STATISTICAL AND COMPUTATIONAL CHALLENGES OF CONSTRAINING GREENHOUSE GAS BUDGETS

Anna M. Michalak

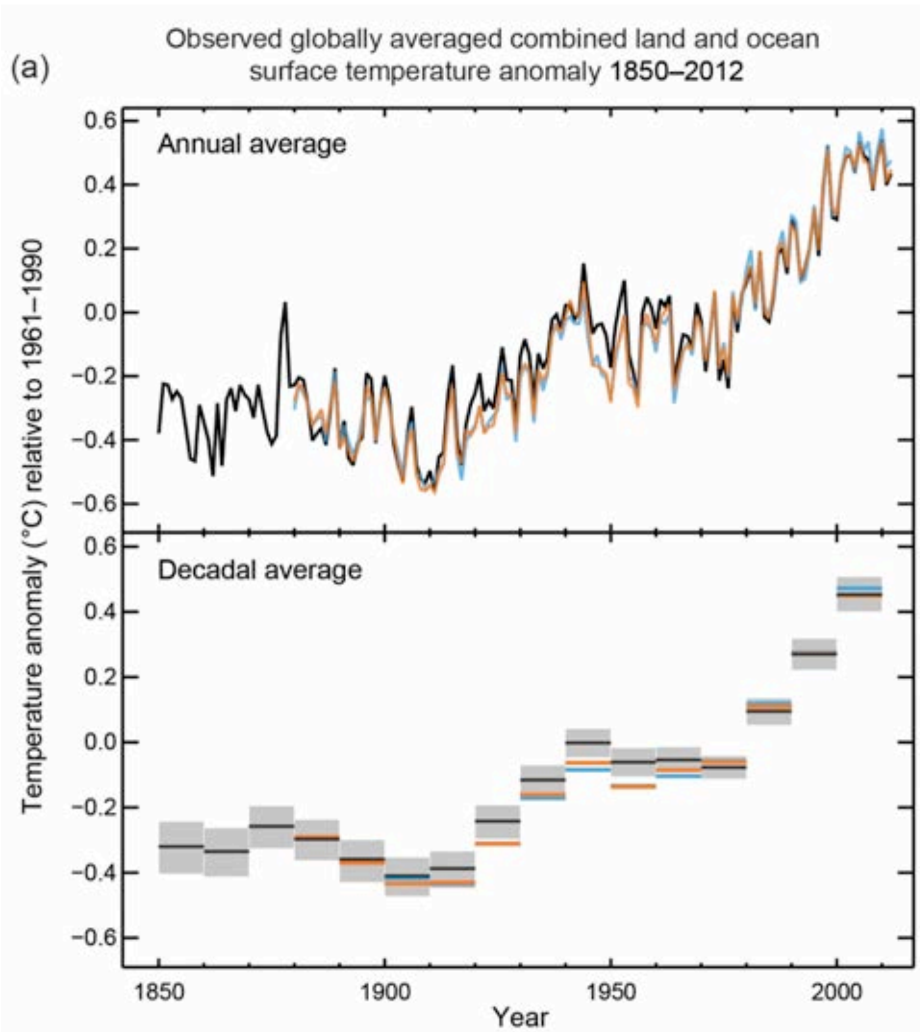
Department of Global Ecology, Carnegie Institution for Science

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
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
1850
1900
1950
2000
Year

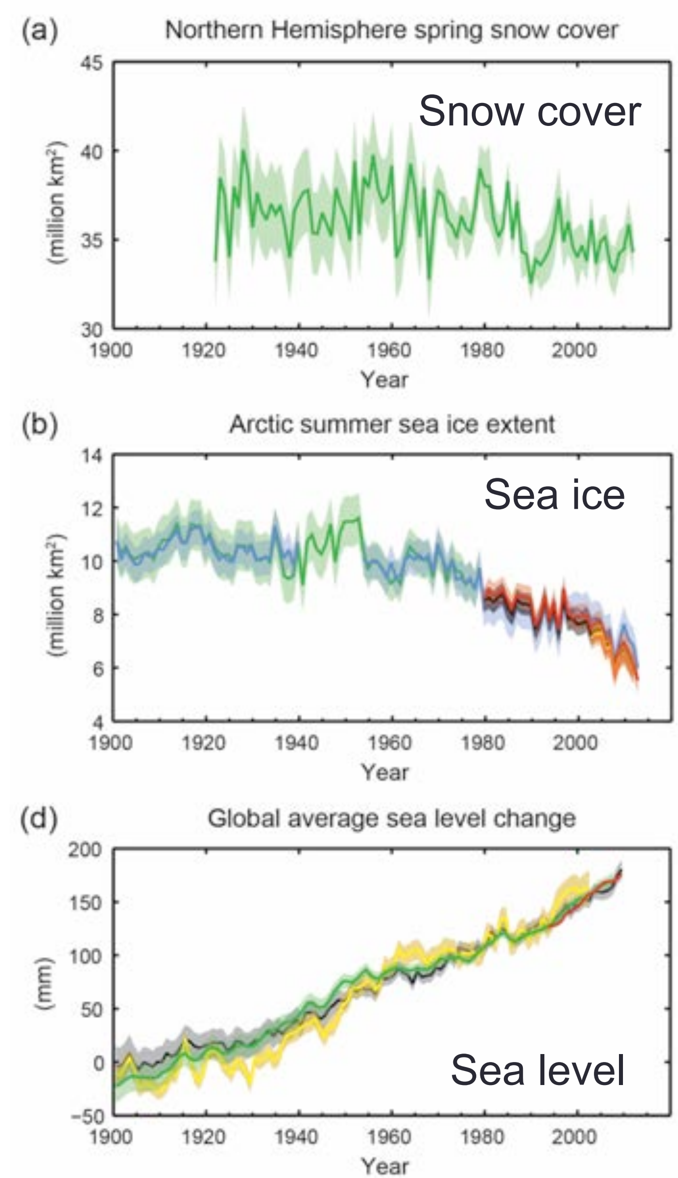
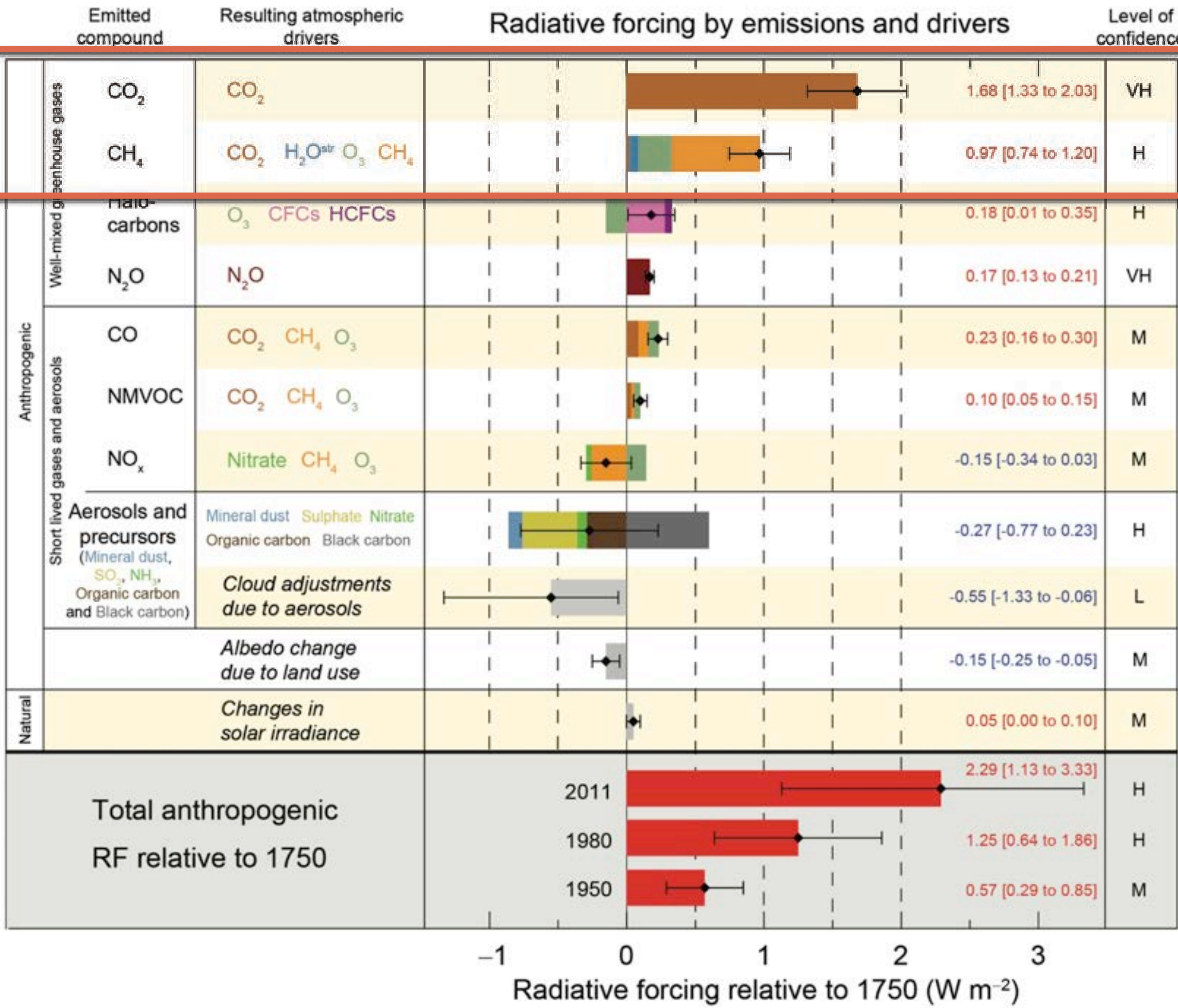
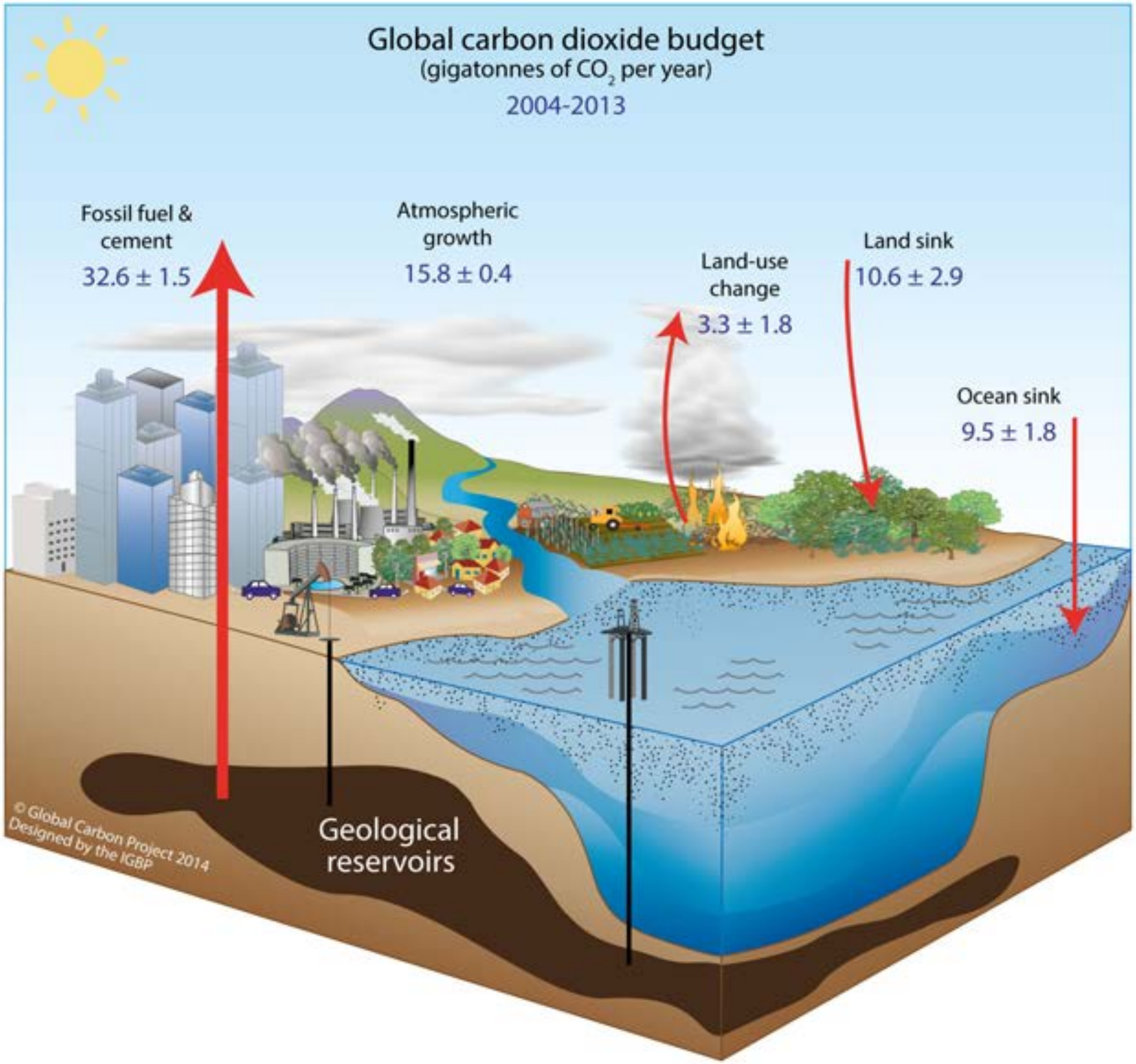

(a)
Northern Hemisphere spring snow cover
Snow cover
(million km²)
45
40
35
30
1900
1920
1940
1960
1980
2000
Year
(b)
Arctic summer sea ice extent
Sea ice
(million km²)
14
12
10
8
6
4
Year
(d)
Global average sea level change
(mm)
200
150
100
50
0
−50
Sea level
Year

Photo courtesy of Bartshé Miller

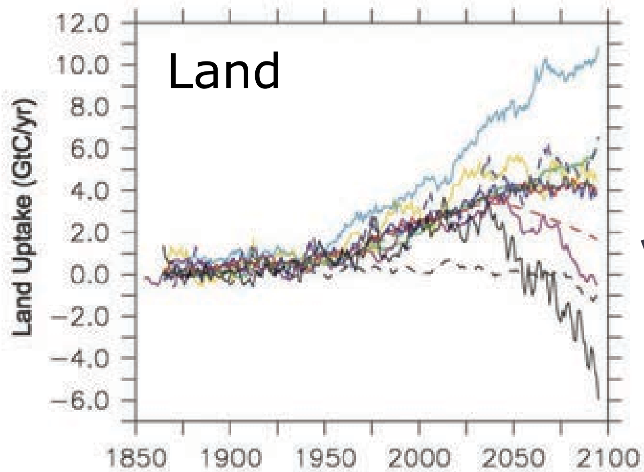
Tioga Pass, January 12 2015

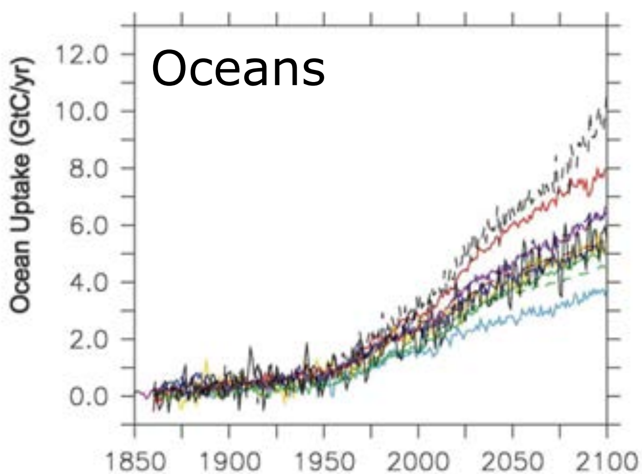

Emitted compound
Resulting atmospheric drivers
Radiative forcing by emissions and drivers
Level of confidence
Anthropogenic
Well-mixed greenhouse gases
CO_2
CO_2
1.68 [1.33 to 2.03]
VH
CH_4
CO_2 H_2O^{str} O_3 CH_4
0.97 [0.74 to 1.20]
H
Halo-carbons
O_3 CFCs HCFCs
0.18 [0.01 to 0.35]
H
N_2O
N_2O
0.17 [0.13 to 0.21]
VH
Short lived gases and aerosols
CO
CO_2 CH_4 O_3
0.23 [0.16 to 0.30]
M
NMVOC
CO_2 CH_4 O_3
0.10 [0.05 to 0.15]
M
NO_x
Nitrate CH_4 O_3
-0.15 [-0.34 to 0.03]
M
Aerosols and precursors (Mineral dust, SO_2, NH_3, Organic carbon and Black carbon)
Mineral dust Sulphate Nitrate Organic carbon Black carbon
-0.27 [-0.77 to 0.23]
H
Cloud adjustments due to aerosols
-0.55 [-1.33 to -0.06]
L
Albedo change due to land use
-0.15 [-0.25 to -0.05]
M
Natural
Changes in solar irradiance
0.05 [0.00 to 0.10]
M
Total anthropogenic RF relative to 1750
2011
2.29 [1.13 to 3.33]
H
1980
1.25 [0.64 to 1.86]
H
1950
0.57 [0.29 to 0.85]
M
–1
0
1
2
3
Radiative forcing relative to 1750 ($W m^{-2}$)



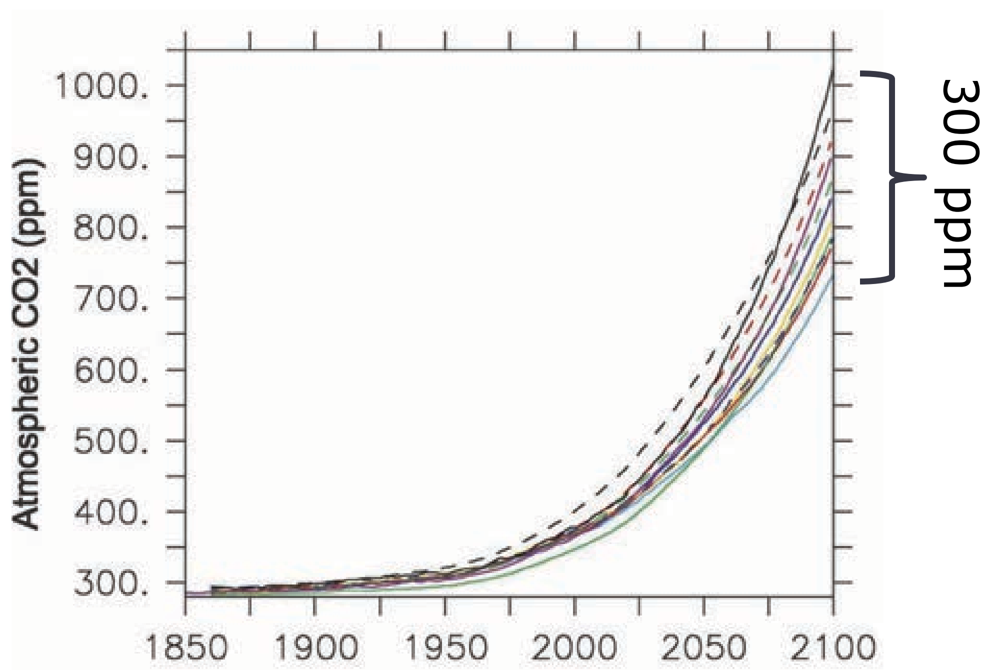
Perturbation of the global carbon cycle caused by anthropogenic activities, averaged globally for the decade 2004–2013 ($GtCO_2$/yr)

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



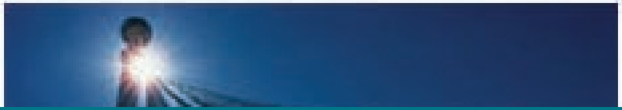
Source: Friedlingstein et al. (2006) showing projections from coupled carbon and climate simulations for several models.

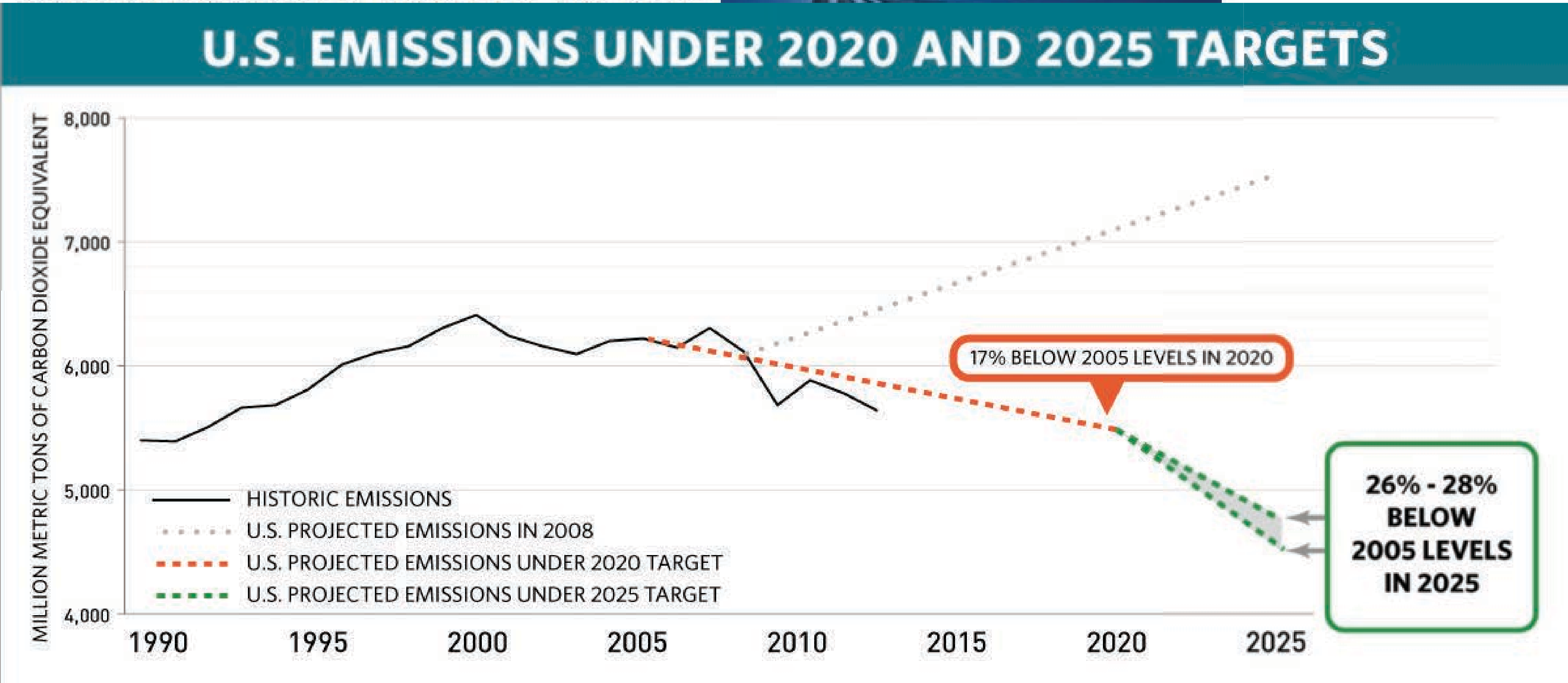
Paris Agreement

Policy | Documentation

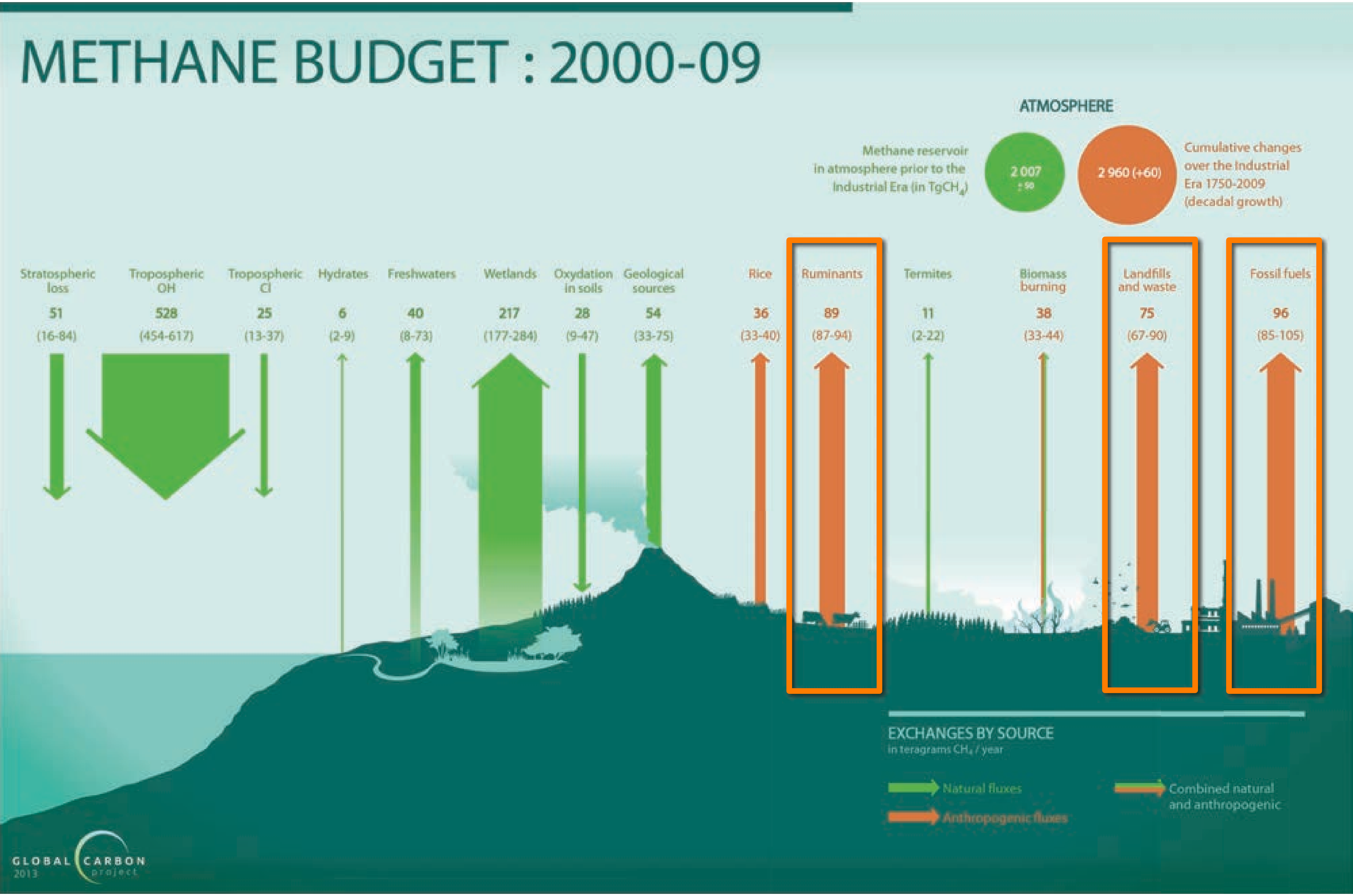
At the Paris climate conference (COP21) in December 2015, 195 countries adopted the first-ever universal, legally binding global climate deal.

The agreement sets out a global action plan to put the world on track to avoid dangerous climate





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



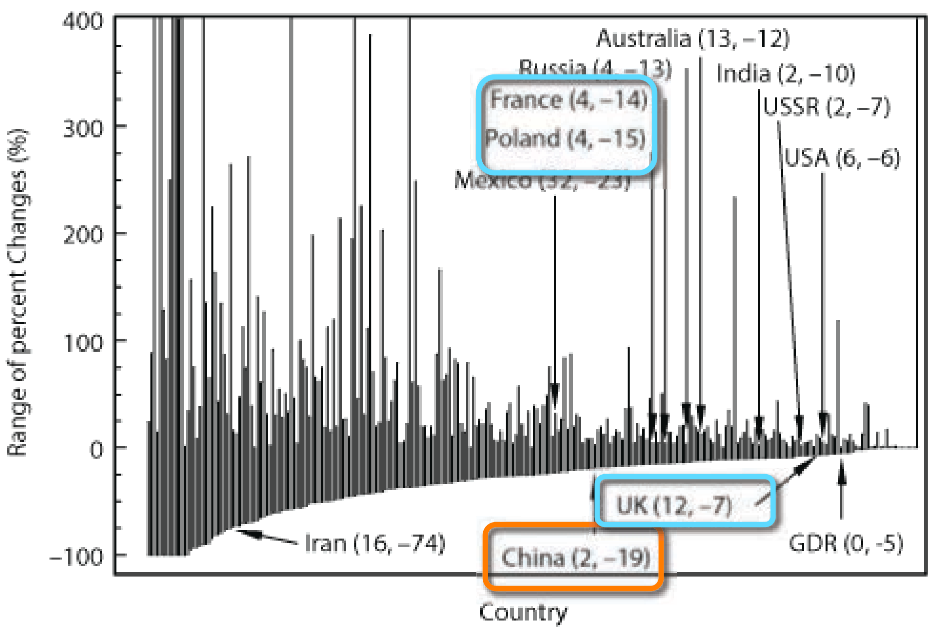
U.S.-CANADA JOINT STATEMENT ON CLIMATE, ENERGY, AND ARCTIC LEADERSHIP

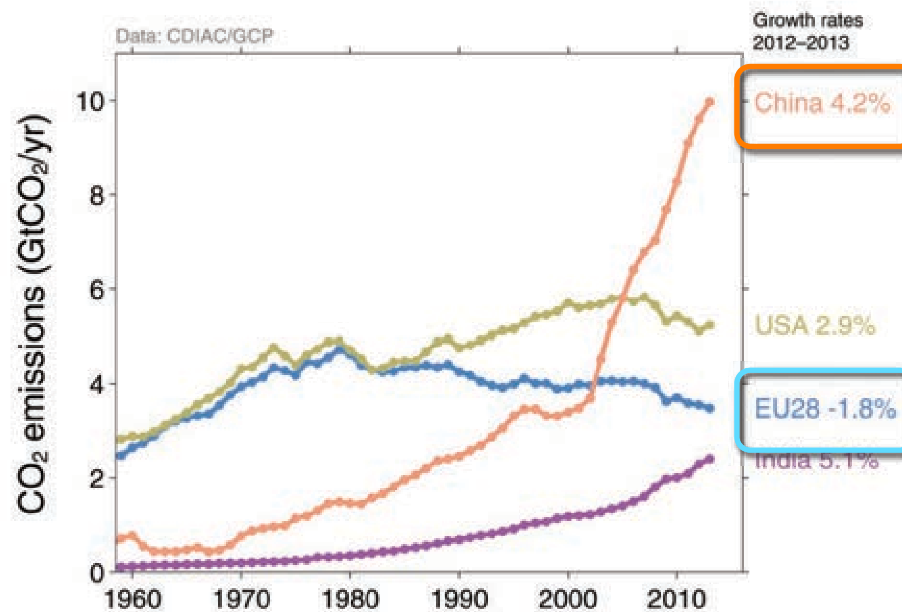
Coordinated domestic climate action

Building on a history of working together to reduce air emissions, Canada and the U.S., commit to take action to reduce methane emissions from the oil and gas sector, the world's largest industrial methane source, in support of achieving our respective international climate change commitments. To set us on an ambitious and achievable path, the leaders commit to reduce methane emissions by 40-45 percent below 2012 levels by 2025 from the oil and gas sector, and explore new opportunities for additional methane reductions. The leaders also invite other countries to join the target or develop their own methane reduction goal. To achieve this target, both countries commit to:

How do we know emissions?

Self reporting

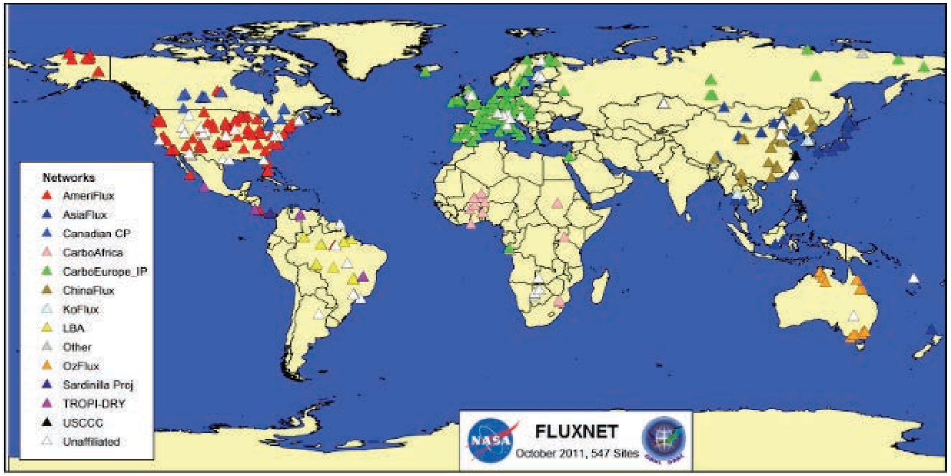


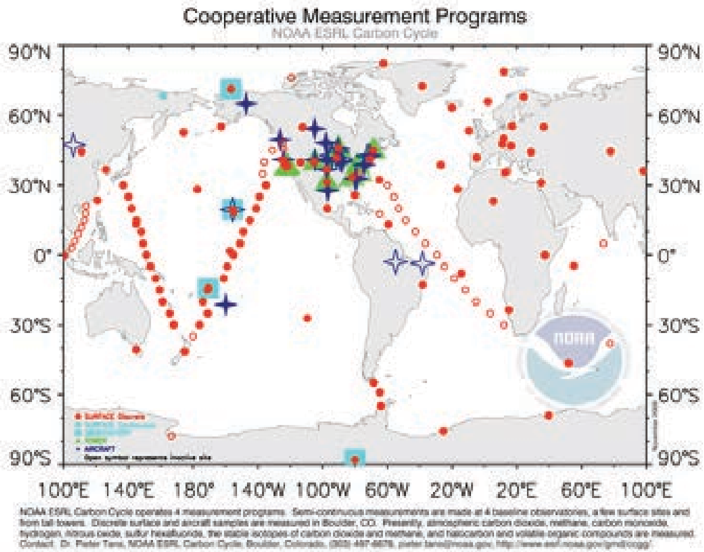


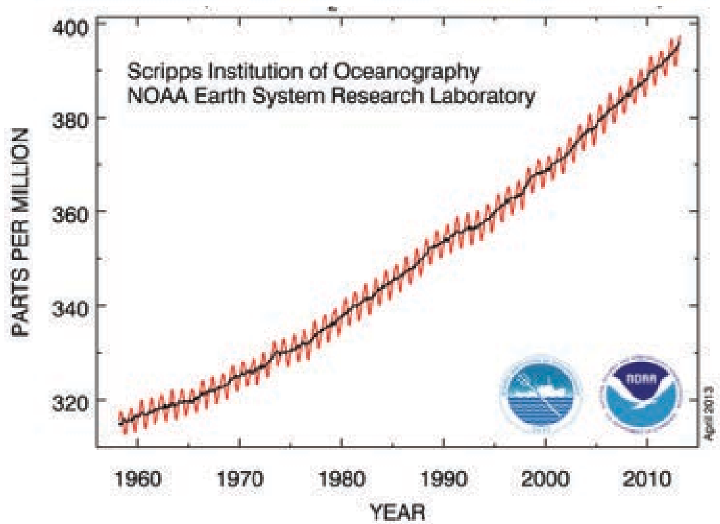
How do we know emissions?

Inventories

How do we know emissions?



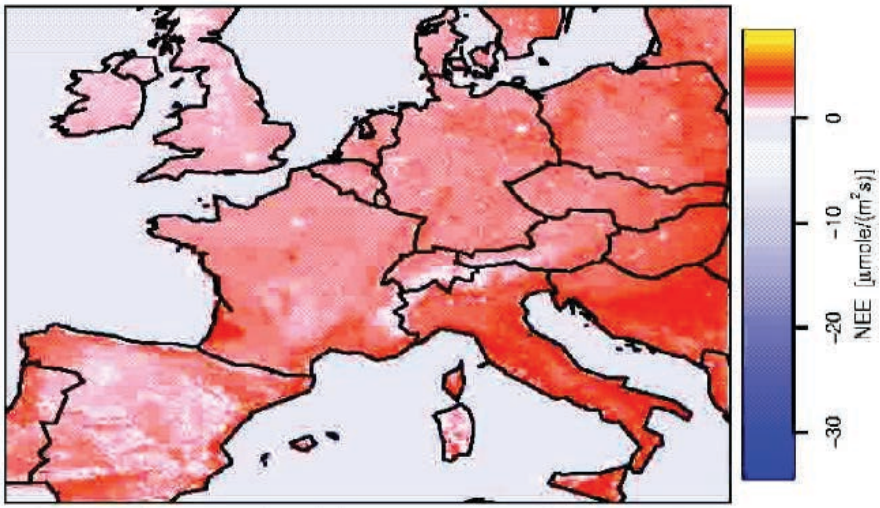




Fluxes (i.e. emissions / uptake)

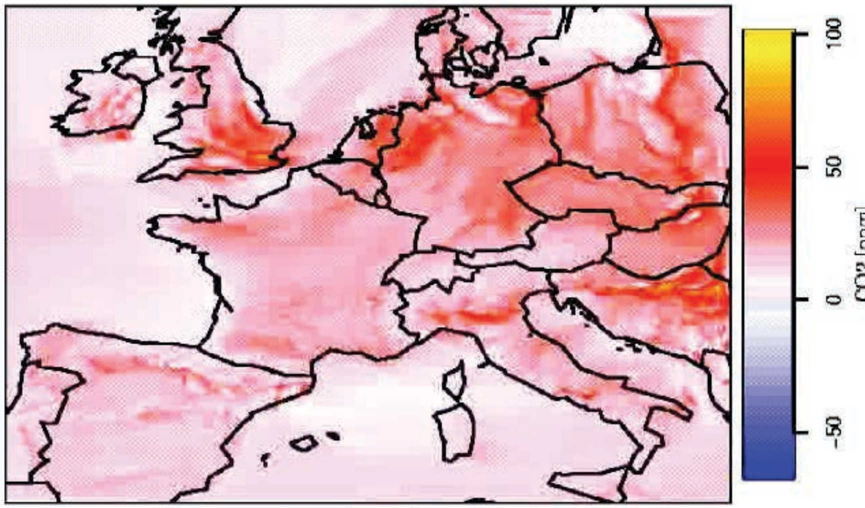
Concentrations

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



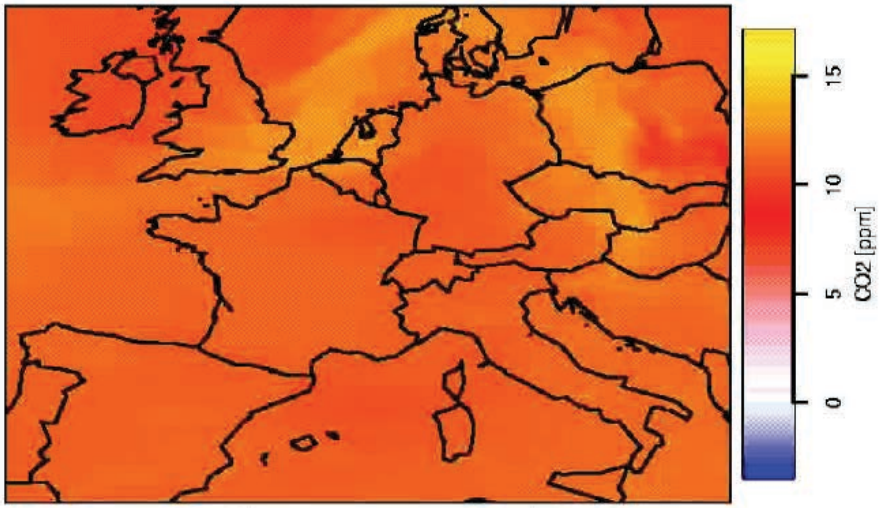
Vegetation–Photosysthesis and Respiration Model, created at MPI–BGC

CO2 at 0.1 km, time 2003–07–02_00:00:00



WRF+CASA+VPRM, created at MPI–BGC

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



WRF+CASA+VPRM, created at MPI–BGC

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

All vary in space and time

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

- 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

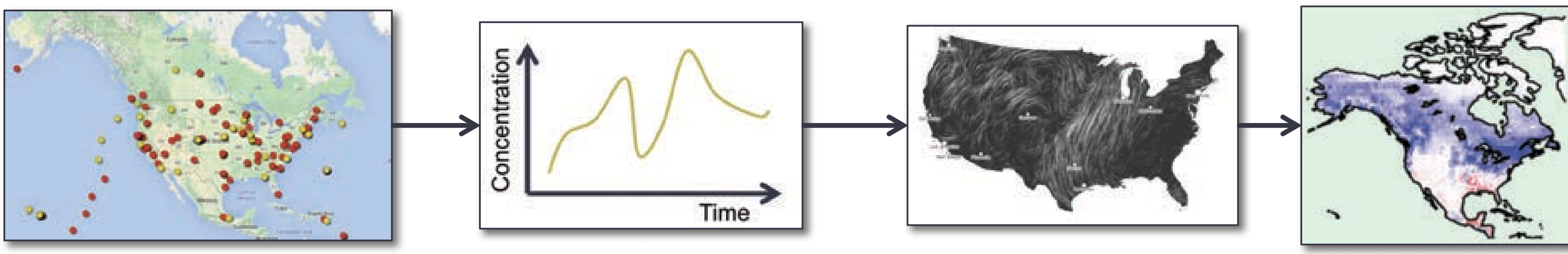
ε_{rep}: "representation" error (finite resolution in y)

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

Overall inverse problem

All vary in space and time

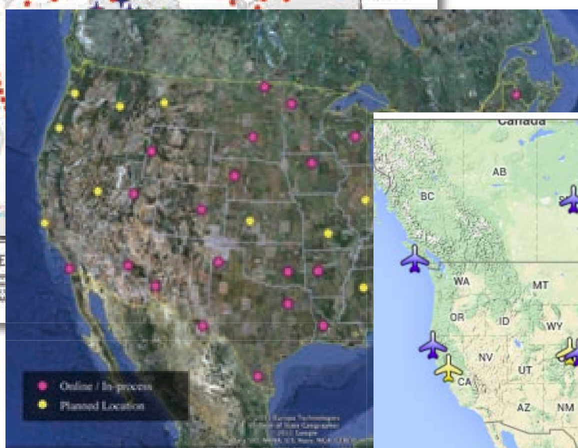
$$y = h(z) + \varepsilon_y + \varepsilon_h + \varepsilon_{rep} + \varepsilon_{agg}$$

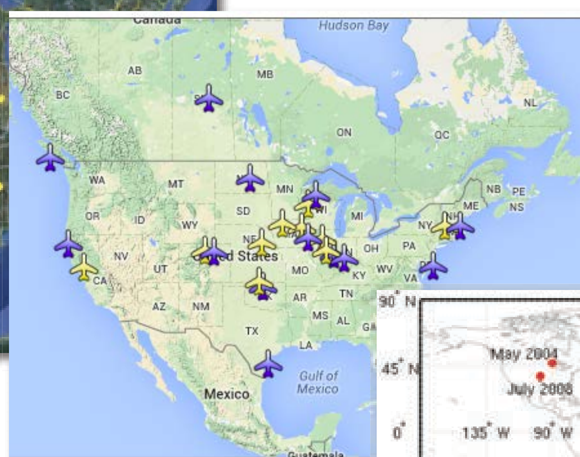


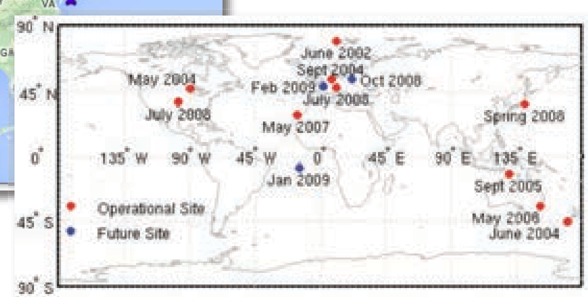
$\underbrace{\qquad\qquad}_{y}$ $\underbrace{\qquad}_{h(.)}$ $\underbrace{\qquad}_{z}$

Causation

Inference







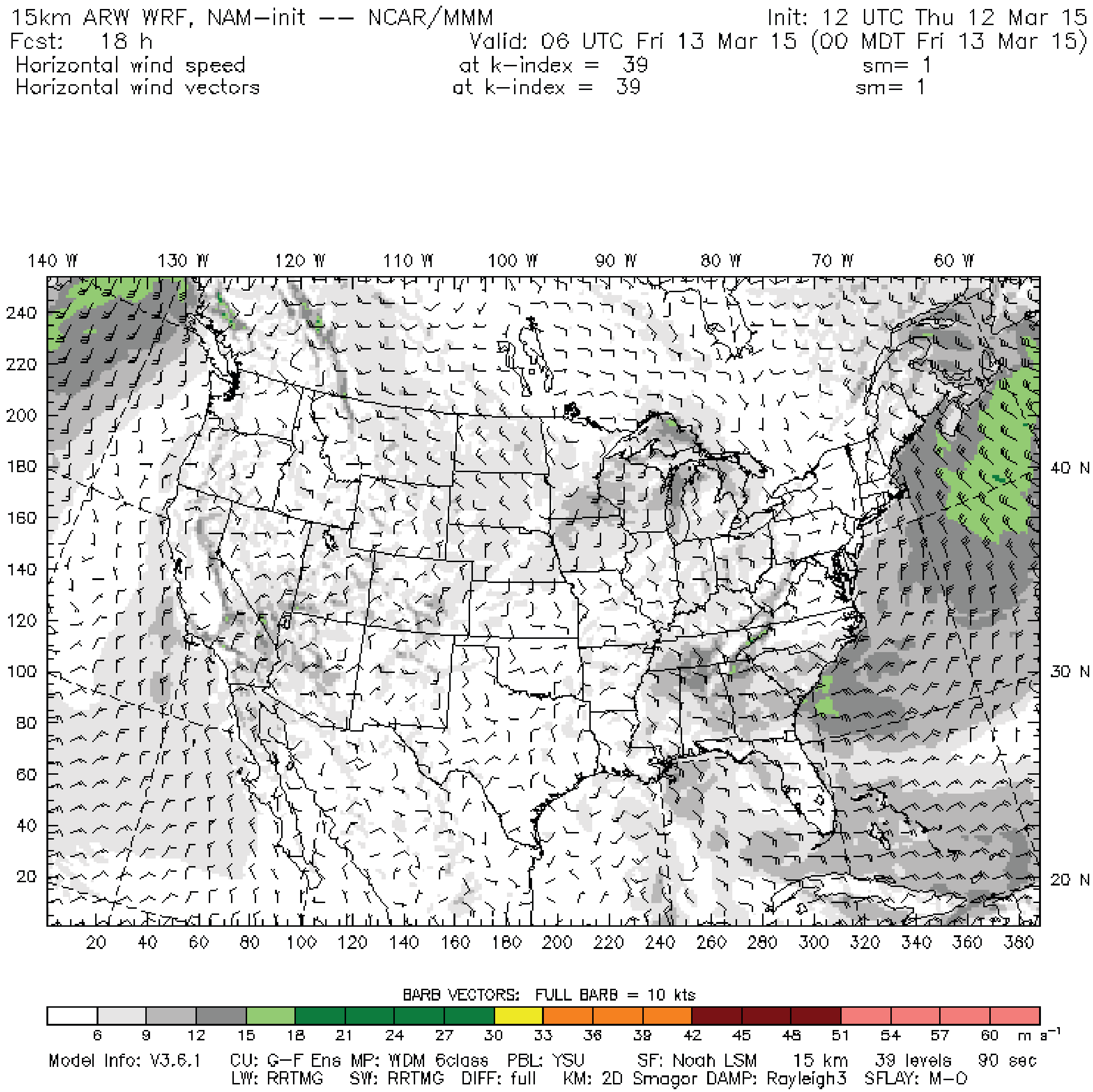
AIRS
Atmospheric Infrared Sounder

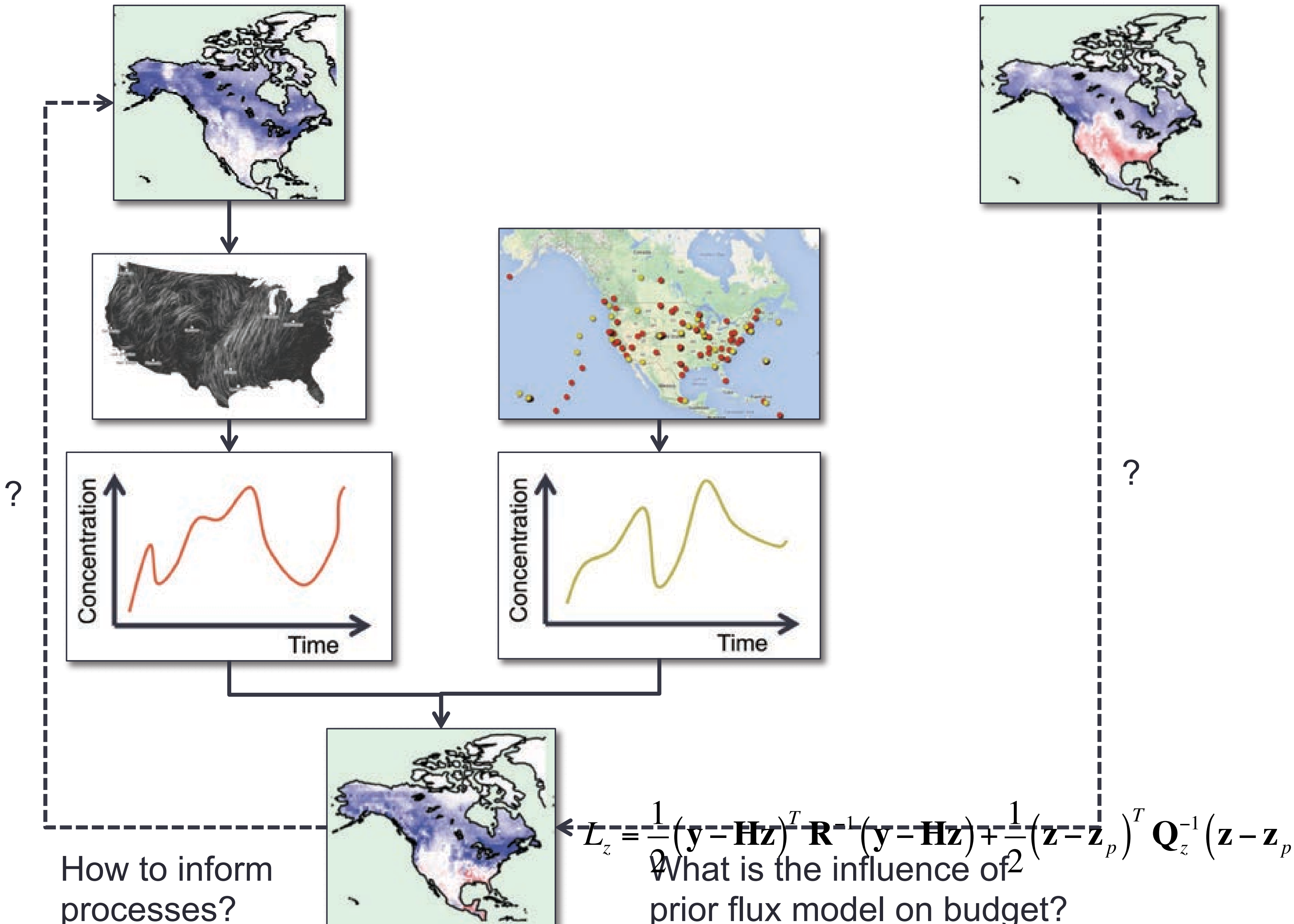
GOSAT
Greenhouse Gases Observing Satellite

OCO-2
ORBITING CARBON OBSERVATORY

Observations, y

Atmospheric transport, *h(.)*





$$L_z = \frac{1}{2}\left(\mathbf{y}-\mathbf{Hz}\right)^T \mathbf{R}^{-1}\left(\mathbf{y}-\mathbf{Hz}\right) + \frac{1}{2}\left(\mathbf{z}-\mathbf{z}_p\right)^T \mathbf{Q}_z^{-1}\left(\mathbf{z}-\mathbf{z}_p\right)$$

Mixed linear model

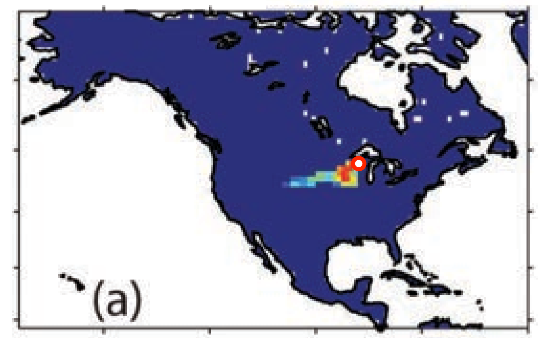
All vary in space and time

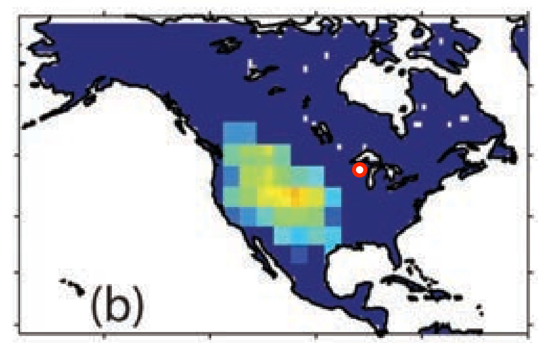
$$y = h(z) + \varepsilon_y + \varepsilon_h + \varepsilon_{rep} + \varepsilon_{agg}$$

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

High spatiotemporal resolution for $\mathbf{z}$

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



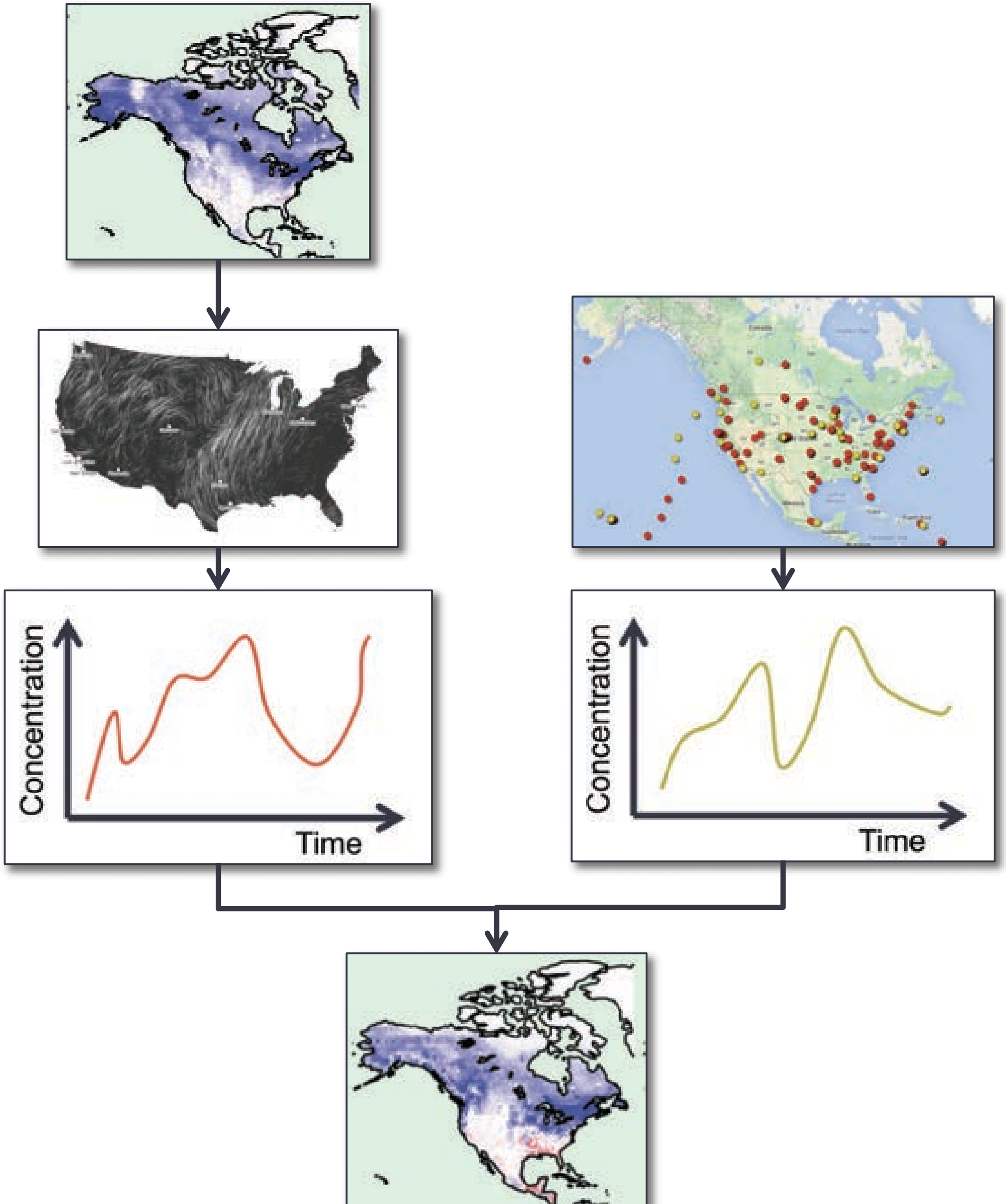


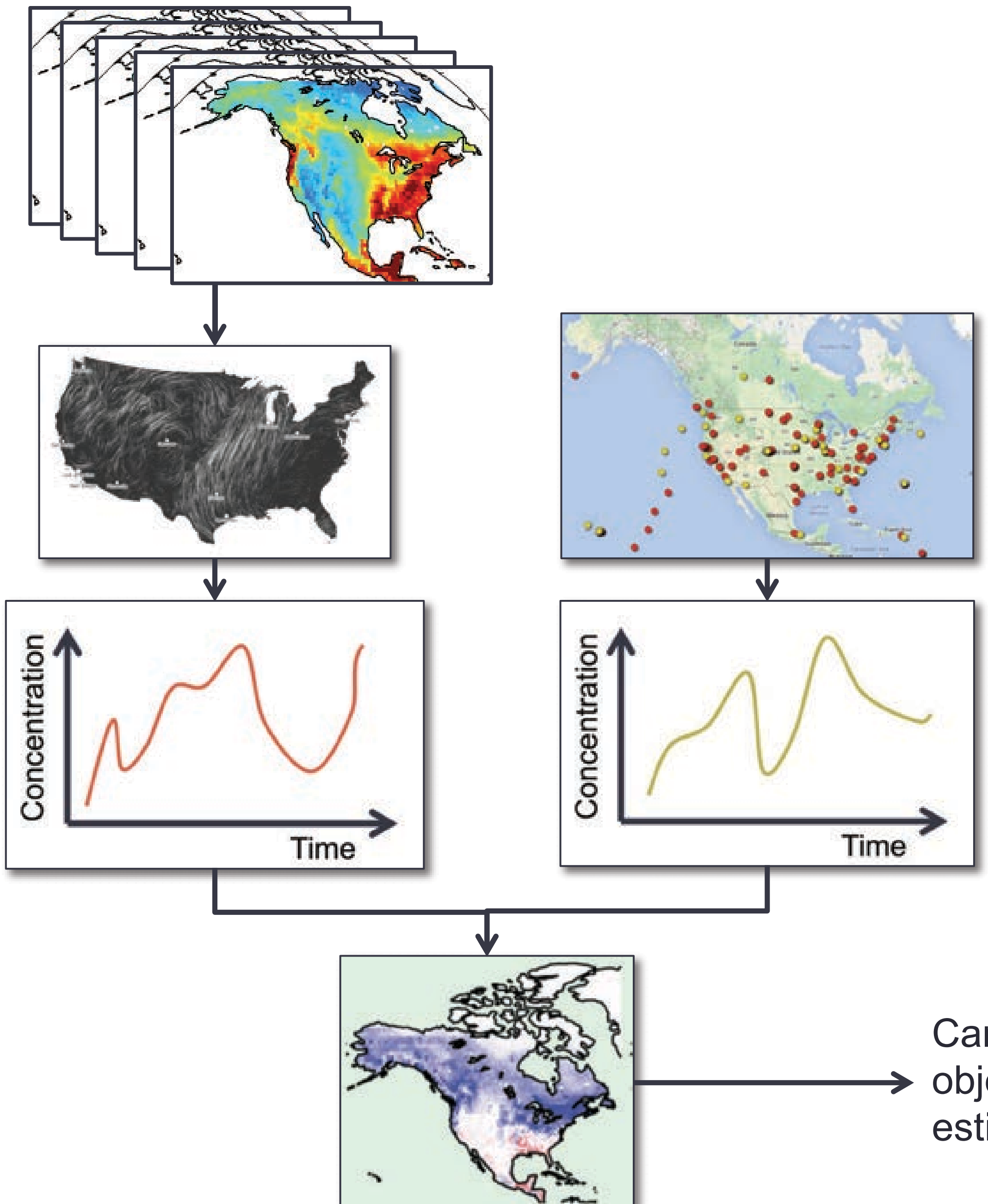
$\mathbf{y} = \mathbf{HX}\beta + \mathbf{H}\xi + \varepsilon$ BIC for model selection (space-time correlated residuals)

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

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

ReML for parameter estimation


Concentration
Time
Concentration
Time



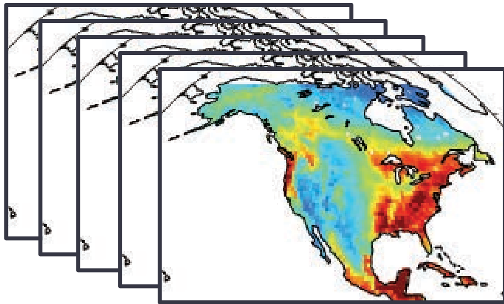
Bigger, better, faster, more!
Michalak et al. (2004), Gourdji et al. (2010), Chatterjee (2012), Yadav (2013a,b), Miller (2014a)

Biospheric CO_2 budgets:
Mueller et al. (2008), Gourdji et al. (2008, 2012)

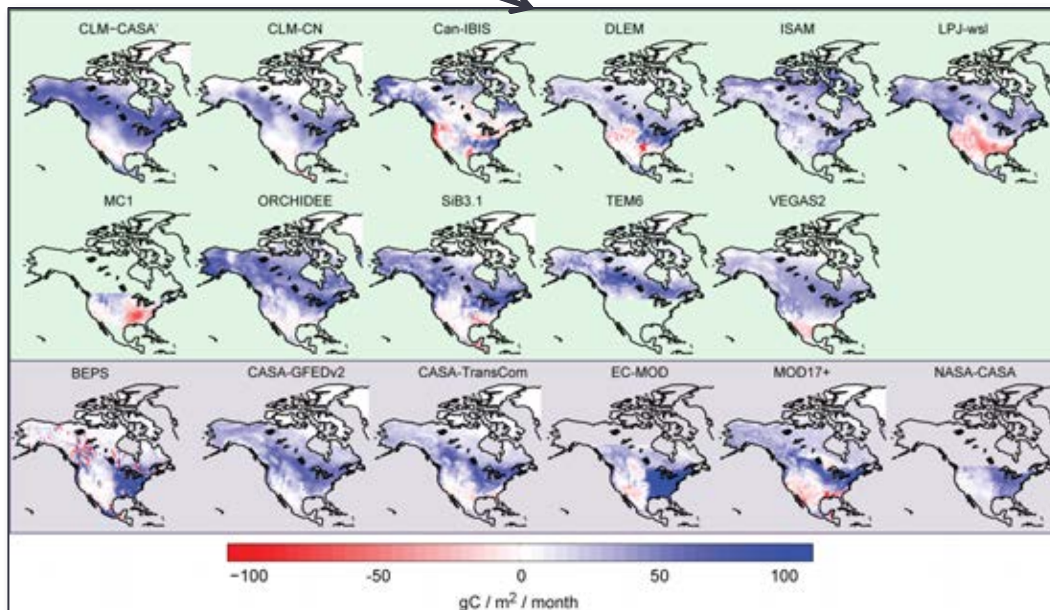
Anthropogenic CO_2 budgets:
Shiga et al. (2014), Yadav et al. (in revision)

CH_4 and N_2O budgets:
Miller et al. (2012, 2013, 2014b)

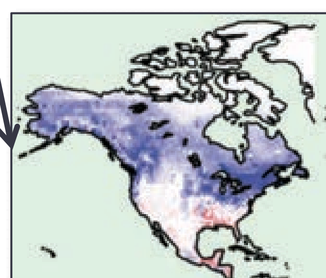
Can provide more objective budget estimates



Can evaluate models' process representations



Can provide process information directly at target scales



Can confront models with independent flux estimate

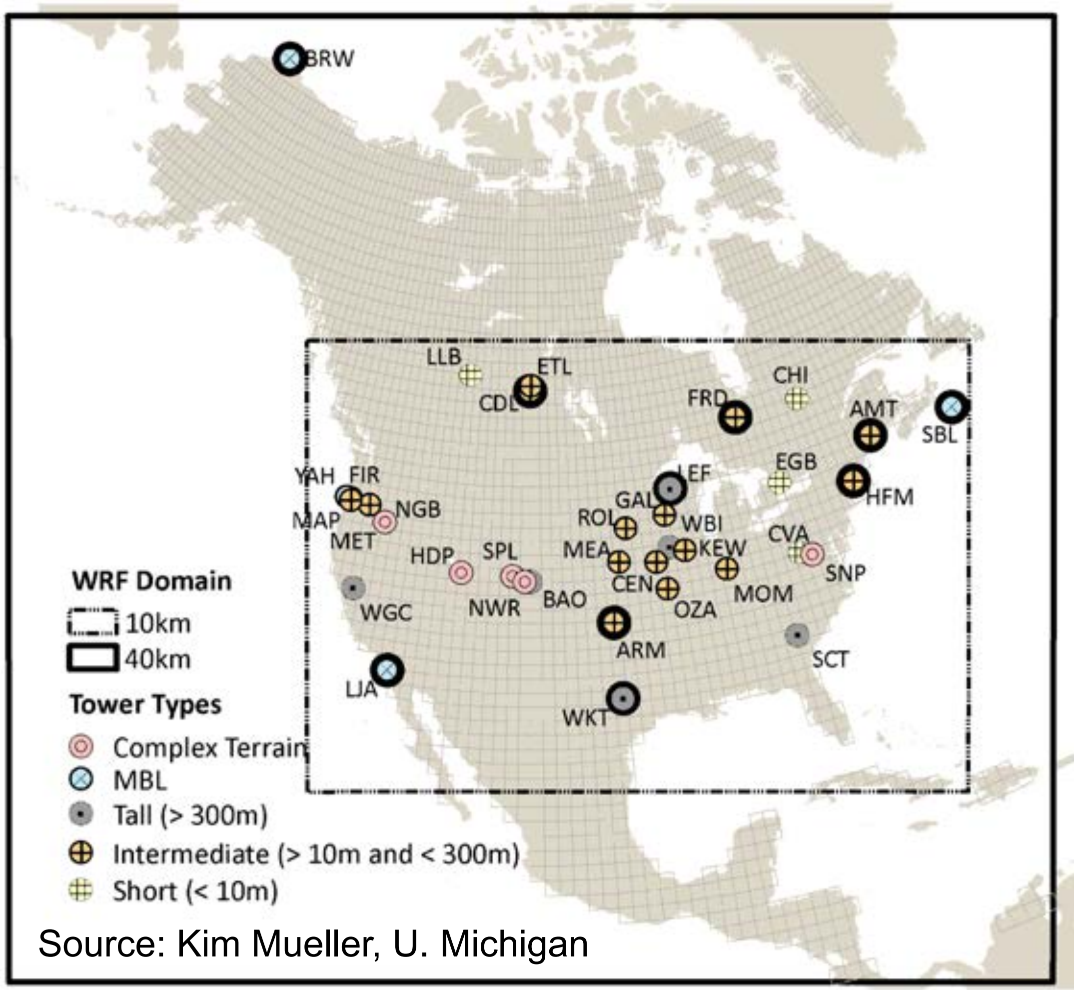
Increasing cost of inversions

Regional CO_2 inversions over North America for one year at 1° x 1°; 3-hourly

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

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

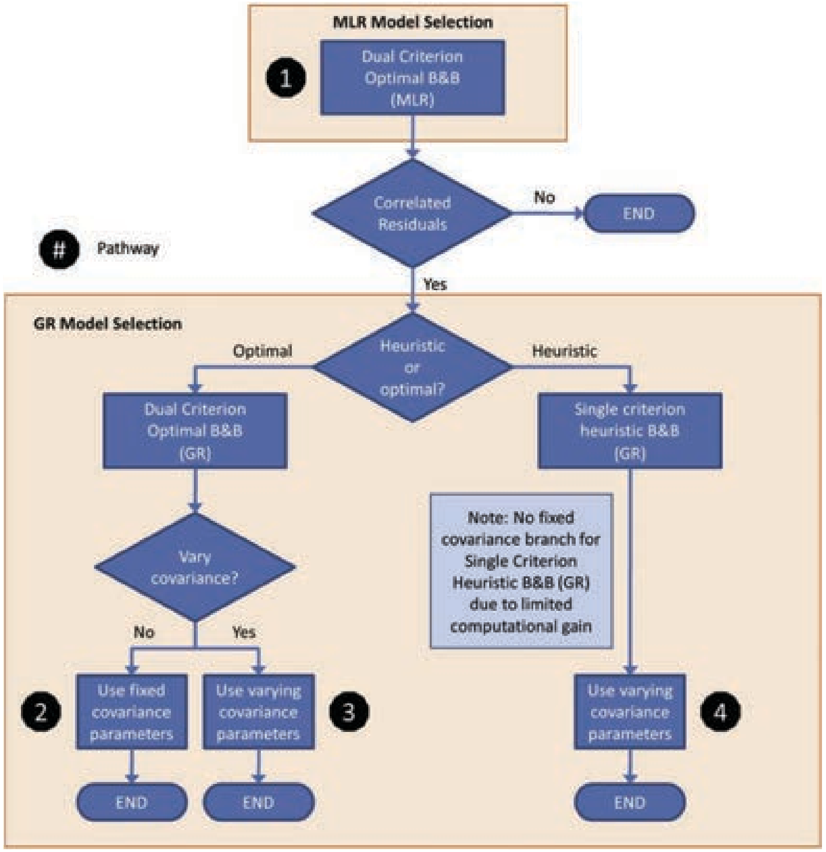
$\mathbf{X}$: ~10^2

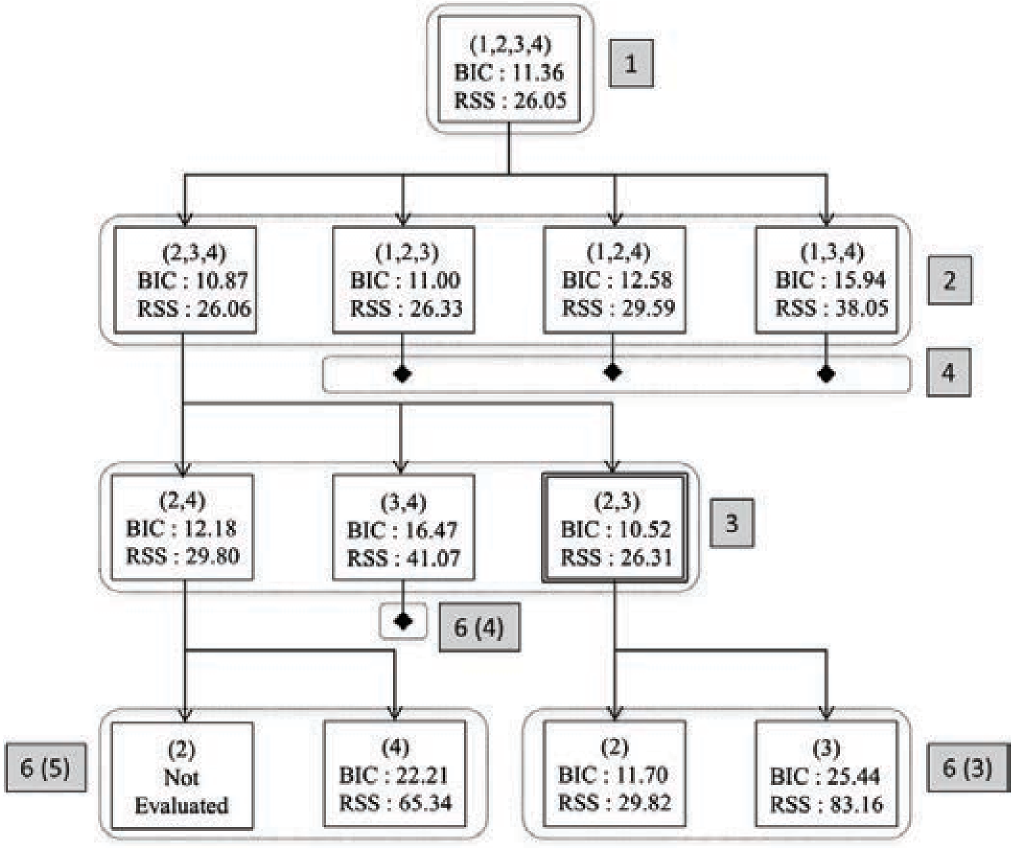


Source: Kim Mueller, U. Michigan

($\mathbf{H}$: ~10^5x10^6; $\mathbf{Q}$: ~10^6x10^6)

Branch & bound algorithm for model selection





k covariate yields 2^k candidate models

Matrix multiplication & posterior covariances

$$\mathbf{y} \sim N\left(\mathbf{HX}\beta, \mathbf{HQH}^T + \mathbf{R}\right)$$

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

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

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

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

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

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

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

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

Ensemble SRF approaches

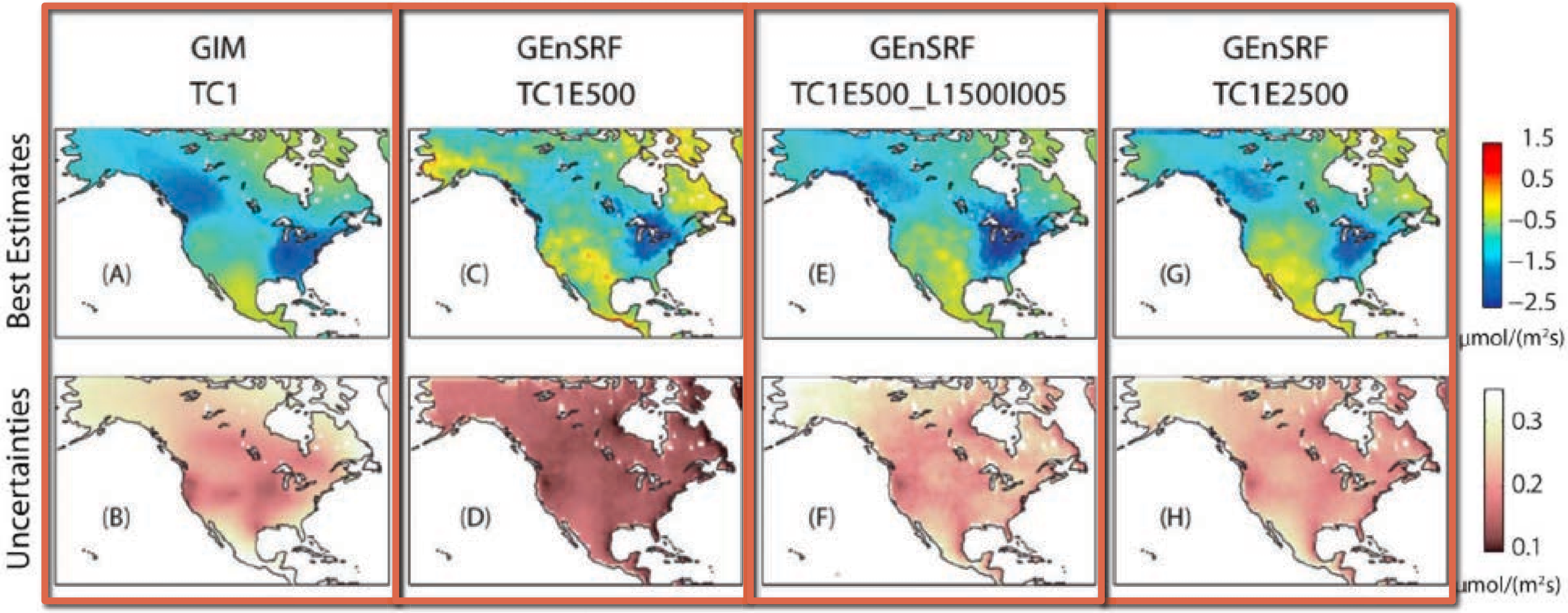
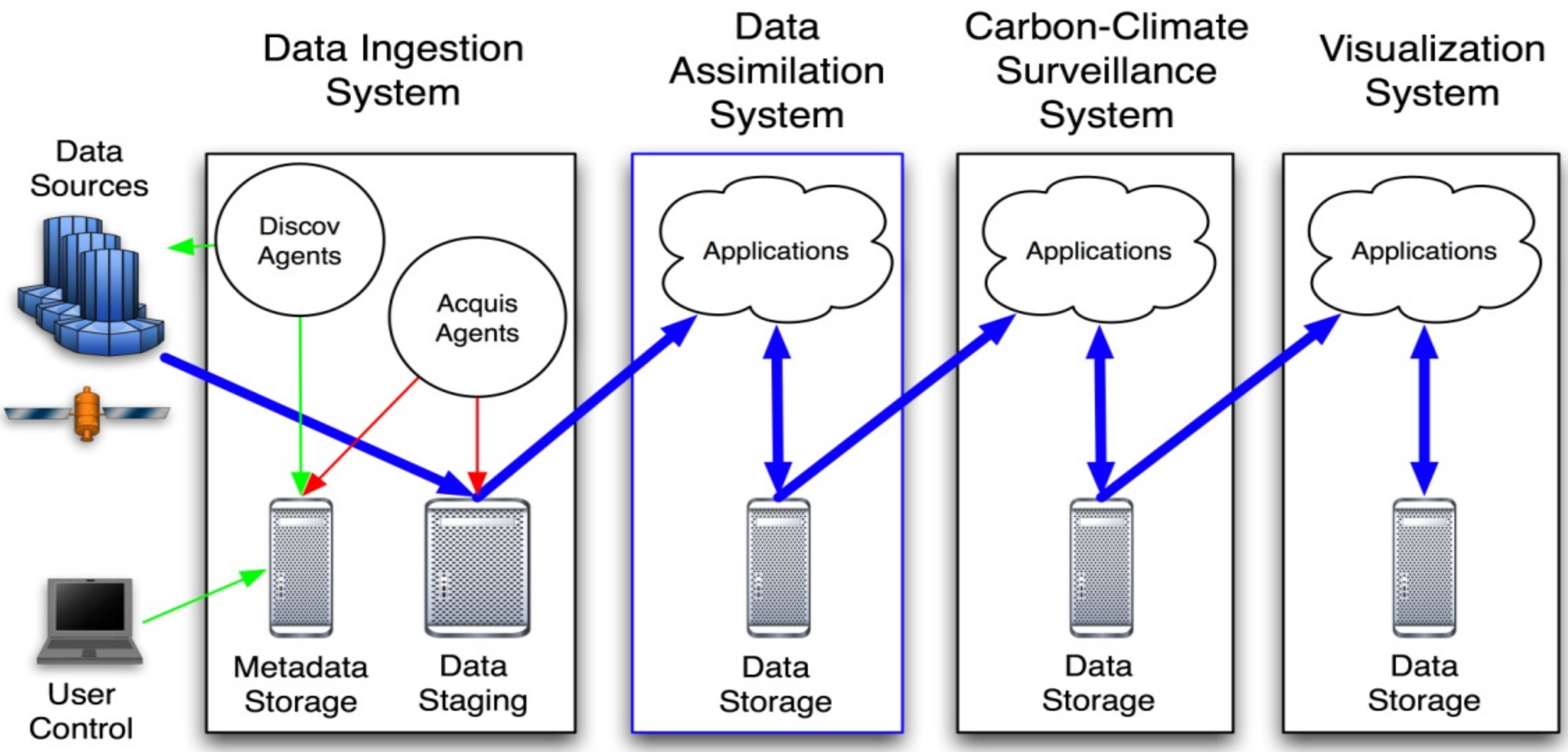


Figure 4. TC1 (top) flux estimates and (bottom) associated uncertainties aggregated to the monthly scale for (a and b) GIM and (c–h) three different GEnSRF runs.

Features:

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

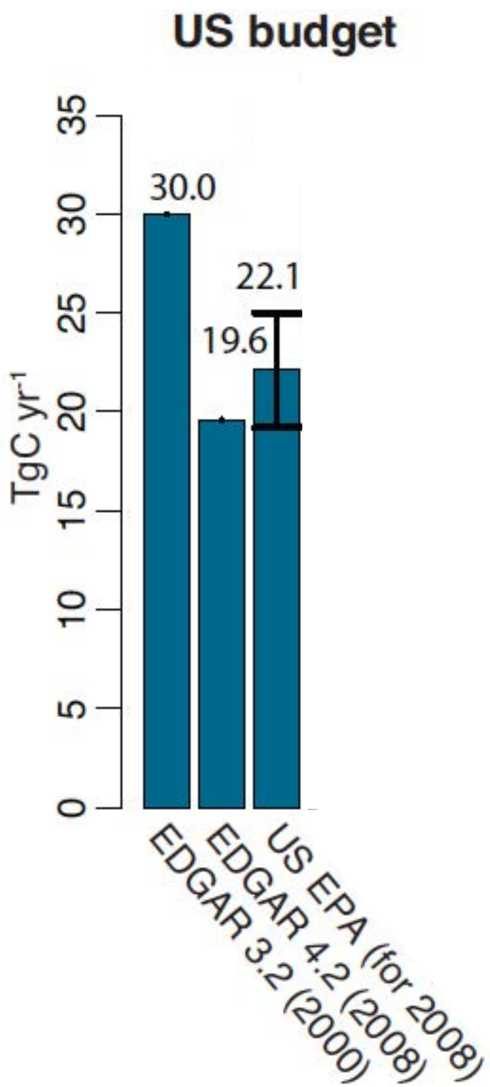
Real-Time Large-Scale Parallel Intelligent CO_2 Data Assimilation System

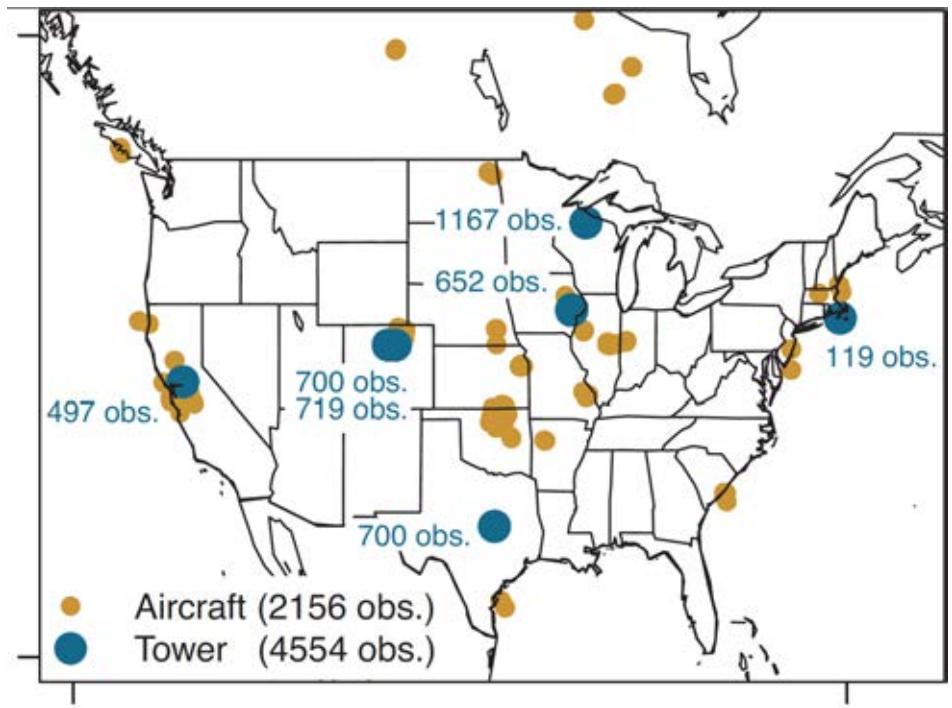


Take home messages

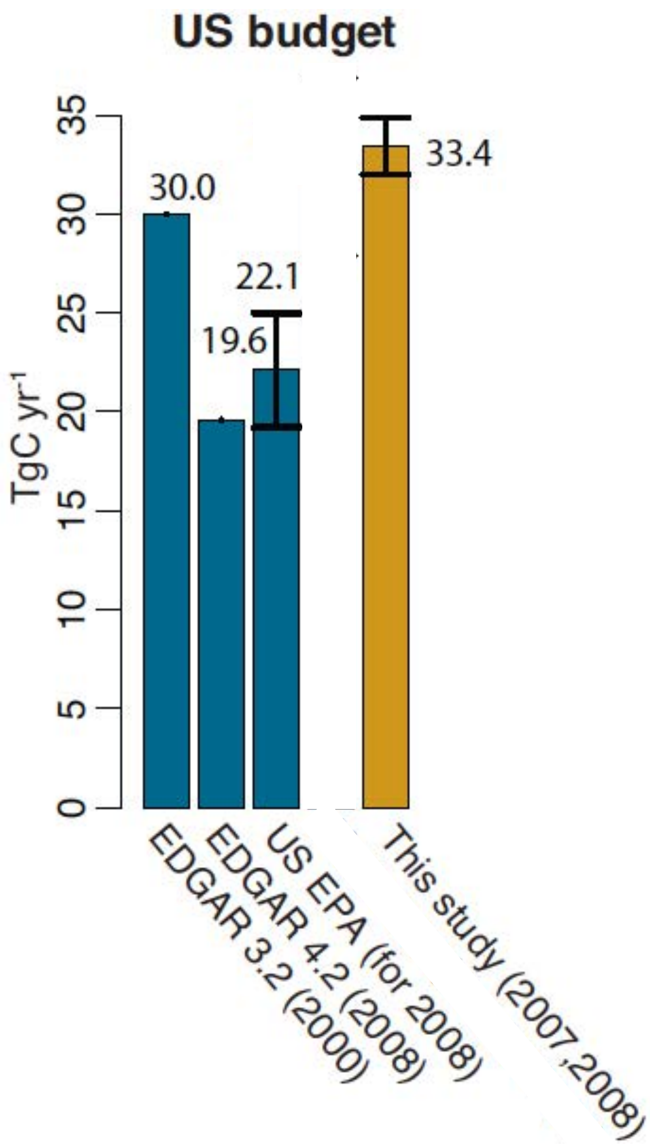
- The need to constrain greenhouse gas budgets inevitably leads to the need for the solution of inverse problems
- These inverse problems:
  - Require (intelligently) choosing among many uncomfortable assumptions
  - Are becoming increasingly statistically sophisticated and computationally demanding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

U.S. anthropogenic methane emissions



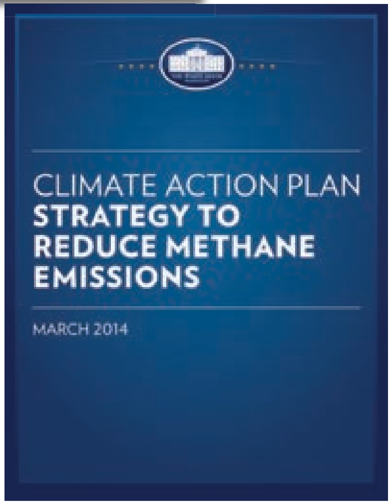


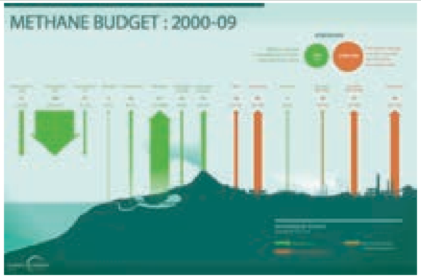
U.S. methane emissions



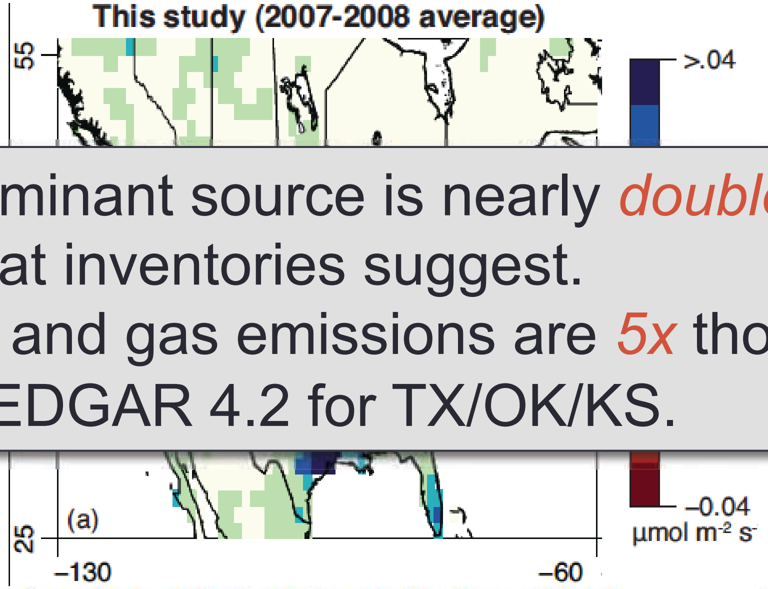
U.S. anthropogenic methane emissions are 50% higher than EPA estimates

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





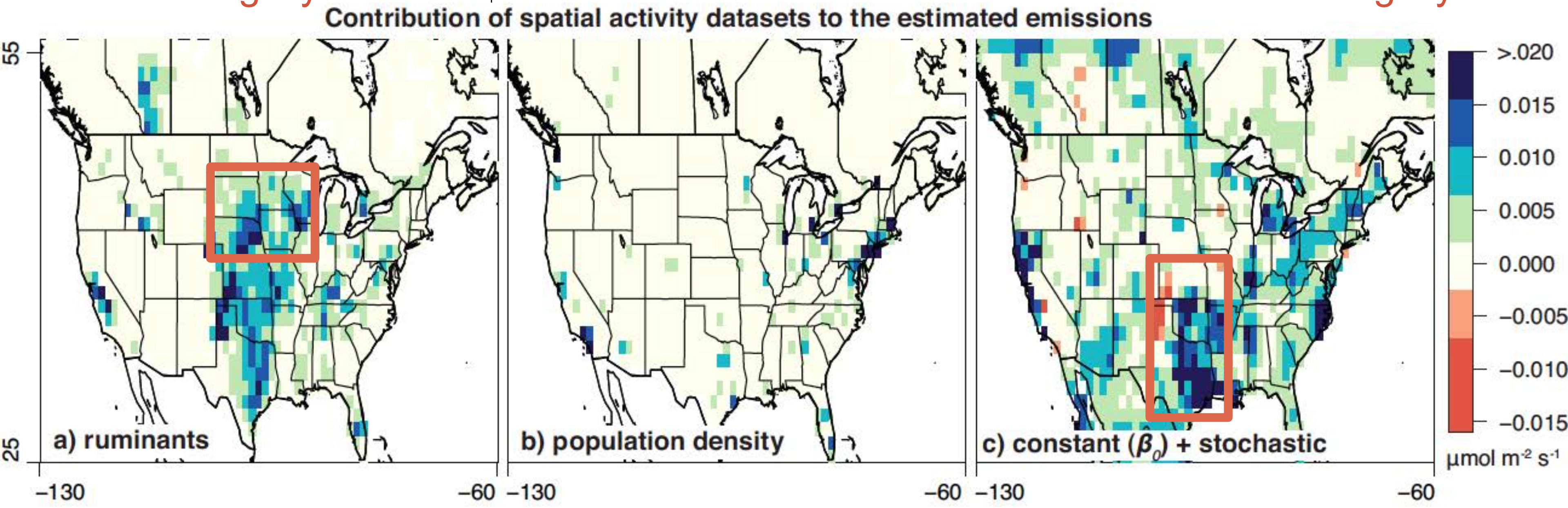
Estimated methane fluxes



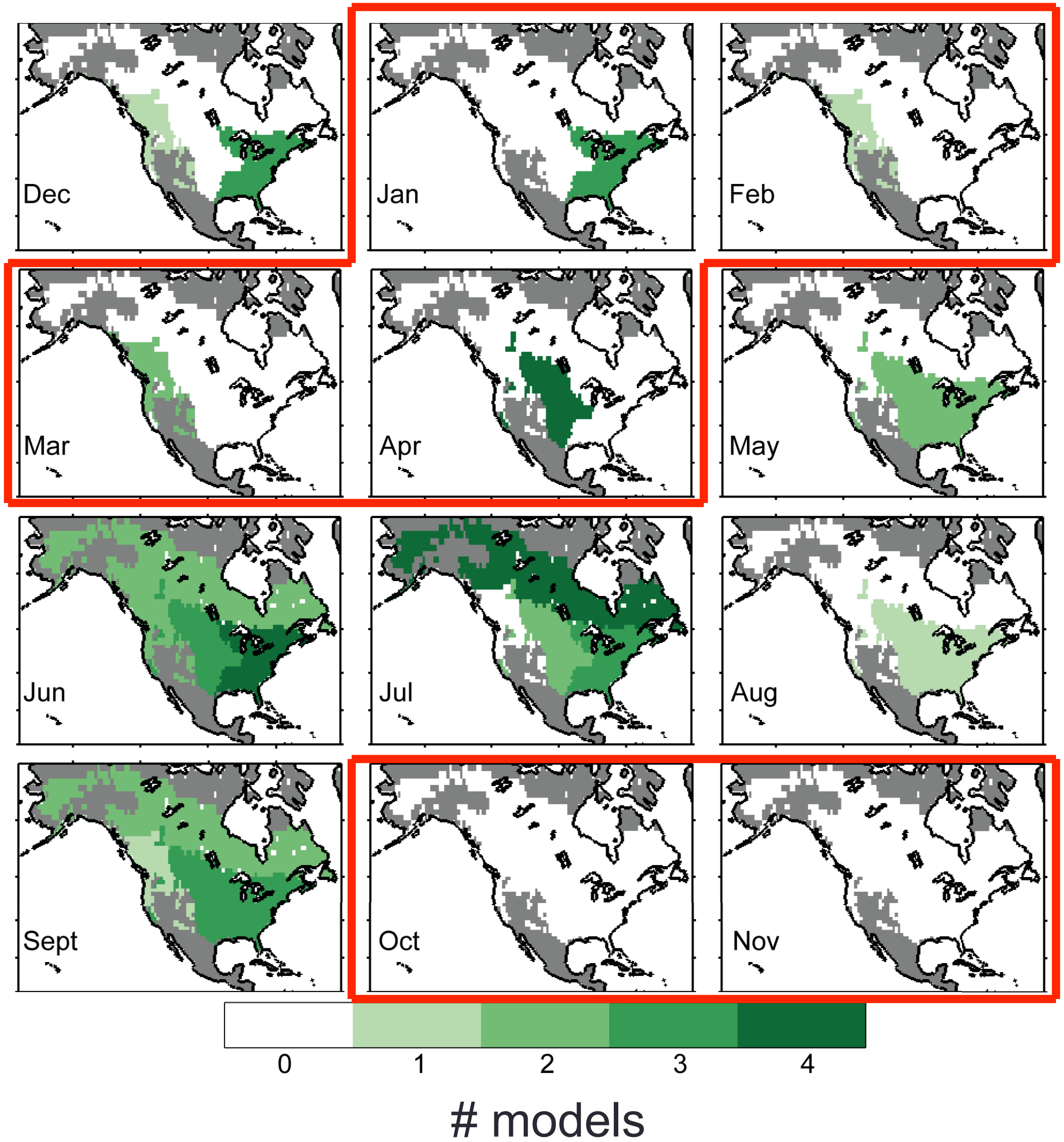
Ruminant source is nearly *double* what inventories suggest.
Oil and gas emissions are *5x* those in EDGAR 4.2 for TX/OK/KS.

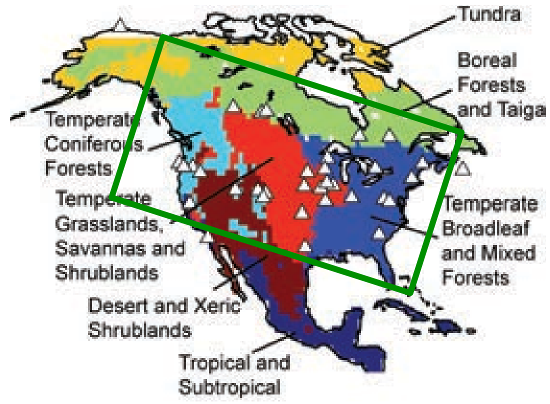
3.4 ± 0.7 TgC yr^{-1}

3.7 ± 2.0 TgC yr^{-1}



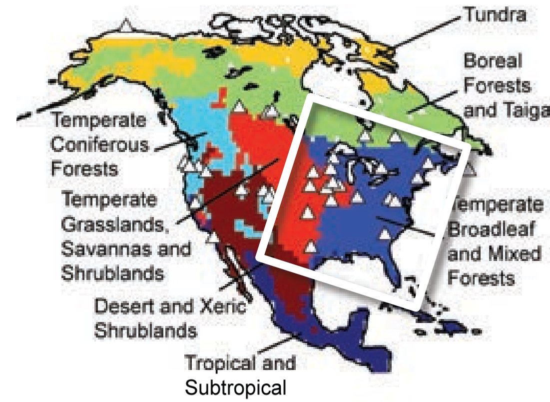
Confronting model flux *patterns* with obs

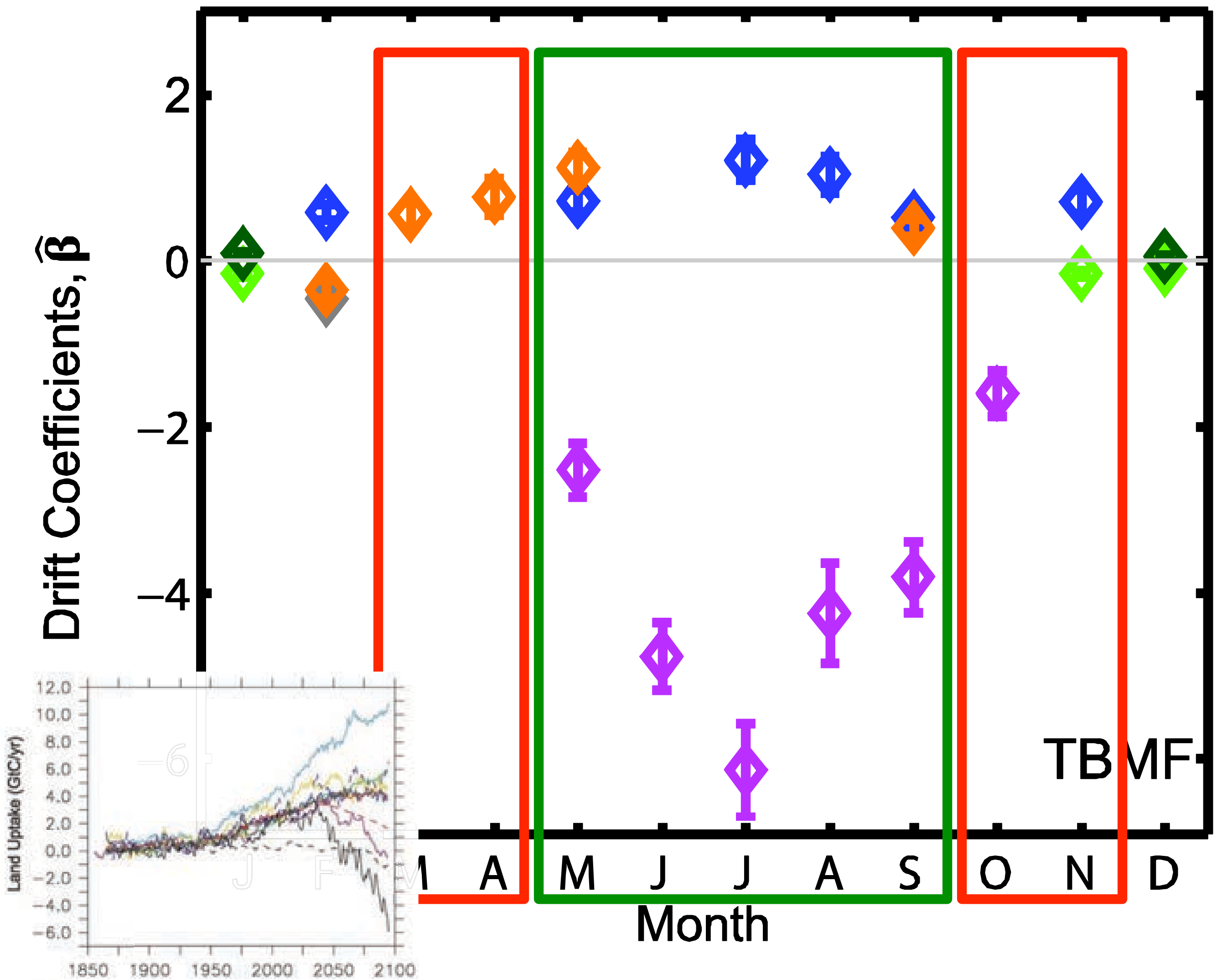




Models' flux patterns do not explain observed variability in atmospheric observations for much of the year, but they do better during growing season.

Providing process information directly at target scales

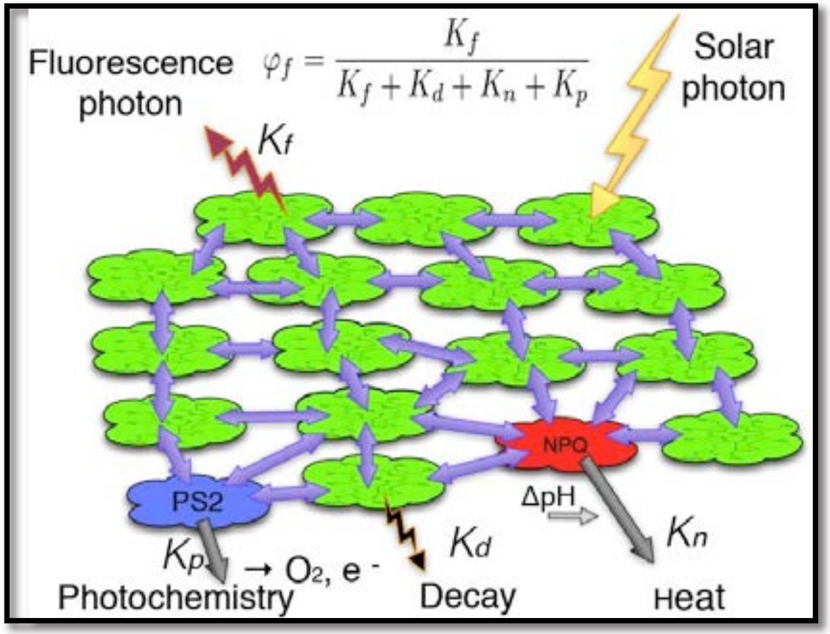




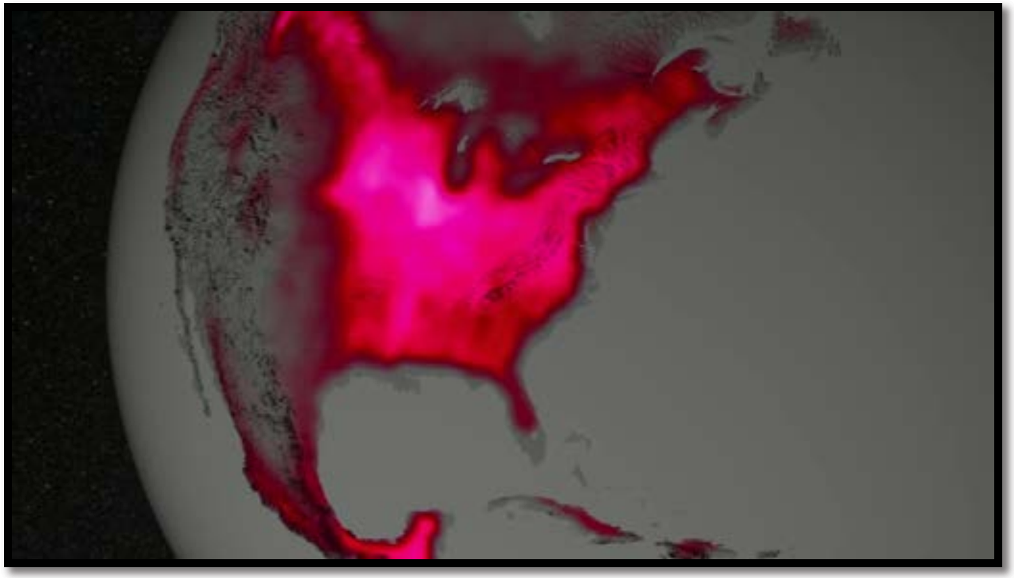
Models explain flux patterns well when flux patterns are dominated by patterns in radiation

Solar Induced Fluorescence

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

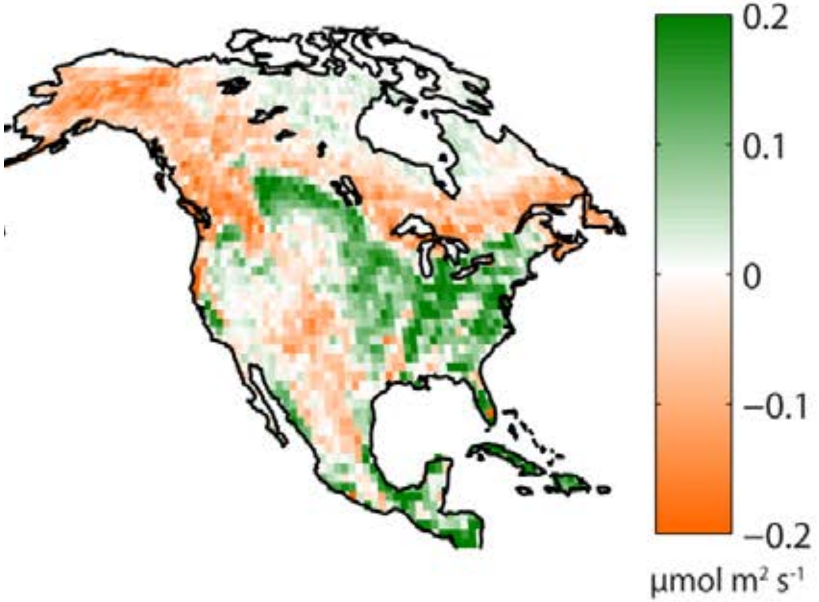


Source: Frankenberg, 2011



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

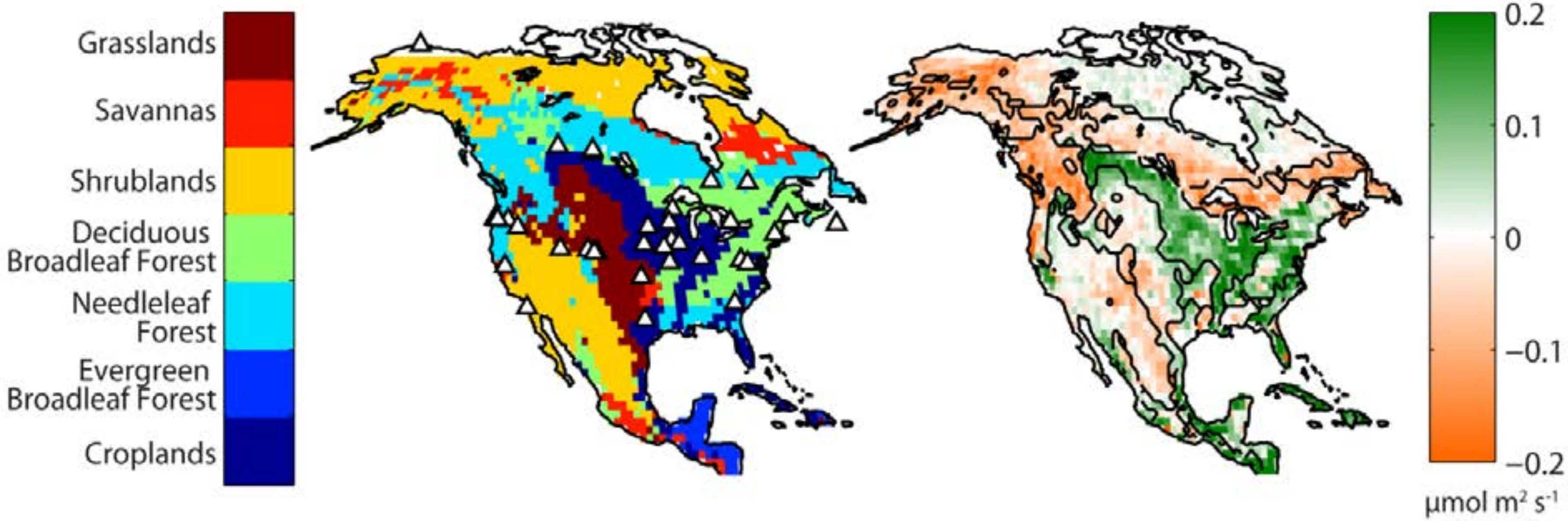
Differences at 1° x 1°, aggregated over March to October



More net uptake

Less net uptake

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



Informing inversions with SIF leads to redistribution of carbon sink, with increased sink in croplands and reduced sink in needleleaf forests

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