STATISTICAL AND COMPUTATIONAL CHALLENGES OF CONSTRAINING GREENHOUSE GAS BUDGETS

Anna M. Michalak

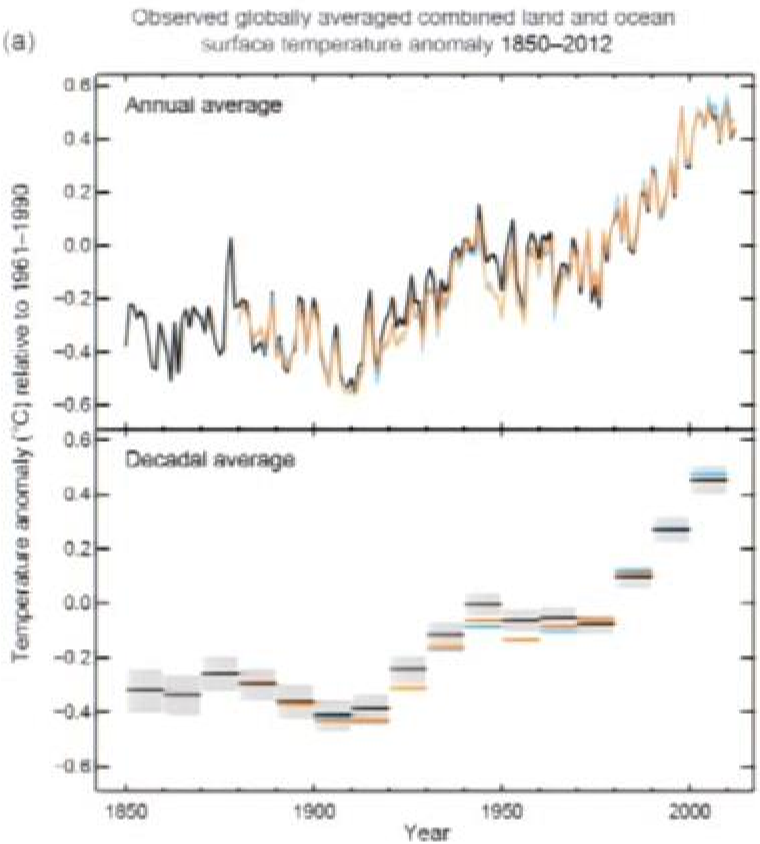
Department of Global Ecology, Carnegie Institution for Science

Department of Environmental Earth Systems Science, Stanford University

CARNEGIE INSTITUTION FOR SCIENCE | DEPARTMENT OF GLOBAL ECOLOGY

Take home messages

- The need to constrain greenhouse gas budgets inevitably leads to the need for the solution of inverse problems
- These inverse problems:
  - Require (intelligently) choosing among many uncomfortable assumptions
  - Are becoming increasingly statistically sophisticated and computationally demanding
  - Done carefully, can lead to fundamental insights with management and policy implications


(a)
Observed globally averaged combined land and ocean surface temperature anomaly 1850–2012
Annual average
Decadal average
Temperature anomaly (°C) relative to 1961–1990
Year

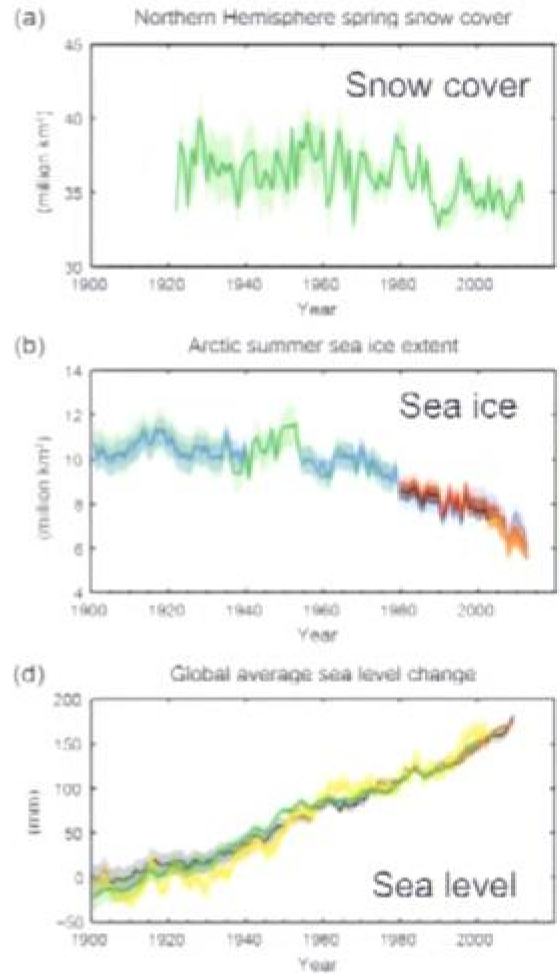
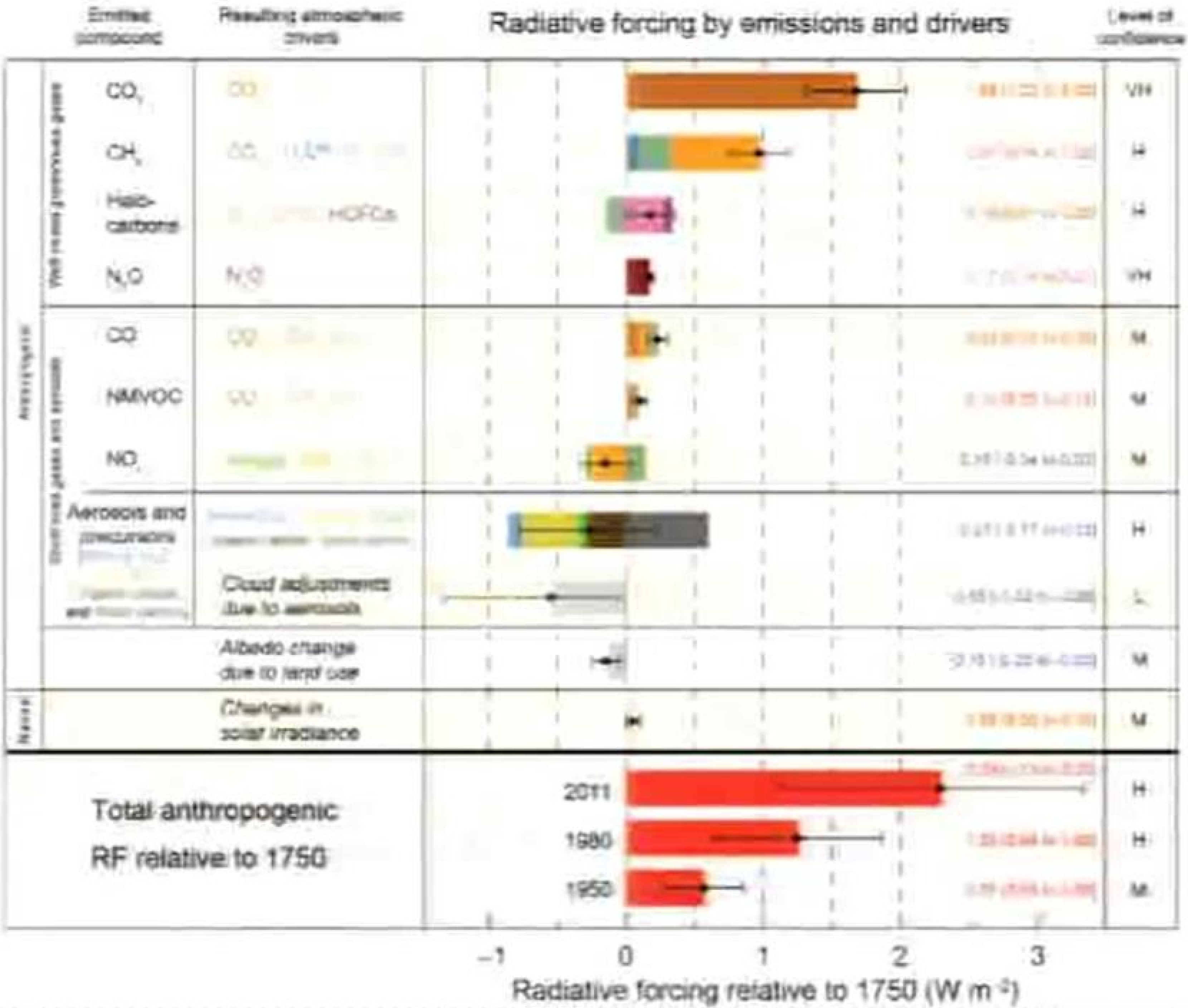
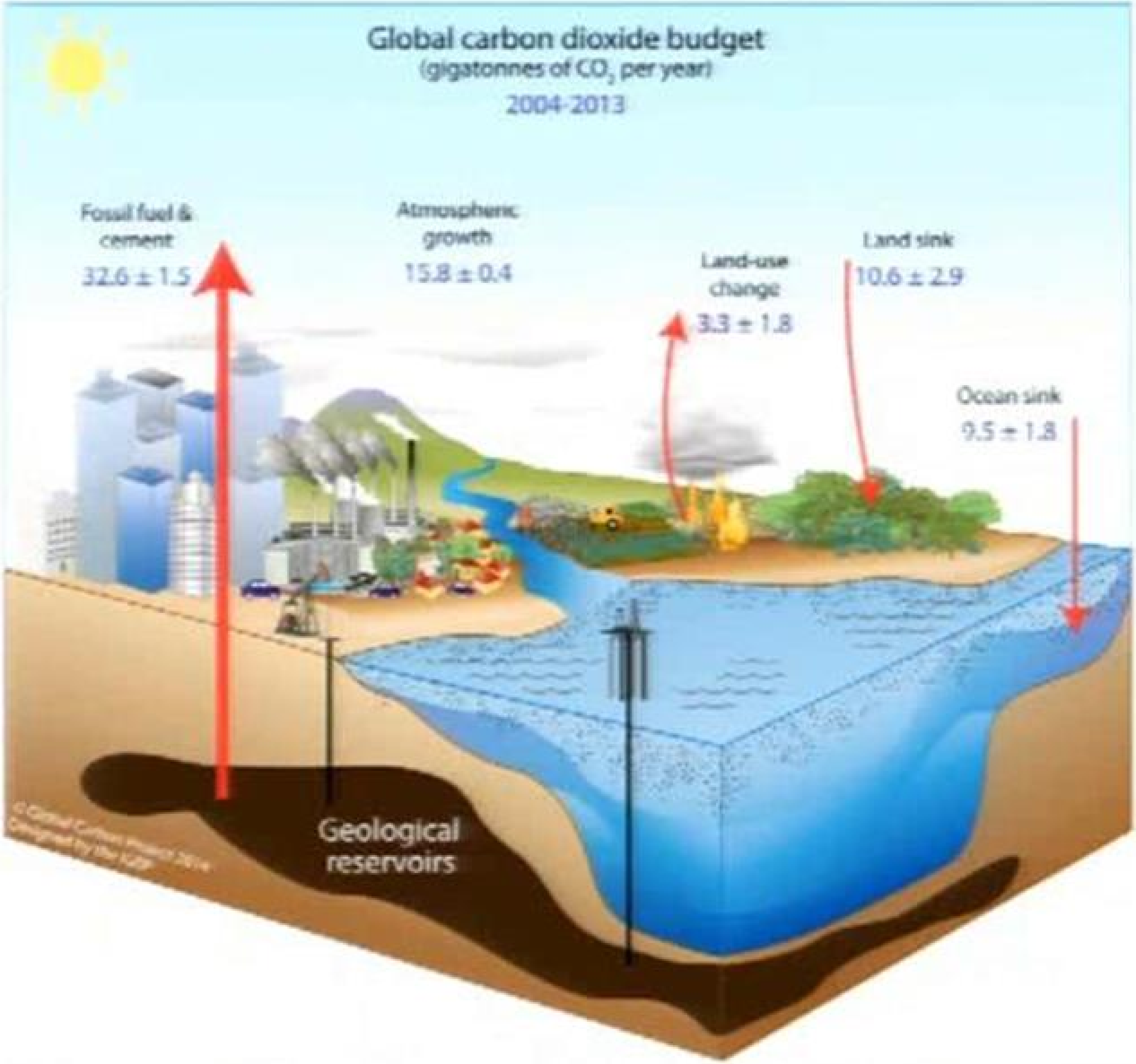

(a)
Northern Hemisphere spring snow cover
Snow cover
(million km²)
Year
(b)
Arctic summer sea ice extent
Sea ice
(million km²)
Year
(d)
Global average sea level change
Sea level
(mm)
Year

Photo courtesy of Bartshé Miller

Tioga Pass, January 12 2015

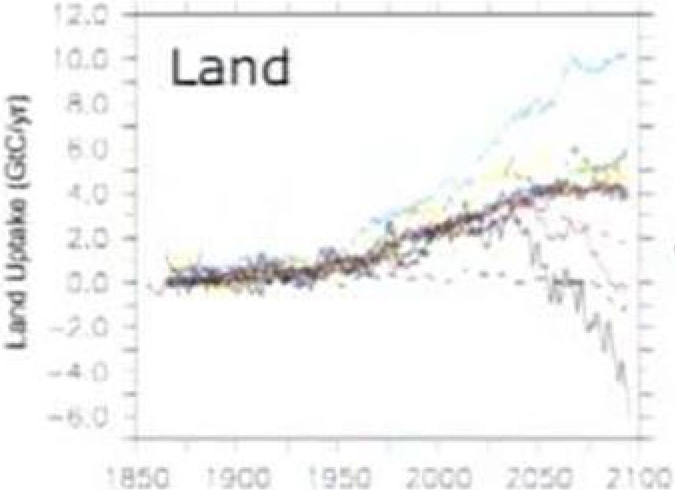

Radiative forcing by emissions and drivers
CO_2
CH_4
Halo-carbons
HCFCs
N_2O
CO
NMVOC
NO_x
Aerosols and precursors
Cloud adjustments due to aerosols
Albedo change due to land use
Changes in solar irradiance
VH
H
M
L
Total anthropogenic RF relative to 1750
2011
1980
1950
−1
0
1
2
3
Radiative forcing relative to 1750 (W m^{-2})

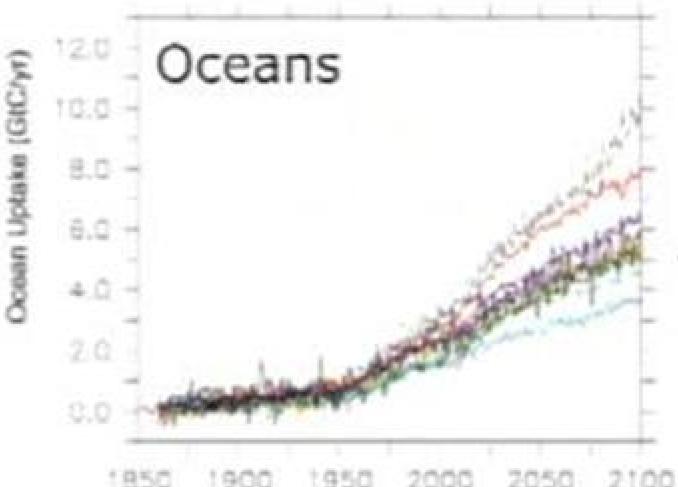
Source: CDIAC; NOAA-ESRL; Le Quéré et al 2014; Global Carbon Budget 2014



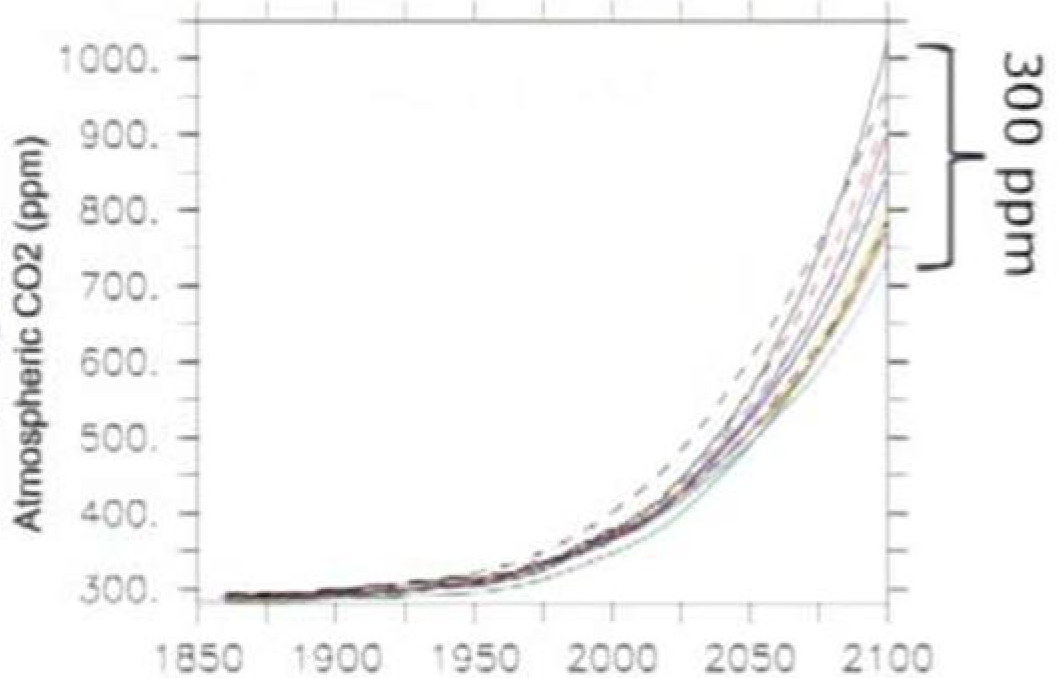
Perturbation of the global carbon cycle caused by anthropogenic activities, averaged globally for the decade 2004–2013 ($GtCO_2$/yr)

The future of natural carbon sinks





Uncertainty associated with the future of natural carbon sinks is one of three major sources of uncertainty in future climate projections



Source: Friedlingstein et al. (2006) showing projections from coupled carbon and climate simulations for several models.

The White House

Office of the Press Secretary

For Immediate Release November 11, 2014

FACT SHEET: U.S.-China Joint Announcement on Climate Change and Clean Energy Cooperation

President Obama Announces Ambitious 2025 Target to Cut U.S. Climate Pollution by 26-28 Percent from 2005 Levels

Building on strong progress during the first six years of the Administration, today President Obama announced a new target to cut net greenhouse gas emissions 26-28 percent below 2005 levels by 2025. At the same time, President Xi Jinping of China announced targets to peak CO_2 emissions around 2030, with the intention to try to peak early, and to increase the non-fossil fuel share of all energy to around 20 percent by 2030.

Global Carbon Project 2013; Figure based on Kirschke et al. 2013, Nature Geoscience

METHANE BUDGET : 2000-09

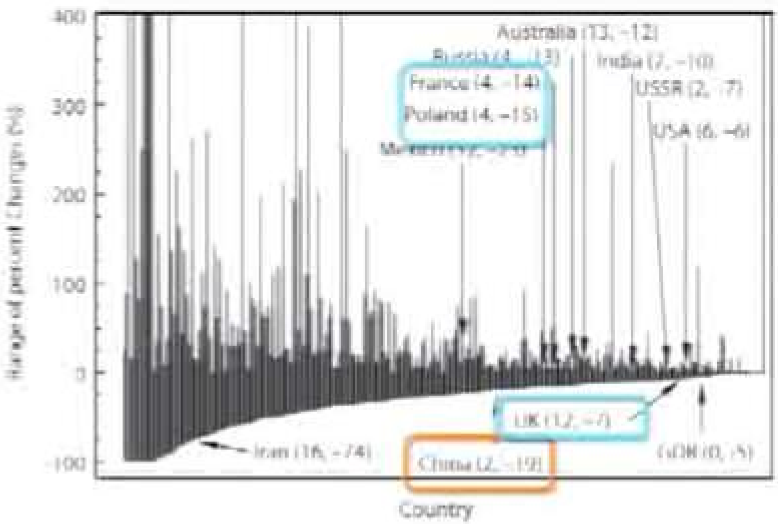
THE WHITE HOUSE
WASHINGTON

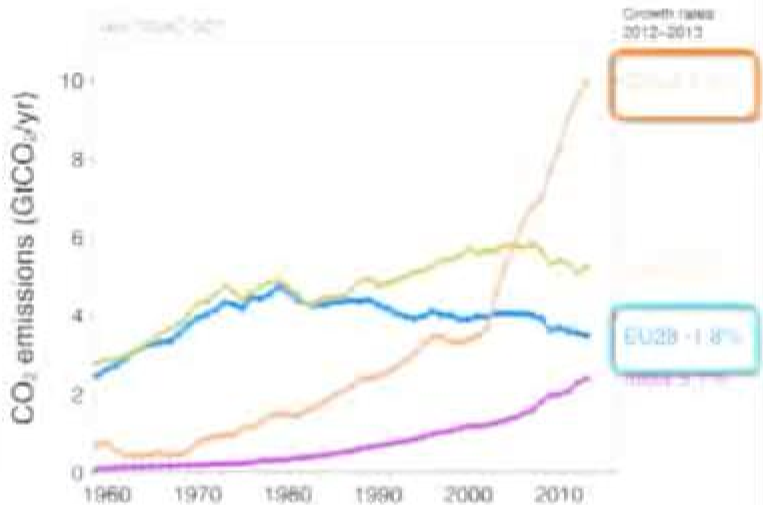
CLIMATE ACTION PLAN
STRATEGY TO REDUCE METHANE EMISSIONS

MARCH 2014

Andres et al. (2014); Global Carbon Project

How do we know emissions?
Self reporting

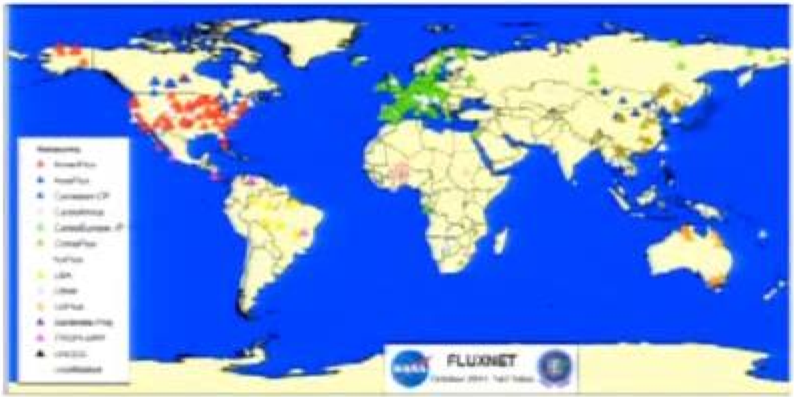


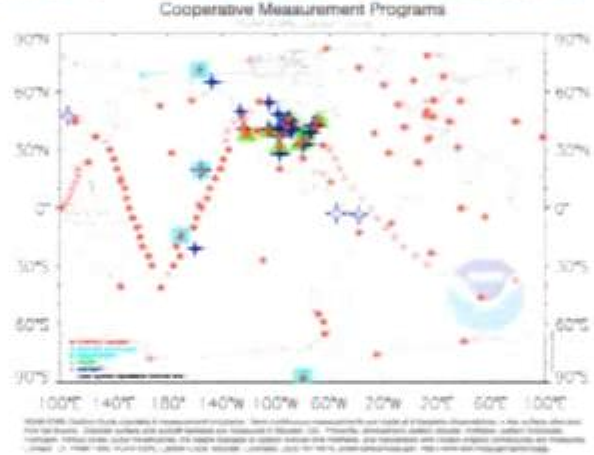


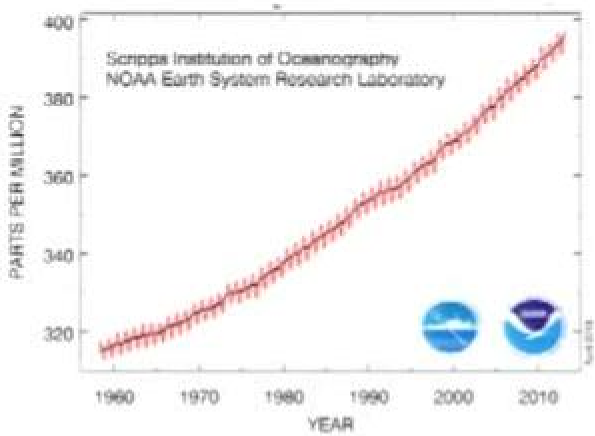
How do we know emissions?

Inventories

How do we know emissions? Observations





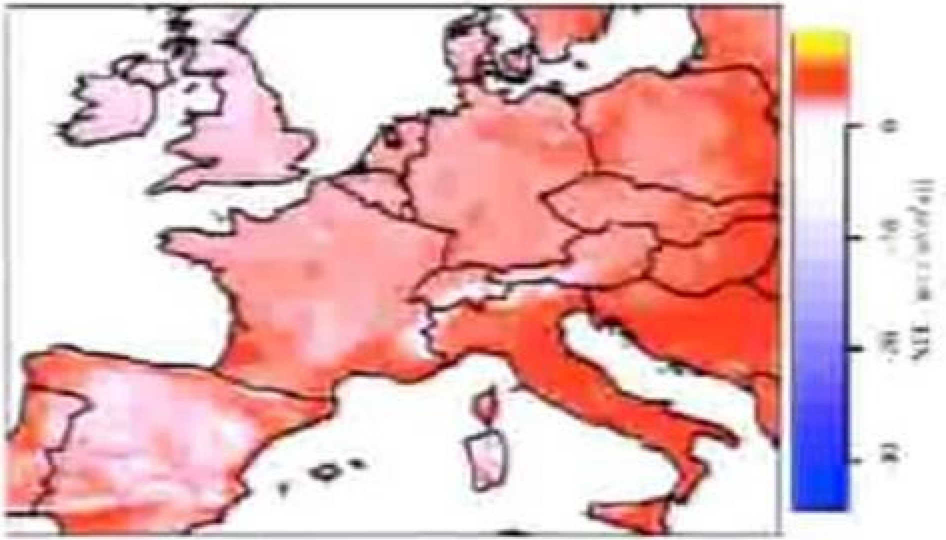


Fluxes (i.e. emissions / uptake)

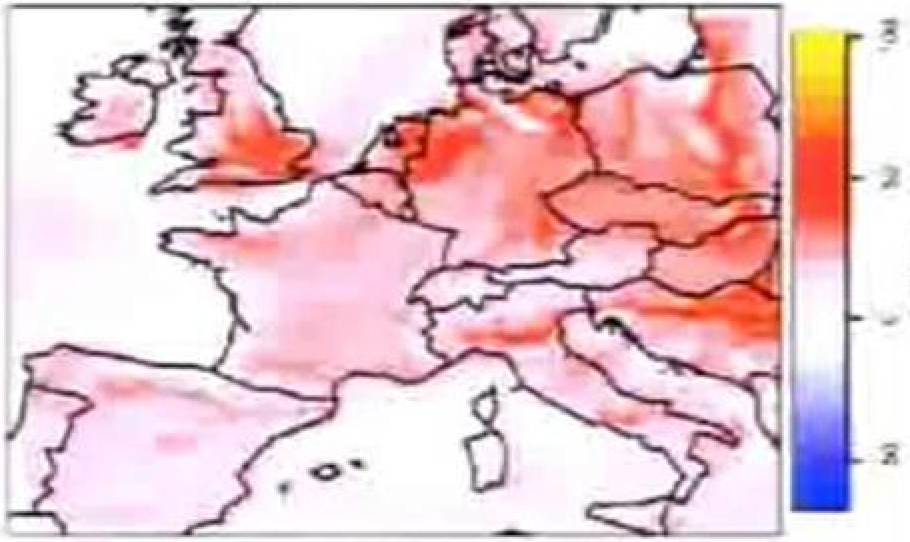
Concentrations

Source: Christoph Gerbig, MPI-BGC

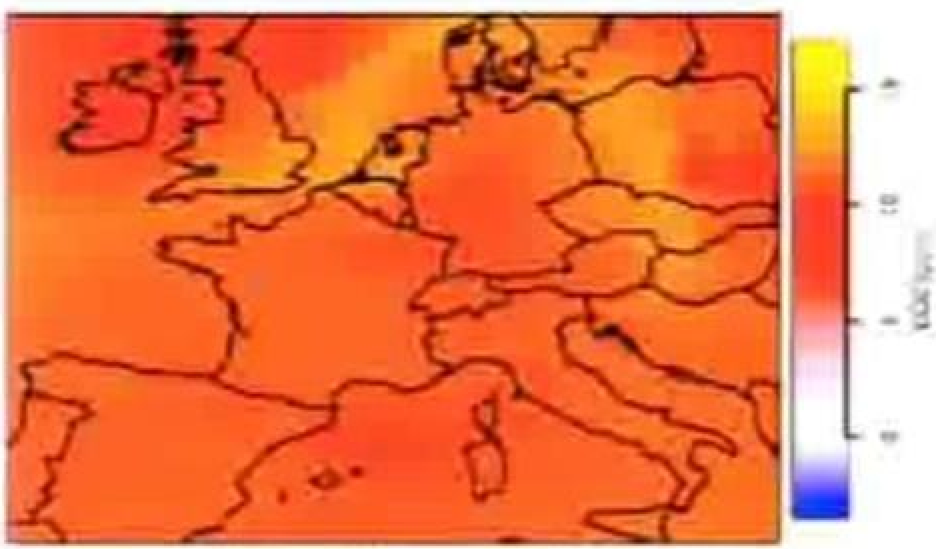
Net Ecosystem Exchange, time 2003-07-02_01:00:00



CO2 at 0,1 km, time 2003-07-02_00:00:00



column average CO2, time 2003-07-02_00:00:00



Take home messages

- The need to constrain greenhouse gas budgets inevitably leads to the need for the solution of inverse problems
- These inverse problems:
  - Require (intelligently) choosing among many uncomfortable assumptions
  - Are becoming increasingly statistically sophisticated and computationally demanding
  - Done carefully, can lead to fundamental insights with management and policy implications

Overall inverse problem

$$y = h(z) + \varepsilon_y + \varepsilon_h + \varepsilon_{rep} + \varepsilon_{agg}$$

- Find z given y, where:

y: atmospheric concentration observations (some places, some times)

z: surface fluxes (everywhere, all the time)

$h(.)$: atmospheric transport

ε_y: measurement error

ε_h: atmospheric transport model error

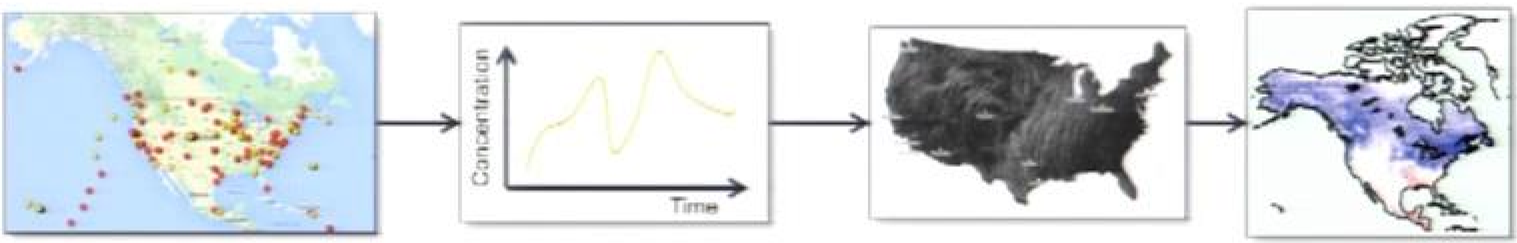
ε_{rep}: "representation" error (finite resolution in y)

ε_{agg}: "aggregation" error (finite resolution in z)

Overall inverse problem

All vary in space and time

$$y = h(z) + \varepsilon_y + \varepsilon_h + \varepsilon_{rep} + \varepsilon_{agg}$$



y $h(.)$ z

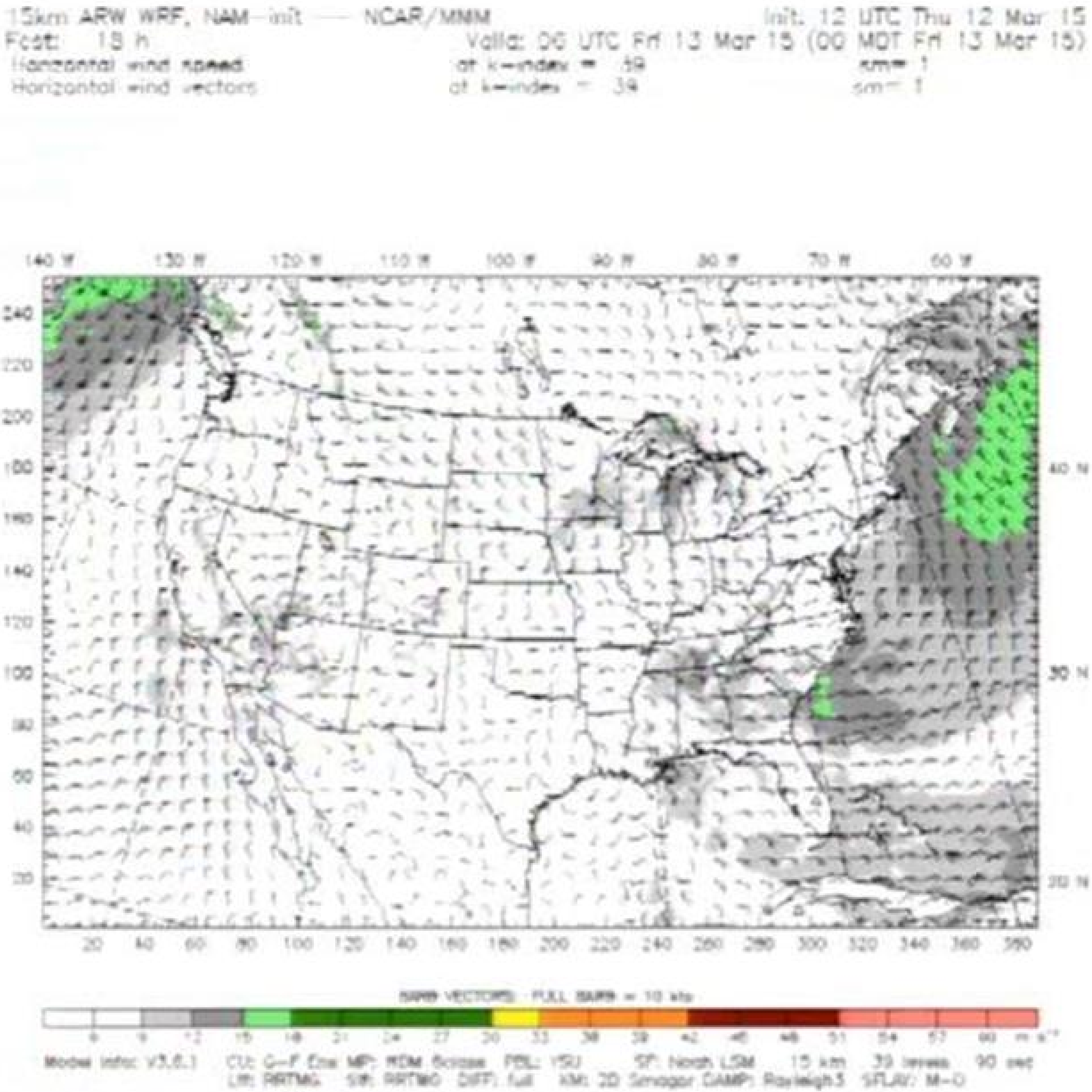
Causation

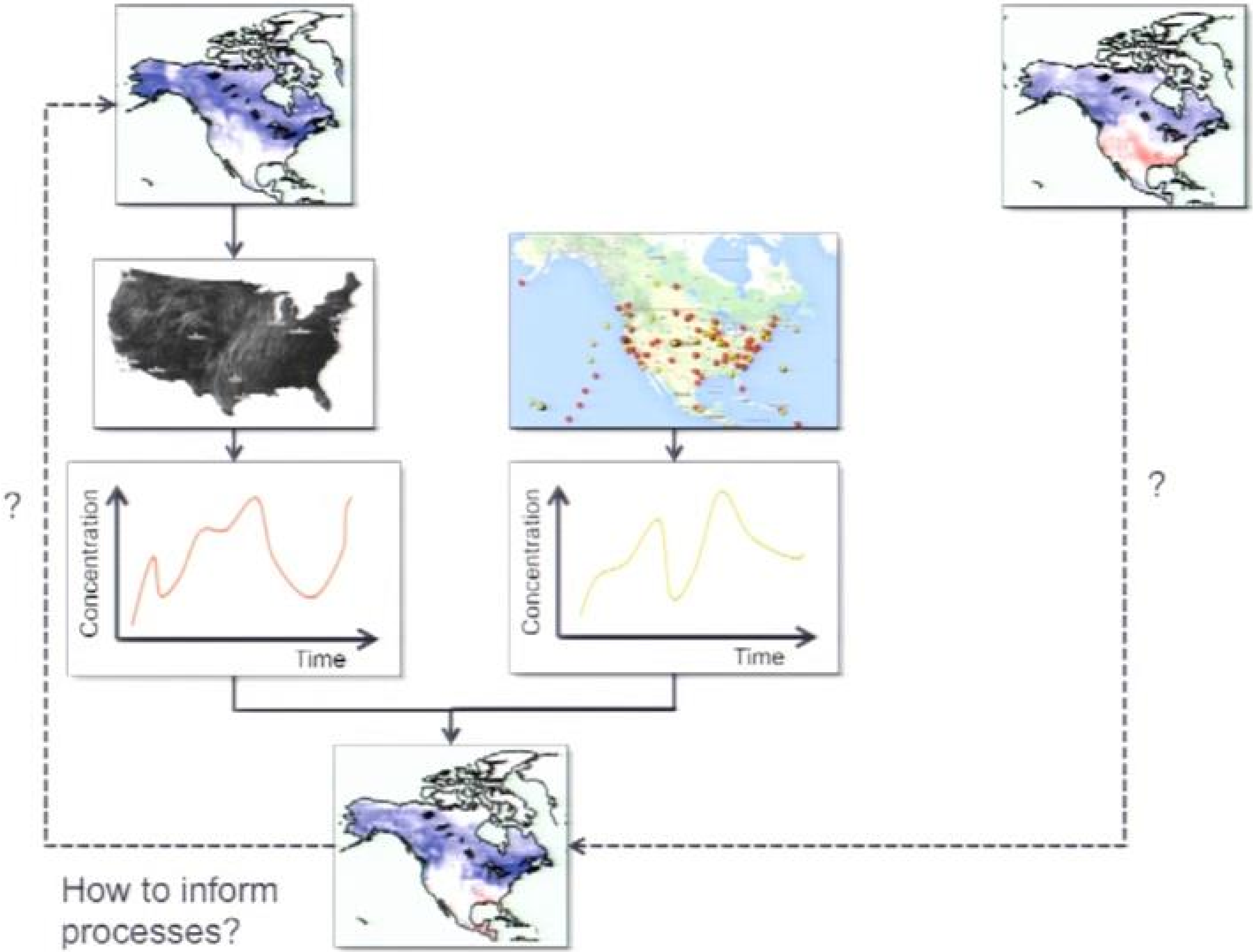




Observations, y

Atmospheric transport, *h(.)*





How to inform processes?

Mixed linear model

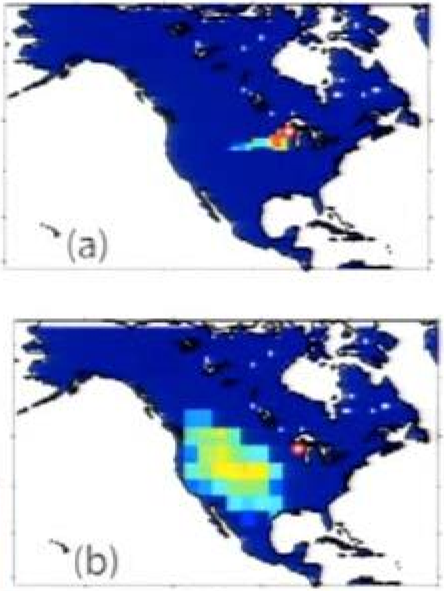
All vary in space
and time

$$y = h(z) + \varepsilon_y + \varepsilon_h + \varepsilon_{rep} + \varepsilon_{agg}$$

$\mathbf{y} = \mathbf{Hz} + \varepsilon$ Linear forward model

High spatiotemporal resolution for $\mathbf{z}$

$\mathbf{z} = \mathbf{X}\beta + \xi$

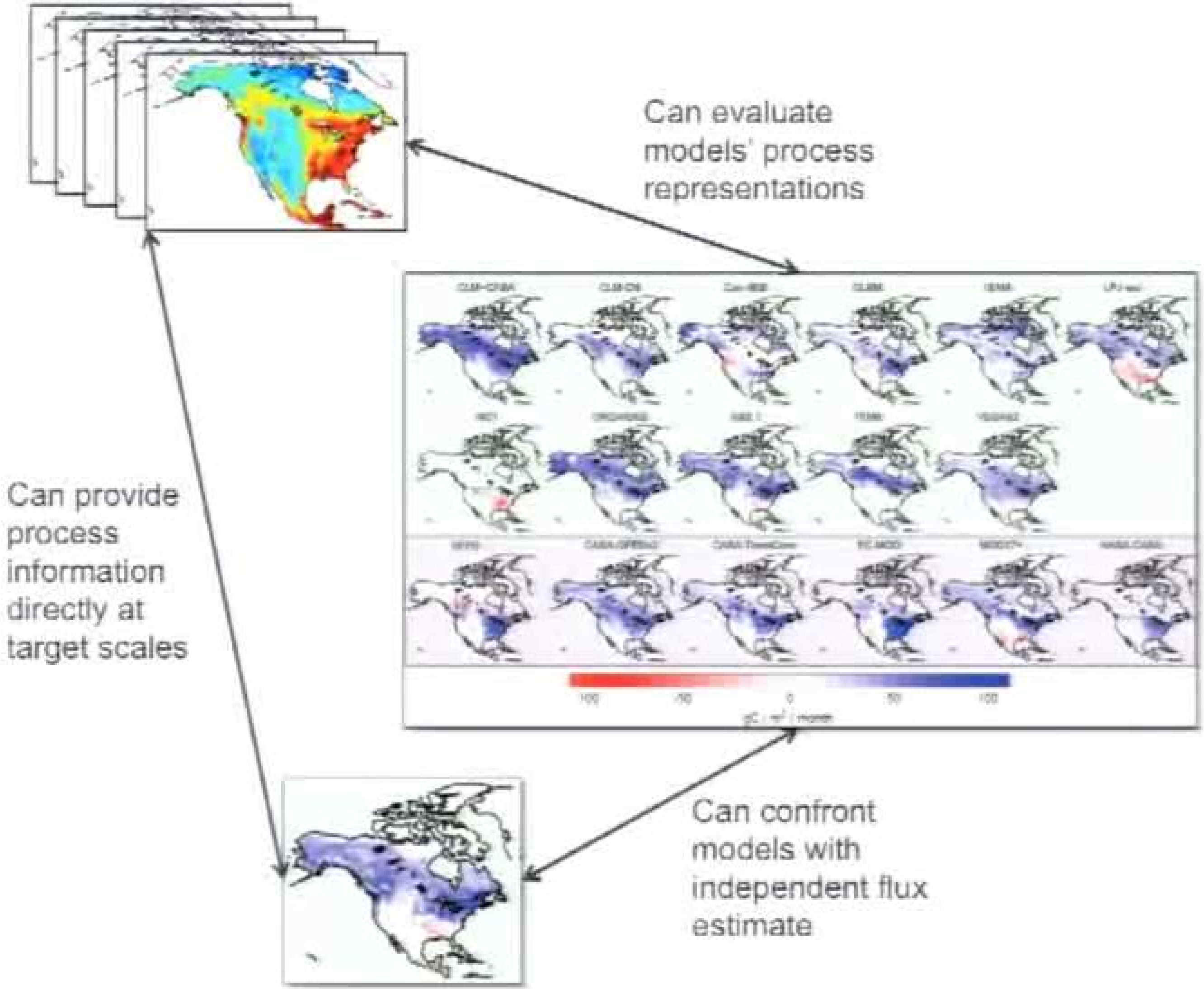


$\mathbf{y} = \mathbf{HX}\beta + \mathbf{H}\xi + \varepsilon$ BIC for model selection (space-time correlated residuals)

$\xi \sim N(0, \mathbf{Q})$ Stationary in space, nonstationary in time, parametric model, not sparse

$\varepsilon \sim N(0, \mathbf{R})$ Independent, variable variance

ReML for parameter estimation


Can evaluate models' process representations
Can provide process information directly at target scales
Can confront models with independent flux estimate

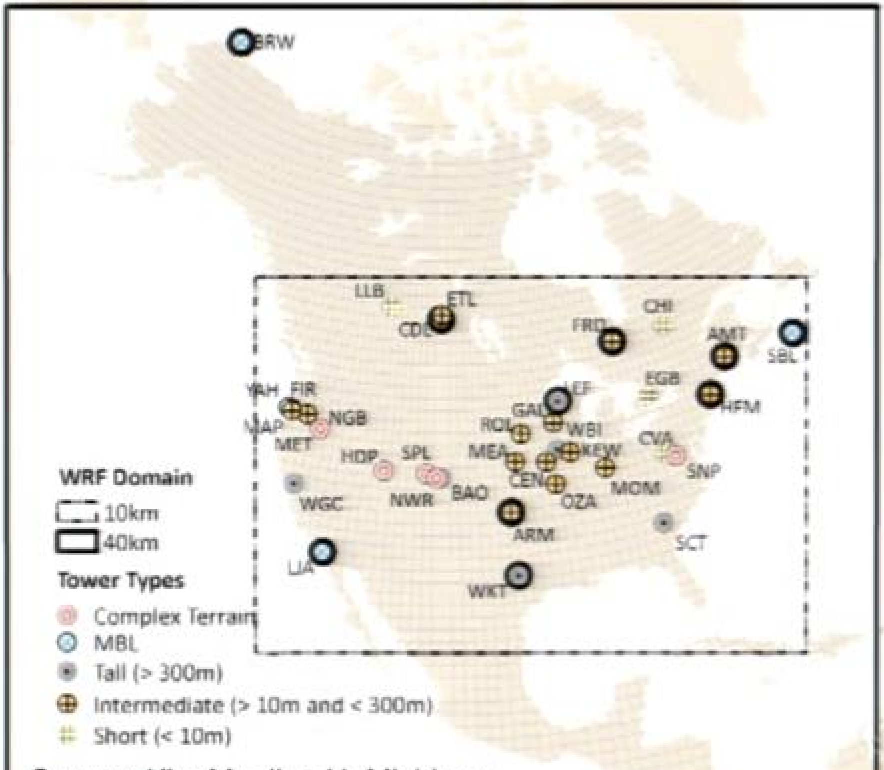
Increasing cost of inversions

Regional CO_2 inversions over North America for one year at $1^\circ \times 1^\circ$; 3-hourly

$\mathbf{y}$: ~10^5

$\mathbf{z}$: ~10^6

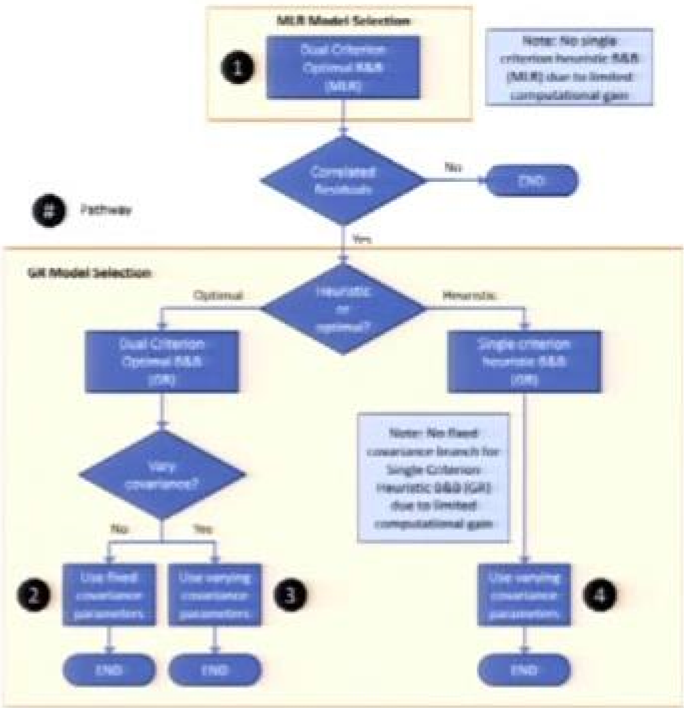
$\mathbf{X}$: ~10^2



Source: Kim Mueller, U. Michigan

($\mathbf{H}$: ~$10^5 \times 10^6$; $\mathbf{Q}$: ~$10^6 \times 10^6$)

Branch & bound algorithm for model selection



k covariate yields 2^k candidate models

Matrix multiplication & posterior covariances

$$\mathbf{y} \sim N\left(\mathbf{HX}\beta, \mathbf{HQH}^T + \mathbf{R}\right) \qquad \hat{\mathbf{z}} \sim N\left(\Lambda\mathbf{y}, \left(\mathbf{H}^T\mathbf{R}^{-1}\mathbf{H} + \mathbf{Q}^{-1}\right)^{-1}\right)$$

$$\mathbf{Q} = \sigma_s^2 \overbrace{\left[\exp\left(-\frac{\mathbf{X}_\tau}{l_\tau}\right)\right]}^{\substack{\text{temporal} \\ \text{covariance}(\mathbf{D})}} \otimes \overbrace{\left[\exp\left(-\frac{\mathbf{X}_s}{l_s}\right)\right]}^{\substack{\text{spatial} \\ \text{covariance}(\mathbf{E})}}.$$

$$\mathbf{Q}_{\text{sum}}(m_s \times m_\tau) = \left(\left(\sum_{j=t_l}^{t_u}\sum_{i=t_l}^{t_u} d_{(i,j)}\right)\mathbf{E}\right).$$

$$\mathbf{H}_{(n \times m_\tau m_s)} = \left(\underbrace{\mathbf{h}_1}_{(n \times m_s)} \; \underbrace{\mathbf{h}_2}_{(n \times m_s)} \; \cdots \; \underbrace{\mathbf{h}m_\tau}_{(n \times m_s)}\right)$$

$$(\mathbf{HQ})_{\text{sum}} = \left(\sum_{j=t_l}^{t_u}\left(\sum_{i=1}^{m_\tau} \boldsymbol{h}_i\, d_{(i,j)}\right)\mathbf{E}\right)_{(n \times m_s)}$$

$$\mathbf{HQ}_{(n \times m_\tau m_s)} = \left(\underbrace{\left(\sum_{i=1}^{m_\tau}\mathbf{h}_i d_{(i,1)}\right)\mathbf{E}}_{(n \times m_s)} \underbrace{\left(\sum_{i=1}^{m_\tau}\mathbf{h}_i d_{(i,2)}\right)\mathbf{E}}_{(n \times m_s)} \cdots \underbrace{\left(\sum_{i=1}^{m_\tau}\mathbf{h}_i d_{(i,m_\tau)}\right)\mathbf{E}}_{(n \times m_s)}\right)$$

$$\bar{\mathbf{V}}_{\hat{\mathbf{s}}} = \frac{\left(\mathbf{Q}_{\text{sum}} - (\mathbf{HQ})_{\text{sum}}^T\left(\mathbf{HQH}^T + \mathbf{R}\right)^{-1}(\mathbf{HQ})_{\text{sum}}\right)}{k^2}.$$

Both algorithms require $O(n^{2.5})$ operations instead of $O(n^3)$ for direct solution.

Ensemble SRF approaches

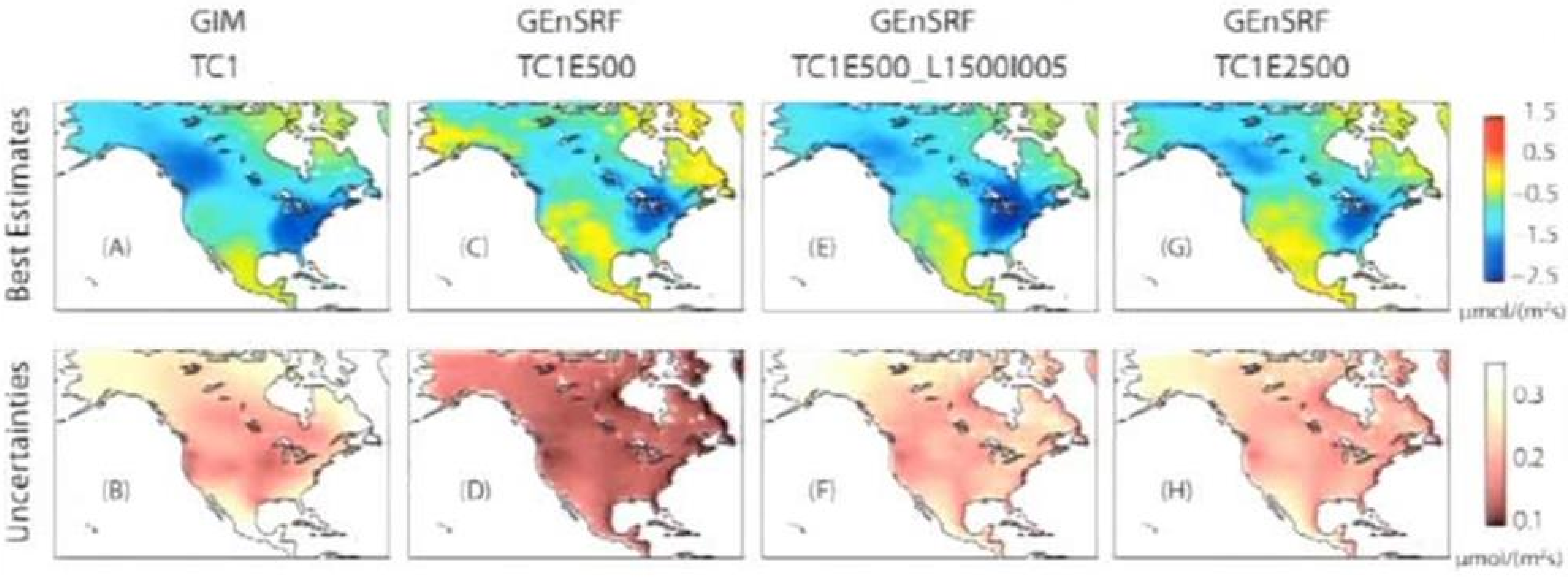
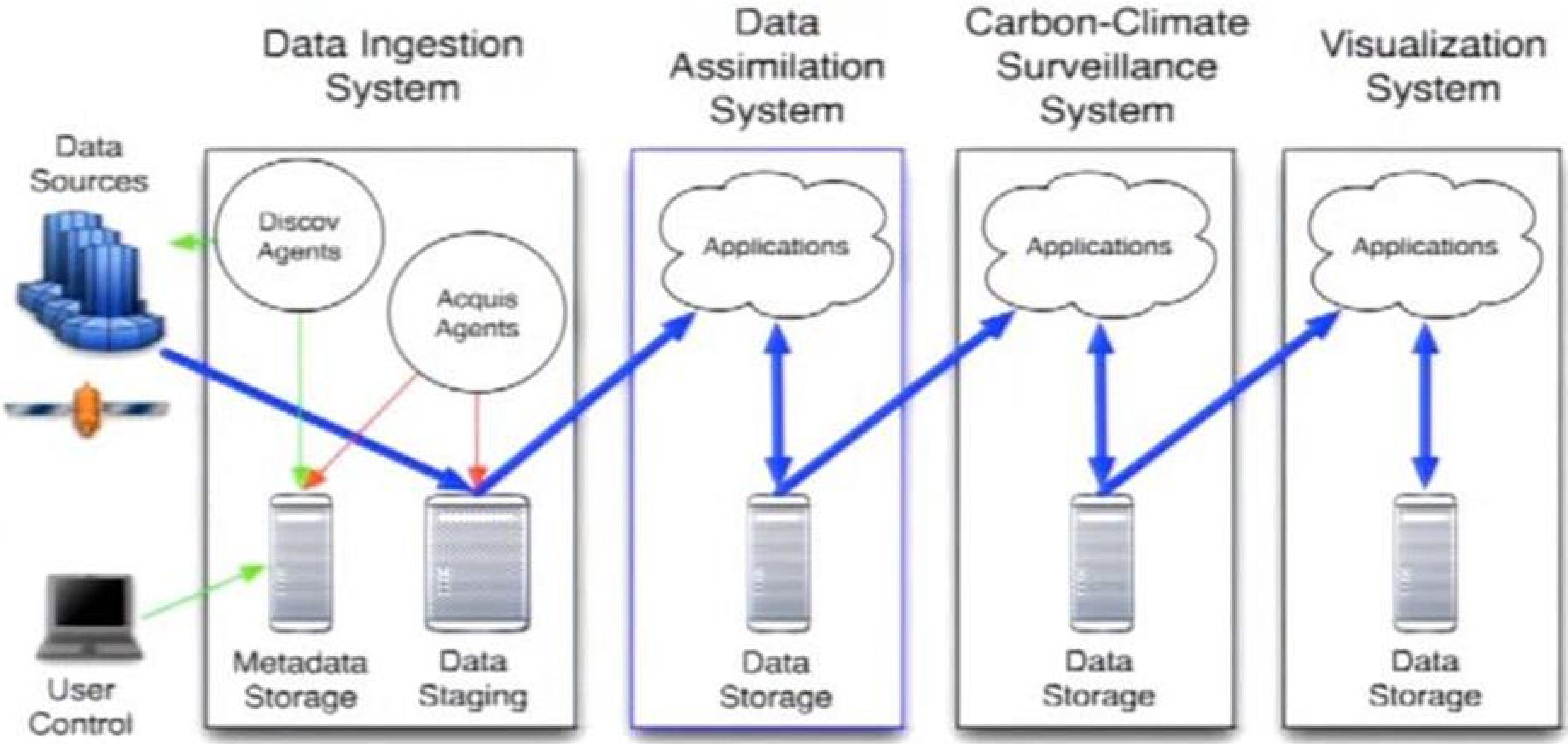


Figure 4. TC1 (top) flux estimates and (bottom) associated uncertainties aggregated to the monthly scale for (a and b) GIM and (c–h) three different GEnSRF runs.

Features:

- No dynamical model
- Kalman smoother
- Heterogeneous (in space and time) observational network

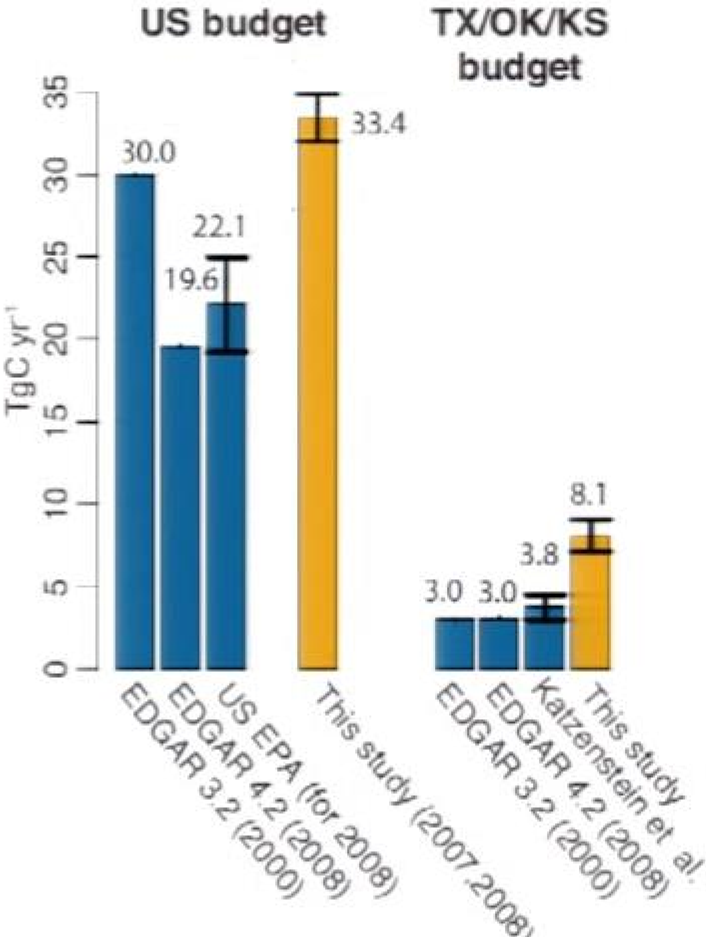
Real-Time Large-Scale Parallel Intelligent CO_2 Data Assimilation System



Take home messages

- The need to constrain greenhouse gas budgets inevitably leads to the need for the solution of inverse problems
- These inverse problems:
  - Require (intelligently) choosing among many uncomfortable assumptions
  - Are becoming increasingly statistically sophisticated and computationally demanding
  - Done carefully, can lead to fundamental insights with management and policy implications

U.S. methane emissions

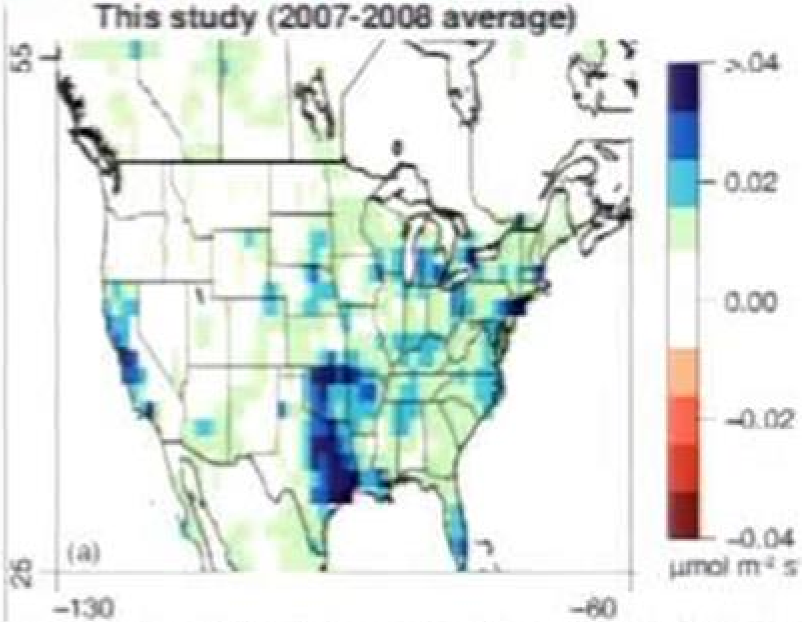


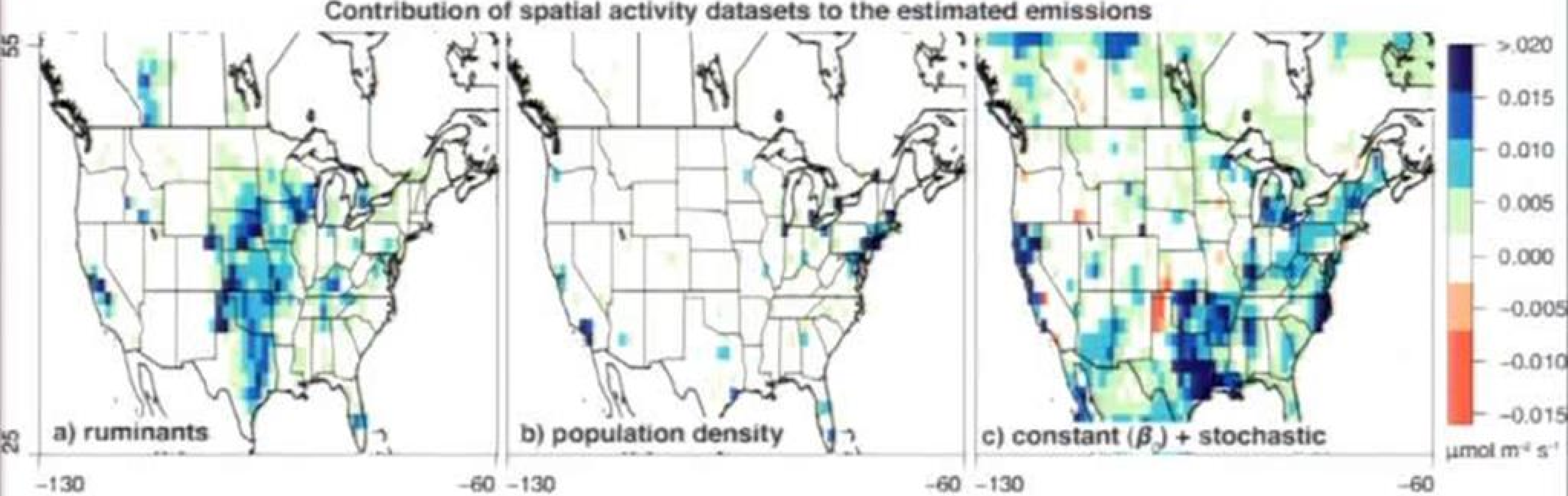
U.S. anthropogenic methane emissions are 50% higher than EPA estimates

Methane emissions in TX / OK / KS are *triple* of what inventories suggest, and a *quarter* of total U.S. emissions

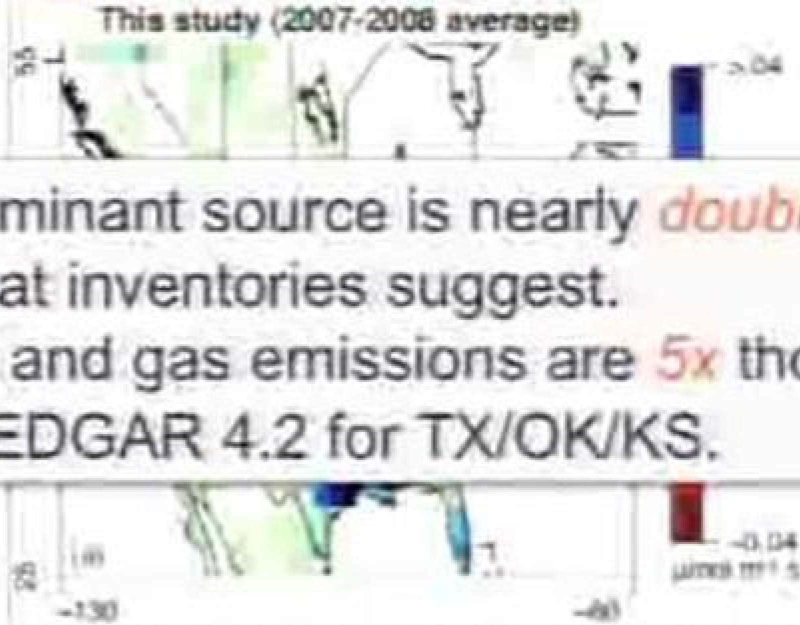
Miller et al. (PNAS, 2013)

Estimated methane fluxes





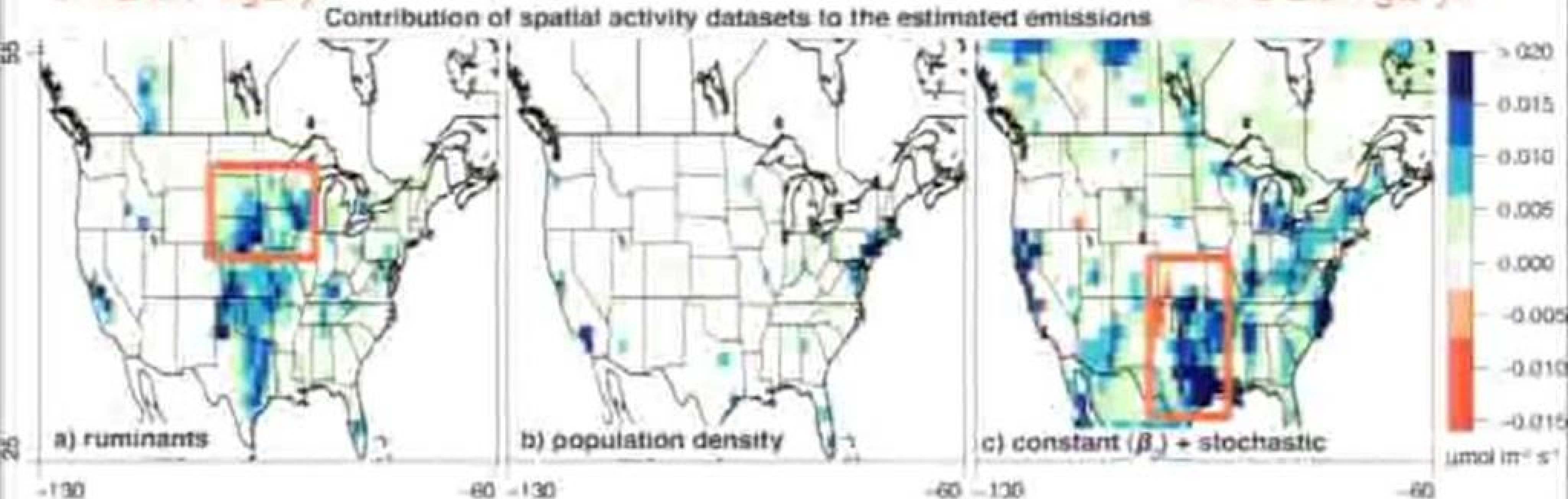
Estimated methane fluxes



Ruminant source is nearly *double* what inventories suggest.
Oil and gas emissions are 5x those in EDGAR 4.2 for TX/OK/KS.

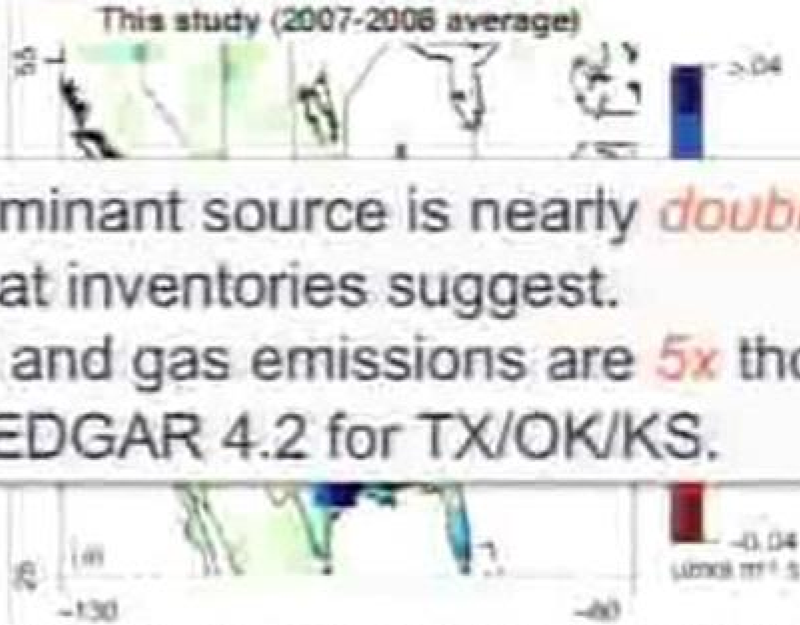
3.4 ± 0.7 TgC yr^{-1}

3.7 ± 2.0 TgC yr^{-1}



Miller et al. (PNAS, 2013)

Estimated methane fluxes

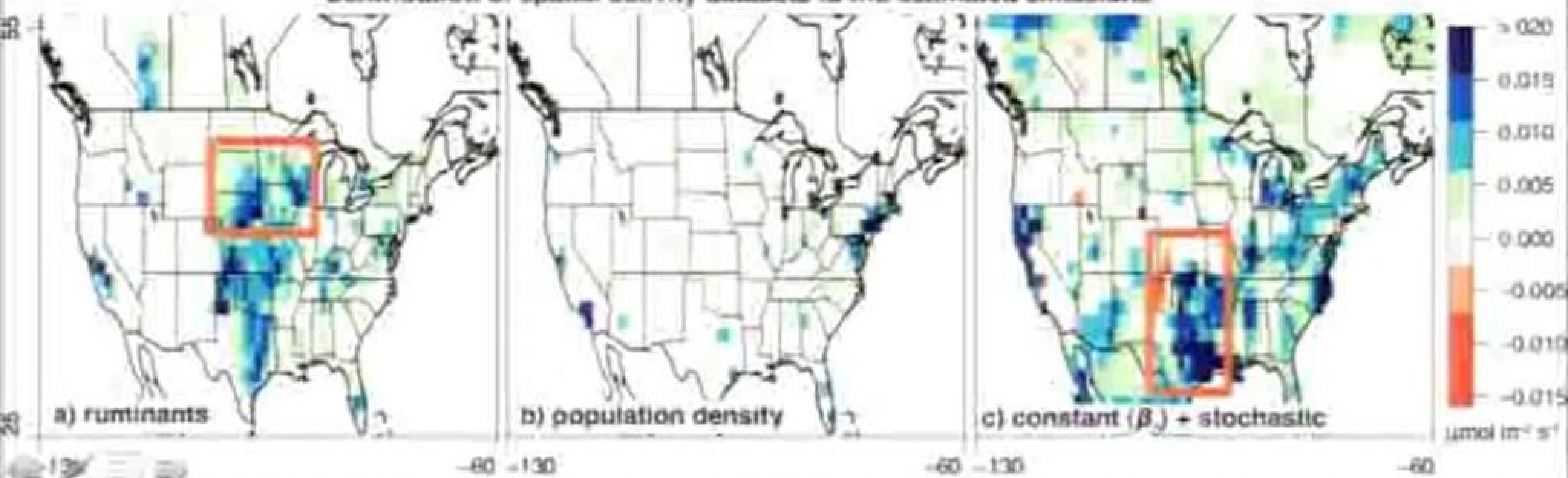


Ruminant source is nearly *double* what inventories suggest.
Oil and gas emissions are 5x those in EDGAR 4.2 for TX/OK/KS.

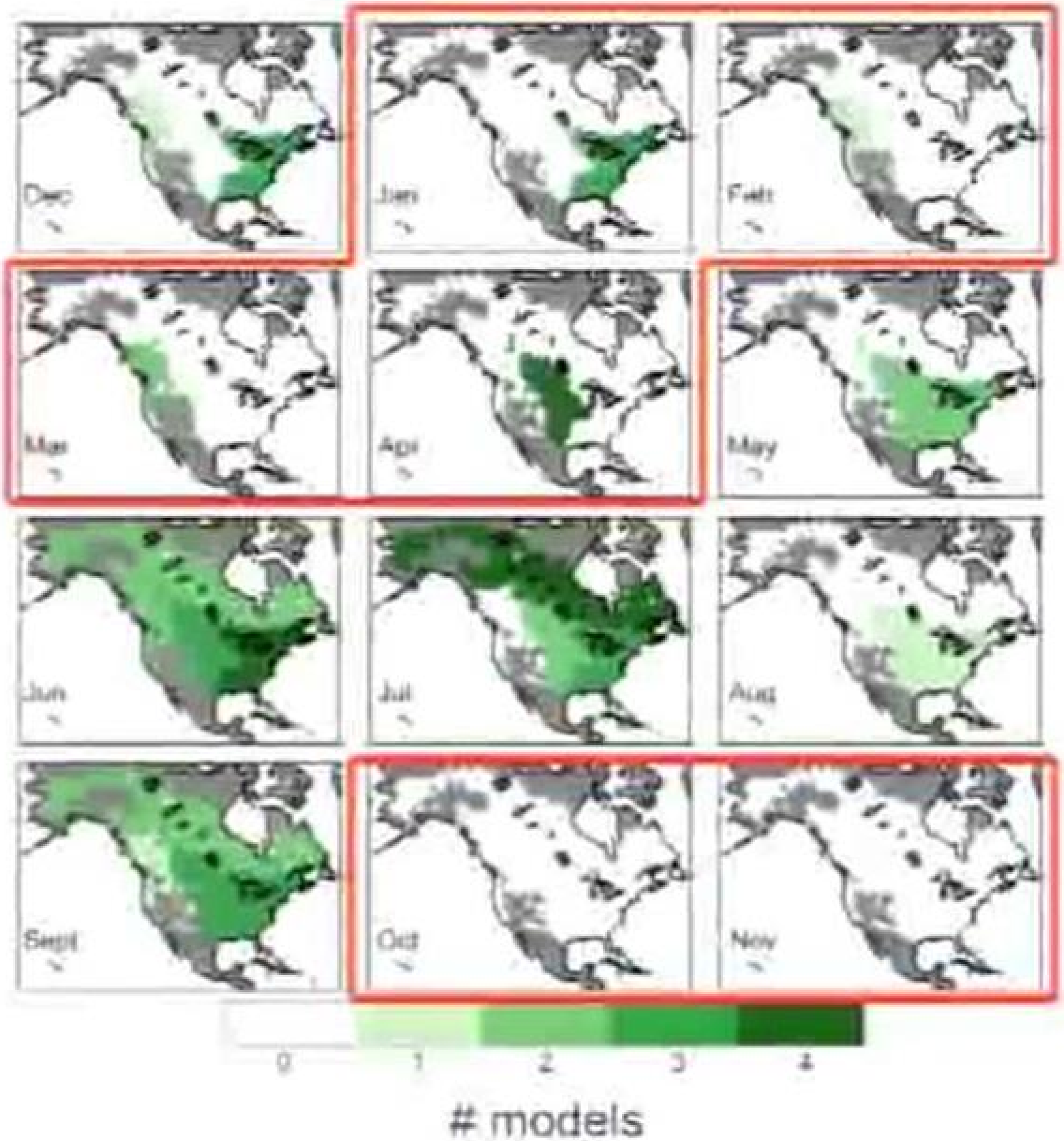
3.4 ± 0.7 TgC yr^{-1}

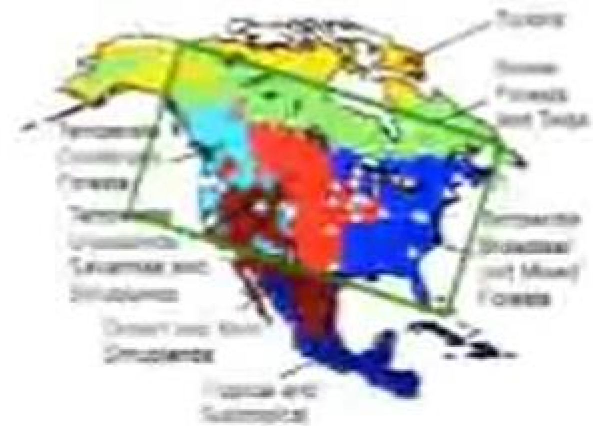
3.7 ± 2.0 TgC yr^{-1}

Contribution of spatial activity datasets to the estimated emissions



Confronting model flux *patterns* with obs

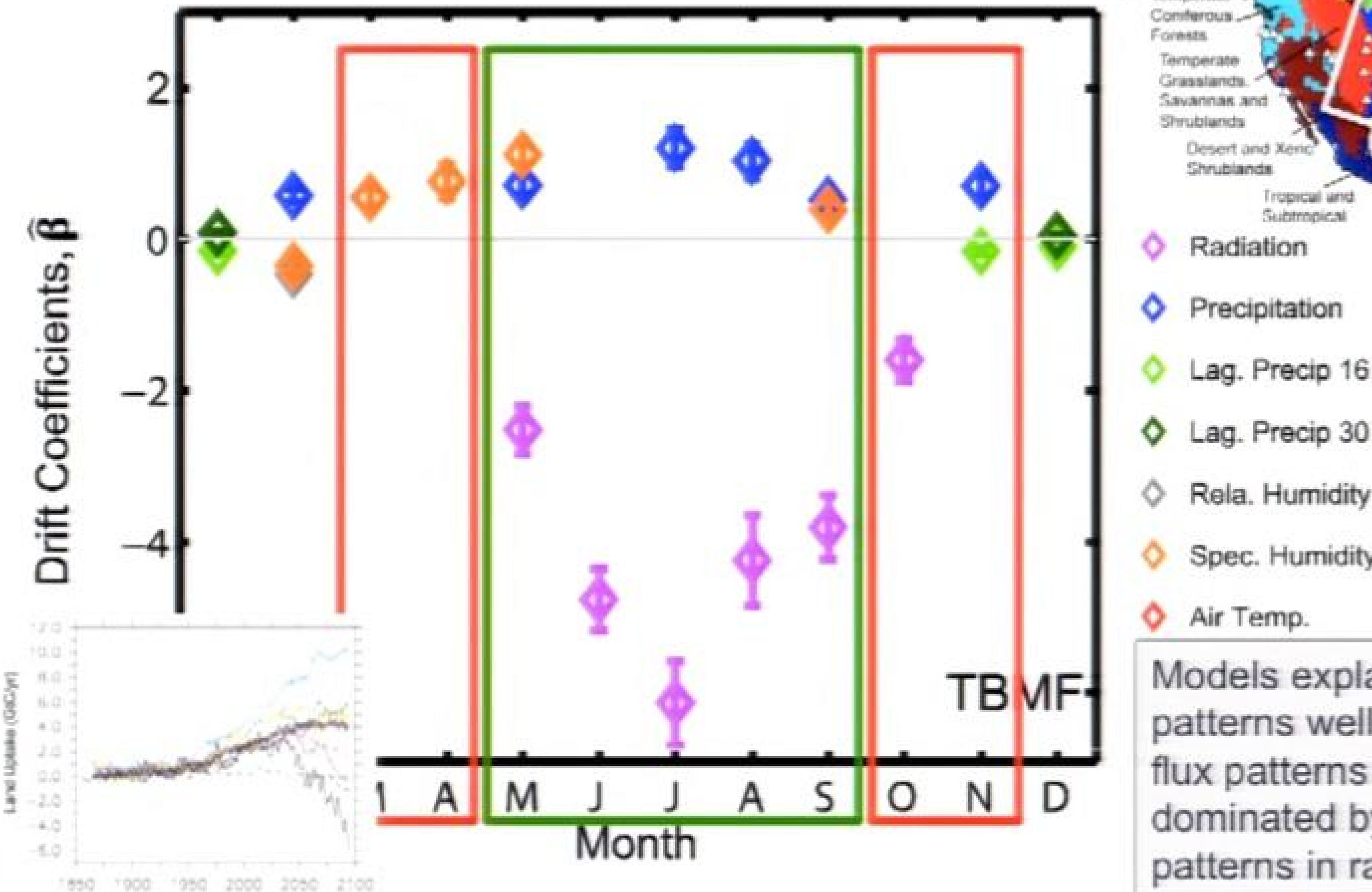




Models' flux patterns do not explain observed variability in atmospheric observations for much of the year, but they do better during growing season.

Fang et al. (GBC *in revision*)

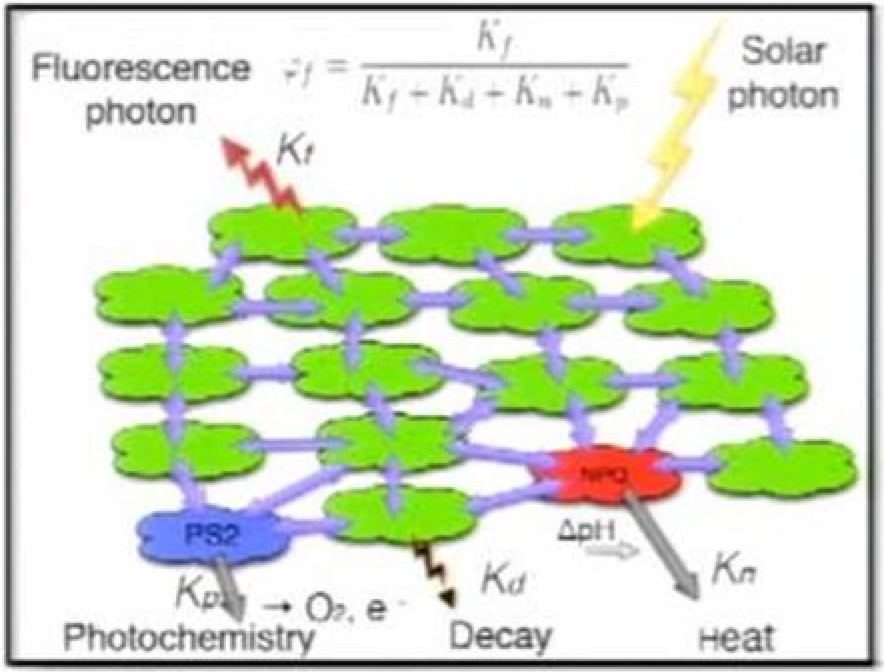
Providing process information directly at target scales



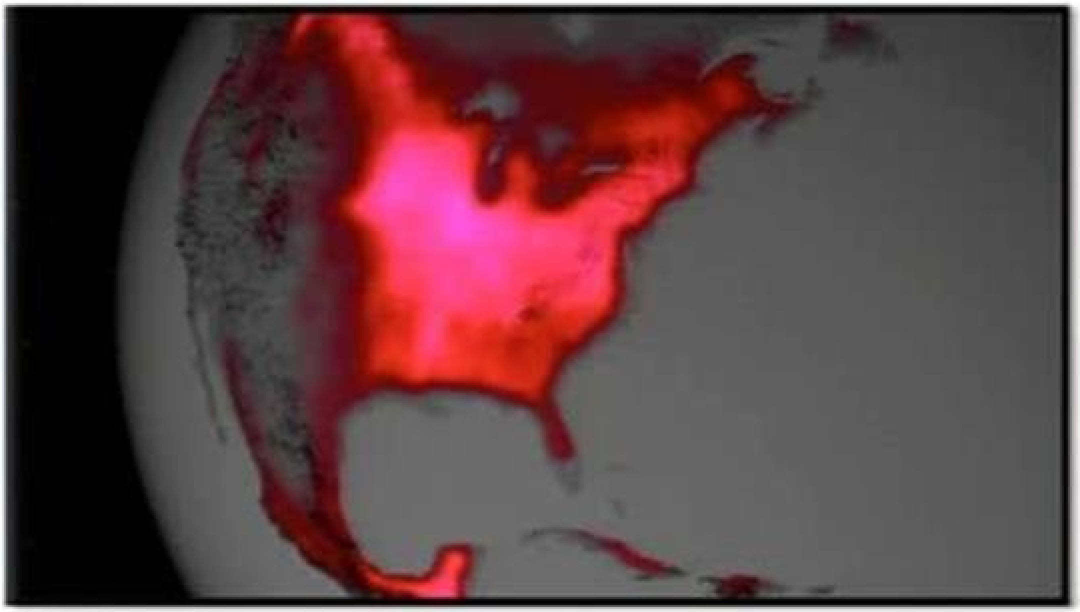
Models explain flux patterns well when flux patterns are dominated by patterns in radiation

Solar Induced Fluorescence

SIF emitted during photosynthesis and is therefore potentially a promising measure of GPP



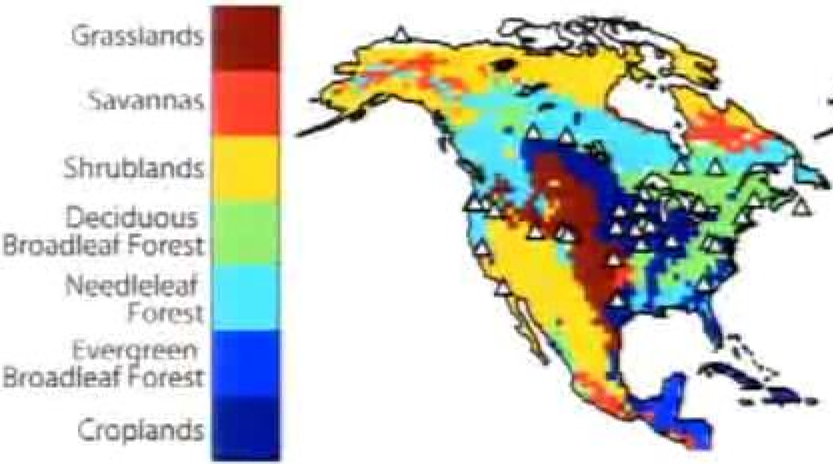
Source: Frankenberg, 2011

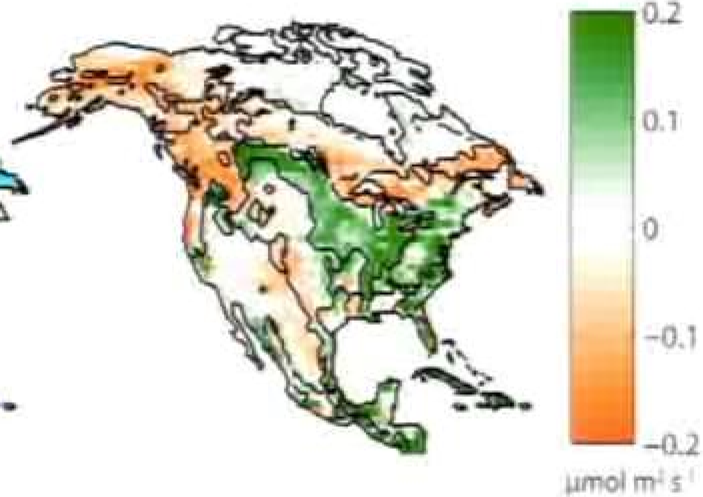


Source: http://www.nasa.gov/press/goddard/2014/march/satellite-shows-high-productivity-from-us-corn-belt/#.U8QK4_ldV8G

Shiga et al. (*in prep*)

Differences at 1° x 1°, aggregated over March to October





Take home messages

- The need to constrain greenhouse gas budgets inevitably leads to the need for the solution of inverse problems
- These inverse problems:
  - Require (intelligently) choosing among many uncomfortable assumptions
  - Are becoming increasingly statistically sophisticated and computationally demanding
  - Done carefully, can lead to fundamental insights with management and policy implications

Acknowledgments

- PUORG: (Current:) Yuanyuan Fang, Yoichi Shiga, Jovan Tadic, (Alumni:) Abhishek Chatterjee, Sharon Gourdji, Debbie Huntzinger, Kim Mueller, Vineet Yadav
- NOAA-ESRL: Pieter Tans, Arlyn Andrews, Gabrielle Petron, Mike Trudeau
- AER: Thomas Nehrkorn, John Henderson, Janusz Eluszkiewicz
- NACP Regional Interim Synthesis Participants
- NOAA-ESRL Cooperative Air Sampling Network
- NASA HEC Project Columbia, Pleiades, and technical support staff

More information: michalak@stanford.edu

http://dge.stanford.edu/michalaklab