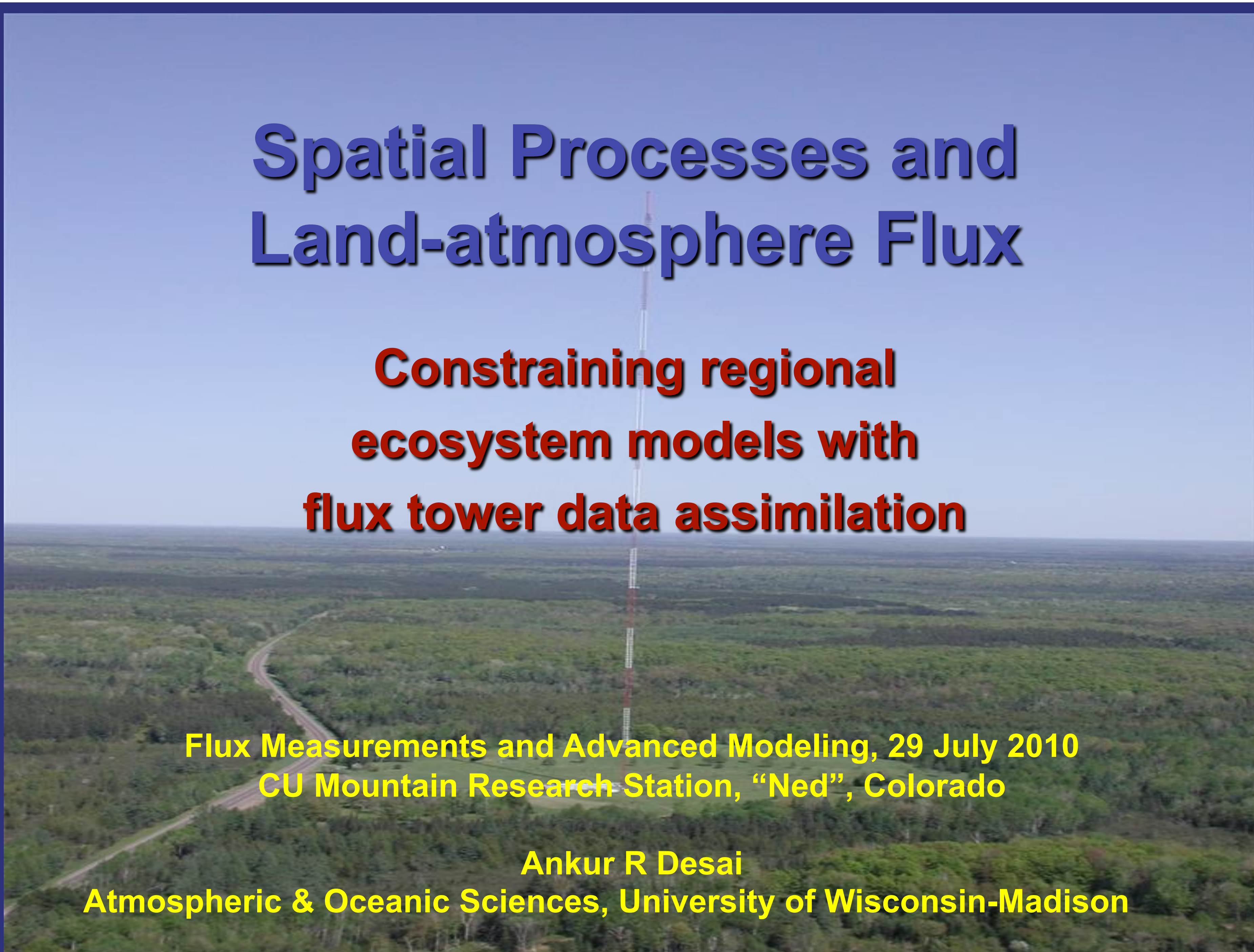


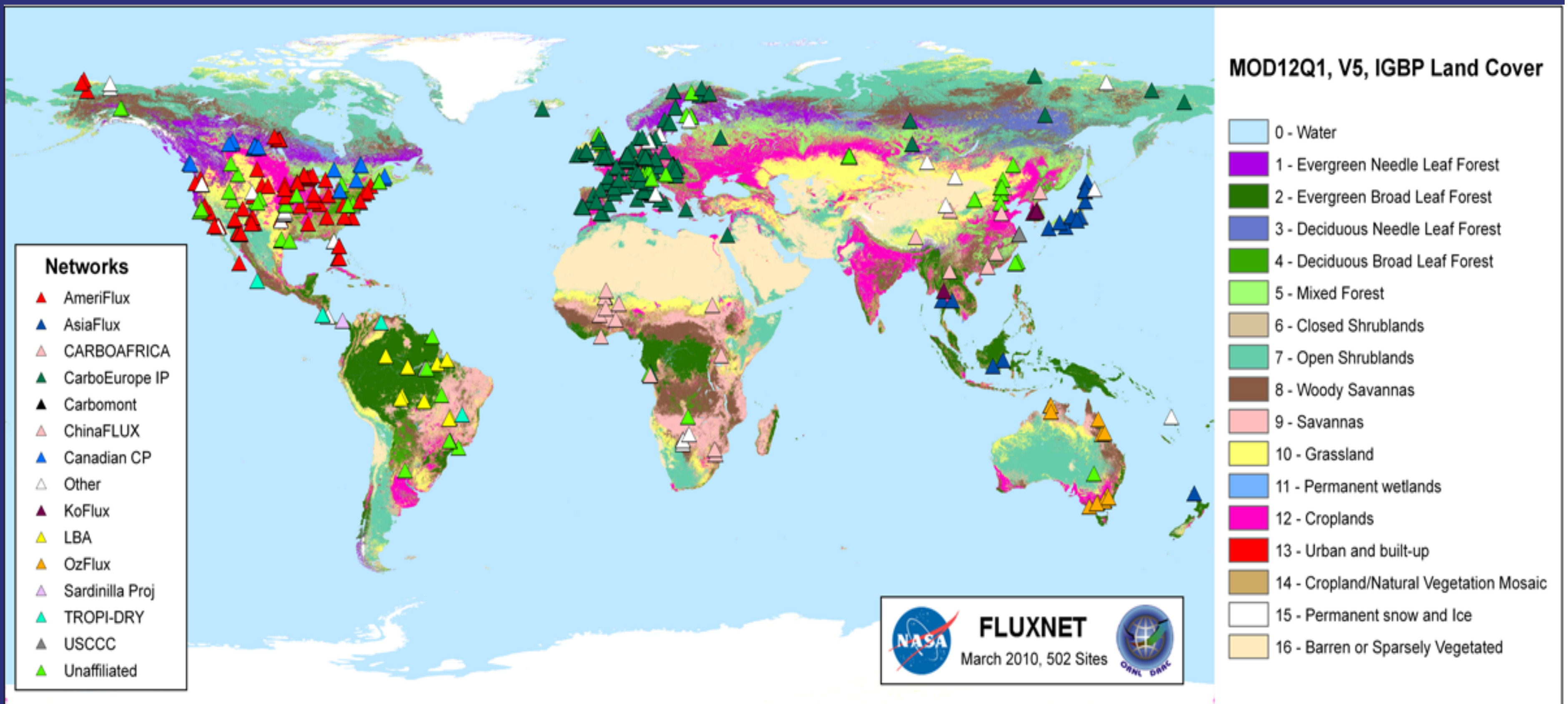
Spatial Processes and Land-atmosphere Flux

Constraining regional ecosystem models with flux tower data assimilation

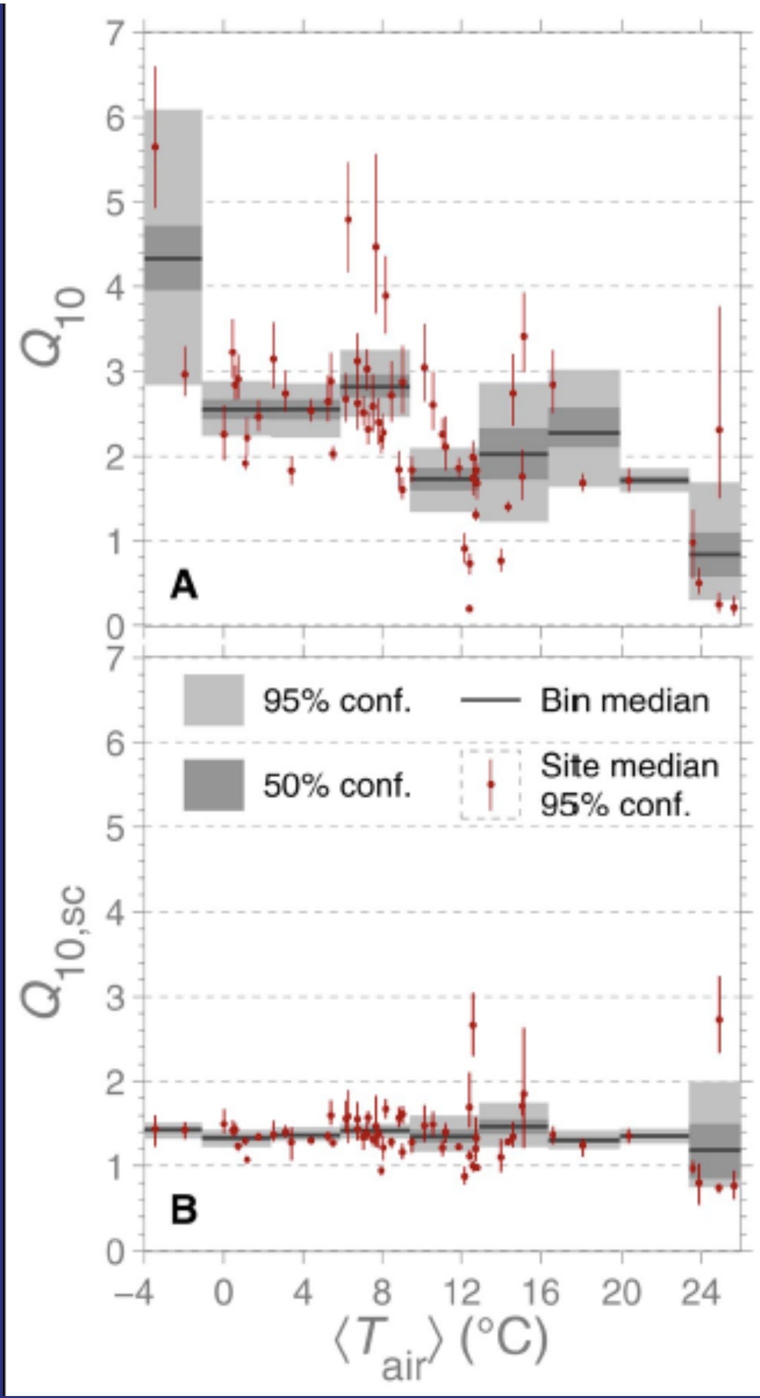
Flux Measurements and Advanced Modeling, 29 July 2010
CU Mountain Research Station, “Ned”, Colorado

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

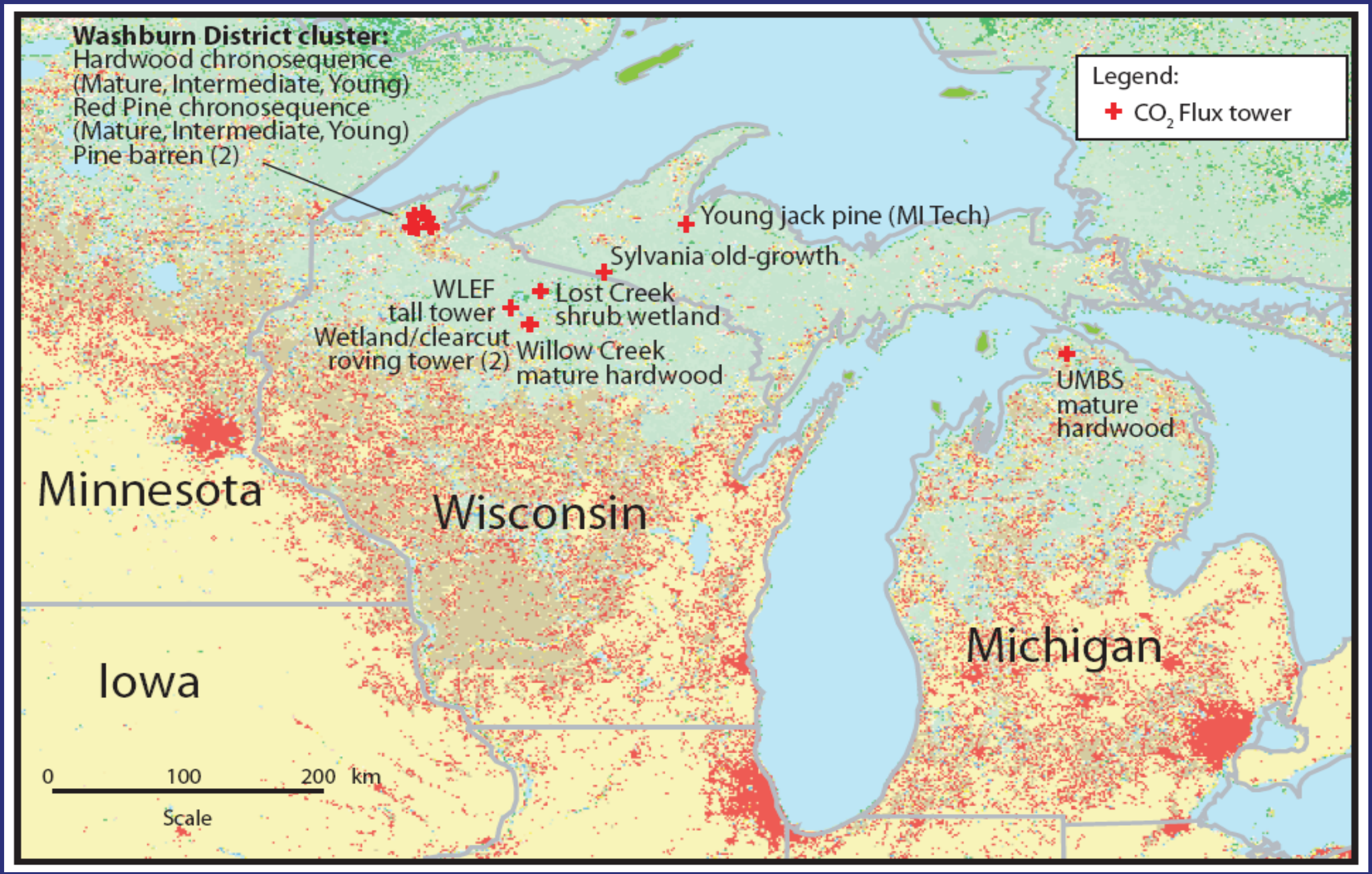
Let's Get Spacey…



- Mahecha et al., 2010, Science
- LaThuile Fluxnet Synthesis Database
  - http://www.fluxdata.org



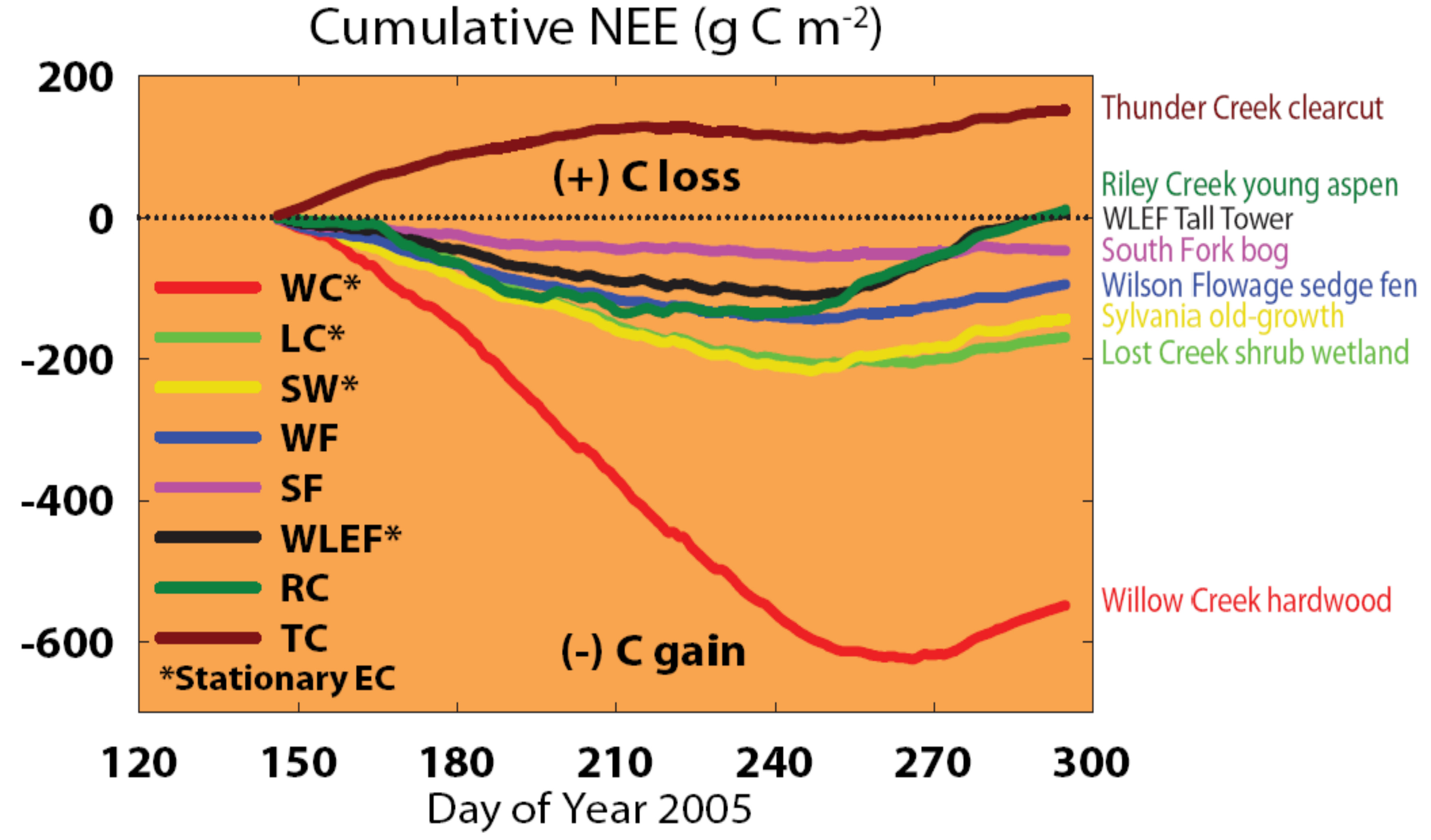
Let’s Get Regional, too!



Why Regional?

- Spatial interpolation/extrapolation
- Evaluation across scales
- Landscape level controls on biogeochem.
- Understand cause of spatial variability
- Emergent properties of landscapes

Why Regional?



Courtesy: Nic Saliendra

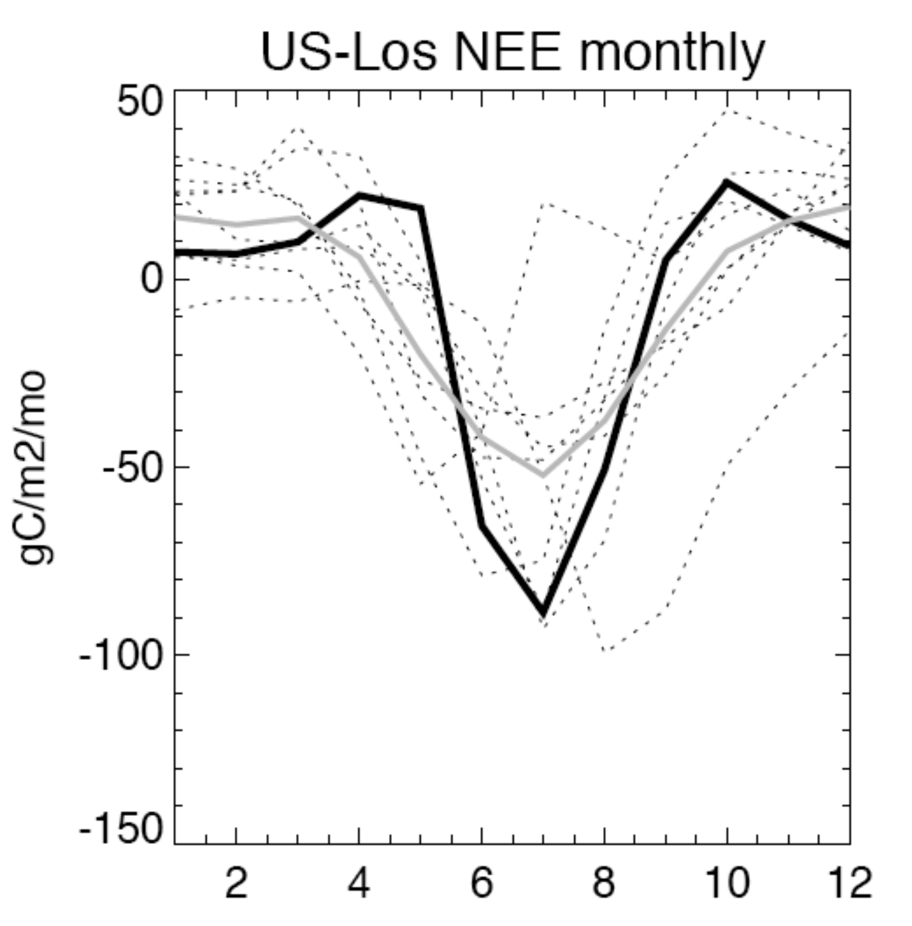
Why Data Assimilation?

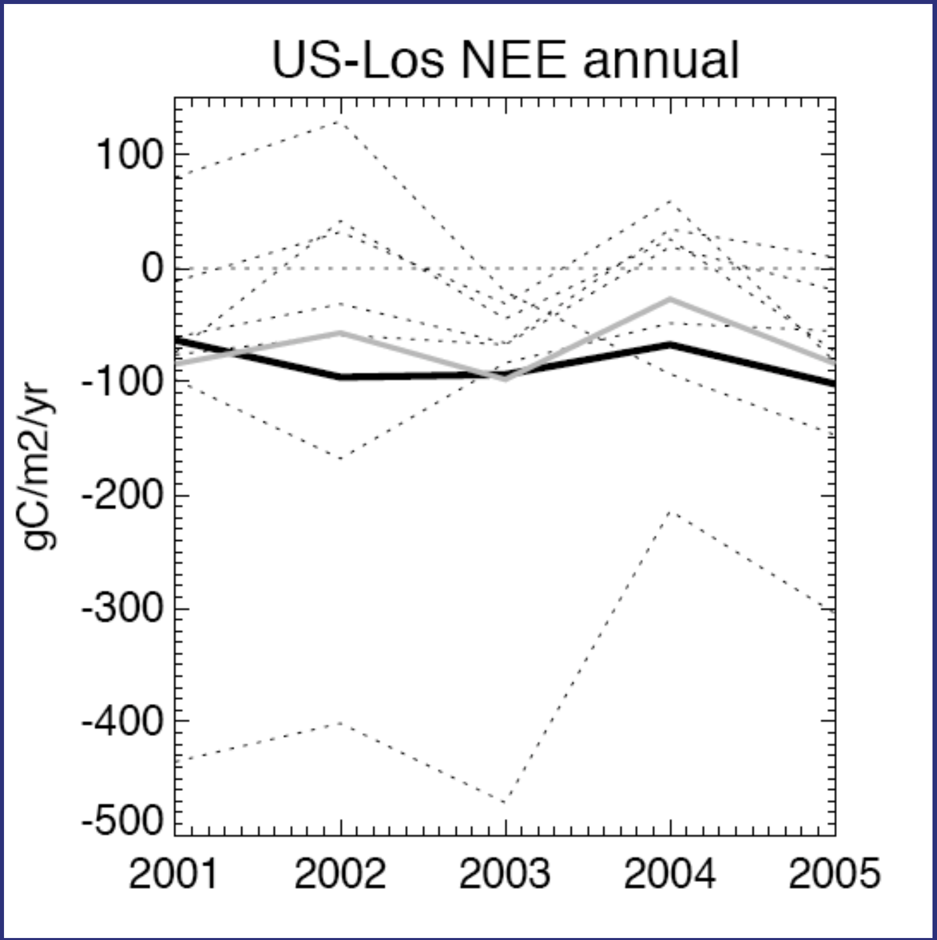
- Meteorological, ecosystem, and parameter variability hard to observe/model
- Data assimilation can help isolate model mechanisms responsible for spatial variability
- Optimization across multiple types of data
- Optimization across space

Why Data Assimilation?

- Old way:
  - Make a model
  - Guess some parameters
  - Compare to data
  - Publish the best comparisons
  - Attribute discrepancies to error
  - Be happy

Discrepancies





Why Data Assimilation?

- New way:
  - Constrain model(s) with observations
  - Find where model or parameters cannot explain observations
  - Learn something about fundamental interactions
  - Publish the discrepancies and knowledge gained
  - Work harder, be slightly less happy, but generate more knowledge

Back to Those Stats…

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

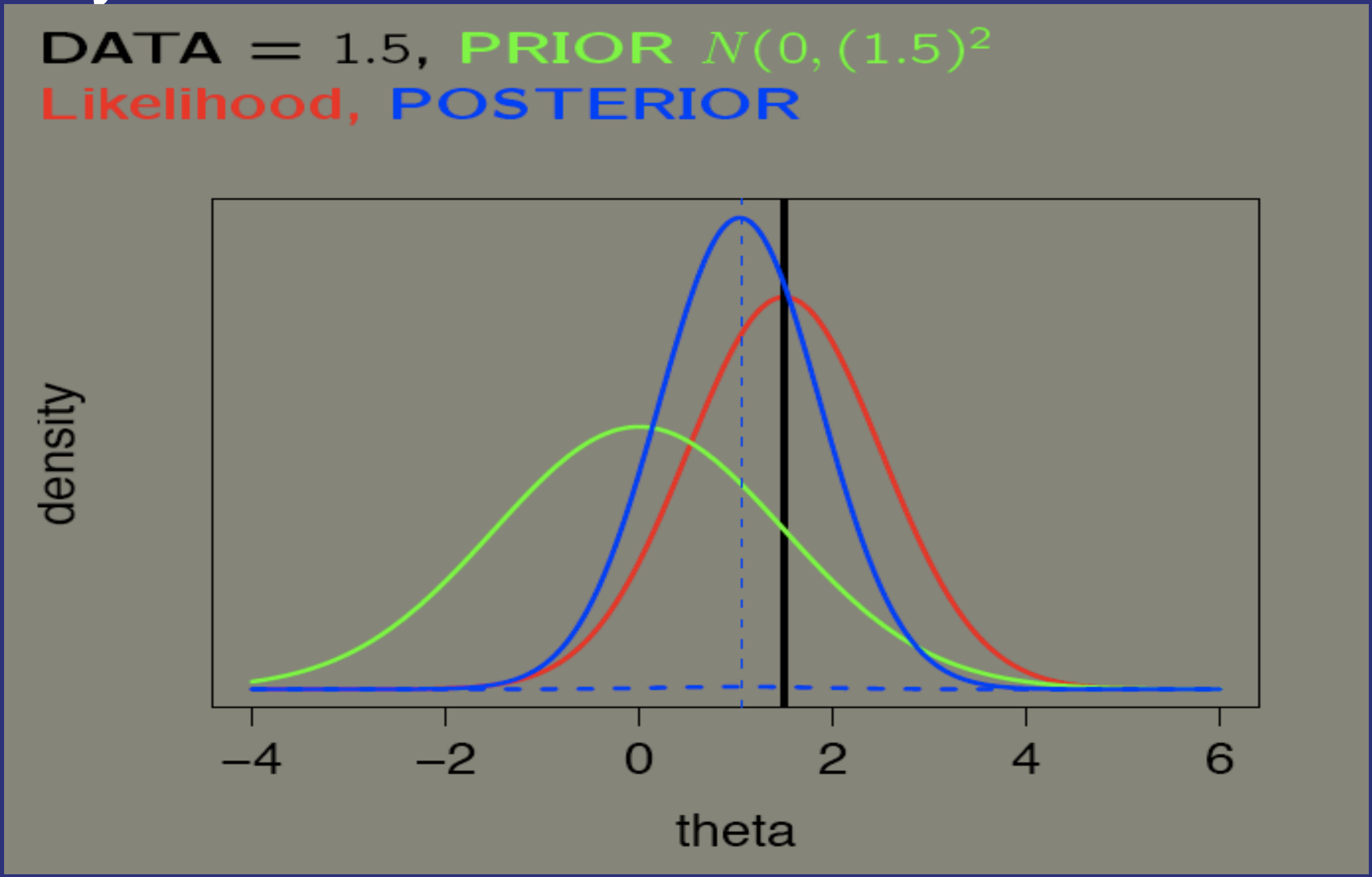
[P|D] = ([D|P] [P]) / [D]

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

Posterior =
(Likelihood x Prior) / Normalizing Constraint

For the Visually Minded

- D Nychka, NCAR



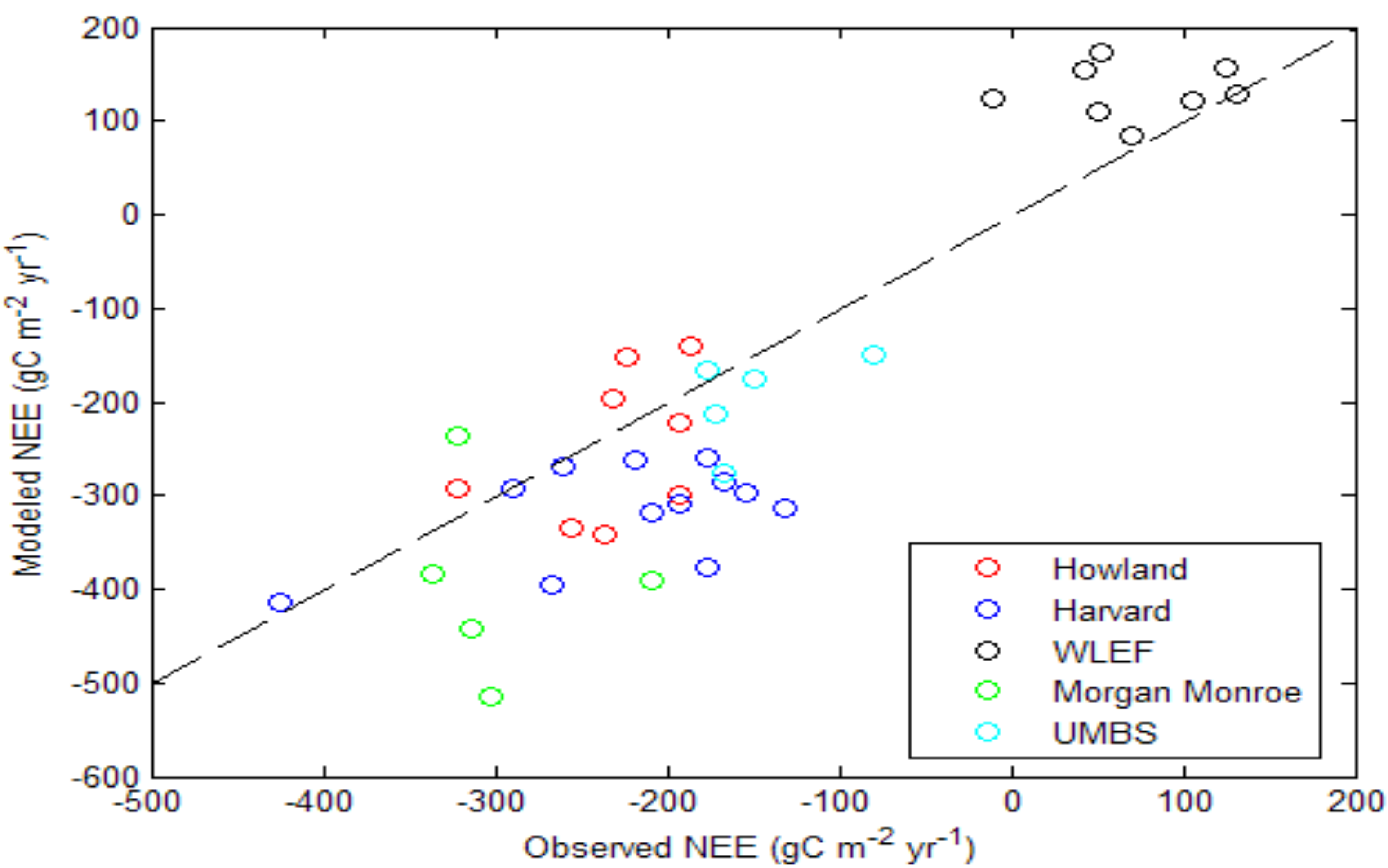
A Case Study

- Coherent Interannual Variability
- Flux Decomposition

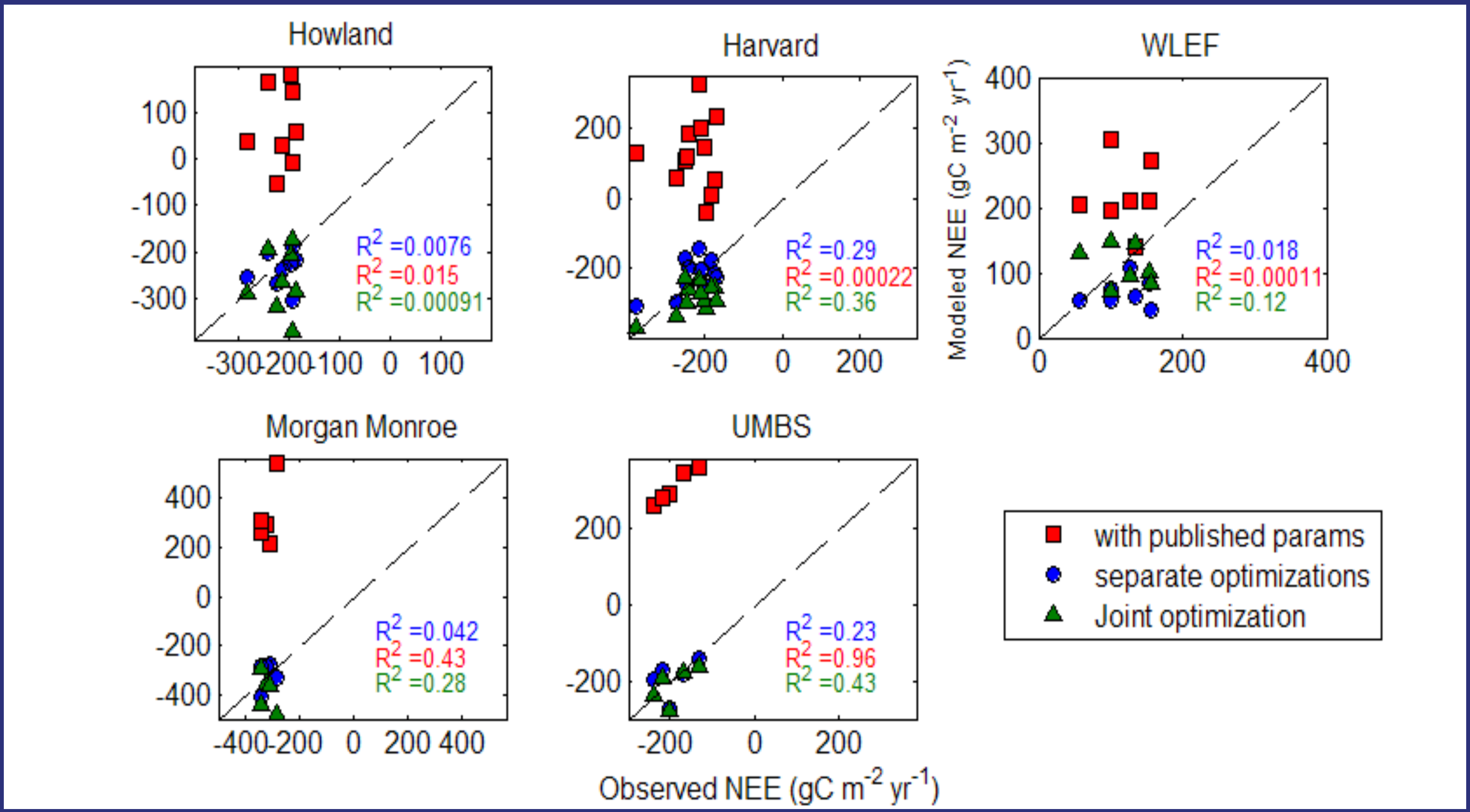
Coherent Interannual Variability

Desai et al., 2010 (accepted) JGR-G

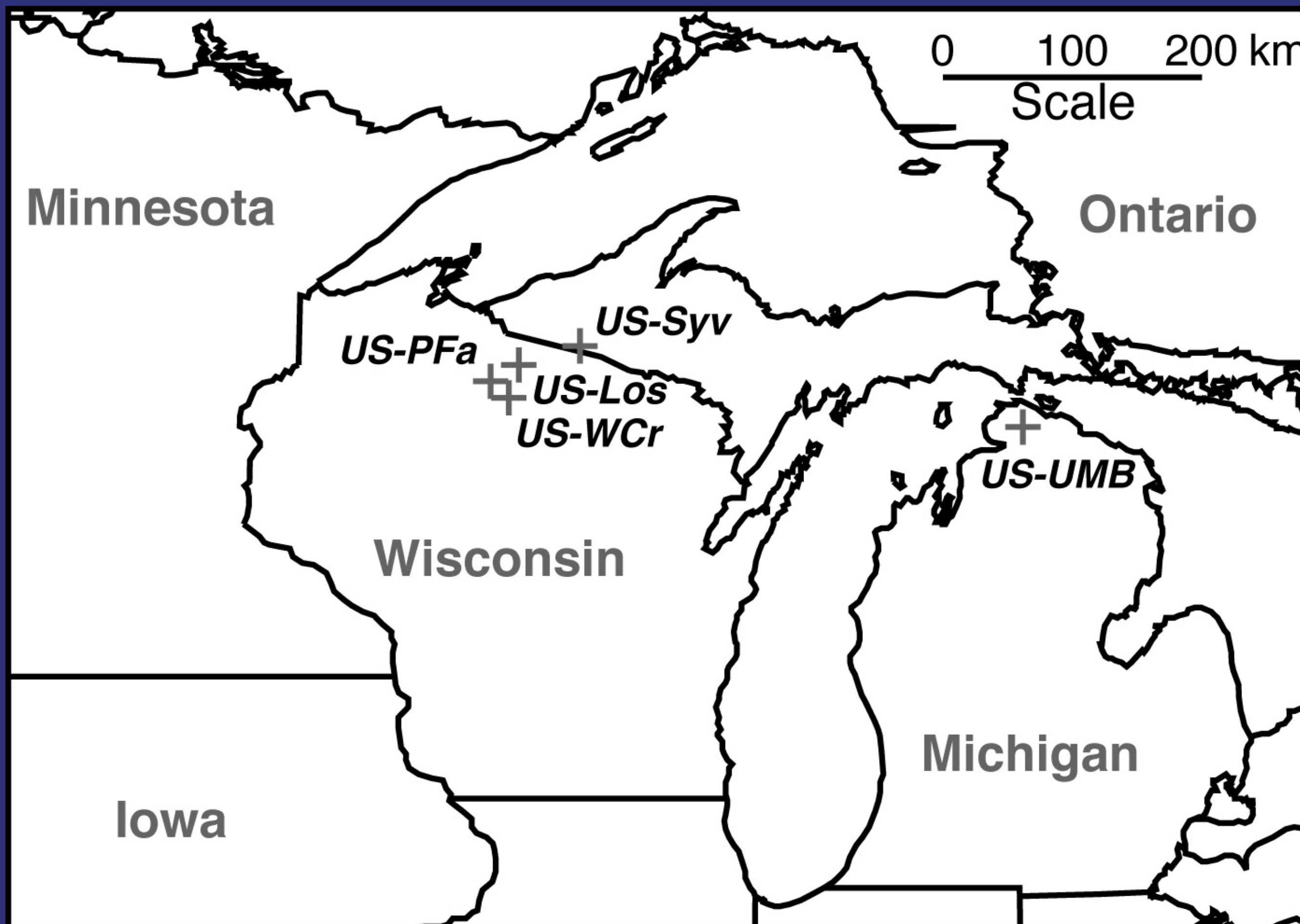
Ricciuto et al.



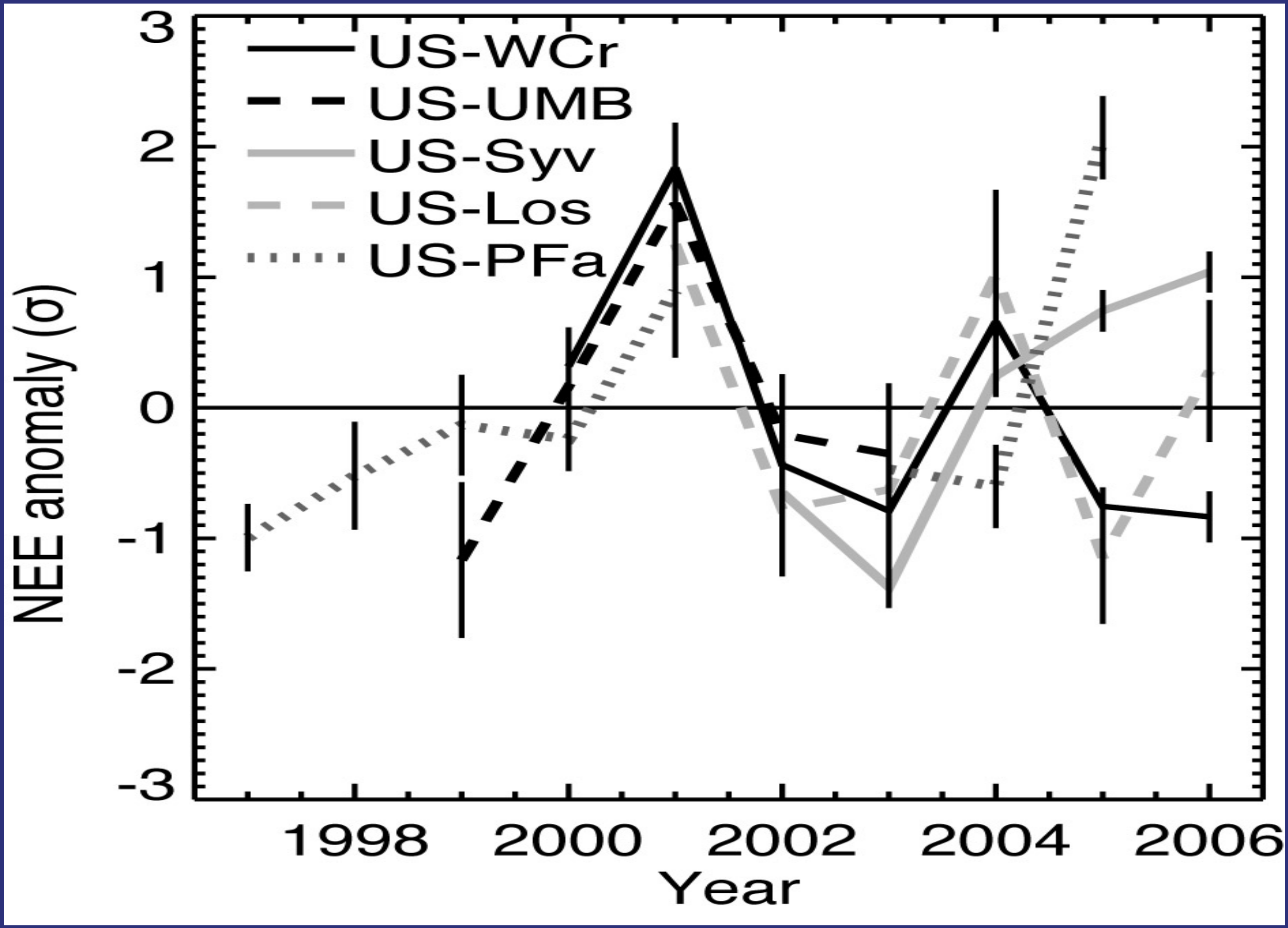
Ricciuto et al.



The Sites


0 100 200 km
Scale
Minnesota
Ontario
US-Syv
US-PFa
US-Los
US-WCr
US-UMB
Wisconsin
Michigan
Iowa

Regional Coherence



IAV

- Does variability in growing season start or end explain IAV in this region?
  - Hypothesis: growing season length explains IAV
- If so, are the controls also coherent across region?
- Steps:
  - Construct a simple ecosystem model
  - Assimilate flux data and information about IAV

Simple Model

- Twice daily model, annually resetting pools
- Driven by PAR, Air and Soil T, VPD
- LUE based GPP model f(PAR,T,VPD)
- Three respiration pools f(Air T, Soil T, GPP)
- Phenology
  - Sigmoidal Threshold GDD (base 10) function for leaf on
  - Sigmoidal Threshold Daily Mean Soil Temp function for leaf off
- 17 parameters, 3 are fixed
  - Output: NEE, ER, GPP, LAI

Parameters

Name	Definition	Value
Fixed parameters		
k	Light extinction coefficient	0.5 FIXED
LAI_{min}	Minimum leaf area	US-WCr 0.0, US-UMB 0.0, US-Syl 0.5, US-Los 0.0, US-PFa 0.5
LAI_{max}	Maximum leaf area	US-WCr 5.3, US-UMB 3.7, US-Syl 4.1, US-Los 4.9, US-PFa 3.7
Phenology parameters		
α	Leaf on (L_{ON}) slope	0.05 (0.05-0.5)
GDD_{thresh}	Growing degree day threshold	200 (10-400)
β	Leaf off (L_{OFF}) slope	0.1 (0.05-0.5)
$TEMP_{thresh}$	Soil temperature threshold	4 (0-20)
Photosynthesis parameters		
LUE	Light use efficiency	0.25 (0-1)
T_{min}	Minimum photosynthetic temperature	0 (-15-10)
T_{opt}	Optimum photosynthetic temperature	15 (5-40)
VPD_{max}	Maximum photosynthetic VPD	3000 (0-20000)
VPD_{min}	Minimum photosynthetic VPD	100 (0-2000)
Respiration parameters		
r_s	Basal maintenance respiration	2 (0.1-5)
r_v	Basal growth respiration	2 (0.1-5)
b_1	Maintenance respiration rate	0.03 (0-0.5)
b_2	Growth respiration rate	0.03 (0-0.25)
b_3	Leaf respiration fraction	0.05 (0-0.25)

Assimilation Method

- MCMC is an method to minimize model-data mismatch
  - Quasi-random walk through parameter space (Metropolis-Hastings)
    - Prior parameters distribution needed
    - Start at many random places (chains)
  - 1. Randomly change parameter from current to a nearby value
    - Use simulated annealing to tune how far you move from current spot
  - 2. Move “downhill” to maximize a likelihood in model-data error
    - Avoid local minima by occasionally performing “uphill” moves in proportion to maximum likelihood of accepted point
  - 3. End chain when % accepted reaches a threshold, or back to 1
  - 4. Pick best chain and continue space exploration
    - Save parameter sets after a “burn-in” period
    - End result – “best” parameter set and confidence intervals
- Any sort of observations could be used, but need a fast model and many iterations

Cost Function

- Original log likelihood computes sum of squared difference at hourly
  - Maybe it overfits hourly data at expense of slower variations?
- What if we also added some information about longer time scale differences to this likelihood?

New Cost Function

Original

$$L_D = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x_i - \mu_i)^2}{2\sigma^2}}$$

Modified

$$\begin{cases} L_y = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x_i^m - \mu_i^m)^2}{2\sigma^2}} \\ x_i^m = \sum_{j=i-DOY}^{i} x_i; \mu_i^m = \sum_{j=i-DOY}^{i} \mu_i \\ L = L_D L_y \end{cases}$$

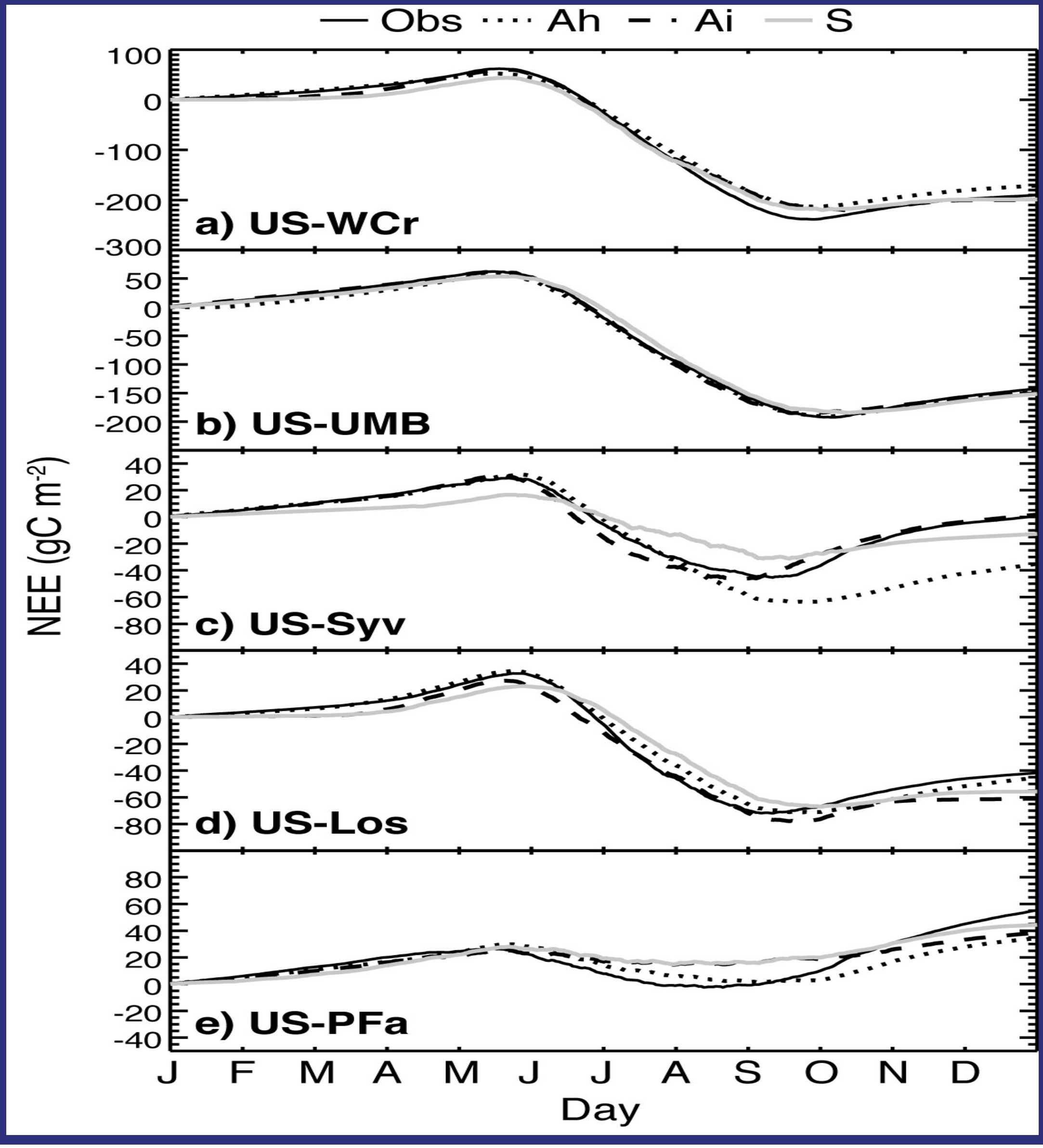
Synchrony Cost Function

- Joint spatial data assimilation for the four phenology parameters
  - If phenology controls IAV coherence, then the joint data assimilation should do as well as the site-level assimilation
  - Method: Concatenate all 20 years of flux data (5 sites x 5 years), estimate 50 independent parameters (5 sites x 10 param) + 4 common parameters, significantly boost # iterations!

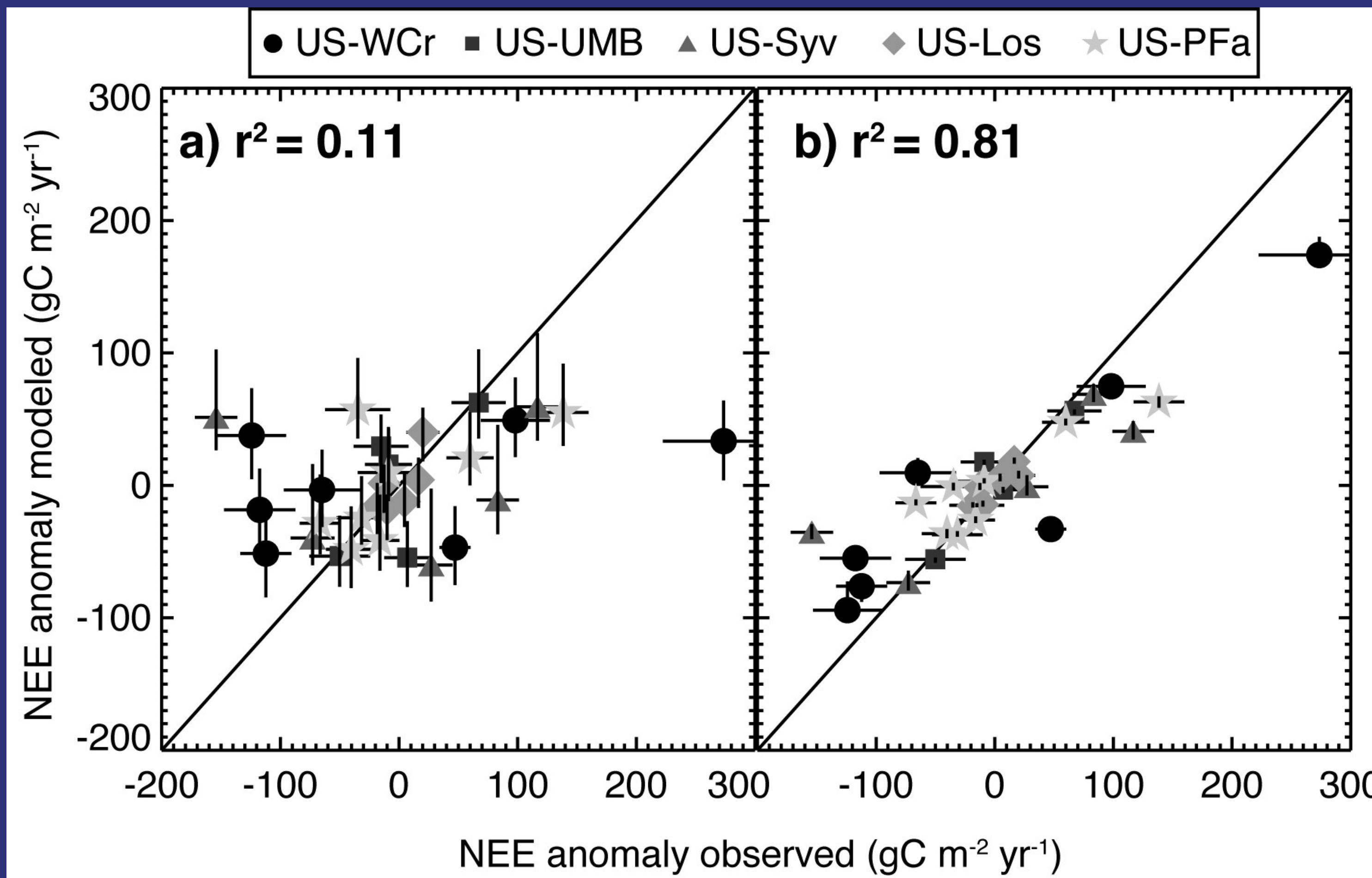
Experiment Design

- Ah Site assimilation, Original CF
- Ai Site assimilation, Modified CF
- S Synchronous assimilation, Modified CF

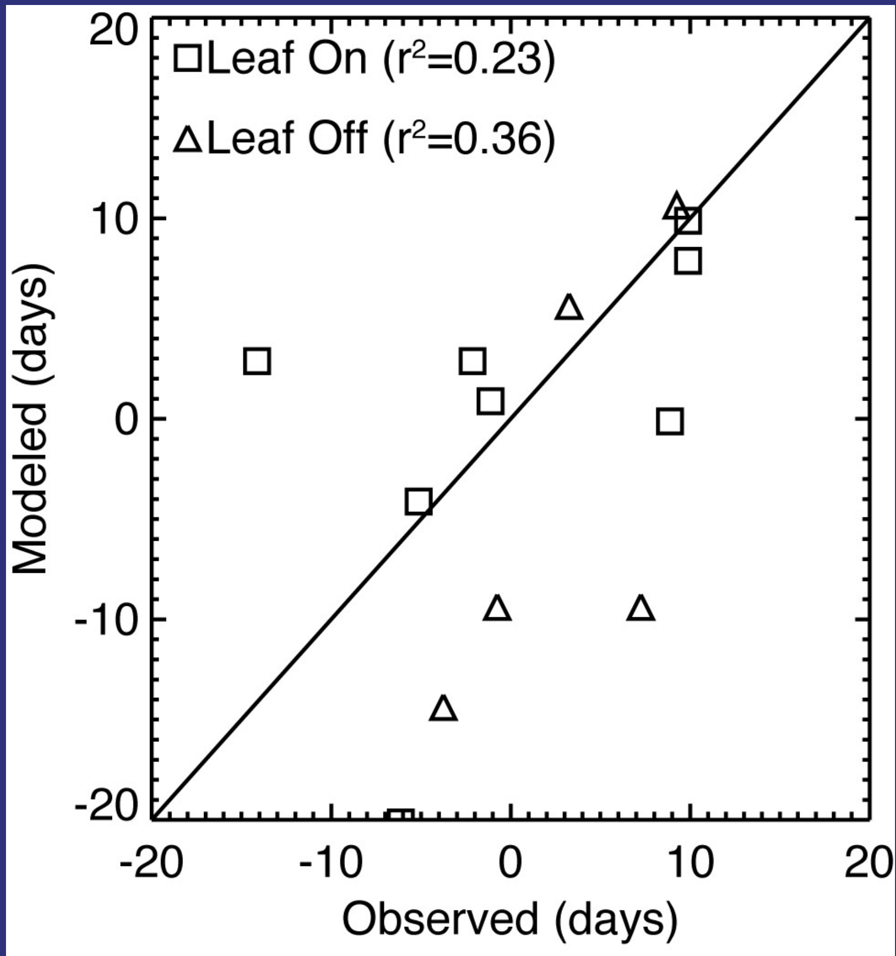
Cumulative NEE



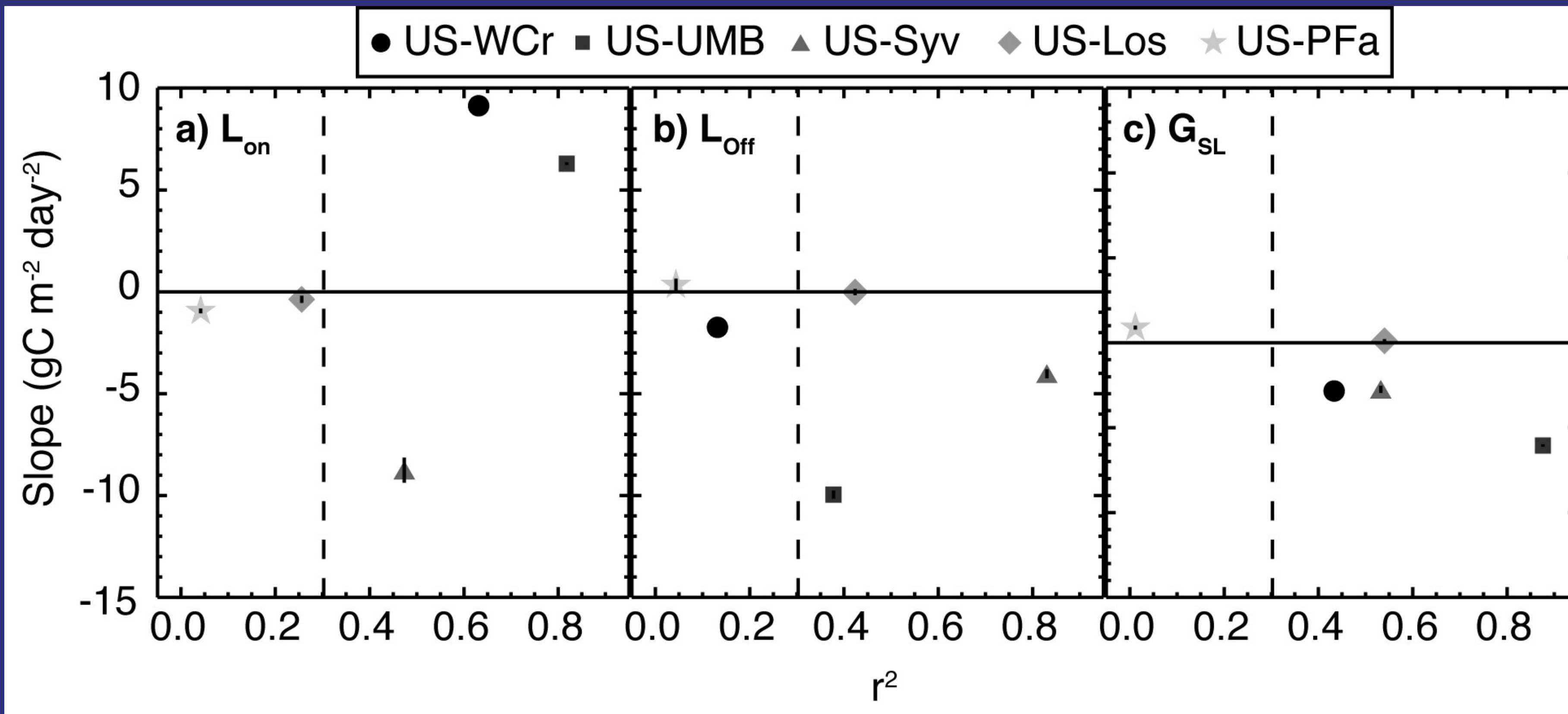
Ah (left) vs Ai (right)



Did We Just Get Lucky?

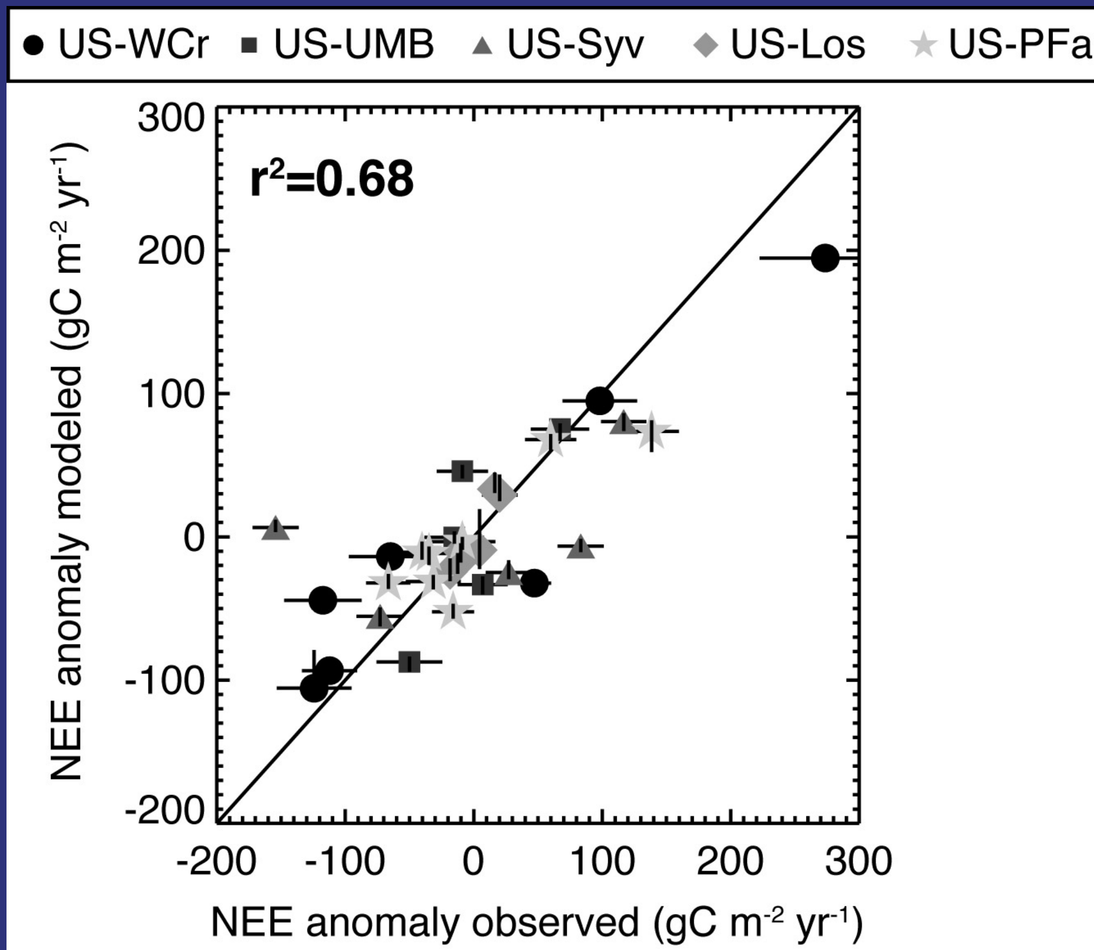


Controls



US-WCr
US-UMB
US-Syv
US-Los
US-PFa
a) L_{on}
b) L_{Off}
c) G_{SL}
Slope (gC m-2 day-2)
10
5
0
-5
-10
-15
0.0 0.2 0.4 0.6 0.8
r2

Synchronous IAV



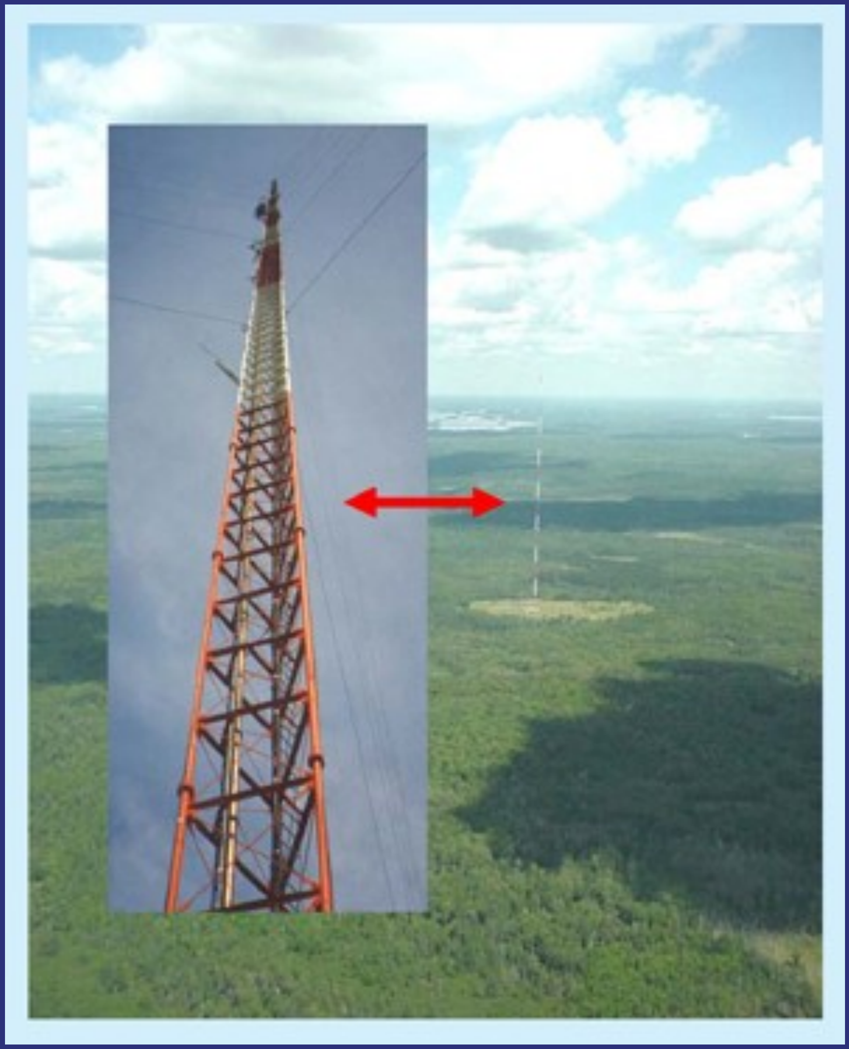
Needs

- Evaluation against independent data
- Cost functions for multiple kinds of data with differing time steps
  - Spectral techniques? (Stoy et al., 2009)
- Testing multiple models
  - Information criteria, different flavors of MCMC

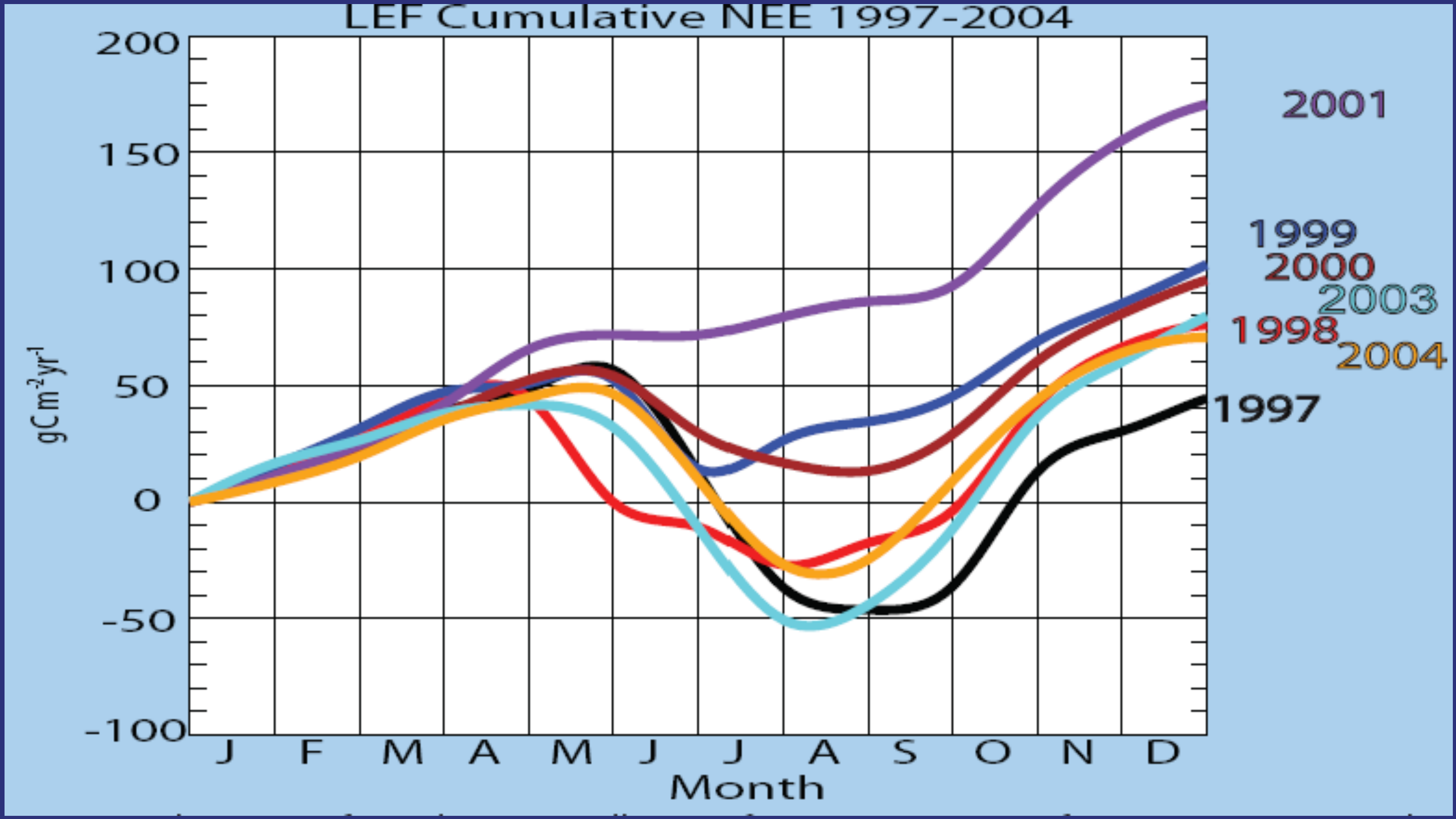
Flux Decomposition

Wang et al., 2006, JGR-G
Desai et al., 2008, Ag For Met

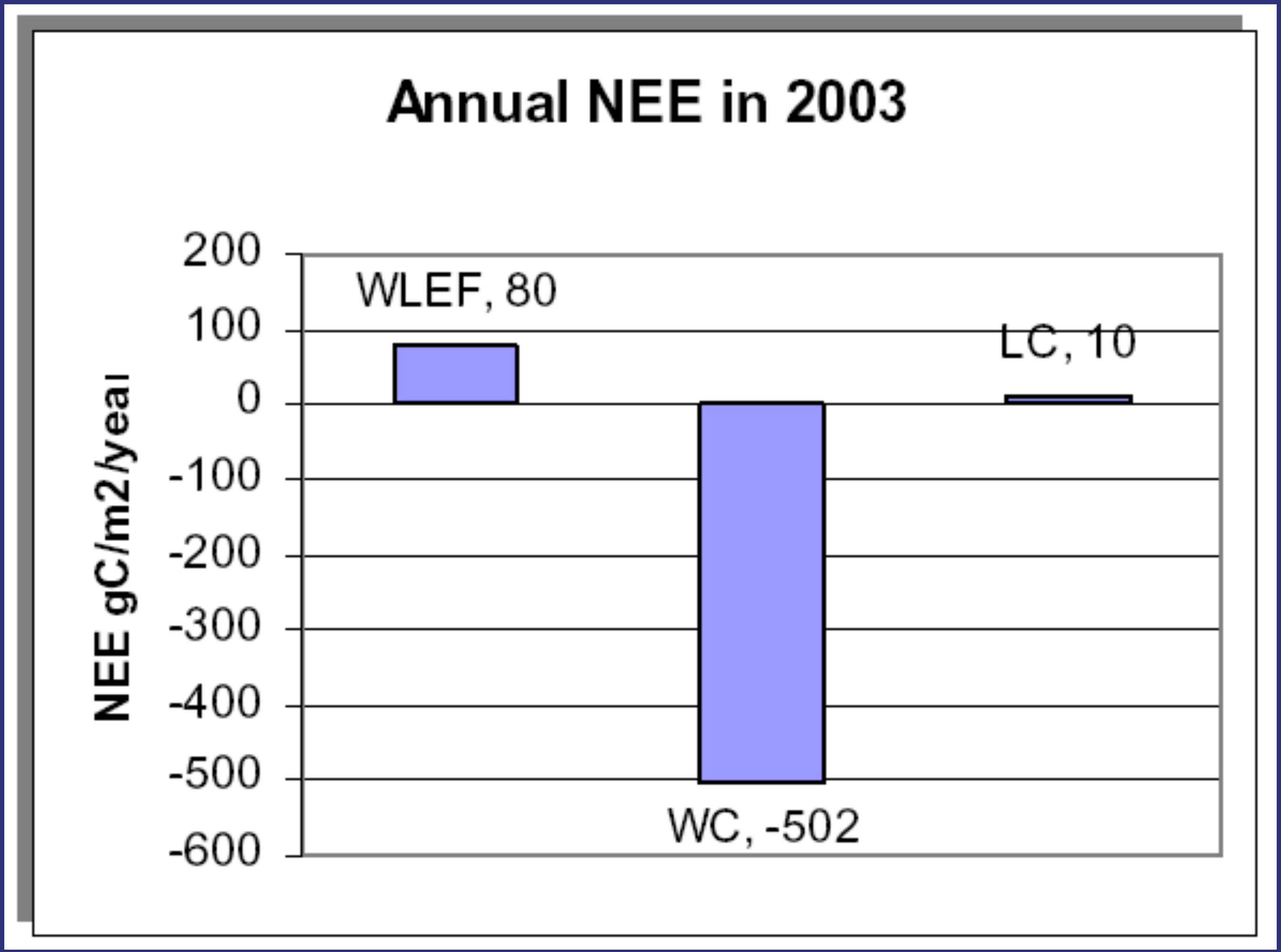
Our Tower is Bigger…



Is This the Regional Flux?



Not Sure



Annual NEE in 2003
NEE gC/m2/year
200
100
0
-100
-200
-300
-400
-500
-600
WLEF, 80
LC, 10
WC, -502

Lots of Variability

Flux Decomposition Recipe

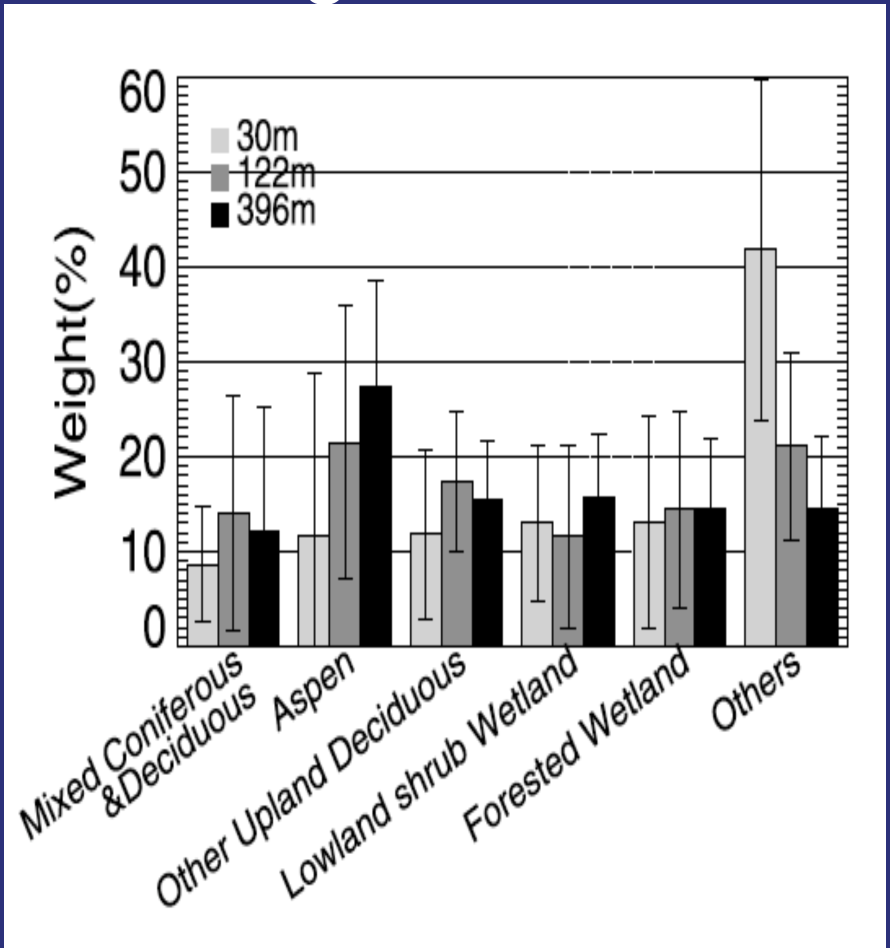
- 1 gridded spatial land cover map
- 1 spatial flux footprint model
- 1 time series of wind velocity, u*, heat flux
- 1 time series of NEE
- 1 time series of T, PAR and VPD
- 1 simple ecosystem model

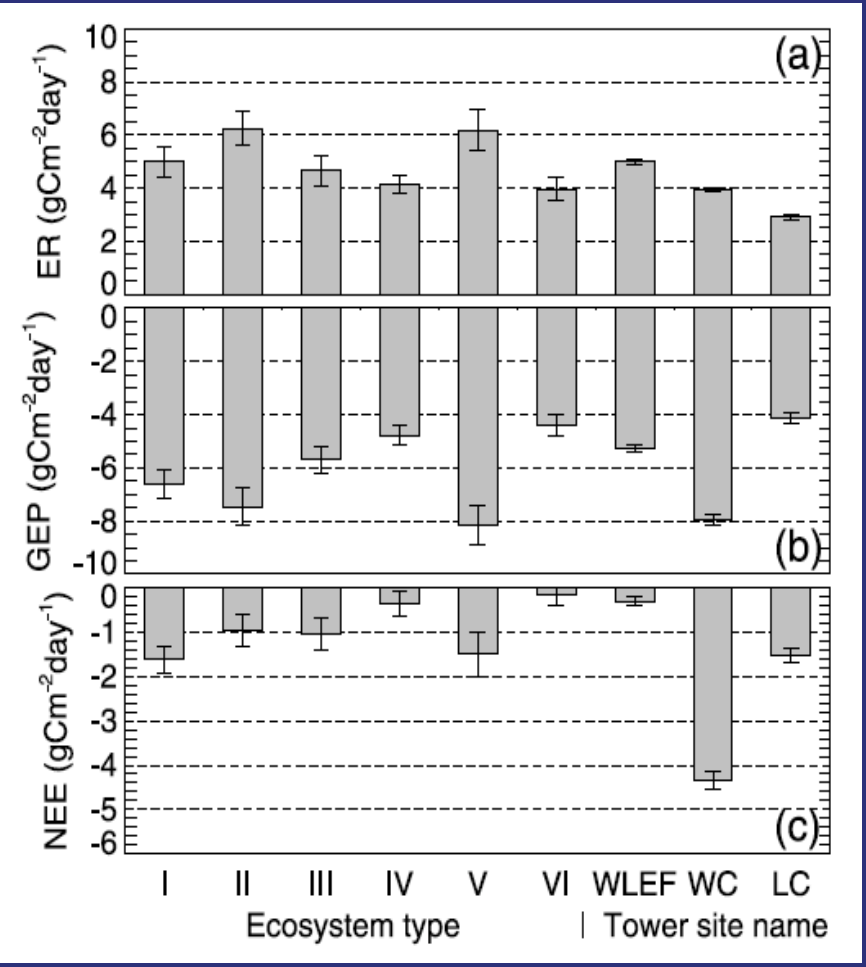
Flux Decomposition Recipe

- 1. Apply velocity, u*, and heat flux to footprint model
- 2. For each time step, derive land cover statistics within each footprint from the land cover map
- 3. Using land cover statistics, run a weighted spatial ecosystem model driven by T, PAR, and VPD
  - Model(cover1)*area1 + Model(cover2)*area2 + … = NEE
- 4. Use assimilation technique to estimate model parameters for each cover type based on tall tower time series of NEE and time varying estimates of proportional land cover at each time step

Voila!

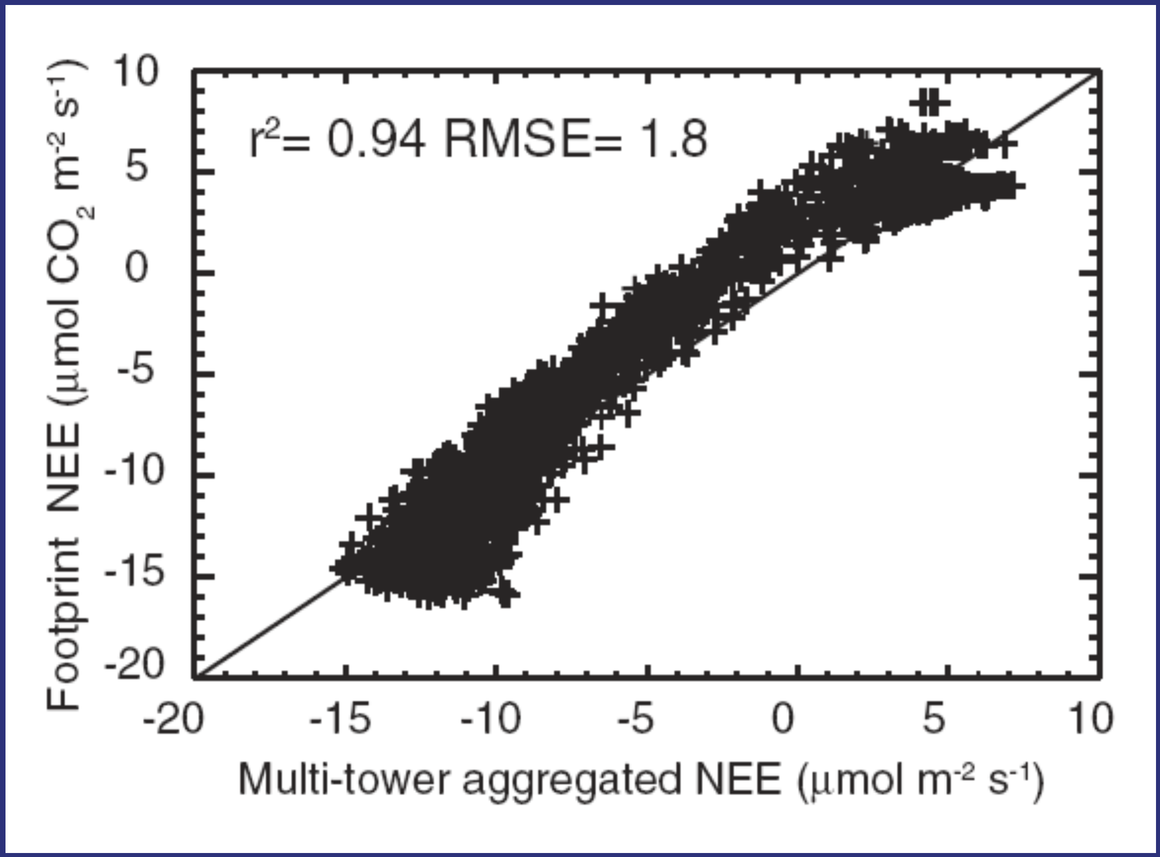
- Wang et al., 2006





Evaluation

- Desai et al., 2008



Enough?

What Did We Learn?

- Spatial prediction, scaling, parameterization all benefit from data assimilation
- Interannual variability has interesting spatial attributes that are hard to model
- You can't build infinite towers, or even a sufficient number
  - Use data assim. to discover optimal design?
- Spatial covariate and uncertainty information needs to be considered in data assimilation
  - "The only thing that makes life possible is permanent, intolerable uncertainty; not knowing what comes next." -- Ursula K. LeGuin

Where is Your Research Headed?

- What questions do you have?
  - Mechanisms, forcings, inference, evaluation, prediction, estimating error or uncertainty
- What kinds of data do you have, can get, can steal?
  - "Method-hopping"
- A model can mean many things…
- Data assimilation can be another tool in your toolbox to answer questions, discover new ones

Data Assimilation Uses

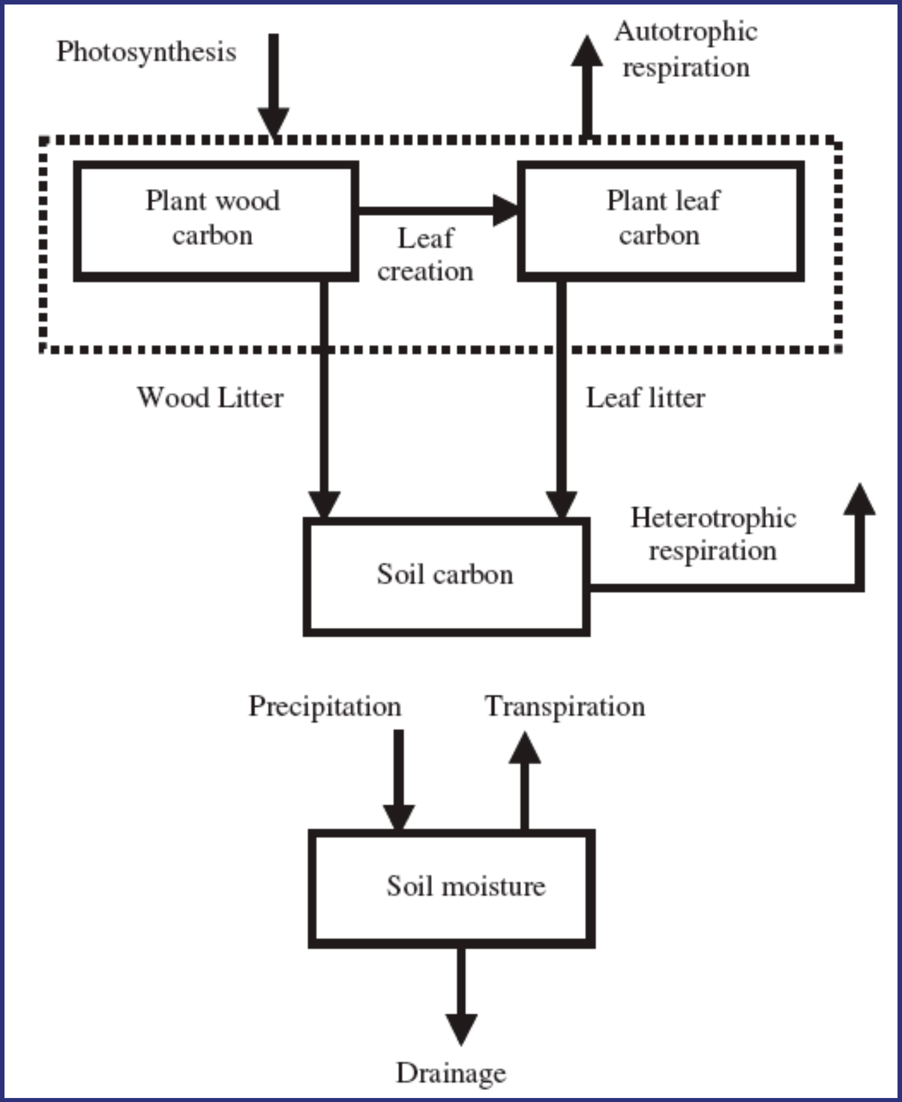
- Not just limited to ecosystem carbon flux models
- E.g. estimating surface or boundary layer values (e.g., z_0), advection, transpiration, data gaps, tracer transport
- Many kinds, for estimating state or parameters

Upcoming Lab Preview

- Sipnet at flux towers
- Parameter estimation with MCMC
- Group projects

Sipnet

- A “simplified” model of ecosystem carbon / water and land-atmosphere interaction
  - Minimal number of parameters
  - Driven by meteorological forcing
- Still has >60 parameters
  - Braswell et al., 2005, *GCB*
  - Sacks et al., 2006, *GCB*
  - Zobitz et al., 2008
  - Moore et al., 2008
  - Hu et al., 2009



Thanks

- Ankur R Desai
- desai@aos.wisc.edu
- http://flux.aos.wisc.edu