

RESEARCH ARTICLE

A spatially explicit inventory scaling approach to estimate urban CO₂ emissions

Kristian D. Hajny^{1,2,*}, Cody R. Floerchinger³, Israel Lopez-Coto^{2,4}, Joseph R. Pitt^{2,5}, Conor K. Gately^{3,6}, Kevin R. Gurney⁷, Lucy R. Hutyra⁸, Thilina Jayarathne^{1,9}, Robert Kaeser¹, Geoffrey S. Roest⁷, Maryann Sargent³, Brian H. Stirm¹⁰, Jay Tomlin¹, Alexander J. Turner¹¹, Paul B. Shepson^{1,2}, and Steven Wofsy³

Appropriate techniques to quantify greenhouse gas emission reductions in cities over time are necessary to monitor the progress of these efforts and effectively inform continuing mitigation. We introduce a scaling factor (SF) method that combines aircraft measurements and dispersion modeling to estimate urban emissions and apply it to 9 nongrowing season research aircraft flights around New York City (NYC) in 2018–2020. This SF approach uses a weighting function to focus on an area of interest while still accounting for upwind emissions. We estimate carbon dioxide (CO₂) emissions from NYC and the Greater New York Area (GNA) and compare to nested inversion analyses of the same data. The average calculated CO₂ emission rates for NYC and the GNA, representative of daytime emissions for the flights, were (49 ± 16) kmol/s and (144 ± 44) kmol/s, respectively (uncertainties reported as $\pm 1\sigma$ variability across the 9 flights). These emissions are within $\sim 15\%$ of an inversion analysis and agree well with inventory estimates. By using an ensemble, we also investigate the variability introduced by several sources and find that day-to-day variability dominates the overall variability. This work investigates and demonstrates the capability of an SF method to quantify emissions specific to particular areas of interest.

Keywords: Urban emissions, Carbon dioxide, Emission quantification, Airborne greenhouse gas measurements, Inventory scaling, New York City

1. Introduction

Carbon dioxide (CO₂) emissions from fossil fuel combustion are the largest source of anthropogenic climate change and urban areas are responsible for a large fraction of these emissions (Intergovernmental Panel on Climate Change [IPCC], 2013). According to 2015 inventory data (Gurney et al., 2020a) and census urban area definitions (U.S. Census Bureau, 2021c), 43.8% of CO₂ emissions across the contiguous United States came from urban areas and the urban share of emissions is expected to increase as populations continue urbanizing (International Energy Agency, 2013; IPCC, 2013; United Nations

Department of Economic and Social Affairs Population Division, 2019). As such, cities have also been at the forefront of greenhouse gas (GHG) reduction efforts (Trencher et al., 2016; Sethi et al., 2020). Many U.S. cities have set goals for significant GHG reductions and several states have passed legislation to invest in renewable energy (Sierra Club, 2020). In particular, New York (NY) state recently passed several laws outlining actions to significantly reduce GHG emissions with set targets across sectors. The Climate Leadership and Community Protection Act sets a goal of net zero emissions from electricity, including purchased electricity, by 2040 and an 85%

¹Department of Chemistry, Purdue University, West Lafayette, IN, USA

²School of Marine and Atmospheric Sciences, Stony Brook University, Stony Brook, NY, USA

³Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA

⁴Special Programs Office, National Institute of Standards and Technology, Gaithersburg, MD, USA

⁵Current address: School of Chemistry, University of Bristol, Bristol, UK

⁶Current address: Massachusetts Metropolitan Area Planning Council, Boston, MA, USA

⁷School of Informatics, Computing, and Cyber Systems, Northern Arizona University, Flagstaff, AZ, USA

⁸Department of Earth and Environment, Boston University, Boston, MA, USA

⁹Current address: Bristol Myers Squibb, New Brunswick, NJ, USA

¹⁰School of Aviation and Transportation Technology, Purdue University, West Lafayette, IN, USA

¹¹Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA

*Corresponding author:
 Email: kristian.hajny@stonybrook.edu

Table 1. Statistics for 3 northeastern U.S. cities (U.S. Census Bureau, 2021b). DOI: <https://doi.org/10.1525/elementa.2021.00121.t1>

Urban Area	Area (km ²)	Population (Million People)	Population Density (People/km ²)	Bottom-Up CO ₂ Emissions (kmol/s)
Greater New York Area (NY-NJ-CT)	8,936	18.4	2,054	103
Boston (MA-NH-RI)	4,852	4.2	862	26
Washington (DC-VA-MD)	3,423	4.6	464	20

Bottom-up emissions are from Vulcan (Gurney et al., 2020a), discussed further in the Methods and Materials section.

decrease across all sectors by 2050 (Kaminsky, 2019). In addition to this, New York City (NYC) has passed the Climate Mobilization Act that includes several bills targeting building emissions, the largest emitting sector for the city when including indirect emissions (Cventure LLC et al., 2017), with large fines for noncompliance (New York City Council, 2019).

Several recent studies focused on the assessment of large northeastern U.S. city emissions have been published (McKain et al., 2015; Ren et al., 2018; Sargent et al., 2018; Plant et al., 2019; Ahn et al., 2020; Lopez-Coto et al., 2020; Floerchinger et al., 2021; Pitt et al., 2022). **Table 1** highlights some key details about 3 of these northeastern U.S. urban areas for context (U.S. Census Bureau, 2021b). It is clear from **Table 1** that the Greater New York Area (GNA), or New York-Newark urban area as named in the Census database (U.S. Census Bureau, 2021c), is significantly larger (more urban sprawl) and denser than the previously studied urban areas. Additionally, almost half of the CO₂ emissions come from NYC, the dense core of the GNA. All urban areas in **Table 1**, including the GNA, are those defined by the U.S. census. Their definition of urbanized areas can be summarized as a central urban core and any densely settled surrounding territory (urban fringe) that together has a minimum of 50,000 people, typically including any contiguous urban fringe with at least 1,000 people per square mile (U.S. Census Bureau, 1995).

As cities enact laws with explicit reduction requirements, it becomes necessary to have high-precision emissions monitoring techniques to inform progress. This is often done using self-reported bottom-up inventories, which involve combining emission factors (e.g., kmol CO₂/m² s for residential housing) with activity data (e.g., total area of residential homes, m²) to calculate emissions across source categories. In addition to the self-reported inventories that cities use, there are published bottom-up inventories that utilize a combination of emission factors and activity data, modeling, monitoring station data, fuel statistics, and so on from agencies like the Energy Information Agency and Environmental Protection Agency to allocate emissions in a gridded field (Gurney et al., 2009; Gurney et al., 2012; Gately & Hutyra, 2017). Gurney et al. (2021) discuss in detail the large differences that can exist between these self-reported inventories and published inventories. There are also published disaggregated inventories that rely instead on proxies to disaggregate reported national total emissions to a gridded product. These can

be generated more regularly at a global scale but have additional uncertainty due to the use of a proxy rather than the relevant activity data (Janssens-Maenhout et al., 2017; Oda et al., 2018; Gurney et al., 2019). These published inventories can inform policy about emissions and are often used as priors in atmospheric transport modeling efforts (Lamb et al., 2016; Gurney et al., 2019).

Bottom-up inventories rely on a vast amount of input information, which makes them difficult and time-consuming to develop. This leads to gaps between emissions represented in inventories and present-day emissions. Top-down studies fill this gap and provide an independent estimate to compare to inventories by using measurements to quantify emissions in near-real time (Cambaliza et al., 2014; Lauvaux et al., 2016; Gourdji et al., 2018; Pitt et al., 2018; Turnbull et al., 2019; Gurney et al., 2021). For example, Turner et al. (2020) and Yadav et al. (2021) quantified CO₂ reductions during the COVID-19 pandemic and identified the economic sectors most affected. Such a task would be impossible without both high-quality inventories and recent high-precision measurements. Timely information like this can be crucial to inform policy makers about the impacts of legislation, so they can adjust policy as needed to meet emission reduction targets.

Three commonly used top-down methods for emission quantification are the mass balance experiment, inverse modeling, and inventory scaling. Mass balance experiments provide a domain-total emission rate that can be calculated relatively quickly based on a simple conceptual model. However, it can be difficult to define an appropriate background concentration and the area which the calculated emission rate represents (Cambaliza et al., 2014; O'Shea et al., 2014; Heimbürger et al., 2017; Ren et al., 2018; Ahn et al., 2020). Inverse modeling techniques use a dispersion model to simulate concentration enhancements at the measurement locations based on an initial emission map (prior), typically an emission inventory. They then typically weight these data based on prescribed error covariances of the transport model, prior, and measurements to reach an optimized emission map (posterior) that reduces the mismatch between the modeled and measured enhancements (Michalak et al., 2004; Mueller et al., 2008; Stohl et al., 2009; Lauvaux et al., 2016; Lauvaux et al., 2020; Lopez-Coto et al., 2020). Scaling approaches follow the same basic premise as an inverse model except they optimize the entire emission

map using a single scaling factor (SF). This assumes that the spatial distribution of emissions in the prior is accurate; however, it allows for a simpler approach to calculate posterior emissions and their sensitivity to model parameters while avoiding the complexities associated with defining error covariances (McKain et al., 2015; Pitt et al., 2018; Sargent et al., 2018; Feng et al., 2019; Kunik et al., 2019; Pitt et al., 2022). Karion et al. (2019) and Neininger et al. (2021) discuss work using several of these approaches and compare the results across them.

In this work, we introduce a variation of the SF approach. This approach requires a prior emission map of surface fluxes, a dispersion model to relate surface emissions to downwind measurements, and measured concentrations against which these modeled enhancements are compared. No other measured quantities are used. As legislation is limited to political boundaries, the focus of an SF approach should be on such discrete areas of interest (AOIs). However, measurements are influenced by all sources upwind and not solely those within an AOI, so modeled enhancements must incorporate all upwind emissions (i.e., use a large modeling domain) to appropriately compare to measured enhancements. As such, defining an appropriate background concentration to separate the AOI from other upwind sources is crucial (Mueller et al., 2018; Pitt et al., 2018; Sargent et al., 2018). Here, we begin with a simple long-range background based on the lowest observed and modeled concentrations. This allows us to relate measured and modeled enhancements as it accounts for the regional background concentration that only affects measurements, but the resulting enhancements represent all upwind emissions within the modeled domain (Lopez-Coto et al., 2020; Pitt et al., 2022). Using these enhancements in an SF approach requires assuming that the SF for the entire upwind region is the same as that of just the AOI. This could be particularly problematic with small AOIs such as cities where only a fraction of the total emissions may originate from the AOI. For example, one can imagine a prior with city emissions lower than true emissions ($SF = 2$) while those in the surrounding semirural area are slightly high ($SF = 0.8$). If this spatial heterogeneity across the domain was ignored, the observational analysis for all upwind emissions would result in a smaller SF than that of city emissions. If this smaller SF is then applied to just the city, the AOI, the posterior city emissions would be biased low. On the other hand, if the upwind sources were unaccounted-for in the simulation domain, the measured enhancements that originated in those sources would be attributed to the AOI, causing the city emissions to be biased high instead. Worse still, flights are influenced by upwind sources to varying degrees, meaning that this bias would vary from flight to flight, increasing daily variability. To minimize these effects, we have developed an SF approach that is specific to an AOI by weighting modeled and measured enhancements by the fraction of the modeled enhancements originated by the emissions only in the AOI, in addition to accounting for the long-range background.

Here, we apply this SF approach to 9 nongrowing-season flights downwind of NYC, the details of which are

available in Supporting Information (SI) Table S1. We calculate posterior emission rates for 2 AOIs: the 5 NYC boroughs (Department of City Planning [DCP], n.d.) and the larger GNA (U.S. Census Bureau, 2021c), which includes NYC, and compare to an inversion analysis using the same AOIs (Pitt et al., 2022). These AOIs are clearly outlined in **Figure 1A and B**. Finally, we investigate multiple methods of calculating the SF and discuss the variability in the resulting posterior emission rates across SF calculation method, flight day, choice of meteorological model outputs (MET) used in dispersion modeling, and choice of prior. The different assumptions inherent in this method make it a good complement to more complex inversion approaches when the spatial distribution of prior emissions is sufficiently accurate to reproduce the measured plume structure.

2. Methods and materials

2.1. Flight design and equipment

All flights were conducted using Purdue University's Airborne Laboratory for Atmospheric Research (ALAR), a modified twin-engine Beechcraft Duchess (Cambahiza et al., 2014; Heimbigner et al., 2017). ALAR is equipped with a GPS/INS (Global Positioning System/Inertial Navigation System) system, a Best Air Turbulence probe for high precision 3-dimensional winds (Garman et al., 2006; Garman et al., 2008), and a Picarro Cavity Ring Down Spectrometer designed for measurements of CO_2 , methane (CH_4), and water vapor (H_2O ; Crosson, 2008), although this work focuses exclusively on CO_2 emissions. Measurements are reported as dry air mole fractions, in units of micromoles per mole of dry air, or parts per million (ppm). Each flight included multiple in situ 3-point calibrations using National Oceanic and Atmospheric Administration-certified standard cylinders of CO_2 and CH_4 (WMO- CO_2 -X2007, WMO- CH_4 -X2004A; Dlugokencky et al., 2005; Tans et al., 2017; see SI section S1 for detailed calibration information).

Over the course of the campaign, we used 2 different G2301-f analyzers and a G2401-m analyzer. The G2301-f analyzers were both run in low-flow mode, which measures each species every approximately 1.2 s and approximately 1.4 s, respectively. The G2401-m analyzer measured each species approximately every 2.3 s. We estimate the precision of the CO_2 measurements to be 0.04 ppm based on the average 1σ variability during calibrations.

This campaign involved 9 flights around NYC in February/March of 2019 and 2020 and November of 2018 and 2019 when the biosphere was less active. Details about the conditions of each flight are available in SI Table S1. The flights are discussed throughout the text using sequential research flight (RF) labels (e.g., RF4). **Figure 1A and B** shows the corresponding flight dates. Airborne measurements are sensitive to emissions distributed over a wide area relative to ground-based observations and can capture any vertical variability in CO_2 concentrations. All flights were conducted by flying multiple transects downwind of the AOIs at different heights within the boundary layer. Within a given flight, the flight track across all transects was kept as constant as possible. Most flights also

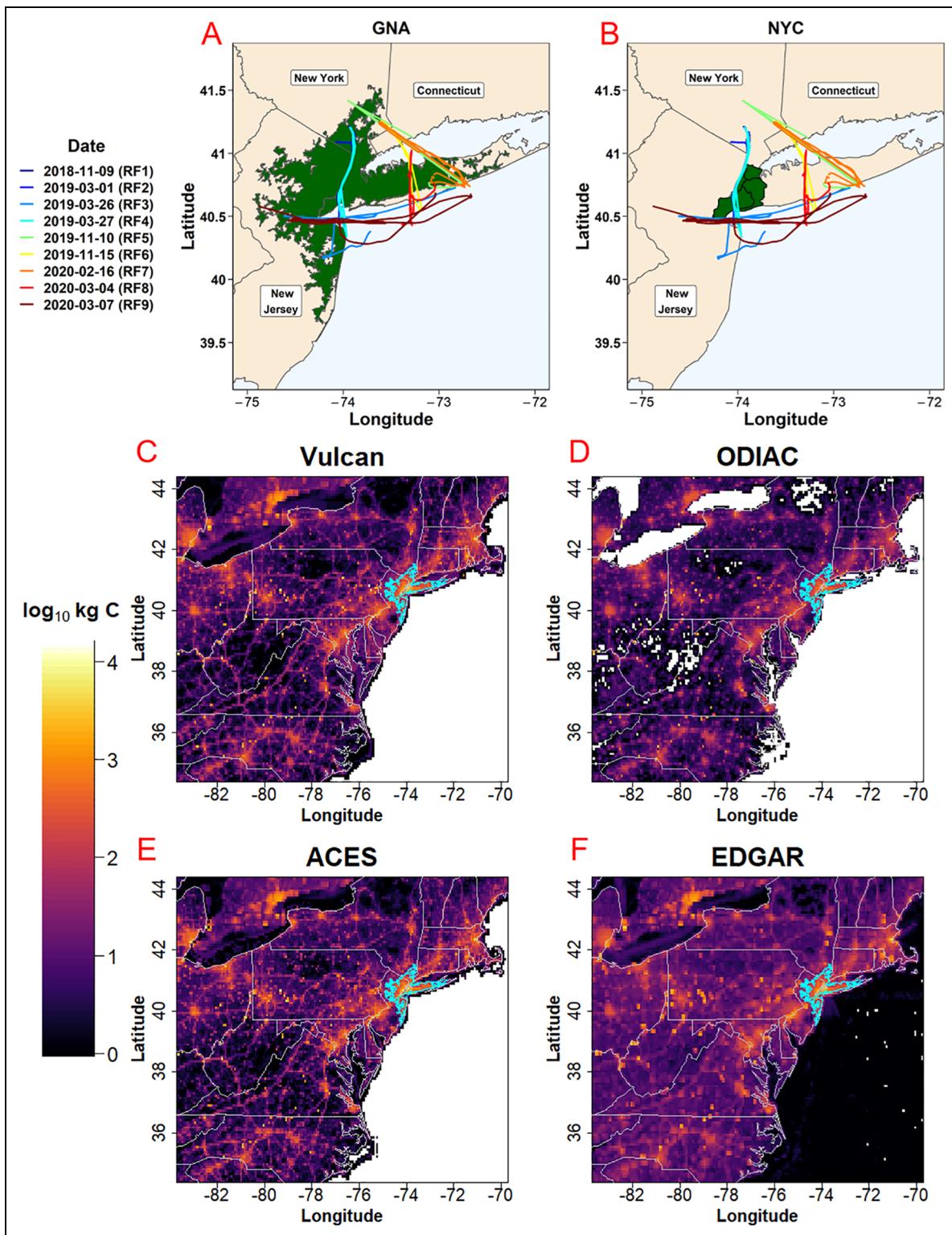


Figure 1. Flight tracks and carbon dioxide (CO₂) priors. (A and B) All flight tracks overlaid on a state map (U.S. Census Bureau, 2021a) with the areas of interest highlighted in green, (A) highlighting the Greater New York Area (GNA) and (B) highlighting New York City. November 9, 2018, March 1, 2019, and March 27, 2019, had similar downwind passes over the Hudson River and thus cannot be distinguished. Winds were from the east on November 9, 2018, March 1, 2019, and March 27, 2019, west on November 15, 2019, and March 4, 2020, southwest on November 10, 2019, and February 16, and north on March 26, 2019, and March 7, 2020. (C–F) Each CO₂ prior across the modeling domain overlaid on a state map, with the GNA outlined in blue. All priors are in log₁₀ (kg C/km² h) with the color scale saturated at log₁₀ (0) kg C/km² hr. Water areas differ as Vulcan and the Anthropogenic Carbon Emissions System are national and thus do not extend beyond the coast, Emission Database for Global Atmospheric Research is global and includes marine emissions, and Open-Source Data Inventory for Anthropogenic CO₂ does not include emissions over water. DOI: <https://doi.org/10.1525/elementa.2021.00121.f1>

Table 2. Names and acronyms for all ensemble members used. DOI: <https://doi.org/10.1525/elementa.2021.00121.t2>

Category	Acronym	Name
Meteorology model outputs (MET)	HRRR	High Resolution Rapid Refresh (National Oceanic and Atmospheric Administration, 2020)
	GFS	Global Forecast System (National Oceanic and Atmospheric Administration, 2020)
	NAMS	North American Mesoscale Forecast System (National Oceanic and Atmospheric Administration, 2020)
	ERA5	European Centre for Medium Range Weather Forecasts Fifth Reanalysis (Copernicus, 2020)
Turbulence parameterizations	KBLT = 2	Kantha–Clayson (Kantha & Clayson, 2000)
	KBLT = 5	Hanna (Hanna, 1982)
Priors	ACES–2017	Anthropogenic Carbon Emissions System, a 1-km ² bottom-up national prior (Gately & Hutyra, 2017)
	Vulcan–2015	A 1-km ² bottom-up national prior (Gurney et al., 2020a)
	EDGAR–2018	Emission Database for Global Atmospheric Research, an approximately 12-km ² disaggregated global prior (European Commission, 2019; Crippa et al., 2019)
	ODIAC–2018	Open-Source Data Inventory for Anthropogenic CO ₂ , a 1-km ² disaggregated global prior (Oda & Maksyutov, 2015)
SF calculation methods	I	Integral
	OLS	Ordinary least squares regression
	MA	Major axis regression
	B	Bayesian optimization
	2B	Bayesian optimization including a simultaneous linear background correction (slope and intercept)

Prior acronyms include the year that the data represent.

included an upwind pass flown just before/after the downwind passes to identify inflowing plumes of elevated GHG concentrations. Flight tracks for the downwind passes for all 9 flights are shown in **Figure 1A** and **B** with the 2 AOIs, NYC, and the GNA, highlighted. Complete flight tracks can be seen in Figure S1. The large urban areas of Philadelphia (39.95 N, 75.18 W) and Washington, DC–Baltimore (39.10 N, 76.85 W) are also clear in **Figure 1B–E** as emission hot spots to the southwest of NYC, making them important potential upwind sources when winds were from the southwest.

This work employs an ensemble approach to estimate the variability in the result introduced by the various inputs (transport models and emission inventories) as well as SF calculation methods and avoids relying entirely on any one particular member. We use a total of 4 meteorology data products (MET), 2 turbulence parameterizations in the dispersion model (making 8 different transport models), 4 bottom-up inventories as priors, and 5 SF calculation methods, all discussed in detail below and summarized in **Table 2**. This results in a total of 160 ensemble members per flight. The distributions of posterior emission rates across these 160 ensemble members times 9 RFs (1,440 individual posterior emission rates) are discussed in the results. Relying on the mean of a plausible group of, for example, MET, mitigates potential errors in any one particular MET while also allowing us to estimate

the contribution of MET uncertainty (quantified as the spread in the posterior emission rate across the 8 transport models) to the overall posterior emission rate uncertainty.

2.2. Transport modeling

The surface influence, or footprint, of flight data was determined using the Hybrid Single Particle Lagrangian Integrated Trajectory Model (HYSPPLIT) v 5.0.0 (Draxler & Hess, 1998; Stein et al., 2016; Loughner et al., 2021). A footprint, or influence function, represents the measurement's sensitivity to surface emissions, providing a quantitative link between measured enhancements in concentration and upwind surface fluxes. Each HYSPPLIT run was repeated using 4 different publicly available MET as an input and 2 different parameterizations of vertical velocity variance, Kantha–Clayson (Kantha & Clayson, 2000) and Hanna (Hanna, 1982), as described in **Table 2** (see Table S2 for the spatiotemporal resolution of each MET). Footprints were generated on a large modeling domain, bounded by 34.4 N, 44.4 N, 83.7 W, and 69.7 W at a resolution of 0.08 × 0.08, as shown in **Figure 1B–E**, to ensure that potentially significant upwind sources were included in the simulation.

Measured concentrations were block-averaged to a 1-min resolution for all flight data within the boundary layer. HYSPPLIT particle releases for each minute consisted

of approximately 1,000 particles distributed across the locations flown during that minute. Modeled anthropogenic CO₂ enhancements along the flight track were then simulated by multiplying footprints (ppm·($\frac{m^2}{\mu mol}$) or mole fraction/flux) from the HYSPLIT dispersion model (Draxler & Hess, 1998; Stein et al., 2016; Loughner et al., 2021) with emissions from one of the 4 priors (converted to units of $\frac{\mu mol}{m^2 s}$). Modeled biogenic CO₂ enhancements were then calculated in the same manner using the vegetation photosynthesis and respiration model (VPRM; Mahadevan et al., 2008; Gourdji et al., 2022) and added to modeled anthropogenic enhancements. This simulates the mixing ratio enhancement in CO₂ (ppm) relative to a long-range background value at the boundary of the domain, and this background must be subtracted from measured concentrations to compare the measured enhancements with these modeled enhancements.

2.2.1. Background calculations

We calculate the long-range background using a percentile method. We define times when we sampled background air using both measured and modeled data to mitigate the potential risk of selecting background times with large modeled enhancements in cases where the modeled plume is mislocated. First, we add the 1-min block-averaged modeled and measured data pointwise. Then, we identify the timestamps with concentrations between the first and fifth percentile of concentrations from this combined data set. These timestamps represent background air sampling for the data sets. The measured background is then defined as the average of all measured data during these timestamps, and the modeled background is the average of all modeled data during these timestamps. As the timestamps chosen to represent background are partially dependent on the modeled data, this process is done separately for every flight and for every unique ensemble member to yield slightly different backgrounds.

2.3. Prior emissions

We use 2 published bottom-up national priors, Anthropogenic Carbon Emissions System (ACES) version 2.0 (Gately & Hutyra, 2017) and Vulcan version 3.0 (Gurney et al., 2020b) as well as 2 global disaggregated priors, Open-Source Data Inventory for Anthropogenic CO₂ (ODIAC) version 2019 (Oda & Maksyutov, 2015), and Emission Database for Global Atmospheric Research (EDGAR) version 5 (Crippa et al., 2019; European Commission, 2019) to calculate modeled anthropogenic CO₂ emissions. Both ACES and Vulcan are hourly, 1-km² national bottom-up CO₂ priors. As bottom-up products, they both incorporate data from emission factors and activity data, modeling, fuel statistics, and so on to allocate emissions in a gridded field. They generally make use of similar methods and data sets, but there are still several methodological differences that can lead to some significant gridcell differences, as discussed in detail by Gurney et al. (2020b). For the analyses using Vulcan and ACES, Canadian emissions within the domain are taken from EDGAR since ACES and Vulcan do not extend beyond the United States. EDGAR (Janssens-Maenhout et al., 2017; Crippa et al., 2019; European

Commission, 2019) is a monthly, 0.1° (approximately 12 km² for the northeast United States) global prior that disaggregates national emissions using nationally specific activity data and emission factors combined with the appropriate proxies for different economic sectors (e.g., human population, road density). It also makes use of the Carbon Monitoring for Action database for point sources. Lastly, ODIAC (Oda & Maksyutov, 2015; Oda et al., 2018) is a monthly, 1 km², global prior that disaggregates the Carbon Dioxide Information Analysis Center national emissions primarily by using the Carbon Monitoring for Action database for point sources and night-light satellite data for nonpoint sources. For all analyses, we use annually averaged emissions for the most recent year available for each prior, which is indicated in **Table 2**. We also regressed all priors to match the extent and resolution of the domain using a conservative regridding scheme provided by the Python package xESMF (Zhuang, 2021).

Measured concentrations represent bulk emissions from all upwind sources, including both anthropogenic emissions and biogenic emissions, but all priors used only represent anthropogenic emissions. To incorporate biogenic CO₂ emissions into the modeled emissions, we also used a modified version of VPRM (Mahadevan et al., 2008), described in detail by Gourdji et al. (2022). VPRM is an empirical model optimized using flux tower data. It calculates net ecosystem exchange (respiration + photosynthesis) as a function of the enhanced vegetation index, land surface water index, temperature, and photosynthetically active radiation using a combination of in situ measurements and remote sensing. We used hourly fluxes for the dates of the flight data, generated at a 0.02° resolution.

2.4. Weighting approach

Figure 2 shows 3 examples of the measured and modeled enhancements, the modeled enhancements originating only from the AOIs, and the weighting functions. It is clear from **Figure 2 (A, C, and E)** that AOI and non-AOI enhancements show concurrent peaks in many cases. As previously mentioned, the long-range background subtracted from both measured and modeled values allows them to be directly compared but does not address sources upwind of the AOI overlapping with those of the AOI (e.g., Philadelphia in RF5). To address this problem, we have developed a weighting function to better isolate the enhancements of specific AOIs.

We calculate a weighting function for 2 different AOIs, the 5 NYC boroughs (including water areas; DCP) and the GNA (U.S. Census Bureau, 2021c), which includes NYC, as highlighted in **Figure 1B** and **A**, respectively. The weighting function is defined pointwise ($Weighting_i$) as the fraction of modeled enhancements, that is, the product of prior and footprint, from the AOI ($Modeled_{AOI_i}$) divided by the modeled enhancements for the entire domain ($Modeled_{total}$) as shown in Equation 1. As the weighting function is calculated using modeled enhancements, it relies somewhat more heavily on the modeling (MET and prior) than simply using a long-range background. This approach, like all scaling approaches, is also entirely

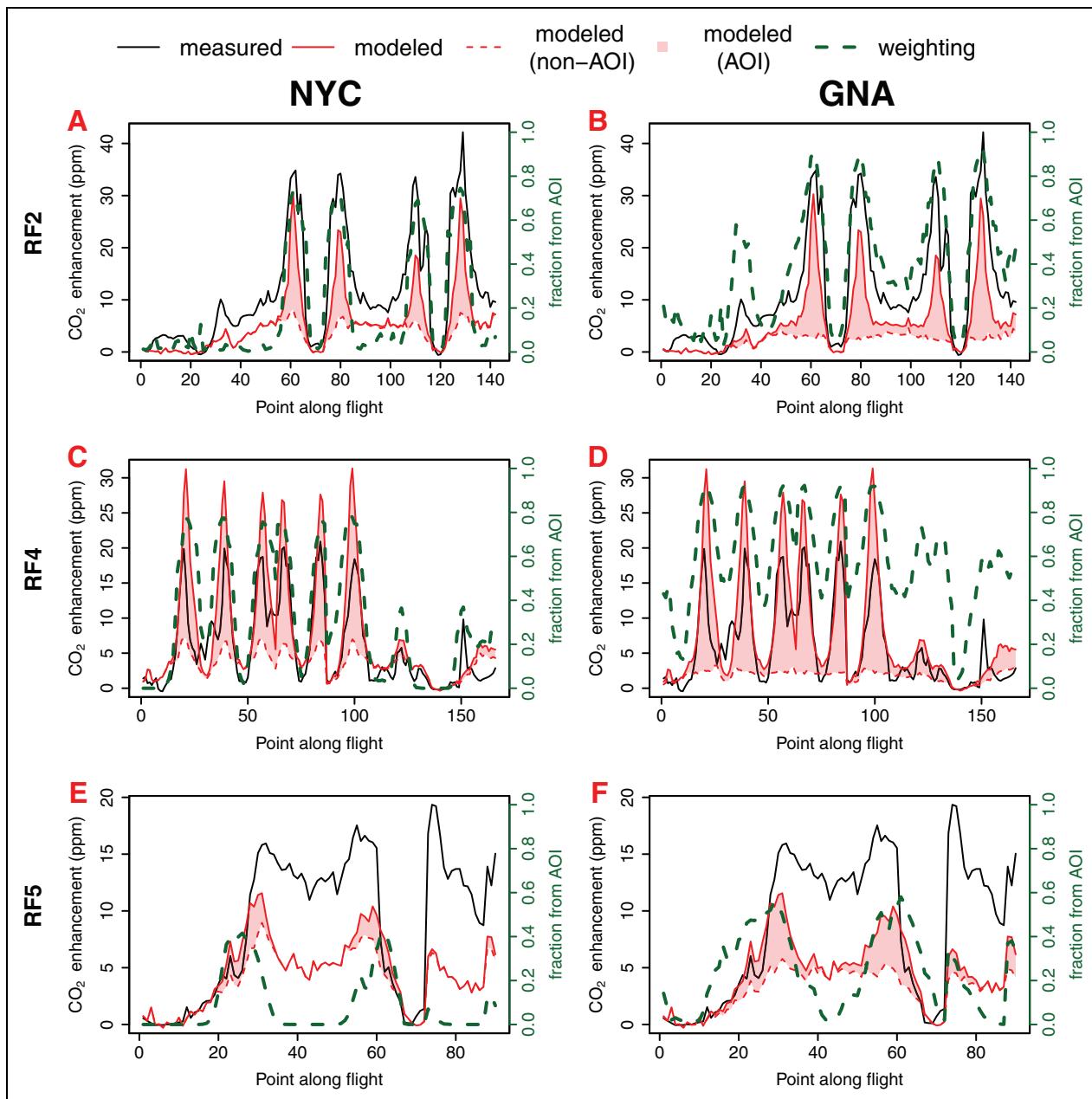


Figure 2. Example time series of measured carbon dioxide (CO₂) enhancements, modeled CO₂ enhancements, and weighting functions. Measured and modeled CO₂ enhancements over time for the flights on (A and B) March 1, 2019, (C and D) March 27, 2019, and (E and F) November 10, 2019, using global forecast system modeled winds and the Vulcan prior. (A, C, and E) Use New York City (NYC) as the area of interest (AOI) and (B, D, and F) use the Greater New York Area. Each valley–peak–valley represents a downwind transect covering the extent of the plume. Upwind data are included for the flights with upwind transects, typically with near 0 weighting for NYC. The shaded regions represent the modeled enhancements that originated from the AOI, red lines represent the modeled enhancements from all upwind sources, and the green dashed lines represent the weighting functions. The flight on November 10, 2019, has smaller weighting and a larger mismatch between measured and modeled enhancements because of the SW winds mixing significant Philadelphia emissions with those from the AOI. DOI: <https://doi.org/10.1525/elementa.2021.00121.f2>

reliant on the prior for the spatial distribution of emissions. Weights are then multiplied with total modeled enhancements, meaning the sum of AOI and non-AOI enhancements (i.e., the solid red line in **Figure 2**) and measured enhancements after subtracting their respective long-range backgrounds. This weighting allows us to minimize the effect of non-AOI sources before calculating an

SF and can be thought of as an additional background subtraction that is specific to the AOI.

$$\text{Weighting}_i = \frac{\text{Modeled}_{\text{AOI}_i}}{\text{Modeled}_{\text{total}_i}}. \quad (1)$$

Flights with significant influence upwind of the AOI, such as the flight on November 10, 2019 (RF5), would be

poorly suited to any SF approach without some way to isolate the AOI, as the measurements are heavily influenced by non-AOI emissions. In such cases, one likely cannot assume that the SF for the AOI is consistent with that for all upwind sources either, and therefore, the influence of these non-AOI emissions could limit the number of usable flights (to seven of the nine in our case) if it is unaccounted-for. However, with the weighting approach, these flights can still be used, as non-AOI emissions are deweighted appropriately to focus the SF primarily on the AOI. Weighting should also reduce the variability from flight to flight as the footprints are more comparable flight-to-flight without the influence of regions farther upwind.

2.5. SF approaches

We calculated SFs, which represent the factor by which the weighted modeled enhancements need to be multiplied to best match the weighted measured enhancements, using 5 methods. We calculated a single SF for each ensemble member for each flight (i.e., not per transect). The ensemble approach allows us to balance the benefits and drawbacks of each technique. For example, regression-based approaches would be more heavily impacted by a poor correlation between measured and modeled enhancements, which could occur if plumes are offset due to prior or transport model error. Bias correction methods would, on the other hand, be most impacted by background errors. Posterior emission totals for the AOIs were then calculated by multiplying these SFs by the total prior emissions within the AOIs, which can be seen in **Figure 1B–E** (see Figure S2 for the priors shown over the AOIs only).

2.5.1. Integral (I) SF method

The first method was an integral ratio SF (I). We integrated weighted enhancements across each flight for both measured and modeled enhancements. The ratio of these integrated enhancements gives an SF for that flight (Equation 2)

$$\text{Integral SF} = \frac{\sum(\text{Weight}_i[\text{CO}_{2\text{obs},i} - \text{BG}_{\text{obs}}])}{\sum(\text{Weight}_i[\text{CO}_{2\text{mod},i} - \text{BG}_{\text{mod}}])}. \quad (2)$$

Here, $\text{CO}_{2\text{obs}}$ and $\text{CO}_{2\text{mod}}$ represent the measured and modeled mole fractions, respectively, while BG_{obs} and BG_{mod} represent the measured and modeled long-range background mole fractions. Weight represents the weighting function and the index i represents the index of the 1 min measurements along the flight.

2.5.2. Regression (ordinary least squares [OLS] and major axis [MA]) SF methods

We also calculated SFs using an OLS and MA fit of weighted modeled enhancements (y) against weighted measured enhancements (x) across a flight. The slope of this fit is the reciprocal of the SF for this approach. This was done with the intercept forced to zero as this can partially mitigate the impact of poor correlation on the slope (see SI Section S2).

2.5.3. Bayes (B and 2B) SF methods

Lastly, we calculated SFs using 2 different approaches that rely on Bayes's Theorem, described more thoroughly in SI section S3. Broadly, these involve minimizing the differences between measured and modeled enhancements, similar to regression approaches, but including a regularization term that considers the uncertainties and prior knowledge of the SF, similar to an inversion. Optimum SFs were obtained by minimizing the cost function J (Enting, 2002; Tarantola, 2005):

$$J(\lambda) = \frac{1}{2}[(\lambda - \lambda_b)^T P_b^{-1}(\lambda - \lambda_b) + (H(\lambda) - y)^T R^{-1}(H(\lambda) - y)], \quad (3)$$

where λ is the state vector (i.e., SF), λ_b is the first guess or a priori state vector, P_b is the a priori error variance matrix, which represents the uncertainties in our a priori knowledge about the state vector, and R is the model-data error covariance matrix, which represents the uncertainties in the modeled enhancements $H(\lambda)$ and the observations y , also known as model-data mismatch. In our case, R is a diagonal covariance matrix that includes background uncertainty, measurement uncertainty, and transport modeling uncertainty added in quadrature.

The first approach (B) optimizes a single SF using the analytical solution to Equation 3. In the second approach (2B), 3 parameters are simultaneously optimized in the same manner. In this case, the first term is the SF, the second term is an offset to account for potential background bias (i.e., intercept), and the third term is the slope of a linear trend to account for spatial gradients in the background. In this work, we chose as prior SFs $\lambda_b = [1 \ 0 \ 0]$ with prior uncertainties $P_b = [1 \ 1 \ 0.1]$ for 2B and for B we use $\lambda_b = [1]$ and $P_b = [1]$.

3. Results and discussion

We report posterior anthropogenic emission rates based on the range of results across the various MET, SF methods, flight days, and priors using the standard deviation across posteriors, one category at a time. These provide information about the variability in posterior emission rates introduced by each of these aspects of the ensemble. We then present campaign-averaged emission rates for both AOIs as NYC represents a more policy-relevant area, while the wider GNA better represents the area of sensitivity for our flights. The campaign average, representative of daytime emissions during our flight days, is the average posterior for the AOI across all 160 ensemble members and 9 flights (for a total of 1,440 calculated posterior emission rates).

In this work, the biogenic fluxes from VPRM are used solely to ensure the best possible comparison between measured and modeled enhancements. The VPRM modeled emissions are generally small and, after background subtraction, have little effect on the modeled enhancements. However, they have a more significant effect on the weighting functions as the rural areas outside of the AOIs have more significant biogenic emissions, thus affecting the ratio between AOI and non-AOI emissions (i.e., the weighting functions). This means that ignoring

the biogenic contribution would have led to an overestimation of the AOIs emission rates. This can be seen in SI Figures S3 and S4.

3.1. Posterior emission rates

The posterior emission rates across the ensemble are shown in **Figure 3** and summarized in **Table 3**. Across all parameters, the campaign-averaged posterior emission rate for NYC is (49 ± 16) kmol/s (mean $\pm 1\sigma$), and for the GNA, it is (144 ± 44) kmol/s, with the bounds representing the 1σ variability across the 9 flights. We rely on the daily variability when discussing the campaign-averaged posterior emission rates because the daily variability incorporates both real variability in emissions and apparent variability caused by methodological sources, including those that are not captured by the ensemble spread, which are discussed in detail below.

The campaign-averaged posterior emission rates are larger than any of the prior emission rates (colored lines in **Figure 3**), but this is as expected given the priors are annual averages while all flights occurred during daytime hours in the nongrowing season when anthropogenic emissions tend to be larger. To more appropriately compare to the inventories, we use emission rates from the hourly ACES (year 2017) and Vulcan (year 2015) inventories for the dates and times that best represent the flights as was done in Pitt et al. (2022; abbreviated as ACI for the Anthropogenic Carbon Emissions System inventory and VUI for Vulcan in **Figure 3**, see SI section S4 for a detailed description). The average emission rates for these representative inventories are summarized in **Table 3**. For ACES, the average representative emission rates are (46 ± 9) kmol/s for NYC (7% less than our posterior estimate) and (145 ± 21) kmol/s for the GNA (<1% greater than our posterior estimate). For Vulcan, the average representative emission rates are (52 ± 10) kmol/s for NYC (6% greater than our posterior estimate) and (124 ± 20) kmol/s for the GNA (14% less than our posterior estimate). Bounds here represent the 1σ variability across 45 representative days. Agreement for both AOIs is within the daily variability of the representative inventories and well within the daily variability of the posterior emission rates. This level of agreement for both AOIs, even though the SFs differ by as much as 50%, supports the argument that using weighting functions helped us to separate AOI emissions from non-AOI emissions. Although it is difficult to directly compare these emission rates to the NYC self-reported inventory (SRI), a recent study by Gurney et al. (2021) compared annual 2015 Vulcan emissions with the 2015 NYC SRI, considering only source categories for which the SRI did not include any indirect emissions (i.e., Scope 2 or 3). That study found that the total NYC CO₂ emissions over these source categories in the SRI were 19% lower than the equivalent Vulcan estimates.

3.2. Sources of variability

The variability in the posterior emission rate with respect to the prior used is 14% for NYC and 8.1% for the GNA

(**Figure 3A and B**). This is an improvement from the variability across the prior emission rates, which show 44% variability in NYC and 35% variability in the GNA across the 4 priors, driven largely by ODIAC. If ODIAC is excluded, then the variability in the posterior emission rate drops to 6.6% for NYC and 7.9% for the GNA. This is still an improvement from the variability in the prior emission rates excluding ODIAC, which are 9.1% for NYC and 13.3% for the GNA. The variability across the transport models (MET and turbulence parameterizations) is 5.4% for NYC and 5.7% for the GNA (**Figure 3C and D**). Additionally, posterior emission rates using the vertical turbulence option based on Hanna (1982; HYSPLIT setting kblt = 5) are consistently lower than those based on Kantha–Clayson (Kantha & Clayson, 2000; HYSPLIT setting kblt = 2), with mean posterior emission rates that are, on average, 3.3% lower. This feature was seen in Pitt et al. (2022) as well. The variability across the 5 SF calculation methods is only 2.2% for NYC and 3.6% for the GNA (**Figure 3E and F**), making this the smallest source of variability. Finally, the daily posterior emission rate variability is 32% for NYC and 31% for the GNA (**Figure 3G and H**), making it a larger source of variability than all other terms combined in quadrature, that is, an estimate of the method's uncertainty (15% for NYC and 10% for the GNA).

These variabilities are summarized in **Table 3** and combined to provide an estimation of the overall variability for the posterior emission rates as in Lopez-Coto et al. (2020) and Pitt et al. (2022). Combining these terms assumes that they are independent sources of variability, which is unlikely to be true. Additionally, the choice of ensemble members (MET, priors, etc.) could impact the ensemble spread across these parameters. Therefore, the estimated variabilities may not represent the true uncertainties, but they do still provide an estimate of the likely variability introduced by the various model choices. If this SF approach is used over larger timescales to assess trends, then the daily variability is also a relevant part of the campaign-averaged posterior variability. This results in a combined variability more comparable to the uncertainties for individual mass balance experiment approaches (Cambaliza et al., 2014; O'Shea et al., 2014; Heimbürger et al., 2017; Ren et al., 2018; Ahn et al., 2020). It is worth noting that while the calculated daily variability is large (approximately 32%), the 95%, 2-sided confidence intervals ($t_{1-\alpha/2,N-1} \times s/\sqrt{N}$, $t_{0.975, 8} = 2.306$) of our campaign mean with 8 degrees of freedom ($N - 1$) are smaller (approximately 25%) and can potentially be further reduced by adding additional flights.

Real variability in the inventories themselves is significant, including a factor of ≥ 1.5 change between summer and winter and day-to-day variability of $\geq 15\%$ (see Figure S6). This intraannual variability alone is much larger than the annual reductions that policies for many cities, including NYC, would require. The daily variability in posterior emission rates is even larger than that of the representative days in ACES (19% for NYC and 15% for the GNA) or Vulcan (19% for NYC and 16% for the GNA). Potential reasons for this increased daily variability with the SF

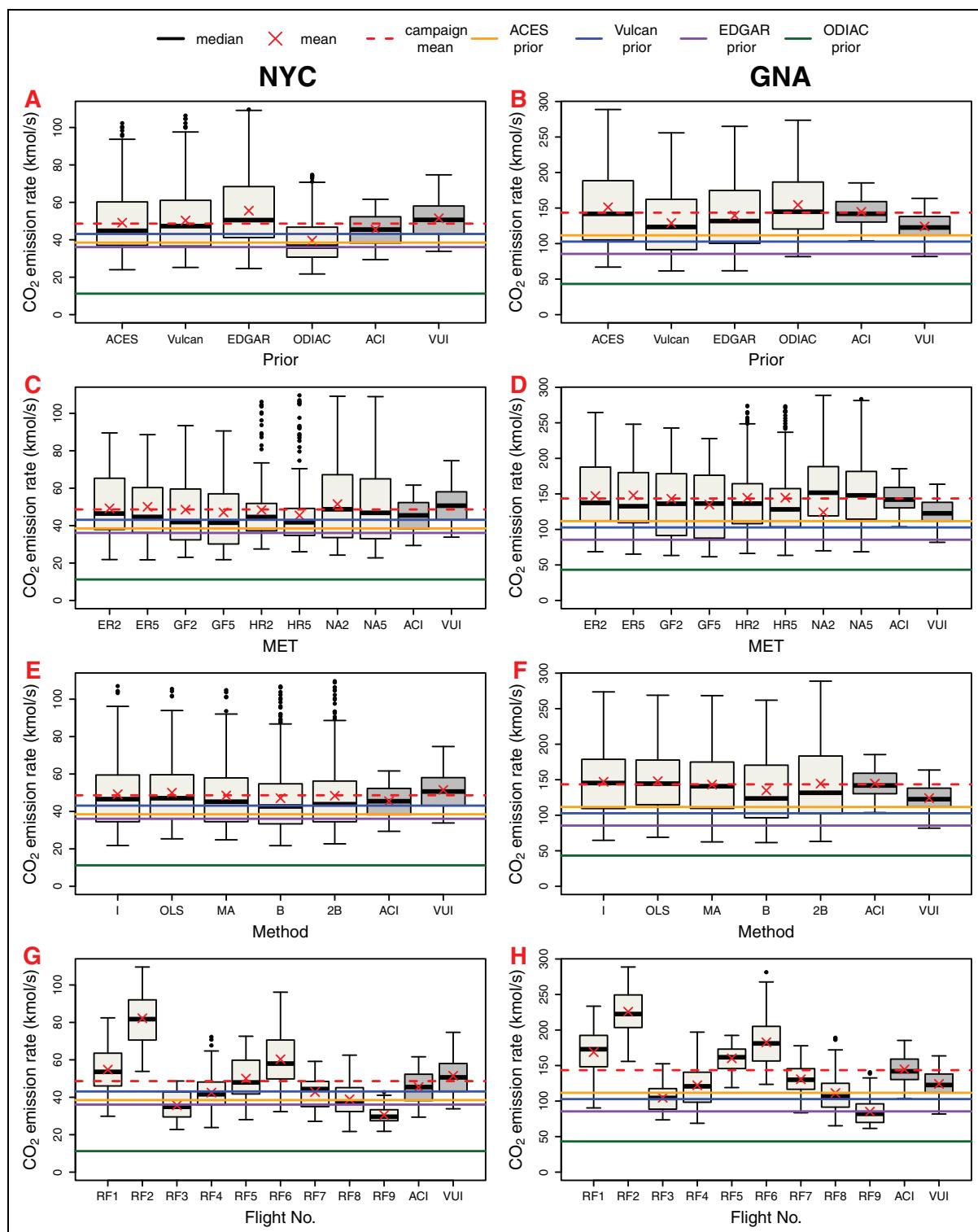


Figure 3. Boxplots of the posterior carbon dioxide (CO₂) emission rates for New York City (NYC) and the Greater New York Area (GNA). Boxplots of the ensemble of posterior CO₂ emission rates for NYC and the GNA, grouped by (A and B) prior, (C and D) meteorological model outputs (MET), (E and F) scaling factor calculation method, and (G and H) flight day. Boxes indicate the 25th–75th percentile, whiskers extend to the most extreme data point within 1.5 times the interquartile range of the box, and circles mark any data outside these bounds. ACI and VUI represent the Anthropogenic Carbon Emissions System and Vulcan inventories for similar days and times as the flights (see SI Section S4 for details). The MET are abbreviated as ER for the European Centre for Medium Range Weather Forecasts Fifth Reanalysis, GF for the Global Forecast System, HR for the High Resolution Rapid Refresh, and NA for the North American Mesoscale Forecast System, while the 2 and 5 represent the Kantha–Clayson (Kantha & Clayson, 2000) and Hanna (Hanna, 1982) turbulence parameterizations, respectively. DOI: <https://doi.org/10.1525/elementa.2021.00121.f3>

Table 3. Posterior emission rates and relative 1σ variabilities across each parameter. DOI: <https://doi.org/10.1525/elementa.2021.00121.t3>

Value	GNA				NYC			
	SF	Inversion	VUI	ACI	SF	Inversion	VUI	ACI
Posterior (kmol/s)	144	125	124	145	48.6	44.6	51.5	45.5
Daily variability	31% (44)	31% (39)	16% (20)	14% (21)	32% (16)	41% (18)	19% (9.8)	19% (8.8)
Prior variability	8.1% (12)	7.9% (9.9)	NA	NA	13.6% (6.6)	5.7% (2.5)	NA	NA
MET variability	5.7% (8.2)	3.8% (4.7)	NA	NA	5.4% (2.6)	4.9% (2.2)	NA	NA
SF method variability	3.6% (5.2)	NA	NA	NA	2.2% (1.1)	NA	NA	NA
Combined variability	33% (47)	33% (41)	NA	NA	35% (17)	41% (18)	NA	NA

The values in parentheses indicate the absolute 1σ variability in kmol/s. GNA = Greater New York Area; NYC = New York City; ACI = Anthropogenic Carbon Emissions System inventory; VUI = Vulcan inventory; SF = scaling factor; NA = not applicable.

approach are largely the same as previously discussed for inversions (Lopez-Coto et al., 2020; Pitt et al., 2022). The likely causes can be summarized as: (1) spatiotemporal aliasing, (2) day-to-day variability in transport model performance, (3) underestimated daily variability in inventories, and (4) interannual variability. The aliasing stems from the fact that each flight samples a somewhat different footprint, that is, is more sensitive to different parts of the AOI. Additionally, some sources have large hourly variability and are sampled at somewhat different times on different flights. These two effects cause spatiotemporal variability in emissions to appear as daily variability in emissions, as discussed in detail in Lopez-Coto et al. (2020). Increasing the number of simultaneous measurements was suggested by Lopez-Coto et al. (2020) as a way to at least partially alleviate this spatiotemporal aliasing. The transport model ensemble cannot account for poor performance across the ensemble and the performance will vary by day. The daily variability of emissions is underestimated by inventories because the temporal profiles of emissions used in inventories rely on some degree of interpolation or proxy data that can smooth out some temporal variability (making it akin to a climatology rather than a weather of emissions). The daily variability for the years that we flew also may have been larger than that of the inventory years, which are 2017 for ACES and 2015 for Vulcan. Lastly, the flights cover 2 winters (November 2018–March 2020), and some interannual variability in emissions may exist.

3.3. Comparison to inversion

These flights were also analyzed using a nested Bayesian inversion with similar ensemble members in Pitt et al. (2022). The data processing (long-range background, modeling parameters, etc.) of this work matches the base case of Pitt et al. (2022). This allows for a direct comparison of the results using these 2 different techniques as summarized in **Table 3**. The campaign-averaged posterior emission rate calculated in this work using an SF approach is 9.1% greater than the inversion-based campaign-averaged posterior emission rate for NYC and 15% greater for the GNA. This is well within the 1σ combined variability of

either method and both approaches show good agreement with representative inventory estimates from ACES and Vulcan, providing confidence in both approaches and in the inventories. Daily average emission rates are also highly correlated between methods, with a Pearson correlation that is over 0.9 for both AOIs. Additionally, the relative contribution of the different sources of variability is quite similar between the methods.

Interestingly, the SF approach shows more consistent relative daily variability (i.e., as a percentage) between the two AOIs than the inversion analysis, which exhibits increased relative variability for NYC. This increased relative variability for NYC is also observed in the representative inventories, although with lower values. However, in absolute terms (kmol/s), the SF approach shows larger variability for the GNA and almost the same variability for NYC. Both this and the larger campaign-averaged posterior emission rate of the SF approach are due to the different assumptions inherent to these approaches. Inversions directly optimize both AOI and non-AOI sources and can adjust the spatial distribution of the prior although they default to the prior value in areas that have no data with footprint influence. The SF approach with a weighting function applies a single SF to the prior AOI, so it assumes that the prior distribution is accurate (as does any scaling approach) and it can only separate AOI from non-AOI emissions by assuming the AOI: non-AOI emission ratio in the prior is accurate. Finally, although the ensemble constructed and methods used were different, Lopez-Coto et al. (2020) also saw similar component variabilities in a study using flight data in an ensemble of inversions for the Washington, DC–Baltimore metropolitan area. The authors saw 33% daily variability, 7%–15% variability across transport models (MET and turbulence parameterizations, including an experimental turbulence parameterization in HYSPLIT), and 11% variability across priors (including a flat prior with no spatial information).

4. Conclusions

The SF ensemble approach used here avoids many of the issues faced when applying a mass balance method or

nonweighted SF methods to calculate urban emissions, by directly relating emissions to a source area and prior using transport modeling. It is also simpler to apply than a full inverse modeling approach making it useful for broad applications. Since the approach relies on different underlying assumptions, it provides a good complement to these alternative methods and has been shown here to provide similar results to an inverse modeling study using the same data. Both approaches also agree well with inventory estimates for both AOIs. This level of agreement, especially given the GNA had SFs as much as 50% larger than NYC SFs, supports the claim that the weighting function is able to reduce the effect of non-AOI emissions. This case adds to the literature showing that CO₂ inventory estimates, as well as transport models, are generally able to represent measurements relatively well (Lauvaux et al., 2016; Sargent et al., 2018; Ahn et al., 2020; Lauvaux et al., 2020; Lopez-Coto et al., 2020). The variability due to meteorology and that due to the priors were both relatively large sources of posterior emission rate variability, which suggests that improvements in both transport modeling (both HYSPLIT and the meteorological models that feed it) and bottom-up inventories are necessary to further reduce uncertainties in quantification methods such as this. This approach also resulted in posterior emission rates that were broadly consistent across priors even though EDGAR is a 10× coarser global product and ODIAC was initially biased quite low for the region. However, the SFs are sensitive to a number of errors as shown by the previous work (Pitt et al., 2018; Sargent et al., 2018; Karion et al., 2019) as well as here and, thus, a careful model-data screening must be done to avoid large excursions in the estimations. Examples include most SF approaches used require a high correlation between modeled and measured enhancements, any SF approach requires an accurate spatial distribution of prior emissions, there must not be significant upwind sources outside the model domain, and sources outside an AOI must be considered.

The inclusion of the upwind sources in the simulation domain (i.e., using a large domain) combined with the weighting function allows this approach to appropriately differentiate between emissions from particular AOIs and emissions from farther upwind, addressing the last two concerns. This allows us to focus on emissions within political and legal boundaries, such as cities, that are of particular interest for emissions quantification and trend analysis. Additionally, the ensemble approach used here can partially minimize the influence of some of the other limitations, and thus, it is preferred as opposed to a single SF method, prior, and MET. Top-down approaches such as this allow us to compare/contrast results with developed emission inventories and quantify emissions in near-real time. Future work more thoroughly comparing this ensemble SF approach to inversion as well as mass balance methods would be of value as these top-down approaches continue to be employed and further refined.

Data accessibility statement

Aircraft measurement data used in this study are available from the Harvard Dataverse at <https://doi.org/10.7910/DVN/UUCQXX>.

Supplemental files

The supplemental PDF file for this article includes the following:

Table S1. Conditions and date of each flight.

Table S2. Spatiotemporal resolutions of the gridded archive meteorological model outputs (MET) files.

Table S3. Details of calibration cylinders used.

Figure S1. Flight tracks and time series of measured carbon dioxide (CO₂) enhancements, modeled CO₂ enhancements, and weighting functions.

Figure S2. Carbon dioxide priors zoomed to the Greater New York Area.

Figure S3. Boxplots of posterior carbon dioxide emission rates demonstrating the impact of biogenic emissions.

Figure S4. Example enhancement time series with and without biogenic emissions.

Figure S5. Example regression scaling factor plots showing the effect of correlation.

Figure S6. Boxplots of inventory emissions showing seasonal and diurnal profiles.

Text S1. Calibration cylinder details.

Text S2. The effect of correlation on regression scaling factors.

Text S3. Detailed Bayesian scaling factor methods.

Text S4. Representative inventory values.

Acknowledgments

The authors thank air traffic control in and around New York City for their support with our flight patterns and the Jonathan Amy Facility for Chemical Instrumentation at Purdue University for continued assistance in maintaining equipment in the Airborne Laboratory for Atmospheric Research.

Funding

We thank the National Institute of Standards and Technology (NIST) for partial support of this work, via Award Nos. 70NANB16H262 and 70NANB19H167. K. R. Gurney acknowledges the support from the National Aeronautics and Space Administration (NASA) grant NNX14AJ20G and the NASA Carbon Monitoring System program, Understanding User Needs for Carbon Information project (sub-contract 1491755).

Competing interests

The authors have no competing interests, as defined by Elementa, that might be perceived to influence the research presented in this article. Paul B. Shepson is an Associate Editor at Elementa. He was not involved in the review process of this article.

Author contributions

Contributed to flight measurements and flight design: KDH, CRF, ILC, JRP, TJ, RK, BHS, JT, PBS.

Contributed to acquisition and proper use of inventory data: CKG, KRG, LRH, GSR.

Contributed to analysis and interpretation of data: KDH, CRF, ILC, JRP, CKG, KRG, LRH, MS, PBS, SW.

Drafted and/or revised this article: KDH, CRF, ILC, JRP, CKG, KRG, LRH, GSR, MS, AJT, PBS, SW.

Approved the submitted version for publication: All Authors.

Disclaimer

Certain commercial equipment, instruments, or materials are identified in this article to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by National Institute of Standards and Technology nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.

References

- Ahn, DY, Hansford, JR, Howe, ST, Ren, XR, Salawitch, RJ, Zeng, N, Cohen, MD, Stunder, B, Salmon, OE, Shepson, PB, Gurney, KR, Oda, T, Lopez-Coto, I, Whetstone, J, Dickerson, RR.** 2020. Fluxes of atmospheric greenhouse-gases in Maryland (FLAGG-MD): Emissions of carbon dioxide in the Baltimore, MD-Washington, D.C. Area. *Journal of Geophysical Research: Atmospheres* **125**(9): e2019JD032004. DOI: <https://dx.doi.org/10.1029/2019JD032004>.
- Cambaliza, MOL, Shepson, PB, Caulton, DR, Stirm, B, Samarov, D, Gurney, KR, Turnbull, J, Davis, KJ, Possolo, A, Karion, A, Sweeney, C, Moser, B, Hendricks, A, Lauvaux, T, Mays, K, Whetstone, J, Huang, J, Razlivanova, I, Miles, NL, Richardson, SJ.** 2014. Assessment of uncertainties of an aircraft-based mass balance approach for quantifying urban greenhouse gas emissions. *Atmospheric Chemistry and Physics* **14**(17): 9029–9050. DOI: <https://dx.doi.org/10.5194/acp-14-9029-2014>.
- Copernicus.** 2020. Copernicus Climate Change Service (C3S). 2017. ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Available at <https://cds.climate.copernicus.eu/cdsapp>. Accessed 1 June 2020.
- Crippa, M, Oreggioni, G, Guzzardi, D, Muntean, M, Schaaf, E, Lo Vullo, E, Solazzo, E, Monforti-Ferrario, F, Olivier, J, Vignati, E, European Commission, Joint Research Centre.** 2019. *Fossil CO₂ and GHG emissions of all world countries: 2019 Report*. Publications Office of the European Union. Available at http://publications.europa.eu/publication/manifestation_identifier/PUB_KJNA29849ENN. Accessed 18 November 2019.
- Crosson, ER.** 2008. A cavity ring-down analyzer for measuring atmospheric levels of methane, carbon dioxide, and water vapor. *Applied Physics B* **92**(3): 403–408. DOI: <https://dx.doi.org/10.1007/s00340-008-3135-y>.
- Cventure, LLC, Pasion, C, Oyenuga, C, Gouin, K.** 2017. City of New York inventory of New York City's greenhouse gas emissions. *Mayor's Office of Sustainability*. Available at https://www1.nyc.gov/assets/sustainability/downloads/pdf/GHG_Inventory_2017.pdf. Accessed 4 February 2020.
- Department of City Planning.** n.d. NYC borough boundaries (water areas included). Available at <https://data.cityofnewyork.us/City-Government/Borough-Boundaries-Water-Areas-Included-/tv64-9x69>. Accessed 1 September 2020.
- Dlugokencky, EJ, Myers, RC, Lang, PM, Masarie, KA, Crotwell, AM, Thoning, KW, Hall, BD, Elkins, JW, Steele, LP.** 2005. Conversion of NOAA atmospheric dry air CH₄ mole fractions to a gravimetrically prepared standard scale. *Journal of Geophysical Research: Atmospheres* **110**(D18). DOI: <https://doi.org/10.1029/2005JD006035>.
- Draxler, R, Hess, G.** 1998. An overview of the HYSPLIT 4 modeling system for trajectories, dispersion, and deposition. *Australian Meteorological Magazine* **47**: 295–308.
- Enting, IG.** 2002. *Inverse problems in atmospheric constituent transport*. Cambridge, UK: Cambridge University Press. Available at <http://public.eblib.com/choice/publicfullrecord.aspx?p=217899>. 24 September 2020.
- European Commission.** 2019. EDGAR v5.0 global greenhouse gas emissions. Available at https://edgar.jrc.ec.europa.eu/overview.php?v=50_GHG. Accessed 18 November 2019.
- Feng, S, Lauvaux, T, Keller, K, Davis, KJ, Rayner, P, Oda, T, Gurney, KR.** 2019. A road map for improving the treatment of uncertainties in high-resolution regional carbon flux inverse estimates. *Geophysical Research Letters* **46**(22): 13461–13469. DOI: <https://dx.doi.org/10.1029/2019GL082987>.
- Floerchinger, C, Shepson, PB, Hajny, K, Daube, BC, Stirm, BH, Sweeney, C, Wofsy, SC.** 2021. Relative flux measurements of biogenic and natural gas-derived methane for seven U.S. Cities. *Elementa: Science of the Anthropocene* **9**(1). DOI: <https://dx.doi.org/10.1525/elementa.2021.000119>.
- Garman, KE, Hill, KA, Wyss, P, Carlsen, M, Zimmerman, JR, Stirm, BH, Carney, TQ, Santini, R, Shepson, PB.** 2006. An airborne and wind tunnel evaluation of a wind turbulence measurement system for aircraft-based flux measurements. *Journal of Atmospheric and Oceanic Technology* **23**(12): 1696–1708. DOI: <https://dx.doi.org/10.1175/JTECH1940.1>.
- Garman, KE, Wyss, P, Carlsen, M, Zimmerman, JR, Stirm, BH, Carney, TQ, Santini, R, Shepson, PB.** 2008. The contribution of variability of lift-induced upwash to the uncertainty in vertical winds determined from an aircraft platform. *Boundary-Layer Meteorology* **126**(3): 461–476. DOI: <https://dx.doi.org/10.1007/s10546-007-9237-y>.
- Gately, CK, Hutyra, LR.** 2017. Large uncertainties in urban-scale carbon emissions: Uncertainties in urban carbon emissions. *Journal of Geophysical*

- Research: Atmospheres* **122**(20): 11242–11260. DOI: <https://dx.doi.org/10.1002/2017JD027359>.
- Gourdji, SM, Karion, A, Lopez-Coto, I, Ghosh, S, Mueller, KL, Zhou, Y, Williams, CA, Baker, IT, Haynes, K, James, W.** 2022. A modified vegetation photosynthesis and respiration model (VPRM) for the eastern USA and CANADA, evaluated with comparison to atmospheric observations and other biospheric models. *Journal of Geophysical Research: Biogeosciences* **127**. DOI: <https://dx.doi.org/10.1029/2021JG006290>.
- Gourdji, SM, Yadav, V, Karion, A, Mueller, KL, Conley, S, Ryerson, T, Nehrkorn, T, Kort, EA.** 2018. Reducing errors in aircraft atmospheric inversion estimates of point-source emissions: The Aliso Canyon natural gas leak as a natural tracer experiment. *Environmental Research Letters* **13**(4): 045003. DOI: <https://dx.doi.org/10.1088/1748-9326/aab049>.
- Gurney, KR, Liang, J, O'Keffe, D, Patarasuk, R, Hutchins, M, Huang, J, Rao, P, Song, Y.** 2019. Comparison of global downscaled versus bottom-up fossil fuel CO₂ emissions at the urban scale in four U.S. urban areas. *Journal of Geophysical Research: Atmospheres* **124**(5): 2823–2840. DOI: <https://dx.doi.org/10.1029/2018JD028859>.
- Gurney, KR, Liang, J, Patarasuk, R, Song, Y, Huang, J, Roest, G.** 2020a. The Vulcan version 3.0 high-resolution fossil fuel CO₂ emissions for the United States. *Journal of Geophysical Research: Atmospheres* **125**(19): e2020JD032974. DOI: <https://dx.doi.org/10.1029/2020JD032974>.
- Gurney, KR, Liang, J, Roest, G, Song, Y, Mueller, K, Lauvaux, T.** 2021. Under-reporting of greenhouse gas emissions in U.S. cities. *Nature Communications* **12**(1): 553. DOI: <https://dx.doi.org/10.1038/s41467-020-20871-0>.
- Gurney, KR, Mendoza, DL, Zhou, Y, Fischer, ML, Miller, CC, Geethakumar, S, de la Rue du Can, S.** 2009. High resolution fossil fuel combustion CO₂ emission fluxes for the United States. *Environmental Science & Technology* **43**(14): 5535–5541. DOI: <https://dx.doi.org/10.1021/es900806c>.
- Gurney, KR, Razlivanov, I, Song, Y, Zhou, Y, Benes, B, Abdul-Massih, M.** 2012. Quantification of fossil fuel CO₂ emissions on the building/street scale for a large U.S. city. *Environmental Science & Technology* **46**(21): 12194–12202. DOI: <https://dx.doi.org/10.1021/es301128z>.
- Gurney, KR, Song, Y, Liang, J, Roest, G.** 2020b. Toward accurate, policy-relevant fossil fuel CO₂ emission landscapes. *Environmental Science & Technology* **54**(16): 9896–9907. DOI: <https://dx.doi.org/10.1021/acs.est.0c01175>.
- Hanna, SR.** 1982. Applications in air pollution modeling, in Nieuwstadt, FTM, van Dop, H eds., *Atmospheric turbulence and air pollution modelling*. Dordrecht, The Netherlands: Springer: 275–310.
- Heimbürger, AMF, Harvey, RM, Shepson, PB, Stirm, BH, Gore, C, Turnbull, J, Cambaliza, MOL, Salmon, OE, Kerlo, AEM, Lavoie, TN, Davis, KJ,**
- Lauvaux, T, Karion, A, Sweeney, C, Brewer, WA, Hardesty, RM, Gurney, KR.** 2017. Assessing the optimized precision of the aircraft mass balance method for measurement of urban greenhouse gas emission rates through averaging. *Elementa: Science of the Anthropocene* **5**(0): 26. DOI: <https://dx.doi.org/10.1525/elementa.134>.
- International Energy Agency.** 2013. *World energy outlook 2013*. Paris, France: OECD Publishing. Available at <https://www.iea.org/reports/world-energy-outlook-2013>. Accessed 4 February 2020.
- IPCC.** 2013. *Climate change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge University Press. Available at <https://www.ipcc.ch/report/ar5/wg1/>. Accessed 9 May 2019.
- Janssens-Maenhout, G, Crippa, M, Guizzardi, D, Muntean, M, Schaaf, E, Dentener, F, Bergamaschi, P, Pagliari, V.** 2017. EDGAR v4.3.2 Global Atlas of the three major greenhouse gas emissions for the period 1970–2012. *Earth System Science Data* **11**(3): 959–1002.
- Kaminsky, T.** 2019. Climate Leadership Community Protection Act. Available at <https://www.nysenate.gov/legislation/bills/2019/s6599>. Accessed 30 January 2020.
- Kantha, LH, Clayton CA.** 2000. *Small scale processes in geophysical fluid flows*. San Diego, CA: Academic Press.
- Karion, A, Lauvaux, T, Lopez Coto, I, Sweeney, C, Mueller, K, Gourdji, S, Angevine, W, Barkley, Z, Deng, A, Andrews, A, Stein, A, Whetstone, J.** 2019. Inter-comparison of atmospheric trace gas dispersion models: Barnett shale case study. *Atmospheric Chemistry and Physics* **19**(4): 2561–2576. DOI: <https://dx.doi.org/10.5194/acp-19-2561-2019>.
- Kunik, L, Mallia, DV, Gurney, KR, Mendoza, DL, Oda, T, Lin, JC.** 2019. Bayesian inverse estimation of urban CO₂ emissions: Results from a synthetic data simulation over Salt Lake City, UT. *Elementa: Science of the Anthropocene* **7**(1): 36. DOI: <https://dx.doi.org/10.1525/elementa.375>.
- Lamb, BK, Cambaliza, MOL, Davis, KJ, Edburg, SL, Ferrara, TW, Floerchinger, C, Heimbürger, AMF, Herndon, S, Lauvaux, T, Lavoie, T, Lyon, DR, Miles, N, Prasad, KR, Richardson, S, Roscioli, JR, Salmon, OE, Shepson, PB, Stirm, BH, Whetstone, J.** 2016. Direct and indirect measurements and modeling of methane emissions in Indianapolis, Indiana. *Environmental Science & Technology* **50**(16): 8910–8917. DOI: <https://dx.doi.org/10.1021/acs.est.6b01198>.
- Lauvaux, T, Gurney, KR, Miles, NL, Davis, KJ, Richardson, SJ, Deng, A, Nathan, BJ, Oda, T, Wang, JA, Hutyra, L, Turnbull, J.** 2020. Policy-relevant assessment of urban CO₂ emissions. *Environmental Science & Technology* **54**(16): 10237–10245. DOI: <https://dx.doi.org/10.1021/acs.est.0c00343>.

- Lauvaux, T, Miles, NL, Deng, A, Richardson, SJ, Cambaliza, MO, Davis, KJ, Gaudet, B, Gurney, KR, Huang, J, O'Keefe, D, Song, Y, Karion, A, Oda, T, Patarasuk, R, Razlivanov, I, Sarmiento, D, Shepson, P, Sweeney, C, Turnbull, J, Wu, K.** 2016. High-resolution atmospheric inversion of urban CO₂ emissions during the dormant season of the Indianapolis flux experiment (INFLUX). *Journal of Geophysical Research: Atmospheres* **121**(10): 5213–5236. DOI: <https://dx.doi.org/10.1002/2015JD024473>.
- Lopez-Coto, I, Ren, X, Salmon, OE, Karion, A, Shepson, PB, Dickerson, RR, Stein, A, Prasad, K, Whetstone, JR.** 2020. Wintertime CO₂, CH₄, and CO emissions estimation for the Washington, DC–Baltimore metropolitan area using an inverse modeling technique. *Environmental Science & Technology* **54**(5): 2606–2614. DOI: <https://dx.doi.org/10.1021/acs.est.9b06619>.
- Loughner, CP, Fasoli, B, Stein, AF, Lin, JC.** 2021. Incorporating features from the stochastic time-inverted Lagrangian transport (STILT) model into the hybrid single-particle Lagrangian integrated trajectory (HYSPIT) model: A unified dispersion model for time-forward and time-reversed applications. *Journal of Applied Meteorology and Climatology* **60**(6): 799–810. DOI: <https://dx.doi.org/10.1175/JAMC-D-20-0158.1>.
- Mahadevan, P, Wofsy, SC, Matross, DM, Xiao, X, Dunn, AL, Lin, JC, Gerbig, C, Munger, JW, Chow, VY, Gottlieb, EW.** 2008. A satellite-based biosphere parameterization for net ecosystem CO₂ exchange: Vegetation photosynthesis and respiration model (VPRM). *Global Biogeochemical Cycles* **22**(2). DOI: <https://dx.doi.org/10.1029/2006GB002735>.
- McKain, K, Down, A, Raciti, SM, Budney, J, Hutyra, LR, Floerchinger, C, Herndon, SC, Nehrkorn, T, Zahniser, MS, Jackson, RB, Phillips, N, Wofsy, SC.** 2015. Methane emissions from natural gas infrastructure and use in the urban region of Boston, Massachusetts. *Proceedings of the National Academy of Sciences* **112**(7): 1941. DOI: <https://dx.doi.org/10.1073/pnas.1416261112>.
- Michalak, AM, Bruhwiler, L, Tans, PP.** 2004. A geostatistical approach to surface flux estimation of atmospheric trace gases. *Journal of Geophysical Research: Atmospheres* **109**(D14). DOI: <https://dx.doi.org/10.1029/2003JD004422>.
- Mueller, KL, Gourdji, SM, Michalak, AM.** 2008. Global monthly averaged CO₂ fluxes recovered using a geostatistical inverse modeling approach: 1. Results using atmospheric measurements. *Journal of Geophysical Research: Atmospheres* **113**(D21). DOI: <https://doi.org/10.1029/2007JD009734>.
- Mueller, KL, Yadav, V, Lopez-Coto, I, Karion, A, Gourdji, S, Martin, C, Whetstone, J.** 2018. Siting background towers to characterize incoming air for urban greenhouse gas estimation: a case study in the Washington, DC/Baltimore area. *Journal of Geophysical Research: Atmospheres* **123**(5): 2910–2926. DOI: <https://dx.doi.org/10.1002/2017JD027364>.
- National Oceanic and Atmospheric Administration.** 2020. READY—Gridded data archives. Available at <https://ready.arl.noaa.gov/archives.php>. Accessed 4 February 2020.
- Neininger, BG, Kelly, BFJ, Hacker, JM, Lu, X, Schwietzke, S.** 2021. Coal seam gas industry methane emissions in the Surat Basin, Australia: Comparing airborne measurements with inventories. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **379**(2210): 20200458. DOI: <https://dx.doi.org/10.1098/rsta.2020.0458>.
- New York City Council.** 2019. Press release: Council to vote on Climate Mobilization Act ahead of earth day. Available at <https://council.nyc.gov/press/2019/04/18/1730/>. Accessed 4 February 2020.
- Oda, T, Maksyutov, S.** 2015. ODIAC fossil fuel CO₂ emissions dataset (ODIAC2019). *Center for Global Environmental Research, National Institute for Environmental Studies*. DOI: <https://dx.doi.org/10.17595/20170411.001>. Accessed 28 November 2019.
- Oda, T, Maksyutov, S, Andres, RJ.** 2018. The open-source data inventory for anthropogenic CO₂, version 2016 (ODIAC2016): A global monthly fossil fuel CO₂ gridded emissions data product for tracer transport simulations and surface flux inversions. *Earth System Science Data* **10**(1): 87–107. DOI: <https://dx.doi.org/10.5194/essd-10-87-2018>.
- O'Shea, SJ, Allen, G, Fleming, ZL, Bauguitte, SJB, Percival, CJ, Gallagher, MW, Lee, J, Helfter, C, Nemitz, E.** 2014. Area fluxes of carbon dioxide, methane, and carbon monoxide derived from airborne measurements around greater London: A case study during summer 2012. *Journal of Geophysical Research: Atmospheres* **119**(8): 4940–4952. DOI: <https://dx.doi.org/10.1002/2013JD021269>.
- Pitt, JR, Allen, G, Bauguitte, S, Gallagher, M, Lee, J, Drysdale, W, Nelson, B, Manning, A, Palmer, P.** 2018. Assessing London CO₂, CH₄ and CO emissions using aircraft measurements and dispersion modeling. *Atmospheric Chemistry and Physics Discussions* 1–22. DOI: <https://dx.doi.org/10.5194/acp-2018-1033>.
- Pitt, JR, Lopez-Coto, I, Hajny, KD, Tomlin, J, Kaeser, R, Jayarathne, T, Stirm, BH, Floerchinger, CR, Loughner, CP, Gately, CK, Hutyra, LR, Gurney, KR, Roest, GS, Liang, J, Gourdji, S, Karion, A, Whetstone, JR, Shepson, PB.** 2022. New York City greenhouse gas emissions estimated with inverse modeling of aircraft measurements. *Elementa: Science of the Anthropocene* **10**(1). DOI: <https://dx.doi.org/10.1525/elementa.2021.00082>.
- Plant, G, Kort, EA, Floerchinger, C, Gvakharia, A, Vimont, I, Sweeney, C.** 2019. Large fugitive methane emissions from urban centers along the U.S. East Coast. *Geophysical Research Letters* **46**(14):

- 8500–8507. DOI: <https://dx.doi.org/10.1029/2019GL082635>.
- Ren, X, Salmon, OE, Hansford, JR, Ahn, D, Hall, D, Benish, SE, Stratton, PR, He, H, Sahu, S, Grimes, C, Heimbürger, AMF, Martin, CR, Cohen, MD, Stunder, B, Salawitch, RJ, Ehrman, SH, Shepson, PB, Dickerson, RR.** 2018. Methane emissions from the Baltimore-Washington area based on airborne observations: Comparison to emissions inventories. *Journal of Geophysical Research: Atmospheres* **123**(16): 8869–8882. DOI: <https://dx.doi.org/10.1029/2018JD028851>.
- Sargent, M, Barrera, Y, Nehrkorn, T, Hutyra, LR, Gately, CK, Jones, T, McKain, K, Sweeney, C, Hegarty, J, Hardiman, B, Wang, JA, Wofsy, SC.** 2018. Anthropogenic and biogenic CO₂ fluxes in the Boston urban region. *Proceedings of the National Academy of Sciences* **115**(29): 7491. DOI: <https://dx.doi.org/10.1073/pnas.1803715115>.
- Sethi, M, Lamb, W, Minx, J, Creutzig, F.** 2020. Climate change mitigation in cities: A systematic scoping of case studies. *Environmental Research Letters* **15**(9): 093008. DOI: <https://doi.org/10.1088/1748-9326/ab99ff>.
- Sierra Club.** 2020. 100% Commitments in cities, counties, & states. Available at <https://www.sierraclub.org/ready-for-100/commitments>. Accessed 31 January 2020.
- Stein, A, Draxler, R, Rolph, G, Stunder, B, Cohen, M, Ngan, F.** 2016. NOAA's HYSPLIT atmospheric transport and dispersion modeling system. *Bulletin of the American Meteorological Society* **96**(12): 150504130527006. DOI: <https://dx.doi.org/10.1175/BAMS-D-14-00110.1>.
- Stohl, A, Seibert, P, Arduini, J, Eckhardt, S, Fraser, P, Greally, BR, Lunder, C, Maione, M, Mühle, J, O'Doherty, S, Prinn, RG, Reimann, S, Saito, T, Schmidbauer, N, Simmonds, PG, Vollmer, MK, Weiss, RF, Yokouchi, Y.** 2009. An analytical inversion method for determining regional and global emissions of greenhouse gases: Sensitivity studies and application to halocarbons. *Atmospheric Chemistry and Physics* **9**(5): 1597–1620. DOI: <https://doi.org/10.5194/acp-9-1597-2009>.
- Tans, PP, Crotwell, AM, Thoning, KW.** 2017. Abundances of isotopologues and calibration of CO₂ greenhouse gas measurements. *Atmospheric Measurement Techniques* **10**(7): 2669–2685. DOI: <https://dx.doi.org/10.5194/amt-10-2669-2017>.
- Tarantola, A.** 2005. *Inverse problem theory and methods for model parameter estimation*. Society for Industrial and Applied Mathematics (Other Titles in Applied Mathematics). DOI: <https://dx.doi.org/10.1137/1.9780898717921>.
- Trencher, G, Castán Broto, V, Takagi, T, Spriggins, Z, Nishida, Y, Yarime, M.** 2016. Innovative policy practices to advance building energy efficiency and retrofitting: Approaches, impacts and challenges in ten C40 cities. *Environmental Science & Policy* **66**: 353–365. DOI: <https://dx.doi.org/10.1016/j.envsci.2016.06.021>.
- Turnbull, JC, Karion, A, Davis, KJ, Lauvaux, T, Miles, NL, Richardson, SJ, Sweeney, C, McKain, K, Lehman, SJ, Gurney, KR, Patarasuk, R, Liang, J, Shepson, PB, Heimbürger, A, Harvey, R, Whetstone, J.** 2019. Synthesis of urban CO₂ emission estimates from multiple methods from the Indianapolis flux project (INFLUX). *Environmental Science & Technology* **53**(1): 287–295. DOI: <https://dx.doi.org/10.1021/acs.est.8b05552>.
- Turner, AJ, Kim, J, Fitzmaurice, H, Newman, C, Worthington, K, Chan, K, Wooldridge, PJ, Köehler, P, Frankenberg, C, Cohen, RC.** 2020. Observed impacts of COVID-19 on urban CO₂ emissions. *Geophysical Research Letters* **47**(22): e2020GL090037. DOI: <https://dx.doi.org/10.1029/2020GL090037>.
- United Nations Department of Economic and Social Affairs Population Division.** 2019. *World urbanization prospects: The 2018 revision*. New York, NY: United Nations.
- U.S. Census Bureau.** 1995. Urban and rural classification. Available at <https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural.html>. Accessed 20 May 2022.
- U.S. Census Bureau.** 2021a. Cartographic boundary files. Available at <https://www.census.gov/geographies/mapping-files/time-series/geo/cartographic-boundary.html>. Accessed 22 May 2022.
- U.S. Census Bureau.** 2021b. Counts and percentages of various characteristics of urban and rural areas in the United States, Puerto Rico and the Island Areas. Available at <https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural/ua-facts.html>. Accessed 7 September 2021.
- U.S. Census Bureau.** 2021c. TIGER/line shapefiles from the geography program. Available at <https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html>. Accessed 18 February 2021.
- Yadav, V, Ghosh, S, Mueller, K, Karion, A, Roest, G, Gourdji, SM, Lopez-Coto, I, Gurney, KR, Parazoo, N, Verhulst, KR, Kim, J, Prinzivalli, S, Fain, C, Nehrkorn, T, Mountain, M, Keeling, RF, Weiss, RF, Duren, R, Miller, CE, Whetstone, J.** 2021. The impact of COVID-19 on CO₂ emissions in the Los Angeles and Washington DC/Baltimore Metropolitan Areas. *Geophysical Research Letters* **48**(11): e2021GL092744. DOI: <https://dx.doi.org/10.1029/2021GL092744>.
- Zhuang, J.** 2021. xESMF: Universal regridder for geospatial data. DOI: <https://dx.doi.org/10.5281/zenodo.1134365>. Accessed 22 February 2021.

How to cite this article: Hajny, KD, Floerchinger, CR, Lopez-Coto, I, Pitt, JR, Gately, CK, Gurney, KR, Hutyra, LR, Jayarathne, T, Kaeser, R, Roest, GS, Sargent, M, Stirm, BH, Tomlin, J, Turner, AJ, Shepson, PB, Wofsy, S. 2022. A spatially explicit inventory scaling approach to estimate urban CO₂ emissions. *Elementa: Science of the Anthropocene* 10(1). DOI: <https://doi.org/10.1525/elementa.2021.00121>

Domain Editor-in-Chief: Detlev Helmig, Boulder AIR LLC, Boulder, CO, USA

Associate Editor: Andrea Pozzer, Max Planck Institute for Chemistry, Mainz, Germany

Knowledge Domain: Atmospheric Science

Published: June 28, 2022 **Accepted:** May 25, 2022 **Submitted:** December 20, 2021

Copyright: © 2022 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.