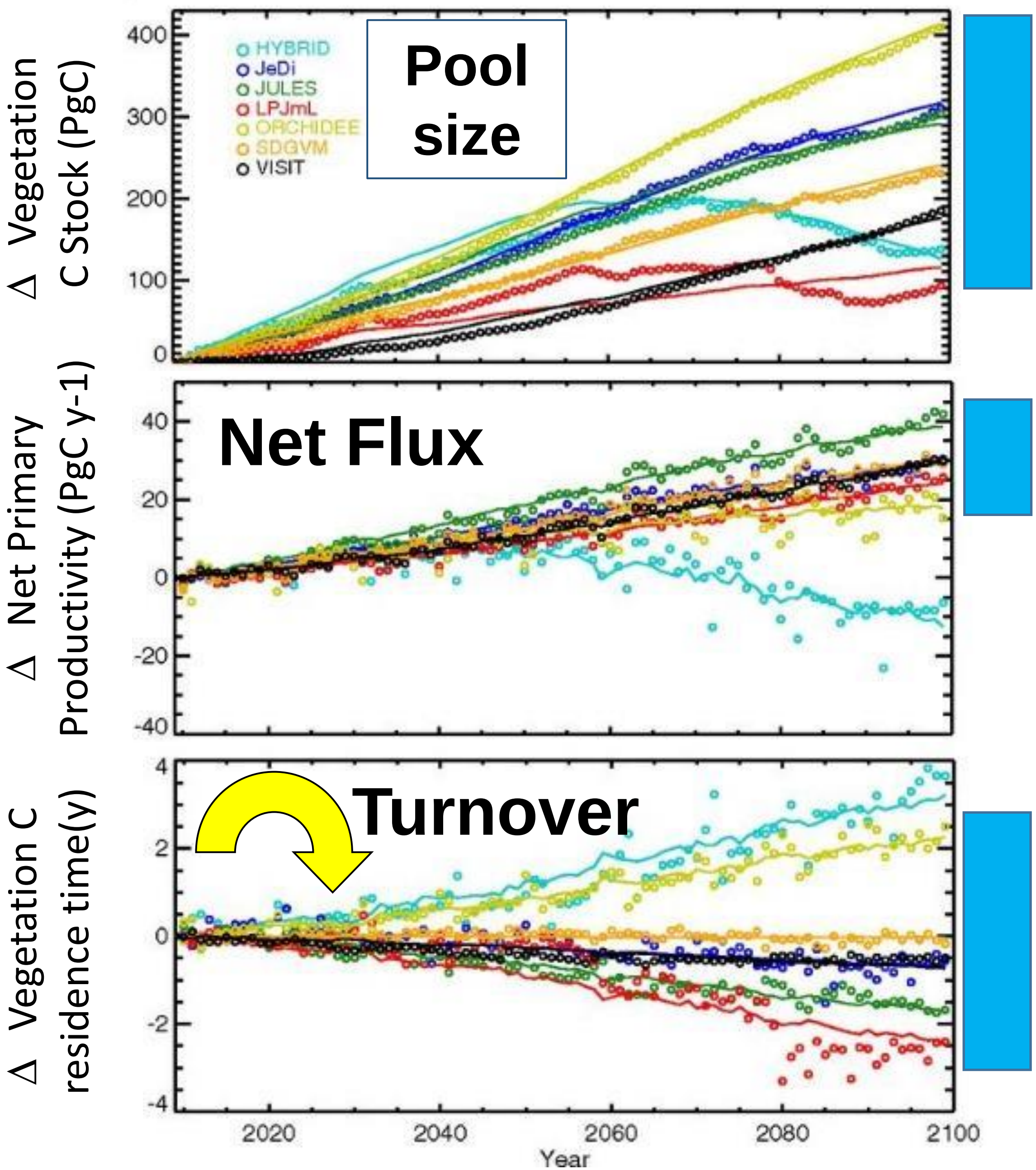
Units are in Gt of Carbon

Ecosystem C Balance

Measurements

1. **Direct flux measurement**
2. **Growth and Turnover**



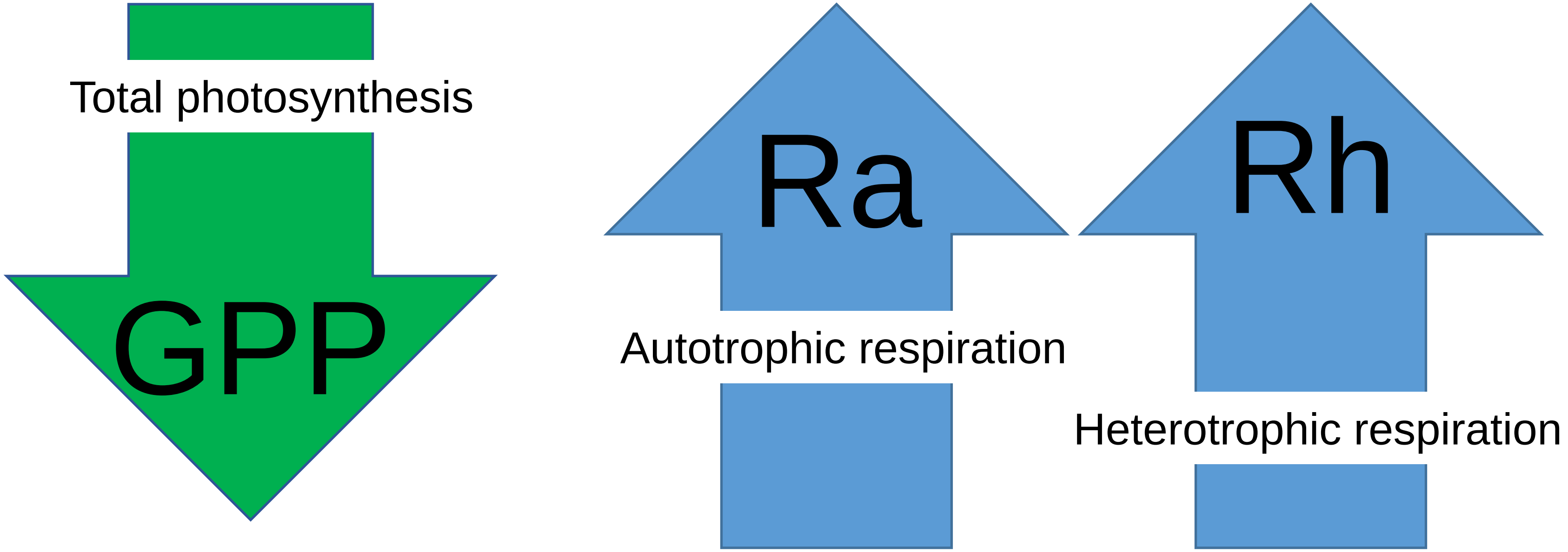


Carbon residence time controls projections of future carbon stored in vegetation

Friend et al. PNAS 2014;111:3280-3285

Net carbon balance is a SMALL difference between LARGE fluxes

Net Ecosystem Exchange
(Net Accumulation in the atmosphere)



Net Ecosystem Productivity NEP
(Net Accumulation in the Ecosystem)

Measuring fluxes

Photos: Ray Leuning

Global Distribution of Eddy Flux Towers - FLUXNET

Global Distribution of Eddy Flux Towers - FLUXNET

How do we extract knowledge from all these sites to give us information about how carbon cycling respond to climate ?

What is data assimilation?

- Systematic combination of data and models
- Taking into account the uncertainties in both
- Process model provides an analytical framework
- If done well:
  - Modeled state becomes more consistent with observations (and hopefully with the truth!)
  - Makes forecasts more accurate (as initial conditions are improved)

Learning from flux data at ecosystem scales

Separation of NEE into GPP and Re → **Contrast between Day (psn) and Night (no psn) allows separation**

(Sacks et al 2006, 2007)

Responses of NEE to precipitation change → **Seasonal co-ordination of GPP and ET (obs) allows a reasonable response to be extracted**

(Moore et al 2008)

Flux data alone does not constrain below ground processes well → **Above and below ground processes confounded in tower based measurements**

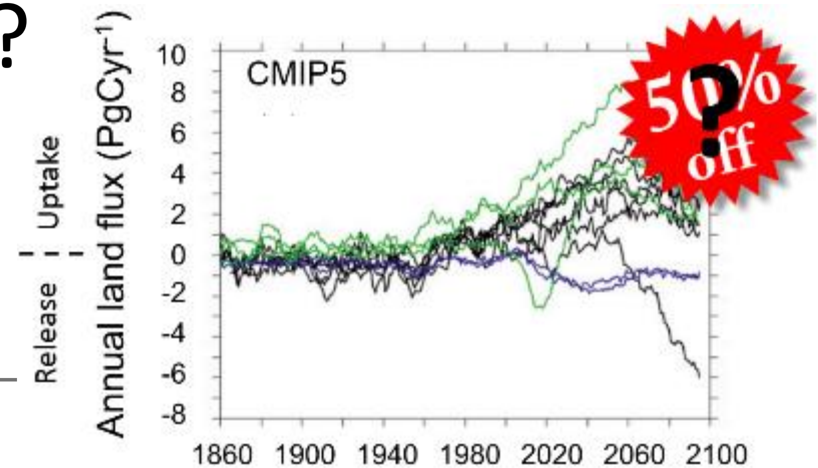
(Zobitz et al 2008)

NEE does not constrain long term processes → **Biomass, Soil Resp, LAI, litterfall can be used to constrain different parameters in the model**

Richardson *et al*. 2010

Making Better Spaghetti?

Fluxes don't help much with longer time scales & model Structure problems



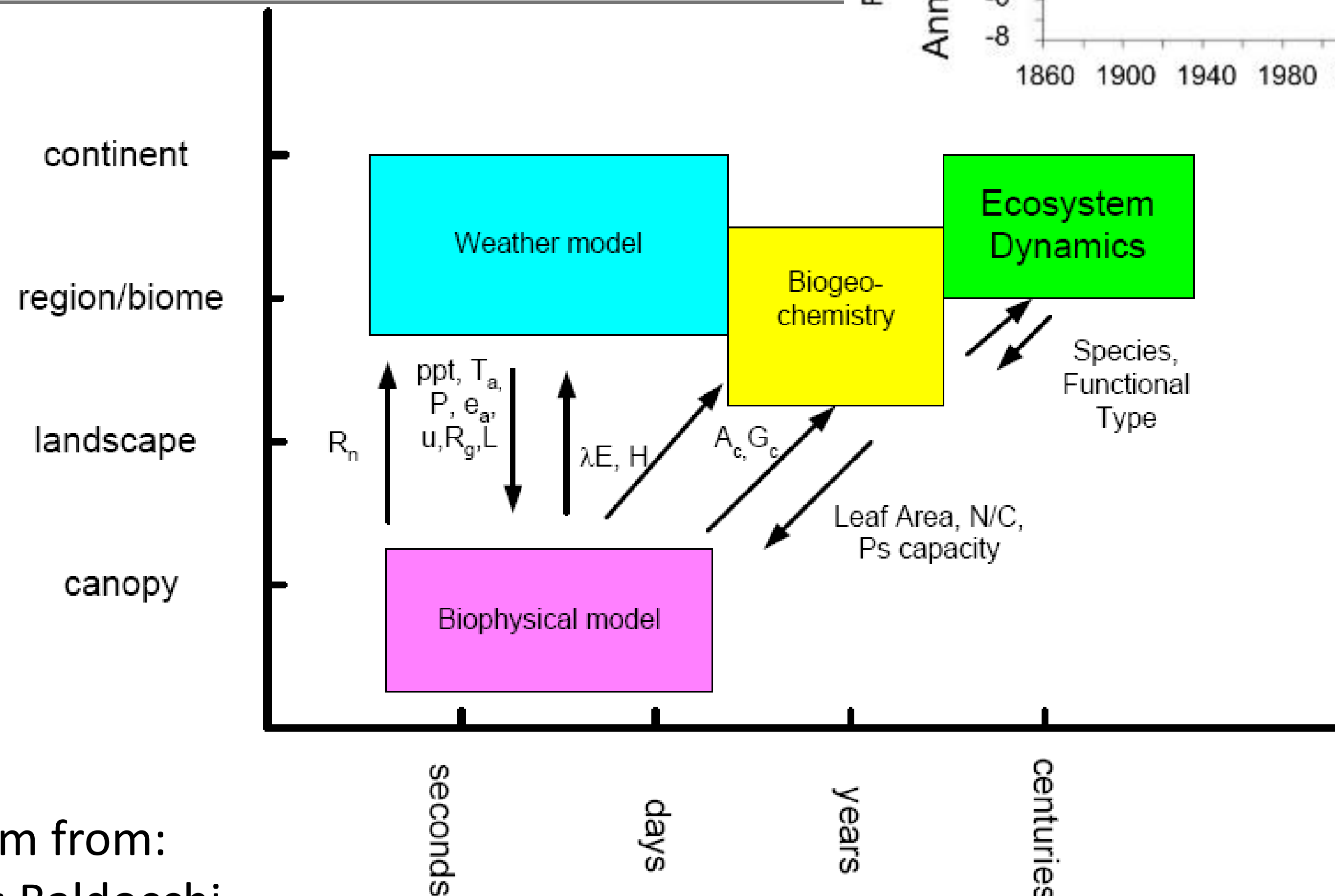
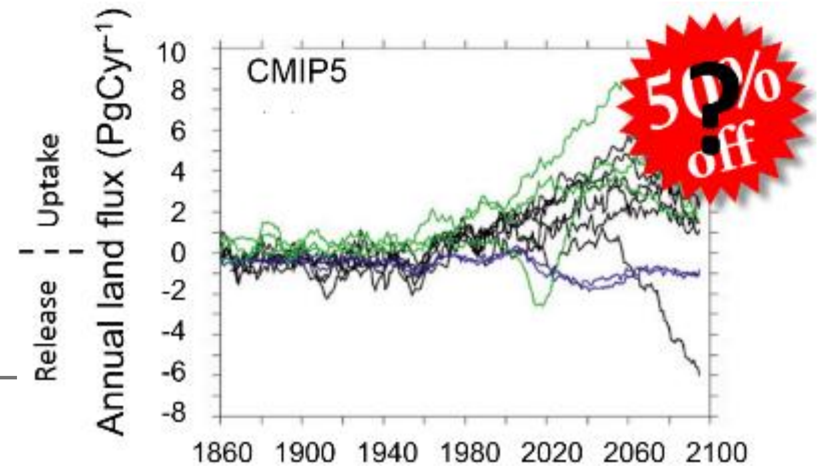


Diagram from:
Dennis Baldocchi

Better Spaghetti?

Fluxes don't help much with longer time scales



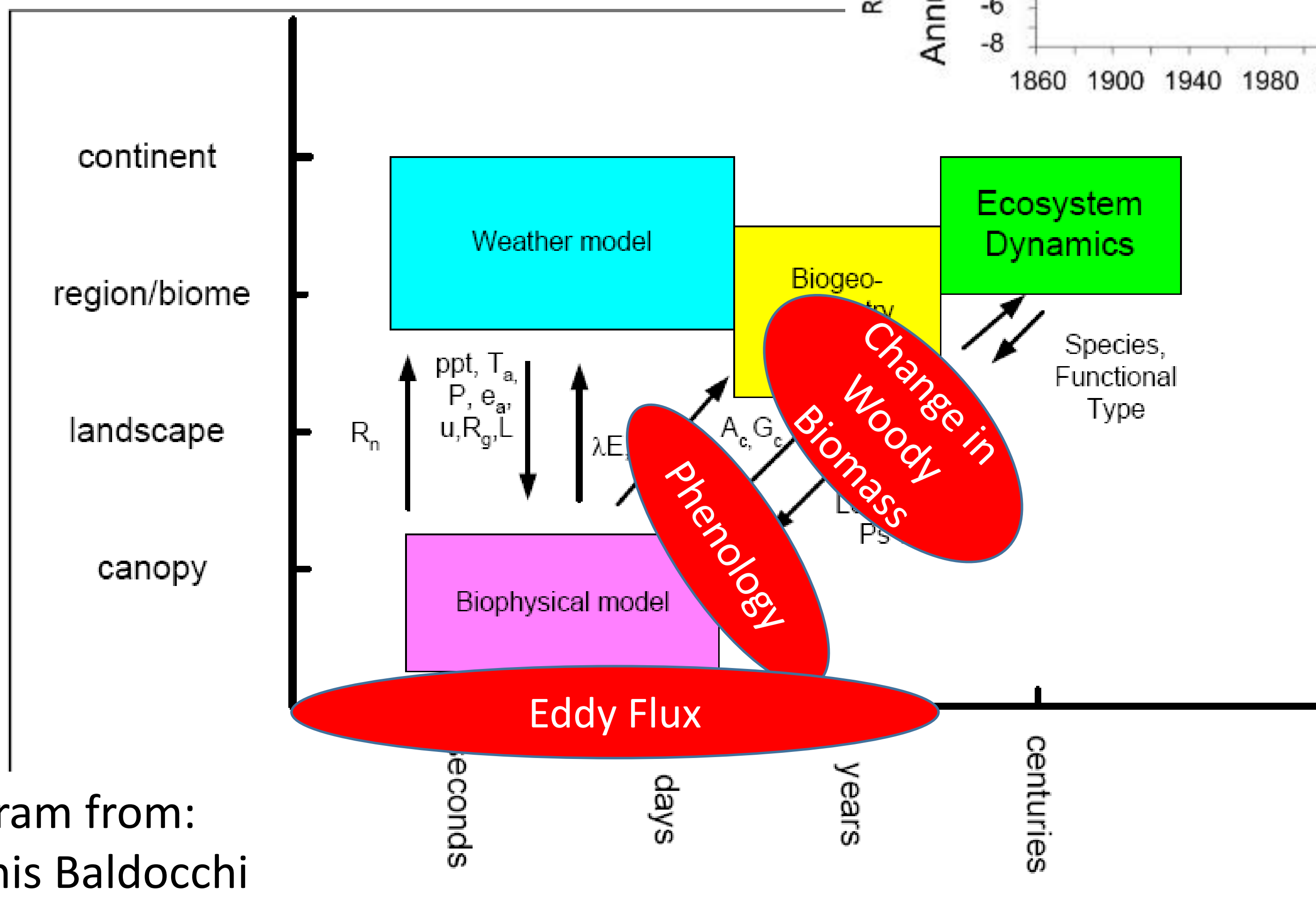
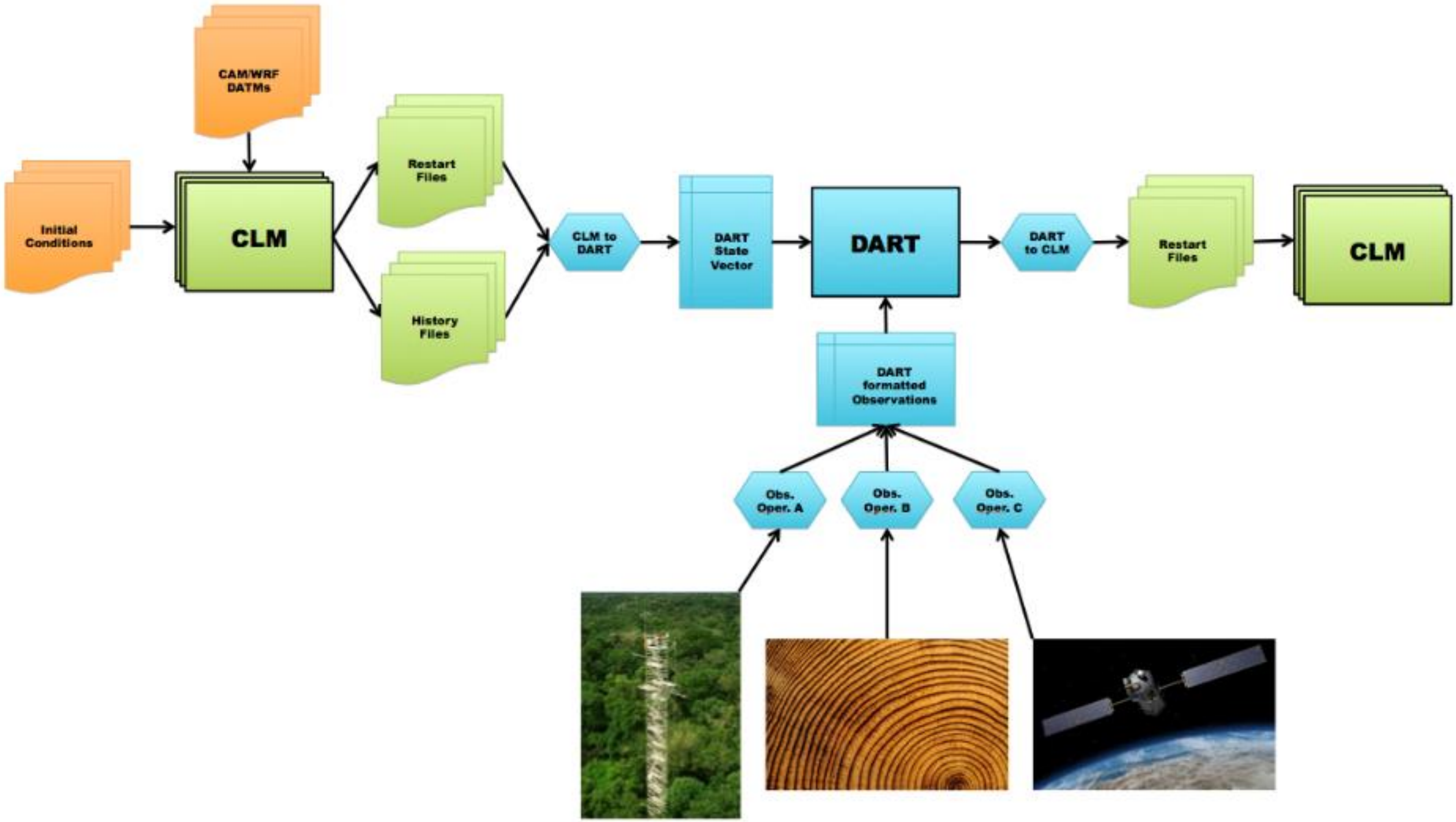


Diagram from:
Dennis Baldocchi

CLM-DART an Earth System Model DA system



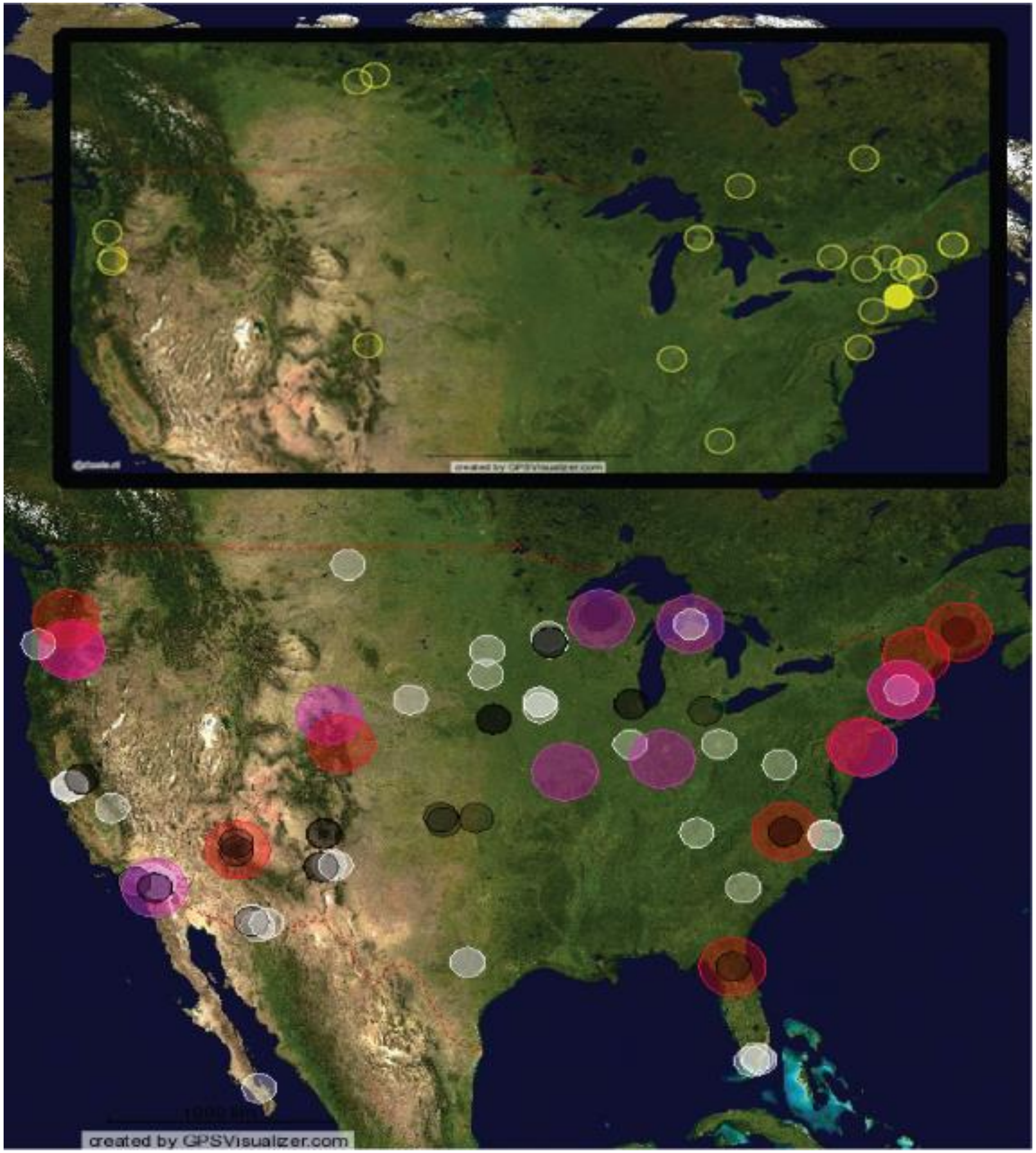
CLM-DART Development Strategy

1. Multi-instance capability in CESM
2. CLM to DART coupling
3. CLM-DART setup scripts
4. Add observation processors
5. Test at site level with synthetic experiments
6. Test at site level with real observations
7. Test globally with synthetic experiments
8. Test globally with real observations
9. Iterate 4-7 as new observations are added

Site Level Data Assimilation Lessons

- Assimilate combinations of different observations
  - MODIS Leaf Area Index Product
  - Plot biomass estimates
  - Flux tower Net Ecosystem Exchange
- Having carried this out at a number of site, we consistently find assimilating LAI tends to reduce fit with NEE
- This suggests am issue with model structure

DOE Regional and Global Climate Modeling DE-SC0016011

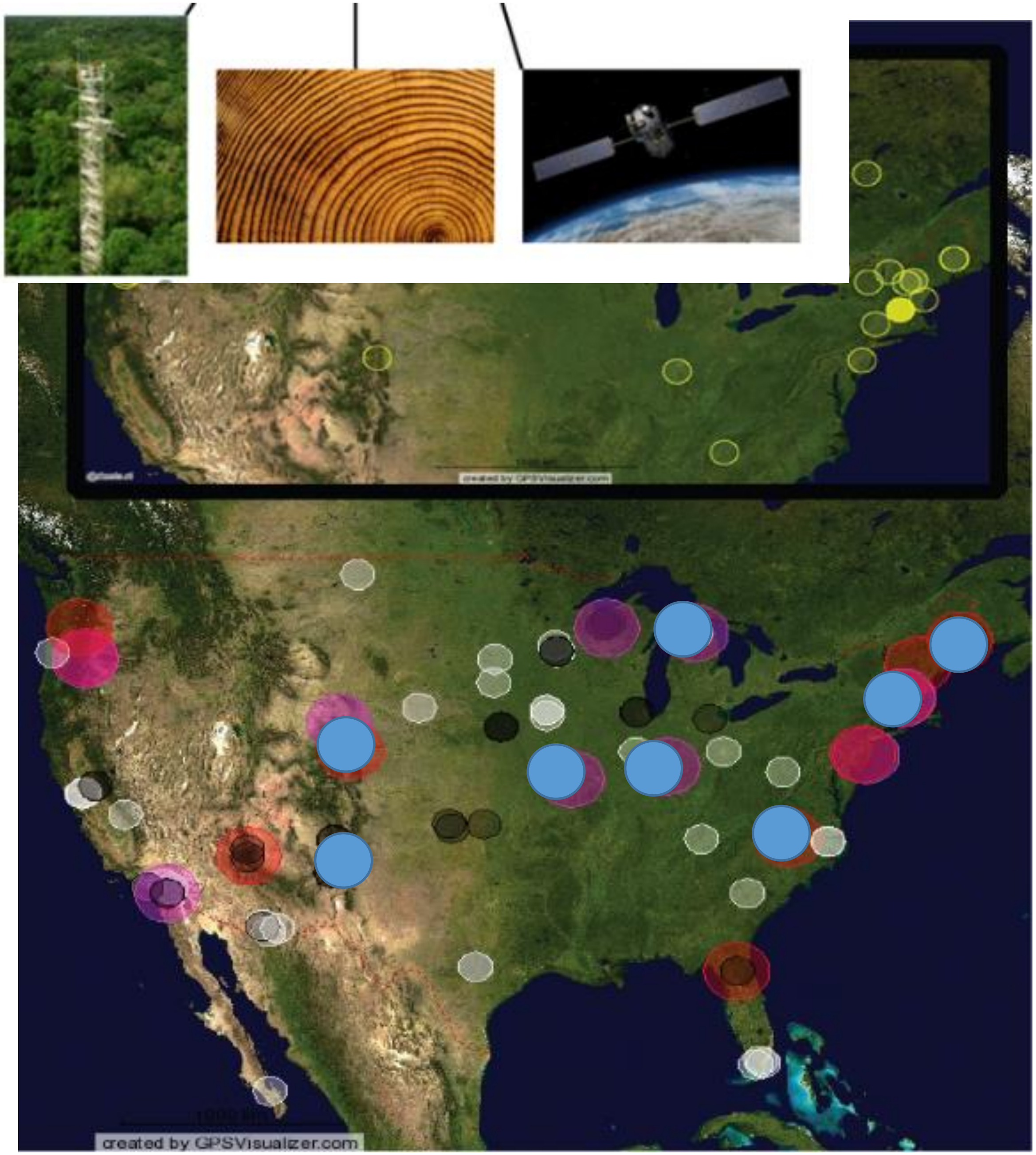


TESTING LAND SURFACE MODELS AT THREE DIFFERENT TIMESCALES USING THE AMERIFLUX NETWORK

YELLOW: PHENOCAM NETWORK

DOTS: AMERIFLUX SITES

BLACK: KNOWN BIOMETRIC DATA
PINK: TREE RING SAMPLING



TESTING LAND SURFACE MODELS AT THREE DIFFERENT TIMESCALES USING THE AMERIFLUX NETWORK

YELLOW: PHENOCAM NETWORK

DOTS: AMERIFLUX SITES

BLACK: KNOWN BIOMETRIC DATA
PINK: TREE RING SAMPLING

For effective data assimilation

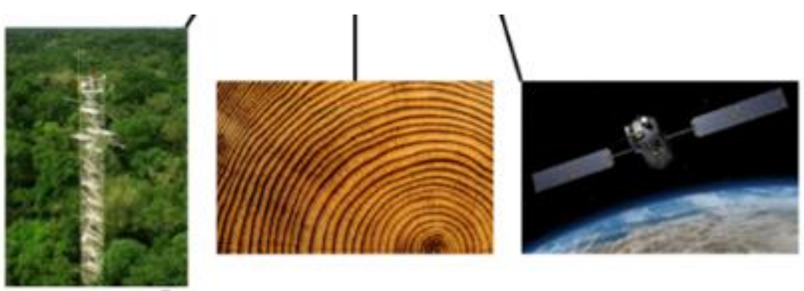
The model and the data must have a common means of communication
- either the model predicts the data type being assimilated
- or we have a way to translating the data or model so that they can be compared statistically

The model should contain the processes that govern the data

the assimilation could fail or
the resulting combination could be spurious

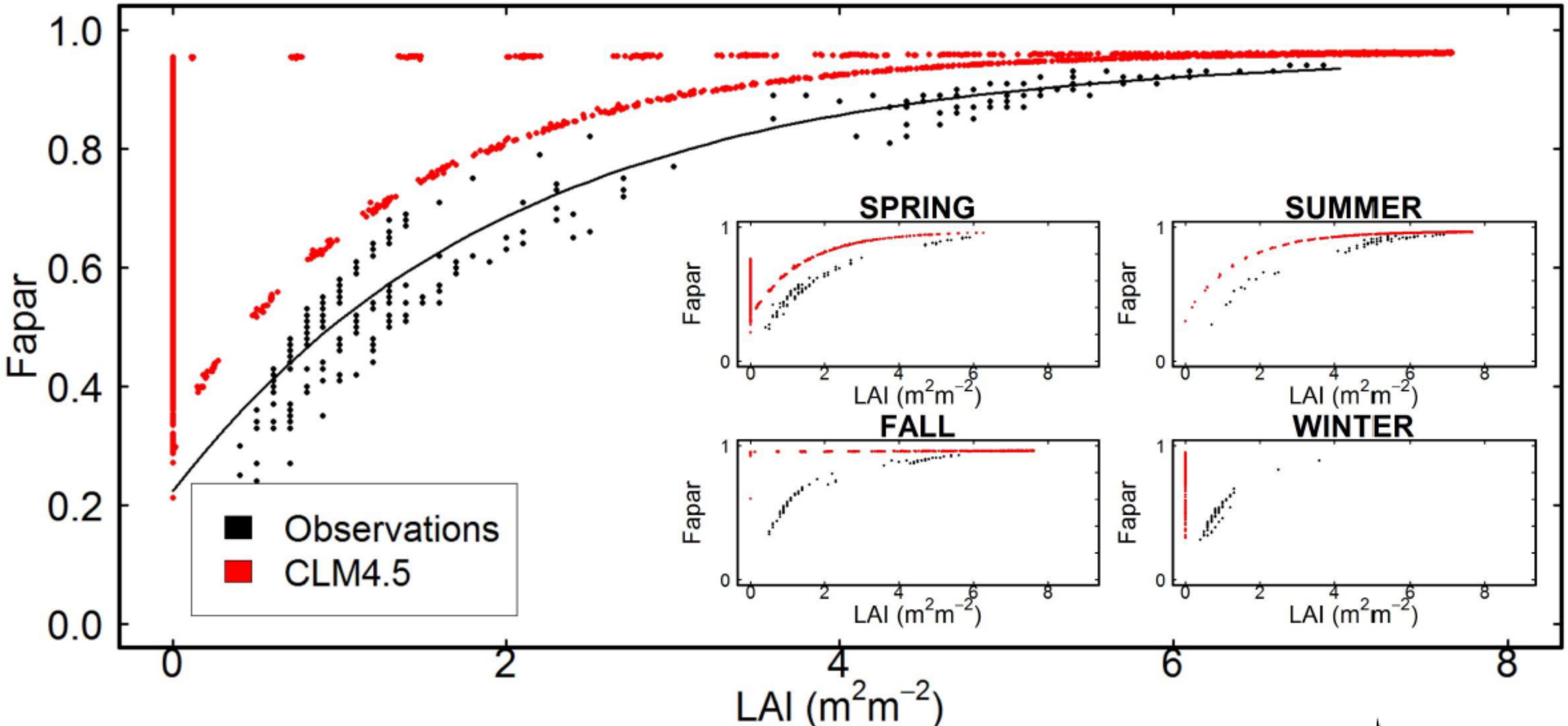
The uncertainty in the dataset should be well characterized

Otherwise either the model or the data will be given too much weight



DOE Regional and Global Climate Modeling DE-SC0016011

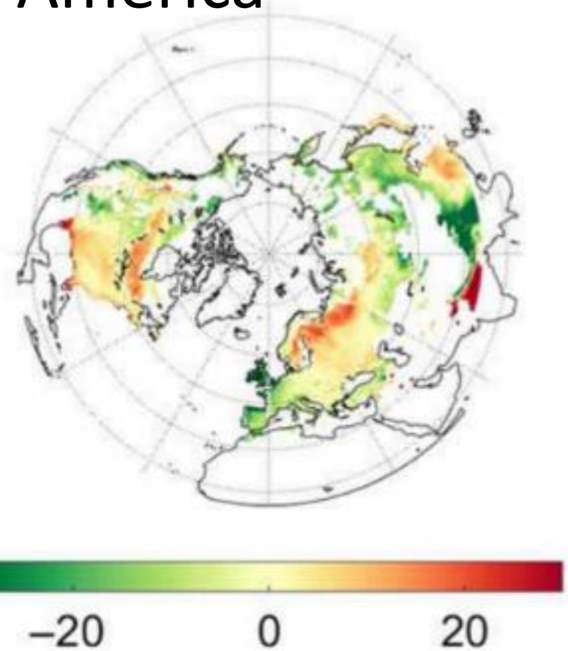
Phenology Problems: CLM fAPAR mismatch in spring with MODIS fAPAR



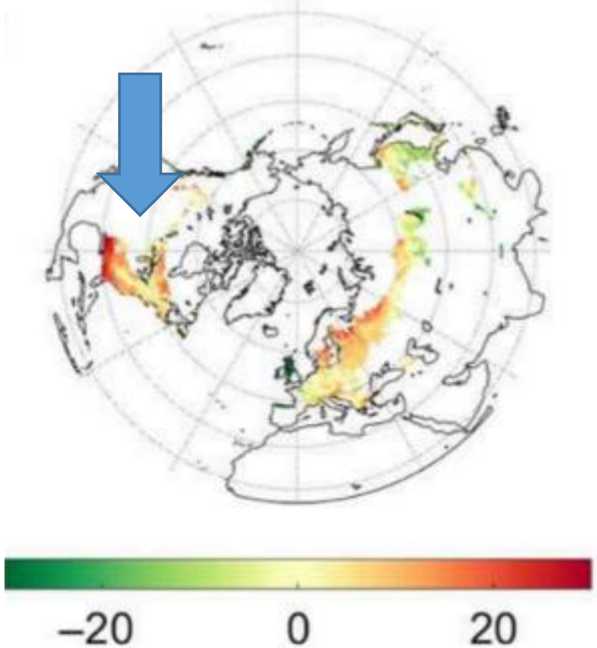
Montane et al. *NACP 2015*

DOE Regional and Global Climate Modeling DE-SC0016011

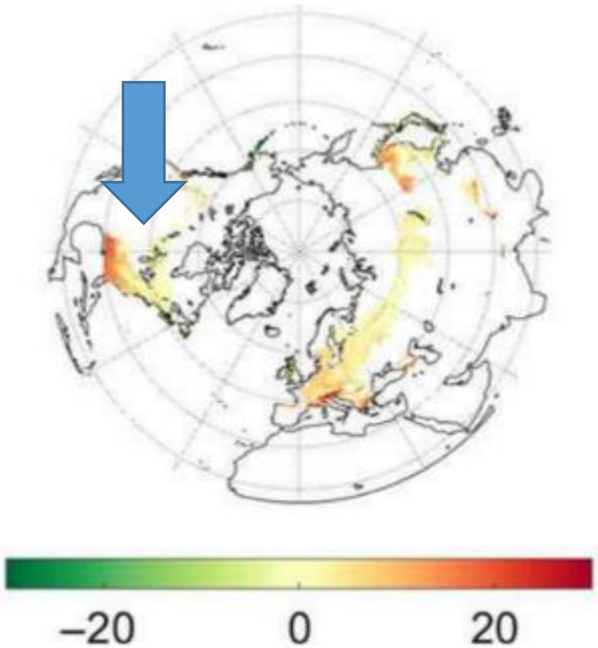
Modification of CLM phenology module: Addition of chilling process improves fit with MODIS data in North America



CLM vs CLMphenocam



CLM vs MODIS



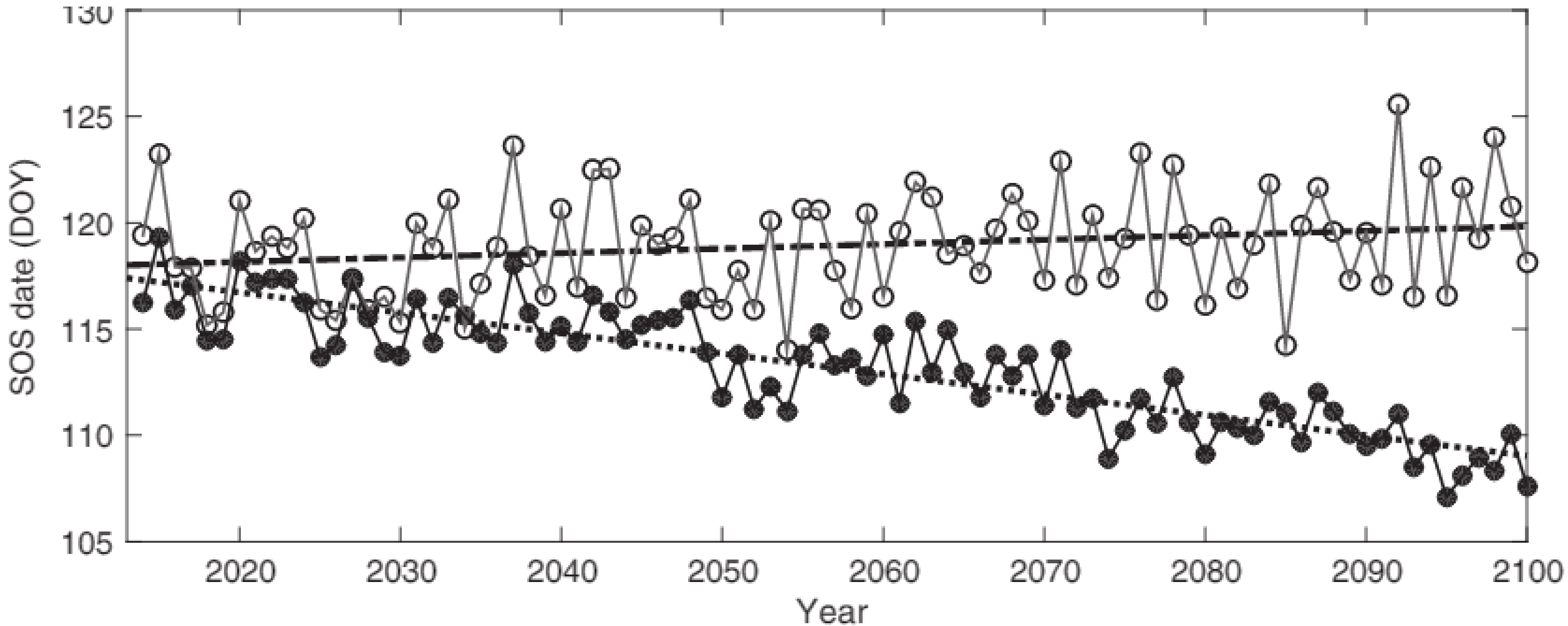
CLMphenocam vs MODIS

New Phenology Module parameterizing based on Phenocam data

Chen *et al*. 2016 Global Change Biology

DOE Regional and Global Climate Modeling DE-SC0016011

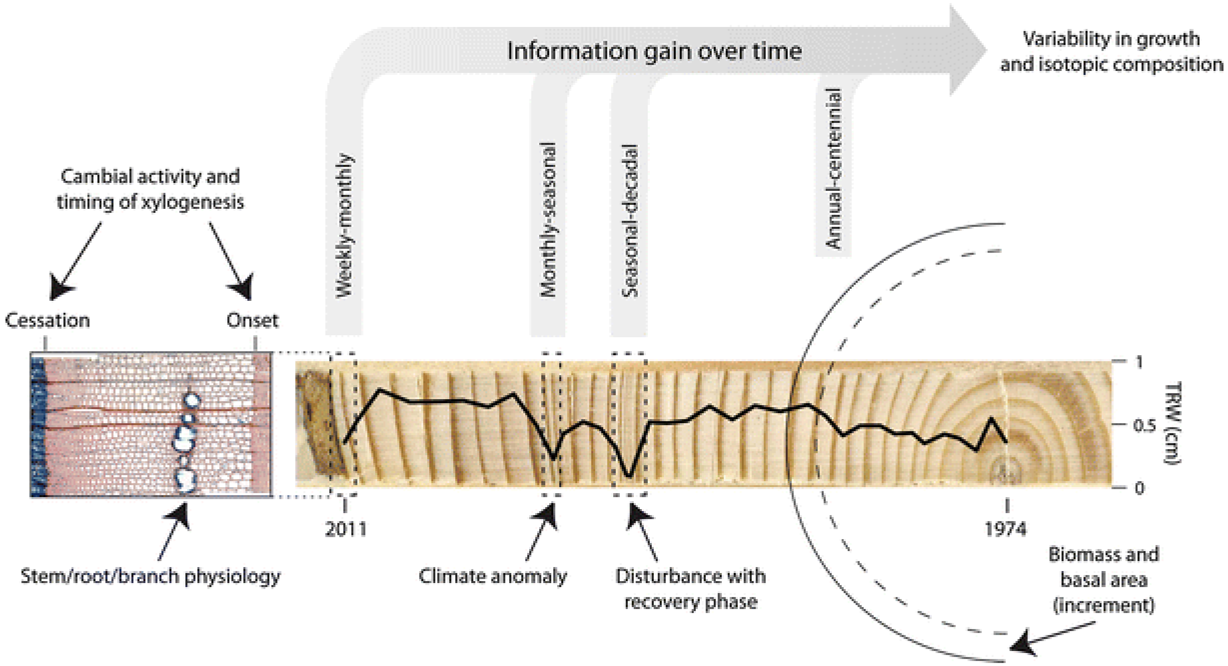
Improved CLM phenology module translates to earlier springs in RCP8.5 projections



Chen *et al*. 2016 Global Change Biology

DOE Regional and Global Climate Modeling DE-SC0016011

Tree rings - Linking the carbon cycle and climate
Decadal-centennial constraints for Earth System Models



Information content of tree increment cores can provide a constraint on Land Surface Models with respect to:

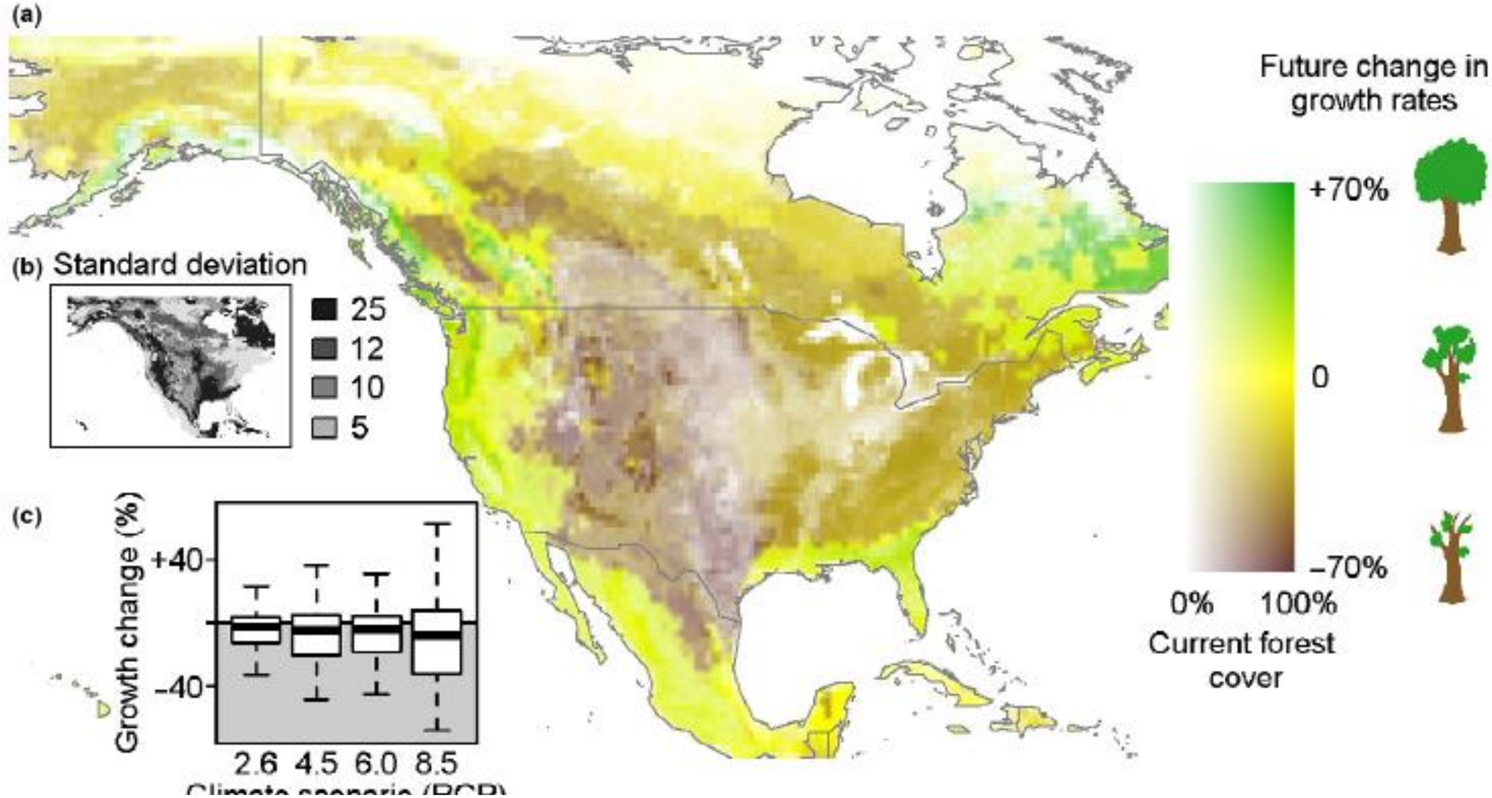
1. growth phenology
2. forest productivity
3. CO_2 fertilization
4. forest disturbances
5. vegetation model evaluation

Increment cores can augment short term metrics like eddy covariance towers at seasonal, inter-annual and decadal-centennial timescales

Babst F, Alexander MR, Szejner P, Bouriaud O, Klesse S, Roden J, Ciais P, Poulter B, Frank D, Moore DJP, Trouet V (2014) A tree-ring perspective on the terrestrial carbon cycle *Oecologia* 176 (2), 307-322 OECO-D-14-00512

DOE Regional and Global Climate Modeling DE-SC0016011

Observed forest sensitivity to climate implies large changes in 21st century North American forest growth

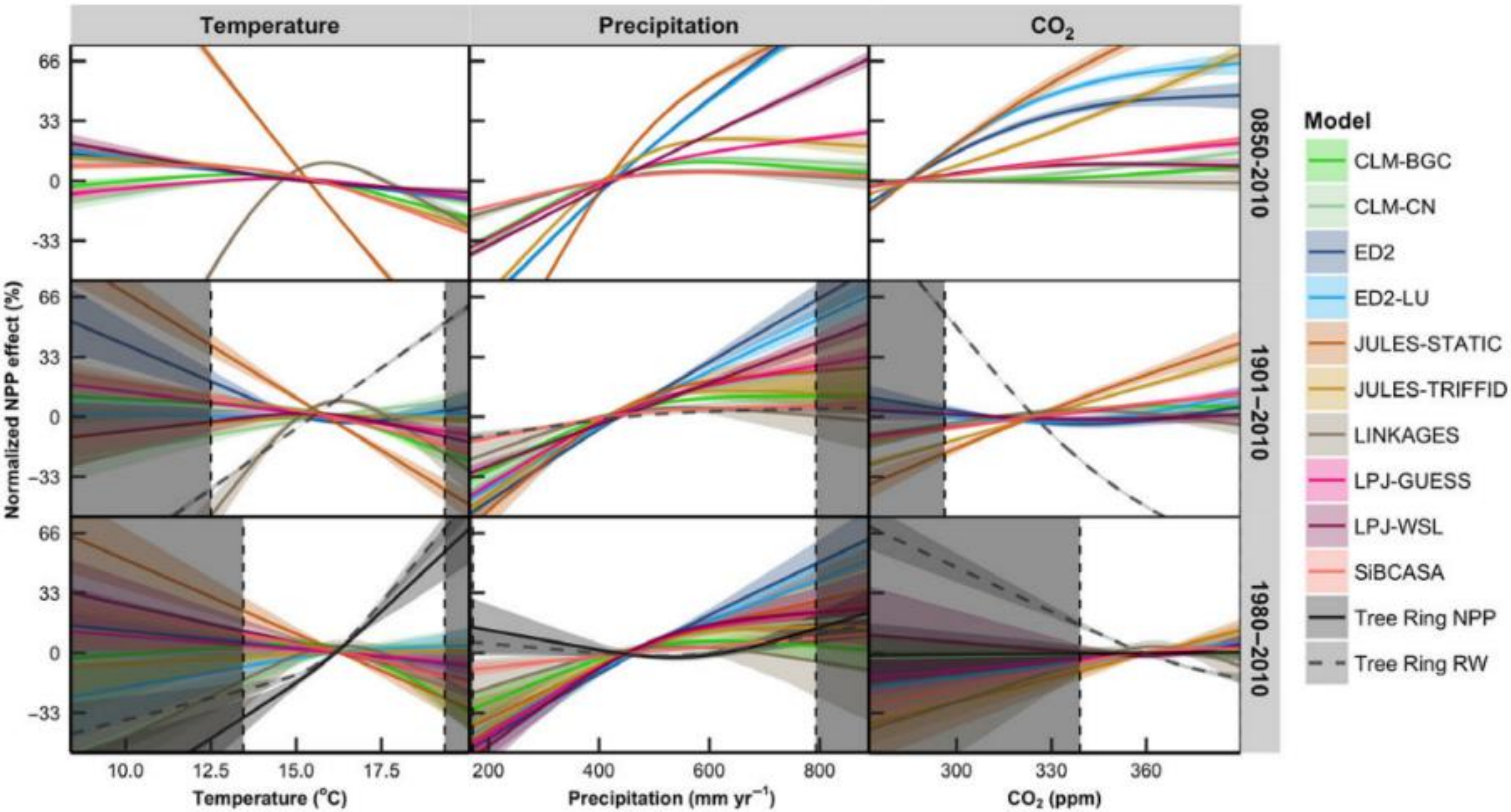


(1) climate change negatively impacted forest growth rates in the interior west and positively impacted forest growth along the western,southeastern and northeastern coasts;
(2) shifting climate sensitivities offset positive effects of warming on high-latitude forests, leaving no evidence for continued ‘boreal greening’;
(3) It took a 72% WUE enhancement to compensate for continentally averaged growth declines under RCP 8.5.

Charney et al 2016 Ecology Letters

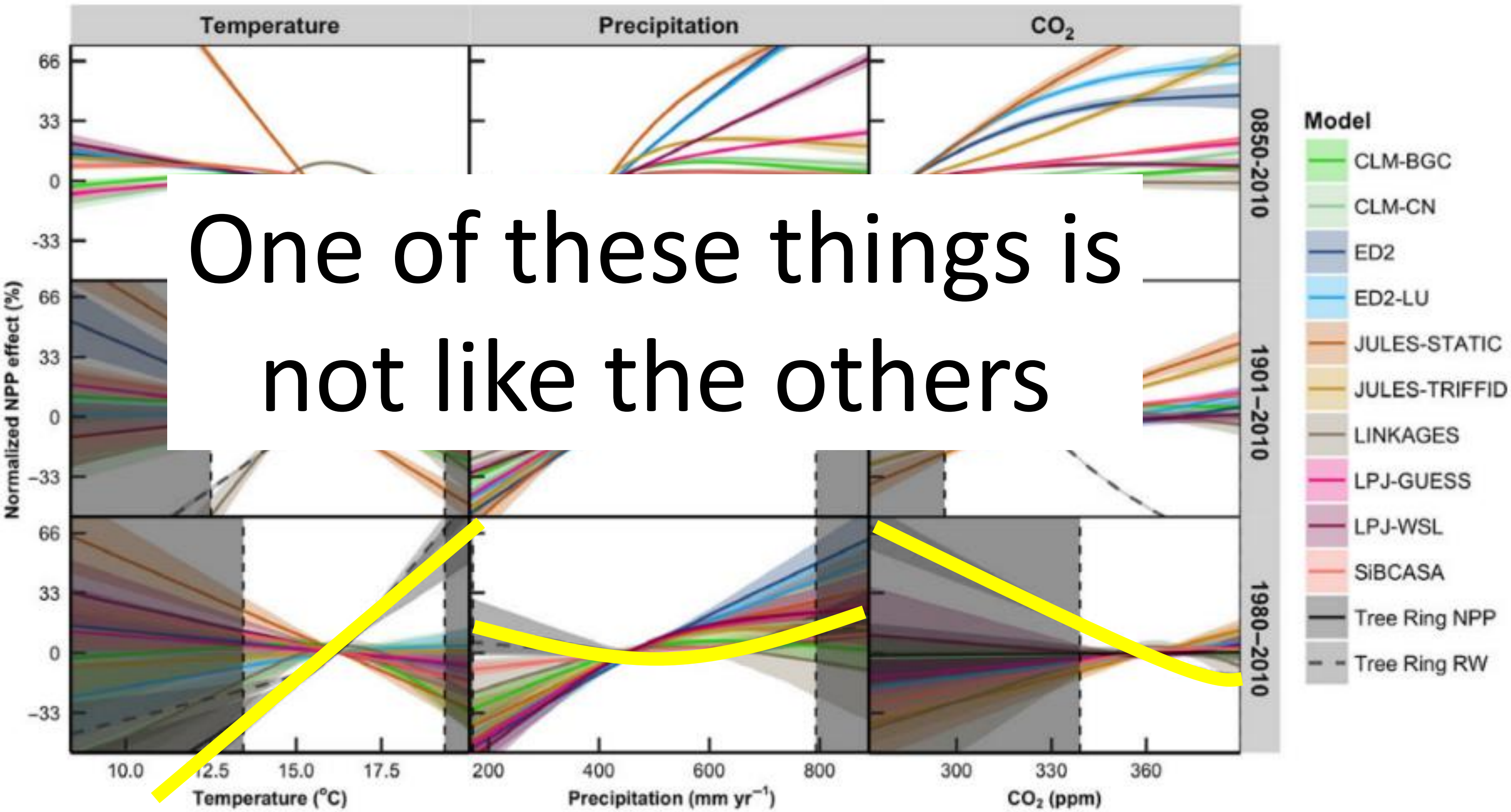
DOE Regional and Global Climate Modeling DE-SC0016011

Tree Rings and Terrestrial Biosphere Models



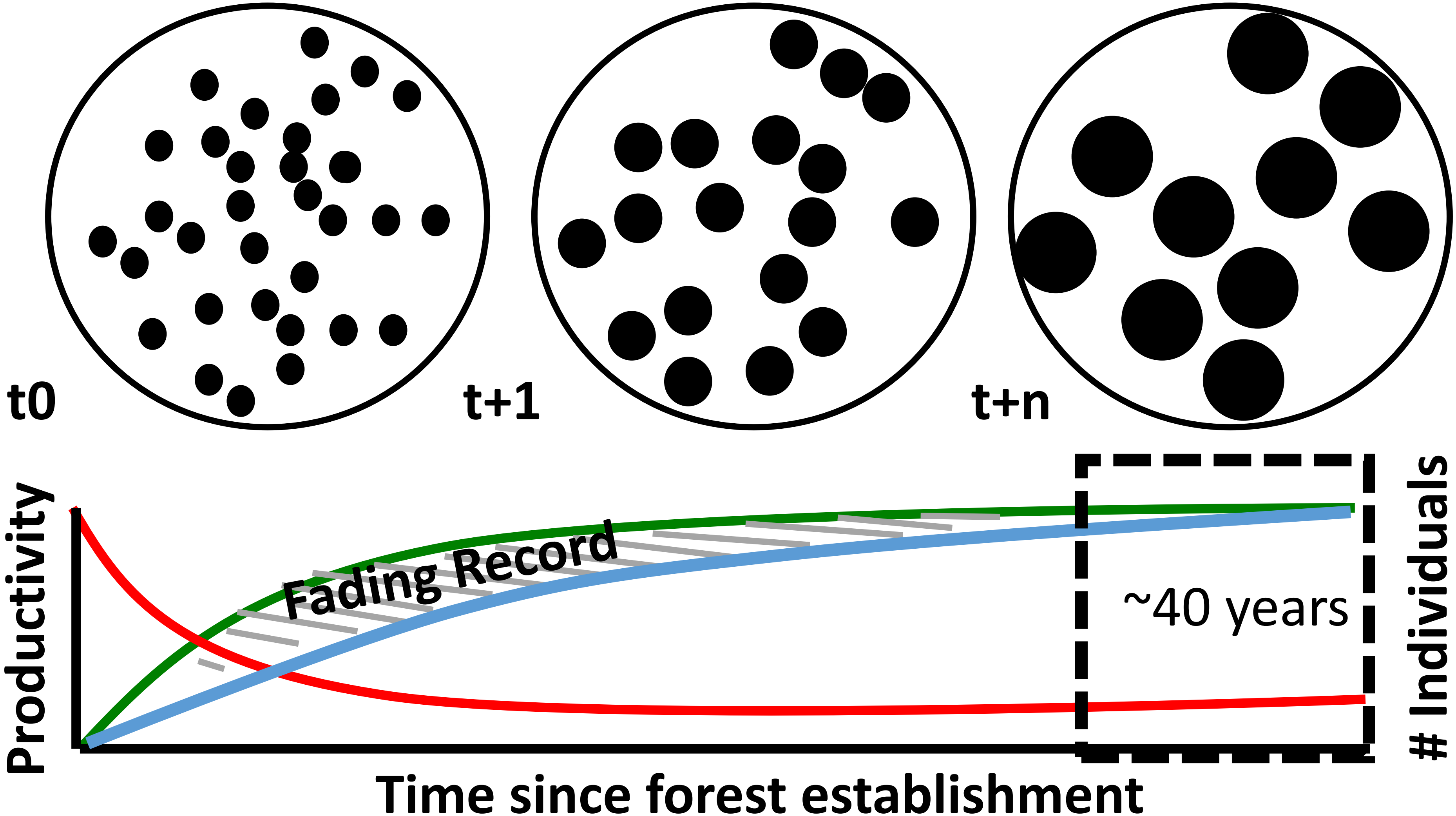
Rollinson et al 2017 *Glob Change Bio*

Tree Rings and Terrestrial Biosphere Models



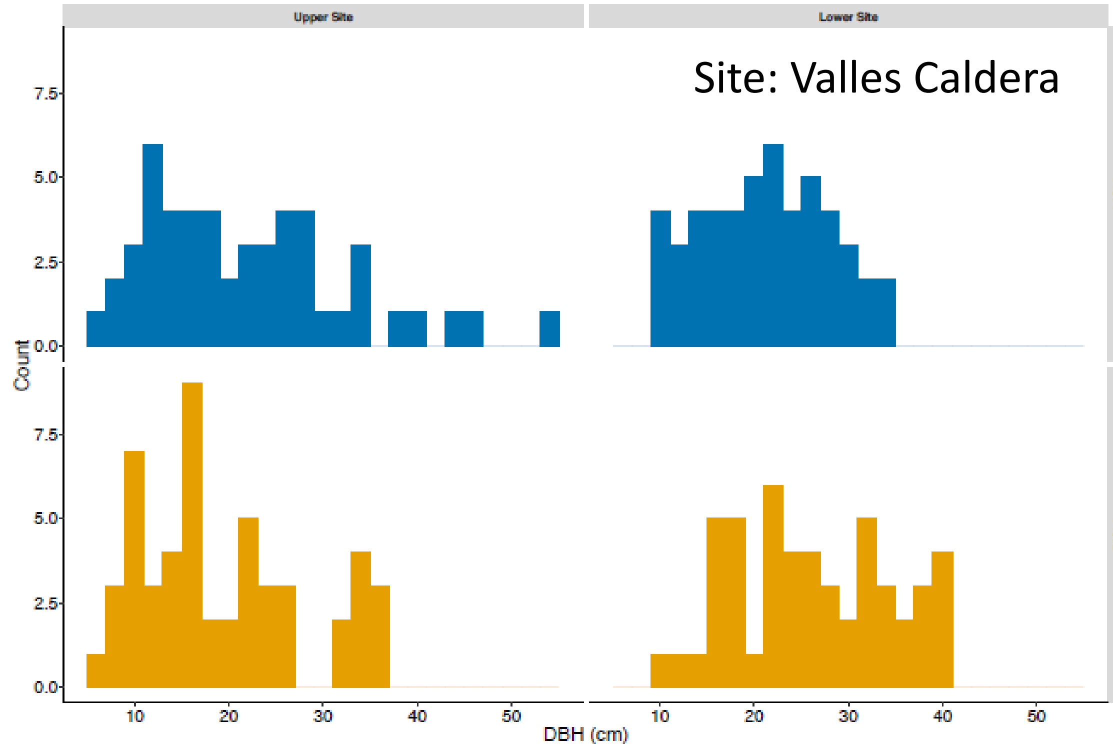
Rollinson et al 2017 *Glob Change Bio*

Reconstructing stand level NPP dynamics



Real forests are complex

Sampling Tree Rings Ecologically



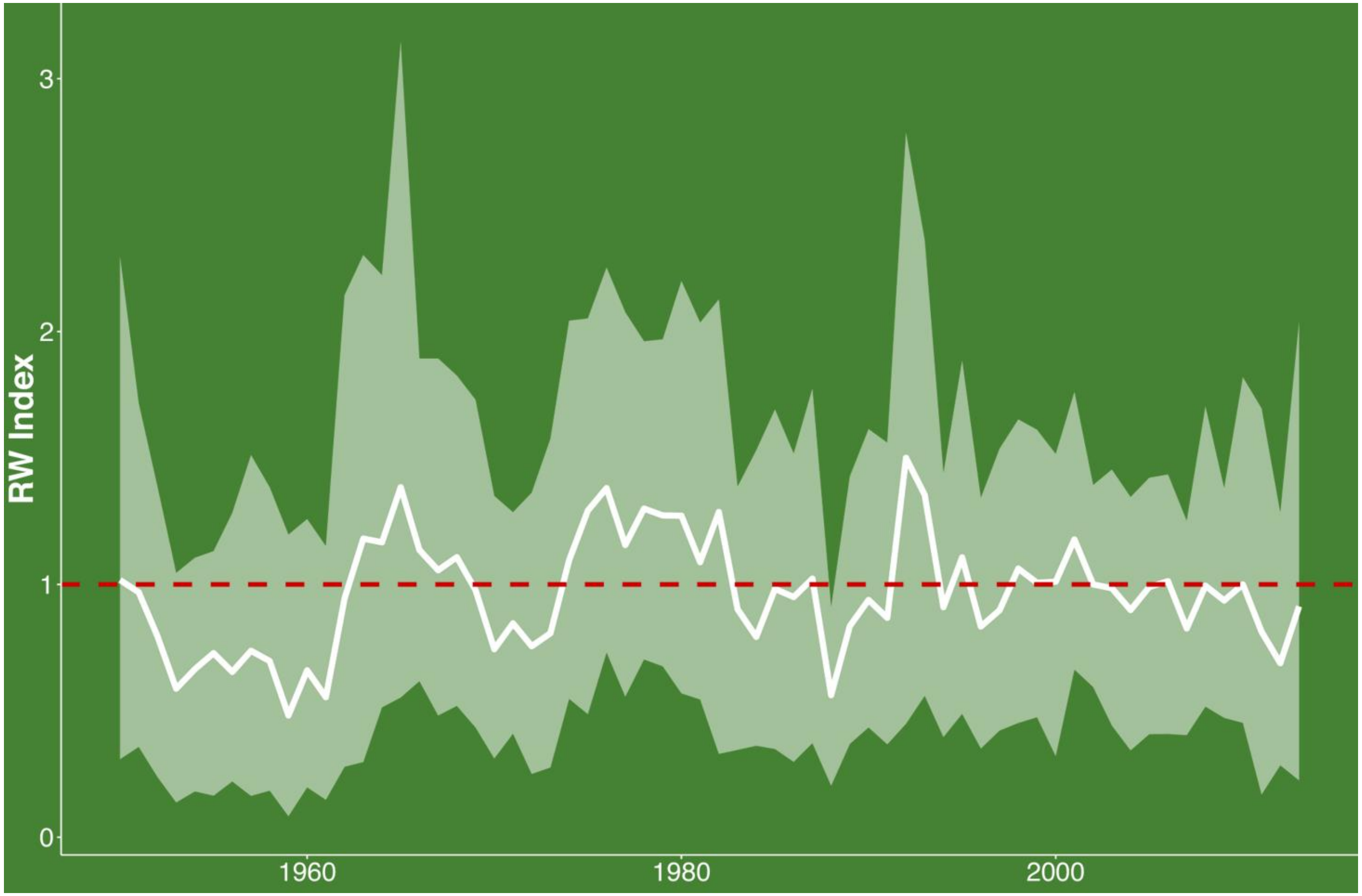
"We collected tree cores from 13 different sites from across the US. We traveled over 10K miles that summer to collect cores from over 1500 trees."

3000+ cores were mounted, sanded, cross dated, measured and statistically analysed.
Ross Alexander, PhD

Alexander et al 2017a *in revision TREES S&F*

DOE Regional and Global Climate Modeling DE-SC0016011

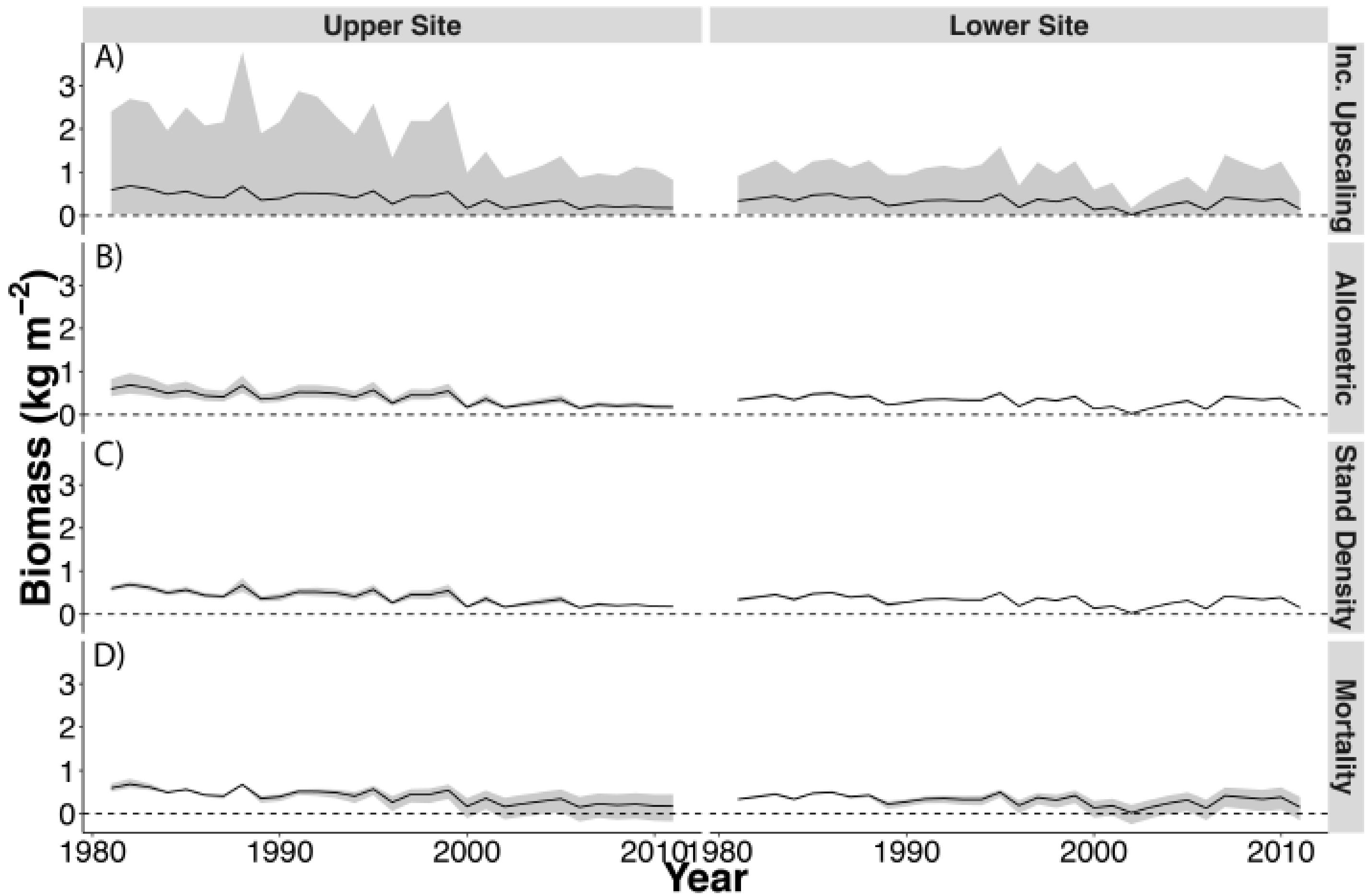
We can reconstruct the record of ring width for trees through time



CLM does not understand ring width index

Alexander et al 2017 *PhD Thesis*

Ring width is translated to biomass increment using allometric relationships

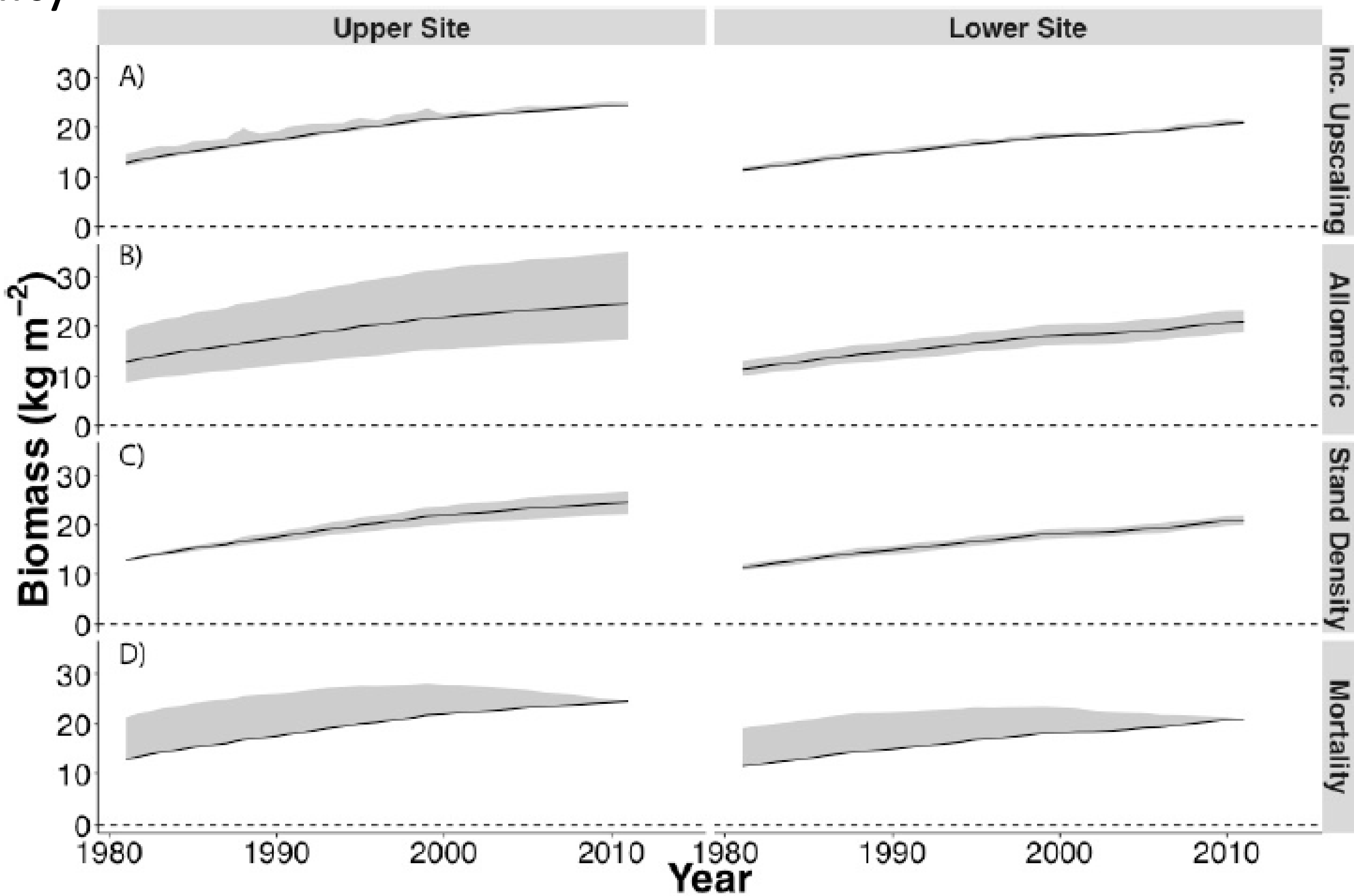


Site: Valles Caldera

Alexander et al 2017a *in revision TREES S&F*

DOE Regional and Global Climate Modeling DE-SC0016011

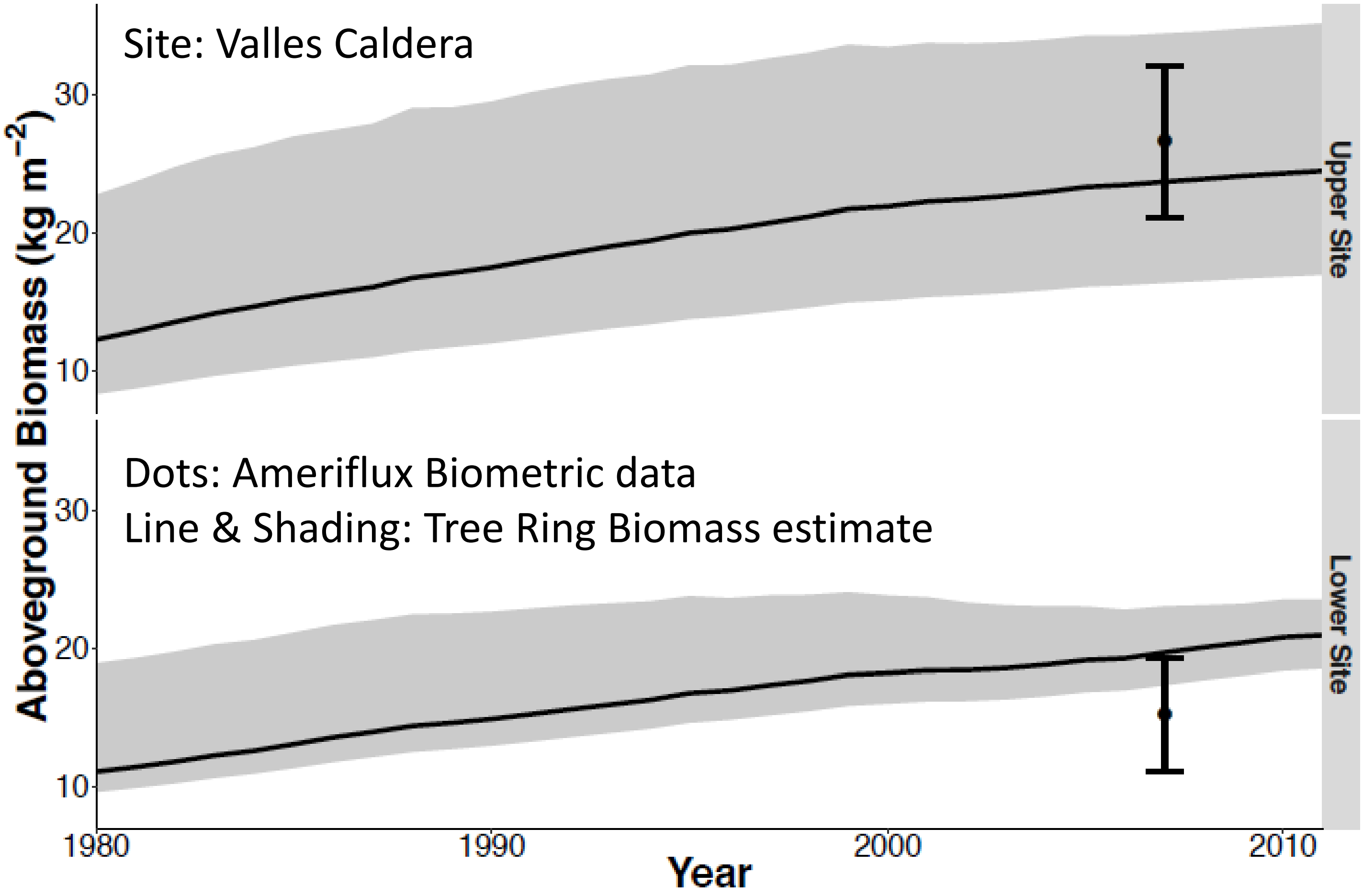
Total woody biomass is calculated for the period 1980 – now With uncertainty from increment, allometry, stand structure & mortality



Site: Valles Caldera

Alexander et al 2017a *in revision TREES S&F*

Total woody biomass is calculated for the period 1980 – now With uncertainty from increment, allometry, stand structure & mortality



Alexander et al 2017a *in revision TREES S&F*

DOE Regional and Global Climate Modeling DE-SC0016011

Challenge the model's structure – how well can CLM replicate biometric observations?

- 4 EVERGREEN FORESTS: Niwot Ridge, Valles Caldera, Howland, and Duke Forest Loblolly Pine

- 5 DECIDUOUS FORESTS: UMBS, Harvard, Missouri Ozark, Morgan Monroe, and Duke Forest Hardwoods

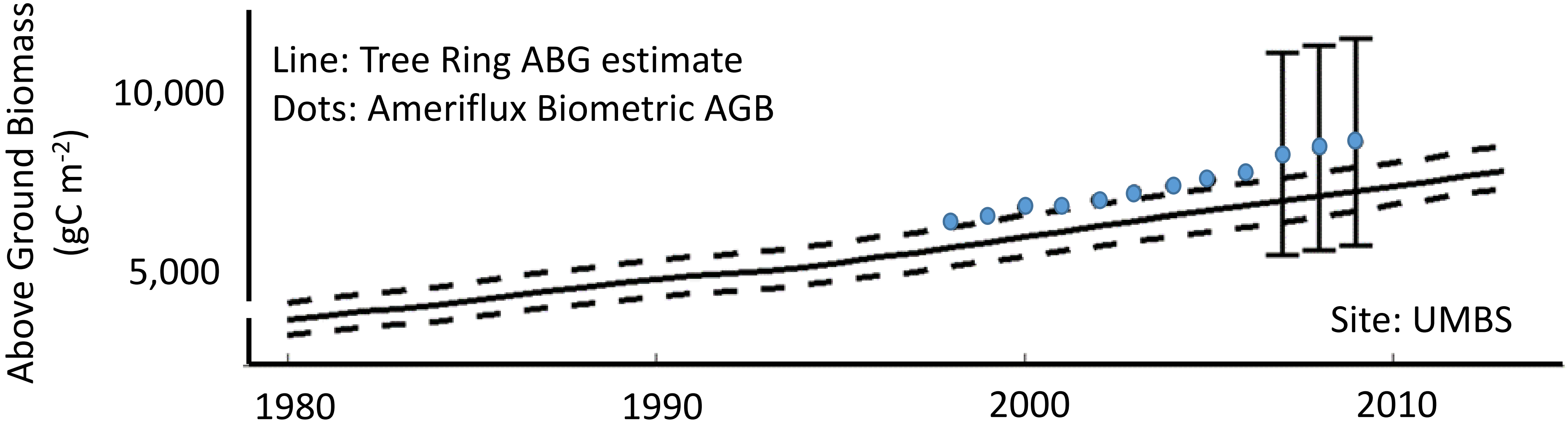
Challenge the model's structure – how well can CLM replicate biometric observations?

Site (ID)	Longitude	Latitude	Reference	C fluxes data	Aboveground biomass (AmeriFlux)	Aboveground biomass (tree-ring)	LAI in-situ data	Leaf C-LAI data	Stem C/Leaf C data
Evergreen									
Niwot Ridge (NR1)	-105.5464	40.0329	Blanken, 2016	1999-2013	2003	1980-2012	n.a.	n.a.	n.a.
Valles Caldera Mixed Conifer (Vcm)	-106.5321	35.8884	Litvak, 2016	2007-2013	2007	1980-2011	n.a.	n.a.	2007
Howland Forest (Ho1)	-68.7402	45.2041	Hollinger, 2016	1996-2004 2006-2013	2003	1980-2012	2006	n.a.	2003
Duke Forest Loblolly Pine (Dk3)	-79.0942	35.9782	Stoy et al., 2016	1998-2005	2001-2005	n.a.	2002-2005	2002-2005	n.a.
Deciduous									
University of Michigan Biological Station (UMB)	-84.7138	45.5598	Gough et al., 2009; Gough et al., 2013; Gough et al., 2016	2005-2013	1998-2011	1980-2013	1997-2013	1998-2009	1998-2009
Harvard Forest (Ha1)	-72.1715	42.5378	Munger, 2016	1992-2013	2006-2008	1980-2012	1998,1999, 2005-2008, 2010	n.a.	n.a.
Missouri Ozark (MOz)	-92.2000	38.7441	Wood and Gu, 2016	n.a.	n.a.	1980-2013	2006-2012	n.a.	n.a.
Morgan Monroe State Forest (MMS)	-86.4131	39.3232	Novick and Phillips, 2016	1999-2013	1999-2005	1980-2013	1999-2006, 2009	1999-2005	1999-2005
Duke Forest Hardwoods (Dk2)	-79.1004	35.9736	Oishi et al., 2016	2001-2005	2002	1980-2013	2006	n.a.	n.a.

Montane et al. *for* GMD

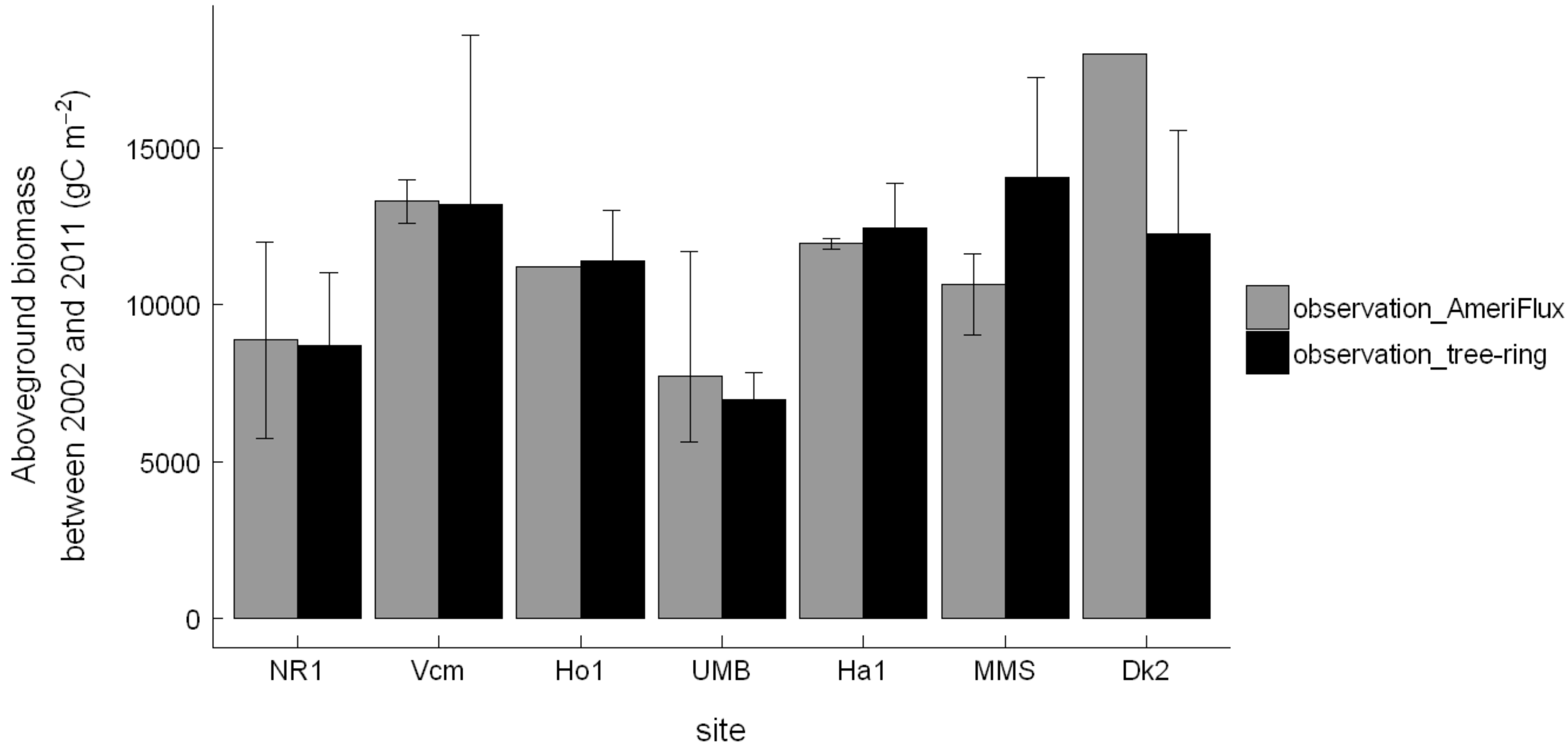
DOE Regional and Global Climate Modeling DE-SC0016011

Reasonable consistency between Ameriflux Biomass (various methods) and our tree ring reconstructed biomass



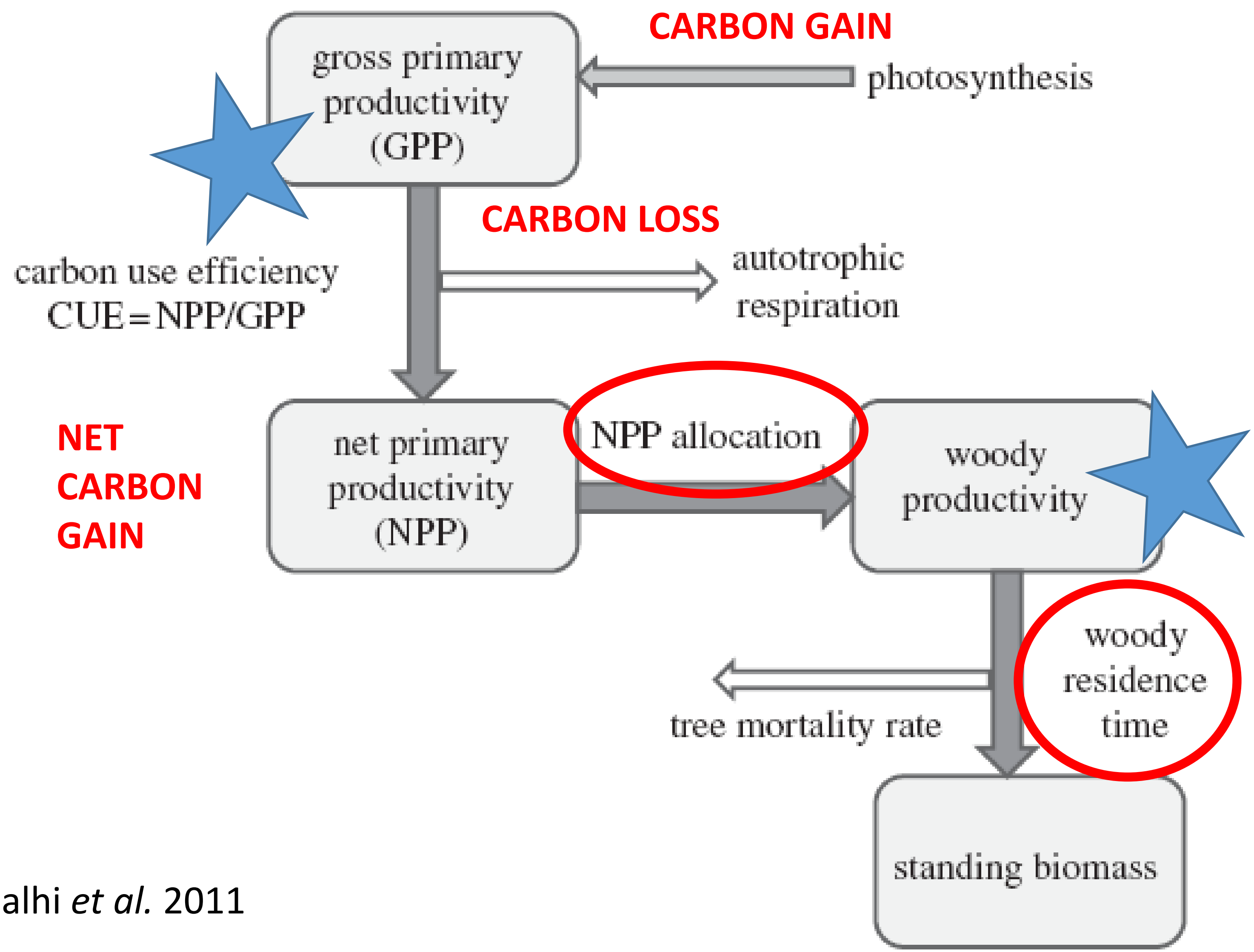
Montane et al. *for* GMD

Variation, but reasonable consistency between Ameriflux Biomass and Tree Ring Reconstructed Biomass



Montane et al. *for* GMD

DOE Regional and Global Climate Modeling DE-SC0016011



Malhi *et al.* 2011

PRODUCTIVITY TURNOVER

$$dB_i / dt = a_i NPP - u_i B_i$$

i= Plant pool i (leaves, stem, coarse roots and fine roots)

$\mathbf{B_i}$= Biomass of plant pool i (kg m^{-2})

dt = 1 year

$\mathbf{a_i}$= allocation coefficent for the plant pool i, and they add to 1.

NPP= Net Primary Productivity (kg m^{-2} $year^{-1}$)

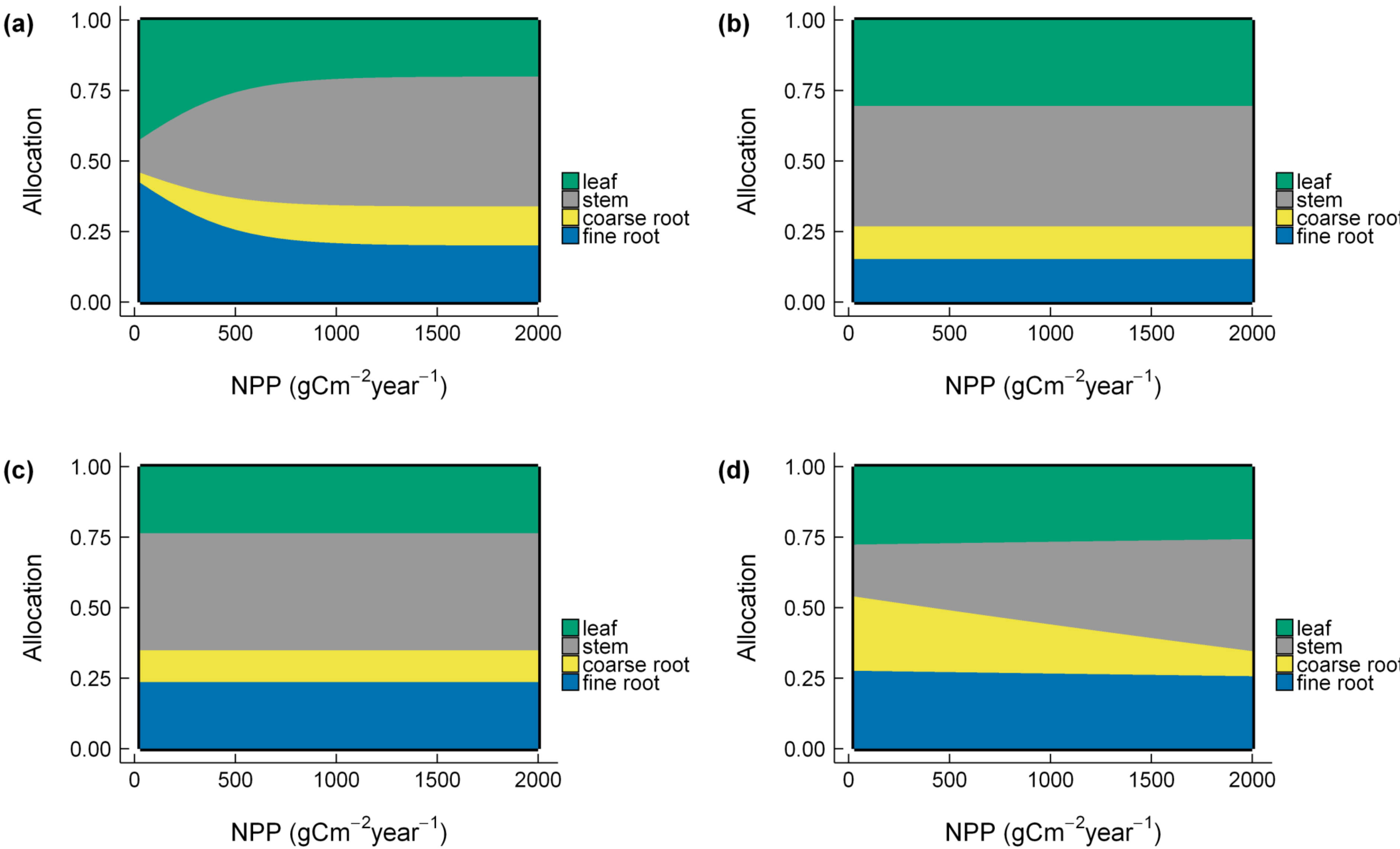
$\mathbf{u_i}$=turnover rate of plant pool i ($year^{-1}$)

PRODUCTIVITY TURNOVER

$$dB_i / dt = a_i NPP - u_i B_i$$

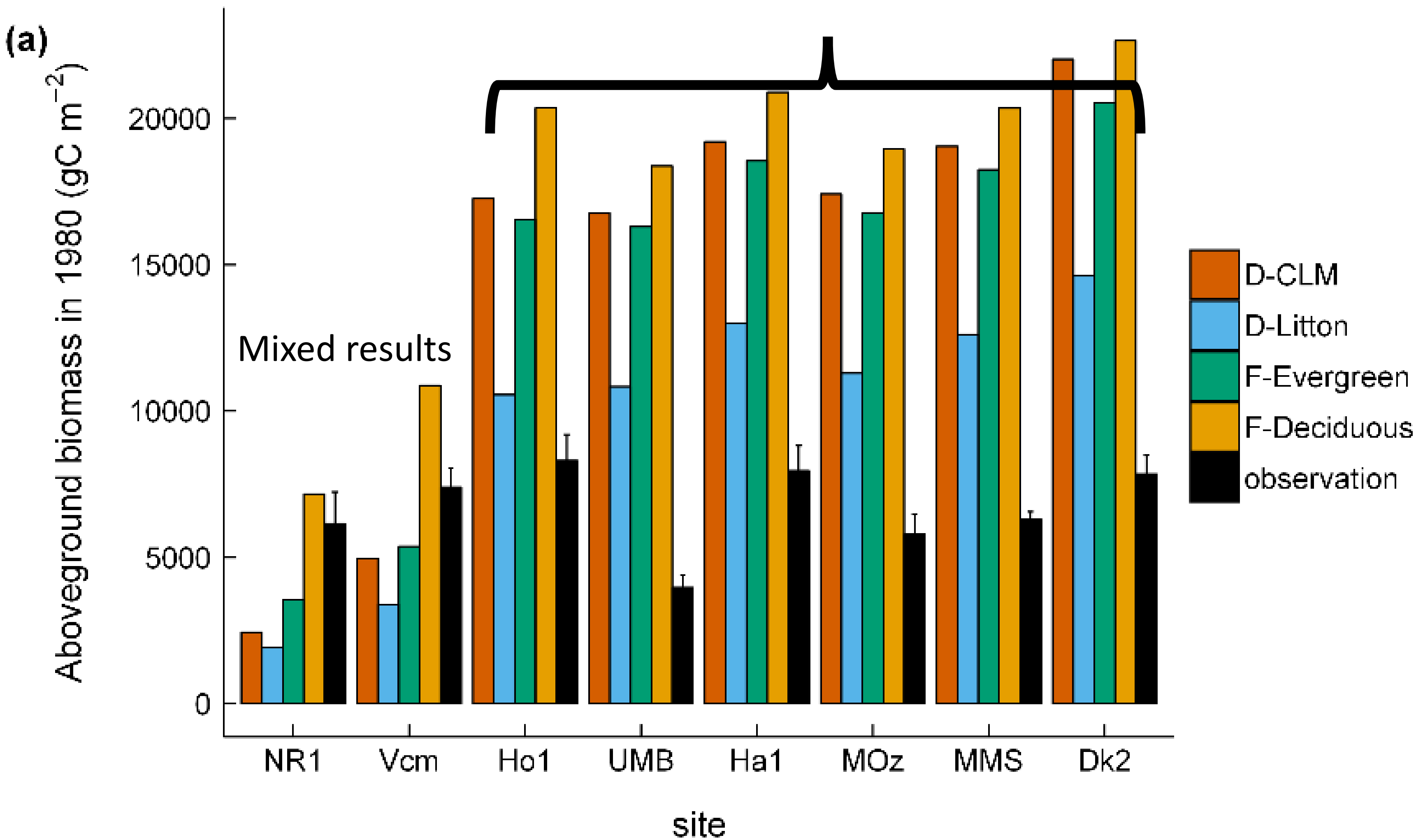
C ALLOCATION SCHEME

C ALLOCATION: ABOVEGROUND PRODUCTIVITY



Above ground biomass at start of the run (1980) mixed story but hints at D-Litton scheme

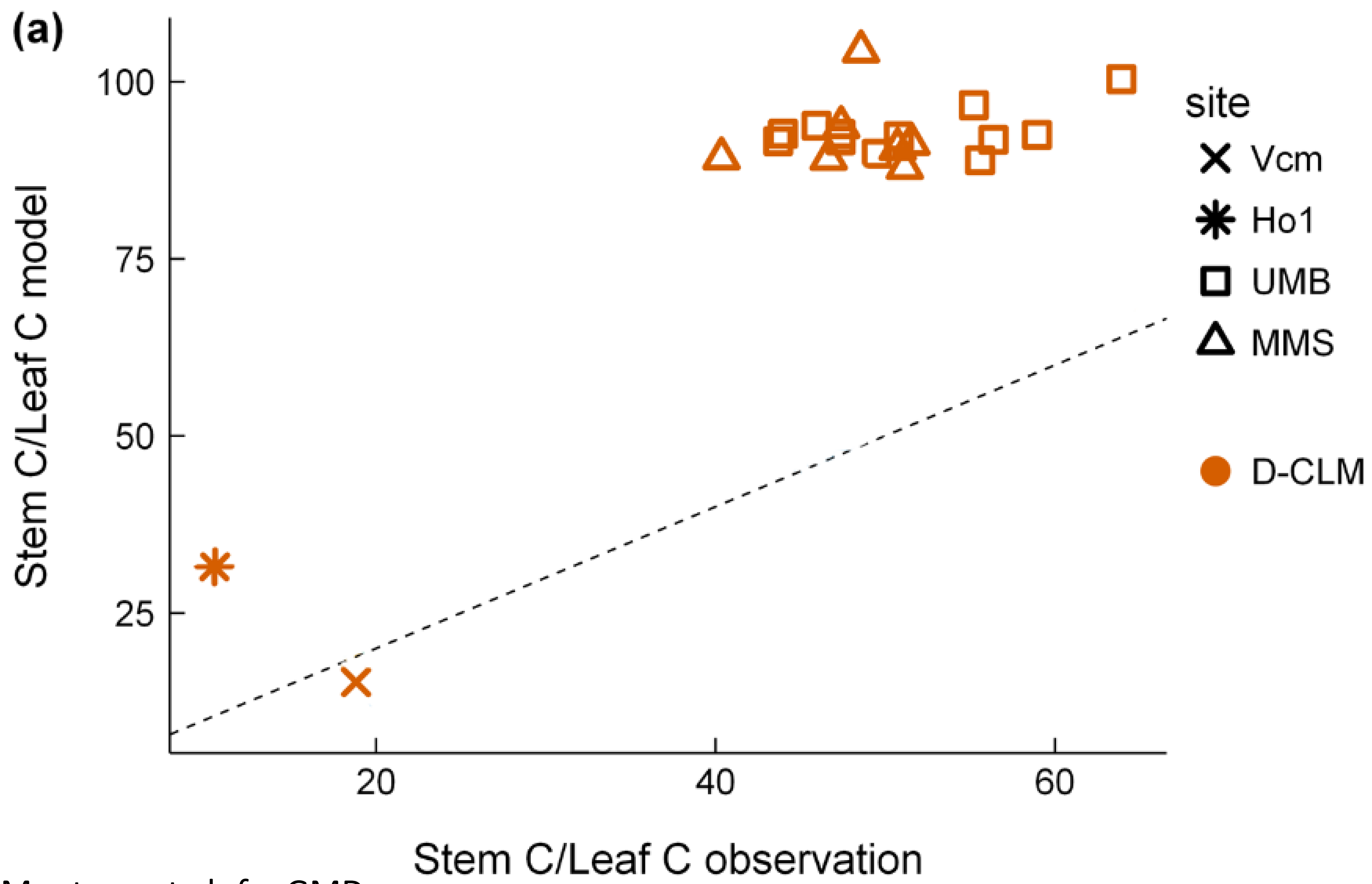
Overestimates



Montane et al. *for* GMD

DOE Regional and Global Climate Modeling DE-SC0016011

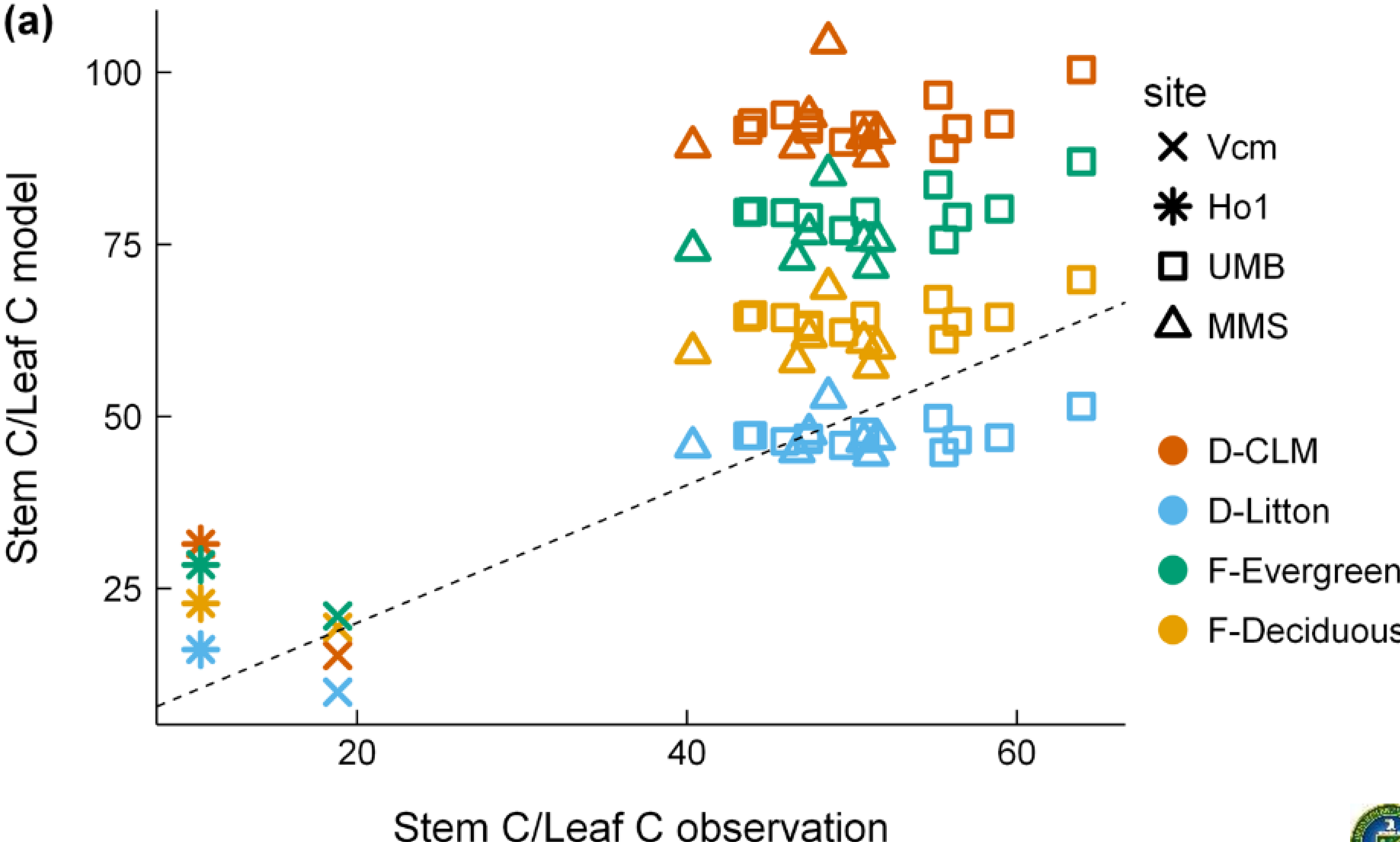
Ameriflux StemC/Leaf C available for 4 sites



Montane et al. *for* GMD

DOE Regional and Global Climate Modeling DE-SC0016011

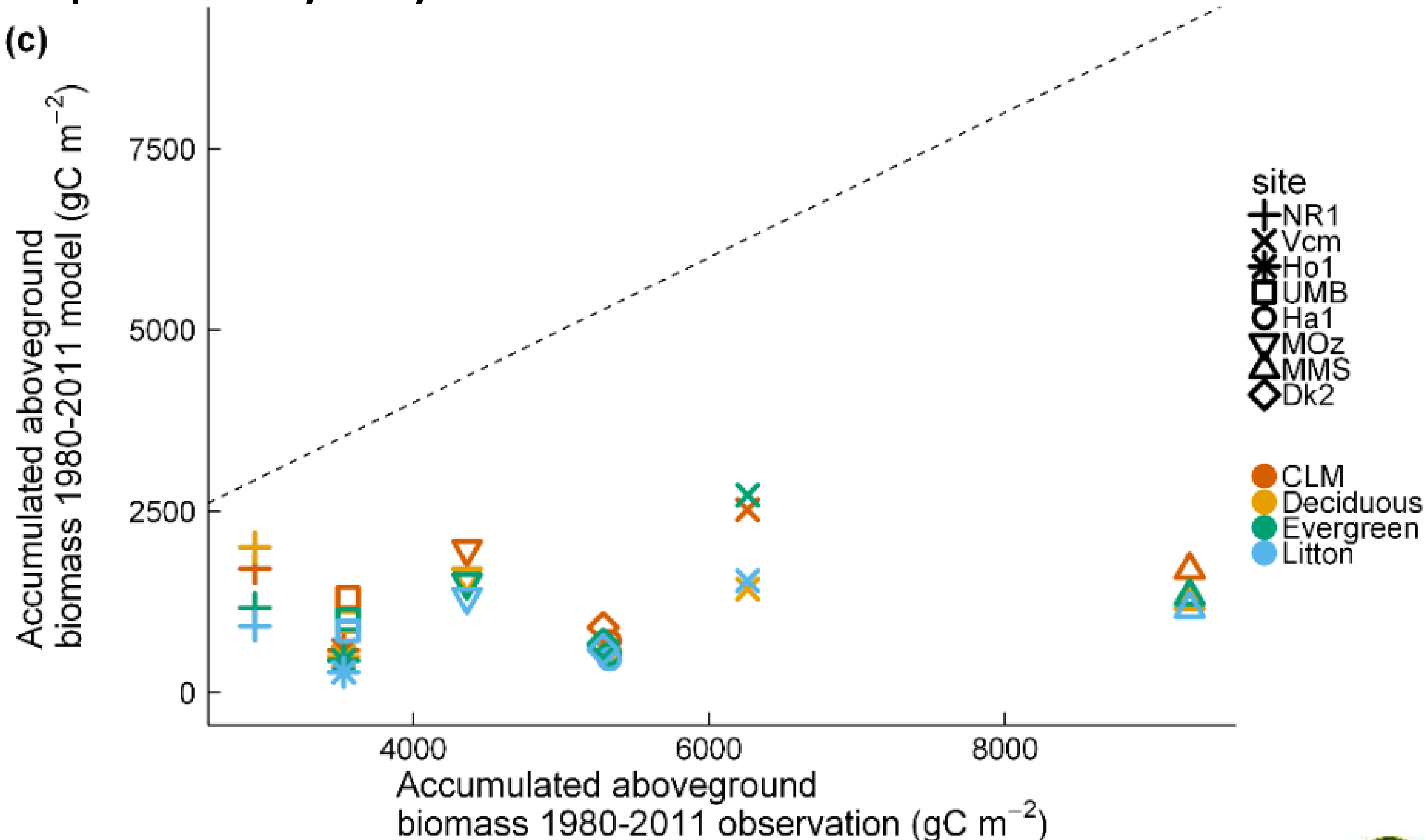
Ameriflux StemC/Leaf C indicates D-Litton scheme works well



Montane et al. *for* GMD

DOE Regional and Global Climate Modeling DE-SC0016011

30 year increase in biomass increment is NOT captured by any scheme



Montane et al. *for* GMD

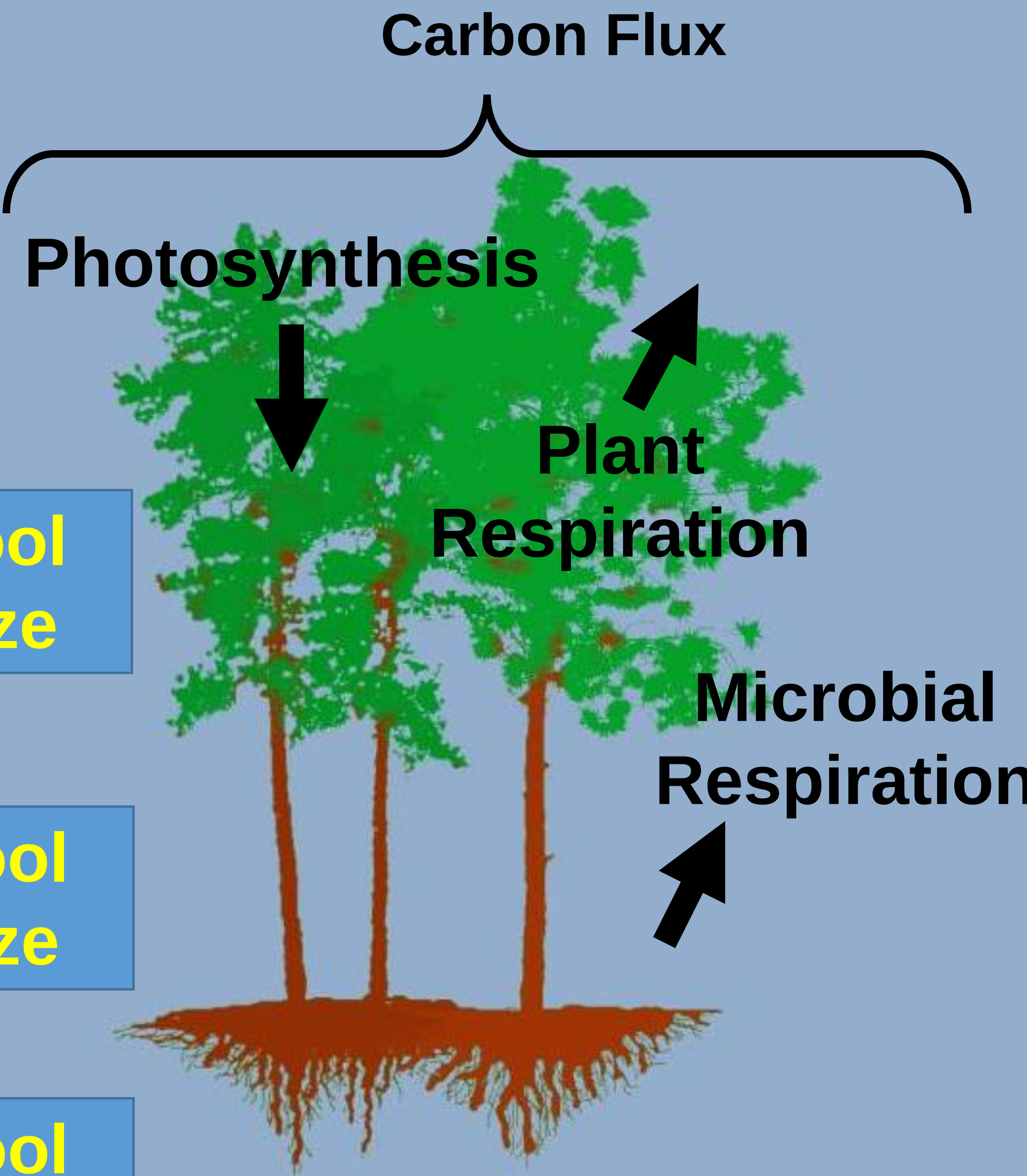
PRODUCTIVITY

TURNOVER

$$dB_i / dt = a_i NPP - u_i B_i$$

C ALLOCATION SCHEME

Strength & duration of carbon sink in forests



Carbon in live biomass — Pool size

Carbon in detritus — Pool size

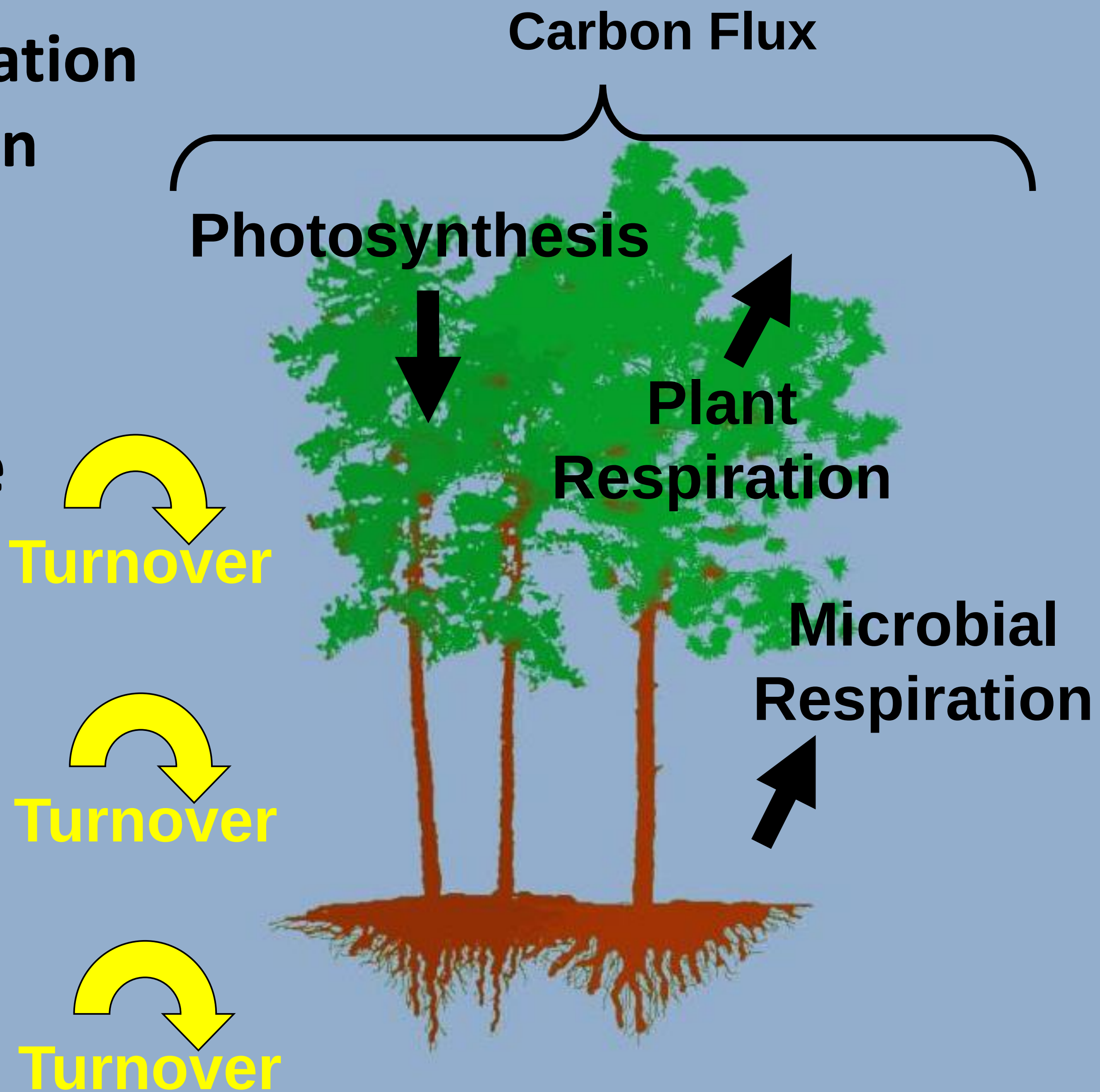
Carbon in soil — Pool size

Strength & duration of carbon sink in forests

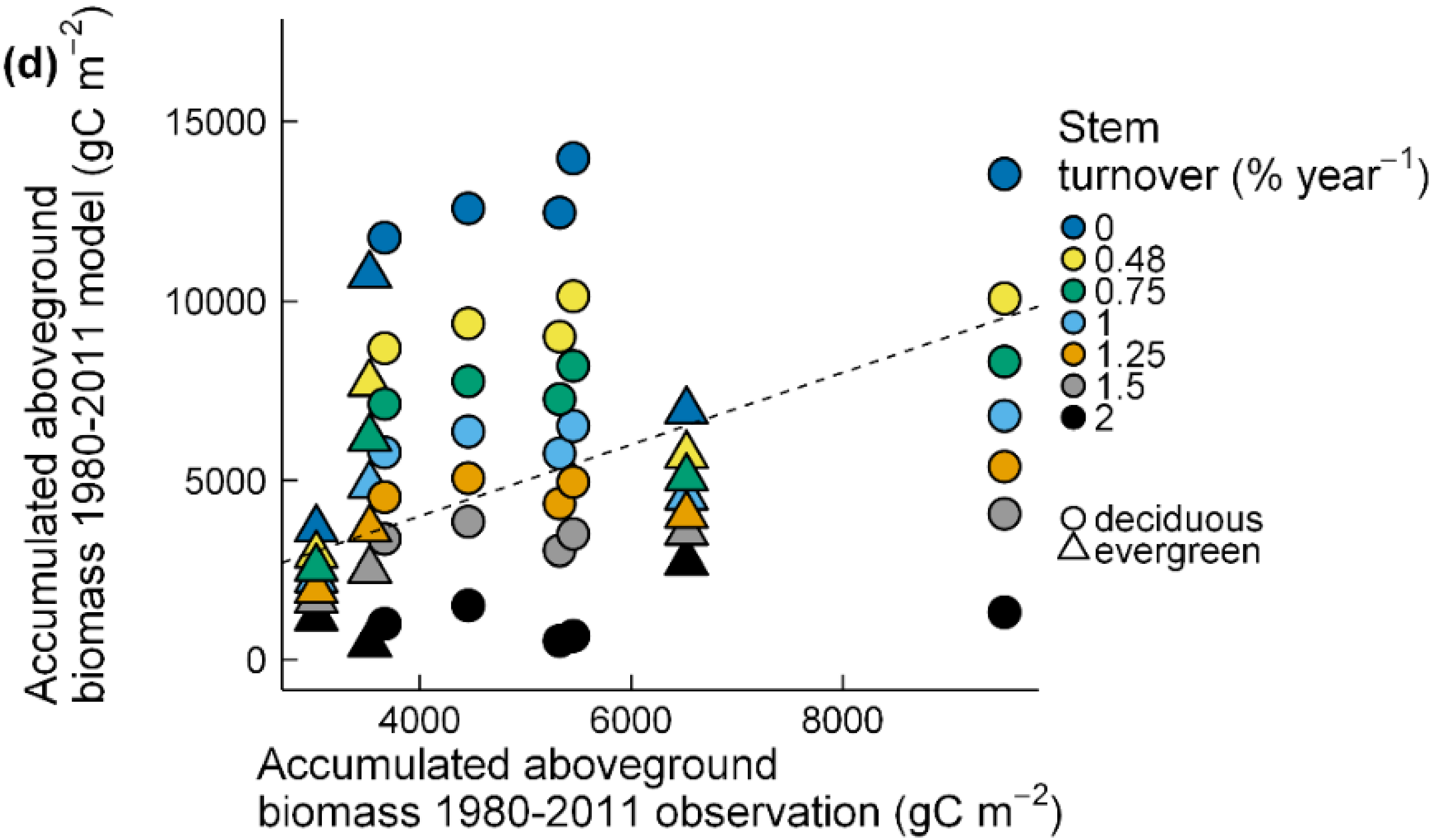
Carbon in live biomass

Carbon in detritus

Carbon in soil



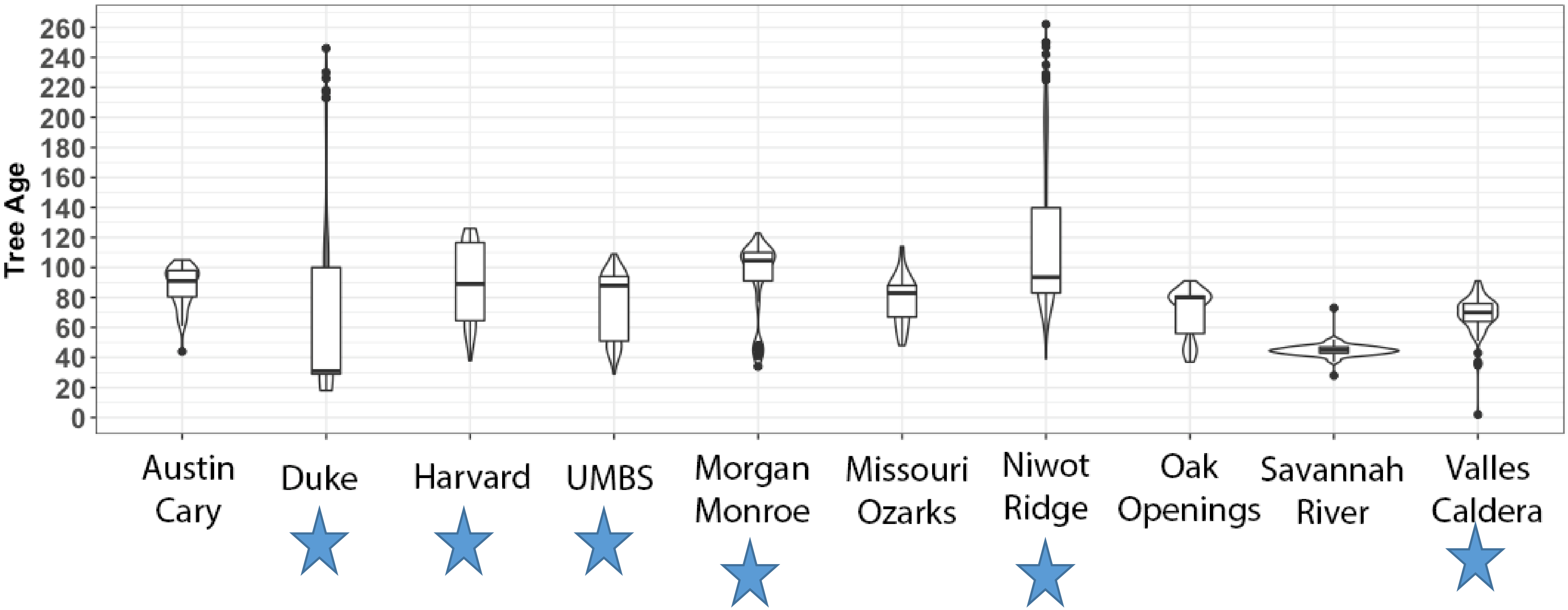
Stem turnover is poorly constrained – reasonable values for forests can account for model-data mismatch



Montane et al. *for* GMD

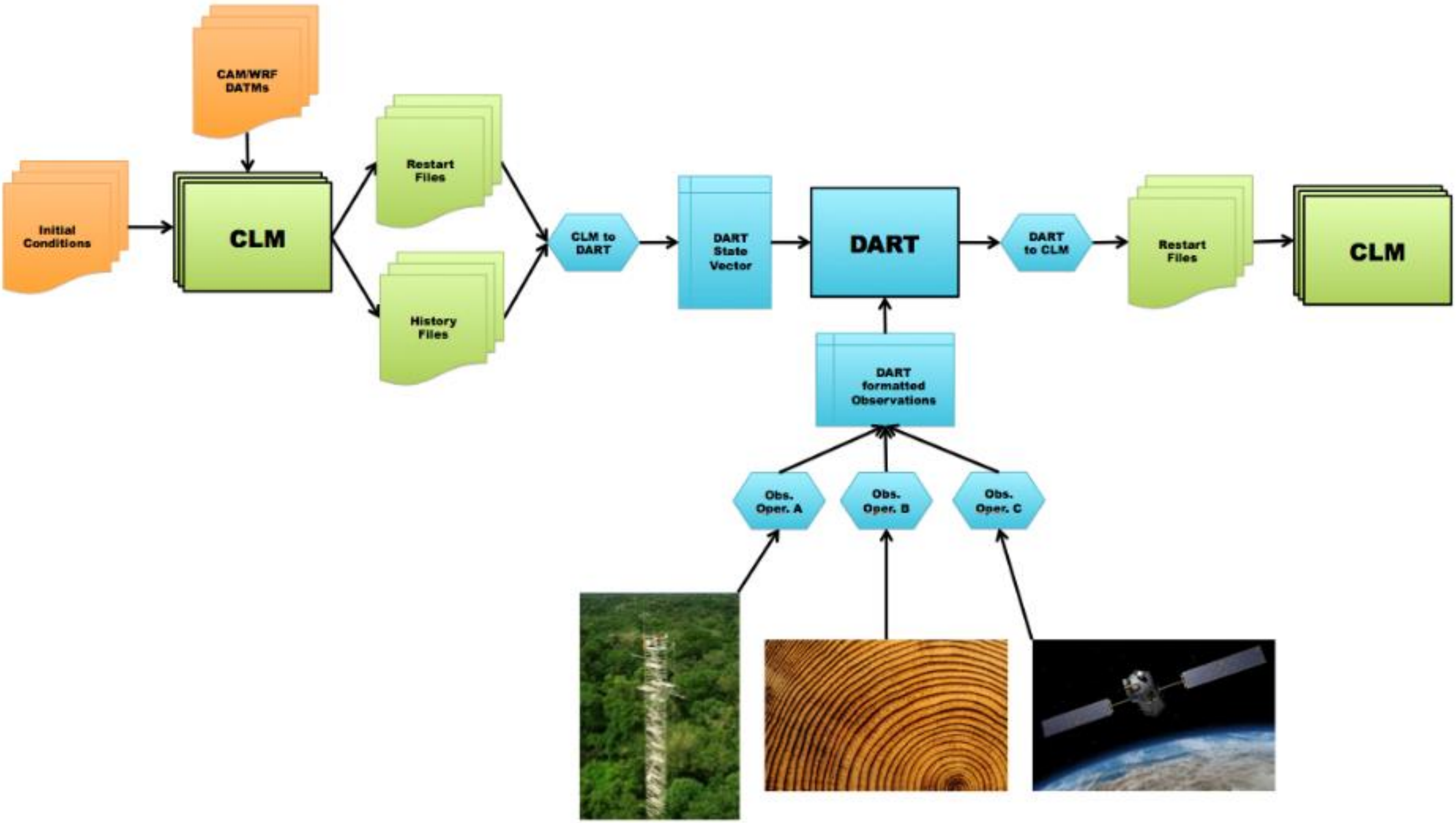
DOE Regional and Global Climate Modeling DE-SC0016011

Sites are not likely at steady state – "geographic average" of 2% is not likely appropriate

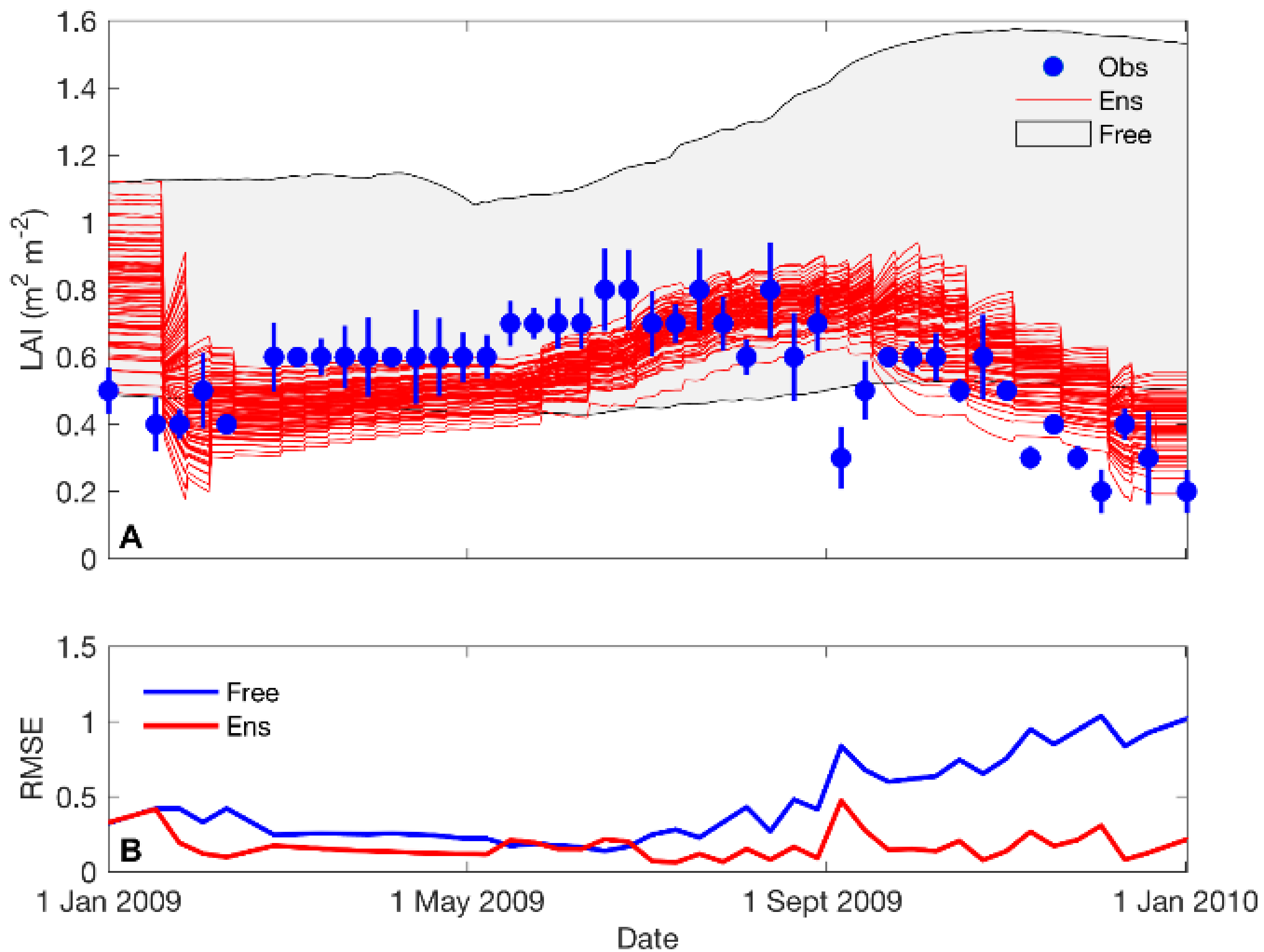


Hudson, Alexander & Moore (unpublished)

CLM-DART an Earth System Model DA system



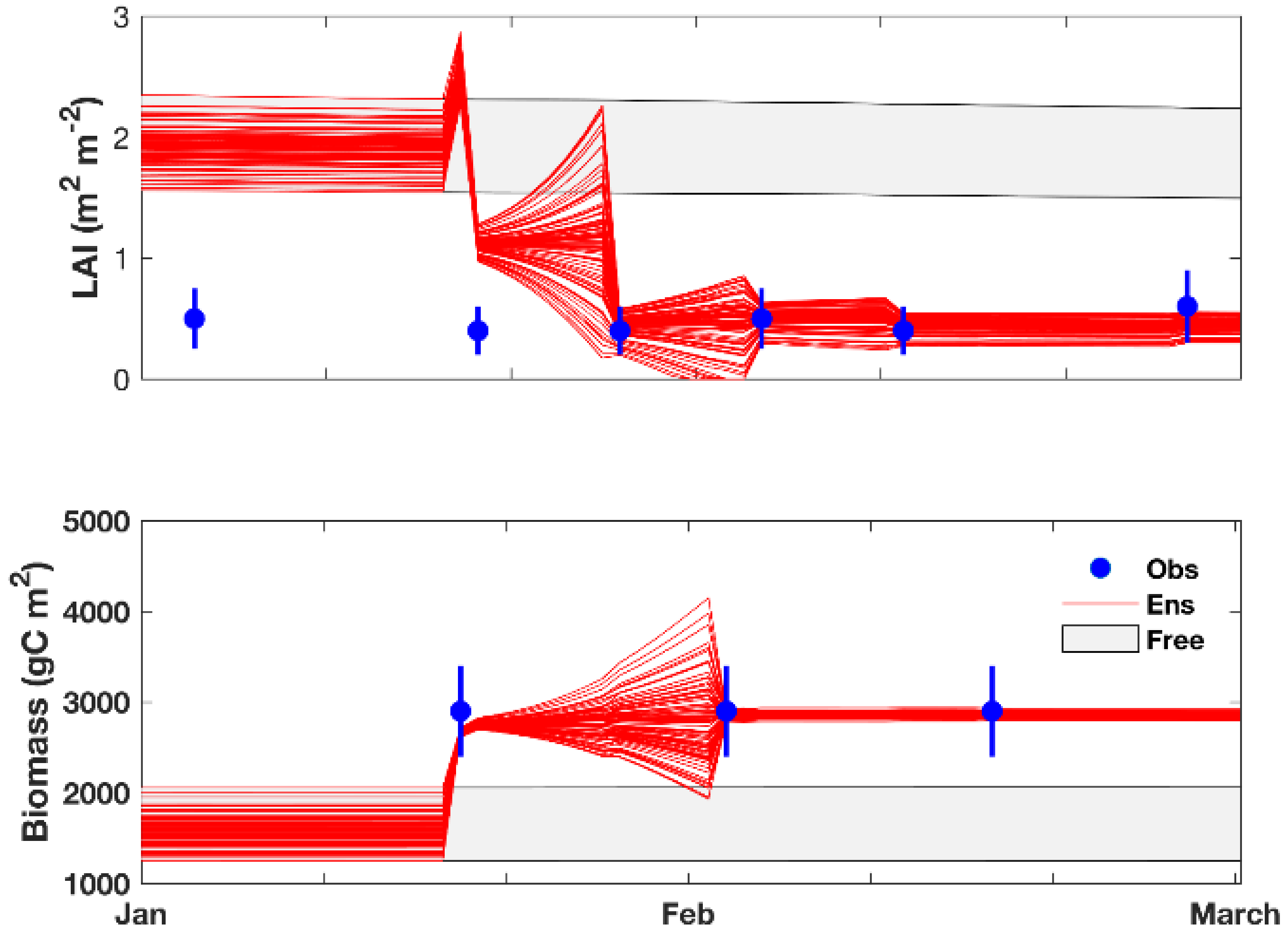
Assimilating LAI from MODIS: Significant reductions in ensemble spread



Data Assimilation Research Testbed

Assimilating LAI & Biometric Biomass

*Currently implementing new CLM model routines

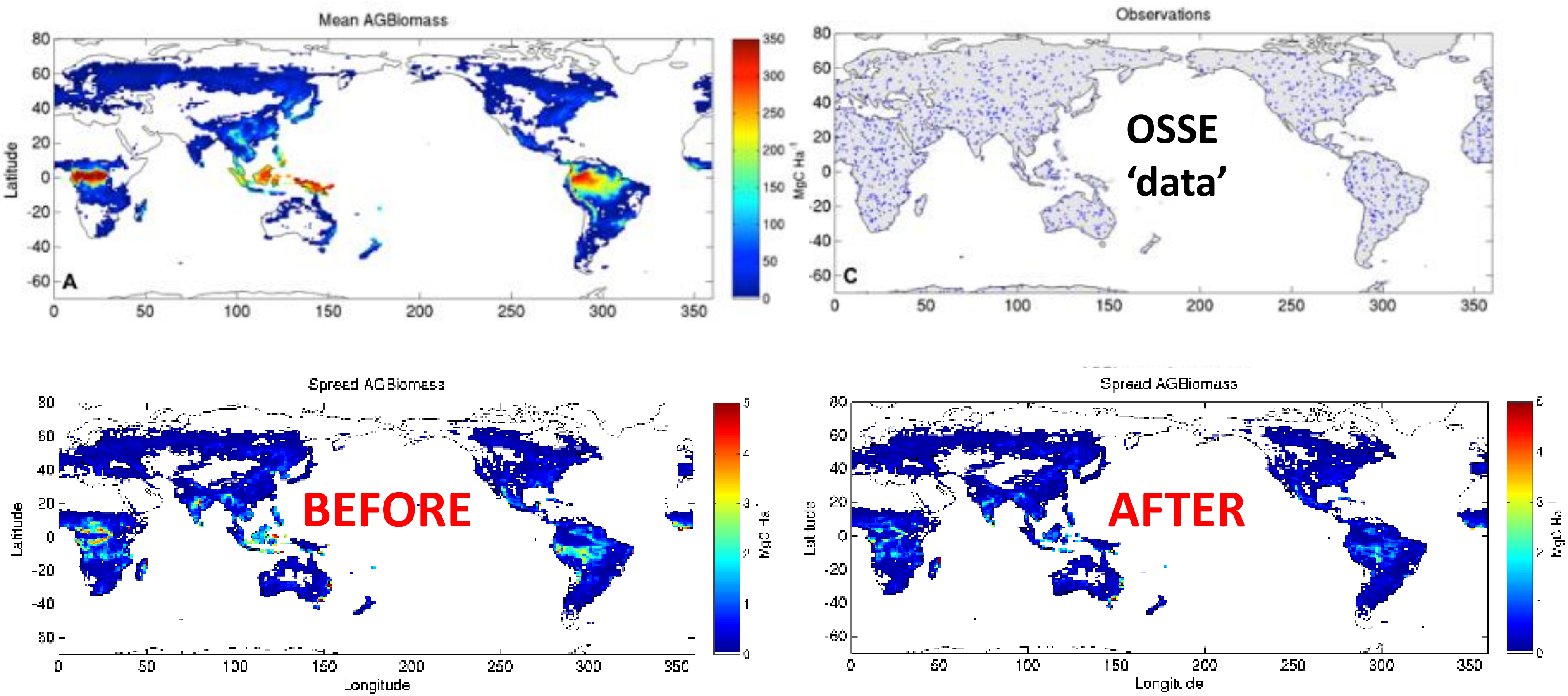


Observing System Simulation Experiments

- Site level OSSEs and real observation testing has shown biomass is a powerful constraint
- In-situ biomass observations are rare – we have 14 sites US ~the same in EU
- On-going remote sensing developments aim to measure biomass from space
- In this example, we test the ability of the CLM-DART DA system to assimilate 20,000 “pseudo-observations” globally

NCAR UNM

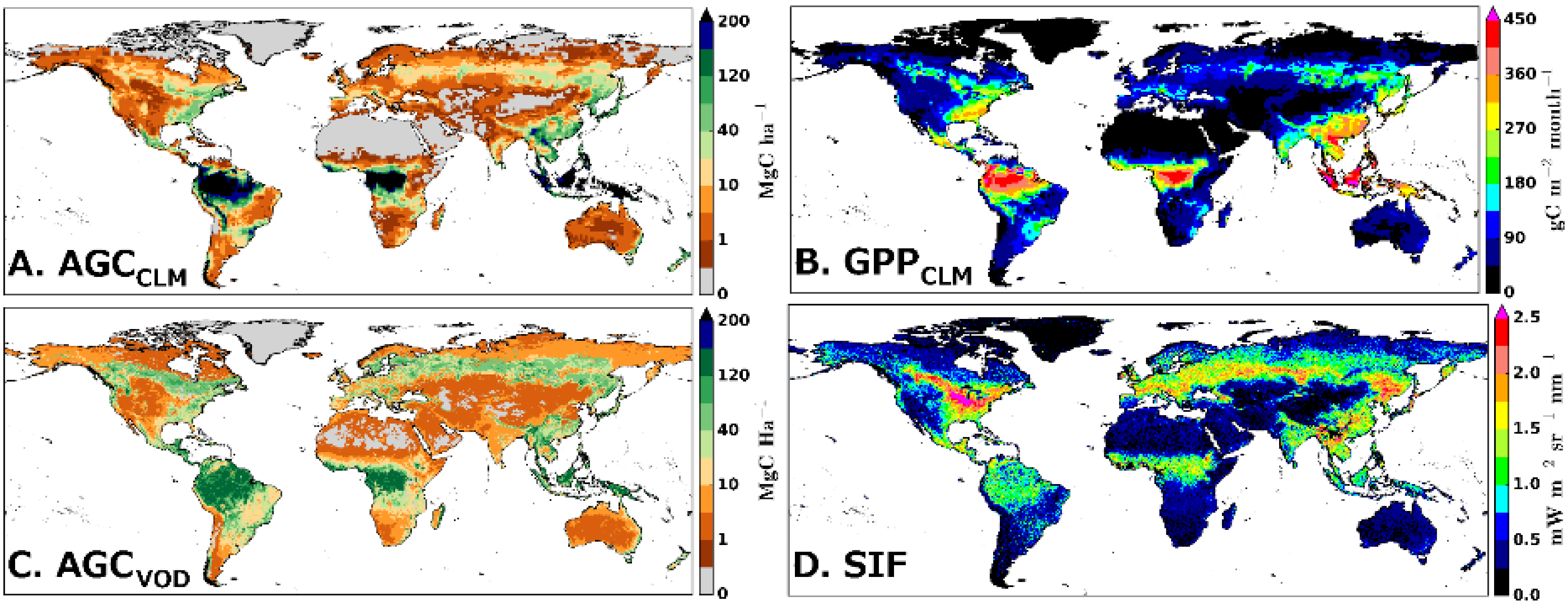
Using pseudo observations: ensemble spread of above ground biomass decreases (global data assimilation works)



Data Assimilation Research Testbed

DOE Regional and Global Climate Modeling DE-SC0016011

Plans to assimilate global datasets in addition to site data.



Courtesy Bill Kolby-Smith, UA

References

Babst, F., Alexander, M. R., Szejner, P., Bouriaud, O., Klesse, S., Roden, J., Moore, DJP & Trouet, V. (2014). A tree-ring perspective on the terrestrial carbon cycle. Oecologia, 176(2), 307-322.

Chen, M., Melaas, E. K., Gray, J. M., Friedl, M. A., & Richardson, A. D. (2016). A new seasonal-deciduous spring phenology submodel in the Community Land Model 4.5: impacts on carbon and water cycling under future climate scenarios. *Global change biology*, *22*(11), 3675-3688.

Charney, N. D., Babst, F., Poulter, B., Record, S., Trouet, V. M., Frank, D., ... & Evans, M. E. (2016). Observed forest sensitivity to climate implies large changes in 21st century North American forest growth. Ecology Letters, 19(9), 1119-1128.

Melaas, E. K., Friedl, M. A., & Richardson, A. D. (2016). Multiscale modeling of spring phenology across Deciduous Forests in the Eastern United States. Global change biology, 22(2), 792-805.

Liu, Y., Hill, M. J., Zhang, X., Wang, Z., Richardson, A. D., Hufkens, K., ... & Schaaf, C. B. (2017). Using data from Landsat, MODIS, VIIRS and PhenoCams to monitor the phenology of California oak/grass savanna and open grassland across spatial scales. Agricultural and Forest Meteorology, 237, 311-325.

Alexander, M.R., C.R. Rollinson, F. Babst, V. Trouet, D.J.P. Moore. Uncertainty in tree-ring based aboveground biomass estimates does not substantially alter growth-climate relationships. Pending Revisions: Trees: Structure and Function ***in revision***

Alexander, M.R., C.R. Rollinson, A. Dye, N. Pederson, D.J.P. Moore, V. Trouet. Differential climate responses exist among canopy strata in temperate forests of the eastern US. Submitted to Journal of Ecology ***in review***

Montane, Fox, Arellano, Alexander, Dye, MacBean, Trouet, Babst, Hessl, Pederson, Bishop, Boher, Gough, Novick, Wood, Moore (***submitted***) Evaluating the effect of alternative carbon allocation schemes in a land surface model on carbon fluxes, pools and turnover in temperate forests Geophysical Model Development.

DOE Regional and Global Climate Modeling DE-SC0016011

Rollinson, C. R., Liu, Y., Raiho, A., Moore, D. J., McLachlan, J., Bishop, D. A., ... & Pederson, N. (2017). Emergent climate and CO2 sensitivities of net primary productivity in ecosystem models do not agree with empirical data in temperate forests of eastern North America. Global Change Biology.

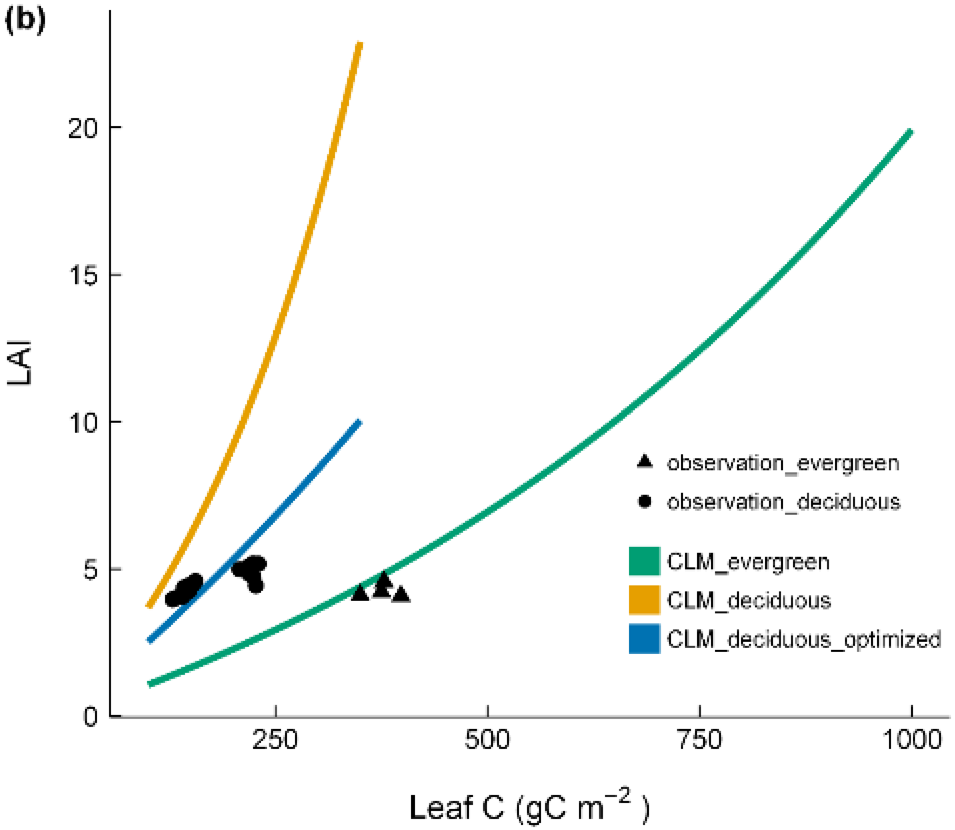
Dye, A., Barker Plotkin, A., Bishop, D., Pederson, N., Poulter, B., & Hessl, A. (2016). Comparing tree-ring and permanent plot estimates of aboveground net primary production in three eastern US forests. Ecosphere, 7(9).

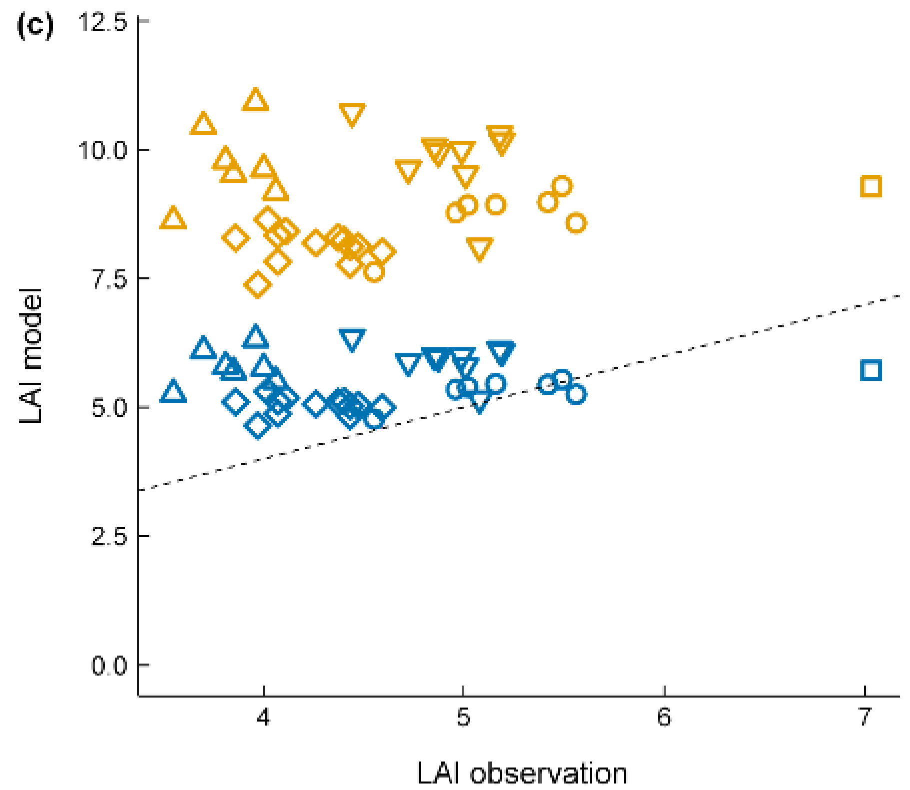
NSF Macrosystems Biology 1241851

Many big questions remain

- How to create initial ensemble spread – how large should it be?
- How to maintain ensemble spread – is climate forcing variability the best approach?
- What do we do about carbon/water balance – its lost at the moment and balance checks are removed?
- What are the most informative observations to use?
- What are the best temporal aggregation strategies for EC flux tower data?
- Can we develop appropriate observation operators to link them with CLM state?
- How can we best use an ensemble DA approach for parameter estimation – we can augment DART state vector with CLM parameters, but which ones?

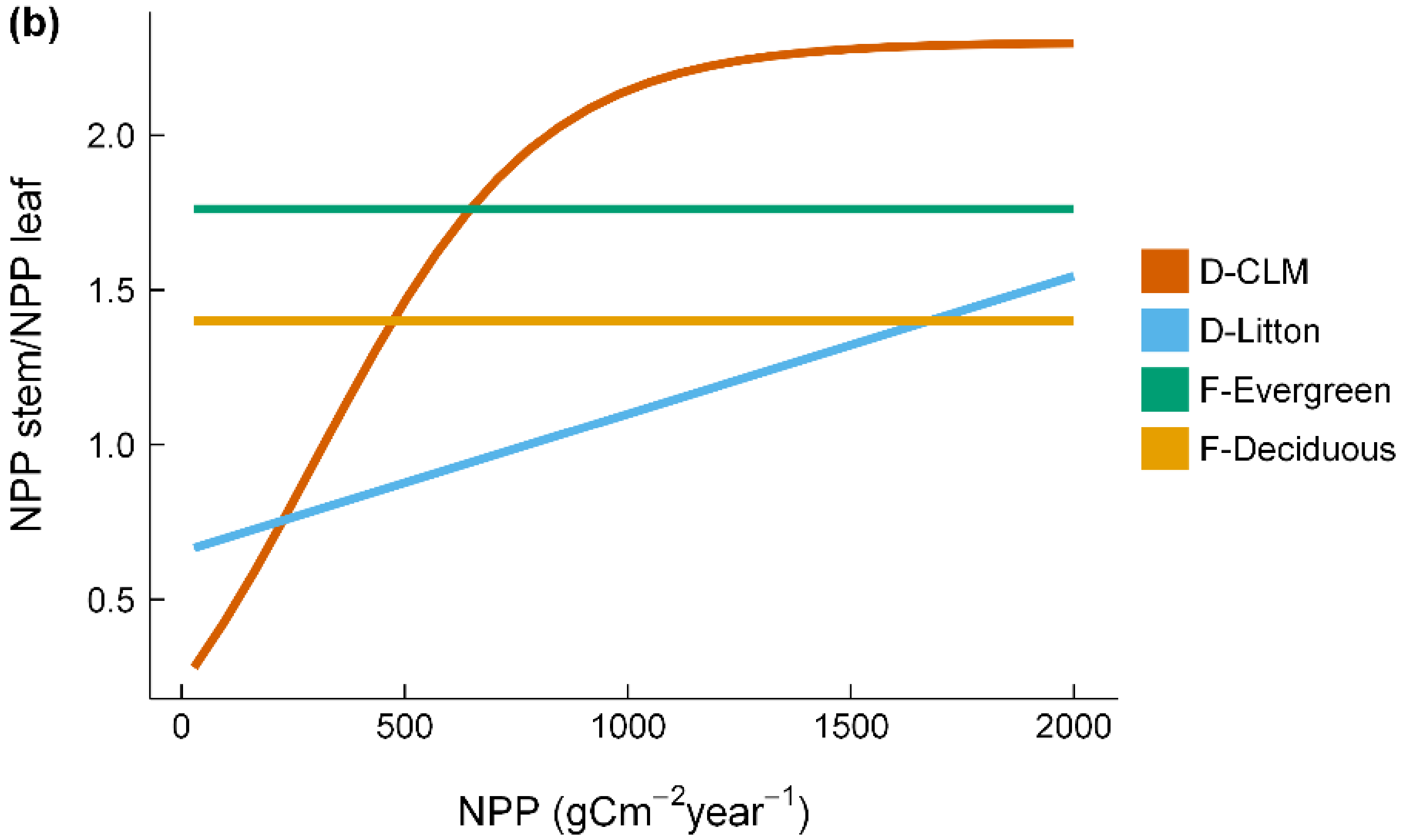
Model Development – Leaf area to carbon ratio incorrectly specified in CLM



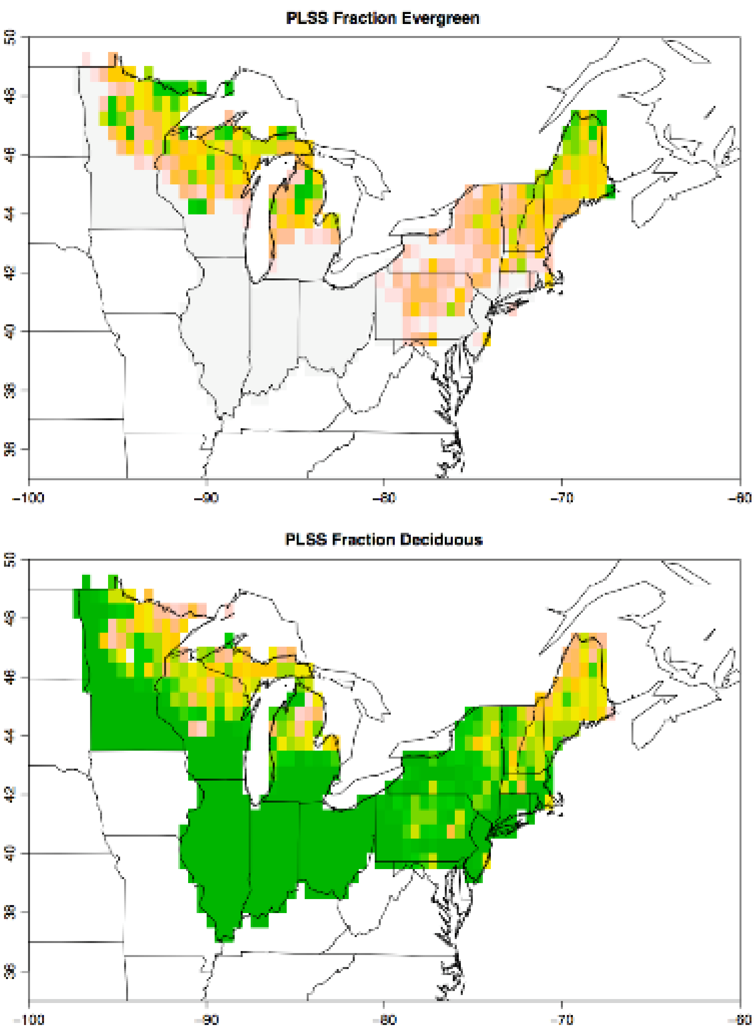


Montane et al. *for* GMD

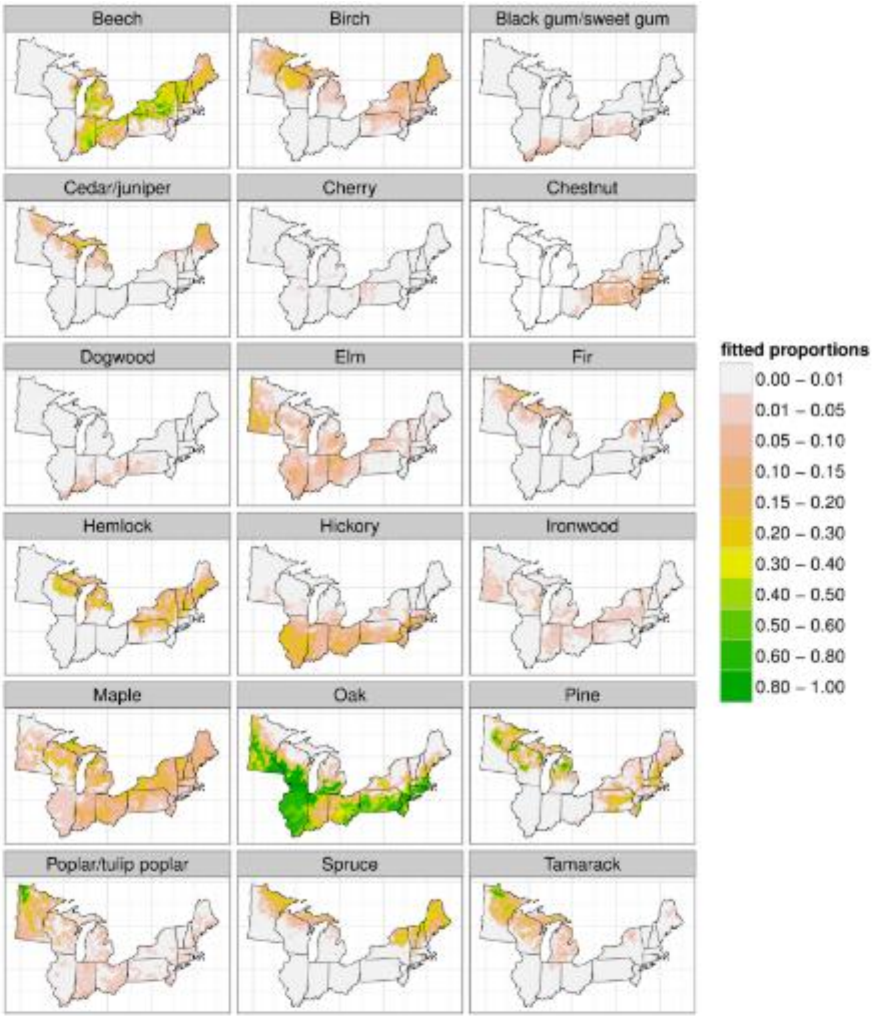
DOE Regional and Global Climate Modeling DE-SC0016011


(b)
NPP stem/NPP leaf
2.0
1.5
1.0
0.5
0
500
1000
1500
2000
NPP (gCm−2year−1)
D-CLM
D-Litton
F-Evergreen
F-Deciduous

Paleon Project "Settlement-Era" vegetation

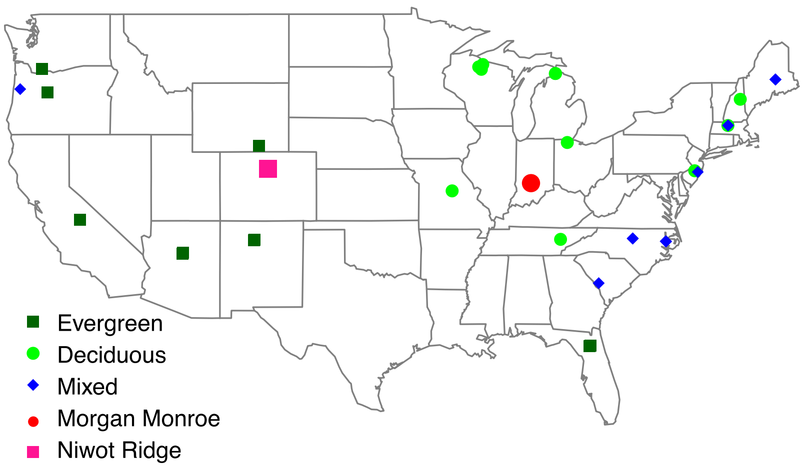


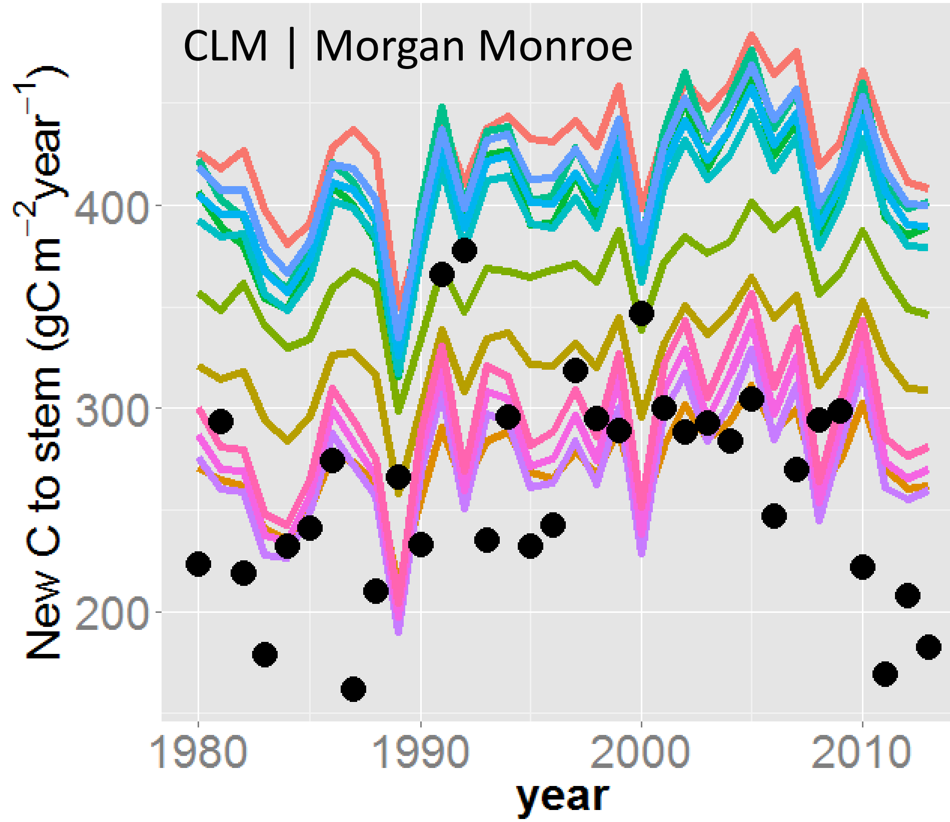
Matthes et al. 2016 JGR-Biogeosciences



Paciorek et al. 2016 PLoS-ONE

Building a data base of aboveground NPP based on tree rings





Tree ring 'biomass' record

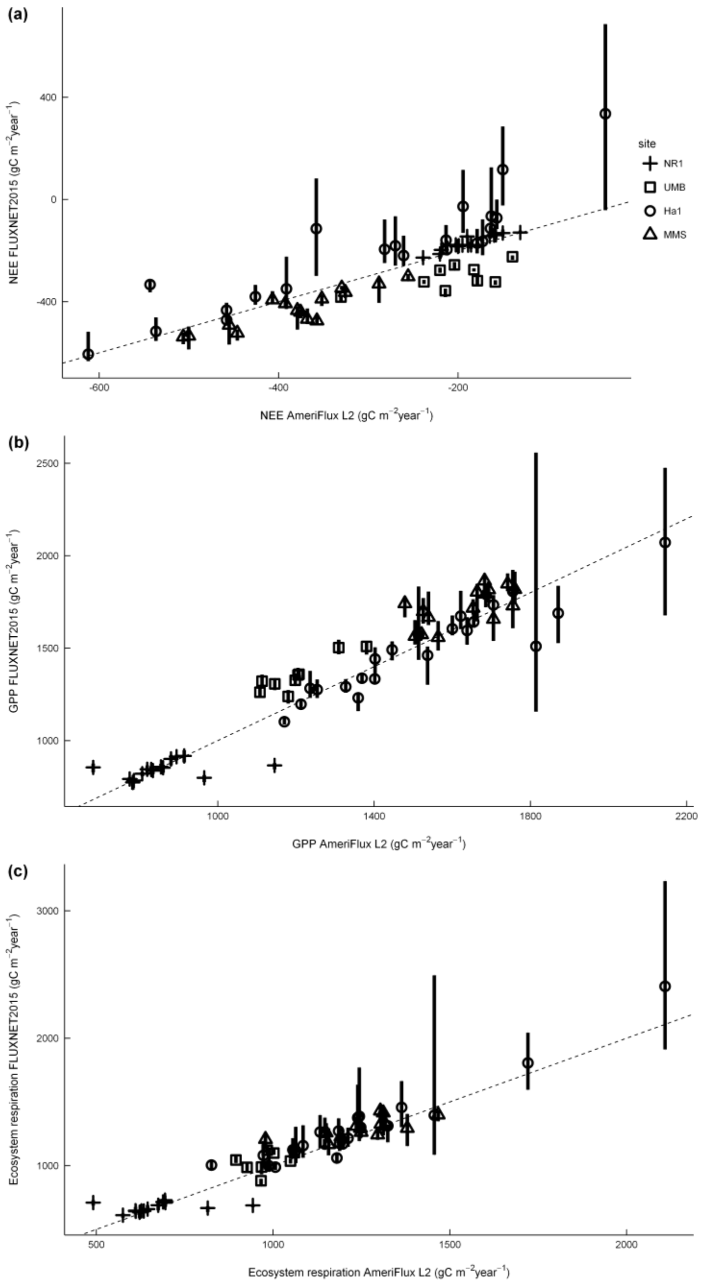
Flux

Colors – CLM (different allocation schemes)

● Reconstructed Biomass

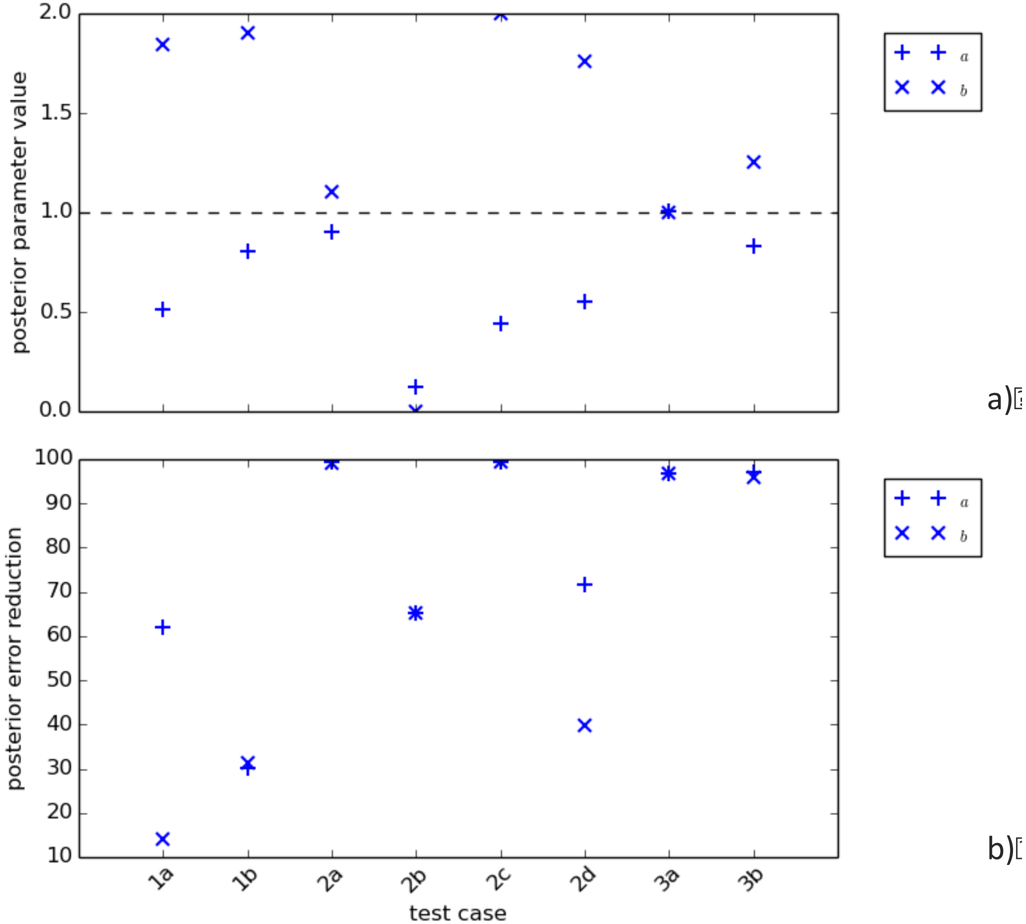
Francesc Montane & Ross Alexander

DOE Regional and Global Climate Modeling DE-SC0016011



(a)
NEE FLUXNET2015 (gC m^{-2}year^{-1})
NEE AmeriFlux L2 (gC m^{-2}year^{-1})
site
NR1
UMB
Ha1
MMS
(b)
GPP FLUXNET2015 (gC m^{-2}year^{-1})
GPP AmeriFlux L2 (gC m^{-2}year^{-1})
(c)
Ecosystem respiration FLUXNET2015 (gC m^{-2}year^{-1})
Ecosystem respiration AmeriFlux L2 (gC m^{-2}year^{-1})

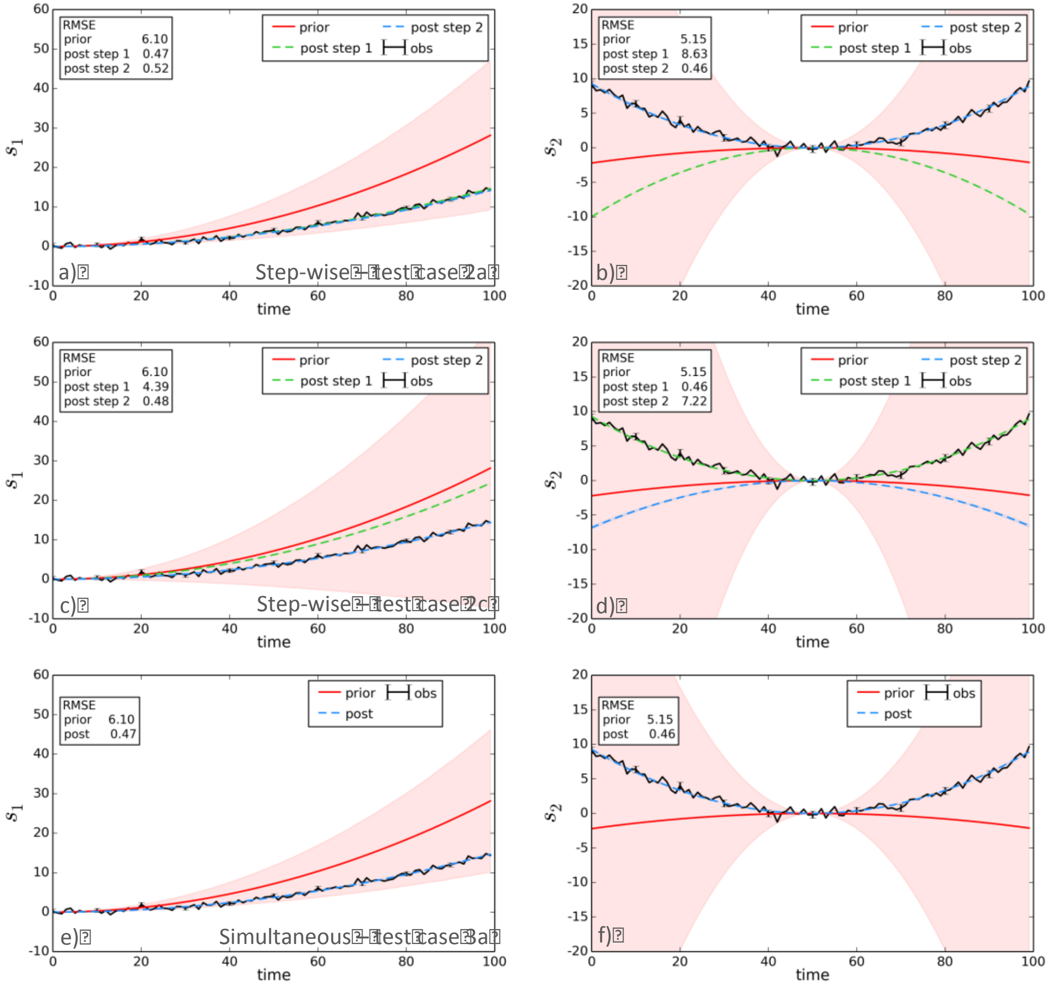
MacBean, N., P. Peylin, F. Chevallier, M. Scholze, M., and G. Schürmann (2016) Consistent assimilation of multiple data streams in a carbon cycle data assimilation system, *Geoscientific Model Development*, 9, 3569-3588, doi:10.5194/gmd-9-3569-2016.



Test case	Step 1	Step 2	Parameter error covariance terms propagated in step 2?
Separate			
1a	s_1	–	–
1b	s_2	–	–
Step-wise			
2a	s_1	s_2	yes
2b	s_1	s_2	no
2c	s_2	s_1	yes
2d	s_2	s_1	no
Simultaneous			
3a	s_1 and s_2	–	–
3b	s_1 and only 1 obs for s_2	–	–

MacBean, N., P. Peylin, F. Chevallier, M. Scholze, M., and G. Schürmann (2016) Consistent assimilation of multiple data streams in a carbon cycle data assimilation system, *Geoscientific Model Development*, 9, 3569-3588, doi:10.5194/gmd-9-3569-2016.

Order of data stream assimilation



s_1 then s_2

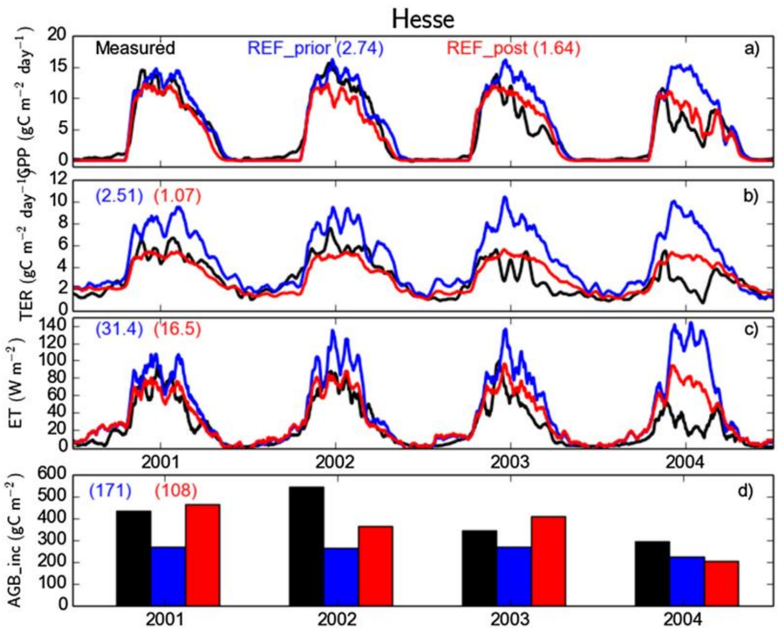
s_2 then s_1

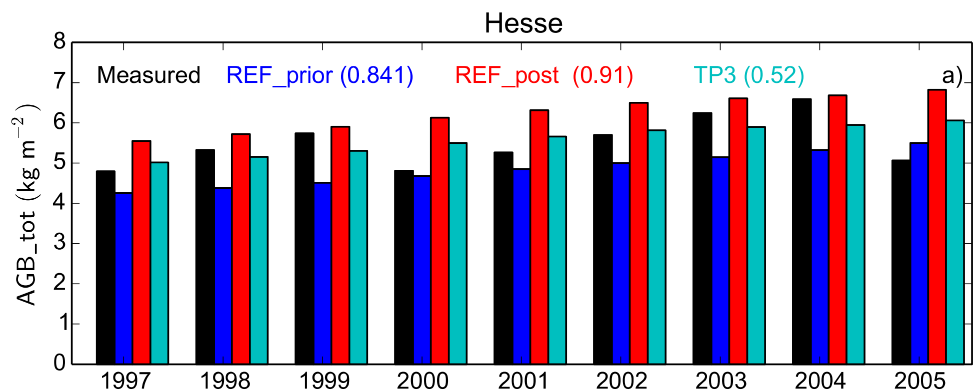
s_1 and s_2

Test case	Step 1	Step 2	Parameter error covariance terms propagated in step 2?
Separate			
1a	s_1	–	–
1b	s_2	–	–
Step-wise			
2a	s_1	s_2	yes
2b	s_1	s_2	no
2c	s_2	s_1	yes
2d	s_2	s_1	no
Simultaneous			
3a	s_1 and s_2	–	–
3b	s_1 and only 1 obs for s_2	–	–

Thum, T., **N. MacBean**, P. Peylin, C. Bacour, D. Santaren, B. Longdoz, D. Loustau and P. Ciais (2017) The potential benefit of using forest biomass data in addition to carbon and water flux measurements to constrain ecosystem model parameters: case studies at two temperate forest sites, *Agricultural and Forest Meteorology*, 234, 48-65.

→ Issues with using aboveground biomass increment vs aboveground biomass





AGB increment always positive…

… leads to worse fit to total AGB …

… can improve after optimization with total AGB …

BUT residence time too low (40 → ~17 years)

Not accounting for disturbance and human activity