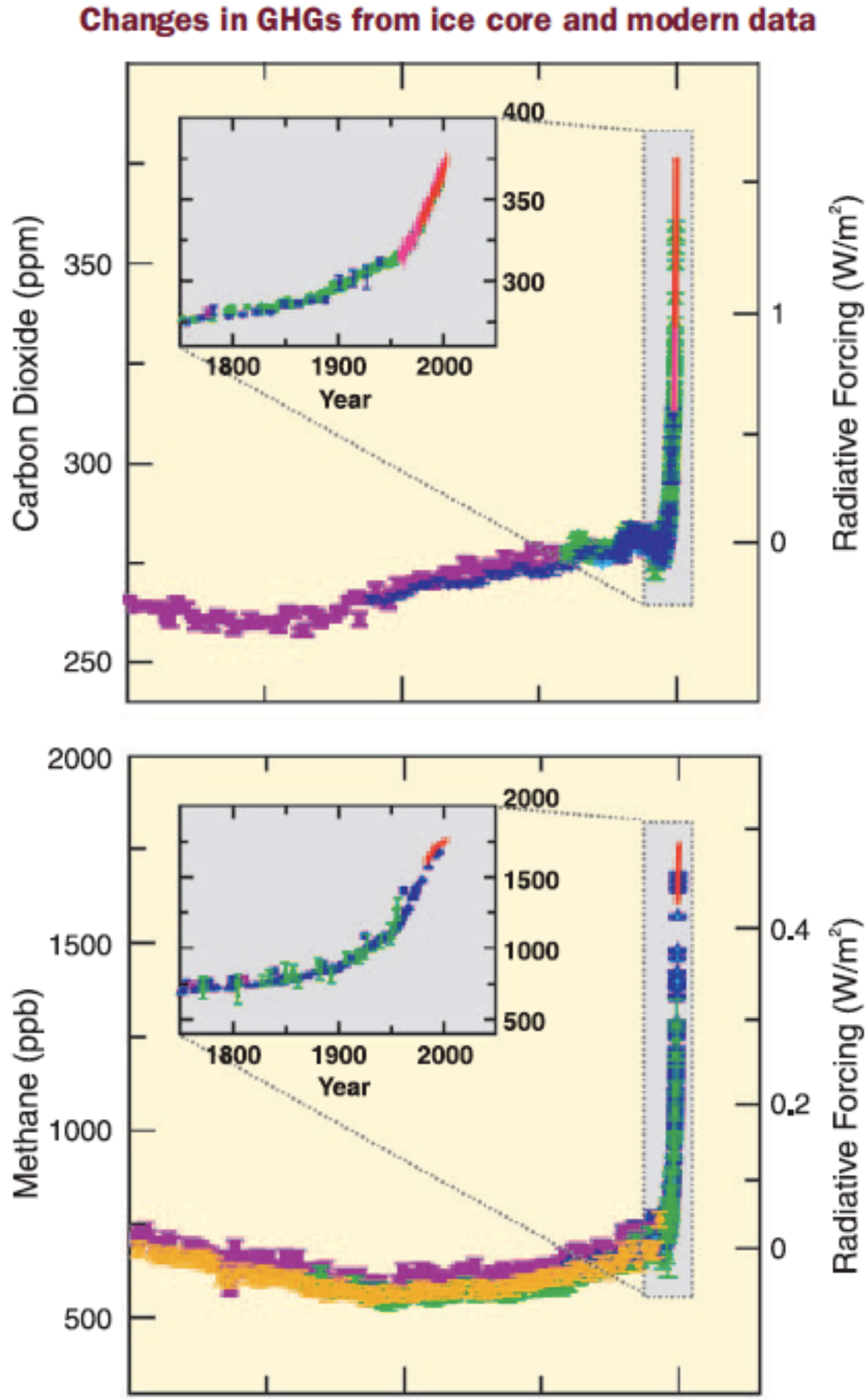
PennState

Big data, big uncertainties and big challenges in carbon cycle science.

Kenneth J. Davis
Professor of Atmospheric and Climate Science
Department of Meteorology and Atmospheric Science and
Earth and Environmental Systems Institute
The Pennsylvania State University

Symposium on Advanced Assimilation and Uncertainty Quantification in Big Data Research for Weather, Climate and Earth System Monitoring and Prediction
State College, Pennsylvania, 23 May, 2016

Carbon cycle science: What's the big deal?



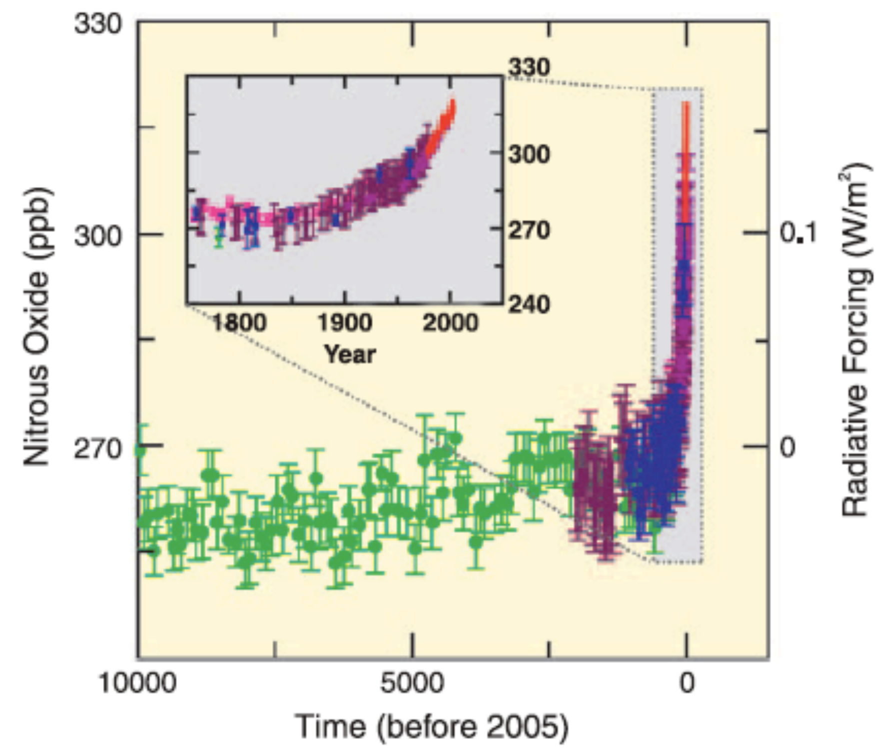
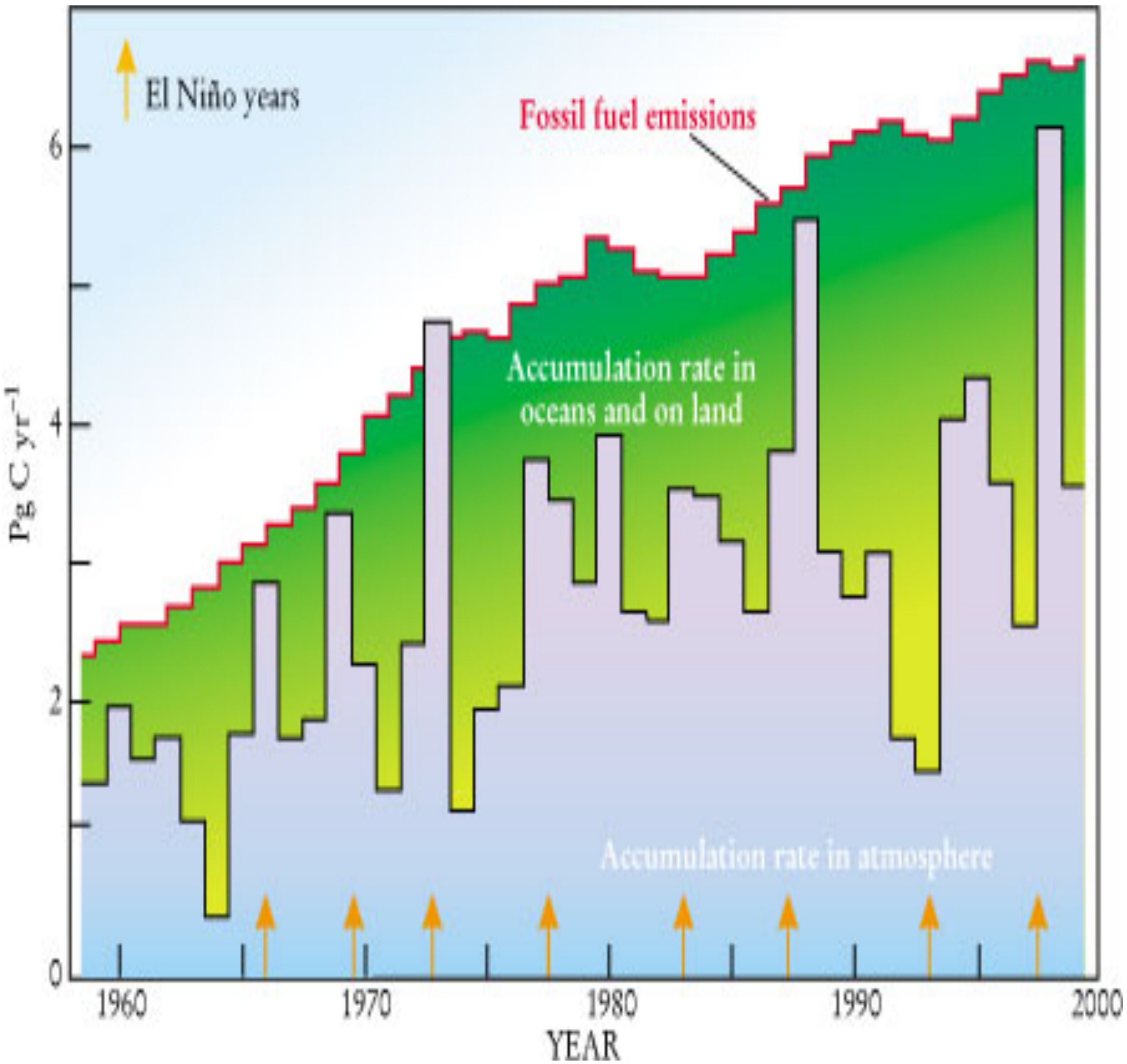


Figure 2.3. *Atmospheric concentrations of CO_2, CH_4 and N_2O over the last 10,000 years (large panels) and since 1750 (inset panels). Measurements are shown from ice cores (symbols with different colours for different studies) and atmospheric samples (red lines). The corresponding radiative forcings relative to 1750 are shown on the right hand axes of the large panels. {WGI Figure SPM.1}*

Ten-thousand year view

IPCC, WG1, AR4

Terrestrial ecosystems are removing large quantities of CO_2 from the atmosphere.

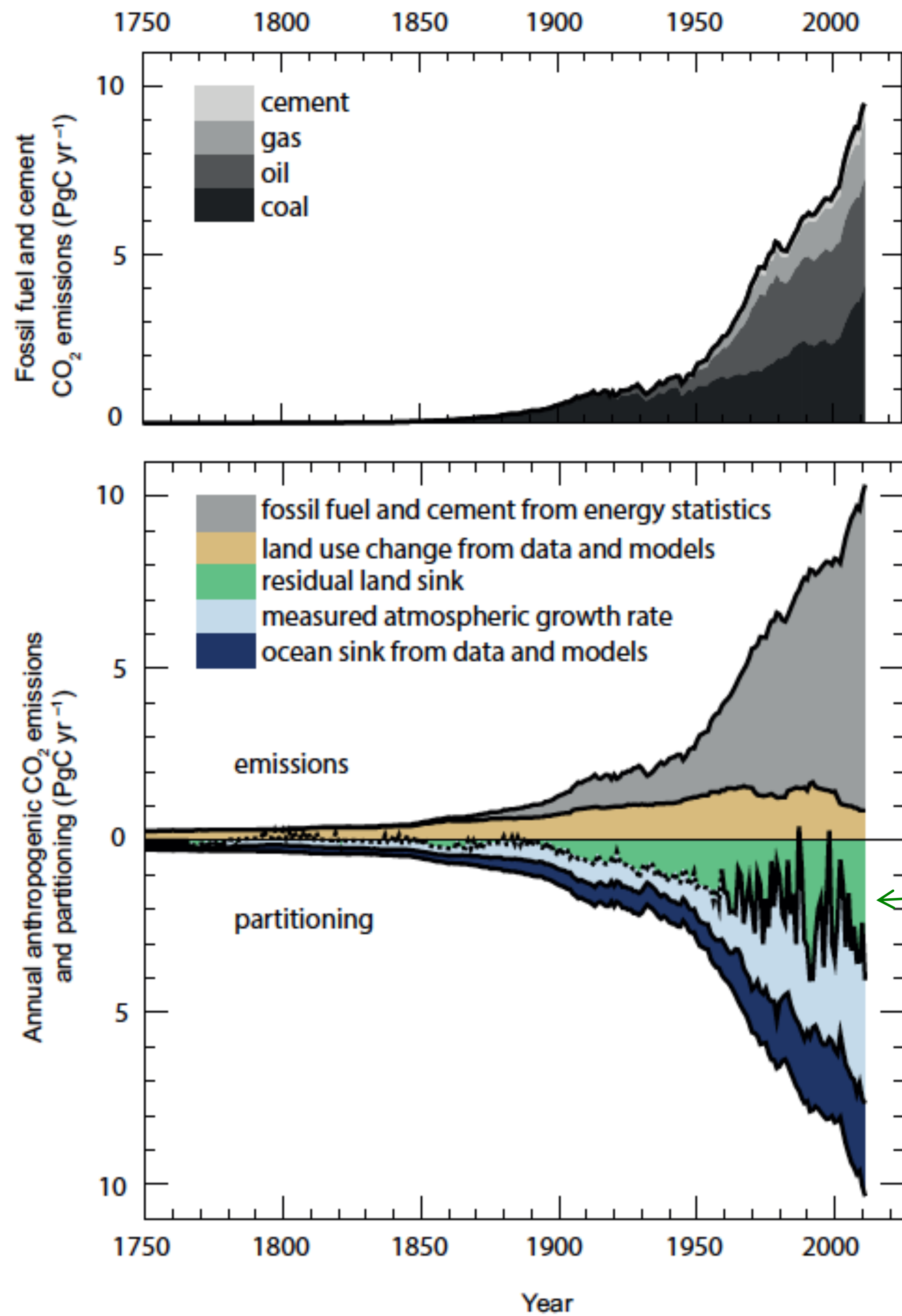


http://www.aip.org/pt/vol-55/iss-8/captions/p30cap2.html

- Terrestrial (and marine) systems are removing a lot of CO_2!
- The terrestrial sink is increasing with time
- The terrestrial sink has large interannual variability, likely related to climate variability.
- *Where is this happening? Why is this happening?*

(Global data – atmospheric sampling.)

Sarmiento and Gruber, 2002



1. Fossil fuel burning is a huge CO_2 source. Must be managed.
2. There is a large terrestrial biosphere sink that is poorly understood and highly variable.

Energy stats

Land use records

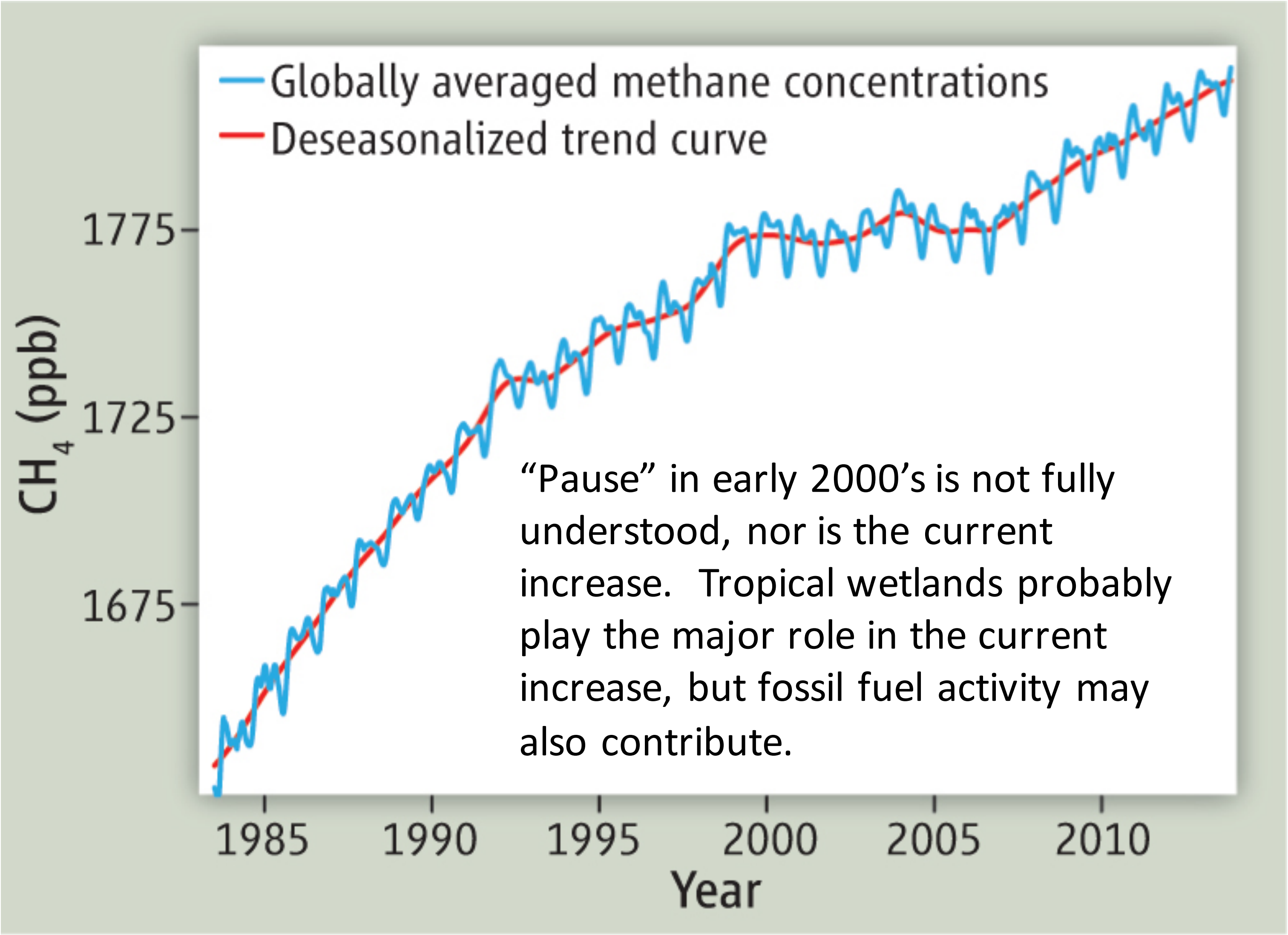
Residual terrestrial sink

Atmospheric CO_2 measurements

Ocean and atmospheric data

IPCC, WG1, AR5

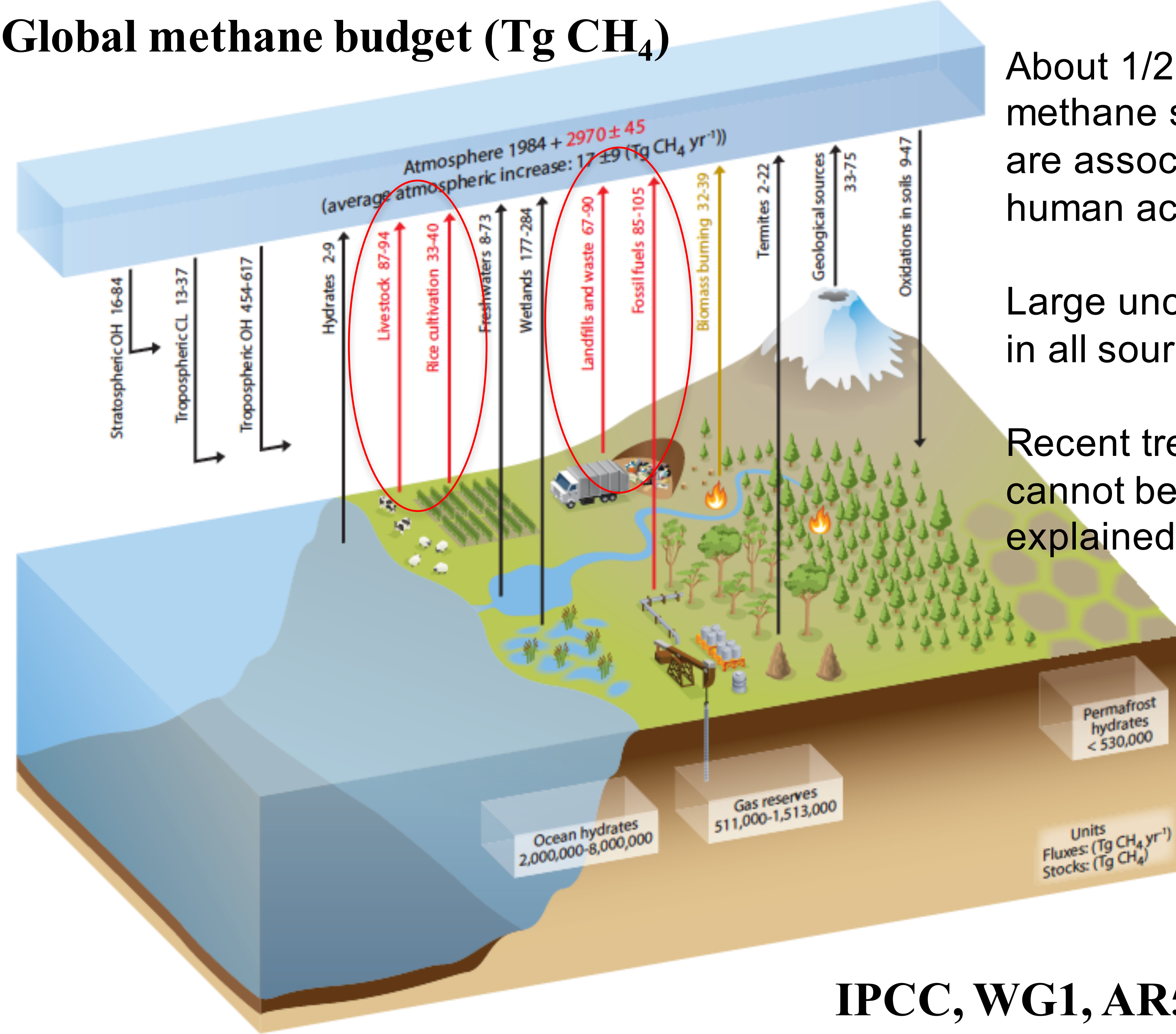
Recent years: Methane concentrations are “on the rise again”



Euan G. Nisbet et al. Science 2014;343:493-495

Published by AAAS

Global methane budget (Tg CH_4)



About 1/2 of methane sources are associated with human activity.

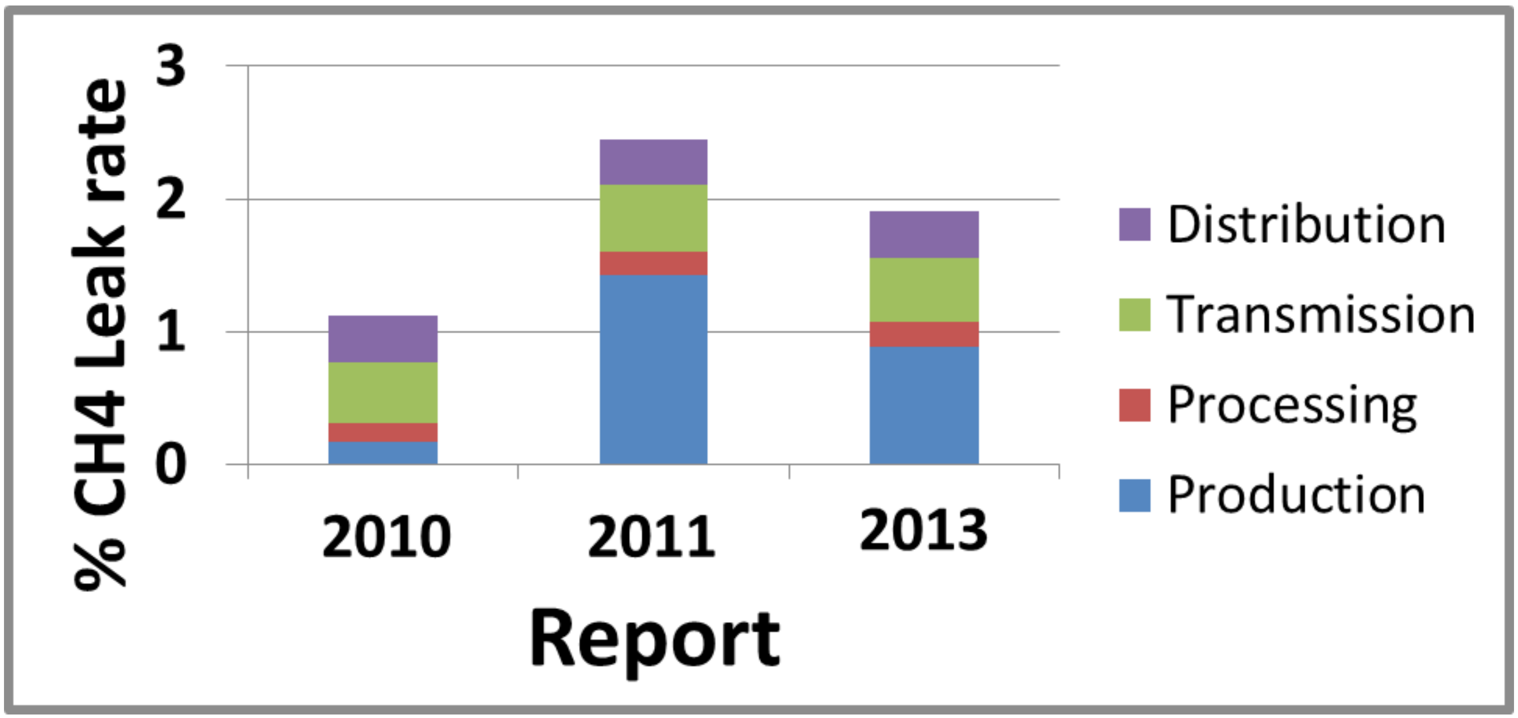
Large uncertainty in all sources.

Recent trends cannot be explained.

IPCC, WG1, AR5, Fig 6.2

Example: Uncertainty in CH_4 emissions from the production of natural gas is large

- Changes in estimated methane leakage as a percentage of production (USEPA 2010, 2011, 2013) are dominated by *changing estimates* of leakage during production (other sectors may also be uncertain – just not revised in these reports).



Diagnoses of current carbon sources/sinks are not very accurate.

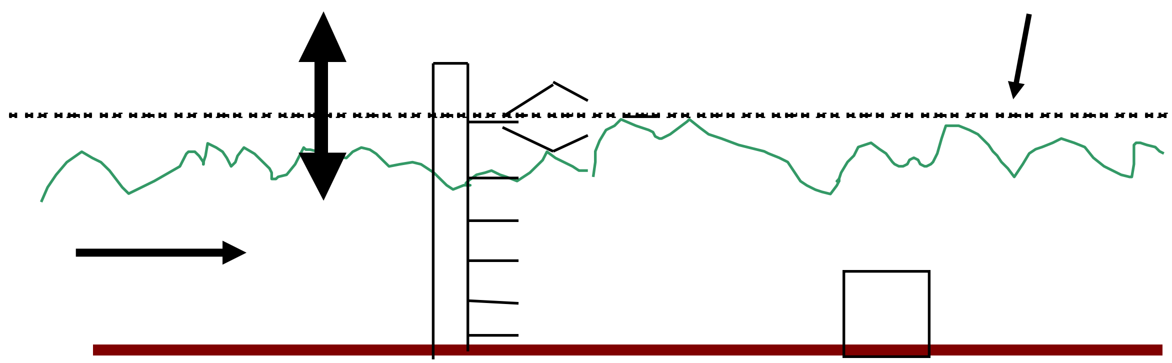
(at “regional” spatial scales)

and “we can’t manage what we can’t measure.”

Methods

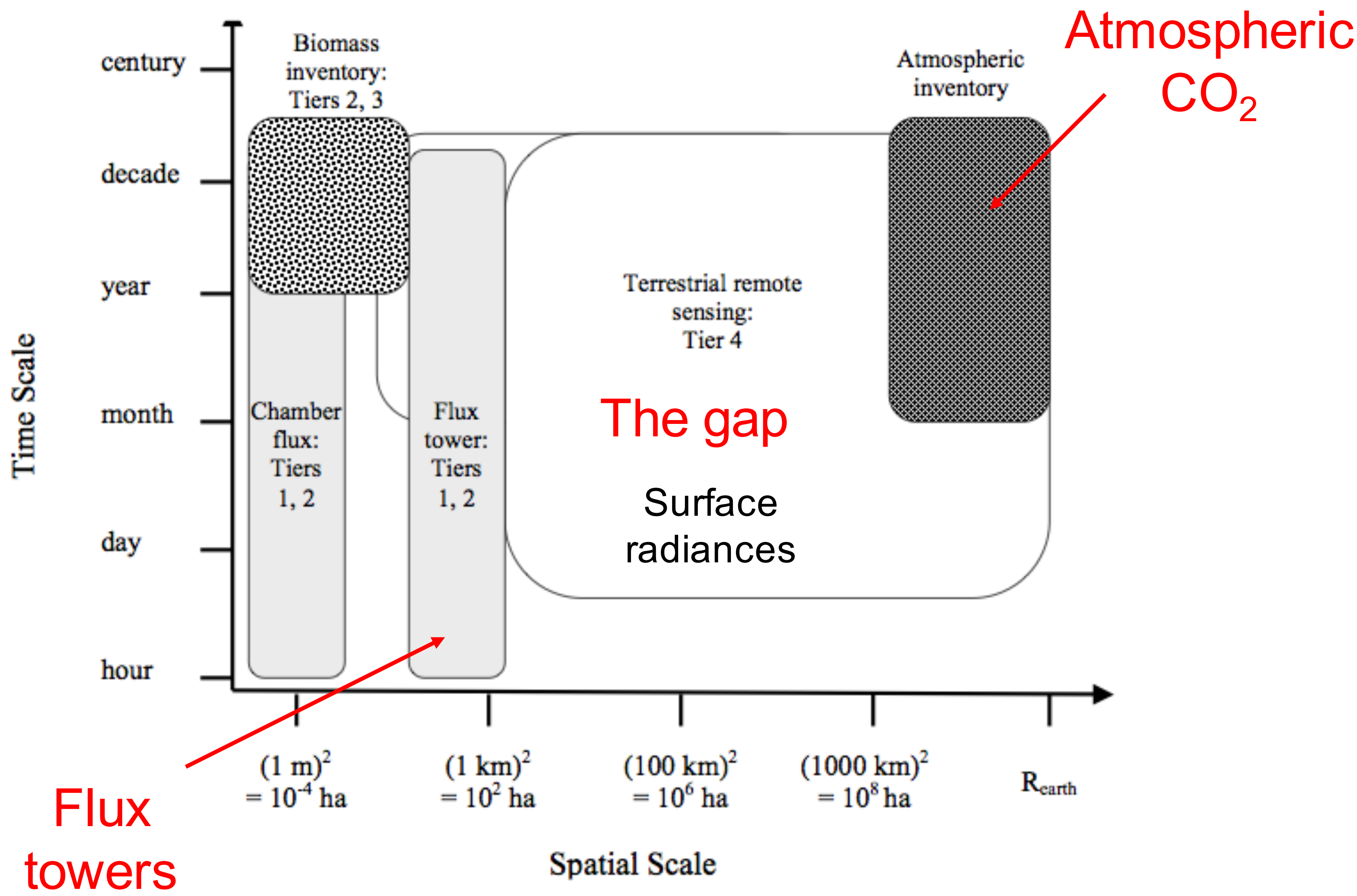
Change in atmospheric concentration of CO_2 over time = inversion or ABL budget approach.

Flux of carbon across this plane = tower or aircraft flux approach



Change in biomass over time = inventory approach

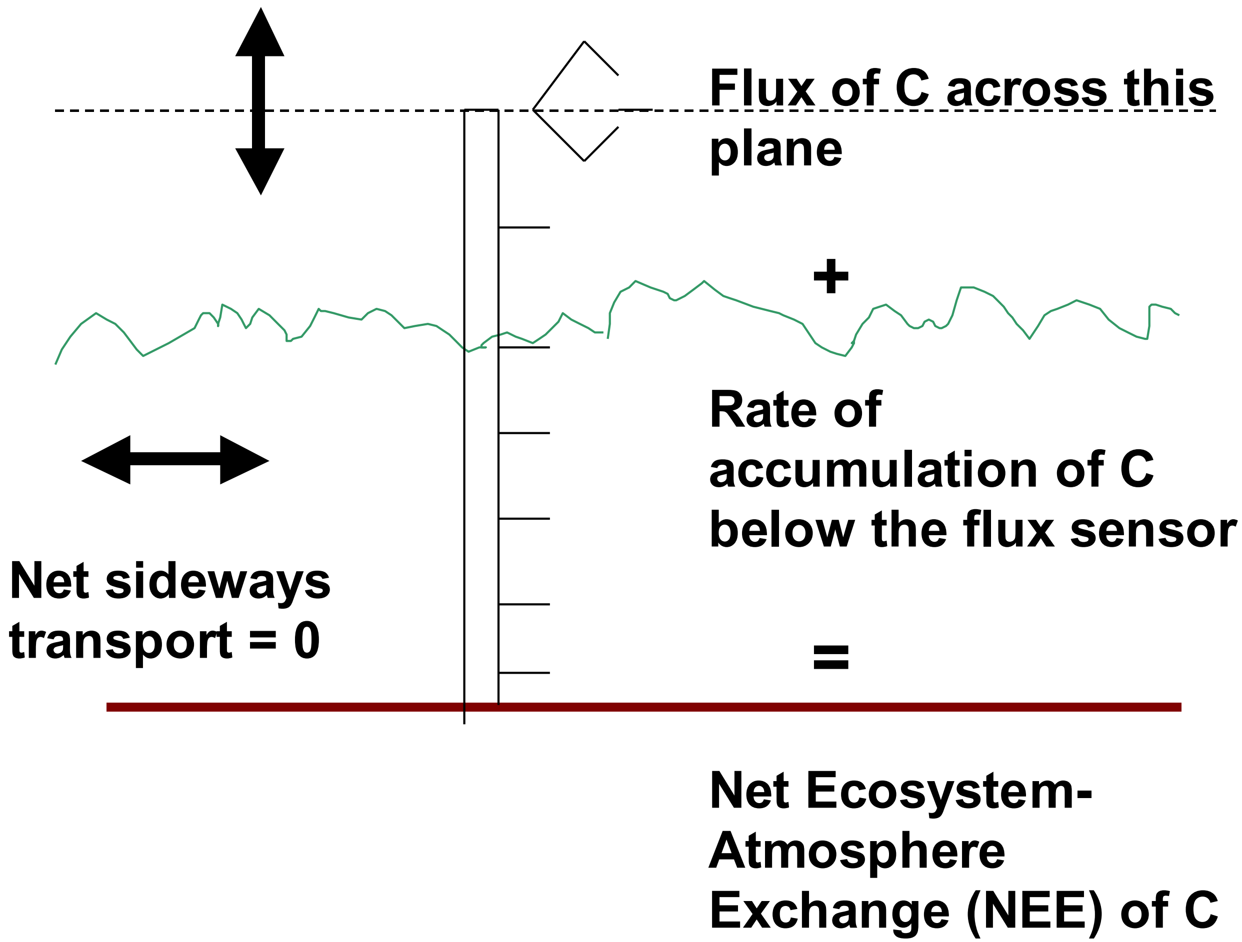
Change in CO_2 concentration in a small box over time = chamber flux approach



Atmospheric CO_2
century
decade
year
month
day
hour
Time Scale
Biomass inventory: Tiers 2, 3
Atmospheric inventory
Chamber flux: Tiers 1, 2
Flux tower: Tiers 1, 2
Terrestrial remote sensing: Tier 4
The gap
Surface radiances
$(1\ m)^2 = 10^{-4}$ ha
$(1\ km)^2 = 10^2$ ha
$(100\ km)^2 = 10^6$ ha
$(1000\ km)^2 = 10^8$ ha
R_{earth}
Spatial Scale
Flux towers

Davis, 2008

Method – eddy covariance

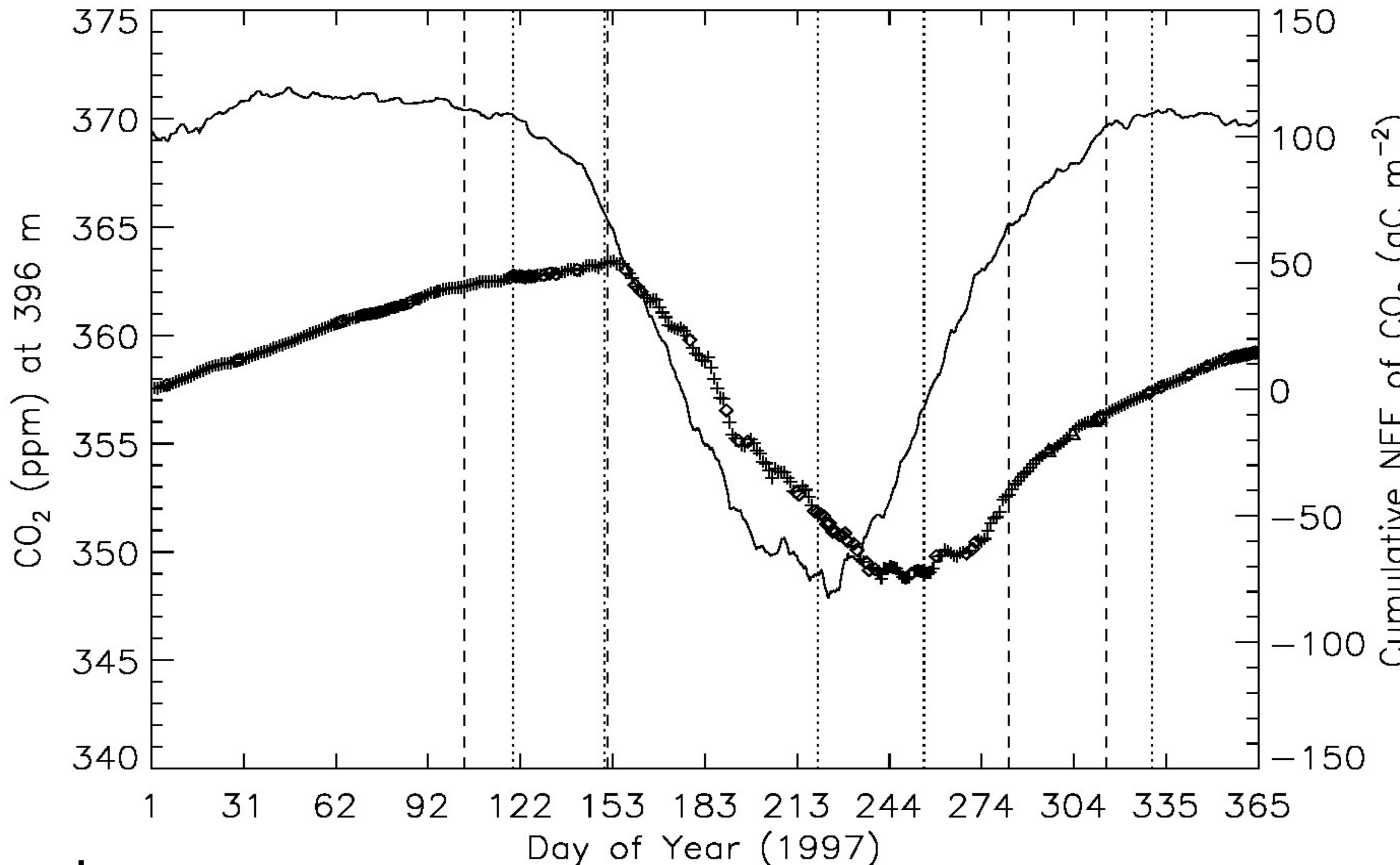


Sonic anemometer

Infrared gas analyzer

Campbell Scientific, Inc.
LI-COR, Inc.

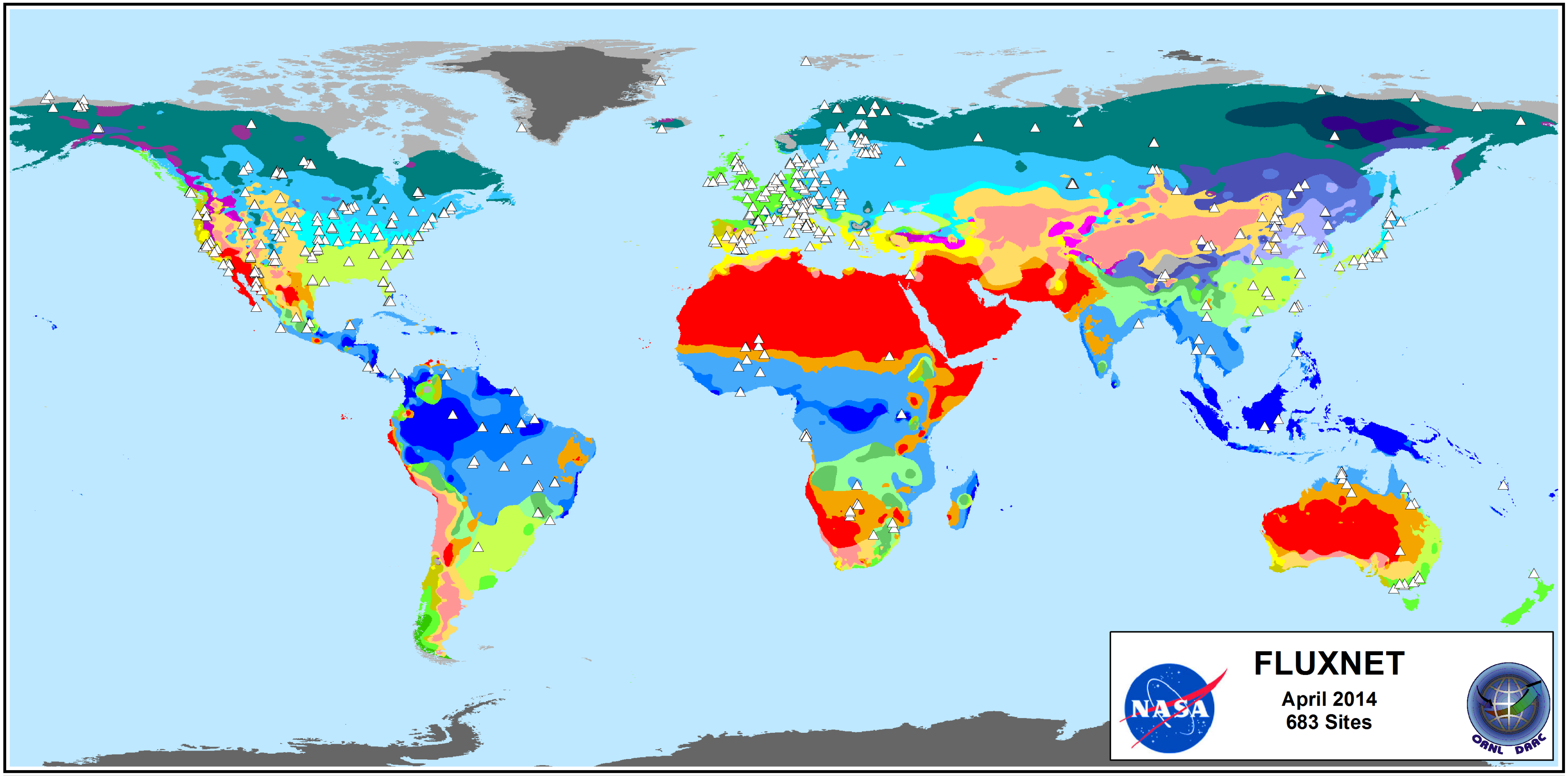
Net ecosystem-atmosphere exchange of CO_2 in northern Wisconsin



Weak carbon source

Davis et al, 2003

Global flux tower co-op: Hundreds of sites

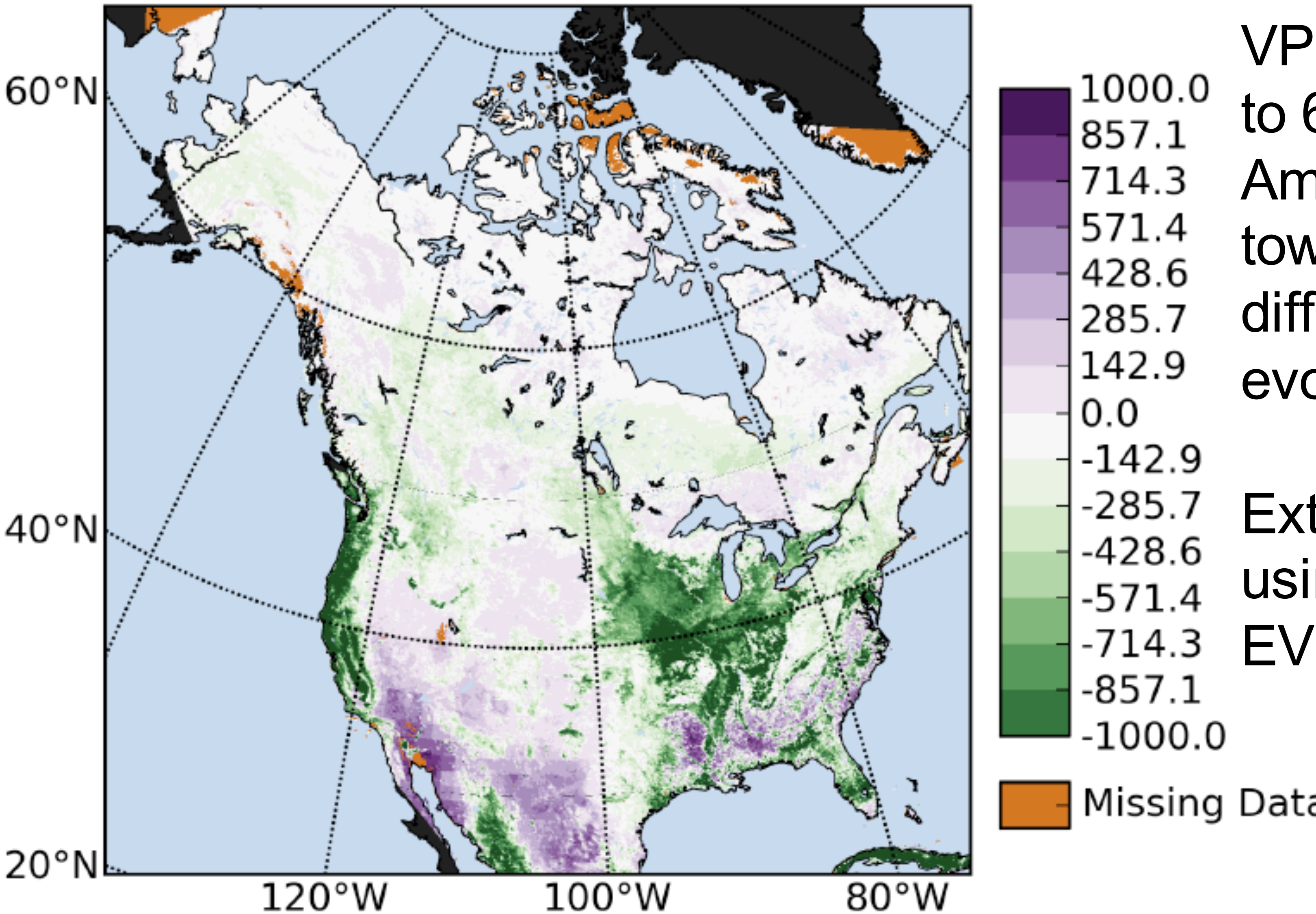


Koppen-Geiger Climate Classification (2006)

- Af - Tropical/Rainforest
- Am - Tropical/Monsoon
- Aw - Tropical/Savannah
- BWh - Arid/Desert/Hot
- BWk - Arid/Desert/Cold
- BSh - Arid/Steppe/Hot
- BSk - Arid/Steppe/Cold
- Temperate/Dry Summer/Hot Summer
- Temperate/Dry Summer/Warm Summer
- Temperate/Dry Summer/Cold Summer
- Cwb - Temperate/Dry Winter/Warm Summer
- Cwc - Temperate/Dry Winter/Cold Summer
- Cfa - Temperate/Without dry season/Hot Summer
- Cfb - Temperate/Without dry season/Warm Summer
- Cfc - Temperate/Without dry season/Cold Summer
- Dsa - Cold/Dry Summer/Hot Summer
- Dsb - Cold/Dry Summer/Warm Summer
- Dsc - Cold/Dry Summer/Cold Summer
- Dsd - Cold/Dry Summer/Very Cold_Winter
- Dwa - Cold/Dry Winter/Hot Summer
- Dwb - Cold/Dry Winter/Warm Summer
- Dwc - Cold/Dry Winter/Cold Summer
- Dwd - Cold/Dry Winter/Very Cold Winter
- Dfa - Cold/Without dry season/Hot Summer
- Dfb - Cold/Without dry season/Warm Summer
- Dfc - Cold/Without dry season/Cold Summer
- Dfd - Cold/Without dry season/Very Cold Winter
- ET - Polar/Tundra
- EF - Polar/Frost
- ET - Polar/Tundra
- EF - Polar/Frost

~~! Beware of closed data policies !~~

A CO_2 flux map for N. America



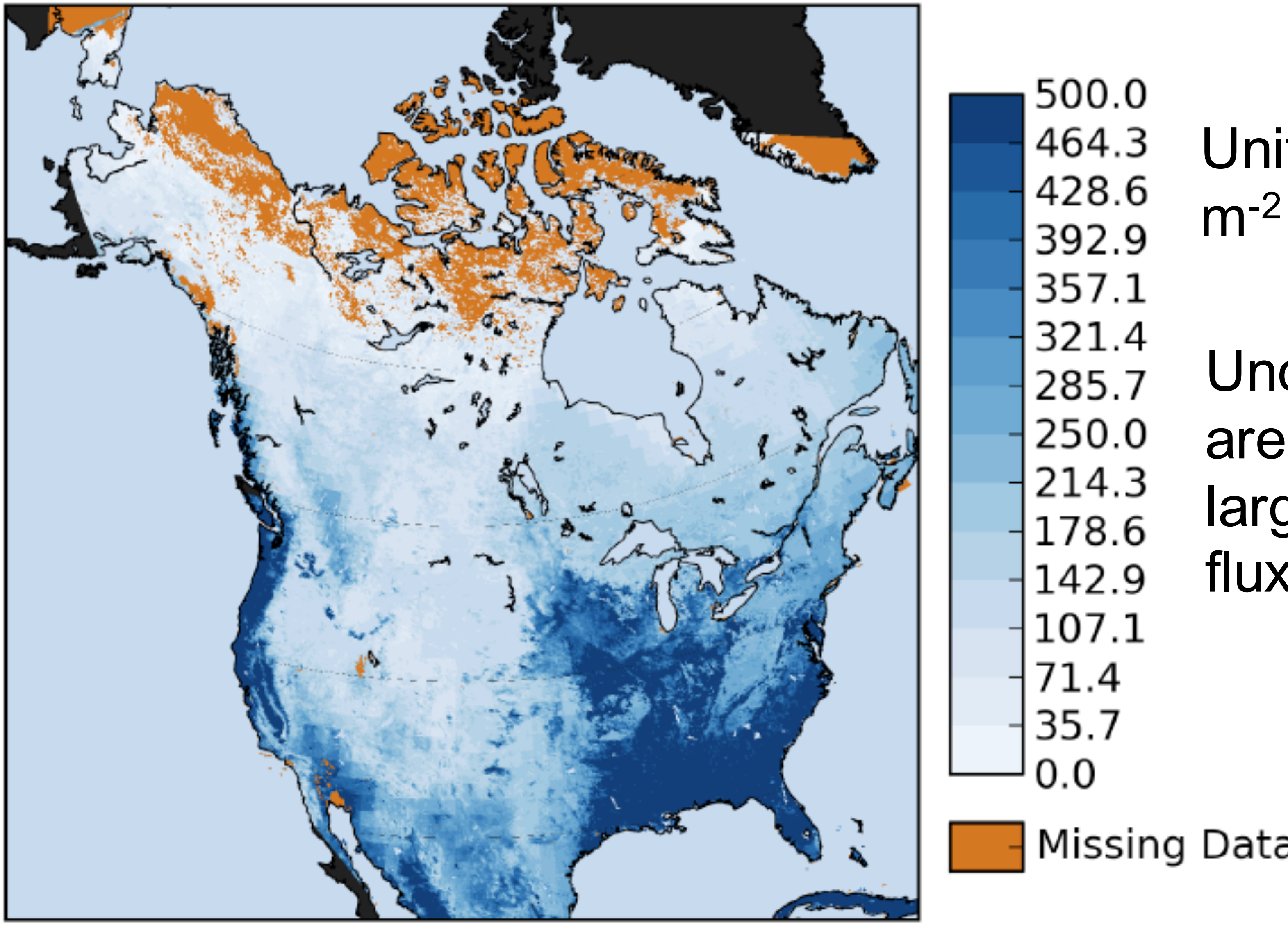
VPRM fitted to 65 N. American flux towers using differential evolution.

Extrapolate using T, PAR, EVI, PFTs.

Annual NEE in gC m^{-2} yr^{-1}, 2002

Hilton et al, 2014, Biogeosciences

Annual NEE error map, 2002



Units are gC m^{-2} yr^{-1}.

Uncertainties are about as large as the fluxes.

Hilton et al, 2014, Biogeosciences

Interannual CO_2 flux variations are very difficult to simulate (and measure?)

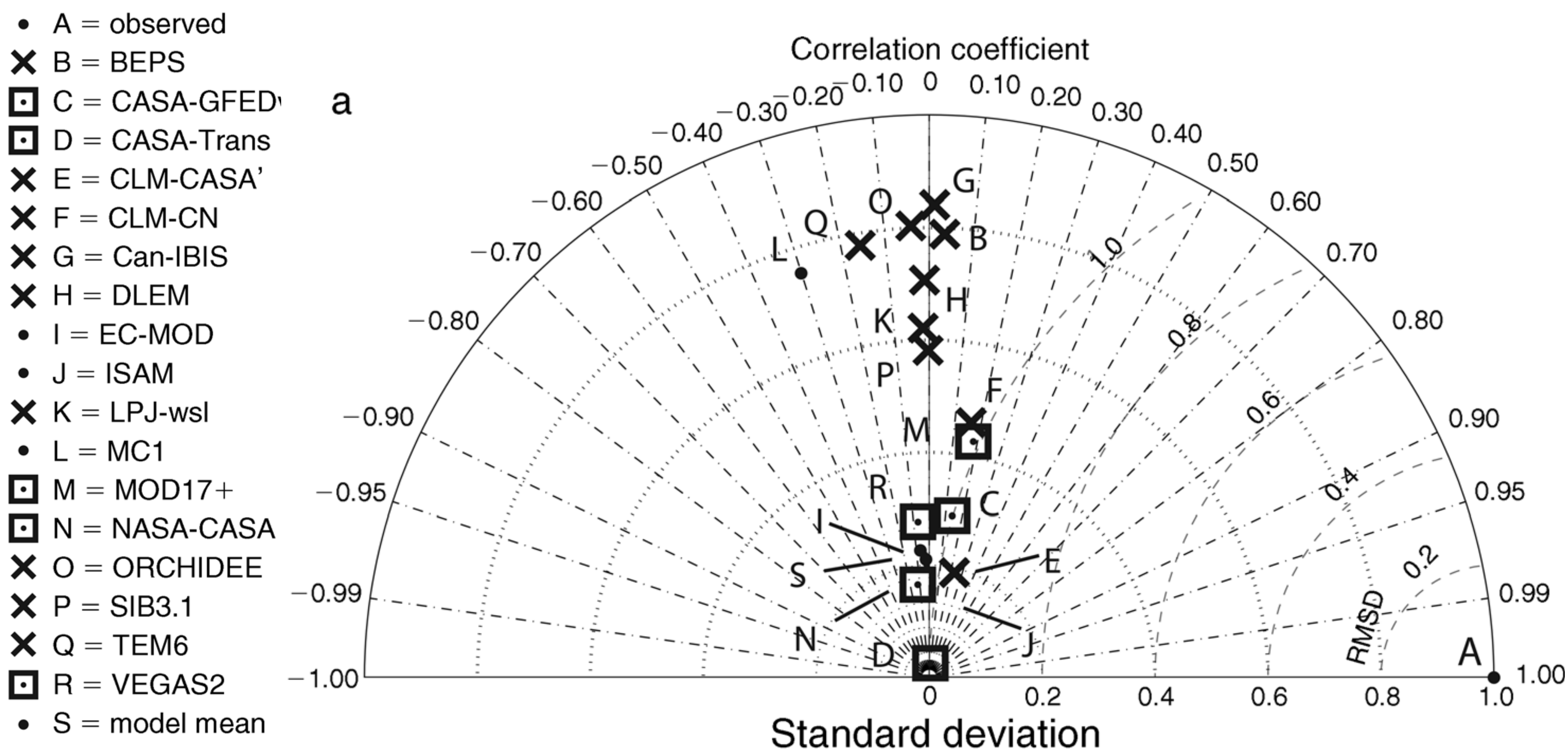


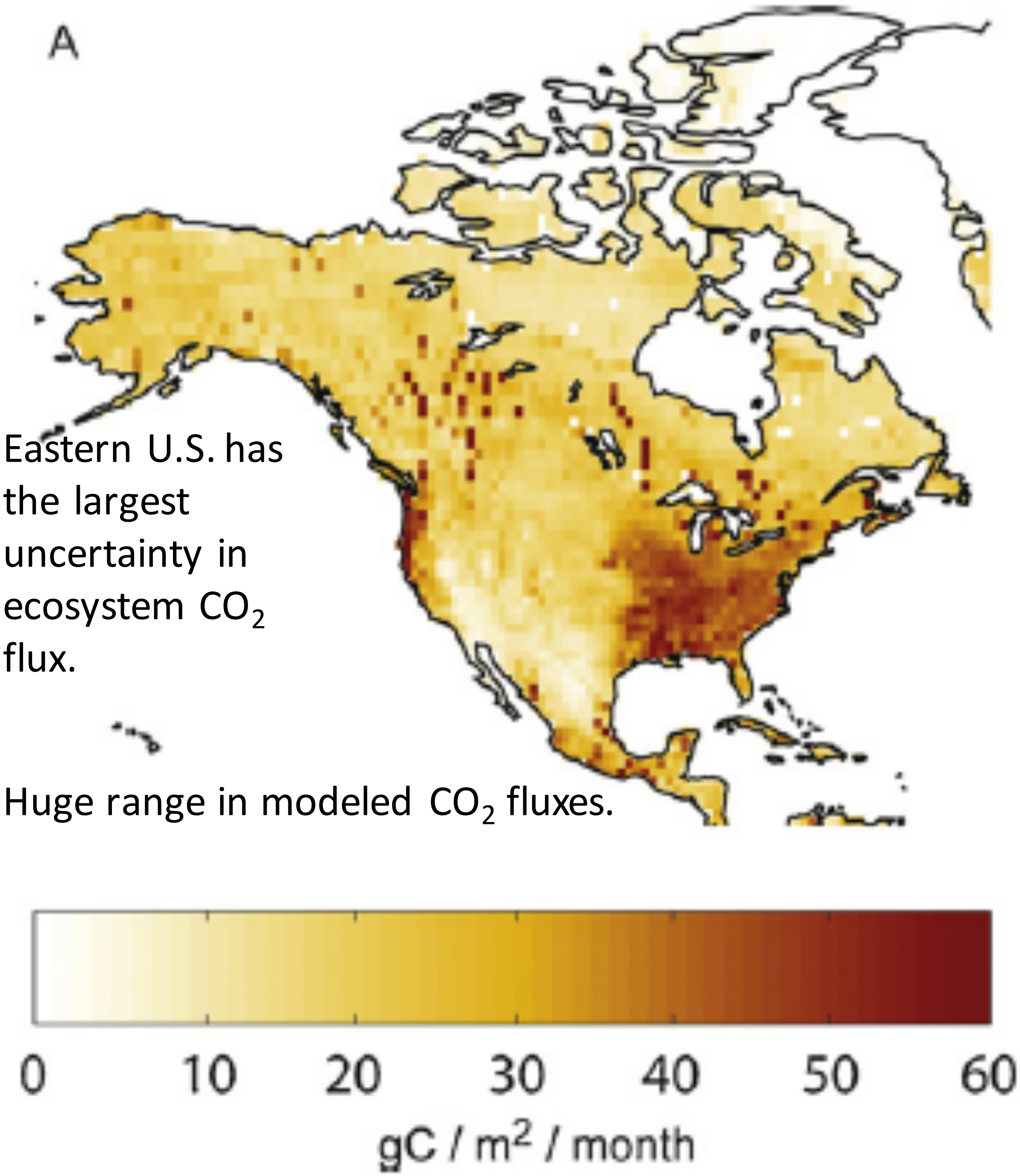
FIG. 4. Annual fluxes for all sites for (a) NEE, (b) GPP, and (c) RE. The statistics of correlation coefficient (black dotted-dashed axis lines), average difference in flux magnitude between the modeled and observed fluxes (RMSD; gray dashed axis lines), and standard deviation (gray dotted axis lines) are calculated from temporal (within-site) modeled variability. Squares represent light-use-efficiency models, X's represent enzyme-kinetic models, and dots represent statistical models (observed and model mean).

North American model – flux tower comparison.

Raczka et al., (2013)

Regional CO_2 fluxes are highly uncertain.

Across-model standard deviation in long-term mean (2000–2005) summer (June, July, August) terrestrial biosphere model estimates of net ecosystem productivity.



Eastern U.S. has the largest uncertainty in ecosystem CO_2 flux.

Huge range in modeled CO_2 fluxes.

"The range in model estimates of net ecosystem productivity (NEP) for North America is much narrower than estimates of productivity or respiration, with estimates of NEP varying between −0.7 and 2.2 PgC yr−1, while gross primary productivity and heterotrophic respiration vary between 12.2 and 32.9 PgC yr−1 and 5.6 and 13.2 PgC yr−1, respectively."

Huntzinger et al. (2012)

Why is it so difficult to simulate ecosystem-atmosphere carbon fluxes?

Why is it so difficult to simulate ecosystem-atmosphere carbon fluxes?

Ecosystem processes are complex and governing equations are highly parameterized.

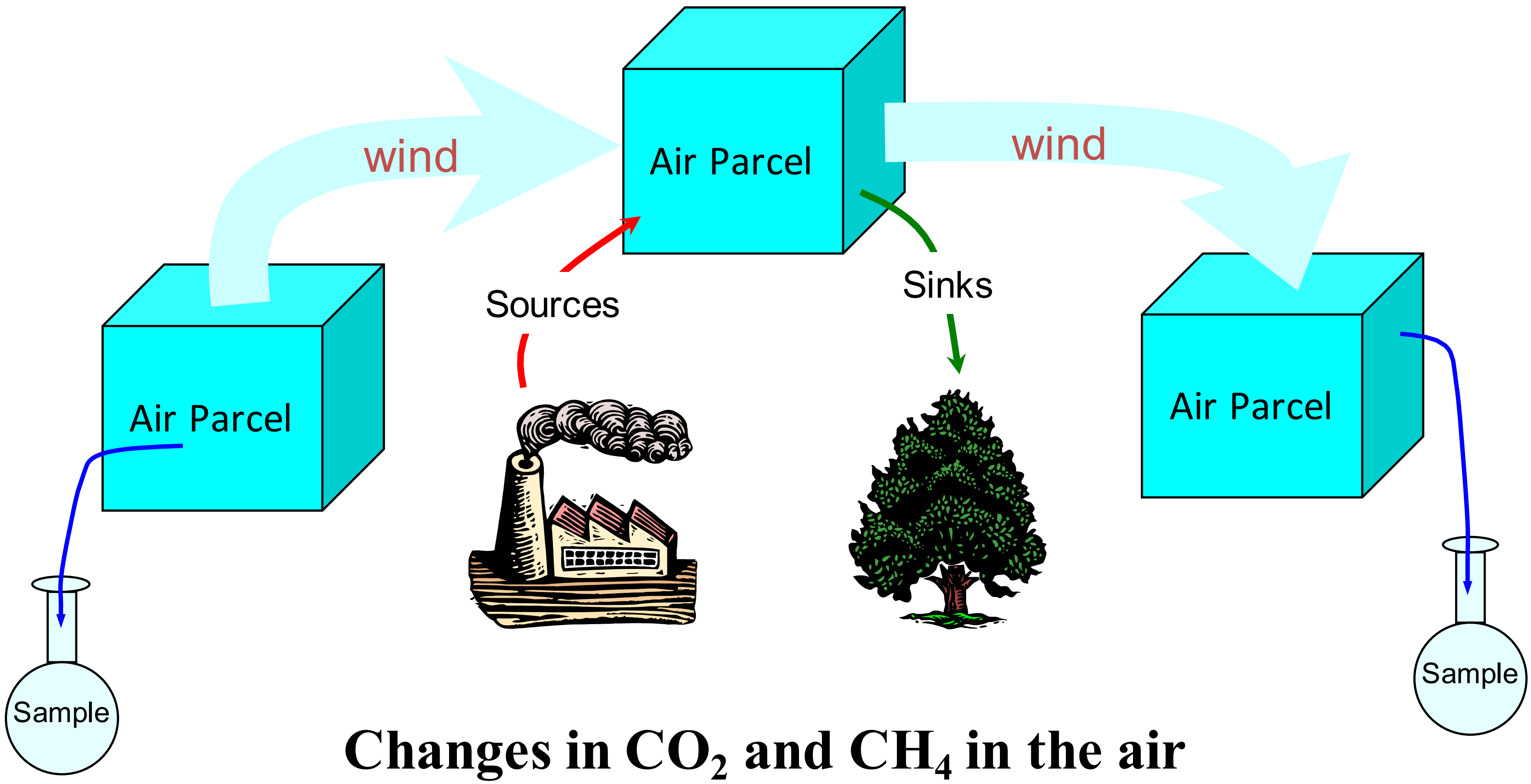
The land surface is heterogeneous down to very small spatial resolution.

Atmospheric inversions have the potential to close this gap

- Measure CO_2 at point A
- Follow air flow to point B
- Measure CO_2 at point B
- Infer sources and sinks of CO_2 in between A and B.

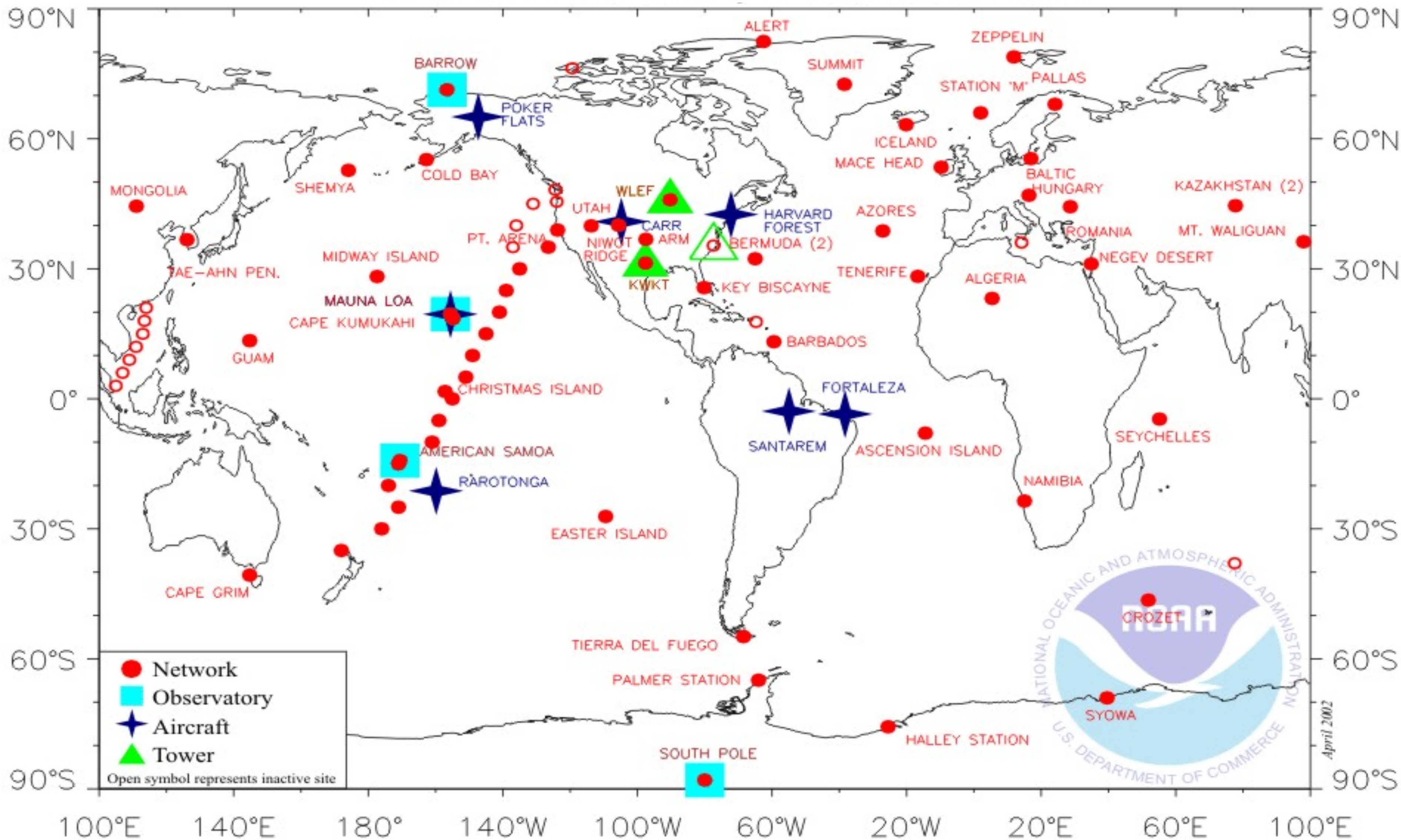
- Requires dense, high-quality atmospheric data, and accurate atmospheric transport.

"Atmospheric inversion"



Changes in CO_2 and CH_4 in the air tell us about sources and sinks

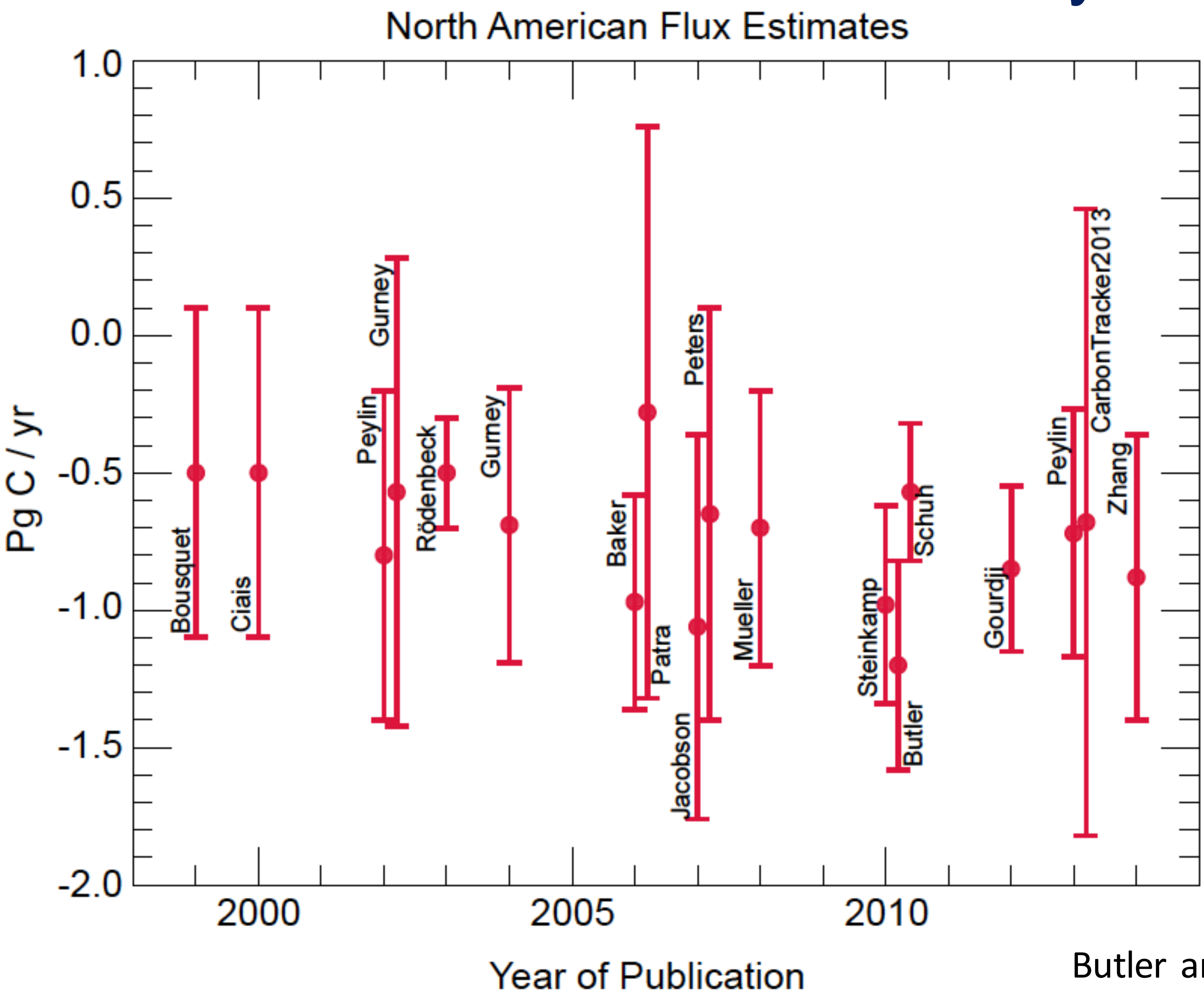
Global atmospheric CO_2 measurement network: 200(?) sites



The NOAA CMDL Carbon Cycle Greenhouse Gases group operates 4 measurement programs. In situ measurements are made at the CMDL baseline observatories: Barrow, Alaska; Mauna Loa, Hawaii; Tutuila, American Samoa; and South Pole, Antarctica. The cooperative air sampling network includes samples from fixed sites and commercial ships. Measurements from tall towers and aircraft began in 1992. Presently, atmospheric carbon dioxide, methane, carbon monoxide, hydrogen, nitrous oxide, sulfur hexaflouride, and the stable isotopes of carbon dioxide and methane are measured. Group Chief: Dr. Pieter Tans, Carbon Cycle Greenhouse Gases, Boulder, Colorado, (303) 497-6678 (ptans@cmdl.noaa.gov, http://www.cmdl.noaa.gov/ccgg).

Old figure…more sites today…and OCO-2

Results from atmospheric inversions: North American terrestrial ecosystem fluxes



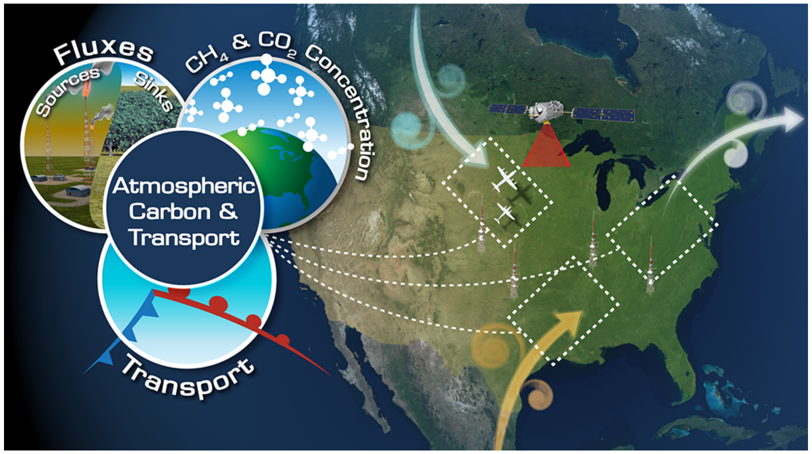
This shows that there is a significant N. American terrestrial sink.

We more or less knew that in 1990.

Butler and Davis, AGU 2014

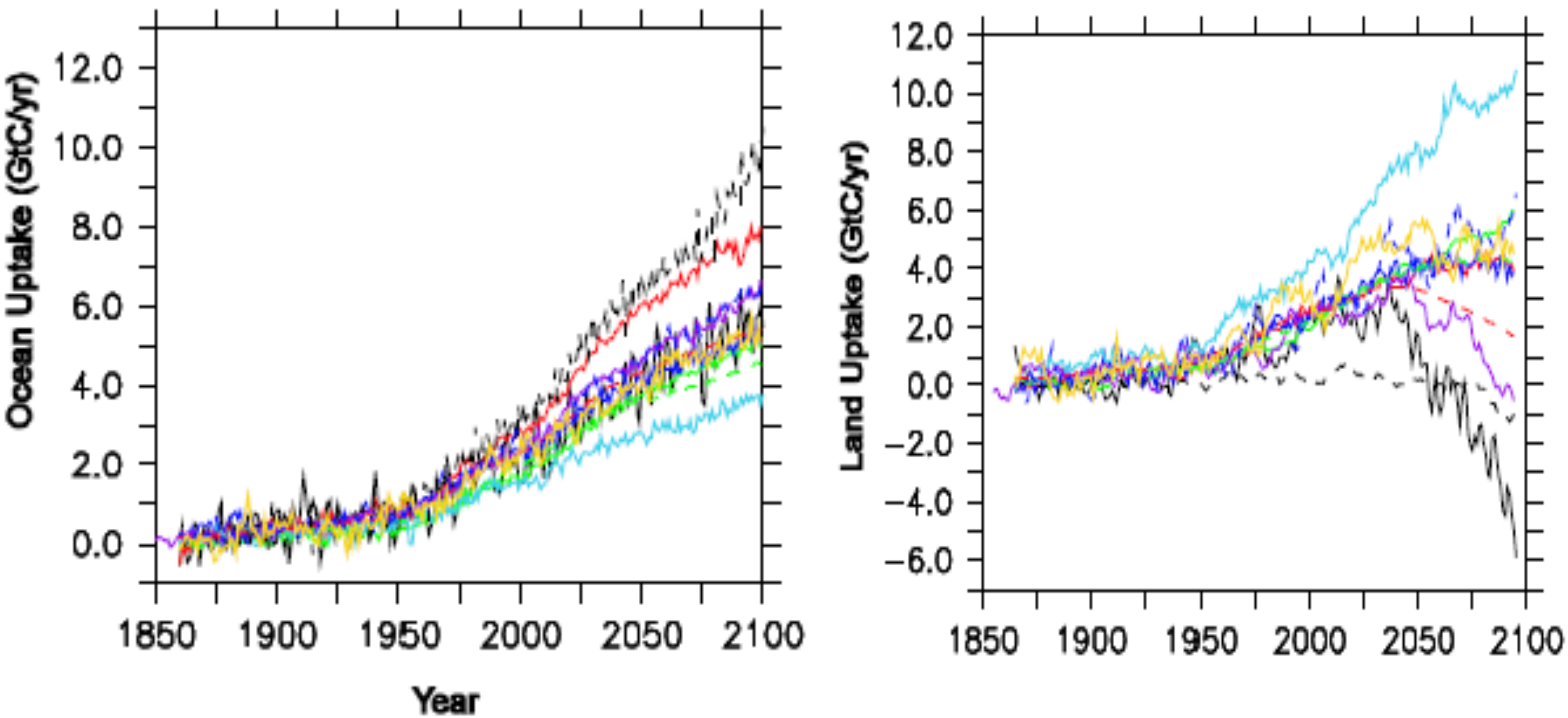
Three primary sources of uncertainty in GHG inverse flux estimates

1. Limited atmospheric data CO_2 and CH_4 data density
2. Uncertain CO_2 and CH_4 prior flux estimates
3. Poor knowledge of atmospheric transport – uncertainties largely unknown



Predicting future carbon fluxes

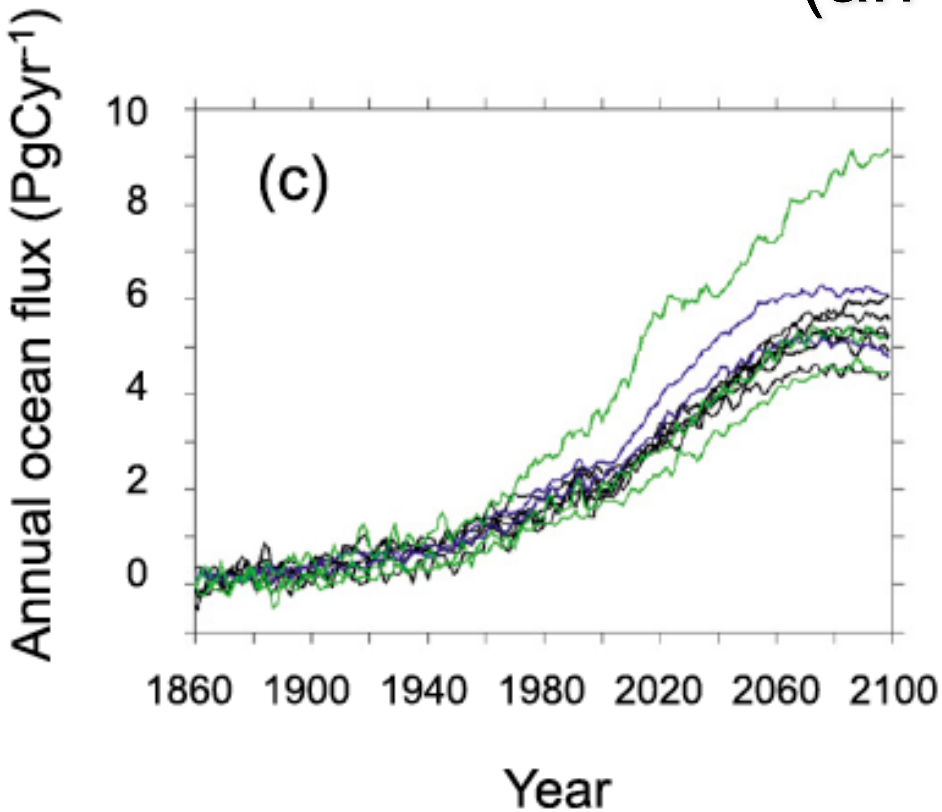
- **C^4MIP: comparison of 10 coupled climate/carbon models**
- **Large range of uncertainty (16 GtC yr^{-1} range in land flux by 2100) in the "natural" sinks buffering climate change. Management challenge!**



Friedlingstein et al., (2006)

Predicting future carbon fluxes
(an update)

CMIP5 results



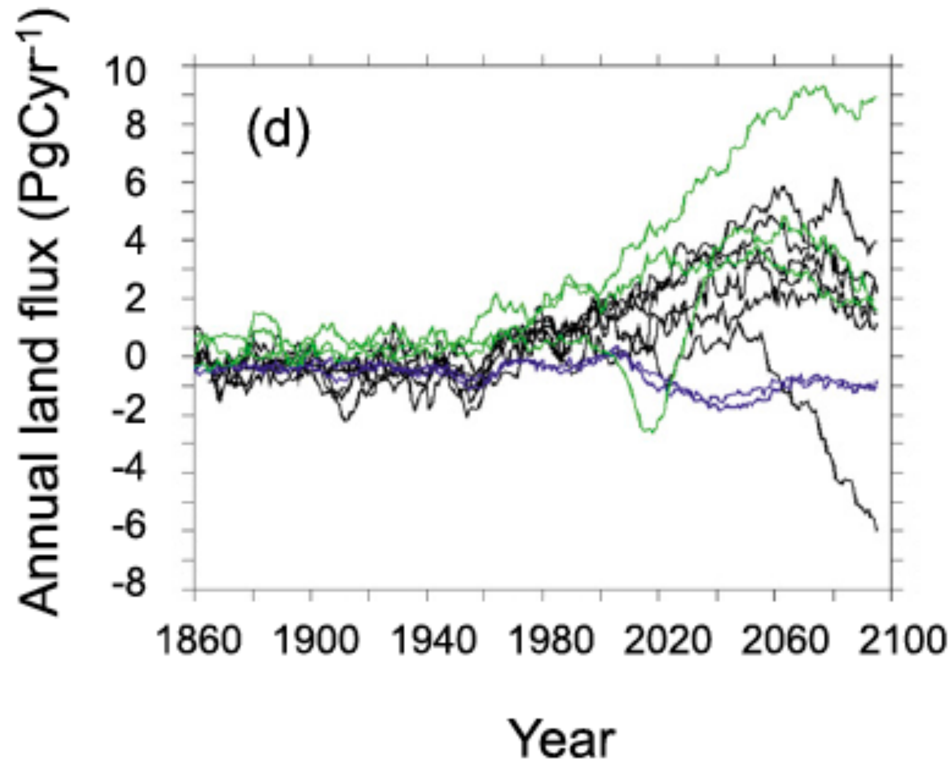


FIG. 4. Range of (a) cumulative global air to ocean carbon flux (PgC), (b) cumulative global air to land carbon flux (PgC) from the 11 ESMs E-driven simulations, (c) the annual global air to ocean carbon flux, and (d) annual global air to land carbon flux. Color code for model types is as in Fig. 1.

Friedlingstein et al., 2014

Observations needed to evaluate and improve these models are lacking.

What can be done?

- Advance process understanding.
- Move towards multi-state data assimilation.
- Close the measurement methods gap.
- Apply more measurements to more models. Enter the era of networked observations and ensemble modeling.

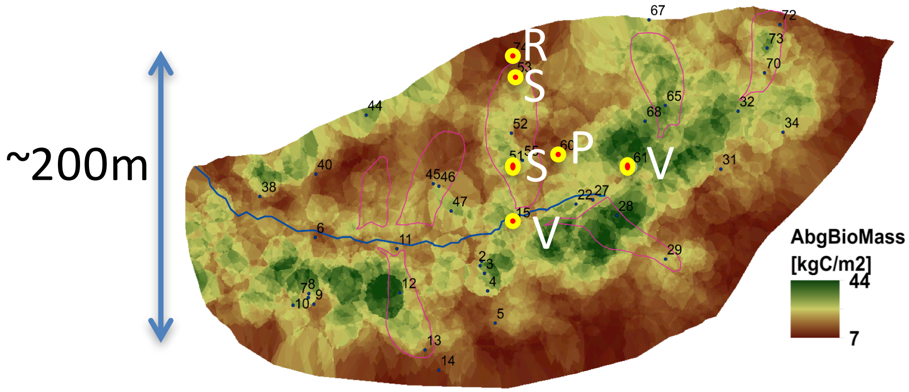
Some examples

- The world isn't flat!
  - Improve models by studying the carbon-water-nitrogen cycles in complex terrain.
- Closing the measurement gap with atmospheric inversions:
  - Increase measurement density.
  - Reduce uncertainty in atmospheric transport.

What can be done?

- Advance process understanding.
- Move towards multi-state data assimilation.
- Close the measurement methods gap.
- Apply more measurements to more models. Enter the era of networked observations and ensemble modeling.

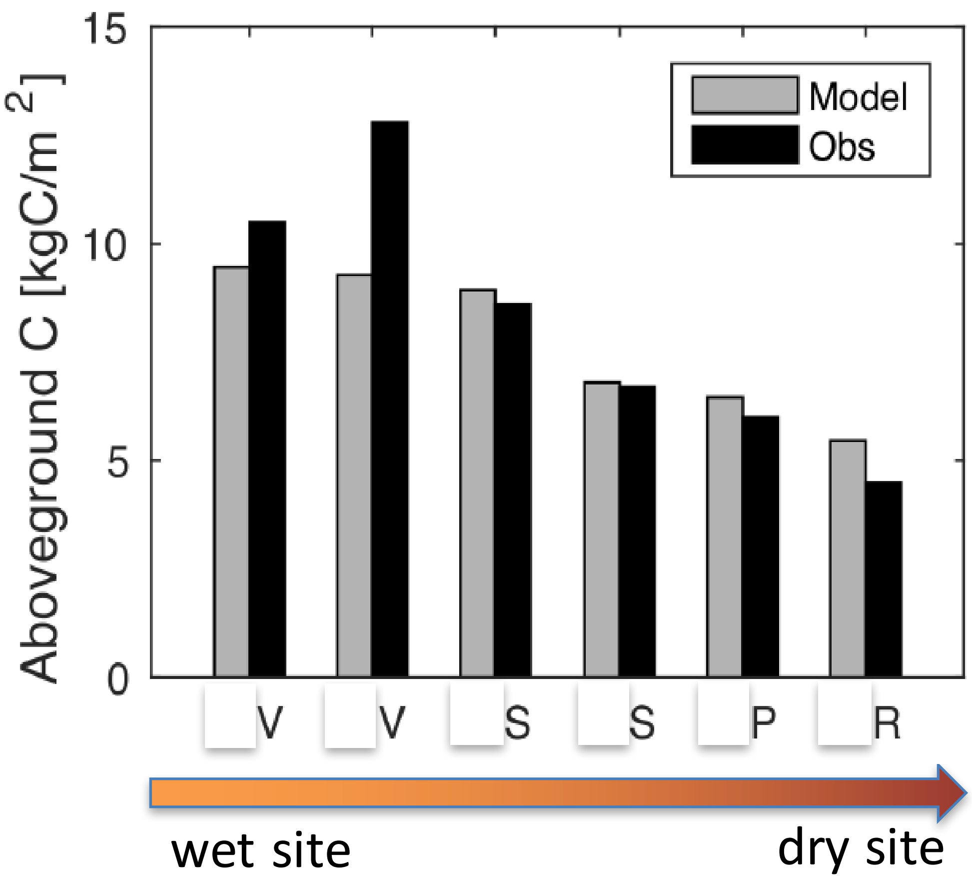
Spatial pattern of aboveground carbon at the Susquehanna Shale Hills Critical Zone Observatory

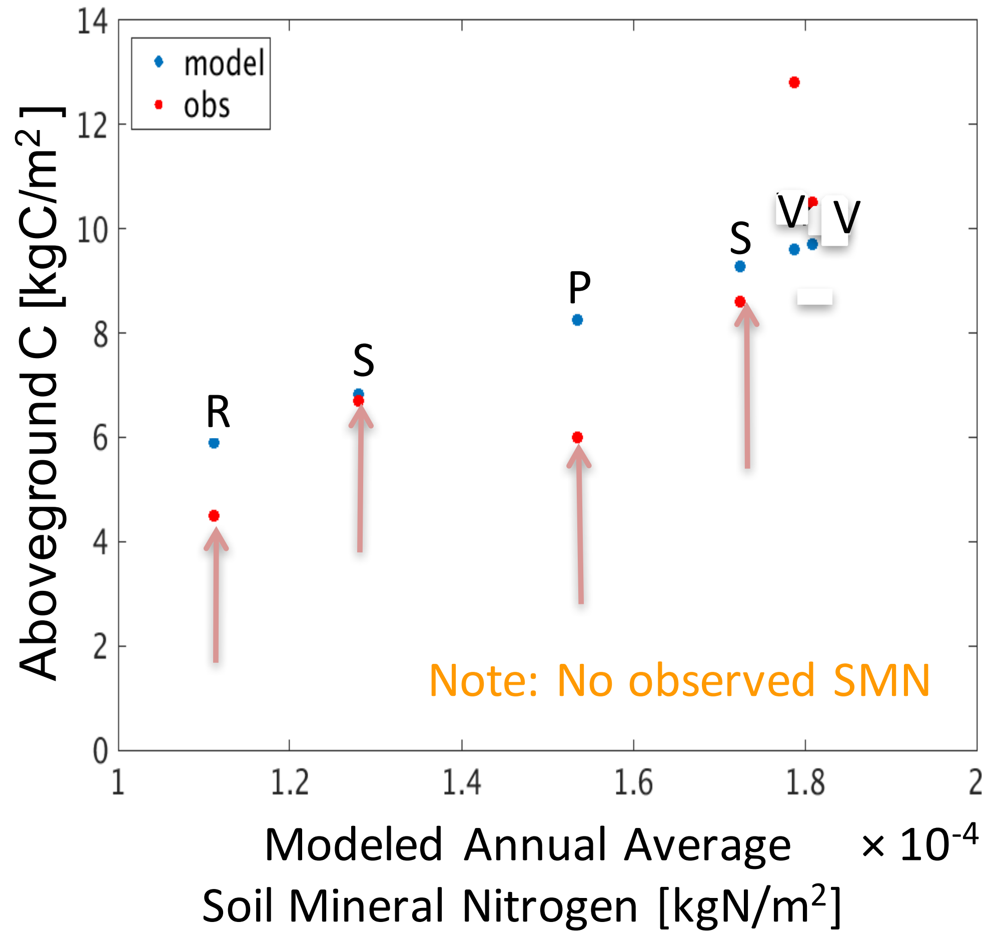


Data Courtesy of Margot Kaye's group

N determines the spatial pattern of aboveground biomass, and is closely linked to watershed hydrology

Assimilated data: Watershed average soil carbon and aboveground carbon turnover rate.





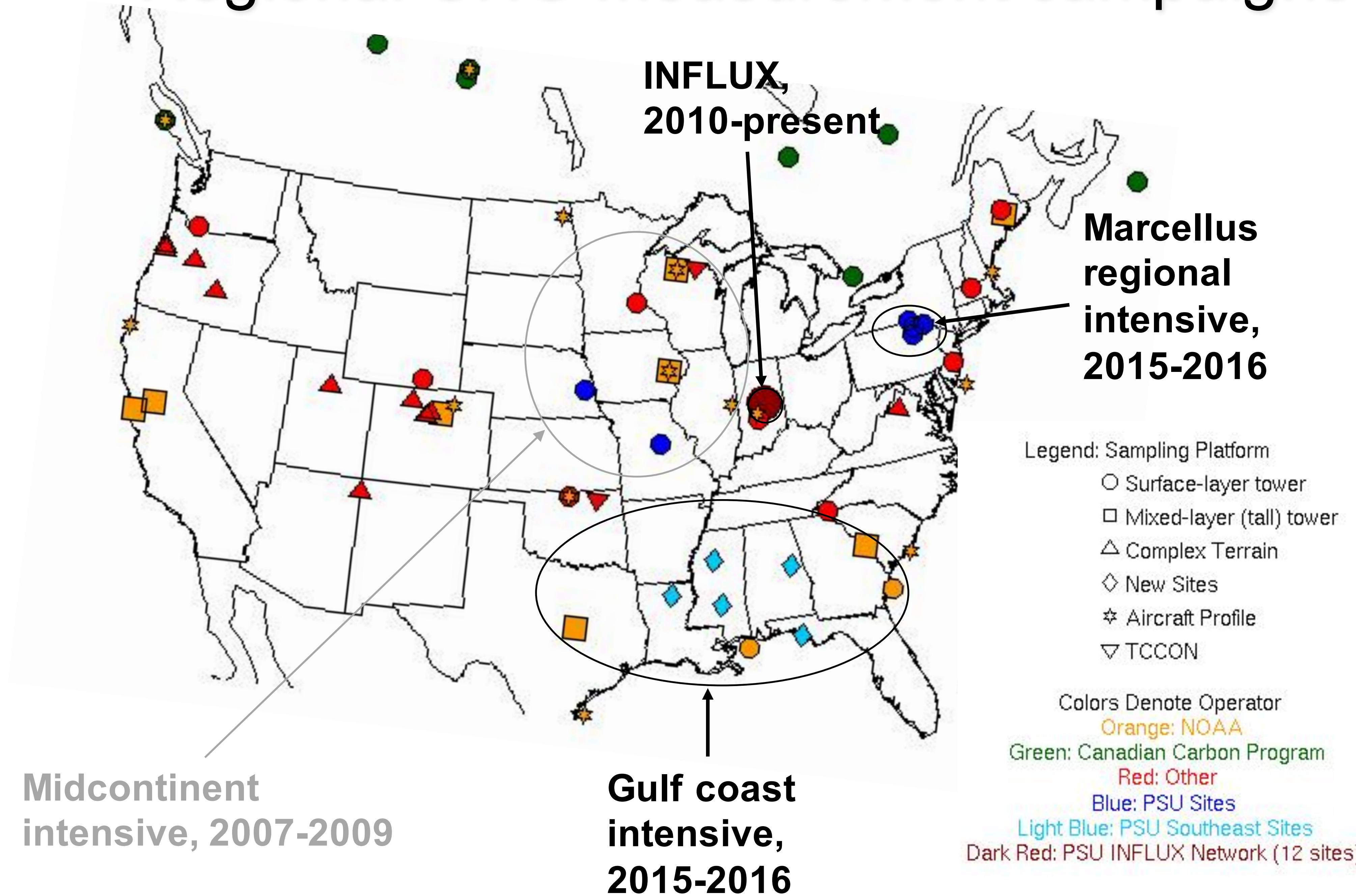
What can be done?

- Advance process understanding.
- Move towards multi-state data assimilation.
- Close the measurement methods gap.
- Apply more measurements to more models. Enter the era of networked observations and ensemble modeling.

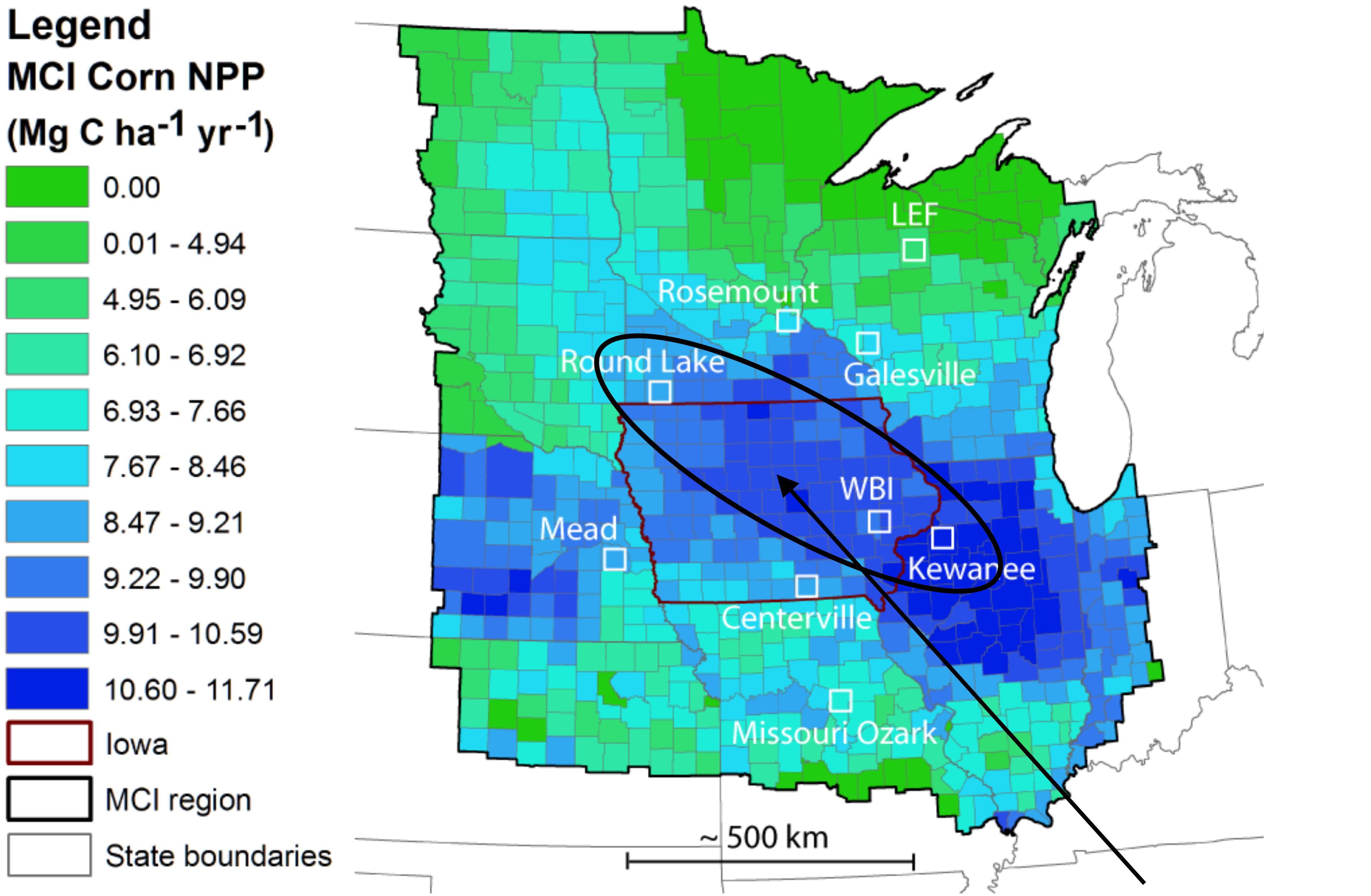
Some examples

- The world isn't flat!
  - Improve models by studying the carbon-water-nitrogen cycles in complex terrain.
- Closing the measurement gap with atmospheric inversions:
  - Increase measurement density.
  - Reduce uncertainty in atmospheric transport.

Regional GHG measurement campaigns

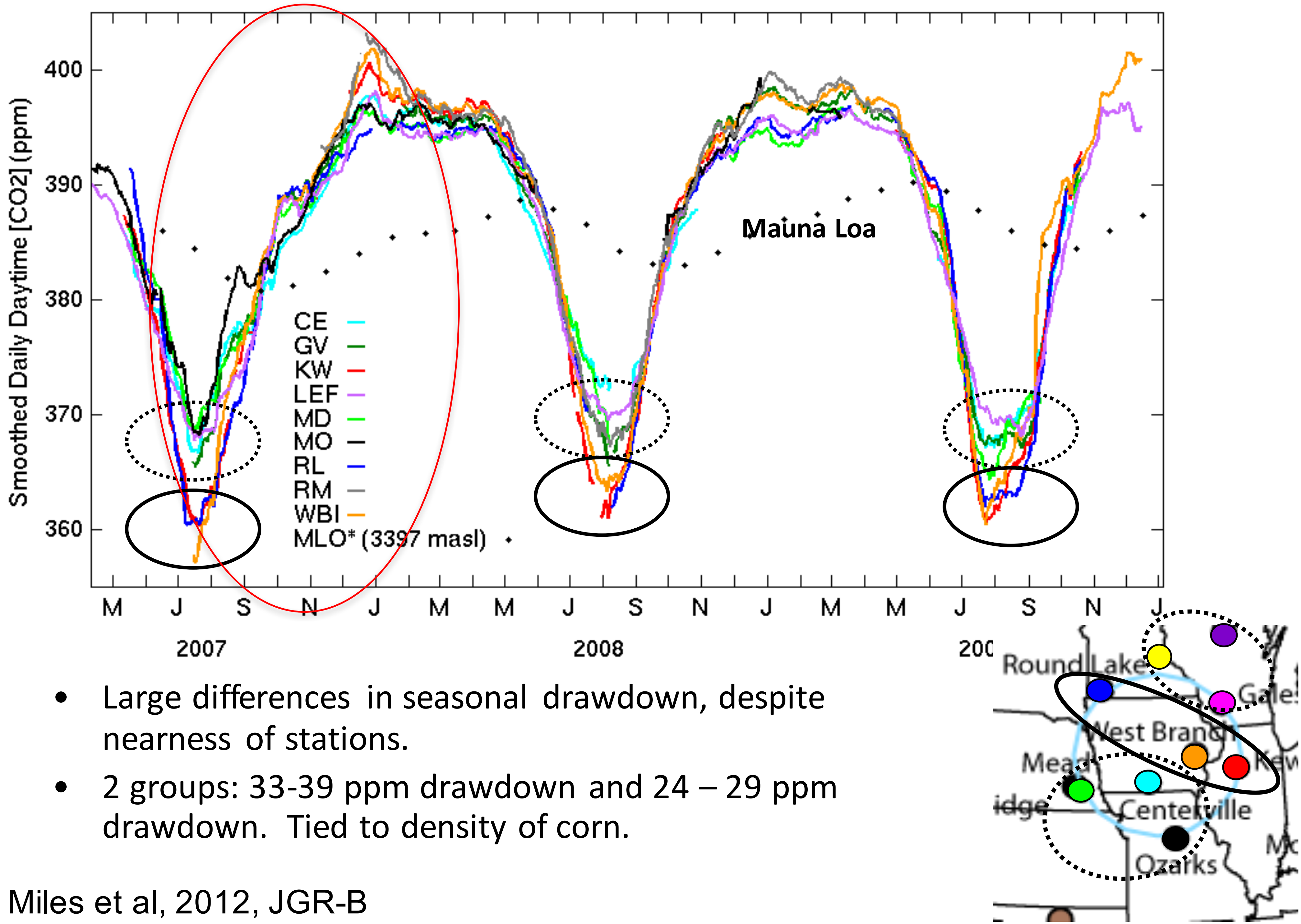


MidContinent Regional Intensive Tower-Based CO_2 Observational Network



Miles et al, 2012, JGR-B

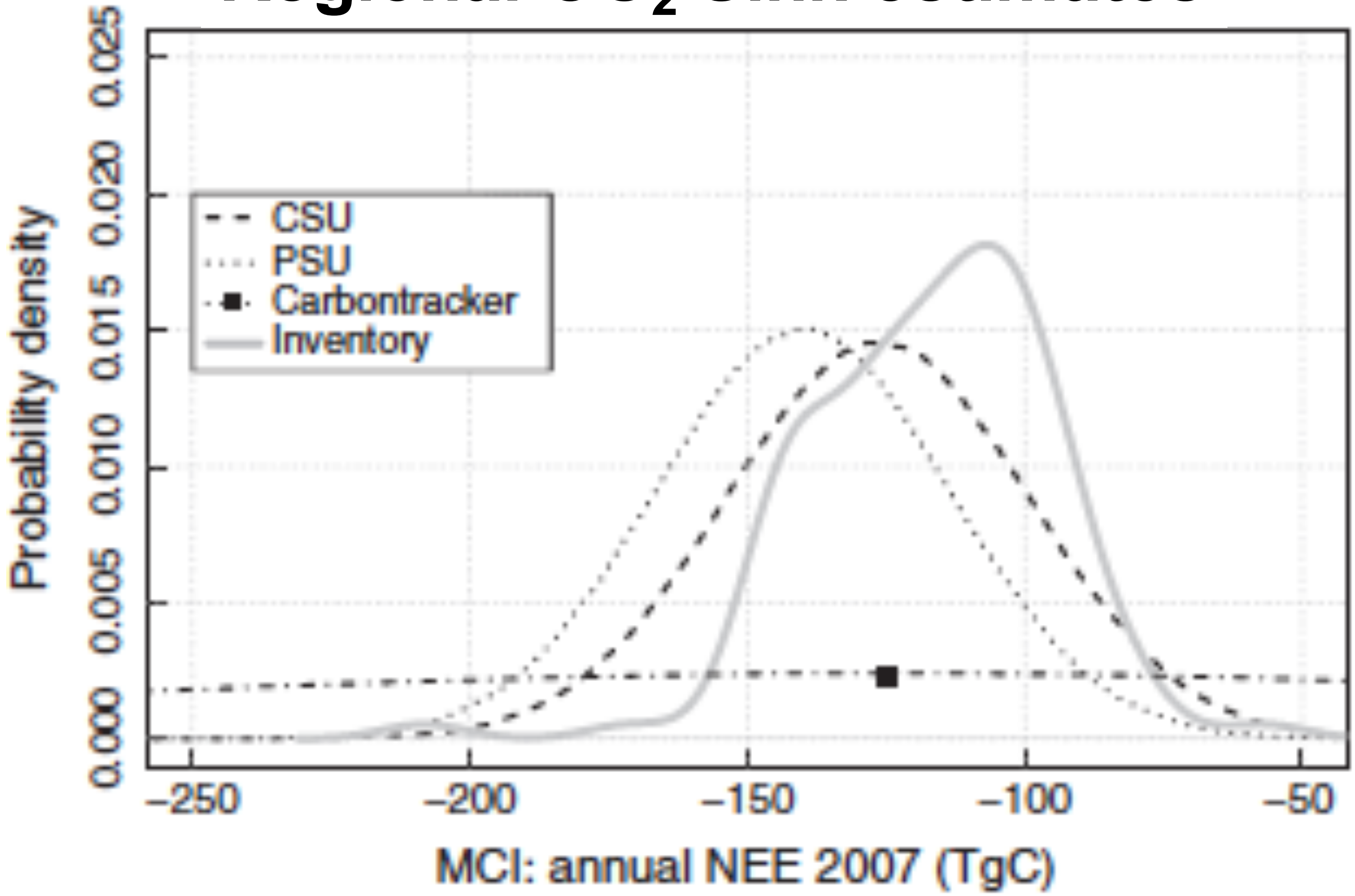
MCI 31 day running mean daily daytime average CO2



- Large differences in seasonal drawdown, despite nearness of stations.
- 2 groups: 33-39 ppm drawdown and 24 – 29 ppm drawdown. Tied to density of corn.

Miles et al, 2012, JGR-B

Regional CO_2 Sink estimates



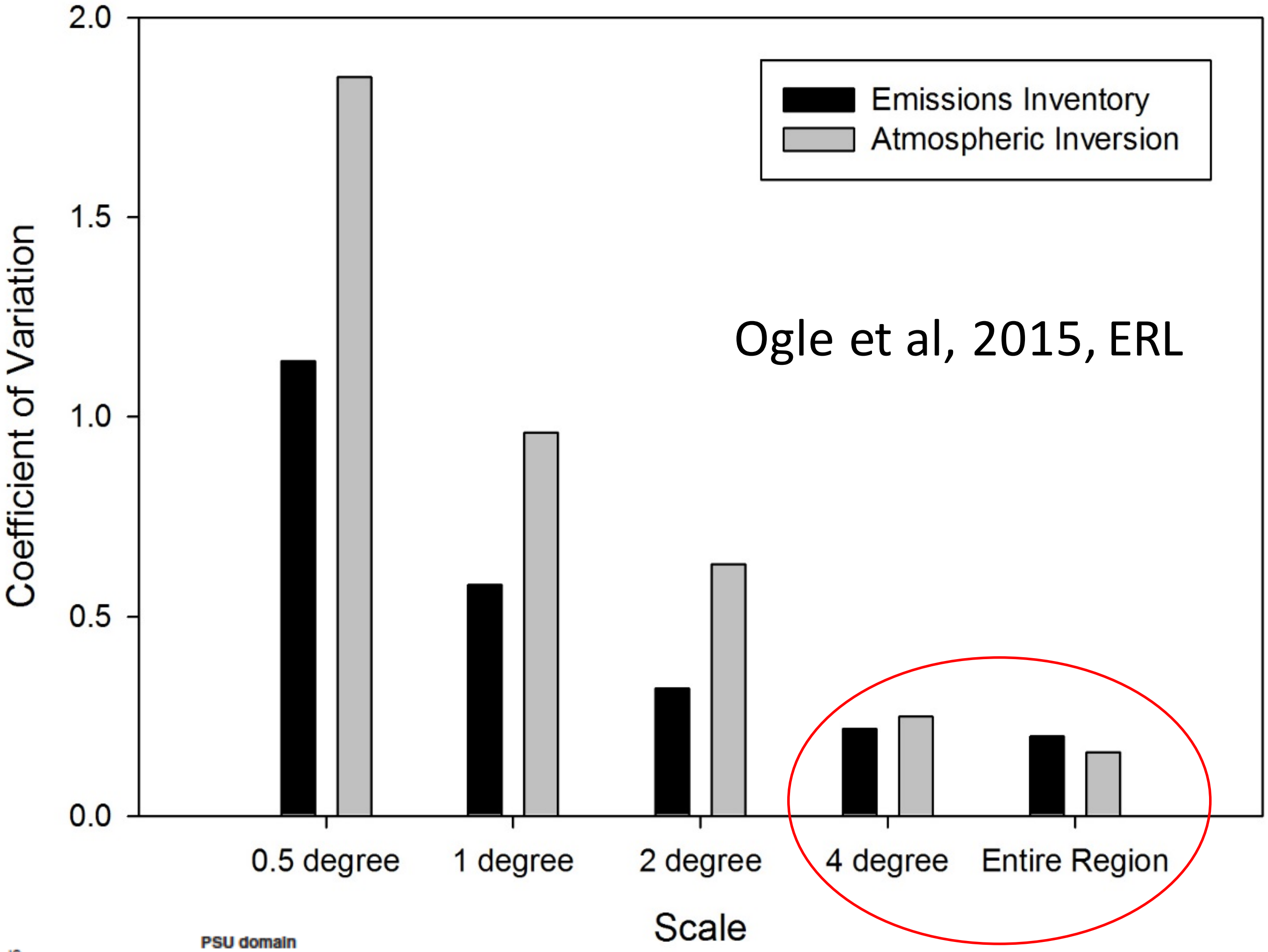
Schuh et al, 2013, GCB

Atmospheric inversions and agricultural inventory agree.

Regional inversions and inventory have similar uncertainty bounds!

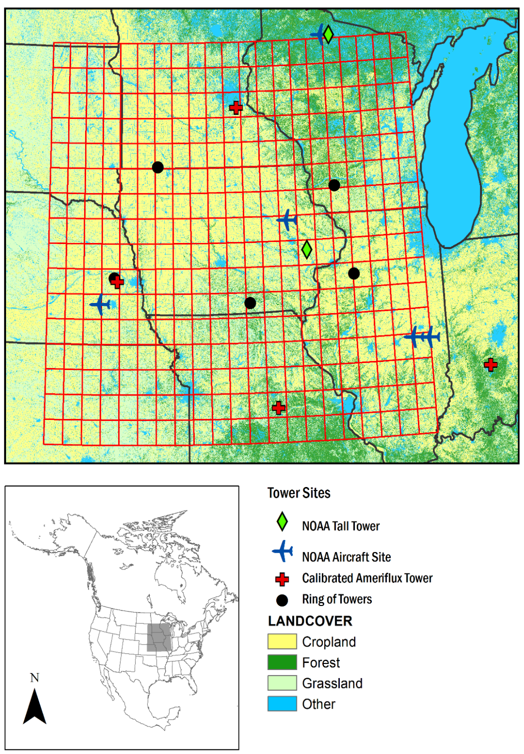
Atmospheric inversions have great potential for carbon balance inference given suitable data density.

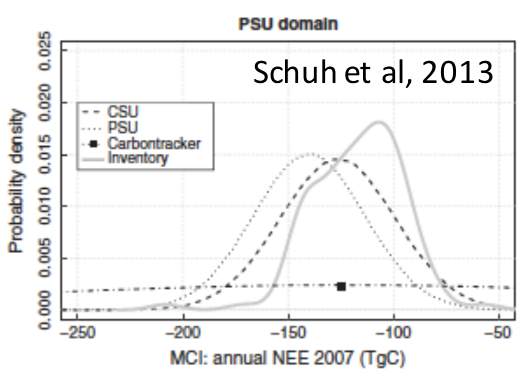
Cross-over point? Inversion vs. inventory



Atmospheric inversions provide great insights at global scale. Emissions inventories are very informative at small scales. Can we bridge the gap?

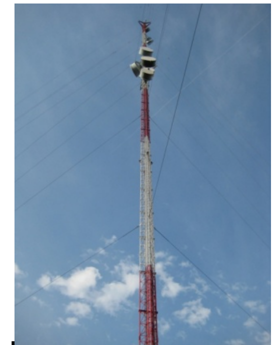
Midcontinent Intensive study area

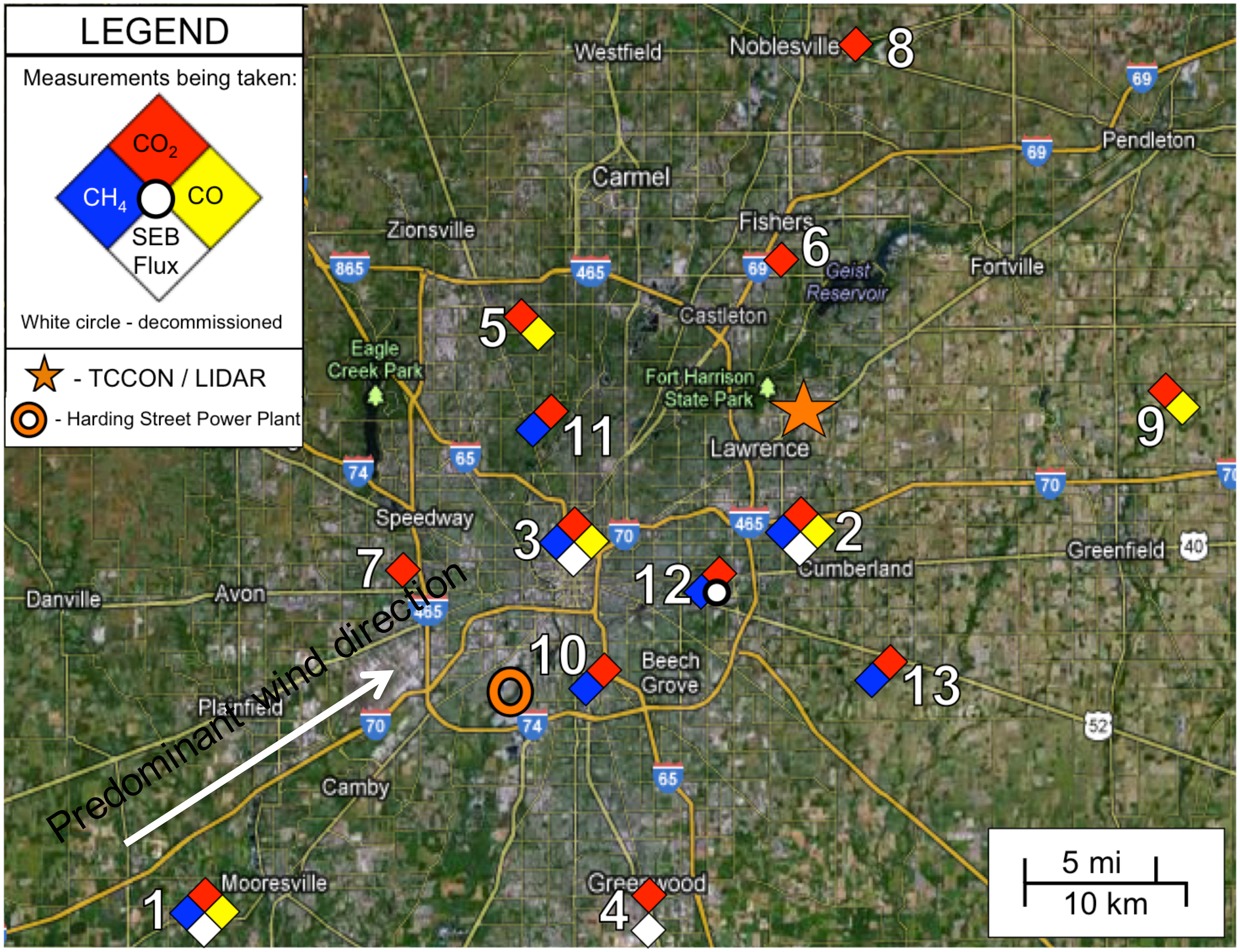




MCI results suggest that uncertainty in an atmospheric inversion equals the uncertainty in an agricultural inventory at (several 100 km)^2 resolution for this inventory and these atmospheric data

INFLUX GROUND-BASED NETWORK



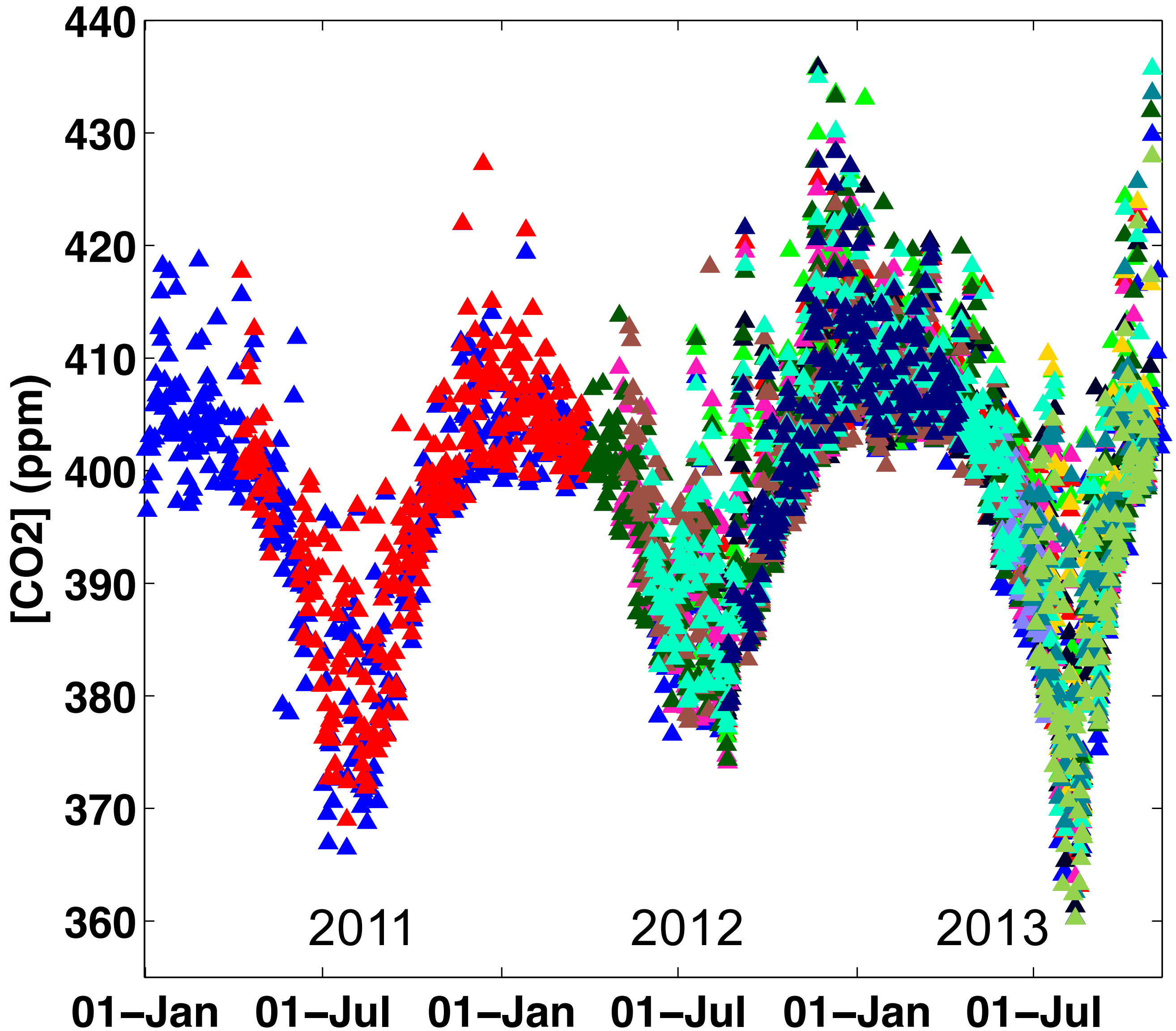


- Communications towers ~100 m AGL
- Picarro, CRDS sensors
- 12 measuring CO2, 11 with CH4, and 5 with CO

- 6 NOAA automated flask samplers
- NOAA LIDAR
- Eddy flux at 4 towers

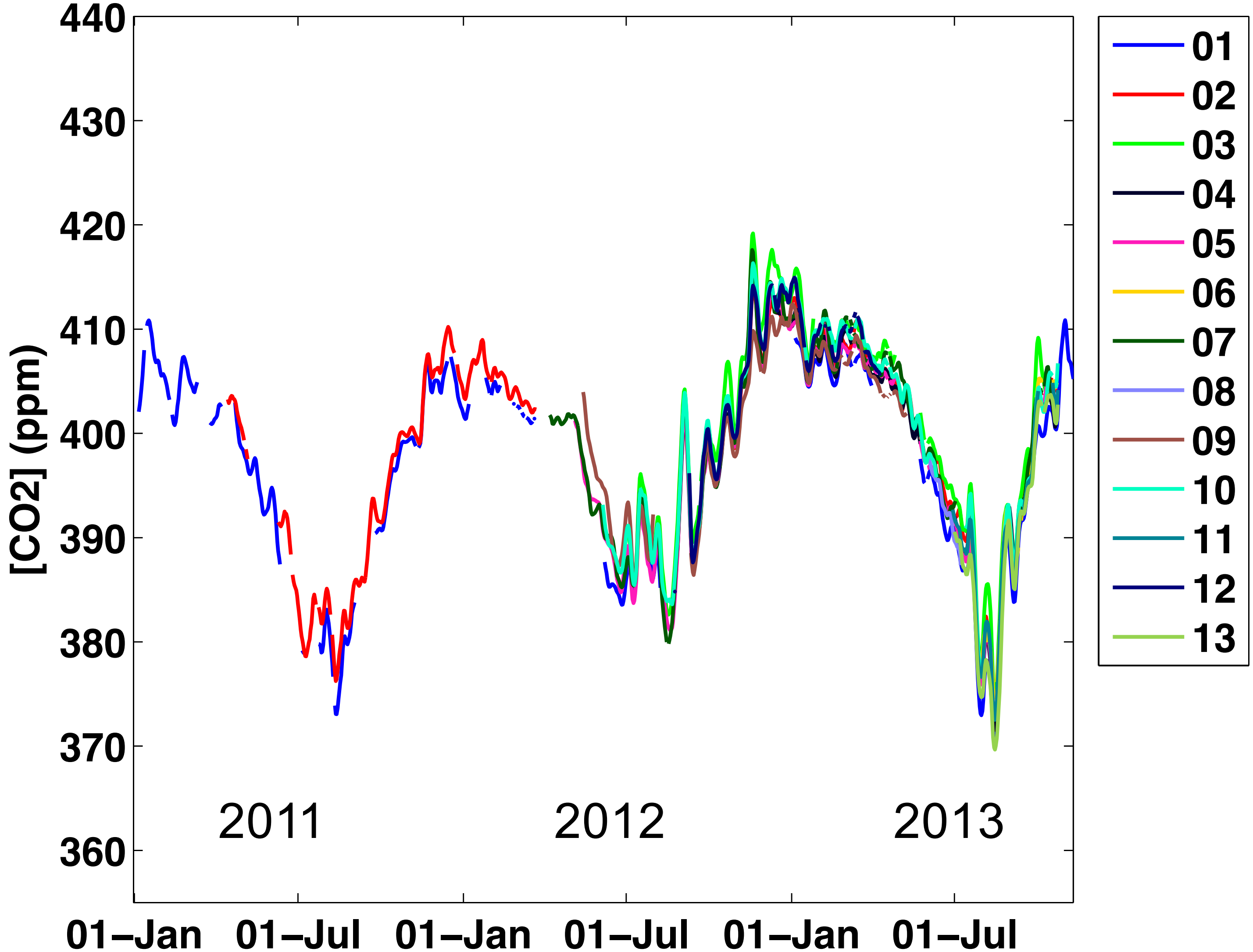
Comparison of $[CO_2]$ at INFLUX sites

- Afternoon daily $[CO_2]$
- Seasonal signal is apparent
- Significant overlap between sites (weather-driven variability)



Miles et al, in prep

- Afternoon [CO_2] with 21-day smoothing
- Seasonal and synoptic cycles are evident
- Site 03 (downtown): high [CO_2]
- Site 01 (background): low [CO_2]



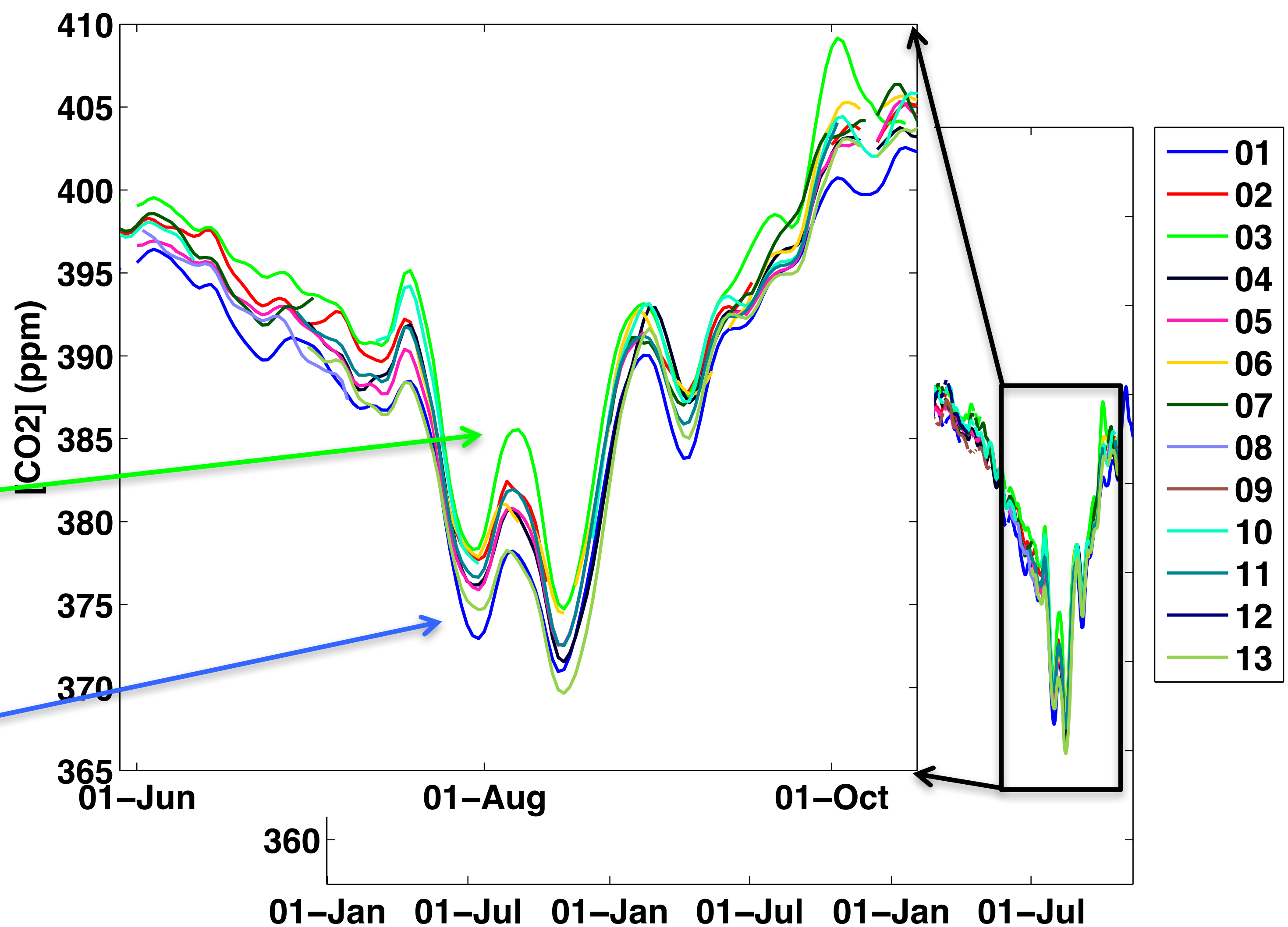
Miles et al, in prep

• Afternoon [CO_2] with 21-day smoothing

• Seasonal and synoptic cycles are evident

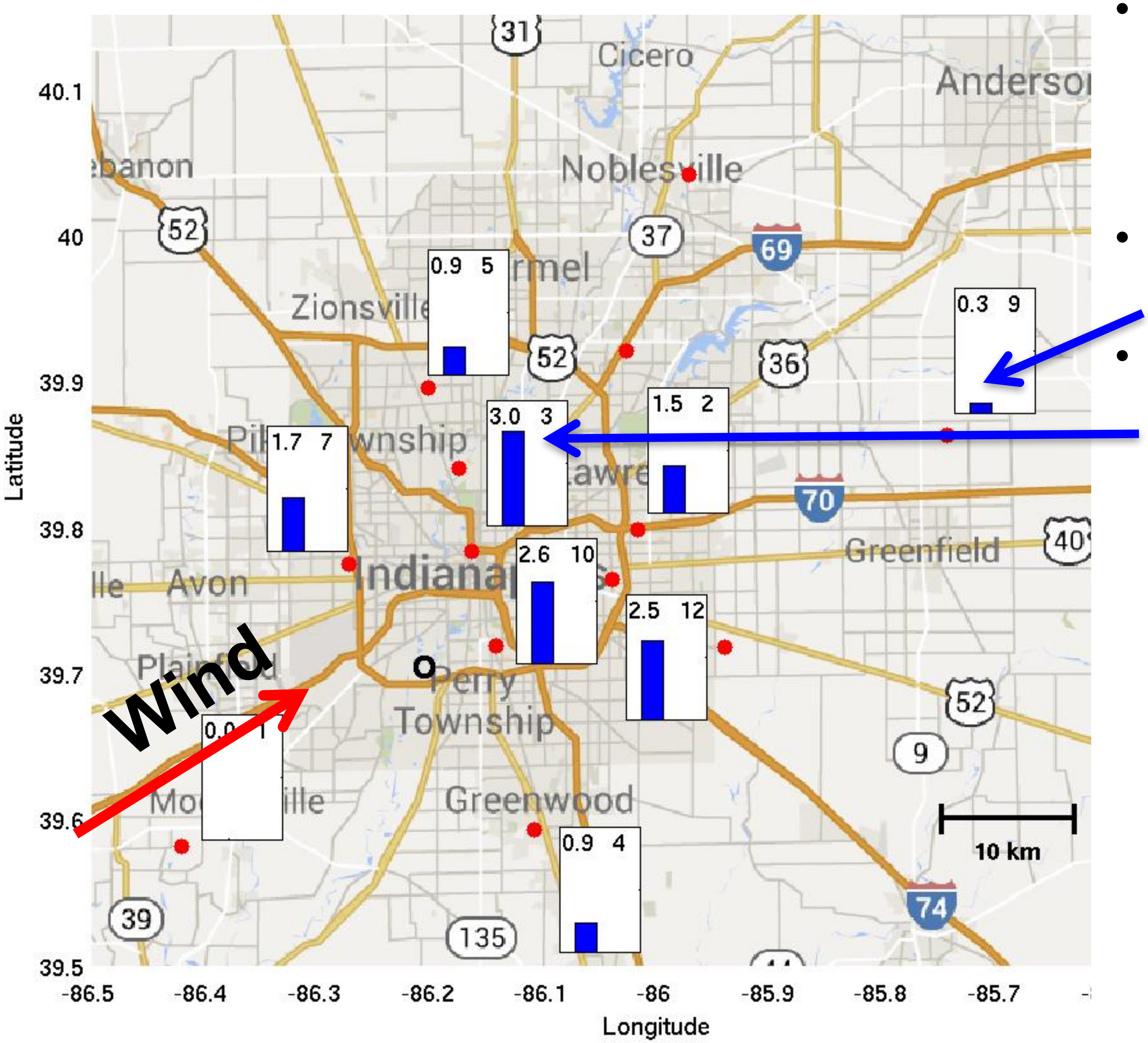
• Site 03 (downtown): high [CO_2]

• Site 01 (background): low [CO_2]



Miles et al, in prep

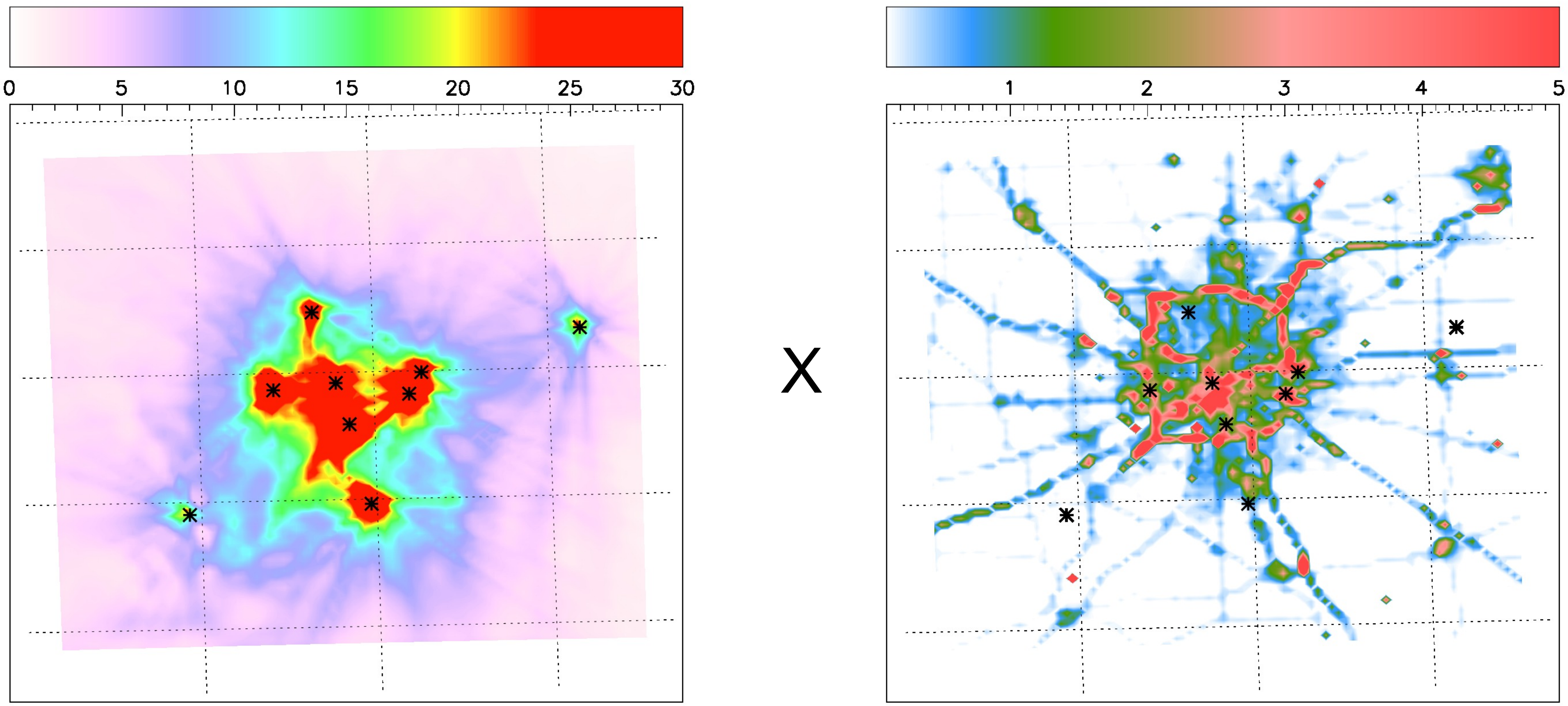
Spatial structure of urban CO_2: observed



- Observed CO_2: afternoon values, averaged Jan-April 2013
- Site 09: 0.3 ppm larger than Site 01
- Site 03: measures larger $[CO_2]$ by 3 ppm

Miles et al, in prep

Modeled CO2 mixing ratios

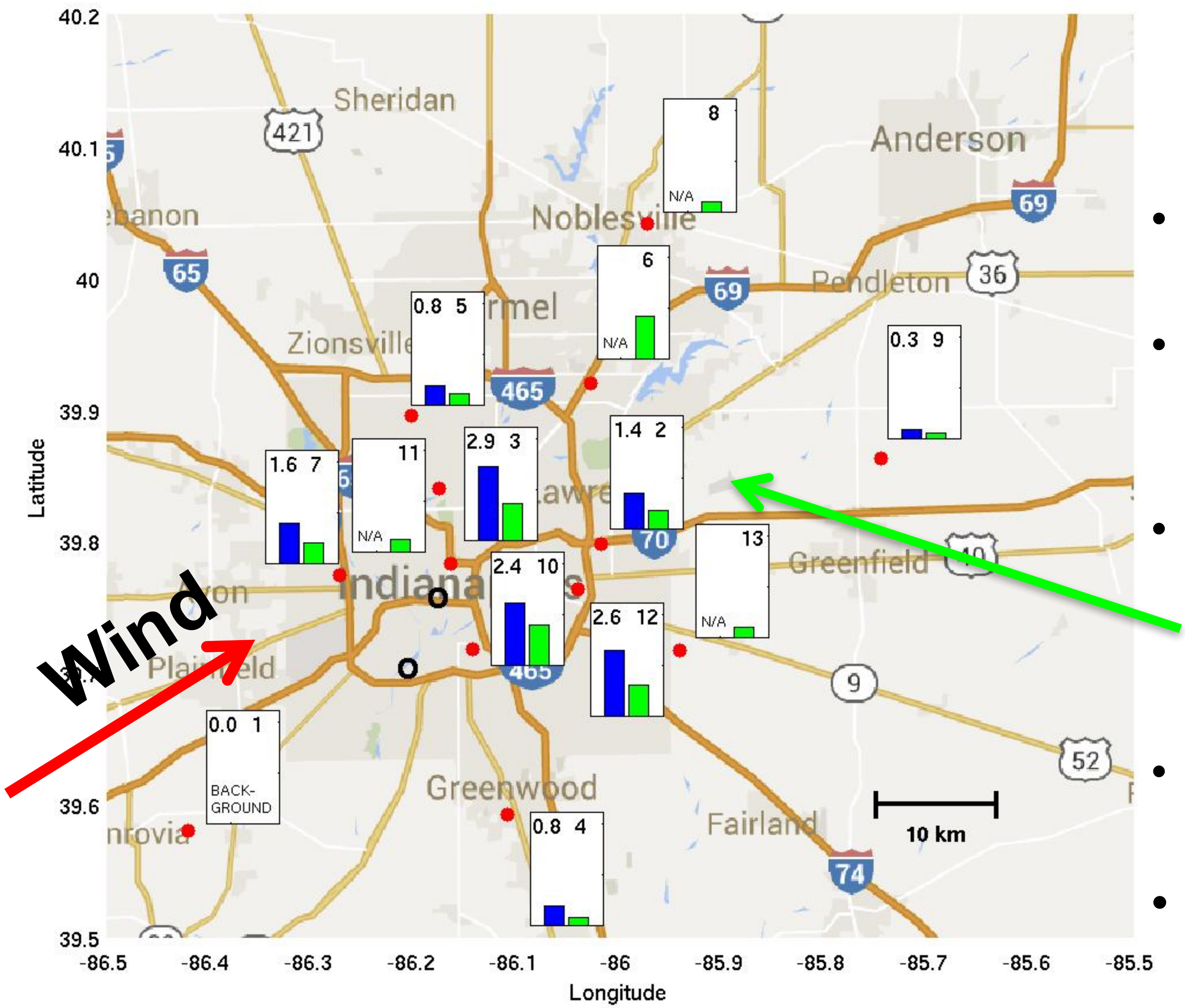


Combination of tower surface footprints with prior CO_2 emissions to generate modeled mixing ratios

Lauvaux et al, submitted

Spatial structure of urban CO_2: observed and modeled

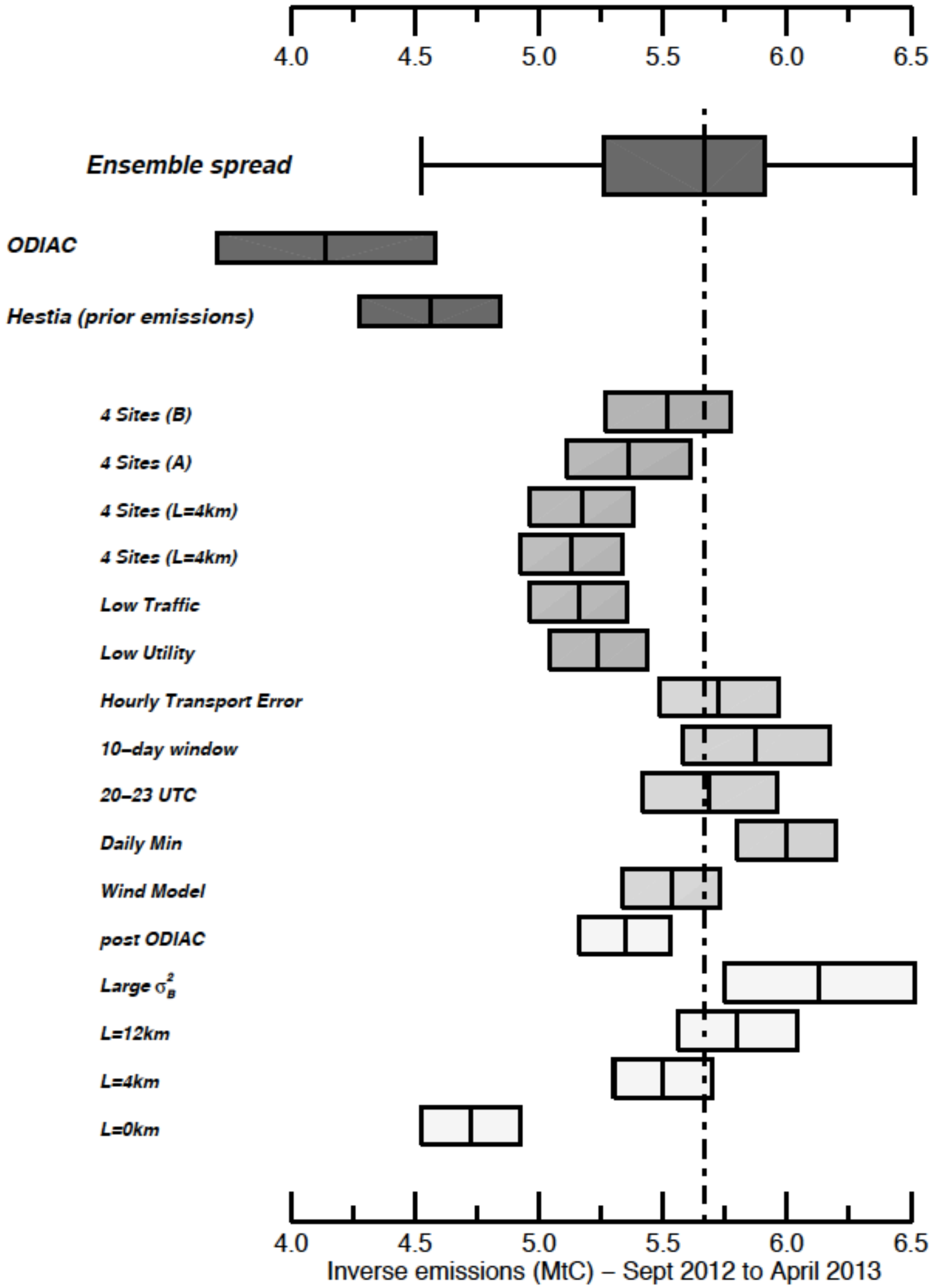
Blue = observed; Green = modeled



(observed CO_2 enhancement, tower number)

- Observed CO_2: afternoon values, averaged Jan-April 2013
- Site 09: 0.3 ppm larger than Site 01
- Site 03: measures larger [CO_2] by 3 ppm
- Modeled CO_2 using LPDM footprints and Hestia emissions
- Overall, the spatial structure is similar
- Miles et al., in prep

Indianapolis whole-city emissions



Sept12 – Apr13 Indianapolis CO2 emissions:

Hestia: 4.6 ktC

Inversion: 5.7 ktC +/- 0.2 ktC

Lauvaux et al, submitted

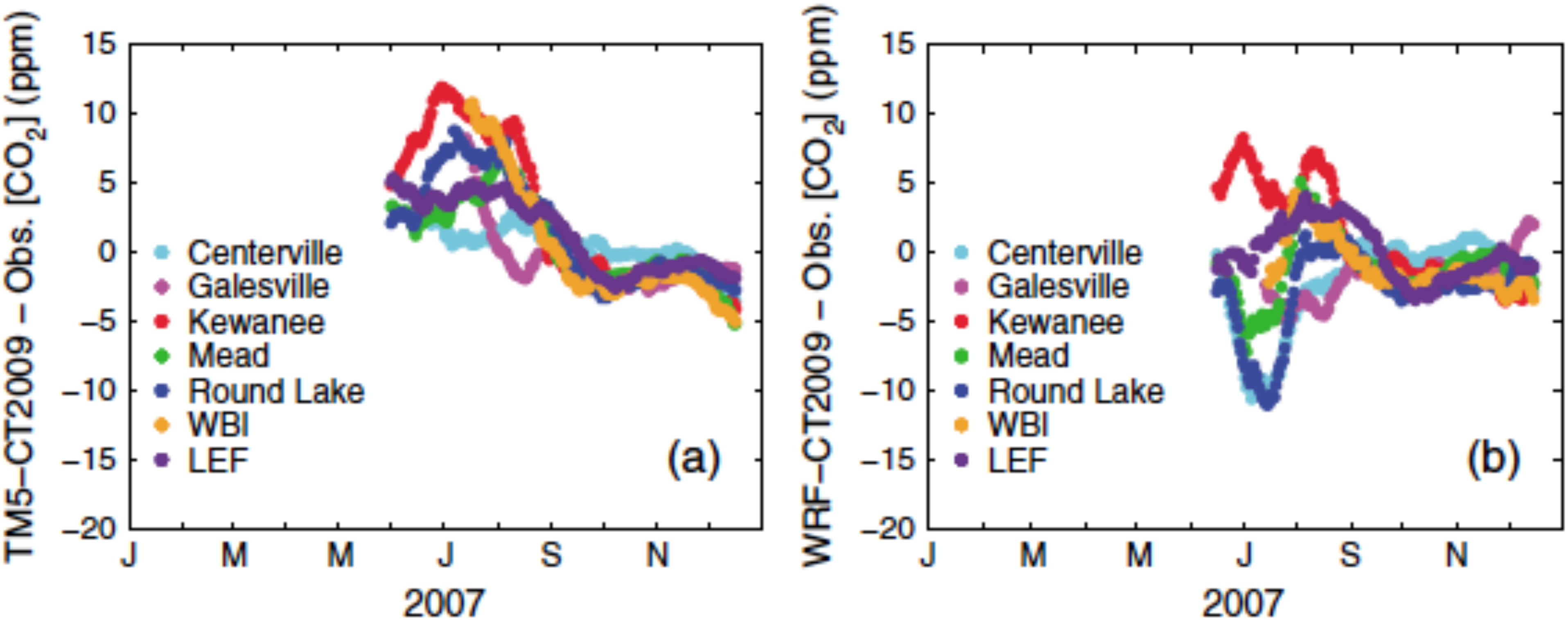
What can be done?

- Advance process understanding.
- Move towards multi-state data assimilation.
- Close the measurement methods gap.
- Apply more measurements to more models. Enter the era of networked observations and ensemble modeling.

Some examples

- The world isn't flat!
  - Improve models by studying the carbon-water-nitrogen cycles in complex terrain.
- Closing the measurement gap with atmospheric inversions:
  - Increase measurement density.
  - Reduce uncertainty in atmospheric transport.

Comparison – TM5 and WRF
How much does transport matter?

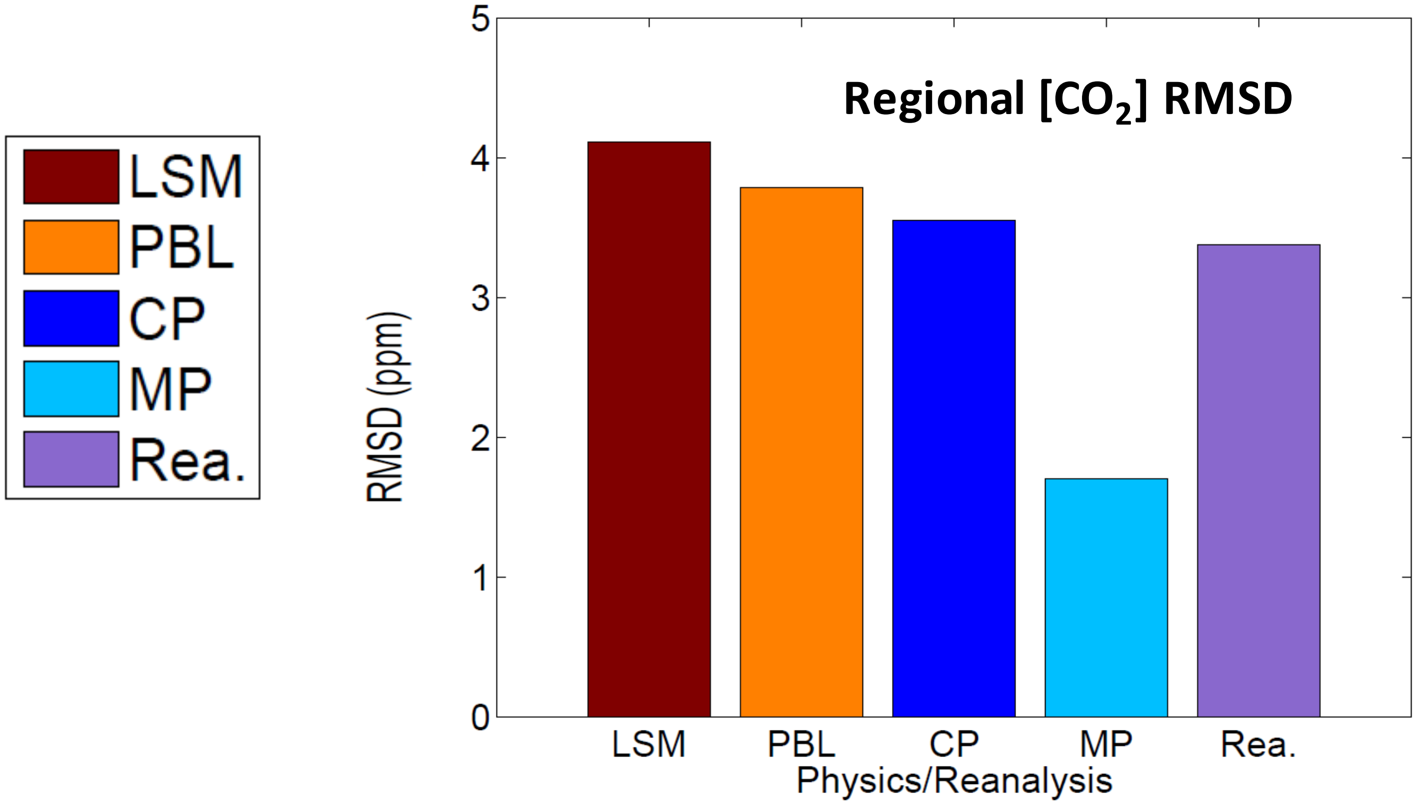


Identical CO_2 fluxes and boundary conditions.

Midsummer, monthly averaged ABL CO_2 differs by as much as 15 ppm due only to atmospheric transport.

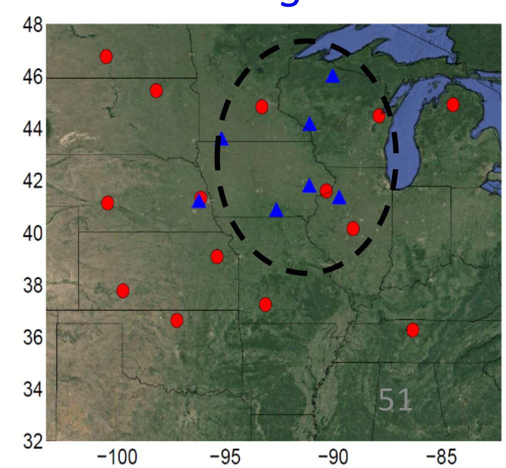
Diaz-Isaac et al, 2014, JGR-A.

Which Physics Parameterization Drives CO_2 Errors?



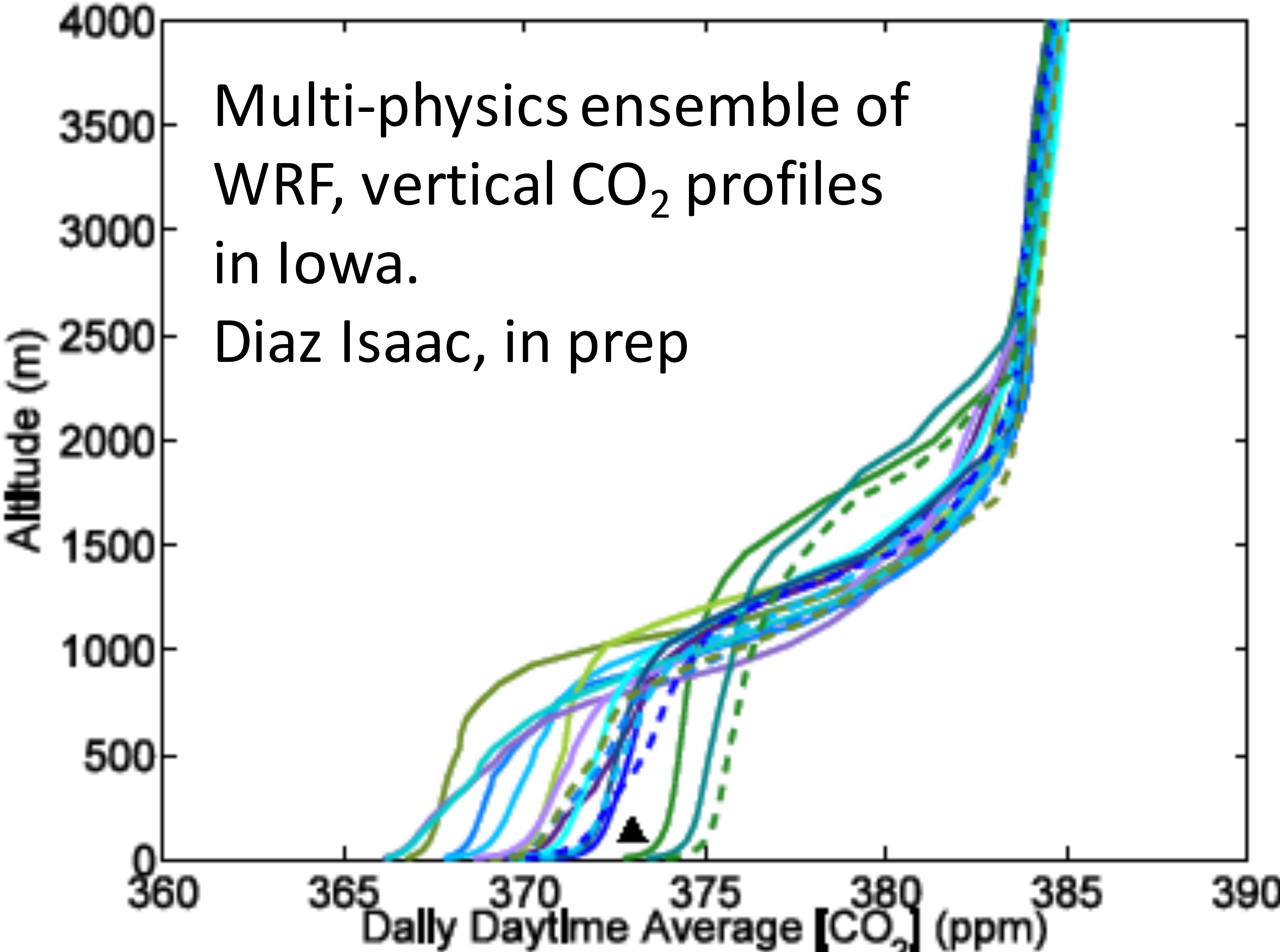
Diaz-Isaac et al, in prep

Sites: *blue triangles*



- *Model-Ensemble mean comparison used to isolate transport errors.*
- ***Local Scale:*** LSMs, PBL schemes and Cumulus parameterizations (CP) all have a big impact in CO_2 mole fraction errors.
- ***Regional scale:*** LSMs, PBL schemes, Cumulus parameterization(CP) and reanalysis have a big impact in CO_2 errors.
- ***PBL physics is not the only physics parameterization that matters.***

How much do CO_2 simulations vary within this ensemble?

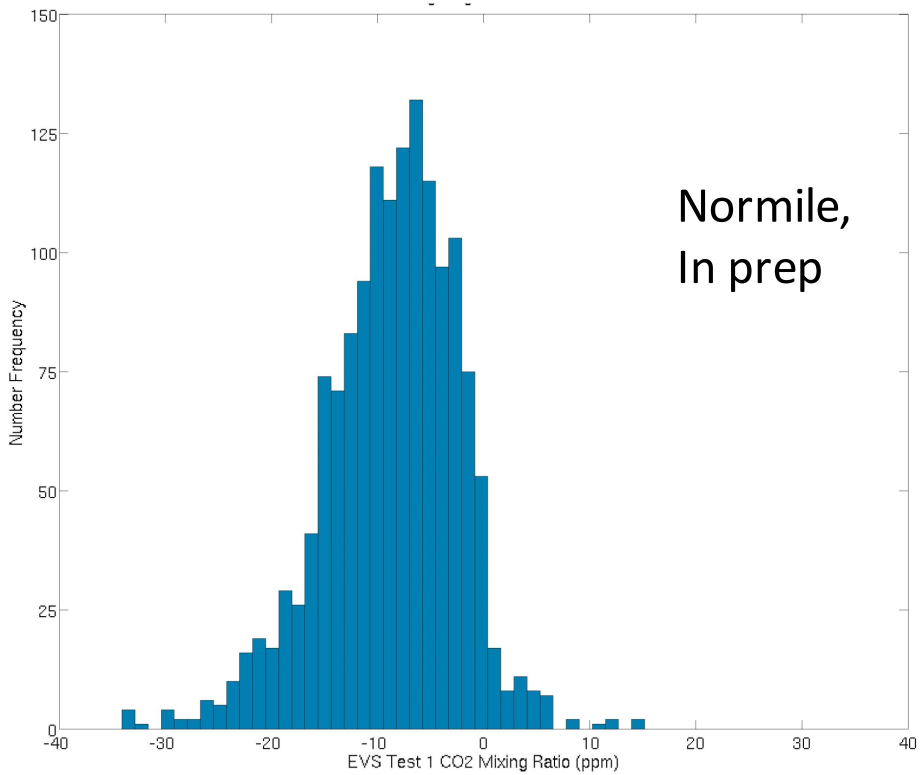


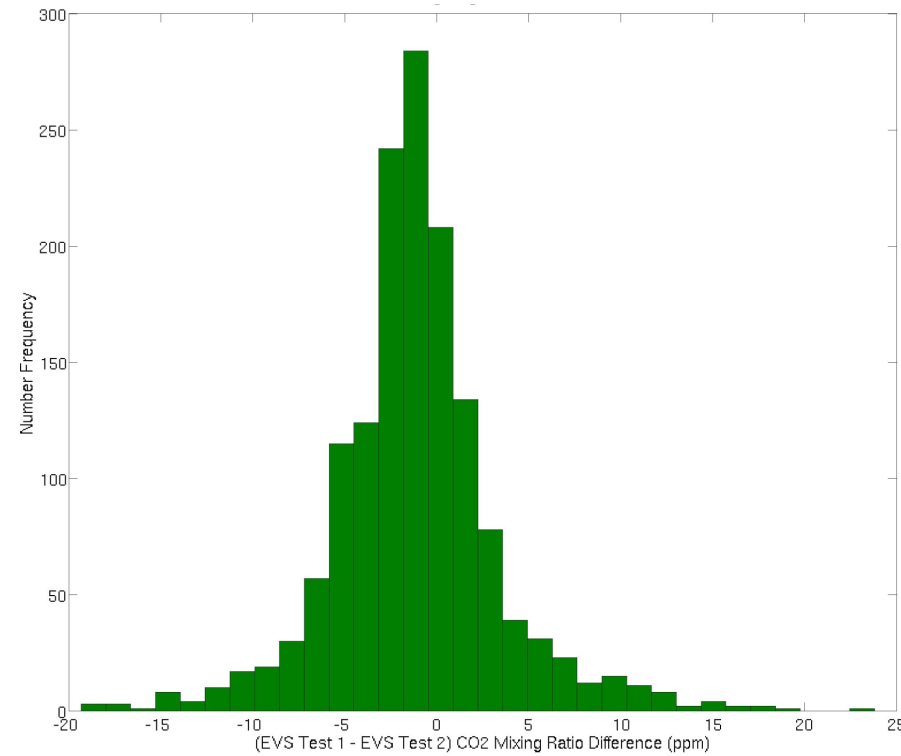
All members of the ensemble yield plausible atmospheric transport (comparison to winds, ABL depth, surface flux observations).

All physics parameterizations contribute significantly.

Diaz-Isaac, in prep

How large are transport differences compared to flux contributions? About 50% of the continental biological CO_2 signal





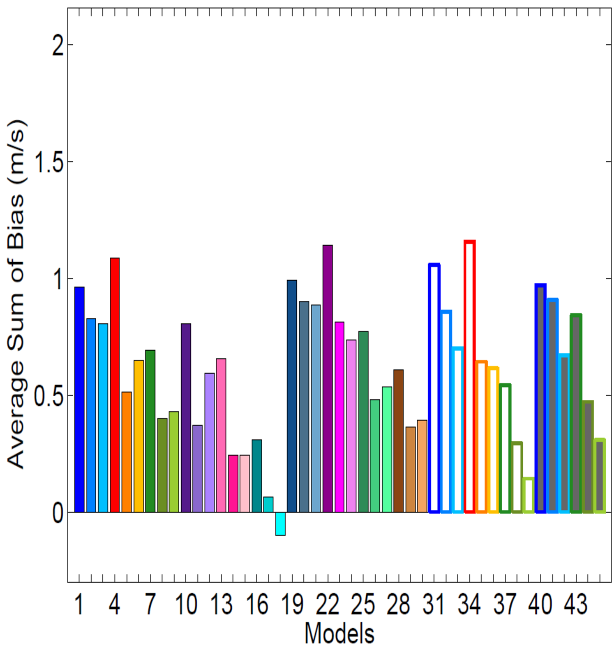
Eastern region site-to-site daytime ABL CO_2 contributions from continental biogenic fluxes. August, 2008. WRF, Carbon Tracker boundaries.

Eastern region site-to-site daytime ABL total CO_2 differences between two transport realizations. August, 2008. WRF, Carbon Tracker boundaries.

OK, transport matters. So what do we do about this?

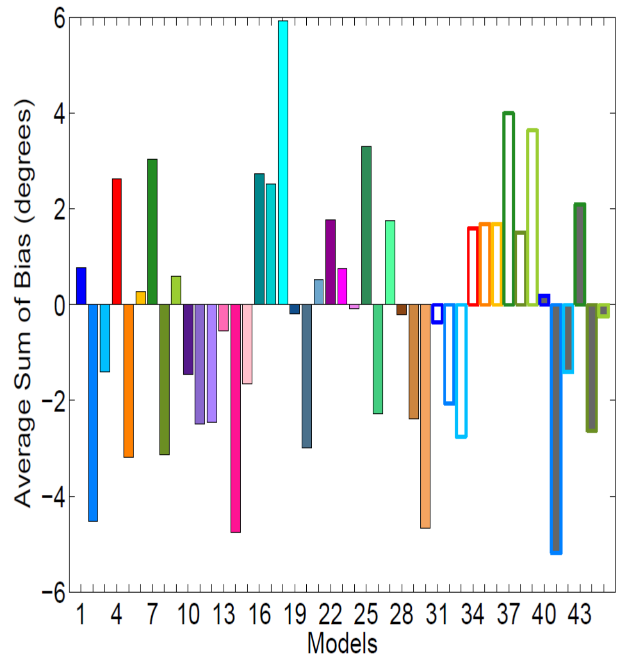
Are the models biased? Examination of a 45-member WRF parameterization ensemble

Wind Speed



- Regional wind speed ME is ***positive for all the configurations*** except one.
- Generally one ***PBL scheme (i.e., YSU)*** shows a ***higher ME*** than the rest.

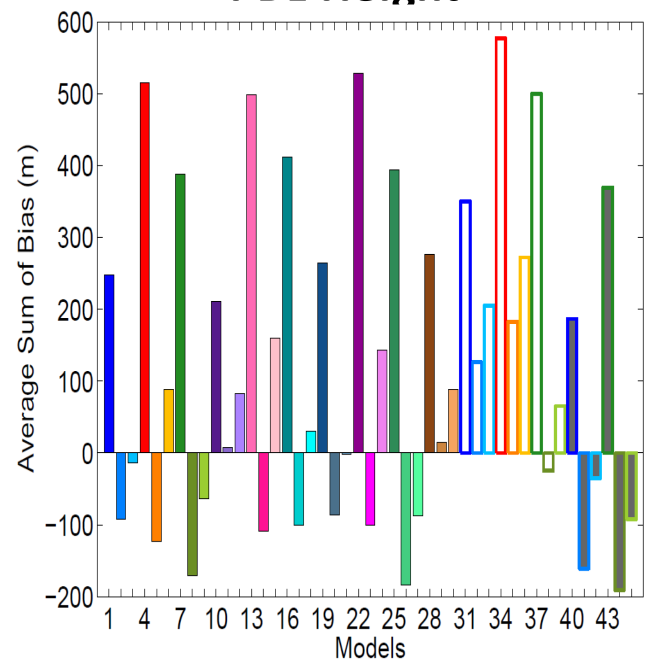
Wind Direction



- Regional wind direction mean error is highly variable across the different model configurations.

Diaz-Isaac et al, in prep

PBL Height



- Regional PBL height ME is highly variable across the different model configurations.
- Generally one ***PBL scheme (i.e., YSU) and LSM (i.e., RUC)*** shows a ***higher ME***.

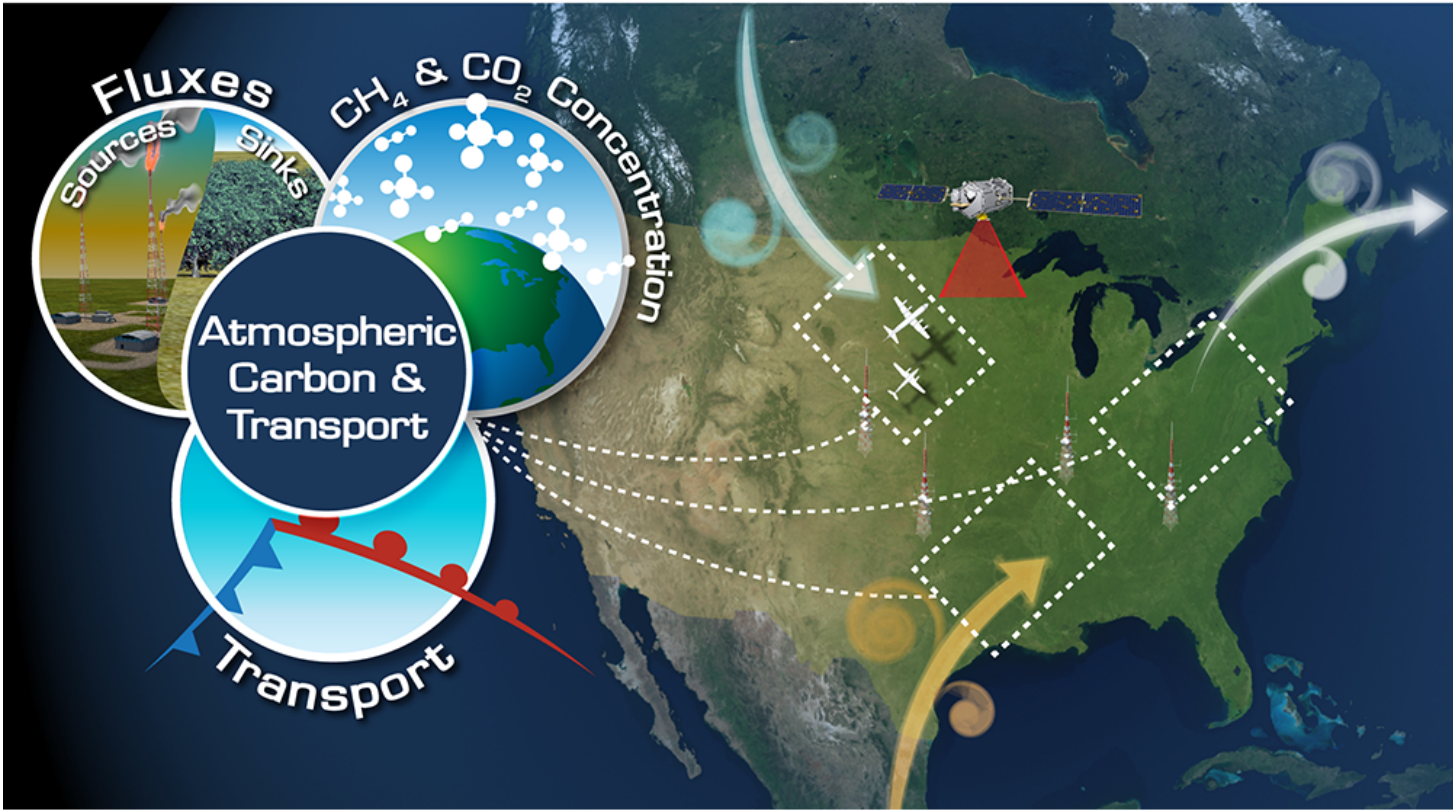
Atmospheric Carbon and Transport – America

A new NASA Earth Venture mission dedicated to improving the accuracy, precision and resolution of atmospheric inverse estimates of CO_2 and CH_4 sources and sinks

Kenneth Davis[1], David Baker[2], John Barrick[3], Joseph Berry[4], Kevin Bowman[5], Edward Browell[3], Lori Bruhwiler[6], Gao Chen[3], George Collatz[7], Robert Cook[8], Scott Denning[2], Jeremy Dobler[9], Syed Ismail[3], Andrew Jacobson[6], Anna Karion[6], Thomas Lauvaux[5], Bing Lin[3], Matt McGill[7], Byron Meadows[3], Anna Michalak[4], Natasha Miles[1], John Miller[6], Berrien Moore[10], Amin Nehrir[3], Lesley Ott[7], Michael Obland[3], Christopher O'Dell[2], Stephen Pawson[7], Gabrielle Petron[6], Andrew Schuh[2], Colm Sweeney[6], Pieter Tans[6], Yaxing Wei[8], and Melissa Yang[3]

[1]The Pennsylvania State University, [2]Colorado State University, [3]NASA Langley Research Center, [4]Carnegie Institution of Stanford, [5]NASA Jet Propulsion Lab, [6]NOAA ESRL/University of Colorado, [7]NASA Goddard Space Flight Center, [8]Oak Ridge National Lab, [9]Exelis, Inc., [10]University of Oklahoma

ACT-America



http://act-america.larc.nasa.gov/

Image credit: Tim Marvel / NASA Langley

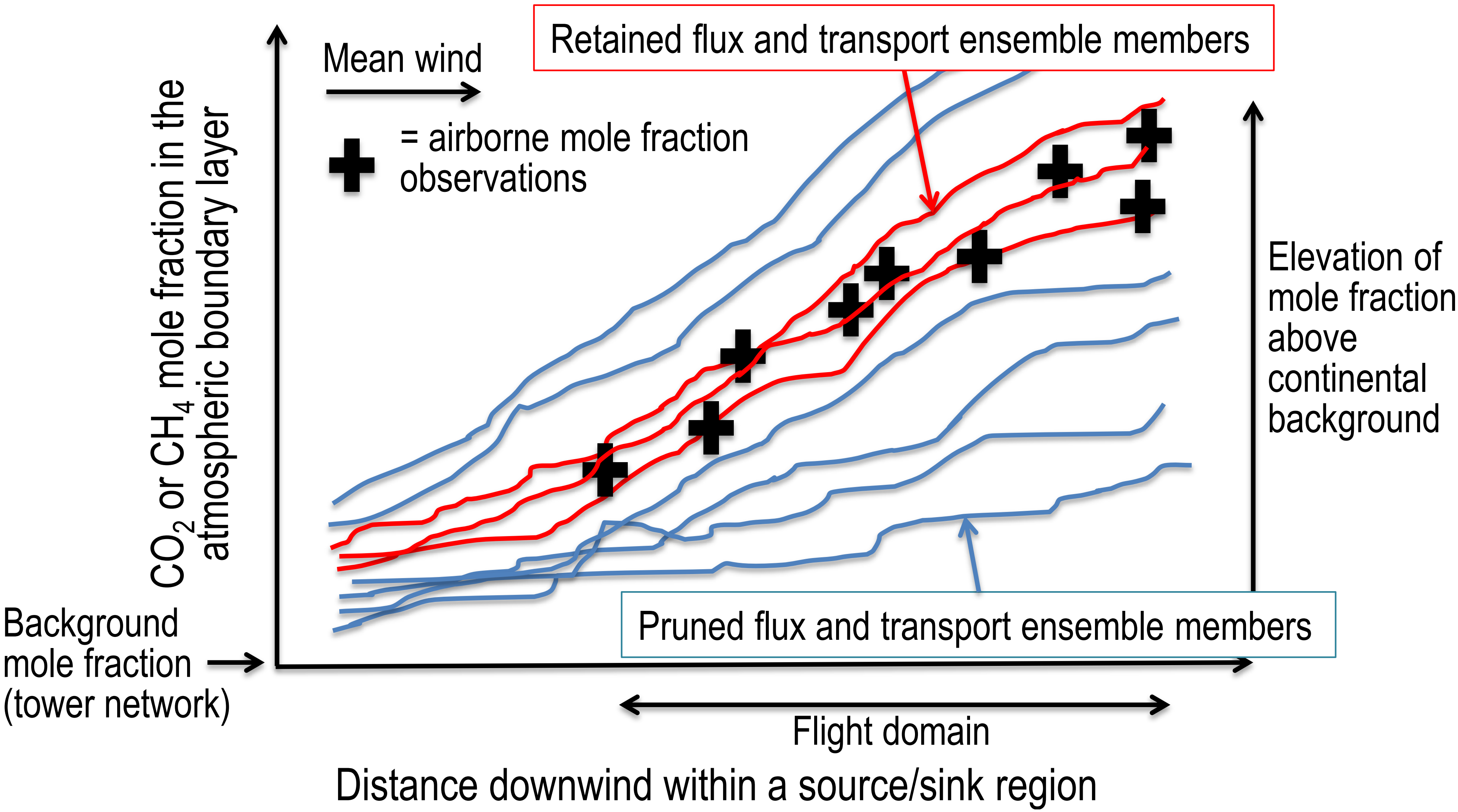
ACT-America Mission Objectives

1. Quantify and reduce atmospheric transport uncertainties
2. Improve regional-scale, seasonal prior estimates of CO_2 and CH_4 fluxes
3. Evaluate the sensitivity of Orbiting Carbon Observatory-2 (OCO-2) column CO_2 measurements to regional variability in tropospheric CO_2

These goals address the three primary sources of uncertainty in atmospheric inversions – transport error, prior flux uncertainty and limited data density

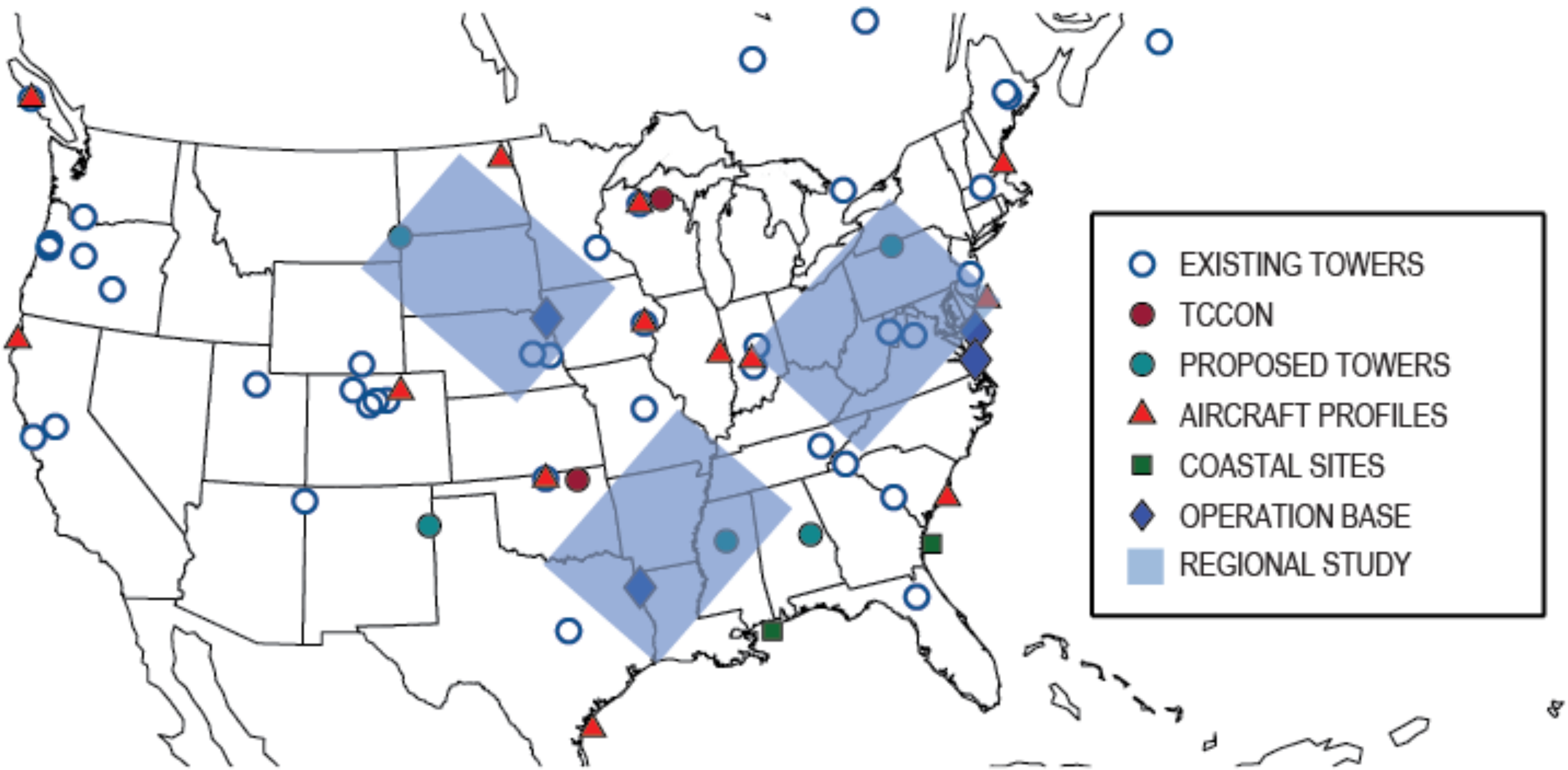
Imagine air flowing across a landscape that is a source of GHGs, and aircraft data tracking the changes in GHG mole fraction across the landscape…

Simplified vision of model (flux and transport) ensemble pruning using airborne observations



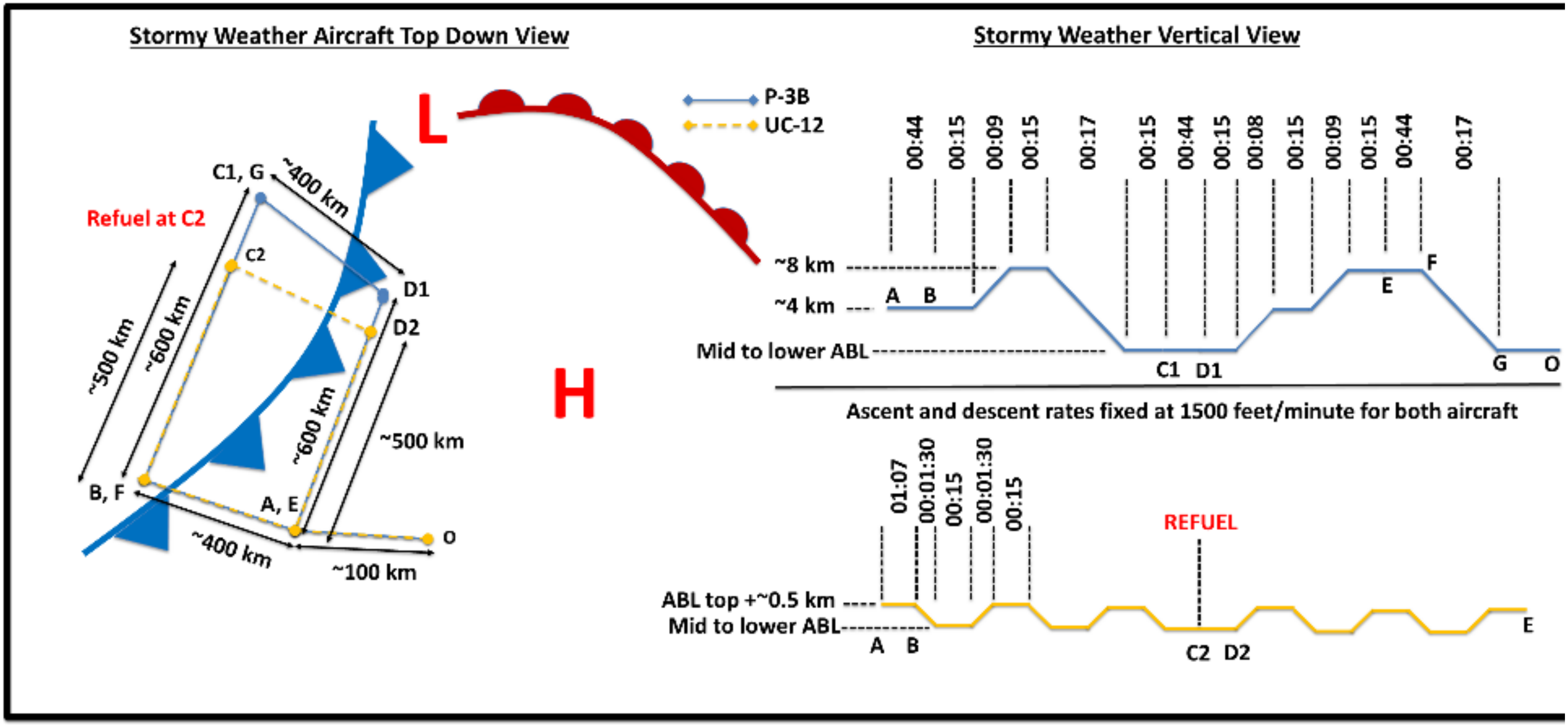
Pruned ensembles lead to more accurate and precise flux inversions using long-term GHG data (towers, flasks, satellite, NOAA airborne profiling.

Where?



The eastern half of the United States, a region that includes a highly productive biosphere, vigorous agricultural activity, extensive gas and oil extraction, dynamic, seasonally varying weather patterns and the most extensive GHG and meteorological observing networks on Earth, serves as an ideal setting for the ACT-America mission.

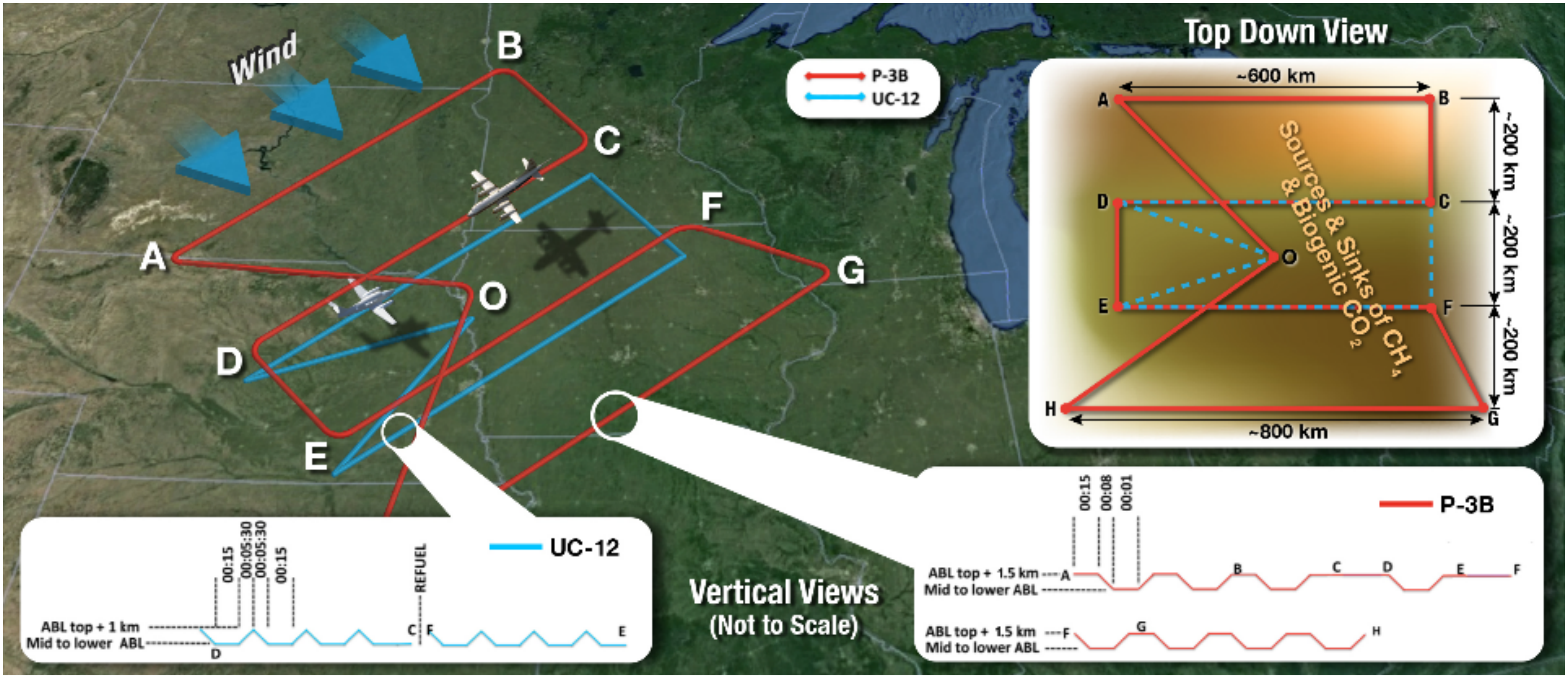
Stormy-weather (transport-dominated) flight plans (objective 1)



- Measure atmospheric state, CO_2, CH_4 and tracers (CO, $^{14}CO_2$, O_3) across and around frontal systems.
- Evaluate atmospheric transport in our model ensemble. Prune transport ensemble.

Fair-weather (flux-dominated) flight plan (objectives 1 and 2)

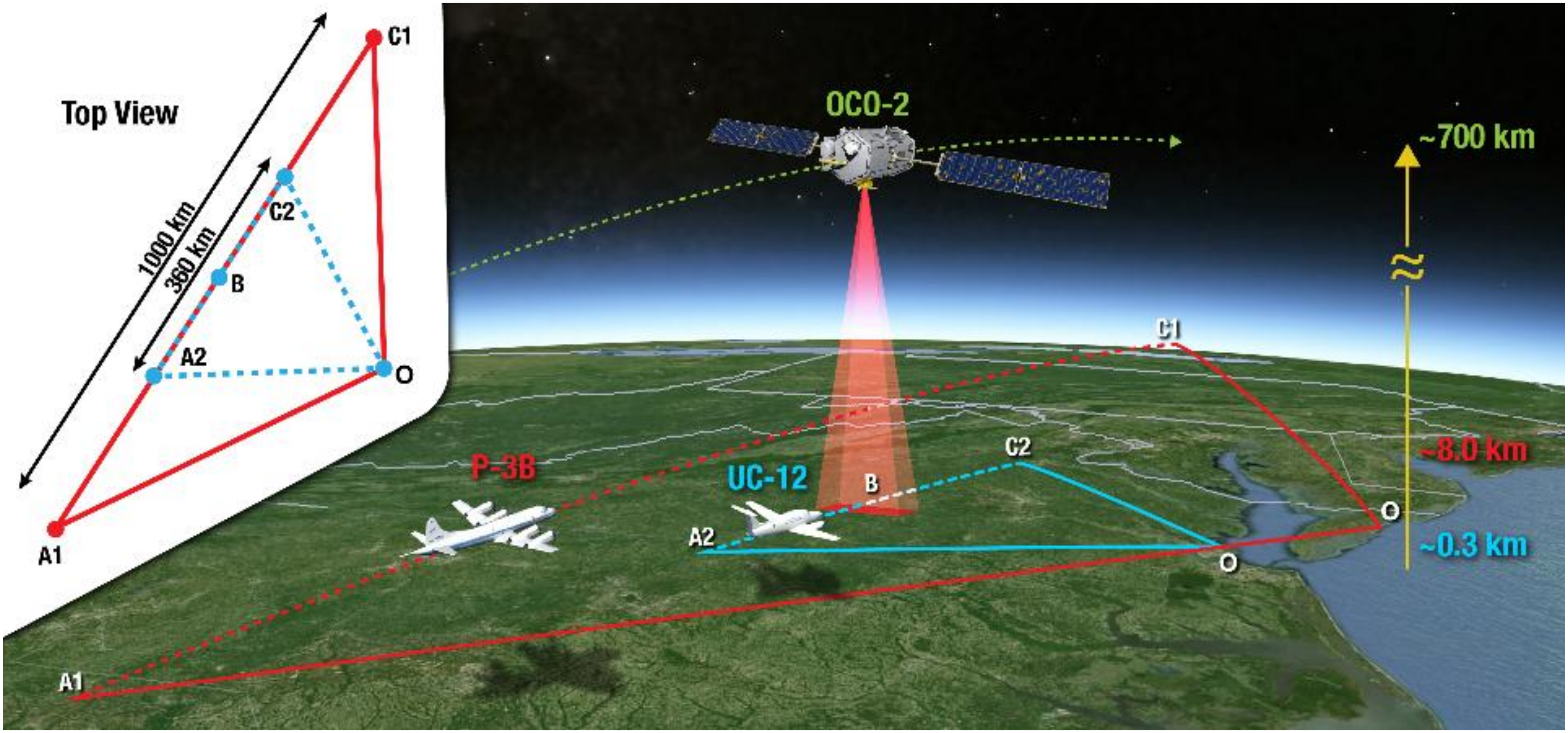
Tim Marvel, NASA Langley



- Measure winds, ABL depth, CO_2, CH_4 and tracers (CO, $^{14}CO_2$, O_3) across 100's of km.
- Solve for regional fluxes for the days of flights directly – prune prior flux estimates.
- Evaluate fair weather meteorology in atmospheric transport ensemble

OCO-2 under-flights (objective 3)

Tim Marvel, NASA Langley



- Measure much of the atmospheric CO_2 column at < 20km horizontal resolution across 100's of km below OCO-2. Also measure aerosols, clouds with lidar.
- Compare spatial variability in airborne CO_2 to OCO-2 CO_2. Evaluate OCO-2 ability to capture tropospheric CO_2 variability along-track.

Flight Campaign Schedules

Season/ Year	Spr 2016	Sum 2016	Fall 2016	Win 2017	Spr 2017	Sum 2017	Fall 2017	Win 2018	Spr 2018	Sum 2018	Fall 2018
First choice	C-130 test flights	X		X		X	X		X		
Fallback option 1	C-130 test flights	X		X			X		X	X	
Fallback option 2	C-130 test flights	X		X		X			X		X
Fallback option 3	C-130 test flights		X		X	X		X		X	

Proposed start date: *Summer 2016, given timeline for C-130 modifications and aircraft access for flight testing*.

Year 1 (2015): Instrument aircraft, integrate modeling systems, perform flight design simulations. *Work with pre-existing aircraft data sets.*

Years 2-4 (2016-18): Flight campaigns and analyses. Goals 1-3.

Year 5 (2019): Wrap up goals 1-3. Apply findings to a multi-year reanalysis of N. American C fluxes using long-term observational assets (i.e., demonstrate new atmospheric inversion system).

End date: 20 Jan, 2020.

Overarching Goal

- The overarching goal of the Atmospheric Carbon and Transport-America (ACT-America) mission is to improve regional to continental scale diagnoses of carbon dioxide (CO_2) and methane (CH_4) sources and sinks.
- The mission will enable and demonstrate a new generation of atmospheric inversion systems for quantifying atmospheric CO_2 and CH_4 fluxes.
- These inverse flux estimates will be able to:
  - Evaluate and improve terrestrial carbon cycle models, and
  - Monitor carbon fluxes to support climate-change mitigation efforts.

conclusions

- Carbon cycle science is in its early stages as a predictive, data-rich science.
- Basic process understanding needs to be improved.
- Multi-state observations and ensemble modeling is being introduced.
- Time is short. The time for management is now.

Thanks!