**High-resolution U.S. methane emissions inferred from an inversion of 2019 TROPOMI satellite data: contributions from states, urban areas, and landfills**

Hannah Nesser1, Daniel J. Jacob1, Joannes D. Maasakkers2, Alba Lorente2, Zichong Chen1, Xiao Lu3, Lu Shen4, Zhen Qu5, Tia Scarpelli6, Melissa P. Sulprizio1, Margaux Winter1, Shuang Ma7, A. Anthony Bloom8, John Worden8, Melissa Weitz9, William Irving9, Cynthia A. Randles10, Bryan K. Mignone10,

1School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA

2SRON Netherlands Institute for Space Research, Utrecht, the Netherlands

3School of Atmospheric Sciences, Sun Yat-sen University, Zhuhai, Guangdong Province, China

4Department of Atmospheric and Oceanic Sciences, School of Physics, Peking University, Beijing 100871, China

5Department of Marine, Earth and Atmospheric Sciences, North Carolina State University, Raleigh, NC, USA

6(Check)

7(Check)

8Jet Propulsion Laboratory, California Institute of Technology, Pasadena, California, USA

9Climate Change Division, Environmental Protection Agency, Washington, District of Columbia 20460, United States

10ExxonMobil Research and Engineering Company, Annandale, NJ, USA

We quantify 2019 methane emissions in the contiguous U.S. (CONUS) at 0.25° × 0.3125° resolution by inverse analysis of atmospheric methane columns measured by the Tropospheric Monitoring Instrument (TROPOMI). A gridded version of the U.S. Environmental Protection Agency (EPA) methane inventory serves as the prior estimate for the inversion. We optimize emissions and quantify observing system information content through analytical minimization of a Bayesian cost function. We achieve high resolution with a reduced-rank characterization of the observing system that preserves information content by optimizing emissions only where observing system errors are sufficiently low. The resulting inversion decreases computational cost by an order of magnitude compared to a full-rank inversion at the same resolution while optimizing 80% of anthropogenic emissions in CONUS. Our optimal (posterior) estimate of anthropogenic emissions in CONUS is 30.9 (30.0 - 31.8) Tg a-1, where the values in parentheses give the spread of an 8-member inversion ensemble. This is a 15% increase from the most recent EPA Greenhouse Gas Inventory (GHGI) estimate for 2019 of 26.8 Tg a-1. We partition emissions by sector, with livestock emissions of 10.4 Tg a-1, oil and natural gas of 10.4 Tg a-1, coal of 1.5 Tg a-1, landfills of 6.9 Tg a-1, and other sources of 1.1 Tg a-1. We find a significant 52% increase and 30% decrease from the GHGI for landfills and coal. We attribute the increase in landfill emissions to overestimated recovery efficiencies at landfill gas collection facilities, underestimated emissions from leaks and other anomalous events, and underestimated emission rates over time. We attribute the decrease in coal emissions to overestimated surface mine emissions. We exploit the high resolution of our inversion to quantify emissions in the 48 states in CONUS. We compare the results to EPA’s new state-level inventories as well as independent inventories produced by state agencies. Our posterior emissions are on average 38% larger than the EPA inventories in the largest 10 methane-producing states, with the largest upward adjustments in states with large oil and natural gas emissions and the largest downward adjustments in states with large contributions from coal emissions. We also calculate emissions for 95 geographically diverse urban areas. We find average urban area emissions increase from the gridded EPA emissions scaled to match the 2019 GHGI by 42% with a 66% increase in the largest 10 methane-producing cities compared to the 15% average correction over CONUS. We attribute the discrepancy to underestimated landfill and oil and natural gas distribution emissions.

**1 Introduction**

All modeled pathways that prevent global warming above 1.5°C require methane emissions reductions (IPCC SR5). The United Nations Framework Convention on Climate Change (UNFCCC) requires member parties to report anthropogenic methane emissions, which are produced mainly from oil and natural gas, livestock, landfills, coal, and wastewater. Wetlands are the dominant natural source of methane emissions (Bloom et al., xxxx). The bottom-up approaches used to generate these emission inventories use information on sectoral activity levels and emission factors, but considerable uncertainty can exist in these values. Top-down evaluations of bottom-up inventories use observations of atmospheric methane together with an atmospheric transport model to infer emissions through inverse analyses. These emission estimates are most useful if they achieve high spatial resolution and maximize the information content of the observation-model system. Here we use column methane observations from the TROPOMI satellite instrument in a reduced-rank analytical inversion to infer methane emissions and the associated information content at 0.25° × 0.3125° (≈25 × 25 km2) resolution over the contiguous U.S. (CONUS) for 2019, allowing detailed analysis of sectoral, state, and urban emissions.

Satellite observations of atmospheric methane column concentrations inferred from measurement of backscattered sunlight in the shortwave infrared have been used extensively in inverse analyses of methane emissions (Streets et al., 2013; Jacob et al., 2022). Past satellite instruments were limited by large pixel sizes (SCIAMCHY, 2003 – 2012) or sparse observations (GOSAT, 2009 – present). The Tropospheric Monitoring Instrument (TROPOMI) aboard the Sentinel-5 Precursor satellite provides daily, global observations of atmospheric methane columns at 5.5 km × 7 km nadir pixel resolution over land (Hu et al., 2018; Lorente et al., 2021) with a 3% success rate limited by cloud cover, optically dark surfaces, and heterogeneous terrain (Hasekamp et al., 2019). Inversions of TROPOMI data allow for high-resolution emission quantification but require understanding the actual information content of the observations.

Inverse analyses optimize methane emissions (the state vector) by fitting observations to simulated concentrations from a chemical transport model (CTM). This is typically done by minimizing a Bayesian cost function regularized by a prior emissions estimate given by a bottom-up inventory. When a linear relationship exists between emissions and concentrations, as in the case of methane, the optimal (posterior) solution and the associated errors and information content can be found analytically. This method also supports the generation of inversion ensembles. However, it requires the computationally expensive construction of the Jacobian matrix that represents the relationship between emissions and concentrations in the CTM. This matrix is typically constructed by conducting a CTM perturbation simulation for each emission element, limiting either the spatial resolution of the optimized emissions or the size of the inversion domain. In Nesser et al. (2021), we demonstrated an alternative method that approximates the Jacobian matrix by perturbing emission patterns that are informed by both the prior emissions and the observations. This approach optimally exploits the information content of the observations, allowing emissions quantification at the highest resolution possible by optimizing emissions only where the observing system provides a constraint and defaulting to the prior estimate elsewhere.

Quantifying methane emissions from the U.S. is of particular interest to validate bottom-up inventories. Several inverse studies of national emissions have found large discrepancies between bottom-up emissions estimates from multiple generations of the U.S. Environmental Protection Agency’s Greenhouse Gas Inventory (GHGI) and inferred emissions from satellite, aircraft, and in situ observations. Wecht et al. (2014) found livestock emissions 40% larger than the 2013 GHGI for the summer of 2004. Miller et al. (2013) inferred emissions 50% larger than the 2013 GHGI for 2007 to 2008, which they attributed to underestimated oil, natural gas and livestock emissions. Turner et al. (2015) found similar results for emissions from 2009 to 2011 compared to the 2014 GHGI. Maasakkers et al. (2021) inferred oil and natural gas emissions 35% and 22% higher than the 2020 GHGI, respectively, for 2010 to 2015. Lu et al. (2022) found mean 2010 – 2017 anthropogenic emissions 10.0 (6.5 - 11.8) Tg a-1 larger than the 2021 GHGI estimate for the same period, which they attributed largely to oil emissions. The coarse resolution of these studies complicated understanding the source of the discrepancy, particularly for livestock and landfill emissions.

Higher resolution regional studies have targeted specific aspects of U.S. emissions, including emissions from different sectors, states, and urban areas. Many focus on the oil and natural gas sector. Karion et al. (2015) found Barnett Shale emissions consistent with the GHGI but larger than those reported by industry to the EPA’s Greenhouse Gas Reporting Program (GHGRP). A series of studies inferred much higher emissions in the Permian Basin than in the Gridded EPA (GEPA) inventory, a spatially disaggregated version of the 2016 GHGI for 2012 (Zhang et al., 2020; Schneising et al., 2020; Liu et al., 2021; Shen et al., 2022; Chen et al., 2022; Varon et al., in review). And, an ensemble of inversions over North America oil and natural gas basins found emissions 80% larger than the 2021 GHGI from May 2018 to February 2020 (Shen et al., 2022). In studies of other sectors and regions, Chen et al. (2018) and Yu et al. (2021) found upper Midwest emissions that were an 80% and 24% increase, respectively, from GEPA, a discrepancy attributed mainly to livestock emissions. Jeong et al. (2016) inferred California emissions 20% to 80% larger than an inventory from the California Air Resources Board. And, Plant et al. (2019) inferred methane emissions from six East Coast urban areas more than two times larger than GEPA. While these regional studies provide insight into some of the sources of the observed discrepancy with EPA estimates, they are hindered by their limited domain.

Here we use the reduced-rank method of Nesser et al. (2021) in an analytical inversion of 2019 TROPOMI observations at 0.25° × 0.3125° resolution over North America using national emission inventories reported to the UNFCCC as prior estimates. Several filters are applied to the TROPOMI data to avoid biases. We focus our analysis on CONUS , with particular attention to emissions from landfills, states, and urban areas. We compare our results to the most recent version of the GHGI (EPA, 2022). This inventory includes for new emissions estimates for individual states, for which our inversion provides the first observational evaluation. We also compare our results to inventories prepared by individual states and cities.

**2** **Data and methods**

We conduct an ensemble of inversions of 2019 TROPOMI methane observations over the North American domain shown in Figure 1 (9.75°N - 60°N, 130°W - 60°W) using the GEOS-Chem CTM as forward model. The TROPOMI observations are fit to simulated GEOS-Chem concentrations to optimize mean methane emissions for 2019 at 0.25° × 0.3125° spatial resolution. This corresponds to state vector with dimension = 23691 including all grid cells with prior methane emissions larger than 0.1 Mg km-2 a-1, which represents over 99% of North American methane emissions. In a subset of the ensemble, we optimize boundary conditions corresponding to each of the four cardinal directions (north, south, east, and west) for the nested GEOS-Chem simulation. Methane chemical and soil sinks are not optimized because they are relatively uniform and slow compared to the ventilation timescale over the domain.

**2.1 Reduced-rank analytical inversion**

The inversion uses observed concentrations in a vector to optimize gridded emissions in the state vector assuming normal errors by minimizing a Bayesian cost function regularized by the prior emissions estimate :

The prior and observing system error covariance matrices and , respectively, are assumed diagonal in the absence of better information. The regularization factor corrects for the absence of covariance in (Chevallier et al., 2014). We generate an inversion ensemble using a range of prior errors and values to capture the inversion’s sensitivity to uncertainty in these parameters (section 2.7). The reduced-rank Jacobian matrix represents the sensitivity of concentrations to emissions in the CTM (Rodgers, 2000). To construct a rank Jacobian matrix for the 0.25° × 0.3125° GEOS-Chem grid at reduced computational cost, we perturb in the CTM the emission patterns that best capture the prior emissions and the information content of the TROPOMI observations as described in section 2.6.

Analytical minimization of the cost function following Rodgers (2000) yields the optimal (posterior) state vector estimate , error covariance matrix , and information content given by the averaging kernel matrix , which describes the sensitivity of the posterior estimate to the true state vector. The diagonal elements of are often referred to as averaging kernel sensitivities and their sum (trace of ) gives the degrees of freedom for signal (DOFS) that represent the number of pieces of information independently quantified by the observing system (Rodgers, 2000). Calculation of the Rodgers (2000) solution, which inverts an matrix that is a function of , produces numerical instabilities due to the rank reduction.

Here we use a reduced-rank approximation of the posterior solution following Bousserez and Henze (2018) to solve the inversion on an orthonormal basis that optimally spans the information content of the satellite–forward model observing system. The basis is given by the eigendecomposition of the prior-preconditioned Hessian of the cost function

where the columns of are the eigenvectors and is a diagonal matrix with entries equal to the eigenvalues. The calculation of requires significant memory for large and . We reduce memory usage using Dask, a Python parallelization package. The reduced-rank approximation is then generated using the largest eigenvalues and the associated eigenvectors :

Here, approximates the full-rank (FR) posterior by minimizing the difference between the two, and and are the optimal posterior error covariance and averaging kernel matrices for an inversion solved with a reduced-rank forward model. We choose to match the rank of the reduced-rank Jacobian matrix, which is chosen to maximize the DOFS within the available computational resources (section 2.8). The reduced-rank inversion and Jacobian matrix do not attempt to optimize emissions in areas with low information content, so we default to the prior estimate for grid cells with averaging kernel sensitivities less than 0.05.

**2.2 Prior estimates and errors**

Figure 1 shows the prior emissions estimates for different sectors. Anthropogenic emissions are given by the spatially disaggregated (gridded) versions of the EPA GHGI for the U.S. for 2012 (GEPA; Maasakkers et al., 2016), the INECC inventory for Mexico for 2015 (Scarpelli et al., 2020), and the ECCC inventory for Canada for 2018 (Scarpelli et al., 2021). We update GEPA oil and natural gas emissions to 2018 following Shen et al. (2022). Because GEPA is known to greatly underestimate oil and natural gas emissions in the Permian basin, the largest production basin in CONUS, we use the Environmental Defense Fund’s inventory for the Permian basin for 2019 (Zhang et al., 2020). We treat oil and natural gas as a single sector in our analysis due to significant source co-location and uncertainty in the partitioning of oil and gas wells. Anthropogenic emissions for Central America and the Caribbean islands are provided by the EDGAR v4.3.2 global emission inventory for 2012. Anthropogenic emissions are assumed aseasonal except for manure management and rice cultivation, for which we apply monthly scaling factors as described by Maasakkers et al. (2016) and Zhang et al. (2016), respectively.

Prior emissions for wetlands are given by the high-performance subset of 9 members of the WetCHARTs ensemble version 1.3.1 that best match global GOSAT inversion results (Ma et al. 2021). Lu et al. (2022) found in an inversion of GOSAT data over North America that the high-performance subset overestimated wetland methane emissions, particularly at high latitudes. We remove from the ensemble the two members (models 1923 and 2913) that are most responsible for this overestimate. Other natural methane emission sources are minor and include open fires, termites, and geological seeps, for which we follow the emissions described in Lu et al. (2022).

We assume uniform relative prior error standard deviations for emissions of between 50% and 100% with no error covariance between grid cells in our inversion ensemble. The 50% relative error lower bound follows previous inversions that optimized methane emissions over North America (Maasakkers et al., 2021; Lu et al., 2022). We inflate errors up to 100% to account for displacement errors and increased error covariance at high resolution (Maasakkers et al. 2016). Errors for each ensemble member are chosen as described in section 2.7.

**2.3 Forward model**

We use the nested version of the GEOS-Chem CTM 12.7 at 0.25° × 0.3125° resolution over North America as the forward model for the inversion. Earlier versions of the methane simulation were described by Wecht et al. (2014) and Turner et al. (2015). The model is driven by GEOS-FP meteorological fields from the NASA Global Modeling and Assimilation Office (GMAO). Methane loss from OH, Cl, soil uptake, and stratospheric oxidation is described in Maasakkers et al. (2019), but are unimportant in our domain because loss by ventilation is much faster. Initial conditions for January 1, 2019 and 3-hourly boundary conditions for the year are specified by methane concentration fields from a global GEOS-Chem simulation at 2° × 2.5° resolution using optimized emissions from a global inversion of TROPOMI observations (Qu et al., 2021). These concentration fields are unbiased with respect to the global TROPOMI data.

**2.4 TROPOMI observations**

TROPOMI has provided daily, global observations of dry column methane mixing ratios at 7 × 7 km2 nadir pixel resolution since May 2018 and at 5.5 × 7 km2 nadir pixel resolution since August 2019 (citation). TROPOMI measures backscattered solar radiation at 2.3 μm from a sun-synchronous orbit with a local overpass time of 13:30 (Veefkind et al. 2012). Methane concentrations are inferred from a full-physics retrieval that can fail due to cloud cover, variable topography, low or heterogeneous albedo, and high aerosol loading (citation). As a result, TROPOMI has a xx% retrieval success rate over North America for 2019. We use retrieval v14 as described by Lorente et al. (2021), which has a -3.4 ± 5.6 ppb bias relative to the Total Carbon Column Observing Network (TCCON). We use only high-quality retrievals as indicated by the quality assessment flag.

Previous analyses of TROPOMI data identified surface artifacts (Barré et al., 2021) and spatially variable biases relative to the more accurate but sparser GOSAT data (Qu et al., 2021; Chen et al., 2023). We filter the data to remove snow- and ice-covered scenes using blended albedo, an empirical parameter developed by Wunch et al. (2011) and suggested for the TROPOMI data by Lorente et al. (2021). We remove scenes with blended albedo > 0.75 in non-summer seasons. We also remove scenes with albedo in the shortwave infrared less than 0.05 following de Gouw et al. (2020), which account for most of the remaining unphysical TROPOMI observations (XCH4 < 1700 ppb), and scenes north of 50°N in winter.

Figure 2 shows the final 2919358 observations used for the inversion on the GEOS-Chem 0.25° × 0.3125° grid, and Figure S1 shows the improved correlation with seasonal averaged GOSAT observations on a 2° × 2° grid. The filters preserve 69% of the high-quality retrievals of TROPOMI v14 and increase the GOSAT – TROPOMI correlation in all seasons, with the largest increases in winter and spring. Seasonal regional biases decrease by between 7% and 21% and are always within the one standard deviation range of both the TROPOMI and GOSAT data. Comparison to a GEOS-Chem simulation driven by the prior emissions shows a mean GEOS-Chem – TROPOMI bias of = 9.1 ppb over North America. This bias can also be fit as a linear function of degrees latitude as . This bias and its latitudinal structure could be caused by errors in the boundary conditions. We correct it in our inversion ensemble members by removing either the continental mean bias or a latitude-dependent correction from the GEOS-Chem concentrations.

**2.5 Observing system errors**

The observing system error covariance matrix (equation 1), includes contributions from the forward model, the instrument, and representation error (Brasseur and Jacob, 2017). We calculate the error variances using the residual error method (Heald et al. 2004). This method assumes that the mean difference between the TROPOMI observations and the prior GEOS-Chem simulation, which we calculate on a seasonal 2° × 2° grid, is caused by errors in emissions that will be corrected by the inversion. The standard deviation of the residual errors after subtracting the mean gridded errors then gives the observational errors. We set a minimum error of 10 ppb, which applies to 32% observations. We find a mean observational error standard deviation of 11.5 ppb, with larger errors in winter and at high latitudes. The resulting variances are the diagonal elements of . Off-diagonal terms are assumed zero in the absence of better information, which we account for by introducing the regularization factor (Chevallier et al., 2007). We describe the choice of in section 2.7.

**2.6 Jacobian matrix**

Constructing the Jacobian matrix for our inversion would typically require conducting a 1-year perturbation simulation for each of the = 23691 grid cells optimized. This is computationally intractable. We construct the Jacobian matrix at substantially decreased computational cost using the reduced-rank method introduced by Nesser et al. (2021), which takes advantage of the heterogeneous information content of the TROPOMI observations. This method updates an initial, low-cost estimate of the Jacobian matrix by perturbing the patterns that best explain the information content of the observing system, constructing a reduced-rank Jacobian matrix while optimally preserving information content.

We construct the initial estimate of the Jacobian matrix using the mass-balance approach introduced by Nesser et al. (2021). We assume that a perturbation of methane emissions in grid cell *j* produces column mixing ratio enhancements in observation grid cell *i* according to

where is a dimensionless coefficient providing a crude representation of turbulent diffusion, and are the molecular weights of dry air and methane, respectively, is a ventilation length scale equal to the square root of the grid cell area, is gravitational acceleration, is the wind speed taken here as 5 km h-1, and is the surface pressure taken here as 1000 hPa. The use of produces off-diagonal structure in , which we found in Nesser et al. (2021) is necessary for an effective first estimate. We apply a simple isotropic turbulent diffusion scheme in which the influence of emissions spreads linearly to concentric rings of grid cells. This is represented as , where gives the absolute value of the difference in latitude or longitude grid cell index between and , 36 is the sum of values, and gives the number of grid cells in the corresponding concentric ring.

We use together with the error covariance matrices and to calculate the initial patterns of information content that are perturbed in the forward model. We calculate the prior pre-conditioned Hessian (equation 2) using and complete its eigendecomposition. The resulting matrix of eigenvectors is related to the patterns of information content via , which is equivalent to the eigenvectors of the averaging kernel matrix calculated with (Bousserez and Henze, 2018). We perturb the = 434 eigenvectors that capture 50% of the DOFS generated with . We then apply an optimal operator that restores the original state dimension and minimizes information content loss to yield an updated reduced-rank Jacobian matrix estimate . We then recompute the eigenvectors, perturb the = 1952 eigenvectors that explain 80% of the initial DOFS, and construct the final reduced-rank Jacobian matrix . This iterative update scheme optimizes the information content of the posterior solution while reducing the computational cost by an order of magnitude (Nesser et al., 2021).

**2.7 Inversion ensemble**

The posterior error covariance matrix that results from Bayesian optimization underestimates errors in the posterior solution by failing to account for systematic observational errors or errors in the inversion parameters (Houweling et al., 2014). We estimate the effect of these errors and reduce the noise generated by the reduced-rank Jacobian matrix by creating a quality-controlled ensemble of inversions, summarized in table 1. We conduct inversions that do and do not include a latitudinal correction to the model – observation difference and boundary condition corrections. For each inversion, we further choose the relative prior error (50%, 75%, or 100%) and regularization factor (between 0.175 and 0.5) so that the prior term of the cost function evaluated at the posterior solution averages to 1 across all grid cells optimized by the reduced-rank inversion, which is the expected value from the narrow chi-square distribution (Lu et al., 2021). This yields an ensemble of 8 quality-controlled inversions with indistinguishable validity. We report the mean posterior emissions for the ensemble, with uncertainty ranges given by the ensemble range.

**2.8 Source attribution**

The high resolution of the inversion facilitates the attribution of the posterior emissions estimates to different source categories including sectors, states, and urban areas. We aggregate the native resolution emissions and errors estimates to the corresponding quantities for each source category using a summation matrix , where is the number of source categories. The rows of are given by the relative contribution of each grid cell to a particular source category. For sectoral attribution, the rows are given by the relative contribution of each sector to a grid cell in the prior emissions estimate. For state and urban area attribution, the rows are given by the fraction of each grid cell within the state or urban area, respectively. The reduced-dimension posterior estimate, posterior covariance matrix, and averaging kernel matrix are then given by

where is the Moore-Penrose pseudo inverse (Caliesi et al., 2005). This approach to source attribution assumes that the prior fractional sectoral contributions are correct in each grid cell and that emission sources are evenly distributed in grid cells that cross state or urban lines. The high resolution of the inversion decreases the chance of source co-location or significant distributional errors across boundaries.

**3 Results and discussion**

Figure 3 shows the ensemble mean posterior correction factors relative to the prior emissions estimate (left) and the corresponding averaging kernel sensitivities (right). Grid cells unoptimized by any inversion (mean averaging kernel sensitivity < 0.05) are left blank. We find 772 (421 - 1279) DOFS across the domain, where the values in parentheses are the ensemble minimum and maximum, respectively. This represents a large increase in information content relative to past inversions over North America: Lu et al. (2022) found 114 DOFS in a joint inversion of data from GOSAT and NOAA’s ObsPack, while Shen et al. (2022) found 201 DOFS in an inversion of TROPOMI observations over 14 oil and natural gas basins. This increase reflects both the improved constraint provided by TROPOMI and the benefit of achieving 0.25° × 0.3125° resolution on the continental scale. Of these DOFS, 37 (15 - 69) are found in Canada, 641 (350 – 1058) in CONUS, and 86 (53 - 134) in Mexico. The high fraction of optimized emissions and DOFS in CONUS reflects the co-location of emissions (figure 1) with the highest density of TROPOMI observations (figure 2). As a result, we focus our discussion on CONUS, where we isolate anthropogenic emissions by removing contributions from wetlands and other natural sources following Section 2.8.

The mean posterior emissions when used in GEOS-Chem decreases the root mean squared error of the simulated observations with the TROPOMI data by xx% relative to the prior GEOS-Chem simulation. [Insert posterior evaluation.] We also compare the prior and mean posterior simulations to in situ surface and aircraft observations from the GLOBALVIEWplus CH4 ObsPack v(xx) database maintained by the National Oceanic and Atmospheric Administration (NOAA) Global Monitoring Laboratory (cite). [Insert independent posterior evaluation.]

**3.1 CONUS sectoral emissions**

We find posterior anthropogenic methane emissions of 30.9 (30.0 - 31.8) Tg a-1 in CONUS, a 15% (12% - 19%) increase from the most recent GHGI estimate for 2019 (henceforth “GHGI”) of 26.8 Tg a-1 (EPA, 2022). Lu et al. (2022) found larger anthropogenic emissions of 36.2 (32.1 - 37.6) Tg a-1 over the same domain for 2017 by optimizing emissions and trends in a joint inversion of GOSAT and in situ observations for 2010 - 2017. Deng et al. (2022) reviewed an ensemble of global inversions and found median U.S. posterior anthropogenic emissions for 2019 of 26.5 (20.8 - 38.7) Tg a-1 with GOSAT data and 31.9 (23.9 - 43.1) Tg a-1 with in situ data.

We allocate our national total to the emission sectors described in section 2.2 using the attribution method described in section 2.8. From the off-diagonal structure of (equation 8), we find very low posterior error correlation between the sectors (mean error correlation coefficients less than 0.2), indicating that we can separate sectoral emissions. Figure 4 and Table 2 summarize the results. Livestock, oil and natural gas, and landfills explain 90% of posterior anthropogenic emissions and all increase relative to the GHGI. We find a significant decrease from the GHGI only for coal. For these four sectors, we find sectoral averaging kernel sensitivities between 0.47 and 0.91, significantly larger than the values found by Lu et al. (2022) from GOSAT and in situ data. We find a small but significant increase in wetland emissions that is consistent with the large range found by Lu et al. (2022). However, the observing system only constrains about half of wetland emissions, with most of the inferred increase limited to the southeastern coast, including in South Carolina, Georgia, and eastern Florida. Here we focus on emissions from landfills, coal, and livestock, which are less studied than CONUS oil and gas.

Landfill emissions show the largest relative and absolute increase from the GHGI for 2019. We find posterior emissions of 6.9 (6.4 - 7.5) Tg a-1, an increase of 52% relative to the GHGI estimate of 4.5 (3.5 – 5.5) Tg a-1, where the parenthetical values represent EPA bottom-up errors Lu et al. (2022) found similar mean posterior landfill emissions of 7.5 (5.9 – 7.7) Tg a-1 for 2017. We attribute the discrepancy to three components of landfill inventory methodologies. First, for landfills with gas recovery systems, the GHGI relies on reported emissions from the GHGRP that use too-high collection efficiencies, particularly for large facilities. Second, the GHGI does not account for site-specific operation effects that may produce anomalous but significant emissions. Third, the GHGI assumes that emissions peak the year after waste is added to a landfill, while methane production rates have been shown to be relatively constant in time (Spokas et al., 2015; NAS, 2018; EPA, 2022). We discuss these causes in detail in Section 3.2.

Coal mining emissions of 1.5 (1.2 - 1.9) Tg a-1 exhibit the largest decrease in sectoral emissions relative to the GHGI estimate of 2.1 (2.0 - 2.5) Tg a-1. Lu et al. (2022) found much larger posterior emissions of 2.9 (2.3 - 3.4) Tg a-1 for 2017 that are consistent with GHGI estimates for 2012. Our lower estimate better reflects the 30% decrease in CONUS coal production from 2012 to 2019 (EIA, 202x), which is also shown in the 30% decrease in GHGI coal emissions over the same period. We attribute the discrepancy with the 2019 GHGI to overestimated surface mine emissions. Of the three largest coal-producing regions in CONUS, Appalachia produces 56% of coal from underground mines and 64% of posterior emissions, the Illinois Basin 30% and 20%, and the Powder River Basin 0% and a negligible fraction. If we decrease the GHGI estimate for surface mine emissions from 0.3 Tg a-1 to 0.1 Tg a-1, we find statistical agreement with our estimate. However, single-pass satellite observations have shown large surface mine emissions (e.g., Sadavarte et al., 2021; GHGSat, 2021). The persistence of such plumes is not documented and could explain the difference between these observations and our annual average.

Livestock emissions show broad agreement with the GHGI, with posterior emissions of 10.4 (10.0 - 10.7) Tg a-1 representing an insignificant 11% increase from the GHGI estimate of 9.4 (8.6 - 10.2) Tg a-1, where the parenthetical values represent the sum in quadrature of the EPA bottom-up errors for livestock subsectors. Lu et al. (2022) found similar mean posterior livestock emissions of 10.4 (8.8 – 11.6) Tg a-1 over CONUS for 2017. Yu et al. (2021) conducted a seasonal inversion of aircraft observations over the north central U.S. and south central Canada to find mean posterior livestock emissions of 5.5 (5.1 - 6.2) Tg a-1, which agrees with our posterior livestock estimate of 5.4 (5.2 - 5.6) Tg a-1 over the same region. Despite agreement with total GHGI livestock estimates, we find a significant increase in manure management emissions from 2.3 Tg a-1 to 3.1 Tg a-1, which would almost entirely explain the observed discrepancy between the mean GHGI and posterior emissions in our work, Lu et al. (2022), and Yu et al. (2021). The increase in manure management emissions is concentrated over the California Central Valley, northern Iowa, and Sampson and Duplin Counties in North Carolina. Notably, Iowa is the largest pork-producing state, and Sampson and Duplin Counties are the two largest pork-producing counties in CONUS (USDA, 2017). Due to variability in manure management practices, we find no correlation between our inferred increase and hog populations.

Posterior oil and natural gas emissions are 10.4 (10.1 - 10.7) Tg a-1, an insignificant 18% increase from the GHGI estimate of 8.8 (6.9 - 11.1) Tg a-1. Lu et al. (2022) found much larger posterior emissions of 4.8 (3.1 - 4.9) Tg a-1 for oil and 8.9 (8.0 - 9.8) Tg a-1 for natural gas in 2017. However, they found decreasing natural gas emissions beginning in 2014. Compared to Lu et al. (2022), Shen et al. (2022) found lower oil and natural gas emissions of 12.6 ± 2.1 Tg a-1 from an inversion of TROPOMI data over 14 basins extrapolated to the national scale for May 2018 to 2020, which is consistent with continued emissions decreases into 2019. These emissions are insignificantly larger than our posterior, and we find consistent basin-level results as shown in Figure S2. All posterior basin emissions are within 0.25 Tg a-1 of Shen et al. (2022) and all but five basins are within 0.10 Tg a-1. Of the basins where posterior emissions exceed the 0.5 Tg a-1 threshold defined by Shen et al. (2022) for successful quantification of basin emissions by TROPOMI, we find significant but still small differences only in the Delaware. We find total emissions in the Permian, including both the Delaware and Midland basins, of 2.9 Tg a-1, which is consistent with other recent studies (Zhang et al., 2020; Schneising et al., 2020; Liu et al., 2021; Varon et al., in review). Basin-level differences with Shen et al. (2022) could result from the different time periods examined or boundary condition treatments, while the difference in total emissions could result from the extrapolation from 14 basins to CONUS.

**3.2 Landfill emissions**

We find a significant 52% increase in landfill emissions relative to the GHGI (EPA, 2022) that we attribute to three emissions underestimates resulting from inventory methodologies. First, for landfills with gas collection systems, the GHGI uses too-large recovery efficiencies. Second, the GHGI does not account for site-specific operation effects that may produce large, often anomalous emissions that are significant for emissions totals. Third, the GHGI underestimates methane production rates over time.

The GHGI uses two methodologies to estimate landfill emissions (EPA, 2022). The first uses a first order decay model based on landfill mass so that emissions peak the year after waste deposition (IPCC, 2006). Recovered methane is removed from the total. However, a survey of 128 California landfills with gas recovery systems found that methane was produced at relatively constant rates over time (Spokas et al., 2015). The second method back-calculates emissions as a function of recovered methane given an estimated collection efficiency based on cover and operation methods. A default efficiency of 0.75 is assumed if cover information is unavailable (40 CFR § 98.343). GHGI estimates also incorporate GHGRP information from 1297 landfills emitting more than 1 Gg a-1 across the U.S., 500 of which have gas recovery systems (EPA LMOP, 2022). GHGRP methane emissions can be estimated with two of four methods depending on whether gas is collected (EPA, 2022; 40 CFR § 98.343). Landfills without gas collection estimate emissions based on landfill attributes and the GHGI first order decay model. Landfills with gas collection must use one of these methods in addition to a back-calculation approach. These methods have high uncertainties and have not been field validated (NAS, 2018).

We examine the source of the observed mismatch with the GHGI estimate by comparing our posterior landfill emissions to GHGRP facilities that report more than 5 Gg a-1 methane. We limit our attention to landfills where TROPOMI provides an averaging kernel sensitivity greater than 0.20. We exclude 13 facilities in grid cells containing multiple landfills because we are unable to separate the emissions. In the grid cells corresponding to the remaining 58 facilities, landfills explain 68 ± 22% of prior emissions (1 standard deviation range) so that we are confident of our ability to separate landfill emissions from other sources.

We validate the results by comparison to aircraft-derived estimates for four facilities. We find good agreement with three of the studies. At the Seneca Meadows Solid Waste Management Facility (NY), Catena et al. (2022) found 27.1 ± 6.2 Gg a-1 based on measurements from November 2021 compared to our posterior emissions of 23.9 (14.1 - 36.3) Gg a-1. Duren et al. (2019) found emissions of 16.8 ± 7.2 Gg a-1 at the Newby Island Landfill (CA) between 2016 and 2018 compared to our value of 15.2 (12.7 - 21.3) Gg a-1. They also found 10.2 ± 4.8 Gg a-1 at the Altamont Landfill and Resource Recovery Facility (CA) compared to our slightly larger but statistically comparable emissions of 17.0 (12.6 - 23.3) Gg a-1. By contrast, at the West Miramar Sanitary Landfill (CA), we find much larger emissions of 24.1 (22.1 - 25.4) Gg a-1 compared to a California Air Resources Board estimate of 13.7 ± 3.1 Gg a-1 in 2020 and a single flight from Duren et al. (2019) that estimated emissions of 7.7 ± 1.6 Gg a-1. The discrepancy may be explained by a November 2019 leak that produced up to 8662 ppm methane (Elmer, 2021).

Figure 5 compares our posterior landfill emissions to GHGRP values for these 58 facilities. We find larger mean facility emissions of 12.9 Gg a-1 compared to the GHGRP mean of 10.0 Gg a-1. We find a very low correlation (R2 = 0.04) between GHGRP emissions and our posterior that persists when we consider only facilities that do (32 landfills, R2 = 0.00) or do not (26 landfills, R2 = 0.12) capture landfill gas. Most of the variability in the posterior emissions occurs for facilities that report smaller (<10 Gg a-1) total emissions, suggesting that emissions from small landfills are less captured by GHGRP methodologies.

For the 32 facilities that recover gas, we use reported avoided methane emissions together with posterior and GHGRP emissions to calculate a posterior and reported recovery efficiency, respectively. We find a very similar mean posterior recovery rate of 0.47 compared to the GHGRP mean of 0.49. Both efficiencies are much smaller than the 0.75 default. Despite similar mean efficiencies, we find a weak correlation (R2 = 0.24), indicating that site-level variability is significant. For the 19 facilities that produce electricity, we compare the recovery rate difference to the duration of operation, facility capacity, and 2019 generation levels, as shown in Figure S2. We find that older or larger facilities tend to be less efficient than reported, while newer or smaller facilities tend to be more efficient. This suggests that efficiencies are overestimated for the facilities producing the most methane. Indeed, facilities with more than 5 MW capacity have mean posterior recovery efficiencies of 0.48 compared to the GHGRP value of 0.60.

We consider in detail the 24 facilities for which posterior emissions show a significant 50% difference from the GHGRP. We find larger emissions for 16 of these facilities. Many of these have a history of air quality or landfill standard violations, suggesting that anomalous events are significant for emissions estimation. At the West Miramar Sanitary Landfill (CA), where emissions of 8662 ppm were recorded in November 2019, we find emissions 3.8 (3.6 - 4.1) times larger than the GHGRP. The Sussex County Landfill (VA) was fined $99000 in 2016 for failing to address cracks in the landfill cover (citation). We find emissions 2.3 (1.6 - 3.4) times larger than reported. The Newby Island Landfill (CA), where our emissions are 2.8 (2.3 - 3.9) times larger than the GHGRP, received 30 violation notices from 2014 to 2020, including for gas leaks (citation). Finally, the Decatur-Morgan County Regional Municipal Solid Waste Landfill (AL) received an order in 2019 for failing to adequately cover new waste (citation). Our emissions are 1.7 (1.5 - 1.9) times larger than reported. The two facilities with the largest posterior - GHGRP differences experienced significant operational changes in the last decade, suggesting that the first order decay model is inadequate to methane emissions over time. Advanced Disposal Orchard Hills Landfill, Inc. (IL) completed a landfill gas recovery system in November 2016 (citation), and reported emissions peaked at 7.4 Gg a-1 in 2019. We find emissions of 39.6 (23.8 - 45.3) Gg a-1. And, the City of Dothan Sanitary Landfill (AL) has been full since 2014 (citation), when reported emissions peaked at 7.4 Gg a-1. We find emissions of 34.6 (28.2 - 42.8) Gg a-1.

Of the eight facilities for which we find smaller emissions than the GHGRP by 50%, seven have gas recovery systems. Four of the eight show a dramatic, likely unphysical decrease in methane emissions from 2019 to 2020 that is consistent with a change in the methodology used to calculate emissions. In three of these cases, the updated emissions are statistically consistent with our estimated emissions. Five of the facilities are also located in or near Appalachia, so that decreased emissions due to coal may be incorrectly attributed to landfill emissions.

**3.3 State emissions**

The EPA recently began to produce an annual inventory of U.S. anthropogenic greenhouse gas emissions disaggregated by state (henceforth “state GHGI”; EPA, 2022). The state GHGI is consistent with the national GHGI so that the same methods are used for emissions estimation and so that the total emissions are the same in both inventories. The EPA relies on existing datasets used by the national GHGI to disaggregate emissions estimates where possible. Elsewhere, proxy data such as population is used to spatially disaggregate the national totals. GHGRP data is included where available. State GHGI estimates are developed independently of state-level greenhouse gas inventories, which may result in discrepancies in sectoral or total values due to different in methods or accounting. The EPA lists state-level inventories for 24 states and Washington, D.C. Of these, we find that eight produce a methane emissions value separate from total greenhouse gases emissions. To our knowledge, there is no comprehensive observation-based evaluation of the state GHGI or independent state inventories.

We partition emissions, excluding offshore emissions, to each of the 48 states in CONUS as described in section 2.8 and compare the results to the state GHGI and to inventories prepared independently by state governments. Figure 6 shows state GHGI and posterior emissions for the 29 states responsible for 90% of posterior CONUS anthropogenic emissions excluding offshore emissions, and Table S1 shows the full results. TROPOMI provides a strong constraint at this resolution, with a state averaging kernel sensitivity larger than 0.5 in most states. The reduced-rank inversion also constrains emissions in most states, optimizing an average of 80% of prior emissions in each of the top 25 methane producing states.

We find a large average increase of 12% in state posterior emissions compared with the state GHGI estimates that is concentrated in the top 10 methane-producing states, where state posteriors are on average 38% larger. These states are responsible for 55% of posterior CONUS methane emissions. We find a significant emissions increase compared to the state GHGI in eight of these ten states. Oil and natural gas emissions have a disproportionate influence in these states, generating 47% of mean posterior emissions and 45% of the observed increase, compared to the national averages of 32% and 36%, respectively. In Texas, New Mexico, Louisiana, and Oklahoma, the oil and natural gas sector explains more than 60% of each state’s posterior emissions, with emissions increases concentrated in the Permian basin (Texas and New Mexico), the Haynesville Shale (Louisiana), and the Anadarko Shale (Oklahoma). Livestock and landfills also play a significant role in these eight states. Emissions in California and Iowa are dominated by the livestock sector, with much of the observed increase attributed to manure management emissions (Section 3.1). Landfills explain 41% of Illinois emissions and 62% of the Florida total. Indeed, Florida is home to three of the ten largest landfills as reported by the GHGRP in 2019 (EPA, 2022). Consistent with our sectoral analysis, the largest absolute negative posterior - state GHGI differences are found in historically coal-producing states, including Wyoming, West Virginia, and Pennsylvania.

We consider in more detail Texas and California, which are responsible for 21% and 6% of posterior CONUS anthropogenic emissions, respectively. Our posterior estimate for Texas is 6.3 (6.1 - 6.5) Tg a-1, a 69% increase from the state GHGI estimate of 3.7 Tg a-1 that is attributed almost entirely to the oil and natural gas sector. Indeed, oil and natural gas emissions account for 69% of posterior emissions compared to 56% in the state GHGI. The Permian basin alone explains almost 40% of Texas’ posterior emissions.

In California, we find posterior emissions of 2.1 (2.0 - 2.1) Tg a-1. Our posterior emissions are a 33% increase from the state GHGI estimate of 1.5 Tg a-1, and a 32% increase from the independent estimate prepared by the California Air Resource Board (CARB) of 1.6 Tg a-1. Our posterior estimate is smaller than but consistent with the estimate of 2.4 ± 0.5 Tg a-1 found by an inversion of in situ observations in California from June 2013 to May 2014 (Jeong et al., 2016). We find good agreement with the sectoral partitioning in the state GHGI, the CARB inventory, and Jeong et al. (2016). Livestock explain 54% of emissions in the posterior, 50% in the state GHGI, 54% in the CARB inventory, and 57% in Jeong et al. (2016), while landfills explain 25%, 23%, 21%, and xx% of emissions, respectively. We find slightly smaller relative contributions from oil and natural gas, which is responsible for 11% of posterior emissions compared to 20%, 17%, and 17% in the state GHGI, the CARB inventory, and Jeong et al. (2016), respectively. This partitioning differs from that found in an inversion of observations from the 2010 CalNex aircraft campaign, where 30% of emissions were attributed to livestock, 38% to landfills, and 22% to oil and natural gas (Wecht et al., 2014). The discrepancy could be due to temporal or seasonal differences, or from the coarser resolution (1/2° × 2/3°) of the CalNex inversion.

We also compare the distribution of posterior emissions in CARB’s air basins to that found by Jeong et al. (2016), as shown in Figure S3. Among basins with posterior emissions greater than 50 Gg a-1, we find good agreement in the San Joaquin Valley, Sacramento Valley, Mojave Desert, and Mountain Counties, which account for 73% of posterior emissions in California. We find lower emissions than Jeong et al. (2016) and the other studies listed in Figure S3 in the San Francisco Bay and South Coast air basins. This underestimate is consistent across the coastal air basins, suggesting that it may be due to errors in the boundary condition. While we correct the boundary condition as part of the inversion in five of eight ensemble members, we find a mean downward correction across these members of only 0.4 ppb. It is possible that higher spatial or temporal resolution would be necessary to fully correct the boundary condition.

We also compare our posterior emissions to available state greenhouse gas inventories for the year closest to 2019 where the observing system provides a strong constraint (state averaging kernel sensitivity greater than 0.5), including Pennsylvania, Louisiana, Iowa, and Colorado. Our posterior agrees only with the Pennsylvania estimate, where we find a source shift from fossil fuels (from 73% in the prior and 76% in the inventory to 63% in the posterior) to landfills and livestock (from 25% in the prior to 34% in the posterior). We find that Louisiana’s state inventory is too low due to underestimated oil and natural gas emissions, while Iowa’s is too low due to underestimated livestock emissions, particularly from manure management (Section 3.1). Colorado’s state inventory is 65% larger than our posterior due to oil and natural gas emissions that are more than twice as large.

**3.4 Urban area emissions**

Urban areas are home to 81% of the U.S. population (U.S. Census, 2010) and are major sources of greenhouse gas emissions, including methane (Gurney et al., 2015; Hopkins et al., 2016). The 280 urban areas with 2010 populations greater than one million are responsible for almost a quarter of GEPA emissions scaled to match the GHGI with post-meter emissions allocated by population (henceforth “scaled GEPA”). In an average city, urban emissions are produced by landfills (39%), natural gas distribution (6%), post-meter emissions (4%), and wastewater (6%). As urban populations grow (Seto et al., 2012), these emissions are likely to increase (Hopkins et al., 2016). Cities are therefore well-positioned to address methane emissions through waste-reduction initiatives, leak-detection programs, and strategic contracts with landfill operators and natural gas utilities. Regulation by air pollution control districts can also aid urban emissions reduction efforts while improving air quality (Hopkins et al., 2016). Indeed, C40, a performance-based coalition of over 100 mayors dedicated to climate mitigation, recommends that cities target a 50% reduction in methane emissions by 2030 (C40, 2022). Numerous cities, including New York City, Philadelphia, Washington, D.C., and San Francisco, are working toward similar reductions through zero-waste programs. And, the U.S. Methane Emissions Reduction Action Plan intends to work with local governments to set up methane monitoring systems to identify and publicize information about municipal gas distribution leaks. The plan also challenges members of the U.S. Climate Mayors to prioritize pipeline abandonment or replacement (The White House, 2022). However, mitigation efforts are complicated by considerable uncertainty in urban methane emissions estimates, with studies generally finding much larger emissions than GEPA (e.g., Sargent et al., 2021; Plant et al., 2019). Attributing observed discrepancies to emission sources is complicated by urban area variability (Plant et al., 2019; Floerchinger et al., 2021; Sargent et al., 2021).

We calculate posterior emissions for 95 urban areas across CONUS, providing the first comprehensive national analysis of urban emissions. We define urban areas using the U.S. Census Topographically Integrated Geographic Encoding and Referencing system (TIGER)/Line Urban Areas, which include all densely developed territory, for cities with 2010 populations greater than one million people (U.S. Census). Urban emissions depend significantly on the defined geographical extent due to the presence of large emitters such as landfills near city limits (e.g., Balashov et al., 2020; Plant et al., 2022). We follow Plant et al. (2022) and others in using the census designation to standardize the definition across CONUS. We attribute all emissions in any overlapping grid cell to the urban area to account for the distribution of emissions, include neighboring landfills, and improve comparability with regional emissions estimates from aircraft or tower observations. We limit our attention to the 95 urban areas where the inversion achieves averaging kernel sensitivities greater than 0.2, a relatively low threshold that accounts for the limited information provided by TROPOMI at this scale.

Urban anthropogenic posterior emissions in these 95 cities are 5.5 (5.0 - 6.1) Tg a-1, 37 (25 - 53) % larger than the scaled GEPA value of 4.0 Tg a-1. Individual urban area emissions, shown in Table S2, increase by an average 42 (30 - 56) %. These increases are much larger than the 15% (12% - 19%) increase we find in total anthropogenic emissions compared to the GHGI. They also represent a considerable 18% of CONUS emissions, equivalent to more than half of all CONUS oil and natural gas emissions. We are unable to attribute the emissions increase to a particular sector due to source co-location within urban areas. However, given that landfills explain 41% of prior emissions in an average urban area and increase 52% relative to the GHGI, it is likely that they are responsible for a large fraction of the observed discrepancy. It is also likely that oil and natural gas emissions, which represent 20% of prior emissions in an average urban area but explain between 32% and 100% of urban methane emissions (Plant et al., 2019; Floerchinger et al., 2021; Sargent et al., 2021), are significantly underestimated. The EPA added post-meter natural gas emissions of 457 Gg a-1 in their 2022 methane inventory, but this is insufficient to account for the observed discrepancy. For both landfill and oil and natural gas emissions, city-specific variability prevents further attribution of urban emissions. Indeed, we find no correlation between the posterior increase and urban area population, population change from 2000 to 2010, population density, or surface area.

We consider in detail the top 10 methane-producing cities as ranked by posterior emissions from landfills, natural gas distribution and post meter systems, and wastewater, as shown in Figure 7. These cities explain 38% (37% - 40%) of urban posterior emissions. We find a mean increase relative to scaled GEPA of 66% that is not explained by population-dependent factors such as post-meter emissions or landfilled mass. Indeed, the two largest cities by population (New York City and Los Angeles) have the smallest per capita methane emissions of these cities. We also compare our posterior emissions to inventories from New York City and Philadelphia, the only available methane emission estimates among these cities. Our emissions are more than twice as large, which likely results from differences in urban area boundaries.

We verify our results by comparison to 12 studies of these urban areas published since 2015, which we list in Figure 6. The studies estimated emissions in only a limited number of urban areas, with most focusing on New York City or Los Angeles. Almost all the studies used larger definitions of urban area extent, with only Pitt et al. (2022) and Plant et al. (2022) using the U.S. Census designation. Most studies used aircraft or tower observations to infer emissions by inverting a CTM (Cui et al., 2015; Jeong et al., 2016; Cusworth et al., 2020; Pitt et al., 2022; Yadav et al., 2019, 2023). Kuwayama et al. (2019) used a mass balance approach instead. Other studies used observed methane to CO2 or CO ratios together with bottom-up inventories of these gases (Wong et al., 2015; Wunch et al., 2016; Plant et al., 2019). Plant et al. (2022) used the same approach with TROPOMI methane to CO emissions.

On average, we find lower but statistically comparable emissions relative to these studies that likely results from our restrictive definition of urban area. The only surveyed study that used aircraft data to estimate emissions within a U.S. Census Urban Area found 314 ± 96 Gg a-1 in New York City, which is very similar to our estimate of 309 (241 - 417) Gg a-1 (Pitt et al., 2022). Plant et al. (2022) also used U.S. Census Urban Areas but relied on TROPOMI methane to CO ratios. They found slightly larger emissions in Atlanta and Philadelphia and much larger emissions in New York City, but their error bars spanned ranges that were almost twice as large as the derived emissions, limiting the utility of the comparison. Plant et al. (2019) also found larger emissions in New York and Philadelphia but used larger definitions of urban areas and produced similarly wide error ranges.

Los Angeles is the only city where we find statistical disagreement with previous studies. All surveyed studies quantified emissions in CARB’s South Coast air basin, and we find a smaller but still significant low bias when we use an equivalent extent as shown in Figure S3. Much of the remaining discrepancy results from decreasing emissions over the periods studied. The 2013 closure of the Puente Hills Landfill, one of the largest landfills in CONUS, decreased methane emissions significantly (CARB, 2018; Yadav et al., 2019), which is not fully reflected in the estimates of Cui et al. (2015), Wong et al. (2015), or Wunch et al. (2016). Yadav et al. (2023) found that Los Angeles emissions decreased an additional 7% from January 2015 to May 2020. Their posterior estimate of 251 ± 5 Gg a-1 for 2019 is still larger than our value of 179 (171 - 193) Gg a-1, which we attribute to a low bias in the western boundary condition that is not addressed at the spatiotemporal resolution of our boundary condition correction (section 3.3).

**4 Conclusions**

[Insert.]