**High-resolution North American methane emissions inferred from an inversion of 2019 TROPOMI satellite data**

Hannah Nesser1, Daniel J. Jacob1, Joannes D. Maasakkers2, Xiao Lu1,3, Lu Shen1,4, Zhen Qu1,5, Zichong Chen1, Tia Scarpelli6, Melissa P. Sulprizio1, Margaux Winter1, Jeremy S. Brandman7, Cynthia A. Randles7, Ashutosh Tewari7, A. Anthony Bloom8, Shuang Ma8, John Worden8…

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

6Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA, USA

7ExxonMobil Research and Engineering Company, Annandale, NJ, USA

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

We quantify 2019 North American methane emissions at 0.25° ⨉ 0.3125° resolution by inverse analysis of atmospheric methane columns measured by the Tropospheric Monitoring Instrument (TROPOMI). Prior methane emission estimates for the inversion are given by gridded versions of the national inventories reported by individual countries under the Paris Agreement in 2015 (Mexico), 2012 (the United States), and 2018 (Canada). We update United States oil and natural gas emission to match 2018 infrastructure and emissions. We generate a 15-member ensemble of improved emissions estimates through analytical minimization of a Bayesian cost function, yielding explicit quantification of the information content of the observing system. We achieve high resolution results with a reduced-rank characterization of the observing system that preserves information content in locations of high emissions and observational density while defaulting to the prior estimate in areas of weak emissions. Doing so decreases computational cost by an order of magnitude while effectively optimizing 70% of national emissions. We find the largest constraint in the contiguous U.S. (CONUS, where we find emissions increase 3.1 (1.5 - 4.7) Tg a-1 relative to the prior for a total of 31.4 (30.0 - 33.0) Tg a-1. We find the largest relative and absolute sectoral change in landfill emissions, reflecting the importance of our high-resolution optimization and the significance of urban emissions. We partition our optimized emissions to the 48 states in the contiguous U.S. (CONUS) and find the consolidation of emissions in the largest methane-producing states, with more than 50% of posterior emissions emitted by 7 states. We also find a consolidation of emissions in urban areas, where we find that our initial emissions estimate increases by 29% (16% - 44%). Although urban areas occupy less than 3% of surface area in CONUS, posterior emissions explain 8.3% (7.7% - 7.2%) of the national total. We also find a significant increase in livestock emissions that correlates with hog and dairy cattle operations, suggesting underestimated manure management emissions. The significance of our results over both point and area sources demonstrates the importance of achieving high resolution at continental scales and the potential to improve the attribution of inferred greenhouse gas emissions to sectors, nations, states, and urban areas.

**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 methane emissions. Bottom-up approaches used to generate these emission inventories use information on sectoral activity levels and emission factors, but there can be considerable uncertainty in this information. Top-down satellite observations of atmospheric methane can evaluate bottom-up inventories through inverse analyses relating the observed concentrations to emissions using an atmospheric transport model. This requires high spatial resolution as well as characterization of the errors and information content in the satellite observations. Here we use methane observations from the TROPOMI satellite instrument to infer methane emissions and the associated information content at 0.25° ⨉ 0.3125° (≈25 × 25 km2) resolution over North America 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). Previous satellite instruments were limited by low information content resulting from coarse pixel resolution (SCIAMCHY, 2003 – 2012) or sparse sampling (GOSAT, 2009 - present). TROPOMI now 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), though with a 3% retrieval success rate limited by cloud cover, optically dark surfaces, and heterogeneous terrain (Hasekamp et al., 2019). The TROPOMI data enable in principle high-resolution quantification of emissions but this requires understanding the actual information content of the observations.

Inverse analyses optimize methane emissions estimates by fitting observations to simulated concentrations from a chemical transport model (CTM). This is commonly done using Bayesian synthesis with prior estimates from a bottom-up inventory, and minimization of the Bayesian cost function to obtain optimized (posterior) emission estimates can be done either numerically or analytically (Jacob et al., 2016). Analytical solution provides closed-form posterior emission estimates and also enables easy generation of inversion ensembles, but it requires explicit computation of the Jacobian matrix relating emissions to concentrations in the CTM. This limits the number of emission elements (state vector) that can be optimized and hence the resolution of the inversion… relationship between emissions and concentrationsit , the optimal (posterior) solution can be found by numerically or analytically minimizing a Bayesian cost function regularized by a prior emissions estimate. The numerical (variational) approach uses gradient descent algorithms together with the adjoint of the CTM. In the case of inversions of satellite data where the number of observations greatly exceeds the number of optimized emissions elements, the computational cost of this approach scales less-than-exponentially with the resolution of the emissions estimate. The computational cost can still be prohibitive at high resolution due to long convergence times. Moreover, the ensemble or low-rank approaches used to estimate the information content of the observing system are computationally expensive and often incomplete. The analytical approach provides closed-form expressions for the posterior emissions and the associated error and information content. However, the computational cost scales more-than-exponentially with the resolution of the optimized emissions due to the cost of constructing the Jacobian matrix that characterizes the linear relationship between emissions and concentrations in the CTM.

Past inversions of satellite data achieved high-resolution at the sectoral and national level using variational approaches, by regridding coarse results, or by reducing the dimension of the optimized emissions. Wecht et al. (2014) used the GEOS-Chem adjoint to optimize methane emissions over North America at 1/2° ⨉ 2/3° resolution using data from the SCIAMACHY satellite, but they estimated errors using only 8 inversions. Deng et al. (2022) partitioned into national emissions 21 inversions from the global methane budget (Saunois et al., 2020), with emissions ranging in resolution from 2.5° ⨉ 2.5° to 6° ⨉ 4° downscaled to 1° ⨉ 1°, but this approach relied on prior information and may have suffered from biases along borders. Cusworth et al. (2021) derived and demonstrated an approach to optimally determine gridded sectoral emissions using the uncertainty structure of the prior estimate, but this approach is limited by the quality of the prior error estimate. Turner et al. (2015) solved for emissions on the basis of Gaussian functions. This technique has been applied in inversions of GOSAT data over North America (Turner et al., 2015; Maasakkers et al., 2021; Lu et al., 2022) and China (Chen et al., 2022) but is limited in its ability to resolve unidentified emission hotspots because of the aggregation scheme’s reliance on the initial estimate. Nesser et al. (2021) proposed solving for emissions along the patterns of maximum information content, but the method has never been implemented in practice. Other inversions reduced the dimension of the emissions by vector by optimizing emissions over selected source regions instead (e.g., Yu et al., 2021; Shen et al., 2022; Plant et al., 2022), but these studies are also limited by prior knowledge of source locations and by their ability to accurately capture source statistics with a small sample size.

High-resolution inversions of TROPOMI satellite observations offer the opportunity to optimize national, regional, and sectoral methane emissions at high resolution. We used TROPOMI observations over North America for 2019 to optimize emissions and information content at 0.25° ⨉ 0.3125° resolution where the observing system provides a constraint (sections 2.1 and 2.6; Nesser et al., 2021). Our prior emission estimate is provided by gridded versions of the inventories prepared by Mexico, the United States, and Canada under the Paris Agreement so that our corrections are policy relevant (section 2.2). We used cleaned TROPOMI data that avoids the largest biases with respect to albedo and snow- and ice-cover (section 2.4) to generate a 15-member, quality-controlled inversion ensemble that improves our estimation of errors (section 2.7). The high-resolution of the results improves our ability to accurately partition the emissions to sectors and source regions (section 2.8), which we demonstrate in our analysis of sectors (section 3.1), states (section 3.2), urban areas (section 3.3), and livestock (section 3.4).

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

We conduct an inversion of 2019 TROPOMI methane observations over North America. The TROPOMI observations are fit to simulated concentrations from the GEOS-Chem chemical transport model (CTM, [www.geos-chem.org](http://www.geos-chem.org)) to optimize mean methane emissions at 0.25° ⨉ 0.3125° spatial resolution. We calculate the optimal emissions and the associated error covariance and information content by finding the analytical minimum of a Bayesian cost function regularized by a prior emissions estimate (section 2.1). Sections 2.2 through 2.6 describe the components of the inversion: section 2.2 describes the state vector, prior emissions, and prior errors; section 2.3 describes GEOS-Chem; section 2.4 describes TROPOMI observations; section 2.5 describes the observing system errors; and section 2.6 describes the novel, reduced-rank method used to calculate the Jacobian matrix. Section 2.7 summarizes the inversion ensemble. Section 2.8 describes the method used to attribute the posterior emissions to sectors, states, and urban areas.

**2.1 Reduced-rank analytical inversion**

We optimize the state vector of gridded emissions assuming normal errors by minimizing a Bayesian cost function

where and are the prior emissions estimate and error covariance matrix, respectively (section 2.4); and are the vector of observations and the error covariance matrix, respectively (section 2.5); is the chemical transport model (CTM) that simulates observations as a function of emissions (section 2.2); and is a regularization factor that accounts for the absence of covariance in (section 2.4) (Brasseur and Jacob, 2017). The nested methane CTM is linear so that where is the Jacobian matrix and is constant. Analytical minimization of the cost function then 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 **A** are commonly referred to as averaging kernel sensitivities and their sum (trace of **A**) gives the degrees of freedom for signal (DOFS), the number of pieces of information independently constrained by the observing system (Rodgers, 2000). We generate an inversion ensemble using a range of prior emissions, prior errors, observations, and observing system errors, giving an additional estimate of inversion error (section 2.7).

The standard analytical solution is numerically unstable for large since it requires inverting a non-sparse matrix. We use reduced-rank approximations that solve the inversion on an orthonormal basis that optimally spans the information content of the satellite–forward model observing system (Bousserez and Henze, 2018). The basis is given by the eigendecomposition of the prior-preconditioned Hessian

where the columns of are the eigenvectors and the diagonal of gives the eigenvalues. Then,

Here, is the full-rank approximation that minimizes error relative to the standard solution; and are the posterior error covariance matrix and averaging kernel matrix associated with the reduced-rank Jacobian matrix (section 2.6), respectively; is the matrix of the first eigenvectors; and is a diagonal matrix containing the largest eigenvalues. We choose = 1952 to match the rank of the Jacobian matrix. To reduce errors associated with the rank reduction, we filter out grid cells with averaging kernel sensitivities less than 0.05. The calculation of the prior pre-conditioned Hessian still requires significant memory for large and . We parallelize the calculation manually and using Dask, a python program that performs parallelization for datasets that exceed available memory.

**2.2 State vector, prior estimate, and prior error**

In all inversions, we optimize emissions in 23,691 grid cells at 0.25° ⨉ 0.3125° resolution over North America, including all grid cells containing land or anthropogenic methane emissions larger than 0.1 Mg km-2 a-1, representing over 99% of methane emissions in North America. We also conduct inversions that optimize boundary four cardinal (north, south, east, and west) boundary condition elements. Methane chemical and soil sinks are not optimized because these loss processes are slow compared to the ventilation timescale.

Figure 1 shows the spatial distribution of major source sectors in the prior emissions estimate. Anthropogenic emissions for the United States, Mexico, and Canada are given by the spatially disaggregated (gridded) versions of the EPA GHGI for 2012 (Maasakkers et al., 2016), the INECC inventory for 2015 (Scarpelli et al., 2020), and the ECCC estimates for 2018 (Scarpelli et al., 2021), respectively. To account for changes in the distribution and magnitude of oil and natural gas emissions in the United States since 2012, we update the distribution of production fields using 2018 DrillingInfo data and scale the total natural gas production, transmission, processing, and distribution emissions to match 2018 emissions as reported in the 2020 GHGI. We also use the Environmental Defense Fund’s high-resolution inventory over the Permian basin for xxxx, one of the largest oil and natural gas producing regions in North America (Zhang et al., 2020). Because this inventory does not differentiate between oil and natural gas emissions due to the challenges of separating the sources, we treat oil and natural gas as a single sector in our analysis. All other anthropogenic emissions in the North American domain 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 seasonal scaling factors as described by Maasakkers et al. (2016) and Zhang et al. (2016), respectively.

Wetlands are the dominant natural source of methane emissions, but large uncertainties exist in their distribution and magnitude (Bloom et al., xxxx). We address this uncertainty by including two wetland inventories in our inversion ensemble. Figure 1 (bottom right) shows the scaled and subsetted versions of the high performance WetCHARTs ensemble version 1.3.1 used (Ma et al. 2021). Previous inversions found a large overestimate of wetland methane emissions in the high-performance ensemble, particularly in the boreal wetlands (Lu et al., 2022). The scaled inventory decreases total wetland emissions by a factor of 4.04 based on a comparison of the ensemble to FLUXNET CH4, a network of eddy covariance tower data (Shuang et al., ????). The subsetted inventory removes two ensemble members that produce anomalously large emissions in the high northern latitudes in summer and fall (Lu et al. 202?). Other natural methane emission sources include open fires, termites, and geological seeps. We use the same emissions for these sources as in Lu et al. (2022).

We assume uniform relative prior error standard deviations for emissions of between 50% and 100% for ach ensemble member. 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 increase errors up to 100% to account for displacement errors and increased error covariance at high resolution (Maasakkers et al. 2016). Ensemble members that optimize the boundary conditions use a base error standard deviation of 10 ppb, which is scaled by between 0.5 and 1 depending on the choice of grid cell prior error. We choose the prior error for each ensemble member as described in section 2.7. Although lognormal errors better represent the distribution of methane emissions and avoid negative posterior emissions (e.g., Maasakkers et al., 20?, Chen et al., 2022), the corresponding solution iteratively updates the Jacobian matrix, which is computationally prohibitive for our large observation and state vector dimensions (section 2.1). We instead use normal errors. In the absence of better information, we assume there is no error covariance.

**2.3 Forward model**

We use the nested version of the GEOS-Chem chemical transport model (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). Initial conditions for January 2019 and 3-hourly boundary conditions for the year are specified with methane concentration fields from a global GEOS-Chem simulation at 2° ⨉ 2.5° resolution using emissions optimized by a global inversion of TROPOMI observations (Qu et al. , 2021). These concentration fields are unbiased with respect to the global TROPOMI data and are informed predominantly by observations outside of North America. We validate GEOS-Chem by comparison to surface and aircraft methane observations for May 2018. We use observations from the Atmospheric Tomography Mission (ATom) i, the Atmospheric Carbon and Transport – America (ACT-America) campaign, and the NOAA Observation Package (ObsPack). We find a mean model-observation bias of 6.36 ppb and a correlation of R = 0.45. We also find no significant latitudinal bias in the model-observation difference, although the surface and aircraft observations provide significant coverage only between 30°N and 50°N.

**2.4 TROPOMI observations**

The Tropospheric Monitoring Instrument (TROPOMI) aboard the Sentinel-5 Precursor satellite 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 is inferred from a full-physics retrieval that can fail as a result of 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 the retrieval 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.

We evaluate the TROPOMI data using methane observations from the Greenhouse gases Observing SATellite (GOSAT). Launched in 2009, GOSAT provides high-precision observations of methane in 10 km diameter nadir pixels separated by ~250 km along- and cross-track. GOSAT measures backscattered solar radiation at 1.65 μm with a local overpass time of about 13:00 and a three-day return time. We use the GOSAT methane retrieval version 9.0 of the University of Leicester obtained by the CO2 proxy method (Parker and Boesch, 2020, last accessed 29 December 2020). We use only high-quality retrievals as indicated by the quality assessment flag. Due to the sparse coverage of GOSAT, we also evaluate the TROPOMI data using a GEOS-Chem simulation run with the prior emissions.

We compare average seasonal TROPOMI and GOSAT methane observations on a 2° ⨉ 2° grid following Lorente et al. (2021). We find large regional biases, defined as the standard deviation of the mean TROPOMI – GOSAT difference, of between 15 ppb (summer) and 20 ppb (winter). The winter-time biases are likely due to snow- and ice-covered scenes (Lorente et al. 2021). We identify these scenes using blended albedo, an empirical parameter that combines shortwave and near-infrared albedo and that correlates with snow- and ice-cover when greater than about 1 (Wunch et al. 2011). We remove scenes with blended albedo > 0.75 in fall, winter, and spring. We also remove scenes with shortwave albedo less than 0.05 following de Gouw et al., 2020. These scenes exhibit large prior GEOS-Chem – TROPOMI biases and disproportionately account for the remaining unphysical TROPOMI observations (XCH4 < 1700 ppb). Finally, we remove scenes north of 50°N in winter, which are likely to correspond with snow- and ice-cover and which exhibit anomalous prior GEOS-Chem – TROPOMI differences.

The applied filters increase the GOSAT-TROPOMI correlation in all seasons, with the largest increases in winter and spring (from R2 = 0.20 to 0.30 and from R2 = 0.32 to R2 = 0.49, respectively), when snow- and ice-cover are most likely to impact the TROPOMI retrieval. Seasonal regional biases decrease by between 7% and 21% and are in all cases less than the standard deviation of both the TROPOMI and GOSAT data. The mean TROPOMI – GOSAT biases are also consistent with the -10.3 ± 16.8 ppb bias found by Lorente et al. (2021).

After filtering the data, we find a residual aseasonal latitudinal bias in the TROPOMI – GEOS-Chem difference. This bias has been noted and corrected previously by Turner et al. (xxxx), Maasakkers et al. (xxxx), and Zhang et al. (xxxx). We define a latitudinal correction term (ppb) for the GEOS-Chem – TROPOMI difference using a first-order polynomial. For the inversion ensemble members that use scaled wetland emissions, we find , where is the degrees latitude. For the members that use the subsetted wetland emissions, we find . We also conduct sensitivity tests without the latitudinal correction. In these cases, we remove a mean bias of 8.53 ppb (scaled wetland emissions) and 9.11 ppb (subsetted wetland ensemble), which we attribute to errors in the boundary conditions.

Figure 2 (top row) shows the 2919358 final observations, regridded onto the GEOS-Chem grid and averaged seasonally, that constitute our observation vector **y**. The bottom row shows the observational density. We preserve 69% of the original high-quality data and find good agreement with the prior GEOS-Chem simulation (R between 0.58 and 0.60, depending on the wetland emissions prior and bias correction). We still find large, seasonally-variable gradients in the prior GEOS-Chem – TROPOMI difference (e.g., in spring over Northern Wisconsin), suggesting the possibility of residual systematic biases in the observations. We account for these biases in our observing system errors.

**2.5 Observing system errors**

Observing system errors include 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 in each grid cell 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 calculate the seasonal mean difference on a 2° ⨉ 2° grid to minimize the effect of systematic albedo biases in the TROPOMI data. In the <100 scenes where the residual standard deviation is less than the reported instrument error standard deviation (Lorente et al., 2021), we use the latter instead. We also set a minimum error of 10 ppb, which applies to between 31 and 32% observations. We find a mean observational error standard deviation of 11.5 ppb in all cases, with larger errors in winter and at high latitudes. The resulting variances are the diagonal elements of our observational error covariance matrix . Off-diagonal terms are assumed zero in the absence of better information. We introduce a regularization factor , described in section 2.7, to account for the lack of covariance (Chevallier et al., 2007).

**2.6 Jacobian matrix**

The relationship between simulated methane concentrations and emissions in the nested version of GEOS-Chem is strictly linear and is described by the Jacobian matrix . The Jacobian matrix is typically constructed by conducting a forward model simulation for each state vector element. While this is an embarrassingly parallel problem, constructing this matrix for the 23691 0.25° x 0.3125° resolution grid cells optimized by this inversion is computationally intractable. We take advantage of the heterogeneous information content of the TROPOMI observations to construct the Jacobian matrix at substantially decreased computational cost using the reduced-rank method introduced by Nesser et al. (2021). 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 to represent the influence of emission in grid cell *j* on grid cell *i*, and are the molecular weights of dry air and methane, 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. We assume that decreases exponentially as = {10, 6, 4, 3, 2.5, 0} for || *i-j* || = {0, 1, 2, 3, 4, >4}, where the norm ||.,.|| is the maximum difference in latitude or longitude grid cell index between *i* and *j*. This models a turbulent diffusion influence in concentric rings around grid cell *i,* with cut-off beyond a certain distance. The values of are chosen so that their sum over all grid cells equals unity. As pointed out by Nesser et al. (2021) Assuming a turbulent diffusion influence is important to decrease the sparsity of and increase its value as a first estimate, as explained by Nesser et al. (2021).

We use to calculate the initial patterns of information content to be perturbed in the forward model. The corresponding averaging kernel matrix captures the dominant patterns of information content through its dependence on the prior error covariance matrix and on the observational density as quantified by the observational error covariance matrix and by the sparsity structure of (Nesser et al., 2021). The initial patterns of information content are then given by the eigenvectors of the averaging kernel matrix calculated as where is the th eigenvector of the prior-preconditioned Hessian (Bousserez and Henze, 2018). We perturb the = 434 eigenvectors that span 50% of the initial information content in the forward model. We then apply an optimal operator that restores the original state dimension and minimizes information content loss to the resulting matrix to yield an updated reduced-rank Jacobian matrix estimate . We then recompute the eigenvectors, perturb the = 1952 eigenvectors that explain 80% of the information content, and construct the updated reduced-rank Jacobian matrix . This 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 yielded by an analytical inversion underestimates errors by failing to account for errors in the inversion parameters (Yu et al., 2021). We estimate these errors and reduce the noise associated with the reduced-rank Jacobian approximation by generating a quality-controlled ensemble of inversions, summarized in table 1. We generate eight base inversions by varying (1) the choice of prior wetland emissions estimate (either the scaled wetland emissions or the modified wetland ensemble), (2) the inclusion of the latitudinal correction to the model – observation difference, and (3) the use of a boundary condition correction. For each of these base inversions, we choose combinations of relative prior errors (50%, 75%, or 100%) and the regularization factor (between 0.001 and 1.0) so that the prior term of the cost function evaluated at the posterior solution averages to 1 across all grid cells where the averaging kernel sensitivities are greater than 0.05. We also require that at least 90% of these optimized grid cells are positive and that at least one grid cell is optimized for each of the 2386 model runs. This yields between one and three ensemble members for each of the base inversions, for a total of 15 members. We report the ensemble mean emissions, 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 sources, including sectors, states, and urban areas, providing a better basis for subsequent aggregation. We aggregate the native resolution emissions estimate and associated errors to find the corresponding quantities for each source category by 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 given 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. Urban areas are defined using the U.S. Census Urban Areas for 2010, which consist of the areas with a population greater than 5104. We consider the 280 urban areas with populations greater than 1e6 according to the 2010 Census. The reduced-dimension posterior estimate and associated posterior covariance and averaging kernel matrices are then given by

where is the Moore-Penrose pseudo inverse (Caliesi et al., 2005).

This approach to source attribution assumes that the distribution of emissions in the prior is correct and that emission sources are evenly distributed in grid cells that cross state or urban lines.

**3 Results and discussion**

Figure 3 shows the ensemble mean posterior correction factors relative to the prior emissions estimate (left) and ensemble mean averaging kernel sensitivities (right). Boundary condition optimization results are shown in table 2. We find 784 (338 - 1278) DOFS across the domain, where the values in parentheses are the ensemble minimum and maximum, respectively. Of these DOFS, 38 (11 - 69) are found in Canada, 648 (281 – 1058) in the contiguous United States (CONUS), and 89 (45 - 134) in Mexico. The high DOFS in CONUS reflects high and widely distributed emissions, while the low DOFS in Canada reflects the low TROPOMI sampling density at high latitudes. We find a large increase in information content relative to past inversions over the same domain: 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 high resolution on continental scales.

We use the averaging kernel sensitivities greater than 0.05 to define the grid cells that are optimized by the reduced-rank inversion. Optimized grid cells explain 21% (11% - 32%) of prior emissions in Canada, 74% (60% - 83%) in CONUS, and 66% (53% - 71%) in Mexico, reflecting the spatially variable information content. Across the optimized grid cells, we find a mean of 258 (40 - 572) negative values out of the 23691 grid cells optimized by the inversion across the ensemble. These values are mostly small and of the same order of magnitude as the soil sink.

We conduct an independent evaluation of the reduced-rank inversion by comparing our mean posterior emissions to a full-rank regional inversion of TROPOMI observations to quantify emissions from major oil and natural gas producing regions (Shen et al., 2022). [Insert]

The mean posterior emissions when used in GEOS-Chem decrease 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 National sectoral emissions**

Across the inversion ensemble, we find that relative to the prior, mean posterior anthropogenic and natural methane emissions decrease -0.8 (-2.2 - 0.1) Tg a-1 in Canada, increase 3.1 (1.5 - 4.7) Tg a-1 in CONUS, and decrease -0.7 (-1.3 - 0.1) Tg a-1 in Mexico. The indeterminate sign of the emissions change in Canada and Mexico results from the low fraction of prior emissions optimized by the reduced-rank inversion, which reflects the low information content available in these regions. As a result, we focus our discussion on CONUS.

We find posterior anthropogenic emissions of 31.4 (30.0 - 33.0) Tg a-1 in CONUS. This represents a 9% (5% - 15%) increase from our prior estimate of 28.7 Tg a-1 and a 17% (12% - 23%) increase from the most recent EPA Greenhouse Gas Inventory (GHGI) estimate of 26.8 Tg a-1 for 2019. Lu et al. (2022) found larger but still consistent anthropogenic methane 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) conducted a survey of GOSAT inversions and found similar median posterior anthropogenic methane emissions for the United States of 26.5 (20.8 – 38.7) Tg a-1 for 2019. The lower estimate in these inversions could reflect changes between 2017 and 2019, different prior emissions, or differences between GOSAT and TROPOMI. Notably, GOSAT has similar accuracy and precision to TROPOMI, but produces two orders of magnitude fewer measurements. Deng et al. (2022) also surveyed inversions of in situ data, which used high accuracy and high precision observations from surface sites across the United States. In this case, they found median posterior anthropogenic methane emissions of 31.9 (23.9 - 43.1) Tg a-1, which closely matches our result and suggests that the lower estimate results at least partly from TROPOMI’s improved observational density. This explanation is further supported by the increase in DOFS we find compared to Lu et al. (2022) and by the narrower posterior emissions ranges we find compared to both Lu et al. (2022) and both inversion ensembles in Deng et al. (2022).

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 Shatred in equation (8), we find very low posterior error correlation between the sectors (mean error correlation coefficients less than 0.2 in all cases), indicating that we can successfully separate sectoral emissions. Figure 4 shows the resulting prior and mean posterior emissions for each sector in Canada, CONUS, and Mexico and table 1 shows the same data for CONUS. Almost all significant changes in sectoral emissions are observed in CONUS, with the largest changes occurring for the biggest anthropogenic methane emission sources, including landfills, livestock, oil and natural gas, and coal. We find little adjustment to wastewater, wetlands, and other sources in part due to the observing system’s inability to correct for regions containing these sources. For the sectors with significant adjustments, we find sectoral averaging kernel sensitivities between 0.49 and 0.92, with an average increase of 64% relative to the equivalent values found by Lu et al. (2022), indicating that TROPOMI provides an improved constraint on sectoral emissions. We also find a narrower or equal ensemble spread compared to Lu et al. (2022). Each sector is also well constrained by the reduced-rank inversion: on average, between 65% and 81% of the prior sectoral emissions are optimized.

Landfill emissions show the largest relative and absolute increase from the prior emissions estimate. We optimize 81% (62% - 90%) of prior landfills emissions and find a sectoral averaging kernel sensitivity of 0.49 (0.31 – 0.66), indicating that on average the observing system informs 49% of the posterior estimate with the remainder provided by the prior. We find posterior emissions of 7.2 (6.4 - 8.2) Tg a-1, an increase of 27% (13% - 45%) relative to the prior and 58% (41% - 81%) relative to the most recent EPA GHGI estimate for 2019. 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 increase to underestimated urban emissions, including both direct landfill emissions and co-located post-meter natural gas emissions, which we discuss in section 3.3.

Livestock emissions show the second largest absolute increase from the prior emissions estimate. We optimize 65% (45% - 76%) of prior emissions and find a sectoral averaging kernel sensitivity of 0.67 (0.50 - 0.77). Posterior emissions of 10.5 (9.9 - 10.9) Tg a-1 represent a increase of 14% (8% - 19%) from the prior and 11% (5% - 16%) from the most recent EPA GHGI estimate for 2019. 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 United States and south central Canada to find mean posterior livestock emissions of 5.5 (5.1 - 6.2) Tg a-1, which agrees well with our posterior livestock estimate of 5.4 (5.1 - 5.6) Tg a-1 over the same region. We attribute the low EPA GHGI estimate to underestimated manure management emissions, which we discuss in section 3.4.

Our posterior emissions for fossil fuel emissions show a significant increase in information content produced by the improved observational density offered by TROPOMI. We optimize 88% (80% - 92%) of prior oil and natural gas emissions and almost all prior coal mining emissions, and our sectoral averaging kernel sensitivities are twice as large as those found in Lu et al. (2022). Posterior oil and natural gas emissions are 10.5 (10.2 - 10.9) Tg a-1, an increase of 11% (8% - 15%) from the prior emissions estimate and 15% (12% - 20%) from the most recent EPA GHGI estimate for 2019. The larger discrepancy from the updated EPA GHGI estimate reflects the inventory’s decreased production emissions despite increases in processing, transmission, and storage emissions and the addition of post-meter emissions. Coal mining emissions exhibit the largest decline in sectoral emissions, decreasing 49% (34% - 61%) relative to the prior and 30% (10% - 46%) relative to the most recent EPA GHGI estimate for 2019. The decrease in emissions is consistent the 30% decrease in CONUS coal production since 2012 (USGS).

We also find broad consistency with past studies of fossil fuel methane emissions. Lu et al. (2022) found much larger fossil fuel posterior emissions of 4.8 (3.1 - 4.9) Tg a-1 for oil, 8.9 (8.0 - 9.8) Tg a-1 for natural gas, and 2.9 (2.3 - 3.4) Tg a-1 for coal for 2017. However, they found decreasing natural gas emissions beginning in 2014 that could explain part of the discrepancy if they continued until 2019. Shen et al. (2022) found slightly lower oil and natural gas emissions of 12.6 ± 2.1 Tg a-1 from an inversion of TROPOMI data over oil and natural gas basins extrapolated to the national scale for May 2018 to 2020, which is consistent with continued emissions decreases. Deng et al. (2022) found much smaller cumulative fossil fuel emissions of 9.8 (8.1 - 13.7) Tg a-1 for GOSAT inversions and 12.6 (8.0 - 16.7) Tg a-1 for in situ inversions in 2019, both of which are consistent with our result.

**3.2 Emission inventories for individual states**

The high resolution achieved by our inversion allows for finer-scale source attribution. Within CONUS, we partition emissions, excluding offshore emissions, to each of the 48 states as described in section 2.8. Figure 5 shows state prior and posterior emissions, table 4 shows the corresponding data for the top 10 methane producing states, and table S1 shows the full results. TROPOMI provides a strong constraint at this resolution, with sensitivity to TROPOMI observations larger than 0.5 in 29 of all 48 states and in 20 of the top 25 methane producing states. Our reduced-rank inversion also constrains emissions in most states: we optimize an average of 80% of prior emissions in each of the top 25 methane producing states.

Emissions are consolidated in the largest methane producing states, with 7 states responsible for 50% of mean posterior emissions in CONUS (8 for prior emissions) and 7 states responsible for more than 54% of inferred emissions increases. Texas alone produces 21% of the mean posterior emissions (18% in the prior) and 25% of the inferred emissions increase. Oil and natural gas emissions account for 66% of the observed increase in Texas and are concentrated over the Permian basin, the largest oil and natural gas producing region in CONUS (cite). We also find significant increases over the Texan portions of the Haynesville shale and over the Houston metropolitan area. More generally, we find significant emission increases in 7 of the top 10 methane producing states, with a significant decrease in New Mexico and no significant change in Pennsylvania and Kansas. Oil and natural gas are responsible for 47% of mean posterior emissions and 44% of the observed increase in these 10 states, compared to the national average of 33% and 36%, respectively.

Shifting trends in methane emission sources since 2012 result in a significant re-ordering in the largest methane producing states from the prior to the posterior. The largest increases in ranking occur in Florida (21st to 8th), Indiana (26th to 15th), Georgia (29th to 19th), and Mississippi (33rd to 24th). Of these, Indiana, Georgia, and Mississippi are sensitive to the observing system, with mean state averaging kernel sensitivities between 0.5 and 0.6, while Florida has a mean sensitivity of 0.32. In each of these three states, the observed ranking increase is driven by increased landfill emissions. The largest decreases in emissions rank all occur in coal-producing states, including Wyoming (14th to 42nd), West Virginia (5th to 20th), Kentucky (17th to 27th), and Virginia (19th to 28th), all of which have state averaging kernel sensitivities greater than 0.5. This reordering reflects a systematic underestimate of urban emissions (section 3.2) and a shift in the dominance of methane emission sources in the last decade.

We compare our posterior emissions to the EPA state emissions inventories for 2019, scaled to match the total mean posterior emissions. On average, our mean posterior emissions are only 4% smaller than the scaled EPA estimates. However, the distribution of the posterior – inventory differences is highly skewed, with the largest relative discrepancies found in the largest methane producing states: in the top 10 methane producing states, we find that the state posteriors are on average 18% larger than the scaled EPA estimates. Consistent with our sectoral analysis, the largest absolute negative posterior – inventory differences are found in historically coal-producing states, including Wyoming, Pennsylvania, and West Virginia. The largest positive differences occur in states with large oil and natural gas emissions, including Texas, Louisiana, and New Mexico.

We also compare our posterior emissions to state greenhouse gas inventories for the year closest to 2019 where available and where the observing system provides a strong constraint (state averaging kernel sensitivity greater than 0.5). Of the inventories of California, Colorado, Iowa, Louisiana, and Pennsylvania, our posterior agrees only with Pennsylvania. While posterior emissions are not significantly different from the prior in Pennsylvania, 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). The inventories of California, Iowa, and Louisiana are smaller than our mean posterior emissions, while Colorado finds larger emissions, with different driving sectors in each state. Our prior and mean posterior estimates are larger than the California inventory by 13% and 32%, respectively, with 78% of the inferred increase attributable to livestock and 14% to oil and natural gas. In Iowa, our prior estimate is only 4% smaller than the inventory value, but we find a mean 33% increase in emissions that is attributed to livestock (70%) and landfills (18%). The Louisiana inventory finds methane emissions half as small as our prior, and we find an additional mean 65% increase in emissions attributed to oil and natural gas (71%) and other sources, predominantly rice (16%). The Colorado inventory is larger by 44% than our prior estimate and 65% than our mean posterior emissions. The difference between the inventory and the inversion is due to oil and natural gas, where the inventory estimates emissions more than twice as large as our prior.

[Add comparison to other state papers, especially states in the top 10]

**3.3 Emission inventories for individual urban areas**

We calculate the mean posterior emissions for the 262 urban areas defined by the U.S. Census Bureau for 2010 (section 2.8) with a 2010 population greater than 1e6 and that are optimized by at least one of the 15 inversion ensemble members. Most past studies of urban areas relied on in situ and aircraft observations and therefore quantified emissions in a small (<10) number of cities. Plant et al. (2022) used TROPOMI data to analyze urban methane emissions in CONUS, but they limited their analysis to eight cities. To our knowledge, this study is the first to comprehensively analyze urban emissions in CONUS.

The 262 analyzed urban areas occupy less than 3% of CONUS surface area but are responsible for 6.8% of prior emissions. The observing system provides a weaker constraint on urban areas than on sectors or states due to smoothing errors, the larger instrument errors incurred by the heterogeneity of the urban grid, and the increase in model transport errors at high resolution. We find mean urban averaging kernel sensitivity values of 0.19 (0.09 - 0.30). Despite the weak constraint, we still find significant corrections: posterior emissions are on average 29% (16% - 44%) larger than the prior so that urban areas are responsible for 8.3% (7.7% - 7.2%) of total emissions.

The observed increase in urban emissions suggests an underestimate of landfill emissions or natural gas distribution or post-meter emissions. The EPA added post-meter emissions of 456 Gg a-1 to their most recent GHGI for 2019, representing 5% of oil and natural gas emissions. This addition explains 69% (43% - 130%) of the inferred net adjustment of 658 (350 - 1066) Gg a-1 to urban areas. We find that we are unable to explain the remaining discrepancy because source co-location results in high sectoral covariance. We also find no correlations between the observed increase in either aggregated urban areas or individual urban grid cells and urban area population for 2010, population change from 2000 to 2010, population density, or surface area. The lack of correlation reflects the variability of methane emission sources and trends between urban areas, as supported by earlier studies. Floerchinger et al. (2021) found “dramatically different” emission profiles in an aircraft survey of methane and ethane over seven U.S. cities, with the fraction of emissions attributable to natural gas ranging from 0.32 to 1.0. Sargent et al. (2021) surveyed 12 studies spanning six cities and found that natural gas contributions to urban emissions ranged from 43% to 88%. Both studies included geographically diverse cities with a range of populations. Plant et al. (2019) used aircraft data over six cities on the East Coast of CONUS and found that on average between 80% and 110% of emissions in each city were explained by natural gas emissions. However, their confidence intervals ranged from a minimum of 45% to a maximum of 170%, reflecting significant uncertainty in the source attribution.

We look in detail at the top 10 methane producing urban areas as ranked by mean posterior emissions, shown in Figure 6. We find a larger constraint in these cities compared to the average, with an average mean urban averaging kernel sensitivity of 0.40. As found in our states analysis, methane emissions are consolidated in the regions with the largest emissions. These top 10 methane producing cities are responsible for 39% of prior methane emissions and 45% (42% - 49%) of mean posterior methane emissions in the 262 urban areas analyzed. We find a much larger mean increase relative to the prior of 78% compared to the average of 29%. Emissions more than double in four of the 10 cities: Detroit (130% increase), Miami (189% increase), Atlanta (110% increase), and Indianapolis (134% increase). The large increase in Miami may be due to aliasing co-located wetlands, which is reflected in the larger range of posterior estimates across the ensemble. We find a decrease in emissions only in Los Angeles.

There is considerable overlap between the top 10 methane producing cities and the 10 largest cities as ranked by the 2010 Census population. This reflects the relationship between population and both landfill emissions and distribution and post-meter natural gas emissions. However, we find no correlation between population and the inferred emissions change. Indeed, the three largest cities by population in CONUS (New York, Los Angeles, and Chicago) have the three smallest per capita methane emissions within the top 10. Washington and Boston, top 10 cities by population, are excluded from the list of the top 10 methane producing cities. Boston is likely excluded from the list due to the low constraint provided by TROPOMI rather than low emissions: we find a mean urban averaging kernel sensitivity of 0.10. Detroit and Indianapolis are included instead, pointing to the importance of smaller, non-coastal population centers for methane emissions. These two cities have the largest per capita posterior methane emissions in the top 10, followed closely by Dallas. Moreover, these cities are understudied compared to the larger urban centers: of 17 studies of urban methane emissions published since 2015, only three considered emissions from cities outside of coastal metropolitan centers. These three studies quantified emissions from two cities: Kansas City (Plant et al., 2022) and Indianapolis (Lamb et al., 2016 and Jones et al., 2021).

We compare our posterior emissions to previous studies. While urban methane emissions tend to be relatively stable over time (e.g., Fairley and Fischer 2015; Sargent et al., 2021), we limit ourselves to studies published since 2015 to minimize the effect of population and infrastructure changes on the comparison. Across the 17 studies surveyed, we find smaller emissions on average, suggesting that our estimate of city methane emissions may be conservative. Some of the underestimate may be result from the specification of urban extent; our restrictive definition may generate smaller total urban emissions. We also incur larger errors by using satellite observations instead of in situ and aircraft data, which are more sensitive to the variable topography and albedo of the urban grid. Moreover, we find the largest discrepancies over cities with complex coastlines, including San Francisco (Fairley and Fischer, 2015; Jeong et al., 2016, 2016; Guha et al., 2022), Washington (D.C.; Huang et al., 2016; Ren et al., 2018; Lopez-Coto et al., 2020; Plant et al., 2019, 2022), Boston (Plant et al., 2019, 2022; Sargent et al., 2021), and Baltimore (Huang et al., 2016; Ren et al., 2018; Lopez-Coto et al., 2020; Plant et al., 2019, 2022), suggesting that TROPOMI may have a coastal bias. We find smaller discrepancies over Philadelphia and Atlanta, where Plant et al. (2022) finds mean emissions that are larger than but still consistent with our posterior using TROPOMI methane : carbon monoxide ratios and carbon monoxide inventories. We find consistent results with all surveyed studies in New York City and Indianapolis (Plant et al., 2019, 2022; Pitt et al., 2022; Lamb et al., 2016; Jones et al., 2021).

We also compare our posterior emissions to methane estimates from city greenhouse gas inventories. Of the 12 largest cities by posterior emissions or by 2010 population, we find that only New York City and Philadelphia estimate methane estimate. In both cases, the inventory value is lower than both our prior and our posterior results. Indeed, the inferred increase in methane emissions in both cities results in posterior emissions more than twice as large as the inventory value.

**3.4 Livestock emissions**

[Insert work on identifying the cause of the livestock increase.]

**4 Conclusions**

[Insert.]