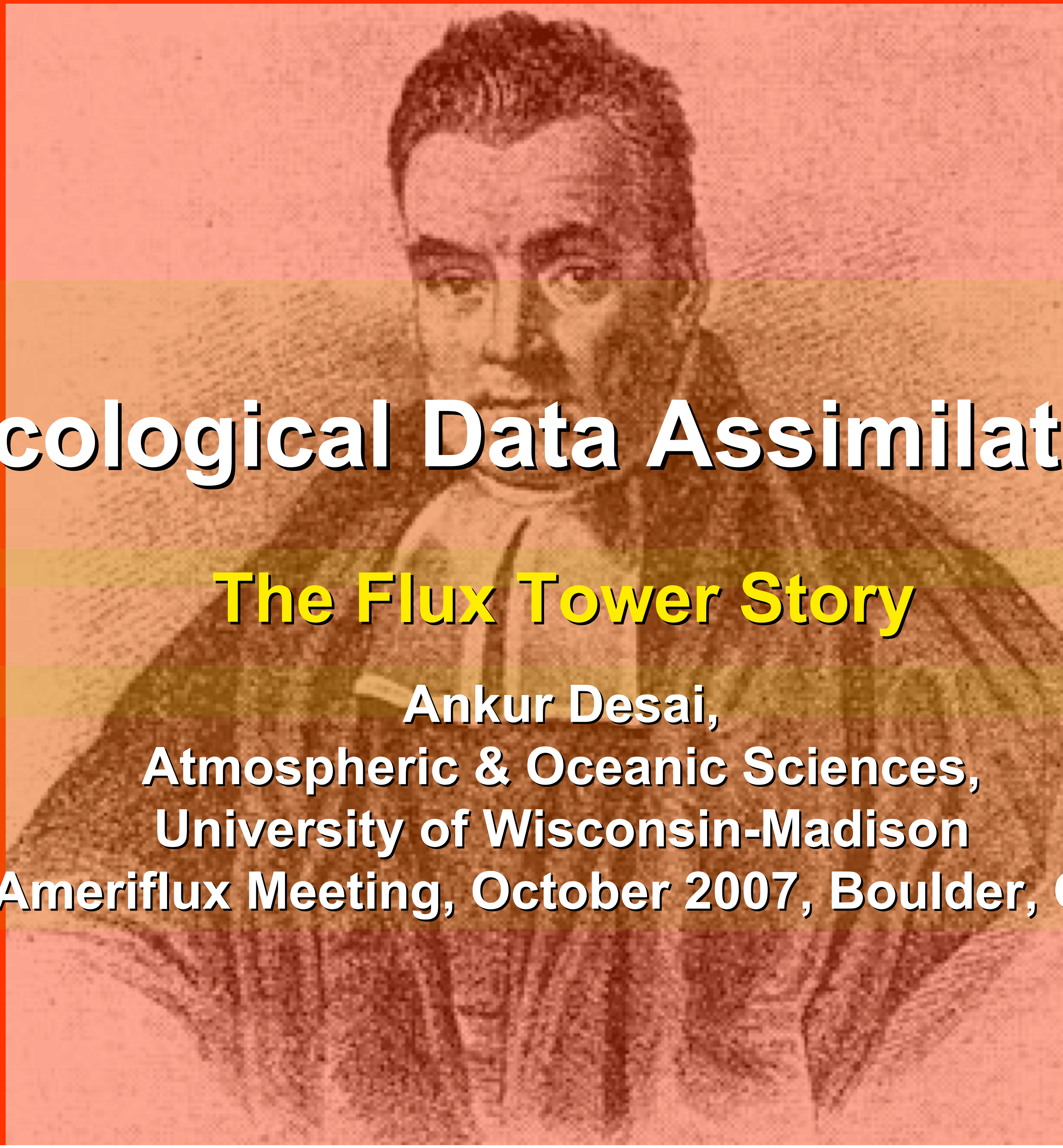


Ecological Data Assimilation

The Flux Tower Story

Ankur Desai,
Atmospheric & Oceanic Sciences,
University of Wisconsin-Madison

Ameriflux Meeting, October 2007, Boulder, CO

The Scene

- An ecosystem or land-atmosphere model
  - With parameters, drivers, fluxes, stocks
  - Probably non-linear, might be chaotic
  - Many parameters are not well known (e.g., Q10)
- Driver data (filled?)
- Some observation you want to reproduce (e.g., CO_2 flux) - data has noise/uncertainty
- How to minimize model-data difference taking noise into account and estimate “true” parameters and their uncertainty?
  - Partly depends on the questions you were exploring with model

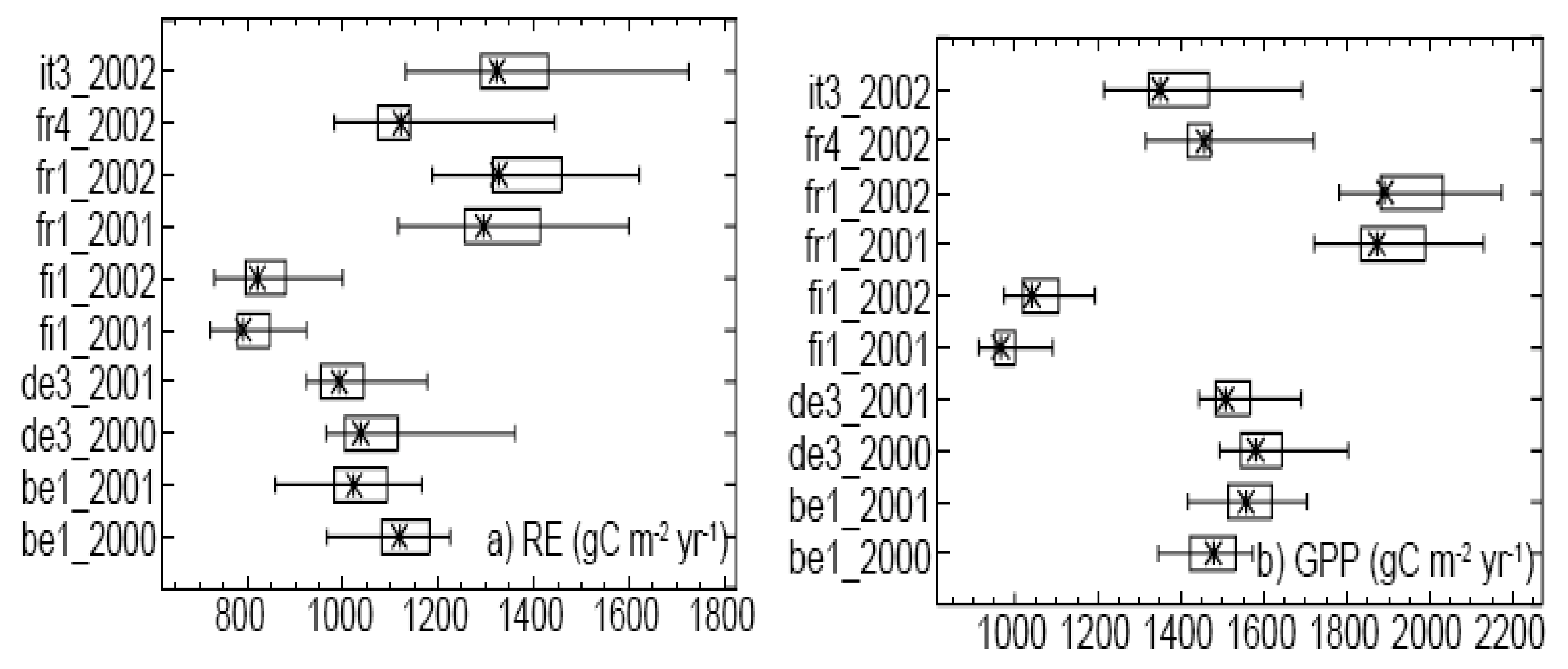
Why are we doing this?

- Prediction / State space exploration
- Spatial scaling (model calibration)
- Parameter estimation / comparison
- Estimate unobserved state variables (GPP)
- Mechanism testing / Model selection
- Observation set consistency / value
- Hypothesis testing

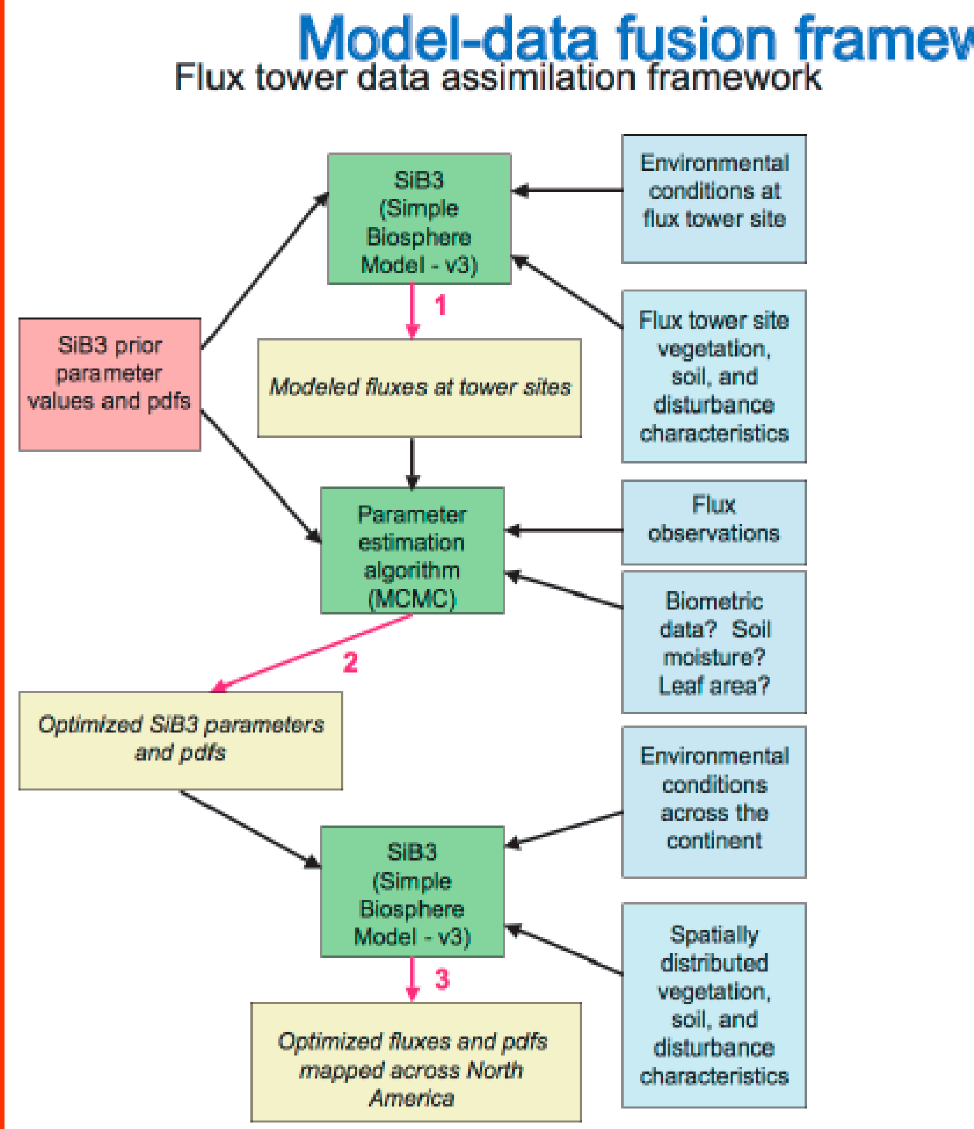
- Flux towers are well suited to helping models do a better job at all of these…

Desai et al., accepted, AgForMet (GPP/RE)
Moffatt et al., in press, AgForMet (Gaps)

- Flux tower gap filling and GPP/RE retrieval are kinds data assimilation



Courtesy T. Hilton and K. Davis



Questions within this framework:

1) How should flux towers be grouped when solving for SiB3 parameters?

2) How should the resulting optimized parameters be mapped across space?

3) How should this framework be linked to the atmospheric framework. Can we solve for model parameters at that stage?

Some solutions (simple data assimilation)

- Manual (guess parameters, run, compare, try again)
  - An army of students helps
- Least squares linear fits
- Maximum likelihood
- Steepest descent and gradient optimizers (e.g., Levenburg-Marquardt, Gauss-Newton)

- Better solution: let’s ask Mathematicians and Meteorologists instead…

Bayes' Theorem to the rescue

[A|B] = [AB] / [B]

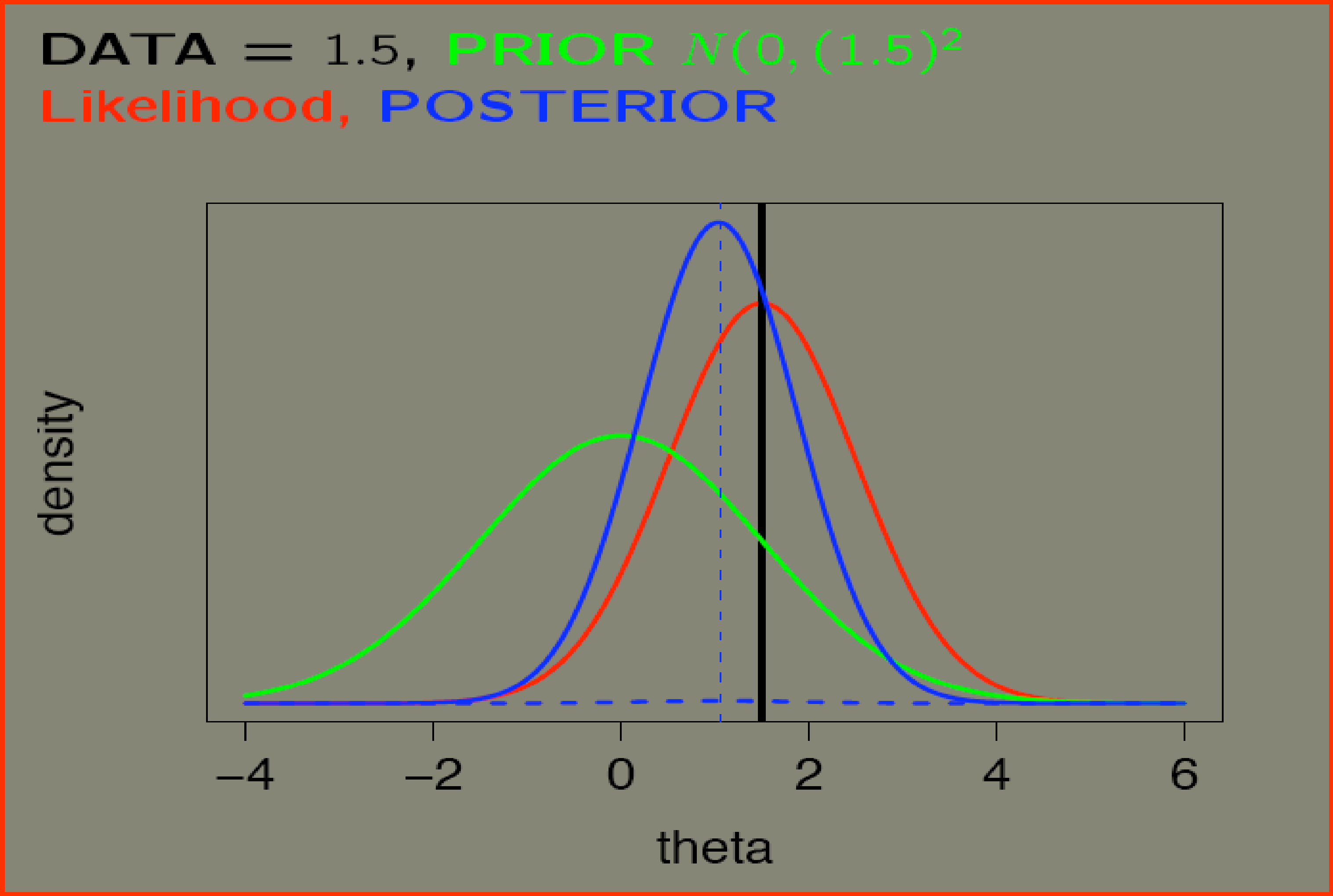
[P|D] = ([D|P] [P]) / [D]

(parameters given data) = [(data given parameters)× (parameters)] / (data)

Posterior = (Likelihood x Prior) / Normalizing Constraint

In the long run, this is least-squares and Gaussian in the basic setup (can be modified). Main things needed for implementation are Forward operator and Likelihood function

Courtesy D. Nychka, NCAR



Leading to advanced data assimilation

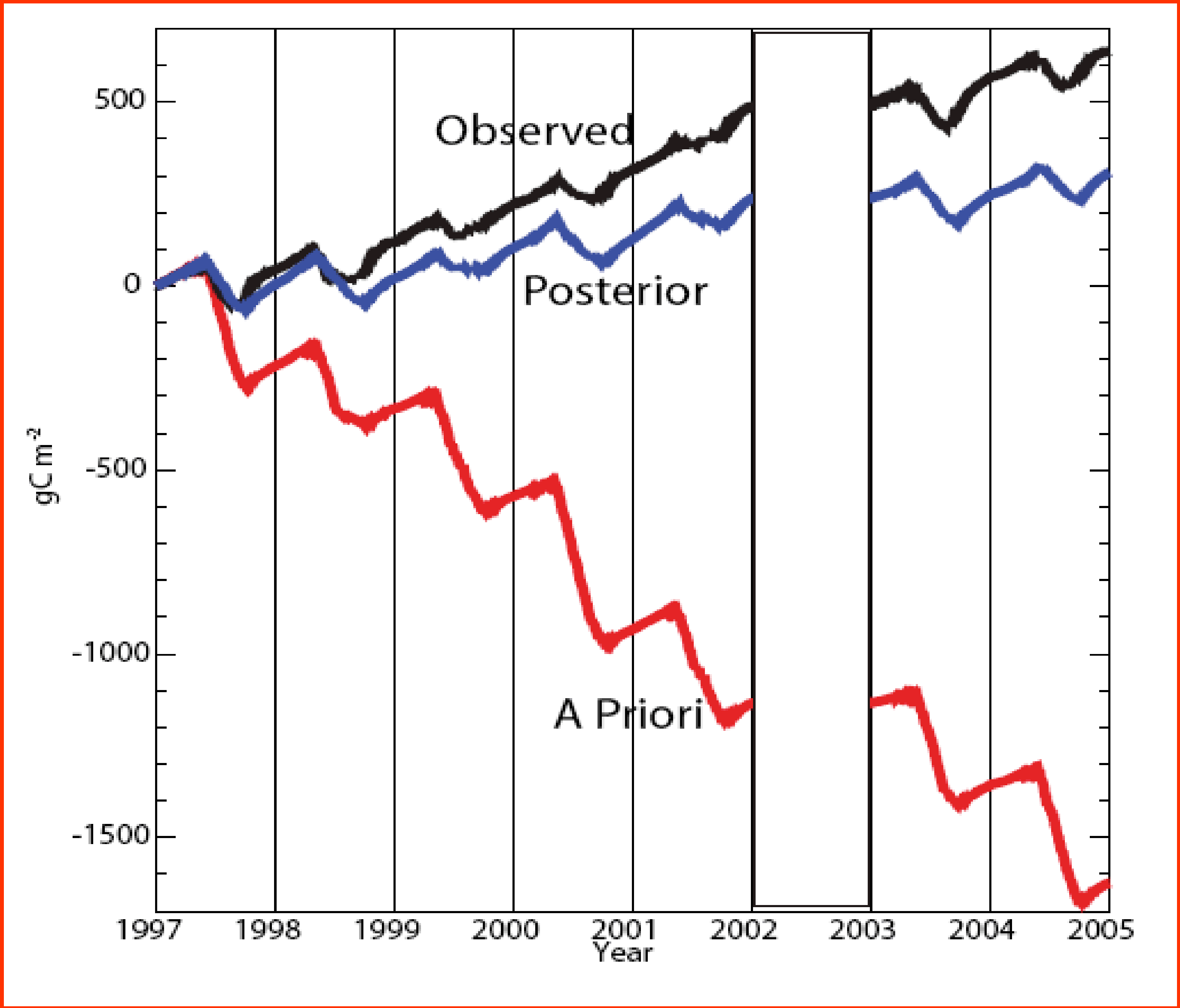
- Direct parameter distribution exploration
- Markov Chain Monte Carlo (MCMC) - Metropolis-Hastings Algorithm (Metropolis et al, 1953) and other stochastic techniques
- (Ensemble) Kalman Filters and Smoothers
  - Good for expensive models, multiple datasets
- Genetic Algorithms - e.g., Stochastic Evolutionary Ranking Strategy (SRES)
- Neural networks
- Variational methods* (need to know adjoint)

- Tests with REFLEX, …

Lots of activity

- Big focus on MCMC
  - Big problem with MCMC is need for many iterations to sample parameter space
    - Recall [D|P]
  - But others methods are gaining
- Some models with data assimilation routines developed or in development include SipNET, TREES,ORCHIDEE, BETHY, TRIFFID, ED, Biome-BGC, LoTEC, SiB3

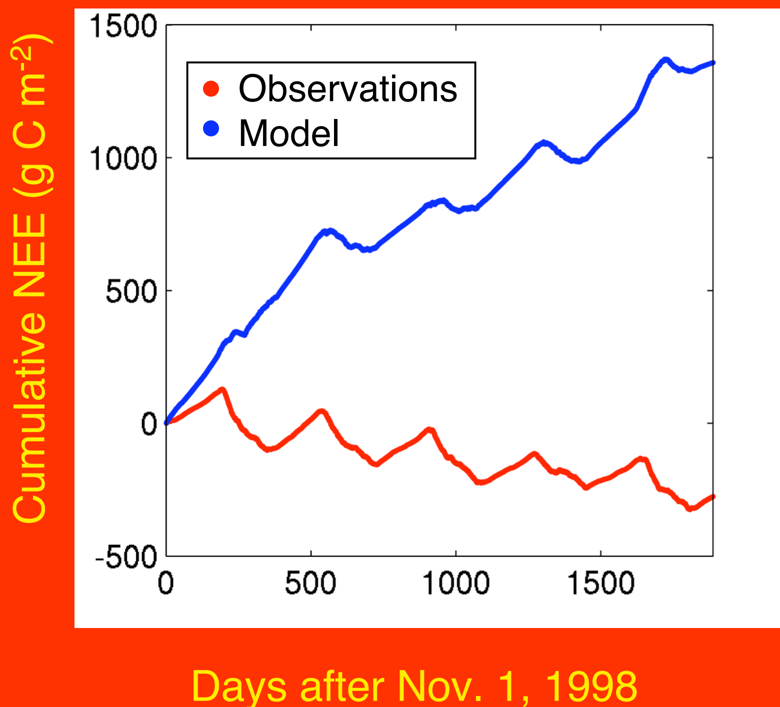
WLEF - Desai et al., in prep

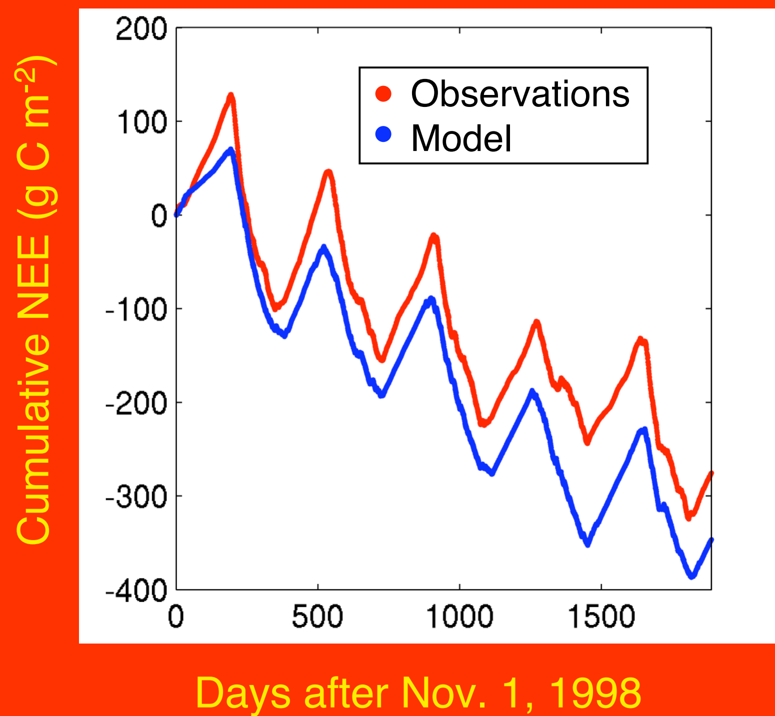


Desai et al, in prep

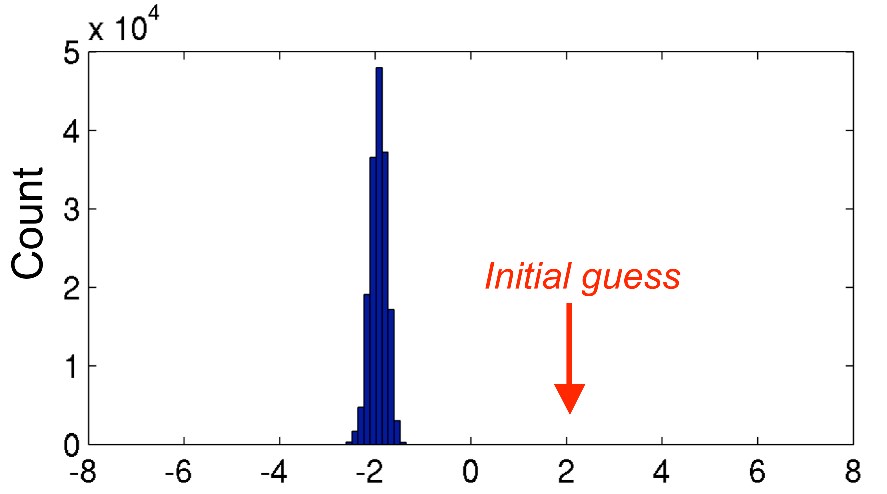
	Prior	Posterior
Growth related parameters		
photosynthetic capacity (amax)	112	58.6 +/- 2.2
growth respiration fraction	0.33	0.34 +/- 0.06
VPD modifier slope	0.05	0.066 +/- 0.009
Half saturation PAR	17	9.0 +/- 0.76
Light attenuation	0.5	0.67 +/- 0.02
WUE factor	10.9	13.4 +/- 0.46*
Decomposition parameters		
Lloyd-Taylor E0	309	448 +/- 121
Lloyd-Taylor T0	-46	-59.5 +/- 10.6
Turnover rate	0.03	0.19 +/- 0.02

Niwot Ridge - Sacks et al., 2006, GCB

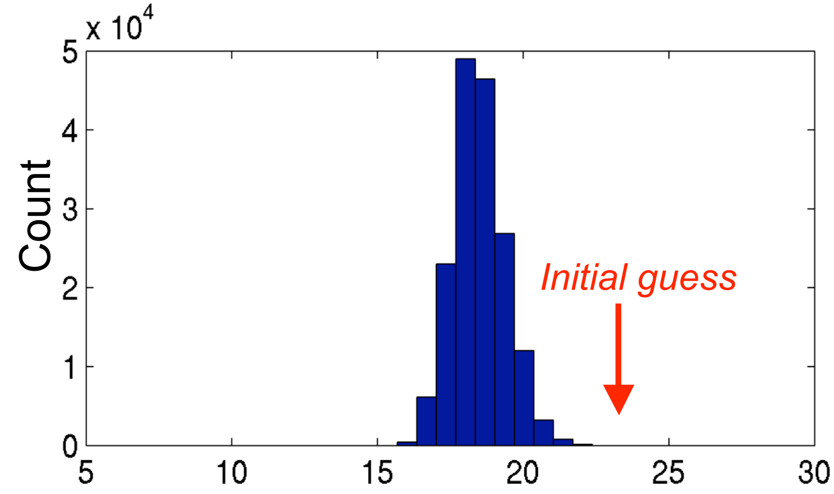




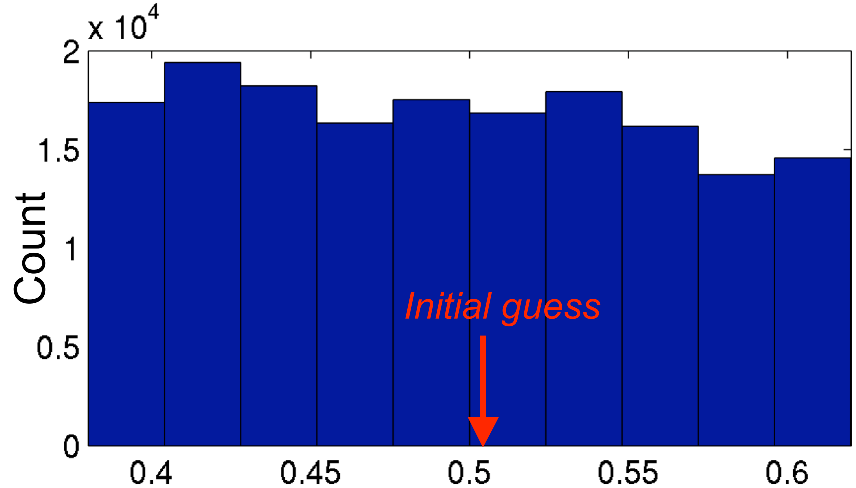
Sacks et al., 2006, GCB



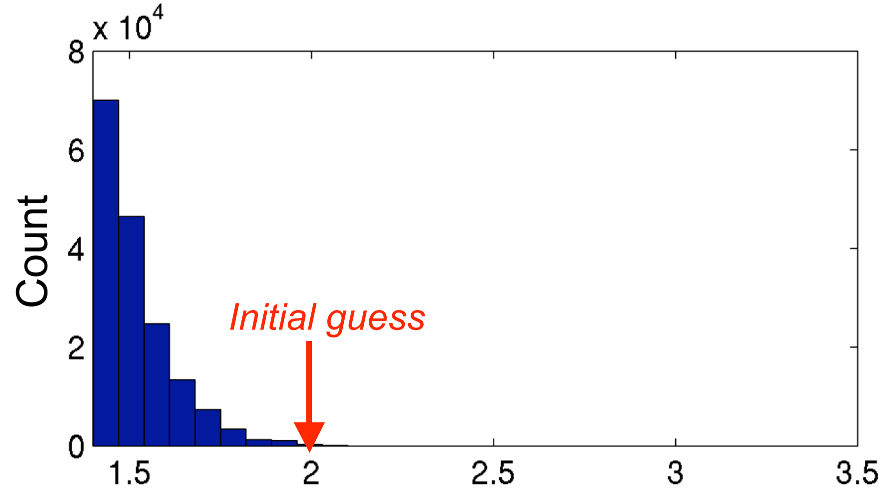
Min. temp. for photosynthesis



Optimum temp. for photosynthesis



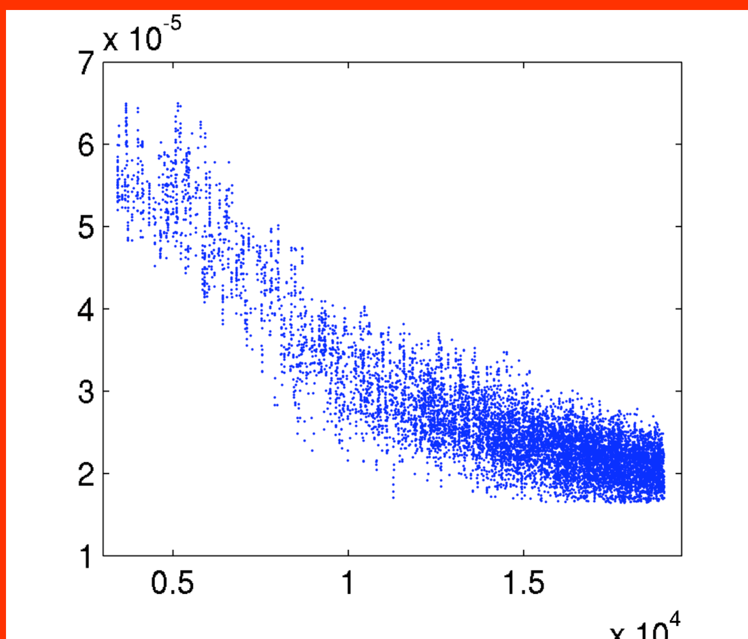
PAR attenuation coefficient



Soil respiration Q_{10}

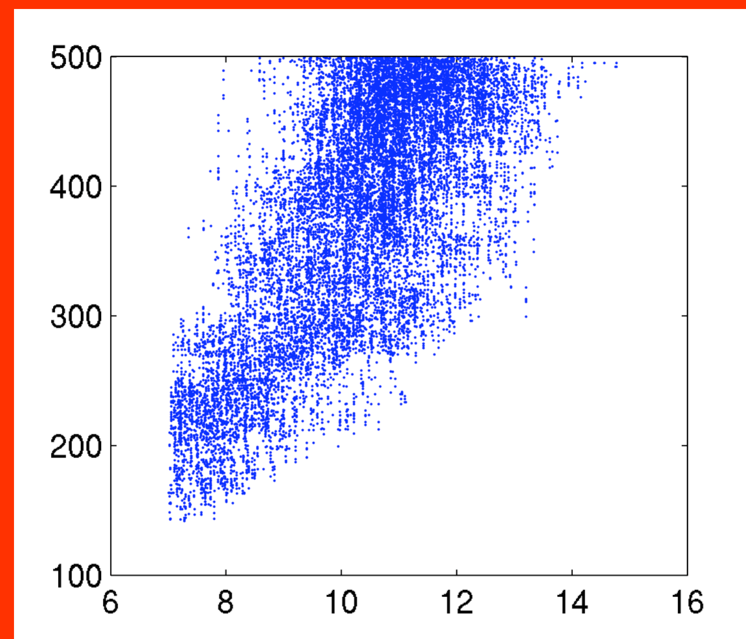
Sacks et al., 2006, GCB
Sacks et al., 2007, Oecologia

Base soil respiration rate
($g\ C\ g^{-1}\ C\ day^{-1}$)

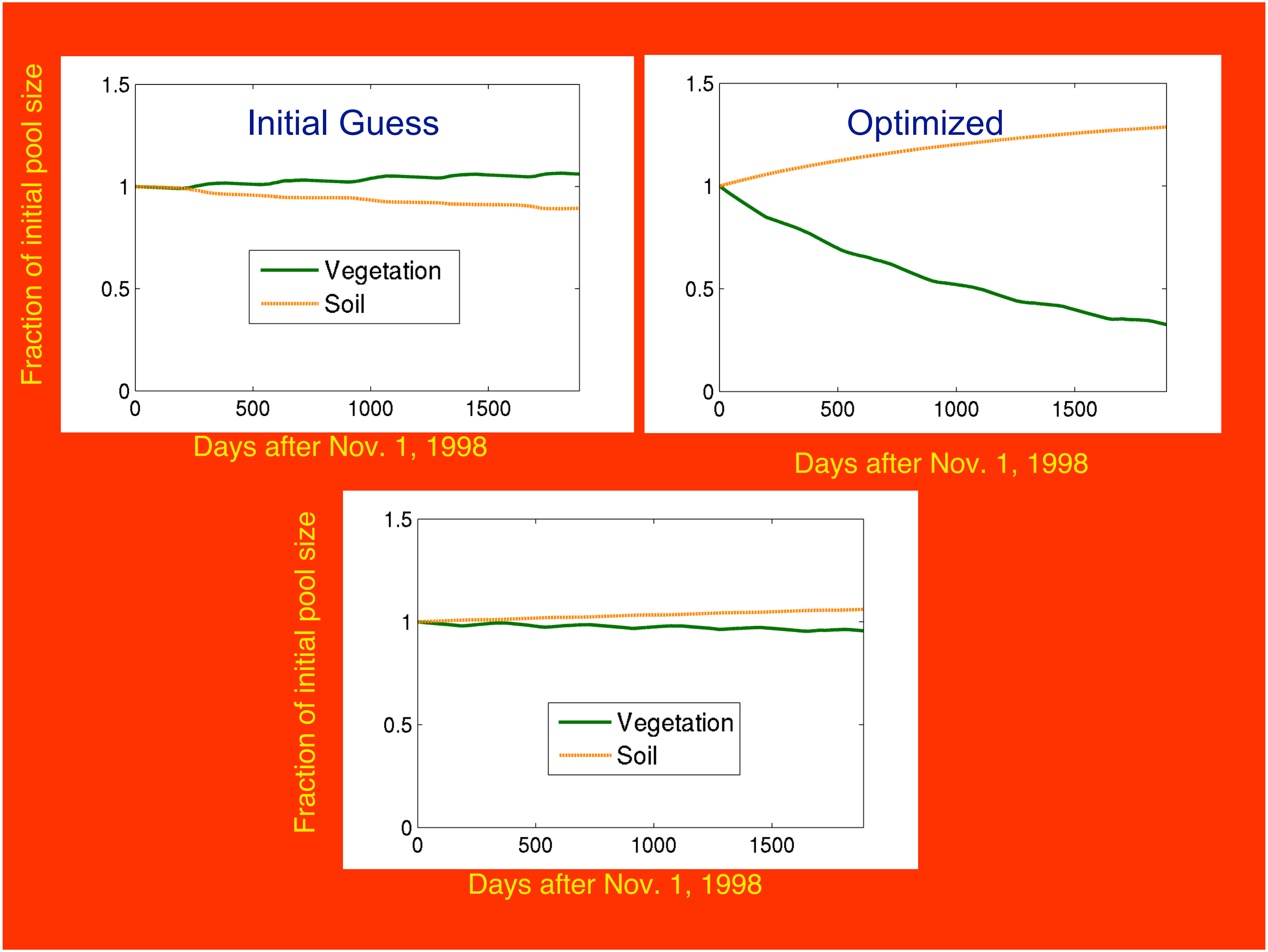


Initial soil C content
($g\ C\ m^{-2}$)

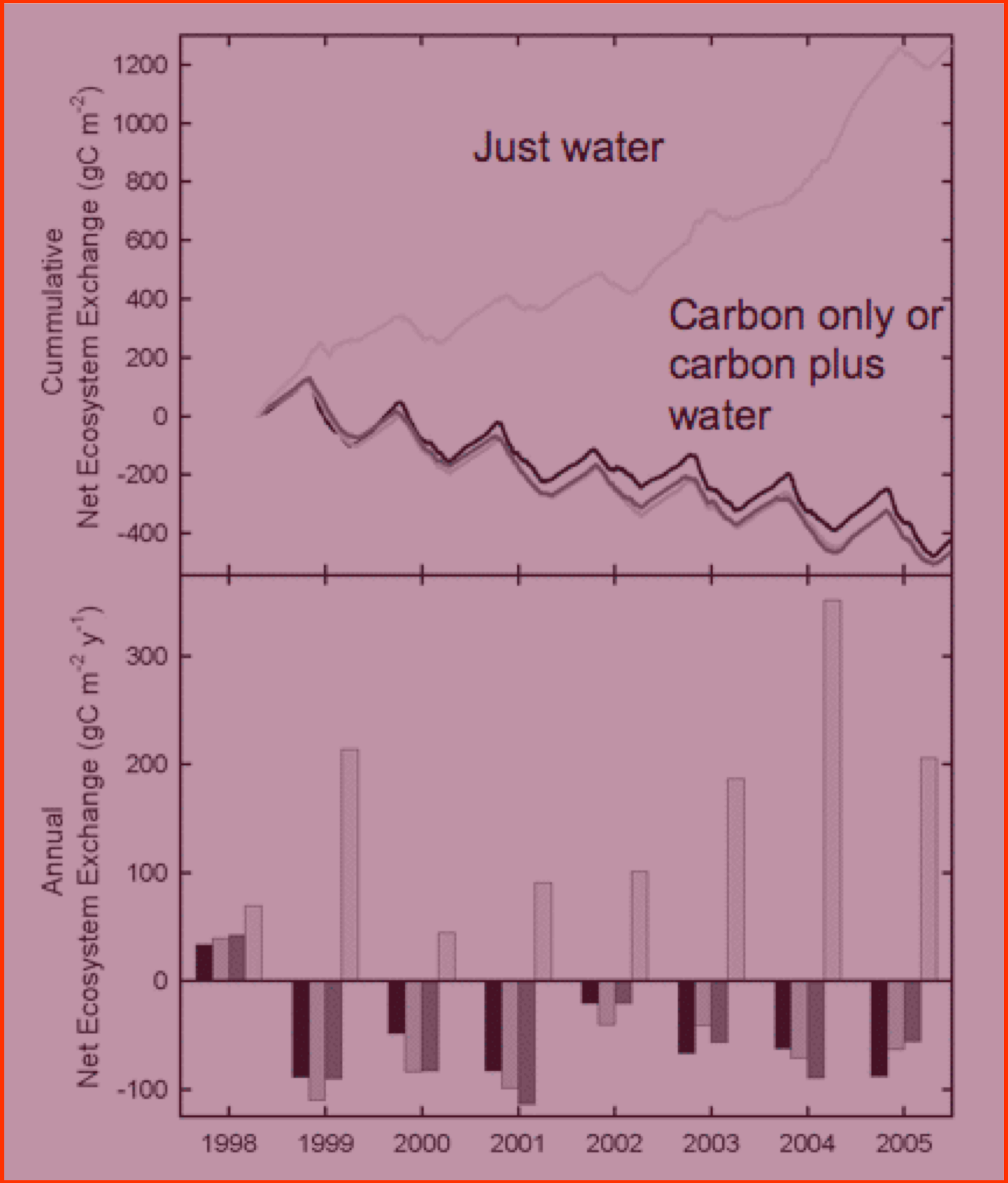
C content of leaves per unit area
($g\ C\ m^{-2}$)

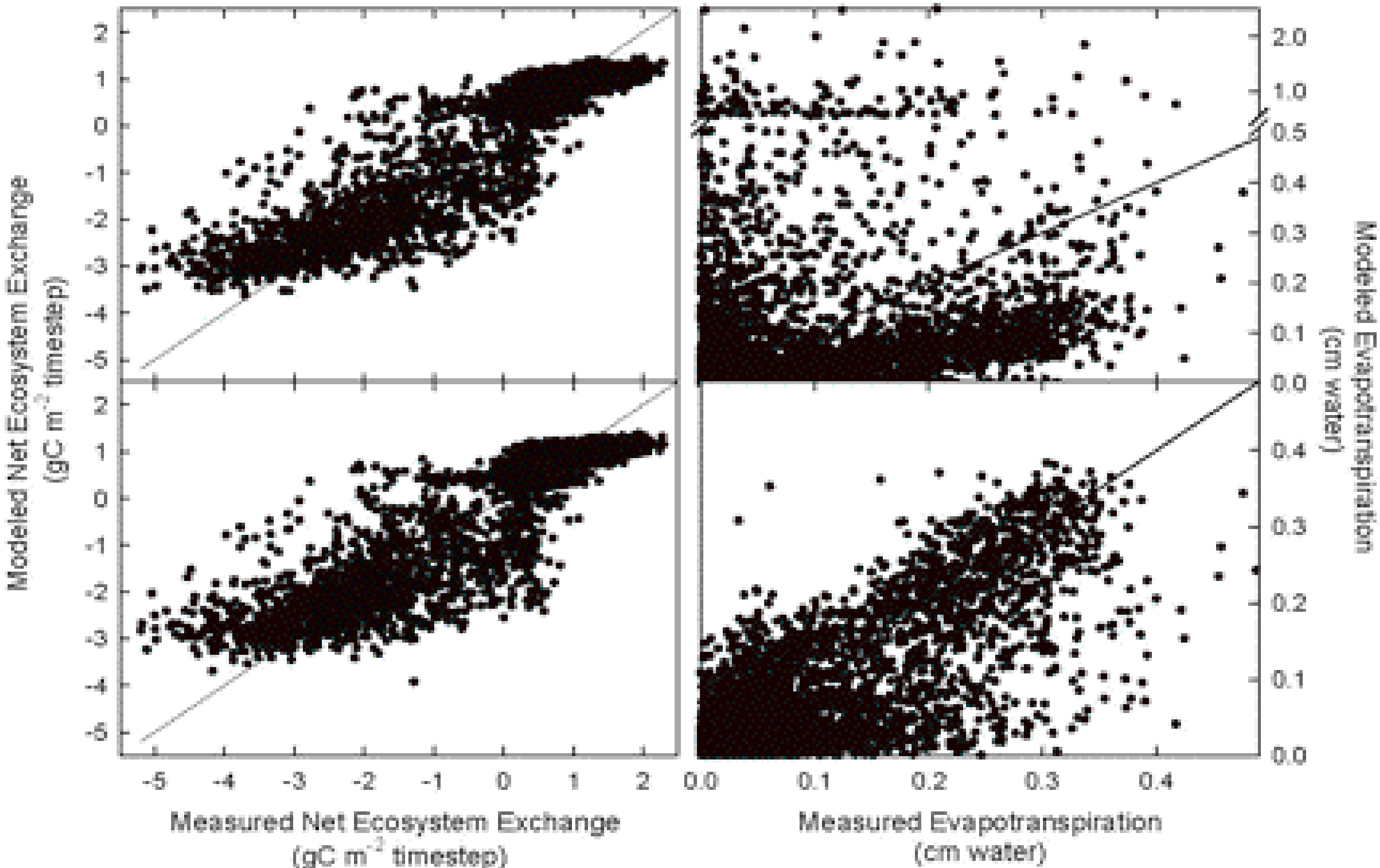


PAR half-saturation point
($mol\ m^{-2}\ day^{-1}$)



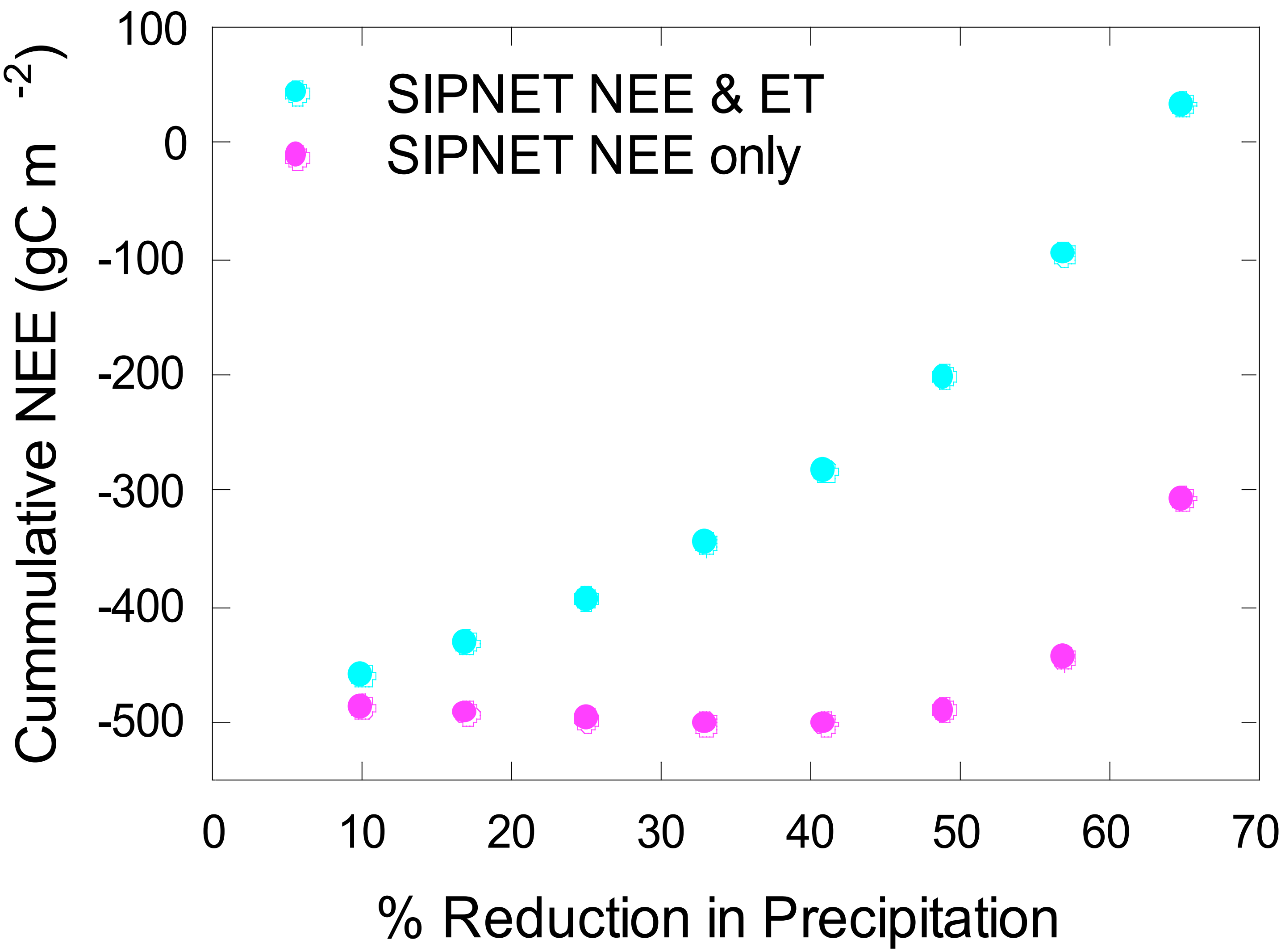
Initial Guess
Optimized
Vegetation
Soil
Vegetation
Soil
Fraction of initial pool size
Fraction of initial pool size
Days after Nov. 1, 1998
Days after Nov. 1, 1998
Days after Nov. 1, 1998


Just water
Carbon only or carbon plus water
Cummulative Net Ecosystem Exchange (gC m-2)
1200
1000
800
600
400
200
0
-200
-400
Annual Net Ecosystem Exchange (gC m-2 y-1)
300
200
100
0
-100
1998
1999
2000
2001
2002
2003
2004
2005

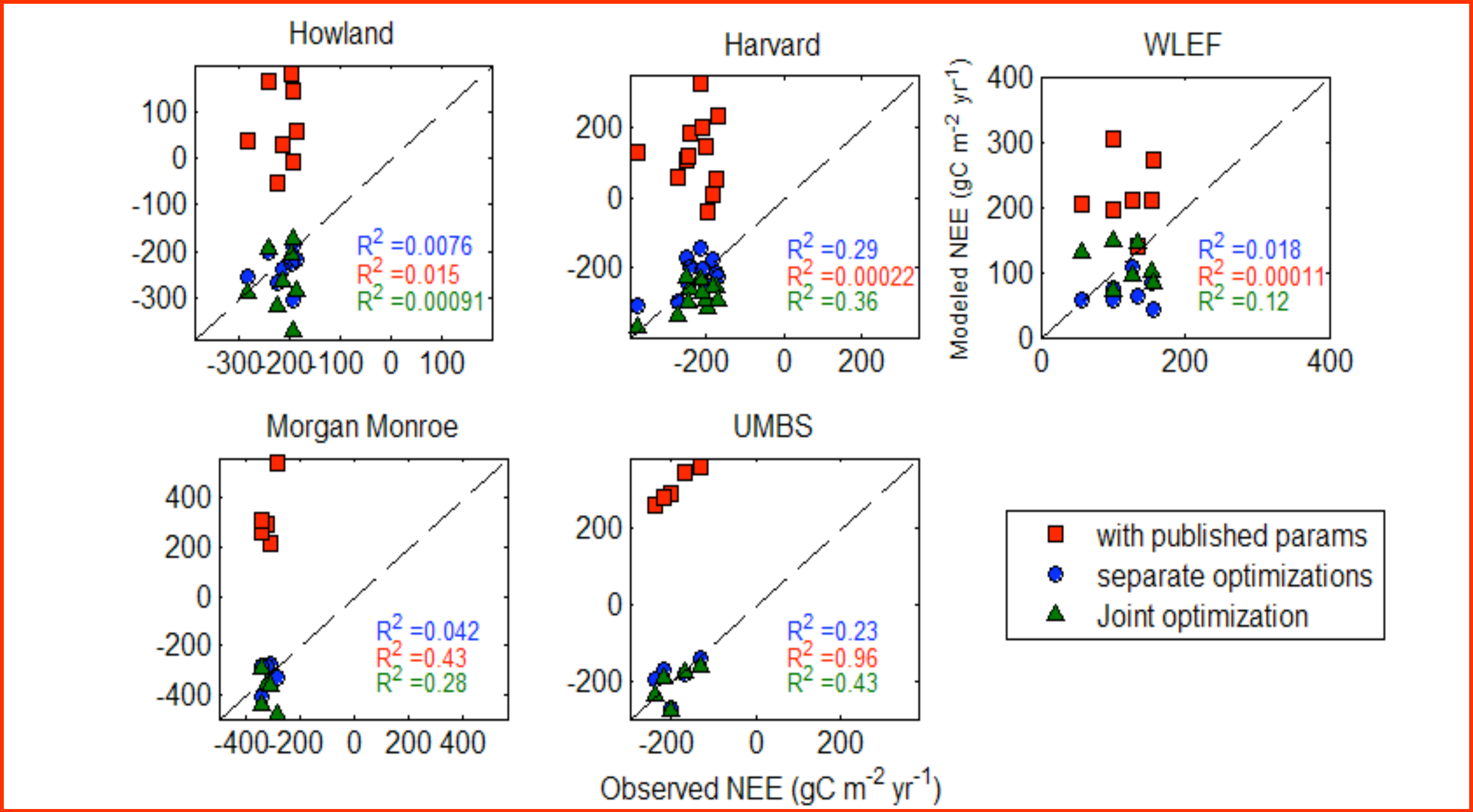


Modeled Net Ecosystem Exchange
(gC m-2 timestep)
Measured Net Ecosystem Exchange
(gC m-2 timestep)
Measured Evapotranspiration
(cm water)
Modeled Evapotranspiration
(cm water)

D. Moore, in prep, AgForMet



Ricciuto et al., in press, AgForMet
Riccituo et al., submitted



Ricciuto et al., submitted

Parameter	Published	HV	HW	WL	MM	UM
CS	10.0	4.81	4.827	**9.81**	10.7	7.37
Q_{10H}	2.0	1.47	**2.394**	**3.28**	1.92	**2.65**
θ_{opt}	0.55	**0.843**	0.717	**0.745**	0.685	**0.764**
θ_{fac}	0.8	1.00	0.320	0.858	1.00	1.00
θ_c	0.30	0.247	**0.129**	0.240	0.285	0.239
θ_w	0.13	0.00	0.0498	0.0534	0.00	0.00
n_{l0BL}	0.036	**0.149**	0.250	**0.130**	**0.100**	**0.112**
n_{l0NL}	0.030	**0.140**	**0.045**	**0.0430**	**0.0500**	0.0516
α_{BL}	0.06	**0.0535**	0.117	**0.0314**	**0.0498**	**0.0353**
α_{NL}	0.06	**0.0274**	**0.021**	**0.483**	0.249	0.340
Q_{10VM}	2.0	**2.23**	**1.74**	**2.36**	**2.15**	**2.04**
$T_{low\ BL}$	-5.0	**1.48**	-40.0	**-9. 70**	**6.06**	**-1.83**
T_{lowNL}	-15.0	**3.55**	**3.41**	**-1.80**	10.0	-20.0
T_{uppBL}	33.0	50.0	50.0	50.0	50.0	50.0
$T_{upp\ NL}$	28.0	**11.66**	50.0	50.0	10.0	49.1
D_c	0.09	**0.0192**	**0.0266**	**0.0249**	**0.0371**	**0.0282**
R_g	0.25	0.107	0.0500	0.122	0.0500	0.0500
R_{dc}	0.015	0.00822	**0.0208**	**0.0183**	0.001 9	**0.00866**
Q_{10RD}	2.0	**1.49**	**1.61**	**2.00**	**1.100**	**1.33**
T_{off}	0.0	**11.6**	**0.80**	**9.30**	**14.3**	**11.4**
β_{off}	0.90	0.678	0.850	**0.672**	0.172	0.517
γ_p	**14.5**	**18.2**	**13.4**	**16.5**	20.0	38.7
-(Log L)	N/A	157209	100989	69511	68554	61217

Well-constrained parameters:

- leaf nitrogen
- quantum efficiency
- T_{lower}
- phenology (T_{off})

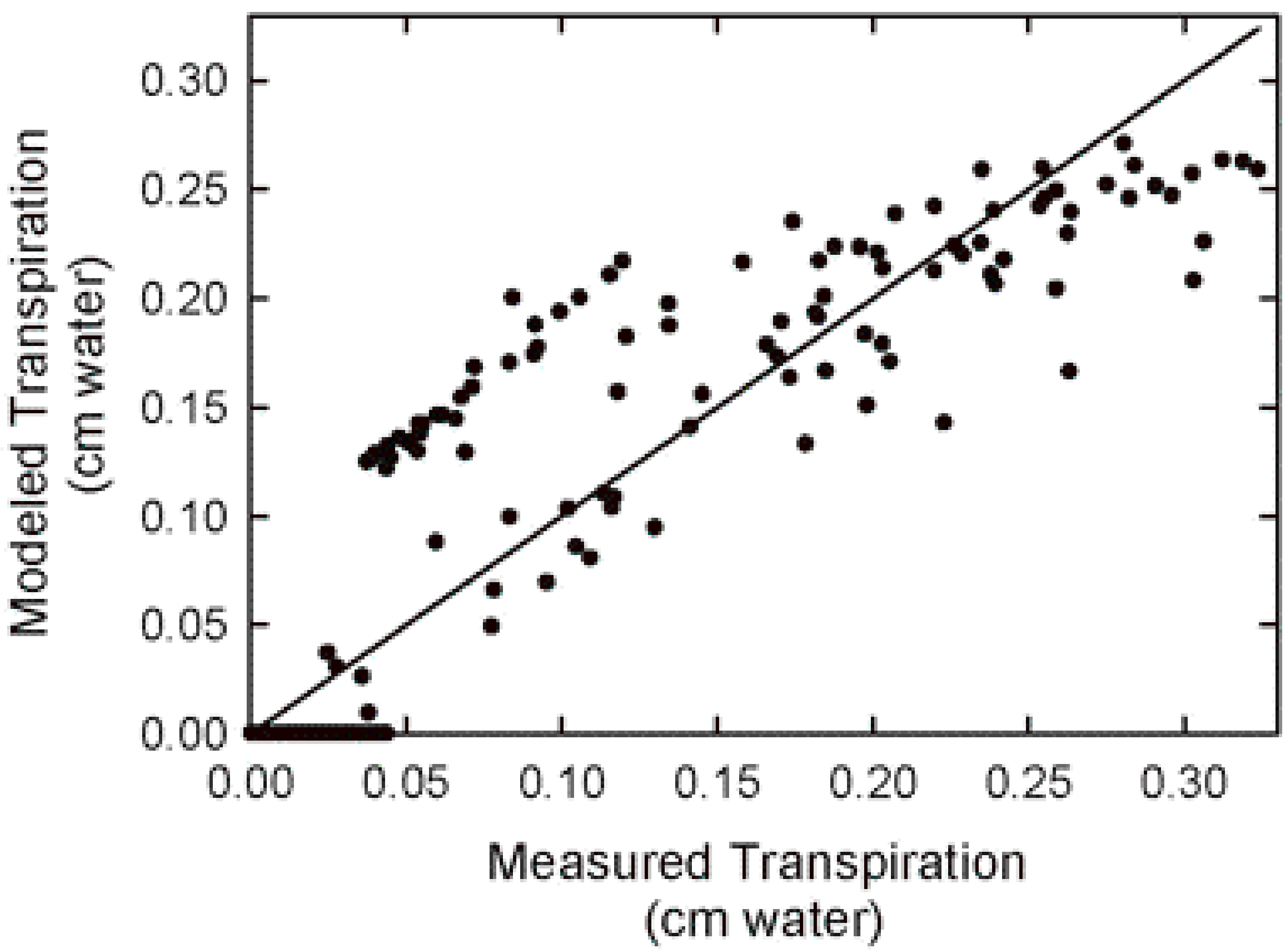
Poorly constrained parameters

- autotrophic respiration
- soil moisture parameters

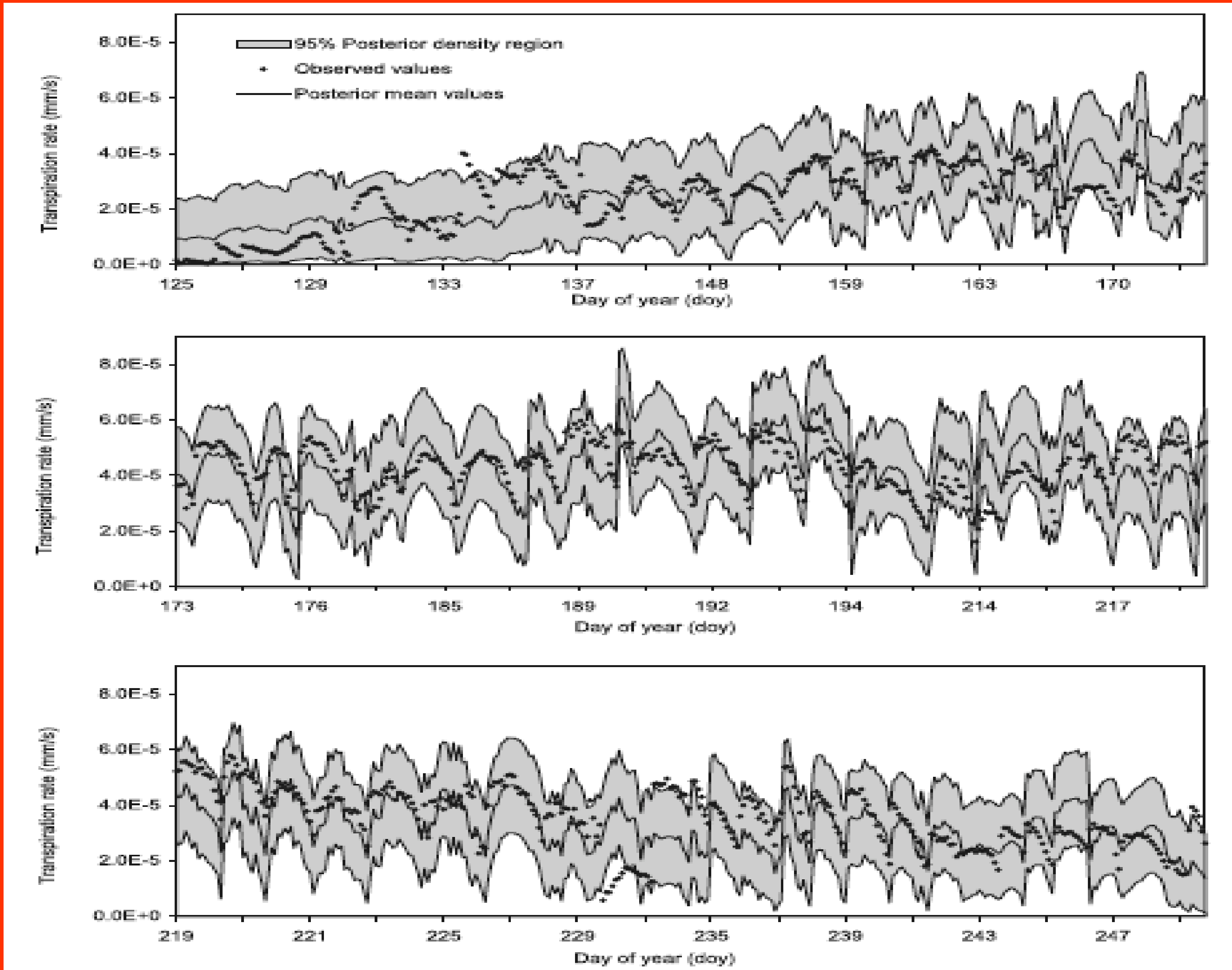
Difference from published

Coherence across space

Moore et al., submitted



Samanta et al., 2007, WaterResourcesRes.



A Request and Future Directions

- For model optimization / data assimilation - missing flux data is not necessarily a big deal
- But: Filled driver data (micromet, soil temp, radiation) are generally needed
- Note: Driving data vs assimilated data is an arbitrary choice (as long as you can build a forward operator)
- We can assimilate not just CO_2/H_2O flux, but also concentrations, micromet, radiation, wind, isotopes, chamber flux, inventory data, sapflux
- Assimilation of multiple towers promising, but consistent data sets needed

Thanks

- Thanks to the PSU/ UW-Madison/ NCAR/ SUNY-Buffalo/ CSU/ ORNL/ King's College / WHRC/ ChEAS flux data assimilation collective and funders, DOE, NICCR, NSF, NASA, USDA