

# UC Berkeley

## UC Berkeley Previously Published Works

### Title

A multi-city urban atmospheric greenhouse gas measurement data synthesis

### Permalink

<https://escholarship.org/uc/item/7md0h3sj>

### Journal

Scientific Data, 9(1)

### ISSN

2052-4463

### Authors

Mitchell, Logan E

Lin, John C

Hutyra, Lucy R

et al.

### Publication Date

2022

### DOI

10.1038/s41597-022-01467-3

Peer reviewed



OPEN

DATA DESCRIPTOR

# A multi-city urban atmospheric greenhouse gas measurement data synthesis

Logan E. Mitchell<sup>1</sup>✉, John C. Lin<sup>1</sup>, Lucy R. Hutyra<sup>2</sup>, David R. Bowling<sup>1</sup>, Ronald C. Cohen<sup>3</sup>, Kenneth J. Davis<sup>4</sup>, Elizabeth DiGangi<sup>5</sup>, Riley M. Duren<sup>6,7</sup>, James R. Ehleringer<sup>1</sup>, Clayton Fain<sup>5</sup>, Matthias Falk<sup>8</sup>, Abhinav Guha<sup>9</sup>, Anna Karion<sup>10</sup>, Ralph F. Keeling<sup>11</sup>, Jooil Kim<sup>11</sup>, Natasha L. Miles<sup>4</sup>, Charles E. Miller<sup>7</sup>, Sally Newman<sup>9</sup>, Diane E. Pataki<sup>1</sup>, Steve Prinzivalli<sup>5</sup>, Xinrong Ren<sup>12</sup>, Andrew Rice<sup>13</sup>, Scott J. Richardson<sup>4</sup>, Maryann Sargent<sup>14</sup>, Britton B. Stephens<sup>15</sup>, Jocelyn C. Turnbull<sup>16,17</sup>, Kristal R. Verhulst<sup>7</sup>, Felix Vogel<sup>18</sup>, Ray F. Weiss<sup>11</sup>, James Whetstone<sup>10</sup> & Steven C. Wofsy<sup>14</sup>

Urban regions emit a large fraction of anthropogenic emissions of greenhouse gases (GHG) such as carbon dioxide (CO<sub>2</sub>) and methane (CH<sub>4</sub>) that contribute to modern-day climate change. As such, a growing number of urban policymakers and stakeholders are adopting emission reduction targets and implementing policies to reach those targets. Over the past two decades research teams have established urban GHG monitoring networks to determine how much, where, and why a particular city emits GHGs, and to track changes in emissions over time. Coordination among these efforts has been limited, restricting the scope of analyses and insights. Here we present a harmonized data set synthesizing urban GHG observations from cities with monitoring networks across North America that will facilitate cross-city analyses and address scientific questions that are difficult to address in isolation.

## Background & Summary

Historically, fossil fuel emissions from energy consumed in cities accounted for an estimated 70% of all energy-related GHG emissions<sup>1,2</sup>. Projections of future urbanization trends suggest that the percentage of people living in cities will increase along with urban fossil fuel GHG emissions<sup>1,3</sup>. However, since the adoption of the Paris Climate Agreement in 2016, there has been substantial interest in sub-national greenhouse gas (GHG) mitigation actions, including at municipal spatial scales. Several networks of urban leaders have formed to build momentum and develop best practices for reducing emissions (e.g., C40 cities <https://www.c40.org/>, Climate Mayors <http://climatemayors.org/>, and the Global Covenant of Mayors for Climate and Energy <https://www.globalcovenantofmayors.org/>). The United Nations Framework Convention on Climate Change (UNFCCC) hosts an online portal for the Non-state Actor Zone for Climate Action (NAZCA) that include urban actions to address climate change (<https://climateaction.unfccc.int/>). At the same time as urban leaders have been making commitments to reduce GHGs, new economically competitive technologies are being developed and deployed at scale demonstrating that emission reductions are feasible<sup>4,5</sup>.

With the adoption of urban GHG emission reduction targets, there has been growing interest in monitoring progress and assessing the efficacy of mitigation policies. Creating a simple accounting of GHG emissions

<sup>1</sup>University of Utah, Salt Lake City, UT, USA. <sup>2</sup>Boston University, Boston, MA, USA. <sup>3</sup>University of California Berkeley, Berkeley, CA, USA. <sup>4</sup>The Pennsylvania State University, University Park, PA, USA. <sup>5</sup>Earth Networks, Inc., Germantown, MD, USA. <sup>6</sup>University of Arizona, Tucson, AZ, USA. <sup>7</sup>Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA. <sup>8</sup>California Air Resources Board, Sacramento, CA, USA. <sup>9</sup>Bay Area Air Quality Management District, San Francisco, CA, USA. <sup>10</sup>National Institute of Standards and Technology, Gaithersburg, MD, USA. <sup>11</sup>Scripps Institute of Oceanography, University of California San Diego, La Jolla, CA, USA. <sup>12</sup>Air Resources Laboratory, National Oceanic and Atmospheric Administration, College Park, MD, USA. <sup>13</sup>Portland State University, Portland, OR, USA. <sup>14</sup>Harvard University, Cambridge, MA, USA. <sup>15</sup>University Corporation for Atmospheric Research, Boulder, CO, USA. <sup>16</sup>GNS Science, Lower Hutt, New Zealand. <sup>17</sup>CIRES, University of Colorado at Boulder, Boulder, CO, USA. <sup>18</sup>Environment and Climate Change Canada, Toronto, Canada. ✉e-mail: [logan.mitchell@utah.edu](mailto:logan.mitchell@utah.edu)

City	# of sites	Species	City GHG emission reduction target	Project website
Boston <sup>30,31</sup>	8	CO <sub>2</sub> , CH <sub>4</sub> , CO	50% <sub>2005</sub> by 2030 & net zero by 2050 <sup>55</sup>	<a href="http://atmos.seas.harvard.edu/lab/index.html">http://atmos.seas.harvard.edu/lab/index.html</a>
Indianapolis <sup>17,18,32-34</sup>	14	CO <sub>2</sub> , CH <sub>4</sub> , CO	Net zero by 2050 <sup>56</sup>	<a href="http://sites.psu.edu/influx/">http://sites.psu.edu/influx/</a>
Los Angeles <sup>35</sup>	12	CO <sub>2</sub> , CH <sub>4</sub>	50% <sub>1990</sub> by 2025, 73% <sub>1990</sub> by 2035, & net zero by 2050 <sup>57</sup>	<a href="https://megacities.jpl.nasa.gov/">https://megacities.jpl.nasa.gov/</a>
Washington D.C./Baltimore <sup>46</sup>	13	CO <sub>2</sub> , CH <sub>4</sub>	Baltimore: 15% <sub>2010</sub> by 2020 <sup>58</sup> D.C.: 50% <sub>2006</sub> by 2032 & 80% <sub>2006</sub> by 2050 <sup>59</sup>	<a href="https://www.nist.gov/topics/northeast-corridor-urban-test-bed">https://www.nist.gov/topics/northeast-corridor-urban-test-bed</a>
Portland <sup>36</sup>	3	CO <sub>2</sub>	50% <sub>1990</sub> by 2030 & net zero by 2050 <sup>60</sup>	<a href="http://web.pdx.edu/~arice/CO2_PDX.html">http://web.pdx.edu/~arice/CO2_PDX.html</a>
Salt Lake City <sup>37-41</sup>	7	CO <sub>2</sub> , CH <sub>4</sub>	50% <sub>2009</sub> by 2030 & 80% <sub>2009</sub> by 2040 <sup>41</sup>	<a href="https://air.utah.edu/">https://air.utah.edu/</a>
San Francisco (BAAQMD) <sup>*42</sup>	4	CO <sub>2</sub> , CH <sub>4</sub> , CO	Net zero by 2050 <sup>62</sup>	<a href="https://www.baaqmd.gov/about-air-quality/air-quality-measurement/ghg-measurement">https://www.baaqmd.gov/about-air-quality/air-quality-measurement/ghg-measurement</a>
San Francisco (BEACO <sub>2</sub> N) <sup>†43</sup>	65	CO <sub>2</sub>	Net zero by 2050 <sup>62</sup>	<a href="http://www.beacon.berkeley.edu/">http://www.beacon.berkeley.edu/</a>
Toronto <sup>44,45</sup>	4	CO <sub>2</sub> , CH <sub>4</sub> , CO	30% <sub>1990</sub> by 2020, 65% <sub>1990</sub> by 2030, & net zero by 2050 <sup>63</sup>	<a href="https://www.canada.ca/en/environment-climate-change/services/climate-change/greenhouse-gases-aerosols-monitoring.html">https://www.canada.ca/en/environment-climate-change/services/climate-change/greenhouse-gases-aerosols-monitoring.html</a>

**Table 1.** CO<sub>2</sub>-USA urban greenhouse gas monitoring networks. The emissions baseline year is listed in subscript for the city GHG emission reduction targets. \*BAAQMD: Bay Area Air Quality Management District. †BEACO<sub>2</sub>N: Berkeley Environmental Air-quality & CO<sub>2</sub> Network.

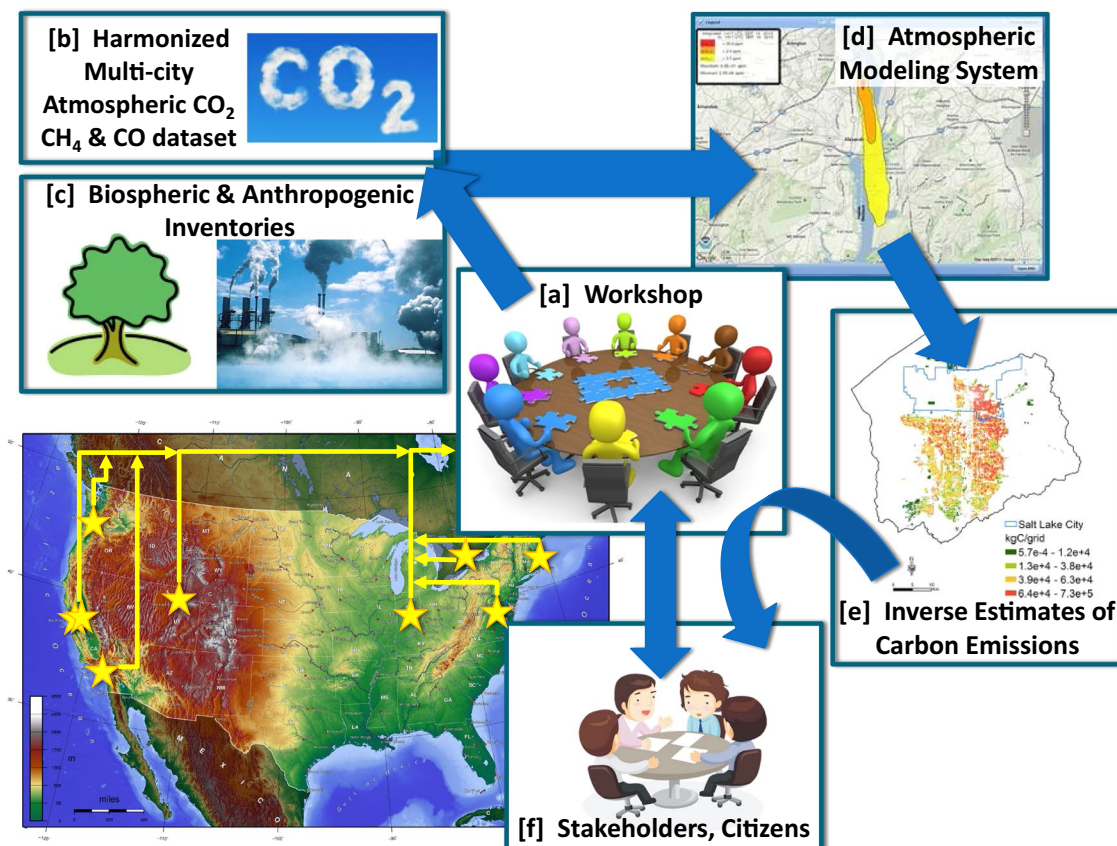
using energy consumption data is extremely difficult because granular fossil fuel consumption data are private, have considerable temporal latency, and may not have relevant geographic distributions<sup>6-8</sup>. Research teams have created fossil fuel emission inventories using a variety of techniques, however these can be incomplete or have missing source sectors<sup>8-15</sup>. Urban atmospheric measurements of GHGs are an important complement to these emissions inventories because they are sensitive to “Scope 1” emissions (i.e. direct emissions within an urban airshed) and can therefore be used to evaluate emission inventories<sup>16-18</sup>. Recognising this need, research teams in several cities established urban GHG monitoring networks over the past two decades (Table 1). As these were established, the scientific objectives and network designs varied across cities and research products primarily emanated from the research teams from each city.

The CO<sub>2</sub> Urban Synthesis and Analysis (CO<sub>2</sub>-USA) network is a synthesis effort primarily supported by the U.S. National Oceanic and Atmospheric Administration, with additional support from National Institute of Standards and Technology. CO<sub>2</sub>-USA was established to build a collaborative network of urban carbon cycle researchers to facilitate data sharing, create analysis frameworks to enable cross-city synthesis analyses<sup>19</sup>, and enable new collaborations tackling objectives that are difficult to address in isolation (Fig. 1). One of the principal objectives of the CO<sub>2</sub>-USA project was to develop a harmonized synthesis data set of atmospheric dry-air mole fraction urban GHG measurements that is readily usable, traceable to international calibration standards, and accessible to a large community: researchers, urban stakeholders, and the public. This synthesis data set could then be compared with remotely sensed or satellite data sets<sup>20-22</sup> and could also be combined with estimates of biospheric fluxes<sup>7,23</sup>, fossil fuel inventories<sup>14,24-27</sup>, and atmospheric transport models<sup>28</sup> to evaluate emissions across cities. Building a multi-city analysis framework will enable studies to quantify and understand similarities and differences in how much, where, and why GHG fluxes differ across cities. Applications could include assessing changes in emissions during the COVID-19 pandemic or creating consistent methodologies to assess emissions across different cities that are pursuing emission reduction policies<sup>19</sup>. This information could then be presented to urban stakeholders and policymakers to evaluate progress towards emission reduction goals (Fig. 1). These efforts will also advance similar efforts underway at the international level through the World Meteorological Organization (<https://ig3is.wmo.int/>).

Here we will present the synthesis GHG mole fraction data product from the CO<sub>2</sub>-USA project incorporating GHG measurements from multiple cities. We will describe the urban monitoring networks, how the data was compiled, how to access it, and how to use it.

## Methods

Construction of the CO<sub>2</sub>-USA synthesis data set follows FAIR data principles<sup>29</sup>, meaning that the data and metadata are Findable, Accessible, Interoperable, and Reusable. The data are findable by having a unique and persistent DOI, and have extensive metadata. They are accessible using standard protocols for retrieval from the web. They are interoperable by using variable naming conventions (Climate and Forecast <http://cfconventions.org/> and Attribute Convention for Data Discovery [https://wiki.esipfed.org/Attribute\\_Convention\\_for\\_Data\\_Discovery\\_1-3](https://wiki.esipfed.org/Attribute_Convention_for_Data_Discovery_1-3)) and include citations to the original references documenting the data sets. They are reusable by meeting community standards for carbon cycle measurements with detailed provenance. The data are also distributed with a ‘Fair Use Data Policy’ to guide users on appropriate use and attribution (discussed further in Usage Notes below).



**Fig. 1** Schematic of the CO<sub>2</sub>-USA project. The map shows the eight cities currently in the CO<sub>2</sub>-USA network (there are two measurement networks in San Francisco). (a) Workshops were held to determine community, stakeholder and policymaker needs. (b) This paper describes the harmonized multi-city dataset of atmospheric CO<sub>2</sub>, CH<sub>4</sub> and CO. (c) Inventories of biospheric and anthropogenic emissions were created. (d) Updates were made to atmospheric models. (e) All of these components can be used together to create estimates of urban carbon emissions using inverse analyses. (f) These results can be presented to the community, stakeholders, and policymakers.

The cities in the dataset include Boston<sup>30,31</sup>, Indianapolis<sup>17,18,32–34</sup>, Los Angeles<sup>35</sup>, Portland<sup>36</sup>, Salt Lake City<sup>37–41</sup>, San Francisco (the Bay Area Air Quality Management District [BAAQMD] network<sup>42</sup> and the Berkeley Environmental Air-quality & CO<sub>2</sub> Network [BEACO<sub>2</sub>N]<sup>43</sup>), Toronto<sup>44,45</sup>, and Washington D.C./Baltimore (the Northeast Corridor)<sup>46</sup> (Table 1). Data providers collected greenhouse gas (GHG) dry-air mole fraction measurements including carbon dioxide (CO<sub>2</sub>), methane (CH<sub>4</sub>), and carbon monoxide (CO) using a variety of instrumentation, experimental setups, and site configurations. All measurements were made using spectroscopy-based instrumentation that includes the Picarro G2301, Picarro G2401, Los Gatos Research Ultraportable Greenhouse Gas Analyzer, Los Gatos Research EP-30, Li-COR 840, LI-COR 6262, and Vaisala CarboCap GMP343. Each of these instruments also measure water vapour to correct for the spectroscopic absorption and dilution of water vapour and reports the mole fraction of the target species on a dry-air mole fraction basis.

Experimental setups varied across cities and sites within cities, but most sites had a gas analyzer, calibration tanks containing reference gas mixtures, a data logger, and tubing to an external inlet<sup>32</sup>. Additional details of the experimental setup and calibrations are described in the Technical Validation section below. The BEACO<sub>2</sub>N network in the San Francisco Bay Area used a modified setup without calibration tanks on site, and instead used a whole network calibration approach to correct for individual instrument drift<sup>43</sup>. Typical site configuration included inlets at each site installed either on towers or on the roofs of buildings. In Los Angeles, Indianapolis, and Washington D.C./Baltimore, several tower sites had up to four inlets at different heights above ground level connected to a single analyzer<sup>34,35,46</sup>. Data providers in each city conducted their own quality assurance and quality control (QA/QC) procedures on the native data set prior to assembling the data in the synthesis data set.

The temporal frequency of the native measurements was another element of experimental design that varied across cities, but each city produced hourly averaged data that is commonly used with atmospheric modelling. In this data product we reported the hourly averages using the “floored hour,” so for example, data from 08:00 to 08:59 were averaged and reported as the hour of 08:00 UTC. In addition, the standard deviation and number of measurements within the hour were reported where possible. Some data providers also calculated a more comprehensive assessment of the analytical measurement uncertainty, incorporating uncertainties from calibration gases, water vapour corrections and other factors<sup>32,35,39,46</sup>. This additional uncertainty was included in the data set where available.

A common approach in the use of urban GHG measurements is to subtract out the mole fraction of the air flowing into the city, usually referred to as the ‘background’ value. The difference between the observed and the background mole fractions results in an important quantity, the enhancement or excess amount that is intended to reflect how fluxes within the city altered the atmospheric composition of the urban air shed. While this is a simple concept, it is complex in practice for several reasons<sup>34,47</sup>, and research groups have used a variety of approaches to construct background mole fractions. The largest difference in approaches is between using a modelled, measured, or hybrid background. Modelled background mole fractions are typically derived by tracking air parcels backward in time and then assigning the air parcel the value from a relatively coarse-scale (~100 km × 100 km) global model such as Carbon Tracker<sup>28,48–50</sup>. Measured background mole fractions can be derived from measurements taken at a measurement site that is upwind of a city<sup>33–35</sup>. A hybrid approach uses measurements and model results to determine a suitable background<sup>31</sup>. An important nuance among these approaches is that the boundary of the city where the background is defined can differ greatly. For example, in a model, the boundary could be 10’s to 100’s of km away from a city depending on the grid spacing within the model. Conversely, an upwind tower could be placed on the edge of a city. For these reasons, the definition of a background is usually dependent on each application.

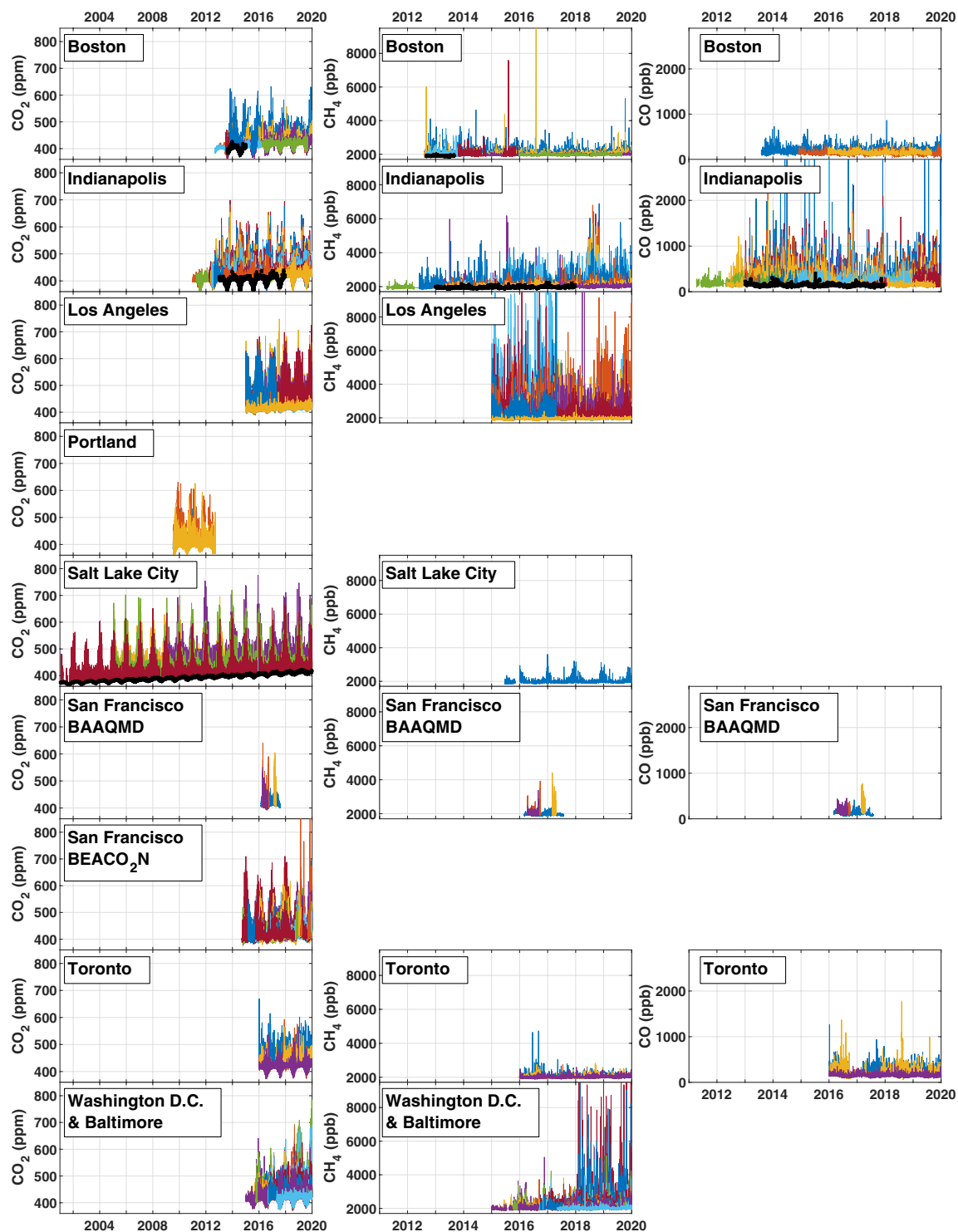
Despite the diversity in background definitions, it is still useful to report them since they represent an expert assessment of the influence of local processes on the background value. In this data synthesis we have therefore included background values at cities that have presented them in their published literature. For these cases, the city’s data set includes a file listed as ‘background.’ It is important to remember that the background is different from the other measurement sites in the data set, and to refer to the ‘references’ in the file header for additional information about how the background was constructed. For cities without a ‘background’ file, typically one or more of the sites have been used to establish background mole fractions and these details are further described in the city references.

To increase accessibility, the data were archived in both NetCDF (<https://www.unidata.ucar.edu/software/netcdf/>) and plain text data formats. The text files were created from the NetCDF files to ensure that all of the self-documented header information was identical between the files. In some cities there were multiple inlet heights and multiple species being measured at a single site. This archive used a separate file for each inlet height and species so that the file format was standardized to include a single site location, inlet height, and species in each data file.

## Data Records

The CO<sub>2</sub>-USA synthesis data set<sup>51</sup> contains measurements of CO<sub>2</sub> (671 years of data from 159 inlets at 130 sites), CH<sub>4</sub> (340 years of data from 78 inlets at 52 sites), and CO (168 years of data from 32 inlets at 22 sites) (Figs. 2 and 3). The code used to build the CO<sub>2</sub>-USA synthesis data files as well as scripts that can extract and plot the data from the data files is maintained on GitHub ([https://github.com/uataq/co2usa\\_data\\_synthesis](https://github.com/uataq/co2usa_data_synthesis)). Each data file is generated for a unique site, inlet height, and species, and also contains a comprehensive, self-documented header with the following global attributes that are formatted using the Climate and Forecast version 1.7 (CF-1.7) and Attribute Convention for Data Discovery version 1.3 (ACDD-1.3) conventions:

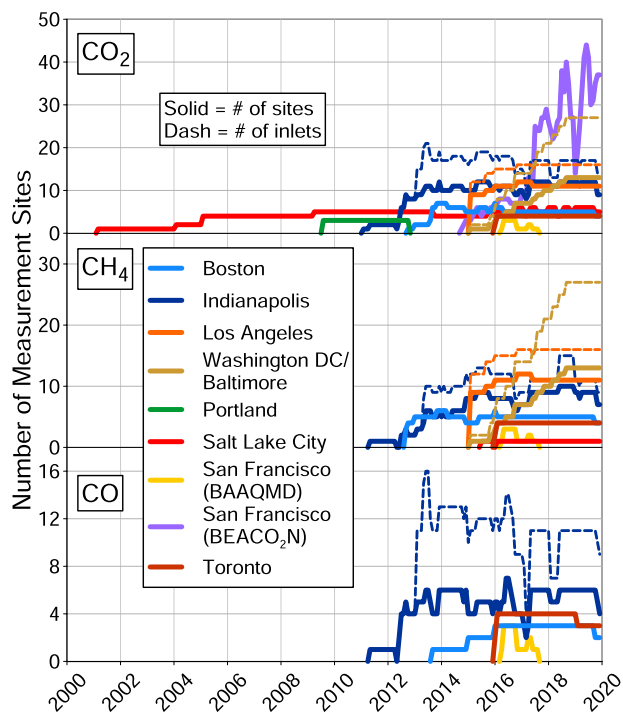
- Title – A short phrase or sentence describing the dataset.
- Summary – A paragraph describing the dataset, analogous to an abstract for a paper.
- Keywords – A comma-separated list of key words and/or phrases.
- Comment – Miscellaneous information about the data or methods used to produce it.
- References – The original citation where the data, ancillary data, and any uncertainty calculations were presented.
- Source – The method of production of the original data. In this data set all of the data were collected using spectroscopy.
- Date created – The date on which this version of the data was created.
- Date issued – The date on which this data (including all modifications) was formally issued.
- Fair use policy – Fair use policy discussed above.
- Site code – Short code used for the site.
- Site name – Long name description for the site.
- Site latitude – Current latitude of the site.
- Site longitude – Current longitude of the site.
- Site elevation – Current elevation of the site.
- Site inlet height – Current inlet height of the site.
- Site UTC2LST – Number of hours between Coordinated Universal Time (UTC) and Local Standard Time (LST). Note: LST does not include Daylight Savings time offsets.
- Dataset parameter – Species contained in the data file.
- Dataset calibration scale – International calibration scale used for the dataset measurements.
- Dataset start date – Time of the first measurement in the data file.
- Dataset end date – Time of the last measurement in the data file.
- Dataset data frequency and units – 1 hour.
- Provider information – A series of entries about the original data providers including project website, address, and email addresses.
- Compilation information – A series of entries about the data compiler including address and email address.
- ID – Digital Object Identifier (DOI) of the CO<sub>2</sub>-USA data set.
- cdm\_data\_type – The data type, as derived from Unidata’s Common Data Model Scientific Data types and understood by THREDDS. (This is a THREDDS “dataType”, and is different from the CF NetCDF attribute



**Fig. 2** Time series of the CO<sub>2</sub>-USA data synthesis data set. The columns show CO<sub>2</sub> (left), CH<sub>4</sub> (middle), and CO (right) while the rows show the CO<sub>2</sub>-USA cities in alphabetical order. Each colored line represents an inlet at a site while the thicker black lines are the background mole fractions (shown for species/cities where a published background is available). BAAQMD: Bay Area Air Quality Management District. BEACO<sub>2</sub>N: Berkeley Environmental Air-quality & CO<sub>2</sub> Network.

‘featureType’, which indicates a Discrete Sampling Geometry file in CF.) For the CO<sub>2</sub>-USA data set, this is a “timeSeries.”

- featureType – Description of a single feature with this discrete sampling geometry. For the CO<sub>2</sub>-USA data set, this is a “timeSeries.”



**Fig. 3** The number of inlets and measurement sites in each city in the CO<sub>2</sub>-USA network over time. BAAQMD: Bay Area Air Quality Management District. BEACO<sub>2</sub>N: Berkeley Environmental Air-quality & CO<sub>2</sub> Network.

- Conventions – A comma-separated list of the conventions that are followed by the dataset. For the CO<sub>2</sub>-USA data set this included: Climate and Forecast version 1.7 (CF-1.7) and Attribute Convention for Data Discovery version 1.3 (ACDD-1.3).

In addition to the global attributes, the following variable attributes were included in each data file:

- UTC time in POSIX format to facilitate a machine-readable time variable.
- UTC time using the ISO 8601 format (e.g., 2016-01-01T00:00:00Z) to provide a human-readable time variable.
- Hourly averaged species mole fractions (e.g., CO<sub>2</sub>, CH<sub>4</sub>, or CO). Hourly averages were reported using the floored hour, so for example, data from 08:00 to 08:59 were averaged and reported under the hour of 08:00 UTC.
- Standard deviation of the native mole fraction measurements within the hour.
- Number of native mole fraction measurements used to calculate the hourly average and standard deviation.
- Uncertainty of the hourly averaged mole fraction measurements determined by the data provider. This variable was included when it was available by the data providers that calculate it. The methods for determining the measurement uncertainty differ between research teams and are not always included.
- Latitude of the site at each hour. This generally stays the same through time, but it can change if a site was moved.
- Longitude of the site at each hour. This generally stays the same through time, but it can change if a site was moved.
- Elevation of the site above sea level at each hour. This generally stays the same through time, but it can change if a site was moved.
- Inlet height above ground level at the site at each hour. This generally stays the same through time, but it can change if a site was moved.

### Technical Validation

Several approaches were used to validate greenhouse gas measurements in each city. At minimum, the research teams in every city calibrated their measurements with working calibration standards in the form of high-pressure gas cylinders that have known mole fractions and are tied to World Meteorological Organization calibration standards, WMO CO<sub>2</sub> X2007<sup>52</sup>, WMO CH<sub>4</sub> X2004A<sup>53</sup>, and WMO CO X2014A<sup>54</sup>. Calibration frequency varies across cities and sites from hourly to daily intervals, with the timing being dependent on several factors that could impact instrument drift (e.g., whether or not the instrument is in a temperature-controlled setting). The BEACO<sub>2</sub>N network uses a modified approach with their low-cost sensor network design<sup>43</sup>. That network has a dedicated site with typical calibrations, and then uses a whole-network calibration approach to correct for site-specific instrument drift. In addition to regular calibrations, some cities have conducted ‘round

robin' calibration tests where one or more calibration cylinders were used to test instruments at sites across a city<sup>32</sup>, or between laboratories in different cities<sup>39</sup>.

## Usage Notes

Maintaining greenhouse gas measurement networks requires a substantial amount of expertise and effort, and there are ongoing quality control efforts in each of these GHG measurement programs. While the data set hosted by the ORNL DAAC is openly shared, without restriction, in accordance with NASA's Earth Science program, we have included a 'Fair Use Data Policy' to guide responsible use of the data set:

*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 producers and will depend on the nature of your work. While this data is available under a CC0 license, responsible use includes properly citing the data. When you start data analysis that may result in a publication, we recommend that you contact the data producers directly since they have primary knowledge of their data and any updates and, if it is appropriate, so they have the opportunity to contribute substantively to the analysis and become a co-author. Data producers reserve the right to make corrections to the data based on scientific grounds (e.g. recalibration or operational issues). This dataset is made freely and openly available, with a goal that the results of work using this data also be made freely and openly available to the greatest extent possible.*

To facilitate dissemination and use of the CO<sub>2</sub>-USA synthesis data set, we are maintaining an open access GitHub code repository: [https://github.com/uataq/co2usa\\_data\\_synthesis](https://github.com/uataq/co2usa_data_synthesis). This repository includes the code used to build the CO<sub>2</sub>-USA synthesis data files as well as scripts that can extract and plot the data from the data files. Currently the CO<sub>2</sub>-USA GitHub repository includes scripts in the R (<https://cran.r-project.org/>), Python (<https://www.python.org/>), and Matlab (<https://www.mathworks.com/>) programming languages. Use the following instructions to access the data set:

- 1) Download the CO<sub>2</sub>-USA synthesis data set<sup>51</sup> from <https://doi.org/10.3334/ORNLLAAC/1916>.
- 2) Download the "co2usa\_load\_netCDF.r", "co2usa\_load\_netCDF.py", or "co2usa\_load\_netCDF.m" scripts for use with the R, Python, or Matlab programming language, respectively.
- 3) Following the instructions in the script, select the cities and species to load, and set the path to the data files on your local computer.
- 4) Executing the script will load and plot the data.

## Code availability

All of the code used to create and extract the CO<sub>2</sub>-USA synthesis data set is maintained in an open access GitHub repository: [https://github.com/uataq/co2usa\\_data\\_synthesis](https://github.com/uataq/co2usa_data_synthesis).

Received: 10 September 2021; Accepted: 9 June 2022;

Published online: 24 June 2022

## References

1. International Energy Agency. *World energy outlook 2008*. (International Energy Agency, 2008).
2. Duren, R. M. & Miller, C. E. Measuring the carbon emissions of megacities. *Nat. Clim. Change* **2**, 560–562 (2012).
3. United Nations Department of Economic and Social Affairs. *World Urbanization Prospects: The 2018 Revision*. <https://population.un.org/wup/Publications/Files/WUP2018-Report.pdf> (2019).
4. He, G. *et al.* Rapid cost decrease of renewables and storage accelerates the decarbonization of China's power system. *Nat. Commun.* **11**, 2486 (2020).
5. Knobloch, F. *et al.* Net emission reductions from electric cars and heat pumps in 59 world regions over time. *Nat. Sustain.* 1–11, <https://doi.org/10.1038/s41893-020-0488-7> (2020).
6. Andres, R. J. *et al.* A synthesis of carbon dioxide emissions from fossil-fuel combustion. *Biogeosciences* **9**, 1845–1871 (2012).
7. Hutyra, L. R. *et al.* Urbanization and the carbon cycle: Current capabilities and research outlook from the natural sciences perspective. *Earths Future* **2**, 473–495 (2014).
8. Gatley, C. K. & Hutyra, L. R. Large Uncertainties in Urban-Scale Carbon Emissions. *J. Geophys. Res. Atmospheres* **122**, 2017JD027359 (2017).
9. Plant, G. *et al.* Large Fugitive Methane Emissions From Urban Centers Along the U.S. East Coast. *Geophys. Res. Lett.* **46**, 8500–8507 (2019).
10. Kille, N. *et al.* Separation of Methane Emissions From Agricultural and Natural Gas Sources in the Colorado Front Range. *Geophys. Res. Lett.* **46**, 3990–3998 (2019).
11. Duren, R. M. *et al.* California's methane super-emitters. *Nature* **575**, 180–184 (2019).
12. Hutchins, M. G., Colby, J. D., Marland, G. & Marland, E. A comparison of five high-resolution spatially-explicit, fossil-fuel, carbon dioxide emission inventories for the United States. *Mitig. Adapt. Strateg. Glob. Change* 1–26, <https://doi.org/10.1007/s11027-016-9709-9> (2016).
13. Gurney, K. R., Huang, J. & Coltin, K. Bias present in US federal agency power plant CO<sub>2</sub> emissions data and implications for the US clean power plan. *Environ. Res. Lett.* **11**, 064005 (2016).
14. Gurney, K. R. *et al.* Comparison of Global Downscaled Versus Bottom-Up Fossil Fuel CO<sub>2</sub> Emissions at the Urban Scale in Four U.S. Urban Areas. *J. Geophys. Res. Atmospheres* **124**, 2823–2840 (2019).
15. Gurney, K. R. *et al.* Under-reporting of greenhouse gas emissions in U.S. cities. *Nat. Commun.* **12**, 553 (2021).
16. Gurney, K. R. *et al.* Climate change: Track urban emissions on a human scale. *Nature* **525**, 179–181 (2015).
17. Turnbull, J. C. *et al.* Synthesis of Urban CO<sub>2</sub> Emission Estimates from Multiple Methods from the Indianapolis Flux Project (INFLUX). *Environ. Sci. Technol.* **53**, 287–295 (2019).
18. Lauvaux, T. *et al.* Policy-Relevant Assessment of Urban CO<sub>2</sub> Emissions. *Environ. Sci. Technol.* **54**, 10237–10245 (2020).



19. Yadav, V. *et al.* The Impact of COVID-19 on CO<sub>2</sub> Emissions in the Los Angeles and Washington DC/Baltimore Metropolitan Areas. *Geophys. Res. Lett.* **48**, e2021GL092744 (2021).
20. Wu, D., Lin, J. C., Oda, T. & Kort, E. A. Space-based quantification of per capita CO<sub>2</sub> emissions from cities. *Environ. Res. Lett.* **15**, 035004 (2020).
21. Crisp, D. *et al.* The on-orbit performance of the Orbiting Carbon Observatory-2 (OCO-2) instrument and its radiometrically calibrated products. *Atmospheric Meas. Tech.* **10**, 59–81 (2017).
22. Eldering, A., Taylor, T. E., O'Dell, C. W. & Pavlick, R. The OCO-3 mission: measurement objectives and expected performance based on 1 year of simulated data. *Atmospheric Meas. Tech.* **12**, 2341–2370 (2019).
23. Pataki, D. E. *et al.* Urban ecosystems and the North American carbon cycle. *Glob. Change Biol.* **12**, 2092–2102 (2006).
24. Oda, T. *et al.* On the impact of granularity of space-based urban CO<sub>2</sub> emissions in urban atmospheric inversions: A case study for Indianapolis, IN. *Elem Sci Anth* **5** (2017).
25. Crippa, M. *et al.* Global anthropogenic emissions in urban areas: patterns, trends, and challenges. *Environ. Res. Lett.* **16**, 074033 (2021).
26. Gurney, K. R. *et al.* The Hestia fossil fuel CO<sub>2</sub> emissions data product for the Los Angeles megacity (Hestia-LA). *Earth Syst. Sci. Data* **11**, 1309–1335 (2019).
27. Patarasuk, R. *et al.* Urban high-resolution fossil fuel CO<sub>2</sub> emissions quantification and exploration of emission drivers for potential policy applications. *Urban Ecosyst.* **19**, 1013–1039 (2016).
28. Fasoli, B., Lin, J. C., Bowling, D. R., Mitchell, L. & Mendoza, D. Simulating atmospheric tracer concentrations for spatially distributed receptors: updates to the Stochastic Time-Inverted Lagrangian Transport model's R interface (STILT-R version 2). *Geosci. Model Dev.* **11**, 2813–2824 (2018).
29. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Sci. Data* **3**, 160018 (2016).
30. McKain, K. *et al.* Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts. *Proc. Natl. Acad. Sci.* 201416261, <https://doi.org/10.1073/pnas.1416261112> (2015).
31. Sargent, M. *et al.* Anthropogenic and biogenic CO<sub>2</sub> fluxes in the Boston urban region. *Proc. Natl. Acad. Sci.* **115**, 7491–7496 (2018).
32. Richardson, S. J. *et al.* Tower measurement network of *in-situ* CO<sub>2</sub>, CH<sub>4</sub>, and CO in support of the Indianapolis FLUX (INFLUX) Experiment. *Elem Sci Anth* **5** (2017).
33. Davis, K. J. *et al.* The Indianapolis Flux Experiment (INFLUX): A test-bed for developing urban greenhouse gas emission measurements. *Elem. Sci. Anthr.* **5** (2017).
34. Miles, N. L. *et al.* Quantification of urban atmospheric boundary layer greenhouse gas dry mole fraction enhancements in the dormant season: Results from the Indianapolis Flux Experiment (INFLUX). *Elem Sci Anth* **5** (2017).
35. Verhulst, K. R. *et al.* Carbon dioxide and methane measurements from the Los Angeles Megacity Carbon Project—Part 1: calibration, urban enhancements, and uncertainty estimates. *Atmospheric Chem. Phys.* **17**, 8313–8341 (2017).
36. Rice, A. & Boström, G. Measurements of carbon dioxide in an Oregon metropolitan region. *Atmos. Environ.* **45**, 1138–1144 (2011).
37. Mitchell, L. E. *et al.* Long-term urban carbon dioxide observations reveal spatial and temporal dynamics related to urban characteristics and growth. *Proc. Natl. Acad. Sci.* **115**, 2912–2917 (2018).
38. Lin, J. C. *et al.* CO<sub>2</sub> and Carbon Emissions from Cities: Linkages to Air Quality, Socioeconomic Activity, and Stakeholders in the Salt Lake City Urban Area. *Bull. Am. Meteorol. Soc.* **99**, 2325–2339 (2018).
39. Bares, R. *et al.* The Utah urban carbon dioxide (UUCON) and Uintah Basin greenhouse gas networks: instrumentation, data, and measurement uncertainty. *Earth Syst. Sci. Data* **11**, 1291–1308 (2019).
40. Pataki, D. E., Bowling, D. R. & Ehleringer, J. R. Seasonal cycle of carbon dioxide and its isotopic composition in an urban atmosphere: Anthropogenic and biogenic effects. *J. Geophys. Res.-Atmospheres* **108**, 1–8 (2003).
41. Pataki, D. E., Bowling, D. R., Ehleringer, J. R. & Zobitz, J. M. High resolution atmospheric monitoring of urban carbon dioxide sources. *Geophys. Res. Lett.* **33**, 1–5 (2006).
42. Newman, S. & Guha, A. Bay Area Air Quality Management District (BAAQMD) Greenhouse Gas Measurement Program. <https://www.baaqmd.gov/about-air-quality/air-quality-measurement/ghg-measurement> (2017).
43. Shusterman, A. A. *et al.* The Berkeley Atmospheric CO<sub>2</sub> Observation Network: initial evaluation. *Atmos Chem Phys* **16**, 13449–13463 (2016).
44. Vogel, F. R. *et al.* Regional non-CO<sub>2</sub> greenhouse gas fluxes inferred from atmospheric measurements in Ontario, Canada. *J. Integr. Environ. Sci.* **9**, 41–55 (2012).
45. Pugliese, S. C. *et al.* High-resolution quantification of atmospheric CO<sub>2</sub> mixing ratios in the Greater Toronto Area, Canada. *Atmospheric Chem. Phys.* **18**, 3387–3401 (2018).
46. Karion, A. *et al.* Greenhouse gas observations from the Northeast Corridor tower network. *Earth Syst. Sci. Data* **12**, 699–717 (2020).
47. Balashov, N. V. *et al.* Background heterogeneity and other uncertainties in estimating urban methane flux: results from the Indianapolis Flux Experiment (INFLUX). *Atmospheric Chem. Phys.* **20**, 4545–4559 (2020).
48. Peters, W. *et al.* An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. *Proc. Natl. Acad. Sci.* **104**, 18925–18930 (2007).
49. Mallia, D. V., Lin, J. C., Urbanski, S., Ehleringer, J. & Nehrkorn, T. Impacts of upwind wildfire emissions on CO, CO<sub>2</sub>, and PM<sub>2.5</sub> concentrations in Salt Lake City, Utah. *J. Geophys. Res. Atmospheres* **120**, 2014JD022472 (2015).
50. Gerbig, C. *et al.* Toward constraining regional-scale fluxes of CO<sub>2</sub> with atmospheric observations over a continent: 2. Analysis of COBRA data using a receptor-oriented framework. *J. Geophys. Res. Atmospheres* **108**, n/a–n/a (2003).
51. Mitchell, L. E. *et al.* NACP: Urban Greenhouse Gases across the CO<sub>2</sub> Urban Synthesis and Analysis Network, V2, ORNL Distributed Active Archive Center, <https://doi.org/10.3334/ORNLDAAC/1916> (2021).
52. Zhao, C. L. & Tans, P. P. Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air. *J. Geophys. Res. Atmospheres* **111**, D08S09 (2006).
53. Dlugokencky, E. J. *et al.* Conversion of NOAA atmospheric dry air CH<sub>4</sub> mole fractions to a gravimetrically prepared standard scale. *J. Geophys. Res.-Atmospheres* **110**, 8 (2005).
54. NOAA Global Monitoring Laboratory. Carbon Monoxide (CO) WMO Scale. [https://www.esrl.noaa.gov/gmd/ccl/co\\_scale.html](https://www.esrl.noaa.gov/gmd/ccl/co_scale.html).
55. City of Boston. *Boston Climate Action* <https://www.boston.gov/departments/environment/boston-climate-action> (2019).
56. City of Indianapolis & Marion County. *Thrive Indianapolis* <https://www.thriveindianapolis.com> (2019).
57. City of Los Angeles. *L.A.'s Green New Deal Sustainability pLAN 2019* <https://plan.lamayor.org/node> (2019).
58. City of Baltimore. *Climate Action Plan* <https://www.baltimore.sustainability.org/plans/climate-action-plan/> (2013).
59. Government of the District of Columbia. *Climate Action Planning* <https://doee.dc.gov/service/climate-action-planning> (2018).
60. City of Portland. *History and key documents of climate planning and action in Portland* <https://www.portland.gov/bps/climate-action/history-and-key-documents> (2020).
61. Salt Lake City Corporation. *Climate Positive 2040* <https://www.slc.gov/sustainability/climate-positive/> (2019).
62. San Francisco Department of the Environment. *San Francisco Climate Action Plan* <https://sfplanning.org/project/san-francisco-climate-action-plan> (2013).
63. City of Toronto. *TransformTO* <https://www.toronto.ca/services-payments/water-environment/environmentally-friendly-city-initiatives/transformto/> (2017).

## Acknowledgements

This study was supported by NOAA Climate Program Office's Atmospheric Chemistry, Carbon Cycle, and Climate Program grants NA17OAR4310084, with additional support from National Institute of Standards and Technology. Part of the research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration (80NM0018D0004). Each of the urban monitoring projects have been supported by multiple agencies and sponsors, and acknowledgments for those can be found in the cited literature for each city. Disclaimer: References are made to certain commercially available products in this paper to adequately specify the experimental procedures involved. Such identification does not imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that these products are the best for the purpose specified.

## Author contributions

L.E.M. wrote the code compiling the data set. J.C.L., L.R.H., and S.C.W. designed and oversaw the overall project. All authors contributed data to the CO<sub>2</sub>-USA data synthesis and to the writing of the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to L.E.M.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022