Inferring the sources driving variations in the contemporary carbon cycle



Alexander J. Turner

Assistant Professor, U. Washington

Advanced Modeling & Simulation (AMS) Seminar Series

Funded by a NASA Early Career Faculty (ECF) Grant

NASA Ames Research Center, June 16, 2022

Inferring the sources driving variations in the contemporary carbon cycle

Work I'll talk about today:

Turner et al., *GRL* (2020): *"Observed impacts of COVID-19 on urban CO_2 emissions"*

Turner et al., *BG* (2021): *"Extreme events driving year-to-year differences in GPP across the US"*

Ongoing work: *Developing computationally efficient emulators for atmospheric transport*

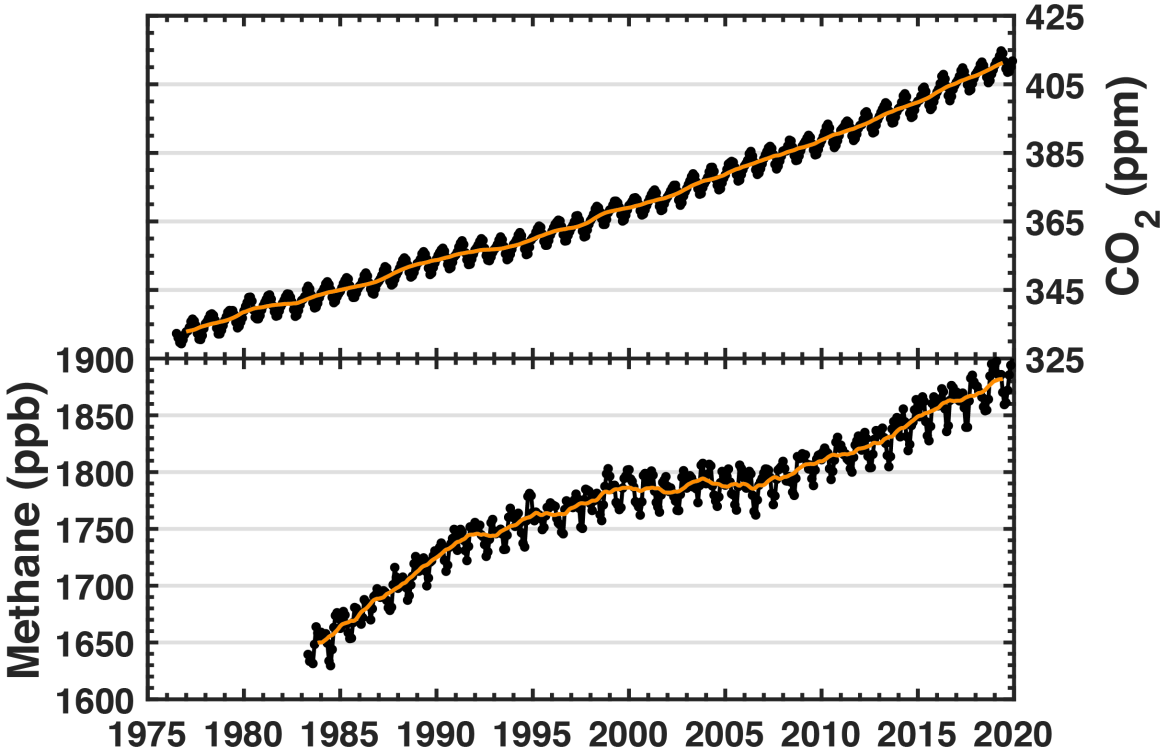
Alexander J. Turner
Assistant Professor, U. Washington

Advanced Modeling & Simulation (AMS) Seminar Series
Funded by a NASA Early Career Faculty (ECF) Grant
NASA Ames Research Center, June 16, 2022

Why do we care about the carbon cycle?

1) Methane & CO_2 are the strongest anthropogenic radiative forcers

- **Account for more than 85% of the GHG forcing since 1750**



2) Recent trends in atmospheric methane are not well understood

Why do we care about the carbon cycle?

1) Methane & CO_2 are the strongest anthropogenic radiative forcers

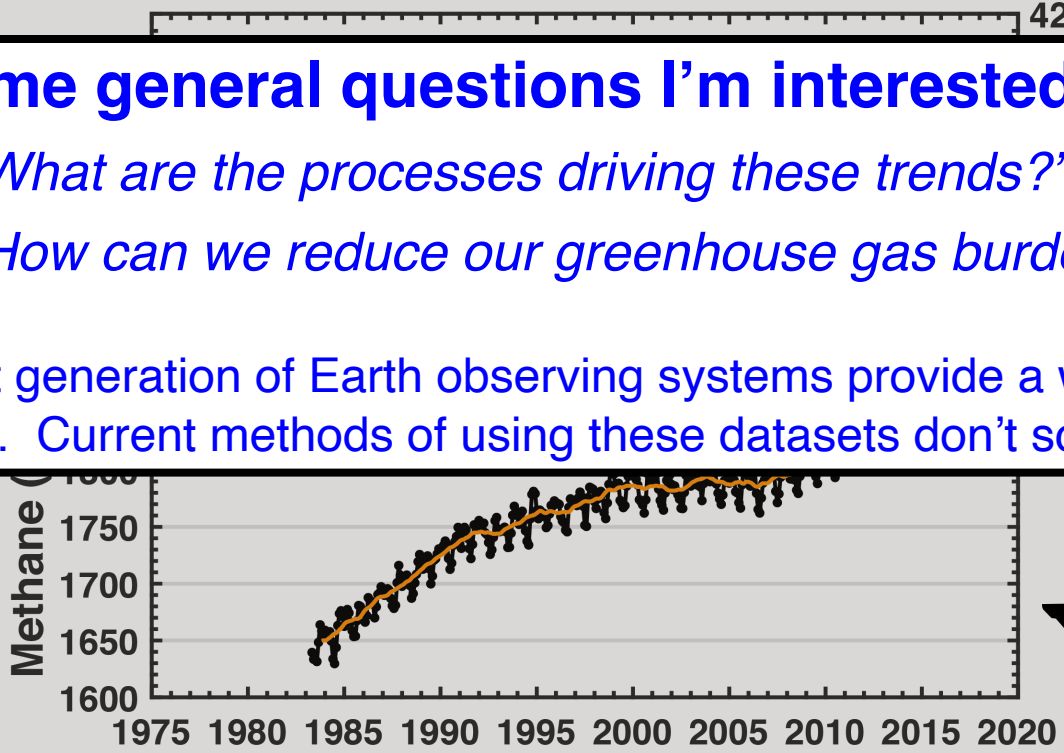
- **Account for more than 85% of the GHG forcing since 1750**

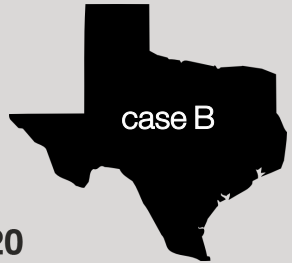
Some general questions I'm interested in:

- *"What are the processes driving these trends?"*
- *"How can we reduce our greenhouse gas burdens?"*

Next generation of Earth observing systems provide a wealth of data. Current methods of using these datasets don't scale well.



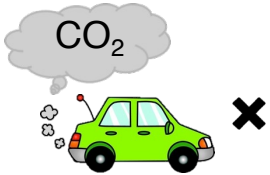




2) Recent trends in atmospheric methane are not well understood

Methods of estimating emissions

Bottom-up



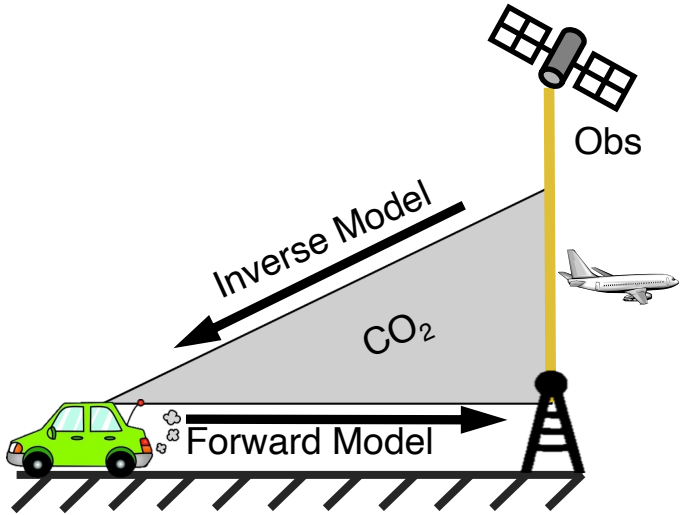
×

Traffic Density



= Map of US traffic emissions

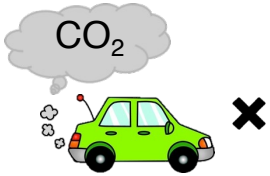
Top-down



Atmospheric observations allow us to improve the bottom-up inventory

Methods of estimating emissions

Bottom-up



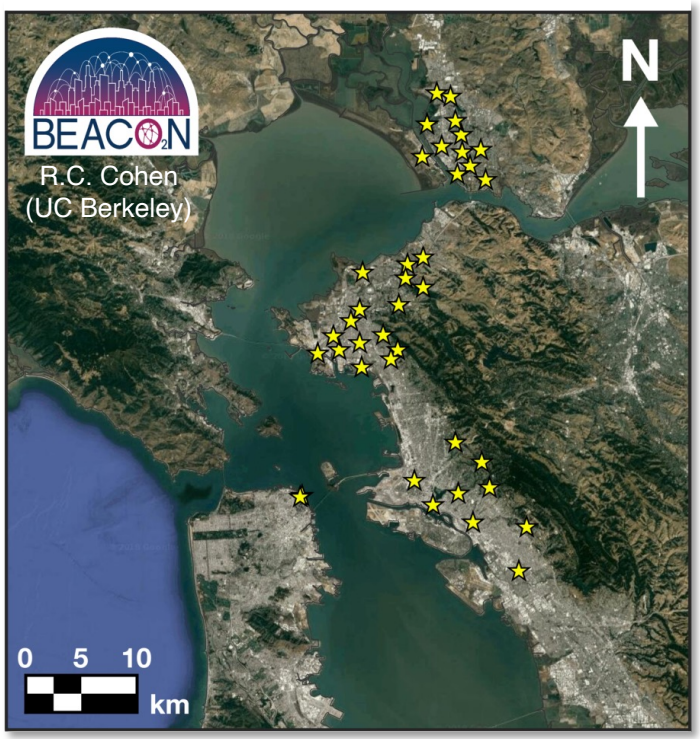
×

Traffic Density



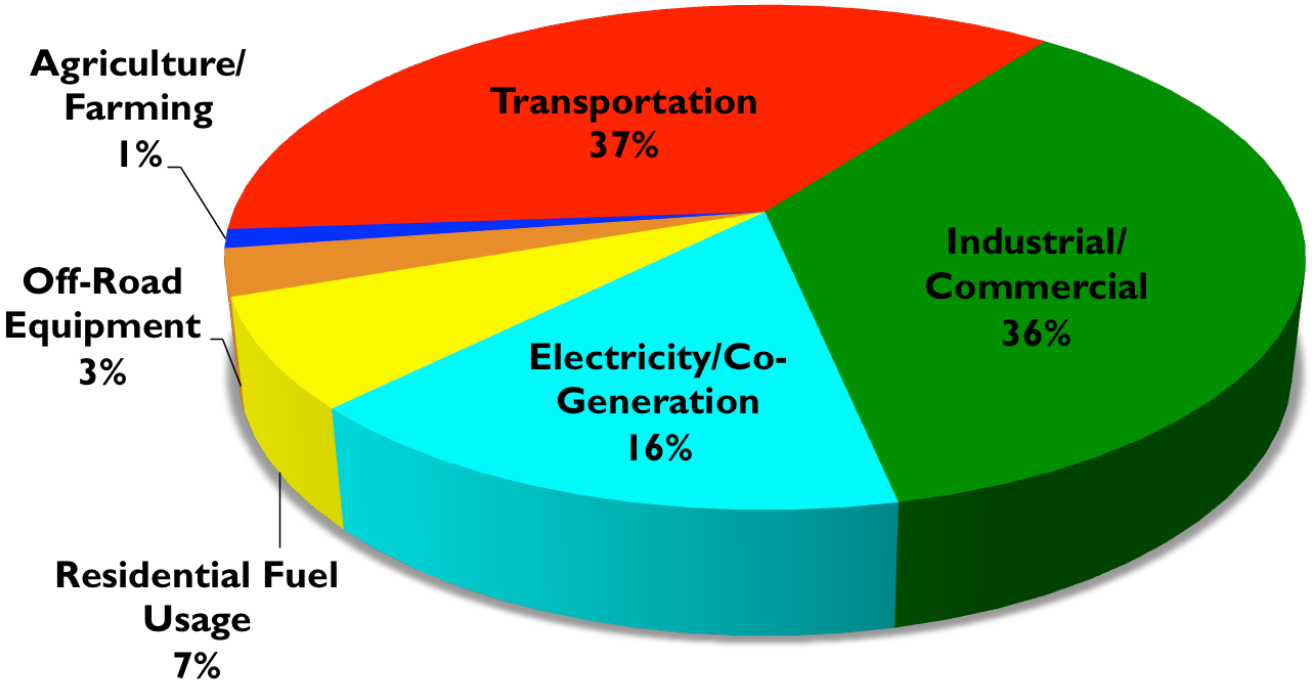
= Map of US traffic emissions

Top-down



Atmospheric observations allow us to improve the bottom-up inventory

Bottom-up estimates of urban CO_2



Bay Area Air Quality Management District (BAAQMD) provides county totals

Bottom-up estimates of urban CO_2

No.	Plant #	Plant Name	Plant Address	City	Zipcode	CO2 Equivalent Emissions (Metric Tons per year)		
						Biogenic	Non-Biogenic	Total
1	11	Shell Martinez Refinery	3485 Pacheco Blvd	Martinez	94553	-	4,976,544	4,976,544
2	10	Chevron Products Company	841 Chevron Way	Richmond	94802	-	4,303,800	4,303,800
3	14628	Tesoro Refining and Marketing Company	150 Solano Way, Avon Refinery	Martinez	94553	-	2,804,678	2,804,678
4	12626	Valero Refining Company - California	3400 E 2nd Street	Benicia	94510	-	2,568,988	2,568,988
5	12095	Delta Energy Center	Arcy Lane	Pittsburg	94565	-	1,895,320	1,895,320
6	16	ConocoPhillips - San Francisco Refinery	1380 San Pablo Ave	Rodeo	94572	-	1,577,872	1,577,872
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
.	.	.	.	.	.	.	.	.
169	18447	PPF OFF One Market Plaza LLC	Tow	San Francisco	94105	-	4,979	4,979
170	2168	Jelly Belly Candy Company	One Jelly Belly Lane	Fairfield	94533	-	4,858	4,858
171	17315	C & H Sugar Company, Inc	830 Loring Avenue	Crockett	94525	-	4,817	4,817
172	10408	County Asphalt	5501 Imhoff Drive	Martinez	94553	-	4,807	4,807

- ☐ **Transportation (37%)**
- ☑ **Industrial/Commercial (36%)** ← Reported by BAAQMD
- ☑ **Electricity/Co-Generation (16%)** ← Reported by BAAQMD
- ☑ **Residential Fuel Usage (7%)** ← map to population and temperature
- ☒ ~~**Off-Road Equipment (3%)**~~ neglecting due to small contribution
- ☒ ~~**Agriculture/Farming (1%)**~~ neglecting due to small contribution

Bottom-up estimates of urban CO_2

High-resolution mapping of motor vehicle carbon dioxide emissions

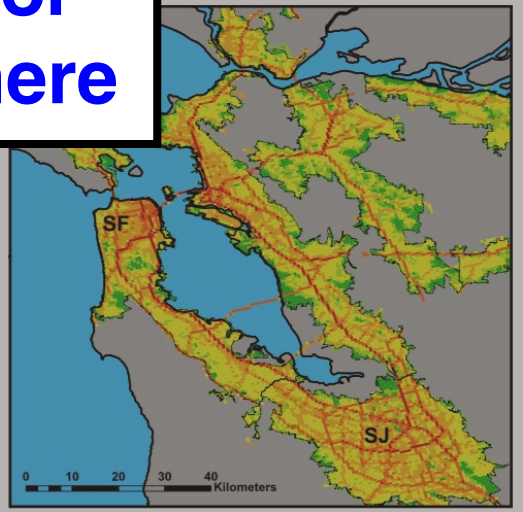
Brian C. McDonald[1], Zoe C. McBride[1], Elliot W. Martin[2], and Robert A. Harley[1]

[1]Department of Civil and Environmental Engineering, University of California, Berkeley, California, USA, [2]Transportation Sustainability Research Center, University of California, Berkeley, California, USA

169	18447	PPF C			
170	2168	Jelly B			
171	17315	C & H			
172	10408	County Asphalt	5501 Imhoff Drive	Martinez	94553

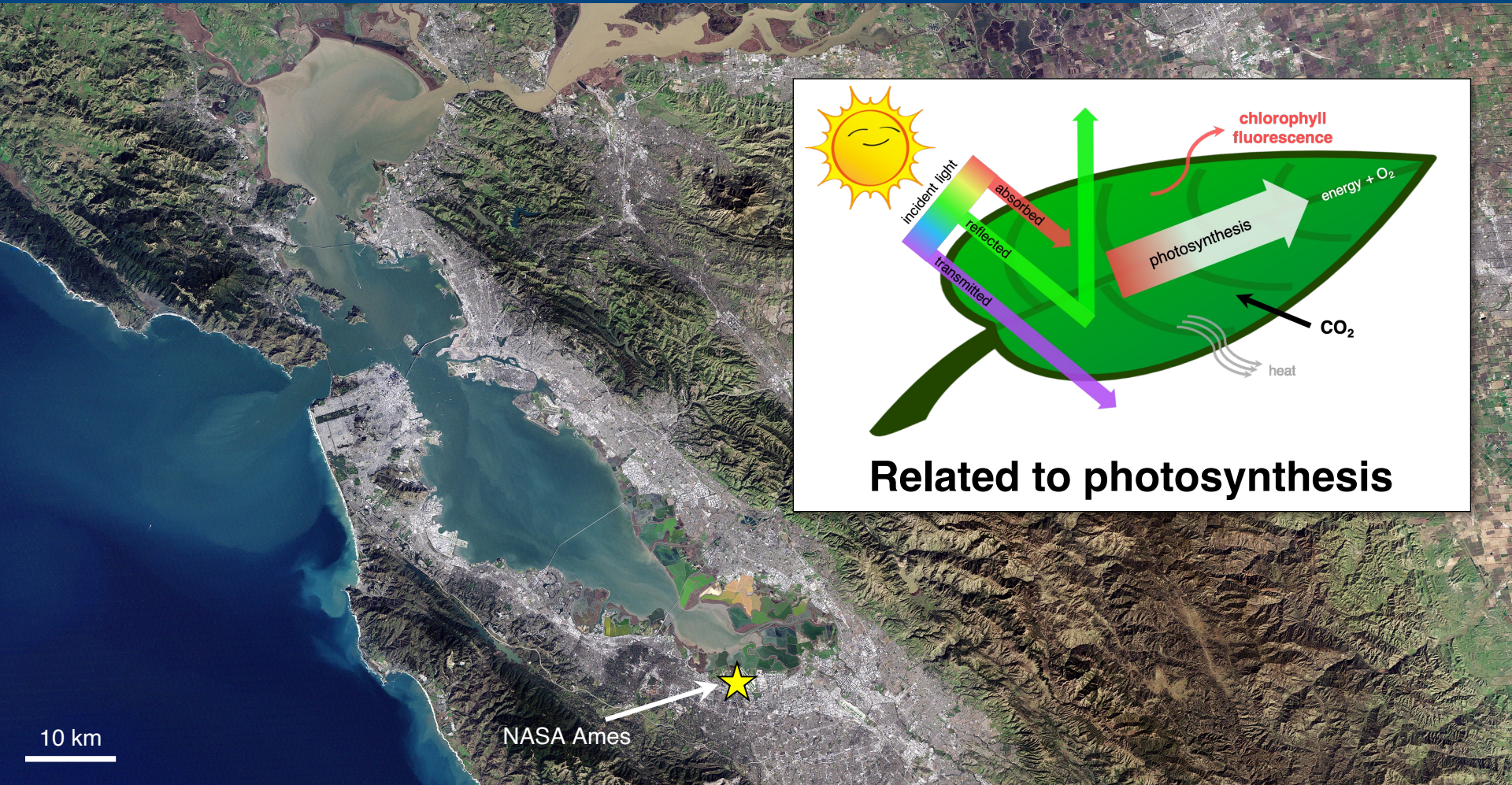
Still need to account for uptake from the biosphere

- ☑ **Transportation (37%)**
- ☑ **Industrial/Commercial (36%)** ← Reported by BAAQMD
- ☑ **Electricity/Co-Generation (16%)** ← Reported by BAAQMD
- ☑ **Residential Fuel Usage (7%)** ← map to population and temperature
- ☐ ~~**Off-Road Equipment (3%)**~~ neglecting due to small contribution
- ☐ ~~**Agriculture/Farming (1%)**~~ neglecting due to small contribution

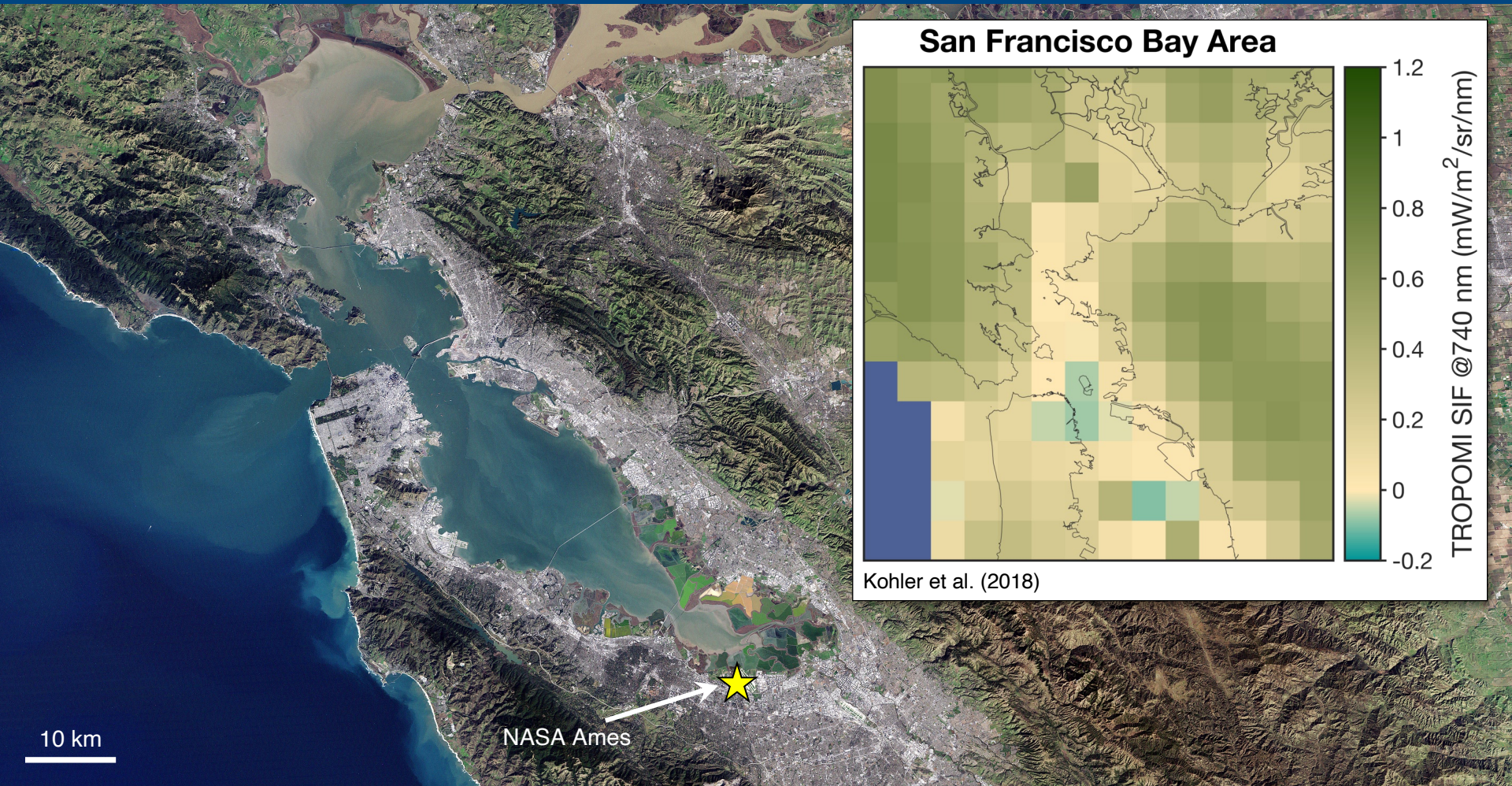


McDonald et al. (2014)

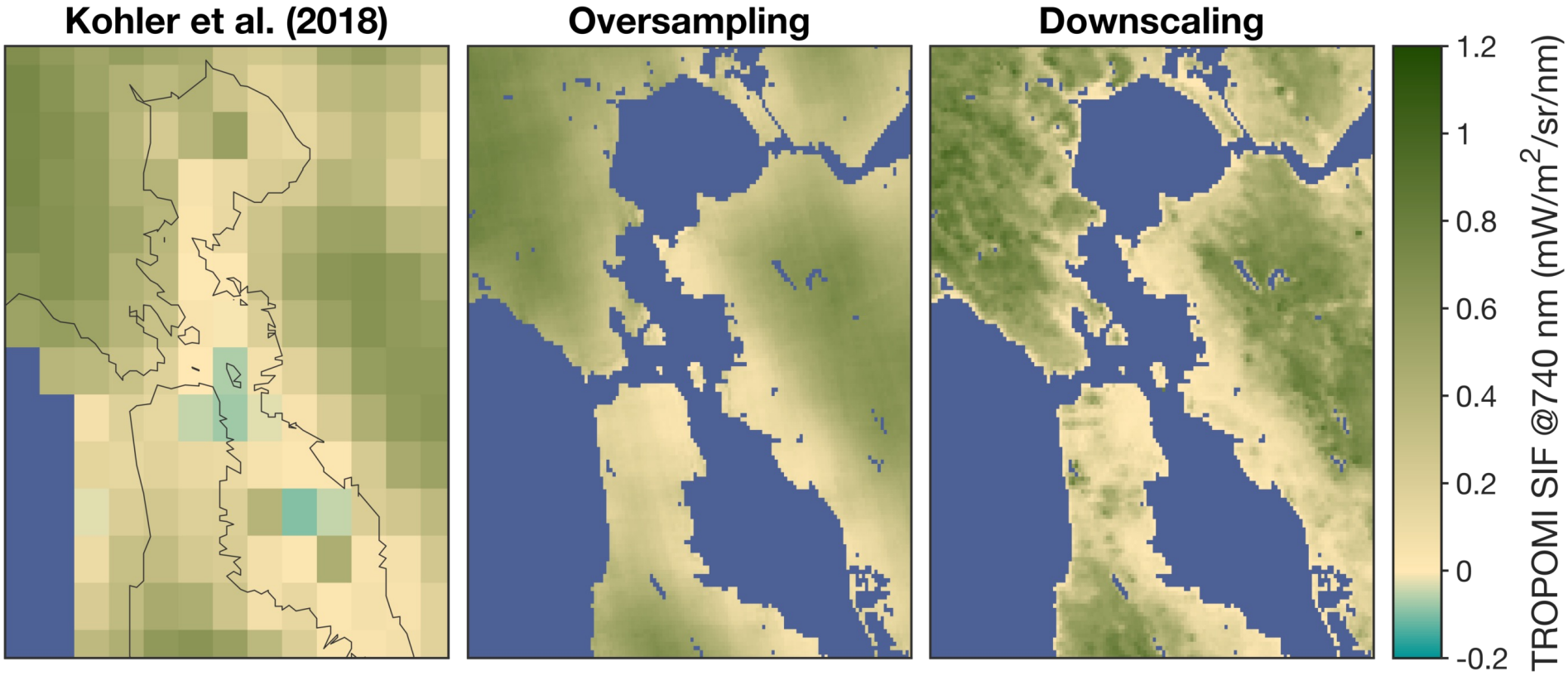
Inferring biospheric uptake of CO_2 using solar-induced chlorophyll fluorescence



Inferring biospheric uptake of CO_2 using solar-induced chlorophyll fluorescence

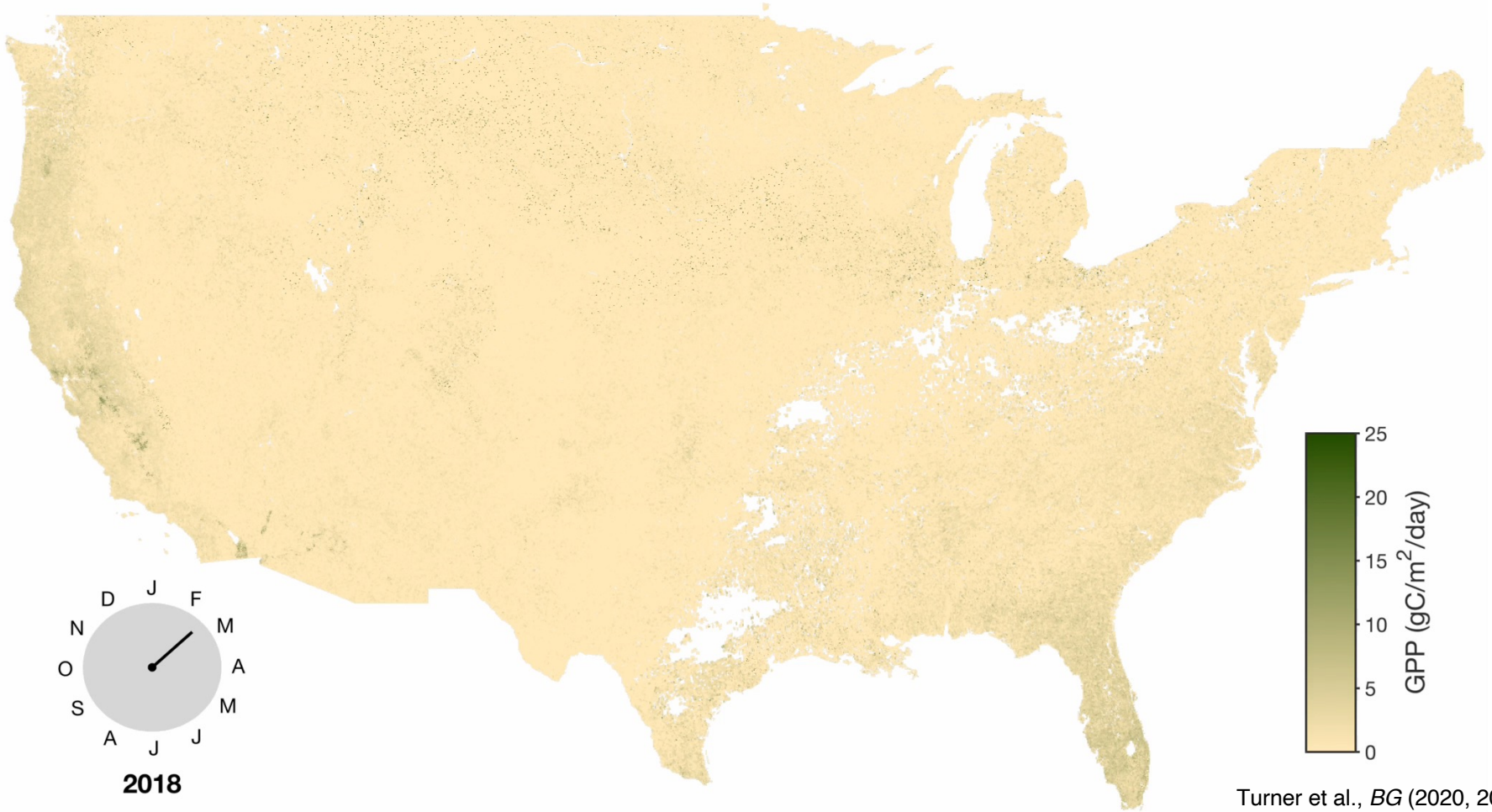


Obtaining high spatial resolution information from SIF



Turner et al., *BG* (2020, 2021)

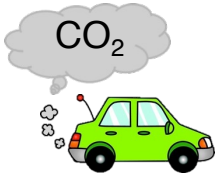
We then extended this over all of CONUS at 500-m resolution



Turner et al., *BG* (2020, 2021)

Putting it all together

Bottom-up



×

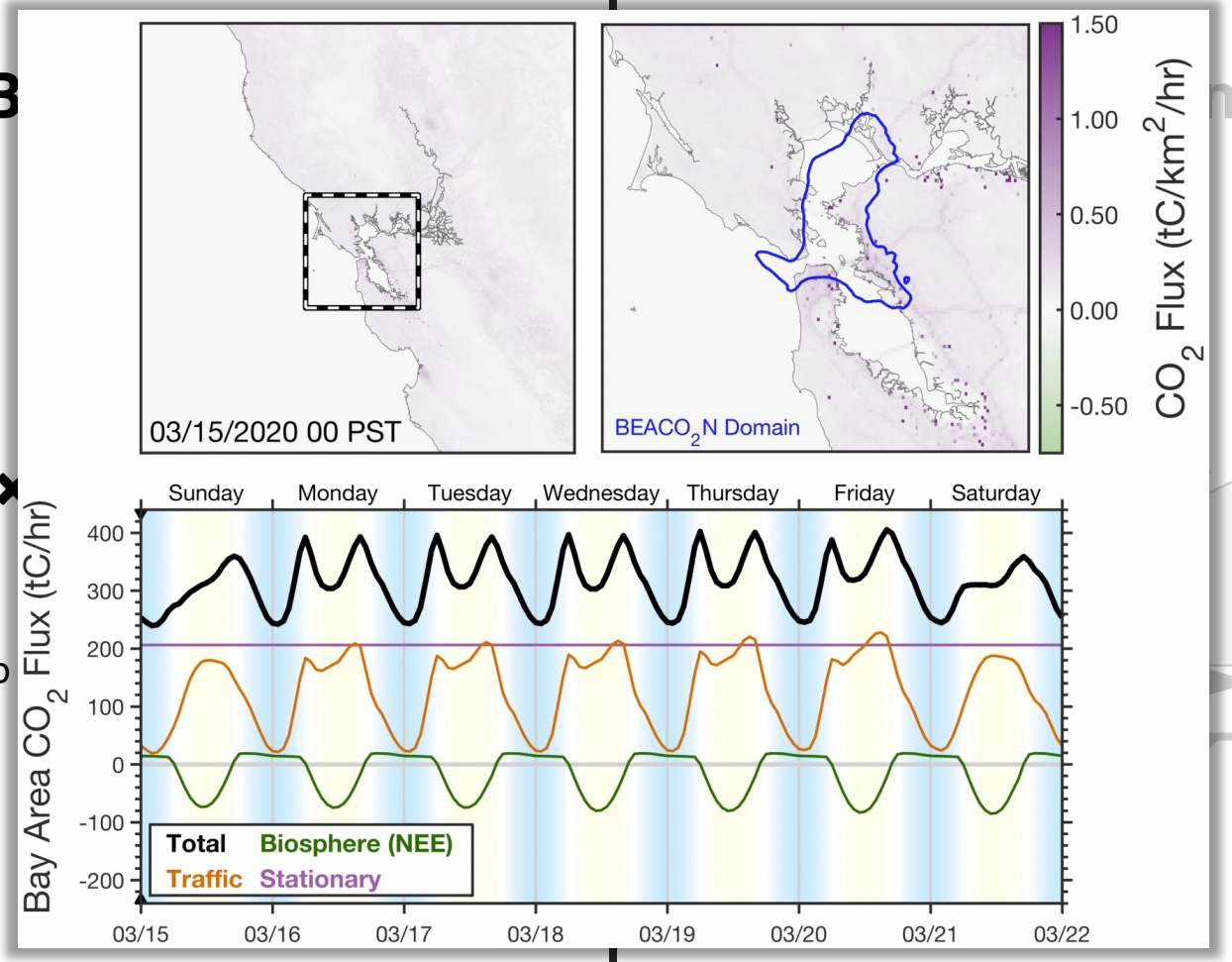
Traffic Density



= Map of US traffic emissions

Top-down

Putting it all together

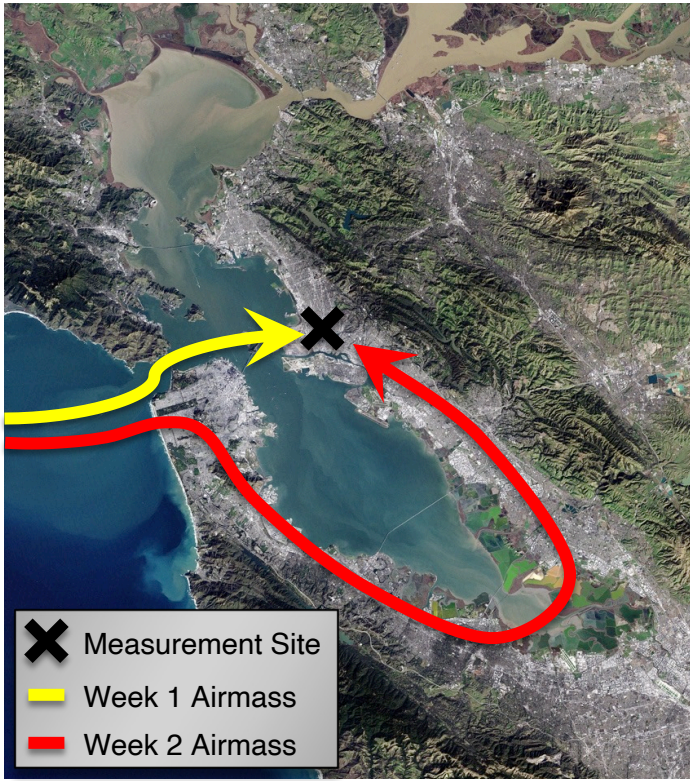


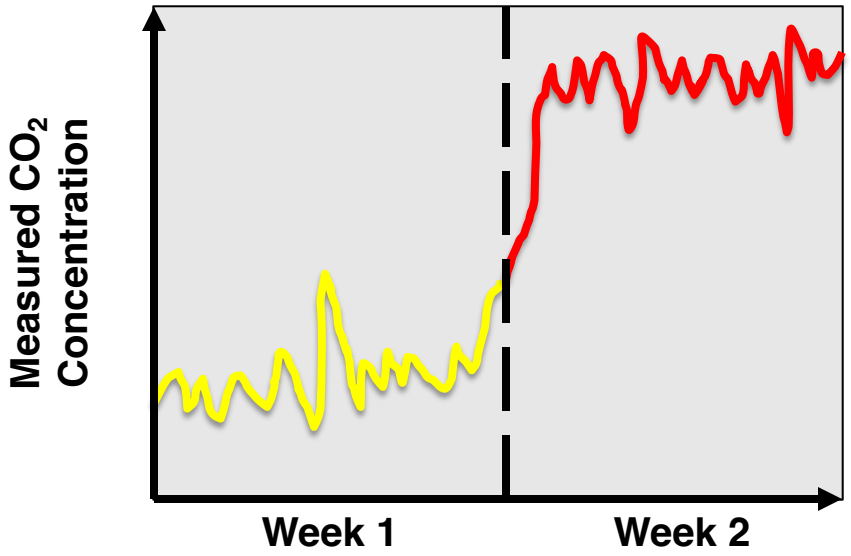
CO_2

Map

Obs

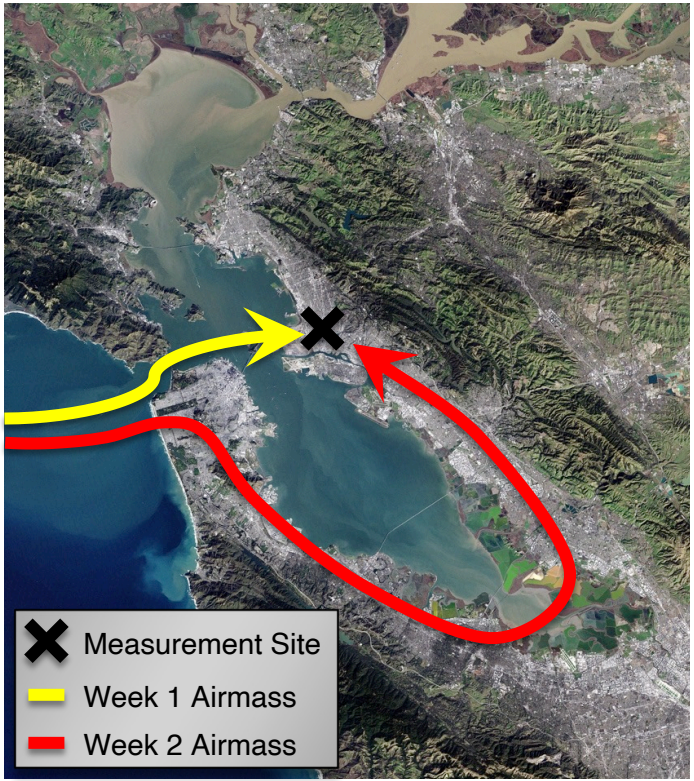
Estimating emissions using atmospheric measurements





Implies there is an urban source

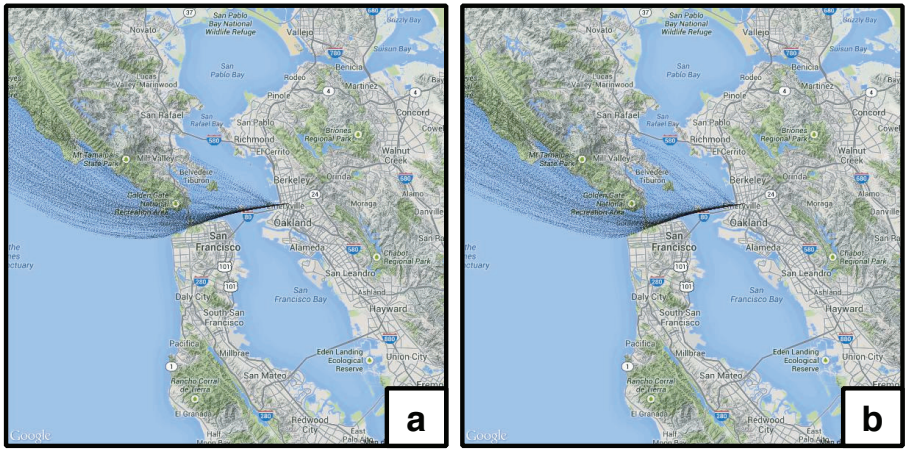
Estimating emissions using atmospheric measurements



$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\epsilon}$$

obs — *transport model* — *surface fluxes* — *error*

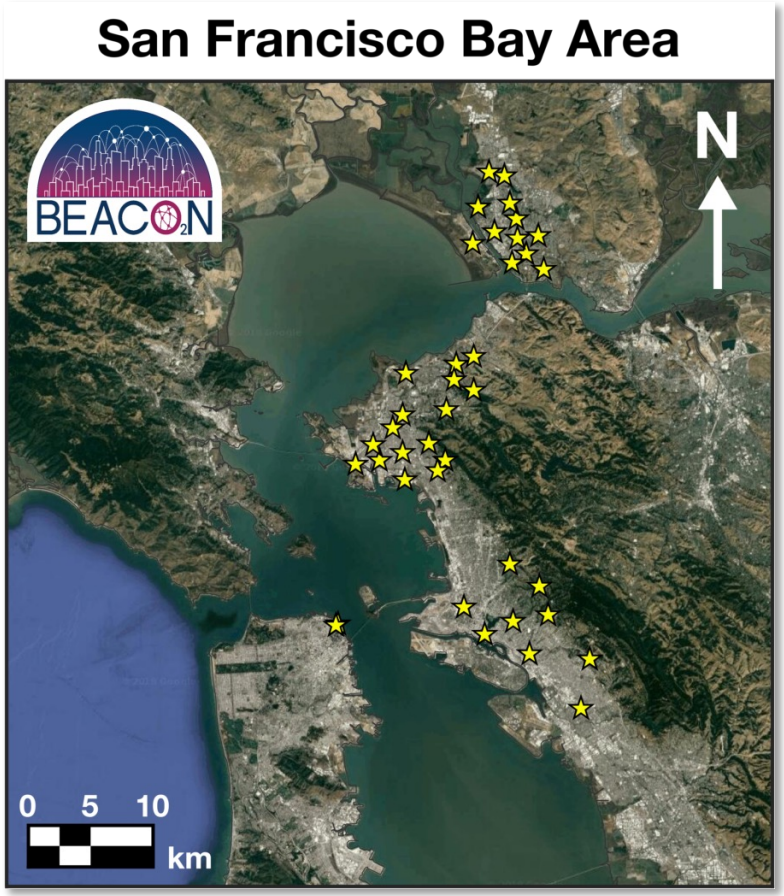
Estimating emissions using atmospheric measurements is an inverse problem



Observed impacts of COVID-19 on urban CO_2 emissions

Turner et al., *GRL* (2021)

Berkeley Environmental Air-Quality & CO_2 Network (BEACO$_2$N)



Established in 2013

- Aim was to “blanket” a city with sensors
- In contrast to most studies using <5 sites

Species observed

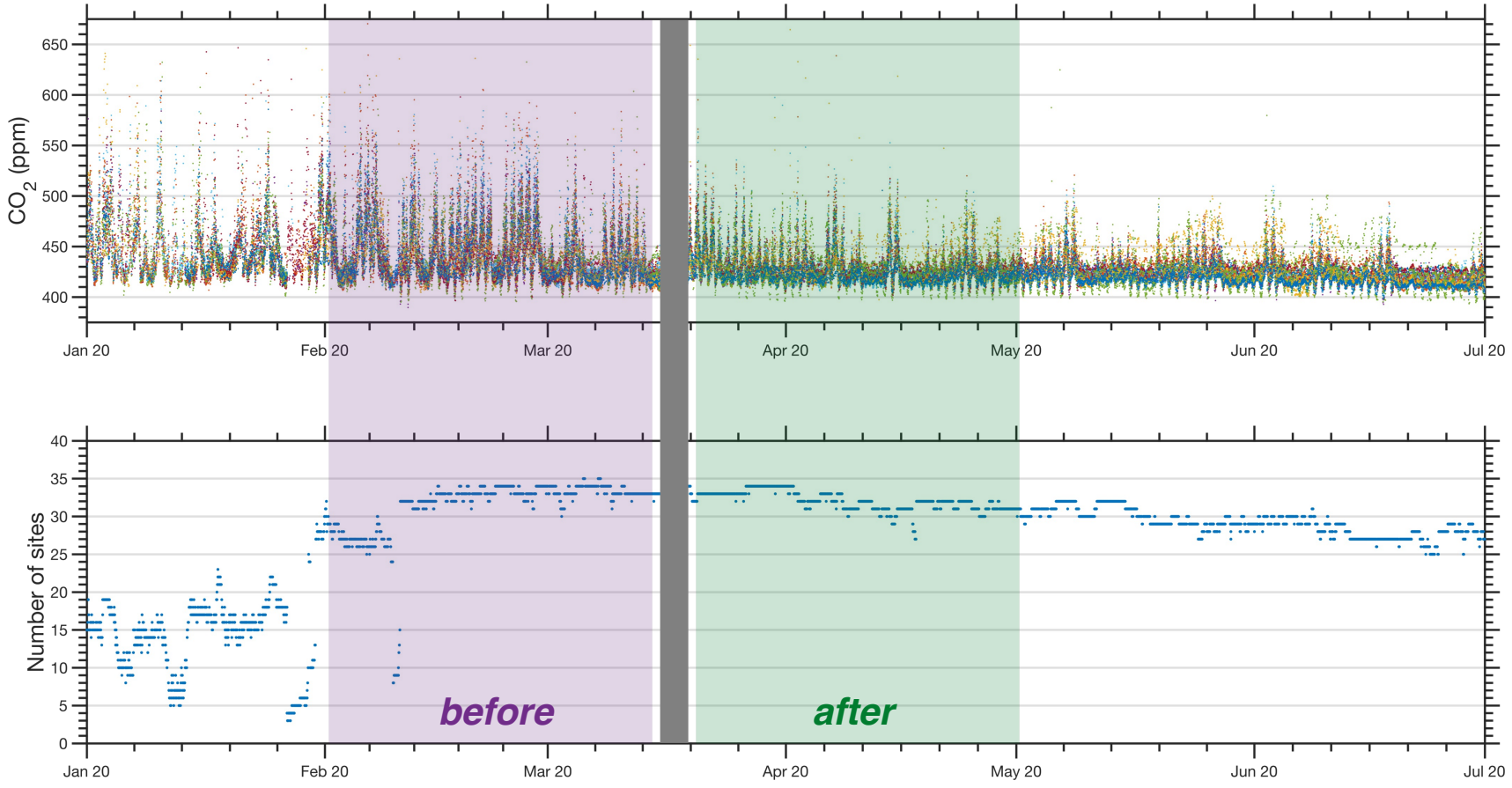
- CO_2 (Vaisala GMP343)
- CO, NO, NO_2, O_3 (AlphaSense)
- PM (Shinyei)

More than 70 sites in the Bay Area

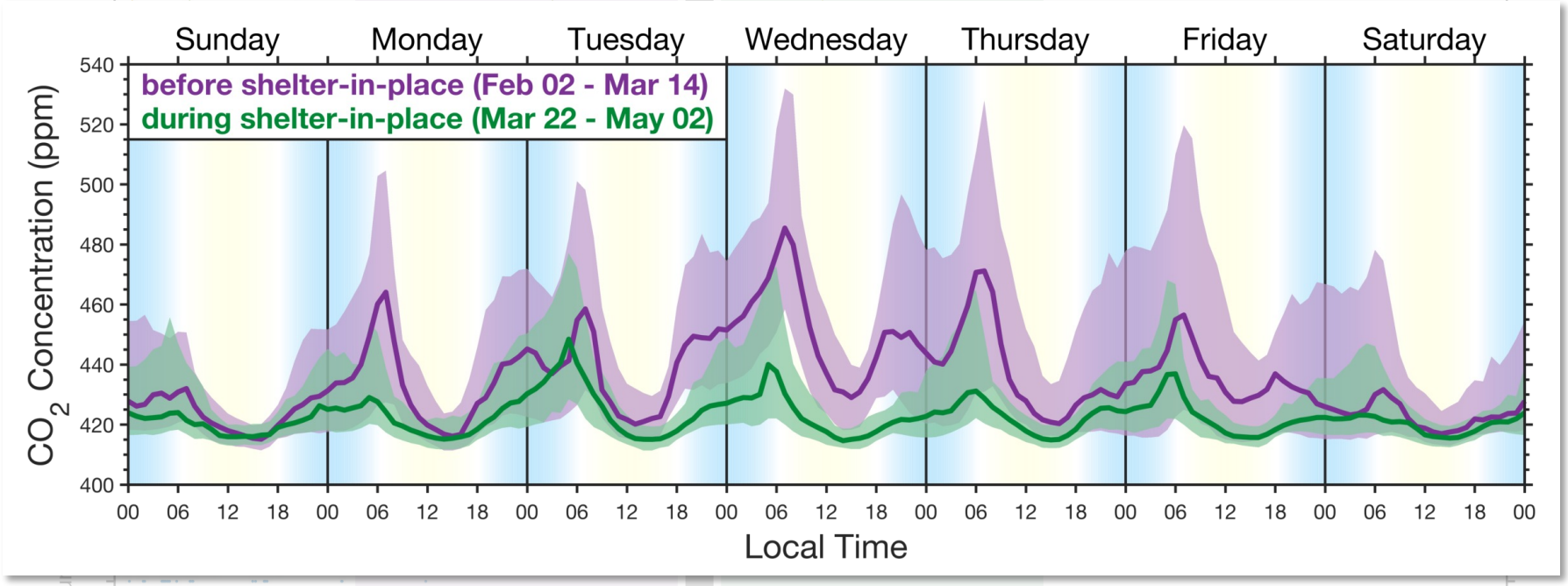
- 35 sites operational during Spring 2020
- Also in Houston, New York, and Glasgow

Shusterman et al. (2016), Turner et al. (2016), Kim et al. (2018), Shusterman et al. (2018)

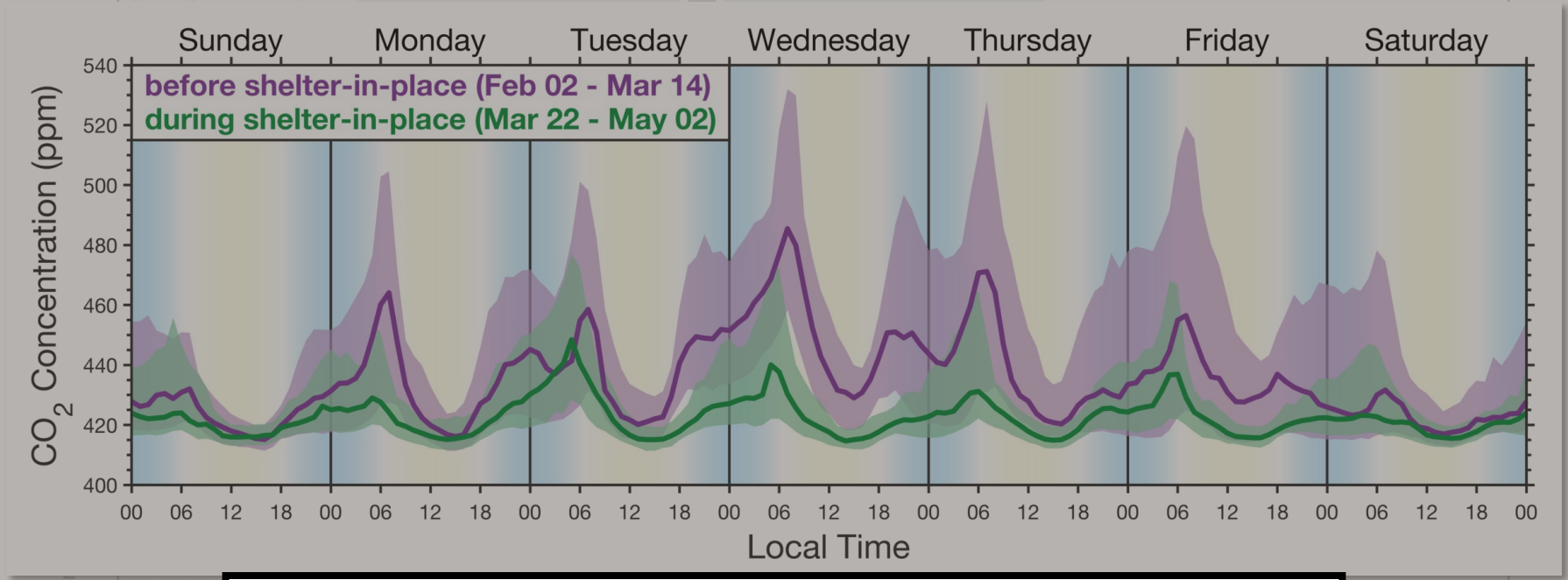
Berkeley Environmental Air-Quality & CO_2 Network (BEACO$_2$N)



Berkeley Environmental Air-Quality & CO_2 Network (BEACO$_2$N)



Berkeley Environmental Air-Quality & CO_2 Network (BEACO$_2$N)

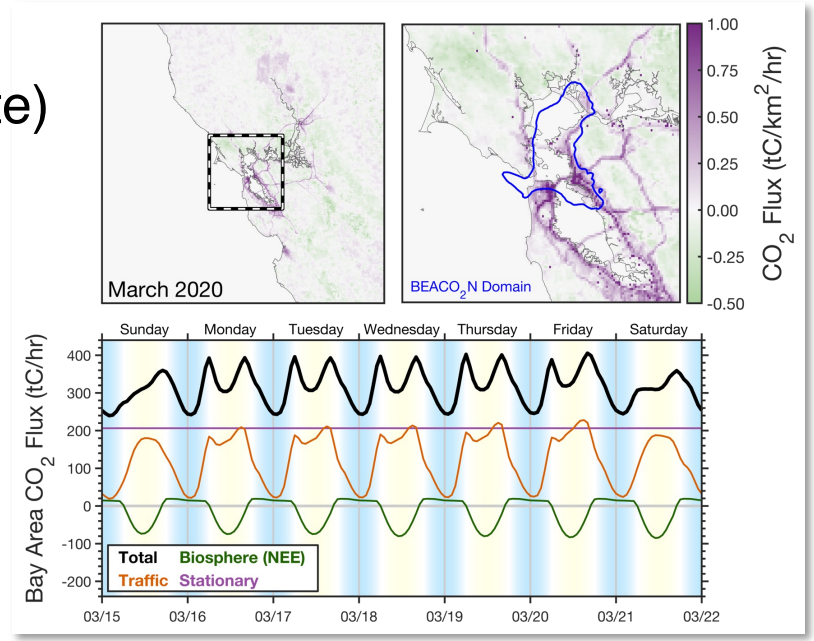


Observe a clear decrease, but how much is attributable to emissions?

Top-down estimate of urban CO_2 emissions from BEACO$_2$N

$$\hat{\mathbf{x}} = \mathbf{x_a} + (\mathbf{HB})^T \left(\mathbf{HBH}^T + \mathbf{R}\right)^{-1} (\mathbf{y} - \mathbf{Hx_a})$$

$\mathbf{x_a}$: prior emissions (bottom-up estimate)



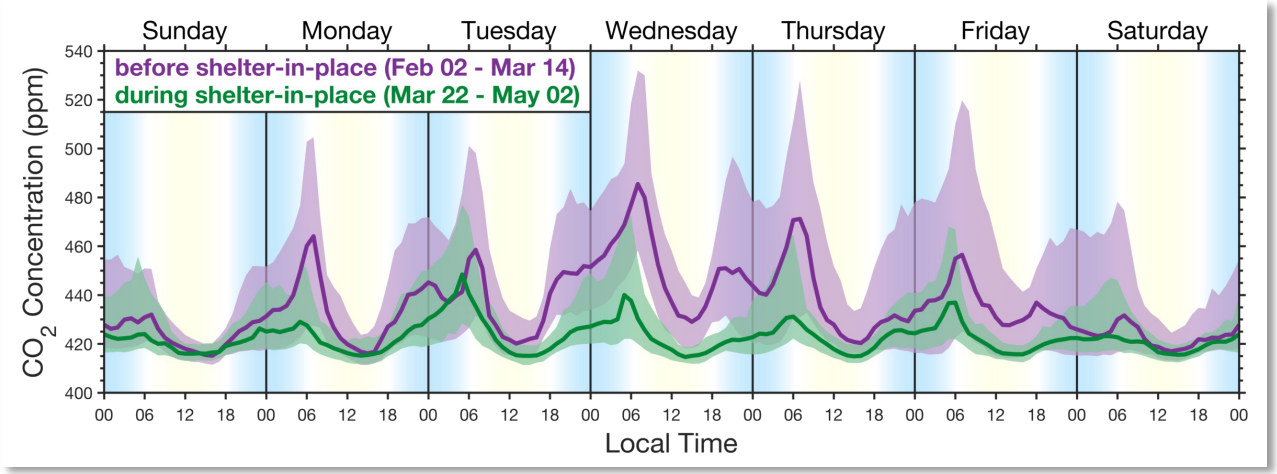
Goal: Estimate hourly CO_2 fluxes at 900-m during COVID-19

Top-down estimate of urban CO_2 emissions from BEACO$_2$N

$$\hat{\mathbf{x}} = \mathbf{x_a} + (\mathbf{HB})^T \left(\mathbf{HBH}^T + \mathbf{R}\right)^{-1} (\mathbf{y} - \mathbf{Hx_a})$$

$\mathbf{x_a}$: prior emissions (bottom-up estimate)

$\mathbf{y}$: observations



Goal: Estimate hourly CO_2 fluxes at 900-m during COVID-19

Top-down estimate of urban CO_2 emissions from BEACO$_2$N

$$\hat{\mathbf{x}} = \mathbf{x_a} + (\mathbf{HB})^T \left(\mathbf{HBH}^T + \mathbf{R} \right)^{-1} (\mathbf{y} - \mathbf{Hx_a})$$

$\mathbf{x_a}$: prior emissions (bottom-up estimate)

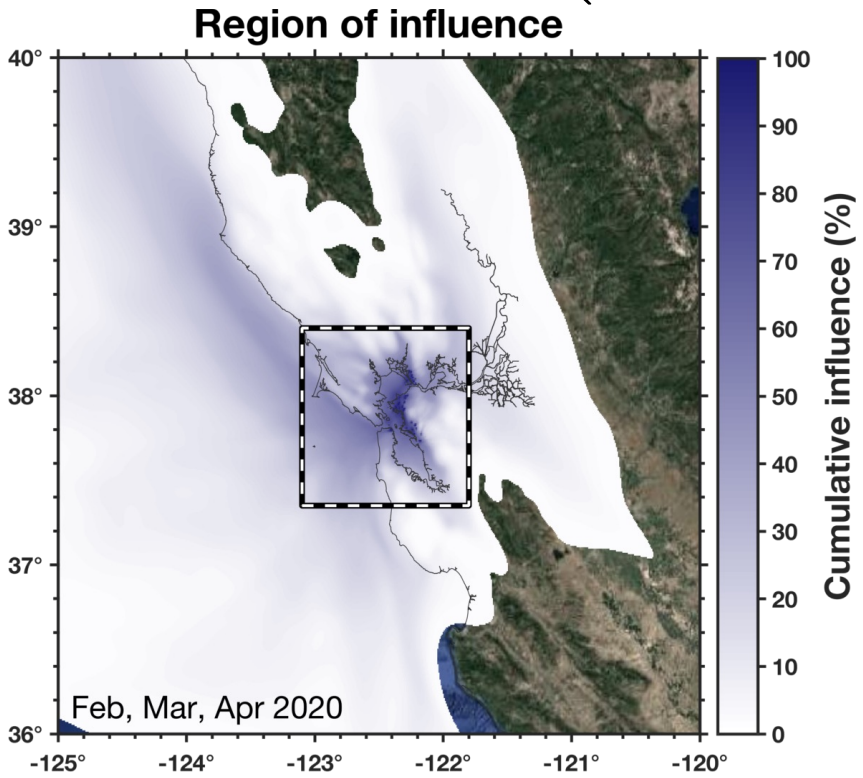
$\mathbf{y}$: observations

$\mathbf{H}$: footprint matrix (represents atmospheric transport)

Goal: Estimate hourly CO_2 fluxes at 900-m during COVID-19

Top-down estimate of urban CO_2 emissions from BEACO$_2$N

$$\hat{\mathbf{x}} = \mathbf{x_a} + (\mathbf{HB})^T \left(\mathbf{HBH}^T + \mathbf{R} \right)^{-1} (\mathbf{y} - \mathbf{Hx_a})$$

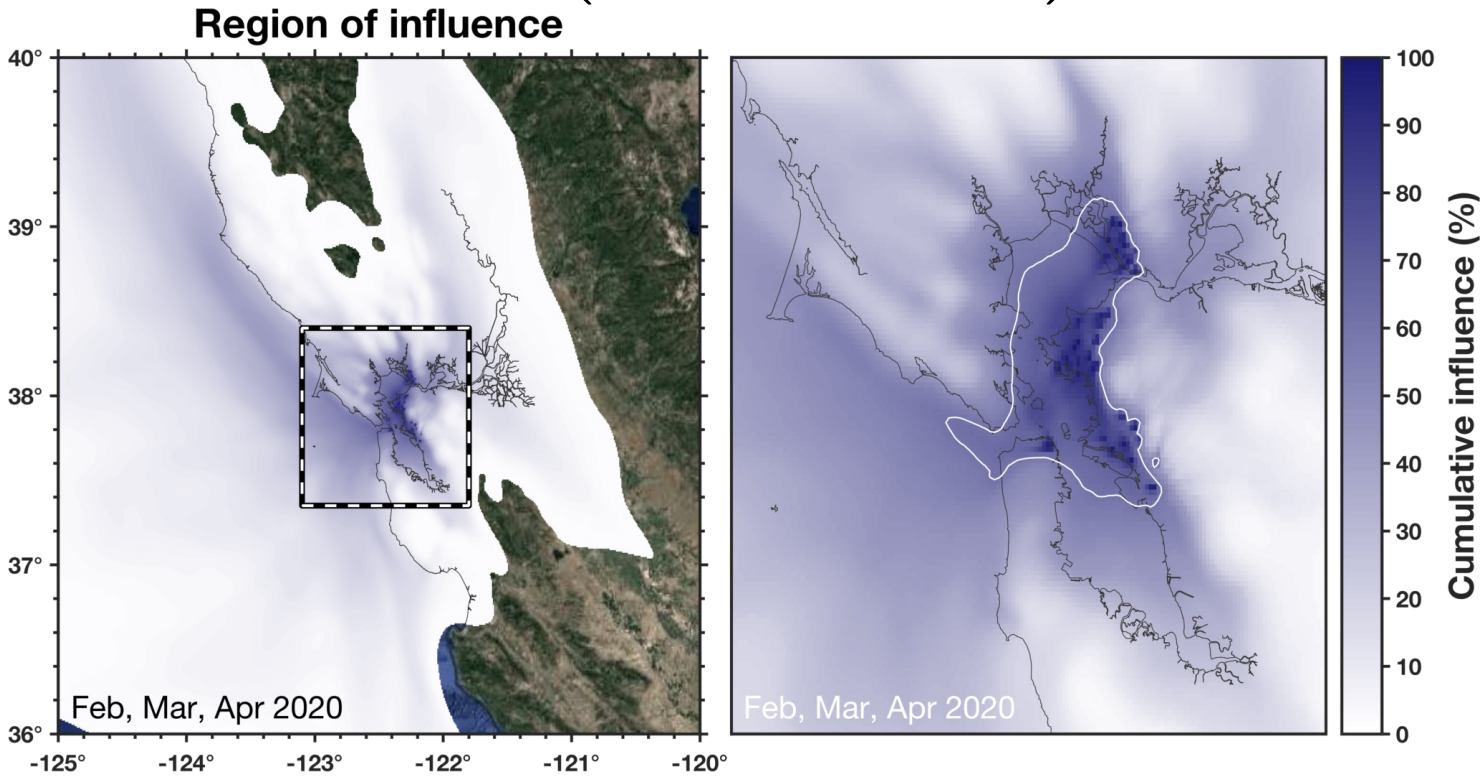




NOAA HRRR meteorology, 72 hour back trajectories with 1000 particles (STILT)

Top-down estimate of urban CO_2 emissions from BEACO$_2$N

$$\hat{\mathbf{x}} = \mathbf{x_a} + (\mathbf{HB})^T \left(\mathbf{HBH}^T + \mathbf{R} \right)^{-1} (\mathbf{y} - \mathbf{Hx_a})$$



NOAA HRRR meteorology, 72 hour back trajectories with 1000 particles (STILT)

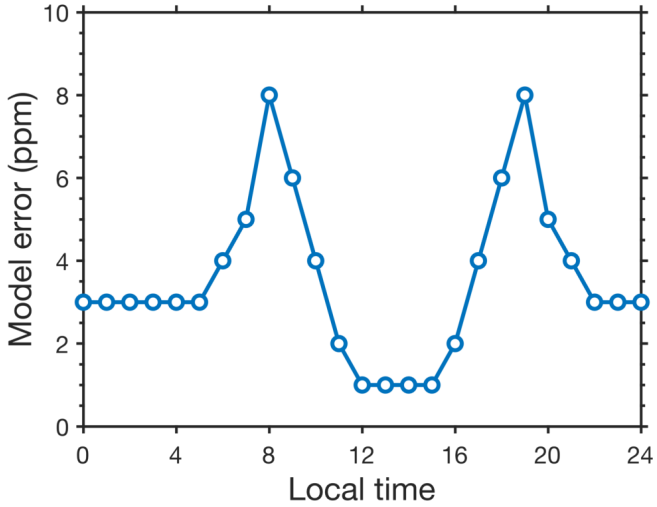
Top-down estimate of urban CO_2 emissions from BEACO$_2$N

$$\hat{\mathbf{x}} = \mathbf{x_a} + (\mathbf{HB})^T \left(\mathbf{HBH}^T + \mathbf{R} \right)^{-1} (\mathbf{y} - \mathbf{Hx_a})$$

$$\mathbf{R} = \mathbf{R}_I + \mathbf{R}_B + \mathbf{R}_M$$

instrument error *background error* *model error*

- Instrument error from std dev of 1-min avg data
  - Lower bound of 1 ppm
- Background error is derived from ensemble of trajectories (~1 ppm)
- Model error varies over the day
- Off-diagonal structure from correlation lengths:
  - τ = 1 km & τ = 2 hours



Top-down estimate of urban CO_2 emissions from BEACO$_2$N

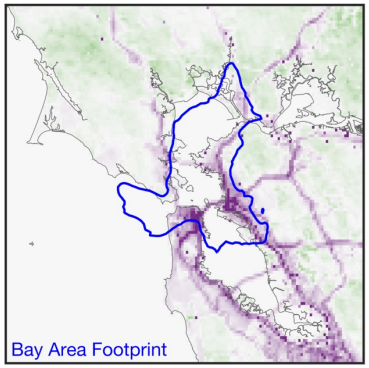
$$\hat{\mathbf{x}} = \mathbf{x_a} + (\mathbf{HB})^T \left(\mathbf{HBH}^T + \mathbf{R} \right)^{-1} (\mathbf{y} - \mathbf{Hx_a})$$

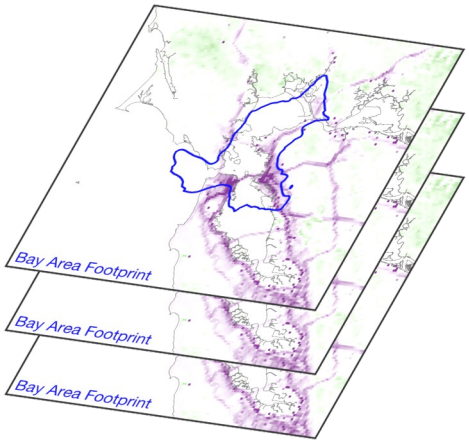
The prior error covariance matrix (**B**) is tricky
We will return to **B** in a couple slides…

Defining the system of equations

Goal: Estimate hourly fluxes at 900-m for 15 days

- **Problem setup**:
  - Horizontal grid is $n_x \times n_y$ where $n_x = 157$ and $n_y = 121$
  - 15 days of hourly fluxes: $n_t = 360$
- Reshape each timestep into a vector ($n_x\, n_y = 18{,}997$ spatial elements)
- "Stack" the timesteps ($m = n_x\, n_y\, n_t = 6{,}338{,}920$ parameters): **55 Mb***!

$\mathbf{x}_1 =$ 

$=$ 
$\mathbf{x} =$ 

$= \begin{pmatrix} \vdots \\ \mathbf{x}_1 \\ \vdots \\ \vdots \\ \begin{pmatrix} \vdots \\ \mathbf{x}_{360} \\ \vdots \end{pmatrix} \end{pmatrix}$

- Prior error covariance matrix is $[m \times m]$ **4.7×10^{13}** elements (**374 Tb***!)

**Assuming double precision: 8 bytes per element*

Prior error covariance matrix (**B**)

- This matrix has **4.7×10^{13}** elements (naïve storage requirement: **374 Tb**)
- Can describe as a Kronecker product ($\otimes$) of two sub-matrices

$$\mathbf{B} = \mathbf{D} \otimes \mathbf{E} = \begin{pmatrix} d_{(1,1)}\mathbf{E} & \cdots & d_{(1,n_T)}\mathbf{E} \\ \vdots & \ddots & \vdots \\ d_{(n_T,1)}\mathbf{E} & \cdots & d_{(n_T,n_T)}\mathbf{E} \end{pmatrix}$$

temporal component

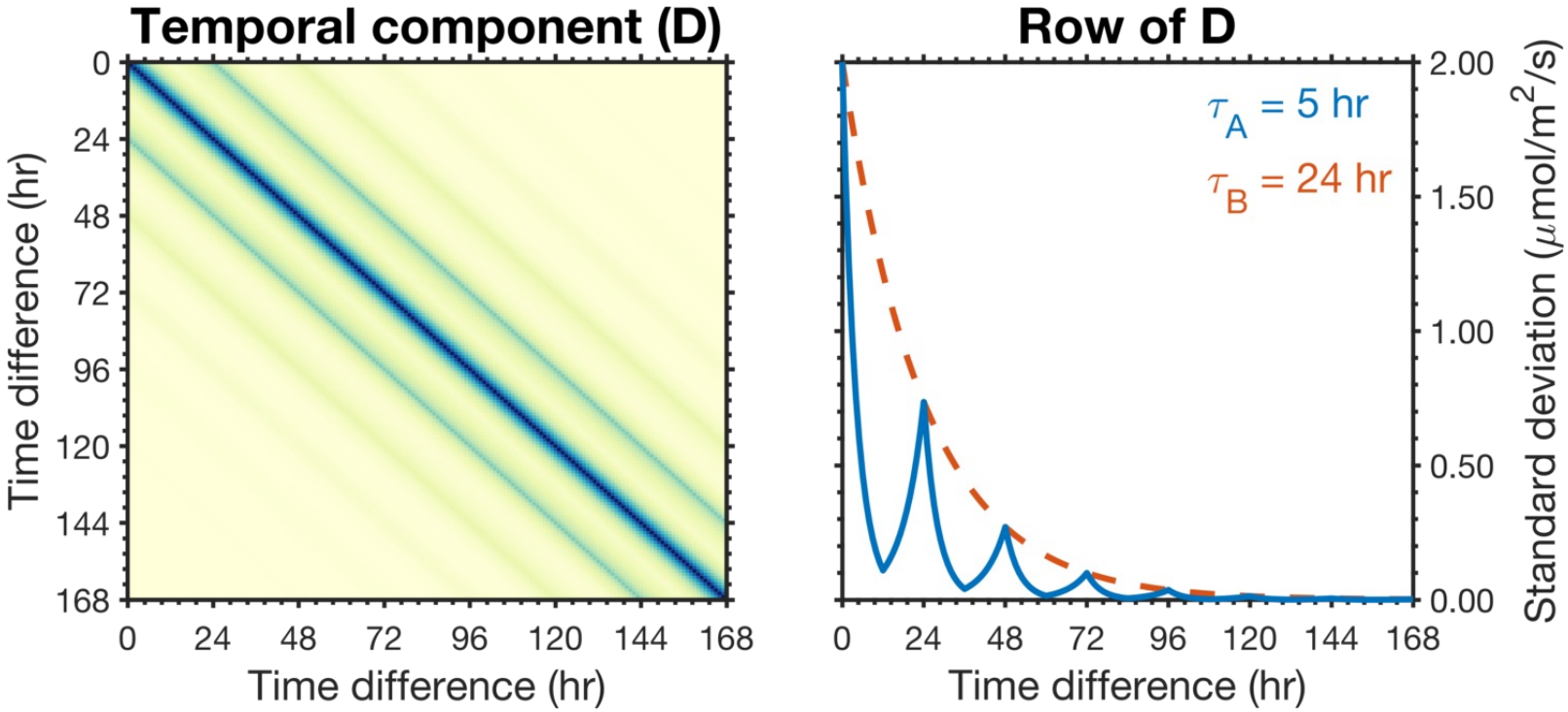
spatial component

Assume the matrix is separable

- never construct the full matrix

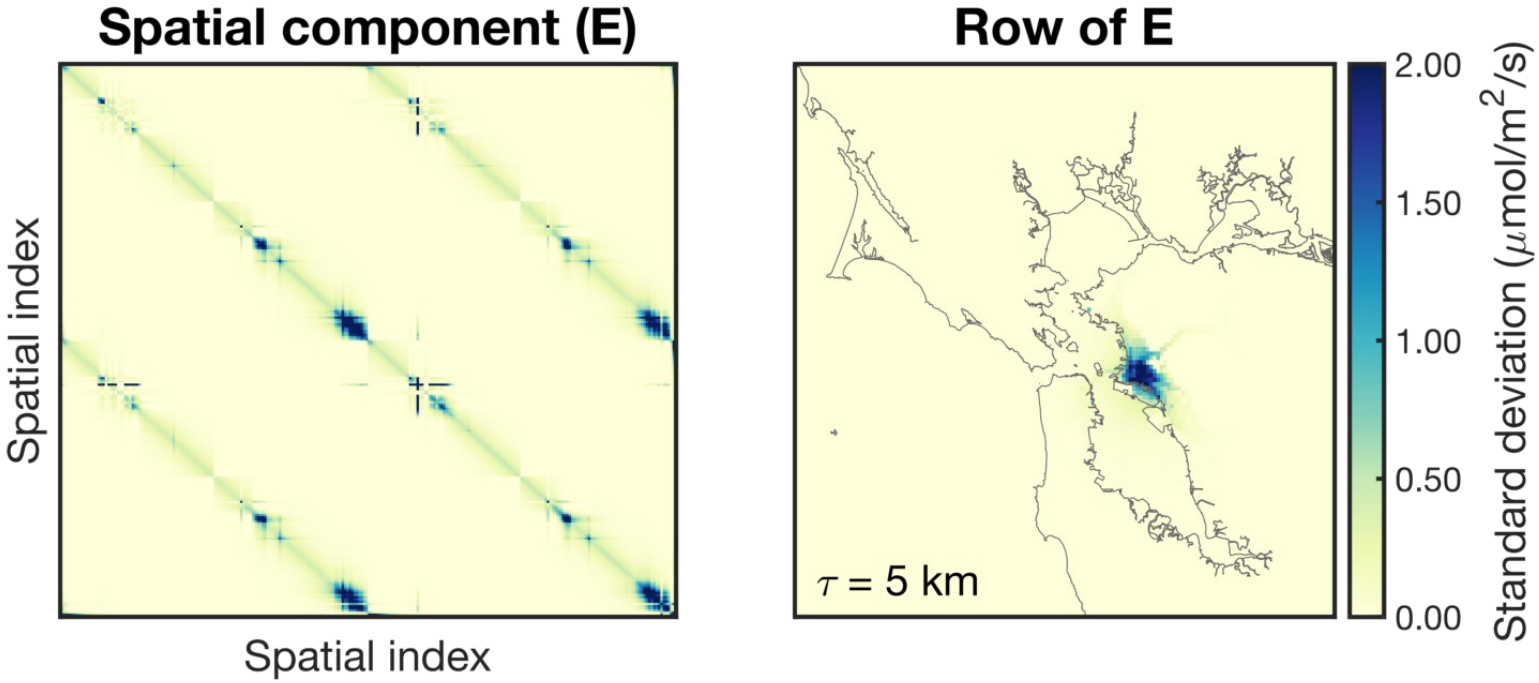
Prior error covariance matrix (**B**)

- This matrix has **4.7×10^{13}** elements (naïve storage requirement: **374 Tb**)
- Can describe as a Kronecker product ($\otimes$) of two sub-matrices



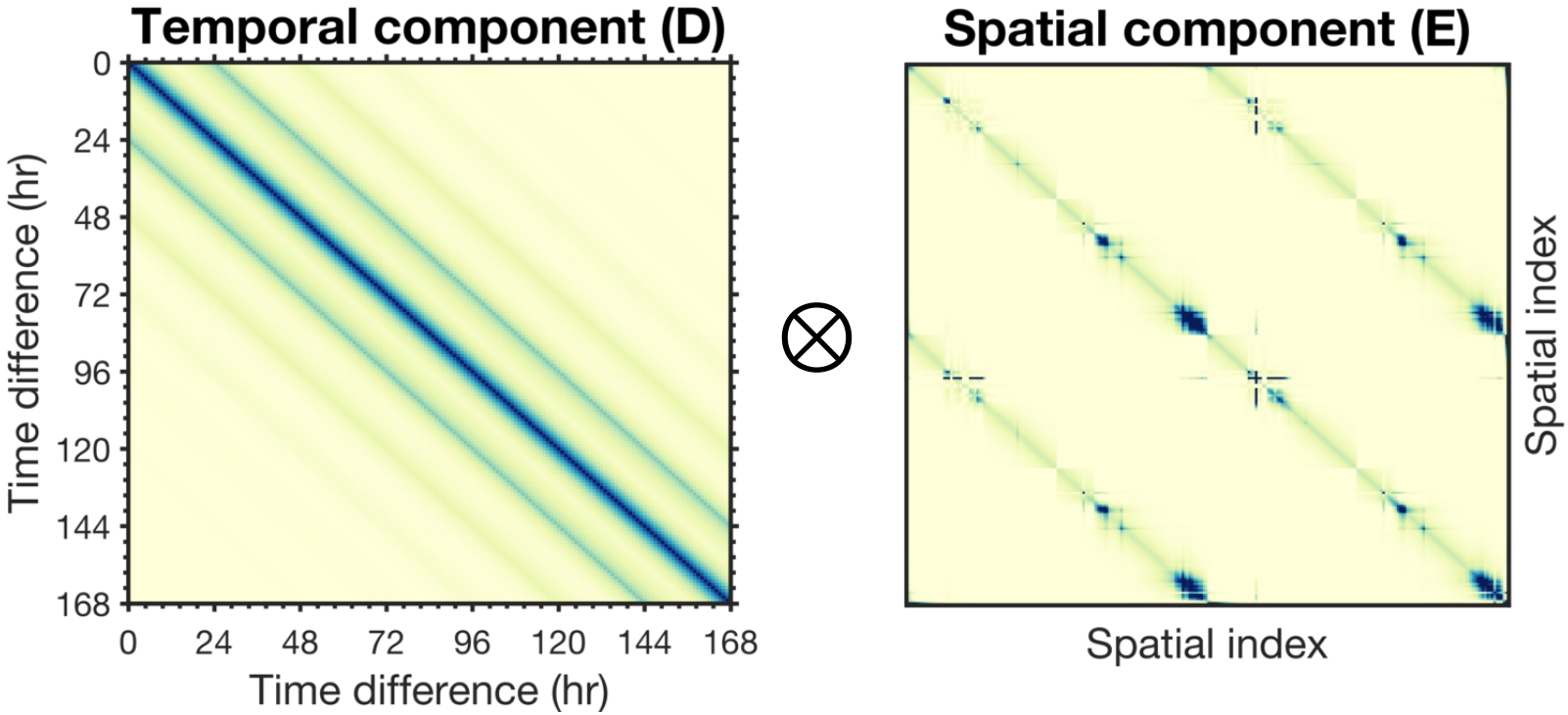
Prior error covariance matrix (**B**)

- This matrix has $\mathbf{4.7 \times 10^{13}}$ elements (naïve storage requirement: **374 Tb**)
- Can describe as a Kronecker product ($\otimes$) of two sub-matrices



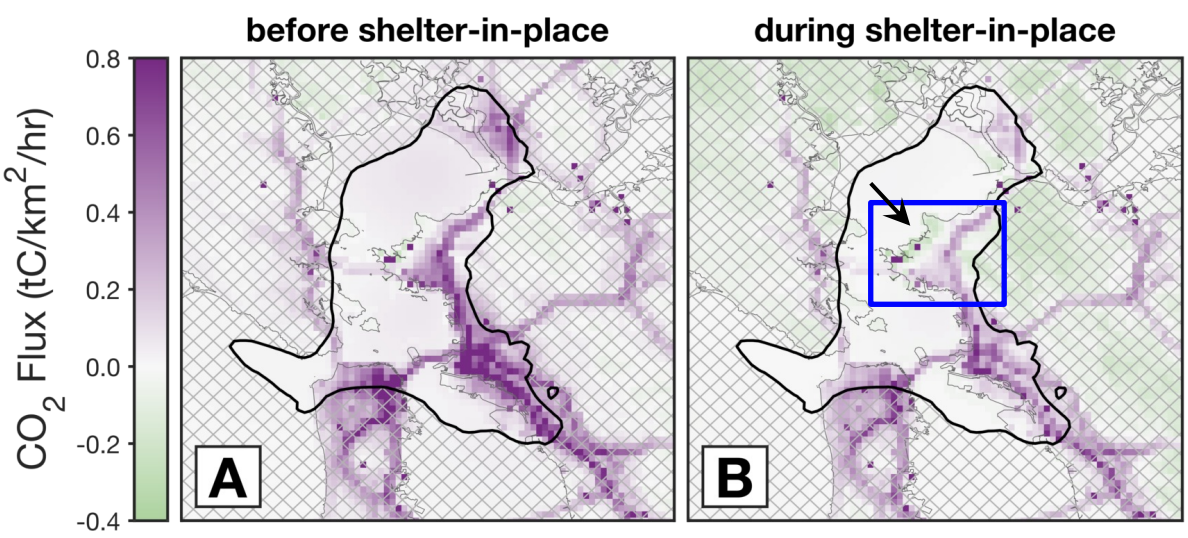
Prior error covariance matrix (**B**)

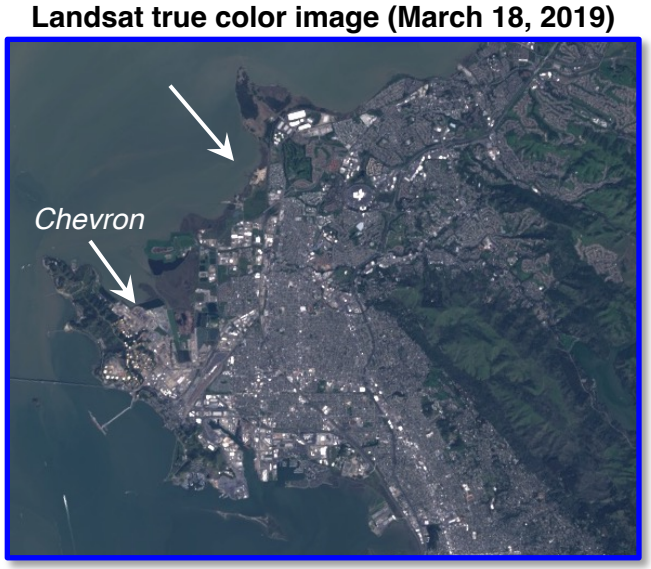
- This matrix has $\mathbf{4.7 \times 10^{13}}$ elements (naïve storage requirement: **374 Tb**)
- Can describe as a Kronecker product ($\otimes$) of two sub-matrices



CO_2 fluxes inferred from atmospheric measurements

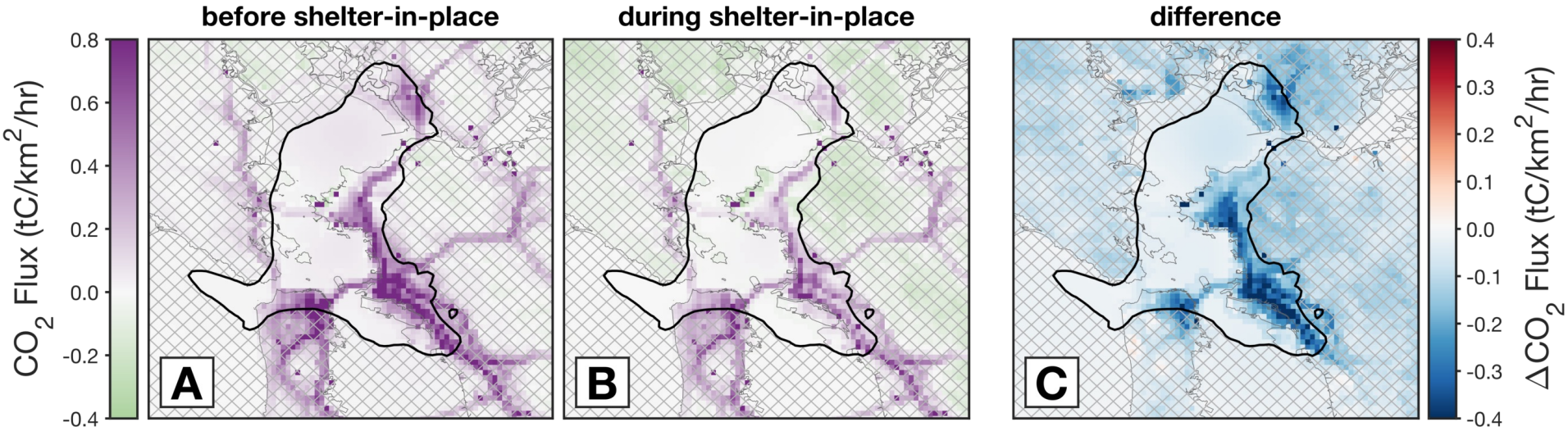
Spatial patterns of the posterior fluxes





- **Hatched area indicates low sensitivity from our network**
  - Sensitive to most of East Bay, part of San Francisco
- **Negligible differences outside hatched domain**
  - Biosphere changes are inferred from TROPOMI SIF

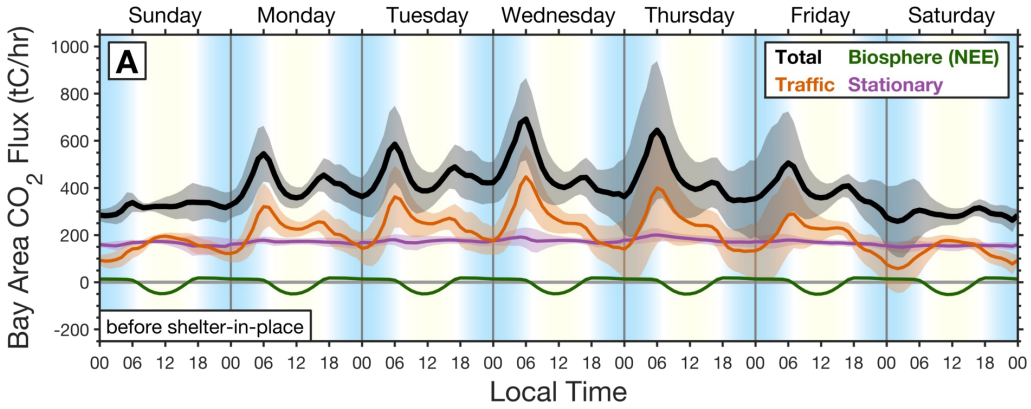
Spatial patterns of the posterior fluxes



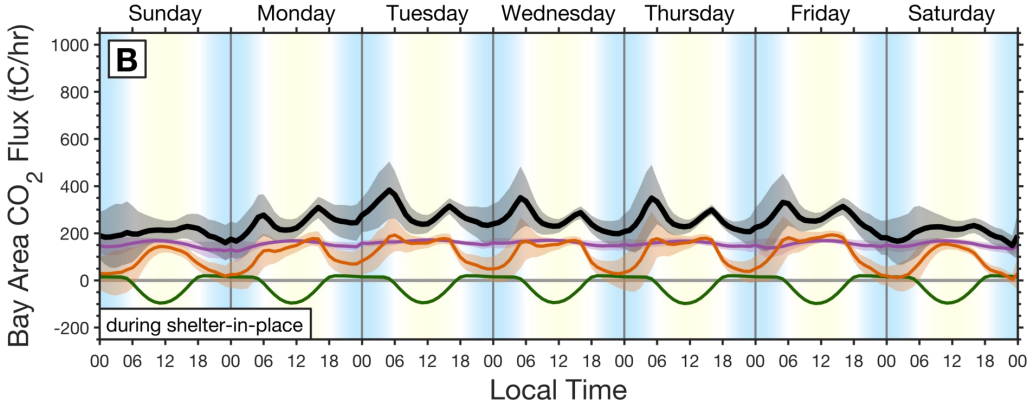
- **Hatched area indicates low sensitivity from our network**
  - Sensitive to most of East Bay, part of San Francisco
- **Negligible differences outside hatched domain**
  - Biosphere changes are inferred from TROPOMI SIF
- **Large changes over urban core and crucial commuter freeways (e.g., I-880)**

Temporal patterns of the posterior fluxes

- **mid-week emission increase**
  - Coincides with morning rush
  - Not driven by the prior



- **30% decrease** in fossil emissions
- **48% decrease** in traffic emissions
- **8% decrease** from other anthro.
- **170% increase** in biosphere
  - Inferred from TROPOMI SIF



Pronounced change to weekly cycle of emissions

Development of a high-fidelity emulator of atmospheric transport for dense observing systems





Nikhil Dadheech (UW)



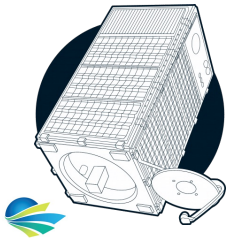
Tai-Long He (UW)

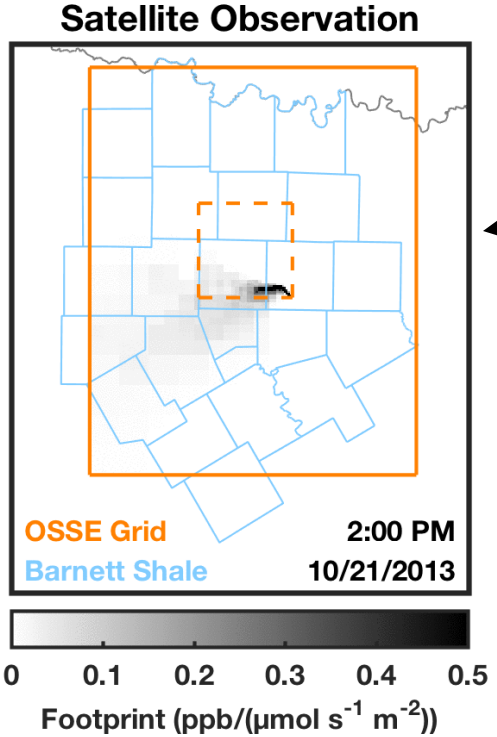
Funded by a NASA Early Career Faculty (ECF) Grant: 80NSSC21K1808

Generating footprints is expensive for dense observing systems

- **Constructed a library of footprints for proposed satellites**
- **Was used in the design of GeoCARB and MethaneSAT**







this is the "**H**" matrix:

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\epsilon}$$

transport model

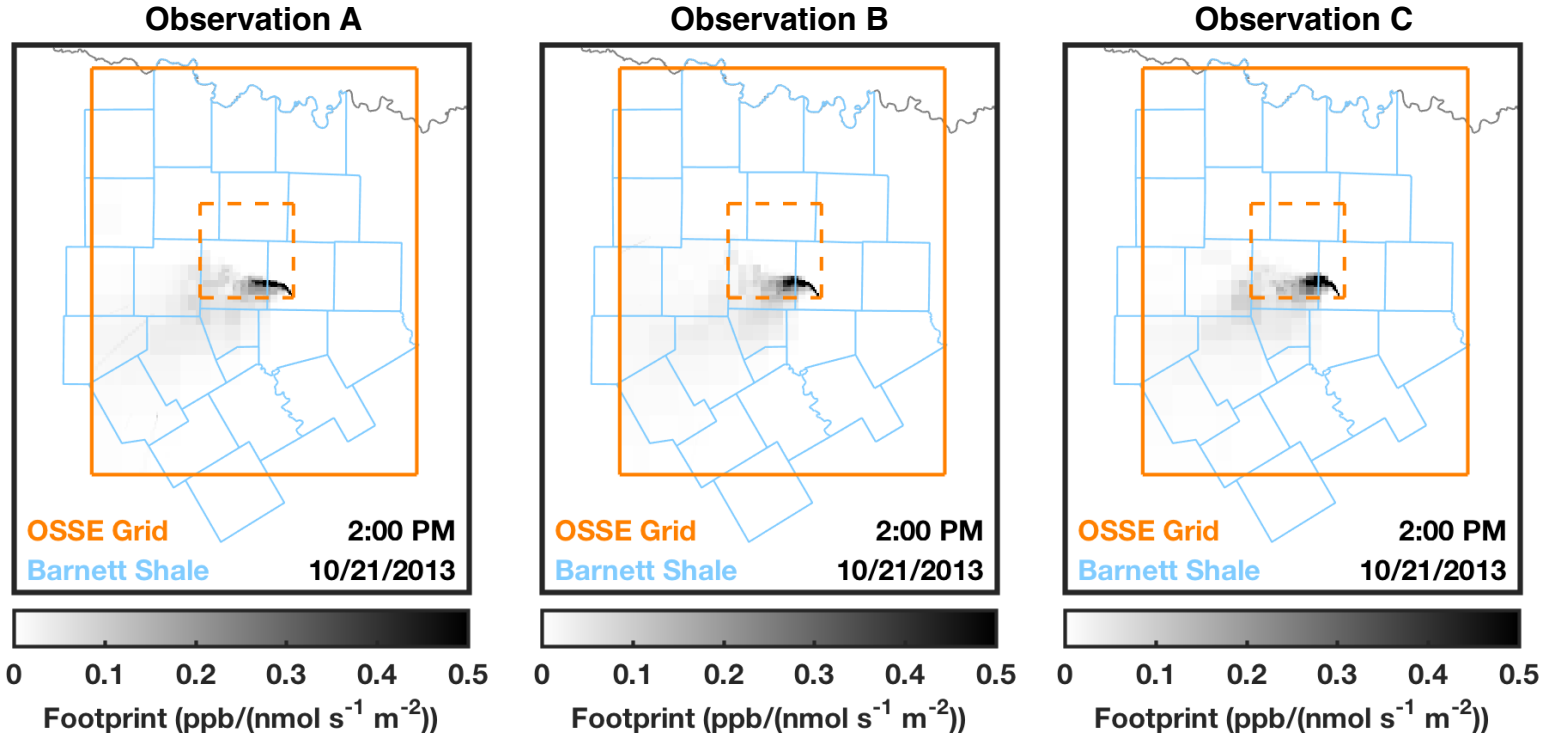
Each observation has a unique mapping to the fluxes

1 week of hourly footprints at 1km resolution are ~4Tb

Took ~6 months to run the 10^8 trajectories

Turner et al., *ACP* (2018)

Strong similarity between the neighboring footprints



Can we develop a reduced order representation for this system to reduce the computational cost?

Developing an efficient method to compute footprints

Computational Approach	Goal
Forward models	Solve for output: $\mathbf{y} = H(x)$
State estimation methods	Solve for input: $y = H(\mathbf{x})$
Model estimation	Solve for model: $y = \mathbf{H}(x)$

Benefits:

1) Reduces computational cost of constructing "**H**"
2) Chemical transport modeling is generally de-coupled from the meteorology, could facilitate joint-estimation of meteorology and emissions
3) Can reduce the memory cost of storing "**H**"

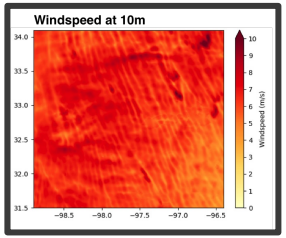
Cons: Induces additional error (should be quantifiable though)

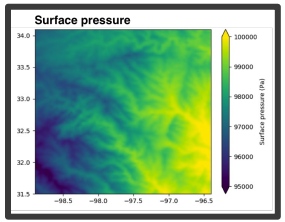
Efficiently computing the footprints

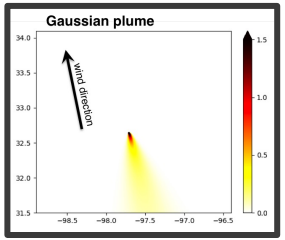
Using a U-net LSTM constructed in PyTorch

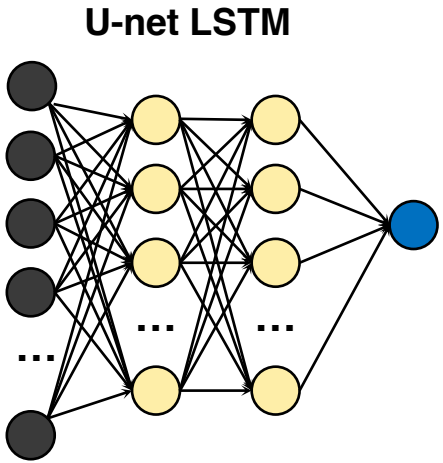
Inputs:

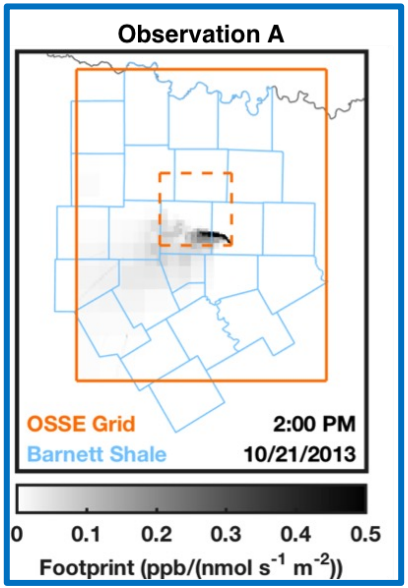
Building blocks for "mental model": winds, topography, …











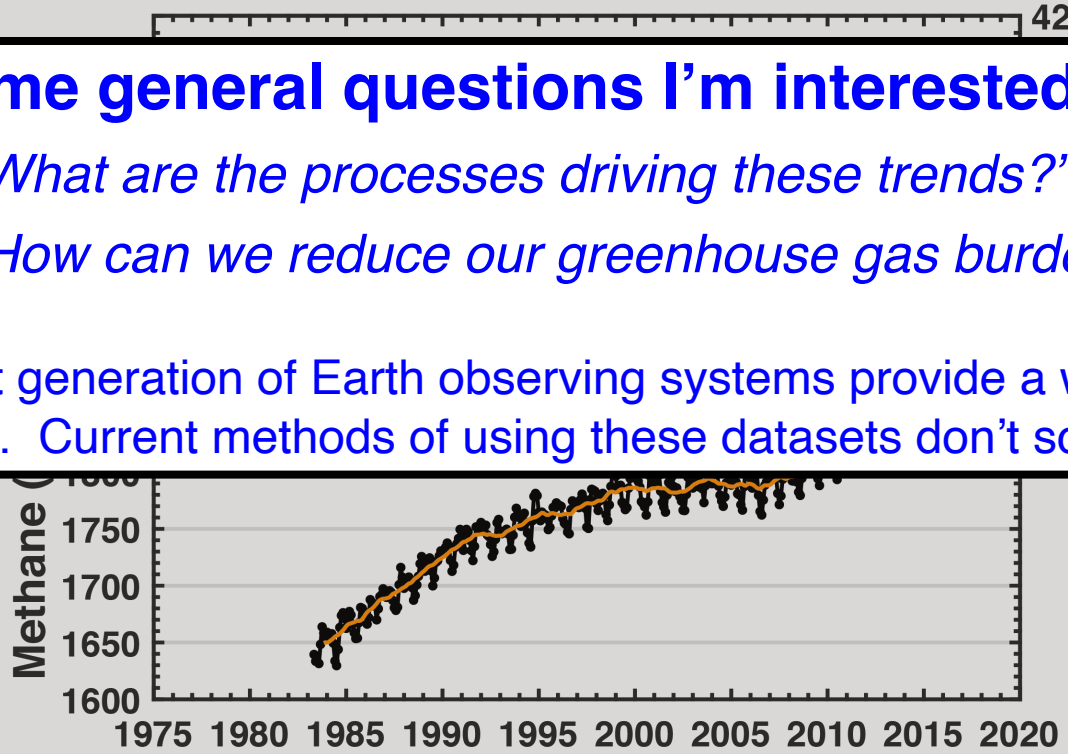
Target to learn:

Surface sensitivity of an atmospheric measurement

Why do we care about the carbon cycle?

1) Methane & CO_2 are the strongest anthropogenic radiative forcers

- **Account for more than 85% of the GHG forcing since 1750**



Some general questions I'm interested in:

- *"What are the processes driving these trends?"*
- *"How can we reduce our greenhouse gas burdens?"*

Next generation of Earth observing systems provide a wealth of data. Current methods of using these datasets don't scale well.

2) Recent trends in atmospheric methane are not well understood