

# A New Harmonized Multi-City Greenhouse Gas Data Product and its Applications

*Logan Mitchell, John C. Lin, Ronald C. Cohen, Russell Dickerson, Ken Davis, Riley M. Duren, Matthias Falk, Gretchen Greene, Abhinav Guha, Anna Karion, Ralph F. Keeling, Jooil Kim, Natasha Miles, Sally Newman, Xinrong Ren, Andrew Rice, Scott Richardson, Maryann Sargent, Jocelyn Turnbull, Kristal R Verhulst, Ray F. Weiss, Steven C. Wofsy, Tamae Wong*

CO2-USA Workshop

October 25, 2018

University of Utah, Salt Lake City, UT



DEPARTMENT OF ATMOSPHERIC SCIENCES | THE UNIVERSITY OF UTAH  
**Land-Atmosphere Interactions Research (LAIR) Group**

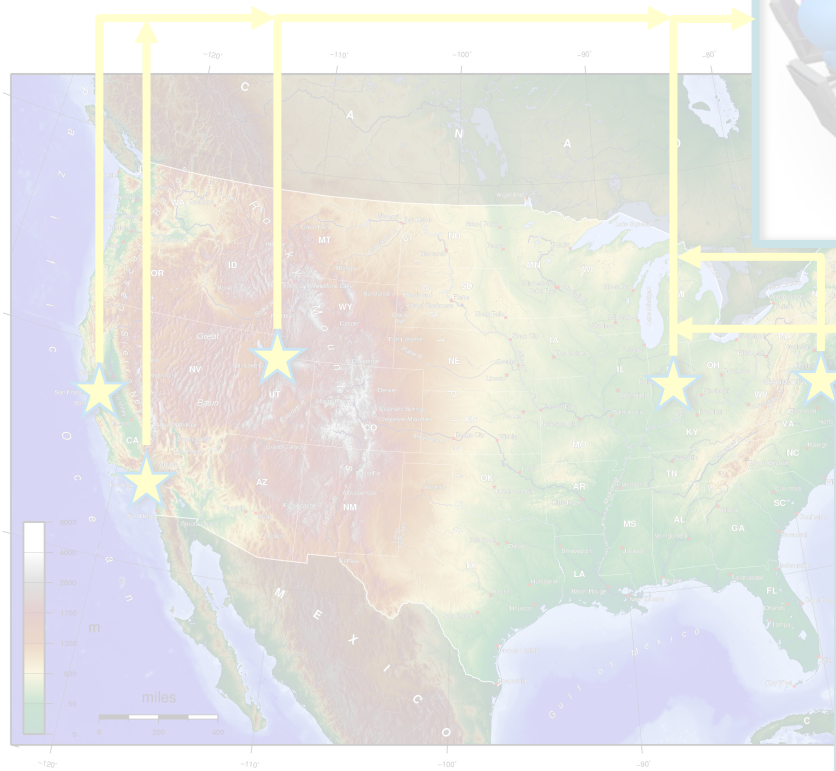
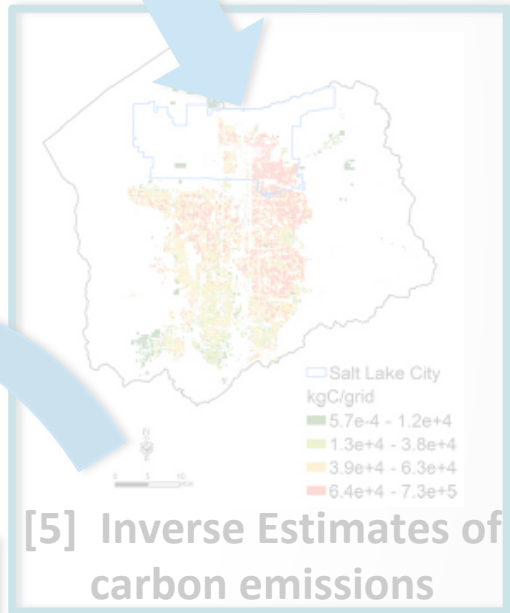
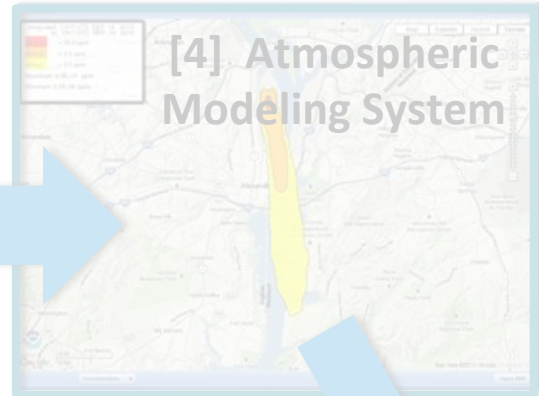
[Logan.Mitchell@Utah.edu](mailto:Logan.Mitchell@Utah.edu)



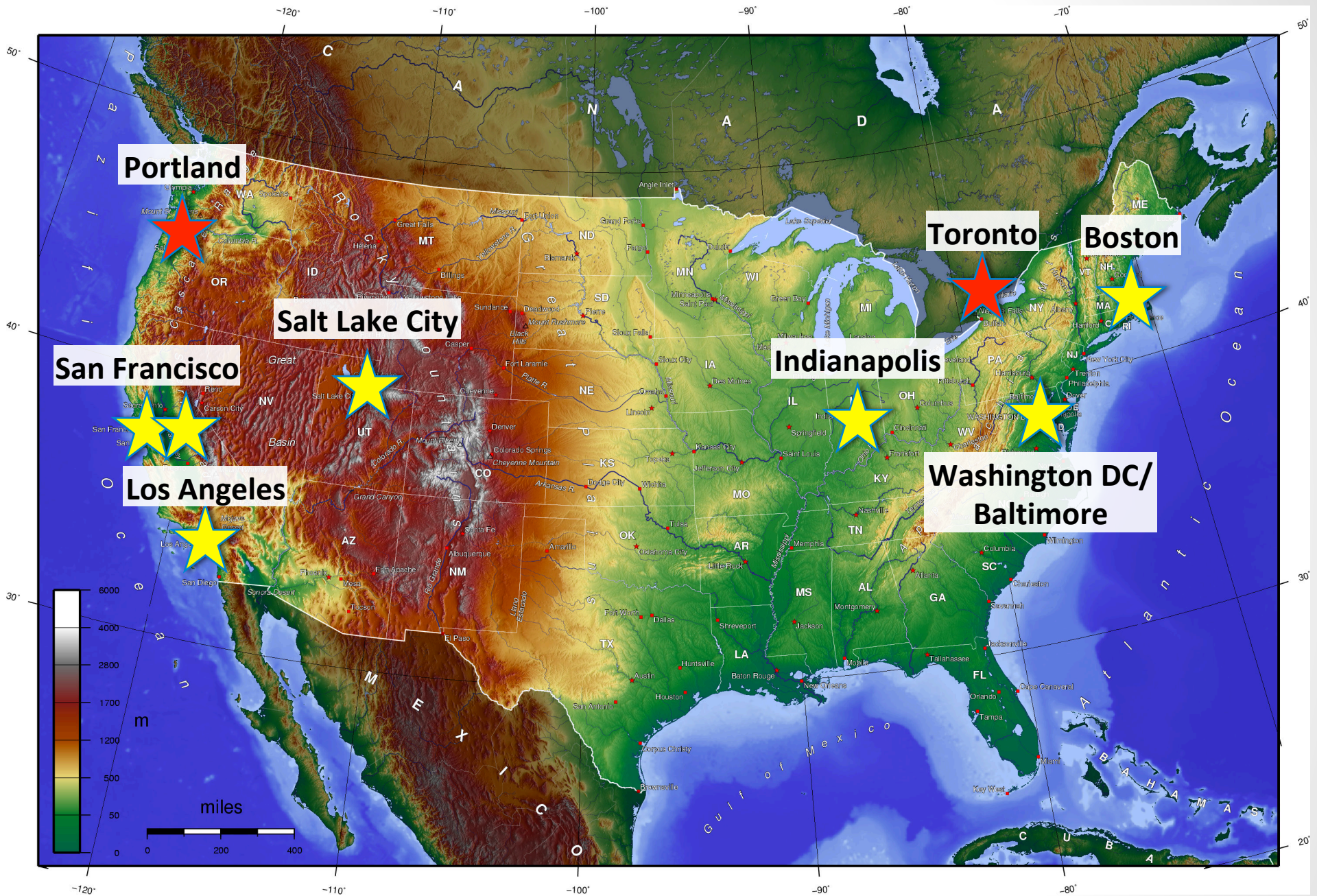
**[2] Harmonized  
Multi-city  
Atmospheric CO<sub>2</sub> &  
CH<sub>4</sub> dataset**

A logo for CO2, with the letters 'CO' in white and '2' in blue, set against a blue background with white cloud-like patterns.

**[3] Biospheric & Anthropogenic  
Inventories**

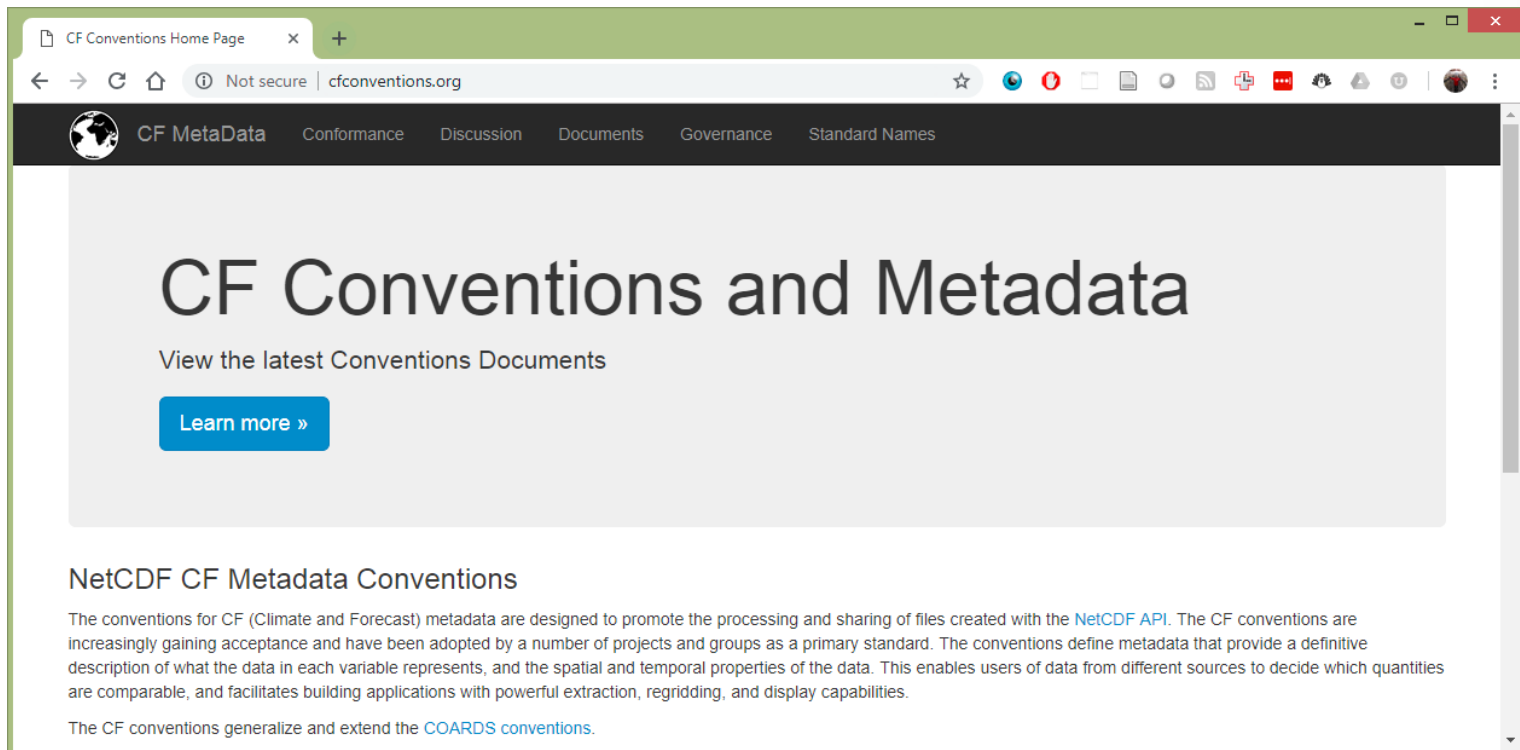
Two images: on the left, a green tree icon representing biospheric inventory; on the right, a photograph of a factory with smokestacks emitting white smoke, representing anthropogenic inventory.





# Harmonized data

- Consistent format for all of the cities to facilitate easier data access.
- Conforms to Climate Forecasting Variable naming conventions to facilitate **human** and **machine readable** data.
- <http://cfconventions.org/>





# Harmonized data

- Consistent format for all of the cities to facilitate easier data access.
  - Conforms to Climate Forecasting Variable naming conventions to facilitate **human** and **machine readable** data.
  - Available in **NetCDF** and **text** data formats.
  - Files contain:
    - Data providers
    - Contact information
    - **Fair use policy**
    - Calibration scales
    - Start/end date & time
    - Site attributes
      - Name
      - Lat/Lon
      - Elevation
      - Inlet height
      - UTC to LST offset
- CO2-USA Data Synthesis Fair Use Data Policy:***
- These cooperative data products are made **freely available** to the public and scientific community **to advance the study of urban carbon cycling and associated air pollutants**.
  - **Fair credit should be given to data contributors** and will depend on the nature of your work. When you start data analysis that may result in a publication, **it is your responsibility to contact the data contributors directly, such that, if it is appropriate, they have the opportunity to contribute substantively and become a co-author.**
  - Data contributors **reserve the right to make corrections to the data based on scientific grounds** (e.g. recalibration or operational issues).
  - Use of the data **implies an agreement to reciprocate by making your research efforts (e.g. measurements as well as model tools, data products, and code) publicly available in a timely manner to the best of your ability.**

# Harmonized data

- Consistent format for all of the cities to facilitate easier data access.
  - Conforms to Climate Forecasting Variable naming conventions to facilitate **human** and **machine readable** data.
  - Available in **NetCDF** and **text** data formats.
  - Files contain:
    - Data providers
    - Contact information
    - Fair use policy
    - Calibration scales
    - Start/end date & time
    - Site attributes
      - Name
      - Lat/Lon
      - Elevation
      - Inlet height
      - UTC to LST offset
1. Time (UTC)
    - POSIX time (seconds since 1970-01-01T00:00:00Z)
  2. Time (UTC)
    - ISO 8601 format (2016-01-01T00:00:00Z)
  3. Hourly averaged mixing ratio
    - CO<sub>2</sub>, CH<sub>4</sub>, or CO
  4. Standard deviation of data within the hour
  5. Number of observations within the hour
  6. Measurement uncertainty (from data provider)
  7. Lat
  8. Lon
  9. Elevation
  10. Inlet height



# NetCDF

# Plain text

Panoply: Sources

File Edit View History Bookmarks Plot Window Help

Create Plot Combine Plot Open Dataset

Datasets Catalogs Bookmarks

Name	Long Name	Type
los_angeles_all_site...	los_angeles_all_...	Local File
background_co2	background_co2	—
CIT_co2_48m	CIT_co2_48m	—
CNP_co2_15m	CNP_co2_15m	—
COM_co2_25m	COM_co2_25m	—
COM_co2_45m	COM_co2_45m	—
FUL_co2_50m	FUL_co2_50m	—
GRA_co2_31m	GRA_co2_31m	—
GRA_co2_51m	GRA_co2_51m	—
IRV_co2_20m	IRV_co2_20m	—
LJO_co2_13m	LJO_co2_13m	—
ONT_co2_25m	ONT_co2_25m	—
ONT_co2_41m	ONT_co2_41m	—
SCI_co2_27m	SCI_co2_27m	—
USC1_co2_50m	USC1_co2_50m	—
USC2_co2_50m	USC2_co2_50m	—
VIC_co2_100m	VIC_co2_100m	—
VIC_co2_50m	VIC_co2_50m	—

Group "CIT\_co2\_48m"

```
dimensions:
  time = 8784;
variables:
  double co2(time=8784);
    :_FillValue = -1.0E34; // double
    :units = "micromol mol-1";
    :standard_name = "mole_fraction_of_carbon_dioxide";
    :long_name = "Average of the co2 mole fraction measured";
    :cell_method = "co2: mean";

  double std_dev(time=8784);
    :_FillValue = -1.0E34; // double
    :units = "micromol mol-1";
    :cell_method = "co2: standard_deviation";
    :long_name = "Standard deviation of the co2 mole fraction measured";

  double n(time=8784);
    :_FillValue = -1.0E34; // double
    :units = "count";
    :long_name = "Number of the co2 mole fraction measurements";

  double lat(time=8784);
    :standard_name = "latitude";
    :axis = "Y";
    :_FillValue = -1.0E34; // double
    :units = "degrees_north";

  double lon(time=8784);
    :standard_name = "longitude";
    :axis = "X";
    :_FillValue = -1.0E34; // double
    :units = "degrees_east";

  double elevation(time=8784);
```

Show: All variables

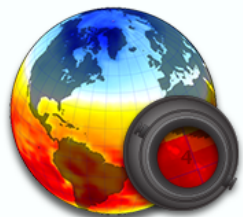
C:\Users\Logan\gcloud.utah.edu\data\co2-usa\synthesis\_output\los\_angeles\txt\_formatted\_files\lo...

File Edit Search View Tools Window Help

los\_angeles\_CIT\_co2\_48m\_1\_hour\_R0\_2017-05-17.txt

```
1 # Number of header lines : 111+
2 #
3 # ----->>>+
4 # GLOBAL ATTRIBUTES+
5 #
6 # title : Hourly averaged atmospheric carbon dioxide (CO2) measurements in Los Angeles.+
7 # summary : Hourly averaged atmospheric CO2 measurements from 16 monitoring sites in Los Angeles.+
8 # keywords : carbon dioxide, methane, carbon monoxide, urban, greenhouse gas+
9 # comment : Observations represent the hourly average CO2 mole fraction with the time stamp
  representing the floored hour. For example: data from 08:00 to 08:59 were averaged and have the
  time stamp of 08:00.+
10 # references : Verhulst, Kristal R., Anna Karion, Jooil Kim, Peter K. Salameh, Ralph F. Keeling,
  Sally Newman, John Miller, et al. "Carbon Dioxide and Methane Measurements from the Los Angeles
  Megacity Carbon Project - Part 1: Calibration, Urban Enhancements, and Uncertainty Estimates."
  Atmospheric Chemistry and Physics 17, no. 13 (July 7, 2017): 8313-41.
  https://doi.org/10.5194/acp-17-8313-2017.+
11 # source : spectroscopy+
12 # date_created : 2017-05-17+
13 # date_issued : 2017-05-17+
14 # fair_use_policy1 : These cooperative data products are made freely available to the public and
  scientific community to advance the study of urban carbon cycling and associated air pollutants.+
15 # fair_use_policy2 : Fair credit should be given to data contributors and will depend on the
  nature of your work. When you start data analysis that may result in a publication, it is your
  responsibility to contact the data contributors directly, such that, if it is appropriate, they
  have the opportunity to contribute substantively and become a co-author.+
16 # fair_use_policy3 : Data contributors reserve the right to make corrections to the data based on
  scientific grounds (e.g. recalibration or operational issues).+
17 # fair_use_policy4 : Use of the data implies an agreement to reciprocate by making your research
  efforts (e.g. measurements as well as model tools, data products, and code) publicly available in
  a timely manner to the best of your ability.+
18 # provider_total_listed : 2+
19 # provider_url : https://megacities.jpl.nasa.gov/+
20 # provider_1_name : Kristal Verhulst+
21 # provider_1_address1 : Jet Propulsion Laboratory M/S 233-300+
22 # provider_1_address2 : 4800 Oak Grove Drive+
23 # provider_1_address3 : Pasadena, CA 91109+
24 # provider_1_country : United States+
25 # provider_1_city : Los Angeles+
26 # provider_1_affiliation : NASA Jet Propulsion Laboratory (JPL)+
27 # provider 1 email : Kristal.R.Verhulst@jpl.nasa.gov+
```

Text Ln 1, Col 31 Western European (Windows)



<https://www.giss.nasa.gov/tools/panoply/>



# GitHub

- A GitHub repository has been created for the data synthesis effort.
- All of the code to process the input data is publicly available
- Will help inform new research teams about the best reporting practices & file formats
  - Lowers the barriers of entry for new cities

The screenshot shows the GitHub repository page for `loganemitchell/co2usa_data_synthesis`. The repository is titled "Urban greenhouse gas network measurements" and is managed by loganemitchell. It has 20 commits, 1 branch, 0 releases, 1 contributor, and is licensed under GPL-3.0. The repository is currently on the master branch. The file list shows various files related to CO2 data synthesis, including LICENSE, README.md, and several .m files for different cities. The README.md file is selected, showing the title "CO2 Urban Synthesis and Analysis (CO2-USA) Data Synthesis" and a description of the project.

loganemitchell / co2usa\_data\_synthesis

Unwatch 3 Star 0 Fork 0

Code Issues 0 Pull requests 0 Projects 0 Wiki Insights Settings

Urban greenhouse gas network measurements

Manage topics

20 commits 1 branch 0 releases 1 contributor GPL-3.0

Branch: master New pull request

Create new file Upload files Find file Clone or download

loganemitchell Google Drive file not needed in this repository Latest commit a4db801 15 days ago

LICENSE	Initial commit	5 months ago
README.md	Update README.md	3 months ago
co2usa_boston_to_netCDF.m	Working dir update	5 months ago
co2usa_create_netCDF.m	Added file management code	20 days ago
co2usa_indianapolis_to_netCDF.m	Update of file format from Tasha on 4/23/2018	20 days ago
co2usa_load_netCDF.m	Loads all of the city data	20 days ago
co2usa_los_angeles_to_netCDF.m	Working dir updates	5 months ago
co2usa_netCDF2txt.m	Added status messages to the script	20 days ago
co2usa_portland_to_netCDF.m	Initial code for Portland OR data	20 days ago
co2usa_san_francisco_baaqmd_to_netCDF.m	Updated a time zone definition	20 days ago
co2usa_san_francisco_beacon_to_netCDF.m	Initial coding	5 months ago

README.md

## CO2 Urban Synthesis and Analysis (CO2-USA) Data Synthesis

Urban greenhouse gas network measurements for cities in the CO2 Urban Synthesis and Analysis network.

More information about the CO2 Urban Synthesis and Analysis project can be found on our main project web page:

<http://sites.bu.edu/co2usa/>

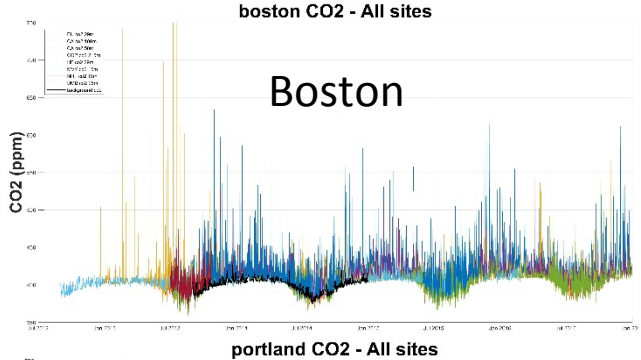
This Git-Hub page contains the code that is being used to generate a harmonized CO2 and CH4 mixing ratio dataset that is readily useable, traceable, and accessible by the research community and the public.

The preliminary synthesis data product can be viewed at:

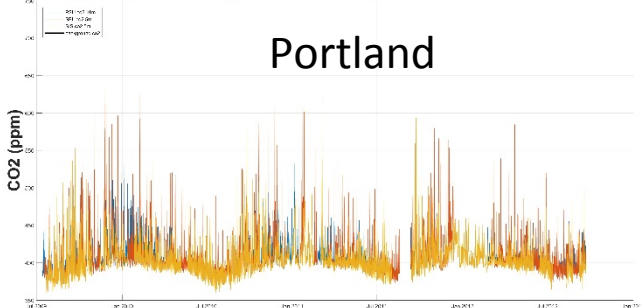
[http://home.chnc.utah.edu/~u0932260/co2-usa/synthesis\\_output/](http://home.chnc.utah.edu/~u0932260/co2-usa/synthesis_output/)



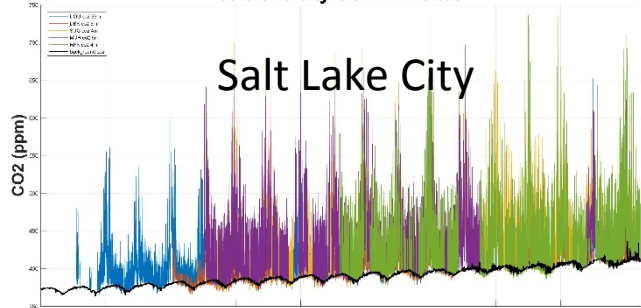
boston CO2 - All sites



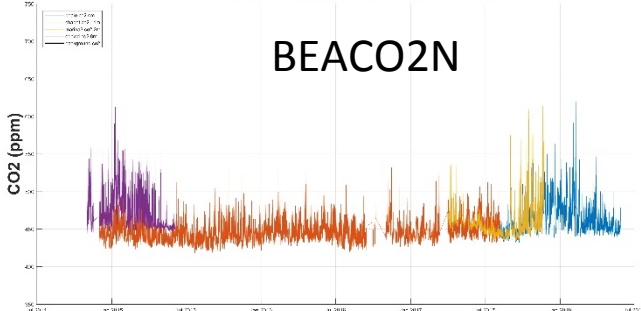
Portland



salt lake city CO2 - All sites



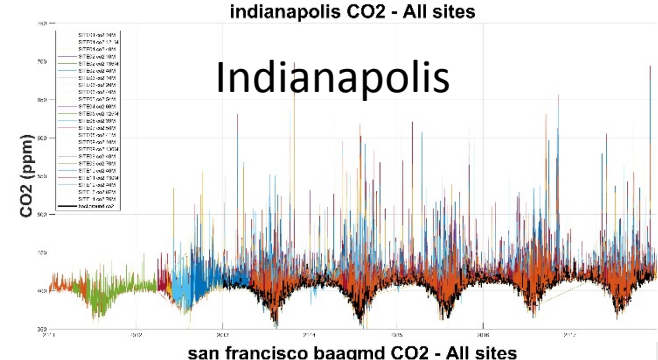
san francisco beacon CO2 - All sites



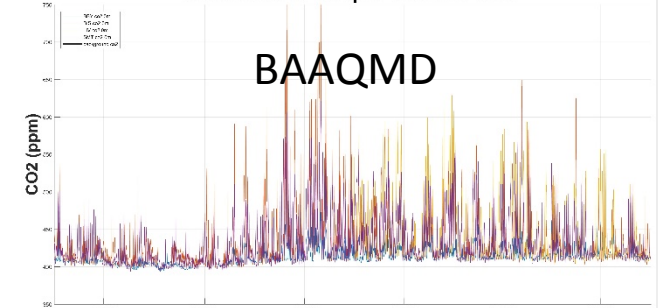
*There is a lot of data!*

- What does a (qualitative) comparison of the data from the cities look like?

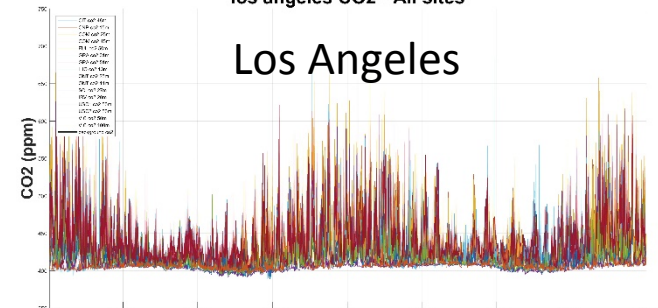
indianapolis CO2 - All sites



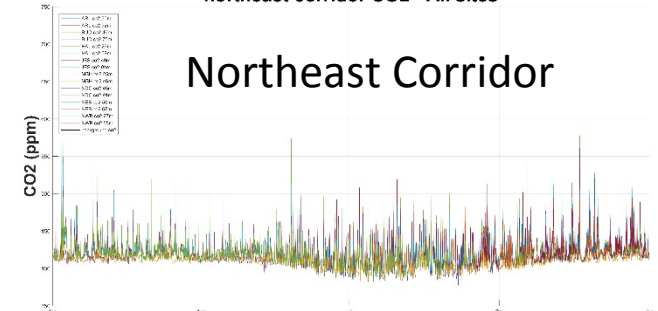
BAAQMD



los angeles CO2 - All sites

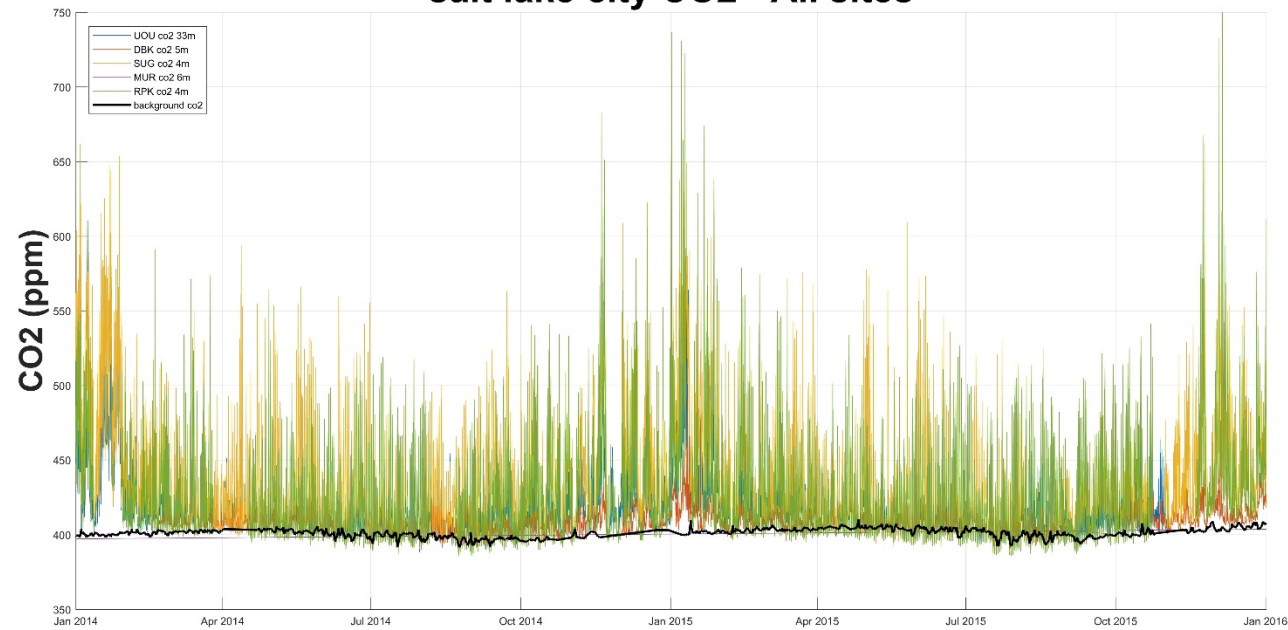


northeast corridor CO2 - All sites

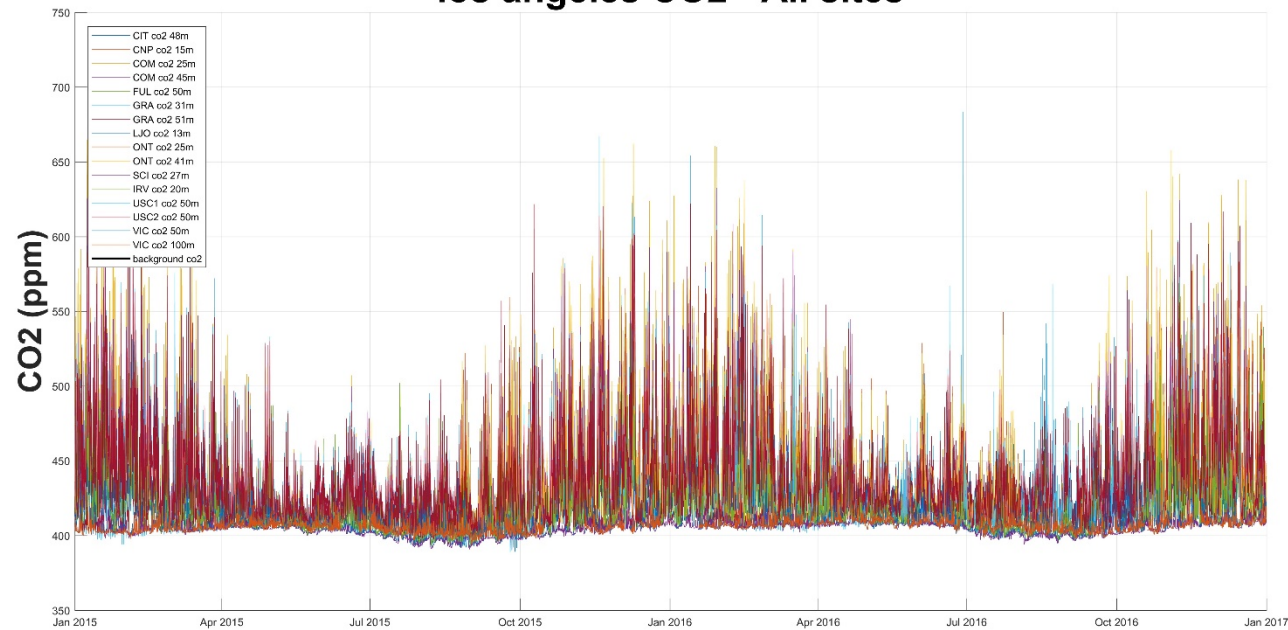


- Looking at a 2-year time period:
- LA has a similar pattern of higher CO<sub>2</sub> in the winter.
- Inlet height is also a factor:
  - SLC's inlets are mostly <10m
  - LA's inlets are all >15m

**salt lake city CO<sub>2</sub> - All sites**



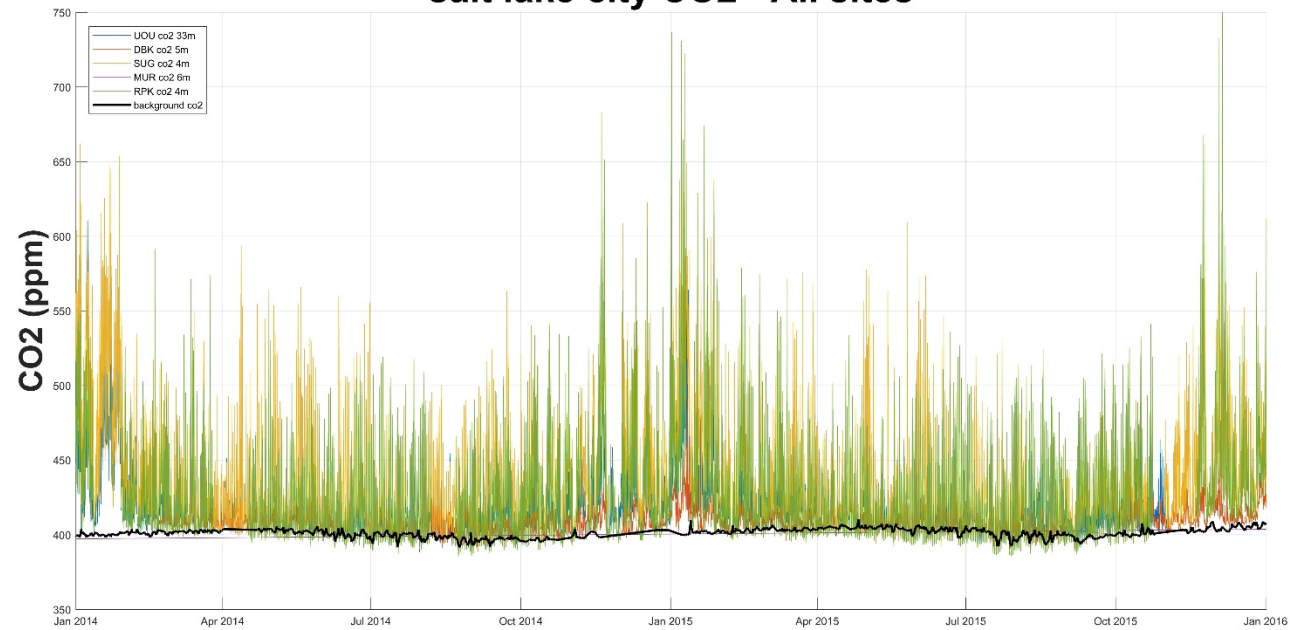
**los angeles CO<sub>2</sub> - All sites**



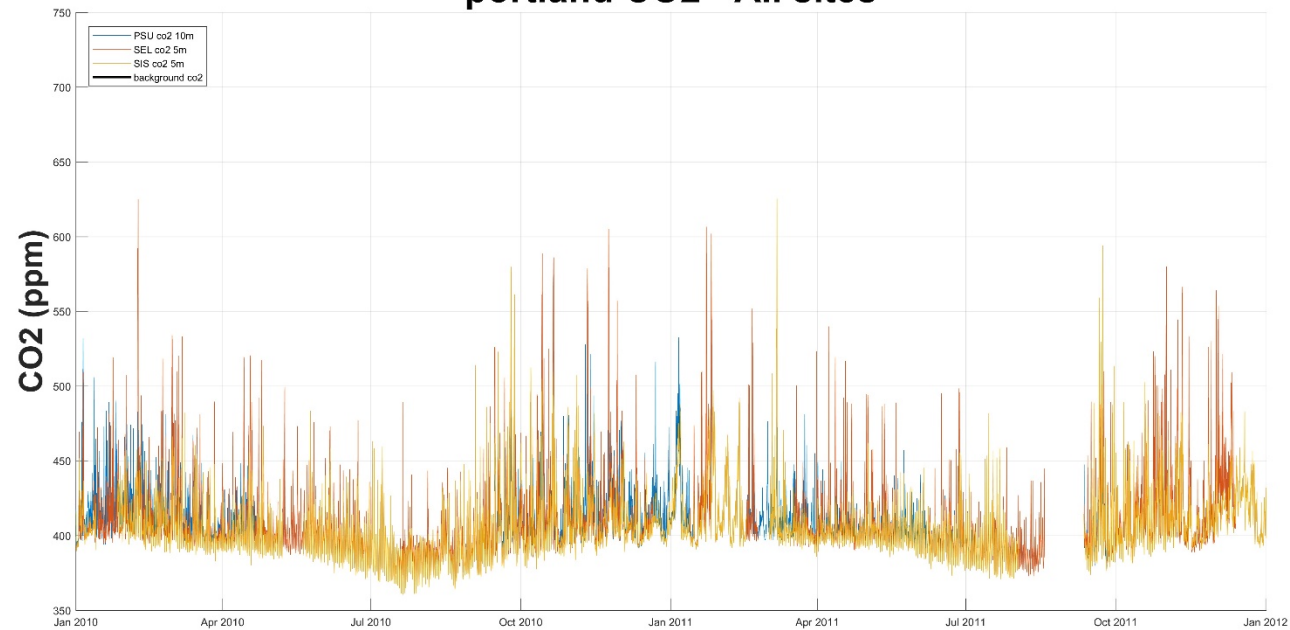


- Looking at a 2-year time period:
- Portland does not have as high of winter time peaks.
- Larger depletion in the summertime.

**salt lake city CO2 - All sites**

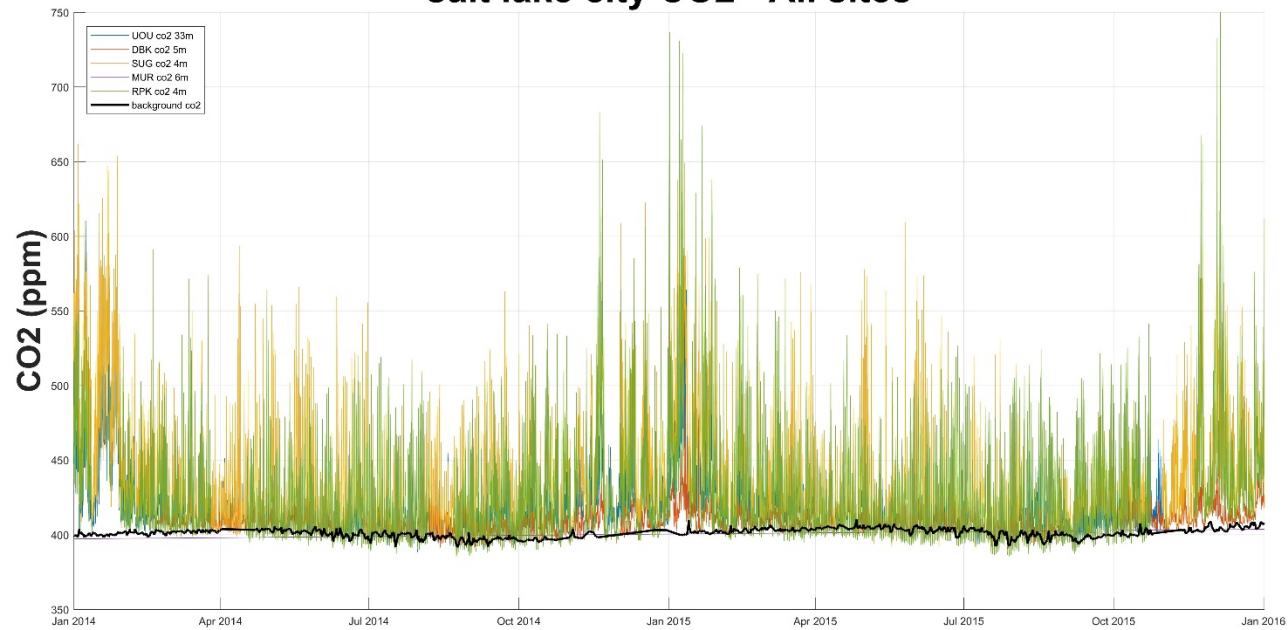


**portland CO2 - All sites**

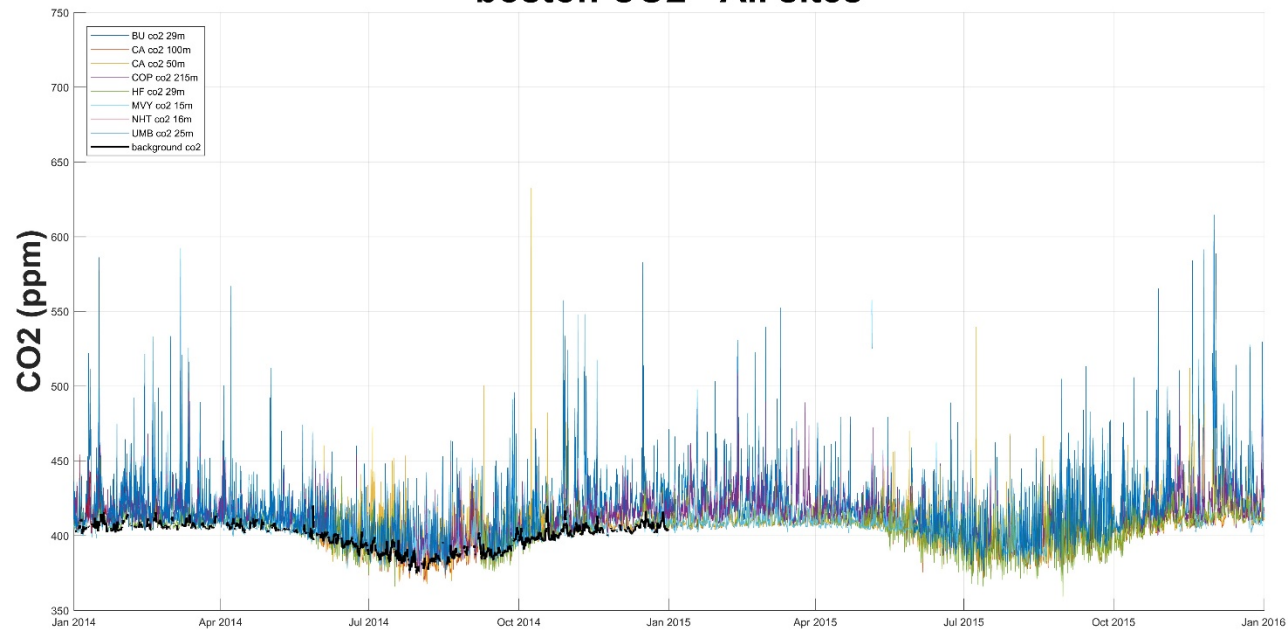


- Looking at a 2-year time period:
- Boston also does not have as high of winter time peaks.
- Also greater summertime depletion.

**salt lake city CO2 - All sites**



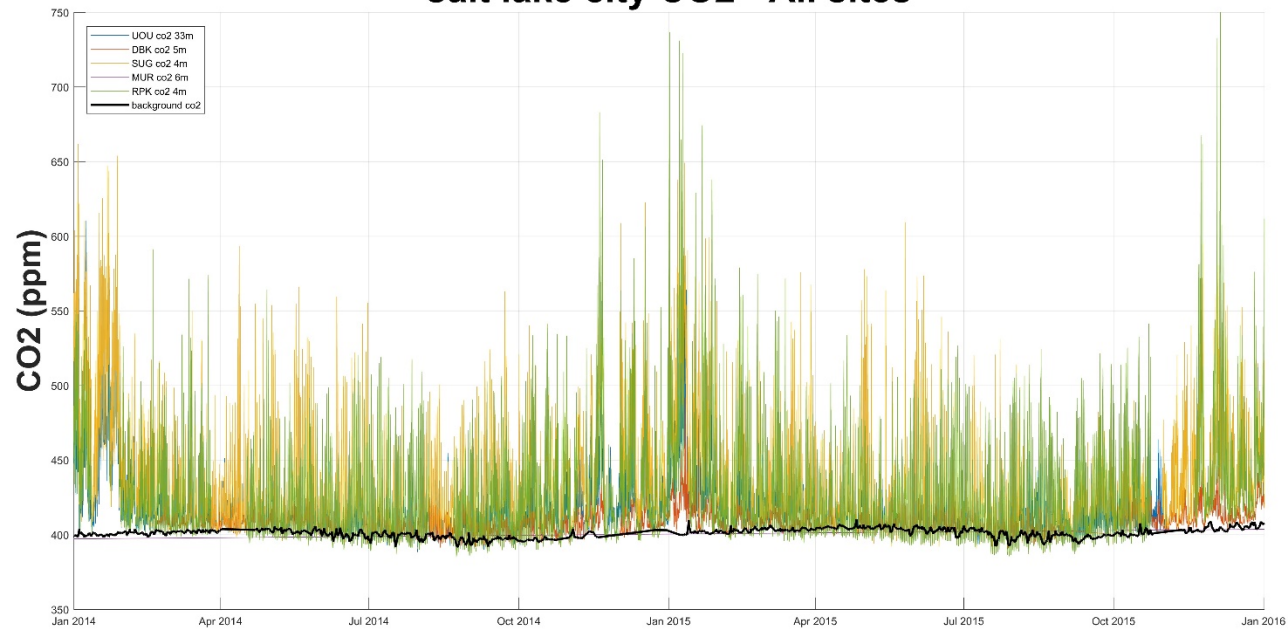
**boston CO2 - All sites**



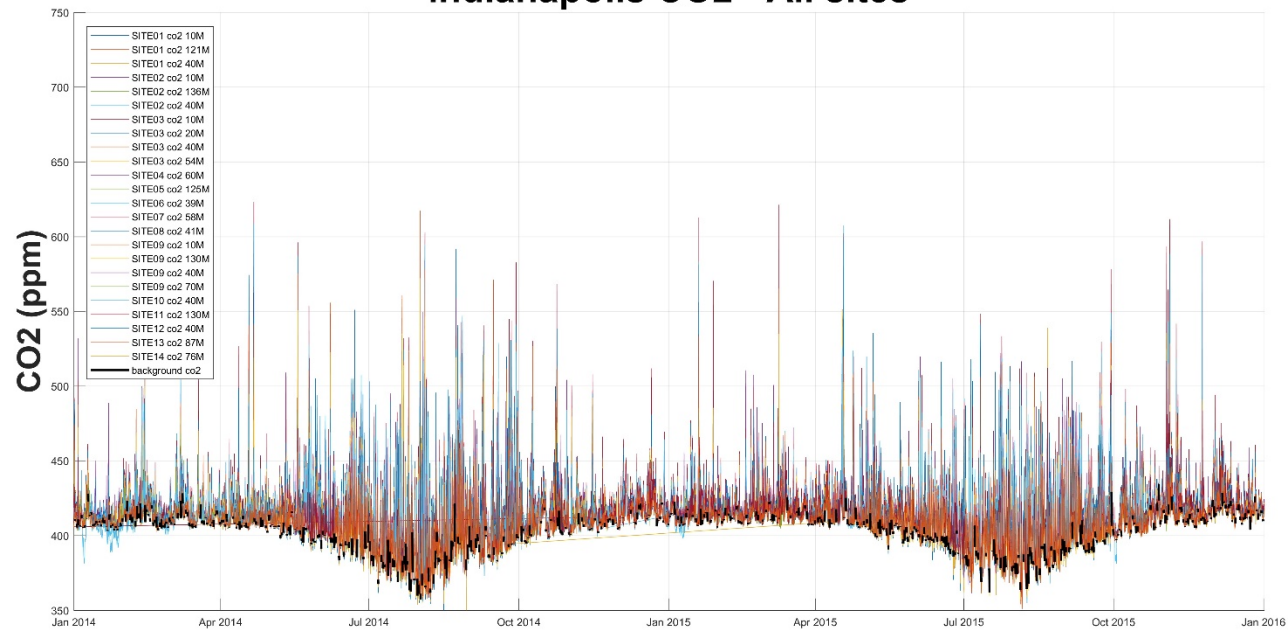


- Looking at a 2-year time period:
- Indianapolis also does not have as high of winter time peaks.
- Much greater summertime depletion, but also large peaks in the summertime.

**salt lake city CO2 - All sites**

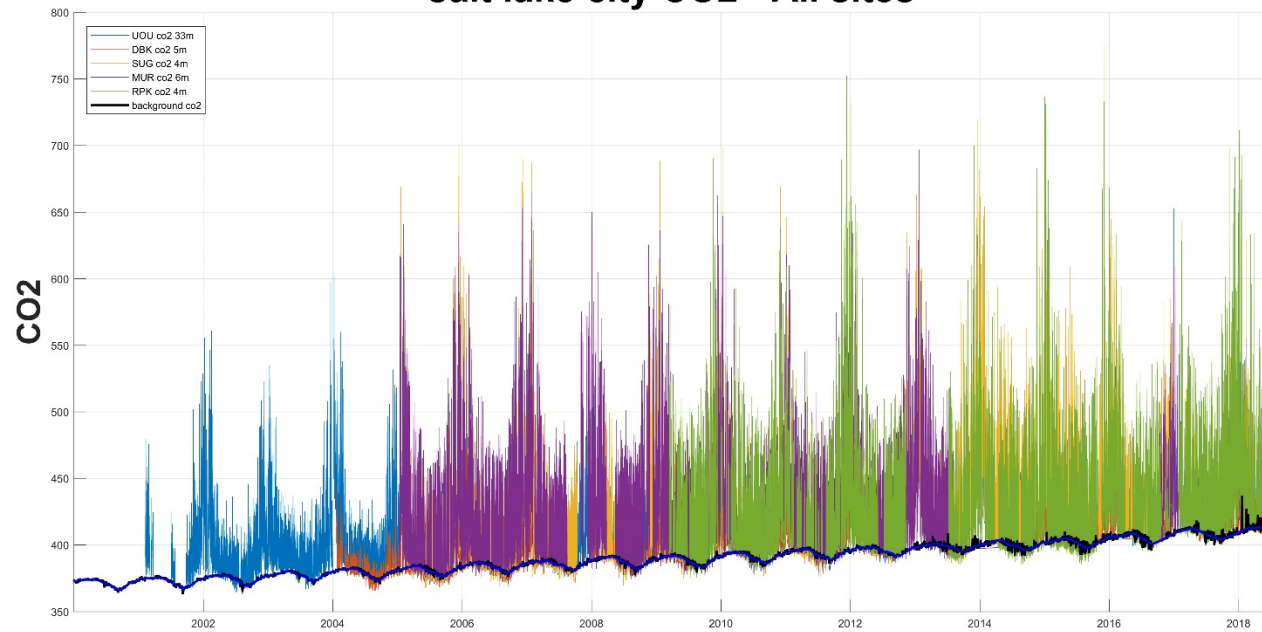


**indianapolis CO2 - All sites**

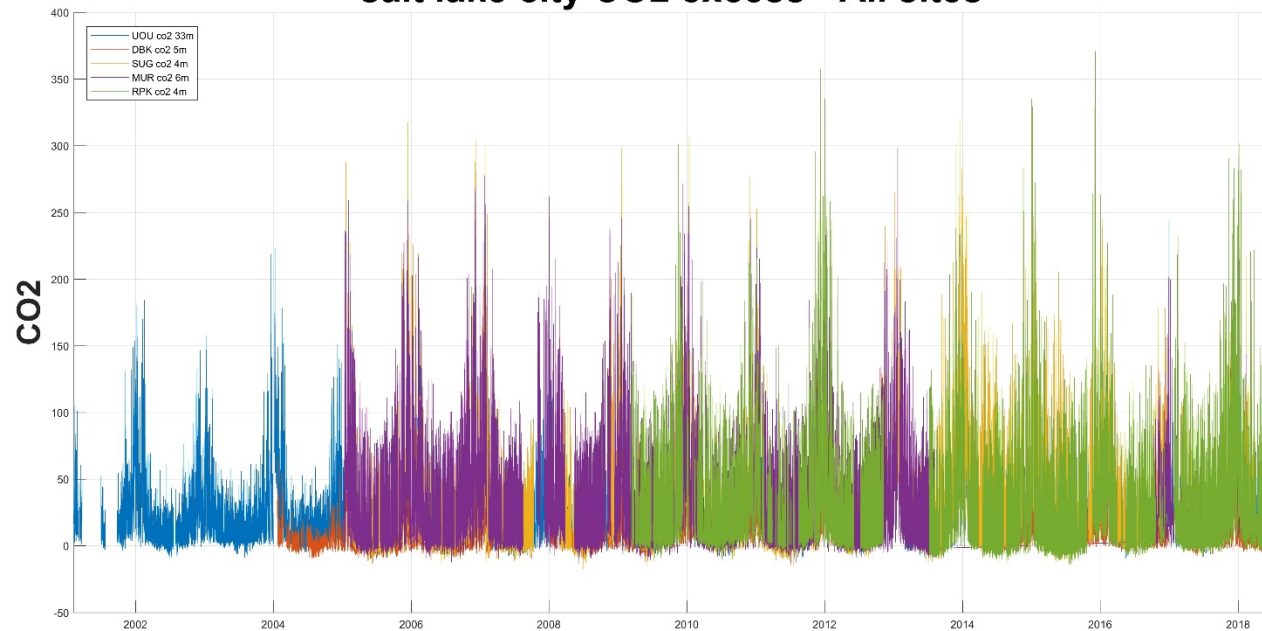


- If you subtract out the background you get CO2 excess.
- IN THEORY:
  - Represents scope 1 emissions within the station “footprint”
- IN REALITY:
  - Its complicated...

**salt lake city CO2 - All sites**

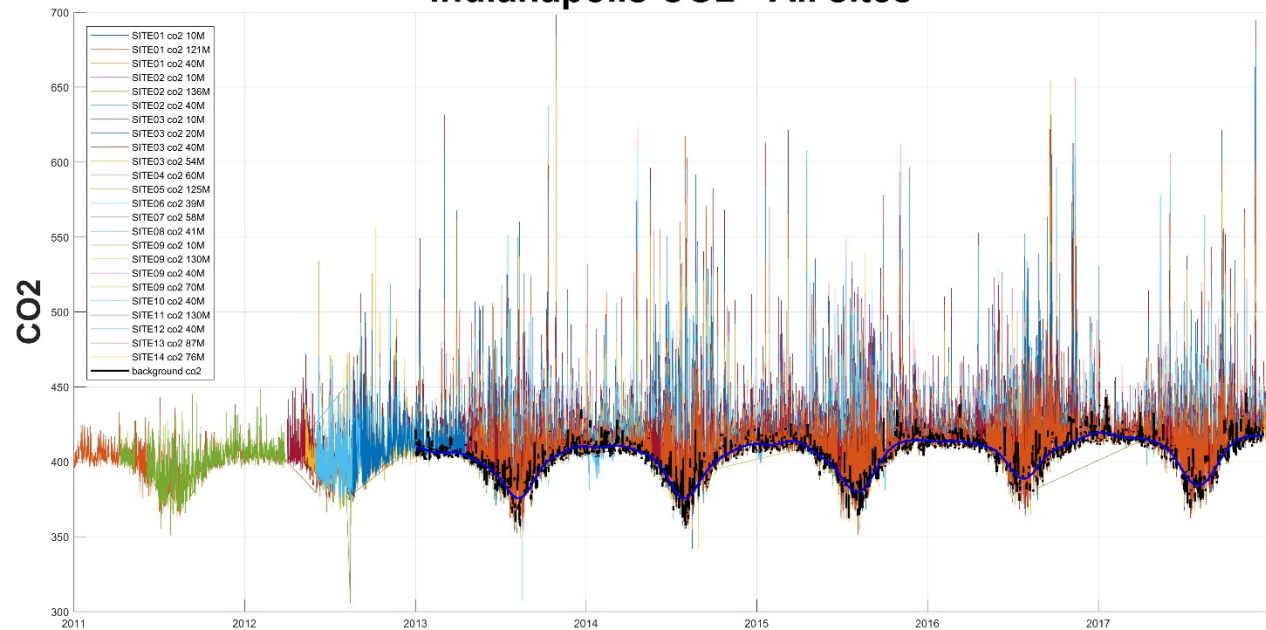


**salt lake city CO2 excess - All sites**

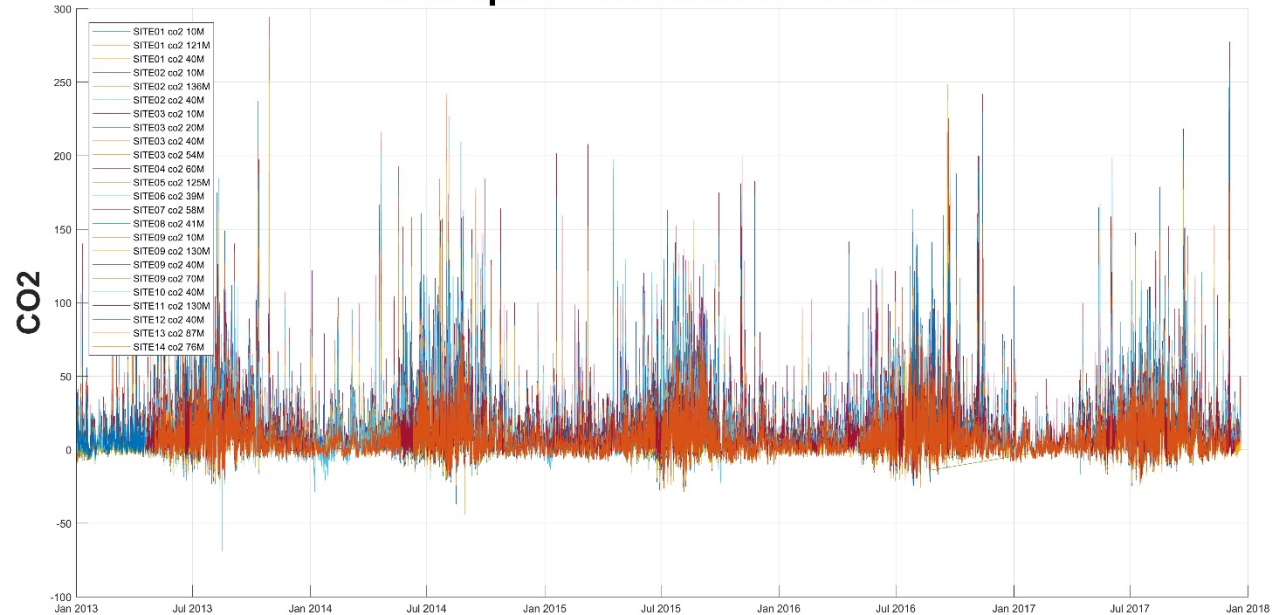


- If you subtract out the background you get CO2 excess.
- IN THEORY:
  - Represents scope 1 emissions within the station “footprint”
- IN REALITY:
  - Its complicated...

indianapolis CO2 - All sites



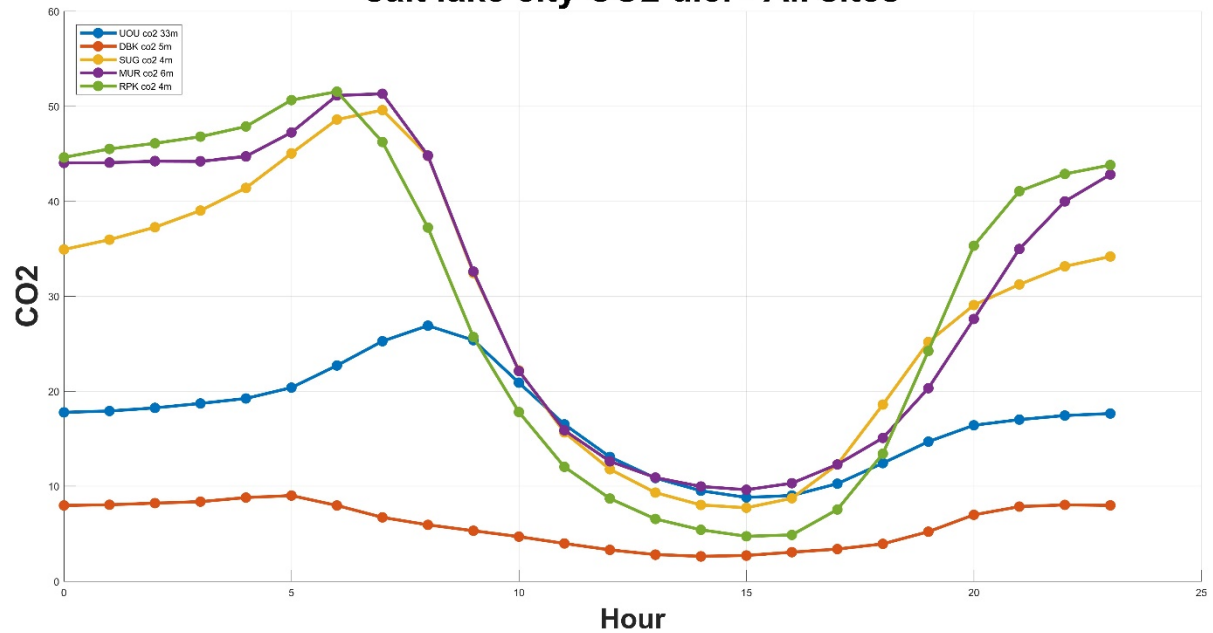
indianapolis CO2 excess - All sites



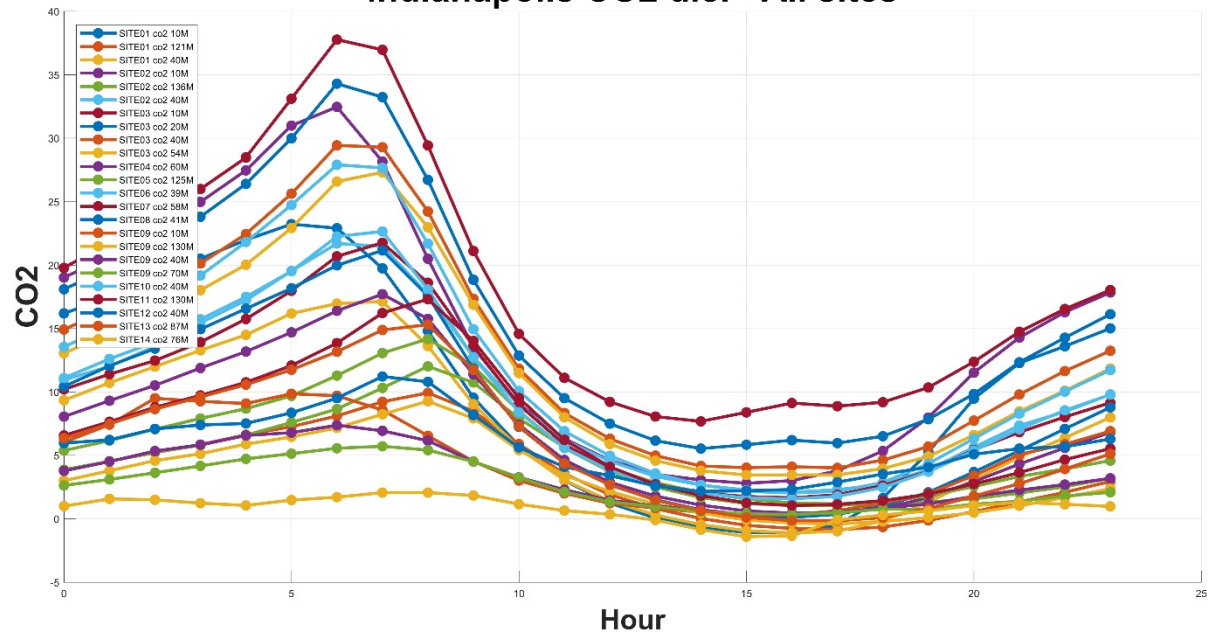


- Diel profiles
- There are meteorologically important differences
- In SLC, CO<sub>2</sub> increases in the early evening and plateaus until the AM bump.
- In Indianapolis CO<sub>2</sub> increases throughout the night.

**salt lake city CO<sub>2</sub> diel - All sites**

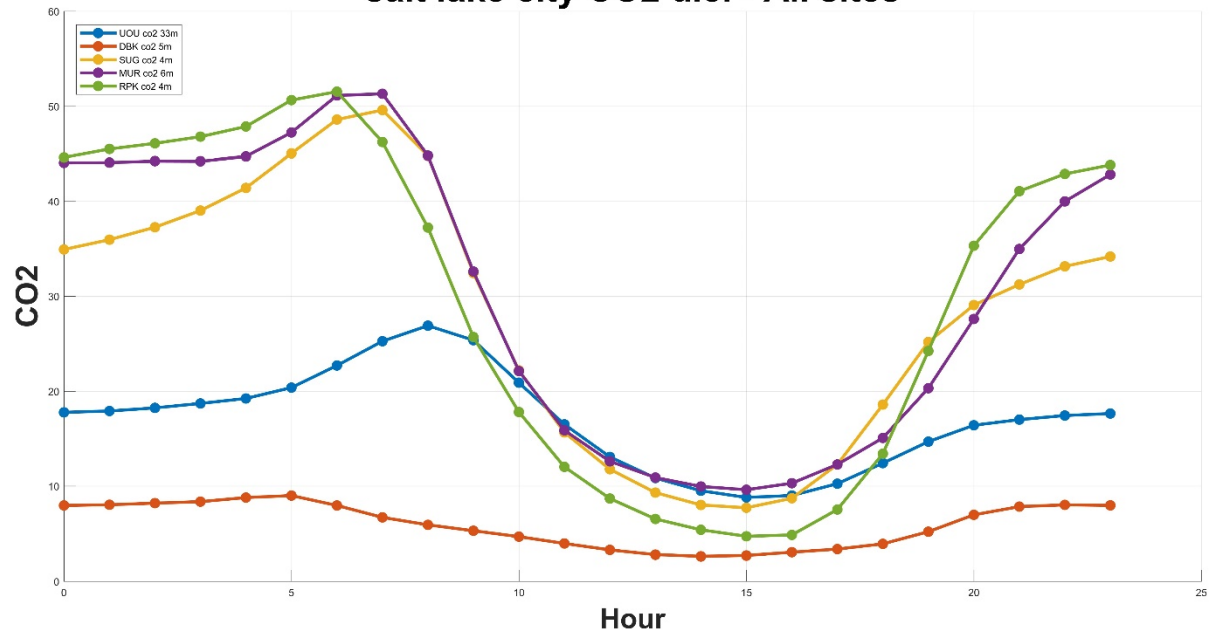


**indianapolis CO<sub>2</sub> diel - All sites**

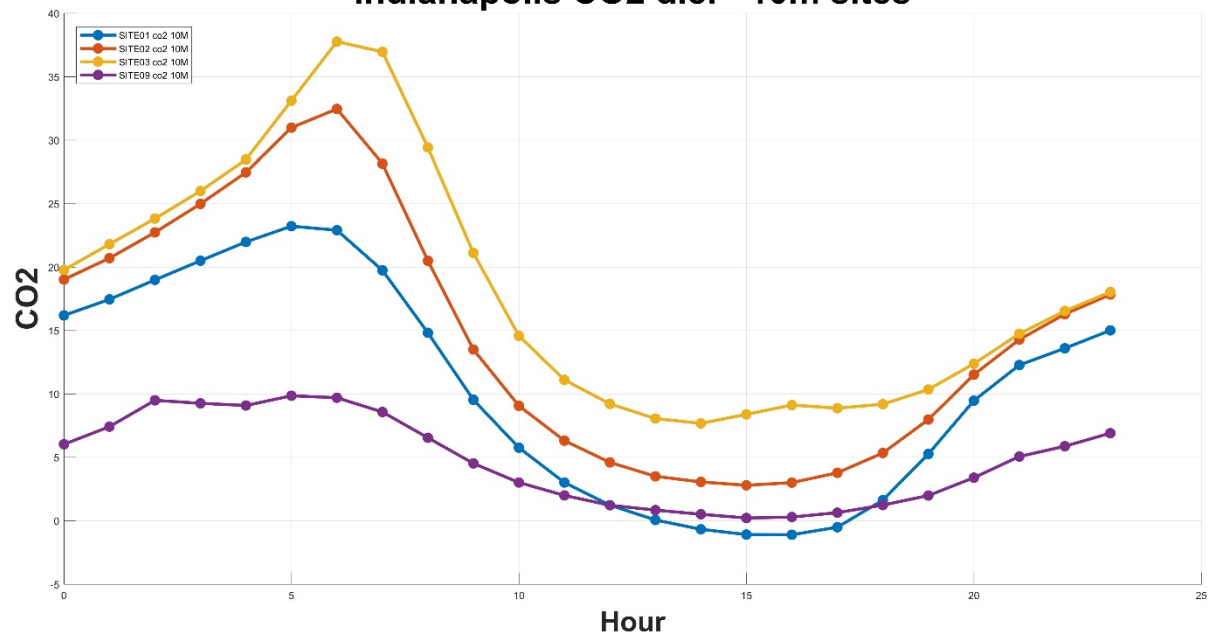


- Diel profiles
- There are meteorologically important differences
- In SLC, CO<sub>2</sub> increases in the early evening and plateaus until the AM bump.
- In Indianapolis CO<sub>2</sub> increases throughout the night.

**salt lake city CO<sub>2</sub> diel - All sites**



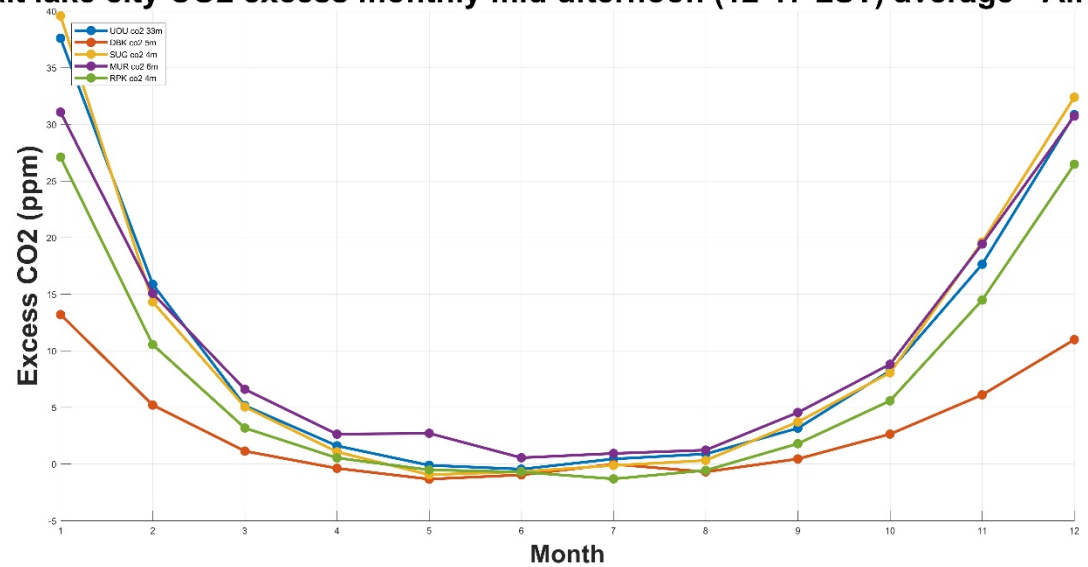
**indianapolis CO<sub>2</sub> diel - 10m sites**



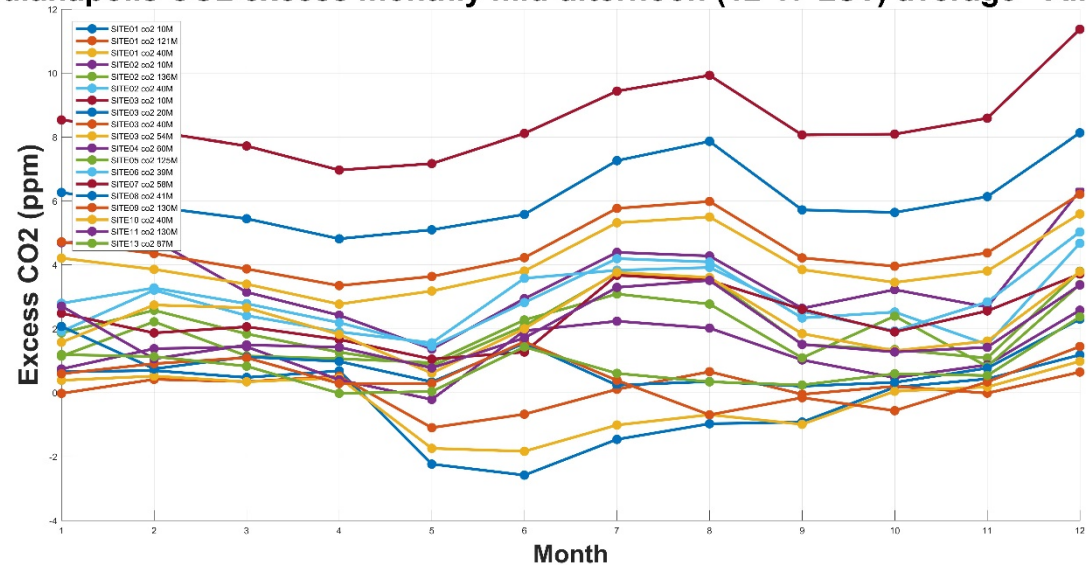


- Seasonal Cycle
- Looking only at mid-afternoon:
- Seasonal cycle is still present in SLC data
- Seasonal cycle practically disappears at Indianapolis

**salt lake city CO2 excess monthly mid-afternoon (12-17 LST) average - All sites**



**indianapolis CO2 excess monthly mid-afternoon (12-17 LST) average - All sites**





# The Background... (a discussion)

- Having a background time series makes the data set a lot more useful for initial comparisons.
  - If you're doing an inverse analysis, you will probably want to come up with your own background. Application and domain specific.
- There are a variety of methodologies being used for quantifying the background.
- Should we include a “portfolio” of background time series? For example:
  - Upwind tower
  - Filtered for high wind speed days
  - Low pass filtered
  - Carbon Tracker
  - Optional background from the data provider (e.g. a hybrid model-data background)
- There needs to be a lot more work on understanding the best practices for background construction and how it affects inverse modeling analyses.
  - Using an inappropriate background could lead to misinterpreting the data!
  - Background uncertainties!

# What's next?



- Determine the background(s) that we want to include for each city.
- Version 1 will be archived at the ORNL DAAC in the next few months.
  - <https://daac.ornl.gov/>
  - We will let the CO2-USA mailing list know when it is available.
- Write a community paper:
  - Describing the synthesis data set
  - Description of background methodologies
  - Initial cross-city comparisons
- Focus for next year is to add:
  - Update city data
  - Mobile data
  - Flask data

## THANK YOU:

- Data contributors
- Data Synthesis Working Group