Mesoscale Data Assimilation for NACP

Scott Denning

Dusanka Zupanski

Marek Uliasz

Evolving in-situ CO_2 Network

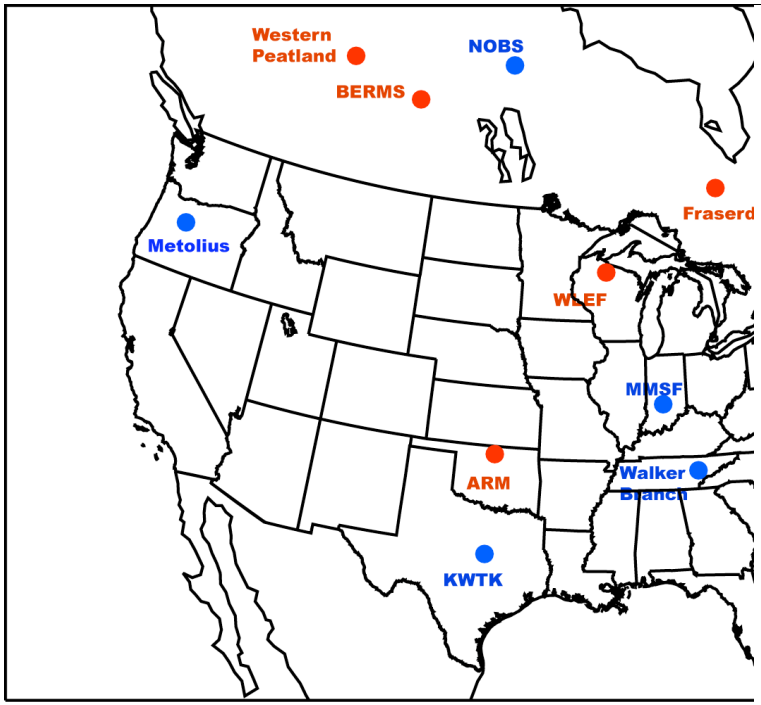


Figure 2: Continuous calibrated CO_2 mixing ratio sites to be us assimilation. Sites for which we currently have hourly data are which we anticipate data soon are in blue.

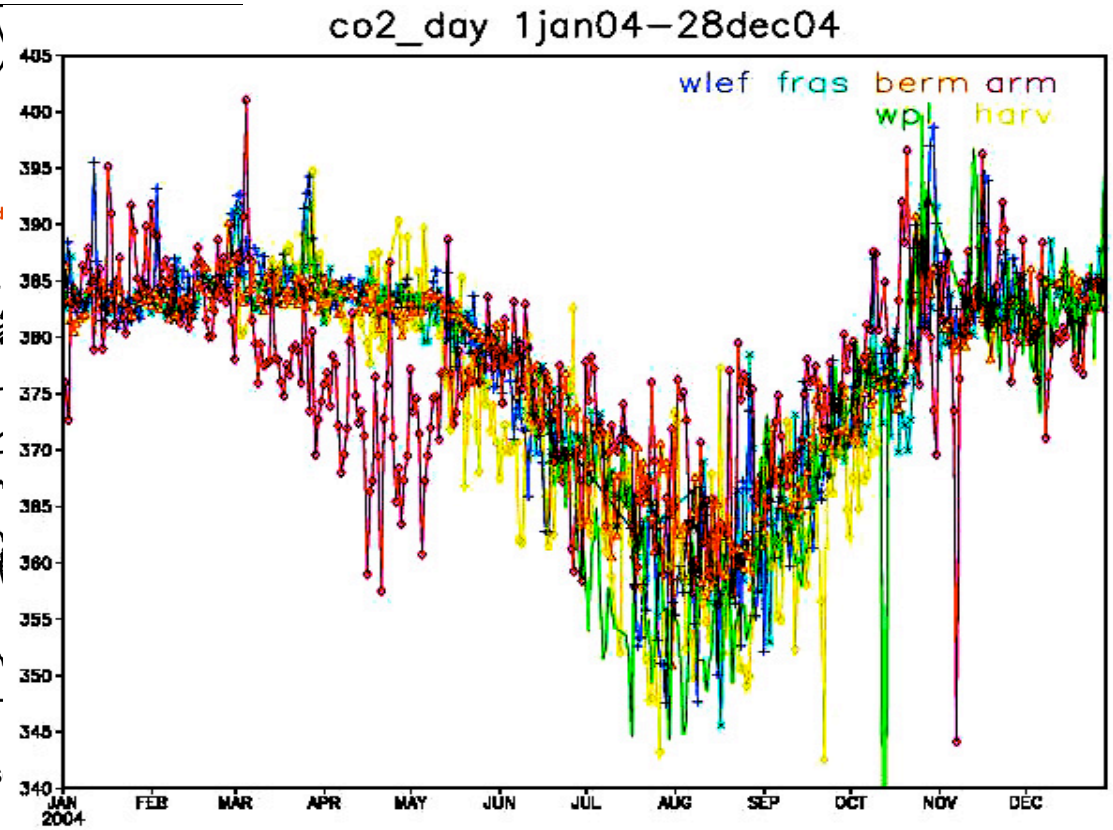
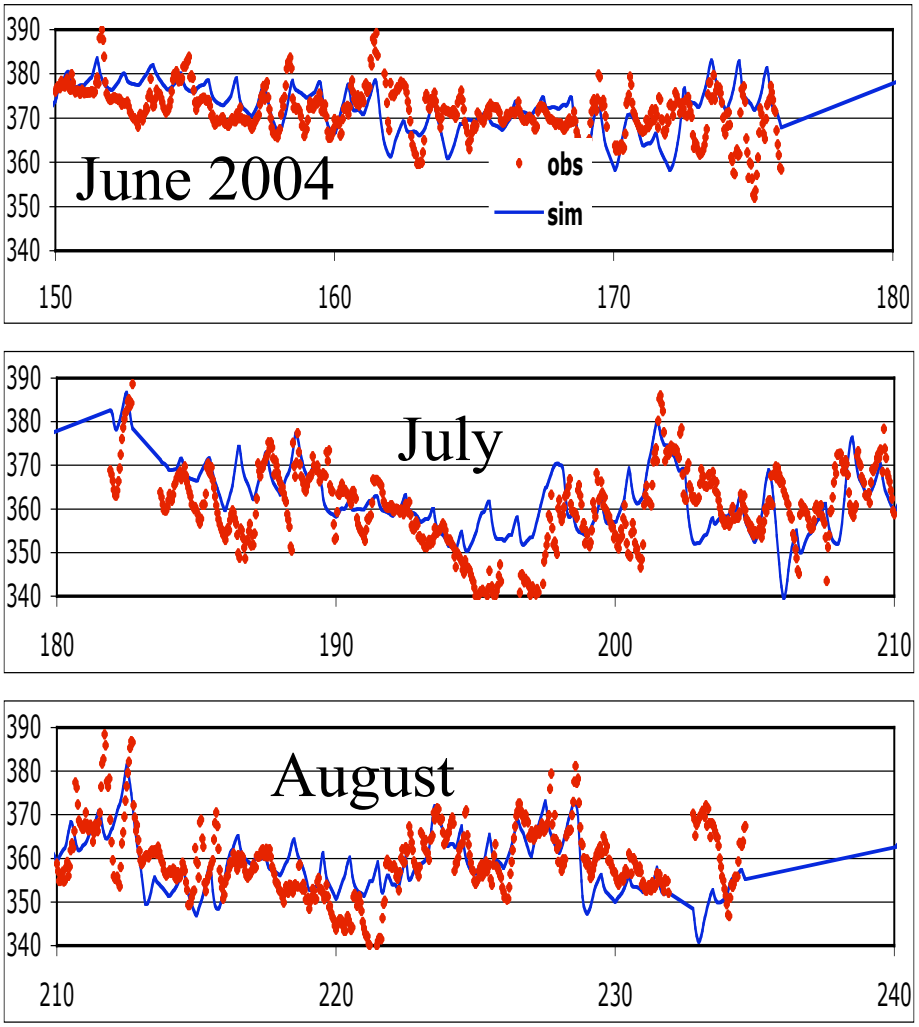


Figure 3: Daily minimum CO_2 mixing ratio for 6 observing sites in 2004.

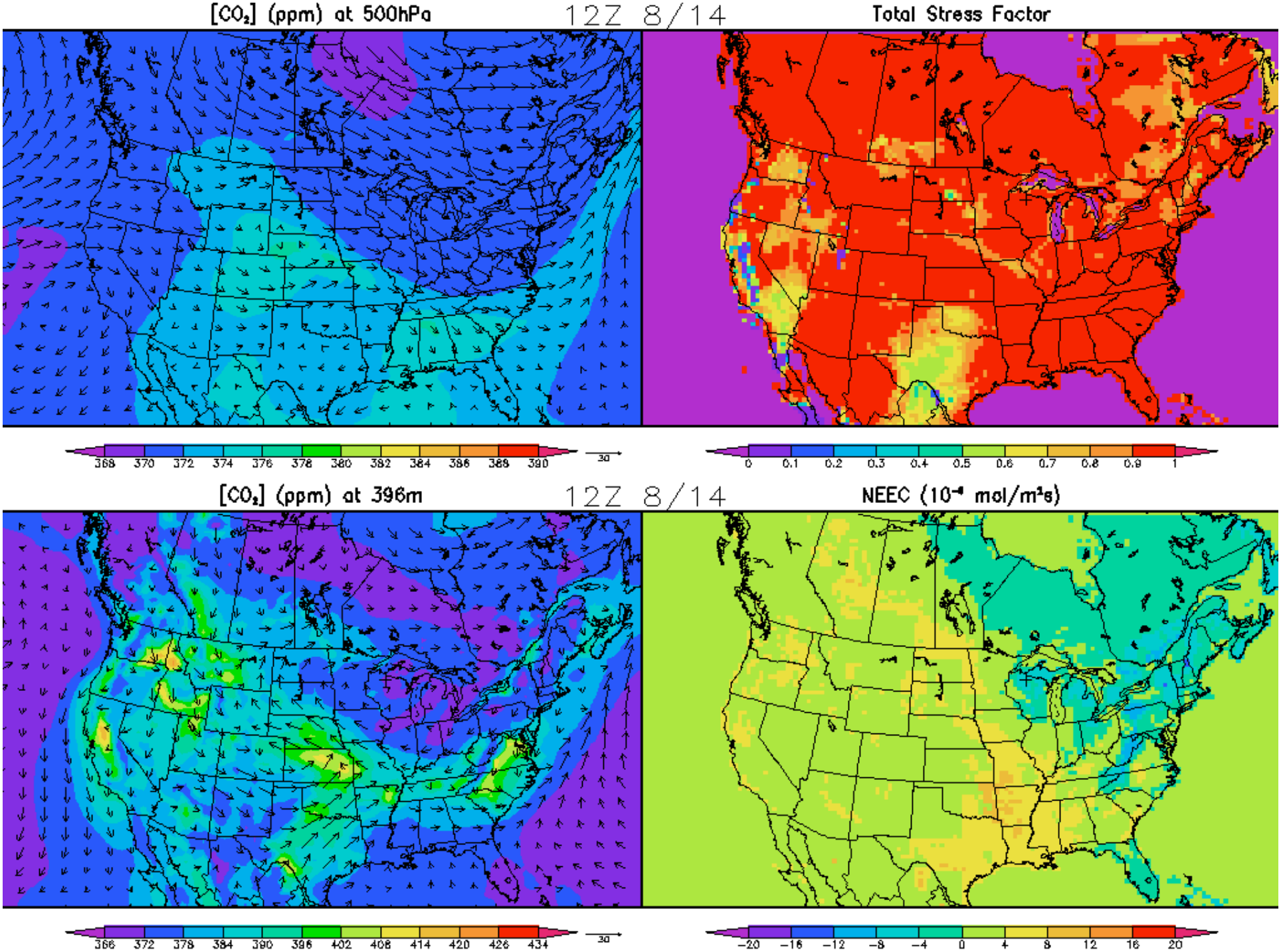
- 13 sites in 2005, maybe 20 by 2007 (NOAA?)
- Big synoptic variations and spatial gradients!

CO_2 Modeling in GEOS4

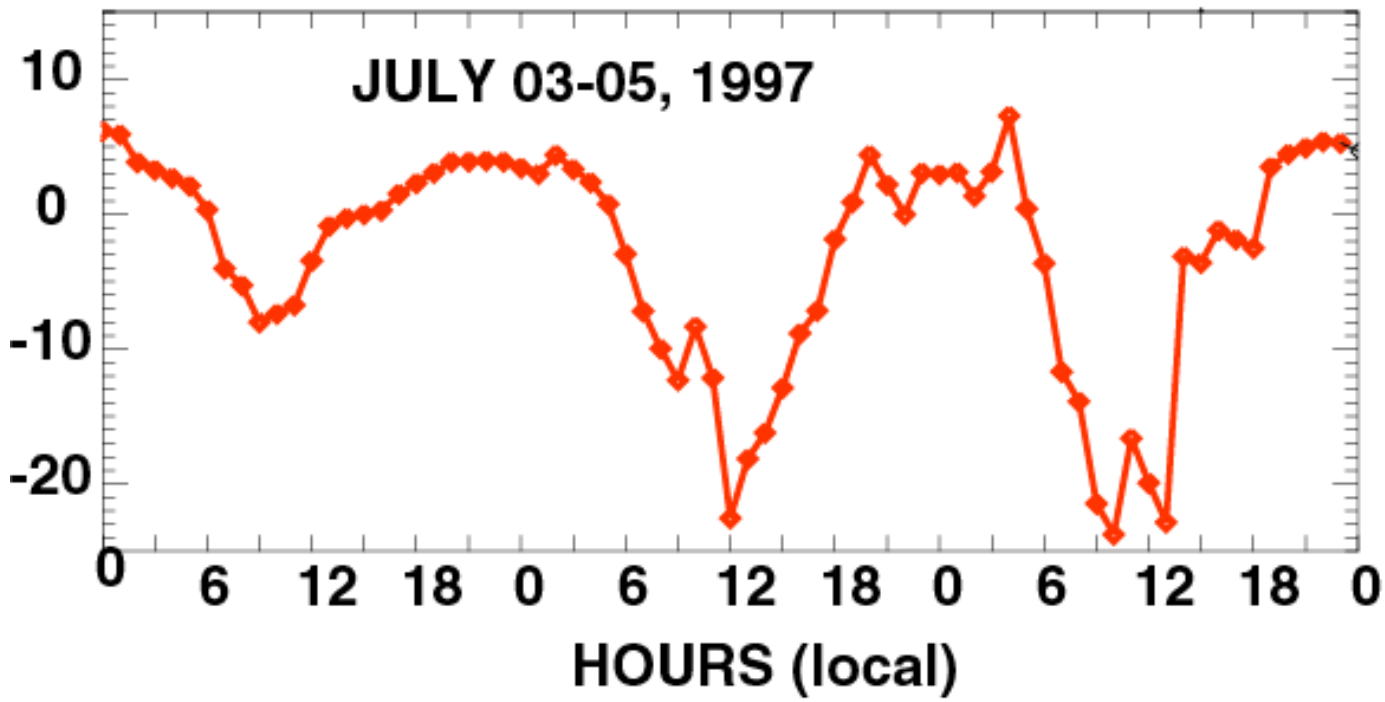


1° x 1.25 ° grid

- Offline coupling: SiB driven by NDVI and GEOS surface weather
- Hourly SiB fluxes prescribed as LBC
- Offline tracer transport using PCTM driven by GEOS winds, convection, and turbulence
- Reasonable variability at seasonal, diurnal, and synoptic time scales (r^2 = .66)


[CO2] (ppm) at 500hPa
12Z 8/14
Total Stress Factor
368 370 372 374 376 378 380 382 384 386 388 390
30
0 0.1 0.2 0.3 0.4 0.5 0.6 0.7 0.8 0.9 1
[CO2] (ppm) at 396m
12Z 8/14
NEEC (10⁻⁴ mol/m²s)
366 372 378 384 390 396 402 408 414 420 426 434
30
−20 −16 −12 −8 −4 0 4 8 12 16 20

Measured NEE of CO_2 (WLEF)



- Coherent diurnal cycles, but …
- Day-to-day variability of ~ factor of 2 due to passing weather disturbances
- How to specify temporal autocorrelation in inversions?

CO_2 Asssimilation

$$F_{CO2}(x,y,t) = \beta_R(x,y)R(x,y,t) - \beta_A(x,y)A(x,y,t))$$

$$C_{k,m} = \sum_{i,j,n}\left(\beta_{R,i,j}R_{i,j,n}C^*_{k,m,i,j,n} + \beta_{A,i,j}A_{i,j,n}C^*_{k,m,i,j,n}\right)\Delta t_f \Delta x \Delta y + C_{IN,k,m}$$

where $k = tower\ \ number$

$m = sampling\ time$

$n = flux\ release\ time$

$i, j = source\ grid\ cells$

- Fine-scale variations (hourly, 20-km pixels) from weather forcing, NDVI as processed by forward model logic
- Multiplicative biases (10-day, 100-km pixels) derived by inversion of hourly [CO_2]

Data Assimilation Set-Up

Discrete stochastic-dynamic model

$$\mathbf{m}: \quad x_k = M(x_{k-1}) + G(x_{k-1})w_{k-1}$$

w_{k-1} – **model error (stochastic forcing)**

M – **non-linear dynamic (RAMS-SiB-BGC) model**

G – **model (matrix) reflecting the state dependence of model error**

Discrete stochastic observation model

$$\mathbf{D}: \quad y_k = H(x_k) + \varepsilon_k$$

ε_k – **measurement + representativeness error**

H – **non-linear observation operator (M → D)**

Dusanka Zupanski, CIRA/CSU
Zupanski@CIRA.colostate.edu

DATA ASSIMILATION EQUATIONS:

(1) State estimate (optimal solution):

MAXIMUM LIKELIHOOD ESTIMATE (VARIATIONAL APPROACH):

$$J = \frac{1}{2}[\boldsymbol{x} - \boldsymbol{x}_b]^T (\boldsymbol{P}^f)^{-1} [\boldsymbol{x} - \boldsymbol{x}_b] + \frac{1}{2}[H(\boldsymbol{x}) - \boldsymbol{y}_{obs}]^T \boldsymbol{R}^{-1} [H(\boldsymbol{x}) - \boldsymbol{y}_{obs}] = \min$$

MINIMUM VARIANCE ESTIMATE (KALMAN FILTER APPROACH):

$$\boldsymbol{x} = \boldsymbol{x}_a = \boldsymbol{x}_b + \boldsymbol{P}\boldsymbol{H}^T (\boldsymbol{H}\boldsymbol{P}\boldsymbol{H}^T + \boldsymbol{R})^{-1} (\boldsymbol{y}_{obs} - \boldsymbol{H}\boldsymbol{x}_b)$$

(2) Estimate of the uncertainty of the solution:

$$\boldsymbol{P}^f = \boldsymbol{M}\boldsymbol{P}^a \boldsymbol{M}^T + \boldsymbol{G}\boldsymbol{Q}\,\boldsymbol{G}^T$$

KALMAN FILTER APPROACH

$$(\boldsymbol{P}^f)^{i,j} = [M(\boldsymbol{x} + \boldsymbol{p}_i) - M(\boldsymbol{x})][M(\boldsymbol{x} + \boldsymbol{p}_j) - M(\boldsymbol{x})]^T$$

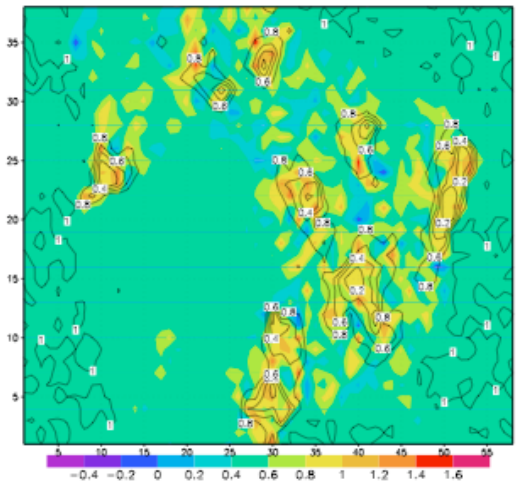
ENSEMBLE KALMAN FILTER or EnsDA APPROACH

In EnsDA solution is defined in ensemble subspace (reduced rank problem) !

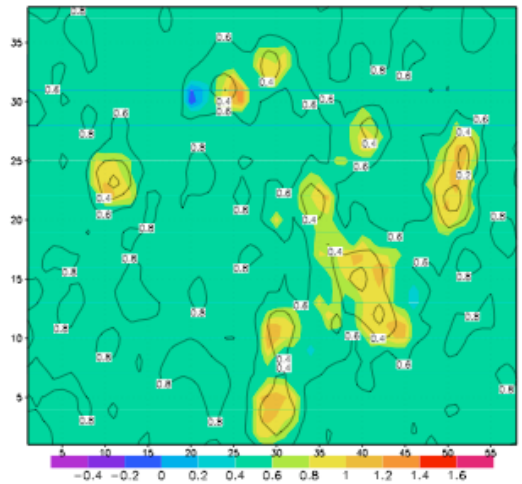
Mesoscale Ensemble CDAS

(synthetic data experiment)

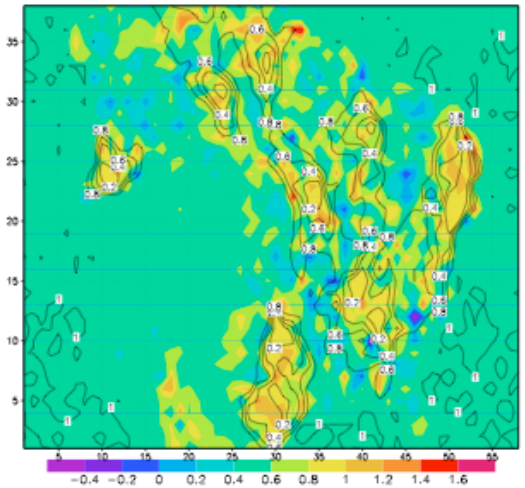
β_{GPP} N=4408



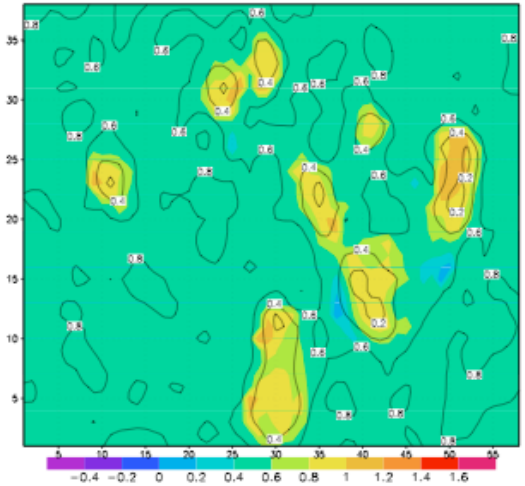
β_{GPP} N=100



β_R N=4408

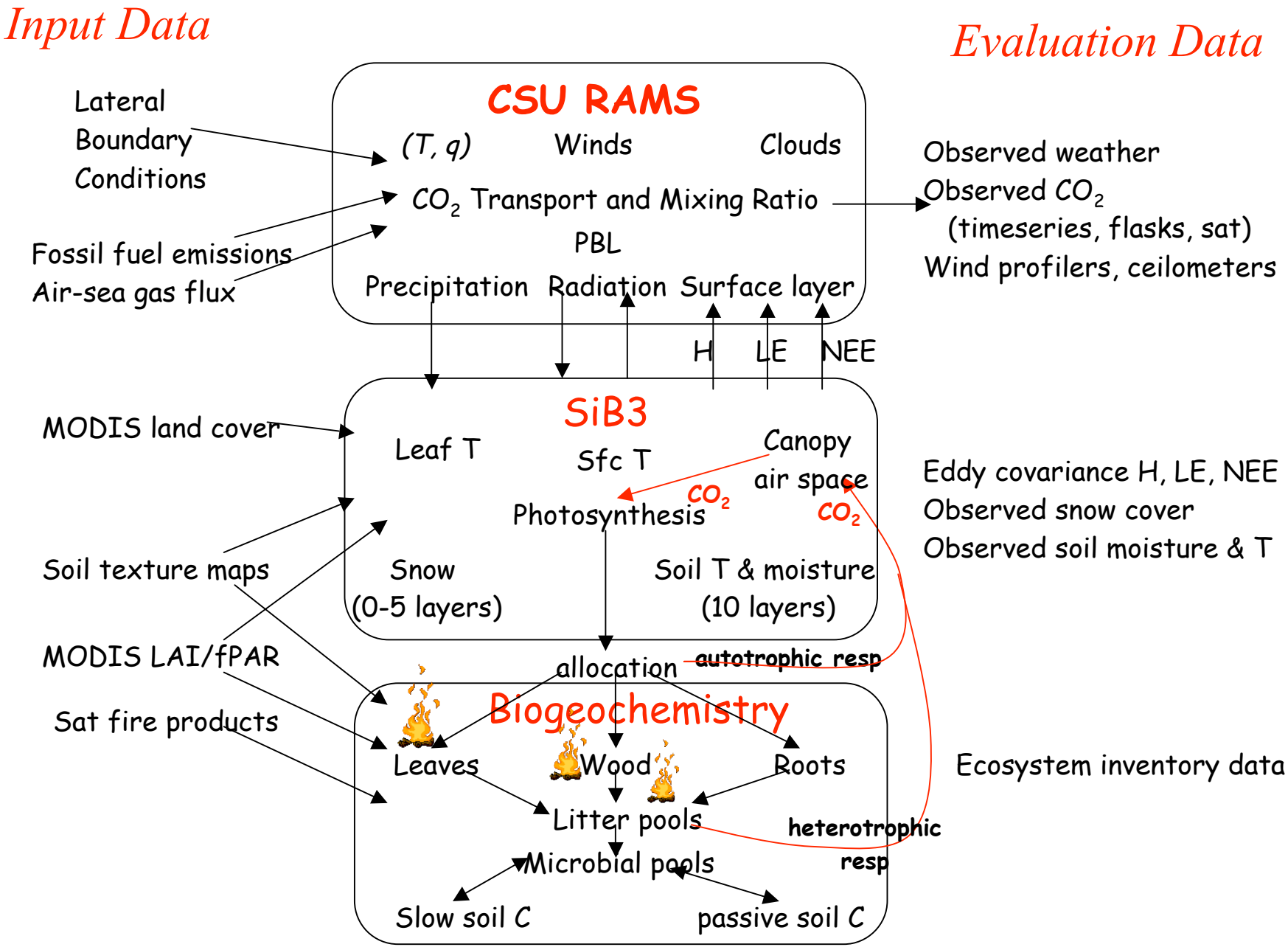


β_R N=100



- SiB-RAMS weather, GPP, respiration, atmospheric [CO_2]
- Sample model atmosphere at 11 towers each hour
- Estimate biases in GPP and resp every 10 days on 100 km grid (5800 x 3800 km domain)
- Smaller ensembles ~ stable w/ covariance localization and 5 Δx smoothing

Coupled Modeling and Assimilation System

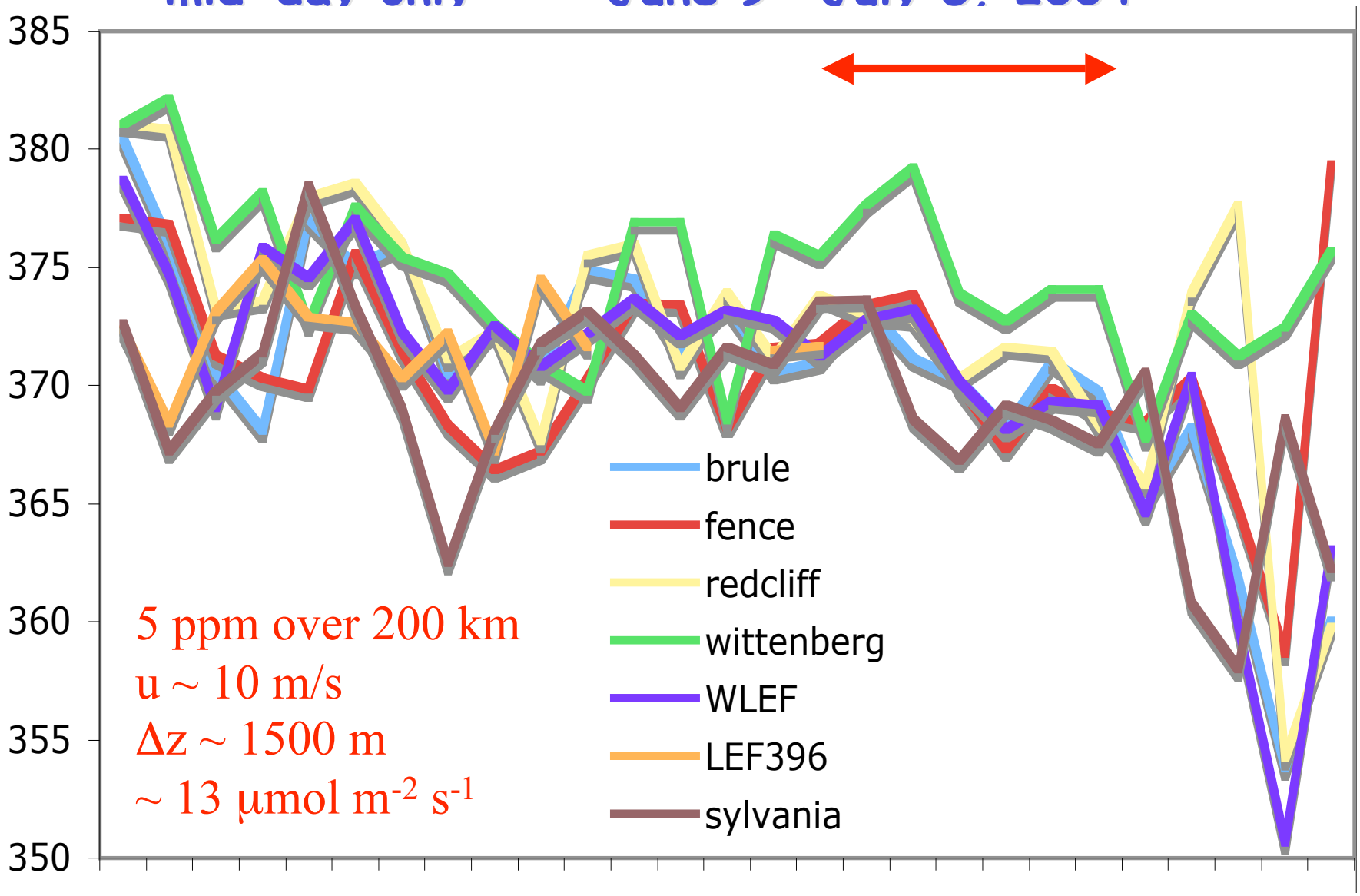


Ring of Towers

- inexpensive instruments deployed on six 75-m towers in 2004
- ~200 km radius
- 1-minute data May-August

Ring of Towers Data

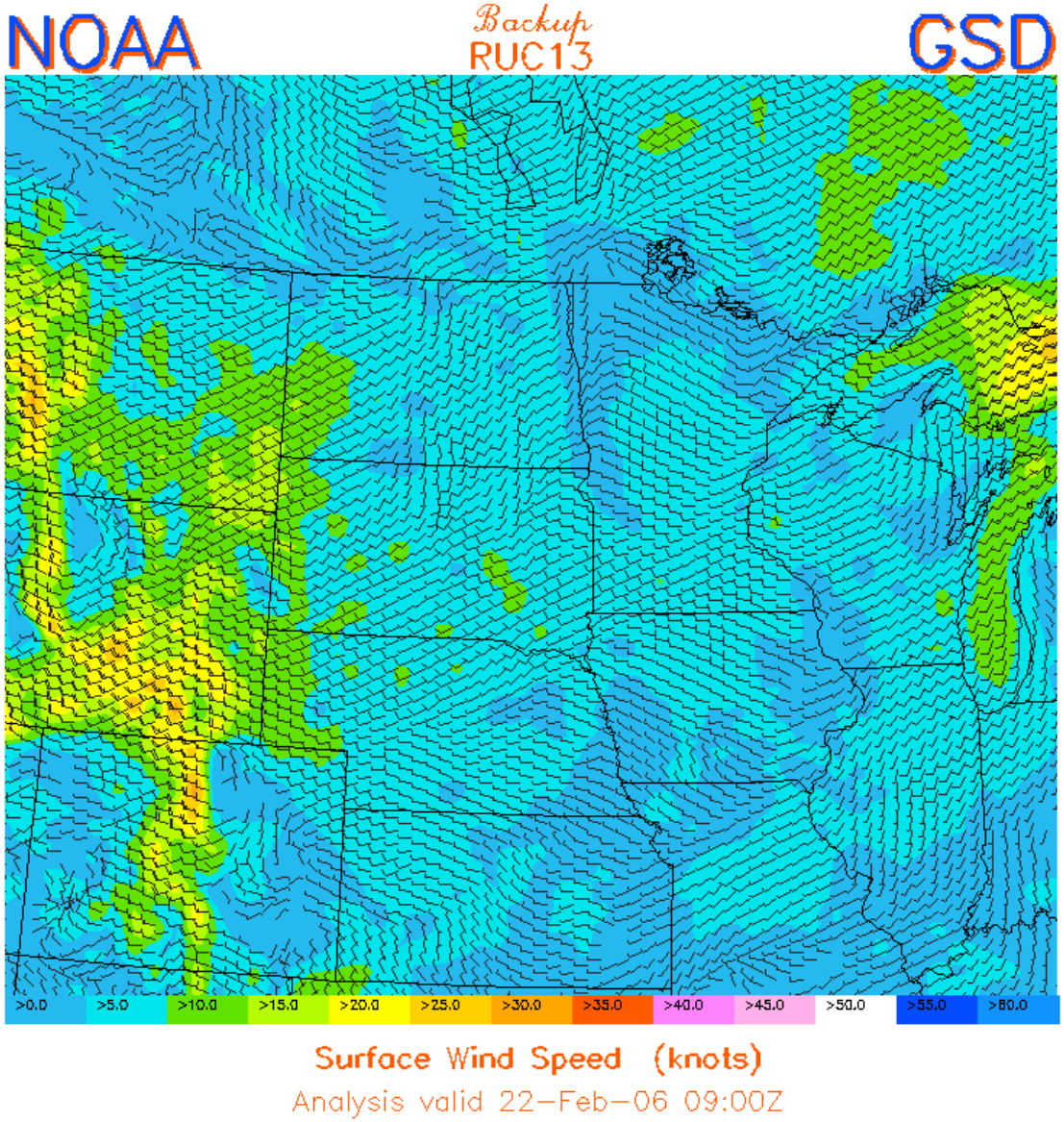
mid-day only June 9- July 5, 2004



Mid-Cont Intensive Plans?

- Technical capability to do high-resolution nested-grid simulations or assimmilations for the MCI
  - E.g., 10 km weather, NEE, CO2 over MCI domain during growing season
  - cloud-resolving (1-km) for ~ a week
- We are not funded to do this, but will propose it to DOE TCO
- Detailed comparison to towers, flux aircraft, MINT airborne data

Footprints from RUC-13



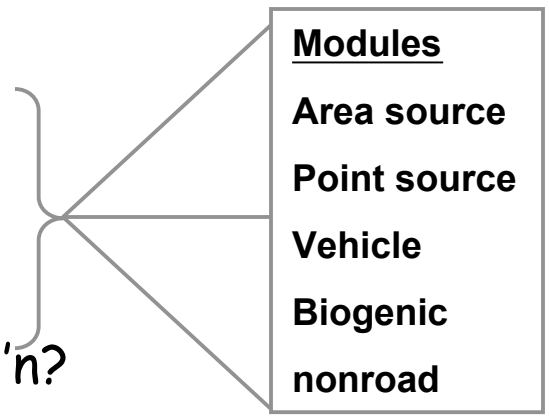
$\forall$ Δx = 13 km,
Δt = 1 hr

- Archiving u, v, w, cloud flux, TKE every hour for whole domain
- LPDM for NACP towers
- Infl function generation via web interface available to all by 2007

Process-Based Fossil Fuel Emissions
(K. Gurney, PI)

- Combine inventory data and process attributes to construct detailed space and time dependent emissions of fossil fuel CO_2 and CO
- Database/model has three classes of inputs:
  - Point sources (e.g., power plants, smelters)
  - Mobile sources (vehicle emissions)
  - Area sources (e.g., residential sources)
- Resolution: 36 km, hourly

  downscale from roads, points, lights, pop'n?



- Optimize parameters in emissions model(s) using transport in RAMS and observed [CO]